Epigenetic differences after prenatal adversity: The Dutch Hunger Winter

Elmar Wouter Tobi

Voor mijn familie, en in het bijzonder zij die hebben geleden onder gevangenschap, dwangarbeid en honger.

Epigenetic differences after prenatal adversity: The Dutch Hunger Winter
Ir. Elmar W. Tobi

Financial support for the printing of this thesis was provided by the Netherlands Consortium for Healthy Ageing (NGI 050060810) and Sequenom Inc.
The picture forming the basis of the cover was kindly provided by: Beeldbank W.O.2 - NIOD

PhD thesis with summary in Dutch
ISBN: 978-90-8891-714-1

Cover design: Proefschriftmaken.nl || Uitgeverij BOXPress
Printed & Lay Out by: Proefschriftmaken.nl || Uitgeverij BOXPress
Published by: Uitgeverij BOXPress, 's-Hertogenbosch

Epigenetic differences after prenatal adversity: The Dutch Hunger Winter

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College van Promoties
te verdedigen op dinsdag 29 oktober 2013
klokke 16:15 uur

door

Elmar Wouter Tobi
Geboren te Delft
in 1981

Promotiecommissie

Promotor:	Prof. Dr. P.E. Slagboom
Co-promotores:	Dr. B.T. Heijmans
	Dr. L.H. Lumey *Columbia University, NY, USA*
Overige leden:	Prof. Dr. Ir. S.M. van der Maarel
	Prof. Dr. B.J. Zwaan *Wageningen University*
	Prof. Dr. R.P.M. Steegers-Theunissen *Erasmus Medical Center*

Content

Introduction

Genes, environment and something else

The central dogma of biology states that heritable information is passed residue-by-residue from DNA to RNA and finally protein. Moreover it states that such information cannot be transferred back to protein or nucleic acid[1]. However, the concept of the gene as the sole container of biological information has been challenged by several observations. For instance, the phenomenon that a gene's activity is determined by its place in the genome[2] is explained by the fact that information is also contained in the positional context of genes in the genome. Another observation is that genetic mutations and environmental perturbations during development can give rise to the same phenotype[3]. More recently an enigmatic 'third component' has been found that causes phenotypic variation during development that persists into adulthood in inbred or even monogenetic populations under a constant environment[4,5]. Indeed, it has proven very hard to produce an identical copy of an organism. When the first cat, called carbon copy (a.k.a. copycat), was cloned it had a completely different coat coloring than the mother[6]. All these observations lead to the question what molecular marks could mediate these positional effects, the influence of the early environment or the seemingly stochastic variations in genetically identical organisms.

It is believed that so-called epigenetic marks which envelop DNA may be the mediator. The study of 'epigenetics' has exploded since the year 2000[7], but different researchers have different definitions, which is due to the fact that the term epigenetics arose multiple times during the previous century (**Box 1**)[8]. In this thesis we adhere to the definition by Jaenisch & Bird, namely that of epigenetics as the study of the molecular mechanisms by which heritable changes in gene expression potential occur that are not caused by changes in DNA sequence[9]. This definition calls for a cell-autonomous nature of epigenetic information that is passed during mitosis and possibly meiosis and excludes sustained expression changes mediated by extracellular signals or by morphology[10]. These epigenetic marks may contain stable genomic information potentially providing the molecular basis to explain part of the phenotypic variation in humans[11]. Research on epigenetic marks in human

populations is focusing in particular on those diseases occurring in adulthood that are linked with disturbances of the early environment[12]. In this thesis we lay the first basis to ultimately elucidate the role of epigenetic marks in early development and disease in humans.

Box1: the origins of epigenetics

The term epigenetics was originally coined by Waddington[72], who in 1939[73] started his attempts to conceptually merge embryology, evolution and genetics. He hypothesized that a cell, tissue or organism is formed through dynamic interconnected networks of genes that interact with the environment during development. He put forward the term 'epigenotype' as the whole of these 'organizing relations' standing between the genotype and the phenotype and 'epigenetics' as the discipline studying the epigenotype. His groundbreaking network view of genetics and biology did originally not envision mechanisms outside the gene and the environment[74].

The second major usage of the term epigenetics was introduced by Nanney[75] and refined by Harris as the study of mechanisms that regulate the expression of the genetic information[76]. This concept of epigenetic control was soon adapted to describe the mechanisms underlying cellular inheritance other than that mediated by DNA[77], and to denote mechanisms underlying cellular differentation[78]. These usages denoted mechanism independent or complementary to genes. This is in contrast to Waddington's original usage, although Waddington noted that both concepts 'do not bite' each other.

It has become clear that the molecular mechanisms regulating the potential of a genomic region to become transcribed may be effectuated by genes[79] and shaped by genetic variation[14,30] and environmental conditions during early development[38]. These new insights are very much in the spirit of Waddington's concept of the epigenotype. This has led to a renewed enthusiasm for Waddington's network view on development and biology, which has proven highly influential for recent refinements of evolutionary theory[80,81]. Moreover, (molecular) geneticists coming from the tradition of the second emergence of the term epigenetics recently proposed a merger of Waddington's original theoretical framework with the expanding body of knowledge on 'epigenetic' gene expression regulation during differentation[82]. Development is back in the heart of Biology, Genetics and Evolution by partly re-inventing the concept of the epigenotype and its study: epigenetics.

Epigenetic marks: wrapping DNA

Epigenetic marks are intimately linked to development, for the same genome expresses different parts of its information depending on the cell type. During development different sections of the genome are activated and deactivated and loop together into foci of relative high or low activity[13]. In all organisms in nature the compaction and relaxation of DNA is mediated by the histone proteins forming so-called nucleosome complexes with DNA (**Figure 1**). Multiple of these DNA-protein structures together from a quaternary structure called chromatin and the modifications on the 'tails' of the histones in the

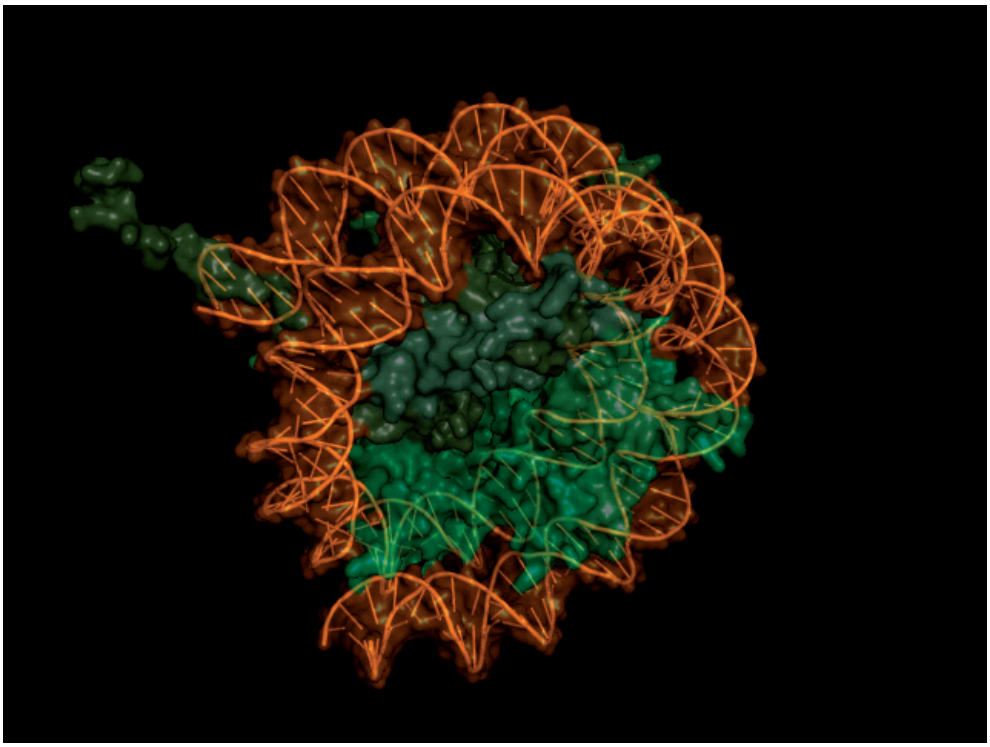


Figure 1. DNA wrapped along a nucleosome core
DNA is wrapped around nucleosomes, complexes of several histone proteins, in the cell core. Visible are the histones in different shades of green, together forming a nucleosome. On several locations the histone proteins protrude over the DNA helix (orange), locking it into place. The strength of this interaction is influenced by modifications placed on these histone tails.

complex help to determine its compactness. DNA positioned within regions with nucleosomes in a more open, euchromatin conformation is more readily transcribed into RNA. DNA positioned within regions where nucleosomes are more densely compressed, in a so-called heterochromatin conformation, is less readily transcribed. The DNA sequence itself influences nucleosome positioning and chromatin formation[14]. However, the correct chromatin state was shown to be faithfully transmitted to daughter cells in yeast despite the deletion of the relevant genetic signals[15] and also the prenatal environment can persistently change chromatin modifications[16,17]. Other, more recently discovered mechanisms may also influence the chromatin landscape. Various forms of non-protein coding RNA is transcribed from the genome and influence chromatin conformations within and between cells[18] and have even been found to influence the epigenetic make-up of the germ line, thereby potentially influencing a following generation[19].

Epigenetic marks: modifying DNA

Several families of the tree of life have also developed covalent modifications of DNA[20]. Methylation of DNA is limited to the cytosine in CpG dinucleotides in adult humans and several other complex species[21]. DNA methylation results in a more compacted double helix[22]. DNA methylation may also inhibit or promote the binding of certain proteins to DNA. Like chromatin, DNA methylation is highly dynamic during development and functions as a regulator of tissue differentiation[23]. During and directly after fertilization DNA is passively demethylated in the maternal contribution and actively demethylated in the paternal contribution to the zygote's new genome and during this time several intermediates of 5-methyl cytosine are formed through not yet completely elucidated mechanisms[24]. However, the genome is quickly remethylated during the earliest stages of development[25]. For instance, the promoters of genes that regulate pluripotency or lineage commitment are methylated upon differentiation, thereby persistently blocking a return to a more pluripotent cell type[26,27]. Secondly, DNA methylation can stably silence specific alleles of so-called imprinted genes dependent on the parent of

origin[28], a phenomenon that defies classical Mendelian inheritance patterns. DNA methylation is influenced by certain DNA sequences[29] and single nucleotide polymorphisms in cis and trans[30]. However, DNA methylation is also influenced by the before mentioned non-coding RNAs and by environmental factors during development[31]. The latter is the topic of this thesis.

Development, environment and epigenetic change: a hint from animal models

It is suggested by the epidemiological literature that early environmental conditions are linked to later life health and disease[32]. The nature of these associations is and has been hotly debated[33-35]. Experiments modulating animal nutrition and stress levels during development revealed that the precise timing of the environmental exposure and the sex of the exposed fetus is important in determining the phenotypic outcome[36]. Most experiments entailed limiting the amount of protein, calories or folic acid to the developing fetus[37]. From 2003 the wider research community became aware of epigenetic marks as a possible molecular link between nutrition during development and adult phenotypes. In two mouse models, with a repetitive element inserted in front of the gene that influences coat color (*agouti*)[38] and tail shape (*axin* fused)[39] respectively, it was handsomely shown that the amount of methyl donors in the maternal diet could shift the amount of DNA methylation at these elements (**Figure 2**). Such shifts in methylation changed the amount of expression of the neighboring gene and thereby the phenotype. Not long afterwards other examples at 'regular' genes were found[31,40-43]. Several of these experiments found epigenetic changes that influenced genes which have been functionally implicated with health effects that arise as a consequence of prenatal malnutrition. For instance, prenatal protein restriction changed the amount of DNA methylation and expression of the *agtr1b* gene, which is implicated in hypertension[44].

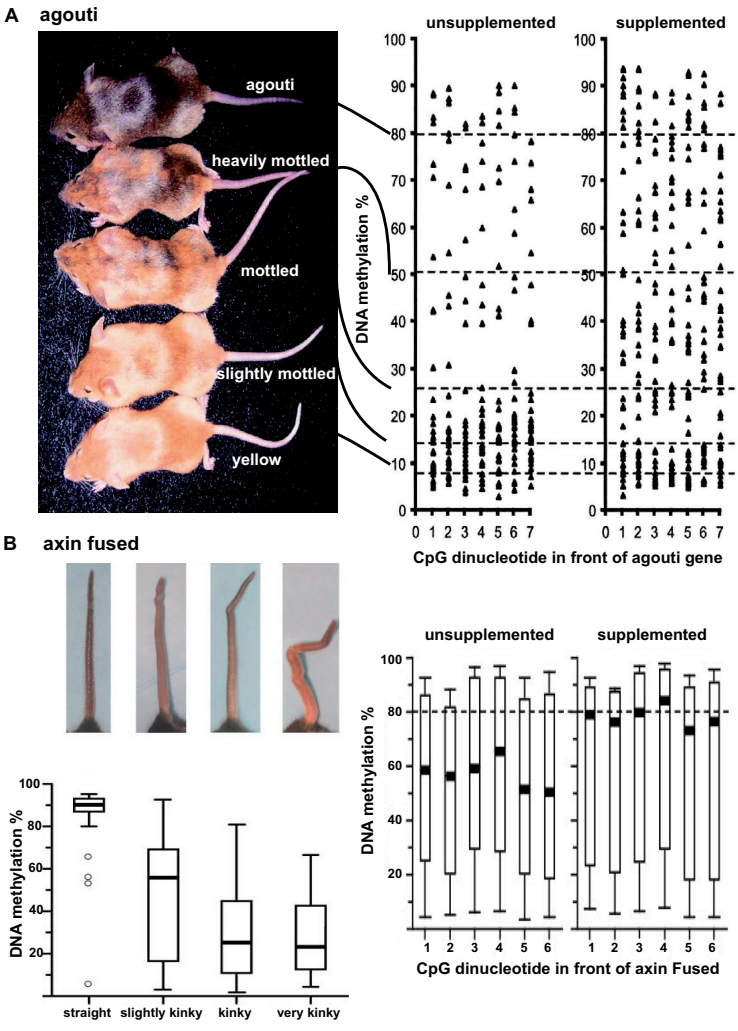


Figure 2. The effect of Supplementation on DNA methylation and phenotype in the *agouti* and *axin* Fused mouse models

A. In a normal litter there are a variety of colors from yellow to agouti (left panel). Upon supplementation of the maternal diet during pregnancy with methyl donors, like folic acid, more offspring is born with an agouti color. This is due to a general increase in the average methylation of 7 CpG dinucleotides at a retrotransposon situated in front of the *agouti* gene (right panel). Copyright© 2003, American Society for Microbiology, *all rights reserved*

B. In a cross resulting in heterozygous *axin Fused* litters, a great incidence of kinked tails is found. Tail form is influenced by DNA methylation in front of the *axin* locus (left panel). Upon supplementation of the maternal diet during pregnancy with folic acid this incidence is greatly decreased (by over 50%). It was found that folic acid supplementation increased the amount of DNA methylation at 6 CpG dinucleotides at this locus (right panel), resulting in litters with more mice with straighter tails. Copyright© 2006, John Wiley and Sons, *all rights reserved*

Development and disease

These animal models were intended to clarify and give body to the hotly debated epidemiological observations in relation to the associations between a low birth weight and uterine growth restriction and an unfavorable body mass index (BMI), type 2 diabetes (T2D) and hypertension[32]. Prenatal growth restriction and a low birth weight are seen in this context as a proxy for malnutrition *in utero*. Recently this view was challenged by observations in large twin cohort studies that have shown that the association between birth weight and BMI or T2D is confounded by genetic factors[45-47] and a large study in parent-offspring trios showing that the association with BMI is not only confounded by genetic but also familial factors[48].

Studies on the consequences of prenatal exposure to famine, which can be analyzed as natural experiments of sorts, show that obesity and diabetes may still arise by a poor prenatal environment and nutrition. An association between prenatal famine exposure and obesity has been found following the Dutch and the Great Leap Forward famines[49-52], where the associations in later life were most prominent in exposed women. The occurrence of diabetes was also found to be increased in two Dutch cohorts[53,54] and in a large population based study in the Ukraine[55]. Another well replicated finding is the association of periconceptional famine exposure with schizophrenia[56]. These replicated findings are either independent of the gestational timing of the famine exposure or specific to early exposure (**Figure 3)**, which unlike mid and late gestational famine exposure is not associated with a reduced birth weight[57]. The famine and twin studies raise questions on the suitability of birth weight and in utero growth restriction as a proxy for prenatal malnutrition, at least in Western cohorts. More defined maternal and environmental characteristics, like maternal BMI, hypertension and gestational diabetes are also associated with the same later life phenotypes associated with a low birth weight[58-60].

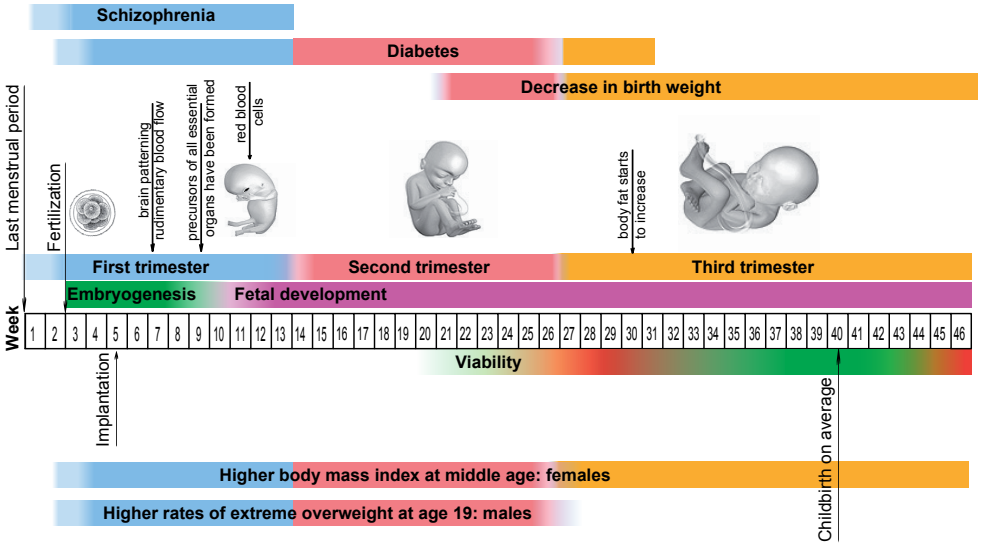


Figure 3. Replicated phenotypic associations with prenatal famine exposure
Human gestation is depicted by week across the 3 trimesters with above the replicated associations found to be independent of gender and below the replicated associations with confirmed or possible gender specificity. *Adaptation from original source on Wikimedia Commons, Mikael Häggström*

Of mice and men: variation and starting material

The data from animal malnutrition studies led to hypothesis that epigenetic processes may underlie the above mentioned link in humans between the prenatal environment and adult health[10,12]. For human studies it was clear that there would be challenges to overcome. First, the animal studies indicated that the differences induced by malnutrition could be small[42] and when the average difference was larger there was often quite some variation, as observed in the *agouti* and *Axin* Fused mouse models[38,39] (**Figure 2A en 2B, right panels**). Secondly, it was shown in animal models that not only the prenatal period but also the postnatal period may be crucial[43,61] and that the postnatal environment may counter effects induced by prenatal conditions[62]. It was also shown in animal experiments that one exposure may counteract the epigenetic effects of another[63,64]. Since human populations are never under controlled homogeneous conditions these latter studies

indicate that confounding will be a serious issue to deal with when studying epigenetic marks in human population studies. Furthermore, both genetic factors[30] as stochastic factors during ageing may influence epigenetic marks in humans[65,66], thereby adding biological noise to studies.
There is also the issue of tissue specificity. Since epigenetic marks are one of the main mechanisms underlying cell differentiation, differences between tissues are bound to be plentiful. In bio banked cohorts often only DNA from whole blood or buccal swaps is available, which are peripheral tissues that may have a different epigenetic pattern to internal tissues[67]. Whole blood is a mixture of highly differentiated cell types which may cause additional non-technical variation, since blood cell populations differ between persons and over time. Both these available sources of DNA are in most cases only suitable for DNA methylation measurements, since the native chromatin structure and non-coding RNA is lost during DNA isolation or storage.

A developmental extreme: The Dutch Hunger Winter Families Study

To reduce variation and increase the chance of success the first genetic studies in humans resorted to family studies, well defined exposures and phenotypic extremes. In this thesis we undertook the same strategy for human epigenetic studies on DNA methylation. In September 1944 the Dutch national railways went on a national strike in support of Allied operation Market Garden, aimed at opening up the main Nazi manufacturing center, the Ruhr region, for a direct assault. This operation stalled after liberating the Southern part of The Netherlands. Little to no food was transported by the Nazi's as military transports were given a higher priority and later on food transports by rail and road were prohibited as a punitive measure, while transportation over water was made difficult by winter conditions and severe fuel shortages. The official government issued rations fell rapidly and went below 1,000 kcals/day by November 1944, reaching as little as 500 kcal/day in April 1945. Before the famine the Dutch population was well fed and the famine rapidly ended after liberation in May by massive Allied relief efforts[68].

This discreteness in time, shorter than the nine months of human gestation, and the fact that the health care system remained working, makes that adults can be traced that were exposed during specific periods of their development *in utero*.

For this thesis we make use of The Dutch Hunger Winter Families study[69], which consists of 658 individuals born between 1943 and 1947 in three hospitals in Holland, located in cities exposed to famine in the winter of 1944-'45. For 313 of these individuals a same-sex sibling could be recruited as an unexposed control group. The disaster of the Dutch Hunger Winter, as captured in this cohort, offers a unique quasi-experimental setting for study. We focused our measurements on the individuals that were exposed either early or late in gestation and with a same-sex siblings available for study. This design (partly) matches for genetics and the early familial environment, reducing some of the variation inherent to human studies. Furthermore, the focus on exposure early or late in gestation offers an interesting contrast in relation to the observations relating to a low birth weight and *in utero* growth restriction. Individuals exposed during the last trimester of pregnancy were much lighter at birth, while individuals conceived during the famine and exposed up to 10 weeks into development were not[57] (**Figure 3**). Furthermore, experiments in animals indicate that the period around and just following conception (periconceptional exposure) is a particularly sensitive period of development[31]. Indeed it may be the period during which some of the epigenetic differences arise that were initially discovered following exposure during the entire pregnancy[63]. Differences induced early in development may be mitotically heritable, thus propagated to multiple tissues[70] and we hypothesize that differences induced during this period may be more readily detectible in whole blood, the tissue collected in this cohort, and is more likely to reflect differences in relevant but inaccessible tissues.

Aims and outline of this thesis

In this thesis we aimed to lay the groundwork needed for epigenetic epidemiology in human populations and take the first steps to ascertain if the associations between early development and adult health may in part be mediated by epigenetic mechanisms, as was indicated in animal models. To this end we first studied the variation, stability and tissue specificity of DNA methylation patterns in human blood and buccal cell DNA of candidate loci chosen for their key roles in development, growth and metabolism. Characterizing DNA methylation and investigating its suitability as a marker for molecular epidemiological studies (**Chapter 2**).

Secondly we aimed to discover if we could find associations between DNA methylation and prenatal famine exposure in humans. In particular we hypothesize that most associations should be found in those exposed during periconception. As a proof-of-principle, we studied whether DNA methylation within the key imprinted developmental gene *insulin like growth factor 2* (*IGF2*) is associated with prenatal exposure to the Dutch famine (**Chapter 3**). We then extended this first measurement in the Dutch Hunger Winter Families Study to the fifteen other candidate loci investigated in chapter two. With the aim to investigate if the associations between DNA methylation and prenatal famine mirror the epidemiological findings with timing independent and sex-specific associations (**Chapter 4**).

We then set our observations in the Dutch Famine in the context of prenatal growth restriction and more contemporary prenatal adversities such as maternal hypertension (**Chapter 5**). For this we resorted to the POPS study, which is also a cohort at a developmental extreme. The POPS study is a nation-wide prospective study including 94% of all live born infants born very preterm (<32 weeks) and/or with a very low birth weight (<1500 gram) in 1983[71]. We compared 113 individuals born preterm, but with a normal size for their gestational age, with 38 individuals born preterm and in uterine growth restricted. This to investigate if DNA methylation differences at loci identified in the Dutch Famine are also detectable after early prenatal growth restriction.

Finally, we set out to characterize the DNA methylation differences associated with prenatal famine exposure by extending our measurements to multiple regulatory regions within the *IGF2* locus and contrast the influence of the famine with that of genetic variation at the same locus (**Chapter 6**). After delving deeper, we broadened our inquiries by extending our measurements to a genome-scale using next generation bisulfite sequencing (**Chapter 7**). Investigating what would constitute the normal effect size and the genomic characteristics of regions sensitive to the prenatal environment and extend our analyses away from single genes and loci to entire regions and pathways. The experimental chapters are followed by a summary of the results (**Chapter 8**) and a general discussion (**Chapter 9**). We conclude with a popular summary of the results in Dutch (**Chapter 10**) and the acknowledgements section.

References

1. Crick F. Central dogma of molecular biology. Nature 1970; 227: 561-3
2. Girton JR, Johansen KM. Chromatin structure and the regulation of gene expression: the lessons of PEV in Drosophila. Adv.Genet. 2008; 61: 1-43
3. Goldschmidt R. The material basis of evolution. New Haven: Yale University Press, 1940
4. Gartner K. A third component causing random variability beside environment and genotype. A reason for the limited success of a 30 year long effort to standardize laboratory animals? Int.J.Epidemiol. 2012; 41: 335-41
5. Vogt G, Huber M, Thiemann M, van den Boogaart G, Schmitz OJ, Schubart CD. Production of different phenotypes from the same genotype in the same environment by developmental variation. J.Exp.Biol. 2008; 211: 510-23
6. McVittie B. The cloned moggy with two mums, 2006
7. Haig D. Commentary: The epidemiology of epigenetics. Int.J.Epidemiol. 2012; 41: 13-6
8. Haig D. The (dual) origin of epigenetics. Cold Spring Harb.Symp.Quant.Biol. 2004; 69: 67-70
9. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet. 2003; 33 Suppl: 245-54
10. Waterland RA, Michels KB. Epigenetic epidemiology of the developmental origins hypothesis. Annu.Rev.Nutr. 2007; 27: 363-88
11. Jablonka E. Epigenetic epidemiology. Int.J.Epidemiol. 2004; 33: 929-35
12. Gluckman PD, Hanson MA, Cooper C, Thornburg KL. Effect of in utero and early-life conditions on adult health and disease. N.Engl.J.Med. 2008; 359: 61-73
13. Beisel C, Paro R. Silencing chromatin: comparing modes and mechanisms. Nat.Rev. Genet. 2011; 12: 123-35
14. Iyer VR. Nucleosome positioning: bringing order to the eukaryotic genome. Trends Cell Biol. 2012; 22: 250-6
15. Wheeler BS, Ruderman BT, Willard HF, Scott KC. Uncoupling of genomic and epigenetic signals in the maintenance and inheritance of heterochromatin domains in fission yeast. Genetics 2012; 190: 549-57
16. Sandovici I, Smith NH, Nitert MD, Ackers-Johnson M, Uribe-Lewis S, Ito Y, et al. Maternal diet and aging alter the epigenetic control of a promoter-enhancer interaction at the Hnf4a gene in rat pancreatic islets. Proc. Natl.Acad.Sci.U.S.A 2011; 108: 5449-54
17. Zheng S, Rollet M, Pan YX. Maternal protein restriction during pregnancy induces CCAAT/enhancer-binding protein (C/EBPbeta) expression through the regulation of histone modification at its promoter region in female offspring rat skeletal muscle. Epigenetics. 2011; 6: 161-70
18. Sabin LR, Delas MJ, Hannon GJ. Dogma derailed: the many influences of RNA on the genome. Mol.Cell 2013; 49: 783-94
19. Daxinger L, Whitelaw E. Understanding transgenerational epigenetic inheritance via the gametes in mammals. Nat.Rev.Genet. 2012; 13: 153-62
20. Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, et al. Conservation and divergence of methylation patterning in plants and animals. Proc.Natl.Acad.Sci.U.S.A 2010; 107: 8689-94
21. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat.Rev. Genet. 2010; 11: 204-20
22. Severin PM, Zou X, Gaub HE, Schulten K. Cytosine methylation alters DNA mechanical properties. Nucleic Acids Res. 2011; 39: 8740-51
23. Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. Science 2001; 293: 1089-93
24. Guibert S, Weber M. Functions of DNA methylation and hydroxymethylation in Mammalian development. Curr.Top.Dev.Biol. 2013; 104: 47-83
25. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, et al. A unique regulatory phase of DNA methylation in the early

mammalian embryo. Nature 2012; 484: 339-44

26. Lee HJ, Hinshelwood RA, Bouras T, Gallego-Ortega D, Valdes-Mora F, Blazek K, et al. Lineage specific methylation of the Elf5 promoter in mammary epithelial cells. Stem Cells 2011; 29: 1611-9

27. Yeo S, Jeong S, Kim J, Han JS, Han YM, Kang YK. Characterization of DNA methylation change in stem cell marker genes during differentiation of human embryonic stem cells. Biochem.Biophys.Res.Commun. 2007; 359: 536-42

28. Reik W, Walter J. Genomic imprinting: parental influence on the genome. Nat.Rev. Genet. 2001; 2: 21-32

29. Lienert F, Wirbelauer C, Som I, Dean A, Mohn F, Schubeler D. Identification of genetic elements that autonomously determine DNA methylation states. Nat.Genet. 2011; 43: 1091-7

30. Gertz J, Varley KE, Reddy TE, Bowling KM, Pauli F, Parker SL, et al. Analysis of DNA methylation in a three-generation family reveals widespread genetic influence on epigenetic regulation. PLoS.Genet. 2011; 7: e1002228

31. Sinclair KD, Allegrucci C, Singh R, Gardner DS, Sebastian S, Bispham J, et al. DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and methionine status. Proc.Natl.Acad.Sci.U.S.A 2007; 104: 19351-6

32. Barker DJ. The origins of the developmental origins theory. J.Intern.Med. 2007; 261: 412-7

33. Adams M, Andersen AM, Andersen PK, Haig D, Henriksen TB, Hertz-Picciotto I, et al. Sostrup statement on low birthweight. Int.J.Epidemiol. 2003; 32: 884-5

34. Joseph KS, Kramer MS. Review of the evidence on fetal and early childhood antecedents of adult chronic disease. Epidemiol.Rev. 1996; 18: 158-74

35. Paneth N, Susser M. Early origin of coronary heart disease (the “Barker hypothesis”). BMJ 1995; 310: 411-2

36. Gilbert JS, Nijland MJ. Sex differences in the developmental origins of hypertension and cardiorenal disease. Am.J.Physiol Regul. Integr.Comp Physiol 2008; 295: R1941-R1952

37. Nathanielsz PW. Animal models that elucidate basic principles of the developmental origins of adult diseases. ILAR.J. 2006; 47: 73-82

38. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol.Cell Biol. 2003; 23: 5293-300

39. 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: 401-6

40. Burdge GC, Lillycrop KA, Phillips ES, Slater-Jefferies JL, Jackson AA, Hanson MA. Folic acid supplementation during the juvenile-pubertal period in rats modifies the phenotype and epigenotype induced by prenatal nutrition. J.Nutr. 2009; 139: 1054-60

41. Lillycrop KA, Slater-Jefferies JL, Hanson MA, Godfrey KM, Jackson AA, Burdge GC. Induction of altered epigenetic regulation of the hepatic glucocorticoid receptor in the offspring of rats fed a protein-restricted diet during pregnancy suggests that reduced DNA methyltransferase-1 expression is involved in impaired DNA methylation and changes in histone modifications. Br.J.Nutr. 2007; 97: 1064-73

42. Lillycrop KA, Phillips ES, Jackson AA, Hanson MA, Burdge GC. Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring. J.Nutr. 2005; 135: 1382-6

43. Weaver IC, Cervoni N, Champagne FA, D’Alessio AC, Sharma S, Seckl JR, et al. Epigenetic programming by maternal behavior. Nat.Neurosci. 2004; 7: 847-54

44. Bogdarina I, Welham S, King PJ, Burns SP, Clark AJ. Epigenetic modification of the renin-angiotensin system in the fetal programming of hypertension. Circ.Res. 2007; 100: 520-6

45. Johansson S, Iliadou A, Bergvall N, de FU, Kramer MS, Pawitan Y, et al. The association between low birth weight and type 2 diabetes: contribution of genetic factors. Epidemiology 2008; 19: 659-65

46. Bergvall N, Lindam A, Pawitan Y, Lichtenstein P, Cnattingius S, Iliadou A. Importance of

familial factors in associations between offspring birth weight and parental risk of type-2 diabetes. Int.J.Epidemiol. 2008; 37: 185-92

47. Oberg S, Cnattingius S, Sandin S, Lichtenstein P, Iliadou AN. Birth weight predicts risk of cardiovascular disease within dizygotic but not monozygotic twin pairs: a large population-based co-twin-control study. Circulation 2011; 123: 2792-8

48. Fleten C, Nystad W, Stigum H, Skjaerven R, Lawlor DA, Davey SG, et al. Parent-Offspring Body Mass Index Associations in the Norwegian Mother and Child Cohort Study: A Family-based Approach to Studying the Role of the Intrauterine Environment in Childhood Adiposity. Am.J.Epidemiol. 2012; 176: 83-92

49. Stein AD, Kahn HS, Rundle A, Zybert PA, van der Pal-de Bruin, Lumey LH. Anthropometric measures in middle age after exposure to famine during gestation: evidence from the Dutch famine. Am.J.Clin.Nutr. 2007; 85: 869-76

50. Ravelli AC, van der Meulen JH, Osmond C, Barker DJ, Bleker OP. Obesity at the age of 50 y in men and women exposed to famine prenatally. Am.J.Clin.Nutr. 1999; 70: 811-6

51. Ravelli GP, Stein ZA, Susser MW. Obesity in young men after famine exposure in utero and early infancy. N.Engl.J.Med. 1976; 295: 349-53

52. Wang Y, Wang X, Kong Y, Zhang JH, Zeng Q. The Great Chinese Famine leads to shorter and overweight females in Chongqing Chinese population after 50 years. Obesity. (Silver.Spring) 2010; 18: 588-92

53. de Rooij SR, Painter RC, Roseboom TJ, Phillips DI, Osmond C, Barker DJ, et al. Glucose tolerance at age 58 and the decline of glucose tolerance in comparison with age 50 in people prenatally exposed to the Dutch famine. Diabetologia 2006; 49: 637-43

54. Lumey LH, Stein AD, Kahn HS. Food restriction during gestation and impaired fasting glucose or glucose tolerance and type 2 diabetes mellitus in adulthood: evidence from the Dutch Hunger Winter Families Study, 2009: 1; S164

55. Vaiserman AM, Khalangot ND, Pisaruk AV, Mekhova LV, Kolyada AK, Kutsenko KYu, et al. Predisposition to Type II Diabetes among those Residents of Ukraine Whose Prenatal Development Coincided with the Famine of 1932-1933, 2011: 362-6

56. Xu MQ, Sun WS, Liu BX, Feng GY, Yu L, Yang L, et al. Prenatal Malnutrition and Adult Schizophrenia: Further Evidence From the 1959-1961 Chinese Famine. Schizophr.Bull. 2009; 35: 568-76

57. Stein AD, Zybert PA, van de BM, Lumey LH. Intrauterine famine exposure and body proportions at birth: the Dutch Hunger Winter. Int.J.Epidemiol. 2004; 33: 831-6

58. McLean M, Chipps D, Cheung NW. Mother to child transmission of diabetes mellitus: does gestational diabetes program Type 2 diabetes in the next generation? Diabet.Med. 2006; 23: 1213-5

59. Lawlor DA, Macdonald-Wallis C, Fraser A, Nelson SM, Hingorani A, Davey SG, et al. Cardiovascular biomarkers and vascular function during childhood in the offspring of mothers with hypertensive disorders of pregnancy: findings from the Avon Longitudinal Study of Parents and Children. Eur.Heart J. 2012; 33: 335-45

60. Kral JG, Biron S, Simard S, Hould FS, Lebel S, Marceau S, et al. Large maternal weight loss from obesity surgery prevents transmission of obesity to children who were followed for 2 to 18 years. Pediatrics 2006; 118: e1644-e1649

61. Waterland RA, Lin JR, Smith CA, Jirtle RL. Post-weaning diet affects genomic imprinting at the insulin-like growth factor 2 (Igf2) locus. Hum.Mol.Genet. 2006; 15: 705-16

62. Gluckman PD, Lillycrop KA, Vickers MH, Pleasants AB, Phillips ES, Beedle AS, et al. Metabolic plasticity during mammalian development is directionally dependent on early nutritional status. Proc.Natl.Acad. Sci.U.S.A 2007; 104: 12796-800

63. Bogdarina I, Haase A, Langley-Evans S, Clark AJ. Glucocorticoid effects on the programming of AT1b angiotensin receptor gene methylation and expression in the rat. PLoS.One. 2010; 5: e9237

64. Downing C, Johnson TE, Larson C, Leakey TI, Siegfried RN, Rafferty TM, et al. Subtle decreases in DNA methylation and gene expression at the mouse Igf2 locus following prenatal alcohol exposure: effects of a methyl-supplemented diet. Alcohol 2011; 45: 65-71

65. Talens RP, Christensen K, Putter H, Willemsen G, Christiansen L, Kremer D, et al. Epigenetic variation during the adult lifespan: cross-sectional and longitudinal data on monozygotic twin pairs. Aging Cell 2012; 11: 694-703

66. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. Proc.Natl.Acad.Sci.U.S.A 2005; 102: 10604-9

67. Heijmans BT, Mill J. Commentary: The seven plagues of epigenetic epidemiology. Int.J.Epidemiol. 2012; 41: 74-8

68. Burger GCE, Drummond JC, Sandstead HR. Malnutrition and Starvation in Western Netherlands, September 1944-July 1945. The Hague: General State Printing Office, 1948

69. Lumey LH, Stein AD, Kahn HS, van der Pal-de Bruin KM, Blauw GJ, Zybert PA, et al. Cohort profile: the Dutch Hunger Winter families study. Int.J.Epidemiol. 2007; 36: 1196-204

70. Morgan HD, Sutherland HG, Martin DI, Whitelaw E. Epigenetic inheritance at the agouti locus in the mouse. Nat.Genet. 1999; 23: 314-8

71. Walther FJ, den Ouden AL, Verloove-Vanhorick SP. Looking back in time: outcome of a national cohort of very preterm infants born in The Netherlands in 1983. Early Hum. Dev. 2000; 59: 175-91

72. Waddington CH. An Introduction to Modern Genetics. New York: MacMillan, 1939,

73. Waddington C.H. The epigenotype, 1942: 18-20

74. Jablonka E, Lamm E. Commentary: The epigenotype--a dynamic network view of development. Int.J.Epidemiol. 2012; 41: 16-20

75. Nanney DL. Epigenetic Control Systems. Proc.Natl.Acad.Sci.U.S.A 1958; 44: 712-7

76. Harris M. Cell Culture and Somatic Variation. New York: Holt, Rinehart & Winston, 1964

77. Ephrussi B. The cytoplasm and somatic cell variation. J.Cell Physiol Suppl 1958; 52: 35-49

78. Abercrombie M. A general review of the nature of differentiation, In: deReuck AVS, Knight J, eds. Cell differentation. Boston: Little, Brown and company, 1967: 3-17

79. Kucharski R, Maleszka J, Foret S, Maleszka R. Nutritional control of reproductive status in honeybees via DNA methylation. Science 2008; 319: 1827-30

80. Gilbert SF, Epel D. Ecological Developmental Biology: integrating Epigenetics, Medicine, and Evolution. Sunderland: Sinauer Associates Inc., 2008

81. Jablonka E, Lamb MJ. Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, And Symbolic Variation In The History Of Life. Cambridge: MIT press Ltd, 2006

82. Pujadas E, Feinberg AP. Regulated noise in the epigenetic landscape of development and disease. Cell 2012; 148: 1123-31

Variation, patterns and temporal stability of DNA methylation: considerations for epigenetic epidemiology

Rudolf P. Talens[1], Dorret I. Boomsma[2], **Elmar W. Tobi**[1], Dennis Kremer[1], J. Wouter Jukema[3], Gonneke Willemsen[2], Hein Putter[4], P. Eline Slagboom[1,5], Bastiaan T. Heijmans[1,5]

1. Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
2. Biological Psychology, VU University Amsterdam, Amsterdam, The Netherlands
3. Cardiology, Leiden University Medical Center, Leiden, The Netherlands
4. Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands
5. The Netherlands Consortium for Healthy Ageing, Leiden, The Netherlands

FASEB J. **2010** Sep:24(9):3135-3144

Abstract

The prospect of finding epigenetic risk factors for complex diseases would be greatly enhanced if DNA from existing biobanks, which is generally extracted from whole blood, could be used to perform epigenetic association studies. We characterized features of DNA methylation at 16 candidate loci, 8 of which were imprinted, in DNA samples from the Netherlands Twin Register biobank. Except for unmethylated or fully methylated sites, CpG methylation varied considerably in a sample of 30 unrelated individuals. This variation remained after accounting for the cellular heterogeneity of blood. Methylation of CpG-sites was correlated within loci and across chromosomes for 4 imprinted loci. In 34 additional individuals, we investigated the DNA methylation of 8 representative loci in 2 longitudinal blood and 2 longitudinal buccal cell samples (follow-up 11-20 and 2-8 years, respectively). 5 of 8 loci were stable over time (ρ>0.75) in both tissues indicating that prospective epigenetic studies may be possible. For 4 loci, the DNA methylation in blood (mesoderm) correlated with that in the buccal cells (ectoderm) (ρ>0.75). Our data suggest that epigenetic studies on complex diseases may be feasible for a proportion of genomic loci provided that they are carefully designed.

Introduction

Epigenetics refers to heritable differences in gene expression potential that are not caused by variation in the DNA sequence[1,2]. Its molecular basis is the chemical modification of either the DNA itself (cytosine methylation in CpG dinculeotides) or the histones that package the chromatin (e.g. methylation, acetylation, phosphorylation)[3-5]. It has frequently been proposed that changes in these epigenetic marks significantly contribute to the risk of complex diseases including cancer, cardiovascular and metabolic disease[1,6-10]. However, with the exception of studies on cancer, empirical data from epidemiological studies supporting these hypotheses are largely absent mainly due to technical and methodological limitations.

Many of the technical limitations have been resolved, in particular with respect to the high throughput measurement of DNA methylation[11,12]. DNA methylation is correlated with other layers of epigenetic marks, particularly histone modifications[13]. DNA methylation may be the most suitable epigenetic mark for large-scale epidemiological studies, since methyl groups are covalently bound to CpG dinucleotides and are not lost during routine DNA extraction, unlike histone modifications. This opens the possibility of exploiting existing DNA biobanks for research purposes, to discover epigenetic risk factors for complex disease.

Epigenetic studies will require the development of data resources analogous to those that facilitated genetic association studies. The resources should include epigenome maps charting DNA methylation marks[14], the description of inter-individual variation in DNA methylation (cf. single nucleotide polymorphisms and copy number variants)[15] and data on the patterns within this variation (cf. linkage disequilibrium)[16]. To guide the development of such epigenome-wide resources, candidate loci may be studied. In this respect differentially methylated regions influencing imprinting[17], transposon-derived sequences[17], CpG island shores[18] and recognition sequences for methylation-dependent transcription factors[19] are of particular interest.

In addition, several issues potentially limiting the use of existing biobanks for epigenetic epidemiology need to be addressed. Firstly, DNA in biobanks

is mostly extracted from whole blood, which, like any tissue, consists of different cell types that may carry different epigenetic marks and whose relative numbers may vary between individuals[20]. Secondly, the stability of DNA methylation over time should be known before the association of DNA methylation with future disease risk can be assessed. Global (or average) DNA methylation has been reported to change over time[21,22], but DNA methylation of specific loci may be more stable[16,23]. Lastly, it will be crucial to address to what extent DNA methylation measured in blood marks in less accessible tissues that are directly involved in disease. Despite scattered reports that this may be the case[24-26], the issue remains largely unresolved.
We assessed whether genomic DNA stored in existing biobanks would be suitable for epigenetic epidemiological studies. To this end we addressed the inter-individual variation in DNA methylation of 16 candidate loci for cardiovascular and metabolic disease, the influence of blood cell heterogeneity on this variation, the stability of DNA methylation over time and its correlation between whole blood (mesoderm) and buccal cells (ectoderm) in individuals from the Netherlands Twin Register (NTR)[27,28].

Results

We quantitatively measured the DNA methylation of 16 candidate loci using DNA samples from the NTR biobank, to estimate its variation in the general population. The DNA was extracted from whole blood. The DNA samples were from 30 unrelated individuals, who were selected to represent the broad range in age and metabolic parameters of the whole biobank. After removing CpGs for which local genetic polymorphisms could interfere with DNA methylation measurements, and CpGs that did not meet the quality criteria, we recorded methylation of 164 CpG sites, distributed over 104 CpG units, of which 62 contained a single CpG site (**Table 1**).

Table 1. Characteristics of methylation assays

Locus	Chromosome	Megabase	Gene function	CpG sites[2]	Single CpG sites[3]	Imprinted	CpG-island	Promoter	Intragenic	Intergenic	Putative mTFBS[4]	Confirmed mTFBS[4]	Transposon	Literature[5]
IL10[1]	01q32.1	205.01	Anti-inflammation	4	2			+			3	6	+	+[43]
NR3C1	05q31.3	142.67	Stress response	20	4		+	+			3	2		+[19]
TNF	06p21.33	31.65	Pro-inflammatory	8	5			+			2			+[45]
IGF2R[1]	06q25.3	160.35	Growth/Apoptosis	10	0	?	+	+	+		1			+[46]
GRB10	07p12.2	50.82	IIS inhibitor	16	5	+	+	+			4	2		+[47]
LEP[1]	07q32.1	127.67	Metabolism	10	5		+	+				3		+[48]
CRH[1]	08q13.1	67.25	Stress response	5	5			+			1	1		+[49]
ABCA1	09q31.1	106.73	Cholesterol transport	16	3		+	+			4	4		+/-[50]
IGF2[1]	11p15.5	2.13	Early growth	5	3	+		+	+					+[24]
INSIGF[1]	11p15.5	2.14	(Embryonic) Growth	4	4	+/-		+			2			+[51]
KCNQ1OT1[1]	11p15.5	2.68	Imprinting control region	14	7	+	+	+						+[52]
MEG3	14q32.2	100.36	Growth suppressor	7	3	+	+	+			1	1		+[53]
FTO	16q12.2	52.38	Development	10	4				+				+	
APOC1[1]	19q13.32	50.11	Metabolism	6	6					+	3			+/-[54]
GNASAS	20q13.32	56.86	Growth/Lypolytic signal	17	3			+				1		+[55]
GNAS A/B	20q13.32	56.90	Growth/Lypolytic signal	12	3	+	+	+						+[44]

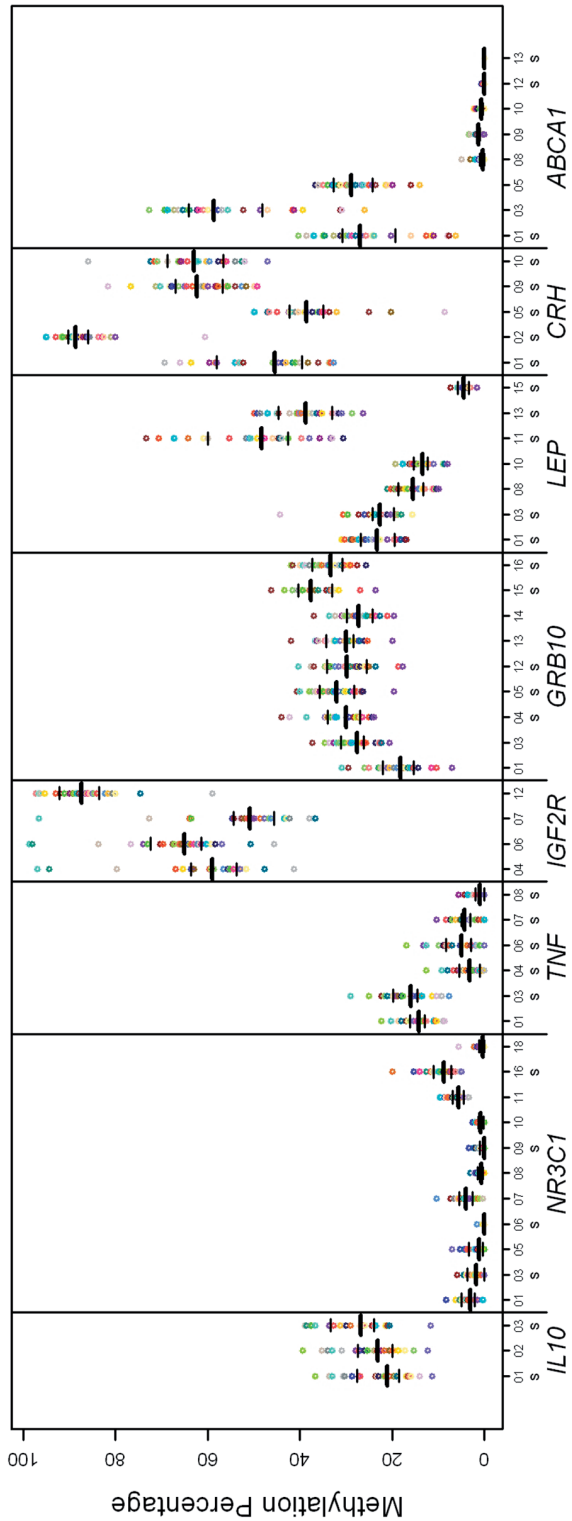
1 With these assays the stability across time and the correlation between tissues was also investigated

2 CpG sites that met the quality criteria described in the methods section

3 CpG sites of which the methylation proportion was measured individually

4 Methylation sensitive transcription factor-binding sites

5 CpG methylation previously reported to associate with gene expression


Methylation Percentage
0
20
40
60
80
100
IL10
NR3C1
TNF
IGF2R
GRB10
LEP
CRH
ABCA1

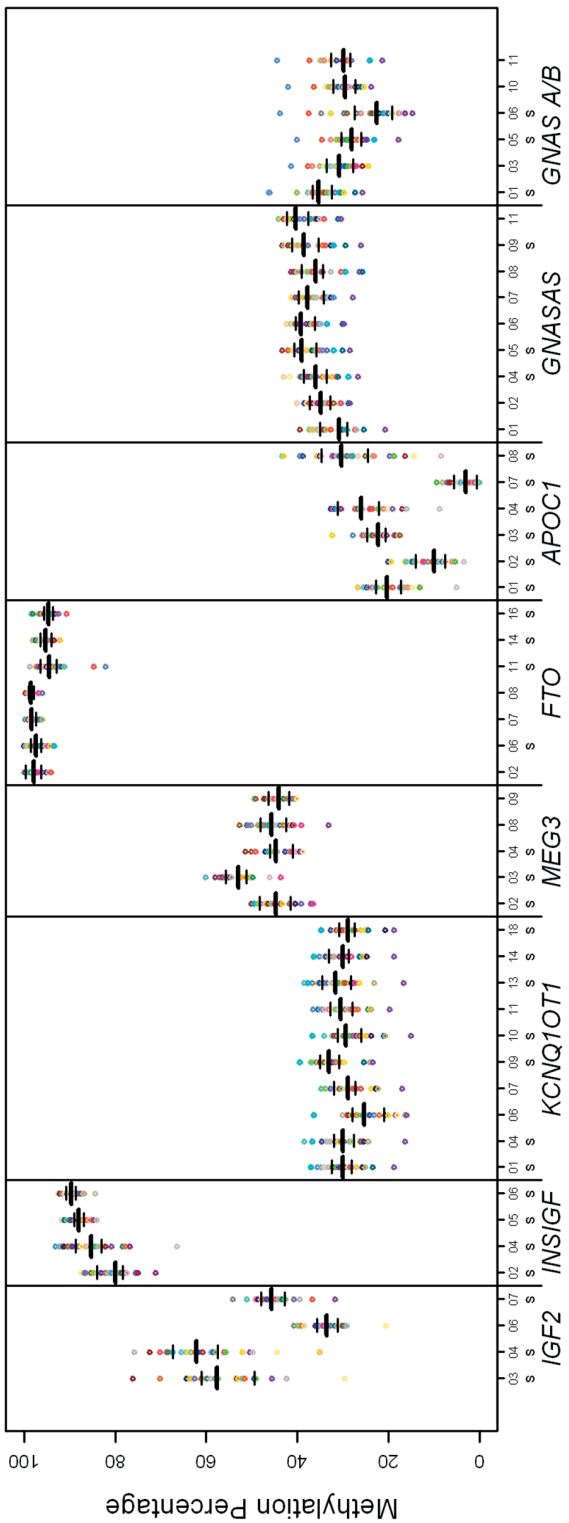


Figure 1. Inter-individual variation in DNA methylation

Methylation percentage, y-axis, at every CpG unit, x-axis, for each of the 30 individuals, coloured dots. The order of the loci is based on their chromosomal location, starting at the lowest designation. CpG unit numbers are counted from the Forward primer onwards. The name of each locus is given below the x-axis. The bold horizontal bar gives the median, the thin horizontal bars show the inter-quartile range for each CpG unit. The vertical lines across the plot separate the loci. The corresponding CpG sites of each CpG unit are given in table S5B. Individually measured CpG sites are marked with an "S" below the unit number.

Inter-individual variation in DNA methylation

The average methylation of the loci studied ranged from 0 % to 98 % (**Figure 1**). Within the majority of loci, the average methylation of CpG sites was similar. Exceptions were *LEP* with a 34% methylation difference between CpG sites located 18 bases apart and *ABCA1* with a 31% difference between sites 26 bases apart. Previous studies reported methylation differences between men and women[29,30]. We could not detect such differences in an explorative analysis. This may be related to the study size and the number of tests performed.

A considerable inter-individual variation in CpG methylation was observed. The variation approximated a normal distribution except for CpG units that showed no or very little variation (average DNA methylation close to 0 % or 100 %). The extent of this variation varied per CpG unit (0% <SD<15%). To exclude the possibility that this difference in variation might merely be due to the fact that DNA methylation is truncated at 0 % and 100 %, a variance stabilizing transformation was applied[31]. The variation remained significantly different between CpG units ($P=10^{-16}$). To validate these findings, we measured the DNA methylation of 8 out of 16 loci (55 CpG sites, 31 CpG units with 1 CpG site) in 34 additional individuals from the NTR biobank, and this yielded similar results (*data not shown*).

Cellular heterogeneity

DNA methylation was measured on genomic DNA extracted from whole blood. As whole blood consists of different cell types, which may display differences in DNA methylation, we tested to what extent the inter-individual variation observed was influenced by cellular heterogeneity as assessed by counting the major cell subclasses. The largest cellular fraction consisted of neutrophils (average proportion 53.7 %; SD=8.8 %) thus contributing the majority of DNA molecules on which CpG methylation was analyzed. For 10 out of 16 loci, the variation in DNA methylation was not associated with this measure of cellular heterogeneity (**Table 2**). For the remaining loci, the

variation in DNA methylation that could be explained by variation in cellular heterogeneity was generally small and associations were of borderline significance. *IL10*, which is highly expressed in leukocytes, was a notable exception: 50% of its variation in DNA methylation could be attributed to the neutrophil cell count. The analysis was repeated using the lymphocyte percentage (average proportion 34.4 %; SD=8.4 %), which was highly correlated with the neutrophil percentage (r = -0.95) and similar results were observed (*data not shown*). No influence of monocyte percentage was observed on DNA methylation (average 8.0 %; SD=1.8 %; correlation with the neutrophil proportion: r = -0.48). The influence of cellular heterogeneity was not affected by gender (*data not shown*).

Table 2. Association of neutrophil proportion with DNA methylation

Locus	Variance explained	p-value of effect
IL10	50.1 %	$3.9*10^{-06}$
NR3C1	0.2 %	0.555
TNF	8.0 %	0.037
IGF2R	5.0 %	0.208
GRB10	0.7 %	0.625
LEP	7.4 %	0.019
CRH	4.0 %	0.022
ABCA1	7.3 %	0.021
IGF2	3.5 %	0.185
INSIGF	0.3 %	0.674
KCNQ1OT1	0.4 %	0.733
MEG3	4.1 %	0.165
FTO	0.1 %	0.714
APOC1	6.2 %	0.026
GNASAS	0.8 %	0.590
GNAS A/B	0.0 %	0.889

To validate these findings, we performed the same test on the 8 loci in an additional 34 individuals (**Supplement I, Table S1**). The loci previously not showing an association were again not associated with the neutrophil percentage (*IGF2R*, *IGF2*, *INSGF* and *KCNQ1OT1*). Two of the associations of borderline significance were not found (*APOC1* and *CRH*), but the modest

association of *LEP* with the neutrophil percentage was replicated ($P=1.0x10^{-4}$). Again, a substantial proportion of the variation in *IL10* methylation could be attributed to the neutrophil percentage (27.9%, $P=8.0x10^{-8}$).

Correlations and patterns of CpG methylation

To investigate patterns of DNA methylation further within and across loci, correlations between the methylation of CpG sites were computed and visualized using a heat map after unsupervised clustering (**Figure 2**). CpG methylation was particularly correlated within loci ($r_{max} = 0.95$) but also across loci ($r_{max} = 0.68$). The cluster of loci correlating irrespectively of chromosomal location included paternally imprinted loci (*MEG3* and *GNASAS*) and maternally imprinted loci (*GRB10, KCNQ1OT1* and *GNAS A/B*). These observations were unaffected by variance-stabilizing transformation or adjustment for cell heterogeneity prior to analysis (*data not shown*). When the correlation was studied again, but now in 60 controls of the Dutch Hunger Winter Families Study that we measured previously[29,32] similar patterns were found (*data not shown*). The correlations were similar for both sexes in both study populations (*data not shown*).

Stability over time

To study the stability of DNA methylation over time, we selected 34 additional individuals from the NTR for whom two blood samples were taken 11-20 years apart. The methylation of 8 loci that were representative of the set of 16 loci was measured (**Table 1**). Overall, DNA methylation was similar at the two time points (**Figure 3A**) and only minor differences were observed (**Table 3**). Similar average methylation levels between the time-points do not indicate stability per se since methylation may increase in some and decrease in other individuals over time. The variation around the average difference was greatest for *IL10*, which also showed the greatest average difference (-2.8% (SD = 9.1); Table 3A). It was lowest for *KCNQ1OT1* (SD = 2.8), indicating relative stability over time. An alternative way to express

stability, which takes into account the differences in inter-individual variation of the loci, is to compute correlation coefficients (**Table 3**). For 5 of the 8 loci, the correlation $\rho > 0.75$ indicated substantial stability between the time points. These loci included *IGF2R* ($\rho = 0.88$) and *APOC1* ($\rho = 0.96$). Note that the correlation was low for *KCNQ1OT1* ($\rho = 0.31$), which can be attributed to the very low level of inter-individual variation. Temporal stability was similar

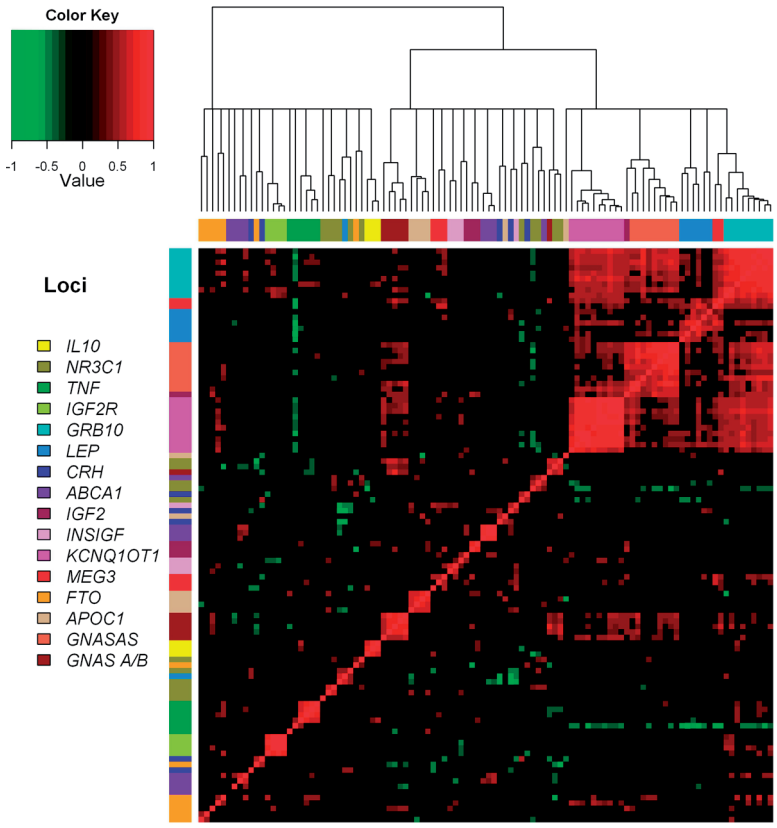


Figure 2. Correlation between CpG sites within and across loci

Heat map depicting correlations between the methylation levels of all CpG units of the 16 loci measured in the first group of 30 individuals. For reference the CpG units are annotated by a colour, based on the locus, in the left and upper margin. The diagonal axis running from the lower left to the upper right corner is the line of symmetry where each CpG unit hypothetically correlates with itself. Full correlation (1) is plotted as the brightest red shade, full inverse correlation (-1) is plotted as the brightest green shade, no correlation (0) is plotted as black. Non-significant correlations are depicted as no correlation. The complete clustering is based on the Euclidean distance.

in both sexes (*data not shown*). From the same 34 individuals, DNA samples from buccal swabs taken 2-8 years apart were available and showed similar results (**Table 4**).

To exclude the possibility that the higher correlations observed were due to sequence variation not present in dbSNP, we used the mass spectra to identify CpG methylation measurements that were suspected to have been influenced by sequence variation[33]. This was the case for 1 or more individuals for 7/41 CpG units. Removal of these CpG measurements did not affect the correlations $\rho > 0.75$ (*IGF2R*, $\rho = 0.87$; *LEP*, $\rho = 0.90$; *IGF2*, $\rho = 0.92$; *CRH*, $\rho = 0.94$; and *APOC1*, $\rho = 0.95$).

Table 3. Comparison of DNA methylation in blood samples of the two time points

	Methylation mean (SD)		Difference	Spearman's
Locus	Old blood	New Blood	mean (SD)	rho (ρ)
IL10	22.4% (9.0)	25.2% (6.6)	- 2.8% (9.1)	0.422
IGF2R	65.8% (16.8)	67.6% (17.7)	- 1.8% (8.1)	0.883
LEP	20.0% (11.5)	21.8% (13.0)	- 1.8% (6.1)	0.895
CRH	63.4% (22.1)	63.8% (21.1)	- 0.4% (6.7)	0.942
IGF2	49.4% (11.8)	49.0% (11.3)	0.4% (4.7)	0.924
INSIGF	86.4% (4.1)	85.5% (4.3)	0.9% (3.7)	0.649
KCNQ1OT1	30.8% (2.2)	31.8% (2.6)	- 1.0% (2.8)	0.307
APOC1	19.7% (11.4)	19.6% (11.7)	0.1% (3.4)	0.956

Table 4. Comparison of DNA methylation in buccal cell samples from 34 individuals at two time points

	Methylation mean (SD)		Difference	Spearman's rho
Locus	Old Buccal cells	New Buccal cells	mean (SD)	(ρ)
IL10	66.9% (13.6)	65.8% (18.4)	1.1% (17.0)	0.613
IGF2R	81.0% (12.9)	82.9% (11.7)	- 1.9% (9.0)	0.701
LEP	10.7% (7.4)	11.0% (7.8)	- 0.3% (7.3)	0.608
CRH	62.7% (21.0)	62.3% (21.5)	0.4% (9.2)	0.888
IGF2	32.5% (8.8)	32.3% (11.0)	0.2% (9.0)	0.686
INSIGF	84.6% (4.4)	84.4% (4.7)	0.2% (3.9)	0.543
KCNQ1OT1	35.8% (3.0)	34.9% (3.0)	0.9% (4.0)	0.097
APOC1	9.4% (7.2)	10.8% (9.7)	- 1.4% (6.1)	0.836

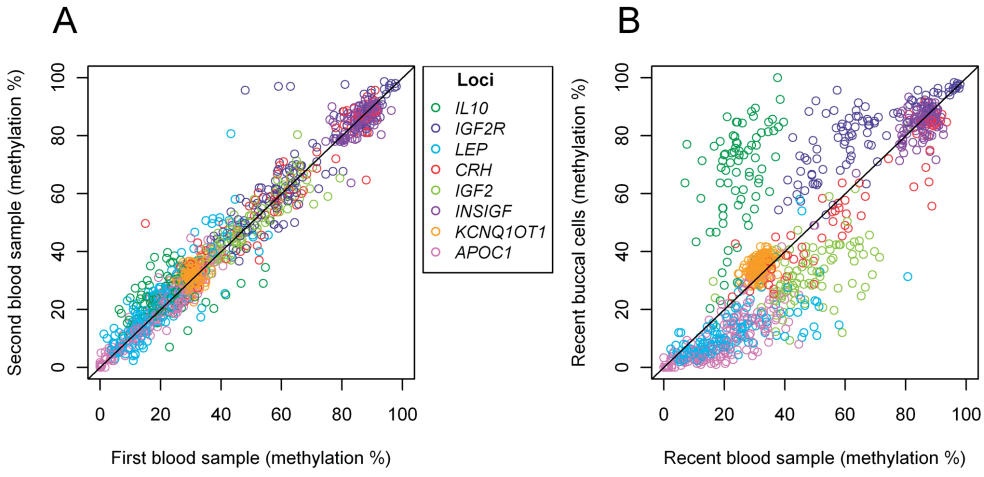


Figure 3. Temporal stability and comparison between blood and buccal cell DNA methylation

Scatter plots for individual comparison of CpG methylation between the DNA samples. CpG units of each individual are annotated by colouring based on the locus. The diagonal x = y line is plotted in black for reference. **A.** CpG methylation in the first blood sample (x-axis) is plotted against methylation in the second, more recent, blood sample (y-axis). Each dot represents one CpG unit of one individual in both DNA samples. **B.** CpG methylation in the recent blood DNA sample (x-axis) is plotted against methylation in the recent buccal swab DNA sample (y-axis). Each dot represents one CpG unit of one individual in both DNA samples.

Correlation of CpG methylation between tissues

To test whether DNA methylation in blood could mark that in other tissues, we studied DNA methylation of the 8 loci in the recent blood (mesoderm) and buccal swab (ectoderm) samples of the individuals in whom stability over time was tested. The average level of DNA methylation was generally different between the two tissues and the extent of the difference depended on the locus (**Figure 3B, Table 5**). The variation around the average difference also varied per locus. Again, *IL-10* showed the highest variation (SD = 17.1) and *KCNQ1OT1* the lowest (SD = 3.0). For all loci, the SD of the average difference between the tissues was higher than that of the average difference between the time points. For 4 of the 8 loci, the correlation of DNA methylation between the two tissues was greater than 0.75. These loci

included *IGF2R* (ρ = 0.83) and *APOC1* (ρ = 0.82). Again, the correlations were similar in both sexes (*data not shown*). After removing CpGs measurements suspected to be influenced by sequence variation not present in dbSNP, similar correlations were found (*IGF2R*, ρ = 0.82; *LEP*, ρ = 0.80; *CRH*, ρ = 0.90; and *APOC1*, ρ = 0.81).

Table 5. Comparison of DNA methylation in buccal cell samples from 34 individuals at two time points

Locus	Methylation, mean (SD)		Difference, mean (SD)	Spearman's rho (ρ)
	Blood	Buccal cells		
IL10	24.8% (7.1)	64.9% (18.8)	- 40.1% (17.1)	0.442
IGF2R	68.3% (17.4)	81.6% (12.4)	- 13.3% (10.6)	0.827
LEP	21.7% (13.7)	11.8% (8.8)	9.9% (9.9)	0.798
CRH	63.8% (21.9)	58.4% (21.4)	5.4% (7.9)	0.905
IGF2	48.8% (11.5)	32.4% (11.0)	16.4% (10.8)	0.557
INSIGF	85.4% (4.3)	84.4% (4.7)	1.0% (4.6)	0.371
KCNQ1OT1	32.0% (2.6)	34.7% (3.1)	- 2.7% (3.0)	0.481
APOC1	19.9% (11.7)	10.7% (9.6)	9.2% (7.6)	0.822

Discussion

Epigenetic risk factors are thought to contribute to the development of common diseases such as cardiovascular and metabolic disease[6,9,10]. Here we investigated whether genomic DNA from existing biobanks is suitable for the identification of these risk factors in epidemiological studies[1,7]. Using genomic DNA from the Netherlands Twin Register biobank[27,28], we first assessed the inter-individual variation in DNA methylation for 16 candidate loci, since the human epigenome map is still in development[14] and epigenome-wide resources on variation (i.e. the epivariome) are lacking.

We observed considerable variation in CpG methylation between individuals, except for loci that are either not methylated or fully methylated. The extent of this variation varied between CpG sites. Earlier reports frequently characterized CpG methylation as hypo-, iso- or hyper-methylation[24,34]. Our

data support our own previous work[16] and that of others[15] in which DNA methylation was more accurately described as a quantitative trait.

Cellular heterogeneity

Secondly, we addressed the possibility that the variation in DNA methylation could simply be attributed to cellular heterogeneity in leukocytes between individuals[20]. Blood, like any tissue, consists of a mixture of different cell types that all may have a cell-specific epigenome[3]. Our results show that for the large majority of candidate loci, inter-individual differences in the cellular composition of the blood sample did not contribute to the variation observed in DNA methylation or explained only a minor proportion of this variation. One notable exception was the *IL10* locus for which cellular heterogeneity explained up to half of the total variation in DNA methylation.
If cell counts are available for whole blood samples stored in a biobank, the potentially confounding influence of cellular heterogeneity can be monitored using standard statistical methods. If no data on cellular heterogeneity are available, it may be necessary to exclude the association of cellular heterogeneity either with the outcome of interest or with methylation of the locus studied. The latter can be addressed for example in a (sub)study for which data on leukocyte populations are available. Our study suggests that no such relationship will be observed for many loci, in which case biobanks without data on cellular heterogeneity may still be useful.

DNA methylation is correlated within and across loci

Thirdly, we investigated patterns in CpG methylation within and across loci. We found that within the locus CpG methylation is highly correlated, except for candidate loci that were not methylated or fully methylated, which corroborates recent findings[16,35]. This observation suggests that assessing the methylation of a subset of CpGs is sufficient to cover the variation in DNA methylation at a locus. This is analogous to genetic association studies in which a small number of tagging SNPs can cover all genetic variation at a

locus due to linkage disequilibrium[36]. Moreover, our results provide the first indication that methylation of CpG sites can also be correlated irrespectively of their chromosomal location. This was observed for a subset of mainly imprinted loci, which may be related to the mechanisms responsible for establishing methylation marks at DMRs[37,38].

Stability over time

Fourthly, since DNA methylation is a reversible process[39], it may not be stable over time. If so, this would preclude conclusions about causality in epidemiological studies, since DNA methylation may change during a follow-up period or the differences in DNA methylation may be the consequence of disease[40]. The majority of loci tested were stable over time in DNA from blood and buccal cells despite possible changes in cellular composition during the follow-up period. The fact that we investigated DNA samples that were taken 11 to 20 years apart implies that these DNA methylation marks may be investigated in most prospective cohort studies in which participants are followed for the development of disease for similar or shorter follow-up periods.

However, for a minority of loci we found that although, on average, there was no difference in DNA methylation between the time points, the correlation was lower, indicating relaxed maintenance of these DNA methylation marks. This data resembles recent results on global DNA methylation studying similar follow-up periods[22]. The age of the individuals in our study was limited to young and middle ages (14 to 62 years old). Therefore, we cannot exclude instability over very long periods of time, nor can we exclude the occurrence of greater changes in old age. Indeed, instability of the DNA methylation marks in old age has been reported for both locus-specific[23,32] and global[41] DNA methylation. Moreover, our study did not address the possible occurrence of changes in DNA methylation as a consequence of disease or processes preceding its clinical manifestation[40].

Existing biobanks

DNA from existing biobanks generally is extracted from easily accessible tissues such as blood. Future studies may reveal DNA methylation patterns in such tissues that mark the risk of disease. As a first step towards establishing a possible causal role, it will be necessary to determine that DNA methylation measured in peripheral tissues is associated with that in tissues directly involved in the disease of interest. Although DNA methylation is thought to be a mechanism driving cell differentiation leading to tissue-specific differentially methylated regions[18], initial reports indicated that DNA methylation measured in blood may be informative. For example, *IGF2* and *ER-α* methylation in blood marked that of colon tissue[24,25]. Also, an autopsy study of 6 subjects and 11 tissues, which did not include blood, suggested that the hypo- and hyper-methylation status of loci is commonly preserved across tissues. Comparing the methylation of candidate loci in blood and buccal cells, we found that for half of the loci tested, DNA methylation measured in blood was a marker for that in buccal cells. These results are promising since blood and buccal cells stem from different germ layers (mesoderm and ectoderm, respectively) and warrant the investigation of correlations with other tissues involved in disease. Genome-scale studies in particular will be informative for defining the (sequence) characteristics of loci showing correlations across tissues. Such studies will be required to interpret the results of epidemiological studies on DNA methylation in blood in a meaningful way.

New biobanking efforts

Our study on the suitability of DNA from existing biobanks for epigenetic studies provides leads for setting up new biobanks specifically aimed at epigenetic epidemiology. Since the correlation between DNA methylation, as measured in DNA from blood and other tissues (directly involved in disease), appears to be complex and locus-dependent, such initiatives should ideally include efforts to sample tissues others than blood (mesoderm), at least for a subgroup representing the cohort. Tissues representing the three germ

layers, for which collection is feasible, include: *tissues from the mesoderm*: biopsies of skeletal muscle, subcutaneous fat and the dermal layer of a skin punch biopsy (fibroblasts); *tissues from the ectoderm*: the epidermal layer of a skin punch biopsy (keratinocytes) and buccal cells; and *tissues from the endoderm*: a urine sample (bladder lining) and a stool sample (colonic mucosa). In addition, follow-up sampling of the various tissues should be included (at least for a sub-group) to assess temporal stability and changes in DNA methylation as a consequence of pathology. To account for the cellular heterogeneity of blood samples, blood cell populations should be counted if whole blood is biobanked. This can easily be done using cheap, routine methods. An alternative approach to reduce the cellular heterogeneity is to store peripheral blood mononuclear cells (PBMCs) instead of whole blood. PBMCs include lymphocytes (T and B cells), and monocytes, while the granulocytes (mainly neutrophils) are lost. To completely remove cellular heterogeneity, cells can be separated using magnetic-activated cell sorting (MACS). However, this is exceedingly costly and will not be feasible for larger numbers in most projects.

Conclusion

Taken together, our results indicate that there are good prospects for the use of existing biobanks for epigenetic studies. Loci that are suitable for testing in epigenetic studies demonstrate inter-individual variation in DNA methylation, stability of this variation in DNA methylation over time and a correlation between DNA methylation as measured blood and the tissue of interest. Our data shows that meeting these criteria is locus-dependent. Therefore, it may be necessary to address this issue for each combination of locus, tissue and disease in new studies.

Methods and Materials

Study populations

The individuals investigated in this study were selected from the Netherlands Twin Register (NTR) biobank[27,28], which includes DNA samples from Dutch twins and their family members (parents, siblings, offspring and spouses). Firstly, unrelated individuals (**Supplement I, Table S2**) were selected (n=30) to study inter-individual variation in DNA methylation, the influence of cell heterogeneity and patterns of DNA methylation. These 30 individuals were selected from the ongoing NTR biobank project for which 9560 individuals were included. Random selection would result in a sample of individuals with characteristics very close to the average in the complete cohort. For this we applied the D-optimality criterion to the Fisher information matrix, which enabled us to select 30 individuals representative of the whole range of phenotypic variation in age and metabolic parameters present in the complete cohort. The age of the individuals selected ranged from 21 to 73 years; metabolic parameters of interest included: waist circumference, fasting blood glucose level, serum LDL and HDL cholesterol. Plasma and serum measurements and cell counts of whole blood were obtained using the standardized methods previously described (missing cell count information for 2/30 individuals). Furthermore, we selected the proportion of males to females and of those who had never been smokers, to former and to current smokers so that it was equal to the proportion in the complete cohort.
Secondly, 34 individuals were selected for assessing the correlation of DNA methylation across time and tissue, and for validating the findings on the group of 30 individuals described above. This group consisted of participants in the NTR biobank project, who also took part in previous NTR projects. This allowed for recent DNA samples from whole blood (with information of cell counts) and buccal cells, as well as previous DNA samples from whole blood drawn 11 to 20 years earlier, and from buccal swabs taken 2 to 8 years earlier (**Supplement I, Table S3**). The age at first sampling ranged from 14 to 62 years. Among the 34 individuals, 17 were male, 26 were unrelated

individuals and 8 were monozygotic twin pairs. DNA from all samples was extracted from whole blood and buccal swabs using standard methods.
Thirdly, the results on within-individual correlation between CpG-units were validated using 60 controls (28 males, mean age 57 years) from the Dutch Hunger Winter Families Study[42]. DNA methylation was measured at the same loci using the same methods as the current study[29,32].

DNA methylation

Loci were selected on the basis of their potential involvement in cardiovascular and metabolic disease through the role of the adjacent candidate gene in growth, lipid metabolism, energy metabolism, inflammation, or stress response. Assay design focused on the regions of these loci that contained features with a potential for epigenetic regulation as observed in human, animal or cell culture experiments[19,24,43-55]. The loci selected included promoter elements, CpG-islands, transposon-derived sequences, methyl-sensitive transcription factor binding sites (mTFBS), imprinted differentially methylated regions (DMR) and regions reported to regulate transcription through DNA methylation. Methylation assays were designed using the methprimer tool[56] on sections of sequence downloaded from the UCSC genome browser[57]. 58 Assays were tested for the reliability of the methylation measurement. 40 assays gave a reliable measurement, and based on the priority given to the associated candidate gene and epigenetic properties (**Table 1**), 16 of these were selected to cover the whole range of possible average methylation levels (0%-100%).
One microgram of genomic DNA was bisulfite-converted using the EZ 96-DNA methylation (Zymo Research). DNA of the 30 individuals, in whom variation in DNA methylation was investigated, was converted on a single 96-well plate. DNA methylation of all 16 loci was measured using the same bisulfite-converted sample. The 4 samples (blood, buccal, recent and old) from individuals selected for testing the correlation over time and across tissues were bisulfite-treated on the same 96-well plate. For this sub-study, two 96-well plates were used to process the 136 samples, each plate with

an equal number of individuals. Methylation of the 8 loci was measured using a single bisulfite-converted DNA sample. Primers used to amplify the region of each assay are given (**Supplement I, Table S4**). DNA methylation was measured using a mass spectrometry-based method (Epityper, Sequenom)[58] whose quantitative accuracy (R^2 duplicate measurements ≥ 0.98) and concordance with clonal PCR bisulfite sequencing was reported previously[59,60]. All measurements were done in triplicate. Quality control consisted of several steps. CpG site containing fragments that had equal or overlapping mass, making them irresolvable by mass spectrometry, and CpG-sites containing fragments whose measurement was confounded by single nucleotide polymorphisms[16], according to dbSNP build 128 were discarded (**Supplement I, Table S7**). Next, at least two of the three replicate measurements had to be successful and the SD of the replicate measurements had to be 0.10 or less. Only CpG sites with a success rate greater than 75 % for the latter two criteria were considered fit for further analysis and the average was calculated for the replicate measurements. With these criteria applied, DNA methylation of 164 CpG-sites, distributed over 104 CpG site containing fragments (CpG units[58]), could be measured in the first sample of 30 individuals (**Supplement I, Table S5**), 62 CpG units contained 1 CpG site, 28 CpG units contained 2 CpG sites, 10 CpG units contained 3 CpG sites and 4 CpG units contained 4 CpG sites. The methylation of CpG sites occurring on a single fragment (CpG unit) cannot be resolved individually. Average CpG methylation for these CpG units was calculated using the RSeqMeth module[59]. The average success rate for the 104 CpG units assessed was 97%. In the second group of individuals, 41 CpG units containing 55 CpG sites could be measured applying the same criteria only to the recent blood samples and 38 CpG units containing 52 CpG sites to all 4 samples (**Supplement I, Table S6**).

To exclude the influence of DNA sequence variation not present in dbSNP on higher correlations observed between DNA methylation measured in a recent and in an old sample and blood and buccal swab, the evaluateSNPs() function of the R-module MassArray was used[33]. The Epityper method for DNA methylation measurements is based on a protocol to resequence

genomic DNA using mass spectrometry (MassCleave, Sequenom[61]). By comparing the mass spectrum observed with the one expected according to a reference sequence, data points can be identified that are suspected of being confounded by sequence variation. This interference can be directly because the sequence variant affects a fragment containing a CpG site or indirectly because a sequence variant changes the mass of a non-CpG fragment so that it overlaps with a CpG-containing fragment. It is noteworthy, that sequencing genomic DNA would deal with direct interference only. Data-points suspected to be affected by unknown sequence variation were excluded and correlations recalculated to examine their influence. CpG measurements were removed for *APOC1*, CpG 1 (4 individuals excluded for all DNA samples (recent, old, blood, buccal)), CpG 10 (8) and CpG 11 (11); *LEP*, CpG 8 (1); *IGF2*, CpG 6&7 (21); *IGF2R*, CpG 4&5 (1) and CpG 11-13 (4); for *CRH* no measurements were excluded.

Statistical analyses

There were considerable differences in the variation between loci due to that some loci and CpG sites with have average methylation levels close to 0% or 100%. Such loci have by definition a truncated variance. In order to circumvent this problem, the following variance-stabilizing transformation was applied[31].

$$transformed\ value = Arctan\left(\left(\frac{Methylation}{(1 - Methylation)}\right)^2\right)$$

Using the transformed values, the equality of the variance of CpG sites was tested with Levene's test.

To test whether variation in DNA methylation was confounded by cellular heterogeneity, nested linear mixed models[62] were applied to the transformed data. The basic model was created as a baseline, to be subtracted from the nested models. It included the CpG site as a fixed effect. The three nested models each had one percentage of a major white blood cell subclass, namely neutrophils, lymphocytes or monocytes, added to the basic model as

an extra fixed effect to test whether variation in this percentage could explain part of the variation in DNA methylation. The actual amount of variation in DNA methylation that could be explained by the percentage of the white blood cell type was calculated as 100% minus the percentage of the residual variance of the nested model with respect to the residual variance of the basic model. The linear mixed model accounts for correlated methylation within individuals and deals with methylation data missing at random without imputation[32]. It may be seen as an extension of the paired T-test: the model will reduce to a paired T-test with identical results if between-group methylation differences are assessed for a single CpG site and if data are complete and all other factors are omitted.

Bivariate Pearson correlation coefficients between CpG sites were calculated after transforming the transformed methylation levels for neutrophil percentage. Non-significant correlations were treated as if there were no correlation (value of correlation set to 0). Patterns in the correlation matrix were visualized with a heat map after unsupervised complete linkage clustering, which utilizes the Euclidean distance (the difference between two points in the matrix squared).

To gain insight in the stability over time the difference between DNA methylation at two time points was calculated per individual for each CpG unit as methylation of the old sample minus methylation of the new sample. Missing values were excluded pair-wise. Spearman's rank correlation coefficient was used to calculate the correlation between the two time points.

The correlation between tissues was addressed the following. The difference between DNA methylation in blood and buccal cells was calculated per individual for each CpG unit as methylation of the recent blood sample minus methylation of the recent buccal cell sample. Missing values were excluded pair-wise. Spearman's rank correlation coefficient was used to calculate the correlation between the two tissues.

Male and female data were analyzed separately in every test. However, in view of the study size the outcomes of these analyses should be considered as purely explorative. All P-values are two-sided and statistical analyses were performed using SPSS 16.0.

Supplementary material

Additional supplementary material, beside that provided in **Supplement I**, is available at *The FASEB Journal* online.

Acknowledgements

We are thankful to the participants of the NTR and we would like to thank Dr. L.H. Lumey, Mailman School of Public Health, Columbia University, New York, for access to the Dutch Hunger Winter Family Study samples used in our additional analysis.

Funding

This work was supported by grants from the Netherlands Heart Foundation [NHS2006B083]; the Netherlands Organization for Scientific Research NWO; Twin-family database for behavior genomics studies [480-04-004, 911–03-016]; Spinozapremie [SPI 56-464-14192]; The Centre for Medical Systems Biology (CMSB) in the framework of the Netherlands Genomics Initiative (NGI) and the Centre for Neurogenomics and Cognitive Research (CNCR-VU); Genomewide analyses of European twin and population cohorts [EU/QLRT-2001-01254]; and the European Union–funded Network of Excellence LifeSpan (FP6 036894).

References

1. Waterland RA, Michels KB. Epigenetic Epidemiology of the Developmental Origins Hypothesis. Annu.Rev.Nutr. 2007; 27: 363-88
2. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet. 2003; 33 Suppl: 245-54
3. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat.Rev.Genet. 2008; 9: 465-76
4. Kouzarides T. Chromatin modifications and their function. Cell 2007; 128: 693-705
5. Kouzarides T. SnapShot: Histone-modifying enzymes. Cell 2007; 128: 802
6. Petronis A. Human morbid genetics revisited: relevance of epigenetics. Trends Genet. 2001; 17: 142-6
7. Foley DL, Craig JM, Morley R, Olsson CJ, Dwyer T, Smith K, et al. Prospects for epigenetic epidemiology. Am.J.Epidemiol. 2009; 169: 389-400
8. Feinberg AP. Epigenetics at the epicenter of modern medicine. JAMA 2008; 299: 1345-50
9. Pons D, de Vries FR, van den Elsen PJ, Heijmans BT, Quax PH, Jukema JW. Epigenetic histone acetylation modifiers in vascular remodelling: new targets for therapy in cardiovascular disease. Eur.Heart J. 2009; 30: 266-77
10. Turunen MP, Aavik E, Yla-Herttuala S. Epigenetics and atherosclerosis. Biochim. Biophys.Acta 2009; 1790: 886-91
11. Shen L, Waterland RA. Methods of DNA methylation analysis. Curr.Opin.Clin.Nutr. Metab Care 2007; 10: 576-81
12. Smith ZD, Gu H, Bock C, Gnirke A, Meissner A. High-throughput bisulfite sequencing in mammalian genomes. Methods 2009; 48: 226-32
13. Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. Cell 2007; 128: 669-81
14. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 2009; 462: 315-22
15. Rakyan VK, Hildmann T, Novik KL, Lewin J, Tost J, Cox AV, et al. DNA methylation profiling of the human major histocompatibility complex: a pilot study for the human epigenome project. PLoS.Biol. 2004; 2: e405
16. Heijmans BT, Kremer D, Tobi EW, Boomsma DI, Slagboom PE. Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus. Hum.Mol.Genet. 2007; 16: 547-54
17. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol.Cell Biol. 2003; 23: 5293-300
18. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat.Genet. 2009; 41: 178-86
19. Weaver IC, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR, et al. Epigenetic programming by maternal behavior. Nat.Neurosci. 2004; 7: 847-54
20. Martin GM. Epigenetic drift in aging identical twins. Proc.Natl.Acad.Sci.U.S.A 2005; 102: 10413-4
21. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. Proc.Natl.Acad.Sci.U.S.A 2005; 102: 10604-9
22. Bjornsson HT, Sigurdsson MI, Fallin MD, Irizarry RA, Aspelund T, Cui H, et al. Intra-individual change over time in DNA methylation with familial clustering. JAMA 2008; 299: 2877-83
23. Ito Y, Koessler T, Ibrahim AE, Rai S, Vowler SL, bu-Amero S, et al. Somatically acquired hypomethylation of IGF2 in breast and colorectal cancer. Hum.Mol.Genet. 2008; 17: 2633-43
24. Cui H, Cruz-Correa M, Giardiello FM, Hutcheon DF, Kafonek DR, Brandenburg S, et al. Loss of IGF2 imprinting: a potential marker

of colorectal cancer risk. Science 2003; 299: 1753-5

25. Ally MS, Al-Ghnaniem R, Pufulete M. The relationship between gene-specific DNA methylation in leukocytes and normal colorectal mucosa in subjects with and without colorectal tumors. Cancer Epidemiol. Biomarkers Prev. 2009; 18: 922-8

26. Byun HM, Siegmund KD, Pan F, Weisenberger DJ, Kanel G, Laird PW, et al. Epigenetic profiling of somatic tissues from human autopsy specimens identifies tissue- and individual-specific DNA methylation patterns. Hum.Mol.Genet. 2009; 18: 4808-17

27. Boomsma DI, de Geus EJ, Vink JM, Stubbe JH, Distel MA, Hottenga JJ, et al. Netherlands Twin Register: from twins to twin families. Twin.Res.Hum.Genet. 2006; 9: 849-57

28. Boomsma DI, Willemsen G, Sullivan PF, Heutink P, Meijer P, Sondervan D, et al. Genome-wide association of major depression: description of samples for the GAIN Major Depressive Disorder Study: NTR and NESDA biobank projects. Eur.J.Hum. Genet. 2008; 16: 335-42

29. Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD, et al. DNA Methylation differences after exposure to prenatal famine are common and timing- and sex-specific. Hum.Mol.Genet. 2009; 18: 4046-53

30. El-Maarri O, Becker T, Junen J, Manzoor SS, az-Lacava A, Schwaab R, et al. Gender specific differences in levels of DNA methylation at selected loci from human total blood: a tendency toward higher methylation levels in males. Hum.Genet. 2007; 122: 505-14

31. Altman DG. Practical statistics for medical research. London; New York: Chapman and Hall, 1991

32. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc. Natl.Acad.Sci.U.S.A 2008; 105: 17046-9

33. Thompson RF, Suzuki M, Lau KW, Greally JM. A pipeline for the quantitative analysis of CG dinucleotide methylation using mass spectrometry. Bioinformatics. 2009; 25: 2164-70

34. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc.Natl.Acad.Sci.U.S.A 1996; 93: 9821-6

35. Bock C, Walter J, Paulsen M, Lengauer T. Inter-individual variation of DNA methylation and its implications for large-scale epigenome mapping. Nucleic Acids Res. 2008; 36: e55

36. The International HapMap Consortium. The International HapMap Project. Nature 2003; 426: 789-96

37. Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. Dnmt3L and the establishment of maternal genomic imprints. Science 2001; 294: 2536-9

38. Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 2007; 448: 714-7

39. Ramchandani S, Bhattacharya SK, Cervoni N, Szyf M. DNA methylation is a reversible biological signal. Proc.Natl.Acad.Sci.U.S.A 1999; 96: 6107-12

40. Heijmans BT, Tobi EW, Lumey LH, Slagboom PE. The epigenome: Archive of the prenatal environment. Epigenetics. 2009; 4: 526-31

41. Bollati V, Schwartz J, Wright R, Litonjua A, Tarantini L, Suh H, et al. Decline in genomic DNA methylation through aging in a cohort of elderly subjects. Mech.Ageing Dev. 2009; 130: 234-9

42. Lumey L, Stein AD, Kahn HS, van der Pal-de Bruin KM, Blauw G, Zybert PA, et al. Cohort Profile: the Dutch Hunger Winter Families Study. Int.J.Epidemiol. 2007; 36: 1196-204

43. Dong J, Ivascu C, Chang HD, Wu P, Angeli R, Maggi L, et al. IL-10 Is Excluded from the Functional Cytokine Memory of Human CD4+ Memory T Lymphocytes. J.Immunol. 2007; 179: 2389-96

44. Hayward BE, Kamiya M, Strain L, Moran V, Campbell R, Hayashizaki Y, et al. The human GNAS1 gene is imprinted and encodes distinct paternally and biallelically expressed G proteins. Proc.Natl.Acad.Sci.U.S.A 1998; 95: 10038-43

45. Sullivan KE, Reddy AB, Dietzmann K, Suriano AR, Kocieda VP, Stewart M, et al. Epigenetic

regulation of tumor necrosis factor alpha. Mol. Cell Biol. 2007; 27: 5147-60

46. Sandovici I, Leppert M, Hawk PR, Suarez A, Linares Y, Sapienza C. Familial aggregation of abnormal methylation of parental alleles at the IGF2/H19 and IGF2R differentially methylated regions. Hum.Mol.Genet. 2003; 12: 1569-78

47. Arnaud P, Monk D, Hitchins M, Gordon E, Dean W, Beechey CV, et al. Conserved methylation imprints in the human and mouse GRB10 genes with divergent allelic expression suggests differential reading of the same mark. Hum.Mol.Genet. 2003; 12: 1005-19

48. Melzner I, Scott V, Dorsch K, Fischer P, Wabitsch M, Bruderlein S, et al. Leptin gene expression in human preadipocytes is switched on by maturation-induced demethylation of distinct CpGs in its proximal promoter. J.Biol.Chem. 2002; 277: 45420-7

49. McGill BE, Bundle SF, Yaylaoglu MB, Carson JP, Thaller C, Zoghbi HY. Enhanced anxiety and stress-induced corticosterone release are associated with increased Crh expression in a mouse model of Rett syndrome. Proc.Natl. Acad.Sci.U.S.A 2006; 103: 18267-72

50. Probst MC, Thumann H, Aslanidis C, Langmann T, Buechler C, Patsch W, et al. Screening for functional sequence variations and mutations in ABCA1. Atherosclerosis 2004; 175: 269-79

51. Adkins RM, Fain JN, Krushkal J, Klauser CK, Magann EF, Morrison JC. Association between paternally inherited haplotypes upstream of the insulin gene and umbilical cord IGF-II levels. Pediatr.Res. 2007; 62: 451-5

52. Mitsuya K, Meguro M, Lee MP, Katoh M, Schulz TC, Kugoh H, et al. LIT1, an imprinted antisense RNA in the human KvLQT1 locus identified by screening for differentially expressed transcripts using monochromosomal hybrids. Hum.Mol.Genet. 1999; 8: 1209-17

53. Rosa AL, Wu YQ, Kwabi-Addo B, Coveler KJ, Reid S, V, Shaffer LG. Allele-specific methylation of a functional CTCF binding site upstream of MEG3 in the human imprinted domain of 14q32. Chromosome.Res. 2005; 13: 809-18

54. Medstrand P, Landry JR, Mager DL. Long terminal repeats are used as alternative promoters for the endothelin B receptor and apolipoprotein C-I genes in humans. J.Biol. Chem. 2001; 276: 1896-903

55. Hayward BE, Bonthron DT. An imprinted antisense transcript at the human GNAS1 locus. Hum.Mol.Genet. 2000; 9: 835-41

56. Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002; 18: 1427-31

57. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002; 12: 996-1006

58. Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. Proc.Natl.Acad.Sci.U.S.A 2005; 102: 15785-90

59. Coolen MW, Statham AL, Gardiner-Garden M, Clark SJ. Genomic profiling of CpG methylation and allelic specificity using quantitative high-throughput mass spectrometry: critical evaluation and improvements. Nucleic Acids Res. 2007; 35: e119

60. Ehrich M, Turner J, Gibbs P, Lipton L, Giovanneti M, Cantor C, et al. Cytosine methylation profiling of cancer cell lines. Proc. Natl.Acad.Sci.U.S.A 2008; 105: 4844-9

61. Stanssens P, Zabeau M, Meersseman G, Remes G, Gansemans Y, Storm N, et al. High-throughput MALDI-TOF discovery of genomic sequence polymorphisms. Genome Res. 2004; 14: 126-33

62. West B, Welch K, Galecki A. Linear Mixed Models: A Practical Guide Using Statistical Software. Chapmann Hall / CRC Press, 2006

Persistent epigenetic differences associated with prenatal exposure to famine in humans

Elmar W. Tobi[1,†], Bastiaan T. Heijmans[1,7,†], Aryeh D. Stein[2], Hein Putter[3], Gerard J. Blauw[4], Ezra S. Susser[5,6], P. Eline Slagboom[1,7], L.H. Lumey[5]

1. Molecular Epidemiology, Leiden University Medical Centre, Leiden, The Netherlands
2. Global Health, Emory University, Atlanta, GA, USA
3. Medical Statistics, Leiden University Medical Centre, Leiden, The Netherlands
4. Gerontology and Geriatrics, Leiden University Medical Centre, Leiden, The Netherlands
5. Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA
6. New York State Psychiatric Institute, New York, NY, USA
7. The Netherlands Consortium for Healthy Ageing, The Netherlands

† These authors contributed equally to this work.

Proc Natl Acad Sci U S A. **2008** Nov 4:105(44):17046-17049

Abstract

Extensive epidemiological studies suggested that adult disease risk is associated with adverse environmental conditions early in development. While the mechanisms behind these relationships are unclear, an involvement of epigenetic dysregulation has been hypothesized. Here we show that individuals who were prenatally exposed to famine during the Dutch Hunger Winter in 1944/45 had, six decades later, less DNA methylation of the imprinted *IGF2* gene than their unexposed same-sex siblings. The association was specific for periconceptional exposure reinforcing that very early mammalian development is a crucial period for establishing and maintaining epigenetic marks. Exposure late in gestation was not related to *IGF2* methylation. These data are the first to contribute empirical support for the hypothesis that early-life environmental conditions can cause epigenetic changes in humans that persist throughout life.

Introduction

Superimposed on the DNA sequence is a layer of epigenetic information that is heritable, particularly during mitosis, and controls the potential of a genomic region to become transcribed[1]. The two main molecular marks that make up this information and regulate chromatin structure and DNA accessibility are methyl groups coupled to a cytosine in CpG dinucleotide and histone modifications that package the DNA[2].

Although generally stable, animal studies indicated that certain transient environmental influences can produce persistent changes in epigenetic marks that have life-long phenotypic consequences[3,4]. Early embryonic development is of special interest in this respect since it is a crucial period in establishing and maintaining epigenetic marks[5]. Indeed, culturing of preimplantation mice embryos showed that epigenetic marks are susceptible to nutritional conditions in the very early stages of mammalian development[6,7].

One of the rare opportunities for studying the relevance of such findings to humans, is presented by individuals who were prenatally exposed to famine during the Dutch Hunger Winter[8]. This period of famine was the consequence of a German imposed food-embargo in the western part of The Netherlands towards the end of World War II in the winter of 1944/45. During this period registries and health care remained intact so that individuals who were prenatally exposed to this famine can be traced. Moreover, the period of famine was clearly defined and official food rations documented. These unique features allow us to assess whether prenatal exposure to famine is associated with persistent epigenetic differences in humans.

One of the best characterized epigenetically regulated loci is insulin-like growth factor II (*IGF2*). *IGF2* is a key factor in human growth and development and is maternally imprinted[9]. Imprinting is maintained through the *IGF2* differentially methylated region (DMR) whose hypo-methylation leads to bi-allelic expression of *IGF2*[10]. We recently studied *IGF2* DMR methylation in 372 twins[11]. *IGF2* DMR methylation was a normally distributed quantitative trait and was largely determined by genetic factors in both adolescence and middle age indicating that the methylation mark is stable up to middle age.

If affected by environmental conditions early in human development, altered *IGF2* DMR methylation may therefore be detected many years later.
Here we use our ongoing Hunger Winter Families Study[8] to investigate whether prenatal exposure to famine is associated with persistent differences in methylation of the *IGF2* DMR. Our primary focus was exposure during periconception, thus ensuring that the exposure was present during the very early stages of development that are critical in epigenetic programming. To further investigate the role of timing, we also studied individuals who were exposed late in gestation.

Results

Our primary goal was to test whether periconceptional exposure to famine was associated with differences in *IGF2* DMR methylation in adulthood. Using a quantitative mass spectrometry-based method[11,12], the methylation of 5 CpG dinucleotides within the *IGF2* DMR was measured. Three CpG sites were measured individually and 2 simultaneously, because they could not be resolved due to their close proximity.

Periconceptional exposure

We selected the 60 individuals from the Hunger Winter Families Study who were conceived during the famine six decades ago. The exposure period thus included the very early stages of development. The exposed individuals were compared to their same-sex sibling to achieve partial genetic matching. All but one CpG site were significantly less methylated among periconceptionally exposed individuals as compared to their siblings ($1.5x10^{-4} \leq P \leq 8.1x10^{-3}$; see **Table 1**). The average methylation fraction of the *IGF2* DMR based on all 5 CpG sites was 0.488 among exposed and 0.515 among unexposed siblings. Periconceptional exposure therefore was associated with -5.2% lower methylation ($P=5.9x10^{-5}$) corresponding to 0.48 standard deviations of the controls. The association was independent of sex ($P_{interaction}=0.20$).

Table 1. *IGF2* DMR methylation among individuals periconceptionally exposed to famine and their unexposed, same-sex siblings.

IGF2 DMR methylation	Mean methylation fraction (SD)				Relative change exposed	Difference in SDs	P[1]
	Exposed (n=60)		Controls (n=60)				
Average	0.488	(0.047)	0.515	(0.055)	-5.2%	-0.48	$5.9x10^{-5}$
CpG 1	0.436	(0.037)	0.470	(0.041)	-6.9%	-0.78	$1.5x10^{-4}$
CpG 2&3	0.451	(0.033)	0.473	(0.055)	-4.7%	-0.41	$8.1x10^{-3}$
CpG 4	0.577	(0.114)	0.591	(0.112)	-2.3%	-0.12	0.41
CpG 5	0.491	(0.061)	0.529	(0.068)	-7.2%	-0.56	$1.4x10^{-3}$

1. P-values were adjusted for age.

Figure 1 displays the difference in *IGF2* DMR methylation within sib ships according to the estimated conception date of the famine-exposed individual. *IGF2* DMR methylation was lowest in the famine-exposed individual among 72% (43/60) of sib ships and this lower methylation was observed in conceptions across the famine period. Official daily rations were set weekly and the same for every individual. The average daily rations during the famine were 667 kcal (SD, 151) (**Figure 1**).

As a technical validation of this finding, *IGF2* DMR methylation was re-measured among 46 out of 60 periconceptionally exposed individuals and their same-sex sibling, repeating the whole procedure from bisulfite treatment to quantification. A similarly lower 5.6% *IGF2* DMR methylation was observed ($P=2.1x10^{-3}$), confirming our initial findings.

Late gestational exposure

To further investigate the influence of timing, we selected the 62 individuals who were exposed to famine late in gestation for at least 10 weeks so that they were born in or shortly after the famine. This time no difference in *IGF2* DMR methylation was observed between exposed individuals and their unexposed siblings (**Table 2** and **Figure 1B**).

To formally test whether the association with lower *IGF2* DMR methylation depended on timing of the exposure, we analyzed the periconceptional

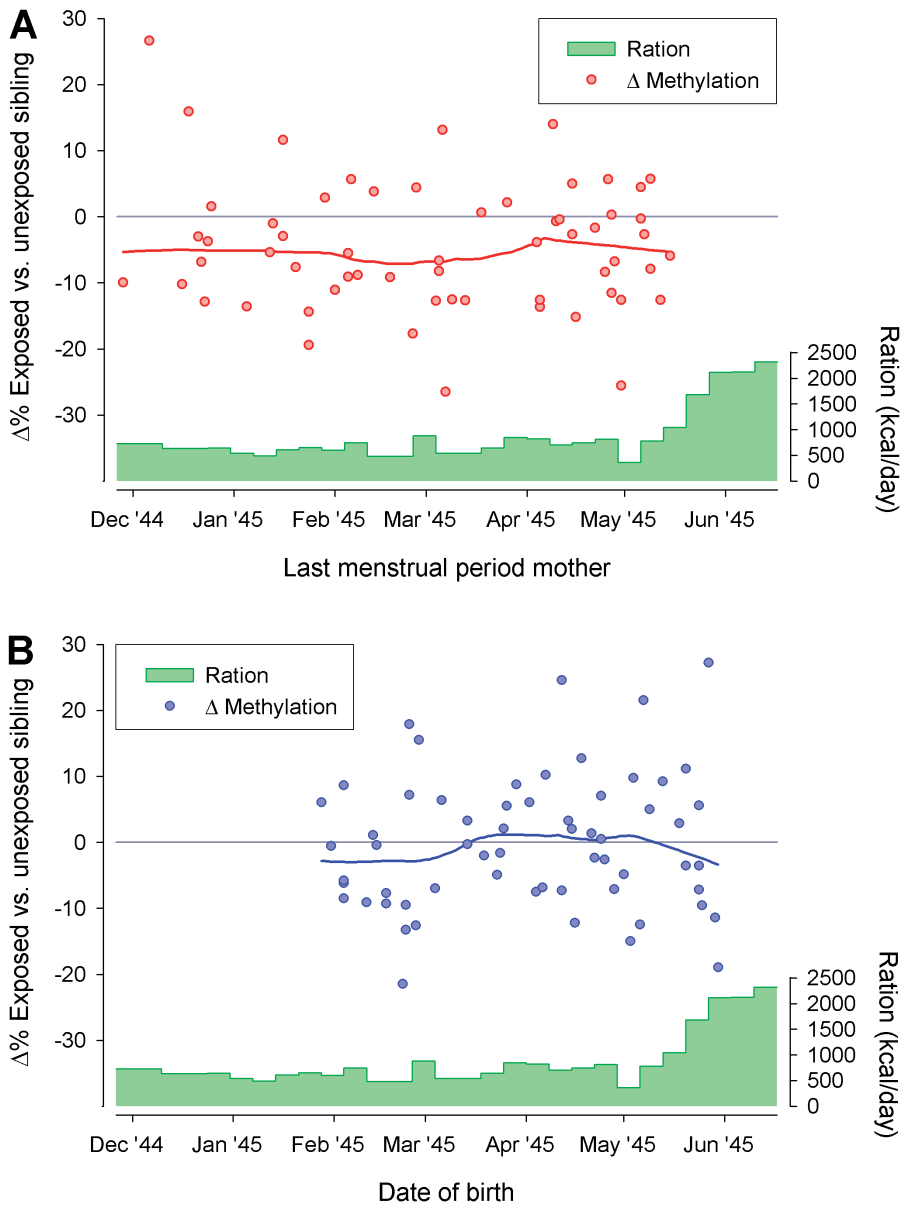


Figure 1. Within sibling pair difference in IGF2 DMR methylation set against the estimate of conception or data of birth of the famine-exposed individual.

A lowes curve is drawn to describe the difference in methylation. In green, the average distributed rations (kcal/day) between December 1944 and June 1945 are depicted. **A.**To describe the difference in methylation according to the data of the last menstrual period of the mother (a common estimate of conception), a lowess curve (red) was drawn. **B.** The lowess curve (blue) describes the difference in methylation according to the birth dates.

and late exposure groups together with all 122 controls in a single model (**Table 3**). Periconceptional exposure was associated with lower methylation ($P=1.5x10^{-5}$) whereas late exposure was not (P=0.69). Furthermore, there was statistically significant evidence for an interaction between timing and exposure ($P_{interaction}=4.7x10^{-3}$) indicating that the association was timing-specific.

Table 2. IGF2 DMR methylation among individuals exposed to famine late in gestation and their unexposed, same-sex siblings.

IGF2 DMR methylation	Mean methylation fraction (SD) Exposed (n=62)		Controls (n=62)		Relative change exposed	Difference in SDs	P[1]
Average	0.514	0.045	0.519	0.036	-0.9%	-0.12	0.64
CpG 1	0.460	0.044	0.464	0.048	-0.9%	-0.09	0.68
CpG 2&3	0.462	0.039	0.471	0.039	-1.7%	-0.21	0.46
CpG 4	0.602	0.085	0.612	0.073	-1.5%	-0.12	0.30
CpG 5	0.529	0.060	0.531	0.060	-0.3%	-0.02	0.77

1. P-values were adjusted for age.

Birth weight

The mean birth weight of the 62 individuals exposed late in gestation was 3126 g (SD, 408), which is 296 g lower (95% CI, -420 to -170 g) than the mean (3422 g; SD, 464) of 324 reference births in 1943 at the same institutions ($P= 4x10^{-6}$)[13]. The lower birth weight underscores the impact of the famine during the Hunger Winter notwithstanding the absence of an association with *IGF2* DMR methylation. The mean birth weight of the 60 individuals periconceptionally exposed was 3612 g (SD, 648) and was not lower compared to the reference births (95% CI, +15 to +365 g; P=0.03). *IGF2* DMR methylation was not associated with birth weight (P=0.39).

Age association

To place the 5.2% lower *IGF2* DMR methylation association with periconceptional famine exposure into perspective, we assessed the relationship between age and *IGF2* DMR methylation in the 122 control individuals. Within the age range studied of 43 to 70 years, a 10 year older age was associated with a 3.2% lower methylation (P=0.037).

Table 3. Timing of famine exposure during gestation, IGF2 DMR methylation and birth weight.

	Periconceptional exposure	Late gestational exposure	all controls
N	60	62	122
Males, %	46.7	45.2	45.9
Mean age, years	58.1 (SD, 0.35)	58.8 (SD, 0.4)	57.1 (SD, 5.5)
Birth weight, g	3612 (SD, 648)	3126 (SD, 408)	-
IGF2 DMR Methylation			
Average	0.488 (SD, 0.047)	0.514 (SD, 0.045)	0.517 (SD,0.047)
$P_{vs\ all\ controls}$ [1]	$1.5x10^{-5}$	0.69	
$P_{interaction}$			$4.7x10^{-3}$

1. P values were adjusted for age.

Discussion

Here we report that periconceptional exposure to famine during the Dutch Hunger Winter is associated with lower methylation of the *IGF2* DMR six decades later. The hypo-methylation we observed is remarkably comparable to that found in offspring of female rats fed a protein deficient diet starting before pregnancy (-5.2% in our human study versus -8.2% and -10.2% of rat *Nr3c1* and *Ppara* genes, respectively)[3]. The similarity substantiates that famine is the main culprit in *IGF2* hypo-methylation. An additional contribution of other stressors like cold and emotional stress [8], however, cannot be ruled out. Our study provides the first evidence that transient environmental

conditions early in human gestation can be recorded as persistent changes in epigenetic information.

In contrast to periconceptional exposure to famine, exposure late in gestation was not associated with *IGF2* DMR methylation. Epigenetic marks may be particularly vulnerable during the very early stages of mammalian development which are crucial in establishing and maintaining epigenetic marks[5]. Experiments in which mouse zygotes were cultured to blastocyst stage favour this hypothesis[6,7]. The timing-dependence of the association we observed may, however, also relate to timing of tissue development[14]. We studied blood and adult blood cells stem from the hematopoietic system that is established in the first stages of development (e.g. day 10.5 of the mouse embryo[15]). Detailed future studies are required to establish whether the susceptibility of epigenetic marks is an intrinsic property of early mammalian development or a general feature of newly developing tissues throughout gestation.

The developmental origins hypothesis states that adverse conditions during development contribute to adult disease risk[16]. While the mechanisms behind these relationships are unclear, it has been proposed that epigenetic dysregulation is involved[16-18]. Our results are a key element in elaborating this hypothesis. Human studies on the developmental origins of health and disease, however, often use low birth weight as a proxy for a compromised prenatal development[16]. Our data indicate that such studies are not necessarily sufficient to test the involvement of epigenetics and thereby extend our previous finding that birth weight is a poor surrogate for nutritional status during gestation[13]. Epigenetic differences were found among individuals who were exposed to famine early in gestation and had a normal birth weight. Exposure late in gestation was associated with low birth weight as expected but not with epigenetic changes. To monitor the crucial stages of early development, assessing maternal life style especially regarding nutrition[19] and embryo growth using 3D ultrasound[20] will be more appropriate than birth weight.

The current study presents a first example of an association between a periconceptional exposure and DNA methylation in humans. It will be

of prime interest to investigate whether other exposures during early development that are more common in modern societies like overnutrition[3] and assisted reproductive technologies[21] give rise to similar associations. Also, the extent to which epigenetic marks at other genomic regions are vulnerable to such exposures remains to be established. A key question of future studies will be to assess the phenotypic consequences of changes in epigenetic marks. Diseases that have been associated with early gestational exposure to famine like schizophrenia[22] and coronary heart disease[23] are of particular interest in this respect. Analogous to current studies in genetic epidemiology[24], such epigenetic epidemiological studies may need to be large and include replication. Understanding how epigenetic control depends on early exposures may shed light on the link between development and health over life time and ultimately suggest new ways to prevent human disease.

Materials and Methods

Study population

Design of and recruitment for the Hunger Winter Families Study was described previously[8]. Individuals pr enatally exposed to famine were recruited by identifying and follow-up of live singleton births in 1945 and early 1946 at three institutions in famine-exposed cities (the midwifery training schools in Amsterdam and Rotterdam and the Leiden University hospital). As controls, same-sex siblings and unrelated individuals from the same institutions were recruited who were born before or conceived after the famine period. Clinical examination including obtaining a blood sample was completed for 311 exposed individuals, 311 same-sex siblings and 349 unrelated controls. Birth weight was abstracted from birth records from the three institution. No birth weight data are available for the same-sex siblings, who were not born at these institutions.

For the current epigenetic study, we focused on exposed individuals and their sibling as control to achieve partial genetic matching in view of the high heritability of *IGF2* DMR methylation[11]. From these, we selected the sib ships

with an individual periconceptionally exposed to famine and those with an individual exposed to famine late in gestation. Periconceptional exposure was defined as a last menstrual period of the mother before conceiving the exposed individual between November 28, 1944 and May 15, 1945. This yielded 60 sib ships. Exposure late in gestation was defined as a birth between January January 28 and May 30 1945 so that the duration of the famine exposure was at least 10 weeks. This yielded 62 sib ships.

DNA methylation

Methylation of the *IGF2* DMR was measured using genomic DNA from whole blood extracted using the salting out method. One microgram of genomic DNA was bisulfite-treated using the EZ 96-DNA methylation kit (Zymo Research). Sib ships were bisulfite treated on the same plate. The region harbouring the *IGF2* DMR was amplified using primers described elsewhere[11]. DNA methylation was quantified using a mass spectrometry-based method (Epityper, Sequenom)[12]. All measurements were done in triplicate and CpG dinucleotides whose measurement was confounded by single nucleotide polymorphisms were discarded as part of quality controls[11]. This resulted in 93% complete data. DNA methylation of 5 CpG dinucleotides could be measured, 3 of which individually and 2 of which as pair because they were directly adjacent and could not be resolved individually.

Statistical analysis

Mean methylation fractions of individual CpGs and their SDs presented in tables are based on raw data. To obtain the average methylation of the whole *IGF2* DMR presented in tables and figures, missing methylation data was first imputed using estimates from the mixed model thus exploiting the correlation between CpG sites[11]. To test for differences between exposed individuals and their unexposed siblings, linear mixed models were used. This analysis accounts for family relations, correlated methylation of CpG dinucleotides and methylation data missing at random (thus p-values were

calculated without imputation of missing data). Exposure status, CpG dinucleotide and age were entered as fixed effects and sib ship as random effect. The model including both the periconceptional and the late exposure groups were extended with a variable indicating timing of the exposure and an interaction term exposure status times exposure time. To test for the association between *IGF2* DMR methylation and birth weight, birth weight was added as a fixed effect. P-values are two-sided and statistical analyses were performed using SPSS 14.0.

Acknowledgements

We wish to thank the participants of the Hunger Winter Families Study, TNO Quality of Life for contact tracing, the staff of the Gerontology and Geriatrics Study Centre at the LUMC for performing the clinical examinations, Marja Kersbergen and Margot van Schie for extracting genomic DNA and Dennis Kremer for technical assistance. This work was financially supported by grants from the Netherlands Heart Foundation (2006B083 to BTH), the US National Institutes of Health (RO1-HL067914 to LHL), The Netherlands Organization for Scientific Research NWO (grant 911-03-016 to PES) and the EU funded Network of Excellence LifeSpan (FP6 036894).

References

1. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet. 2003; 33 Suppl: 245-54
2. Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. Cell 2007; 128: 669-81
3. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol.Cell Biol. 2003; 23: 5293-300
4. Sinclair KD, Allegrucci C, Singh R, Gardner DS, Sebastian S, Bispham J, et al. DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and methionine status. Proc.Natl.Acad.Sci.U.S.A 2007; 104: 19351-6
5. Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. Science 2001; 293: 1089-93
6. Doherty AS, Mann MR, Tremblay KD, Bartolomei MS, Schultz RM. Differential effects of culture on imprinted H19 expression in the preimplantation mouse embryo. Biol. Reprod. 2000; 62: 1526-35
7. Morgan HD, Jin XL, Li A, Whitelaw E, O'Neill C. The Culture of Zygotes to the Blastocyst Stage Changes the Postnatal Expression of an Epigentically Labile Allele, Agouti Viable Yellow, in Mice. Biol.Reprod. 2008; DOI:10.1095/biolreprod.108.068213:
8. Lumey LH, Stein AD, Kahn HS, van der Pal-de Bruin KM, Blauw GJ, Zybert PA, et al. Cohort Profile: the Dutch Hunger Winter Families Study. Int.J.Epidemiol. 2007; 36: 1196-204
9. Smith FM, Garfield AS, Ward A. Regulation of growth and metabolism by imprinted genes. Cytogenet.Genome Res. 2006; 113: 279-91
10. Cui H, Cruz-Correa M, Giardiello FM, Hutcheon DF, Kafonek DR, Brandenburg S, et al. Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. Science 2003; 299: 1753-5
11. Heijmans BT, Kremer D, Tobi EW, Boomsma DI, Slagboom PE. Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus. Hum.Mol.Genet. 2007; 16: 547-54
12. Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. Proc.Natl.Acad.Sci.U.S.A 2005; 102: 15785-90
13. Stein AD, Zybert PA, van de BM, Lumey LH. Intrauterine famine exposure and body proportions at birth: the Dutch Hunger Winter. Int.J.Epidemiol. 2004; 33: 831-6
14. Waterland RA, Lin JR, Smith CA, Jirtle RL. Post-weaning diet affects genomic imprinting at the insulin-like growth factor 2 (Igf2) locus. Hum.Mol.Genet. 2006; 15: 705-16
15. Dzierzak E, Speck NA. Of lineage and legacy: the development of mammalian hematopoietic stem cells. Nat.Immunol. 2008; 9: 129-36
16. Waterland RA, Michels KB. Epigenetic Epidemiology of the Developmental Origins Hypothesis. Annu.Rev.Nutr. 2007; 27: 363-88
17. Cutfield WS, Hofman PL, Mitchell M, Morison IM. Could epigenetics play a role in the developmental origins of health and disease? Pediatr.Res. 2007; 61: 68R-75R
18. McClellan JM, Susser E, King MC. Maternal famine, de novo mutations, and schizophrenia. JAMA 2006; 296: 582-4
19. Verkleij-Hagoort AC, de Vries JH, Stegers MP, Lindemans J, Ursem NT, Steegers-Theunissen RP. Validation of the assessment of folate and vitamin B12 intake in women of reproductive age: the method of triads. Eur.J.Clin.Nutr. 2007; 61: 610-5
20. Verwoerd-Dikkeboom CM, Koning AH, van der Spek PJ, Exalto N, Steegers EA. Embryonic staging using a 3D virtual reality system. Hum.Reprod. 2008; 23: 1479-84
21. DeBaun MR, Niemitz EL, Feinberg AP. Association of in vitro fertilization with Beckwith-Wiedemann syndrome and epigenetic alterations of LIT1 and H19. Am.J.Hum.Genet. 2003; 72: 156-60
22. Susser E, Neugebauer R, Hoek HW, Brown AS, Lin S, Labovitz D, et al. Schizophrenia

after prenatal famine. Further evidence. Arch. Gen.Psychiatry 1996; 53: 25-31

23. Painter RC, de R, Sr., Bossuyt PM, Simmers TA, Osmond C, Barker DJ, et al. Early onset of coronary artery disease after prenatal exposure to the Dutch famine. Am.J.Clin.Nutr. 2006; 84: 322-7

24. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447: 661-78

DNA Methylation differences after exposure to prenatal famine are common and timing- and sex-specific

Elmar W. Tobi[1], L.H. Lumey[3,5], Rudolf P. Talens[1], Dennis Kremer[1], Hein Putter[2], Aryeh D. Stein[4], P. Eline Slagboom[1,6], Bastiaan T. Heijmans[1,6]

1. Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
2. Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands
3. Endocrinology, Leiden University Medical Center, Leiden, The Netherlands
4. Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA
5. Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA
6. The Netherlands Consortium for Healthy Ageing, The Netherlands

Hum. Mol. Genet. **2009** Nov 1:18(21):4046-4053

Abstract

Prenatal famine in humans has been associated with various later-life consequences depending on the gestational timing of the insult and the sex of the exposed individual. Epigenetic mechanisms have been proposed to underlie these associations. Indeed, animal studies and our early human data on the imprinted *IGF2* locus indicated a link between prenatal nutritional and DNA methylation. However, it remains unclear how common changes in DNA methylation are and whether they are sex- and timing-specific paralleling the later-life consequences of prenatal famine exposure. To this end we investigated the methylation of 15 loci implicated in growth and metabolic disease in individuals who were prenatally exposed to a war-time famine in 1944-45. Methylation of *INSIGF* was lower among individuals who were periconceptionally exposed to the famine (n=60) as compared with their unexposed same-sex siblings ($P=2x10^{-5}$), whereas methylation of *IL10*, *LEP*, *ABCA1, GNASAS* and *MEG3* were higher (all $P<10^{-3}$). A significant interaction with sex was observed for *INSIGF*, *LEP* and *GNASAS*. Next, methylation of 8 representative loci was compared between 62 individuals exposed late in gestation and their unexposed siblings. Methylation was different for *GNASAS* ($P=1.1x10^{-7}$) and, in men, *LEP* (P=0.017). Our data indicate that persistent changes in DNA methylation may be a common consequence of prenatal famine exposure and that these changes depend on the sex of the exposed individual and the gestational timing of the exposure.

Introduction

Adverse environmental conditions during specific windows of mammalian development can have lasting effects on metabolic pathways and physiology, thereby influencing the susceptibility to chronic diseases[1]. An extensive epidemiologic literature has reported associations between characteristics of early development and health outcomes later in life[2,3]. Historical famines provide a quasi-experimental setting in which the long-term consequences of adverse conditions during development can be studied in humans. Studies of the Dutch Hunger Winter, a severe wartime famine at the end of WWII affecting the western part of The Netherlands, suggest that famine exposure *in utero* can lead to various adverse metabolic or mental phenotypes, depending on the sex of the exposed individual and the timing of the exposure during gestation[4-7].

The period around conception may be especially sensitive to famine exposure[8]. Exposure to famine in this period is associated with diverse phenotypic outcomes such as an increased risk of adult schizophrenia[7] and *spina bifida* at birth in men[9]. However, the effects of prenatal famine exposure are not limited to this developmental period. An example of a general, sex-specific late life effect is the increase in body mass index[6,10] and various lipids in blood[4] among famine-exposed women, irrespective of the precise gestational timing of the exposure. Also, increases in cerebro-cardiovascular related deaths have been reported among individuals exposed to seasonal food shortages independent of the gestational timing in a historical cohort[11], although preliminary results from the Dutch Famine indicated that an increased risk of coronary artery disease is specific for exposure to famine early in gestation[12]. Animal experiments confirm the importance of timing and sex[13,14]. Thus, prenatal exposure to famine can have different long-term effects that depend on the timing of the exposure and the sex of the exposed individual.

Persistent epigenetic changes induced by environmental factors are a plausible molecular mechanism underlying the relationship between early development and later life disease[15,16]. Experiments in animal models provide

strong supporting evidence. Manipulation of the maternal diet during pregnancy lead to a persistent shift in average DNA methylation levels of specific genes in offspring resulting in permanent changes in coat color or tail shape[17,18]. A proof-of-principle for complex diseases was reported by Bogdarina *et al*., who showed in rats that a low protein diet during pregnancy was associated with decreased DNA methylation of the *Agtr1b* gene promoter in offspring, which may explain the increase in blood pressure among these animals[19]. We recently showed that similar mechanisms may be operative in humans; periconceptional exposure to famine was associated with a decrease in DNA methylation of the insulin-like growth factor 2 (*IGF2*) differentially methylated region (DMR)[20]. However, to be a valid candidate mechanism in humans, the effect of prenatal famine on DNA methylation should be widespread, mirror the epidemiologic findings with sex- and timing-specific associations and affect genes in relevant pathways.

To further explore associations between prenatal famine and DNA methylation, including the role of the timing of exposure and the sex of the prenatally exposed individual, we assessed the methylation state of loci in 15 candidates genes involved in metabolic and cardiovascular disease and growth with diverse epigenetic features in our ongoing Hunger Winter Families Study[21].

Results

We studied methylation of loci implicated in the transcriptional regulation of 15 candidate genes for metabolic and cardiovascular disease. The loci studied included imprinted loci (*GNASAS*, *GNAS A/B*, *MEG3*, *KCNQ1OT1, INSIGF* and *GRB10*), a putatively imprinted locus (*IGF2R*) and non-imprinted loci (*IL10*, *TNF*, *ABCA1*, *APOC1*, *FTO*, *LEP*, *NR3C1* and *CRH*). The selection of the loci measured was based on a combination of factors, including binding of methylation sensitive transcription factors, associations of DNA methylation with gene expression and the presence of differentially methylated regions for imprinted loci (a detailed overview is given in **Chapter 2, Table 1**). The characteristics of these loci, like the variation in DNA methylation, stability

over time and correlation with a tissue of a different embryonic lineage (buccal cells) was previously detailed (**Chapter 2**)[22].

Periconceptional exposure

DNA methylation of these 15 loci was measured in 60 individuals conceived during the Dutch Famine (i.e. exposed periconceptionally) and compared with their unexposed same-sex sibling to minimize the possible confounding effects of familial environment and genetic background. For six of the fifteen loci, significant differences in DNA methylation were observed (**Table 1**). DNA methylation was increased among famine exposed individuals for the imprinted genes *GNASAS* ($P=3.1x10^{-6}$) and *MEG3* ($P=8.0x10^{-3}$) and the non-imprinted *IL10* ($P=1.8x10^{-6}$), *ABCA1* ($P=8.2x10^{-4}$) and *LEP* ($P=2.9x10^{-3}$) proximal promoters. DNA methylation was decreased for the imprinted *INSIGF* promoter ($P=2.3x10^{-5}$), which is part of the proximal promoter of *INS*[23]. For the remaining 9 loci there was no association with periconceptional exposure to famine. All associations remained statistically significant after Bonferroni correction for multiple testing (15 loci) with the exception of *MEG3* ($P_{Bonferroni}$= 0.12). When analysed separately, individual CpG dinucleotides showed similar associations with famine exposure as the complete loci (**Figure 1A**). The loci affected did not share obvious features with respect to sequence, epigenetic features or biological function.

Late gestational exposure

Epigenetic modulation may also occur during other developmental windows[24,25]. We therefore also studied 62 individuals who were exposed to famine late in gestation together with their unexposed, same-sex siblings. We measured DNA methylation at four loci that were associated with periconceptional famine exposure (*IL10*, *GNASAS*, *INSIGF* and *LEP*) and at four that were not (*IGF2R*, *APOC1*, *KCNQ1OT1* and *CRH*). These loci include four imprinted ones and diverse epigenetic features.

No associations were observed except for a significant reduction in methylation at the *GNASAS* locus ($P=1.1x10^{-7}$, $P_{Bonferroni}= 8.8 x10^{-7}$) (**Table 2**). This association was consistent for the individual CpG sites within the locus (data not shown). The direction of the association was opposite to what was observed for periconceptional exposure. We then combined all periconceptional and late pregnancy exposed individuals and their controls in a single analysis to test for a statistical interaction between the famine

Table 1. DNA methylation and periconceptional exposure to famine.

Gene locus[1]	Average methylation % (SD)		Within pair difference (Δ%)[2]	Effect size (SD units)[3]	P-value[4]	P-value Bonferroni corrected[6]
IL10	20.8	(6.5)	2.4	0.37	$1.8x10^{-6}$	$2.7x10^{-5}$
GNASAS	48.8	(4.7)	1.1	0.24	$3.1x10^{-6}$	$4.7x10^{-5}$
INSIGF	84.8	(2.6)	-1.6	-0.61	$2.3x10^{-5}$	$3.5x10^{-4}$
LEP	28.6	(4.9)	1.2	0.24	$2.9x10^{-3}$	$4.4x10^{-2}$
MEG3	54.0	(2.4)	0.5	0.21	$8.0x10^{-3}$	0.12
ABCA1	19.9	(4.2)	0.7	0.17	0.017	0.26
ABCA1 meth[5]	36.9	(8.2)	1.7	0.21	$8.2x10^{-4}$	0.012
KCNQ1OT1	30.1	(1.5)	-0.2	-0.16	0.053	n.s.
GRB10	47.2	(4.6)	0.4	0.08	0.091	n.s.
GNASAB	40.3	(5.0)	0.6	0.11	0.092	n.s.
APOC1	16.7	(3.1)	-0.5	-0.17	0.13	n.s.
IGF2R	84.1	(6.9)	-0.7	-0.10	0.29	n.s.
FTO	97.3	(0.8)	-0.5	-0.61	0.28	n.s.
CRH	58.9	(6.0)	0.4	0.07	0.61	n.s.
TNF	9.6	(1.7)	0.1	0.06	0.63	n.s.
NR3C1	4.8	(1.1)	0.0	-0.01	0.79	n.s.

1. Table sorted on P-value.
2. Average absolute difference in DNA methylation between exposed and unexposed siblings.
3. Observed within pair difference divided by the standard deviation in the sibling controls.
4. Two-sided P-value resulting from a linear mixed model accounting for family relations, bisulfite batch and age at blood draw.
5. Separate analysis of methylated CpGs in *ABCA1*. In contrast to the 3' CpGs, 5' CpG dinucleotides (n=13) were methylated. The methylation of the methylated 5' CpGs was highly correlated (R >0.87) amongst themselves, but not with methylation of the 3' CpG dinucleotides, which had little to no methylation (<4.5%).
6. The Bonferroni corrected P-values. Results that were already not significant before Bonferroni correction are shown as "n.s." (non significant).

associations with DNA methylation and the precise gestational timing of the exposure. The DNA methylation differences found for *IL10*, *GNASAS* and *INSIGF* were timing-specific (**Table 2**), but for *LEP* the test for interaction was not significant.

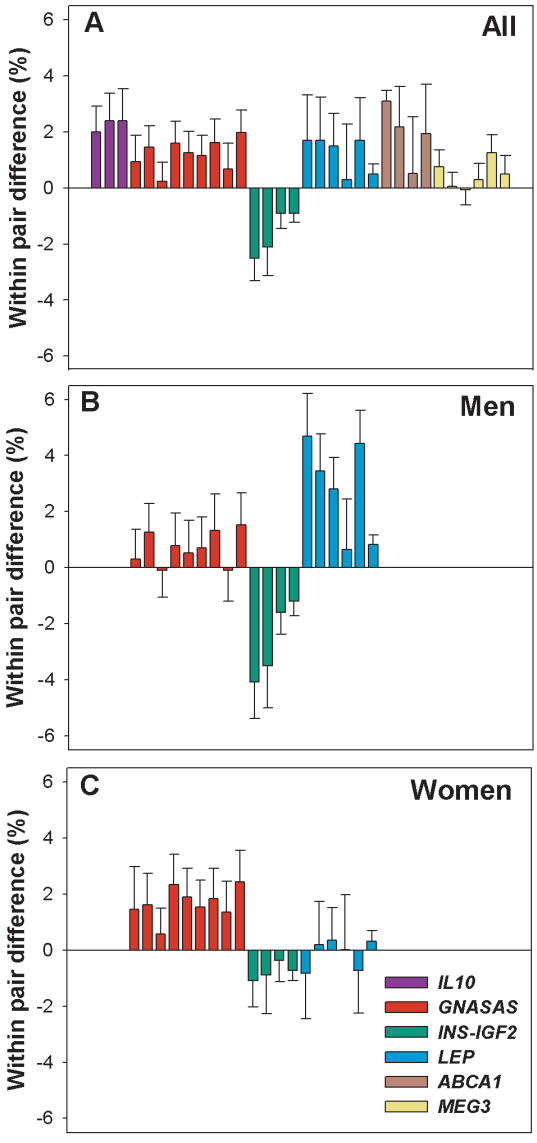


Figure 1. Difference in DNA methylation of CpG dinucleotides in siblings discordant for periconceptional exposure to famine.

Bars in the figures represent the average absolute within pair difference in DNA methylation and their standard errors for CpG dinucleotides. A positive difference indicates a higher methylation level among exposed individuals

A. The absolute within pair difference for CpG dinucleotides for which a significant overall difference of the locus in DNA methylation was observed.

B. As in panel A but for men. Only loci showing a significant interaction between sex and exposure are depicted.

C. As in panel B, but for women only.

Table 2. DNA methylation and late gestational exposure to famine.

Gene locus[1]	Average methylation % (SD)[2]	Within pair difference (Δ%)[3]	Effect size (SD units)[4]	P-value[5]	P-value Bonferroni corrected[6]	P-value[7] timing specificity
IL10	20.7 (5.0)	-0.2	-0.04	0.76	n.s.	$1.2x10^{-3}$
GNASAS	48.8 (4.2)	-1.1	-0.26	$1.1x10^{-7}$	$8.8x10^{-7}$	$3.1x10^{-12}$
INSIGF	84.7 (2.8)	0.0	0.0	0.95	n.s.	$3.2x10^{-4}$
LEP	28.7 (4.6)	0.4	0.09	0.18	n.s.	0.13
KCNQ1OT1	30.2 (1.7)	0.2	0.12	0.17	n.s.	0.058
APOC1	16.7 (3.1)	-0.6	-0.19	0.22	n.s.	0.90
IGF2R	84.0 (4.4)	0.0	0.0	0.88	n.s.	0.25
CRH	58.9 (4.8)	0.5	0.10	0.51	n.s.	0.86

1. This table has been given the same order as table 1.
2. The batch corrected average methylation for the unexposed late gestation sibling controls and the standard deviation.
3. Average absolute difference in DNA methylation between exposed and unexposed siblings.
4. Observed within pair difference divided by the standard deviation in the sibling controls.
5. Two-sided P-value resulting from a linear mixed model accounting for family relations, bisulfite batch and age at blood draw. This test was performed on the late gestational exposed sibships (N=62)
6. The Bonferroni corrected P-values. Results that were already not significant before Bonferroni correction are shown as “n.s.” (non significant).
7. Two-sided P value resulting from the test for timing specificity. The timing specificity was calculated by joining the datasets for both the periconceptional and late gestational siblings and their unexposed same-sex siblings and introducing an interaction term for gestational timing times exposure status. This test thus includes all 122 pairs.

Table 3. Sex-specific associations of DNA methylation with periconceptional exposure to famine.

Gene locus	P sex interaction[1]	Sex	Within pair difference (Δ%)[2]	Effect size (SD units)[3]	P-value[4]
GNASAS	0.027	♂	0.7	0.15	0.013
		♀	1.5	0.33	$1.1x10^{-5}$
INSIGF	$8.5x10^{-3}$	♂	-2.6	-0.99	$6.5x10^{-6}$
		♀	-0.8	-0.29	0.11
LEP	$2.3x10^{-4}$	♂	2.8	0.57	$3.6x10^{-7}$
		♀	-0.2	-0.04	0.70

1. The two sided P-value from the test for sex-specificity of the observed periconceptional effect of prenatal famine. This was tested by entering an interaction term of sex times the exposure status in the linear mixed model.
2. Average absolute difference in DNA methylation between exposed and unexposed siblings.
3. Observed within pair difference divided by the standard deviation in the sibling controls.
4. Two-sided P-value resulting from a linear mixed model accounting for family relations, bisulfitebatch and age at blood draw.

Sex-specific associations

Previous work found basal DNA methylation differences between men and women[26]. We therefore first tested for sex differences in DNA methylation for the measured loci combining the two sibling control groups (men N=56, women N=66). DNA methylation was higher in men for *IGF2R* (+2.6%, P=0.019) and lower in men for *LEP* (-2.6%, $P=3.0x10^{-3}$), *IL10* (-2.9%, P=0.015), and *APOC1* (-1.5, P=0.015) as compared to women. No significant differences were found for the other 4 loci (*GNASAS*, *INSIGF*, *KCNQ1OT1* and *CRH*).
Next, we tested if the observed significant associations with prenatal famine were sex-specific. The interaction between sex and periconceptional famine exposure was significant for *LEP* ($P=2.3x10^{-4}$), *INSIGF* ($P=8.5x10^{-3}$) and *GNASAS* (P=0.027) (**Table 3**). For *LEP* and *INSIGF* the association of famine exposure with DNA methylation was restricted to men ($P_{LEP, ♂}=3.6x10^{-7}$, $P_{INSIGF, ♂}=6.5x10^{-6}$) (**Figure 1B and 1C**). For *GNASAS* the association was significant in both sexes but most pronounced in women ($P_{men}=0.013$; $P_{women}=1.1x10^{-5}$). The association between late gestational famine exposure and *GNASAS* methylation was independent of sex.

Timing independent association

For *LEP* there was no indication for a significant interaction between the famine association with DNA methylation and the gestational timing of the exposure, even though the methylation difference was significant only following periconceptional famine exposure. Since, this association was later found to be male-specific, we tested for an interaction with sex in the late exposure group. Indeed, a significantly higher *LEP* methylation was found for men exposed late in gestation (P=0.017). Analysis of the whole cohort (N=244) including both exposure groups, revealed a significant association between prenatal exposure to famine irrespective of the precise gestational timing (P=0.003). Further analysis suggested that this association was male-specific ($P_{interaction}$= $1.3x10^{-6}$; men: +2.2% (0.83SD), P=$7.5x10^{-8}$; women: P=0.47).

Discussion

We studied the DNA methylation levels of 15 loci for their association with prenatal exposure to the Dutch Famine at the end of WWII. For six of the loci studied, we observed significant differences in DNA methylation after famine exposure during periconception (*INSIGF*, *GNASAS*, *MEG3, IL10*, *LEP* and *ABCA1*). This association differed by sex for three loci (*INSIGF*, *GNASAS* and *LEP*). Of the eight loci tested, exposure to famine late in gestation was associated with methylation for *GNASAS* and for *LEP,* which was specific for men. Of interest, the differences in DNA methylation included both increases and decreases, in one case even at the same locus after exposure during different gestational periods. Together with our previous finding that the *IGF2* DMR is associated with periconceptional exposure[20] our data indicate that an adverse prenatal environment may trigger widespread and persistent changes in DNA methylation.

Our current and previous[20] observations suggest that the periconceptional period may be an especially sensitive exposure period in humans. This might be inherent to mammalian development[27,28] and this hypothesis is also

supported by detailed animal studies[29,30]. The association of *GNASAS* and *LEP* with late gestational exposure, however, suggests that environmentally induced DNA methylation changes may not be limited to the periconceptional period. This is in line with findings that methylation of the glucocorticoid receptor promoter depends on postnatal circumstances (e.g., maternal care in rats[25] and child abuse in humans[24]). In addition to timing-specific associations, we observed sex-specific associations for three of the six loci for which DNA methylation was significantly associated with prenatal exposure to famine. In men, *LEP* methylation was associated with prenatal famine irrespective of the timing of exposure. Our observation that the methylation changes in relation to the prenatal environment may be sex-specific is in agreement with the sex-specific methylation changes found in offspring of sheep that were folate and vitamin B_{12} restricted during periconception[30]. How such sex-specific associations can arise is currently unknown, but interactions between sex hormones and the expression of DNA methyltransferases may be a factor[31].

The differences in DNA methylation observed here are comparable, although slightly smaller, than we previously found for *IGF2* (absolute difference of 2.7% (5.2% relative to the mean DNA methylation level in the population)[20]. The smaller average differences may be related to the inherent stochastic nature of epigenetic processes[32] leading to a large variability in responses. The stochastic nature is strikingly illustrated by the large variation observed in the response of *agouti* gene methylation on maternal methyl donor supplementation even though the mice are inbred and the environmental conditions are highly controlled[17]. In human studies, genetic and environmental heterogeneity may further obscure the full impact of prenatal famine. In our study, we tried to minimize the heterogeneity caused by these factors by comparing exposed individuals with their unexposed, same-sex siblings. Another potential source of heterogeneity is the cellular diversity of whole blood, the tissue we studied. This heterogeneity is less likely to play a role for the affected imprinted loci *IGF2*[20], *GNASAS* and *MEG3*, since methylation is generally cell-type independent for imprinted loci[33] and the observed differences for non-imprinted loci between exposed individuals and controls

were similar. Indeed, with the exception of *IL10* we previously found that cell heterogeneity had little to no influence on the DNA methylation levels of the reported loci[22].

With respect to our finding that *GNASAS* methylation was associated with exposure to famine late in gestation, it should be mentioned that a compromised late gestational development was hypothesized to lead to an immature immune system. The absolute numbers of lymphocytes were reported to be lower in children with a shorter gestation in a large, population-based study[34]. However, since the proportions of the different cell types were not affected, it is unlikely that this has contributed to our findings.

The changes observed are comparable to those found in the liver of rats prenatally exposed to a protein-deficient diet, where promoter methylation of the *Pparα* promoter was decreased from 6.1% to 4.5% with individual CpG dinucleotides affected up to 5% and explaining up to 43% of the variance in gene expression of *Pparα*[35]. This small absolute decrease of *Pparα* DNA methylation corresponded to a large ~26% relative change. It may be hypothesized that modest absolute changes in DNA methylation may lead to significant changes in gene expression for loci with a relatively low methylation level so that relative changes are substantial. The larger relative changes of *LEP* in men (2.8/27.1=10.3%) and *IL10* (2.4/20.8=11.5%) may be more likely to have functional consequences than that for *INSIGF* in men (2.6/84.8=3.1%).

Similar to animal studies of methionine restriction[30], we observed both increases and decreases in DNA methylation, depending on the locus studied. Our results cannot be readily explained by damage due to a deficiency in dietary methyl donors due to the famine and may thus be part of an adaptive response. To definitely prove or disprove the existence of an adaptive response it will be necessary to characterize epigenetic responses of entire relevant pathways. Epigenetic mechanisms may contribute to the development of a thrifty phenotype[1,15] and sex differences in this respect are increasingly well described[13]. It has been hypothesized that a thrifty phenotype may result from the combined effects of smaller epigenetic changes across the genome shifting metabolic networks[36]. *INS*[37] and *LEP*[36], both affected by

prenatal famine exposure, were suggested to be particular relevant in this respect. Our observations provide empirical evidence for these hypotheses in humans.

Our finding that the association of DNA methylation with prenatal conditions may depend on timing and sex matches the specificity in phenotypic outcomes observed for prenatal famine exposure, including neonatal outcomes[9] and psychiatric[7] and metabolic traits[4,6,10]. It remains to be determined if the observed differences in DNA methylation in blood mark functional differences in relevant tissues for these traits. Previous work on *IGF2* DMR has shown that methylation in blood can mark the methylation level in a relevant tissue[38]. Also, the CpG sites studied in the *LEP* promoter were previously reported to show similar methylation in peripheral blood and adipocytes *in vivo*[39] and to influence *LEP* expression *in vitro*[40]. Changes in DNA methylation induced early in development, like by prenatal famine, may be propagated soma-wide contributing to the observed correlation in DNA methylation across tissues[17,29]. A preliminary analysis of our data did not reveal significant associations between DNA methylation and plasma lipids and BMI that we previously reported to depend on prenatal exposure to famine in women in this cohort[4,6]. But DNA methylation at most genes relevant to these particular complex phenotypes was found to be affected in men only, and not women. Similar to genetic studies, epigenetic studies linking variation in DNA methylation and complex disease will most likely require large study samples. Such studies may be particularly promising for *IL10,* which we found to be sensitive to periconceptional famine exposure. Genetic variation influencing *IL10* expression at this locus is associated with schizophrenia[41,42], a phenotype also particular to the periconceptional period[7].

In summary, our study shows that exposure to famine in pregnancy may cause persistent changes in DNA methylation levels of multiple imprinted and non-imprinted genes with diverse biological functions. Our data support the hypothesis that associations between early developmental conditions and health outcomes later in life may be mediated by changes in the epigenetic information layer. Understanding how disturbances early in

human development are linked to later life disease may suggest new ways to prevent disease.

Materials and Methods

Study population

Study participants are a subset of the population from our ongoing Hunger Winter Families Study, whose design and recruitment was described previously[21]. In short, this study includes individuals exposed to famine during different periods in pregnancy and time-controls born before or after the famine. Study subjects were selected from births between 1943-1947 at three institutions in famine-exposed cities (the midwifery training schools in Amsterdam and Rotterdam and the Leiden University Medical Center). The daily rations distributed by the authorities during the famine period from November 28, 1944 to May 15, 1945 period contained an average energy equivalent of 667 kcal (SD, 151). During the famine there was little variation in the percentage of calories from proteins (12%), fat (19%) and carbohydrates (69%)[43]. In addition to time controls from these three institutions, we recruited whenever possible a same-sex unexposed sibling of these individuals to serve as sibling control. Clinical examinations, including blood collections, were completed for 311 births in these institutions and 311 sibling controls. Ethical approval for the study was obtained from the participating institutions, and all participants provided written informed consent.

The subset selected for the present study includes two exposure groups with their unexposed sibling controls. The first exposure group comprises births conceived during the height of the famine, with periconceptional and early pregnancy exposure. The second exposure group comprises births during the height of the famine, with late pregnancy exposure. Periconceptionally exposed individuals were defined as births with a mother's last menstrual period between November 28, 1944 and May 15, 1945. This group includes 60 individuals (age 58.1y, SD 0.35), of whom 28 were male. Individuals exposed in late pregnancy were defined as births between January 28 and

May 30, 1945. This group includes 62 individuals (age 58.8y, SD 0.40), of whom 28 were male. As controls for each exposure group we used a matched same-sex sibling (age 57.1y, SD 5.50). The total study population therefore includes 244 individuals.

DNA methylation measurements

Genomic DNA was isolated from whole blood using the salting out method. One microgram of genomic DNA was bisulfite treated using the EZ 96-DNA methylation kit (Zymo Research). All samples were bisulfite treated on a total of three 96-well plates. Sibling pairs were on the same plate and periconceptionally and last trimester exposed pairs were equally distributed over the plates. DNA methylation of CpG dinucleotides were measured by a mass spectrometry based method (Epityper, Sequenom). The method determines the amount of DNA methylation by interrogating thousands of DNA copies assuming that 1 ng of genomic DNA equals ~300 copies. The quantitative nature, accuracy and reproducibility of this method has been shown extensively[20,44,45]. All biochemical steps inherent to the methodology were performed according to the manufacturers' protocol. Bisulfite converted DNA specific PCR primers used to amplify the 15 investigated regions are summarized in **Supplement I, Table S4**. This table includes the precise genomic location of the amplified regions and an overview of the CpG sites quantized in each amplicon. All individuals exposed early in gestation and their sibling controls were measured in triplicate on a single 384 well plate for each locus and the same was true for the late gestational exposed individuals and their sibling controls. Data quality control and filtering were done as previously described[46]. Data filtering consisted of the removal of CpG dinucleotides of which the measurement could be confounded by single nucleotide polymorphisms and of CpG dinucleotides of which the measurement success rate was below 80%. Common causes of a lower success rate include fragments bordering on the upper and lower limits of the mass range that can be detected and cases of fragments of which the base of the peak signal in the mass spectrum overlapped another fragment.

The success rate for the CpG containing fragments that could be measured within the limits of the methodology was 93.3%.

Statistical analysis

The analyses were performed within sib ships to minimize the possible confounding effects of differences in familial environment and genetic. We applied linear mixed models on the raw data without imputation of missing values to calculate exposure specific differences between sibling pairs. All the analyses account for family relations, age at examination, bisulfite plate and the correlation between CpG dinucleotides. Sib ship was entered as a random effect. Age at examination, bisulfite plate, exposure status and CpG dinucleotide, and where appropriate sex, were entered as fixed effects. The test for the timing specificity of the association between exposure and DNA methylation levels was done by adding a variable indicating the timing of the exposure and merging the periconceptional and last trimester datasets: timing specificity was tested by adding an interaction term of exposure status times the timing of the exposure to the linear mixed model. Testing the sex specificity of the associations between famine exposure and DNA methylation levels was done by adding a term for sex times the exposure status of the individuals in the linear mixed model.

The linear mixed model was chosen over a standard paired t-test because it allows for the analysis of multiple individual CpG dinucleotides in one test, accounts for the correlation between adjacent CpG dinucleotides, includes relevant adjustments within the model on the raw data, and uses all available data. The linear mixed model reduces to a test with identical outcome to a paired t-test if the within family difference is assessed for a single CpG dinucleotide, if no adjustments are performed and if there are no incomplete data for the sib pairs.

The basal difference in DNA methylation between men and women in the controls was calculated using ANOVA. All P-values reported are two-sided. All analyses were performed using SPSS 16.0.

Acknowledgements

We thank the participants of the Hunger Winter Families Study, TNO Quality of Life for contact tracing, the staff of the Gerontology and Geriatrics Study Center at the Leiden University Medical Center for performing the clinical examinations and the Central Clinical Chemical Laboratory for extracting genomic DNA. This work was supported by grants from the Netherlands Heart Foundation (2006B083 to B.T.H); the U.S. National Institutes of Health (R01-HL067914 to L.H.L); The Netherlands Organization for Scientific Research NWO (911-03-016 to P.E.S); and the European Union funded Network of Excellence LifeSpan (FP6 036894). Dr. Lumey was supported by a 2008-2009 Lorentz fellowship from the Royal Netherlands Academy of Sciences.

References

1. Hales CN, Barker DJ. The thrifty phenotype hypothesis. Br.Med.Bull. 2001; 60: 5-20
2. Barker DJ. The origins of the developmental origins theory. J.Intern.Med. 2007; 261: 412-7
3. Whincup PH, Kaye SJ, Owen CG, Huxley R, Cook DG, Anazawa S, et al. Birth weight and risk of type 2 diabetes: a systematic review. JAMA 2008; 300: 2886-97
4. Lumey LH, Stein AD, Kahn HS, Romijn JA. Lipid profiles in middle-aged men and women after famine exposure during gestation: the Dutch Hunger Winter Families Study. Am.J.Clin.Nutr. 2009; 89: 1737-43
5. Roseboom T, de RS, Painter R. The Dutch famine and its long-term consequences for adult health. Early Hum.Dev. 2006; 82: 485-91
6. Stein AD, Kahn HS, Rundle A, Zybert PA, van der Pal-de Bruin, Lumey LH. Anthropometric measures in middle age after exposure to famine during gestation: evidence from the Dutch famine. Am.J.Clin.Nutr. 2007; 85: 869-76
7. Susser E, Neugebauer R, Hoek HW, Brown AS, Lin S, Labovitz D, et al. Schizophrenia after prenatal famine. Further evidence. Arch. Gen.Psychiatry 1996; 53: 25-31
8. Kyle UG, Pichard C. The Dutch Famine of 1944-1945: a pathophysiological model of long-term consequences of wasting disease. Curr.Opin.Clin.Nutr.Metab Care 2006; 9: 388-94
9. Brown AS, Susser ES. Sex differences in prevalence of congenital neural defects after periconceptional famine exposure. Epidemiology 1997; 8: 55-8
10. Ravelli AC, van Der Meulen JH, Osmond C, Barker DJ, Bleker OP. Obesity at the age of 50 y in men and women exposed to famine prenatally. Am.J.Clin.Nutr. 1999; 70: 811-6
11. Bygren LO, Edvinsson S, Brostrom G. Change in food availability during pregnancy: Is it related to adult sudden death from cerebro- and cardiovascular disease in offspring? Am.J.Hum.Biol. 2000; 12: 447-53
12. Painter RC, de R, Sr., Bossuyt PM, Simmers TA, Osmond C, Barker DJ, et al. Early onset of coronary artery disease after prenatal exposure to the Dutch famine. Am.J.Clin.Nutr. 2006; 84: 322-7
13. Gilbert JS, Nijland MJ. Sex differences in the developmental origins of hypertension and cardiorenal disease. Am.J.Physiol Regul. Integr.Comp Physiol 2008; 295: R1941-R1952
14. Nathanielsz PW. Animal models that elucidate basic principles of the developmental origins of adult diseases. ILAR.J. 2006; 47: 73-82
15. Gluckman PD, Hanson MA, Cooper C, Thornburg KL. Effect of in utero and early-life conditions on adult health and disease. N.Engl.J.Med. 2008; 359: 61-73
16. Waterland RA, Michels KB. Epigenetic epidemiology of the developmental origins hypothesis. Annu.Rev.Nutr. 2007; 27: 363-88
17. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol.Cell Biol. 2003; 23: 5293-300
18. 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: 401-6
19. Bogdarina I, Welham S, King PJ, Burns SP, Clark AJ. Epigenetic modification of the renin-angiotensin system in the fetal programming of hypertension. Circ.Res. 2007; 100: 520-6
20. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc. Natl.Acad.Sci.U.S.A 2008; 105: 17046-9
21. Lumey LH, Stein AD, Kahn HS, van der Pal-de Bruin KM, Blauw GJ, Zybert PA, et al. Cohort profile: the Dutch Hunger Winter families study. Int.J.Epidemiol. 2007; 36: 1196-204
22. Talens RP, Boomsma DI, Tobi EW, Kremer D, Jukema JW, Willemsen G, et al. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. FASEB J. 2010; 24: 3135-44
23. Monk D, Sanches R, Arnaud P, Apostolidou S, Hills FA, bu-Amero S, et al. Imprinting of IGF2 P0 transcript and novel alternatively spliced

INS-IGF2 isoforms show differences between mouse and human. Hum.Mol.Genet. 2006; 15: 1259-69

24. McGowan PO, Sasaki A, D'Alessio AC, Dymov S, Labonte B, Szyf M, et al. Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. Nat.Neurosci. 2009; 12: 342-8

25. Weaver IC, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR, et al. Epigenetic programming by maternal behavior. Nat.Neurosci. 2004; 7: 847-54

26. El-Maarri O, Becker T, Junen J, Manzoor SS, az-Lacava A, Schwaab R, et al. Gender specific differences in levels of DNA methylation at selected loci from human total blood: a tendency toward higher methylation levels in males. Hum.Genet. 2007; 122: 505-14

27. Morgan HD, Santos F, Green K, Dean W, Reik W. Epigenetic reprogramming in mammals. Hum.Mol.Genet. 2005; 14 Spec No 1: R47-R58

28. Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. Science 2001; 293: 1089-93

29. Morgan HD, Jin XL, Li A, Whitelaw E, O'Neill C. The culture of zygotes to the blastocyst stage changes the postnatal expression of an epigentically labile allele, agouti viable yellow, in mice. Biol.Reprod. 2008; 79: 618-23

30. Sinclair KD, Allegrucci C, Singh R, Gardner DS, Sebastian S, Bispham J, et al. DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and methionine status. Proc.Natl.Acad.Sci.U.S.A 2007; 104: 19351-6

31. Yamagata Y, Asada H, Tamura I, Lee L, Maekawa R, Taniguchi K, et al. DNA methyltransferase expression in the human endometrium: down-regulation by progesterone and estrogen. Hum.Reprod. 2009; 24: 1126-32

32. Bjornsson HT, Fallin MD, Feinberg AP. An integrated epigenetic and genetic approach to common human disease. Trends Genet. 2004; 20: 350-8

33. Reik W, Walter J. Genomic imprinting: parental influence on the genome. Nat.Rev. Genet. 2001; 2: 21-32

34. Duijts L, Bakker-Jonges LE, Labout JA, Jaddoe VW, Hofman A, Steegers EA, et al. Fetal growth influences lymphocyte subset counts at birth: the Generation R Study. Neonatology. 2009; 95: 149-56

35. Lillycrop KA, Phillips ES, Torrens C, Hanson MA, Jackson AA, Burdge GC. Feeding pregnant rats a protein-restricted diet persistently alters the methylation of specific cytosines in the hepatic PPAR alpha promoter of the offspring. Br.J.Nutr. 2008; 100: 278-82

36. Stoger R. The thrifty epigenotype: an acquired and heritable predisposition for obesity and diabetes? Bioessays 2008; 30: 156-66

37. Prentice AM, Rayco-Solon P, Moore SE. Insights from the developing world: thrifty genotypes and thrifty phenotypes. Proc.Nutr. Soc. 2005; 64: 153-61

38. Cui H, Cruz-Correa M, Giardiello FM, Hutcheon DF, Kafonek DR, Brandenburg S, et al. Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. Science 2003; 299: 1753-5

39. Stoger R. In vivo methylation patterns of the leptin promoter in human and mouse. Epigenetics. 2006; 1: 155-62

40. Melzner I, Scott V, Dorsch K, Fischer P, Wabitsch M, Bruderlein S, et al. Leptin gene expression in human preadipocytes is switched on by maturation-induced demethylation of distinct CpGs in its proximal promoter. J.Biol.Chem. 2002; 277: 45420-7

41. Chiavetto LB, Boin F, Zanardini R, Popoli M, Michelato A, Bignotti S, et al. Association between promoter polymorphic haplotypes of interleukin-10 gene and schizophrenia. Biol. Psychiatry 2002; 51: 480-4

42. Ozbey U, Tug E, Namli M. Interleukin-10 gene promoter polymorphism in patients with schizophrenia in a region of East Turkey. World J.Biol.Psychiatry 2009; 1-8

43. Burger GCE, Drummond JC, Sandstead HR. Malnutrition and Starvation in Western Netherlands, September 1944-July 1945. The Hague: General State Printing Office, 1948

44. Coolen MW, Statham AL, Gardiner-Garden M, Clark SJ. Genomic profiling of CpG methylation and allelic specificity using quantitative high-throughput mass spectrometry: critical evaluation and improvements. Nucleic Acids Res. 2007; 35: e119

45. Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. Proc.Natl.Acad.Sci.U.S.A 2005; 102: 15785-90

46. Heijmans BT, Kremer D, Tobi EW, Boomsma DI, Slagboom PE. Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus. Hum.Mol.Genet. 2007; 16: 547-54

DNA methylation of IGF2, GNASAS, INSIGF and LEP and being born small for gestational age

Elmar W. Tobi[1], Bastiaan T. Heijmans[1,5], Dennis Kremer[1], Hein Putter[2], Henriette A. Delemarre-van de Waal[3], Martijn J.J. Finken[3,4], Jan M. Wit[3], P. Eline Slagboom[1,5]

1. Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
2. Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands
3. Pediatrics, Leiden University Medical Center, Leiden, The Netherlands.
4. Pediatrics, VU medical center, Amsterdam, The Netherlands.
5. The Netherlands Consortium for Healthy Ageing, The Netherlands

Epigenetics. **2011** Feb:6(2):171-176

Abstract

Being born small for gestational age (SGA), a proxy for intrauterine growth restriction (IUGR), and prenatal famine exposure are both associated with a greater risk of metabolic disease. Both associations have been hypothesized to involve epigenetic mechanisms. We investigated whether prenatal growth restriction early in pregnancy was associated with changes in DNA methylation at loci that were previously shown to be sensitive to early gestational famine exposure. We compared 38 individuals born preterm (<32 weeks) and with a birth weight too low for their gestational age (<-1SDS; "SGA") with 75 individuals born preterm but with a birth weight appropriate for their gestational age (>-1SDS) and a normal postnatal growth (>-1SDS at 3 months post term; "AGA"). The SGA individuals were not only lighter at birth, but also had a smaller length ($P=3.3x10^{-13}$) and head circumference at birth ($P=4.1x10^{-13}$). The DNA methylation levels of *IGF2*, *GNASAS*, *INSIGF* and *LEP* were 48.5%, 47.5%, 79.4% and 25.7% respectively. This was not significantly different between SGA and AGA individuals. Risk factors for being born SGA, including preeclampsia and maternal smoking, were also not associated with DNA methylation at these loci. Growth restriction early in development is not associated with DNA methylation at loci shown to be affected by prenatal famine exposure. Our and previous results by others indicate that prenatal growth restriction and famine exposure may be associated with different epigenetic changes or non epigenetic mechanisms that may lead to similar later health outcomes.

Introduction

The developmental origins hypothesis states that adverse environmental conditions during specific time windows of mammalian development can have a lasting effect on metabolic pathways and physiology influencing chronic disease susceptibility[1]. Intrauterine growth restriction (IUGR) is considered to be the result of a poor intrauterine environment and has been associated with diverse adverse health outcomes later in life, including type 2 diabetes and hypertension[2-5]. More than 5% of all pregnancies in the Western world result in infants being born small for gestational age (SGA)[6], an often used proxy for IUGR[7].

In animal models IUGR is modeled by inducing placental insufficiency by artificially reducing placental perfusion or by limiting the maternal nutrients supply with protein or caloric restriction[8]. In humans, SGA is associated with both placental insufficiency and suboptimal prenatal nutrition. For instance, preeclampsia which changes placental perfusion[9] is one of the major risk factor for SGA[10] and the risk to develop preeclampsia is reduced by early gestational micronutrient supplementation[11]. Furthermore, micronutrient supplementation during pregnancy has been found to increase birth weight and the risk for severe SGA is decreased by iron-folic acid intake alone[12,13].

The induction of persistent epigenetic change by prenatal environmental conditions may be a mechanism contributing to the associations between early development and later life health in humans. For example, extensive work in animals has shown that placental insufficiency or restricting the maternal diet of protein, folic acid or other micronutrients can persistently alter DNA methylation and other epigenetic marks and may contribute to the development of diabetes and hypertension[14-19]. In humans, periconceptional exposure (e.g. around conception and the first trimester) to the Dutch Famine, a famine at the end of WWII, is associated with persistent differences in DNA methylation of various important loci involved in growth and metabolism, including *IGF2*, *GNASAS, INSIGF* and *LEP*[20,21]. Further work indicated that *IGF2* methylation is also sensitive to maternal periconceptional folic acid use[22].

These loci are also relevant in relation to prenatal growth restriction. *IGF2* is a major driver of embryonic growth[23] and in concordance with this role genetic variation on the paternal allele of the *IGF2-INS* region was found to influence the risk of being born SGA[24]. *GNAS* and *LEP* have both been found to be differentially expressed between placentas of IUGR and normal children[25]. Both the *GNAS* region and leptin are similarly involved in early growth and glucose metabolism as the *IGF2-INS* region[26,27].

SGA and prenatal famine exposure are associated with similar later life phenotypic consequences[3,28], but it is unclear to what extent this is due to the same mechanism. Here we investigated whether growth restriction during early and/or mid gestation (<32 weeks) is associated with differences in DNA methylation at these 4 loci we found to be sensitive to environmental conditions early in development. We selected individuals from the Dutch nation-wide Project On Preterm and Small-for-gestational age infants (POPS) cohort[29] and measured DNA methylation of the *IGF2* differentially methylated region (DMR) and *INSIGF*, *GNASAS* and *LEP* promoters. We compared levels of DNA methylation between preterm born SGA individuals with individuals born preterm but with a birth weight appropriate for their gestational age and with a normal postnatal growth. In addition we explored possible associations with major risk factors for SGA, including preeclampsia and prenatal smoking.

Results

Child, pregnancy and maternal characteristics

In this study, 38 individuals born pre-term and SGA and 75 individuals born preterm but with a birth weight appropriate for their gestational age (AGA), were compared. The SGA individuals were not only small in terms of birth weight, the selection criterion, but also smaller in terms of birth length and head circumference at birth (**Table 1**), compatible with IUGR[30]. The SGA

Table 1. Characteristics at birth and pregnancy.

Characteristics		SGA	AGA	P_{value} [1]
Number of individuals	N	38	75	
Male	%	39.5	44.0	0.65
Adult height	SDS[2](SD)	-1.02(0.99)	-0.17(0.97)	$3.3x10^{-5}$
At Birth				
Gestational age at birth	weeks(SD)	30.6 (1.1)	30.1(1.5)	0.053
Birth weight	SDS[3](SD)	-1.86(0.50)	0.31(0.73)	$1.1x10^{-31}$
Birth head circumference	SDS(SD)	-1.34(0.79)	0.20(0.92)	$4.1x10^{-13}$
Birth length	SDS(SD)	-1.83(0.87)	0.19(1.12)	$3.3x10^{-13}$
Obstetric data				
First child (parity)	%	68%	52%	0.097
Maternal age	years(SD)	27.8(4.8)	27.9(5.9)	0.92
Problematic obstetric history[4]	%	18.4	17.3	0.89
Socio-economic status[5]	SES	3.53(1.50)	3.59(1.55)	0.82
Maternal height	cm(SD)	165.6(5.4)	167.3(6.1)	0.14
Maternal diabetes mellitus	%	5.3	4.0	0.57
Chorioamnionitis[6]	%	5.3	25.3	0.009
Smoking during pregnancy	%	28.9	35.7	0.029
Pre-existing hypertension	%	10.7	2.7	0.08
Preeclampsia	%	65.8	10.7	$3.4x10^{-11}$

1. P value resulting from an unpaired T-Test between the SGA and AGA groups.
2. Standard deviation score from the reference population mean.
3. The birth weight in grams (SD) for the SGA and AGA groups were 963(149) and 1508(301) respectively.
4. Percentage of mothers with previous pregnancies and/or births with complications
5. Socio-economic status of the family given on a 1 to 6 scale. With 1 being the poorest score and 6 being the highest score
6. Intrauterine infection followed by a prolonged rupture of the membrane and preterm labor

individuals remained relatively short (-1.02 SDS), while the AGA group was similar in adult height to the Dutch reference values (-0.17 SDS). The greatest differences in the obstetric data for an SGA child as compared to an AGA child were a higher prevalence of preeclampsia (65.8% vs. 10.7%) and a lower occurrence of chorioamnionitis, an intrauterine infection followed by a prolonged rupture of the membrane and preterm labor (5.3% vs. 25.3%). In addition, smoking during pregnancy was less common in the pregnancies leading to a SGA child (28.9% vs. 35.7%).

Comparison between SGA and AGA

We measured gene specific methylation for *IGF2*, *GNASAS*, *INSIGF* and *LEP* in whole blood. The average DNA methylation levels were 48.5%, 47.5%, 79.4% and 25.7% for *IGF2*, *GNASAS*, *INSIGF* and *LEP*, respectively (**Table 2**). DNA methylation levels in the SGA group were not significantly different from the AGA group. The results were similar for the individual CpG dinucleotides (*data not shown*). The variance in DNA methylation was also not significantly different between the two groups (Levene's test P>0.14). We repeated the analyses using other frequently used cut-offs for birth weight SDS scores to define growth restriction. Using a cut-off of < -1.3 SDS (the 10th percentile, N=34 vs. N=75) or < -2SDS (frequently used by pediatric

Table 2. Methylation difference between SGA and AGA.

	AGA (SD)	SGA - AGA[1]	P_{value}[2]
IGF2	48.5% (3.5)	-0.2%	0.81
GNASAS	47.5% (4.6)	-0.7%	0.41
INSIGF	79.4% (3.2)	-0.2%	0.78
LEP	25.7% (5.3)	-1.3%	0.24

1. The difference in DNA methylation between the small for gestational age and appropriate for gestational age groups. A negative difference means that the SGA group has a lower methylation level.
2. A two-sided P-value resulting from a linear mixed model corrected for the correlation between individual CpG dinucleotides, bisulfite batch and sex between the SGA and AGA groups.

endocrinologists, N=13 vs. N=75) and stratifying all performed analyses by sex did not change the outcome (*data not shown*).

Preeclampsia and other risk factors

The risk to develop preeclampsia is influenced by nutrition in the same period of gestation[11] as our previous studies[20-22]. To reduce the influence of heterogeneity, we first restricted our analysis to the individuals born SGA after a pregnancy with preeclampsia with those born AGA and without (25 vs. 67). No significant differences were found for these loci (data not shown). Next we tested for an association in all measured individuals between DNA methylation and the factors with the greatest difference between the SGA and AGA groups. Preeclampsia and maternal smoking during pregnancy were not associated with DNA methylation at these loci (**Table 3**). A nominally significant association was observed for *LEP* and chorioamnionitis (P=0.033), which would no longer be significant after accounting for the number of tests performed. Factors reported to be associated with increased risk of developing a SGA child, but not found in the current study, namely gestational age, a first pregnancy and maternal hypertension before pregnancy, were not associated with DNA methylation.

Table 3. The relation between DNA methylation and risk factors.

	IGF2		*GNASAS*		*INSIGF*		*LEP*	
	β[1]	P[2]	*β*	P	*β*	P	*β*	P
Preeclampsia	0.8%	0.34	0.9%	0.29	0.0%	0.99	-0.2%	0.88
Chorioamnionitis	0.8%	0.36	0.8%	0.42	0.5%	0.52	2.8%	0.033
Smoking[3]	-1.5%	0.054	-0.9%	0.29	-0.4%	0.53	-1.9%	0.98

1. The beta from a linear mixed model corrected for the correlation between individual CpG dinucleotides, bisulfite batch and sex. The investigated variable was entered as a fixed effect.
2. A two-sided P-value resulting from a linear mixed model corrected for the correlation between individual CpG dinucleotides, bisulfite batch and sex.
3. Smoking during pregnancy by the mother.

Discussion

We tested for the association of being born SGA before 32 weeks of gestation with DNA methylation of *IGF2*, *GNASAS*, *INSIGF* and *LEP* genes for which we previously showed an association with prenatal famine exposure and, for *IGF2*, folic acid supplementation[20-22]. We did not observe differences in DNA methylation at these genes between individuals who were born preterm and growth restricted and individuals born preterm but with a weight appropriate for their gestational age and a normal post natal growth. Preeclampsia was also not associated with DNA methylation levels.

The loci tested for DNA methylation differences may be regarded as markers for prenatal nutritional conditions. Our results are compatible with the interpretation that SGA and preeclampsia do not have a nutritional component in our western and thus well nourished cohort. Other studies on individuals born SGA at term also did not find an association with DNA methylation around the *IGF2* locus[31,32]. Our data does not exclude the possibility that a similar study in developing countries would yield different results for the loci studied, as malnutrition can be expected to play a more prominent role in those countries[33].

In western cohorts SGA may more readily be associated with placental insufficiency and an insufficient transfer of oxygen to the child, which is known to contribute to growth restriction and prenatal programming[34] and shown to influence DNA methylation patterns in animal models[35]. Indeed, epigenetic differences may be still be present in humans born SGA, but at other loci than those influenced by prenatal famine, as is suggested by work by Einstein *et al*.[36]. Beside maternal and environmental factors, however, genetic predisposition may play a role. Twin studies show that some of the associations between birth weight and later health are confounded by genetic factors[2,37]. Indeed, genetic variation influencing birth weight also contributes to the risk of diabetes[38] and genetic variation in the glucocorticoid receptor was found to influence both growth and later glucose homeostasis in children born preterm and SGA[39].

The current study focuses on the influence of conditions during early and mid-gestation to account for the observation that DNA methylation at these loci may be less sensitive during late gestation[20-22]. One may consider the possibility that all very pre-term born children irrespective of prenatal growth experienced an adverse development. In that case, DNA methylation changes may have occurred in both groups studied. However, chorioamnionitis, which is a generally more acute complication of pregnancy, was more prevalent among children born AGA and is not associated with DNA methylation at these loci. Furthermore, the height of AGA individuals at 19 years was not different from the Dutch reference values indicating that prenatal birth per se did not compromise postnatal growth. This supports the interpretation that there are persistent phenotypic differences between SGA and AGA individuals born very preterm, which were not explained by differences in DNA methylation at the measured loci. Also, the association between birth weight and cardiovascular disease was found to be independent of gestational length, suggesting a link with prenatal growth and not preterm birth for fetal programming[3]. A comparison of very preterm children with children at term may not be sufficient to solve this issue because of the possible influence of the intensive neonatal treatments on epigenetic marks.

These results, together with the findings by others[31,32,36], suggest that SGA is not associated with similar epigenetic changes as prenatal famine exposure in Western populations. If so, the etiology of the similar later life consequences associated with these early life conditions, diabetes and cardiovascular disease, may be different. More detailed studies of the epigenetic changes associated with human and animal growth restriction are warranted to gain insight into the link between development and disease. Animal models will be important to elucidate the basic principles, but care may have to be taken when extrapolating epigenetic studies to humans, since it may be possible that animal models implementing nutritional restrictions early in gestation may better simulate human famine exposure than IUGR. Studies in humans will require extensive and detailed phenotyping of prenatal growth, maternal and environmental factors and genetic variation. Most likely such studies will require a relatively large initial study size in which homogeneous subselections

can be made to overcome the complexity and variation inherent to clinical cohorts of prenatal growth restricted humans.

Subjects and Methods

Study population

The Dutch Project On Preterm and Small-for-gestational age infants (POPS) is a nation-wide prospective study, encompassing 94% of all live born infants born very preterm (<32 weeks) and/or with a very low birth weight (<1500 gram) in 1983. The recruitment, details of measurements, and physical and psychosocial outcomes have been reported previously in detail[29,40]. The anthropometric data at birth has been transformed into standard deviation scores (SDS) based on the Swedish references for very preterm infants[41]. The Swedish references were chosen because the Dutch references lack data on birth length and head circumference, while being highly similar[42]. All other anthropometric data has been transformed using the Dutch reference values[43]. The study was approved by the medical ethics committees of all participating centers and written informed consent was obtained from all participants.

Selection for current study

From the POPS cohort we had 413 individuals available who were born before <32 weeks of gestation. We excluded non white participants (excluding 53), twins (excluding 86), individuals treated with glucocorticoids (dexa- or beclomethasone) during the prenatal and/or neonatal period (excluding 71) and individuals with chromosomal abnormalities or inborn errors in metabolism (excluding 2). We defined small for gestational age (SGA) as individuals born with a birth weight of <-1 SDS. As a control group we selected individuals with a birth weight >-1 SDS and a weight at 3 months of >-1 SDS (AGA). From the 201 remaining individuals 42 met our SGA and 92 met our

AGA definition. Four SGA and seventeen AGA individuals had not enough genomic DNA available and were excluded. This resulted in a selection of 38 small for gestational age individuals and 75 individuals with a birth weight appropriate for their gestational age and a normal postnatal growth, which extended up to age 19y (height -0.17SDS).

DNA methylation measurements

Genomic DNA was isolated from whole blood drawn at age 19 using the Qiagen mini kit. Half a microgram of genomic DNA was bisulfite treated using the EZ 96-DNA methylation kit (Zymo Research) using the standard over night bisulfite treatment protocol. The 113 individuals were bisulfite treated on two 96-well plates. SGA and AGA individuals were equally distributed on the plates. The distribution of men and women was also similar on the 2 plates. DNA methylation for individual CpG dinucleotides of *IGF2*, *GNASAS*, *INSIGF* and *LEP* was determined by a mass spectrometry based method (Epityper, Sequenom), for which the reproducibility and accuracy has been shown extensively[20,44,45]. Details of the measured amplicons, including details of functional relevance were published before[46]. In short, *IGF2* DMR hypomethylation was associated with biallelic *IGF2* expression[47] and *INSIGF* locus measured is the DMR located in the promoter of the imprinted *INSIGF* transcript which originates from the *INS* promoter[48]. The *GNASAS* amplicon is part of the *GNAS* DMR2 and is located at the proximal promoter of this imprinted RNA antisense transcript of *GNAS*[49] overlapping the binding site of several transcription factors according to ENCODE[50] CHIP-seq data. The *LEP* amplicon also covers the proximal promoter and includes several CpG sites of which the methylation status influences transcription[51]. DNA methylation was measured in samples from 19 year old individuals which were assumed to provide information on potential epigenetic differences induced during prenatal development. The stability of the methylation marks at the four loci investigated during the life course was suggested by their association with prenatal famine 60 decades post exposure[20,21]. In addition, comparing blood samples taken 10-20 years apart indicated

the stability of the methylation of *IGF2* DMR, *LEP* and, to a lesser extent, *INSIGF* (*GNASAS* was not studied)[46]. Data for the four loci was acquired and processed as previously described[20,21,46]. The PCR and the subsequent steps were performed in triplicate and performed according to the manufacturers' protocol. Each locus was measured on the same 384 well plate for all 113 individuals studied. Data quality control and filtering consisted of the removal of triplicate measurements for which less than two measurements were successful or for measurements with a standard deviation larger than 0.1. CpG dinucleotides of which the measurement could be confounded by single nucleotide polymorphisms and CpG dinucleotides of which the success rate after filtering was below 75% were removed. Details about the primers and the CpG sites included and biological relevance is provided in **Supplement I, Table S4 and Table S5**.

Statistics

Unpaired T-tests were used for the analyses of the anthropometric and pregnancy characteristics. We applied linear mixed models on the raw data without imputation of missing values to calculate differences in DNA methylation for each locus between the SGA and AGA groups. All group analyses account for bisulfite plate, sex and the correlation between CpG dinucleotides. Person identifier was added as random effect and bisulfite batch, sex and group identifier (e.g. being SGA or AGA) were entered as fixed effects. The linear mixed model is preferred above more standard tests because it allows the incorporation of multiple individual CpG dinucleotides in one test, accounts for the correlation between adjacent CpG dinucleotides, incorporates the relevant adjustments within the model on the raw data, and uses available but incomplete data for individuals. All analyses were also performed using <-1.3 SDS (the 10th percentile) and <-2 SDS birth weight as cut-offs to define SGA status. The analyses were also performed for individual CpG sites. The test for associations between birth characteristics or risk factors with DNA methylation was performed by adding the respective variable to the mixed model as a fixed effect. To test for differences in the

variance in DNA methylation between the groups we used the Levene's test statistic for homogeneity of variance from the one-way ANOVA test in PASW 17.0. All analyses were performed using PASW Statistics 17.0, previously known as SPSS. All P-values reported are two sided.

Acknowledgements

We thank the participants, staff and members of the POPS-19 Collaborative Study Group*. This work was supported by the European Union-funded Network of Excellence LifeSpan (FP6 036894), The Netherlands Organization for Scientific Research NWO (91103016 to P.E.S), The Netherlands Consortium for Healthy Ageing (NCHA, 05060810) in the framework of The Netherlands Genomics Initiative (NGI)/NWO. The establishment of the POPS study at 19 years of age was supported by grants from the Netherlands Organization for Health Research and Development (ZonMw), the Edgar Doncker Foundation, Foundation for Public Health Fundraising Campaigns, Phelps Foundation, Swart-van Essen Foundation, Foundation for Children's Welfare Stamps, TNO prevention and Health, NWO, Dutch Kidney Foundation, Sophia Foundation for Medical Research, Stichting Astmabestrijding and Royal Effatha Guyot Group.

*Participants of the Dutch POPS-19 collaborative Study Group:
TNO Prevention and Health, Leiden (ETM Hille, CH de Groot, H Kloosterboer-Boerrigter, AL den Ouden, A Rijpstra, SP Verloove-Vanhorick, JA Vogelaar); Emma's Children's Hospital AMC, Amsterdam (JH Kok, A Ilsen, M van der Lans, WJC Boelen-van der loo, T Lundqvist, HSA Heymans); Univeristy Hospital Groningen, Beatrix Children's Hospital, Groningen (EJ Duiverman, WB Geven, ML Duiverman, LI Geven, EJLE Vrijlandt); University Hospital Maastricht, Maastricht (ALM Mulder, A Gerver); University Medical Center St Radboud, Nijmegen (LAA Kollée, L Reijmers, R Sonnemans); Leiden University Medical Center, Leiden (HA Delemarre-van de Waal, JM Wit, FW Dekker, MJJ Finken); Erasmus MC - Sophia Children's Hospital, University Medical Center Rotterdam (N Weisglas-Kuperus, MG Keijzer-Veen, AJ van der Heijden, JB van Goudoever); VU University Medical Center, Amsterdam

(MM van Weissenbruch, A Cranendonk, L de Groot, JF Samsom); Wilhelmina Children's Hospital, UMC, Utrecht (LS de Vries, KJ Rademaker, E Moerman, M Voogsgeerd); Máxima Medical Center, Veldhoven (MJK de Kleine, P Andriessen, CCM Dielissen-van Helvoirt, I Mohamed); Isala Clinics, Zwolle (HLM van Straaten, W Baerts, GW Veneklaas Slots-Kloosterboer, EMJ Tuller-Pikkemaat); Royal Effatha Guyot Group, Zoetermeer (MH Ens-Dokkum); Association for Parents of Premature Babies (GJ van Steenbrugge).

References

1. Gluckman PD, Hanson MA, Cooper C, Thornburg KL. Effect of in utero and early-life conditions on adult health and disease. N.Engl.J.Med. 2008; 359: 61-73
2. Bergvall N, Iliadou A, Johansson S, de FU, Kramer MS, Pawitan Y, et al. Genetic and shared environmental factors do not confound the association between birth weight and hypertension: a study among Swedish twins. Circulation 2007; 115: 2931-8
3. Godfrey KM, Barker DJ. Fetal nutrition and adult disease. Am.J.Clin.Nutr. 2000; 71: 1344S-52S
4. Ross MG, Beall MH. Adult sequelae of intrauterine growth restriction. Semin. Perinatol. 2008; 32: 213-8
5. Whincup PH, Kaye SJ, Owen CG, Huxley R, Cook DG, Anazawa S, et al. Birth weight and risk of type 2 diabetes: a systematic review. JAMA 2008; 300: 2886-97
6. Lee PA, Chernausek SD, Hokken-Koelega AC, Czernichow P. International Small for Gestational Age Advisory Board consensus development conference statement: management of short children born small for gestational age, April 24-October 1, 2001. Pediatrics 2003; 111: 1253-61
7. Bertino E, Milani S, Fabris C, De CM. Neonatal anthropometric charts: what they are, what they are not. Arch.Dis.Child Fetal Neonatal Ed 2007; 92: F7-F10
8. Nathanielsz PW. Animal models that elucidate basic principles of the developmental origins of adult diseases. ILAR.J. 2006; 47: 73-82
9. Costa J, Rice H, Cardwell C, Hunter A, Ong S. An assessment of vascularity and flow intensity of the placenta in normal pregnancy and pre-eclampsia using three-dimensional ultrasound. J.Matern.Fetal Neonatal Med. 2010; 23: 894-9
10. Zetterstrom K, Lindeberg SN, Haglund B, Hanson U. Chronic hypertension as a risk factor for offspring to be born small for gestational age. Acta Obstet.Gynecol.Scand. 2006; 85: 1046-50
11. Catov JM, Nohr EA, Bodnar LM, Knudson VK, Olsen SF, Olsen J. Association of periconceptional multivitamin use with reduced risk of preeclampsia among normal-weight women in the Danish National Birth Cohort. Am.J.Epidemiol. 2009; 169: 1304-11
12. Shah PS, Ohlsson A. Effects of prenatal multimicronutrient supplementation on pregnancy outcomes: a meta-analysis. CMAJ. 2009; 180: E99-108
13. Haider BA, Bhutta ZA. Multiple-micronutrient supplementation for women during pregnancy. Cochrane.Database.Syst.Rev. 2006; CD004905
14. Bogdarina I, Welham S, King PJ, Burns SP, Clark AJ. Epigenetic modification of the renin-angiotensin system in the fetal programming of hypertension. Circ.Res. 2007; 100: 520-6
15. Lillycrop KA, Phillips ES, Torrens C, Hanson MA, Jackson AA, Burdge GC. Feeding pregnant rats a protein-restricted diet persistently alters the methylation of specific cytosines in the hepatic PPAR alpha promoter of the offspring. Br.J.Nutr. 2008; 100: 278-82
16. Park JH, Stoffers DA, Nicholls RD, Simmons RA. Development of type 2 diabetes following intrauterine growth retardation in rats is associated with progressive epigenetic silencing of Pdx1. J.Clin.Invest 2008; 118: 2316-24
17. Raychaudhuri N, Raychaudhuri S, Thamotharan M, Devaskar SU. Histone code modifications repress glucose transporter 4 expression in the intrauterine growth-restricted offspring. J.Biol.Chem. 2008; 283: 13611-26
18. Sinclair KD, Allegrucci C, Singh R, Gardner DS, Sebastian S, Bispham J, et al. DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and methionine status. Proc.Natl.Acad.Sci.U.S.A 2007; 104: 19351-6
19. 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: 3652-64
20. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, et al. Persistent

epigenetic differences associated with prenatal exposure to famine in humans. Proc. Natl.Acad.Sci.U.S.A 2008; 105: 17046-9

21. Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD, et al. DNA Methylation differences after exposure to prenatal famine are common and timing- and sex-specific. Hum.Mol.Genet. 2009; 18: 4046-53

22. Steegers-Theunissen RP, Obermann-Borst SA, Kremer D, Lindemans J, Siebel C, Steegers EA, et al. Periconceptional maternal folic acid use of 400 microg per day is related to increased methylation of the IGF2 gene in the very young child. PLoS.One. 2009; 4: e7845

23. Constancia M, Hemberger M, Hughes J, Dean W, Ferguson-Smith A, Fundele R, et al. Placental-specific IGF-II is a major modulator of placental and fetal growth. Nature 2002; 417: 945-8

24. Adkins RM, Krushkal J, Klauser CK, Magann EF, Morrison JC, Somes G. Association between small for gestational age and paternally inherited 5' insulin haplotypes. Int.J.Obes.(Lond) 2008; 32: 372-80

25. McMinn J, Wei M, Schupf N, Cusmai J, Johnson EB, Smith AC, et al. Unbalanced placental expression of imprinted genes in human intrauterine growth restriction. Placenta 2006; 27: 540-9

26. Forhead AJ, Fowden AL. The hungry fetus? Role of leptin as a nutritional signal before birth. J.Physiol 2009; 587: 1145-52

27. Weinstein LS, Xie T, Zhang QH, Chen M. Studies of the regulation and function of the Gs alpha gene Gnas using gene targeting technology. Pharmacol.Ther. 2007; 115: 271-91

28. Kyle UG, Pichard C. The Dutch Famine of 1944-1945: a pathophysiological model of long-term consequences of wasting disease. Curr.Opin.Clin.Nutr.Metab Care 2006; 9: 388-94

29. Hille ET, Weisglas-Kuperus N, van Goudoever JB, Jacobusse GW, Ens-Dokkum MH, de GL, et al. Functional outcomes and participation in young adulthood for very preterm and very low birth weight infants: the Dutch Project on Preterm and Small for Gestational Age Infants at 19 years of age. Pediatrics 2007; 120: e587-e595

30. Euser AM, de Wit CC, Finken MJ, Rijken M, Wit JM. Growth of preterm born children. Horm.Res. 2008; 70: 319-28

31. Guo L, Choufani S, Ferreira J, Smith A, Chitayat D, Shuman C, et al. Altered gene expression and methylation of the human chromosome 11 imprinted region in small for gestational age (SGA) placentae. Dev.Biol. 2008; 320: 79-91

32. Tabano S, Colapietro P, Cetin I, Grati FR, Zanutto S, Mando C, et al. Epigenetic modulation of the IGF2/H19 imprinted domain in human embryonic and extra-embryonic compartments and its possible role in fetal growth restriction. Epigenetics. 2010; 5: 313-24

33. Peterson K. Viewpoint: childhood undernutrition: a failing global priority. J.Public Health Policy 2009; 30: 455-64

34. Zhang L. Prenatal hypoxia and cardiac programming. J.Soc.Gynecol.Investig. 2005; 12: 2-13

35. Patterson AJ, Chen M, Xue Q, Xiao D, Zhang L. Chronic Prenatal Hypoxia Induces Epigenetic Programming of PKC{epsilon} Gene Repression in Rat Hearts. Circ.Res. 2010; 107: 365-73

36. Einstein F, Thompson RF, Bhagat TD, Fazzari MJ, Verma A, Barzilai N, et al. Cytosine methylation dysregulation in neonates following intrauterine growth restriction. PLoS. One. 2010; 5: e8887

37. Johansson S, Iliadou A, Bergvall N, de FU, Kramer MS, Pawitan Y, et al. The association between low birth weight and type 2 diabetes: contribution of genetic factors. Epidemiology 2008; 19: 659-65

38. Freathy RM, Mook-Kanamori DO, Sovio U, Prokopenko I, Timpson NJ, Berry DJ, et al. Variants in ADCY5 and near CCNL1 are associated with fetal growth and birth weight. Nat.Genet. 2010; 42: 430-5

39. Finken MJ, Meulenbelt I, Dekker FW, Frolich M, Romijn JA, Slagboom PE, et al. The 23K variant of the R23K polymorphism in the glucocorticoid receptor gene protects against postnatal growth failure and insulin resistance after preterm birth. J.Clin.Endocrinol.Metab 2007; 92: 4777-82

40. Walther FJ, den Ouden AL, Verloove-Vanhorick SP. Looking back in time: outcome

of a national cohort of very preterm infants born in The Netherlands in 1983. Early Hum. Dev. 2000; 59: 175-91

41. Niklasson A, Ericson A, Fryer JG, Karlberg J, Lawrence C, Karlberg P. An update of the Swedish reference standards for weight, length and head circumference at birth for given gestational age (1977-1981). Acta Paediatr.Scand. 1991; 80: 756-62
42. Visser GH, Eilers PH, Elferink-Stinkens PM, Merkus HM, Wit JM. New Dutch reference curves for birthweight by gestational age. Early Hum.Dev. 2009; 85: 737-44
43. Fredriks AM, van BS, Burgmeijer RJ, Meulmeester JF, Beuker RJ, Brugman E, et al. Continuing positive secular growth change in The Netherlands 1955-1997. Pediatr.Res. 2000; 47: 316-23
44. Coolen MW, Statham AL, Gardiner-Garden M, Clark SJ. Genomic profiling of CpG methylation and allelic specificity using quantitative high-throughput mass spectrometry: critical evaluation and improvements. Nucleic Acids Res. 2007; 35: e119
45. Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. Proc.Natl.Acad.Sci.U.S.A 2005; 102: 15785-90
46. Talens RP, Boomsma DI, Tobi EW, Kremer D, Jukema JW, Willemsen G, et al. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. FASEB J. 2010; 24: 3135-44
47. Cui H, Onyango P, Brandenburg S, Wu Y, Hsieh CL, Feinberg AP. Loss of imprinting in colorectal cancer linked to hypomethylation of H19 and IGF2. Cancer Res. 2002; 62: 6442-6
48. Monk D, Sanches R, Arnaud P, Apostolidou S, Hills FA, Abu-Amero S, et al. Imprinting of IGF2 P0 transcript and novel alternatively spliced INS-IGF2 isoforms show differences between mouse and human. Hum.Mol.Genet. 2006; 15: 1259-69
49. Hayward BE, Bonthron DT. An imprinted antisense transcript at the human GNAS1 locus. Hum.Mol.Genet. 2000; 9: 835-41
50. Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 2007; 447: 799-816
51. Stoger R. In vivo methylation patterns of the leptin promoter in human and mouse. Epigenetics. 2006; 1: 155-62

Prenatal famine and genetic variation are independently and additively associated with DNA methylation at regulatory loci within IGF2/H19

Elmar W. Tobi[1], P. Eline Slagboom[1,3], Jenny van Dongen[1,4], Dennis Kremer[1], Aryeh D. Stein[5], Hein Putter[2], Bastiaan T. Heijmans[1,3,*],L. H. Lumey[6,*]

1. Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
2. Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands
3. The Netherlands consortium for Healthy Ageing, Leiden University Medical Center, Leiden, The Netherlands
4. Biological Psychology, VU University Amsterdam, Amsterdam, The Netherlands
5. Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA
6. Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA

* These authors contributed equally to this work

PLoS One. 2012;7(5):e37933

Abstract

Both the early environment and genetic variation may affect DNA methylation, which is one of the major molecular marks of the epigenome. The combined effect of these factors on a well-defined locus has not been studied to date. We evaluated the association of periconceptional exposure to the Dutch Famine of 1944-45, as an example of an early environmental exposure, and single nucleotide polymorphisms covering the genetic variation (tagging SNPs) with DNA methylation at the imprinted *IGF2/H19* region, a model for an epigenetically regulated genomic region.

DNA methylation was measured at five differentially methylated regions (DMRs) that regulate the imprinted status of the *IGF2*/*H19* region. Small but consistent differences in DNA methylation were observed comparing 60 individuals with periconceptional famine exposure with unexposed same-sex siblings at all *IGF2* DMRs (P_{BH}<0.05 after adjustment for multiple testing), but not at the *H19* DMR. *IGF2* DMR0 methylation was associated with *IGF2* SNP rs2239681 (P_{BH}=0.027) and *INS* promoter methylation with *INS* SNPs, including rs689, which tags the *INS* VNTR, suggesting a mechanism for the reported effect of the VNTR on *INS* expression (P_{BH}=3.4x10^{-3}). Prenatal famine and genetic variation showed similar associations with *IGF2*/*H19* methylation and their contributions were additive. They were small in absolute terms (<3%), but on average 0.5 standard deviations relative to the variation in the population.

Our analyses suggest that environmental and genetic factors could have independent and additive similarly sized effects on DNA methylation at the same regulatory site.

Introduction

The epigenome consists of inter-related layers of molecular marks on the DNA that represent non-genetic, but stable and mitotically heritable information determining the gene-expression potential of a genomic region[1]. Studies in animal models show that environmental factors during early development can cause persistent epigenetic changes in DNA methylation that are associated with disease-related phenotypes[2,3]. This suggests that the prenatal environment ('nurture') can persistently influence the expression of DNA sequences ('nature')[4]. Recent studies stress that variation in DNA methylation is primarily influenced by genetic variation[5] and that the DNA sequence itself dictates the DNA methylation state of a locus[6].

Although there is evidence for distinct environmental and genetic influences on DNA methylation, it is not clear how both factors may interact and determine the DNA methylation levels at a particular locus. We at least are not aware of any such studies. Insight in these matters is of interest for the interpretation of epigenome-wide association studies (EWASs)[7] and studies investigating the developmental origins hypothesis[4]. We address this issue by further evaluating the interplay between environmental and genetic factors with respect to DNA methylation for selected regulatory loci within the *IGF2/H19* region.

The *IGF2/H19* imprinted region is one of the best-understood epigenetically controlled loci involving the methylation of various differentially methylated regions (DMRs). Previous studies reported that DNA methylation at the *IGF2* DMR0 is associated with genetic factors[8-10] and the prenatal environment, including periconceptional exposure to the Dutch Famine at the end of WW2[11] and maternal folic acid supplementation[12].Therefore the methylation at selected loci in the *IGF2*/*H19* region in individuals exposed to prenatal famine may offer a special opportunity to evaluate the interplay between genetics and environment on DNA methylation.

The correct mono-allelic expression of genes in *IGF2*/*H19* region in somatic cells is regulated by several DMRs (Figure 1)[13,14]. Going from centromere to telomere, the first DMR is the imprinted insulin promoter (*INS*)[15], which also

influences the neighboring insulin-like growth factor 2 (*IGF2*) gene[16]. *INS* forms a fusion transcript between *INS* and *IGF2* during early development, called *INSIGF*[17] and DNA methylation at this locus is correlated with *INS* transcription[18]. The next DMR is *IGF2* DMR0 (alternate name *IGF2* DMR) at which abnormal DNA methylation is associated with bi-allelic expression of *IGF2*[19,20]. The *IGF2* DMR1, within a large CpG island overlapping the *IGF2AS* promoter (alternate name *PEG8*), is reported to have an insulator function and bind CTCF[21]. *IGF2* DMR2 was also reported to act as an insulator and to bind CTCF and aberrant DNA methylation at the locus is associated with a loss of imprinting[22]. The final DMR is located in the promoter of the *H19* transcript that directly flanks the imprinting control region. Aberrant DNA methylation at this DMR is correlated with a loss of imprinting and over-expression[23].

Here, we present an in-depth characterization of DNA methylation differences at nine regulatory loci within five DMRs across the *IGF2/H19* region between 60 individuals exposed periconceptional to the Dutch Famine and 60 same-sex sibling controls without prenatal famine exposure. All individuals are part of our ongoing Dutch Hunger Winter Families Study[24]. We examined if the famine associations are locus specific or extend to multiple functional loci. We also examined a measure of global methylation to compare the locus-specific associations with possible overall genomic effects after famine exposure. Moreover, we evaluated the association between *IGF2/H19* methylation and common genetic variation in the sibling pairs by genotyping tagging SNPs. Finally, we tested if the associations between famine exposure and genetic variation are independent and contrasted the effect sizes of these associations to describe the relative contribution of ‘nature’ and ‘nurture’ to variation in DNA methylation at *IGF2/H19*.

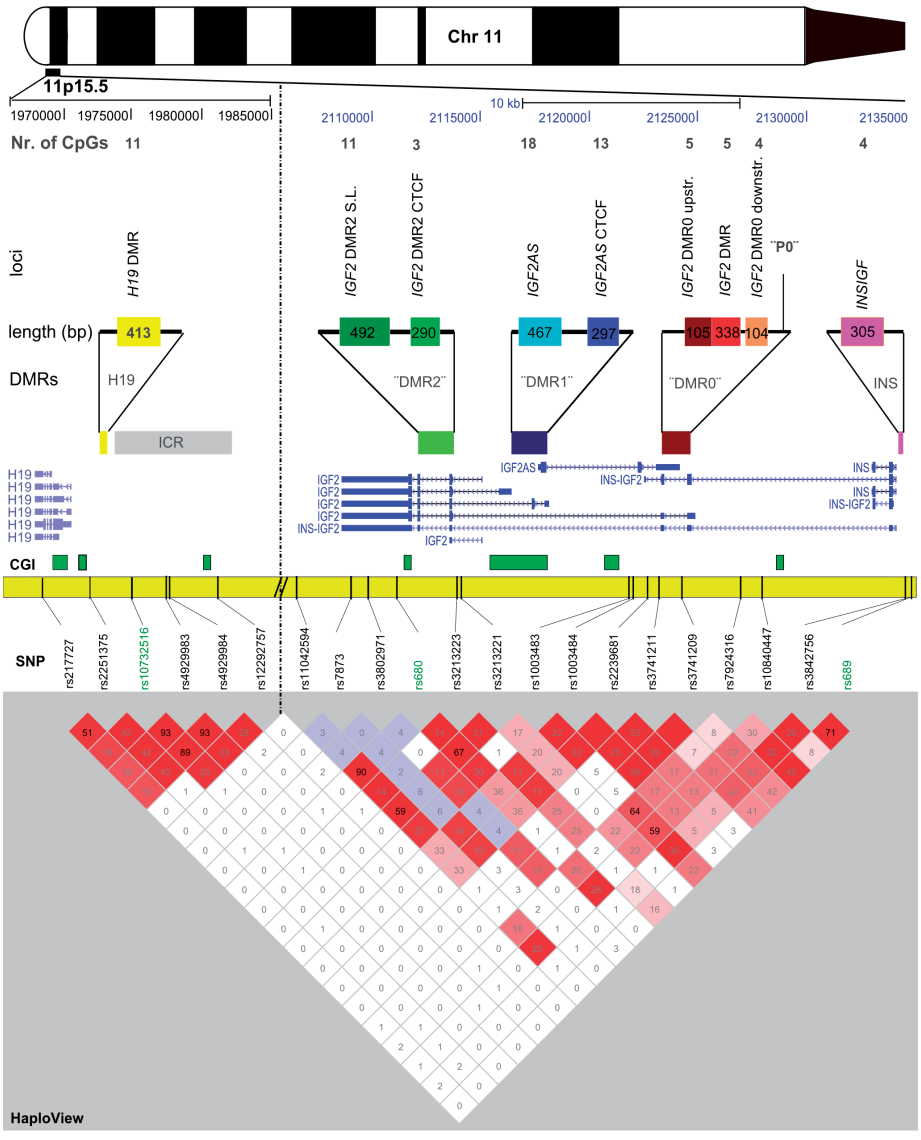


Figure 1. Schematic overview of the IGF2/H19 region, measured loci and genetic variation covered

The colored boxes in the loci pane represent the DNA methylation measurements as distributed over the various functional differentially methylated regions, also defined by unique coloring (*H19* DMR, *IGF2* DMR2, DMR1, DMR0 and the *INS* promoter). The number of CpG sites measured per locus is given above the locus names. The gene structure, as defined by Refseq, is given together with the CpG islands ("CGI", bright green). The yellow bar presents the chromosome, with the various measured SNPs marked by bars. In the HaploView pane the D' between SNPs is given in the color scale, while the R-squared is given in numeric values in the boxes.

Results

Analysis of IGF2/H19 methylation

Within the five DMRs, nine methylated loci were reported to regulate imprinting and expression of *INS*, *INSIGF*, *IGF2* and *H19* (**Figure 1**)[9,17-22,25-27]. We analyzed DNA methylation at one locus in the *INS* promoter (*INSIGF*), three in *IGF2* DMR0 (*IGF2* DMR0 downstr., *IGF2* DMR and *IGF2* DMR0 upstr.), two in *IGF2* DMR1 (*IGF2AS* CTCF and *IGF2AS*) and two in DMR2 (*IGF2* DMR2 CTCF and *IGF2* DMR2 S.L.) and one in the *H19* DMR (*H19* DMR). Information on the functionality of these loci is provided in the materials and methods section. The precise genomic locations are given in **Supplement I, Table S8**. Information on the individual CpG dinucleotides measured within each locus is given in **Supplement I, Table S9**. We measured DNA methylation at these loci in 60 individuals with periconceptional famine exposure and 60 unexposed, same-sex siblings. DNA methylation was quantitatively assessed by mass spectrometry (Epityper)[28], which quantifies the number of methylated and unmethylated fragments following bisulfite PCR and base specific cleavage.

Inspection of DNA methylation patterns showed that DNA methylation at different loci assayed within a DMR was correlated (**Figure 2A**), except for DMR2. In DMR2, methylation at the *IGF2* DMR2 CTCF locus (a CTCF binding site)[21] was not correlated with the *IGF2* DMR2 S.L. locus (a DNA stem loop structure)[27]. In view of the high within DMR correlation, the three loci assayed for *IGF2* DMR0 and the two for *IGF2* DMR1 were also analyzed as a single locus. Positive correlations were observed between DMRs, in particular between DMR0, DMR1, DMR2 CTCF and *INSIGF*. Interestingly these loci are located up to 10kb apart. To a lesser extent, correlations were also observed between *H19* DMR and DMR2 S.L.

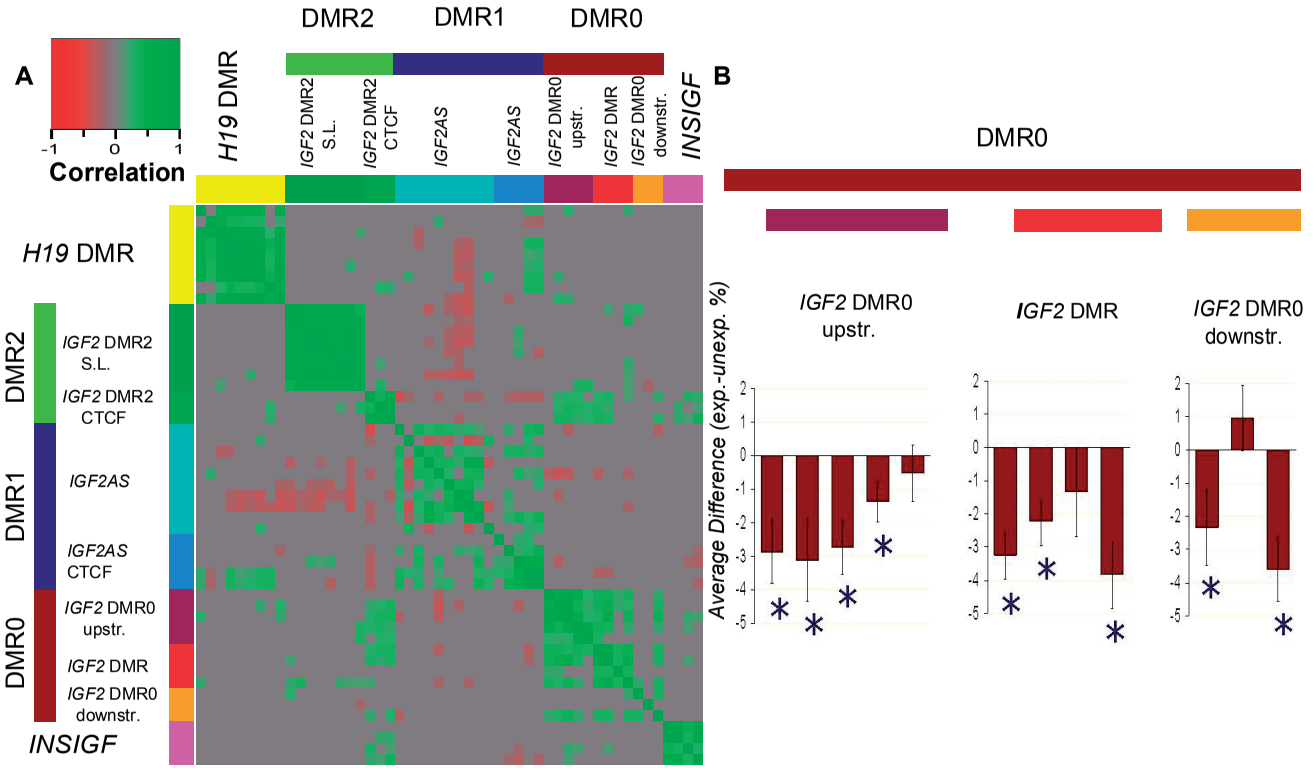


Figure 2. DNA methylation across multiple loci

The colored bars present DNA methylation measurements of the various loci and their grouping in the final analyses. **A**. The correlation of the DNA methylation of CpG dinucleotides within *IGF2/H19*.Each square block represents the pair wise correlation between two CpG dinucleotides in 120 individuals, the 60 individuals exposed periconceptional to famine and their same-sex siblings. Only significant correlations are shown in a color gradient from red (ρ=-1) to gray (ρ=0 or N.S.) to green (ρ=+1). **B**. The average within pair difference in DNA methylation (%) between the famine exposed and their same-sex sibling controls for the 3 amplicons measured in the *IGF2* DMR0. A * denotes if the individual fragment containing one or multiple CpG sites is significantly different between the exposed and unexposed (P<0.05).

Prenatal famine exposure and IGF2/H19 methylation

We previously reported on the association of methylation at the *INS* promoter locus *INSIGF* with periconceptional famine exposure[29]. The *INSIGF* methylation was 1.5% lower in exposed individuals as compared to the unexposed siblings (P_{BH}=0.015 after Benjamini–Hochberg adjustment for multiple testing). Expressed relative to the standard deviation in controls of 2.6%, this difference corresponds to a standardized effect size of -0.6 SD units (**Table 1**).

IGF2 DMR0 methylation was lower in the exposed siblings (Δ=-2.0%; P_{BH}=2.9x10^{-3}), corresponding to a standardized effect size of -0.6 SD units, similar to what was also observed for *INSIGF* (**Table 1**). When analyzed separately, all 3 loci measured within the *IGF2* DMR0 (a locus previously analyzed in this study population[11] and two newly measured loci flanking that locus), were similarly associated with prenatal famine exposure (**Figure 2B and Table 1**). Subsequent analysis of individual CpG dinucleotides in these loci showed a significant association for nine out of twelve CpG containing fragments (**Supplement I, Table S9 and Figure 2B**).

Methylation at the *IGF2* DMR1 was higher in exposed individuals as compared with controls (P_{BH}=0.020), but the absolute difference in DNA methylation was very small (Δ=+0.5%). The difference corresponds to a standardized effect size of 0.5 SD units, similar to that observed for *INSIGF* and *IGF2* DMR0, which is related to the lower inter-individual variation at DMR1 (**Table 1**). DNA methylation at the two individual loci measured within *IGF2* DMR1 was likewise modestly higher in those exposed periconceptional (P_{BH}=0.049). In contrast to other associated DMRs, only a small minority of CpG dinucleotides within the two DMR1 loci (3/24) were statistically significant (**Supplement I, Table S9**).

DNA methylation of the two loci measured in *IGF2* DMR2 was not correlated and therefore analyzed separately. The *IGF2* DMR2 CTCF locus showed a significant association with famine exposure (Δ=-1.2%, P_{BH}=0.02). With an effect size of -0.4 SD units this association was similar to those found for the other investigated *IGF2* DMRs (**Table 1**). All three individual CpG

sites showed a lower methylation level in the exposed compared to the controls and for two out of three the difference was statistically significant (**Supplement I, Table S9**). The *IGF2* DMR2 S.L. locus was not associated with famine exposure (P_{BH}=0.78) and DNA methylation at the *H19* DMR was also not significantly associated with prenatal famine exposure (P_{BH}=0.39).

Table 1. The associations between periconceptional famine exposure and DNA methylation

Locus	controls Methylation (sd) %[1]	Exp. – Unexp. (%)[2]	Effect size[3]	P	P_{BH}[4]
INSIGF	*84.8 (2.6)*	*-1.5*	*-0.6*	*$4.0x10^{-3}$*	*0.015*
IGF2 DMR0	*53.3(3.3)*	*-2.0*	*-0.6*	*$2.7x10^{-4}$*	*$2.9x10^{-3}$*
IGF2 DMR0 downstr.	71.2(3.3)	-1.6	-0.6	0.024	0.038
IGF2 DMR	51.5(5.5)	-2.4	-0.5	$5.3x10^{-4}$	$2.9x10^{-3}$
IGF2 DMR0 upstr.	44.2(4.0)	-1.9	-0.4	$6.0x10^{-3}$	0.017
IGF2 DMR1	*6.4(0.8)*	*0.4*	*+0.5*	*$9.0x10^{-3}$*	*0.020*
IGF2AS CTCF	4.3(0.9)	0.3	+0.3	0.040	0.049
IGF2AS	8.6(1.0)	0.4	+0.4	0.038	0.049
IGF2 DMR2					
IGF2 DMR2 S.L.	49.8(6.3)	0.4	+0.1	0.78	0.78
IGF2 DMR2 CTCF	50.8(2.7)	-1.2	-0.4	0.012	0.022
H19 DMR	*30.6 (2.6)*	*-0.5*	*-0.2*	*0.36*	*0.39*

1. The average DNA methylation in the unexposed sibling controls and the standard deviation of this average, both given in %.
2. The within pair difference in DNA methylation resulting from a linear mixed model corrected for age at blood drawl, correlations between CpG sites, bisulfite conversion batch and with a random effect for sib ship and a random slope for exposure status.
3. The effect size of the within pair difference in relation to the standard deviation in the population.
4. Two-sided P-value, Benjamini-Hochberg ('FDR') corrected for 11 tests.

To evaluate whether the generally lower DNA methylation at *IGF2* DMRs was related to an overall lower genomic DNA methylation, we measured LINE-1 methylation, an estimate of global methylation[30]. LINE-1 methylation was 61.2% (SD 1.4%) in controls and this was not different from individuals with periconceptional famine exposure (Δ=-0.4%, P=0.15, **Supplement**

I; Table S9). This result confirms our previous report that prenatal famine had no effect on three other measures of global methylation in this study population[31] and indicates the absence of a general trend towards either reduced or increased genomic DNA methylation.

Genetic variation and IGF2/H19 methylation

To capture common genetic variation at the *IGF2*/*H19* locus, 21 SNPs were genotyped. The SNPs were selected as tagging SNPs from the HAPMAP CEU panel or selected from literature (**Figure 1 and Supplement I Table S10 and S11**). Linkage disequilibrium (LD) analysis indicated that 16 of the 21 SNPs captured the common genetic variation marked by these SNPs (R^2>0.9). Of these sixteen SNPs, four were located in the *H19* region and twelve in the *IGF2-INS* region; no LD was observed between the two regions (**Figure 1**). Genotype frequencies were similar in exposed individuals and unexposed siblings (P>0.13, without multiple testing correction).

We then explored which of these sixteen tagging SNPs was associated with DNA methylation at the *IGF2*/*H19* DMRs in the sixty sib ships (N=120). DNA methylation at *INSIGF* was significantly associated with SNPs in *IGF2* (rs3741211 [β=-1.5% per minor allele, $P_{BH}=3.4x10^{-3}$]) and *INSIGF* (rs3842756 [β=-2.0%, $P_{BH}=3.9x10^{-4}$] and rs689 [β=-2.3%, $P_{BH}=7.1x10^{-6}$]) (**Table 2**). The standardized effect size of associations increased with decreasing distance from the DMRs (from -0.6 to -0.9 SD units per minor allele, **Figure 3**). The largest effect size was observed for the association of rs689, which is in perfect LD with the *INS* VNTR I/III alleles in Caucasian populations[32], with *INSIGF* methylation (-0.9 SD; $P_{BH}=7.1x10^{-6}$). Other nominally significant associations with *INSIGF* (P<0.05 and P_{BH}>0.05) are reported in **Figure 3**.

For *IGF2* DMR0, rs2239681 was associated with DNA methylation (β=-1.3%, P_{BH}=0.027; **Figure 3**). For two SNPs nominally significantly associations were observed, which included rs680 (i.e. "*Apa*I", β=-1.1%, P=0.017 [P_{BH}=0.17]), for which we reported an association previously[8].

For *IGF2* DMR1 no statistically significant associations were observed after accounting for multiple testing. One of the three nominally significant

Table 2. The significant associations between SNPs and DNA methylation

Locus	Meth. (sd) in %	Effect of genotype on DNA methylation[3]				
INSIGF	84.8(2.6)	SNP	*B*[1]	effect size[2]	P	P_{BH}[4]
		rs3741211	-1.5	-0.6	$1.1x10^{-4}$	$3.4x10^{-3}$
		rs3842756	-2.0	-0.8	$8.2x10^{-6}$	$3.9x10^{-4}$
		rs689	-2.3	-0.9	$7.4x10^{-8}$	$7.1x10^{-6}$
IGF2 DMR0	53.3(3.3)	SNP	*B*	effect size	P	P_{BH}
		rs2239681	-1.3	-0.4	$1.1x10^{-3}$	0.027

1. The change in average DNA methylation in % with each minor allele. From a linear mixed model corrected for age at blood drawl, correlations between CpG sites, bisulfite conversion batch and with a random effect for sib ship and a random slope for exposure status. The genotype was added as continues variable.
2. The effect size of the beta in relation to the variation in DNA methylation in the population.
3. The associations that survive multiple testing correction; a complete overview off all the results is given in supplemental tables S5A-C.
4. Two-sided P-value, after Benjamini-Hochberg correction.

associations was rs689 marking the *INS* VNTR located near *INSIGF* (**Figure 3**; β=0.3%, P=0.012 [P_{BH}=0.14]).

The two loci in *IGF2* DMR2 (CTCF and S.L.) were analyzed separately in contrast to the loci comprising the other DMRs because the methylation levels were not correlated (Figure 2A). No associations were observed for *IGF2* DMR2 CTCF and S.L. when accounting for multiple testing. Two nominally significant association were found for the *H19* SNPs rs2251375 and rs4929983 (tagging rs10732516 in the ICR's sixth CTCF binding site)[33] and *IGF2* DMR2 S.L. DNA methylation (**Figure 3**; β=-2.4%, P=0.028 [P_{BH}=0.21]; β=-2.6%, P=$8.6x10^{-3}$, [P_{BH}=0.13]). For *H19* DMR, no associations were observed in line with earlier observations[8].

Prenatal famine exposure and genetic variation

Expressed as standardized effect sizes, the average DNA methylation difference between exposed individuals and unexposed siblings was 0.5 SD units for significantly associated DMRs (P_{BH}<0.05). A similar average of 0.5 SD per minor allele was observed for associated SNPs (P_{BH}<0.05; **Figure 3**).

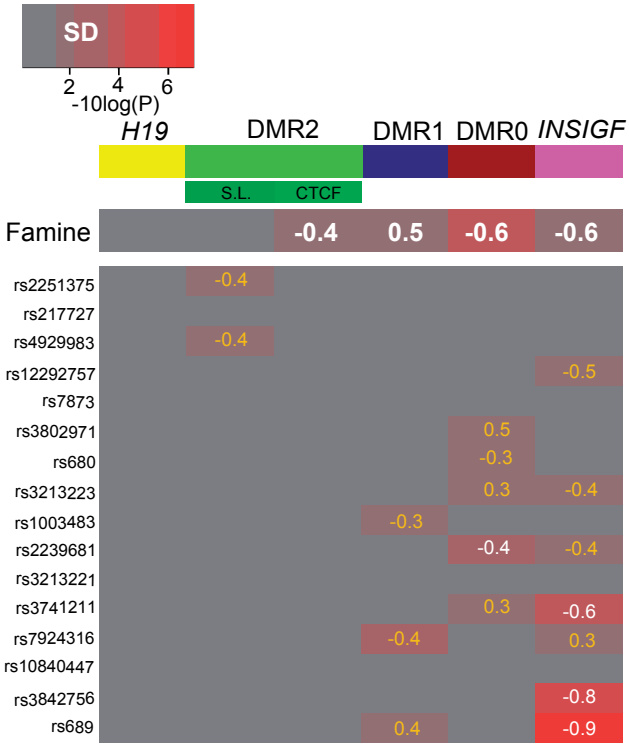


Figure 3. The association of prenatal famine and genetic variation in cis with IGF2/H19 methylation

The associations between famine or genotype with DNA methylation. The p-value of the association (-log10 scale) is given in a color scale from non significant (gray) to highly significant (bright red). DMRs are denoted by the colored bars in top of the figure. The effect sizes are given in standardized SD-scores. In a white font are the associations significant after multiple testing correction. The nominally significant associations are denoted in orange.

Since the methylation at several loci was associated with both famine exposure and SNPs in these sibling pairs, the analyses for associations between prenatal famine and DMR methylation were repeated with adjustment for the SNPs significantly associated with DMR methylation (either nominally or after correction for multiple testing). All famine associations remained statistically significant and the effect sizes remained similar (**Supplement I, Table S12**). Similarly, the genetic associations were not affected after adjustment for prenatal famine exposure (**Supplement I, Table S13**).

Next, we tested for possible interactions between famine exposure and genetic variation with respect to the DNA methylation levels at *IGF2*/*H19*. No interaction was significant after control for multiple testing. Of all tested interactions, only the interactions between prenatal famine exposure and *INSIGF* SNPs rs3842756 (P=0.048) and rs689 (P=0.016) in relation to *IGF2* DMR1 methylation were nominally significant. The effect of prenatal famine exposure and genetic variation on DNA methylation at the DMRs therefore appear to be additive.

Discussion

We studied the relations between periconceptional famine exposure, genetic variation and DNA methylation at DMRs in the imprinted *IGF2/H19* region. Famine exposure was associated with widespread but modest differences in DNA methylation across multiple DMRs within the *INS* and *IGF2* transcribed region. Associations of SNPs with DNA methylation at the *IGF2*/*H19* DMRs were likewise common and modest. When expressed relative to the variation in DNA methylation in the population, prenatal famine and SNPs were associated with similar effect sizes, around 0.5 SD units. Periconceptional famine exposure and genetic variation were associated with DNA methylation at the same DMRs and these associations were independent of each other.

We previously reported a decrease in *IGF2* DMR0 methylation after periconceptional exposure to famine[11]. DNA methylation at two loci directly flanking this locus was similarly associated, extending the affected region in DMR0 to a region of 1.5 kb. Significant differences in DNA methylation were also observed for *IGF2* DMR1, DMR2 and *INSIGF*. Despite being widespread and statistically significant, the absolute differences between the exposed and their siblings varied and were modest (<3.6 percent points) and represent a moderate change when set out against the inter-individual variation (~0.5 SD units). Long-term functional consequences of such small absolute changes, which were also reported for other exposures[12,34-36], remain to be established.

Wide-spread small changes were suggested to be a plausible mechanism by which epigenetic fine-tuning of pathways may occur[37]. In this respect it is of interest to note that the magnitude of the effect sizes was similar for all DMRs. We are the first to show that a prenatal environmental exposure may influence DNA methylation at multiple distinct regulatory sites within the same gene. Small variations in DNA methylation at particular loci associated with prenatal human environmental exposures[12,29,34,35,38], risk factors for disease[39,40], or clinical end-points[41], may represent variation at multiple sites. The similarity of the effect sizes among the DMRs may also reflect a molecular remnant of differences in gene expression during the periconceptional period among famine exposed individuals. A study in mouse oocytes showed that active transcription influences the DNA methylation deposition at *Gnas* and other imprinted DMRs[42]. In addition, changes in gene expression during late mouse liver development also influences DNA methylation at various genes[43]. In view of these studies, the decrease in DMR0 and DMR2 methylation may reflect a temporary decrease in *IGF2* transcription. These differences may initially have been larger than currently observed, as they were measured six decades after the actual exposure. After the exposure, the differences may have been diluted by other environmental influences[44] and stochastic changes accumulating during ageing[45]. However, the subtle nature of the DNA methylation differences on a population level may also be inherent to the noise in the epigenetic response to environmental exposures[46].

Another aspect potentially contributing noise is the cellular heterogeneity of whole blood, the sample type currently studied. However, *IGF2* DMR0 methylation was shown not to be influenced by cellular heterogeneity, nor was there a difference between buccal cells and blood for this DMR[45]. Although this does not exclude differences between blood cell types for loci within *IGF2*/*H19*, including DMR0[10], this implies that the combination of the small size of these differences between cell types in combination with the limited variation in proportions of these cell types in blood between individuals is unlikely to have influenced our observations for this imprinted locus. Moreover, animal studies showed that both an exposures during gametogenesis[3] and the early post conception stage[47] can both lead to epigenetic differences observed in

multiple tissues in adults, presumably because they were propagated soma-wide. Therefore, if differences were induced early in human development, for example by periconceptional famine exposure, they could likewise be propagated soma-wide and be present across cell-types and tissues[48]. We currently do not have access to other (internal) tissues in our cohort to test this hypothesis, an important issue in epigenetic epidemiology[49].

Previous studies suggest that 95% of reported associations between SNPs and DNA methylation occurred for SNPs located within 149 kb of the CpG dinucleotides[50] with a peak at a physical distance of just 45bp[51]. Indeed, associations were most frequent between DMRs and adjacent SNPs in our study. The effect sizes we found are smaller than those reported in some genome-wide studies[52], but comparable to those reported by Bell *et al*. for SNPs influencing DNA methylation at *FTO*[53]. Our study is comparable in size to these studies. We confirm an association for SNPs and *IGF2* DMR methylation as seen in a study among mono- and dizygotic twins[8] and our findings are also in line with results from two larger studies in unrelated individuals and twins for *IGF2* DMR and *H19*[33,54].

The most significant association between genetic variation and *IGF2*/*H19* DNA methylation was observed for rs689 and *INS* promoter methylation (*INSIGF*) (effect size of -0.9 SD units, or -2.3% percent points per minor allele, ~140 bp distance). SNP rs689 is a perfect proxy for the class I and III alleles of the *INS* VNTR in European populations[32] and associations between the *INS* VNTR and type 2 diabetes (T2D), metabolic syndrome and early growth have been frequently reported by some[55,56], but also refuted by others[56-58]. The *INS* VNTR was reported to influence insulin gene expression[59] and DNA methylation at the *INS* promoter were found to be correlated with *INS* expression, HbA_{1c} levels and T2D[18]. Our findings that the *INS* VNTR is associated with *INS* promoter methylation may shed new light on the association of the VNTR with *INS* expression and the metabolic syndrome[60]. This assumes of course that our findings in blood cells extend to relevant tissues directly involved in these conditions. Tissue specificity is not only an issue for associations between the (prenatal) environment and DNA methylation[49], but also for SNP-DNA methylation associations: a sizeable

number of associations between SNPs and DNA methylation in a large study on DNA methylation from four different brain regions were found to be tissue specific[51].

In conclusion, our observations that *INSIGF* and *IGF2* DMR0 DNA methylation levels are independently associated with genetic and early environmental factors is relevant for the design and interpretation of epigenetic association studies involving *IGF2*/*H19*. Our analyses indicate that Mendelian randomization approaches are feasible to infer causality for associations observed between DNA methylation and disease phenotypes[61]. Secondly, our results will be relevant for the interpretation of epigenome-wide association studies as genomic and environmental forces may act in tandem through the epigenome on the phenotype of interest. There may be similar and additive effects of 'nurture' and 'nature' on DNA methylation within *IGF2*/*H19*. Thus, for some loci, epigenetics may be the information layer in which the classical contrast between 'nurture' and 'nature' comes to a modern molecular synthesis[62].

Materials and Methods

Study population

The characteristics and detailed recruitment of the Hunger Winter Families Study were described previously[24]. In short, study subjects were selected from births between 1943-1947 at three institutions in famine-exposed cities (the midwifery training schools in Amsterdam and Rotterdam and the Leiden University Medical Center) and include singleton births exposed to famine *in utero*, same-sex sibling controls not exposed during *in utero* development and time controls conceived and born either just before or just after the famine. Ethical approval for the study was obtained from the participating institutions and all participants provided written informed consent.

Despite the war, nutrition in the Netherlands had generally been adequate until October 1944[63]. Thereafter, supplies became increasingly scarce. By the end of November, the level of official supplementary rations, which

eventually consisted of little more than bread and potatoes, had fallen below 1,000 kcal per day, and by April 1945 they were as low as 500 kcal per day[64]. Since the Dutch population was well fed before and after the famine and since the famine period was shorter than the nine months of human gestation, individuals can be defined by exposure during specific periods of their development in uterus.

In this study we use a subset of the 313 singleton births in the larger cohort[24] who were exposed to the Dutch famine *in utero* and who completed a clinical examination together with a same-sex sibling without prenatal famine exposure. Whole blood was collected from all individuals for DNA extraction during this examination. We focused on individuals whose mother was exposed to famine around the moment of conception and in the first 10 weeks of gestation. These 'periconceptionally' exposed individuals were defined as births with a mother's estimated last menstrual period between November 28, 1944 and May 15, 1945. This group includes 60 individuals of whom 28 are male and 32 are female (age at examination and blood draw 58.1y [SD, 0.35y]). As controls we used their unexposed same-sex sibling for (partial) genetic and gender matching (age at examination and blood draw 57.0y [SD, 5.9y]). 24 controls were conceived and born before the famine (11 male, 13 female) and 36 individuals were conceived and born after the famine (17 male and 19 female). The studied population includes 120 individuals in total.

DNA methylation assay design

We used BLAT against genome build 36 in the UCSC genome browser[65] to find the locations in 11p15.5 mentioned in the various original articles[9,17-22,25-27]. We provide an additional .BED file at the website of PLOS ONE, showing the various locations assayed and the location of the elements from the original articles on which they were based (PLOS ONE online, ***BED S1***). From the centromere outwards the loci were chosen as follows (**Figure 1**).

The *INSIGF* locus was previously measured by us[29,45,66], is imprinted[17] and DNA methylation is correlated with expression[18]. Three loci were designed

for the *IGF2* promoter region, *IGF2* DMR0, at which hypo-methylation is associated with bi-allelic expression[19,20]. *IGF2* DMR upstream (upstr.) and *IGF2* DMR downstream (downstr.) directly flank the *IGF2* DMR locus that we measured previously in this cohort[11]. *IGF2* DMR downstr. is located next to the "P0" promoter, which interacts with the imprinting control region[26]. Two loci were designed to cover the promoter region of the *IGF2*AS transcript, which we name *IGF2* DMR1. One of these loci shows CTCF binding activity (*IGF2*AS CTCF) and one locus demonstrated insulator activity (*IGF2*AS)[21]. Two loci overlap the *IGF2* DMR2, of which aberrant DNA methylation has been linked to loss of imprinting (LOI)[22] and male fertility[25]. *IGF2* DMR2 CTCF overlaps a CTCF binding site[21], while *IGF2* DMR2 S.L. overlaps a highly conserved DNA stem loop structure[27]. Last, the *H19* DMR locus was previously designed[8] to measure part of the *H19* promoter at which aberrant DNA methylation was found to correlate with LOI and over expression[23]. Several primer pairs for the sixth CTCF binding site in the ICR from literature and from our own design were tested, but gave a-specific PCR products or amplification of genomic, non-bisulfite treated DNA in our automated work-flow. We also estimated global methylation using an assay for *LINES*-1[30], based on the same technique.

Primers were designed using Methprimer[67]. The resulting primer and amplicon locations were checked against the latest version of dbSNP and for their spectrum characteristics with the R package RSeqMeth[68]. The sequences of the primers used in our study and the genomic locations they amplify are given in **Supplement I, Table S8**.

DNA methylation measurements

Genomic DNA from whole blood was isolated using the salting-out method. Bisulfite treatment on 500ng of genomic was performed with the EZ 96-DNA methylation kit (Zymo Research) with overnight bisulfite incubation according to the supplier's protocol. The 60 sibling pairs were randomly distributed over two 96 well plates with similar proportions of male and female pairs on each plate. DNA methylation was quantitatively assessed for each locus using

the mass spectrometry based Epityper assay (Sequenom, USA) in triplicate using the manufacturers' protocol on one 384 well plate. PCR was performed with the following cycling protocol: 15 minutes at 95°C, four rounds of 20 seconds at 95°C, 30 seconds at 65°C, 1 minute at 72°C; followed by forty rounds, 20 seconds at 95°C, 30 seconds at 58°C and 1 minute at 72°C; ending with 3 minutes at 72°C. Processing of the Epityper data has been described in detail previously[11,29,45,66]. In short only measurements for CpG dinucleotides containing fragments for which 2 out of 3 measurements were successful, the standard deviation (SD) was smaller than 10% and for which the overall measurement success rate in the population was higher than 75% were included in the final analyses. Before data filtering the SD between the triplicate measurements ranged from 2% to 5.4%, after data filtering this measure ranged from 1.5% to 3.5%. We used the average of these triplicate measurements for the analyses. For each measurement we incorporated non-bisulfite converted genomic DNA and negative controls to check for a-specific amplification and PCR artifacts. None were found. Bisulfite conversion was assessed using the MassArray R package[69], which uses fragments containing a TpG and a cytosine to assess the conversion. No indication for an incomplete bisulfite conversion or PCR amplification of non-bisulfite converted DNA was observed.

SNP selection and genotyping

From the combined HapMap phase I, II and III data[70] the CEU genotype data were downloaded for the region of the Refseq *H19* and the *IGF2* and *INSIGF* transcripts with an additional 150kb at both the 5' and 3' ends. These data were visualized in Haploview[71] for both regions separately. Based on the linkage disequilibrium (LD) structure one or more HaploView defined LD blocks were selected, covering the entire region for which DNA methylation was measured. For *INS* and *IGF2* this resulted in a region stretching from rs11042594 to rs3842748 (NCBI 36 chr11:2,073,729-2,137,971) and for *H19* in a region stretching from rs3741219 till rs3890907 (NCBI36 chr11:1,973,195-1,984,719). In addition, we selected 16 SNPs in these regions that have been

associated with relevant phenotypes such as being born small for gestational age[54,72], birth weight[73-75], body mass index[55,56,60,76,77], type two diabetes[60], postnatal growth[75] and *IGF2* levels[74,78] or with DNA methylation at *IGF2* DMR or *H19* DMR[8]. Twelve of these SNPs were also in the CEU HapMap set. A complete overview is given in **Supplement I**, **Tables S10** and **S11**. We used HaploView pairwise tagging (r^2>0.8) and used force include on the candidate SNPs if they were part of HapMap to obtain a set of tagging SNPs for the region. We used only SNPs with a minor allele frequency higher than 0.1 because of the limited sample size of our cohort. For *H19* rs10732516 and for *INSIGF* rs680, rs3213223 and rs1003484 were added to this list since they are not part of the CEU HapMap set. We thus obtained 10 SNPs for *H19* and 23 SNPs for *IGF2* and *INSIGF*. Genotyping was performed using Sequenom MassARRAY iPLEXGold with the exception of rs10732516. This latter SNP was measured using an ABI 3710 because of the highly repetitive nature of this region. The forward and reverse primers were as follows for this assay: Forward 5'- ACG TTT CCA CGG GCG A -3', Reverse 5'-GCC CTA GTG TGA AAC CCT TCT-3'. This amplifies hg18 region chr11:1977715-1977936. Amplification was performed with the following conditions: 15 minutes at 95^0C, thirty-five times 30 seconds 94^0C, 60 seconds 55^0C, 30 seconds 72^0C with a final step of 3 minutes at 72^0C.

The complete list of SNPs, their biological significance, success rate, the minor allele frequency and the test for Hardy-Weinberg equilibrium is given in supplemental table S3 and S4. In short, for four *H19* and three *INSIGF* SNPs no iPLEX probe design was possible due to the close proximity of other SNPs. Three SNPs were not polymorphic in this Dutch population, one SNP had a lower than 95% success rate, and two SNPs were out of Hardy-Weinberg (P<0.002) according to HaploView and were thus discarded from the analyses. Since not meeting the Hardy-Weinberg criterion can be a sign of selection we tested these two SNPs (rs4320932 and rs4341514) for frequency differences between the exposed and unexposed, but found none (P>0.2). After checking the LD structure in this population of the successfully measured SNPs (**Figure 1**), several SNP were found to be in very high pair-wise LD (R^2>0.9), allowing us to restrict the number of SNPs to test. This

resulted in a final set of four *H19* and twelve SNPs in *IGF2* and *INS* that captured the common genetic variation at *IGF2/H19.* The sixteen SNPs either occurred in CpG dinucleotides themselves ('CG SNPs') or were in LD with such SNPs limiting the possibilities to contrast CG SNPs with non-CG SNPs.

Statistics

We tested for within-pair differences in DNA methylation between exposed individuals and their non-exposed same-sex sibling by applying linear mixed models. With these models the correlation between adjacent CpG sites can be taken into account and all available raw but incomplete data can be used for modeling and control for possible confounders. For an amplicon the difference in DNA methylation between siblings was tested by entering as fixed effects the exposure status (exposed vs. unexposed), a unique identifier for each CpG site within the locus, age at blood draw and the bisulfite batch. To specify a within-sib-pair design, a family (pair) identifier was included as a random effect with intercept. To model the correlation in DNA methylation within an individual we make use of the fact that each family consists of an exposed and same-sex sibling control, therefore adding an exposure status to the random effect as a random slope. This in effect functions as if the individual identifier was added as an additional random effect in our design. This model option allows us to use the same model for both multiple CpG sites and single CpG sites, allowing a unified statistical analysis on all data. The REML likelihood method was used for the model fitting. The difference in DNA methylation for individual CpG sites was calculated with the same model but without the identifier for CpG site. This single CpG linear mixed model yields the same outcome as a paired t-test if no data are missing and no correction for covariates such as age are applied. Some DMRs show an association with age[11], since half of the pairs have a sibling born before the war and half after we could correct for this possible confounder.
The average DNA methylation level of loci was computed using imputed estimates for missing values, since calculating the average with missing

values can skew the average and estimate of the normal variation in the population because of the sometimes large differences in DNA methylation levels between different CpG dinucleotides within a locus. Imputed values were retrieved from the same linear mixed models, which can estimate methylation of CpG sites if data is missing, using information of other CpG sites and the covariates (bisulfite batch, sex and age). The imputed values were never used for any of the analyses. We did not observe significant differences in variance in DNA methylation between the exposed and unexposed (Levene's test for homogeneity $P>0.05$).

To test for associations between DNA methylation and a genotype, the model was extended with the genotype as a fixed effect with the genotype coded as 0 (for common allele homozygous), 1 (heterozygous) or 2 (homozygous for the rare allele) and added as a continuous variable. Finally, to test for interactions between famine exposure and genetic variation an interaction term was added to the model as a fixed effect. In all cases the main effects were also included in the model. We also tested for an interaction between prenatal famine exposure and sex on DNA methylation of all DMRs, but no significant interactions were found, except for *INSIGF* as previously reported[29].

Differences in genotype frequency between the exposed individuals and their unexposed siblings were evaluated using Chi-square test. All analyses were performed in SPSS version 17.0. To visualize correlations in DNA methylation between individual CpG containing fragments and the significance of associations, heatmaps were generated in R version 2.12.1 using the "heatmap.2" function of the gplots package. Multiple testing correction was performed according to the method developed by Benjamini and Hochberg, better known as 'FDR' (false discovery rate) correction using the R base 'p.adjust' function.

Acknowledgements

We express our gratitude to the participants of the Dutch Hunger Winter Families study, the staff of TNO Quality of Life for contact tracing, the staff of the Gerontology and Geriatrics Study Center at the Leiden University Medical Center for performing the clinical examinations and the Central Clinical Chemical Laboratory for extracting genomic DNA.

References

1. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet. 2003; 33 Suppl: 245-54
2. Bogdarina I, Welham S, King PJ, Burns SP, Clark AJ. Epigenetic modification of the renin-angiotensin system in the fetal programming of hypertension. Circ.Res. 2007; 100: 520-6
3. Carone BR, Fauquier L, Habib N, Shea JM, Hart CE, Li R, et al. Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. Cell 2010; 143: 1084-96
4. Waterland RA, Michels KB. Epigenetic epidemiology of the developmental origins hypothesis. Annu.Rev.Nutr. 2007; 27: 363-88
5. Gertz J, Varley KE, Reddy TE, Bowling KM, Pauli F, Parker SL, et al. Analysis of DNA methylation in a three-generation family reveals widespread genetic influence on epigenetic regulation. PLoS.Genet. 2011; 7: e1002228
6. Lienert F, Wirbelauer C, Som I, Dean A, Mohn F, Schubeler D. Identification of genetic elements that autonomously determine DNA methylation states. Nat.Genet. 2011; 43: 1091-7
7. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nat.Rev.Genet. 2011; 12: 529-41
8. Heijmans BT, Kremer D, Tobi EW, Boomsma DI, Slagboom PE. Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus. Hum.Mol.Genet. 2007; 16: 547-54
9. Murrell A, Heeson S, Cooper WN, Douglas E, Apostolidou S, Moore GE, et al. An association between variants in the IGF2 gene and Beckwith-Wiedemann syndrome: interaction between genotype and epigenotype. Hum. Mol.Genet. 2004; 13: 247-55
10. Ollikainen M, Smith KR, Joo EJ, Ng HK, Andronikos R, Novakovic B, et al. DNA methylation analysis of multiple tissues from newborn twins reveals both genetic and intrauterine components to variation in the human neonatal epigenome. Hum.Mol.Genet. 2010; 19: 4176-88
11. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc. Natl.Acad.Sci.U.S.A 2008; 105: 17046-9
12. Steegers-Theunissen RP, Obermann-Borst SA, Kremer D, Lindemans J, Siebel C, Steegers EA, et al. Periconceptional maternal folic acid use of 400 microg per day is related to increased methylation of the IGF2 gene in the very young child. PLoS.One. 2009; 4: e7845
13. Edwards CA, Ferguson-Smith AC. Mechanisms regulating imprinted genes in clusters. Curr.Opin.Cell Biol. 2007; 19: 281-9
14. Ideraabdullah FY, Vigneau S, Bartolomei MS. Genomic imprinting mechanisms in mammals. Mutat.Res. 2008; 647: 77-85
15. Moore GE, Abu-Amero SN, Bell G, Wakeling EL, Kingsnorth A, Stanier P, et al. Evidence that insulin is imprinted in the human yolk sac. Diabetes 2001; 50: 199-203
16. Ferguson LA, Docherty HM, MacKenzie AE, Docherty K. An engineered zinc finger protein reveals a role for the insulin VNTR in the regulation of the insulin and adjacent IGF2 genes. FEBS Lett. 2009; 583: 3181-6
17. Monk D, Sanches R, Arnaud P, Apostolidou S, Hills FA, Abu-Amero S, et al. Imprinting of IGF2 P0 transcript and novel alternatively spliced INS-IGF2 isoforms show differences between mouse and human. Hum.Mol.Genet. 2006; 15: 1259-69
18. Yang BT, Dayeh TA, Kirkpatrick CL, Taneera J, Kumar R, Groop L, et al. Insulin promoter DNA methylation correlates negatively with insulin gene expression and positively with HbA(1c) levels in human pancreatic islets. Diabetologia 2011; 54: 360-7
19. Cui H, Onyango P, Brandenburg S, Wu Y, Hsieh CL, Feinberg AP. Loss of imprinting in colorectal cancer linked to hypomethylation of H19 and IGF2. Cancer Res. 2002; 62: 6442-6
20. Cui H, Cruz-Correa M, Giardiello FM, Hutcheon DF, Kafonek DR, Brandenburg S, et al. Loss of IGF2 imprinting: a potential marker

of colorectal cancer risk. Science 2003; 299: 1753-5

21. Du M, Beatty LG, Zhou W, Lew J, Schoenherr C, Weksberg R, et al. Insulator and silencer sequences in the imprinted region of human chromosome 11p15.5. Hum.Mol.Genet. 2003; 12: 1927-39

22. Dejeux E, Olaso R, Dousset B, Audebourg A, Gut IG, Terris B, et al. Hypermethylation of the IGF2 differentially methylated region 2 is a specific event in insulinomas leading to loss-of-imprinting and overexpression. Endocr. Relat Cancer 2009; 16: 939-52

23. Takai D, Gonzales FA, Tsai YC, Thayer MJ, Jones PA. Large scale mapping of methylcytosines in CTCF-binding sites in the human H19 promoter and aberrant hypomethylation in human bladder cancer. Hum.Mol.Genet. 2001; 10: 2619-26

24. Lumey LH, Stein AD, Kahn HS, van der Pal-de Bruin KM, Blauw GJ, Zybert PA, et al. Cohort profile: the Dutch Hunger Winter families study. Int.J.Epidemiol. 2007; 36: 1196-204

25. Boissonnas CC, Abdalaoui HE, Haelewyn V, Fauque P, Dupont JM, Gut I, et al. Specific epigenetic alterations of IGF2-H19 locus in spermatozoa from infertile men. Eur.J.Hum. Genet. 2010; 18: 73-80

26. Nativio R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, Woodfine K, et al. Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. PLoS.Genet. 2009; 5: e1000739

27. Otte K, Choudhury D, Charalambous M, Engstrom W, Rozell B. A conserved structural element in horse and mouse IGF2 genes binds a methylation sensitive factor. Nucleic Acids Res. 1998; 26: 1605-12

28. Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. Proc.Natl.Acad.Sci.U.S.A 2005; 102: 15785-90

29. Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD, et al. DNA Methylation differences after exposure to prenatal famine are common and timing- and sex-specific. Hum.Mol.Genet. 2009; 18: 4046-53

30. Wang L, Wang F, Guan J, Le J, Wu L, Zou J, et al. Relation between hypomethylation of long interspersed nucleotide elements and risk of neural tube defects. Am.J.Clin.Nutr. 2010; 91: 1359-67

31. Lumey LH, Terry MB, Delgado-Cruzata L, Liao Y, Wang Q, Susser E, et al. Adult global DNA methylation in relation to pre-natal nutrition. Int.J.Epidemiol. 2012; 41: 116-23

32. Bennett ST, Lucassen AM, Gough SC, Powell EE, Undlien DE, Pritchard LE, et al. Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. Nat.Genet. 1995; 9: 284-92

33. Coolen MW, Statham AL, Qu W, Campbell MJ, Henders AK, Montgomery GW, et al. Impact of the genome on the epigenome is manifested in DNA methylation patterns of imprinted regions in monozygotic and dizygotic twins. PLoS.One. 2011; 6: e25590

34. Breton CV, Byun HM, Wenten M, Pan F, Yang A, Gilliland FD. Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation. Am.J.Respir.Crit Care Med. 2009; 180: 462-7

35. Hoyo C, Murtha AP, Schildkraut JM, Jirtle RL, Demark-Wahnefried W, Forman MR, et al. Methylation variation at IGF2 differentially methylated regions and maternal folic acid use before and during pregnancy. Epigenetics. 2011; 6: 928-36

36. Waterland RA, Kellermayer R, Laritsky E, Rayco-Solon P, Harris RA, Travisano M, et al. Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. PLoS.Genet. 2010; 6: e1001252

37. Stoger R. The thrifty epigenotype: an acquired and heritable predisposition for obesity and diabetes? Bioessays 2008; 30: 156-66

38. Murphy SK, Adigun A, Huang Z, Overcash F, Wang F, Jirtle RL, et al. Gender-specific methylation differences in relation to prenatal exposure to cigarette smoke. Gene 2012; 494: 36-43

39. Bouchard L, Thibault S, Guay SP, Santure M, Monpetit A, St-Pierre J, et al. Leptin gene epigenetic adaptation to impaired glucose metabolism during pregnancy. Diabetes Care 2010; 33: 2436-41

40. Hoyo C, Fortner K, Murtha AP, Schildkraut JM, Soubry A, Demark-Wahnefried W, et

al. Association of cord blood methylation fractions at imprinted insulin-like growth factor 2 (IGF2), plasma IGF2, and birth weight. Cancer Causes Control 2012; 23: 635-45

41. Talens RP, Jukema JW, Trompet S, Kremer D, Westendorp RG, Lumey LH, et al. Hypermethylation at loci sensitive to the prenatal environment is associated with increased incidence of myocardial infarction. Int.J.Epidemiol. 2012; 41: 106-15

42. Chotalia M, Smallwood SA, Ruf N, Dawson C, Lucifero D, Frontera M, et al. Transcription is required for establishment of germline methylation marks at imprinted genes. Genes Dev. 2009; 23: 105-17

43. Waterland RA, Kellermayer R, Rached MT, Tatevian N, Gomes MV, Zhang J, et al. Epigenomic profiling indicates a role for DNA methylation in early postnatal liver development. Hum.Mol.Genet. 2009; 18: 3026-38

44. Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. Am.J.Hum.Genet. 2011; 88: 450-7

45. Talens RP, Boomsma DI, Tobi EW, Kremer D, Jukema JW, Willemsen G, et al. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. FASEB J. 2010; 24: 3135-44

46. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol.Cell Biol. 2003; 23: 5293-300

47. Morgan HD, Jin XL, Li A, Whitelaw E, O'Neill C. The culture of zygotes to the blastocyst stage changes the postnatal expression of an epigentically labile allele, agouti viable yellow, in mice. Biol.Reprod. 2008; 79: 618-23

48. Heijmans BT, Tobi EW, Lumey LH, Slagboom PE. The epigenome: archive of the prenatal environment. Epigenetics. 2009; 4: 526-31

49. Heijmans BT, Mill J. Commentary: The seven plagues of epigenetic epidemiology. Int.J.Epidemiol. 2012; 41: 74-8

50. Zhang D, Cheng L, Badner JA, Chen C, Chen Q, Luo W, et al. Genetic control of individual differences in gene-specific methylation in human brain. Am.J.Hum.Genet. 2010; 86: 411-9

51. Gibbs JR, van der Brug MP, Hernandez DG, Traynor BJ, Nalls MA, Lai SL, et al. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. PLoS. Genet. 2010; 6: e1000952

52. Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol. 2011; 12: R10

53. Bell CG, Finer S, Lindgren CM, Wilson GA, Rakyan VK, Teschendorff AE, et al. Integrated genetic and epigenetic analysis identifies haplotype-specific methylation in the FTO type 2 diabetes and obesity susceptibility locus. PLoS.One. 2010; 5: e14040

54. Nagaya K, Makita Y, Taketazu G, Okamoto T, Nakamura E, Hayashi T, et al. Paternal allele of IGF2 gene haplotype CTG is associated with fetal and placental growth in Japanese. Pediatr.Res. 2009; 66: 135-9

55. Gaunt TR, Cooper JA, Miller GJ, Day IN, O'Dell SD. Positive associations between single nucleotide polymorphisms in the IGF2 gene region and body mass index in adult males. Hum.Mol.Genet. 2001; 10: 1491-501

56. Heude B, Petry CJ, Pembrey M, Dunger DB, Ong KK. The insulin gene variable number of tandem repeat: associations and interactions with childhood body fat mass and insulin secretion in normal children. J.Clin. Endocrinol.Metab 2006; 91: 2770-5

57. Bennett A, Sovio U, Ruokonen A, Martikainen H, Pouta A, Taponen S, et al. No association between insulin gene variation and adult metabolic phenotypes in a large Finnish birth cohort. Diabetologia 2005; 48: 886-91

58. Maas JA, Mook-Kanamori DO, Ay L, Steegers EA, van Duijn CM, Hofman A, et al. Insulin VNTR and IGF-1 promoter region polymorphisms are not associated with body composition in early childhood: the generation R study. Horm.Res.Paediatr. 2010; 73: 120-7

59. D'Aleo V, Del GS, Groves C, Lupi R, Tancredi M, McCarthy M, et al. INS VNTR class genotype and the function of isolated human islets. Nutr.Metab Cardiovasc.Dis. 2011; 21: e9-11

60. Le SC, Fallin D, Schork NJ, Bougneres P. The insulin gene VNTR is associated with fasting insulin levels and development of juvenile obesity. Nat.Genet. 2000; 26: 444-6

61. Davey SG, Leary S, Ness A, Lawlor DA. Challenges and novel approaches in the epidemiological study of early life influences on later disease. Adv.Exp.Med.Biol. 2009; 646: 1-14

62. Petronis A. Epigenetics as a unifying principle in the aetiology of complex traits and diseases. Nature 2010; 465: 721-7

63. Trienekens G. The food supply in the Netherlands during the Second World War, In: Smith DF, Phillips, eds. Food, science, policy and regulation in the twentieth century. International and comparative perspectives. London: Routledge, 2000: 117-33

64. Burger GCE, Drummond JC, Sandstead HR. Malnutrition and Starvation in Western Netherlands, September 1944-July 1945. The Hague: General State Printing Office, 1948

65. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002; 12: 996-1006

66. Tobi EW, Heijmans BT, Kremer D, Putter H, Delemarre-van de Waal HA, Finken MJ, et al. DNA methylation of IGF2, GNASAS, INSIGF and LEP and being born small for gestational age. Epigenetics. 2011; 6: 171-6

67. Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002; 18: 1427-31

68. Coolen MW, Statham AL, Gardiner-Garden M, Clark SJ. Genomic profiling of CpG methylation and allelic specificity using quantitative high-throughput mass spectrometry: critical evaluation and improvements. Nucleic Acids Res. 2007; 35: e119

69. Thompson RF, Suzuki M, Lau KW, Greally JM. A pipeline for the quantitative analysis of CG dinucleotide methylation using mass spectrometry. Bioinformatics. 2009; 25: 2164-70

70. Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, et al. Integrating common and rare genetic variation in diverse human populations. Nature 2010; 467: 52-8

71. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005; 21: 263-5

72. Adkins RM, Krushkal J, Klauser CK, Magann EF, Morrison JC, Somes G. Association between small for gestational age and paternally inherited 5' insulin haplotypes. Int.J.Obes.(Lond) 2008; 32: 372-80

73. Adkins RM, Somes G, Morrison JC, Hill JB, Watson EM, Magann EF, et al. Association of Birth Weight with Polymorphisms in the IGF2, H19 and IGF2R Genes. Pediatr.Res. 2010; 68: 429-34

74. Petry CJ, Ong KK, Barratt BJ, Wingate D, Cordell HJ, Ring SM, et al. Common polymorphism in H19 associated with birthweight and cord blood IGF-II levels in humans. BMC.Genet. 2005; 6: 22

75. Petry CJ, Rayco-Solon P, Fulford AJ, Stead JD, Wingate DL, Ong KK, et al. Common polymorphic variation in the genetically diverse African insulin gene and its association with size at birth. Hum.Genet. 2009; 126: 375-84

76. Heude B, Ong KK, Luben R, Wareham NJ, Sandhu MS. Study of association between common variation in the insulin-like growth factor 2 gene and indices of obesity and body size in middle-aged men and women. J.Clin. Endocrinol.Metab 2007; 92: 2734-8

77. O'Dell SD, Bujac SR, Miller GJ, Day IN. Associations of IGF2 ApaI RFLP and INS VNTR class I allele size with obesity. Eur.J.Hum.Genet. 1999; 7: 821-7

78. Adkins RM, Fain JN, Krushkal J, Klauser CK, Magann EF, Morrison JC. Association between paternally inherited haplotypes upstream of the insulin gene and umbilical cord IGF-II levels. Pediatr.Res. 2007; 62: 451-5

DNA methylation signatures link prenatal famine exposure to growth and metabolism

Elmar W.Tobi[1], Jelle J.Goeman[3] , Ramin Monajemi[3], Hongcang Gu[8,9], Hein Putter[3], Yanju Zhang[1], Fabian Müller[4] , Erik. W. van Zwet[3], Christoph Bock[5,6], Alexander Meissner[8,9], L.H.Lumey[1,2] , P.Eline Slagboom[1,7], Bastiaan T.Heijmans[1,7]

1. Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
2. Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA
3. Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands
4. Computational Biology and Applied Algorithmics, Max Planck Institute for Informatics, Saarbrücken, Germany
5. Research Center for Molecular Medicine (CeMM), Austrian Academy of Sciences, Vienna, Austria
6. Department of Medical and Chemical Laboratory Diagnostics, University of Vienna, Vienna, Austria
7. The Netherlands consortium for Healthy Ageing, Leiden University Medical Center, Leiden, The Netherlands
8. Harvard University of Stem cell and Regenerative Biology, Cambridge, MA, USA
9. The Broad Institute of MIT and Harvard, MA, USA

Manuscript submitted

Abstract

Environmental conditions during embryonic development, such as malnutrition, have been linked to adult health. It has been suggested that epigenetic mechanisms may form the molecular link between development and later health. For instance, we have shown that prenatal famine exposure to the Dutch Hunger Winter, a severe war-time famine at the end of WW2, is associated with DNA methylation changes at candidate loci. However, a genome-wide assessment of which genomic annotations and pathways are sensitive to prenatal famine and the relation of such differential methylated regions (DMRs) with phenotypes of interest in relation to prenatal famine exposure is currently lacking.

We generated genome-scale DNA methylation data at 1.2M CpG sites for 24 individuals exposed to the Dutch Famine during early gestation and 24 same-sex unexposed sibling controls. Differential DNA methylation after famine exposure was observed at regions with a regulatory potential. Of the 181 prenatally induced differentially methylated regions (P-DMRs) 60.7% were located in gene-bodies and entailed differences of ~5% between exposed individuals and controls. Validation of six P-DMRs in all available individuals exposed during early gestation (N=60) and unexposed same-sex sibling controls (N=60) showed an attenuation of differential methylation in individuals conceived during April and May 1945, the last 2 months of the famine. The modest differences at individual regions extended to their biological pathways. Methylation differences were predominantly related to the regulation of growth and lipid and cholesterol metabolism. In our study group with early gestational famine exposure, DNA methylation of *INSR* was also associated with birth weight and *CPT1A* methylation with LDL cholesterol.

Prenatal exposure to famine may affect regulatory regions. Although the observed differences at individual regions are modest, they extend across biologically relevant pathways and could potentially explain phenotypes previously associated with prenatal famine exposure.

Introduction

It has been hypothesized that epigenetic mechanisms, including DNA methylation, underlie the link between environmental exposures during early development and adult phenotypes[1]. Experiments in animal models showed that prenatal malnutrition during early development can result in persistent changes in DNA methylation at candidate regions including promoters[2], retrotransposons[3] and imprinted regions[4]. These can subsequently be transmitted during cell differentiation and growth and explain phenotypic variation. It remains unclear whether prenatally-induced differential methylation contributes to the risk of disease in humans.

The quasi-experimental setting of well-documented historical famines represents a powerful opportunity to study this question in humans. Exposure to the Dutch Hunger Winter, a severe war-time famine at the end of WW2, is associated with an adverse metabolic profile consisting of a decreased insulin response, higher body mass index (BMI)[5], and elevated total and LDL cholesterol[6,7], but also an increased risk on schizophrenia[8]. The study of the Dutch Famine, also called the Hunger Winter, is of interest as it offers a clear developmental exposure[5]. More over, its relevance may extend to modern exposures since maternal smoking[9] and gestational diabetes[10] have been associated to similar health outcomes as prenatal famine exposure.

We previously linked prenatal famine exposure to differential DNA methylation at candidate regions, including promoters and imprinted regions[11,12]. Of interest, DNA methylation at a subset of these regions was also sensitive to maternal smoking[13] and gestational diabetes[14]. In line with the observation that the epigenome is particularly dynamic around conception and embryo implantation[15], differential methylation was predominantly observed after exposure to famine during early gestation. To further understand the potential role of prenatally-induced differential methylation in human disease, a genome-wide approach is required to uncover the characteristics of affected genomic regions and annotate the function of the genes and biological pathways involved.

We used Reduced Representation Bisulfite Sequencing[16] to generate DNA methylation data on 1.2M individual CpG dinucleotides in individuals exposed to famine in early gestation and an unexposed same-sex sibling as control. Using a step-wise analysis strategy involving genomic annotation and technical and biological validation, we identified genomic characteristics of prenatally-induced differentially methylated regions (P-DMRs)[17]. Moreover, we also identified pathways affected by prenatal famine exposure and describe individual P-DMRs that are associated with phenotypic outcomes associated with early gestational famine exposure.

Results

We studied 24 same-sex sibling pairs (N=48) from a larger group of 60 pairs of which one sibling was exposed to the Dutch Famine around conception and up to 10 weeks into development (i.e. periconceptional exposure). The subset was constrained to siblings within the ongoing Dutch Hunger Winter Families Study[18] with an age difference of less than 5 years and included an equal number of male and female pairs as well as an equal number of pairs with the control siblings conceived and born before or after the Famine (**Supplement II, Table S1**). We employed Reduced Representation Bisulfite Sequencing (RRBS) on DNA from whole blood to obtain single nucleotide high resolution DNA methylation data on a genome-wide scale[16]. Sequencing was performed on an Illumina GAIIx, utilizing one sequencing lane per individual. This resulted in 25.6 million high quality reads on average per sample (**Supplement II, Table S1 and S2**). Exclusion of CpG sites with a low or extremely high coverage (≤5-fold or >200) and those that were uninformative (mapping to a random chromosome, median methylation=0% or 100%) resulted in 1.206.161 unique CpG sites with an average median sequencing depth of 28 and the data for a CpG site was on average for 99.8% complete (**Supplement II, Table S2, Figure S1A and Figure S1B**). The data showed the classical bimodal nature of DNA methylation (**Supplement II, Figure S2**)

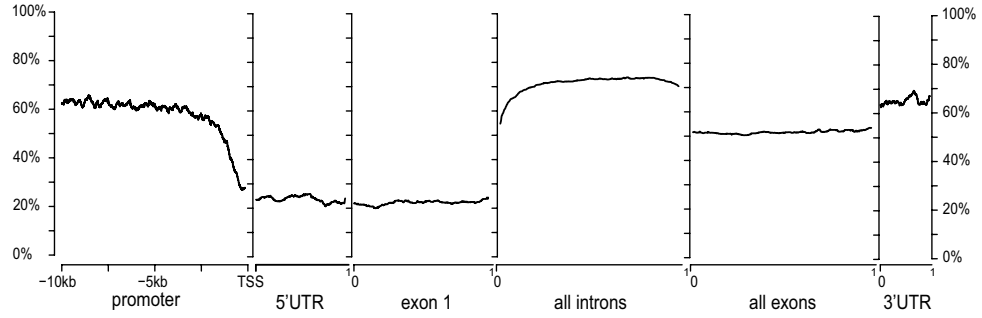


Figure 1. The methylation level of genes in the RRBS dataset

The methylation level across a gene. A lowess has been fit across all data for all entrez genes. The width of the gene elements represents the relative amount of data for such elements in the total dataset.

with promoter and intragenic methylation levels matching those observed in the comprehensive blood methylome[19] (**Figure 1**).

Genome-scale data corroborates previous findings

In previous studies of prenatal famine exposure[20], in particular during early gestation[21], we did not observe differences in global methylation between exposed individuals and sibling controls. Again no difference was found as the mean global methylation extracted from all sequenced reads was 42.55% in the exposed and 42.64% in the sibling controls (Paired t-test, P=0.91, **Supplement II, Figure S3**). Since the average methylation of retrotransposons is regularly taken as a measure of global methylation, we also tested the average DNA methylation of 993 RepBase Update database[22] repeat sequences, again finding no associations (P_{FDR}>0.26). The genome-scale data covered the 16 candidate regions we reported on previously[11,12]. The regions originally found to be differentially methylated after periconceptional famine exposure were also associated in the current data set (P=9.0x10^{-3}). Likewise, no association was observed for regions at which no differential methylation was observed previously (P=0.19), validating our genome-scale approach.

Differential methylation occurs at open chromatin regions

We selected 28 genomic annotations from the literature, mapped the CpG dinucleotides to these annotations and evaluated at which of these annotations DNA methylation was susceptible to early gestational famine exposure. DNA methylation at 5 annotations was associated with famine exposure, namely non-CGI promoters with an open chromatin conformation ('bonafide non-CGI promoters')[23], enhancers[24], exons, DNaseI/FAIRE-seq open chromatin regions, and enhancers active during the pre- and peri-implantation period[43] (P_{FDR}<0.05; **Table 1**). Apart from exons these regions represent potential regulatory regions. Annotations not associated with famine exposure included CGIs, putative meta-stable epialleles[25] and highly variable DNA methylation

Table 1. Outcome of genome-scale testing of genomic annotations

Genomic annotations*	Regions in genome	Covered	$P_{nominal}$	P_{FDR}
Non-CGI, non 'bonafide' promoters[1]	7,014	2,024 (28%)	$9.1x10^{-4}$	0.026
Enhancers[2]	59,466	6,207	$1.9x10^{-3}$	0.026
DNaseI/FAIRE-seq regions[3]	590,252	79,728	$4.4x10^{-3}$	0.036
Middle exons	17,848	1,570	$5.8x10^{-3}$	0.036
Developmental enhancers type I[4]	5,118	922	$6.5x10^{-3}$	0.036
'bonafide' CGI – shores[5]	88,871	27,688	0.012	0.053
Non-coding RNA[6]	718	59	0.015	0.053
Conserved regions[7]	165,937	1,386	0.016	0.053
CGI –shores	319,509	67,811	0.017	0.053
3'UTR	21,004	2,909	0.035	0.085
Non genic CGI[8]	129,049	41,023	0.036	0.085
'Bonafide' CGI – border	88,074	22,777	0.036	0.085
Developmental enhancer type II	2,287	320	0.078	0.15
CGI	343,925	113,673	0.078	0.15
Introns	201,640	61,816	0.080	0.15
hESC bivalent chromatin domains	1,797	1,741	0.16	0.28
Bonafide CGI	44,439	35,271	0.20	0.32
cell-type specific gene promoters	2,372	2,106	0.21	0.32
First exons	51,497	13,507	0.25	0.36
Promoters	23,689	16,904	0.26	0.36
HSC bivalent chromatin domains	2,910	2,779	0.28	0.36
Imprinted promoters	46	42	0.29	0.36
'Bonafide' CGI promoter	16,674	14,880	0.32	0.37
CTCF insulators from CD4+ cells	28,661	4,396	0.32	0.37
Imprinted DMRs	14	6	0.33	0.37
Putative Metastable epialles	38	29	0.43	0.47
Variably methylated regions	227	56	0.55	0.57
Promoters – cancer genes	888	795	0.63	0.63

1. Promoters without CGIs but with a relatively open chromatin state[23]
2. Enhancers characterized by H3K4me1, non-overlapping with promoters[24]
3. Regions with an open Chromatin state as defined by DNaseI and FAIRE-seq signals (UCSC track ENCODE)
4. Enhancers active during first stages of blastocyst development[44]
5. Shores of so-called bonafide CGI, CGI island with an ubiquitously open chromatin structure; oe>0.6, GC%>50% & length >700bp[23]
6. Body of various type of non-coding RNAs
7. Conserved regions outside promoters, CGIs, exons and UTR
8. CGI >10kb from gene

* Details on the genomic annotations can be found in the materials and methods.

regions (VMR)[26]. There was no difference between the average GC content of the 5 significant annotations and the non-significant genomic annotations (t-test, P=0.79).

Characteristics of individual DMRs associated with prenatal famine

Next, the 90,451 individual regions within the 5 associated genomic annotations were tested (median size= 1.4 kb; mean number of CpG dinucleotides= 6.7). This analysis identified 181 regions as prenatally induced differentially methylated regions (P-DMRs, P_{FDR}<0.05). The difference in DNA methylation between exposed individuals and their same-sex sibling was variable at these regions (up to >10%), but generally moderate (median 4.6%; **Figure 2**). DNA methylation at P-DMRs was more commonly higher (60.8%) than lower among exposed individuals which suggests that differential methylation observed is not a simple one-to-one consequence of a possible methyl donor deficiency in the rations during the Famine.

Of the 181 P-DMRs, 60.7% was located in gene bodies, 11.6% in promoters, 10.5% in upstream genic regions, 10.0% in downstream genic regions and 7.2% were intergenic. Further characterization with EPIGRAPH[27] showed an increased co-occurrence with histone marks associated with active

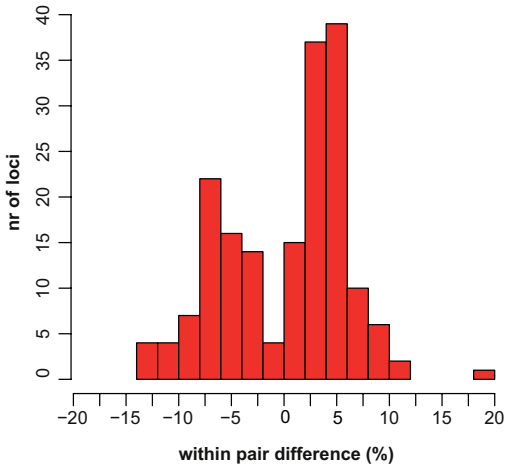


Figure 2. The average within pair difference for the 181 loci surviving multiple testing

enhancers, transcribed gene bodies, active regulatory sites and expressed exons (H3K4me1-3, H2AZ, H3K9me1, PolII, H3K79me1, H3K27me1)[28-30], while a lower co-occurrence with SINES and a lower overall repeat score was observed (P_{FDR}<0.05). Interestingly, although the putative P-DMRs were identified on the basis of whole blood samples, the genes nearest to the P-DMRs were not enriched for tissue-specific gene expression, including blood and bone marrow (OR=0.98 [95% CI:0.59-1.64] and OR=1.01 [95% CI:0.55-1.87], respectively). In contrast, the P-DMRs were enriched for genes differentially expressed during the pre-implantation stage of development (OR=4.95 [95% CI:3.53-6.49])[31] and organogenesis (OR=4.76 [95% CI: 3.11- 7.30])[32] in humans.

Validation of genome-scale measurements

Findings from genome-scale DNA methylation studies require validation with an independent technology. From the 181 putative P-DMRs, we prioritized 11 regions with the lowest p-values and 9 regions according to consistency of DNA methylation differences across CpG dinucleotides and mapping to genes with known functions. For 19/20 regions a working assay could be designed using the mass spectrometry based method Epityper (**Supplement II, Table S5**)[33]. The overall correlation between the average DNA methylation measured using the genome-wide method RRBS and the locus-specific method Epityper was good (r=0.81) (**Figure 3**), despite the fact that the length of RRBS regions defined by genome annotations were larger (>=1 kb) than the regions targeted by Epityper assays (<560bp [as 600bp is the technical limit]). Of the 19 regions, 13 were again associated with famine (P<0.05, **Supplement II, Table S3**). A correspondence between the RRBS and the Epityper measurements of individual regions was defined as a Pearson's correlation greater than 0.7 and was found for 6 loci mapping to *CDH23*, *SMAD7*, *INSR*, *CPT1A*, *KLF13* and *RFTN1* (**Figure 3**). These 6 Epityper assays were taken further for validation in the complete set of 60 periconceptionally exposed individuals and their same-sex sibling controls.

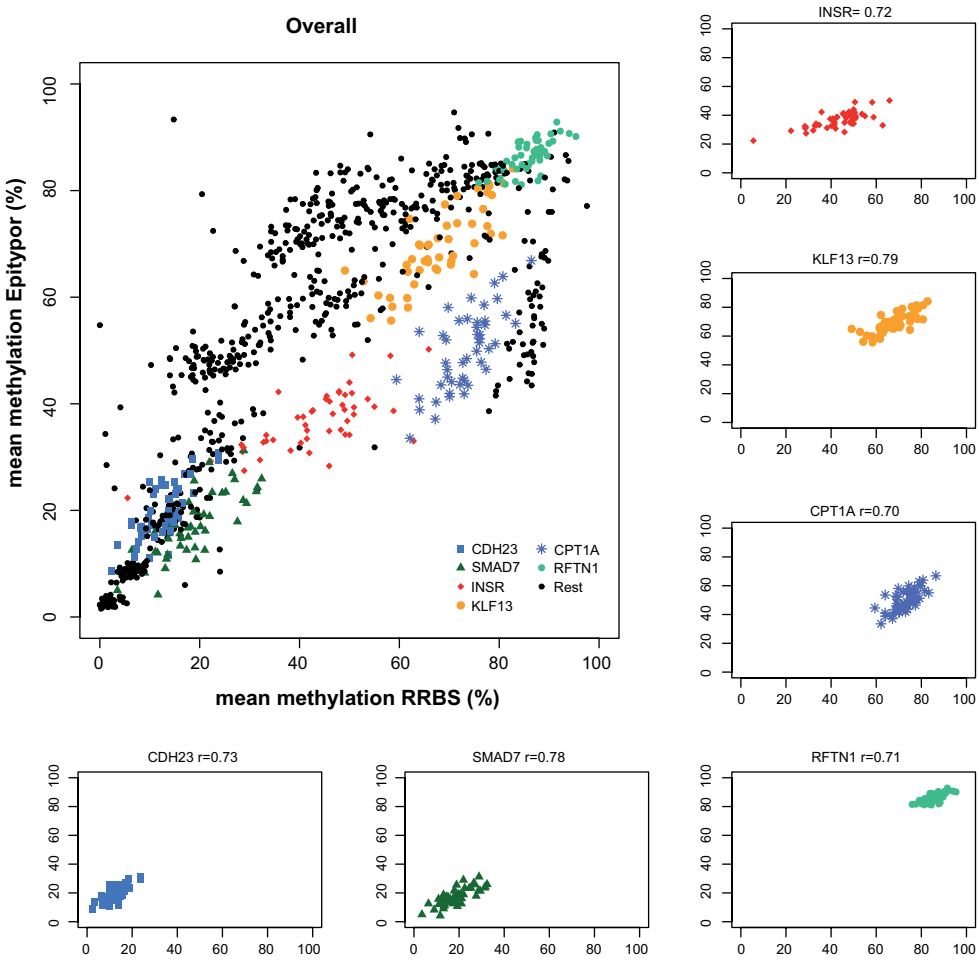


Figure 3. The relationship between the average methylation of the RRBS regions and the Epityper measurements.

The correlation between the average methyation of the RRBS regions and the Epityper measurements of smaller subsets of these regions. Individual regions with a Pearson correlation >0.7 are denoted in color and plotted separately along the main figure. The correlation of the other loci can be found in Supplement II, Table S3.

The associations were consistent for individual CpG dinucleotides across the P-DMRs (**Supplement II, Table S4**), although the within pair differences observed in these 60 sibling pairs using the Epityper assay were somewhat attenuated as compared with those observed in the genome-scale data (**Table 2**). The associations were not affected by age, smoking (neither

current smoking nor package years), socio-economic status (SES) and current diet (kcal/day nor the percentage of fat, carbohydrates or protein in the diet), indicating that the DMRs were independent of current, post-natal environmental exposures of the individuals studied. Lastly, we excluded cellular heterogeneity of whole blood samples as a confounding factor by exploring the potential association of DNA methylation at the 6 regions with blood cell counts in an independent sample set (**Supplement II, Table S7**).

Table 2. Validation of RRBS associations with Epityper

Annotation		Genome-wide (N=48)				Epityper Validation (N=120)	
Type*	Nearest Gene (kb)	Meth. Controls (%)	Within pair diff. (%)	P	P_{FDR}	Within pair diff. (%)	P
1 & 2	*SMAD7* (+25)	21.3	4.2	$1.0x10^{-7}$	$1.1x10^{-3}$	3.2	$2.5x10^{-3}$
1 & 2	*CDH23* (0)	12.4	4.0	$1.3x10^{-7}$	$1.1x10^{-3}$	2.2	$6.3x10^{-3}$
1	*INSR* (0)	43.3	8.1	$3.9x10^{-6}$	0.010	2.0	0.031
1 & 3	*RFTN1* (0)	86.3	-2.3	$3.2x10^{-5}$	0.030	-0.9	0.09
1	*CPT1A* (0)	67.0	4.5	$4.0x10^{-5}$	0.031	2.0	0.05
1	*KLF13* (0)	67.1	-7.9	$6.1x10^{-5}$	0.042	-3.1	0.014

* Type of genomic annotation: 1 = Open chromatin, 2= enhancer, 3 = exon.

Critical window of exposure

It is generally assumed that the influence of a prenatal exposure on DNA methylation depends on its exact timing during development[1]. The gestational timing of the famine exposure is based on the mother's last menstrual period (LMP), a commonly used proxy for the start of pregnancy. To test for an influence of timing on differential methylation we tested for an interaction between the month of the start of pregnancy (month of LMP) and famine exposure.

There was an influence of timing ($P_{interaction}$=0.016) therefore the average within sibling pair difference in DNA methylation was plotted for the 6 P-DMRs against the start of the mother's pregnancy of the prenatally exposed (**Figure 4A**). DNA methylation differences were present throughout the first 4 months

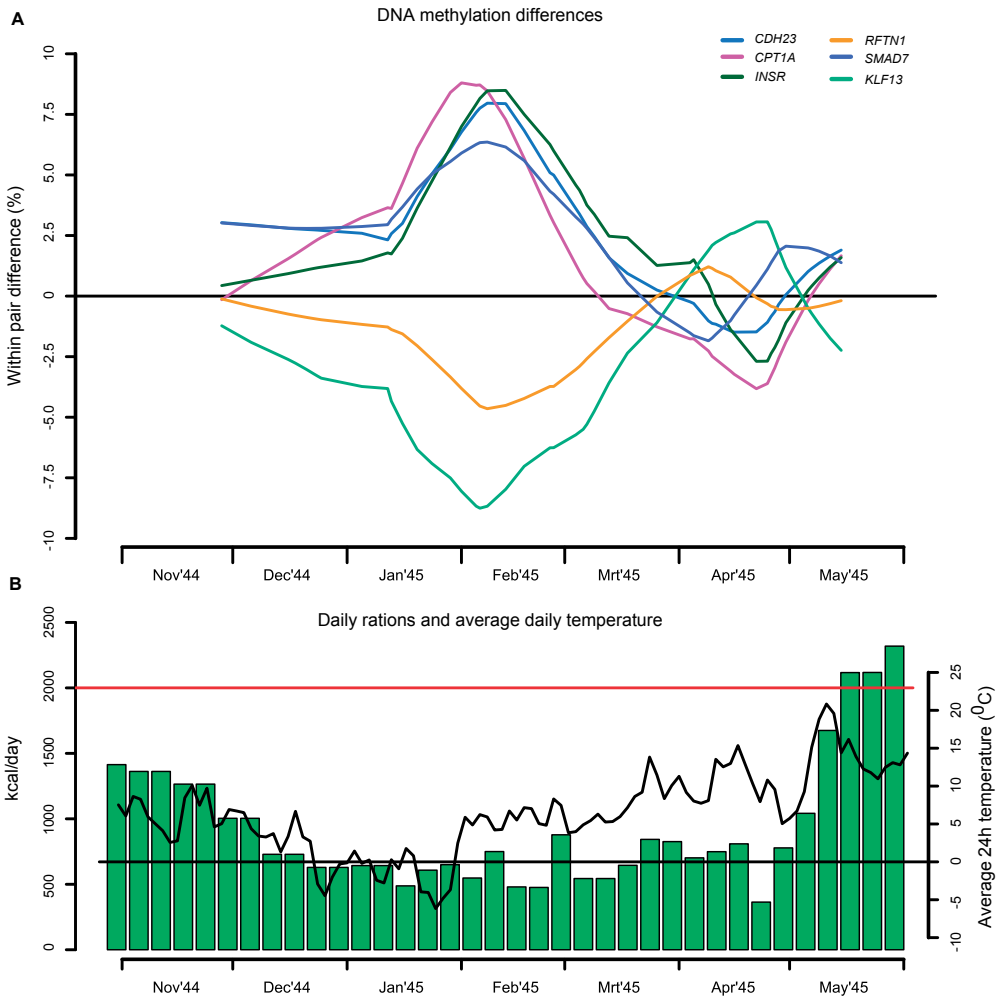


Figure 4. Results across the Famine period

A. A lowess depicting the average within pair difference (y-axis) ass stratified by the estimate of the moment of conception (last menstrual period; x-axis). Each colored line represents a locus.

B. The official daily rations (kcal/day) for each calendar week (green bars) and a lowess depicting the average 24h temperature (black line; source KNMI; DeBilt weather station). The daily requirement of non-pregnant women of 2000kcal/day is denoted in red

of the Famine, but waned towards the end to become virtually absent by the start of April 1945 (test for interaction pre-April and later pregnancies, $P_{interaction}=4.6x10^{-3}$). Re-examination of candidate loci identified as P-DMRs in earlier studies of the same individuals[11,12,21], showed the same attenuation

of DNA methylation differences in pregnancies starting after April 1945 ($P_{interaction}$=1.8x10^{-3}).

These associations with famine exposure for pairs with an exposed sibling conceived before April was consistent. Inspection of the RRBS data in exposed individuals from pre-April pregnancies and sibling controls (n=36 individuals) showed that the DNA methylation differences were virtually identical to those in the complete set of discovery pairs (n=48). These pre-April associations were reliably validated using an independent technology (Epityper). Crucially, the effect sizes in the pre-April exposed individuals and controls not used for discovery (n=36) were remarkably similar to those in the discovery set obtained using the same technology (**Table 3**). Indeed, the replication was high (95% CI:3/6-6/6), confirming the consistency of the data. Conversely, no DNA methylation difference was observed at these regions in exposed individuals conceived later during the Famine as compared to their sibling controls (N=48, P>0.20).

This attenuation may be explained by changes in the conditions towards the end of the Famine, but the rations remained low during the entire famine period and also the daily temperature was not very different in April and May from March (**Figure 4B**). Alternatively, the data may be compatible with intrinsic features of early development and may suggest that DMRs are induced after conception during periods at which large scale DNA methylation changes occur[15] or slowly accumulate during a more prolonged exposure.

The P-DMRs

In all pre-April exposed individuals and sibling controls (n=72), the average within pair differences were 4.2% for *SMAD7* (P=6.8x10^{-4}), 3.6% for *CDH23* (P=3.3x10^{-4}), 3.2% for *INSR* (P=7.1x10^{-3}), -2.1% for *RFTN1* (P=3.6x10^{-3}), 4.1% for *CPT1A* (P=1.9x10^{-3}) and -6.0% for *KLF13* (P=1.7 x10^{-4}). The P-DMRs were located in regions defined by their chromatin state to be an enhancer or isolator and all were located in open chromatin regions according to ENCODE data, indicating these loci may have a regulatory function (**Supplement II, Figure S4A-D**). For instance, the *INSR* P-DMR is located

Table 3. The outcome of a stratified analysis for timing

	Genome-wide <April '45 (N=36)[1]		Technical validation (idem to RRBS)[2]		Replication (N=36)[3]		Epityper April & May '45 (N=48)[4]	
P-DMR	Diff	P	Diff	P	Diff	P	Diff	P
SMAD7	4.3	$2.2x10^{-7}$	3.8	0.010	4.4	0.027	1.6	0.39
CDH23	4.2	$4.5x10^{-6}$	3.6	$1.4x10^{-3}$	3.6	0.034	0.1	0.91
INSR	9.0	$7.4x10^{-7}$	2.5	0.16	3.8	0.016	0.5	0.77
RFTN1	-3.7	$1.1x10^{-7}$	-2.3	$3.5x10^{-3}$	-1.9	0.13	0.8	0.21
CPT1A	4.9	$1.1x10^{-5}$	4.4	0.041	3.8	0.021	-1.1	0.49
KLF13	-8.6	$1.6x10^{-5}$	-5.5	$7.8x10^{-3}$	-6.0	$7.5x10^{-3}$	1.0	0.57

1. Outcome of the famine association for the pairs with one sibling conceived from Nov'44 to Mar'45 (18 pairs, 36 individuals) in the RRBS data.
2. Outcome of the famine associations for the Epityper measurements on the same pairs as measured by RRBS and with a sibling conceived between Nov'44 and Mar'45.
3. Outcome of the famine association for the Epityper measurements for pairs *not* measured by RRBS but with one sibling conceived between Nov'44 and Mar'45 (18 pairs, 36 individuals).
4. Outcome the famine association for the Epityper measurements for all pairs with one sibling conceived in April and May '45 (24 pairs, 48 individuals).

in an intronic enhancer of *INSR* according to ENCODE data and marks a DNaseI hypersensitivity site in 31 different tissues and cell lines (**Figure 6B**) and the P-DMR at *CPT1A* overlapped a weak enhancer in the blood derived GM12787 line and H1 embryonic stem cells. Moreover, it overlaps a strong binding site of the BAF155 transcription factor which is an repressor of self renewal[34] and vital for early liver development[35] (**Figure 6D**).
Indeed the genes in which the P-DMRs are located or are closest have been implicated in development, like eye development (*CDH23* and *RFTN1*), forebrain formation (*SMAD7*[36]), growth (*INSR*) and sustaining early pregnancy (*KLF13*[37]). But beside roles in development most genes also had metabolic functions, including insulin signaling (*INSR*), pancreatic beta cell functioning (*SMAD7*[38]), fatty acid oxidation (*CPTIA*[39]) and cholesterol metabolism (*KLF13*[40]). All in all the P-DMRs are interesting subjects for study.

Famine associations extend into pathways

However, the average within pair differences were modest. It is hypothesized that epigenetic modulation may occur through multiple smaller changes across a gene network[41], a hypothesis that recently gained empirical support[21,42]. So we tested if the association with prenatal famine exposure and DNA methylation extended to the pathways of which the genes belonging to the P-DMRs are a part of. So we re-visited the genome-wide DNA methylation dataset for the pre-April sibling pairs (N=36) and tested for an additional association along the MSigDB pathways and GO Biology terms to which the 6 genes belong to by removing all methylation data mapping to these 6 genes, including the P-DMRs.
DNA methylation of 25 out of 101 gene-sets was associated with famine exposure ($P_{FDR}<0.05$). The three most significant pathways were the GO terms *positive regulation of growth* ($P_{FDR}=5.5x10^{-3}$), *response to activity* ($P_{FDR}=5.6x10^{-3}$) and *regulation of embryonic growth* ($P_{FDR}=0.021$). Four MSigDB terms were significant, namely the KEGG insulin signaling pathway ($P_{FDR}=0.029$), the BIOCARTA *HDAC* pathway ($P_{FDR}=0.032$; involved in myogenesis and cardiac development), the REACTOME *IRS related events*

(P_{FDR}=0.032; encompassing the insulin signaling cascade) and REACTOME *metabolism of lipids and lipoproteins* (P_{FDR}=0.039).

The significant GO terms were clustered based on their relatedness and redundant terms were removed by REVIGO[43]. Visualization of these results showed that the largest cluster was formed from the GO term *regulation of growth* (**Figure 5).** Multiple clusters contained pathways related to lipid and cholesterol metabolism. Indeed, beside the REACTOME term *metabolism of lipids and lipoproteins*, multiple GO terms related to lipid metabolism showed significantly different DNA methylation between exposed individuals and sibling controls (GO: *positive regulation of lipid metabolic process*, P_{FDR}=0.028; *lipid homeostasis*, P_{FDR}=0.042; *triglyceride metabolic process*, P_{FDR}=0.049).

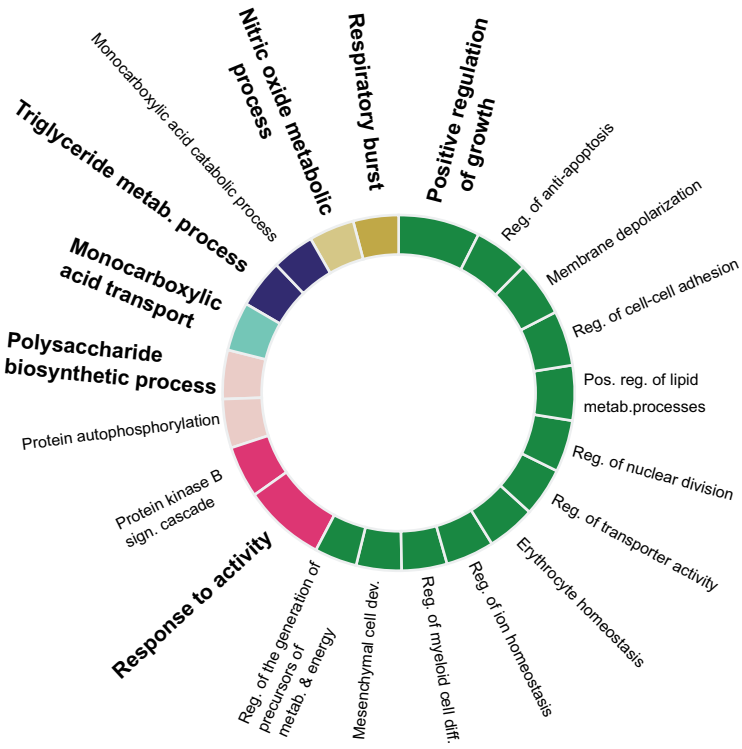


Figure 5. Revigo analysis of the significant pathways

A sunburst graph of the non-redundant clustered FDR significant GO terms associated with prenatal famine exposure. The size of the circular boxes represents the statistical evidence. In bold are the dominant terms of the clusters, which are denoted in different colors.

P-DMRs and phenotypic outcomes

The pathways and P-DMRs are of interest to the reported phenotypic associations with early developmental famine exposure and development in general. Individuals exposed early in gestation were reported to have a higher birth weight than controls[11] and replicated later-life phenotypic outcomes include a higher body mass index (BMI) and an altered glucose response[5] and elevated LDL and total cholesterol levels[6,7]. Phenotypic data for all outcomes except schizophrenia were available and associations between these five phenotypes and methylation at the 6 P-DMRs were explored.

Birth weight data were available for exposed individuals (N=60), but not their sibling controls. In this study group with early gestation exposure, *INSR* methylation was positively correlated with birth weight (R-squared=0.111, $P=9.0x10^{-3}$; **Figure 6A**). The association remained statistically significant after multivariate analysis and after accounting for multiple testing (6 P-DMRs and 5 phenotypic outcomes), gestational age, age at examination, and adult characteristics as SES, diet and smoking ($\beta_{exp\text{-}birth\ weight}$=3.9%/1kg, P_{FDR}=0.033).

The later-life phenotypic outcomes tested were BMI, plasma glucose 120 minutes after an oral glucose tolerance test and LDL and total cholesterol levels. This information was available in both the exposed individuals and the unexposed same-sex siblings (N=120; 60 pairs). *CPT1A* methylation was positively correlated with LDL cholesterol levels (R-squared=0.077, $P=3.5x10^{-3}$, **Figure 6C**), in line with the association of early gestational famine with *CPT1A* methylation and LDL. The statistical association with *CPT1A* methylation and LDL cholesterol remained after correction for multiple testing and further adjustments for age, sex, BMI, SES, smoking and current diet (β_{LDL}=2.4%/mmol*l^{-1}, P_{FDR}=0.033). The observed effects were almost identical in the prenatally exposed individuals and their unexposed same-sex siblings (β_{exp}=2.4%/mmol*l^{-1}; β_{sibs}=2.3%/mmol*l^{-1}), suggesting that *CPT1A* may be a quantitative trait locus for LDL. Exclusion of individuals (N=10) using lipid lowering medication did not affect the association (β_{LDL}=2.5%/mmol*l^{-1}, $P=2.5x10^{-3}$). Similar results were obtained for total cholesterol, but

could be attributed to LDL cholesterol, which was highly correlated with total cholesterol (r=0.92, P=2.2x10^{-16}).

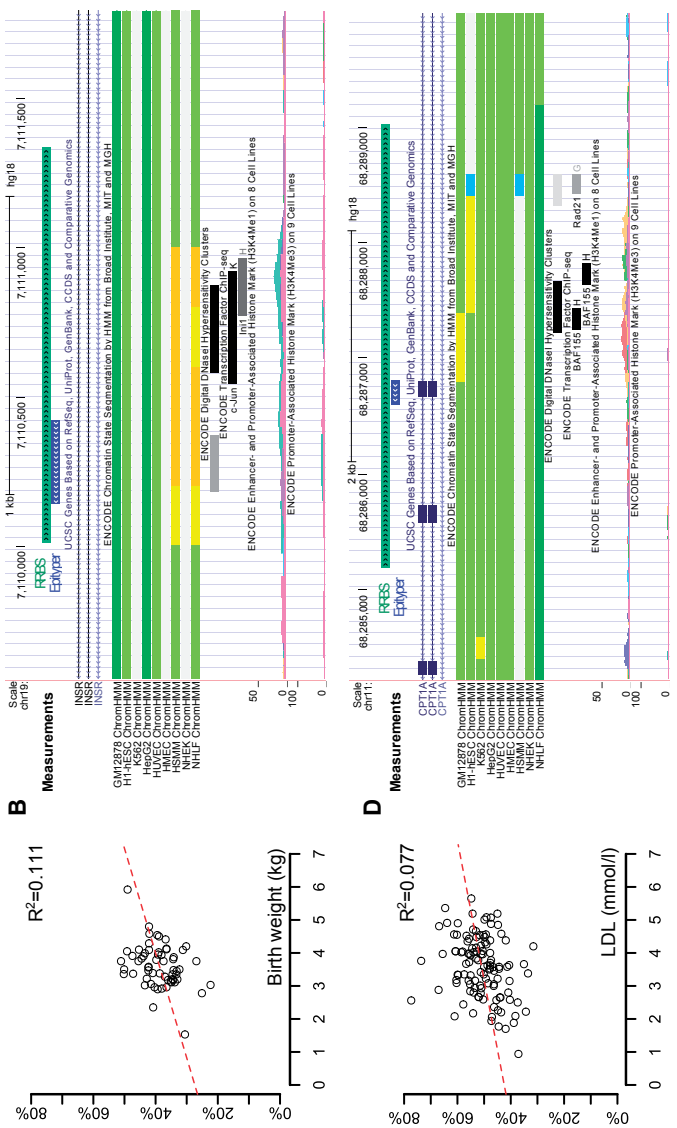


Figure 6. The INSR and CPT1A P-DMRs

A. Scatterplot between birth weight (x-axis) and the average DNA methylation of *INSR* (y-axis) in the 60 prenatally exposed individuals.

B. The genomic annotation of *INSR*. The P-DMR overlaps an enhancer in the HSMM and NHLF cell lines and a DNaseI hypersensitivity cluster in over 30 cell lines.

C. Scatterplot between LDL (x-axis) and the average DNA methylation of *CPT1A* (y-axis) in all 120 siblings.

D. The genomic annotations of *CPT1A*. The P-DMR overlaps an enhancer in the blood derived GM12878 and embryonic stem cell line H1 and a DNaseI hypersensitive cluster in over 30 cell lines. Furthermore, the BAF155 transcription factor binds in this region.

Discussion

We studied the association between famine exposure from peri-conception and into the first trimester and DNA methylation at middle age on a genome-wide scale using reduced representation bisulfite sequencing (RRBS)[16]. DNA methylation at non-CGI promoters[23], DNaseI/FAIRE-seq regions, exons, enhancers, and those enhancers active during the pre- and peri-implantation period[44] was associated with prenatal famine exposure. From these 5 annotations 181 regions were individually associated with famine. These regions were mainly located in gene bodies, co-occurred with histone marks of active chromatin, and were not enriched for tissue specific genes while being enriched for genes active during development. We found evidence that the exact timing of the famine exposure during this developmental time-frame has bearing on DNA methylation levels. We validated P-DMRs mapping to *SMAD7*, *CDH23*, *INSR*, *RFTN1*, *CPT1A* and *KLF13* by an independent methodology and found that associations for individual regions may extend to the pathways they belong to. Indeed, pathways involved in growth and lipid and cholesterol metabolism were associated with famine exposure. We found tentative associations between DNA methylation at an intronic enhancer of *INSR* with birth weight, and intragenic methylation of *CPT1A* with LDL cholesterol.

Genomic annotations affected were not the expected ones

We applied next generation sequencing of bisulfite converted DNA, generating data for 1.2M CpG dinucleotides in 24 exposed individuals and 24 unexposed same-sex sibling controls. This has allowed us to comprehensively study a large array of genomic annotations. Our approach led to discovery that regions with a regulatory potential were especially sensitive to early prenatal famine exposure. The associated regions were mainly located within gene bodies and enhancers, thus located outside the regions and annotations generally queried in epigenetic studies and previously hypothesized to be especially sensitive to prenatal environmental conditions[45].

The nature of the P-DMRs

It is unclear what mechanisms underlie the observed differences in DNA methylation after prenatal famine exposure. We think it is reasonable to entertain a biological explanation. While the prenatal environment may induce P-DMRs that in turn persistently affect gene expression[3], P-DMRs may also reflect altered transcriptional activity during development as it was shown that differential expression during development may influence DNA methylation levels[46,47]. This explanation is favored by the observation that the majority of P-DMRs were intragenic and enriched for genes changing expression during early development. Another question is whether P-DMRs merely represent genomic scars of prenatal adversity or mark an adaptive response to cope with adversity. We found that DNA methylation differences at individual regions extended towards pathways involved and growth and lipid metabolism. This finding is compatible with optimization for energy production and growth under nutritional constraint through epigenetic fine-tuning[41].

As an alternative explanation, we should consider that the observed differences in DNA methylation may have no functional meaning but merely reflect specific characteristics of the selected individuals who survived the famine. Even under non-famine conditions, early pregnancy loss may be as high as 70%[48]. During the Hunger Winter, the number of births decreased by 50%[49], presumably in part because of increased levels of early pregnancy loss. Furthermore, undernutrition during pre-implantation can cause reduced cell numbers in pre- and post implantation rat embryos[50]. Embryos or subsets of cells with an increased growth potential may have had a higher chance of survival and this may explain why we observed a P-DMR at *INSR*, a key regulator of prenatal growth. Even among individuals with exposure in early gestation, we observed an association with birth weight, and methylation changes in multiple pathways related to growth. If our findings are based on selective survival, the P-DMRs would not reflect environmentally-induced changes but characteristics that follow selection on epigenetic variation between embryos.

LDL and CPT1A

Individuals with early gestational famine exposure have a less favorable metabolic profile[5], including a higher LDL cholesterol[6]. For LDL cholesterol, data were available for both the famine exposed individuals and their unexposed sibling controls. In these groups, we found a tentative association between *CPT1A* gene body methylation and LDL. The association between *CPT1A* methylation and LDL was almost identical in the prenatally exposed individuals and their unexposed same-sex siblings. If replicated, *CPT1A* methylation may thus prove to be a LDL quantitative trait locus contributing to the higher LDL levels in the exposed. *CPT1A* is involved in shuttling fatty acids to the mitochondria and as such is the rate limiting enzyme in fatty acid oxidation[39]. The association may with LDL may be related to the observation that free fatty acids drive hepatic LDL production[51], therefore it may be of interesting to relate *CPT1A* methylation with free fatty acids levels, for which we do not have data, rather than LDL.

Tissue specificity

The famine associations along pathways and with birth weight and LDL all relate to processes taking place in other tissues than blood. However, we have not investigated whether these associations extend to other relevant tissues than blood for such samples are simply not readily available in cohorts of generally healthy individuals[17]. We do not exclude the possibility that some of the 181 regions identified entail tissue specific associations, but we did not find enrichment for P-DMRs mapping to tissue specific genes. Since we find associations arising in early gestation the DNA methylation differences in blood may reflect those in other tissues due to mitotic inheritance[52], as was shown in animal studies[3,53].

Concluding

We identified regions with a regulatory potential that may be especially sensitive to early famine exposure. We observed DNA methylation changes in middle age among individuals exposed to environmental disturbances during early human development that extend along pathways. To avoid unmeasured sources of confounding at the family level, we used unexposed same sex siblings as study controls. Our study thereby makes a significant contribution to the further systematic exploration of the relation between prenatal adversity and adult health.

Materials and Methods

Subjects

The recruitment process of the Hunger Winter Families study has been described in detail elsewhere[18]. Study subjects were selected from births between 1943-1947 at three institutions in famine-exposed cities (the midwifery training schools in Amsterdam and Rotterdam and the University Medical Center in Leiden). The series includes singletons born between February 1945-March 1946 who had been exposed to famine in utero, births in 1943 without in-utero exposure serving as pre-famine time controls, and births in 1947 without such exposure serving as post-famine time controls. Whenever possible, we recruited an unexposed same-sex sibling of each individual in this birth series to serve as a family control. This provides sibling pairs, some of which include one famine exposed individual and an unexposed control. Ethical approval for the study was obtained from the participating institutions, and all participants provided written informed consent and additional explicit approval for the current set of genome scale measurements was given by the Leiden University Medical Center medical ethics committee.

The food supply and nutrition had been generally adequate during WWII in the Netherlands[54]. However, after October 1944, supplies became increasingly

scarce and the energy content of official rations fell below 1,000 kcal/day by the end of November 1944. By April 1945, just before liberation, the official rations were as low as 500 kcal/day[55] and the daily rations did not reach 2,000 kcal/day until May 20th. Since the famine period was shorter than the nine months of human gestation, and the population was well fed just before and directly after the famine, individuals can be identified that were exposed during specific periods of their development *in uterus*.

Within the Hunger Winter Families Study there are 313 same-sex sibling pairs who completed clinical examination. Sixty pairs included a sibling exposed to the famine around conception for up to 10 weeks into development (e.g. "periconceptional" exposure), as defined by a mothers' estimated last menstrual period between November 28, 1944 and May 15, 1945 (exposed: age 58.1y, SD 0.35; unexposed same-sex siblings: age 57.1y, SD 5.50). We included for analysis all 38 pairs of this group who had an age difference of less than 5 years at the time of examination. From these pairs we randomly selected 12 male and 12 female sibling pairs. Half of the female and half of the male pairs consisted of a sibling conceived and born after the famine as to prevent a consistent age difference between the prenatally exposed and controls and minimize a possible influence of early childhood famine exposure on the analyses (**Supplement II, Table S1**).

Birth weights were only available for the individuals in the hospital series and were taken from the birth records from the three institutions at which these individuals were born, meaning that we only had birth weight for the sixty individuals conceived during the famine and not their same-sex sibling controls. A telephone interview was performed with all participants and included questions on socio-demographic characteristics such as education, health history, health behaviors such as smoking and drinking, and medications for diabetes, cholesterol and blood pressure. The medical examinations were scheduled early in the morning and included the measurement of height (to the nearest 1mm by portable stadiometer [Seca, Hamburg, Germany]) and body weight (to the nearest 100g by portable scale [Seca, Hamburg, Germany]). The body mass index was calculated from these measurements. Participants had been asked to fast overnight before the clinic visit and were

offered conventional glucose tolerance testing with a 75-mg oral glucose challenge provided in the fasting condition. Glucose was assayed in promptly separated serum by the hexokinase reaction on a Modular P800 (Roche, Boehringer Mannheim, Germany); the interassay coefficient of variation was 1.3-1.8%. The lowest detectable amount was 3mU/L. Type 2 diabetes was defined by a diabetes history with insulin treatment or a fasting glucose value of ≥126 mg/dl (7 mmol/l) or a 2-h post-challenge glucose value ≥ 200mg/dl (11.1 mmol/l). Serum total cholesterol, HDL cholesterol and triglycerides were collected from the baseline fasting blood draw and immediately measured by standard methods[56-58] and LDL cholesterol was calculated for individuals with a triglyceride concentration lower than 400mg/dL using the Friedewald formula[59].

Library generation

Genomic DNA was isolated using the salting out method. We used Reduced Representation Bisulfite Sequencing (RRBS) of which the characteristics and a detailed protocol was published[60,61]. In short, half a microgram of genomic DNA was digested with MspI (NEB). The digestion was end-repaired using Klenow 3'-5' exo$^-$ (NEB) with a mix of 10mM methylated dCTP (Trilink), 10mM dGTP and 80mM dATP (Invitrogen). Sequencing adapters (ATDbio) with all cytosines replaced by methylcytosines were ligated to the end-repaired digestion with T4 ligase (NEB). Next, the adapter ligated digestion was run on a 2.5% 3:1 Nusieve gel (Lonza) and all fragments between 150 and 220 bp were cut out and cleaned (~40-120bp MspI fragment size). Two rounds of bisulfite conversion were performed using the Epitect 5h FFPE protocol (Qiagen). After the consecutive bisulfite treatments the 150-220 bp library was PCR amplified using Pfu Turbo Cx polymerase (Agilent Technologies) using Illumina GAIIx PCR primers (5 min at 95^{o}C, *n* x [30sec. at 95 ^{o}C, 20sec. at 65 ^{o}C, 30sec at 72 ^{o}C], 7 min at 72 ^{o}C, with n ranging from 11-19 cycles) and the resulting product was cleaned and then again put on a 2.5% 3:1 Nusieve gel (Lonza) for a final cleaning and the necessary removal of the PCR primers.

Sequencing and DNA methylation calling

Each of the 48 individual libraries was sequenced on a single Illumina GAIIx lane in the Broad Institute. An average of 25.6 million high quality 36bp single end reads were obtained for each individual for which on average 74.1% (SD 10.4%) could be uniquely aligned to the bisulfite converted human genome (hg19) using custom scripts described previously[61]. Cytosines outside the CpG dinucleotide context were used to assess the bisulfite conversion rate. The average bisulfite conversion rate was 98.9% (SD 0.68%). The number of unmethylated and methylated cytosines per measured CpG dinucleotide was determined from the .BAM alignment files using a custom python script[61]. Global mean methylation was assessed by calculating the average DNA methylation of all CpGs in all accepted reads. We also aligned all high quality reads to the prototypic repeat sequences in the RepBase Update database and again determined the amount of methylated and total number of reads per consensus sequence CpG using custom software[62]. For each repeat type the number of methylated and unmethylated reads were summed, repeats with a median coverage higher than 5 reads over all the measured sibling pairs were included in our analysis.

Data grouping to genomic annotations

Particularly when smaller effect sizes are anticipated and the study size is limited, the focus on annotated regions encompassing multiple CpG dinucleotides is superior to testing single CpG dinucleotides. For it enhances the biological relevance and interpretability of results and is statistically more efficient by reducing the number of tests and increasing the precision of DNA methylation estimates[63]. The location of CpG dinucleotides was first lifted to hg18 (NCBI36) since at the start of the study most annotations were available for hg18. All CpG dinucleotides were then mapped for an overlap with genomic features such as promoters and 'bonafide' CpG islands[23], which are CpG islands (CGIs) with a ubiquitously open chromatin conformation. Most of the genomic feature annotations were taken from *Gu et al.*[16] (available

from http://rrbs-techdev.computational-epigenetics.org). These annotations were supplemented with genome-wide annotations of H3K4me1 and H3K4me3 associated enhancers[24], CTCF binding sites[29] and some additions of particular interest for early development or prenatal environmental exposure. We included the annotations of bivalent chromatin domains from human embryonic stem cells (hESCs)[64] and human hematopoietic stem cells (HSCs)[65], highly variable regions[26], loci hypothesized to be sensitive to early nutrition (putative metastable epialleles)[25] and enhancer regions shown to be associated with genes involved in pre- and peri-implantation development (dev. Enhancers Type I) or early differentiation stages (dev. Enhancers type II)[44].

Individual CpG sites were mapped to a specific genomic locus contained within the annotations of genomic features (e.g. mapped to a particular promoter or other feature) when they had an overlap in terms of their genomic location. Each locus was subsequently mapped to the nearest entrez gene identifier within 100kb. Regions were denoted as intergenic if the distance the nearest gene was larger than 100kb.

GlobalTest for genomic annotation and pathway analysis

We used the R package GlobalTest[66] to test genomic annotation as a whole and to test groups of individual regions mapped to gene-sets. For this test DNA methylation values within a given genomic region were transformed to account for missing values and to account for the differences in the total coverage (thus accuracy of the measurement) and the number of CpG sites per locus. The transformation shrinks the fraction of methylated reads towards the average methylation fraction in the population of 48 subjects. This shrinkage is especially pronounced in subjects with few reads. The transformation was performed as follows:

$M_{i,j}$ Denotes the number of methylated reads for individual *i*, CpG dinucleotide *j* for a particular region.

$T_{i,j}$ Total number of reads for individual *i*, CpG dinucleotide *j* for a particular region.

k The number of CpG sites for a given region (=j_{max})

$M_{i\bullet} = \sum_{j=1}^{k} M_{i,j}$ Denotes the sum of all methylated reads for individual *i* for a particular region.

$T_{i\bullet} = \sum_{j=1}^{k} T_{i,j}$ Denotes the sum of all reads for individual *i* for a particular region.

$$region = \frac{M_{i\bullet} + 5 \times \left(\frac{\sum_{i=1}^{i=48} M_{i\bullet}}{\sum_{i=1}^{i=48} T_{i\bullet}} \right)}{T_{i\bullet} + 5} \times \sqrt{\frac{median(T_{1\bullet},..,T_{48\bullet})}{k}}$$

As part of the validation of this approach we tested the transformed values of the regions that were associated to prenatal famine exposure and a set of loci that were not or only nominally associated in our earlier candidate studies in this population[11,12] by GlobalTest. The positive set was again associated with prenatal famine exposure ($P=9x10^{-3}$) and the negative set was not (P=0.19).

Epityper data generation and pre-processing

Primers were designed using Methprimer[67]. The resulting primer and amplicons locations were checked against the latest version of dbSNP for SNPs. Their spectrum characteristics were checked with the R package RSeqMeth[68]. The sequences of the primers used in our study and the genomic locations they amplify are given in **Supplement II, Table S5**. The regions measured are necessarily smaller than the regions identified by RRBS, the

genomic overlap and the number of identical CpG dinucleotides measured is given in **Supplement II, Table S6.**

One microgram of genomic DNA isolated using the salting-out method from whole blood was bisulfite treatment EZ 96-DNA methylation kit (Zymo Research) with overnight bisulfite incubation according to the supplier's protocol. The 60 sibling pairs were randomly distributed over two 96 well plates with similar proportions of male and female pairs on each plate and in similar proportions for the pairs also measured with RRBS and those who were not. DNA methylation was quantitatively assessed for each locus using the mass spectrometry based Epityper assay (Sequenom, USA) in triplicate using the manufacturers' protocol on one 384 well plate. PCR was performed with the following cycling protocol: 15 minutes at 95^{o}C, 4 rounds of 20 seconds at 95^{o}C, 30 seconds at 65^{o}C, 1 minute at 72^{o}C; followed by 40 rounds, 20 seconds at 95^{o}C, 30 seconds at (see **Supplement II, Table S5**) ^{o}C and 1 minute at 72^{o}C; ending with 3 minutes at 72^{o}C. Processing of the Epityper data has been described in detail previously[11,12,69,70]. A bisulfite converted DNA is processed on the same 384 plate in triplicate. From the bisulfite conversion onward we perform all steps inherent to the Epityper protocol 3 times. Fragments containing CpG dinucleotides ("CpG units") that have a mass within the mass range that do not overlap other CpG units were considered. Fragments were also discarded if dbSNP indicated the possible presence of a SNP in individuals of European decent with a minor allele frequency higher than 5%. In short, measurements for CpG dinucleotides containing fragments for which 2 out of 3 measurements were successful, the standard deviation of these multiple measurements was smaller than 10% and for which the overall measurement success rate in the 60 pairs was higher than 75% were included in the final analyses. Averages for these triplicate measurements were used for the analyses. For each measurement non-bisulfite converted genomic DNA and negative controls were incorporated to check for a-specific amplification and PCR artifacts, none were found. Bisulfite conversion was assessed using direct bisulfite sequencing on an ABI sequencer and was ≥98% for both 96 well plates.

Statistical tests for individual regions

We tested for within-pair differences in DNA methylation between exposed individuals and their same-sex siblings for individual loci by applying generalized linear mixed models on the sequencing data and linear mixed models for the Epityper data. With these models the correlation between adjacent CpG sites in an individual can be taken into account and all available raw but incomplete data can be used for modeling and control for possible confounders. The R programming environment was used for all analyses[71].
For the bisulfite sequencing data we used logistic mixed models using the glmer() function from the lme4 package[72] with a binomial distribution, weighing by the sequencing depth per individual observation. The dependent variable was the DNA methylation fraction. This in effect is the same as modeling each individual read as either 0 or 1 (unmethylated or methylated) in the dependent without weighing for coverage depth, hence the application of a logistic model. Exposure status (exposed vs. unexposed) and a unique identifier for each CpG dinucleotide were entered as fixed effects. To specify a within-sib-pair design, the (family) pair identifier was included as a random effect with intercept. To model the correlation in DNA methylation within an individual we make use of the fact that each family consists of an exposed individual and a same-sex sibling, therefore adding the exposure status to the model as a random slope, possibly correlated to the random intercept. This model is equivalent to one in which the individual identifier was added as a random effect. This model option allows us to use the same model for both multiple CpG sites and single CpG sites, allowing a unified analysis pipeline.
For the continuous Epityper data we used a linear mixed model based on the lmer() function from the same lme4 package, applying the same model used for the RRBS data, but now without the necessity to weigh for coverage and using a normal distribution. The models were fitted by REML or, when model fits were compared, by ML. The outcome of this model is identical to a paired t-test when an individual CpG site is assessed, there is no missing data and no corrections are applied.

We added bisulfite batch as fixed effect, since the 60 pairs were distributed across 2 96 well plates for bisulfite treatment (keeping pairs on the same plate). Since the age difference is larger for some of the pairs in this set, age at blood drawn was also entered as fixed effect. We are able to effectively correct for age because the siblings without prenatal famine exposure are equally distributed between being born before and after the war period. Additional corrections were performed by adding these respective variables as fixed effects to the model. For current diet we had data on the amount of kcal/day consumed and the percentage of fat, carbohydrates and protein in the diet. Interactions were tested by adding an interaction term as a fixed effect, always including the main terms. Models incorporating DNA methylation data from multiple loci were extended by removing exposure status as intercept and instead adding a random effect for individual with a nested random effect denoting the region. Normality of the Epityper data was checked by histograms of the raw data and the lmer() model residuals. Model fits were diagnosed by plotting the residuals against the fitted values and comparing the variance of the residuals across the exposed and unexposed and other factorial covariates.

Multiple testing correction was performed according to the method developed by Bejamini and Hochberg, better known as 'FDR' (false discovery rate) correction using the R base 'p.adjust()' function. All p-values reported are two-sided. Reported confidence intervals are at 95%, without adjustment for multiple testing. The replication rate was calculated using the closed testing procedure based on the Simes inequality as described by Goeman and Solari[73]. The validation rate is an estimate of π_1, the number of correctly rejected null hypotheses, and since its confidence interval is the most informative only this was reported in the Results section. The π_1 itself was 6/6.

Acknowledgements

We express our gratitude to the participants of the Dutch Hunger Winter Families study and the staff of TNO Quality of Life for contact tracing. At the Leiden University Medical Center we wish to acknowledge the staff of the department of Gerontology and Geriatrics Study Center for the physical examinations and the Central Clinical Chemical Laboratory for extracting DNA.

This study was supported by a Horizon Breakthrough grant in the framework of the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research [93518027 to BTH], the European Union-funded Network of Excellence LifeSpan [FP6 036894 from which support was given to PES, BTH and EWT], the U.S.A.'s National Institute of Aging [R01AG042190 to LHL and BTH], The European Union's Seventh Framework Program IDEAL [259679 to PES], by the U.S. National Institutes of Health [R01-HL067914 to LHL] and the Netherlands Consortium for Healthy Ageing [#05060810 to PES]. The funders had no role in study design, data collection, analysis, decision to publish or preparation of the manuscript.

References

1. Gluckman PD, Hanson MA, Bateson P, Beedle AS, Law CM, Bhutta ZA, et al. Towards a new developmental synthesis: adaptive developmental plasticity and human disease. Lancet 2009; 373: 1654-7
2. Bogdarina I, Haase A, Langley-Evans S, Clark AJ. Glucocorticoid effects on the programming of AT1b angiotensin receptor gene methylation and expression in the rat. PLoS.One. 2010; 5: e9237
3. Morgan HD, Jin XL, Li A, Whitelaw E, O'Neill C. The culture of zygotes to the blastocyst stage changes the postnatal expression of an epigentically labile allele, agouti viable yellow, in mice. Biol.Reprod. 2008; 79: 618-23
4. Zhang S, Rattanatray L, MacLaughlin SM, Cropley JE, Suter CM, Molloy L, et al. Periconceptional undernutrition in normal and overweight ewes leads to increased adrenal growth and epigenetic changes in adrenal IGF2/H19 gene in offspring. FASEB J. 2010; 24: 2772-82
5. Lumey LH, Stein AD, Susser E. Prenatal famine and adult health. Annu.Rev.Public Health 2011; 32: 237-62
6. Lumey LH, Stein AD, Kahn HS, Romijn JA. Lipid profiles in middle-aged men and women after famine exposure during gestation: the Dutch Hunger Winter Families Study. Am.J.Clin.Nutr. 2009; 89: 1737-43
7. Lussana F, Painter RC, Ocke MC, Buller HR, Bossuyt PM, Roseboom TJ. Prenatal exposure to the Dutch famine is associated with a preference for fatty foods and a more atherogenic lipid profile. Am.J.Clin.Nutr. 2008; 88: 1648-52
8. Susser E, Neugebauer R, Hoek HW, Brown AS, Lin S, Labovitz D, et al. Schizophrenia after prenatal famine. Further evidence. Arch. Gen.Psychiatry 1996; 53: 25-31
9. Mamun AA, O'Callaghan MJ, Williams GM, Najman JM. Maternal smoking during pregnancy predicts adult offspring cardiovascular risk factors - evidence from a community-based large birth cohort study. PLoS.One. 2012; 7: e41106
10. Dabelea D, Pettitt DJ. Intrauterine diabetic environment confers risks for type 2 diabetes mellitus and obesity in the offspring, in addition to genetic susceptibility. J.Pediatr.Endocrinol. Metab 2001; 14: 1085-91
11. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc. Natl.Acad.Sci.U.S.A 2008; 105: 17046-9
12. Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD, et al. DNA Methylation differences after exposure to prenatal famine are common and timing- and sex-specific. Hum.Mol.Genet. 2009; 18: 4046-53
13. Soubry A, Murphy S, Huang Z, Murtha A, Schildkraut J, Jirtle R, et al. The effects of depression and use of antidepressive medicines during pregnancy on the methylation status of the IGF2 imprinted control regions in the offspring. Clin.Epigenetics. 2011; 3: 2
14. Bouchard L, Thibault S, Guay SP, Santure M, Monpetit A, St-Pierre J, et al. Leptin gene epigenetic adaptation to impaired glucose metabolism during pregnancy. Diabetes Care 2010; 33: 2436-41
15. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, et al. A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature 2012; 484: 339-44
16. Gu H, Bock C, Mikkelsen TS, Jager N, Smith ZD, Tomazou E, et al. Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. Nat.Methods 2010; 7: 133-6
17. Heijmans BT, Tobi EW, Lumey LH, Slagboom PE. The epigenome: archive of the prenatal environment. Epigenetics. 2009; 4: 526-31
18. Lumey LH, Stein AD, Kahn HS, van der Pal-de Bruin KM, Blauw GJ, Zybert PA, et al. Cohort profile: the Dutch Hunger Winter families study. Int.J.Epidemiol. 2007; 36: 1196-204
19. Li Y, Zhu J, Tian G, Li N, Li Q, Ye M, et al. The DNA methylome of human peripheral blood mononuclear cells. PLoS.Biol. 2010; 8: e1000533
20. Lumey LH, Terry MB, Delgado-Cruzata L, Liao Y, Wang Q, Susser E, et al. Adult global DNA

methylation in relation to pre-natal nutrition. Int.J.Epidemiol. 2012; 41: 116-23

21. Tobi EW, Slagboom PE, van DJ, Kremer D, Stein AD, Putter H, 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: e37933

22. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. Repbase Update, a database of eukaryotic repetitive elements. Cytogenet.Genome Res. 2005; 110: 462-7

23. Bock C, Walter J, Paulsen M, Lengauer T. CpG island mapping by epigenome prediction. PLoS.Comput.Biol. 2007; 3: e110

24. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 2009; 459: 108-12

25. Waterland RA, Kellermayer R, Laritsky E, Rayco-Solon P, Harris RA, Travisano M, et al. Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. PLoS.Genet. 2010; 6: e1001252

26. Feinberg AP, Irizarry RA, Fradin D, Aryee MJ, Murakami P, Aspelund T, et al. Personalized epigenomic signatures that are stable over time and covary with body mass index. Sci. Transl.Med. 2010; 2: 49ra67

27. Bock C, Halachev K, Buch J, Lengauer T. EpiGRAPH: user-friendly software for statistical analysis and prediction of (epi) genomic data. Genome Biol. 2009; 10: R14

28. Andersson R, Enroth S, Rada-Iglesias A, Wadelius C, Komorowski J. Nucleosomes are well positioned in exons and carry characteristic histone modifications. Genome Res. 2009; 19: 1732-41

29. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. Cell 2007; 129: 823-37

30. Pekowska A, Benoukraf T, Ferrier P, Spicuglia S. A unique H3K4me2 profile marks tissue-specific gene regulation. Genome Res. 2010; 20: 1493-502

31. Zhang P, Zucchelli M, Bruce S, Hambiliki F, Stavreus-Evers A, Levkov L, et al. Transcriptome profiling of human pre-implantation development. PLoS.One. 2009; 4: e7844

32. Yi H, Xue L, Guo MX, Ma J, Zeng Y, Wang W, et al. Gene expression atlas for human embryogenesis. FASEB J. 2010; 24: 3341-50

33. Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. Proc.Natl.Acad.Sci.U.S.A 2005; 102: 15785-90

34. Schaniel C, Ang YS, Ratnakumar K, Cormier C, James T, Bernstein E, et al. Smarcc1/Baf155 couples self-renewal gene repression with changes in chromatin structure in mouse embryonic stem cells. Stem Cells 2009; 27: 2979-91

35. Kleger A, Mahaddalkar PU, Katz SF, Lechel A, Joo JY, Loya K, et al. Increased reprogramming capacity of mouse liver progenitor cells, compared with differentiated liver cells, requires the BAF complex. Gastroenterology 2012; 142: 907-17

36. Ozair MZ, Noggle S, Warmflash A, Krzyspiak JE, Brivanlou AH. SMAD7 Directly Converts Human Embryonic Stem Cells to Telencephalic Fate by a Default Mechanism. Stem Cells 2012; 1: 35-47

37. Heard ME, Pabona JM, Clayberger C, Krensky AM, Simmen FA, Simmen RC. The reproductive phenotype of mice null for transcription factor kruppel-like factor 13 suggests compensatory function of family member kruppel-like factor 9 in the peri-implantation uterus. Biol.Reprod. 2012; 87: 115

38. Smart NG, Apelqvist AA, Gu X, Harmon EB, Topper JN, MacDonald RJ, et al. Conditional expression of Smad7 in pancreatic beta cells disrupts TGF-beta signaling and induces reversible diabetes mellitus. PLoS.Biol. 2006; 4: e39

39. Zammit VA, Ramsay RR, Bonomini M, Arduini A. Carnitine, mitochondrial function and therapy. Adv.Drug Deliv.Rev. 2009; 61: 1353-62

40. Natesampillai S, Fernandez-Zapico ME, Urrutia R, Veldhuis JD. A novel functional interaction between the Sp1-like protein

KLF13 and SREBP-Sp1 activation complex underlies regulation of low density lipoprotein receptor promoter function. J.Biol.Chem. 2006; 281: 3040-7

41. Stoger R. The thrifty epigenotype: an acquired and heritable predisposition for obesity and diabetes? Bioessays 2008; 30: 156-66
42. Suter M, Ma J, Harris AS, Patterson L, Brown KA, Shope C, et al. Maternal tobacco use modestly alters correlated epigenome-wide placental DNA methylation and gene expression. Epigenetics. 2011; 6: 1284-94
43. Supek F, Bosnjak M, Skunca N, Smuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS.One. 2011; 6: e21800
44. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. Nature 2011; 470: 279-83
45. Waterland RA, Jirtle RL. Early nutrition, epigenetic changes at transposons and imprinted genes, and enhanced susceptibility to adult chronic diseases. Nutrition 2004; 20: 63-8
46. Chotalia M, Smallwood SA, Ruf N, Dawson C, Lucifero D, Frontera M, et al. Transcription is required for establishment of germline methylation marks at imprinted genes. Genes Dev. 2009; 23: 105-17
47. Waterland RA, Kellermayer R, Rached MT, Tatevian N, Gomes MV, Zhang J, et al. Epigenomic profiling indicates a role for DNA methylation in early postnatal liver development. Hum.Mol.Genet. 2009; 18: 3026-38
48. Macklon NS, Geraedts JP, Fauser BC. Conception to ongoing pregnancy: the 'black box' of early pregnancy loss. Hum.Reprod. Update. 2002; 8: 333-43
49. Stein ZA, Susser M, Saenger G, Marolla F. Famine and Human Development: The Dutch Hunger Winter of 1944-1945. New York: Oxford University Press, 1975
50. Kwong WY, Wild AE, Roberts P, Willis AC, Fleming TP. Maternal undernutrition during the preimplantation period of rat development causes blastocyst abnormalities and programming of postnatal hypertension. Development 2000; 127: 4195-202
51. Julius U. Influence of plasma free fatty acids on lipoprotein synthesis and diabetic dyslipidemia. Exp.Clin.Endocrinol.Diabetes 2003; 111: 246-50
52. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nat.Rev.Genet. 2011; 12: 529-41
53. McKay JA, Xie L, Harris S, Wong YK, Ford D, Mathers JC. Blood as a surrogate marker for tissue-specific DNA methylation and changes due to folate depletion in post-partum female mice. Mol.Nutr.Food Res. 2011; 55: 1026-35
54. Trienekens G. The food supply in the Netherlands during the Second World War, In: Smith DF, Phillips, eds. Food, science, policy and regulation in the twentieth century. International and comparative perspectives. London: Routledge, 2000: 117-33
55. Burger GCE, Drummond JC, Sandstead HR. Malnutrition and Starvation in Western Netherlands, September 1944-July 1945. The Hague: General State Printing Office, 1948
56. Rifai N, Cole TG, Iannotti E, Law T, Macke M, Miller R, et al. Assessment of interlaboratory performance in external proficiency testing programs with a direct HDL-cholesterol assay. Clin.Chem. 1998; 44: 1452-8
57. Stinshoff K, Weisshaar D, Staehler F, Hesse D, Gruber W, Steier E. Relation between concentrations of free glycerol and triglycerides in human sera. Clin.Chem. 1977; 23: 1029-32
58. Sugiuchi H, Uji Y, Okabe H, Irie T, Uekama K, Kayahara N, et al. Direct measurement of high-density lipoprotein cholesterol in serum with polyethylene glycol-modified enzymes and sulfated alpha-cyclodextrin. Clin.Chem. 1995; 41: 717-23
59. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin.Chem. 1972; 18: 499-502
60. Gu H, Smith ZD, Bock C, Boyle P, Gnirke A, Meissner A. Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. Nat.Protoc. 2011; 6: 468-81
61. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, et al. Genome-scale

DNA methylation maps of pluripotent and differentiated cells. Nature 2008; 454: 766-70

62. Bock C, Tomazou EM, Brinkman AB, Muller F, Simmer F, Gu H, et al. Quantitative comparison of genome-wide DNA methylation mapping technologies. Nat.Biotechnol. 2010; 28: 1106-14

63. Bock C. Analysing and interpreting DNA methylation data. Nat.Rev.Genet. 2012; 13: 705-19

64. Zhao XD, Han X, Chew JL, Liu J, Chiu KP, Choo A, et al. Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell Stem Cell 2007; 1: 286-98

65. Cui K, Zang C, Roh TY, Schones DE, Childs RW, Peng W, et al. Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. Cell Stem Cell 2009; 4: 80-93

66. Goeman JJ, van de Geer SA, de KF, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. Bioinformatics. 2004; 20: 93-9

67. Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002; 18: 1427-31

68. Coolen MW, Statham AL, Gardiner-Garden M, Clark SJ. Genomic profiling of CpG methylation and allelic specificity using quantitative high-throughput mass spectrometry: critical evaluation and improvements. Nucleic Acids Res. 2007; 35: e119

69. Talens RP, Boomsma DI, Tobi EW, Kremer D, Jukema JW, Willemsen G, et al. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. FASEB J. 2010; 24: 3135-44

70. Tobi EW, Heijmans BT, Kremer D, Putter H, Delemarre-van de Waal HA, Finken MJ, et al. DNA methylation of IGF2, GNASAS, INSIGF and LEP and being born small for gestational age. Epigenetics. 2011; 6: 171-6

71. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, 2012

72. Bates D, Maechler M, Bolker B. lme4: Linear mixed-effects models using S4 classes, 2012

73. Goeman JJ, Solari A. Multiple testing for exploratory research. Statistical Science 2011; 26: 584-97

Summary of Results

Introduction

Early environmental conditions and growth parameters at birth are associated with adult human disease risk in epidemiological studies. Generally, a low birth weight is taken in such studies as a proxy for an adverse environment during development, a view that is hotly debated and has lead to much discussion on the nature of these associations. The study of historic famines and in particular the Dutch Famine has shown that exposure during specific developmental time-frames is associated with a higher risk on metabolic disease and a poorer mental health later in life. Most of these associations are independent of birth weight. Indeed, also other defined prenatal exposures, like smoking and maternal obesity, are associated with variation in traits which have been related to famine exposure during early development.
It is hypothesized that persistent epigenetic differences induced by environmental challenges partly underlie such epidemiological observations. Animal studies have shown the basic principle; prenatal nutrition was found to modulate histone and DNA methylation marks at genes functionally implicated with disease. In this thesis, we addressed the contribution of epigenetic mechanisms to the link between environmental challenges in early life and adult health in humans by studying DNA methylation in individuals prenatally exposed to the Dutch Famine and an unexposed same-sex sibling as control. We also address if individuals with in uterine growth restriction show similar DNA methylation patterns as prenatal famine exposure.

Chapter 2: variation, patterns and stability

First we described variation of DNA methylation in whole blood, which is arguably the most widely available source for epigenetic research in well-characterized biobanks that are currently available. We studied a selection of 16 candidate loci that are epigenetically regulated and have roles in metabolism and development. We discovered that DNA methylation i) is variable in the population ii) is relatively stable over time and iii) is frequently not confounded by the cellular heterogeneity of whole blood samples.

Between CpG dinucleotides a high correlation was found, making it likely that a limited number of CpG dinucleotides can be measured to gauge the methylation status of a region (cf. genetic linkage disequilibrium). We also found a high covariance for a subset of loci between blood and buccal cells, which stem from a different embryonic lineage, indicating that there are genomic regions for which DNA methylation in blood may be used as a proxy for DNA methylation in other tissues.

Chapter 3: A persistent DNA methylation difference

No epigenetic differences associated with (prenatal) environmental conditions were reported on in humans; therefore we wished to set a proof-of-principle. The Dutch Famine, a severe 6 month famine at the end of WWII, provides a defined and severe prenatal environmental exposure. Therefore we measured DNA methylation of the *insulin like growth factor* 2 (*IGF2*), a crucial regulator of fetal growth, in whole blood of 60 individuals exposed early and 62 individuals exposed late in gestation to the Dutch Famine. For each individual we also measured a prenatally unexposed same-sex sibling as control. We hypothesized that early gestation would be the most sensitive period. Work in animals show that this period is the most dynamic period for establishing epigenetic marks and DNA methylation differences induced early in development may be passed on soma-wide.
Early gestational famine exposure was associated with a decrease in *IGF2* DNA methylation six decades post exposure. This was the first evidence that the epigenome may be persistently altered during early development by the environment, providing a candidate mechanism linking development and later disease in humans. Late gestational famine exposure was not associated with *IGF2* DNA methylation, which was in line with our hypothesis.

Chapter 4: Epigenetic differences associated with Famine are common and time- and sex-specific

Next we explored if the *IGF2* DMR association was unique, or that DNA methylation is frequently altered by famine exposure and includes associations with non-imprinted regions. Moreover, associations should mirror the epidemiological literature and should also include sex-specific and timing independent associations if epigenetic change is to be the mechanism underlying the link between development and disease. We therefore extended our investigation in exposure discordant sibling pairs to the additional 15 loci characterized in chapter 2.

DNA methylation was associated with prenatal famine exposure at additional imprinted (*INSIGF*, *GNASAS* and *MEG3*) and non-imprinted regions (*LEP*, *ABCA1* and *IL10*). Overall, the DNA methylation differences were modest on a molecular scale (<4%), but sizeable relatively to the variation in the population and included increases in DNA methylation. Most associations were restricted to early gestational famine exposure and sex-specificity of the associations was common. One association, with DNA methylation at the *leptin* promoter, was independent of the gestational timing of the famine exposure. These results give strength to our *IGF2* DMR study and position epigenetics as a candidate mechanism to explain the association between early development and later disease in humans.

Chapter 5: Generalizability, no association with prenatal growth restriction

A low birth weight, or in uterine growth restriction (IUGR) has been linked to an increased risk of adult metabolic and cardiovascular disease. IUGR is often taken as a proxy for prenatal malnutrition or seen as a sign of a suboptimal prenatal environment. Therefore we tested the relevance of famine-associated DNA methylation differences in contemporary cases of IUGR. We resorted to the developmental extreme of preterm birth (<32 weeks), because most DNA methylation differences were found after early

developmental famine exposure. We measured DNA methylation at *IGF2* DMR, *GNASAS*, *LEP* and *IL10* in preterm IUGR and non-growth restricted individuals. No differences were found. Our results add to a growing literature showing that IUGR may not necessarily find its basis in prenatal malnutrition. IUGR may arise as a result of many different causes and may therefore represent a heterogenic etiology.

Chapter 6: nature & nurture both influence DNA methylation

DNA methylation is also influenced by genetic variation. Therefore we wised to investigate the influence of genetic variation (Nature) on DNA methylation and contrast its influence with that of the environment (Nurture). First, we extended our *IGF2* DMR finding to five regulatory regions across the imprinted *IGF2*/*H19* locus. DNA methylation at most regulatory sites was associated with prenatal famine exposure and their methylation was correlated, indicating that epigenetic fine-tuning may involve larger regions. Next, we measured the genetic variation around *IGF2* and *H19*. DNA methylation at some regulatory sites was also associated with single nucleotides polymorphisms and these associations were similar in effect size to the famine associations. The associations of DNA methylation with prenatal famine and genetic variation were independent and additive. We were the first to show that Nature's and Nurture's influence on DNA methylation may co-exist. Findings from epigenetic association studies may thus have an environmental and genetic component and we should therefore not rush to exclude an influence of environmental factors on DNA methylation levels if a SNP is found to influence DNA methylation at a particular locus.

Chapter 7: epigenome-wide characterization

We aimed to learn at which genomic annotations DNA methylation differences associated with early famine exposure occur the most and if associations do not only extend across larger regions, but also across pathways. DNA

methylation of 1.2 million CpG dinucleotides was assessed by next generation sequencing in 24 individuals exposed to the Dutch Famine in early gestation and 24 unexposed same-sex sibling controls of which one sibling was conceived during the famine. Famine associated DNA methylation changes clustered at regions with a regulatory potential. Increases in DNA methylation were commonly observed and most differentially methylated regions were in gene-bodies. Even within the time-window of the first trimester we found evidence of an effect of the timing of the exposure on DNA methylation. DNA methylation at loci mapping to *CDH23*, *SMAD7*, *INSR*, *CPT1A*, *RFTN1* and *KLF13* were associated to famine in individuals conceived before April 1945, two months before the famine's end. The differences were smaller than 5%, but the associations extended to the pathways these genes belong to, which were pathways related to growth and lipid metabolism. Moreover, we found tentative associations between methylation at *INSR* with birth weight and methylation at *CPT1A* with LDL cholesterol levels. The latter association was almost identical in the prenatally exposed and their unexposed same-sex siblings. If replicated, *CPT1A* methylation may thus prove a LDL quantitative trait locus contributing to the higher LDL levels in the exposed. Modest DNA methylation differences thus extend across biologically relevant pathways and may be linked to phenotypes of interest in relation to early famine exposure.

General Discussion

Main Aims

It is hypothesized that persistent epigenetic changes induced by (early) environmental conditions may partly underlie the association between development and disease[1]. Animal studies showed that prenatal nutrition may induce a persistent change in DNA methylation which in turn affects the expression level of genes known to be implicated in the disease mechanism[2,3]. The main aim of this thesis was to investigate if the prenatal environment is also associated with DNA methylation changes in humans at genes and pathways implicated in disease.

To this end we first characterized the normal variation, patterns and stability of DNA methylation at sixteen loci in blood and buccal swaps (**Chapter 2**), as to learn if DNA methylation patterns in whole blood are suitable for study in human cohorts. To test for an influence of the environment on DNA methylation we first performed a proof-of-principle study by measuring the imprinted *insulin like growth factor* 2 differentially methylated region (*IGF2* DMR) in the Dutch Hunger Winter Families Study[4]. We measured DNA methylation in individuals exposed early or late in gestation to famine and a same-sex sibling as control (**Chapter 3**). We extended this measurement to all sixteen characterized loci to learn which period of development is the most sensitive and to investigate if famine exposure may be associated in a sex-specific and timing dependent fashion with DNA methylation, similar to the phenotypic associations (**Chapter 4**). Intra-uterine growth restriction (IUGR) is hypothesized to be a sign of prenatal malnutrition. Therefore we measured DNA methylation at four loci associated with famine in young adults born following IUGR from the Dutch national cohort of growth restricted and preterm born children (the POPS study)[5] to test if our findings in the Dutch Famine relate to this pregnancy outcome (**Chapter 5**).

Besides relating our findings to more contemporary complications during pregnancy, we also wished to contrast the influence of early environment (nurture) and genetic variation (nature) on DNA methylation. We therefore extended our analyses to regulatory regions within and flanking the *IGF2* locus in The Dutch Hunger Winter Families Study (**Chapter 6**). Finally,

we characterized the regions at which DNA methylation is sensitive to prenatal famine by extending our measurements to a genome-scale, which subsequently also allowed us to comprehensively test DNA methylation along whole pathways of related genes and investigate the relationship between DNA methylation and phenotypes affected by famine exposure (**Chapter 7**).

The first steps: candidate loci

We started with characterizing the variation, patterns and stability of DNA methylation in whole blood (**Chapter 2**). We selected 16 loci involved in metabolism, development and growth with some loci situated around retrotransposons or within imprinted regions. These latter features were put forward as especially liable following experiments on prenatal nutrition in animals[6]. An additional important selection criterion was that CpG dinucleotides either overlapped important DNA methylation sensitive transcription factor binding sites or were shown in literature to correlate with gene expression *in vitro* or *in vivo*. DNA methylation was a quantitative trait and we observed correlation both within and between loci, highly analogues to linkage disequilibrium for genetic variation. Interestingly, DNA methylation in blood was correlated with that in buccal for some loci and more importantly, was stable over time and for most loci not dependent on blood cell heterogeneity. DNA methylation is suitable for study in human epidemiological studies.

IGF2 DMR methylation is associated with early gestational famine exposure

IGF2 regulates the amount of nutrients transferred over the placenta to the developing embryo[7], making it an attractive candidate for epigenetic modifications in relation to nutritional disturbances. Moreover, DNA methylation at the *IGF2* DMR in blood was found to mark DNA methylation patterns in buccal cells (**Chapter 2**) while others found that it marked that in colon, where it is also linked to *IGF2* expression[8,9]. We hoped that this covariance stemmed from the phenomenon that differences induced early in

development may be transmitted to other tissues[10]. Since early development is epigenetically the most dynamic period[11], we hypothesized that famine exposure during early development is associated with DNA methylation at *IGF2* DMR.

We were the first to report for humans a DNA methylation difference associated with a prenatal environmental exposure (**Chapter 3**). A decrease in methylation at *IGF2* DMR was found in individuals exposed from conception up to 10 weeks into development as compared to their prenatally unexposed same-sex siblings (**Figure 1A**). To further test our hypothesis we also tested sibling pairs of which one sibling was exposed during the last trimester of pregnancy. In correspondence with our hypothesis that early development is the most sensitive period we found no effect on DNA methylation as a result of famine exposure during the last trimester (**Figure 1B**)**.** Interestingly, not all sibling pairs showed differential DNA methylation following early gestational famine exposure of one of the siblings (**Figure 1A**). This is analogues to results from experiments in the *agouti* and *axin fused* mice models where not all mice exposed to prenatal folic acid supplementation showed a change in DNA methylation associated with prenatal environmental exposure (**Figure 1C and 1D**). These findings in mice, and ours in humans, indicate either a stochastic or individual-specific component to the epigenetic response to an environmental challenge.

Since the publication of this finding others have linked *IGF2* methylation to birth weight[12,13] and brain morphology[14], which is of interest since early gestational exposure to famine does not lead to a lower birth weight (**Chapter 3**) while it is associated with brain morphology[15], schizophrenia[15] and a poorer mental health[16].

Most associations are limited to early gestation

Next, we wished to see if famine generally affects DNA methylation of genes including those in non-imprinted genomic regions. Moreover, if the associations between famine exposure and DNA methylation are to be relevant they should mirror the phenotypic associations with sex-specific and

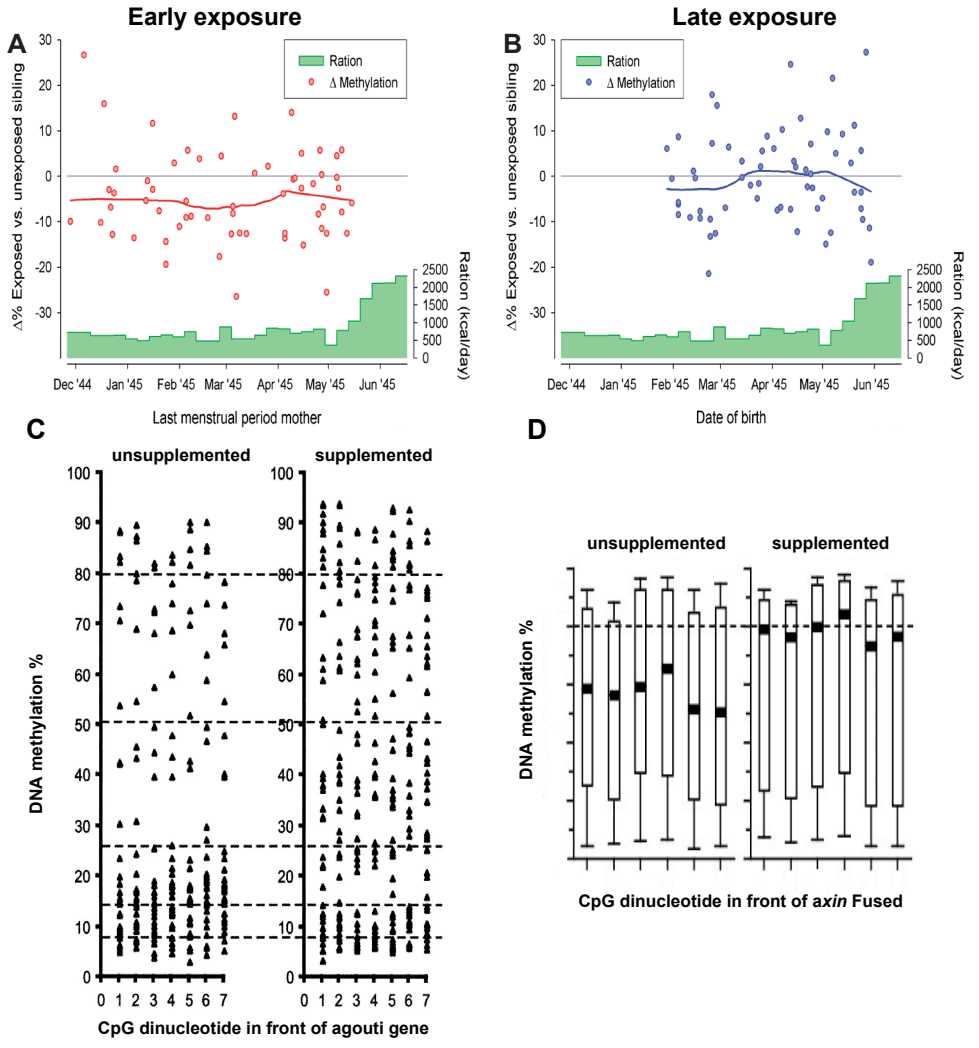


Figure 1. The stochastic nature of famine associated DNA methylation differences
In panels (A) and (B) the within pair differences is plotted along the estimate of conception of the prenatally or birth of the exposed sibling of the pair. **A.** Early exposure leads to a shift in the average within pair difference below the zero line. Lots of variation remains and several pairs do not show a decrease in DNA methylation in the exposed pairs. **B.** No overall difference in within pair methylation is seen after late gestational famine exposure. **C.** The pattern observed after early famine exposure (A) is similar to the influence of folic acid supplementation on the *agouti* locus. **D.** Again also for the axin fused locus a similar pattern is found to *IGF2* (A) and *agouti* (C). Not all mice in a litter show an increase in DNA methylation after folic acid supplementation during pregnancy at the agouti and axin fused loci, lots of variation remains, just as seen in the famine pairs.

timing independent associations. We extended our *IGF2* DMR study to the additional 15 candidate loci characterized in **Chapter 2**. In total, seven out of the sixteen candidate loci showed an association with prenatal famine exposure (**Chapter 4**). Most associations were found with early gestational famine exposure, but we identified several sex specific associations and one timing independent association between famine exposure and DNA methylation, also a first in human studies. Moreover, imprinted and non-imprinted regions were affected. All seven genes (*IGF2*, *GNASAS*, *INSIGF*, *LEP*, *IL10* and *ABCA1*) are involved in development and implicated in disease. Thus, our findings mirror animal studies[17,18] and the patterns observed in the epidemiological literature for famine exposure[19]. Our observations give credence to the hypothesis that the associations between development and disease may be explained by epigenetic mechanisms[1].

IUGR and prenatal malnutrition

We asked ourselves the question if our findings may be extrapolated to other more contemporary exposures and pregnancy complications (**Chapter 5**). Others have successfully found associations of DNA methylation with folic acid supplementation at the *IGF2* DMR[13] and with gestational diabetes at the *LEP*[20] gene. In the epidemiological literature IUGR and a low birth weight are regarded as a proxy for a poor prenatal nutrition[21] and associated to adult health. However, early famine exposure does not lead to a lower birth weight[22] while the epidemiological findings in famine cohorts often entail associations that are either independent on the gestational timing or restricted to early gestational famine exposure[23]. Therefore we tested DNA methylation at loci associated with prenatal famine exposure in the context of IUGR, to examine whether the etiology of IUGR and prenatal famine exposure may have a similar component.

DNA methylation at famine loci is not associated with IUGR

Since our results indicated that early gestation is the most sensitive period we resorted to the developmental extreme of early growth restriction. From the POPS[5] cohort we compared DNA methylation at 19 years of age of growth restricted and non-growth restricted individuals that were born before week 32 of gestation. We did not find differences between the two groups for DNA methylation at *IGF2*, *GNASAS*, *LEP* and *INSIGF* at which DNA methylation was associated with prenatal famine exposure during early or both early and late gestation.

Although we took special care to avoid individuals treated with glucocorticoids, which changes the transport of folic acid and choline over the placenta[24] and may influence DNA methylation[26], pregnancies were affected by intoxications with drugs or alcohol, which may have confounded our analyses[25]. Furthermore, the study was limited in the number of subjects studied and since the individuals were unrelated we had less power than in the famine study to detect DNA methylation differences. However, the contrast in birth weight and body size was considerable and others have also studied the *IGF2* locus in the context of prenatal growth restriction and found no differences[26]. An explanation for our negative findings may also lie in the heterogeneity of causes underlying IUGR in the POPS cohort, which included intoxications, hypertension, placental insufficiencies and (acute) infections.

Different etiologies of IUGR?

Recently it was shown in rats that IUGR induced by placental insufficiency or malnutrition is associated with different effects on *igf2* and *lep* expression[29]. Also in the phenotypic consequences nuance differences can be found. A large meta-analysis of rat studies on hypertension using either caloric or protein restriction to induce IUGR found that protein restriction affected diastolic blood pressure, while caloric restriction affected both diastolic and systolic blood pressure[27]. Such studies indicate that IUGR caused by different exposures may result in a different etiology of the IUGR.

Also in humans, the epidemiological observations relating birth weight and IUGR with adult disease[21,28] may be the result of a mixture of developmental complications, as they were done in developed economies where malnutrition is unlikely to be a major underlying cause of IUGR. Indeed, twin studies have shown for developed economies that associations between birth weight and disease are driven by a considerable genetic component[29-32]. Some have argued that twin pregnancies are not suitable to study in this context[33], since both siblings may be considered growth restricted and have a higher pre- and postnatal mortality rate. However, recent studies have shown that adult all-cause mortality and cardiovascular and diabetes specific mortality is not different in twins, nor is the incidence of disease[37,38].

Future successful studies may need to focus on IUGR as a result of one clear factor. Indeed, focused studies in animals[39] and humans[40] on IUGR as a result of placental insufficiency have successfully discovered epigenetic changes. The epigenetic consequences of the diverse exposures and complications leading to IUGR may then be compared to discover common mechanisms leading to disease. Studying the unique consequences on both the epigenetic and phenotypic level will be likewise insightful.

Characterizing loci sensitive to prenatal famine exposure

Therefore we set out to characterize the effect of early famine exposure on DNA methylation in more detail. We focused our measurements on the 60 sibling pairs of which one sibling was exposed to famine during early gestation. We expanded our measurements to most of the regulatory regions in and around *IGF2* (**Chapter 6**) and to a genome-scale (**Chapter 7**).

Nature and Nurture may both influence DNA methylation

It is known that DNA sequence variation may influence DNA methylation levels[41] and we described before the onset of this thesis that DNA methylation at *IGF2* DMR is heritable[34], while *IGF2* DMR methylation is also sensitive to

the early prenatal environment (**Chapter 3**). Therefore we measured both DNA methylation and genetic variation to contrast the influence of genetic variation (Nature) with that of prenatal famine exposure (Nurture) (**Chapter 6**). DNA methylation at most regulatory regions in *IGF2*/*H19* was associated both with early famine exposure and single nucleotide polymorphisms (SNPs). We showed that the associations were independent of each other although they may occur at the same CpG dinucleotides. The effect sizes of the associations with prenatal famine exposure and genetic variation were highly similar. Findings from epigenetic association studies may therefore relate to genetic or environmental influences and to both.

Characterizing regions sensitive to early famine exposure

In **Chapter 4** we found associations between prenatal famine exposure and DNA methylation at imprinted and non-imprinted regions and we wondered if there are certain genomic regions more sensitive to prenatal famine exposure. To this end we expanded our measurements to 1.2M CpG sites in 24 of the 60 sibling pairs of which one sibling was exposed during early gestation (**Chapter 7**). Genomic regions with a regulatory potential were especially sensitive. Promoters devoid of CpG islands, exons, enhancers (especially those active around implantation) and regions with an open chromatin status, as identified by DNaseI and FAIRE-seq[35], were associated with prenatal famine exposure. Regions with an open chromatin structure are often regulatory and the associations are therefore more likely to be functionally relevant[36].

Imprinted regions, retrotransposon insertion sites and regions of extreme epigenetic variability have been suggested to be especially sensitive to environmental conditions during development[6,45]. We found no evidence that associations with early famine exposure are restricted to these genomic features (**Chapter 4 and 7**). Recently, systematic investigations of imprinted regions following prenatal malnutrition in mice also found no evidence for imprinted regions as particularly sensitive[37,38].

Modest effects of exposure

The more thorough investigations in-depth (**Chapter 6**) and in scope (**Chapter 7**) concurred with the candidate gene findings that within pair DNA methylation differences are small on a molecular scale. Although some larger DNA methylation differences were identified in our genome-wide dataset (**Chapter 7**), the majority was smaller than 5%. Others did also find some larger DNA methylation differences (>20%) among 27K CpG dinucleotides as a result of prenatal micronutrient supplementation and smoking[39-41], but these studies did not validate these findings by an independent method and performed their analyses on sub divisions of already small datasets. All-in-all the evidence points towards small effects of prenatal environmental exposures on DNA methylation and it is sensible to discuss how such small DNA methylation differences may be biologically relevant.

Modest effects: epigenetic fine-tuning?

It was suggested that small epigenetic changes across genes and pathways (epigenetic 'fine-tuning') may accumulatively act to shift gene-expression[42]. The methylation effects associated with early prenatal famine exposure across the regulatory regions of *IGF2*/*H19* were correlated (**Chapter 6**) and in our genome-wide measurement we found associations that extended towards entire pathways (**Chapter 7**). Recently a study across 27K CpG dinucleotides on prenatal tobacco smoke exposure made similar observations[43]. DNA methylation differences associated with prenatal tobacco exposure were modest (<10%) and were enriched across pathways that were found to be differentially expressed. This study, together with **Chapters 6 and 7**, gives strong empirical support for the hypothesis that modest epigenetic effects of an exposure may act accumulatively in genomic regions and pathways.

Modest effects: phenotypic associations

Such accumulative effects of exposure may explain why a limited variation and <10% differences in DNA methylation at regions studied in Chapters 3 to 6 have been related to type 1 and 2 diabetes (*INSIGF*)[44,45], HDL cholesterol, hypercholesterolemia and prior cardiovascular disease (*ABCA1*)[46], birth weight (*IGF2*)[12,13], serum levels (*IGF2*, *LEP*)[12,20] and the risk of myocardial infarction (*GNASAS*, *INSIGF*)[55] in other populations.

Indeed, in **Chapter 7** we found associations for DNA methylation at *CPT1A* and *INSR* with early gestational famine exposure that extended to multiple lipid and cholesterol (for *CPT1A*) and growth related (for *INSR*) pathways. *CPT1A* methylation was also tentatively associated with *LDL* cholesterol and *INSR* methylation with birth weight. These associations may be argued to stem from the multiple effects in DNA methylation along the relevant pathways, a hypothesis we were underpowered to test in the genome-wide data of **Chapter 7** because we only had genome-wide data for 48 individuals and the within pair differences were only present in the 18 pairs with a sibling conceived before April '45 during the Famine, reducing the studies' power even further.

For *CPT1A* the association with LDL was almost identical in the 60 individuals exposed during early gestation and the 60 unexposed same-sex siblings. Therefore the *CPT1A* P-DMR may be a quantitative trait locus (QTL) for LDL and the higher DNA methylation levels following early gestational famine exposure at *CPT1A* may therefore have contributed to the higher LDL cholesterol levels in the exposed[47,48]. For the P-DMR at *INSR* we only had birth weight data for the exposed. It is promising that the direction of the association between *INSR* methylation and birth weight within the exposed is compatible with the interpretation that the increased methylation at *INSR* following early famine exposure may have contributed to the possibly higher birth weight in these individuals (**Chapter 3**). However, this latter inference is highly speculative.

The two associations were small on a molecular scale (*CPT1A*-LDL: 2.4%/1 mmol*l^{-1}; *INSR*-birth weight: 0.4%/100 gram), but sizeable when expressed

relative to the normal variation in the population (0.3 SD/1 SD; 0.5 SD/1 SD) and explained up to 7.7% and 11.1% of the phenotypic variation. Such effect sizes are comparable to those reported for DNA methylation in cord blood at one CpG site in *RXRA* associating with prenatal maternal carbohydrate intake and childhood fat mass[49] and *IGF2* DMR methylation with prenatal folic acid supplementation and birth weight[13].

Modest effects make epigenetic association studies challenging

The modest size of the DNA methylation differences and coefficients of the associations reported by us and others show a real future challenge. The techniques to measure DNA methylation are based on bisulfite conversion, which degrades DNA and can easily fail to deliver a complete conversion, which subsequently influences the quantification of DNA methylation levels. Getting the technical aspects wrong for just a subset of measurement can easily lead to false positive and false negative findings. Detailed descriptions of the technical aspects of the study design and execution should be included in publications as to gauge the possible robustness of the reported findings.

Methodological issues: The Dutch Hunger Winter Families Study

The Dutch Hunger Winter Families study[4] includes same-sex sibling pairs of which one sibling was exposed during either early or late gestation and presents a powerful epidemiological design and a clear extreme environmental exposure to investigate the influence of early and late gestational malnutrition on DNA methylation. This design matches for early familial conditions, sex and partially for genetic variation. However, when interpreting findings several aspects should be considered.

Post-natal exposure?

Our current analyses included prenatally unexposed same-sex siblings as controls. Roughly half of them were born before the famine, and thus most likely exposed as a child. Recently it was shown that self-reported childhood exposure to the Dutch Famine was associated with an increased risk for obesity and diabetes[50,51]. Chinese famine studies have also reported that famine exposure during early childhood is associated with the metabolic syndrome[61,62] and BMI[52], although no effect on postnatal exposure was found on T2D risk[53]. Results are sometimes not replicated between Chinese famine studies[65,66]. Nonetheless, it is prudent in light of these studies to take post-natal exposure into consideration in analyses.

We have stratified our analyses for *IGF2* and the other candidate genes, finding no effect of post-natal exposure on DNA methylation or on the associations reported (*data not shown*). These analyses are, however, limited due to the small number of individuals with postnatal famine exposure in these datasets. Moreover, the age range of these individuals at the time of the famine was large (0-20y), while the postnatal associations are mostly found in a younger age range. Indeed, post-natal growth has been shown to have a marked effect on the associations between birth weight and later health[54]. The effect of early childhood famine exposure remains an interesting issue to explore in the larger Dutch Hunger Winter Families study.

An appetite for DNA methylation?

Since the onset of this thesis it was found that adult DNA methylation levels can be influenced by diet[55] and longitudinal variations in BMI[56]. Nutrient restriction and famine exposure during early gestation may lead to a greater appetite for more energy dense foods in animals and possibly humans[57]. The association between early famine exposure and nutrition preferences awaits further investigation, as the AMC Dutch Famine cohort did[71] and the Dutch Hunger Winter Families study did not find evidence for a higher intake of energy dense foods following early famine exposure[72]. In **Chapter 7** we

controlled the famine associations for the amount of calories, fat, protein and carbohydrates in the current diet as to gain insight in this issue. The associations were not influenced by this possible confounder even though the percentage of fat and carbohydrates was associated with DNA methylation at five out of six regions chosen for validation (*data not shown*).

Time of origin: direct or accumulating over the life-course

Animal experiments have shown that besides arising immediately post-exposure, differences may accumulate following the exposure across the life-time[73]. Other studies in humans on the loci identified by us in **Chapter 3 and 4** suggest that the differences may have arisen at or closely after exposure, rather than accumulated during life. Gestational diabetes was associated with *LEP* promoter methylation[20] and maternal folic acid intake and prenatal smoking with DNA methylation at various DMRs around *IGF2*/*H19*[13,58-60] in newborns and young children. Furthermore, analyses of sibling pairs with a smaller age difference resulted in finding larger famine associated DNA methylation differences (**Chapter 3** & *data not shown*), indicating that the differences do not become more pronounced with increasing age. However, measuring DNA from new samples from the same sibling pairs will be necessary to rule out the possibility that the DNA methylation differences may slowly accumulate across the life-course.

Time of origin: early gestation or tissue specificity?

Our results suggest that periconceptional period is the period at which most differential methylation following famine exposure may be found (**Chapter 4**). This period may even be confined to the period after conception (**Chapter 7**). Animal studies show that the blastocyst period is a very sensitive period and that DNA methylation changes induced here may be passed to tissues not yet formed during this developmental stage[10,61-64]. Because of this mitotic inheritance during development caution needs to be taken when designating the early developmental period as the most sensitive in humans. It is possible

that famine exposure late in gestation has a larger effect on DNA methylation, but perhaps only in specific tissues. To further support our hypothesis that early development is the most sensitive period of development in humans we would need to measure DNA methylation from different tissues in both the early and late gestational exposure groups.

Time of origin: pin-pointing

No influence on DNA methylation was found in individuals conceived during April and May 1945, the last two months of the famine (**Chapter 7**). However, we are not able to pin-point the exact developmental time-point at which famine exposure is associated with DNA methylation. The official rations were steadily increasing during May and we do not know when the amount of calories was sufficient. Furthermore, the used estimate of conception, the last menstrual period, has an uncertainty of 2 weeks.
The DNA methylation data also does not allow us to pin-point a certain developmental event. The differentially methylated regions in **Chapter 7** were enriched for genes changing expression during the first week of human development and for those during organogenesis (>5 weeks). Moreover, enhancers active around implantation (week 1-2) were also associated with famine exposure as a group. Experiments on developing blastocysts[10] will be required to gain more precise insight in this matter.

The nature of the exposure

The Dutch Famine offers a quasi-experimental setting to study the effects of well-defined extreme nutritional changes during gestation. Although malnutrition is arguably the largest component to the exposure, the other possible contributors should not be excluded, like the lack of heating during winter and stress. Dutch women normally gave birth at home, while the prenatally exposed individuals studied were born in hospital[4], indicating that their home situation may have been deemed unsuitable for child birth at the time of the famine.

Moreover, the frequent increases in DNA methylation observed (**Chapter 4, 6 and 7**) also hint to the fact that the effect on DNA methylation may not be a simple result of nutritional shortages in methyl-donors and essential co-factors. This may hint that maternal characteristics, like available fat reserves and size, may be important. Indeed, it may be hypothesized that increases in certain energy carriers freed from the fat reserves as a result of starvation, like free fatty acids and 3-hydroxybutyrate[65], may be related to some of the observed DNA methylation differences between the prenatally exposed and unexposed siblings. *In vitro* experiments in embryonic stem cell during differentiation[66], embryoid bodies[67]and on bovine blastocyst development[68], may be performed to test the effect of shortages and surpluses of certain nutritional compounds and metabolites on DNA methylation.

5-hydroxymethylation

During the earliest embryonic stages massive active demethylation of the paternal genome leads to various oxidation products of 5-methylcytosine (mC), the main of which is 5-hydroxy methylcytosine (5hmC), but also higher oxidation products in the form of 5-formyl- and 5-carboxycytosine have been discovered. 5hmC is a stable base and is abundant in embryonic stem cell and is rapidly depleted as the genome is remethylated during blastocyst development and implantation[69]. It is believed that 5hmC is an intermediate for complete demethylation and rapid remethylation[70] and present at low levels in non-neural adult cells[71]. With bisulfite treatment you cannot distinguish between 5hmC and mC. It may therefore be possible that some of the effects reported may be related to differences in 5hmC levels rather than mC, but considering the low levels of 5hmC in adult blood it should have a limited bearing on our results.

Rather than a possible confounder, the influence of nutritional compounds on the mechanism of 5hmC formation and depletion during development may be interesting to study. Recently it was shown that 5hmC levels in embryonic stem cells are influenced by vitamin C levels *in vitro*[72] as one of the main enzymes implicated in 5hmC formation, *tet1*, is driven to a higher efficiency

by higher vitamin C levels[73]. The effect of nutrient shortages and surpluses on *tet1* and its functional partners are prime targets for detailed functional studies on the possible molecular mechanism behind the observed DNA methylation patterns induced by the prenatal environment.

Bringing epigenetic studies in humans to the next level

Epigenetic epidemiology will surely benefit from larger study sizes and the scrutiny of replication, as was recently shown with the first epigenome-wide association study (EWAS) with multiple replications on the effect of smoking[84,85]. But population size alone will not be enough to discover robust associations, since our and the above mentioned studies are well powered considering that the effect sizes discovered were around 0.5 standard deviations.

Epigenetic epidemiology is like epidemiology

The epigenome, including DNA methylation, is highly dynamic. Recently it was shown that DNA methylation at genes involved in immunity can fluctuate within as little as four days[74]. This means that any analysis within and between cohorts can be easily confounded by a myriad of factors, like the time between DNA sampling and phenotyping and dietary differences between populations and ethnic groups. This touches on a dilemma of epidemiology in general, whether it is more efficient to perform studies on several smaller but well-characterized cohorts or on larger cohorts in which important covariates may be missing, a brute force strategy successfully employed in genetic epidemiology. For epigenetic studies the first approach may prove the most suitable one.

Considerations on the ideal cohort study for EWAS

For epigenetic epidemiology it would help to know a lot about the current and past environmental conditions of individuals within a cohort study, as environmental conditions are among the most likely confounders. Secondly the availability of other *–omics* datasets, in particular for genome-wide expression, is essential. Such datasets may help in assessing the functionality of epigenetic marks associated in EWAS, as it is currently unclear at which genomic regions epigenetic variation has biological consequences.
The ideal cohort would also include some longitudinal and tissue sampling[75]. The first can help with building a case for the causality of epigenetic marks within a disease process, as epigenetic variation may merely reflect disease progression. Tissue sampling is required to ascertain if an association relates to variation in a tissue relevant to the phenotypic association. Sampling should ideally also entail the collection of viable cells for cell culture, since such samples may help to elucidate the molecular mechanism underlying an association. Fibroblasts can be isolated from skin biopsies and could be used to test for an influence of epigenetic marks on the binding of transcription factors and the activity of a genomic region. More mechanistic insight may be gained from mesenchymal stem cells isolated from cryo-preserved muscle biopsies[76]. These mesenchymal stem cells can be differentiated into muscle, fat, cardiomyocytes and even ectodermal tissues[77], and offer room for extensive experimentation. Such experiments are of interest, as both animal studies and human studies have yet to uncover the mechanism underlying the epigenetic changes described.

Epigenetic fine-tuning has consequences for the data analyses

If epigenetics acts mainly by small changes across large regions and along entire pathways more focus should be placed on how to analyze data. The current statistical models for grouped analyses are complicated and still limited in their ability to capture all the correlations in DNA methylation data. We grouped DNA methylation on new genome-wide annotations coming out

of various genome wide annotation efforts, including the ENCODE project[36], but still relied on a simplification of the data to perform pathway level tests (**Chapter 7**). More focus will need to be placed on statistical methodologies for grouped and pathway based analyses to tackle epigenome datasets.

Evolutionary considerations

The differentially methylated regions we discovered map to genes and their pathways with key roles in growth and energy allocation, including *INSR*, *LEP*, *IGF2* and *INS*, arguing for an evolutionary component to our findings.

Thrifty genes: selection?

Some of these particular genes were hypothesized to be involved in metabolic disease in the thrifty genotype hypothesis. This hypothesis states that our ancestors evolved under a frugal diet and that our metabolism is thus genetically ill adapted to deal with our modern energy rich diet, causing obesity and metabolic disease[31,90,91]. However, no links between the increase in metabolic disease and genetic variation has been found in populations undergoing a shift from traditional to affluent “Western” diets[78]. Recently this hypothesis was overhauled[79], stating that under adverse conditions during early gestation embryos with a thrifty genotype are selected. The author refers to our Dutch Hunger Winter Families study as the ideal setting to test this hypothesis. Indeed we have already taken some steps in addressing the issue of genetic selection. In **Chapter 6** we searched for differences in genotype frequency in the *IGF2*/*H19* region in exposed and non-exposed individuals, finding no differences. We have also measured SNPs taken from GWAS on body composition and various glucose and lipid traits and as of yet found no differences between prenatally exposed and unexposed individuals (*data not shown*). To comprehensively address this issue SNP genotyping arrays could be run to search for evidence of genetic selection[79], using algorithms developed for genetic anthropology.

There is little doubt that selection may have occurred as the number of new pregnancies dropped 50% during the famine[80]. However, selection on another genomic level should also be considered in light of the increasing number of studies finding epigenetic differences following early gestational exposure. Under nutrition reduces cell numbers in pre- and post implantation rat embryos[81]. Upon famine exposure cells may be selected with a favorable epigenetic signature for growth. We found that differential methylation was enriched along growth related pathways (**Chapter 7**). Moreover we found associations between early gestational famine exposure and key growth and insulin signaling genes *IGF2* (**Chapter 3**), *INS* (**Chapter 4**) and *INSR* (**Chapter 7**). Again, experiments in animal blastocyst models should give insight in this matter by applying more or less nutrients and measuring markers for apoptosis and the final DNA methylation signatures of the cultured blastocyst.

Thrifty phenotypes: phenotypic plasticity

The thrifty genotype theory has been largely replaced by the thrifty phenotype hypothesis[82] that led to the so-called developmental origins hypothesis of adult disease (DOHaD), which introduces modern ideas on phenotypic plasticity into Medicine[28]. The most modern version of the DOHaD hypothesis states that developmental time-frames exist during which the fetus makes persistent adaptations to the perceived environment (predictive adaptive plasticity), disease follows when the adult environment does not match that during development[96]. This hypothesis is well founded in observations in shorter lived animal models and also supported by some observations in humans. Individuals exposed during early development to famine in China, but adhere to a ‘frugal’ diet as adults are less likely to develop disease than those adhering to an affluent Western diet[83,84]. The same principle is seen in the Gambia following seasonal food shortages[85]. Moreover, individuals from under privileged backgrounds show a less severe phenotype upon starvation in childhood than those from a more wealthy background[86].

Epigenetic fine-tuning of metabolic networks is seen as a candidate mechanism for adaptive phenotypic plasticity, as metabolic and growth related gene networks are subject to intense selection, which may have resulted in genetic canalization and an extreme robustness for genetic mutations[42]. Indeed, *leptin* was put forward as a prime candidate for epigenetic fine-tuning and we found DNA methylation at the *leptin* promoter associated with prenatal famine exposure (**Chapter 4**). Moreover, small changes across regulatory and gene networks were hypothesized to be the mechanism by which this fine-tuning occurs, which matches with our observations (**Chapter 6 and 7**).

Phenotypic inertia

However, *predictive* adaptive phenotypic plasticity makes little evolutionary sense in a long lived species as *homo sapiens* and it could only have arisen when famine is a very common phenomenon[87]. This is a debated issue, as famine most likely only became frequent after the advent of agriculture[88], a period too short for predictive adaptive plasticity to have evolved in man[78] considering our long lifespan and reproductive cycle[89]. Furthermore, predictive adaptive plasticity would logically entail that the epigenetic code is reset every generation, while evidence is accumulating that epigenetic adaptations to environmental circumstances may sometimes be passed to the next generation[47,103,104], even when the environmental circumstance that induced the change is no longer present. Also on the epidemiological level the evidence for transgenerational inheritance is building, although still inherently weak[90].

Kuzawa *et al*. shows that there may have been a selective advantage for long-living primates to be able to adapt to gradual differences in climate over generations[89,91]. The envisioned mechanism would allow one generation to adapt to the environment, but never over-act. Key in this hypothesis of 'phenotypic inertia' is the data collected on the buffering capacity of the womb[92] and the transgenerational information contained in the maternal body composition of mammals (maternal constraint), which would allow for

a slow transgenerational adaptation to robust environmental differences[93]. Epigenetic change in response to environmental circumstances is hypothesized to be one of the mechanisms by which the adaptation and buffering may occur[91]. Measurements in sperm of prenatally exposed men and in the children of all prenatally exposed may help to gain insight in this phenomenon in *Homo sapiens.*

Don't forget about mum

For individuals conceived during the famine, like those studied extensively in this thesis, the mothers may be where the selection has acted, resulting in a selected offspring group. They may have had more fat reserves at the time of the famine, and the BMI of the mother at the time of pregnancy has an influence on the child's later adiposity[94]. These mothers may also have had access to better nutrition. Furthermore, the mothers of the individuals conceived during the famine may be more fertile, although investigations on this issue have been largely inconclusive[23].
For the studies in this thesis it is good to emphasize that we compared the individuals conceived during the famine with a same-sex sibling, so that genetic and familial factors are at least partially similar. Furthermore, there are a numbers of animal studies that show that epigenetic differences may arise in the fetus as a result of a nutritional challenge. Both may argue for an interpretation that the associations presented have arisen, at least in part, as a result of the famine exposure, and not solely as a result of maternal characteristics.

Conclusion and Future prospects

We successfully identified regions at which the DNA methylation levels are associated with prenatal famine exposure six decades post exposure. We discovered that these associations may be timing dependent and sex specific, mimicking the reported epidemiological findings. The regions identified map to biologically relevant genes and pathways for the phenotypes associated with

prenatal famine and we discovered tentative associations with phenotypes also associated with prenatal famine exposure.

The Dutch Hunger Winter Families study offers one of the best epidemiological designs to serve as a discovery cohort for DNA methylation variation linked to both developmental adversity and adult disease. We aim to scale our DNA methylation measurements to the entire cohort and test associations in other cohorts. These cohort scale analyses have now become possible by the 450K Illumina DNA methylation micro-array[95], which allows the cost-effective measurement of 450K CpG dinucleotides in large sample sizes. Associations should be robust and consistent and taken further. For we need to relate associations in the Dutch Hunger Winter Families study to other prenatal exposures, measure longitudinal samples to infer causality, investigate relevant tissues and set up models to uncover the molecular mechanisms behind the associations if we are to understand the biology behind the associations and move on to designing possible interventions.

In the introduction of this thesis we referred to the work by Waddington (**Chapter 1, Box 1**), who tried to conceptually merge embryology, evolution and genetics and who coined the term 'epigenetics' to refer to this interdisciplinary endeavor. The Dutch Hunger Winter Study offers the unique opportunity to study epigenetics in its most modern sense, which is very much in the spirit of Waddington. We should try to link embryology, genetics, cell biology and evolution and through this interdisciplinary approach, as a true 'epigenetic' endeavor, enhance our insight in the mechanisms underlying disease and the continuing evolution of *Homo sapiens* in its modern environment.

References

1. Waterland RA, Michels KB. Epigenetic epidemiology of the developmental origins hypothesis. Annu.Rev.Nutr. 2007; 27: 363-88
2. Bogdarina I, Welham S, King PJ, Burns SP, Clark AJ. Epigenetic modification of the renin-angiotensin system in the fetal programming of hypertension. Circ.Res. 2007; 100: 520-6
3. Chmurzynska A. Fetal programming: link between early nutrition, DNA methylation, and complex diseases. Nutr.Rev. 2010; 68: 87-98
4. Lumey LH, Stein AD, Kahn HS, van der Pal-de Bruin KM, Blauw GJ, Zybert PA, et al. Cohort profile: the Dutch Hunger Winter families study. Int.J.Epidemiol. 2007; 36: 1196-204
5. Walther FJ, den Ouden AL, Verloove-Vanhorick SP. Looking back in time: outcome of a national cohort of very preterm infants born in The Netherlands in 1983. Early Hum. Dev. 2000; 59: 175-91
6. Waterland RA, Jirtle RL. Early nutrition, epigenetic changes at transposons and imprinted genes, and enhanced susceptibility to adult chronic diseases. Nutrition 2004; 20: 63-8
7. Constancia M, Hemberger M, Hughes J, Dean W, Ferguson-Smith A, Fundele R, et al. Placental-specific IGF-II is a major modulator of placental and fetal growth. Nature 2002; 417: 945-8
8. Cui H, Cruz-Correa M, Giardiello FM, Hutcheon DF, Kafonek DR, Brandenburg S, et al. Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. Science 2003; 299: 1753-5
9. Cui H, Onyango P, Brandenburg S, Wu Y, Hsieh CL, Feinberg AP. Loss of imprinting in colorectal cancer linked to hypomethylation of H19 and IGF2. Cancer Res. 2002; 62: 6442-6
10. Morgan HD, Jin XL, Li A, Whitelaw E, O'Neill C. The culture of zygotes to the blastocyst stage changes the postnatal expression of an epigentically labile allele, agouti viable yellow, in mice. Biol.Reprod. 2008; 79: 618-23
11. Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. Science 2001; 293: 1089-93
12. Hoyo C, Fortner K, Murtha AP, Schildkraut JM, Soubry A, Demark-Wahnefried W, et al. Association of cord blood methylation fractions at imprinted insulin-like growth factor 2 (IGF2), plasma IGF2, and birth weight. Cancer Causes Control 2012; 23: 635-45
13. Steegers-Theunissen RP, Obermann-Borst SA, Kremer D, Lindemans J, Siebel C, Steegers EA, et al. Periconceptional maternal folic acid use of 400 microg per day is related to increased methylation of the IGF2 gene in the very young child. PLoS.One. 2009; 4: e7845
14. Pidsley R, Dempster EL, Mill J. Brain weight in males is correlated with DNA methylation at IGF2. Mol.Psychiatry 2010; 15: 880-1
15. Susser ES, Lin SP. Schizophrenia after prenatal exposure to the Dutch Hunger Winter of 1944-1945. Arch.Gen.Psychiatry 1992; 49: 983-8
16. Stein AD, Pierik FH, Verrips GH, Susser ES, Lumey LH. Maternal exposure to the Dutch famine before conception and during pregnancy: quality of life and depressive symptoms in adult offspring. Epidemiology 2009; 20: 909-15
17. Nathanielsz PW. Animal models that elucidate basic principles of the developmental origins of adult diseases. ILAR.J. 2006; 47: 73-82
18. Sinclair KD, Allegrucci C, Singh R, Gardner DS, Sebastian S, Bispham J, et al. DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and methionine status. Proc.Natl.Acad.Sci.U.S.A 2007; 104: 19351-6
19. Roseboom T, de RS, Painter R. The Dutch famine and its long-term consequences for adult health. Early Hum.Dev. 2006; 82: 485-91
20. Bouchard L, Thibault S, Guay SP, Santure M, Monpetit A, St-Pierre J, et al. Leptin gene epigenetic adaptation to impaired glucose metabolism during pregnancy. Diabetes Care 2010; 33: 2436-41
21. Barker DJ. The origins of the developmental origins theory. J.Intern.Med. 2007; 261: 412-7

22. Stein AD, Zybert PA, van de BM, Lumey LH. Intrauterine famine exposure and body proportions at birth: the Dutch Hunger Winter. Int.J.Epidemiol. 2004; 33: 831-6

23. Lumey LH, Stein AD, Susser E. Prenatal famine and adult health. Annu.Rev.Public Health 2011; 32: 237-62

24. Wyrwoll CS, Kerrigan D, Holmes MC, Seckl JR, Drake AJ. Altered placental methyl donor transport in the dexamethasone programmed rat. Placenta 2012; 33: 220-3

25. Downing C, Johnson TE, Larson C, Leakey TI, Siegfried RN, Rafferty TM, et al. Subtle decreases in DNA methylation and gene expression at the mouse Igf2 locus following prenatal alcohol exposure: effects of a methyl-supplemented diet. Alcohol 2011; 45: 65-71

26. Guo L, Choufani S, Ferreira J, Smith A, Chitayat D, Shuman C, et al. Altered gene expression and methylation of the human chromosome 11 imprinted region in small for gestational age (SGA) placentae. Dev.Biol. 2008; 320: 79-91

27. van Abeelen AF, Veenendaal MV, Painter RC, de Rooij SR, Thangaratinam S, Van Der Post JA, et al. The fetal origins of hypertension: a systematic review and meta-analysis of the evidence from animal experiments of maternal undernutrition. J.Hypertens. 2012; 30: 2255-67

28. Gluckman PD, Hanson MA. The conceptual basis for the developmental origins of health and disease, Developmental origins of health and disease. Cambridge: Cambridge University Press, 2006: 33-50

29. Bergvall N, Lindam A, Pawitan Y, Lichtenstein P, Cnattingius S, Iliadou A. Importance of familial factors in associations between offspring birth weight and parental risk of type-2 diabetes. Int.J.Epidemiol. 2008; 37: 185-92

30. Fleten C, Nystad W, Stigum H, Skjaerven R, Lawlor DA, Davey SG, et al. Parent-Offspring Body Mass Index Associations in the Norwegian Mother and Child Cohort Study: A Family-based Approach to Studying the Role of the Intrauterine Environment in Childhood Adiposity. Am.J.Epidemiol. 2012; 176: 83-92

31. Johansson S, Iliadou A, Bergvall N, de FU, Kramer MS, Pawitan Y, et al. The association between low birth weight and type 2 diabetes: contribution of genetic factors. Epidemiology 2008; 19: 659-65

32. Oberg S, Cnattingius S, Sandin S, Lichtenstein P, Iliadou AN. Birth weight predicts risk of cardiovascular disease within dizygotic but not monozygotic twin pairs: a large population-based co-twin-control study. Circulation 2011; 123: 2792-8

33. Phillips DI, Davies MJ, Robinson JS. Fetal growth and the fetal origins hypothesis in twins--problems and perspectives. Twin.Res. 2001; 4: 327-31

34. Heijmans BT, Kremer D, Tobi EW, Boomsma DI, Slagboom PE. Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus. Hum.Mol.Genet. 2007; 16: 547-54

35. Giresi PG, Kim J, McDaniell RM, Iyer VR, Lieb JD. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. Genome Res. 2007; 17: 877-85

36. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, et al. The accessible chromatin landscape of the human genome. Nature 2012; 489: 75-82

37. Ivanova E, Chen JH, Segonds-Pichon A, Ozanne SE, Kelsey G. DNA methylation at differentially methylated regions of imprinted genes are resistant to developmental programming by maternal nutrition. Epigenetics. 2012; 7: 1200-10

38. Radford EJ, Isganaitis E, Jimenez-Chillaron J, Schroeder J, Molla M, Andrews S, et al. An unbiased assessment of the role of imprinted genes in an intergenerational model of developmental programming. PLoS.Genet. 2012; 8: e1002605

39. Borghol N, Suderman M, McArdle W, Racine A, Hallett M, Pembrey M, et al. Associations with early-life socio-economic position in adult DNA methylation. Int.J.Epidemiol. 2012; 41: 62-74

40. Khulan B, Cooper WN, Skinner BM, Bauer J, Owens S, Prentice AM, 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: 2086-101

41. Waterland RA, Kellermayer R, Laritsky E, Rayco-Solon P, Harris RA, Travisano M, et al. Season of conception in rural gambia

affects DNA methylation at putative human metastable epialleles. PLoS.Genet. 2010; 6: e1001252

42. Stoger R. The thrifty epigenotype: an acquired and heritable predisposition for obesity and diabetes? Bioessays 2008; 30: 156-66

43. Suter M, Ma J, Harris AS, Patterson L, Brown KA, Shope C, et al. Maternal tobacco use modestly alters correlated epigenome-wide placental DNA methylation and gene expression. Epigenetics. 2011; 6: 1284-94

44. Fradin D, Le FS, Mille C, Naoui N, Groves C, Zelenika D, et al. Association of the CpG methylation pattern of the proximal insulin gene promoter with type 1 diabetes. PLoS. One. 2012; 7: e36278

45. Yang BT, Dayeh TA, Kirkpatrick CL, Taneera J, Kumar R, Groop L, et al. Insulin promoter DNA methylation correlates negatively with insulin gene expression and positively with HbA(1c) levels in human pancreatic islets. Diabetologia 2011; 54: 360-7

46. Guay SP, Brisson D, Munger J, Lamarche B, Gaudet D, Bouchard L. ABCA1 gene promoter DNA methylation is associated with HDL particle profile and coronary artery disease in familial hypercholesterolemia. Epigenetics. 2012; 7: 464-72

47. Lumey LH, Stein AD, Kahn HS, Romijn JA. Lipid profiles in middle-aged men and women after famine exposure during gestation: the Dutch Hunger Winter Families Study. Am.J.Clin.Nutr. 2009; 89: 1737-43

48. Lussana F, Painter RC, Ocke MC, Buller HR, Bossuyt PM, Roseboom TJ. Prenatal exposure to the Dutch famine is associated with a preference for fatty foods and a more atherogenic lipid profile. Am.J.Clin.Nutr. 2008; 88: 1648-52

49. Godfrey KM, Sheppard A, Gluckman PD, Lillycrop KA, Burdge GC, McLean C, et al. Epigenetic gene promoter methylation at birth is associated with child's later adiposity. Diabetes 2011; 60: 1528-34

50. van Abeelen AF, Elias SG, Bossuyt PM, Grobbee DE, van der Schouw YT, Roseboom TJ, et al. Famine Exposure in the Young and the Risk of Type 2 Diabetes in Adulthood. Diabetes 2012; 61: 2255-60

51. van Abeelen AF, Elias SG, Roseboom TJ, Bossuyt PM, van der Schouw YT, Grobbee DE, et al. Postnatal acute famine and risk of overweight: the dutch hungerwinter study. Int.J.Pediatr. 2012; 2012: 936509

52. Huang C, Li Z, Wang M, Martorell R. Early life exposure to the 1959-1961 Chinese famine has long-term health consequences. J.Nutr. 2010; 140: 1874-8

53. Li Y, He Y, Qi L, Jaddoe VW, Feskens EJ, Yang X, et al. Exposure to the Chinese famine in early life and the risk of hyperglycemia and type 2 diabetes in adulthood. Diabetes 2010; 59: 2400-6

54. Tanner JM. Childhood epidemiology. Physical development. Br.Med.Bull. 1986; 42: 131-8

55. Milagro FI, Campion J, Cordero P, Goyenechea E, Gomez-Uriz AM, Abete I, et al. A dual epigenomic approach for the search of obesity biomarkers: DNA methylation in relation to diet-induced weight loss. FASEB J. 2011; 25: 1378-89

56. Feinberg AP, Irizarry RA, Fradin D, Aryee MJ, Murakami P, Aspelund T, et al. Personalized epigenomic signatures that are stable over time and covary with body mass index. Sci. Transl.Med. 2010; 2: 49ra67

57. Portella AK, Kajantie E, Hovi P, Desai M, Ross MG, Goldani MZ, et al. Effects of in utero conditions on adult feeding preferences, 2012: 140-52

58. Ba Y, Yu H, Liu F, Geng X, Zhu C, Zhu Q, et al. Relationship of folate, vitamin B12 and methylation of insulin-like growth factor-II in maternal and cord blood. Eur.J.Clin.Nutr. 2011; 65: 480-5

59. Murphy SK, Adigun A, Huang Z, Overcash F, Wang F, Jirtle RL, et al. Gender-specific methylation differences in relation to prenatal exposure to cigarette smoke. Gene 2012; 494: 36-43

60. Soubry A, Murphy S, Huang Z, Murtha A, Schildkraut J, Jirtle R, et al. The effects of depression and use of antidepressive medicines during pregnancy on the methylation status of the IGF2 imprinted control regions in the offspring. Clin.Epigenetics. 2011; 3: 2

61. Bogdarina I, Haase A, Langley-Evans S, Clark AJ. Glucocorticoid effects on the programming of AT1b angiotensin receptor gene methylation and expression in the rat. PLoS.One. 2010; 5: e9237

62. Dolinoy DC, Weidman JR, Waterland RA, Jirtle RL. Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. Environ. Health Perspect. 2006; 114: 567-72

63. Dolinoy DC, Huang D, Jirtle RL. Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development. Proc.Natl.Acad.Sci.U.S.A 2007; 104: 13056-61

64. McKay JA, Xie L, Harris S, Wong YK, Ford D, Mathers JC. Blood as a surrogate marker for tissue-specific DNA methylation and changes due to folate depletion in post-partum female mice. Mol.Nutr.Food Res. 2011; 55: 1026-35

65. Korbonits M, Blaine D, Elia M, Powell-Tuck J. Metabolic and hormonal changes during the refeeding period of prolonged fasting. Eur.J.Endocrinol. 2007; 157: 157-66

66. Shiraki N, Higuchi Y, Harada S, Umeda K, Isagawa T, Aburatani H, et al. Differentiation and characterization of embryonic stem cells into three germ layers. Biochem.Biophys.Res. Commun. 2009; 381: 694-9

67. Vassilieva S, Goh HN, Lau KX, Hughes JN, Familari M, Rathjen PD, et al. A system to enrich for primitive streak-derivatives, definitive endoderm and mesoderm, from pluripotent cells in culture. PLoS.One. 2012; 7: e38645

68. Bonilla L, Luchini D, DeVillard E, Hansen PJ. Methionine requirements for the preimplantation bovine embryo. J.Reprod. Dev. 2010; 56: 527-32

69. Nabel CS, Manning SA, Kohli RM. The curious chemical biology of cytosine: deamination, methylation, and oxidation as modulators of genomic potential. ACS Chem.Biol. 2012; 7: 20-30

70. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. Nat.Rev.Genet. 2013; 14: 204-20

71. Globisch D, Munzel M, Muller M, Michalakis S, Wagner M, Koch S, et al. Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. PLoS. One. 2010; 5: e15367

72. Blaschke K, Ebata KT, Karimi MM, Zepeda-Martinez JA, Goyal P, Mahapatra S, et al. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature 2013; 500: 222-6

73. Yin R, Mao SQ, Zhao B, Chong Z, Yang Y, Zhao C, et al. Ascorbic Acid enhances tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. J.Am.Chem.Soc. 2013; 135: 10396-403

74. Byun HM, Nordio F, Coull BA, Tarantini L, Hou L, Bonzini M, et al. Temporal stability of epigenetic markers: sequence characteristics and predictors of short-term DNA methylation variations. PLoS.One. 2012; 7: e39220

75. Heijmans BT, Mill J. Commentary: The seven plagues of epigenetic epidemiology. Int.J.Epidemiol. 2012; 41: 74-8

76. Zheng B, Chen CW, Li G, Thompson SD, Poddar M, Peault B, et al. Isolation of myogenic stem cells from cultures of cryopreserved human skeletal muscle. Cell Transplant. 2012; 21: 1087-93

77. Wu X, Wang S, Chen B, An X. Muscle-derived stem cells: isolation, characterization, differentiation, and application in cell and gene therapy. Cell Tissue Res. 2010; 340: 549-67

78. Prentice AM, Rayco-Solon P, Moore SE. Insights from the developing world: thrifty genotypes and thrifty phenotypes. Proc.Nutr. Soc. 2005; 64: 153-61

79. Edwards MJ. Genetic selection of embryos that later develop the metabolic syndrome. Med.Hypotheses 2012; 78: 621-5

80. Stein ZA, Susser M, Saenger G, Marolla F. Famine and Human Development: The Dutch Hunger Winter of 1944-1945. New York: Oxford University Press, 1975

81. Kwong WY, Wild AE, Roberts P, Willis AC, Fleming TP. Maternal undernutrition during the preimplantation period of rat development causes blastocyst abnormalities and programming of postnatal hypertension. Development 2000; 127: 4195-202

82. Hales CN, Barker DJ. The thrifty phenotype hypothesis. Br.Med.Bull. 2001; 60: 5-20

83. Li Y, Jaddoe VW, Qi L, He Y, Lai J, Wang J, et al. Exposure to the Chinese famine in early life and the risk of hypertension in adulthood. J.Hypertens. 2011; 29: 1085-92

84. Li Y, Jaddoe VW, Qi L, He Y, Wang D, Lai J, et al. Exposure to the chinese famine in early

life and the risk of metabolic syndrome in adulthood. Diabetes Care 2011; 34: 1014-8

85. Moore SE, Halsall I, Howarth D, Poskitt EM, Prentice AM. Glucose, insulin and lipid metabolism in rural Gambians exposed to early malnutrition. Diabet.Med. 2001; 18: 646-53

86. Forrester TE, Badaloo AV, Boyne MS, Osmond C, Thompson D, Green C, et al. Prenatal factors contribute to the emergence of kwashiorkor or marasmus in severe undernutrition: evidence for the predictive adaptation model. PLoS.One. 2012; 7: e35907

87. Baig U, Belsare P, Watve M, Jog M. Can Thrifty Gene(s) or Predictive Fetal Programming for Thriftiness Lead to Obesity? J.Obes. 2011; 2011: 861049

88. Speakman JR. Thrifty genes for obesity, an attractive but flawed idea, and an alternative perspective: the 'drifty gene' hypothesis. Int.J.Obes.(Lond) 2008; 32: 1611-7

89. Kuzawa CW. Fetal origins of developmental plasticity: are fetal cues reliable predictors of future nutritional environments? Am.J.Hum. Biol. 2005; 17: 5-21

90. Morgan DK, Whitelaw E. The case for transgenerational epigenetic inheritance in humans. Mamm.Genome 2008; 19: 394-7

91. Kuzawa CW, Thayer ZM. Timescales of human adaptation: the role of epigenetic processes. Epigenomics. 2011; 3: 221-34

92. Prentice AM, Goldberg GR. Energy adaptations in human pregnancy: limits and long-term consequences. Am.J.Clin.Nutr. 2000; 71: 1226S-32S

93. Price KC, Coe CL. Maternal constraint on fetal growth patterns in the rhesus monkey (Macaca mulatta): the intergenerational link between mothers and daughters. Hum. Reprod. 2000; 15: 452-7

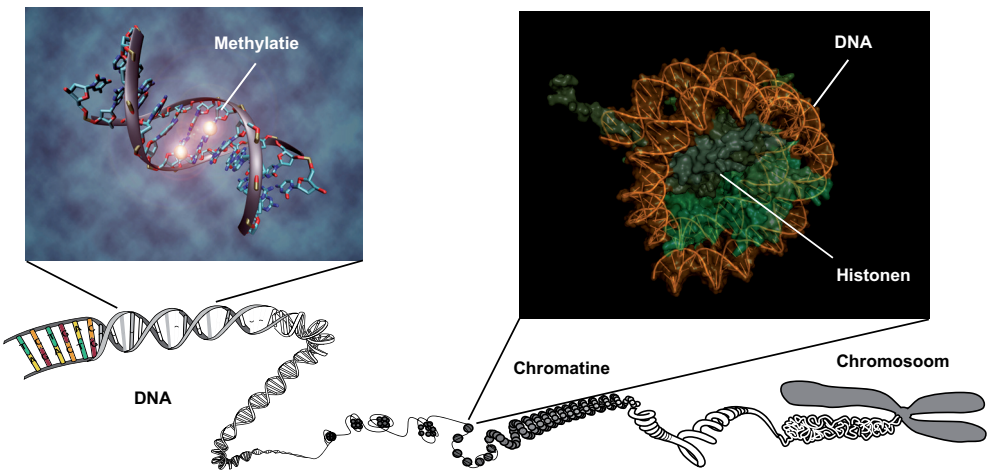
94. Kral JG, Biron S, Simard S, Hould FS, Lebel S, Marceau S, et al. Large maternal weight loss from obesity surgery prevents transmission of obesity to children who were followed for 2 to 18 years. Pediatrics 2006; 118: e1644-e1649

95. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. Epigenomics. 2011; 3: 771-84

Samenvatting

Introductie

Ons DNA bevat alle informatie die nodig is om alle cellen in het lichaam aan te maken en te laten functioneren. Gedurende de embryonale ontwikkeling worden delende cellen steeds meer in een keurslijf gedwongen totdat elke cel één functie krijgt, bijvoorbeeld die van een spiercel of een levercel. Hiervoor moeten delen van de genetische code die de juiste informatie bevatten voor dit celtype aan worden gezet en andere delen uit gezet. De meest bestudeerde mechanismen die dit bewerkstelligen zijn de methylatie van DNA en de aanpassing van histonen. DNA methylatie is een chemische verandering van de CG base paar combinaties in onze erfelijke code en histonen zijn eiwitten waar het DNA omheen gewikkeld is (**Figuur 1**). Beiden beïnvloeden ze hoe compact ons DNA in de celkern is opgeborgen en daarmee hoe toegankelijk dit is voor enzymen en eiwitten om de code 'aan de praat te krijgen' en de expressie van genen te bewerkstelligen. Ze zijn de software die bepalen welk deel van de supercomputer, het DNA, actief is.

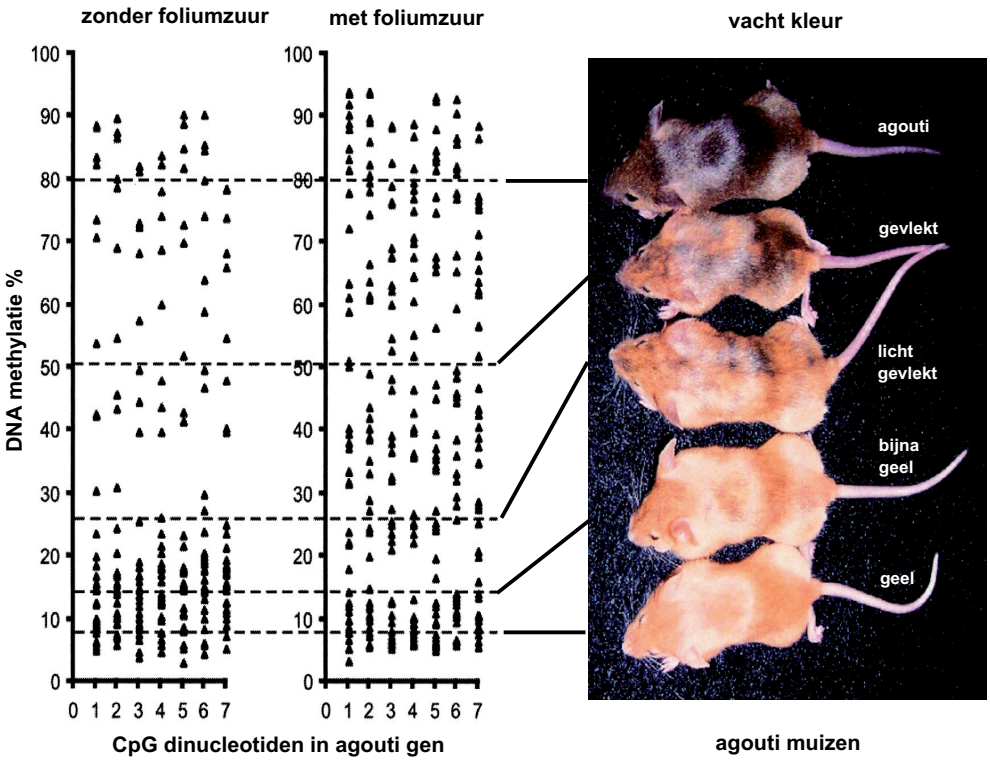


Figuur 1.
DNA methylatie beïnvloedt hoe dicht DNA wordt opgevouwen, doordat er dan wel of niet eiwitten en enzymen aan DNA kunnen plakken. DNA zit normaal opgerold om histonen, grote stukken eiwit, deze histonen kunnen met elkaar dichte rollen vormen (chromatine) die weer opgerold worden in grotere structuren, de chromosomen. ©C.Bock 2012, ©R.Breggen 2012, ©wikimedia

Al heel lang is bekend dat omgeving tijdens de zwangerschap en jeugd invloed heeft op de gezondheid in het latere leven. Dit zijn voorbeelden van biologische fenomenen die niet *1-2-3* te verklaren zijn met de erfelijkheidswetten. Dierstudies wijzen naar een mechanisme om deze relaties te verklaren. *Agouti* muizen kunnen een kleur hebben van geel naar bruin gevlekt tot geheel bruin. Deze kleur erft altijd over via de moederlijke lijn. Ook de hoeveelheid foliumzuur in het dieet van de zwangere muis kan de gemiddelde vachtkleur doen veranderen in het toekomstige nestje. In 2003 werd ontdekt dat deze vachtkleur bepaald wordt door de hoeveelheid DNA methylatie van het vachtkleur gen en dat was afhankelijk van het foliumzuur in het dieet van de moeder. Er wordt geen gele kleur aangemaakt als er veel DNA methylatie op het gen zit: de muis wordt bruin (**Figuur 2**). Deze verschillen konden ook nog eens naar een volgende generatie worden doorgegeven. De mogelijke moleculaire schakel tussen omgevingsfactoren tijdens de vroege ontwikkeling en gevolgen in het latere leven was ontdekt. Andere voedingsexperimenten bij muizen en ratten toonden vervolgens aan dat veranderingen van DNA methylatie ook plaats vindt in genen die betrokken zijn bij de 'ziekte-eigenschappen' van hypertensie en diabetes. De hypothese werd geformuleerd dat dit principe ook wel eens bij de mens kon werken. De basis voor dit proefschrift was geboren.

De Honger Winter

Veel mensen in Nederland hebben geleden onder de Honger Winter van '44-45. Mensen hadden zo weinig te eten dat ze op het dieptepunt minder dan 500 kilocalorieën per dag binnen kregen. Na de oorlog begon langzaam het onderzoek naar de gevolgen van de Honger Winter op de volksgezondheid. Allereerst door het bestuderen van gegevens verzameld door het Centraal Bureau voor de Statistiek en de medische keuring van jonge rekruten. Dit laatste leidde tot een belangrijke publicatie van Dr. Ravelli, Stein en Susser in het gezaghebbend *The New England journal of Medicine* in 1976. Zij stelden vast dat rekruten die waren blootgesteld aan de Honger Winter tijdens de eerste helft van de zwangerschap vaker extreem overgewicht hadden op 19



Figuur 2.

Hier zie je dat met foliumzuur er meer plekken zijn in het DNA van de muis (1 t/m 7) waar de DNA methylatie hoog is. Als de DNA methylatie hoger is, dan is het muisje bruiner van kleur. Met foliumzuur in de voeding van de moedermuis zijn er meer muisjes in het uiteindelijke nestje met een hogere methylatie en dus bruinere kleur. Figuur aangepast uit Waterland *et al*. *MCB*, 2003, ©American Society of Microbiology 2003.

jarige leeftijd. Later zijn er studies opgezet door onderzoekers in Amsterdam, zoals Dr. Lumey, om specifiek die mensen te bestuderen die de Honger Winter hebben doorgemaakt in de baarmoeder. Het bleek dat deze mensen niet alleen een hoger risico hebben op zwaarlijvigheid, maar ook op een ongunstige cholesterol spiegel en een verhoogde kans op diabetes. We weten dankzij geboortegegevens wanneer en hoelang specifieke mensen in de baarmoeder zijn blootgesteld aan extreme ondervoeding. Dit is van belang omdat blootstelling gedurende bepaalde periodes geassocieerd is met een hogere kans op specifieke gezondheidsklachten later in het leven.

Zo hebben mensen die verwekt zijn tijdens de Honger Winter een verhoogd risico op schizofrenie.
Tijdens de Honger Winter functioneerden de ziekenhuizen, waardoor we van sommige mensen aan de hand van hun geboortegegevens met zekerheid kunnen achterhalen gedurende welke periode ze zijn blootgesteld. Verder was de blootstelling heftig en van toepassing op de hele populatie, dit maakt dat de ramp van de Honger Winter een schat van informatie kan opleveren over de invloed van prenatale ondervoeding op de gezondheid in het latere leven en de moleculaire mechanismen die hier een rol bij spelen.

De basis: wat te meten, en de eerste stappen

Koelkasten en vriezers vol, veel onderzoeksinstituten hebben DNA opgeslagen, maar monsters waar nog histonen in gemeten kunnen worden zijn helaas heel schaars. Ook is de technologie om DNA methylatie te meten heel erg verbeterd zodat er nu veel meer genen en mensen tegelijkertijd kunnen worden gemeten. We zijn aan het begin van dit proefschrift eerst gaan onderzoeken of er verschillen zijn in DNA methylatie tussen normale mensen en hoe stabiel deze DNA methylatie is (**Hoofdstuk 2**).
Daarna wilden we weten of omgevingsfactoren tijdens de ontwikkeling invloed kunnen hebben op DNA methylatie van belangrijke genen in de mens. We kozen voor het groei-gen *IGF2.* Dit gen speelt een cruciale rol tijdens de vroege ontwikkeling. *IGF2* regelt de overdracht van voeding via de placenta naar de foetus. Verder erft DNA methylatie van dit gen via de moeder over. Onze hypothese was dat vooral blootstelling rond de conceptie en het eerste trimester van invloed kon zijn op DNA methylatie van dit gen. De Honger Winter Familie studie bood ons de mogelijkheid dit uit te zoeken. Voor het epigenetisch onderzoek hebben we DNA verkregen uit bloed van 122 mensen die verwekt waren of al in het laatste trimester van de embryonale ontwikkeling waren tijdens de Honger Winter. We vergeleken de DNA methylatie van het *IGF2* gen tussen de mensen die in de baarmoeder zijn blootgesteld en hun broer of zus die niet in de baarmoeder zijn blootgesteld (**Hoofdstuk 3**). Degenen die heel vroeg in de ontwikkeling waren blootgesteld

hadden minder DNA methylatie dan hun broer of zus, terwijl degenen die laat in de baarmoeder waren blootgesteld (doordat de moeder al zwanger was toen de Hongerwinter begon) geen verschil met hun broer of zus hadden. Op deze manier konden we als eerste onderzoeksgroep aantonen dat ondervoeding vroeg in de ontwikkeling de DNA methylatie in mensen kan beïnvloeden.

Hoe verhoudt dit zich tot de medische bevindingen?

Mensen die blootgesteld zijn aan de Honger Winter in de baarmoeder hebben 60 jaar later een hoger risico op zwaarlijvigheid, een ongunstige cholesterol spiegel en diabetes. Deze associaties zijn vaak specifiek voor een specifieke prenatale blootstellingsperiode of zijn alleen gevonden in vrouwen of mannen. Volgen DNA methylatie veranderingen dezelfde patronen? Daarom keken we vervolgens in nog eens 15 kandidaat genen die een rol spelen in de vroege ontwikkeling en de stofwisseling in dezelfde mensen als voor *IGF2* (**Hoofdstuk 4**). Wederom vonden we genen met een andere DNA methylatie na een Honger Winter blootstelling. Nu vonden we echter ook toenames in DNA methylatie, die vaak ook nog eens alleen in mannen of alleen in vrouwen bleken voor te komen. Ook was één gen, leptine (*LEP*), zowel verhoogd in de vroeg als laat in de zwangerschap blootgestelde mensen. De meeste genen waren gevoelig voor de vroege blootstellings periode, maar de DNA methylatie veranderingen volgden dezelfde patronen als de epidemiologische bevindingen over o.a. diabetes.

Relevantie voor de huidige dag?

Er bestaat veel epidemiologisch onderzoek waarin men een verband aantoont tussen een laag geboortegewicht of een prenatale groei beperking en dezelfde problemen op middelbare leeftijd als bij mensen die in de baarmoeder aan de Honger Winter waren blootgesteld. Er wordt dan ook gesuggereerd dat een laag geboortegewicht wijst op een slechte overdracht van voeding van de moeder naar het kind. Een redelijk controversiële

aanname. Echter, diermodellen die deze relaties moeten nabootsen omvatten vaak toch interventies in de voeding van zwangere proefdieren. Echter, een verstoring in de bloedsomloop in de placenta kan ook prenatale groei beperking veroorzaken. Deze 'placentale insufficiëntie' heeft echter andere gevolgen op genen dan voedingsinterventies in dieren.

We vroegen ons dan ook af of prenatale groei beperking lijkt op ondervoeding qua DNA methylatie veranderingen in de mens. We vergeleken de DNA methylatie van 4 genen waarvan we in ons onderzoek hadden gevonden dat deze gevoelig waren voor de Honger Winter blootstelling in het DNA van een groep jong volwassenen die te vroeg geboren waren en die meededen in de zogenaamde POPS studie. De ene helft had een prenatale groei beperking en de andere helft niet (**Hoofdstuk 5**). We bestudeerden te vroeg geboren jong volwassenen omdat blootstelling vroeg in de zwangerschap het meeste gevolgen leek te hebben op DNA methylatie. We vonden geen DNA methylatie verschillen tussen mensen met en zonder prenatale groei beperking.

Kunnen deze DNA methylatie verschillen dan niks zeggen over moderne zwangerschappen en hun complicaties? Andere onderzoekers hebben DNA methylatie verschillen gevonden met andere blootstellingen. Zo is DNA methylatie van leptine (**Hoofdstuk 4**) ook geassocieerd met zwangerschaps diabetes en met een lage sociaal economische status van de moeder en de duur van borstvoeding geven. *IGF2* methylatie (**Hoofdstuk 3**) van het kind is ook geassocieerd met de hoeveelheid vitamine B_{12} in het bloed van de moeder en haar rookgedrag vroeg in de zwangerschap. Deze laatste bevinding is al in 2 studies opnieuw gedaan. De studie van de Honger Winter lijkt dus wel degelijk relevant voor het hier en nu in Nederland. Daarnaast komen hongersnoden helaas nog vaak voor in de wereld.

Klein maar veel; Nurture & Nature

De meeste DNA methylatie verschillen zijn klein en we wilden graag weten hoe dit genen zou kunnen beïnvloeden. Zijn ze misschien met velen? Daarom breidden we onze metingen uit in de 60 mensen met een Honger

Winter blootstelling vroeg in de zwangerschap. We keken naar bijna alle stukken DNA rond het *IGF2* gen (**Hoofdstuk 6**) waarvan bekend is dat ze de werking van het *IGF2* gen beïnvloeden. Verder bekeken we ook hoe groot de invloed was van de genetische variatie op DNA methylatie van *IGF2*. Want is klein wel klein en niet gewoon normaal? DNA methylatie van bijna alle stukken DNA was geassocieerd met de Honger Winter blootstelling. Dit kan mogelijk verklaren hoe DNA methylatie de activiteit van genen beïnvloedt: veel kleine veranderingen maken samen het verschil. De invloed van variatie in de DNA code zelf bleek even groot als de invloed van de Honger Winter blootstelling. DNA methylatie van sommige stukken rond het bestudeerde gen was geassocieerd met zowel DNA variaties als de honger winter blootstelling. Deze associaties waren onafhankelijk van elkaar. Nature en Nurture lijken elkaar niet in de weg te zitten.

Het hele genoom

Op welke stukken DNA is DNA methylatie nu gevoelig voor omgevings invloeden tijdens de (vroege) ontwikkeling? Om deze vraag te beantwoorden hebben we metingen verricht aan een veel groter gedeelte van het DNA (een veel groter deel van ons genoom zeggen we dan) dan de paar genen die we tot nu toe hadden bestudeerd. Deze opschaling kon worden bewerkstelligd door de allernieuwste techniek te gebruiken om DNA methylatie te meten. We konden 24 van de 60 vroeg blootgestelde mensen en hun broer of zus meten (**Hoofdstuk 7**). Op iets meer dan 1.2 miljoen plekken werd DNA methylatie gemeten.

Uit de analyse van deze data bleek dat de blootstelling aan de Honger Winter vooral een effect had op DNA methylatie in stukken van het genoom die meer en vaker ‘aan’ staan en een regelende functie hebben. Ook delen van het genoom die actief zijn rond en na de innesteling van het embryo in de baarmoeder waren aangedaan. Verrassend genoeg bleken de meeste aangedane delen midden in genen te zitten, terwijl onderzoekers tot nu toe vaak alleen de delen vlak voor een gen bestuderen. We deden nog een opmerkelijke ontdekking: de mensen die verwekt waren laat in de Honger

Winter bleken de gevonden DNA methylatie verschillen niet te hebben. In deze mensen bleek de Honger Winter geen spoor op hun DNA achter te laten.
De DNA methylatie veranderingen waren verrijkt in groepen genen die samen de insuline en vet huishouding regelen en betrokken zijn in de vroege groei en ontwikkeling. Dit is een interessante observatie omdat een vroege prenatale Honger Winter blootstelling niet leidt tot een afname in geboortegewicht, terwijl het risico later in het leven op zwaarlijvigheid, een ongunstige cholesterol spiegel en diabetes wel hoger is. DNA methylatie in het *CPT1A* gen bleek geassocieerd met zowel de Honger winter blootstelling als 'slecht' LDL cholesterol en DNA methylatie in de insuline receptor (*INSR*) was geassocieerd met zowel de Honger Winter blootstelling als geboortegewicht. De relatie tussen *CPT1A* methylatie en LDL was hetzelfde in de blootgestelde individuen en de controle broers en zussen. De hogere DNA methylatie bij *CPT1A* zou dus potentieel kunnen bijdragen aan het risico op een hoger LDL cholesterol bij de mensen die prenataal zijn blootgesteld aan de Honger Winter.

Toekomst

We weten nu meer over hoe de Honger Winter DNA methylatie beïnvloedt van mensen die zijn blootgesteld in de baarmoeder. We weten hoe groot deze verschillen zijn en in wat voor stukjes DNA ze het vaakst voorkomen en dat DNA waarschijnlijk pas na de bevruchting gevoelig is. Maar we begrijpen nog niet hoe het komt dat de honger heeft bewerkstelligd dat er ook toenames in DNA methylatie worden gevonden. De karige rantsoenen bevatte ten slotte mogelijk te weinig vitamines en foliumzuur, noodzakelijk om DNA te methyleren, dus we verwachtten eerder een afname in DNA methylatie. Experimenten waarbij we zich ontwikkelende cellen tijdelijk blootstellen aan tekorten en te veel suikers en essentiële vetten, aminozuren en vitamines moeten hier uitsluitsel over geven.
We weten ook nog niet genoeg over de relatie van de DNA methylatie verschillen met de verhoogde risico's op overgewicht, diabetes en schizofrenie.

Onze allereerste associaties met LDL cholesterol en geboortegewicht zijn dan wel interessant, maar moeten nog herhaald en verder uitgeplozen worden. Hiervoor moeten we veel meer mensen meten om genoeg kracht te hebben in onze analyses en de bevindingen moeten bevestigd worden in onafhankelijke studies. Verder maken de Honger Winter kinderen zich zorgen of er een effect kan worden doorgegeven aan hun eigen kinderen. We willen dan ook graag gaan onderzoeken of er ook DNA methylatie verschillen te vinden zijn in deze kinderen. Met al deze zaken zijn we inmiddels al druk in de weer; de wetenschap is nooit klaar.

Dankwoord
Curriculum Vitae
Publicaties

Dankwoord

Deze weg leg je niet alleen af, al moet je het wel zelf doen.
Geachte professor Slagboom, beste Eline, bedankt voor de mogelijkheid een kijkje te nemen in de keuken van de wetenschap. Dan Dr. Heijmans, mijn co-promotor en dagelijks begeleider. Beste Bas, je vermogen naast een leermeester ook een toegankelijk mede wetenschapper en collega te zijn is mede de reden geweest dat mijn stage een promotie werd en mijn promotie een postdoc. Daarnaast is er ook nog mijn andere co-promotor Dr. Lumey. Bedankt voor het vertrouwen en je advies bij het werk aan je magnus opus.
In het lab en daarbuiten waren Dennis en Eka van onmisbare hulp en steun. Ook aan Peter heb ik zeer veel steun gehad met zijn vele tips, connecties en gesprekken. Mijn meest directe wetenschappelijk collega was Rudolf. Bedankt voor je luisterend oor, samen stonden we aan de start en beiden halen we de eindstreep. Daarnaast wil ik ook Erik bedanken. Onze gesprekken bevestigden altijd dat, ondanks de vele nadelen, het ook fijn is wetenschapper te zijn. Het werk deze afgelopen jaren was niet hetzelfde of mogelijk geweest zonder alle andere ‘malle eppies’. Daarbij wil ik expliciet de kerstmiskamer c.q. de apenrots bedanken en de mede borrelaars.
Mijn werk werd ondersteund door statistici en een programmeur. Professor Putter, Jelle en Erik van Zwet bedankt voor het werk aan de statistische vraagstukken en Ramin voor de hulp bij de RRBS data. Ook studenten zijn een steuntje in de rug. Het was een geluk Jenny als eerste studente te hebben. Daarna kwam Suzanne. Beste Suzanne, je project was zwaarder dan gedacht en het werk viel je ook zwaarder dan gedacht. Bedankt voor je harde werk. Mijn respect voor je doorzettingsvermogen is groot. Op het staartje kwam Arthur. Ik vond het fijn je als student te hebben en je eerste stapje in de wetenschap te begeleiden. Daarnaast wil ik ook professor Wit, professor Delemarre-van de Waal en Dr. Finken bedanken voor het werk aan het POPS cohort. Professor van der Maarel en Professor Zwaan ben ik erkentelijk voor hun adviezen en de wetenschappelijke discussies.
It was good to satisfy my hunger for knowledge within a large EU project. I wish to thank professor Westendorp for this support and mentorship. It was

wonderful to work with Vincencio, Maris, Agnieszka and Charlotte. Our EU workshop format may perhaps stand the test of time. I am grateful for the trust Randi, Anastasia, Lene and Kaare put in me during the (pilot) projects on their samples and studies. From so many other LIFESPAN participants I learned and even continue to learn much, thank you Joost, Sanne, Sven and the many others.

With the support of LIFESPAN I visited Harvard and the Broad. I thank Dr. Alex Meissner for welcoming me to his group, Dr. Hongcang Gu for his mentorship in the lab and Dr. Christoph Bock for his invaluable help with the sequencing data. I thank Fabian for being my guide at Harvard, for his friendship and help with analyses.

Mijn ouders en zus waren een grote kracht achter het kiezen voor dit traject, aan hen heb ik mijn brede interesses te danken. Als ik 's nachts deze laatste paragraaf schrijf, roept mijn vriendin vanuit bed in de kamer naast mij er een punt achter te zetten. Hermien, mijn grootste steun en toeverlaat, dat was en ben jij. PUNT.

Curriculum Vitae

Elmar Wouter Tobi was born on the 6th of May 1981 in Delft. In the year 2000 he finished his A-levels at the 'Christelijk Lyceum Delft' and went on to the undergraduate program Life Science & Technology, a joint program by Delft University of Technology and Leiden University. His first research internship was at the department of environmental biotechnology at the Kluyver Laboratory in Delft. The topic of this internship was to optimize a method to detect the expression of metabolic genes of sulfate reducing bacteria in complex environmental samples. His main research internship as an undergraduate was at the Leiden University Medical Center, at the department of molecular epidemiology. Here he validated and optimized a technology to measure DNA methylation. After graduating at the end of 2006 (with honours) he started his PhD at the same department under supervision of Professor Slagboom and Dr. Heijmans. His PhD research was performed within the interdisciplinary European project LIFESPAN, which aimed to boost interdisciplinary research into development and ageing. Beside multiple pilot experiments within LIFESPAN his main goal was to study the epigenetic consequences of prenatal famine exposure. The results of this research are outlined in this Thesis. Currently he is employed as a researcher at the same department, studying the epigenetic consequences of the Dutch Famine further and performing epigenome-wide association studies.

List of publications

Heijmans BT, Kremer D, **Tobi EW**, Boomsma DI, Slagboom PE. Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human *IGF2*/*H19* locus. *Hum*. *Mol*. *Genet*. 2007: 547-54

Heijmans BT & **Tobi EW**, Stein AD, Putter H, Blauw GJ, Susser ES, Slagboom PE, Lumey LH. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *PNAS*. 2008: 17046-9

Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD, Slagboom PE, Heijmans BT. DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum*. *Mol*. *Genet.* 2009: 4046-53

Heijmans BT, **Tobi EW**, Lumey LH, Slagboom PE. The epigenome: archive of prenatal environment. *Epigenetics*. 2009: 526-31.

Talens RP, Boomsma DI, **Tobi EW**, Kremer D, Jukema JW, Willemsen G, Putter H, Slagboom PE, Heijmans BT.Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. *FASEB J*. 2010:3135-44

Tobi EW, Heijmans BT, Kremer D, Putter H, Delemarre-van de Waal HA, Finken MJ, Wit JM, Slagboom PE. DNA methylation of *IGF2*, *GNASAS*, *INSIGF* and *LEP* and being born small for gestational age. *Epigenetics*. 2011: 171-6.

Tobi EW, Slagboom PE, van Dongen J, Kremer D, Stein AD, Putter H, Heijmans BT, Lumey LH. Prenatal famine and genetic variation are independently and additively associated with DNA methylation at regulatory loci within *IGF2*/*H19*. *PLoS One*. 2012: e37933.

Obermann-Borst SA, Heijmans BT, Eilers PHC, **Tobi EW**, Steegers EAP, Slagboom PE, Steegers-Theunissen RPM. Periconception maternal smoking and low education are associated with methylation of *INSIGF* in children at the age of 17 months. *J*. *DOHaD*. 2012:315-20

Obermann-Borst SA, Eilers PHC, **Tobi EW**, de Jong F, Slagboom PE, Heijmans BT, Steegers-Theunissen RPM. Duration of breastfeeding and gender are associated with methylation of the *LEPTIN* gene in very young children. *Pediatr*. *Res*. 2013: 344-9

Thijssen PE, **Tobi EW**, Balog J, Schouten SG, Kremer D, El Bouazzaoui F, Henneman P, Putter H, Slagboom PE, Heijmans BT, van der Maarel SM. Chromatin remodeling of human subtelomeres and TERRA promoters upon cellular senescence: commonalities and differences between chromosomes. *Epigenetics*. 2013: 512-21

Tobi EW, Goeman JJ, Monajemi R, Gu H, Putter H, Zhang Y, Müller F, van Zwet EW, Bock C, Meissner A, Lumey LH, Slagboom PE, Heijmans BT. DNA methylation signatures link prenatal famine exposure to growth and metabolism. *Manuscript submitted*

Supplement I

Table S1. Association of neutrophil proportion with DNA methylation of the recent blood samples of 34 additional individuals from the NTR biobank

Locus	Variance explained	p-value of effect
IL10	27.9 %	$8.0*10^{-08}$
IGF2R	1.2 %	0.312
LEP	1.8 %	$1.0*10^{-04}$
CRH	0.0 %	0.993
IGF2	0.6 %	0.378
INSIGF	0.1 %	0.677
KCNQ1OT1	0.0 %	0.808
APOC1	0.1 %	0.763

Table S2. Characteristics of 30 individuals, selected from the NTR biobank

Male, no. (%)	12	(40 %)
Female, no. (%)	18	(60 %)
Smoker, no. (%)	3	(10 %)
Ex-smoker, no. (%)	15	(50 %)
Non-smoker, no. (%)	12	(40 %)
Age at biobanking, mean (SD)	48	(16)
Waist circumference in cm, mean (SD)	98	(22)
LDL-cholesterol (in mmol/L) , mean (SD)	2.47	(1.63)
HDL-cholesterol (in mmol/L) , mean (SD)	1.51	(0.94)
Glucose (in mmol/L) , mean (SD)	9.10	(4.83)

Table S3. Age and longitudinal sampling of 34 individuals from the NTR biobank

Age at first sampling (blood)	Years to first follow-up (buccal swab)	Years to second follow-up (blood and buccal swab)
62	6	13
62	6	13
60	9	12
54	6	14
52	6	13
52	6	13
48	6	13
48	6	13
46	10	12
46	10	12
45	10	12
45	10	12
43	6	13
43	6	12
42	6	11
41	6	13
41	10	12
41	6	12
41	10	12
39	6	11
39	9	12
36	10	12
36	10	14
36	10	12
34	9	12
21	13	15
19	10	16
17	15	17
17	14	16
16	11	16
16	13	16
14	17	19
14	16	18
14	15	19

Table S4. Primers used in bisulfite PCR

Locus	Forward primer[1]	Reverse primer[2]
IL10	TGATTGGTTGAATATGAATTTTTGTAT	CACCCCCTCATTTTTACTTAAAAA
NR3C1	GATTTGGTTTTTTTGGGG	TCCCTTCCCTAAAACCT
TNF	GGGTATTTTTGATGTTTGTGTGTT	CAATACTCATAATATCCTTTCCAAAAAA
IGF2R	AGGTAGAAAAAGGTTTTGGAAG	CAAATCTTAAAAACTAACTAAAAACC
GRB10	GGAATTTTAGGATTAAATTTATGTGA	AACTTCCAAAAAAAACCTCTCC
LEP	GTTTTTGGAGGGATATTAAGGATTT	CTACCAAAAAAAACCAACAAAAAAA
CRH	TGGTTGTTGTTTTTTTGGTAGG	AATTTCTCCACTCCAAAACCTAAA
ABCA1	ATTTTATTGGTGTTTTTGGTTGT	ATCAAAACCTATACTCTCCCTCCTC
IGF2	TGGATAGGAGATTGAGGAGAAA	AAACCCCAACAAAAACCACT
INSIGF	GTTTTGAGGAAGAGGTGTTGA	ACCTAAAATCCAACCACCCTAA
KCNQ1OT1	TTTGGTAGGATTTTGTTGAGGAGTTTT	CTCACACCCAACCAATACCTCATAC
MEG3	TTTTTTTTAATAGTATTTTGATTTTTG	AAATAATCCCCACACACATACC
FTO	GTTTGTAATTTTAGTATTTTGGGAGGT	TTTATTTCCATTTATCCATTCTCAAA
APOC1	GGAGGAGGGAGATTAATATTAATTTGT	ACCCCAAACCTATAACCACCTT
GNASAS	GTAATTTGTGGTATGAGGAAGAGTGA	TAAATAACCCAACTAAATCCCAACA
GNAS A/B	ATGATTTAATTAAGGTTTTAGGAAAGG	TAAAAATACAAAACCTCCCCTACTC
10-mer tag	AGGAAGAGAG + primer	
T7 tag		CAGTAATACGACTCACTATAGGGAGAAGGCT + primer

1. Forward primer, an additional tag is added for Sequenom Epityper PCRs which is denoted below
2. Reverse primer, an additional tag is added for Sequenom Epityper PCRs which is denoted below

Table S5. CpG-sites per fragment of the loci that were analyzed for variation

Locus	CpG-sites* analyzed
IL10	1, 2&3, 4
NR3C1	1&2, 4, 7&8, 9, 10&11, 12&13, 14, 15&16, 17-20, 31, 33&34
TNF	1-3, 5, 6, 9, 10, 11
IGF2R	4&5, 8-10, 11-13, 20&21
GRB10	1&2, 4-6, 7, 8, 17, 18-21, 22&23, 24, 25
LEP	1, 8, 16%17, 19-21, 22, 25, 27
CRH	1, 2, 5, 9, 10
ABCA1	1, 3&4, 6-9, 15&16, 17&18, 19-21, 24, 25
IGF2	3, 4, 6&7, 8
INSIGF	2, 4, 5, 6
KCNQ1OT1	1, 6, 8&9, 10-12, 15, 16, 17&18, 20, 21, 25
MEG3	2, 3, 4, 8&9, 10&11
FTO	2&3, 7, 8&9, 10&11, 14, 17, 19
APOC1	1, 2, 3, 4, 10, 11
GNASAS	1&2, 3&4, 6, 7, 8&9, 10-12, 13&14, 15, 17-19
GNAS A/B	1, 3&4, 7, 8, 13-15, 16-19

* CpG-site number is counted from the forward primer onward

Table S6. CpG-sites per fragment of the loci that were analyzed for stability

Locus	CpG-sites* analyzed
IL10	1, 2&3, 4
IGF2R	4&5, 8-10, 11-13, 20&21, 22†
LEP	1, 8, 16&17, 19-21, 22†, 25, 27
CRH	1†, 2, 5, 9
IGF2	4, 6&7, 8
INSIGF	2, 4, 5, 6
KCNQ1OT1	1, 6, 10-12, 15, 16, 17&18, 20, 21, 25
APOC1	1, 2, 3, 4, 10, 11

* CpG-site number is counted from the forward primer onward

† CpG-site measurement met the quality criteria only in the recent blood samples

Table S7. All CpG-units of the 16 loci

Locus_CpG-unit*	CpG-sites*	Reason for removal prior to quality control
*IL10*_01	CpG_1	
*IL10*_02	CpG_2&3	
*IL10*_03	CpG_4	
*NR3C1*_01	CpG_1&2	
*NR3C1*_02	CpG_3	
*NR3C1*_03	CpG_4	
*NR3C1*_04	CpG_5&6	
*NR3C1*_05	CpG_7&8	
*NR3C1*_06	CpG_9	
*NR3C1*_07	CpG_10&11	
*NR3C1*_08	CpG_12&13	
*NR3C1*_09	CpG_14	
*NR3C1*_10	CpG_15&16	
*NR3C1*_11	CpG_17-20	
*NR3C1*_12	CpG_21	Mass overlap with unit 14
*NR3C1*_13	CpG_22-28	
*NR3C1*_14	CpG_29	Mass overlap with unit 12
*NR3C1*_15	CpG_30	
*NR3C1*_16	CpG_31	
*NR3C1*_17	CpG_32	
*NR3C1*_18	CpG_33&34	
*NR3C1*_19	CpG_35-41	High Mass
*NR3C1*_20	CpG_42	rs5871844 and rs34027900
*TNF*_01	CpG_1-3	
*TNF*_02	CpG_4	
*TNF*_03	CpG_5	
*TNF*_04	CpG_6	
*TNF*_05	CpG_7&8	
*TNF*_06	CpG_9	
*TNF*_07	CpG_10	
*TNF*_08	CpG_11	
*IGF2R*_01	CpG_1	Mass overlap with unit 8
*IGF2R*_02	CpG_2	Mass overlap with unit 3
*IGF2R*_03	CpG_3	Mass overlap with unit 2
*IGF2R*_04	CpG_4&5	
*IGF2R*_05	CpG_6&7	Mass overlap with unit 11
*IGF2R*_06	CpG_8-10	
*IGF2R*_07	CpG_11-13	
*IGF2R*_08	CpG_14	Mass overlap with unit 1
*IGF2R*_09	CpG_15&16	Mass overlap with unit 10 and rs677882 and rs8191722
*IGF2R*_10	CpG_17	Mass overlap with unit 9
*IGF2R*_11	CpG_18&19	Mass overlap with unit 5 and rs8191721 and rs8191720
*IGF2R*_12	CpG_20&21	
*IGF2R*_13	CpG_22	

Table S7. (Continued A): All CpG-units of the 16 loci

Locus_CpG-unit*	CpG-sites*	Reason for removal prior to quality control
*GRB10*_01	CpG_1&2	
*GRB10*_02	CpG_3	Mass overlap with unit 9
*GRB10*_03	CpG_4-6	
*GRB10*_04	CpG_7	
*GRB10*_05	CpG_8	
*GRB10*_06	CpG_9&10	Mass overlap with unit 8
*GRB10*_07	CpG_11	Low mass
*GRB10*_08	CpG_12	Mass overlap with unit 6
*GRB10*_09	CpG_13	Mass overlap with unit 2
*GRB10*_10	CpG_14&15	
*GRB10*_11	CpG_16	
*GRB10*_12	CpG_17	
*GRB10*_13	CpG_18-21	
*GRB10*_14	CpG_22&23	
*GRB10*_15	CpG_24	
*GRB10*_16	CpG_25	
*LEP*_01	CpG_1	
*LEP*_02	CpG_2-7	High Mass and rs791620
*LEP*_03	CpG_8	
*LEP*_04	CpG_9&10	Mass overlap with unit 9
*LEP*_05	CpG_11	Mass overlap with units 12 and 6
*LEP*_06	CpG_12&13	Mass overlap with units 5 and 12
*LEP*_07	CpG_14&15	Mass overlap with unit 14
*LEP*_08	CpG_16&17	
*LEP*_09	CpG_18	Mass overlap with unit 4
*LEP*_10	CpG_19-21	
*LEP*_11	CpG_22	
*LEP*_12	CpG_23&24	Mass overlap with units 5 and 6
*LEP*_13	CpG_25	
*LEP*_14	CpG_26	Mass overlap with unit 7
*LEP*_15	CpG_27	
*LEP*_16	CpG_28	
*LEP*_17	CpG_29	rs2167270
*LEP*_18	CpG_30-32	High Mass
*CRH*_01	CpG_1	
*CRH*_02	CpG_2	
*CRH*_03	CpG_3	
*CRH*_04	CpG_4	
*CRH*_05	CpG_5	
*CRH*_06	CpG_6	Mass overlap with unit 7
*CRH*_07	CpG_7	Mass overlap with unit 6
*CRH*_08	CpG_8	
*CRH*_09	CpG_9	
*CRH*_10	CpG_10	

Table S7. (Continued B): All CpG-units of the 16 loci

Locus_CpG-unit*	CpG-sites*	Reason for removal prior to quality control
*ABCA1*_01	CpG_1	
*ABCA1*_02	CpG_2	
*ABCA1*_03	CpG_3&4	
*ABCA1*_04	CpG_5	
*ABCA1*_05	CpG_6-9	
*ABCA1*_06	CpG_10-13	rs2246298
*ABCA1*_07	CpG_14	
*ABCA1*_08	CpG_15&16	
*ABCA1*_09	CpG_17&18	
*ABCA1*_10	CpG_19-21	
*ABCA1*_11	CpG_22&23	rs13306071
*ABCA1*_12	CpG_24	
*ABCA1*_13	CpG_25	
*ABCA1*_14	CpG_26&27	rs2740483
*IGF2*_01	CpG_1	rs3741208 and rs17883577
*IGF2*_02	CpG_2	rs3741209
*IGF2*_03	CpG_3	
*IGF2*_04	CpG_4	
*IGF2*_05	CpG_5	rs4930041
*IGF2*_06	CpG_6&7	
*IGF2*_07	CpG_8	
*INSIGF*_01	CpG_1	Low mass
*INSIGF*_02	CpG_2	
*INSIGF*_03	CpG_3	
*INSIGF*_04	CpG_4	
*INSIGF*_05	CpG_5	
*INSIGF*_06	CpG_6	
*KCNQ1OT1*_01	CpG_1	
*KCNQ1OT1*_02	CpG_2	
*KCNQ1OT1*_03	CpG_3-5	Mass overlap with unit 8
*KCNQ1OT1*_04	CpG_6	
*KCNQ1OT1*_05	CpG_7	Mass overlap with units 15 and 19
*KCNQ1OT1*_06	CpG_8&9	
*KCNQ1OT1*_07	CpG_10-12	
*KCNQ1OT1*_08	CpG_13&14	Mass overlap with unit 3
*KCNQ1OT1*_09	CpG_15	
*KCNQ1OT1*_10	CpG_16	
*KCNQ1OT1*_11	CpG_17&18	
*KCNQ1OT1*_12	CpG_19	
*KCNQ1OT1*_13	CpG_20	
*KCNQ1OT1*_14	CpG_21	
*KCNQ1OT1*_15	CpG_22	Mass overlap with units 5 and 19
*KCNQ1OT1*_16	CpG_23	rs7940500
*KCNQ1OT1*_17	CpG_24	rs379976
*KCNQ1OT1*_18	CpG_25	
*KCNQ1OT1*_19	CpG_26&27	Mass overlap with units 5 and 15

Table S7. (Continued C): All CpG-units of the 16 loci

Locus_CpG-unit*	CpG-sites*	Reason for removal prior to quality control
*MEG3*_01	CpG_1	
*MEG3*_02	CpG_2	
*MEG3*_03	CpG_3	
*MEG3*_04	CpG_4	
*MEG3*_05	CpG_5	Mass overlap with unit 6
*MEG3*_06	CpG_6	Mass overlap with unit 5
*MEG3*_07	CpG_7	
*MEG3*_08	CpG_8&9	
*MEG3*_09	CpG_10&11	
*MEG3*_10	CpG_12-14	High Mass
*FTO*_01	CpG_1	Mass overlap with Unit 5
*FTO*_02	CpG_2&3	
*FTO*_03	CpG_4	Low mass
*FTO*_04	CpG_5	
*FTO*_05	CpG_6	Mass overlap with Unit 1
*FTO*_06	CpG_7	
*FTO*_07	CpG_8&9	
*FTO*_08	CpG_10&11	
*FTO*_09	CpG_12	Mass overlap with Unit 17
*FTO*_10	CpG_13	
*FTO*_11	CpG_14	
*FTO*_12	CpG_15	
*FTO*_13	CpG_16	
*FTO*_14	CpG_17	
*FTO*_15	CpG_18	
*FTO*_16	CpG_19	
*FTO*_17	CpG_20	Mass overlap with Unit 9
*APOCI*_01	CpG_1	
*APOCI*_02	CpG_2	
*APOCI*_03	CpG_3	
*APOCI*_04	CpG_4	
*APOCI*_05	CpG_5&6	rs402204
*APOCI*_06	CpG_7-9	High Mass and rs5111
*APOCI*_07	CpG_10	
*APOCI*_08	CpG_11	
*GNASAS*_01	CpG_1&2	
*GNASAS*_02	CpG_3&4	
*GNASAS*_03	CpG_5	
*GNASAS*_04	CpG_6	
*GNASAS*_05	CpG_7	
*GNASAS*_06	CpG_8&9	
*GNASAS*_07	CpG_10-12	
*GNASAS*_08	CpG_13&14	
*GNASAS*_09	CpG_15	
*GNASAS*_10	CpG_16	rs45596642
*GNASAS*_11	CpG_17-19	

Table S7. (Continued D): All CpG-units of the 16 loci

Locus_CpG-unit*	CpG-sites*	Reason for removal prior to quality control
*GNAS A/B*_01	CpG_1	
*GNAS A/B*_02	CpG_2	Mass overlap with unit 4
*GNAS A/B*_03	CpG_3&4	
*GNAS A/B*_04	CpG_5&6	Mass overlap with unit 2
*GNAS A/B*_05	CpG_7	
*GNAS A/B*_06	CpG_8	
*GNAS A/B*_07	CpG_9&10	
*GNAS A/B*_08	CpG_11	Low mass
*GNAS A/B*_09	CpG_12	
*GNAS A/B*_10	CpG_13-15	
*GNAS A/B*_11	CpG_16-19	
Totals		
# of amplicons		16
total # units		191
# units outside detection range		9
# units with equal or overlapping mass		36
# units with potential SNP		12
Total # CpG-units removed		**87**

* CpG-unit and CpG-site numbers are counted from the forward primer onward

Table S8. Primers used in Chapter 6

Locus	Location NCBI36/hg18	Strand	Amplicon (bp)	Forward primer[1] 5'-3'	Reverse primer[2] 5'-3'
*H19*DMR	chr 11: 1975948-1976360	-	413	GGGTTTGGGAGAGTTTGTGAGGT	ATACCTACTACTCCCTACCTACCAAC
*IGF2*DMR2	chr11: 2111300-2111791	+	492	GGAAAGGGGTTTAGGATTTTTAT	AACCACTCCCATTATAAACCTTTAAT
*IGF2*DMR2 CTCF	chr11 2112023-2112312	+	290	TAGTAATGTTTAGTTGGAAGGGGAA	ACTACTTAACTCTAAAAACCCCTACCC
*IGF2*AS	chr11: 2117482-2117948	+	467	TTTTAGAGAATTTAGGGGTTTTATTT	CCATACAAATAAAATTTAAACTATATTTCC
*IGF2*AS CTCF	chr11: 2118126-2118422	-	297	GGTTGGAGGGTTTTAAAGTGG	AAAAAAACACTATAATTTTTACCAAATCAA
*IGF2*DMR0 upstr.	chr11: 2125961-2126065	-	105	GTTGTGTGTTTAGTGGTTTTTGTTG	AAAAAATTTACCTAAAAAAAACTTCCC
*IGF2*DMR	chr11: 2126035-2126372	-	338	TGGATAGGAGATTGAGGAGAAA	AAACCCCAACAAAAACCACT
*IGF2*DMR0 downstr.	chr11: 2127117-2127220	-	104	GATGAGGTTTTTTTATTTGTAGGGG	AAAACCAAAATCCTAACAACTACCC
INSIGF	chr11:2138912-2139216	-	305	GTTTTGAGGAAGAGGTGTTGA	ACCTAAAATCCAACCACCCTAA
LINES-1	X58075: 335–767	-	432	GTGTGAGGTGTTAGTGTGTTTTGTT	ATATCCCACACCTAACTCAAAAAAT
Additional	For Sequenom Epityper			AGGAAGAGAG + primer	CAGTAATACGACTCACTATAGGGAGAAGGCT + primer

1. Forward primer, an additional tag is added for Sequenom Epityper PCRs which is denoted below
2. Reverse primer, an additional tag is added for Sequenom Epityper PCRs which is denoted below

PCR was performed with the following cycling protocol: 15 minutes at 95°C, 4 rounds of 20 seconds at 95°C, 30 seconds at 65°C, 1 minute at 72°C; followed by 40 rounds, 20 seconds at 95°C, 30 seconds at 58°C and 1 minute at 72°C; ending with 3 minutes at 72°C.

Table S9. Details of CpG units measured in Chapter 6

Amplicon	CpGsite	Reason for exclusion	Success rate	Mean methylation[2]	SD[3]	Exp-Unexp[4]	P diff[5]
H19 DMR	CpG1	Mass-overlap with fragment CpG 16	95.0				
	CpG2		95.0	28.7	5.8	-0.2	0.91
	CpG3.4.5	Mass-overlap with fragment CpG 11	95.0				
	CpG6	low success rate	65.8				
	CpG7		89.2	32.2	3.2	-0.7	0.23
	CpG8		94.2	32.1	2.0	0.0	0.99
	CpG9.10		94.2	26.8	2.6	0.0	0.99
	CpG11	Mass-overlap with fragment CpG 3.4.5	95.8				
	CpG12	rs12292822	95.0	26.6	2.3	0.0	0.87
	CpG13		95.0	29.0	3.1	-0.2	0.61
	CpG14.15	rs12292818	95.0	31.5	2.5	0.0	0.98
	CpG16	Mass-overlap with fragment CpG 1	95.0				
	CpG17	rs35592994	94.2	29.0	2.4	-0.3	0.56
	CpG18.19		94.2	30.4	3.2	-0.2	0.71
	CpG20		95.0	30.0	3.0	-0.2	0.73
	CpG21	Low success rate	4.2				
	CpG22		86.7	34.0	8.0	-1.7	0.12
	CpG23	Low success rate	3.3				
	CpG24	Low success rate	65.0				
	CpG25		95.0	31.9	3.4	-0.7	0.21

Table S9. (Continued A) Details of CpG units measured in Chapter 6

Amplicon	CpGsite	Reason for exclusion[1]	Success rate	Mean methylation[2]	SD[3]	Exp-Unexp[4]	P diff[5]
IGF2 DMR2	CpG1	Low mass	0.0				
S.L.	CpG2.3	Mass-overlap with fragment CpG 9	65.8				
(DMR2)	CpG4		90.8	48.0	8.9	2.2	0.17
	CpG5.6		95.8	37.2	5.1	-0.3	0.88
	CpG7		94.2	49.8	9.3	0.0	0.96
	CpG8		96.7	56.7	8.2	0.5	0.74
	CpG9	Mass-overlap with fragment CpG 2.3	67.5				
	CpG10	Low mass	0.0				
	CpG11.12		95.0	47.7	5.6	-0.2	0.91
	CpG13		93.3	44.3	7.4	-0.3	0.83
	CpG14.15		96.7	55.5	8.2	0.7	0.68
	CpG16	Mass-overlap with fragment CpG 18 and 21	95.8				
	CpG17	Low success rate	3.3				
	CpG18	Mass-overlap with fragment CpG 16 and 21	95.8				
	CpG19	Low mass	0.0				
	CpG20		95.8	59.4	6.7	-1.0	0.43
	CpG21	Mass-overlap with fragment CpG 16 and 18	95.8				
	CpG22	High mass	0.0				
IGF2 DMR2 CTCF	CpG1		95.0	53.9	3.8	-1.7	0.040
(DMR2)	CpG2	Low success rate	70.8				
	CpG3		95.0	35.0	4.0	-1.4	0.056
	CpG4		95.0	63.5	2.4	-1.1	0.045

1. CpG containing fragments (e.g. 'CpG units'): excluded were fragments containing possible SNPs in CEU (by HAPMAP or 1000genomes), a measurement success rate below <75% or (partial) overlap with other units.
2. Mean methylation in %, based on the raw data.
3. the variation (in %) in the controls
4. The average within pair difference from a Linear Mixed Model, corrected for age and bisulfite batch.
5. The P value belonging to the within pair difference.

Table S9. (Continued B) Details of CpG units measured in Chapter 6

Amplicon	CpGsite	Reason for exclusion[1]	Success rate	Mean methylation[2]	SD[3]	Exp-Unexp[4]	P diff[5]
*IGF2*AS	CpG1.2.3.4	High mass	0.0				
(DMR1)	CpG5		97.5	8.3	2.8	0.7	0.16
	CpG6		85.0	16.4	3.0	3.1	0.070
	CpG7		96.7	5.6	2.6	-0.4	0.43
	CpG8.9.10		96.7	5.4	0.9	-0.2	0.40
	CpG11.12	High mass	0.0				
	CpG13.14.15.16	High mass	0.0				
	CpG17		95.0	5.2	1.5	0.3	0.25
	CpG18	Mass-overlap with fragment CpG 28	97.5				
	CpG19.20.21		97.5	2.5			
	CpG22	Low success rate	59.2				
	CpG23.24.25.26.27	High mass	0.0				
	CpG28	Mass-overlap with fragment CpG 18	97.5				
	CpG29		95.8	13.0	2.4	0.3	0.63
	CpG30		95.0	7.3	2.9	0.1	0.96
	CpG31.32.33		95.0	14.9	2.9	0.7	0.24
	CpG34.35.36.37		97.5	8.5	1.0	0.1	0.55
	CpG38	Low mass	0.0				
	CpG39.40		81.7	11.7	2.2	0.5	0.43
	CpG41		97.5	3.1	0.8	0.5	0.0030

Table S9. (Continued C) Details of CpG units measured in Chapter 6

Amplicon	CpGsite	Reason for exclusion[1]	Success rate	Mean methylation[2]	SD[3]	Exp-Unexp[4]	P diff[5]
*IGF2*AS CTCF	CpG1		98.3	2.5	1.7	0.2	0.61
(DMR1)	CpG2		98.3	0.6	0.6	0.0	0.99
	CpG3.4		92.5	14.9	2.9	0.5	0.42
	CpG5.6.7.8		98.3	4.1	0.9	0.1	0.37
	CpG9.10		95.8	5.1	0.8	0.3	0.11
	CpG11.12		97.5	5.7	1.7	0.6	0.11
	CpG13.14.15.16		97.5	6.9	2.5	0.8	0.063
	CpG17.18.19		91.7	6.4	3.4	0.9	0.20
	CpG20		95.8	1.6	0.8	0.4	0.0054
	CpG21	Low mass	0.0				
	CpG22		96.7	2.0	0.7	0.3	0.019
	CpG23.24		97.5	1.1	0.6	0.0	0.71
	CpG25.26.27	High mass	0.0				
	CpG28.29.30.31	High mass	0.0				
	CpG32		98.3	4.5	1.9	0.5	0.33

Table S9. (Continued D) Details of CpG units measured in Chapter 6

Amplicon	CpGsite	Reason for exclusion[1]	Success rate	Mean methylation[2]	SD[3]	Exp-Unexp[4]	P diff[5]
IGF2 DMR0 upstream	CpG1		94.2	46.6	5.4	-2.9	0.0056
(DMR0)	CpG2		92.5	51.6	7.3	-3.1	0.027
	CpG3		95.0	40.1	5.2	-2.7	0.0056
	CpG4		95.8	50.4	4.0	-1.4	0.049
	CpG5		91.7	32.6	4.3	-0.5	0.57
*IGF2*DMR0	CpG1	Low succes rate	56.7				
downstream	CpG2	Low succes rate	67.5				
(DMR0)	CpG3	Mass-overlap with fragment CpG 4	8.3				
	CpG4	Mass-overlap with fragment CpG 3	8.3				
	CpG5.6	Mass-overlap with fragment CpG 10	84.2				
	CpG7	Low succes rate	44.2				
	CpG8		77.5	82.7	5.7	-2.3	0.05
	CpG9		75.8	60.6	5.8	1.0	0.35
	CpG10	Mass-overlap with fragment CpG 5.6	82.5				
	CpG11	Low succes rate	25.0				
	CpG12.13	rs11601832, but not present in CEU	82.5	70.5	4.9	-3.6	8.5E-4
LINES-1	CpG_1		98.3	64.2	2.6	-1.8	5.8E-4
	CpG_2		98.3	60.0	1.4	-0.8	0.004
	CpG_3		98.3	71.6	2.1	-0.3	0.36
	CpG_4		96.7	36.9	4.6	-0.4	0.28
	CpG_5		97.5	35.3	1.3	-0.1	0.54
	CpG_6.7		98.3	69.4	2.0	0.1	0.13
	CpG_8.9		98.3	68.7	1.9	0.1	0.12
	CpG_10	Low mass					
	CpG_11.12		98.3	83.6	2.5	-0.4	0.31

Table S10. The genotyping results for the H19 LD block

SNP	Source[1]	Success rate[2]	included?[3]	MAF(obs.)	MAF CEU		HW Pval[4]	associations (Pubmed [uid]: type)
rs217727	both	100	YES	0.204	0.15	G:A	0.16	15885138: birthweight and newborn IGF2 levels
rs2839701	tagging	no design poss.						
rs2067051	both	94.2	No, below <95%	0.5	0.482	C:C	0.89	20639793: association with birth weight
rs2251375	both	98.3	YES	0.297	0.292	C:A	0.74	20639793: association with birth weight
rs10732516	candidate	98.3	covered by rs4929983	0.496	NA, 0.44 in Brazilians	T:C	1	In core binding motif 6th CTCF ICR
rs11042170	tagging	no design poss.						
rs2735971	tagging	no design poss.						
rs12417375	tagging	no design poss.						
rs4929983	tagging	100	YES	0.488	0.397	T:C	0.74	
rs4929984	both	98.3	covered by rs4929983	0.47	0.486	A:C	0.86	20639793: association with birth weight
rs12292757	tagging	98.3	YES	0.212	0.125	G:A	0.47	

1. Several SNPs were chosen from the HAPMAP CEU panel as tagging SNPs for the region, also several candidate SNPs were added. Some were both candidate as HAPMAP tagging SNPs.
2. Success rate of the genotyping.
3. Several SNPs could not be measured, one SNP had a low success rate and two SNPs were in perfect LD (r^2>0.9) with another SNP in these individuals and thus not included in the final analysis.
4. The P value resulting from a test for Hardy-Weinberg disequilibrium, significant threshold is P<0.002 because of multiple testing.

Table S11. The genotyping results for the INSIGF LD blocks

SNP	Source[1]	Success rate[2]	included?[3]	MAF(obs.)	MAF CEU		HW Pval[4]	associations (Pubmed [uid]: type)
rs11042594	tagging	98.3	no, covered by rs680 (r^2=0.90)	0.305	0.341	G:A	0.046	-
rs10840356	tagging	0	below <95%	-	-	-	-	-
rs4341514	tagging	100	out of HW P<0.002	0.471	0.442	T:C	3.14E-21	-
rs7873	tagging	100	YES	0.079	0.102	A:G	0.92	-
rs3802971	tagging	100	YES	0.092	0.099	C:T	0.69	-
rs680	candidate	100	YES	0.321	0.33 CEU 1000genomes	G:A	0.076	11448941: BMI adult men, 17289909: muscle functioning, 19434426: birth length, 10573016: body weight in men
rs3213223	candidate	100	YES	0.238	0.199	C:T	1.0	-
rs3213221	both	100	YES	0.412	0.434	C:G	0.11	17289909 :loss of strength following excercise, 17339271:association with IGF2 DMR methylation
rs3213216	tagging	99.2	no variance	0	0.345	G:G	1.0	-
rs3741212	tagging		no design poss.	-	-	-	-	-
rs11603378	tagging	0	below <95%	-	-	-	-	-
rs1003483	both	100	YES	0.424	0.46	T:G	0.0399	19390492: no association CTCF6 and H19DMR methylation, marginal association with paternal haplotype and SGA and placental growth, 17339271: IGF2 DMR methylation
rs1003484	candidate	100	covered by rs2239681 (r^2=1.0)	0.3	CEU 0.25 1000genomes	G:A	0.0035	17339271:methylation IGF2DMR
rs2239681	tagging	100	YES	0.3	0.27	G:A	0.0035	-
rs3741211	both	100	YES	0.374	0.389	A:G	0.0063	19546867: association IGF1BP1 levels, 21078522: endometrial cancer risk, 11448941: adult BMI, 17488802: adult height; 19390492: no association with CTCF6 and H19DMR methylation, marginal association with paternal haplotype transmission and SGA and placental growth

Table S11. (Continued A) The genotyping results for the INSIGF LD blocks

SNP	Source[1]	Success rate[2]	included?[3]	MAF(obs.)	MAF CEU		HW Pval[4]	associations (Pubmed [uid]: type)
rs3741209	candidate	100	covered by rs3741211 (r^2=1.0)	0.379	0.375	G:C	0.0043	18955703: abolishes CpG site in IGF2 DMR
rs3741206	tagging	no design poss.	-	-	-	-	-	
rs4320932	tagging	100	out of HW P<0.002	0.238	0.204	T:C	9.00E-04	-
rs10840442	tagging	no design poss.	-	-	-	-	-	
rs7924316	both	100	YES	0.438	0.465	T:G	0.32	17289909: strength loss following excercise
rs10840447	tagging	100	YES	0.392	0.376	C:T	0.017	-
rs3842756	tagging	100	YES	0.292	0.243	G:A	0.53	12610512: prostate cancer risk
rs689	candidate	100	YES	0.367	0.242	A:T	0.85	19434426: postnatal growth, 16608900: BMI in children, 17667841: paternal transmission associates with newborn IGF2 levels, 17700581: association with SGA risk, 15047631: head circumference at birth, newborn IGF2 levels, 10573016: body weight in men, 11101842: T2D, 11528401: paternal transmission with child BMI and insulin secretion, 9590300: birth size
rs3842738	candidate	100	no variance	0	0.0 CEU 1000genome	C:C	1	17667841: paternal haplotype transmission associates with newborn IGF2 levels, 17700581: paternal haplotype transmission associates with SGA risk

Table S12. Famine Associations corrected for genetic variation

Locus	Exp. –Unexp. (%)[2]	P
INSIGF	*-1.2*	*0.027*
IGF2 DMR0	*-1.9*	*$6.8x10^{-6}$*
IGF2 DMR1	*0.3*	*0.028*
IGF2 DMR2		
IGF2 DMR2 S.L.	0.4	0.78
IGF2 DMR2 CTCF	No	SNPs
H19 DMR	*No*	*SNPs*

Table S13. SNP associations with and without famine exposure correction

Association between	with famine exposure		without famine exposure	
***DMR* – SNP**	**beta**	**P**	**beta**	**P**
IGF2 DMR0-rs2239681	-1.3	$1.1x10^{-3}$	-1.4	$9.9x10^{-4}$
INSIGF-rs3842756	-2.0	$8.2x10^{-6}$	-2.1	$1.4x10^{-5}$
INSIGF-rs689	-2.3	$7.4x10^{-8}$	-2.4	$4.0x10^{-8}$

Supplement II

Table S1. Population and sequencing characteristics

Variable	Quantity
Individuals sequenced	48
Same-sex sibling controls	24
Age (SD)	58.3y (2.1)
Percentage of males	50%
Number of pre-war born sibling	12
Male pre-war born siblings	6
Median quality score reads (SD)	35.3 (2.0)
High quality reads million (SD)	25.6 (7.3)
Reads mapped uniquely (SD)	74.1%(10.7)
Bisulfite conversion (SD)	98.9%(0.7)

Table S2. Data filtering steps

Filtering steps	Total CpGs	CpGs matching
Total unique CpGs	3.174.757	
Random chromosome		3.195
Median Coverage <=5		1.296.450
Median Coverage >200		2.935
Median methylation = 0%		1.400.875
Median methylation = 100%		252.439
Total included unique CpGs*	1.206.149	

*Considerable numbers of CpG dinucleotides match multiple filtering criteria, resulting in a final number of included CpG dinucleotides higher than the subtraction of the 'CpGs matching' column form the total unique CpGs.

Figure S1. Histograms of the median coverage and success rate per CpG dinucleotide
Histogram of the median coverage (e.g. sequencing depth) over the 48 individuals of the CpG dinucleotides included in the analyses (N=1.206.149).
Histogram of the success rate for each CpG CpG dinucleotides included in the analyses for the 48 individuals.

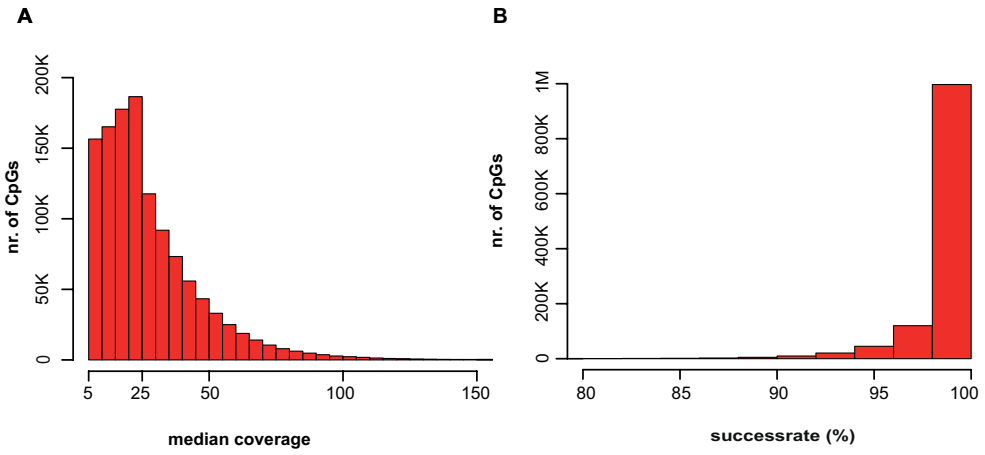


Figure S2. Density plot of the average methylation of the CpG dinucleotides

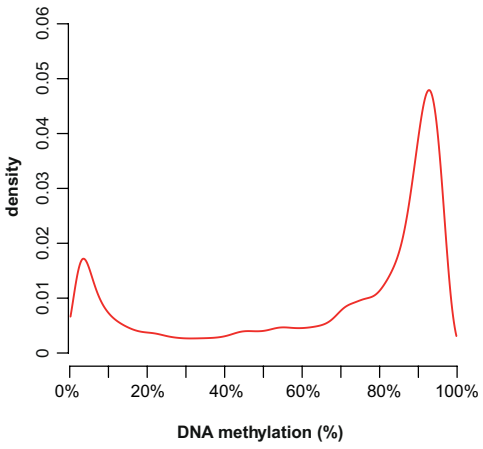
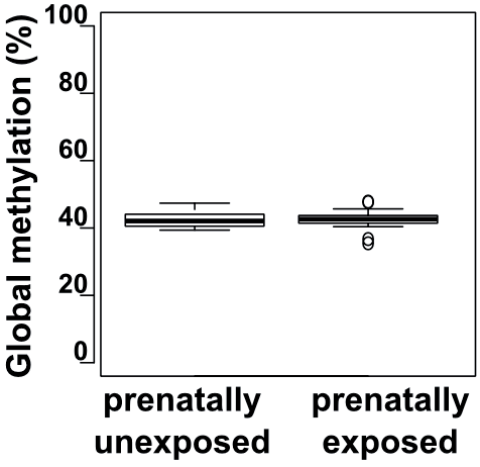


Figure S3. Global methylation



Boxplot depicting the global methylation levels of the prenatally exposed and unexposed siblings.

Table S3. Epityper associations with prenatal famine

Annotation		RRBS (N=48)				Epityper (N=48)	
Feature type[1]	Nearest Gene	Within pair diff. (%)	P[2]	P_{FDR}	correlation[3]	Within pair diff. (%)	P[4]
1	*DHRS4L2*	1.8	$3.0x10^{-9}$	$2.5x10^{-4}$	0.27	0.4	0.71
2	*EHD1*	-5.2	$5.7x10^{-9}$	$2.5x10^{-4}$	0.50	-4.0	$2.2x10^{-4}$
2	*DAPK2*	4.5	$1.1x10^{-8}$	$2.6x10^{-4}$	0.57	1.9	$6.4x10^{-3}$
2 & 3	*LOC554202*	6.5	$1.7x10^{-8}$	$2.6x10^{-4}$	0.52	3.5	$3.2x10^{-3}$
2	*LSM14B*	3.5	$1.7x10^{-8}$	$2.6x10^{-4}$	0.17	1.1	0.24
2	*ASS1*	-4.9	$4.8x10^{-8}$	$6.2x10^{-4}$	0.22	-0.5	0.42
2	*SMAD7*	4.2	$1.0x10^{-7}$	$1.1x10^{-3}$	0.78	4.7	$3.0x10^{-4}$
2 & 3	*CDH23*	4	$1.3x10^{-7}$	$1.1x10^{-3}$	0.73	3.9	$8.1x10^{-5}$
2	*MIR4315-1*	8.9	$2.1x10^{-7}$	$1.7x10^{-3}$	0.47	2.0	0.077
2	*ZIC1*	1.4	$3.2x10^{-7}$	$2.1x10^{-3}$	0.49	0.6	0.011
2	*INSR*	8.1	$3.9x10^{-6}$	0.01	0.72	3.9	0.018
2	*HOXD3*	3.2	$4.3x10^{-6}$	0.01	0.26	1.2	0.066
2	*RPTOR*	1.7	$9.1x10^{-6}$	0.02	0.63	8.6	$2.4x10^{-4}$
2 & 4	*RFTN1*	-2.3	$3.2x10^{-5}$	0.03	0.71	-1.9	$2.4x10^{-3}$
2	*STX1A*	3.7	$3.2x10^{-5}$	0.03	-0.17	1.1	0.17
2	*CPT1A*	4.5	$4.0x10^{-5}$	0.03	0.70	4.0	0.015
2	*SCARB1*	-10.5	$5.3x10^{-5}$	0.04	0.55	-4.9	$1.7x10^{-4}$
2	*KLF13*	-7.9	$6.1x10^{-5}$	0.04	0.79	-5.7	$7.9x10^{-4}$
2	*KLF6*	-6.2	$7.9x10^{-5}$	0.046	0.52	-2.6	$2.6x10^{-4}$

1 Type of genomic feature, 1 = non-CGI 'bonafide' promoter, 2 = Open chromatin, 3= enhancer, 4 = exon.

2 Two-sided P value resulting from a generalized linear mixed model where the analysis was weighted for sequencing depth.

3 The Pearson correlation between the average methylation of the RRBS region and the average of the CpG dinucleotides measured in the smaller region measured by Epityper.

4 The P value coming from a linear mixed model.

Table S4. Validation of the within pair differences in all 60 sibships with Epityper.

CDH23[1]	Meth. Exp(%)	SD(%)	Meth. Unexp(%)	SD(%)	Success rate (%)[2]	Diff.[3]	P[4]
CpG_1	24.8	5.6	22.4	6.1	100	2.3	0.0079
CpG_2	22.6	6.2	20.5	6.1	100	2.1	0.019
CpG_3	19.8	5.1	17.5	5.2	100	2.2	0.0048
CpG_4	22.5	6.1	20.1	6.3	100	2.3	0.0078
SMAD7							
CpG_1	23.7	6.9	20.6	7.1	95	3.0	0.0052
CpG_2	22.6	6.4	20.3	5.9	96.7	2.3	0.014
CpG_3	18.9	7.4	15.6	7.3	96.7	3.1	0.0077
CpG_4	14	5.8	10.6	5.9	94.2	3.3	0.0034
CpG_5.6	20.7	5.7	18.3	5.9	96.7	2.4	0.013
CpG_7	23	7.1	20.5	7.8	91.7	2.6	0.046
INSR							
CpG_2	33.5	5.6	30.6	5.8	95.8	2.9	0.0031
CpG_3	40	7.2	38	6	86.7	1.7	0.13
CpG_5	43.4	7.3	41.7	6.5	91.7	1.7	0.15
KLF13							
CpG_2	75.8	9.5	79.4	11.3	77.5	-3.7	0.07
CpG_4	65.4	7.8	68.4	8.6	99.2	-3.0	0.019
CpG_5	67.2	7.1	70.2	8.5	100	-2.9	0.016
CpG_6	64.2	7	67.4	9	100	-3.2	0.015
CpG_7	58.2	6.7	61.1	8.2	98.3	-2.8	0.020
CpG_9	62.3	7	65.5	8.5	98.3	-3.1	0.013
RFTN1							
CpG_1	89.6	3	90.3	3.1	100	-0.7	0.13
CpG_2	85	3.6	85.6	3.5	100	-0.7	0.19
CpG_3	81.5	3.7	82.8	3.5	100	-1.2	0.046
CpG_4	90.5	4.7	91.6	3.3	86.7	-0.9	0.23
CpG_5	82.3	2.8	83.2	2.7	100	-0.9	0.054
CpG_6.7	92.9	2.8	93.5	2.6	100	-0.6	0.21
CpG_8.9	84.1	3.8	85.2	5.1	100	-0.9	0.21
CpG_10	81.6	4.8	82.8	4.9	100	-1.2	0.12
CpG_12	81.8	4	83	4.2	100	-1.2	0.07
CPT1A							
CpG_2.3	41.3	10.9	39.1	11.2	99.2	2.4	0.09
CpG_5.6	58.1	8.9	56.3	8.9	99.2	1.9	0.11
CpG_8.9	76.5	3.8	75.4	4	99.2	1.1	0.05
CpG_10	49.1	8.8	46.9	8.7	99.2	2.4	0.032
CpG_12	34	8	31.7	8.8	99.2	2.4	0.037

1 The locus and individual CpG sites measured with Epityper after data filtering. The CpG dinucleotides are measured from the forward primer onward.

2 The success rate of the measurement in the 120 individuals

3 The average within pair differences in the 60 sib ships

4 The two-sided P-value resulting from a linear mixed model

Table S5. Epityper primers

NR[1]	Locus		Sequence
0	HUGO	Strand	FORWARD[3] (5'-3')
	coordinates(hg18)	T (°C)[2]	REVERSE[4] (5'-3')
1	DHRS4L2	-	GAGGATAGGGGTATTGGAGGTAAAG
	chr14:23,527,972-23,528,213	53	AAACCCAAACTTACTAATCTAATCCATA
2	EDH1	+	TTTGTTGTGAGGGAAATATAGTGATTG
	chr11:64,389,159-64,389,510	48	CCCTACCTTAATAAATACCAAAACCTAAA
3	DAPK2	-	TGATGATTTAATTTTGTGGGTTTGTGT
	chr15:62,062,839-62,063,043	53	AAATCCTAAAAACCCCACTCACAACT
4	LOC554202	+	AATTATTGGAGTGTAAAATTTTTTT
	chr9:21,595,548-21,595,759	48	CTATTTCCTAATAACCCATATTTAC
5	LSM14B	+	TTGTTTAGGAGGGTTTATTTTATGGTT
	chr20:60,141,366-60,141,715	53	ATAACAAACACTAACTCCCAAACTTCTAAC
6	ASS1	-	TGTTTTAGGGTGGGTATAGTTTAGGT
	chr9:132,344,966-132,345,523	46	ACCAAACCTCCTTAAAACTCTTCATA
7	SMAD7	+	TTGGGTTATATTTTATGTGTGGTGT
	chr18:44,677,194-44,677,679	53	CAACAACTAACTCTTTCCTACATCTAACT
8	CDH23	+	ATAGGGGAAGTTAGGTTTGGTAGAT
	chr10:73,227,653-73,227,914	53	ACTAAATAACCCTAATAAAACCCTC
9	miR4315-1	+	TTTTTTTTGTTTTTGATAGGGTTATG
	chr17:55,579,661-55,579,969	53	CCCAATATTCTAAATTCAAATCTTTACTCT
10	HIF1AN	-	AAAAGTGTTGATAGGGTTAGGAGAG
	chr10:102,319,419-102,319,893	Not working	TTTTATATATAACTTAATATAACAACATCA
11	ZIC1	+	TTTGGGTTTTTTGTTTTTAAGAGGT
	chr3:148,612,363-148,612,675	53	ACTTCCACCTAACTCCTAATTTCTAATTTT
C1	INSR	-	TTTTTAGGAGGTTTTTAGAGTTTTTAGATT
	chr19:7,110,140-7,110,418	44	CTAACCTCAAAATAATCCACCCCAC
C2	KLF13	-	AGGTAGGTATTTGTATAGAGGGGTTTA
	chr15:29,425,223-29,425,563	53	AAACTAACCACACCAAAACTTAATATACTT
C3	STX1A	+	GGTGAGGGGTTATAGATTAGGAGGT
	chr7:72,759,326-72,759,710	53	AAACAAACTAACCAAACAAACAAAACT
C4	RPTOR	-	GTTGGATGAGTAGGTTTTGGATGG
	chr17:76,479,050-76,479,293	46	CAATTCATAAAACAAAAAAAATTTAAAATA
C5	CPT1A	-	TTTAGGATATGGGTAAGTTTTGTTTTATAT
	chr11:68,286,598-68,286,810	48	AAATAATAACCTCCAAAAACCTTTTAAAAA
C6	RFTN1	+	TATGTATTTTTAAGGGGTTGTTTTT
	chr3:16,394,247-16,394,578	48	TAATATTATACCTCAATACCATTCCTCTAT
C7	HOXD3	+	TTTGGTGGTTTAATTTTGGTTAATT
	chr2:176,735,594-176,735,816	48	ATAAAAACATCCCCTCAAAAAAAA
C8	KLF6	-	AATAGTTTGAATTTAGATGTTAGTAG
	chr10:3,813,737-3,814,216	46	CCAAAACTAATACAATACAATAAC
C9	SCARB1	+	GGTGGTTAGGGTTAGTAAGAGAAGTA
	chr12:123,789,477-123,789,580	53	CCTATAACTCAAACTTCAAAAAAAAC

1 Primer pair number corresponding to the lowest p-value in RRBS (nr1-11; no reliable PCR was possible for nr10, chr10:102319110-102321355 [*HIF1AN*]) , several regions were chosen (C1-C9). Sequence of the forward PCR primer: for epiTYPER a tag with the following sequence is added 5': 5'-AGGAAGAGAG-sequence.

2 The annealing temperature in the PCR program: 15 min at 95°c, 15 minutes at 95°C, 4 rounds of 20 seconds at 95°C, 30 seconds at 65°C, 1 minute at 72°C; followed by 40 rounds, 20 seconds at 95°C, 30 seconds at ***Ann.T*** and 1 minute at 72°C; ending with 3 minute s at 72°C. The sequence of the reverse primer, for epiTYPER a tag with the following sequence is added 5': 5'-CAGTAATACGACTCACTATAGGGAGAAGGCT-sequence

Table S6. Number of CpG sites measured, and overlap between RRBS and Epityper measurement

Nearest Gene	location RRBS[1]	Location Epityper[2]	Nr CpGs RRBS	Nr CpGs Epityper	Nr CpGs overlapping
DHRS4L2	chr14:23,526,866-23,528,866	chr14:23,527,972-23,528,213	5	8	5
EHD1	chr11:64,374,355-64,390,875	chr11:64,389,159-64,389,510	26	9	8
DAPK2	chr15:62,060,557-62,063,275	chr15:62,062,839-62,063,043	7	12	2
LOC554202	chr9:21,594,650-21,595,650	chr9:21,595,548-21,595,759	3	3	2
LSM14B	chr20:60,141,215-60,146,975	chr20:60,141,366-60,141,715	7	6	3
ASS1	chr9:132,344,794-132,346,579	chr9:132,344,966-132,345,523	12	14	3
SMAD7	chr18:44,676,775-44,678,655	chr18:44,677,194-44,677,679	7	7	4
CDH23	chr10:73,227,550-73,228,550	chr10:73,227,653-73,227,914	3	4	3
miR4315	chr17:55,579,535-55,580,790	chr17:55,579,661-55,579,969	3	6	3
ZIC1	chr3:148,611,075-148,612,495	chr3:148,612,363-148,612,675	21	14	5
INSR	chr19:7,110,011-7,111,334	chr19:7,110,140-7,110,418	2	3	2
KLF13	chr15:29,423,875-29,427,520	chr15:29,425,223-29,425,563	8	6	6
STX1A	chr7:72,758,075-72,760,255	chr7:72,759,326-72,759,710	7	8	4
RPTOR	chr17:76,471,363-76,483,344	chr17:76,479,050-76,479,293	45	6	3
CPT1A	chr11:68,285,186-68,289,024	chr11:68,286,598-68,286,810	15	8	7
RFTN	chr3:16,394,228-16,394,613	chr3:16,394,247-16,394,578	10	11	5
HOXD3	chr2:176,734,540-176,735,745	chr2:176,735,594-176,735,816	7	6	2
KLF6	chr10:3,810,815-3,816,315	chr10:3,813,737-3,814,216	20	18	10
SCARB1	chr12:123,788,495-123,792,067	chr12:123,789,477-123,789,580	6	6	4

1 Genomic location of the genomic feature

2 Genomic location of the Epityper PCR amplicon

Table S7. Neutrophil variation and DNA methylation at the P-DMRs

P-DMR	Beta[1] (%methylation/ %neutrophils)	P[2]	SD change required to explain famine association[3]
SMAD7	-0.19	0.17	-5.1
CDH23	-0.35	$3.1x10^{-3}$	-1.9
INSR	0.02	0.91	31.4
CPT1A	-0.51	0.11	-1.2
KLF13	0.33	0.08	-2.8
RFTN1	0.00	0.99	∞

1 The effect size of the association of the percentage of neutrophils in blood with methylation in 44 unrelated individuals from the control population in the Leiden Longevity Study.

2 Two-sided p-value for the association between DNA methylation and the percentage of neutrophils in blood.

3 The size of the change in neutrophil percentage in blood, expressed in standard deviations (SD), required to explain in full the association between DNA methylation and prenatal famine exposure (3.3% change in neutrophils is 1 SD change).

4 Only CDH23 is affected by blood cell heterogeneity, but the change in blood cell composition to explain the observed famine association is so large (~2SD) that this is a highly unlikely explanation.