Title: Paternal grandparental exposure to crop failure or surfeit during a childhood slow growth period and epigenetic marks on third generation´s growth-, glucoregulatory and stress genes

Authors: Lars Olov Bygren[1,2]*, Patrick Müller[1], David Brodin[1], Gunnar.Kaati[1], Jan-Åke Gustafsson[1], John G. Kral[3]

Affiliations:

[1] Department of Biosciences and Nutriotion, Karolinska Institutet, SE_14183 Huddinge, Sweden.

[2] Department of Community Medicine and Rehabilitation, Umeå University, SE_90187, Umeå, Sweden.

[3] SUNY Downstate Medical Center, 450 Clarkson Avenue, Box 40, Brooklyn, NY 11203, USA.

* Correspondence to: lars.olov.bygren@ki.se

Abstract: This latest in our series of papers describes transgenerational methylation related to mid-childhood food availability in 19th century Överkalix, Sweden. Failed vs.bountiful crops differentially influenced methylation in grandchildren of exposed paternal grandparents. In this case study of 8 tracked cases with differential ancestral exposure, adult progeny differed in methylated CpGs in three Amigo site gene pathways: "insulin processing", "adipose tissue growth" and "hypothalamic development", reflecting famine, excess food and reaction to food-insecurity stress. This is the first demonstration of transgenerational epigenetic inheritance in humans following grandparental childhood exposure, an early developmental origin of adult disease.

Non-technical summary: Paternal grandparent food supply preceding their prepubertal growth spurt induced epigenetic marks in three gene pathways reflecting famine, excess food and reaction to food-insecurity stress.Our epidemiological findings of adverse transgenerational effects of ancestral overnutrition and conversely, beneficial effects of famine during SGP[1] prompted us to study individuals whose paternal grandparents were randomly exposed, during their Slow Growth Period (SGP) - a sensitive period preceding the pre-pubertal growth spurt - to failed or bumper crops potentially engendering epigenetic marks on mechanistic molecular gene pathways.

For detailed analyses of specific gene ontology (GO) pathways, we tested the three posited structural classes of genes associated with each of 40 preselected GO terms (http://amigo.geneontology.org) for differentially methylated CpG positions (DMRs) between the experimental groups. We selected GO pathways with average differences in methylation between groups >0.1 and p-values <0.05 and performed detailed analyses of those differing > 5%. The posited pathways comprised glucoregulatory genes that were seen as immediate sensors of decreased glycogen stores, whereas lipids and ketones were posited to reflect the duration and magnitude of the fast or responses to overnutrition. Individual reactions to food insecurity were assumed to have activated the HPA-axis. We focused on two putative response lines in direct descendants from a) grandfather to son with grandson and b) paternal grandmother to son with granddaughter, adjusted for crop size during famine or surfeit. We found 39 DMRs in grandchildren after grandparental differential exposure to food availability in GO-pathways associated with famine, excess food and reactions to environmental stressors. Nine of them were large, three GOs exhibiting remarkable DMRs in grandchildren (Table 1).

Results

Exploiting a natural experiment, we studied the paternal lineage transgenerational responses in grandsons and granddaughters and assessed their DNA-methylations. We found 39 DMRs amoung 40 posited pathways in Islands, North and South shores, North and South shelves related to famine, excess and stress, 9 pathways with DMRs above 5%, and 3 above 14%.

Paternal grandparents´exposure vs. methylated genes in the adult grandchild methylation.

The main metabolic pathways implicated in grandpaternal SGP exposure to food-insecurity stress transferred to grandsons were “appetite”, “insulin processing”, “ketone body catabolic process” and “hypothalamic development”. Differential grandmaternal exposure to feast or famine were associated with marks in “adipogenesis”, “insulin-like growth factor reception (IGF-R) binding”, “ketone body catabolic process” and “adrenal gland development” (Table 1).

Three sentinel metabolic pathways implying transgenerational epigenetic inheritance.

GO:0030070 ”indulin processing” influences insulin from proteolysis of the precursor preproinsulin via C-peptide.This is the first signal sequence elevated from preproinsulin. Proinsulin is then cleaved to release the C peptide. This pathway had a DMR in the male line found in the Southern Shore involving 2 genes: PCSK2, proprotein convertas (also called NE2 HUMAN) and CPE Carboxypeptidase E (also called CPBE HUMAN). The pathway is related to GO 1901142 “insulin metabolic process”.

GO:1904179 , “adipose development” includes gene NR1H4 (also called BAR;FXR;HRR1; RIP14) the nuclear hormone receptor pthr 24082, “bile acid receptor”. Other genes are SIRT1 (also called SIR2L1) and chromatin regulatory protein sir2. phtr 11085

GO 21854 “hypothalamus development”, involves GSX1(GShomebox 1), PTX2 (Pituitary homebox 2), UBB (Polyubiquitin-B) and CRH (Corticoliberin)

Parental age data and selection

Average age at birth of first child was 32 years, 31 years and 27 years in the three generations when grandparents had been exposed to famine and 34 years, 27 years and 27 years when grandparents had been exposed to feast. All seven index persons were alive at the age of 74,

one died at the age of 63. For grandparents and parents exposed to famine during SGP, the average age of death was 61 years and 78 years whereas when correspondingly exposed to feast it was 87 years and 71 years (Table 2).

Discussion

We present a case series of epigenomes of 8 adult grandchildren with archival data exhibiting their paternal grand-paternal or -maternal exposure to crop failure or large crops during their SGP linked to differential methylation of genes associated with energy balance and hypothalamic development. The salient novel finding is that grandparental childhood exposure was reflected in grandchild profiles of CpG methylation of key genes.

In earlier papers we described a unique cohort of descendants of grandparents exposed randomly to these weather-dependent variations in food availability. Data emanated from detailed archival records enabling correlations between demographic and crop yield statistics[1,2,3,4,5,6] . Our SGP findings have been confirmed in Germany for food deprivation[7] and in Sweden for parental loss[8]. Our current paper adds mechanistic information based on blood sampling of eight 75-year-old grandchildren enabling determination of candidate epigenetic markers reflecting differential methylation in gene pathways posited to be affected by exposure of their grandparents to crop failure versus bountiful harvest during the grandparental childhood SGP preceding the prepubertal peak stature growth.

We propose that known adaptations to energy deficiency or excess in cells, organs or organisms are expressed transgenerationally appearing as epigenetic marks in humans. Famine might induce marks around promoters in at least two types of pathways: one general or nutrient-specific, related to diminished substrate stores of glycogen, lipid or protein, and/or excess, nutritoxic exposure , the other to environmental stresses of famine through activation of the hypothalamo-pituitary-adrenal (HPA) axis. Based on our epidemiological findings of transgenerational inheritance we posited that changes in DNA methylation levels of cytosine and guanine (CpG probes) differentially methylated to form 5-methylcytosine affecting genes in one generation can explain transgenerational epigenetic inheritance in the third generation[9].

A 2015 review concluded that evidence suggesting that acquired epigenetic marks are passed to the next generation was limited and many other have noted the absence of any mechanism by which gene-regulatory information is transferred from somatic cells to germ cells in the

study of transgenerational epigenetic inheritance, reviewed in Nagy and Turecki[10]. Our epidemiological and epigenetic results demonstrate sex differences, e.g. in development of preproinsulin as has been seen in human pancreatic islets[11]. The sex-bound epigenetic inheritance has been interpreted simply as epigenetic actions and disease manifestations being sex specific, as is the case with many conventional genetic variants[12]. The transcriptions are also not fully understood. Most confusion emanates from not knowing how exposed somatic cells can communicate their exposure to the germline which induces changes lasting for generations.

It is difficult to get an insight into epigenetic inheritance if one believes in a real rather than just a theoretical barrier between somatic and germline cells. One route overcoming the barrier might be extracellular nano-vesicles of neighboring to germ cells shed by somatic cells containing the material required for transcription able to bypass the barrier[13]. In the SGP the male primordial germ cells form active spermatozoa whereas human ovarian stem cells that can be modified in the SGP probably are required.

Presently the preponderance of evidence suggests transgenerational cumulative effects of exposures: diet, behavior, environmental chemicals, activity and the microbiome. For reviews see Sales et al.[9] and Vaiserman et al.[14].

We have focused on the paternal line to discern transgenerational epigenetic inheritance disentangled from maternal-fetal and maternal-infant influences. The pathways in the paternal line are probably similar in the maternal line. We found three interesting gene pathways influenced by paternal grandparents´exposure resulting in grandchildren´s DMRs, pathways induced by famine, overnutrition and food insecurity stress. When the paternal grandfather had been exposed to famine, the grandson exhibited DMRs of insulin processing. When the paternal grandmother had been exposed the granddaughter had DMRs of GOs related to the environmental stress such as “hypothalamus development”(influencing the HPA-axis) related to increased female susceptibility to stress, yet explaining the benefit to mental health recently found in a different setting[7].

A strength of this study is that it exploits a natural secular phenomenon viz. variable food security, the effects of which are documented in a homogeneous well documented 19th century population, allowing exceptional tracking of ancestors. Few founders colonized the area 500 years earlier whereby the area metaphorically became an island of native speakers of Swedish surrounded by Sami- and Finnish-speaking neighbors. Furthermore the index

persons, the grandchildren are dispersed all over Sweden[5] and have neither been exposed to famine nor, being born before WW2, experienced current food excesses during their childhood.

A second strength is the focus on two lines of inheritance that emerged in the epidemiological studies, the line from the paternal grandmother´s son and his daughter and the line from the paternal grandfather with son and grandson. Thirdly our *a priori* selection of gene ontologies, GOs, based on our earlier epidemiological findings of cardiovascular morbidity related to food security reduces the risk of spurious correlations. A confound of multiple testing could be suspected having 40 GOs but the ontologies were chosen for their evidence based responses to only three potential pathways: feast, famine and stress.

Our earlier published studies in the community and correlation of age at the birth of the first child, number of children and variance of survival over three generations ruled out that variable availability of food during SGP was caused selectively[2]. In human studies measures are very crude but the figures for the eight subjects do not demonstrate any biased selection owing to ancestors´random exposure during the SGP. The main weakness is having only 8 index cases causing low statistical power. Furthermore, these 8 were survivors aged 75 years during which their own exposures might have induced methylations confounding those inherited from the grandparents and parents. On the other hand epimutations analogous to mutation most often do not appear on the clinical horizon until late in life.

Confounding could have occurred through differences in the exposure to the feast and famine between index cases owing to social circumstances such as being member of a family with better food resources, or genes affecting sensitivity to undernutrition, or cellular heterogeneity differentially affecting allele-specific or other variables´ methylation[12] . However, the natural experiment exposed all families, showing little variation in poverty in the 19th century mitigating such confound, supported by our earlier research in the community[4] . The isolation of the community during the centuries up to the present might have resulted in an in-bred cohort and diminished assortative mating, both benefitting the epigenetic analysis.

Human transgenerational inheritance still awaits more studies in other settings. Two questions are pressing. How many generations can the epigenetic marks in humans persist? Our own recent unpublished studies in a different setting indicate that it persists at least for four generations. The effects and mechanisms of interaction between genetic and epigenetic inheritance also remain to be determined.

We describe in a small case study of 8 subjects the presence of epigenetic changes attributable to environmental influences potentially engendering several multifactorial diseases and conditions pontentially providing guidance for preventive and therapeutic interventions targeting methylation.

Methods

Classification of exposure to food availability in the environment

A tradition of using a scale for crop estimates was introduced for the years 1816-1849 in Statistics Sweden (Tabellverket) by the 19th century statistician Hellstenius[15]. It was a seven-degree scale running from total failure to good or bountiful crop and has been used since in demographic research. For the years 1865-1902, time of our 8 grandparents´ births, statistical tables of crop yields were the primary source, validated against food price price statistics[16], and qualitative reports from a gubernatorial office (Hushållningssällskapet) of crops and government aid. Using these sources, we created an ordinal seven-item "Hellstenius scale". Over the years the criteria for the items of the scale have changed apace with changes in the Swedish language. The current translation is as follows: Total failure (0), Sparse general growth (1), Weak or Small (2), Below average (3), Average or mediocre (4), Above average (5) and Good or bountiful (6) crop. We defined 0-2 as famine, 3-4 as mediocre, and 5-6 as excess food availability. We empirically used May 1 of the year following crop failure as the time of least food availability and November 1 of the year following good harvest and slaughtering of pigs and cattle as the time of greatest food availability (excess). Grandparent age on those dates was used to relate food availability to his or her SGP.

SGP, The slow growth period

Two periods in our first analysis were posited to differ in demand for food during famines and consequently to differ in transgenerational responses: the prepubertal growth peak and the precedimg slow growth period. These two periods were derived by superimposing the stature growth velocity curves of Prader et al[17] and the ages of 19th century pubarche described by Tanner[18]. The periods for ancestors in the 19th century were set at the ages 8-11 years for female ancestors and at 9-12 years of age for male ancestors. Excess food was not considered to be an issue in the 19th century Överkalix but to analyze potential transgenerational responses we were obliged to focus on the SGP were we indeed found mismatches reflecting transgenerational responses not only to famine but also to excess food availability. Thus, we

199 introduced the concept of the "slow growth period"[1], the causes and mechanisms of which we
200 discuss here in depth.

201 **Samples and pedigrees**

202 Eight subjects, alive at 75 years of age and consenting to blood sampling were selected from
203 the 1935 birth cohort in an original epidemiologic study of transgenerational responses to
204 variable availability of food during ancestors' SGP in Överkalix, Sweden, described in
205 Tinghög et al[5]. Two paternal grandmothers were exposed during SGP to famine (grade 0-2) in
206 1867 and 1900 respectively, and two were exposed to surfeit (grade 5-6) in 1871 and 1879.
207 Two paternal grandfathers were exposed to famine (grade 0-2) in1867 and 1877 respectively
208 and two were exposed to surfeit (grade 5-6) in 1887 and 1881.These subjects represented four
209 pairs of grandchildren of paternal grandparents exposed to feast or famine (Fig1).

210 Whole blood samples in EDTA were drawn at county health services units close to the
211 subjects´ residence, frozen at 70^{o} C and sent to the project center. Pedigrees were originally
212 tracked at parish offices in Sweden and Finland and presently in the National archive, the
213 Regional archive of north Sweden, the Research archive of Umeå University, the National tax
214 bureau and digitalized genealogical net-based registers.

215 **Metabolic pathways posited**

216 The gene ontologies chosen *a priori* described processes related to cardiovascular risks and
217 were related to hunger, excess food and reactions to environmental stress.
218 (http://amigo.geneontology.org). Posited processes during *famine* were "glucose homeostasis"
219 and "glucose transport", "ketone body biosynthesis and catabolic processes", "cellular ketone
220 body metabolic pathways", and "cholesterol homeostasis". In *hunger and excess* "insulin
221 processing and binding", insulin-like growth factor", "insulin receptor signaling and
222 "autophagy pathways" were posited. Individuals´ reactions to disasters were assumed to
223 influence hypothalamus and adrenal gland development. Availability and absorption of folate
224 stimulating methyl donors was of obvious interest. Other posited GOs include appetite,
225 ghrelin, leptin, calcium signaling, eating behavior, adipose and fat cell development, insulin
226 secretion, and lipids.

227 Altogether these added up to 40 pathways, many of them overlapping and covering three
228 kinds of pathways reflecting the three posited reactions.

229 **Genome-wide methylation**

230 Genomic DNA extraction from whole-blood (in one case from buffy coat) was carried out
231 using GeneCatcher Blood kit (Invitrogen, Carlsbad, CA, USA). 500 ng of genomic DNA was
232 bisulfite converted with EZ-96 DNA Methylation kit (Zymo Research, Irvine, CA, USA) and
233 genome wide DNA methylation analysis was carried out using the Infinium MethylationEPIC
234 Bead Chip (Illumina, San Diego, CA, USA). The Laboratory procedures were done according
235 to the manufacturers´ protocol.

236 The array was designed for genome wide methylation analysis with coverage across gene
237 regions with sites in the promoter region, 5UTR, enhancer and gene body, and interrogating
238 more than 850 000 methylation sites at single nucleotide resolution. This technique uses two
239 different probe types (Infinium 1 and 2) with different characteristics, thus requiring
240 normalization to reduce technical bias. For analysis, visualization and extraction of
241 methylation data GenomeStudio software version 2011.1 (Illumina Inc) was used.

242 Methylation levels (beta values) were estimated as the ratio of signal intensity of the
243 methylated alleles to the sum of methylated and unmethylated intensity signals. The beta-
244 values vary from 0 (no methylation) to 1 (100% methylation).

245 The chip covers CpG islands, shores, shelves and the promoted genes. GOs containing genes
246 and biological pathways were assigned from the literature, primarily from the Gene Ontology
247 project (geneontology.org) chosen from their known intra-generational relation to excess and
248 lack of food.

249 The assay protocol combined bisulfite conversion of genomic DNA and whole-genome
250 amplification with array-based capture and scoring of the CpG loci. Signal intensity was
251 measured by scanner to generate beta values, the degree of methylation at a locus. Allele-
252 specific single base extension of the probes incorporated a biotin nucleotide or a dinitrophenyl
253 labeled nucleotide. Signal amplification of the incorporated label further improved the overall
254 signal-to-noise ratio of the assay.

255 Human genomes punctuated by DNA sequences with high frequencies of CpG sites, >200
256 bases, with CpG content of 50% were termed “islands” (CGIs). Differential methylation
257 induced by environment visible in CGIs or in “northern- and southern-*shores*”, within 2 kilo
258 base-pairs (<2kbp) of the islands, and “northern- and southern *shelves*” <2 kbp from the

shores were studied in CGIs, shores and shelves in grandchildren of grandparents having experienced extremes of food availability during their childhood SGP.

Statistical analysis

The association between genes and GO terms was retrieved from Gene Ontology Association (UniProt-GOA) Database, http//www.ebi.ac.uk/GOA.

Number of genes associated with a GO term and represented with at least one probe on the EPIC chip and number of probes associated with the genes were analyzed among index cases, the grandchildren.

Differentially methylated female index cases whose paternal grandmothers were exposed to bumper harvests in their mid-childhood SGP and to crop failure respectively were recorded as well as male index cases whose paternal grandfather had the same extreme exposures.

The criteria for differential methylation was a p-value < 0.05 and an average difference in beta-value between the group averages >0.1. The share of differentially methylated probes was given as percent.

Tables were prepared after beta-mixture quantile normalization (BMIQ) correcting probe design bias. Tables were prepared with the BMIQ-normalised beta values and annotations. The analyses were carried out in Bioconductor-package ChAMP package default probe filtering based on detection p-value, minimum bead count and probes possibly confounded by SNPs and cross hybridisation (https//www.ncbi.nim.nih.gov/pubmed/24063430) left 813007 probes for further analysis.

References

1.Bygren LO, Kaati G, Edvinsson E. Longevity determined by paternal ancestors´nutrition during their slow growth period. Acta Biotheoretica 2001;49:53-59.

2.Kaati G, Bygren LO, Edvinsson S. Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period. Eur J Hum Genet 2002;10:682–8.

3.Pembrey ME, Bygren LO, Kaati G, Edvinsson S, Northstone K et al. Sex-specific, male-line transgenerational responses in humans. Eur J Hum Genet 2006;14:159-166

4.Kaati G, Bygren LO, Pembrey M, Sjostrom M. Transgenerational response to nutrition, early life circumstances and longevity. Eur J Hum Genet 2007;15:784–90.

5.Tinghög P, Carstensen J, Kaati G, Edvinsson S, Sjöström M, Bygren LO. Migration and mortality trajectories: a study of individuals born in therural community of Överkalix. Soc Sci Med. 2011;73:744-51.

6.Bygren LO, Tinghög P, Carstensen J, Edvinsson S, Kaati G, Pembrey ME, Sjöström M. Change in paternal grandmothers' early food supply influenced cardiovascular mortality of the female grandchildren. BMC Genet 2014;15:12.

7.Van den Berg GJ, Pinger PR. Transgenerational effects of childhood conditions on third generation health and educational outcomes. EconHumBiol 2016;23:103-120.

8.Vågerö D, Rajaleid K. Does childhood trauma influence offspring´s birth characteristics?IntJEpidemiol 2016;

9.Sales VM, Ferguson-Smith AC, Patti ME. Epigenetic mechanisms of transmission of metabolic disease across generations. Cell Met 2017;25:559-571.

10.Nagy C, Turecki G.Transgenerational epigenetic inheritance: an open discussion. Epigenomics 2015;7:781-790.

11.Hall E, Volkov P, Dayeh T, Esguerra JLS, Salö S et al. Sex difference in the genome-wide DN A methylation pattern and impact on gene expression, microRNA levels and insulin secretion in human pancreatic islets. Genome Biology 2014;15:522-544.

12.Schaefer S, Nadeau JH. The genetics of epigenetic inheritance: Modes, molecules and mechanisms Volume 90, Quart Review Biology 2015;90: Pages

13.Earon SA, Jayasooriah N, Buckland ME, Buckland ME, Martin DI, Cropley JE, Suter CM. Roll over Weismann: extracellular vesicles in the transgenerational transmisson of environmental effects Epigenomics 2015; 7:1165-71.

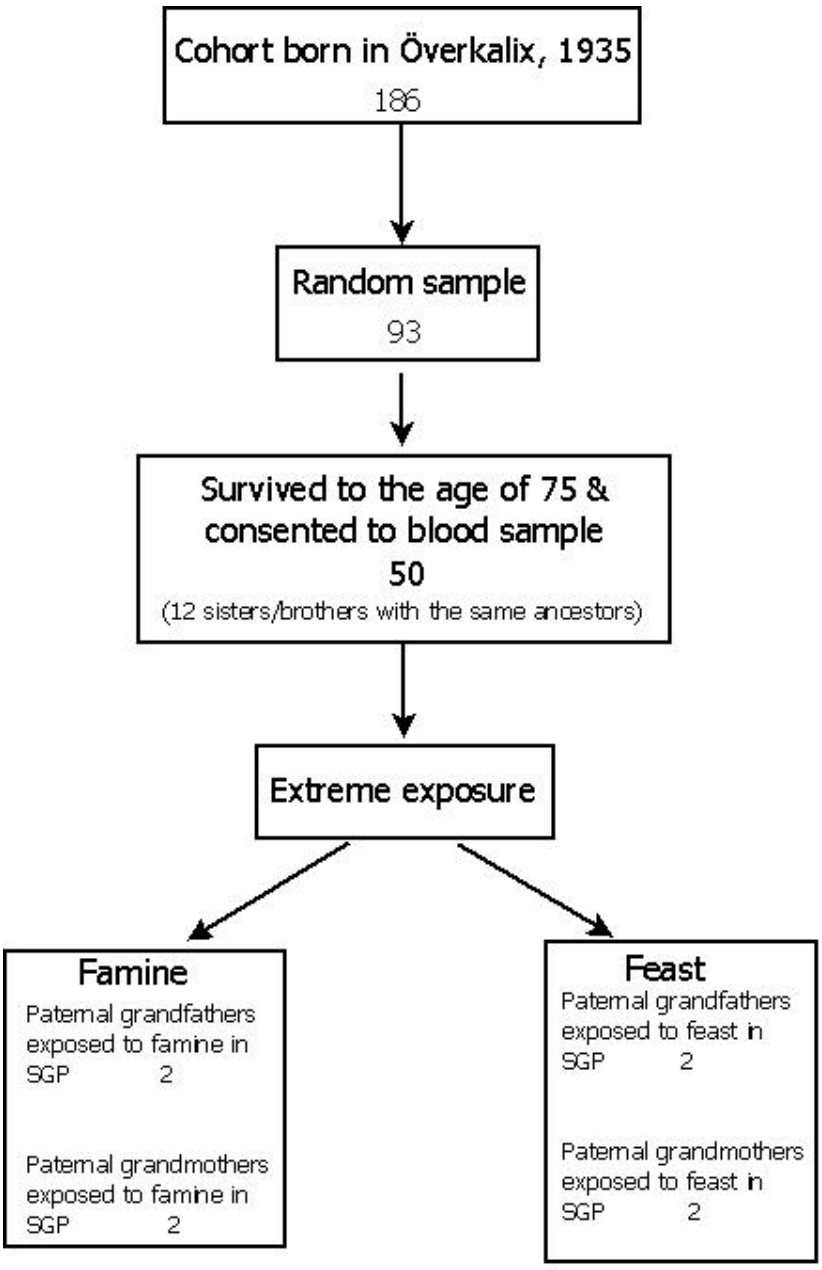
14.Vaiserman AM, Koliada AK, Jirtles RL. Non-genomic transmission of longevity between generations: potential mechanisms and evidence across species. Epigenetics Chromatin 2017;DOI 10.1186/s 13072-017-0145-1

15.Hellstenius, J. Skördarna i Sverige och deras verkningar. [Harvests in Sweden and their repercussions]. Statistisk Tidskrift 77-119. Stockholm, Sweden 1871 (sic).

16.Jörberg, L. (1972). A history of prices in Sweden 1732–1914. CWK Gleerup, Lund, Sweden.

17.Prader, A., R.H. Largo, L. Molinari and C. Issler (1989). Physical growth of Swiss children from birth to 20 years of age. First Zurich Longitudinal Study of Growth and Development. Helvetica Paediatrica Acta 43:Suppl 52: 1-125

18.Tanner, J.M. A History of the Study of Human Growth. Cambridge University Press, Cambridge 1981



328 Figure 1. Flowchart

Table 1. Gene ontology pathways describing reactions to hunger, excess food and stress at paternal grandparent´s exposure to variable food availability, followed by DMRs in grandchildren

Gene Ontology pathway	Genes	CpGs	MethylCpG	DMR%	Shore, Shelf[3]
Paternal grandfather with son and grandson					
32100 Positively regulating appetite	3	163	10	6%	Northern shore
30070 Insulin processing[1]	2	4	1	25%	Southern shore
46952 Ketone body catabolic process	4	16	1	6%	Southern shore
21854 Hypothalamus development[2]	6	33	5	15%	Southern shelf
Paternal grandmother with son and granddaughter					
1904179 Positive adipose development	2	11	2	18%	Northern shore
70341 Fat cell proliferation[4]	5	16	3	18%	Southern shelf
51599 IGF receptor binding[5]	9	94	12	13%	Northern shelf
46952 Ketone body catabolic process	4	16	1	6%	Southern shore
30325 Adrenal gland development	12	57	4	7%	Southern shelf

1. The formation of mature insulin by proteolysis of the precursor preproinsulin. The signal sequence is first elevated from preproinsulin, proinsulin is then cleaved to release the Cpeptide, leaving the A and B chain of mature insulin.
2. The progression of the hypothalamus region of the forebrain from its initial formation to its mature state.
3. CpG sites >200 bases with CpG content of 50% are termed “islands” (CGIs). Differential methylation induced by environment are most often seen in “northern- and southern-shores” <2kb from the islands, and “northern- and southern shelves” <2kb apart of the shores.
4. The multiplication or reproduction of fat cells resulting in expansion of their population.
5. Interacting selectively and non-covalently with the insulin-like growth factor receptor.

343 Table 2. Reproductive fitness following grandparental famine and feast in Slow Growth Period

Index person	Age at first child´s birth			Number of children		Longevity		
Exposure	**Grandparent**	**Father**	**Grandchild**	**Father**	**Grandchild**	**Grandparent**	**Father**	**Grandchild**
Paternal grandfathers exposure								
Famine								
Index person No 15	39	38	NA	4	NA	60	69	74+
Index person No 53	27	31	27	7	1	49	98	74+
Feast								
Index person No 14	28	23	30	6	3	80	70	74+
Index person No 23	33	26	NA	11	NA	94	83	63
Paternal grandmothers exposure								
Famine								
Index person No 59	39	38	34	3	2	80	83	74+
Index person No 86	24	20	22	6	2	55	64	74+
Feast								
Index person No 3	36	29	23	10	4	86	65	74+
Index person No 22	35	31	29	4	2	89	68	74+

344 NA. Not available