DOI:10.14753/SE.2018.2176

TRANSCRIPTIONAL AND GENOME-WIDE EPIGENETIC CHANGES UPON ENVIRONMENTAL STRESS

PhD thesis

Borbála Vető

Molecular Medicine Doctoral School

Semmelweis University

Supervisor: Tamás Arányi, MD, Ph.D

Official reviewers: Zsófia Nemoda, MD, Ph.D
Máté Varga, Ph.D

Head of the Final Examination Committee: János M. Réthelyi, MD, Ph.D

Members of the Final Examination Committee: Krisztina Takács, Ph.D
Attila Patócs, MD, Ph.D

Budapest

2018

DOI:10.14753/SE.2018.2176

TABLE OF CONTENTS

DOI:10.14753/SE.2018.2176

DOI:10.14753/SE.2018.2176

DOI:10.14753/SE.2018.2176

1. LIST OF ABBREVIATIONS

2-HG: 2-hydroxyglutarate

5azadC: 5-aza-2'-deoxycytidine

5caC: 5-carboxylcytosine

5fC: 5-formylcytosine

5hmC: 5-hydroxymethyl-cytosine

5mC: 5-methyl-cytosine

AACS: Acetoacetyl-CoA synthetase

ABCC6: ATP-binding cassette sub-family C member 6

ACOT: Acyl-CoA thioesterase

ACOX1: Acyl-CoA oxidase 1

Akt: PKB (Protein kinase B)

AML: Acute myeloid leukemia

AMP: Adenosine monophosphate

AMPK: AMP activated kinase

APOA1: Apolipoprotein A1

ARGFX: Arginine-fifty homeobox

ARNTL: Aryl hydrocarbon receptor nuclear translocator-like protein 1

ATCC: American type culture collection

ATP: Adenosine triphosphate

BARE: Bile acid response element

BDH1: 3-Hydroxybutyrate Dehydrogenase 1

BER: Base excision repair

bHLH: Basic helix-loop-helix

BLVRA, B: Biliverdin A, B

bp: base pair

bZip: Basic leucine zipper

cAMP: Cyclic adenosine monophosphate

CBP: CREB-binding protein

CEBPα, β: CCAAT/Enhancer binding protein alpha, beta

CGI: CpG islands

ChIP: Chromatin immunoprecipitation

ChREBP: carbohydrate response element binding protein

CpG: Cytosine-phosphate-guanine

CPT1: Carnitine O-palmitoyltransferase 1

CREB: cAMP responsive element binding proteins

CRX: Cone-rod homeobox

CT: control

CTCF: CCCTC binding factor

CYP7A1: Cholesterol 7α-hydroxylase

DBD: DNA-binding domain

DMEM: Dulbecco's Modified Eagle's Medium

DMR: Differentially methylated region

DMS: Differentially methylated site

DMSO: Dimethyl sulfoxide

DNA: Deoxyribonucleic acid

DNMT: DNA methyl-transferase

DPRX: Divergent paired-related homeobox

EDTA: Ethylenediaminetetraacetic acid

EGR-1: Early growth response gene-1

ERK1/2: Extracellular signal-regulated protein kinases 1 and 2

FAS: Fatty acid synthase

FBP1: Fructose-1,6-bisphosphatase 1

FBPase: Fructose 1-6-bisphosphotase

FBS: Fetal bovine serum

FDA: Food and Drug Administration

FOXO1: Forkhead box protein O1

FPKM: Fragments per kilobase per million reads

FXR: Farnesoid X receptor

G6Pase: Glucose-6-phosphatase

G6PC3: Glucose-6-phosphatase 3

GK: Glucokinase

GLUI: Glutamine synthetase

GLUT2: Glucose transporter 2

GR: Glucocorticoid receptor

GS: Glycogen synthase

GST: Glutathione S-transferases

GTF: General transcription factor

HAT: Histone acetyl-transferase

HBE: HNF4α-binding element

HDAC: Histone deacetylase

HDM: Histone Demethylase

hESCs: Human embryonic stem cells

HFD: High fat diet

hMeDIP: Hydroxymethylated DNA immunoprecipitation

HMGCS2: 3-Hydroxy-3-Methylglutaryl-CoA Synthase 2

hMSC: Human mesenchymal stem cell

HMT: Histone Methyltransferase

HNF4α: Hepatocyte nuclear factor 4 alpha

HOX: Homeobox

HPD: 4-hydroxyphenylpyruvate dioxygenase

HTH: Helix-turn-helix

IDH: Isocitrate dehydrogenase

IGV: Integrative genomics viewer

INSIG2: Insulin-induced gene 2 protein

iPSC: Induced pluripotent stem cells

Jmj: Jumonji domain

kb: kilobase

KDM: Histone lysine demethylase

KEGG: Kyoto encyclopedia of genes and genomes

LBD: ligand binding domain

LC-MS/MS: Liquid chromatography coupled with mass spectrometry

LEUTX: Leucine-twenty homeobox

LPK: Liver-type pyruvate kinase

LXR: Liver X receptor

MeDIP: Methylated DNA immunoprecipitation

NAD: Nicotinamide adenine dinucleotide

NADP: Nicotinamide adenine dinucleotide phosphate

NPC: Neuronal progenitor cell

NR: Nuclear receptor

OTX: Orthodenticle

PACAP: Pituitary adenylate cyclase-activating polypeptide

PAX: Paired box

PBS: Phosphate buffered saline

PC: Pyruvate carboxylase

PCR: Polymerase chain reaction

PEPCK/PCK: Phosphoenolpyruvate carboxykinase

PGC: Primordial germ cell

PGC-1α: Peroxisome proliferator-activated receptor gamma co-activator 1 alpha

PIC: Preinitiation complex

PITX: Paired Like Homeodomain

PKA: Protein kinase A

PKC: protein kinase C

PKLR: Pyruvate kinase, liver and red blood cell

PMSF: Phenylmethylsulphonyl fluoride

PPARα: Peroxisome proliferator activated receptor alpha

PRD: Paired

PTSD: Post-traumatic stress disorder

qPCR: quantitative PCR

RNA: Ribonucleic acid

RPKM: Reads per kilobase per million mapped reads

RPMI: Roswell park memorial institute medium

RRBS: Reduced representation bisulfite sequencing

RXR: Retinoic acid receptor

SAH: S-adenosyl-homocysteine

SAM: S-adenosyl-methionine

SD: Standard deviation

SDS: Sodium dodecyl sulfate

SDS-PAGE: Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SMILE: Small heterodimer partner-interacting leucine zipper protein

SNP: Single nucleotide polymorphism

SNV: Single nucleotide variant

SREBP-1c: Sterol regulatory element binding protein 1c

TDG: Thymine DNA glycosylase

TET: Ten-eleven translocation

TF: Transcription factor

TFBS: Transcription factor binding site

THF: Tetrahydrofolate

TPRX: Tetrapeptide repeat homeobox

TSH: Thyroid stimulating hormone

TSS: Transcription start site

UHRF1: Ubiquitin-like, containing PHD and RING finger domains 1

WT: Wild type

DOI:10.14753/SE.2018.2176

2. INTRODUCTION

2.1. Transcriptional regulation of gene expression

2.1.1. Introduction to transcriptional regulation

Eukaryotic gene expression is regulated by transcription. This step is controlled by proteins binding specific gene-regulatory sequences that modify the activity of the RNA polymerase. The tight regulation of gene expression is achieved by the intricate and complex network of these regulatory proteins. General or basal transcription factors (GTFs) bind to specific DNA sequences in order to control transcription rate. They form a large transcription preinitiation complex (PIC) and help to position RNA polymerase II to the promoter. In contrast, specific TFs control the expression of specific genes. They are responsible for the tissue-specific expression of their target genes. Mammalian gene expression is controlled by specific sequences: promoters and enhancers located near or far away from the transcription start site, respectively. The latter includes specific, conserved consensus sequences upstream from the TATA box, such as CCAAT sequences, which bind the CCAAT/Enhancer binding protein (C/EBP). C/EBPs can further bind other factors to open the chromatin structure and recruit GTFs causing transcription activation. A multitude of TFs can bind to *cis*-acting promoters and enhancers and they can interact with the RNA polymerase and the proteins of the transcriptional machinery by DNA looping (**Figure 1**). By the combination of the seemingly redundant activity of individual regulatory sequencing, a higher and synergistically acting unit is set up. Proteins regulating gene expressions may be transcriptional activators, which stimulate transcription *via* interacting with GTFs. In contrast, transcriptional repressors have inhibitory effect since they inhibit the interaction between RNA polymerase and GTFs or they block the binding of transcriptional activators. Moreover, the activators and repressors also cause changes in chromatin structure, making it accessible or inaccessible. Together, they are responsible for tissue-specific gene expression in complex multicellular organisms (summarized in [1]).

DOI:10.14753/SE.2018.2176

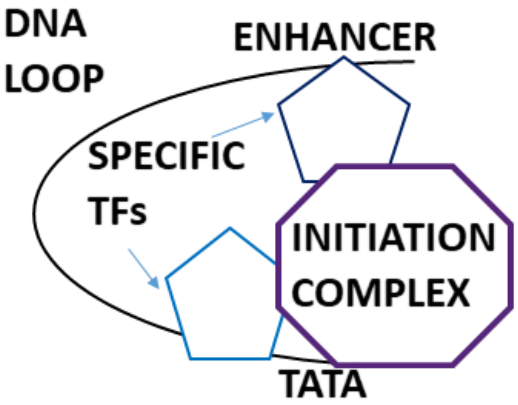


Figure 1. DNA looping. Transcription factors bound at enhancers interact with the general transcription factors at the promoter *via* DNA loops.

The basic function of TFs is binding DNA. It is carried out by the DNA-binding domain (DBD). In addition, TFs can be categorized based on their DBD (**Table 1**) [2]. Firstly, the family of homeobox proteins contain helix-turn-helix (HTH) motif. Vertebrate homeobox genes are highly conserved in structure and function, moreover, they play fundamental roles in animal development and differentiation. Homeobox (HOX) genes are essential in early axis formation, whereas Paired box (Pax) genes play important roles in neural growth, e.g. eye development [3]. Secondly, the typically dimeric Basic helix-loop-helix (bHLH) TFs contain two α-helices connected by a loop. Aryl hydrocarbon receptor nuclear translocator-like protein 1 (ARNTL/BMAL)-Circadian Locomotor Output Cycles Kaput (CLOCK) are prominent members playing a pivotal role in regulating molecular circadian clock [3]. Thirdly, basic leucine zipper (bZip) TFs possess a leucine zipper that dimerizes two DNA binding regions. The bZip family member CEBPβ takes part in the regulation of target hepatic gene expression in interaction with Hepatocyte nuclear factor 4α (HNF4α) [4]. Lastly, TFs families that contain zinc finger domain are nuclear receptors (NRs), which will be introduced in **Section 1.2.** In summary, the role of TFs is multi-faceted ranging from cell cycle control to pathogenesis, including the above mentioned functions.

Table 1. Families and subfamilies of transcription factors based on their DNA-binding domains. (modified from [2])

Family	Representative transcription factors	Some functions
1. Homeodomain:		
Hox	Hoxa-1, Hoxb-2	Axis formation
POU	Pit-1, Oct-2	Pituitary development; neural fate
Pax	Pax1, 2, 3	Neural specification; eye development
2. Basic helix-loop-helix (bHLH)	BMAL1-Clock	Molecular circadian clock
3. Basic leucine zipper (bZip)	C/EBPβ	Liver differentiation
4. Zinc finger:		
Nuclear hormone receptors	HNF4α	Differentiation, metabolism

First of all, some fundamentally important members of the homeobox-containing transcription factors will be discussed.

2.1.2. Homeobox-containing transcription factors

One major and evolutionary highly conserved family of transcription factors is encoded by various homeobox genes. Homeobox genes are found in almost all eukaryotes. They contain a distinctive homeodomain. These transcription factors are responsible for determining cell fates and the formation of many body structures during early embryonic development. Homeobox genes include approximately 100 gene families in which the members underwent extensive gene duplication and divergence during evolution. The most elaborated families further expanded *via* genome duplications in vertebrates (reviewed in [5]).

Hox genes play major roles in the formation of main body parts during early embryonic development. They are responsible for producing individual organs, as well. For example, limb development in mouse are regulated by HoxD and HoxA. Mutations in human Hox genes cause limb defects, for instance polydactyly [6].

The PRD class is the second largest of the homeobox gene classes, the name of which is originating from the Drosophila Paired (Prd) gene. It includes - for instance - Pax genes, Orthodenticle (Otx) and Paired Like Homeodomain (Pitx) genes, which are extremely important in the development of the mesoderm, the nervous system and the gut [5]. A 'new' group of PRD class homeodomains – Arginine-fifty homeobox (Argfx), Divergent paired-related homeobox (Dprx), Leutx (Leucine-twenty homeobox) and Tetrapeptide repeat homeobox (Tprx) – arose from tandem duplication from the Crx (Cone-rod homeobox) gene in a diversity of placental mammals. This parental gene is a member of the Otx gene family, and it is mainly expressed in photoreceptors of mammalian eyes. These daughter genes were formed by tandem duplications from the Crx gene, and they further diverged. Furthermore, some were transposed along (Dprx and Leutx) or to another chromosome (Argfx) throughout asymmetric evolution [7]. The human and mouse genomes have different numbers of homeobox genes due to both gene gain and gene loss events. Argfx, Dprx, Leutx and Tprx were lost in rodent but not in primate evolution [8]. In contrast to the retinal-expressed parental gene, the putative daughter genes are not expressed in photoreceptors, but in testes and human embryonic stem cells (hESCs). By far, the highest expression is found in human embryos at the 4-cell and 8-cell morula stages prior to cell fate restrictions. Morula cells are totipotent, thus they are equivalently capable of developing any cell type. However, it has never been proved that these proteins are actual transcription factors localized in the nucleus and they were never endowed with regulatory roles in totipotent embryonic cells. Therefore, determining their subcellular localisation and analysing their transcription profiles would shed light on their roles activating and repressing developmentally important genes leading to fine-tuning of cell fate decisions. In general, the link of genomic change to phenotypic evolution needs be established.

Secondly, the TF family of nuclear receptors (NRs) will be discussed.

2.1.3. Hepatocyte nuclear factor 4α

Members of the NR superfamily regulate transcription in response to binding their lipophilic ligands, which can be hormones (steroid, thyroid hormones), vitamin D, retinoic acid and also fatty acids and phospholipids. They are particularly important

because they turn the chemical signals into changes in tissue-specific gene expression programs. NRs comprise a ligand-binding and a DNA-binding domain (LBD and DBD, respectively). The ligands bind in the LBD of the receptors with great selectivity based on their special conformations. They can further recruit transcriptional coregulators; repressors and activators. NRs also play an important role in responding to various environmental stimuli, thereby controlling development, homeostasis and metabolism, as well (reviewed in [9]). The major role of these receptors is to maintain the nuclear homeostasis of their ligands. For example, peroxisome proliferator activated receptors (PPARs) regulate fatty acid metabolism by binding fatty acids and Liver X receptor (LXR) controls cholesterol metabolism by binding cholesterol metabolites, respectively. The ligands are synthesized and degraded when required for the adequate hormonal response. Retinoic acid receptor (RXR) regulates multiple metabolic pathways. It can form heterodimers with other NRs, for instance PPAR [10] regulating fatty acid metabolism. LXR, PPAR and also RXR are considered as the main players in whole-body lipid homeostasis and hepatic lipid metabolism. These TFs are highly cooperative, moreover LXR and PPARα demonstrate extensive crosstalk [11]. NRs are pharmacological targets, as well. For instance, PPARα is a major target for therapy in metabolic diseases, including insulin sensitivity, type 2 diabetes and cancer [12].

HNF4α is a member of the nuclear receptor superfamily. The TF is required for liver development and liver cell differentiation, and the HNF4α knockout mouse model is embryonic lethal [13]. HNF4α is expressed in the liver, pancreas, the kidney and the intestines [14] and it is the master regulator in hepatocytes [15]. HNF4α regulates genes involved in lipid and glucose metabolism [16] and other transcription factors [17, 18]. As a characteristic of several nuclear hormone receptors, they are able to react to environmental stresses, such as fasting or feeding and altered hormone levels. Point mutations of HNF4α can cause Maturity Onset Diabetes in the Young (MODY), a rare subtype of diabetes with autosomal-dominant inheritance [19].

HNF4α has long been considered as an orphan receptor, but its ligands are fatty acyl Co-enzyme A thioesters [20] according to crystallization studies [21] (**Figure 2**). Although HNF4α seems to be constitutively active [22] by its constitutive activator ligand, the transcription factor is not constitutively activating its target genes. Therefore, its post-translational modifications and interacting partners became an area of interest [23, 24].

Firstly, the interaction of HNF4α with different transcription factors (e.g. farnesoid X receptor FXR) [25], co-activators and co-repressors was discovered [25, 26]. In addition, mass spectrometry analyses of post-translational modifications revealed that HNF4α can be phosphorylated at several sites in the hepatocellular carcinoma HepG2 cell line [23]. The HNF4α protein was shown to be phosphorylated by different kinase cascades at specific residues. Serine 78 (which is S87 in human) is the most significant site targeted by protein kinase C (PKC) phosphorylation having an inhibitory effect [27]. Protein kinase A (PKA) [28] and p38 [29] were also reported to phosphorylate the protein. (However, these results appeared to be controversial.) Furthermore, AMP activated kinase (AMPK) can phosphorylate the protein at serine 313 [30], which leads to decreased activity of the transcription factor, however, lesser than PKC [27].

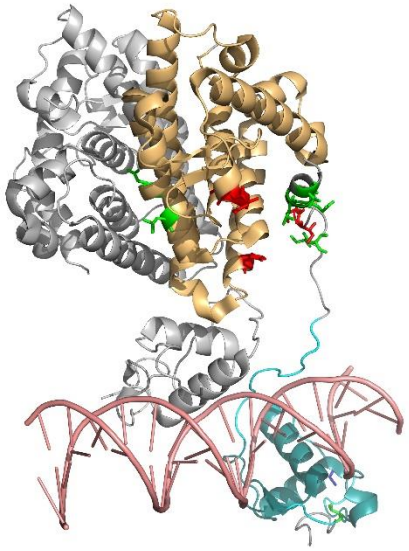


Figure 2. Crystallized structure of HNF4α as a homodimer. (Mutations coloured with the help of Tamás Hegedűs). (modified from [31])

Our group has previously identified various *cis*-regulatory elements at the promoter and a primate-specific sequence at the first intron of one of its target genes, *ABCC6* [4]. We have found that these *cis*-regulatory elements are occupied by HNF4α, CEBPα and β forming a regulatory network (**Figure 3**) [32], where HNF4α orchestrates tissue- and cell-type specific gene expression [4, 33]. Furthermore, we have previously observed that the activation of several kinase cascades (PKC, AMPK and ERK1/2) inhibits the expression of *ABCC6 via* HNF4α in human cell lines [33], where HNF4α was transfected in a

plasmid and gene expression was insensitive to ERK1/2. Therefore, we aim to discover whether HNF4α is a direct target of ERK1/2 phosphorylation and if yes, which phosphorylation site is responsible for the inhibitory effect.

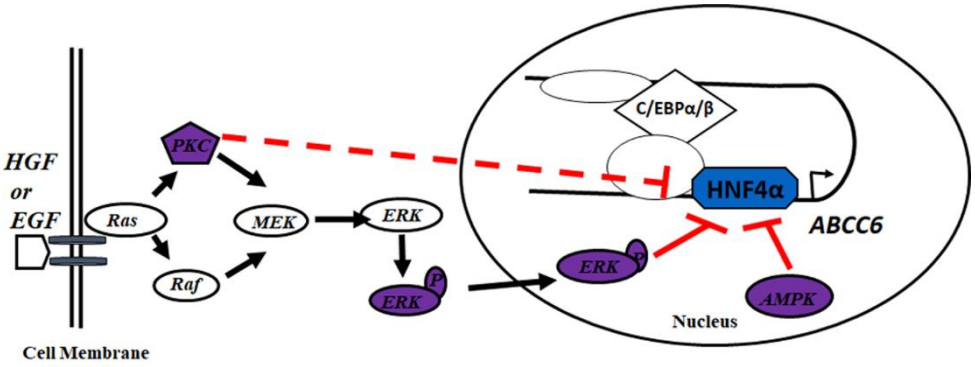


Figure 3. Transcription regulation of the *ABCC6* gene by various signalling pathways and a network of transcription factors including HNF4α. (modified from [32])

As we have seen above, TFs profoundly influence gene expression. However, the structure of DNA (e.g. DNA loops) and its surrounding genomic and chromatin environment in the nuclei of cells are also fundamental. Gene expression is regulated at two interconnected levels: transcriptional and epigenetic. It is now well known that cells can alter their gene expression in response to various environmental stimuli. The underlying processes are epigenetic.

2.2. Genome-wide epigenetic changes

2.2.1. Epigenetics, a definition

Epigenetics is the science of the hereditary information located in the nucleus but not encoded in the genome. Epigenetics involves various modifications of the chromatin structure in the nuclei of cells. Chromatin, a macromolecular complex includes DNA and nuclear proteins, which forms the nucleosome. It contains the four core histones (H3, H4,

H2A and H2B) twice in an octamer wrapped around by DNA. The N-terminal tails of the histones often undergo different modifications (e.g. acetylation and methylation). Moreover, DNA itself can be methylated (**Figure 4**). These covalent, epigenetic modifications affect changes in both chromatin structure and gene transcription. These changes can even be carried on by cell divisions through generations ('epigenetic inheritance') [34].

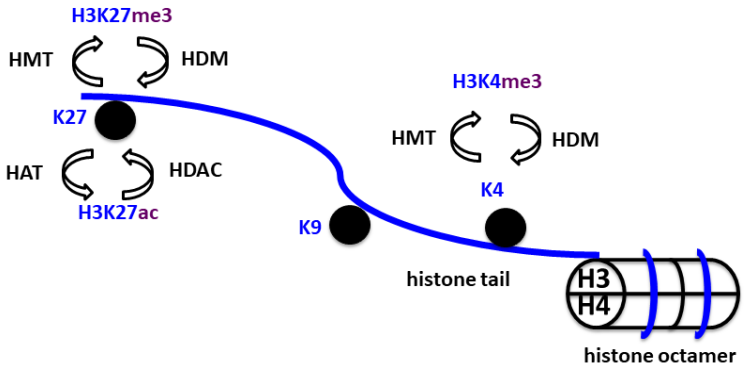


Figure 4. Overview of common covalent epigenetic modifications. Lysine 27 of histone 3 can be acetylated and methylated, as well. H3K27ac is an activating epigenetic mark, whereas H3K27me3 is an inhibitory one. In addition, lysine 4 can also be methylated, which is an activating epigenetic mark. HMT: Histone Methyltransferase; HDM: Histone Demethylase; HAT: Histone acetyl-transferase; HDAC: Histone deacetylase. (modified from [35])

DNA methylation has long been considered to be a stable epigenetic modification. However, cells were shown to actively adapt to the changing environmental conditions. It has recently been reported that DNA methylation can globally change upon various stress factors both in cell lines and in human hematopoietic stem cells [36, 37]. Therefore, these dynamic epigenetic modifications are the interface between the genome and the environment.

In the following, firstly histone modifications, then DNA methylation will be discussed.

2.2.2. Histone Modifications

Nucleosomes are built up from highly conserved histones having N-terminal tails, which can undergo post-translational modifications [38]: acetylation, methylation and phosphorylation. Histones play important roles in different nuclear processes, for instance replication, DNA repair and transcription [39, 40].

Acetylation targets the lysine residues of H3 and H4 histone tails performed by histone acetyl-transferases (HAT) and histone deacetylases (HDAC) [41, 42]. Some TF co-activator complexes can have HAT activity (e.g. p300 and CBP - CREB-binding protein), which recruit other transcription factors and RNA polymerase II [43]. Mutation in the transcriptional co-activator can cause the Rubinstein-Taybi syndrome [44], which is due to reduced histone acetylation and aberrant chromatin regulation. The disease is characterized by facial abnormalities and mental retardation. Histone acetylation characterizes transcriptionally active gene regulatory regions (euchromatin).

Histone methylation can target both arginines and lysines. In addition, arginines can acquire mono- or dimethylation, while lysines can acquire mono-, di- or tri-methylation [40, 45, 46]. A great number of enzymes are responsible for the establishment and removal of these epigenetic modifications. The executing enzymes are capable of selectively implementing one or two specific reactions (such as mono- and di-methylation of a specific lysine).

2.2.3. DNA methylation

DNA methylation is the covalent modification of cytosines by the addition of a methyl group. Methylation occurs at CpG dinucleotides in mammalian cells [47]. CpG dinucleotides are distributed in the genome to GC-poor regions [48], which tend to be methylated in all cell types. Short CpG-rich sequences (CpG islands – CGI) are mostly found in gene regulatory regions, while CpG-poor regions are a feature of the non-coding DNA [49]. Generally, CGIs exhibit tissue-specific methylation patterns [50].

DNA methylation influences gene expression. Methylation of a regulatory region implies gene silencing. TFs function based on the methylation status of their target sequence. Some do not bind methylated sequences (e.g. CTCF [51]), whereas others bind if the DNA is methylated (e.g. Methyl-CpG-binding protein 2 MeCP2 [52]). MeCP2 mediates transcriptional repression together with the Sin3 histone deacetylase [53]. MeCP2

mutations cause Rett syndrome, a progressive neurodevelopmental disease with mental retardation and motor impairment [54]. DNA methylation plays an important role in cell differentiation, X chromosome inactivation and other nuclear processes. DNA methylation is a reversible modification, thus methylation patterns can be established *de novo* and modified genome-wide or at specific loci, reflecting environmental conditions. DNA methylation is implemented by DNA methyl-transferase (DNMT) enzymes, which can establish or maintain the required methylation pattern. The maintenance DNA methyl-transferase DNMT1 preserves methylation patterns and preferentially binds hemi-methylated DNA. Ubiquitin-like, containing PHD and RING finger domains 1 (UHRF1) is a protein which recruits DNMT1 to the hemi-methylated DNA during replication. Therefore, methylation marks are protected during cell division. In contrast, the *de novo* DNA methyl-transferases DNMT3A and DNMT3B create 'new' methylation patterns (**Figure 5**).

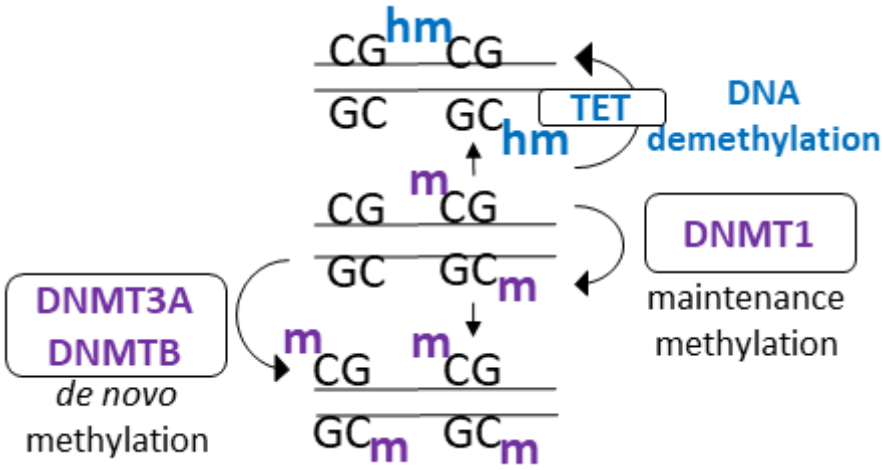


Figure 5. DNA methylation and demethylation performed by DNMT and TET enzymes. DNMT: DNA methyltransferase. TET: Ten eleven translocation.

DNA can be methylated genome-wide or locally in response to environmental stimuli, for example hormonal and metabolic stress. However, these changes are reversible. Embryogenesis is strictly regulated at the level of epigenetic modifications. During mammalian development, vast DNA methylation changes – erasure and re-establishment - occur in the zygote after fertilization and then in primordial germ cells (PGCs) [55]. An

example of *de novo* methylation is ESC differentiation into the three germ layers, which is accompanied by an elevation of methylated promoter regions, which has a critical role in the repression of germ cell-specific genes [56]). For instance, lentiviral transductions and infections can cause genome-wide methylation changes [37, 57]. Furthermore, DNA methylation can also be modulated upon stress or neuronal activity in the adult mouse brain, which is a proof of neuronal DNA methylome plasticity [58]. Post-traumatic stress disorder (PTSD) can also be connected to methylation changes in the Pituitary adenylate cyclase-activating polypeptide (PACAP)-encoding gene detected in peripheral blood [59]. Moreover, the behaviour of rat mothers (licking and grooming) can alter the DNA methylation patterns in the promoter of the glucocorticoid receptor (GR) in the hippocampus of the offspring affecting anxiety and stress response [60]. Epigenetic changes due to metabolic alterations and 'transgenerational inheritance' will be discussed in **Sections 3.1-3.3.**

2.2.4. DNA hydroxymethylation

5-hydroxymethyl-cytosine is an epigenetic modification of cytosines formed by Ten-eleven translocation (TET) (**Figure 5**), thereby taking part in demethylation. 5mC levels are unchanging in various tissues because of CpG methylation (3-7% of all genomic cytosines), whereas 5hmC levels are variable with the highest level measured in brain (approximately 1%) and 10-100 times less (0.01-1% of all cytosines) in other cell types. It has been reported that 5hmC does not have an obvious inhibitory effect on transcription [61, 62].

2.2.5. The function and action of TET enzymes

TET enzymes convert 5mC into 5hmC both *in vitro* and *in vivo* [63] *via* oxidation. In addition, they can further oxidize 5hmC to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) [64]. Passive DNA demethylation refers to the inability of the cells to maintain DNA methylation patterns during cell divisions. In contrast, active DNA demethylation is performed by a thymine-DNA-glycosylase (TDG)-base excision repair (BER)-dependent pathway mediated by TET. TETs have a catalytic core region,

moreover, TET1 and TET3 contain a CXXC zinc finger domain that can bind DNA [65]. Mouse embryonic stem cells (ESCs) have high 5hmC levels. Tet1 was shown to have a role in maintaining pluripotency in ESCs [66]. Knocking out Tet3 in ESCs lead to impaired neuronal differentiation [67] and Tet1/2/3 triple knockout ESCs are incapable of forming embryonic bodies [66].

Furthermore, TET1 mutation and associated low 5hmC levels were first demonstrated in acute myeloid leukemia (AML) [68]. TET2 mutations also resulted in impaired 5hmC production in haematological cancers [69]. Moreover, aberrant DNA methylation and TET-mediated loss of 5hmC can be seen as a hallmark of several human cancers. TETs are 2-oxoglutarate, oxygen- and iron- dependent demethylases, so α-ketoglutarate, a key player of the citric acid cycle can have epigenetic effect on genome activity and therefore influence health and disease. Studying some tumours have revealed mutations in the isocitrate dehydrogenase (IDH) genes disrupt the catalytic activity of TET proteins [70]. Moreover, IDH1/2 mutations were observed to be mutually exclusive with TET2 mutations in the case of AML patients [71]. Furthermore, ascorbate was also shown to take part in the oxidative demethylation of DNA *via* reducing Fe^{3+} to Fe^{2+}. It is required for the full TET catalytic activity [72, 73].

2.2.6. Cancer and epigenetics

Traditionally, cancer was considered driven by accumulated genetic mutations. However, it has recently been found that epigenetic regulatory mechanisms have a fundamental role, as well [74]. Before and during tumour progression, the epigenome is highly changed, for instance *via* genome-wide hypomethylation and increased promoter methylation at CpG islands of tumour suppressor genes. Promoter hypermethylation of tumour suppressor genes, for instance BRCA1/2 cell cycle and DNA repair gene, are widespread in several cancer types. Some other genes undergo loss-of-function because of promoter hypermethylation (reviewed in [75]). Cytosine methylation in the germline results in T:G mismatches created by the deamination of 5mC to T, which are not repaired by the DNA repair system. It can lead to mutations, single nucleotide polymorphisms and variants (SNPs and SNVs) at CpG sites. The most well-known example is the tumour suppressor p53 gene, where 50% of point mutations occur at these sites [76].

DNA methylation can be affected *via* mutations in the methylation machinery. DNMT1 mutations were observed in colorectal cancer [77] and DNMT3A mutations were noted in AML [78]. In addition, they can be overexpressed in various cancer types, which can cause hypermethylation elsewhere in the genome [79]. Moreover, the DNA demethylating enzyme TET2 can be mutated. For further modifications, see Table 1 in [75]. Mutations in histone modifier enzymes (e.g. HDACs) are also known to contribute to tumorigenesis (reviewed in [75]). Several anti-cancer drugs have been developed; HDAC inhibitors [80] or DNMT inhibitors, for instance Decitabine [81]. Therefore, these epigenetic changes, including defects in DNA methylation and histone modifications, can lead to genomic instability and gene expression changes, which might predispose to cancer. Since a cancer is characterized by several genetic and epigenetic aberrations, effective therapies aim at combining "standard" and epigenetic strategies.
2

2.2.7. Investigation of DNA methylation

DNA modifications can be detected by various methods (reviewed in[82]). Bisulfite conversion is the 'gold standard' for detecting 5mC nucleotides [83]. Nevertheless, this chemical conversion is unable to discriminate 5mC and 5hmC [84]. Sequencing of genomic DNA immunoprecipited with anti-5mC and anti-5hmC antibodies (MeDIP and hMeDIP) is used for this purpose. 5mC and 5hmC can also be distinguished at site-specific level by glucosylation of 5hmC followed by MspI digestion, where the glucosylated DNA is resistant to digestion [85]. However, these methods are not adequate for exact quantitative analysis [86].

2.2.8. LC-MS/MS (Liquid Chromatography coupled with Mass Spectrometry)

Global DNA methylation can be accurately quantified by the highly reliable and simple LC-MS/MS. This method is capable of simultaneously measuring global 5mC and 5hmC levels in genomic DNA [87, 88]. The advantage of this method is the ability to distinguish between 5mC and 5hmC. Moreover, global DNA methylation is measured sensitively and accurately and the method is highly quantitative and reproducible. The limit of detection

is very low (below 1 fmol) [88]. Genomic DNA is liberated into nucleobases by formic acid instead of digestion by a multienzyme cocktail. In addition, calibration is performed by using chemical standards (dCTP, 5mdCTP, 5hmdCTP) [89]. Using this method, we have previously shown that DNA methylation is highly dynamic [36].

2.3. Epigenetic implications of metabolism

2.3.1. Epigenetics in intermediary metabolism

The specific transcriptional networks are established and maintained in every cell by precise molecular programs under different nutritional conditions. These cell-specific programs are regulated by epigenetic mechanisms. These activities are further regulated by the actual nutritional state of the cell. The activity of chromatin modifications are influenced by key metabolites and cofactors originating from intermediary metabolism, for instance acetyl-CoA, NAD+, S-adenosyl-methionine (SAM) and α-ketoglutarate (α-KG) [90]. Here, only the link between metabolic sensors and DNA methylation will be discussed (**Figure 6**). Firstly, DNA methylation requires a methyl donor, which originates from the SAM- S-adenosylhomocysteine (SAH) and the folate cycle. The changing levels of SAM have a direct effect on H3K4 trimethylation and gene expression [91]. Furthermore, H3K4me3 levels are maintained at decreased SAM levels by negatively regulating KDM5B/JARID1B histone demethylase activity [91]. SAH, produced from SAM, is an inhibitor of DNA methyltransferases, thus the SAM/SAH ratio serves as a biosensor of actual metabolic status and nutrient availability.

Secondly, α-KG is generated from isocitrate in the mitochondrial citric-acid cycle by IDH2, which contributes to DNA demethylation (reviewed in [35]). α-KG is also formed in the cytoplasm *via* IDH1 and from amino acids, for example arginine and glutamate. Mutations in the IDH enzymes lead to the production of the oncometabolite 2-hydroxyglutarate (2-HG), an inhibitor of α-KG [92], which will disrupt α-ketoglutarate dependent TeT enzyme activity, causing tumours and DNA hypermethylation [71]. Furthermore, the activity of both TET enzymes and Jumonji domain containing histone demethylases (Jmj-KDMs) can be impeded by subsequent metabolites of the citric acid

cycle, for instance fumarate and succinate [93]. Thus, metabolic rate, DNA methylation and chromatin dynamics are tightly connected.

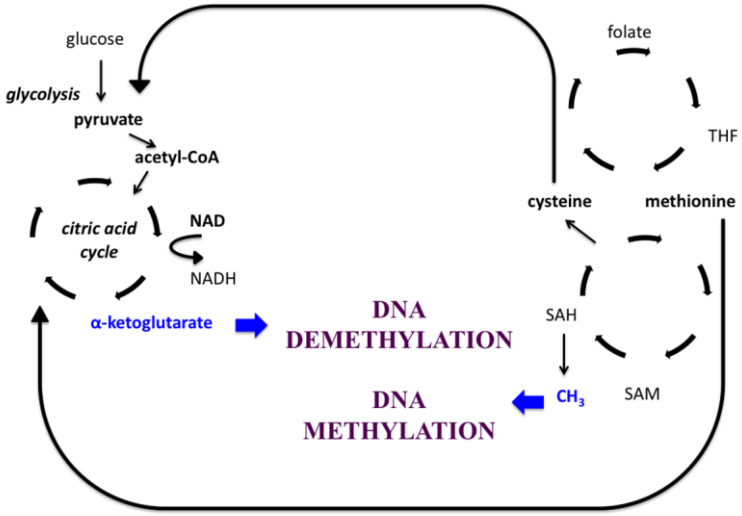


Figure 6. The relationship between DNA (de)methylation and intermediary metabolism. DNA can acquire methylation by DNMT enzymes *via* a methyl group originating from the SAM-SAH and folate cycle. DNA can be demethylation by TeT enzymes *via* α-ketoglutarate. NAD: Nicotinamide adenine dinucleotide; THF: Tetrahydrofolate; SAM: S-adenosyl-methionine; SAH: S-adenosyl-homocysteine. (modified from [35])

2.3.2. Introduction to nutritional challenges

Nutritional challenges accompanied mankind throughout history. During most of the time, lack of sufficient food caused and causes problems. More recently, eating habit alterations (diets, irregular calorie intake, eating junk food) are more and more frequent. Exaggerated calorie intake - or so-called 'Western diet' rich in macronutrients, fat, carbohydrates and protein - leads to obesity, which is nowadays becoming epidemic. It is a well-known risk factor for a lot of diseases, for instance diabetes, cardiovascular diseases (hypertonia, stroke, etc.) and different types of cancer [94, 95].

Nutritional stress can be divided into 2 groups: chronic and acute. By chronic nutritional stress or obesity, almost 2 billion people are affected worldwide

[http://www.who.int/en/96Access date: 06/02/2018]. Moreover, cardiovascular diseases are among the leading causes of death nowadays. Interestingly, both human and rodent studies strongly suggest the role of epigenetic mechanisms in the development of this disease., obesity can affect the following generations, which is much less investigated in human [97] than in rodents [98, 99].

On the other hand, a frequent form of acute nutritional stress is fasting. Since stable blood sugar level is required for glucose homeostasis and metabolic equilibrium, several biochemical pathways have evolved to maintain its concentration above a certain limit. Although these mechanisms are well-known, regulation of the implicated genes is still only partially understood.

2.3.3. Previous studies of nutritional stress

2.3.3.1. Studies on chronic nutritional stress

In vivo studies demonstrating the dynamic changes of chromatin modifications upon environmental challenges are getting more and more frequent. In rodents, the common models investigate different setups of nutritional challenge. Thousands of studies were published on the effect of diets in rodents. It became clear that high-fat, high carbohydrate, calorie restricted, folate and protein-deficient diets disrupt the metabolic balance of the offspring.

Diets have been proved to affect the following generations, as well. Adequate nutrition including macronutrients and micronutrients during pregnancy is important for the normal growth of the embryo and foetus. Perturbations (e.g. not sufficient maternal nutrition) can lead to oxidative stress and DNA damage and mortality [100, 101]. In addition, this may result in cardiovascular diseases, insulin resistance, and increased body fat in adulthood. In rats, for example, HFD lead to early reproductive maturation and later obesity in the offspring [102]. Furthermore, *in utero* exposure to maternal diabetes results in insulin intolerance and high blood pressure in the offspring [103].

Deficiency in an essential nutrient, folate is a known risk factor for liver injury. Maternal methyl donor (folate) supply during pregnancy has long-term effects on the metabolism of the offspring. Folate protects hepatocytes of hyperhomocysteinemia mice from

apoptosis [104]. Low selenium and folate in maternal diet during gestation and lactation has more effects on gene expression in offspring post-weaning in the mouse liver [105]. These experiments point out the importance of the maternal diet (e.g. folate and other methyl donors) during pregnancy on DNA methylation patterns and the metabolism of the offspring. Epigenetics seems to be the link between early life nutrition with later health [106]. We have seen from the above investigations, that nutritional challenges are often accompanied by general disturbances of covalent chromatin modifications.

2.3.3.2. Studies on acute nutritional stress

The changes in transcription factors and enzymes related to metabolism are quite well-studied upon fasting for different, but short durations. However, there are only a very few number of studies investigating epigenetic alterations triggered by metabolic alterations. Namely, differential methylation at the promoter of the *Pepck* (*Phosphoenolpyruvate carboxykinase*) gluconeogenic gene was observed in the post-weaning rats with folic-acid supplemented diet [107]. Furthermore, fasting induced gluconeogenesis, glycogenolysis, β-oxidation and ketogenesis. 12-hour fasting dramatically decreased serum glucose levels in mice and 16-hour fasting increased serum β-OH butyrate concentration [108]. Hepatic glucose production relies on Glucose-6-phosphatase and PEPCK expression, which were also elevated upon 16-hour fasting. Ketogenesis increased, shown by accumulation of liver triacylglycerol. Enhanced ketogenesis was also characterized by increased Carnitine O-palmitoyltransferase 1 (Cpt1), 3-Hydroxybutyrate Dehydrogenase 1 (Bdh1) and 3-Hydroxy-3-Methylglutaryl-CoA Synthase 2 (Hmgcs2) mRNA levels. Refeeding for 1-2 hours after fasting leads to decreased G6Pase and PEPCK mRNA expression. In addition, Cpt1, Bdh1 and Hmgcs2 were no longer elevated. Therefore, refeeding immediately blocked hepatic ketogenesis. The pancreatic hormones regulate the response of the liver to metabolic stress. The insulin/glucagon ratio is decreased upon fasting and restored upon refeeding. Glucagon elevated cAMP concentrations, which directly regulates the transcription of the above mentioned genes. In addition, FoxO1 was dephosphorylated and gluconeogenesis was upregulated [108].

Another study reported decreased hepatic Fatty acid synthase (Fas) and Srebp1c mRNA

level upon 24 hours' fasting, which was restored upon refeeding. Srebp1c is a TF regulating lipogenesis, and it regulates the expression of Fas, which is a key enzyme involved in fatty acid synthesis. Increased insulin levels during refeeding enhanced the activity of Srebp1c. At the same time, elevated RXRβ, γ and Liver X receptor α and β (LXRα and β) mRNA levels were reported [109]. LXR promoted the transcriptional activation of Srebp1c [110]. Furthermore, the mRNA level of Insig2 - encoding a protein, which inhibits the activation of Srebp – was increased when mice were fasted and decreased when refed [111]. Insulin-mediated decrease of Insig2 allows fatty acid synthesis.

Despite the detailed investigations of TFs and enzymes playing a role in metabolic adaptation to nutritional stress, these challenges are less studied in regard to chromatin changes and especially to DNA methylation alterations. In addition, the role of chromatin structure changes and epigenetic factors in these processes need detailed investigations. The switches between periods of scarcity and excess of nutrients require fast adaptive regulation, which includes hormonal responses, gene expression and chromatin changes. By investigating the molecular mechanisms of fasting and refeeding, understanding the fundamental process of physiology has important clinical relevance. Furthermore, the epigenetic component of acute fasting and refeeding is still only poorly understood.

2.3.4. Glucose homeostasis under physiological and stress conditions

For the metabolic balance of mammalian hepatocytes, the role of glucose is unquestionable. The process of maintaining blood glucose at a steady-state level is called glucose homeostasis. For instance, diabetes mellitus - characterized by the inability to maintain glucose homeostasis - is the clinical manifestations of a long-term metabolic abnormality. The result of impaired glucose homeostasis is hyperglycemia. In order to avoid hyperglycemia and fasting hypoglycemia, the body can adjust glucose levels by various cellular mechanisms.

2.3.4.1. Glucose metabolism under physiological conditions

Glucose is an essential energy-providing source for most mammalian tissues. Its

metabolism is tightly regulated, the rate of which is controlled by the consumption of carbohydrates and its absorption in the intestines. The equilibrium among the regulating hormones, namely glucagon, insulin and cortisol is well established in all the key metabolic tissues, such as the brain and the skeletal muscle, however, the critical organ is the liver (**Figure 7**). Glucose is taken up and can be converted to pyruvate by glycolysis (*via* the action of the tissue-specific liver-type pyruvate kinase LPK). Excess of glucose is transported by Glut2 and stored as glycogen *via* glycogenesis and the activation of glycogen synthase (GS). Several other metabolic pathways, for instance lipogenesis, fatty acid biosynthesis and triglyceride synthesis play a role in storage, and at the end, white adipose tissue serves as the depo. Several transcription factors take part in the regulation of these pathways, such as sterol regulatory element binding protein 1c (SREBP-1c) and carbohydrate response element binding protein (ChREBP) (reviewed in [112].

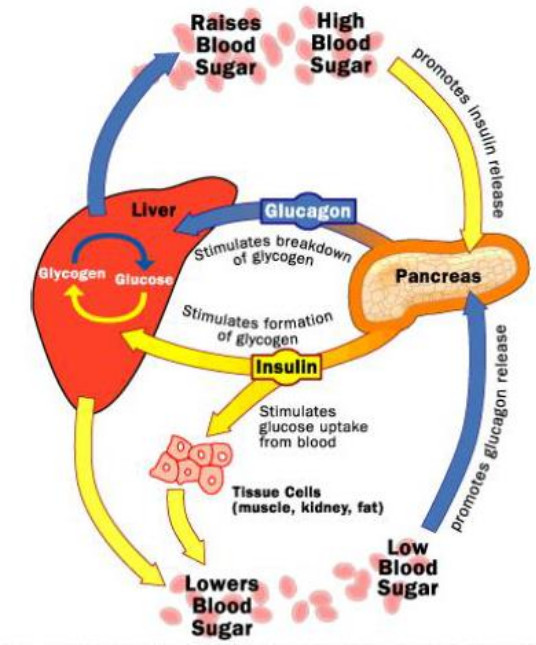


Figure 7. Glucose metabolism and the insulin cycle. (modified from [113])

2.3.4.2. Glucose metabolism in fasting

Upon fasting, glycogenolysis and gluconeogenesis are activated in order to maintain glucose levels *via* its production and release. In fasting conditions, glucagon levels are

elevated, while insulin levels are decreased. Glycogen is broken down *via* glycogenolysis. Long-term fasting results in *de novo* glucose synthesis *via* gluconeogenesis. The enzymes participating in this pathway are: glucose-6-phosphatase (G6Pase), fructose 1-6-bisphosphotase (Fbpase), pyruvate carboxylase (PC) and phosphoenolpyruvate carboxykinase (PEPCK). Hormone signals are also important in generating the desired outcome. The major catabolic hormone glucagon and the stress hormone cortisol activate several transcription factors.

2.3.5. Transcription factors regulating glucose homeostasis in fasting

Fasting, a significant perturbation of mammalian metabolism is very well described at the level of hormones, molecules and enzymes, but until now, only little is known about the underlying complex transcription factor network. Indeed, fasting is highly regulated at the transcriptional level leading to dynamic adaptation to metabolic needs. Upon fasting, transcription factor binding occurs to specific target sequences.

The fasting-related TF families are reported in the literature, which are the following: nuclear receptors (e.g. hepatocyte nuclear factor 4 alpha - HNF4α, glucocorticoid receptor - GR, peroxisome proliferator activated receptor alpha - PPARα, retinoic acid receptor - RXR), cAMP responsive element binding proteins (e.g. CREB1), forkhead box proteins (e.g. FoxO1) and CCAAT enhancer binding proteins (e.g. CEBPα, β). Essential co-activators of gluconeogenesis are the CREB binding protein (CBP)/p300 and the peroxisome proliferator-activated receptor gamma co-activator 1 alpha (PGC-1α) [114].

2.3.5.1. HNF4α in fasting

The maintenance of glucose homeostasis upon fasting and feeding (or hormone) is tightly regulated, already at the transcriptional level, for instance through the action of HNF4α. Glucokinase (GK) and glucose-6-phosphatase (G6Pase) are the key enzymes of glycolysis and gluconeogenesis, respectively. During fasting, GK transcription is inhibited, while feeding or insulin activates it [115] (**Figure 8**). It has been reported that HNF4α and FoxO1 play important roles in these processes, moreover, they directly interact with each other [116]. Upon fasting, FoxO1 represses HNF4α-potentiated GK

transcription, whereas it activates G6Pase gene activation synergizing with HNF4α [117]. Under this condition, FoxO1 localizes in the nucleus and it represses GK promoter activity. At the same time, FoxO1 and HNF4α synergistically activate the G6Pase promoter. There are HNF4α-binding elements (HBEs) on both the GK and the G6Pase promoter. The former contributes to the activation of GK upon insulin stimulation, whereas the latter to the repression of G6Pase upon insulin. During feeding (or insulin treatment), FoxO1 is phosphorylated by Akt and released from HNF4α [116], which can activate the GK promoter *via* the HBE shifting the balance to glycolysis. Thus, FoxO1 can have both positive and negative influence on target gene activity. HNF4α reciprocally regulates GK and G6Pase in that FoxO1 translates the insulin signal. These two processes together shift the metabolic balance towards gluconeogenesis [117] (**Figure 8**). It is the combination of TF binding to a promoter that determines the actual transcriptional outcome. In addition, the promoter region of L-PK (liver-type pyruvate kinase), an important enzyme having a regulatory role in glycolysis can also have FoxO1-regulated, HNF4α-potentiated transcriptional activity [118].

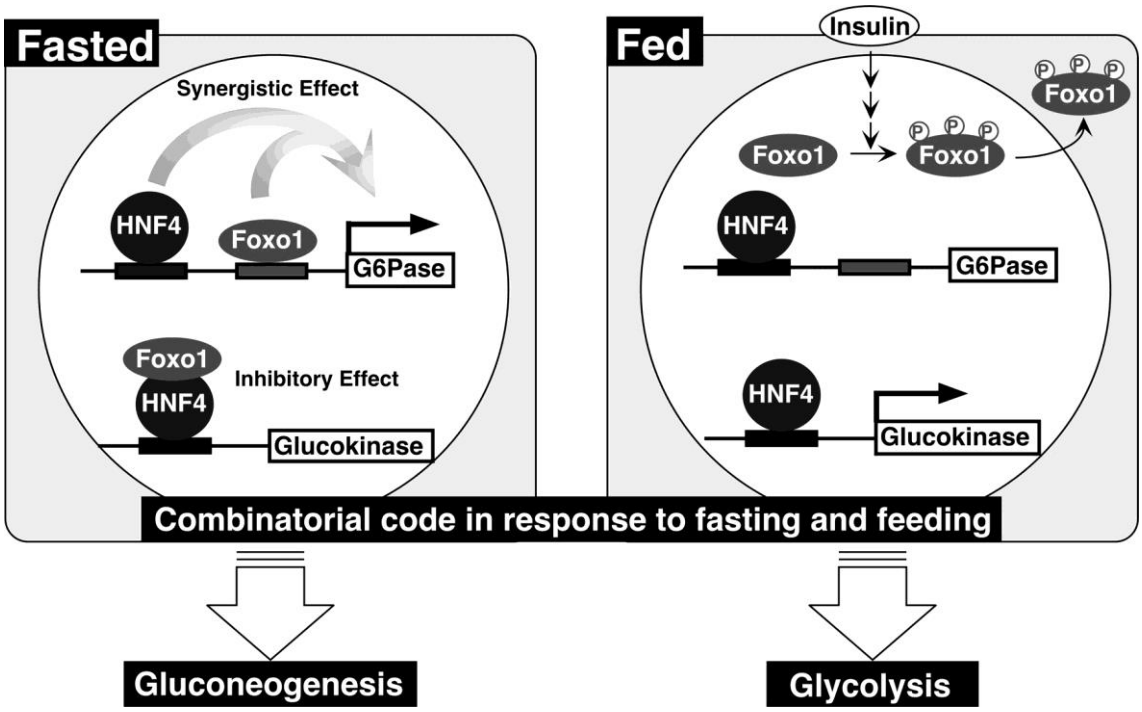


Figure 8. Regulation of glucose homeostasis by FoxO1 and HNF4α [117]. Upon

fasting, FoxO1 binding to HNF4α synergistically activate the gluconeogenic G6Pase and Glucokinase is inhibited. On the contrary, upon feeding, FoxO1 is phosphorylated and HNF4α is free to activate the glycolytic Glucokinase. HNF4α: Hepatocyte nuclear factor 4 alpha; Fox1: Forkhead box protein O1; G6Pase: Glucose-6-phosphatase. Permission to reuse is obtained from the journal.

2.3.5.2. Interaction between HNF4α and PGC-1α

PGC-1α also interacts with HNF4α. In mouse primary hepatocytes lacking HNF4α, PGC-1α cannot activate gluconeogenic genes. PGC-1α also activates the genes of β-oxidation and ketogenesis, however, it was found to be independent of HNF4α activity [119].

A total knockout of HNF4α is embryonic-lethal, but liver-specific knockout mice are viable, however, they have defective lipid homeostasis. HNF4α is required for PGC-1α action (**Figure 9**). To prove that, PGC-1α was expressed in primary hepatocytes lacking HNF4α, which eliminated PEPCK and G6Pase mRNA levels. PGC-1α interacts with HNF4α on the response element of PEPCK promoter. Furthermore, HNF4α itself is not capable of activating the G6Pase promoter, only in the presence of PGC-1α. Finally, not every HNF4α binding site is targeted by PGC-1α co-activation. For instance, PGC-1α can induce β-oxidation even in the absence of HNF4α [119].

DOI:10.14753/SE.2018.2176

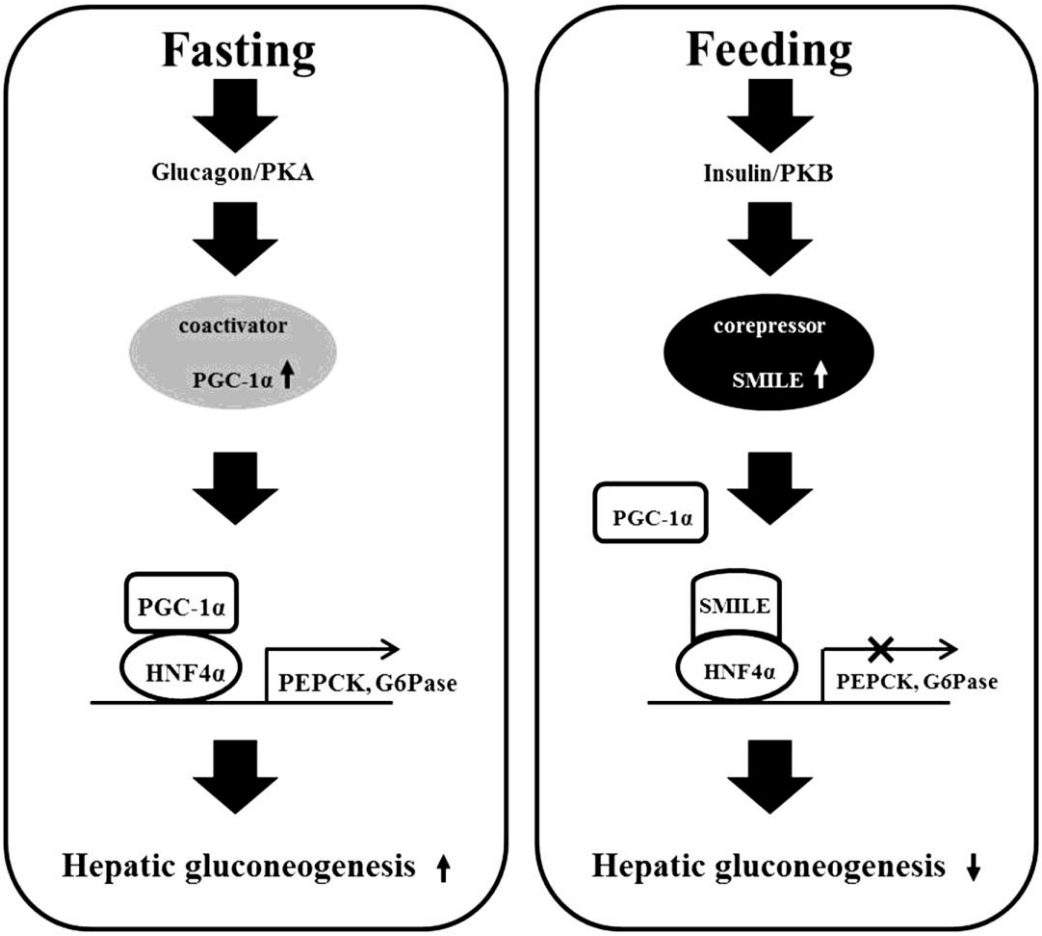


Figure 9. Regulation of hepatic gluconeogenesis by HNF4α and PGC-1α [120]. Fasting induces glucagon/PKA, which activates PGC-1α. In complex with HNF4α, hepatic gluconeogenic genes PEPCK and G6Pase are activated. Feeding has the opposite effect through the action of insulin/PKB and SMILE. PKA: Protein kinase A; PGC-1α: Peroxisome proliferator-activated receptor gamma co-activator 1 alpha; HNF4α: Hepatocyte nuclear factor 4 alpha; PEPCK: Phosphoenolpyruvate carboxykinase; G6Pase: Glucose-6-phosphatase; SMILE: Small heterodimer partner-interacting leucine zipper protein. Permission to reuse is obtained from the journal.

2.3.6. Cholesterol homeostasis

2.3.6.1. Cholesterol 7α-hydroxylase (CYP7A1)

Cholesterol homeostasis depends on the conversion of cholesterol to bile acids in the

liver. This mechanism eliminates excess of cholesterol. CYP7A1 is the enzyme catalyzing the first step of the pathway. It regulates bile acid synthesis and lipid homeostasis. CYP7A1 expression is controlled by various factors, for instance bile acids, nutrients and hormones. A binding site for HNF4α has been identified in the human CYP7A1 promoter, which regulates its transcription [121]. It binds to the bile acid response element (BARE). It has been shown recently that bile acids block HNF4α recruitment of the PGC-1α co-activator to both CYP7A1 and PEPCK genes therefore inhibiting their transcription. The fasting-feeding cycle regulates bile acid production, and fasting induces CYP7A1 expression together with the activation of the gluconeogenic PEPCK *via* PGC-1α [122]. Reduction of bile acid synthesis (by fasting) leads to increased cholesterol levels in the liver. Nutritional status affects CYP7A1 through the regulation of insulin and glucagon hormones. Fasting represses CYP7A1 expression in rats [123]. In addition, it has been reported that glucagon and cAMP inhibit CYP7A1 mRNA expression [124]. Furthermore, cAMP can reduce CYP7A1 gene expression in human primary hepatocytes *via* the action of PKA, which phosphorylates HNF4α. It was shown that the glucagon/cAMP/PKA signalling is responsible for the reduced binding and trans-activational capacity of HNF4α on CYP7A1. However, there is contradiction in the literature as bile acids were reported to inhibit PEPCK [122] and that fasting and cAMP induce both CYP7A1 and PEPCK expression [125].

2.3.7. Temporal organization of transcriptional regulation

On the whole, transcriptional regulatory events are temporally organized during fasting [126]. Firstly, glucagon is secreted shortly upon blood glucose level decrease, which leads to CREB1 activation. TFs regulating fuel production are strictly organized in two major modules: the gluconeogenic and ketogenic module. As the first stage of fasting response, corticosterone is secreted, and GR-activated CREB1 results in the activation of gluconeogenesis. Gluconeogenesis begins approximately 4-6 hours after the beginning of fasting. It is activated to maintain blood glucose levels. At this stage, GR binds to new regulatory sites and increases accessibility to CREB1 binding sites [126]. Therefore, hepatic glucose production is vastly elevated.

Additionally, GR-induced PPARα leads to the activation of ketogenesis. Accordingly, the

delayed ketogenic gene expression results in the production if ketone bodies from fatty-acid oxidation, and it will be the dominant fuel-producing pathway. Prolonged fasting activates fatty acid activation in order to provide ATP for the liver [126]. The formation and export of ketone bodies provide alternative source of glucose for the brain. Hyperglycemia is the major cause of type II diabetes.

2.3.8. Refeeding: hormonal and transcriptional responses

During refeeding, insulin represses glucagon secretion in the pancreas and its action in the liver. Insulin activates the Akt kinase cascade, which activates several pathways including hepatic glucose uptake of hepatocytes. LXR, a nuclear receptor also activates lipogenic genes and activates another transcription factor Srebp1c [127]. Srebp1c was reported to be diminished in fasted mice, whereas increased upon refeeding [128]. Altogether, glycogenolysis, gluconeogenesis and fatty acid oxidation are inhibited, whereas glucose degradation, citric acid cycle, fatty acid and triglyceride synthetic pathways are activated. As mentioned earlier, the DNA methylation changes, their role and the mechanism of their appearance have not been studied yet in detail during the fasting-refeeding cycle.

DOI:10.14753/SE.2018.2176

3. OBJECTIVES

I intended to answer the following questions during my PhD:

I. What is the subcellular localisation and the role during human embryonic development of the group of human PRD-class proteins (Argfx, Dprx, Leutx and Tprx)?

II. Does ERK1 phosphorylate HNF4α? If yes, what is the result of ERK1/2-phosphorylated HNF4α on target gene transcription and target gene DNA-binding?

III. What are the macroscopic, protein level and methylation changes occurring upon short-term fasting and refeeding *in vivo*?

IV. How does treatment with DNA methylation inhibitor or ascorbate affect genomic 5mC and 5hmC levels in cell lines?

DOI:10.14753/SE.2018.2176

4. METHODS

Firstly, those methods and techniques will be described which I performed by myself. The only exception is the "Mouse nutritional stress timeline", which was designed and accomplished together with Tamás Arányi and Flóra Szeri.

4.1. Cloning of homeobox genes

Human Argfx coding sequence was synthesized by GenScript. Dprx and Leutx were cloned by ligating PCR products. 5' Tprx region was ligated to Tprx1-specific 3' region amplified from human DNA. The coding sequences were cloned in-frame with a C-terminal V5 tag in pSF-CMV-COOH-V5 vectors (Oxford Genetics). I transfected the vectors into HeLa cells.

4.2. Immunofluorescence

Immunofluorescence was performed after 48h culture. Cells were fixed with 4% paraformaldehyde. Cells were then permeabilized and BSA blocked (1%, 30 min). Cells were stained with anti-Argfx, anti-Dprx (Abcam) or anti-V5 (ThermoFisher) specific antibodies for 1 h and stained with secondary antibody (Alexa Fluor 488 or 594, ThermoFisher). Cells were co-stained with phalloidin and DAPI (ThermoFisher).

4.3. Cell culture

HepG2 human hepatoma cell line was purchased from ATCC (ATCC HB-8065) and kept in Advanced MEM (ThermoFisher) supplemented with 10% FBS, 2mM L-glutamine, 100 U/ml penicillin and 100 mg/ml streptomycin. HeLa cells were obtained from ATCC and cultured according to the manufacturer's instructions (DMEM supplemented with 10% FBS, 2mM L-glutamine, 100 U/ml penicillin and 100 mg/ml streptomycin). Treatment was performed in serum-free medium for 30 minutes or 24 hours with human recombinant epidermal growth factor (Sigma–Aldrich) at 100 ng/ml final concentration.

4.4. HNF4α mutations

PcDNA5-FRT/TO expression vector containing wild-type, full length human HNF4α gene was obtained from Addgene. Amino acid numbering here refers to mutations of the human HNF4α gene, and numbering of the rat HNF4α gene are adjusted to the human numbering. Mutations were created for serine or threonine phosphorylation sites in order to have phosphomimetic (glutamate or aspartate) or neutral (alanine) mutations: S87D, T166A/S167D, S313D, S451E, T451A/T459E. If two phosphorylation sites were adjacent or very close to each other, both were mutated. Gene synthesis and site-directed mutagenesis were performed by the biotechnology company GenScript. The HNF4α coding cassette was re-cloned from pcDNA3+ into pcDNA5-FRT/TO plasmid.

4.5. Transfection and luciferase experiments

HeLa cells were plated onto 96-well plates in 10.000 cells/well amounts. For transfection, FuGENE HD reagent (Promega), serum-free medium and 2 μg plasmid DNA was used for the cells. Triple co-transfection was performed with the phABCC6(-332/+72)Luc construct ((see [33]) containing the *ABCC6* promoter fragment (-332/+72) cloned upstream of the luciferase coding cassette in pGL3-Basic vector (Promega)), pcDNA5-FRT/TO plasmid encoding HNF4α mutants (GenScript) and pRL-TK Renilla luciferase Control Reporter Vector (Promega). Cells were harvested and lysed after 48 hours. Luciferase activity was determined by Victor luminometric plate reader (Perkin Elmer) utilizing the DualGlo Luciferase system (Roche). Results were normalized firstly for the background noise, then for transfection efficiency by the co-transfected Renilla control reporter vector.

4.6. ChIP (chromatin immunoprecipitation) assay

Formaldehyde - in a final concentration of 1% - was added to the culture media containing approximately 10 million HepG2 cells. After 10 min incubation at room temperature, fixation was stopped by ice-cold glycine at 125 mM final concentration, then washed three times with ice-cold phosphate buffered saline (PBS). Cells were scraped and washed

again with PBS, then centrifuged. Pellets were resuspended in Lysis buffer and incubated vortexed at 4°C for 15 mins. Lysed cells were centrifuged, resuspended in Sonication buffer and sonicated on ice with an MSE sonicator (6 pulses of 15s each at 15% amplitude with 30s off between each pulse). The sonicated fragments were approximately 500 bp long, as determined in advance. After centrifugation at 13,000 x g, the supernatant was diluted 10 times with IP buffer and supplemented with cOmplete Mini cocktail tablets. For each 50 μL extract, a mixture of 60 μL of Dynabeads protein A and 60 μL of Dynabeads protein G containing 2 μg anti-HNF4α mouse monoclonal antibody (Abcam ab41898) or H3K27ac antibody (Abcam ab4729) was added. The mix was incubated at 4°C overnight, then the beads were washed three times with different buffers at 4°C with constant rotation for 5 mins. After washing twice with TE buffer, samples were eluted in 200 μL freshly prepared Elution Buffer. Reverse crosslinking was performed overnight at 65°C. Samples were treated with RNAse at 37°C for 1 hour, then proteinase K was performed for 2 hours at 45°C. DNA was purified using High pure PCR template preparation kit.

4.7. Quantitative PCR

Immunoprecipitated, purified DNA was used as a template for quantitative PCRs. The primers used were designed with BiSearch [129] and their sequences are shown in **Table 2**. PCR amplicons were designed between 100 and 200 bp. QPCR was performed in 20 μL containing SYBR green mix (Roche), 5 μL of ChIP-enriched DNA and 250 nM primers in a 96-well plate. Standard curves were generated from sonicated genomic DNA samples at different dilutions, and their relative amounts were calculated by extrapolating from the dilution curves. All standards and samples were run in duplicate. Plates were run in a LightCycler 480 real-time PCR machine (Roche). Enrichment of a given DNA fragment was calculated by comparing its relative concentration to concentration to the input (which is the total amount of sonicated DNA without immunoprecipitation).

Table 2. Primer sequences for ChIP-qPCR.

PRIMER NAME	PRIMER SEQUENCE
ABCC6 promoter F	AGCCCATTGCATAATCTTCTAAGT
ABCC6 promoter R	ATGGAGACCGCGTCACAG
ABCC6 exon31 F	AAGTACACACAGCATGGCAG
ABCC6 exon31 R	AGGACCTAGCAATACACAGG
β-globin F	AGGACAGGTACGGCTGTCATC
β-globin R	TTTATGCCCAGCCCTGGCTC
APOA1 F	ATTGCAGCCAGGTGAGGAGAA
APOA1 R	TTAGAGACTGCGAGAAGGAG
BLVRA F	TTGTTTTGGAATGGGGGTGG
BLVRA R	AAAAGGGAAGGCTGTGGCAA
BLVRB F	CACCCTTTACCCTCTTTACC
BLVRB R	GCCTGTGCTTTTGTGTTTAC
HPD F	GATAGGGAAAACAGCCACCA
HPD R	TTGGATGATGAGGACACAGG
PKLR F	GTGGCTTACATGCTGTGGCT
PKLR R	TAGGTGGGTTTTGGAGAGGA

4.8. Cell culture

A375 melanoma, A2058 melanoma, HepG2 hepatocarcinoma, HeLa cervix carcinoma, MES-SA uterine sarcoma, H1650 bronchoalveolar carcinoma and HTR8 placenta cell lines were purchased from ATCC and cultured according to the manufacturer's instructions, either in Dulbecco's modified Eagle's medium, Advanced MEM or RPMI-1640 (Sigma) supplemented with 10% fetal bovine serum (FBS), 1% glutamine, 1% penicillin/streptomycin. DT40 cells were cultured in RPMI-1640 medium supplemented with 7% FBS, 3% chicken serum, 50 μM β-mercaptoethanol and penicillin/streptomycin.

4.9. Drug treatments

Cells were treated with 5-aza-2′-deoxycytidine in DMSO at a final concentration of 1 μg/ml or with vehicle for 48 hours. DNA was extracted according to the Puregene (Qiagen) protocol.

DOI:10.14753/SE.2018.2176

4.10. Statistical analyses

For the luciferase experiments, Tukey-HSD test was performed.
For the ChIP-qPCR experiments evaluating the effect of short-term EGF treatment performed in various experiments one-sample t-test was performed.
For the LC-MS/MS and mRNA expression measurements, Student's t-test was performed.

4.11. Mouse nutritional stress timeline

6 male, 8 weeks' old, littermate C57BL/6 mice were used for each group of the fasting-refeeding experiments. All the animals in the groups had drinking water *ad libitum*. We have fasted 6 animals per group for 24 hours (FASTED 24h), for 16 hours (FASTED 16h) or 8 hours (FASTED 8h). Each group consisted of 6 animals. The FASTED 24h group was fasted for 24 hours overnight starting from 6 p.m until 6 p.m. the following day. The FASTED 16h group was fasted for 16 hours from 6 p.m. until 10 a.m. the following day. The FASTED 8h group was fasted for 8 hours starting from 6 p.m. until 2 a.m. the following day. In contrast, refeeding was performed for a group of animals for 8 hours (REFED 16+8h) or for 4 hours (REFED 16+4h) starting immediately after the end of 16 hours' fasting (10 a.m.) until 6 p.m or 2 p.m. during daytime, respectively. Moreover, we intended to have samples from every 4 hours. Therefore, we also had a 'CT 4h' and a 'CT 12h' group. The 7 conditions resulted in 7 groups, each including 6 animals, which is altogether 42 samples. The chart of the timeline and the corresponding groups and samples can be observed on **Figure 10**.

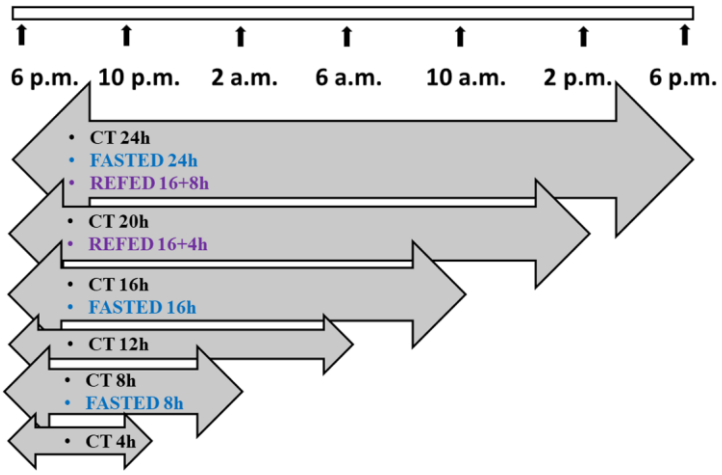


Figure 10. Summary of groups and animals in the fasting-refeeding experiment. Fasting was set to start at 6 p.m. Refeeding was set to start at 10 a.m. Each group contained 6 animals.

4.12. Mouse experiments

C57BL/6J mice were derived from mice purchased from The Jackson Laboratories. Mice were kept under routine laboratory conditions in an approved animal facility. The RCNS, Hungarian Academy of Sciences Institutional Animal Care and Use Committees approved the animal studies. Mouse livers were rapidly perfused in order to wash out blood, immune and other cells from the heterogeneous liver and only investigate hepatocytes. Perfusion was done with RT PBS for 1 minute, and the liver was cleared out and free of blood. Mouse liver was then taken out freshly from the animals and snap frozen in liquid nitrogen. Genomic DNA was prepared with DNeasy Blood & Tissue Kit (Qiagen).

4.13. The Reduced Representation Bisulfite Sequencing (RRBS) method

Reduced Representation Bisulfite Sequencing (RRBS) is a simple and efficient method to analyse DNA methylation at single nucleotide resolution. It is much more cost-effective than whole genome sequencing, but it provides a reduced representation of the genome. The RRBS method is highly efficient (above 99% bisulfite conversion rate) and

contains minimal bias since amplification is reduced to the minimum. Samples then undergo multiplex next-generation sequencing. With the RRBS technique, a coverage of up to 4 million CpGs can be reached for the human genome. The technique is validated and already tested on several species [130, 131].

The RRBS kit and protocol was obtained from the biotechnology company Diagenode (C01030033). Genomic DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen) from freshly obtained mouse liver. 100 ng genomic double-stranded DNA was digested by MspI restriction enzyme, which cuts at CCGG sites, either methylated or unmethylated. Then the flanking ends were prepared using an Ends Preparation Enzyme, where the added bases end in A. At this step, unmethylated and methylated spike-in controls were added to the samples to ensure to control for bisulfite conversion efficiency. It was followed by adaptor ligation and size selection performed by AMPure XP Beads (Beckman Coulter, Inc.). During the adaptor ligation step, a thymine was added to the complementer A. The addition of Illumina 6 base pair-long barcodes enabled future distinction among the samples. The range of the fragments was size selected to be between 200-1200 bp. Sample concentrations were quantified by quantitative PCR, which permitted them to be pooled together by comparing their relative concentrations to each other. Samples were pooled together using the Pooling Aid provided by Diagenode, which compared Ct values, and therefore relative concentrations were calculated with the equation 2^(-dCt). The RRBS kit is optimized to pool together 6 samples. After the volumes of the samples were reduced and adjusted to the right volumes, bisulfite conversion was performed already on 6 samples in 1 pool, and the reaction was carried out overnight. The bisulfite conversion converted unmethylated Cs into Us and methylated Cs remained Cs with very high efficiency. It was followed by quantitative PCR in order to determine the optimal cycle number for the enrichment PCR. Enrichment PCR was conducted with MethylTaq Plus Polymerase - an enzyme without proofreading activity -, and it created Ts complementer to A, which are Us in fact. Lastly, PCR amplification was followed by a clean-up step with the help of magnetic beads. Samples were sequenced with Illumina HiSeq2000 platform with single end 50 bp reads resulting in 120M reads per pool. **Figure 11** shows the flowchart of the technique.

DOI:10.14753/SE.2018.2176

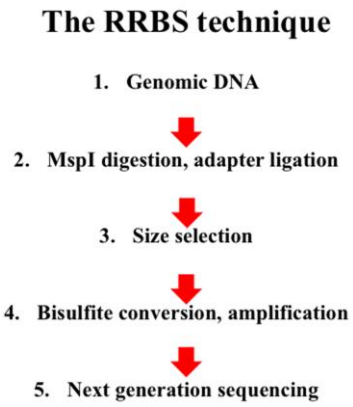


Figure 11. Flowchart of the RRBS technique. MspI digestion site is CCGG. Fragment size is selected for 200-1200 bp.

4.14. RT-PCR

RNA was isolated using the miRNeasy Micro Kit (Qiagen) from freshly obtained mouse liver. Reverse transcription was performed from 1 μg RNA (RevertAid, ThermoFisher). Published primer sequences were used for expression analysis (**Table 3**). QPCR was performed with SYBR green mix in a BioRad CFX96 PCR machine. Standard curves were generated from cDNA at different dilutions, and their relative amounts were calculated by extrapolating from the dilution curves. Enrichment of a given cDNA fragment was calculated by comparing its relative concentration to concentration to the housekeeping 18S RNA.

Table 3. Primer sequences for RT-PCR.

PRIMER NAME	PRIMER SEQUENCE
18S RNA F	GGCCGTTCTTAGTTGGTGGAGCG
18S RNA R	CTGAACGCCACTTGTCCCTC
PCK1 F	CTGCATAACGGTCTGGACTTC
PCK1 R	CAGCAACTGCCCGTACTCC
HNF4a F	GCGGAGGTCAAGCTACGAG
HNF4a R	CAATCTTCTTTGCCCGAATGTC
FAS F	CCTGGATAGCATTCCGAACCT
FAS R	AGCACATCTCGAAGGCTACACA
G6P F	CGACTCGCTATCTCCAAGTGA
G6P R	GTTGAACCAGTCTCCGACCA

The following Methods were not performed by me. These experiments are described in detail in the respective publications.

4.15. Confocal microscopy

Immunofluorescent images were taken with an inverted IX81 motorized microscope equipped with FV1000 Point scanning laser and Becker and Hickel FLIM system.

4.16. Fibroblast RNA-Seq and data analysis

Primary human fibroblasts were combined with constructs and electroporated. Total protein was extracted RIPA buffer supplemented with cOmplete mini protease inhibitor. 1 μg of RNA was used for mRNAseq library preparation utilizing the TruSeq RNA kit. Sequencing was performed at an Illumina HiSeq4000 75 bp paired end platform. Reads were aligned to the human reference genome GRCh38.p2 (NCBI) with the STAR RNA-seq aligner. Differential gene expression and FPKM values were calculated. Gene expression data from 7 developmental time points were used for creating temporal profiles of expression.

4.17. *In vitro* phosphorylation assay

The phosphorylation assay was performed in a mixture including kinase buffer, 500 ng ERK1 kinase (Sigma), HNF4α human recombinant protein with GST-tag at N-terminal (Abnova), and 20 uM ATP including 1 μCi [γ-^{32}P] ATP. Phosphorylation was started by the addition of ATP. The reaction was stopped after 30 minutes at 30°C by adding SDS sample buffer. Samples were run on SDS-PAGE using 10% running gels. Then gels were subjected to autoradiography for 2–12 hours.

4.18. Phosphomapping

The HNF4α protein obtained from the gel band was reductively alkylated with DTT and

iodoacetamide. Then it was digested with trypsin in 20 mM ammonium bicarbonate buffer. An aliquot was run a Thermo/Dionex Ultimate RSLC nano system using a 75um x 15 cm C 18 PepMap column (Thermo/Dionex) coupled to a Thermo LTQ Orbitrap Velos Pro. TOP 15 MS method (65 min linear gradient from 5-40% B (80% acetonitrile in 0.1% formic acid) was used with multi-stage activation. Data was analysed by the Mascot search engine against the sequence of HNF4α. Phosphopeptides were assigned to peptides above a mascot ion score of 20. The same peptide could be present several times by the detection system. However, the higher the ion score for the peptide is, the more likely that the assignment is correct.

4.19. ChIP-Seq data analysis

Raw sequence files of the ChIP-seq samples were analyzed (hg19 reference genome). ChIP-seq peaks were predicted by HOMER. Artefacts were excluded according to the blacklisted genomic regions of the Encyclopedia of DNA Elements using BEDTools. RPKM (Reads Per Kilobase per Million mapped reads) values were calculated on the summit -/+50 bp region of the peaks for HNF4α, or on the whole region of the histone signal for H3K27ac. Motif enrichment analysis was performed. Pathway analysis was obtained using from the KEGG database. The average read density was determined. Read distribution and average density heat maps were made from Java TreeView. Histogram and box plot were obtained using GraphPad Prism.

4.20. Western blot

Cells were washed with PBS and lysed in Harvest Buffer containing 1 mM phenylmethylsulphonyl fluoride (PMSF), 25 μg/ml each of Pepstatin A, trypsin inhibitor and aprotinin. Lysates were centrifuged for 10 min at 4°C. Sample buffer was added to the supernatants, and the samples were boiled for 3 min. Samples were subjected to SDS-PAGE using 7.5, 10 or 12.5% running gels. PVDF membranes were used. Membranes were blocked and incubated for 60 min with the appropriate primary antibodies overnight at 4°C. Monoclonal anti-HNF4α, PCK1 and CEBPα antibodies were used (Abcam). After several washing steps, membranes were incubated for 30 min with a horseradish

peroxidase-conjugated secondary antibody and washed again. Reacting antigens were visualized with the enhanced chemiluminescence detection reagents. Quantification and statistical analysis was performed using densitometry.

4.21. RRBS data analysis

Illumina reads were quality checked with the FastQC software. The reads were adapter trimmed and an additional two nucleotides were removed from their 3' ends. After trimming, reads were mapped to mm9 genome using Bismark tolerating one non-bisulfite mismatch per read. After mapping the sorted sam files were subjected to the methylKit for further analysis. Firstly, methylated and unmethylated Cs in CpG context were read in. Similarity of samples was checked by calculating pairwise Pearson's correlations. Histograms of %methylation per cytosine and histograms of read coverage per cytosine were generated to further assess similarity of samples. Differentially methylated sites were extracted (q <=0.01, minimum difference >= 0%) and annotated according to the type of genomic region they are located (promoter, exon, intron, intergenic or CpGisland, shore, other). Hypo- and hypermethylated chromosomal positions reported by methylKit were further annotated using HOMER package to obtain the nearest ENTREZ and RefseqIDs and Gene Names.

4.22. LC-MS/MS

DNA was hydrolyzed to nucleobases using formic acid. 100% formic acid was added to the samples and put into a 2 ml glass vial. The vial was incubated at 130°C for 90 min. After nitrogen evaporation the samples were reconstituted in acetonitrile:water:formic acid 49.5:49.5:1 solution. Chemical standards (dCTP, 5mdCTP, 5hmdCTP) were used to obtain the highest sensitivity. A Sciex 6500 QTrap mass spectrometer including a turboV ion source was used. Perkin Elmer Series200 system (including a binary pump, autosampler and column oven) was used for separation. Water containing formic acid in 0.1% and acetonitrile containing formic acid in 0.1% was used for separation. Gradient elution was performed. An Agilent RX-Sil column (250 x 4.6 mm, 5um) was used for the separation. The flow rate was 1 ml/min with the injection of 40 µl of samples. The column

temperature was room temperature, samples were at 5°C in the autosampler. Source conditions in mass spectrometric measurements were: spray voltage was 5000 V, evaporation temperature: 500°C, curtain, evaporation and drying gases were 45, 45 and 50 instrument units. 50 msec dwell time and 5 msec pause times were set for the MRM transitions. Collision energy was 30eV.

4.23. Calibration

Calibration was done on nucleotides. The same sample preparation was as described above. A calibration curve consisting of 10 points in the range of 1-10% of mC and 9 point calibration curve in the range of 0.01-1% of hmC relative to C was used by mixing the nucleotide base solutions (dCTP, 5mdCTP, 5hmdCTP). Due to the high dynamic range of the mass spectrometer the relative concentration values did not depend on the absolute amount of mixed nucleotides. The area ratios of 5mC/C and 5hmC/C were measured, but the lowest calibration points were not identical with the lowest limit of quantitation. The LOQ and limit of detection values are dependent on the absolute amount of 5hmC in the sample. Extrapolated ratios were used based on relatively strong peaks for integration even if the area ratios were out of the range.

DOI:10.14753/SE.2018.2176

5. RESULTS

5.1. Transcriptional regulation by homeobox-containing TFs

As forecasted in the **Objectives** section, we intended to answer to following question for homeobox genes:

I. What is the subcellular localisation and the role during human embryonic development of the group of human PRD-class proteins (Argfx, Dprx, Leutx and Tprx)?

In answer to this question, my co-author publication includes the results described below [132].

5.1.1. Subcellular localisation of the group of proteins from the human homeobox genes Argfx, Dprx, Leutx and Tprx

First of all, I cloned the homeobox genes Argfx, Dprx, Leutx and Tprx into mammalian expression vectors or vectors containing C-terminal V5-tags with PCR. I transfected the constructs in HeLa cells in order to express them. I stained the cells with anti-Argfx, anti-Dprx or anti-V5 specific antibodies. Following immunocytochemistry, I took immunofluorescent images (not shown), and confocal images were also taken. The results revealed that these proteins localise predominantly to the nucleus with some exceptions of cytoplasmic staining. In the lack of high transcfection efficiency, I did not perform statistical analysis of nuclear localisation staining %. Some examples of immunocytochemistry images are shown on **Figure 12**. Furthermore, the V5-tagged constructs were also transfected to primary human fibroblasts by Thomas L. Dunwell. Immunocytochemistry followed by confocal microscopy revealed that the proteins in question are also located in the nucleus, although some show other subcellular staining (**Figure 13**). More precisely, Argfx and Tprx1 showed clear nuclear localisation. In contrast, Dprx exhibited nuclear and cytoplasmic staining and Leutx showed nuclear staining, but cells seemed to be disrupted. This prominent nuclear localisation is a characteristic of transcription factors.

DOI:10.14753/SE.2018.2176

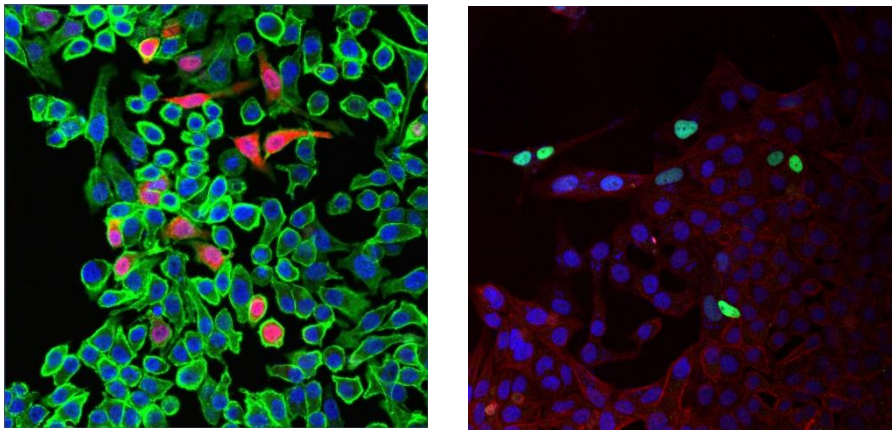


Figure 12. Nuclear localisation of Argfx and Dprx transfected in HeLa cells. Images show nuclei (blue DAPI), ectopic protein (red for Argx on left panel and green for Dprx on right panel) and actin cytoskeleton (green for Argfx and red for Dprx phalloidin) in merged images.

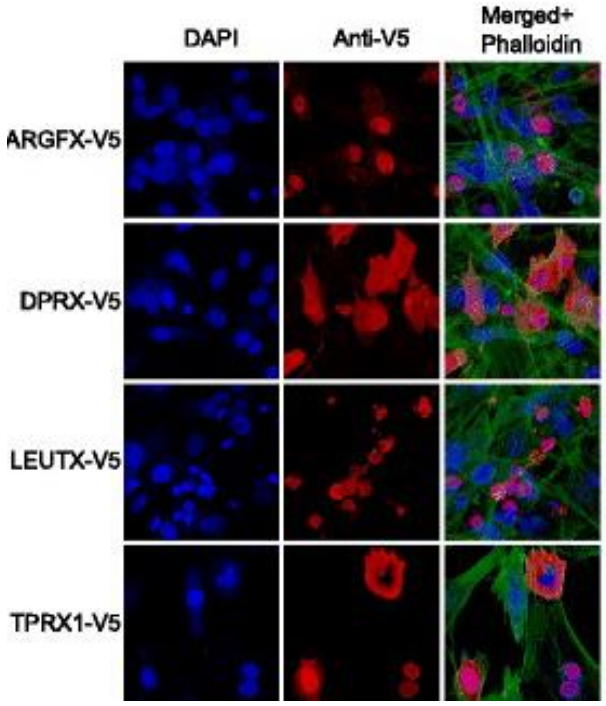


Figure 13. Nuclear localisation of V5-tagged in primary human fibroblasts. (modified from [132])

In addition, transcriptome analysis was performed by Thomas L. Dunwell. Primary human fibroblasts cultured for 48 hours underwent RNA-seq using Illumina platform. A

great number of significantly up- and down-regulated genes were detected by the bioinformaticians (Ignacio Maeso, Thomas L. Dunwell and Chris D. R. Wyatt). The experiments verified the transcriptional activity of these homeobox proteins. Moreover, temporal clustering has revealed that these genes are expressed between the oocyte and blastocyst stages in human embryonic development. When a set of 50 human genes were investigated, Argfx, Leutx and Tprx1 exhibited a sharp transition from low or zero expression until the 4-cell stage to a high expression at 8-cell and morula stages with a steep decline before the blastocyst stage. It clearly shows that these genes are characterized with a sharp switch-on and off expression pattern and they are expressed immediately before cell fate determination (**Figure 14**).

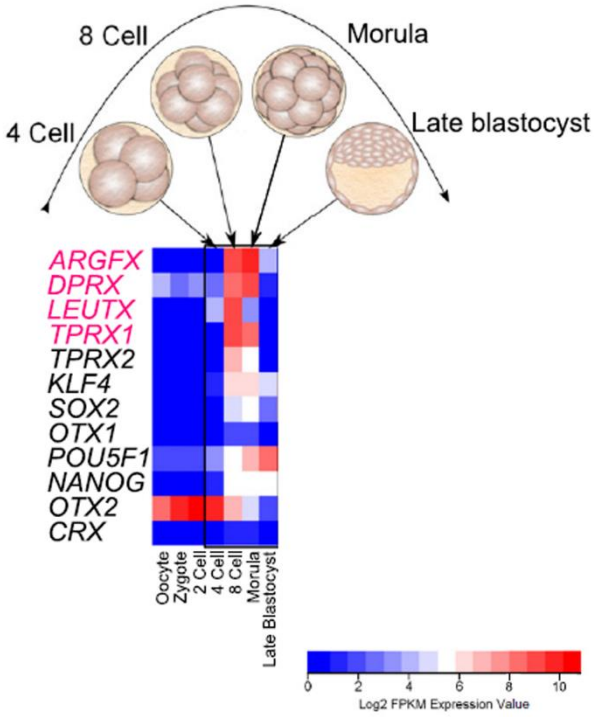


Figure 14. Heatmaps showing expression profiles of human homeobox genes (Argfx, Dprx, Leutx and Tprx) and other stem cell markers. FPKM: fragments per kilobase per million reads on a log scale (red: high, blue: low expression). (modified from [132])

More importantly, a downstream effector has been found, which is the HIST1H2BD histone H2 variant.

DOI:10.14753/SE.2018.2176

5.2. Transcriptional regulation by HNF4α

As we have seen in the **Objectives**, I intended to find answers to the following question: **II.** Does ERK1 phosphorylate HNF4α? If yes, what is the result of ERK1/2-phosphorylated HNF4α on target gene transcription and target gene DNA-binding?

5.2.1. Transfection efficiency and nuclear localisation of the HNF4α protein and its mutant form

Before answering the question raised above, I intended to find out if there is a difference in transfectional efficiency or a defect in the nuclear localisation of the phosphorylation mutant HNF4α protein compared to the wild-type (for detailed explanation, see **Introduction**). From the representative immunofluorescent images (**Figure 15**), a vector containing wild-type HNF4α protein (**left panel**) and that of the S313 phosphomimetic mutant (**right panel**) showed that there is no difference in transfection efficiency or defected nuclear localisation between the two proteins. The other phosphomimetic mutants (for detailed description, see **Results 2.4.**) exhibited no difference in these two factors compared to the unmodified protein.

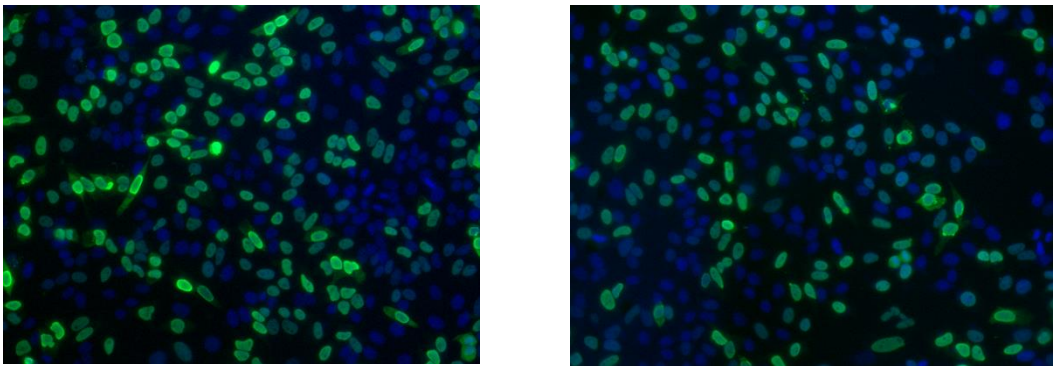


Figure 15. Transfection efficiency and nuclear localisation of HNF4α in HeLa cells transfected with a vector containing the wild-type (left) or the S313 mutant (right) protein. Images show nuclei (blue DAPI) and HNF4α protein (green) in merged images.

In the following, the results from my shared first-author publication will be described [133].

DOI:10.14753/SE.2018.2176

5.2.2. HNF4α phosphorylation by ERK1 *in vitro*

First of all, we performed *in vitro* phosphorylation assay of ERK1 on HNF4α. The assay was performed by Györgyi Vermes and Tamás Arányi. *In vitro* translated human recombinant HNF4α protein with N-terminal GST-tag, ERK1 kinase and radioactively labelled [γ-^{32}P] ATP was utilized in the assay. We ran the samples on SDS-PAGE and we analysed them with autoradiography. **Figure 16** shows that ERK1 kinase can be autophosphorylated, [134], however, it can also phosphorylate the HNF4α protein (see the band in the middle).

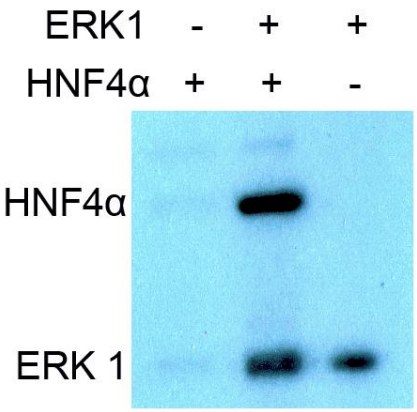


Figure 16. *In vitro* phosphorylation assay of HNF4α by ERK1. (modified from [133])

5.2.3. Specific phosphorylation sites of HNF4α phosphorylated by ERK1

Next, we intended to identify the phosphorylated serine/threonine residues. Thus, we cut the ERK1-phosphorylated, but not labelled HNF4α sample from the gel and mass spectrometry analysis was performed. We have found numerous phosphorylated amino acid residues, nevertheless, two adjacent sites could not be discriminated (**Figure 17** and **Table 4**). These phosphorylation sites could be found in the DNA binding domain, the hinge, the ligand-binding domain and also at the C-terminus. Interestingly, we did not find the phosphorylation site S87 of the human protein (corresponding to rat S78) in our

assay, in spite of previous reports [27]. In conclusion, ERK1/2 can indeed phosphorylate HNF4α at a number of previously described sites (S138/T139, S142/S143, S147/S148, S151, T166/S167, S313) and new ones discovered by us (S95, S262/S265, S451, T457/T459). Furthermore, the ERK1 targets the same positions as other kinases, for example PKA, p38 and AMPK (see **Introduction** and **Discussion**).

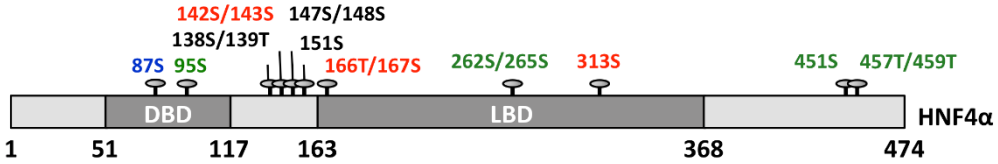


Figure 17. Phosphorylated sites on HNF4α protein by ERK1 kinase detected by mass spectrometry. Sites having an inhibitory effect on target gene transcription are indicated in red. Site 87S – which also has an inhibitory effect – was not identified here. Sites newly identified in this experiment are marked in green. DBD: DNA binding domain, LBD: ligand-binding domain. (modified from [133])

Table 4. Phosphorylated amino acid residues of HNF4α identified by mass spectrometry. DBD: DNA binding domain, LBD: ligand-binding domain. (modified from [133])

PHOSPHORYLATION SITES IDENTIFIED	PART OF HNF4α
S95	DBD
S138/T139	hinge
S142/S143	hinge
S147/S148	hinge
S151	hinge
T166/S167	LBD
S262, S265	LBD
S313	LBD
S451, T457/T549	C-terminus

DOI:10.14753/SE.2018.2176

5.2.4. Phosphorylation site(s) with inhibitory effect on target gene transcription

Next, we were interested which phosphorylation site might have an effect on target gene transcription. Therefore, five selected phosphorylation sites were examined in luciferase reporter gene assay. I designed mutations for either serine or threonine phosphorylation sites resulting in phosphomimetic (glutamate or aspartate) or neutral (alanine) mutants: S87D, T166A/S167D, S313D, S451E, T457A/T459E and S451E/T457A/T459E triple mutant. If two phosphorylation sites were next or close to each other, both were mutated. We changed only one site into a phosphomimetic mutation and mutate the other to a neutral one.

Serine 87 mutation was chosen as a positive control, because this site is a target of PKC, which suppresses HNF4α activity [27]. T166/S167 was described to be phosphorylated by p38α or p38β MAP kinases [23, 29]. The site S313D is targeted by AMPK phosphorylation [24]. Finally, the sites S451 and T457/T459 were newly identified by us, therefore, we intended to examine these C-terminal sites participating in transcriptional regulation of target genes.

After gene synthesis performed by TargetGenes biotechnology company, I co-transfected the different HNF4α mutants and luciferase reporter vector containing the promoter of the target gene *ABCC6* and into HeLa cells (which do not express HNF4α endogenously). I performed the luciferase assays. I normalized the results for the background noise, then for transfection efficiency by the control reporter vector. As shown on **Figure 18**, wild type HNF4α shows similar activity to T166A/S167D, S451E, T457A/T459E and S451E/T457A/T459E (data not shown). Thus, these sites – when phosphorylated – do not have an effect on target gene transcriptional activity. In contrast, both S87D (positive control) and S313D have significant effect on *ABCC6* promoter activity: they inhibit their activity. In summary, both ERK1 and AMPK target the phosphorylation site S313, which has an inhibitory effect on target gene transcription.

DOI:10.14753/SE.2018.2176

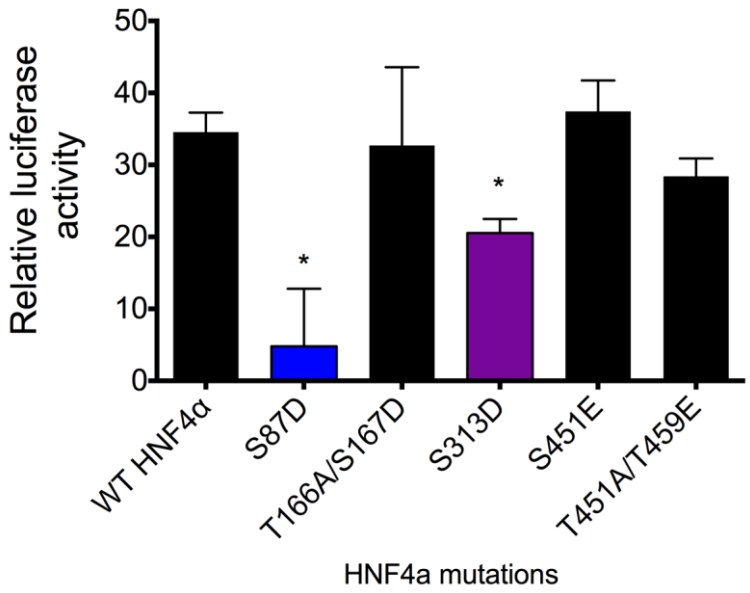


Figure 18. Luciferase assay measuring *ABCC6* promoter activity of HNF4α phosphomimetic mutants in HeLa cells. Mutations targeted selected phosphorylation sites. Triple transfection was performed with phACCC6(-332/+72)Luc promoter construct, pcDNA5-FRT/TO plasmid containing HNF4α mutants and pRL-TK Renilla control reporter vector. Luciferase activity was normalized for background noise and transfection efficiency. Tukey-HSD test was performed. S.D. is indicated. *p<0.05. (modified from [133])

5.2.5. Overlap between HNF4α and the active enhancer histone mark H3K27ac at genomic levels

In order to detect the active HNF4α binding sites at genomic level and select some target loci to examine the effect of ERK1, I performed chromatin immunoprecipitation followed by sequencing (ChIP-seq) with specific antibodies. In a parallel with an anti-HNF4α antibody in the HepG2 cells, I carried out an experiment with an antibody against acetylated histone 3 lysine 27 (H3K27ac). H3K27ac is a covalent epigenetic modification of the chromatin marking active regulatory regions, often enhancers. ChIP was followed by next generation sequencing. The bioinformatic analysis was performed by Dóra Bojcsuk. Altogether, 8748 transcription factor binding sites (TFBSs) could be identified for HNF4α. The overlap between loci bound by both HNF4α and H3K27ac was

remarkable, suggesting that a great number of HNF4α sites are active regulatory regions (**Figure 19**). Many of the HNF4α TFBSs are located near genes associated with PPAR and insulin signalling, fatty acid metabolism and ABC-transporters (**Table 5**).

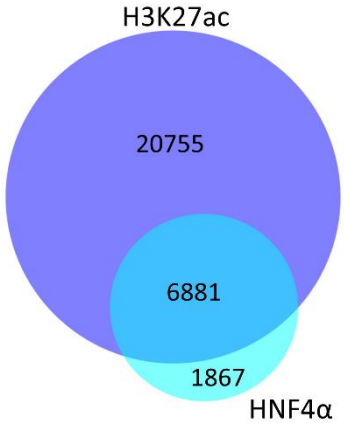


Figure 19. Venn diagram showing the number of HNF4α peaks (8,748) and H3K27ac signals and their overlap. (modified from [133])

Table 5. The most relevant biological pathways related to the HNF4α binding sites. Data from the KEGG database. (modified from [133])

Top 15 pathway terms	log P-value
Peroxisome	-11,593351
PPAR signaling pathway	-9,2559481
Insulin signaling pathway	-9,1645536
Glycine, serine and threonine metabolism	-8,616929
Primary bile acid biosynthesis	-7,1526082
ABC transporters	-5,5569263
Fatty acid metabolism	-5,5569263

For the subsequent experiments, I selected the following HNF4α target genes: *4-hydroxyphenylpyruvate dioxygenase* (*HPD*), *Pyruvate kinase, liver and red blood cell* (*PKLR*). We also intend to examine the *ABCC6* gene since we and others have shown

that HNF4α binds the *ABCC6* promoter [33, 135]. Our ChIP-seq results also revealed that *Apolipoprotein A1* (*APOA1*), another HNF4α target, is occupied by both HNF4α and H3K27ac, showing that it contains an active regulatory region. In addition, it plays a role in PPAR signalling according to our KEGG pathway analysis. Lastly, we have identified two targets of HNF4α – *Biliverdin A* (*BLVRA*) and *Biliverdin B* (*BLVRB*) -, which are closely connected to heme oxygenase in heme catabolism, which is essential in anti-oxidative and anti-inflammatory defence mechanisms. The negative control region was chosen to be the *β-globin* promoter, which lacks both H327ac and HNF4α binding (**Figure 20**).

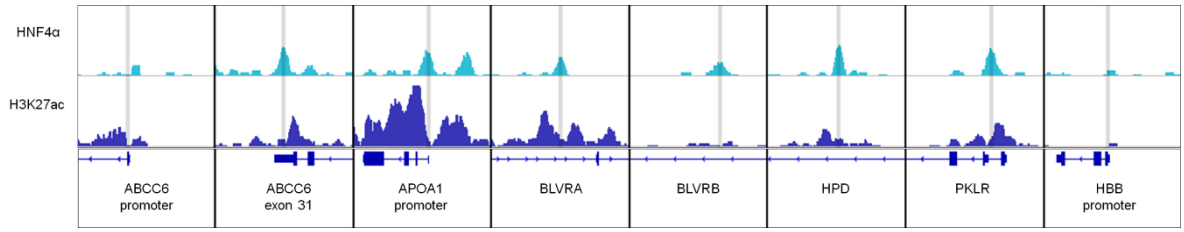


Figure 20. IGV snapshot of HNF4α peaks and H3K27ac ChIP-seq signals showing several HNF4α target genomic regions. ABCC6: ATP-binding cassette subfamily C, member 6; APOA1: Apolipoprotein A1; BLVRA, B: Biliverdin A, B; HPD: 4-hydroxyphenylpyruvate dioxygenase; PKLR: Pyruvate kinase, liver and RBC and HBB: Hemoglobin negative control region. (modified from [133])

5.2.6. Effect of extracellular activation of the ERK pathway on the binding of HNF4α to specific genomic regions

In order to investigate the effect of ERK activation on TF binding, I examined HNF4α binding to several selected genomic target regions upon ERK1/2 induction in HepG2 cells. Phosphorylation is a fast process, and it can happen in minutes [136], therefore, I performed short-term (30 minutes) treatment with epidermal growth factor (EGF), which activates the ERK1/2 signalling pathway as an extracellular ligand.

I examined enrichment of target fragments immunoprecipitated by anti-HNF4α antibody by ChIP-qPCR, in relation to either input (whole fragmented chromatin) (**Figure 21A**) or

negative control region (β-globin) (**Figure 21B**), where HNF4α does not bind. 30 minutes treatment diminished HNF4α binding to its target sites.

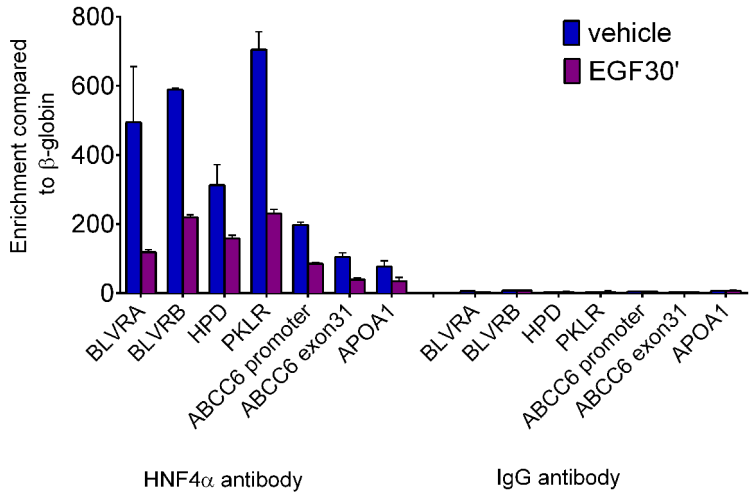


Figure 21A. HNF4α occupancy on hepatic HNF4α target genomic regions in HepG2 cells compared to β-globin. ChIP-qPCR results are shown from a representative experiment. (N=7) Enrichment was compared to % input. HepG2 cells were treated with vehicle (blue) or EGF (purple) for 30 minutes. BLVRA, B: Biliverdin A, B; HPD: 4-hydroxyphenylpyruvate dioxygenase; PKLR: Pyruvate kinase, liver and RBC; ABCC6: ATP-binding cassette subfamily C, member 6 and APOA1: Apolipoprotein A1. S.D. is indicated.

In the experiment presented on **Figure 21**, I used HepG2 cells treated with either vehicle or EGF for 30 min. I prepared chromatin from each sample (1/condition/experiment). Altogether, I performed 7 independent experiments. In each experiment, in order to investigate the chromatin binding of HNF4α upon EGF treatment, I performed ChIP-qPCRs on 7 target genomic regions known to bind HNF4α in HepG2 cells under control conditions. (However, not all of the genes were tested in all experiments, only 45 tests were run instead of 7x7=49.) Short EGF treatment was proved to be effective in most of the cases: short EGF treatment leads to a loss of approximately 50-80% of HNF4α binding compared to the vehicle treatment.

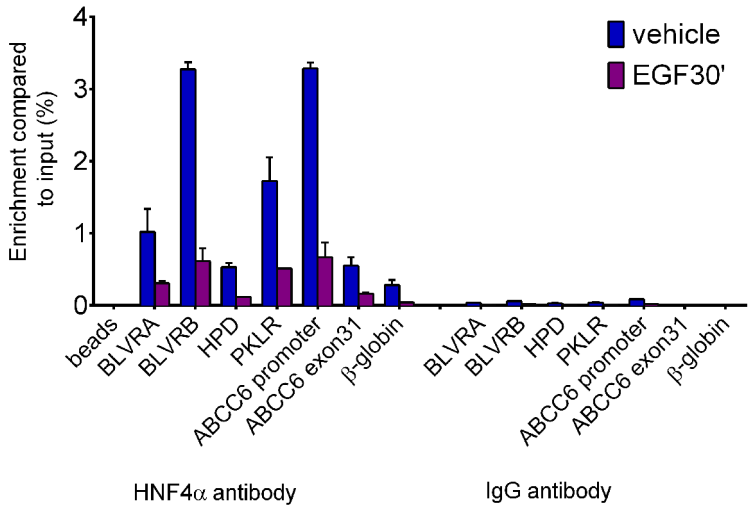


Figure 21B. HNF4α occupancy on hepatic HNF4α target genomic regions in HepG2 cells compared to % input. ChIP-qPCR results are shown from a representative experiment. (N=7) Enrichment was compared to % input. HepG2 cells were treated with vehicle (blue) or EGF (purple) for 30 minutes. BLVRA, B: Biliverdin A, B; HPD: 4-hydroxyphenylpyruvate dioxygenase; PKLR: Pyruvate kinase, liver and RBC; ABCC6: ATP-binding cassette subfamily C, member 6; APOA1: Apolipoprotein A1 and Beta globin: negative control region. S.D. is indicated. (modified from [133])

I also performed statistical tests to demonstrate the effect of EGF. I used the average of the technical qPCR duplicates, then I normalized each measurement of the EGF-treated group to its respective control group. Then I performed a one-sample t-test, where the null hypothesis was that the population mean is equal to 1. I observed a highly significant ($p<0.004$) decrease of HNF4α chromatin binding. I used the same approach to test the different target regions separately, as well (not shown). In conclusion, ERK1/2 phosphorylates HNF4α, resulting in reduced HNF4α DNA-binding capacity to target sequences.

5.3. Effects of short-term nutritional stress

As forecasted in the **Objectives** section, we intended to answer to following question:

III. What are the macroscopic, protein level and methylation changes occuring upon short-term fasting and refeeding *in vivo*?

In answer for the above mentioned question, the results of the fasting-refeeding experiments will be summarized and discussed. These results are not yet published and the project has not been completely finished yet. As described in detail in the **Methods** section, the following groups were formed for the investigations (**Figure 10 in Methods**). We were interested in the short-term effects of acute nutritional challenges (fasting and refeeding) *in vivo* happening in mouse liver. Indeed, we have observed vast changes in the physiology and metabolism of the mice. We intended to characterize the changes on different levels and find mechanisms that drive them. Therefore, we have examined 3 layers:

i) **macroscopic changes** (changes in body weight and blood glucose level),
ii) **protein level changes** (metabolic enzyme or transcription factor abundance kinetics) and
iii) **changes in genome-wide methylation** (methylome) measured by two independent methods

Firstly, macroscopic changes - the physiological consequences of fasting and refeeding - will be considered. Handling food for the mice (fasting and refeeding) were done by me, Tamás Arányi and Flóra Szeri. Animal experiments - including liver perfusion - were performed by Flóra Szeri, who was aided by me and Tamás Arányi. Experiments in relation to macroscopic changes were performed by me, Flóra Szeri and Tamás Arányi.

5.3.1. Weight and blood glucose comparison of groups

5.3.1.1. Weight comparison among groups

We measured the body weight of all animals from each group. The average weight of all control groups was above 21 g. Fasting for 8 hours did not, but fasting for 16 and 24 hours lowered body weight compared to their control group (Student's t-test, $p<0.05$) (**Figure**

22). Refeeding for 8 hours could not restore body weight. Body weight of the 4 hours' refed and its control group does not seem to be statistically different, but it can be attributed to the high SD of the 2 group averages. In conclusion, both fasting and refeeding have a dramatic effect on the body weight of the mice.

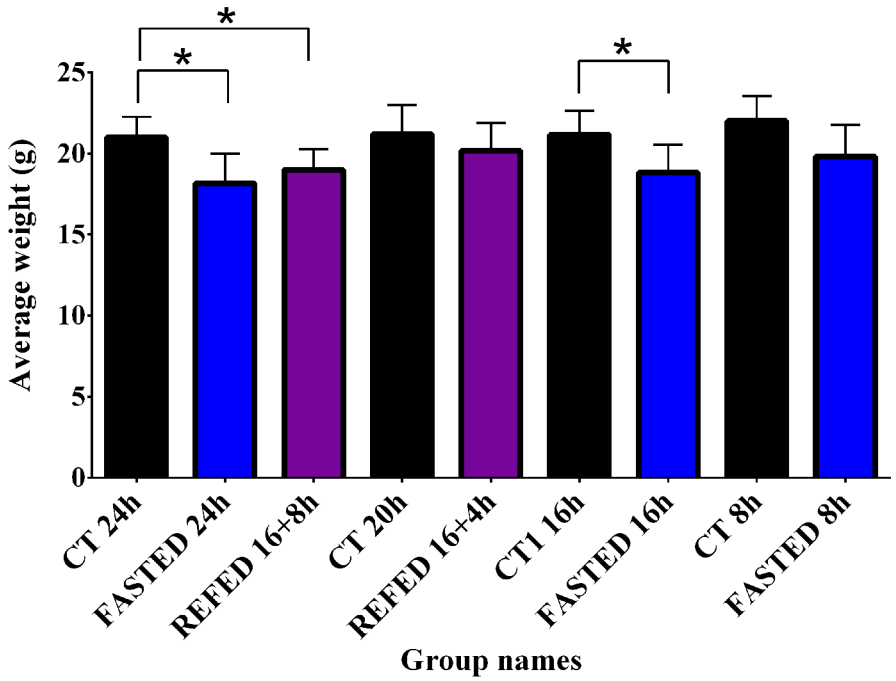


Figure 22. Average of body weight (g) of groups before sacrifice. Control groups are indicated with black, groups undergoing fasting are indicated with blue and groups undergoing refeeding are indicated with purple columns, respectively. Durations of fasting and refeeding are indicated in the group names. SD. * $p<0.05$. N=6.

5.3.1.2. Blood glucose comparison among groups

We also measured the blood glucose levels of all animals from each group. The average blood glucose levels were similar among the control groups, between 7,5 and 8,5, although the blood glucose level of the group sacrificed at 2 p.m. was significantly higher than that of 2 a.m. (**Figure 23**). Furthermore, fasting (8h, 16h, 24h) drastically lowered blood glucose levels (Student's t-test, $p<0.05$) (**Figure 24**), however, the duration of

fasting did not have an effect. Refeeding for 8 hours daytime could restore blood glucose levels.

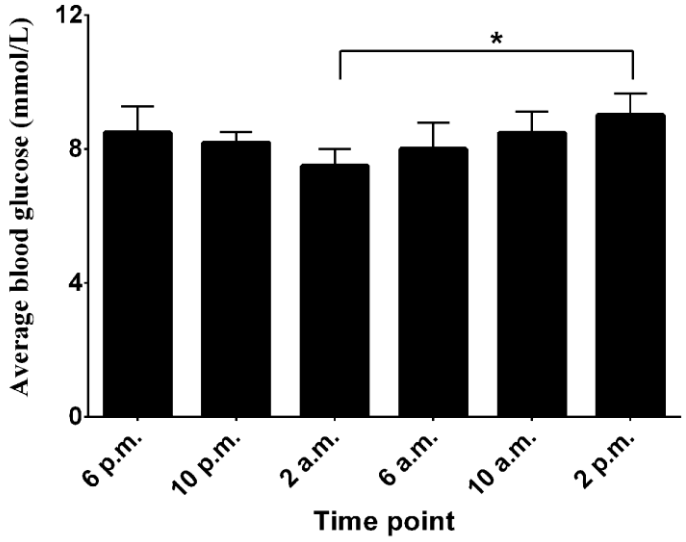


Figure 23. Average of blood glucose levels (mmol/L) of the CT groups. Animals fed *ad libitum* were sacrificed at every 4 hours indicated in the group names. SD. * p<0.05.

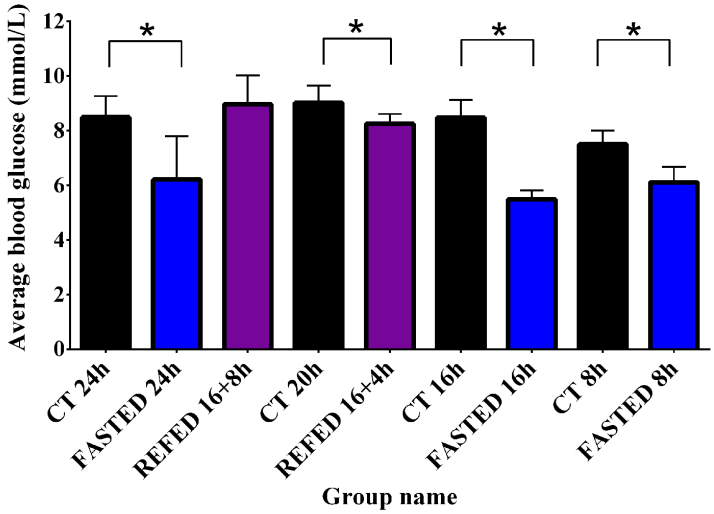


Figure 24. Average of blood glucose levels (mmol/L) of groups before sacrifice. Control groups are indicated with black, groups undergoing fasting are indicated with blue and groups undergoing refeeding are indicated with purple columns, respectively. Durations of fasting and refeeding are indicated in the group names. SD. * p<0.05. N=6.

DOI:10.14753/SE.2018.2176

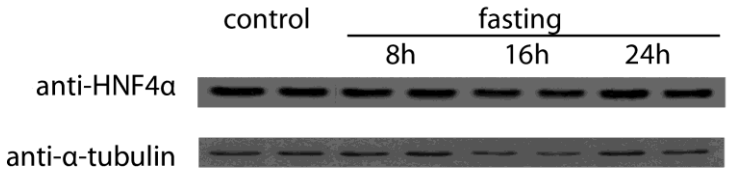
5.3.2. Protein level changes

Secondly, we investigated changes of protein levels. We were interested in different proteins playing an important role in metabolic adaptation of the liver to acute environmental stress, for instance short-term fasting. These are either metabolic enzymes closely related to carbohydrate or glucose metabolism (e.g. gluconeogenesis) or transcription factors involved in responding to nutritional stress in the liver (e.g. HNF4α as discussed above). We have performed Western blot analyses with 4 parallel samples on different proteins. In addition, densitometry and two-tailed T-test was performed on the samples. Western blot experiments were done by Kitti Koprivanacz, Metta Dülk and Ágnes Sárközi.

5.3.2.1. HNF4α protein levels

The role and mechanism of action of HNF4α as a master metabolic regulator in hepatocyte has been described above. Fasting is known to disrupt glucose homeostasis, since HNF4α has a prominent role in glucose metabolism. Therefore, we were interested if the protein levels of the transcription factor change upon fasting.

Our experiments have revealed that the protein level of HNF4α does not change significantly upon acute metabolic stress, i.e. 8 hours', 16 hours' and 24 hours' fasting (**Figure 25**).



DOI:10.14753/SE.2018.2176

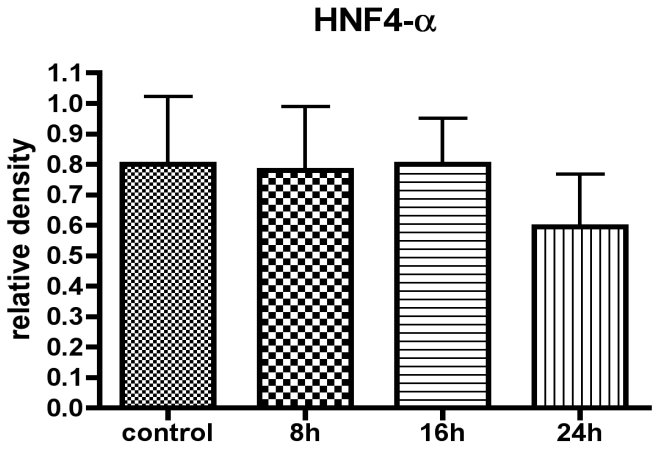


Figure 25. Western blot analysis of the HNF4α and the α-tubulin proteins and its analysis by densitometry. α-tubulin served as loading control. HNF4α is a 53 kDa protein, α-tubulin is a 50 kDa protein. The duration of fasting is indicated in the name of the groups.

5.3.2.2. CEBPα protein levels

It has been reported that fasting induces CEBPα [126], therefore we also investigated the amount of this protein in physiological and fasting conditions (**Figure 26**). Fasting for 24 hours significantly elevated CEBPα protein levels.

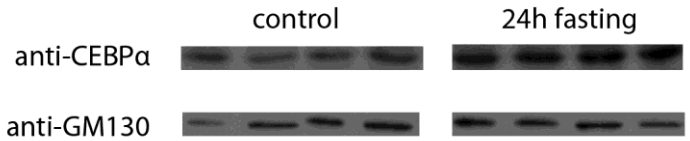


Figure 26. Western blot analysis of the CEBPα and the GM130 proteins. Golgi marker (GM) 130 served as loading control. CEBPα is a 55 kDa protein, GM130 is a 130 kDa protein. The duration of fasting is 24 hours.

5.3.2.3. PCK1 protein levels

Furthermore, PCK1 is a well-known enzyme and key player in gluconeogenesis. Therefore, we examined the protein level changes upon fasting (**Figure 27**). Fasting for 16 hours significantly elevated PCK1 protein levels.

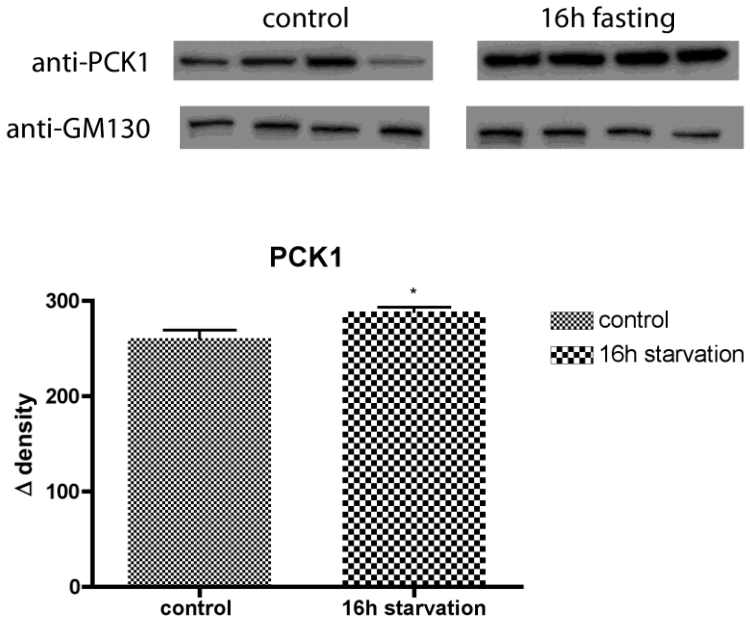


Figure 27. Western blot analysis of the PCK1 and the GM130 proteins and its analysis by densitometry. Golgi marker (GM) 130 served as loading control. PCK1 is a 72 kDa protein, GM130 is a 130 kDa protein. The duration of fasting is 16 hours.

5.3.3. Analysis of sequencing data

Thirdly, we hypothesized that short-term nutritional stress can cause changes in DNA methylation. Since the long-standing conception that DNA methylation is stable has been disproved, interest has been thriving in investigating the dynamic nature of DNA methylation. Rapid DNA methylation changes can occur, for example in human cell lines [36] or as a response to environmental stress factors [37]. Here, we hypothesized that nutritional stress can cause vast methylation changes, as well. Furthermore, we intended to characterize the genome-wide methylation changes. We have investigated the methylation changes occurring upon 16 hours' fasting ('FASTED 16h') and 16 hours' fasting followed by 8 hours' refeeding ('REFED 16+8h') in order to explore the effect of short-term nutritional challenge on global and site/region-specific methylation levels. For

the investigations, we used the RRBS technique. Genomic DNA extraction from the mouse livers and RRBS was performed by me.

5.3.3.1. Sequencing statistics: read number and coverage

In the following, the bioinformatic analyses of the RRBS libraries will be described. These investigations were performed by Piroska Dévay, István Likó, Ábel Fóthi and Csenge Halász.

Firstly, sequencing data filtering was performed, the main steps of which are summarized on **Figure 28**. For the further analyses, the minimum read number (coverage) for next sequencing data was set for 10. Finally, there were approximately 200,000 CpGs at the end which were present in 4 samples out of 6 in all the 4 groups investigated: 'CT 16h', 'FASTED 16h', 'CT 24h' and 'REFED 16+8h'.

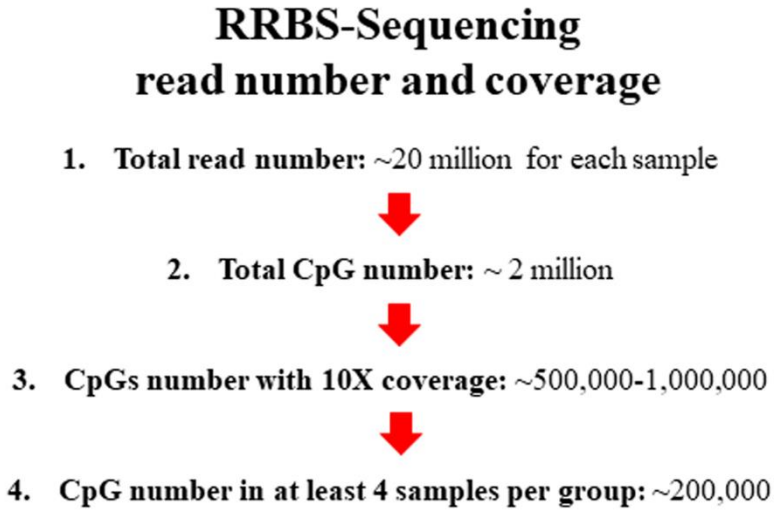


Figure 28. RRBS-sequencing read number and coverage.

5.3.3.2. Histogram of % CpG methylation and CpG coverage

Next, the histogram of % CpG methylation and CpG coverage was plotted. One typical example is shown on **Figure 29**. Most of the CpGs have originally very low methylation

(0-5%) or very high methylation (95-100%) (top panel). This implies that CpGs tend to be either fully methylated or unmethylated. The bottom panel illustrates CpG coverage. In this typical sample, a big portion of the CpGs have 10 reads, however, there is a remarkable number of CpGs with around 10-40 reads. There are some CpGs with as high as 100 reads.

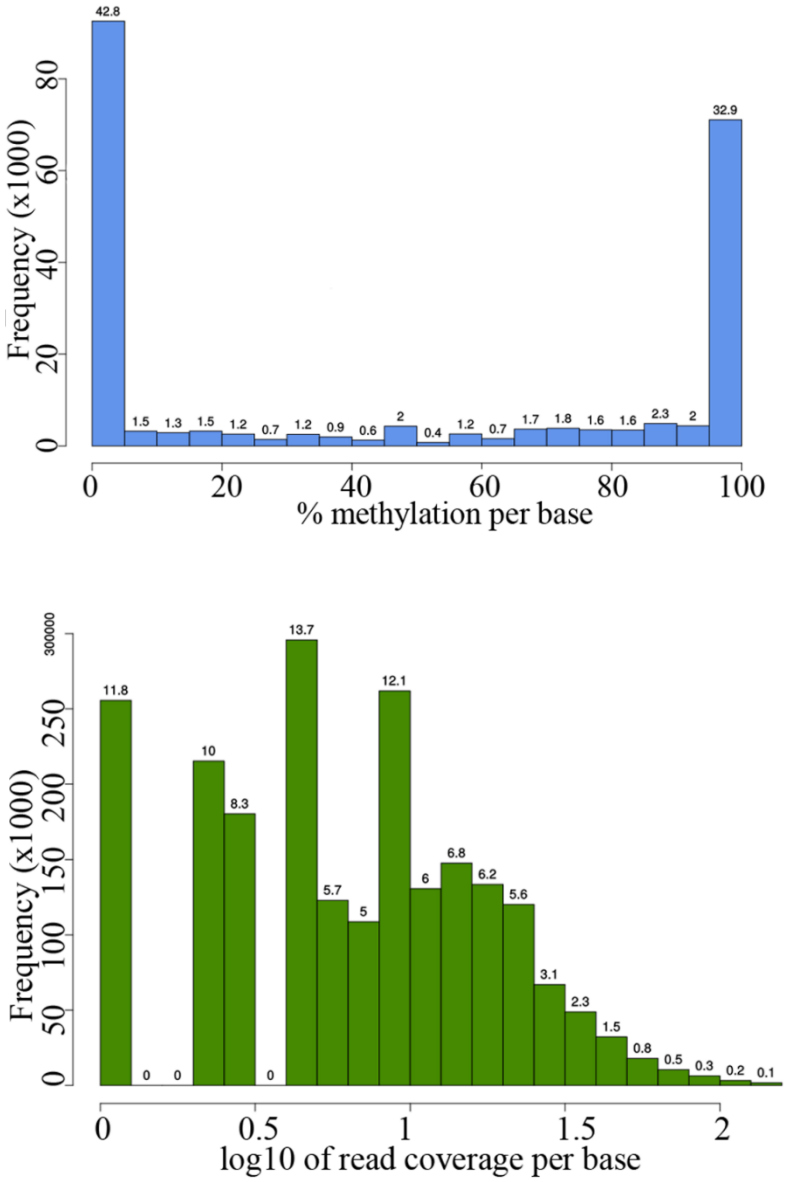


Figure 29. Histogram of % CpG methylation and histogram of CpG coverage.

5.3.4. Analysis of methylation % distributions

Furthermore, methylation percentages were calculated. The average methylation for the 4 groups is plotted on **Figure 30**. We have found that the most diverse group is the 'REFED 16+8h' group; some samples have high and others have low methylation. This suggests that the 6 animals respond to refeeding very differently in terms of methylation. Altogether, there is no significant difference of average methylation among the groups, although the fasted group has almost significantly higher mean methylation compared to its control group (p=0.053), and the refed group has almost significantly lower mean methylation compared to its control group (p=0.059).

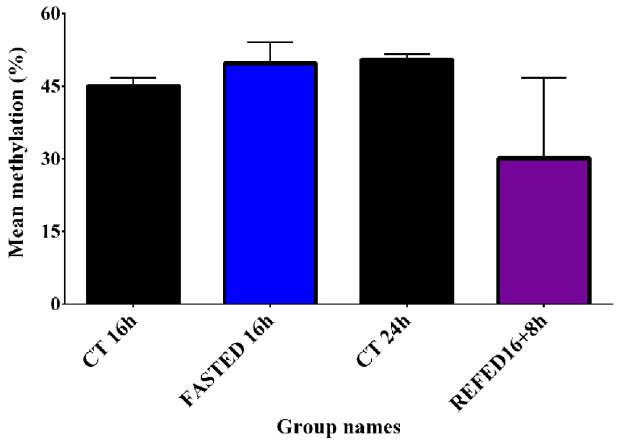


Figure 30. Average of mean methylation % in the 4 groups.

We intended to analyse the methylome sequencing data in more detail, thus we have performed several types of bioinformatic analysis on DNA methylation analyses concerning specificity and sensitivity. We intended to capture all the possible methylation changes without finding any false positives. Therefore, we first applied a stringent method. The results of this stringent approach will be discussed in the following.

5.3.5. Analysis of differential methylation with the stringent analysis

In order to capture the loci of the most important methylation changes, we investigated both differentially methylated individual CpG sites (**DMS**s) and differentially methylated regions (**DMR**). On the one hand, the DMS analysis enables to specifically point out the exact CpGs in the genome where the most relevant methylation changes occur. However,

it is difficult to connect unambiguously the methylation change of an individual C to gene expression change, therefore, it is challenging to draw a parallel between methylation and transcription. On the other hand, in the DMR analysis, we have determined a fixed length for a genomic region in which there must be a minimum number for CpGs changing. In addition, we have investigated promoters. In this manner, the DMR analysis provides information on several adjacent CpGs.

5.3.5.1. Number of differentially methylated sites (DMSs)

For this very stringent, but very specific analysis, we have created a common pool of investigated CpGs, which are present in 4 animals out of 6 in all the 4 groups. Altogether, we investigated 208,031 (not overlapping) CpGs in total and this pool served as the background (all). After calculating differential methylation (methylation change compared to the original methylation level) for individual sites, the number of DMSs of the 4 different comparisons can be observed on **Table 6**.

Table 6. Number of hypo- and hypermethylated CpGs upon fasting and refeeding. DM: differential methylation. CF: control vs fasting. FR: fasting vs refeeding. CR: control vs refeeding. CC: control vs control.

DM direction	CF_hyper	FR_hyper	CR_hyper	CC_hyper
Number	470	720	266	227
DM direction	CF_hypo	FR_hypo	CR_hypo	CC_hypo
Number	221	2101	612	160

As it is striking from the comparisons, fasting resulted in almost 500 CpGs undergoing hypermethylation, whereas less than half undergoing hypomethylation. In contrast to fasting, refeeding had the opposite effect on methylation change: we found more hypermethylated sites than hypomethylated (**Table 5**). Controls did not show difference in the number of hyper- and hypomethylated sites. Thus, 16 hours' overnight fasting resulted in global hypermethylation, while 16 hours' overnight fasting followed by 8 hours' refeeding lead to global hypomethylation. However, it is necessary to emphasize

that both hyper- and hypomethylation happen at the same time in the same liver cells both upon fasting or refeeding.

5.3.5.2. Methylation differences and q values of CpGs

The changing (and not changing) CpGs can be visualized on volcano plots, as well (**Figure 31**). The red dots above the horizontal black lines on the panels represent significantly changing CpGs (q=0.01), which were further analysed. The range of colours shows the count of CpGs. The vast majority of CpGs have around 0% methylation difference. Moreover, the methylation differences are relatively small; there are only a few CpGs with more than 50% methylation change. As it is evident from the volcano plots, there was more hypermethylation than hypomethylation happening in the livers of the fasting mice (left panel). In contrast, clear and conspicuous hypomethylation was present in the samples of mice that underwent refeeding (right panel).

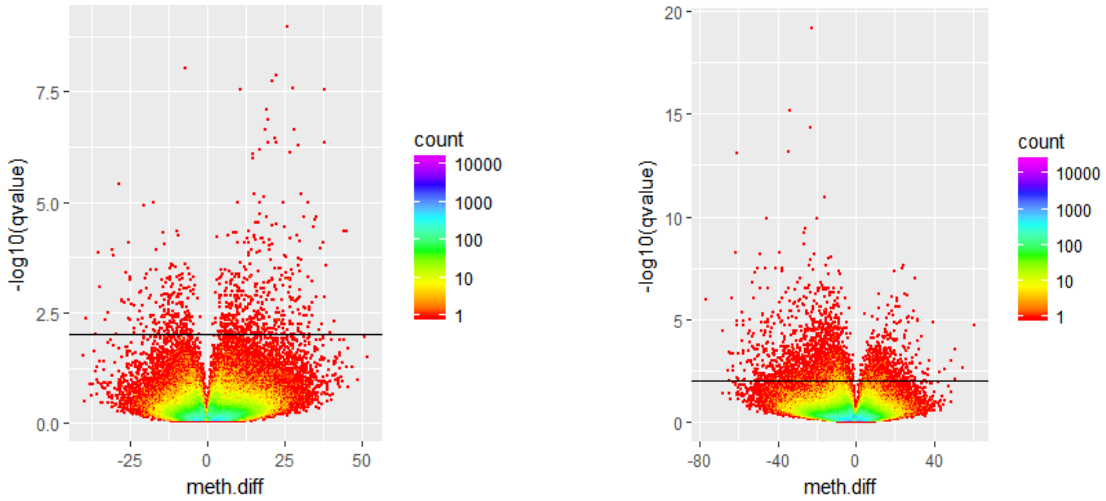


Figure 31. Volcano plots for the comparisons CF (control vs fasting) and FR (fasting versus refeeding). Each changing CpG is plotted based on its methylation difference % and its q value (corrected p value). The range of colours shows the number of CpGs. The horizontal line represents the threshold of statistical significance.

5.3.5.3. CpG distribution around CpG islands

In the following, CpG distribution around CpG islands was investigated (see **Figure 32**). CpG islands are short stretches of CG-rich sequences often located upstream from the TSS, CpG shores are the 2000bp flanking region on each side of CpG islands. Background ('ALL') illustrates all the CpGs present in all the samples (**left panel**). Regarding the distribution of hypermethylation, CpG islands and CpG shores are underrepresented (**middle panel**), but the regions beyond them are overrepresented. Refeeding is characterized by the most distinct pattern from the background for DNA methylation change. Concerning the distribution of hypomethylation, CpG islands are mildly affected and CpG shores are overrepresented compared to the background (**right panel**). In conclusion, hypermethylation occurs mainly outside CpG islands and shores, whereas hypomethylation greatly affects CpG shores. Although, both fasting and refeeding are characterized by methylation changes in the two opposite directions, the distributions of them are unambiguously distinct.

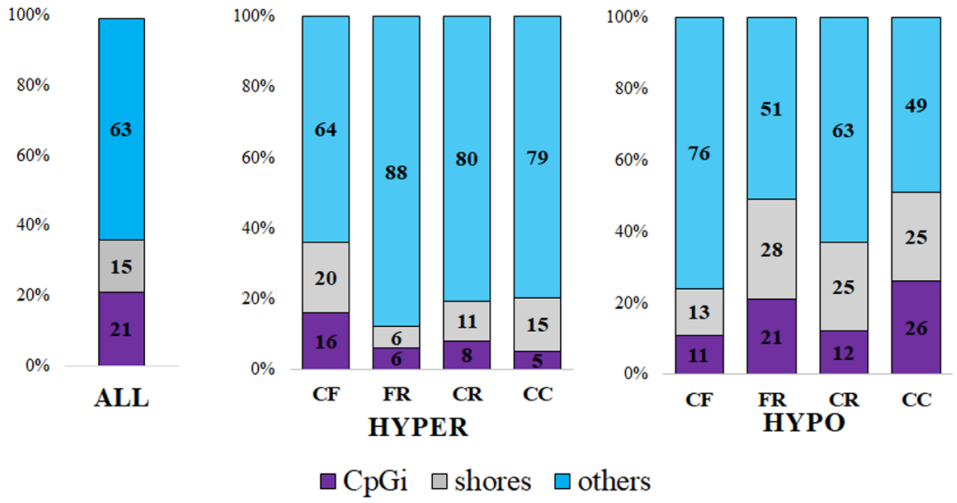


Figure 32. Column representation of CpG distributions around CpG islands for the CpGs undergoing hypo- or hypermethylation upon fasting and refeeding. All: All the CpGs investigated. CF: control vs fasting. FR: fasting vs refeeding. CR: control vs refeeding. CC: control vs control.

5.3.5.4. CpG distributions around genes

The CpG distributions around different regions of a gene can be observed on **Figure 33**. Background ('ALL') illustrates all the CpGs present in all the samples (**left panel**). Regarding the distribution of hypermethylation, promoters and exons are mildly underrepresented. At the same time, introns and intergenic regions are overrepresented (**middle panel**). Refeeding is characterized by the most distinct pattern from the background for DNA methylation change. Concerning the distribution of hypomethylation, DNA methylation changes primarily affect promoters compared to the other regions. Moreover, fasting and refeeding can be compared, as well. Upon fasting, the distribution of the CpGs does not change substantially. However, refeeding leads to profound changes in the distribution of differential methylation compared to the background.

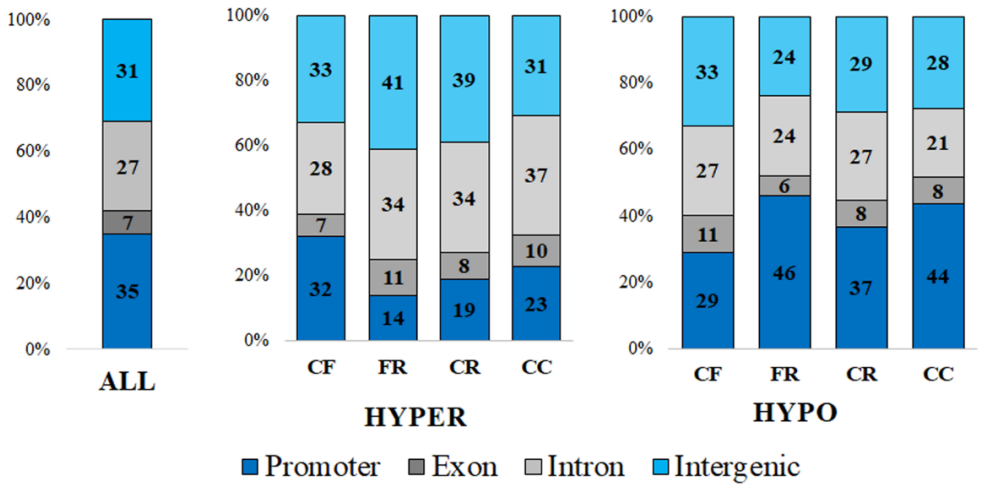


Figure 33. Column representation of CpG distributions around genes for the CpGs undergoing hypo- or hypermethylation upon fasting and refeeding. All: All the CpGs investigated. CF: control vs fasting. FR: fasting vs refeeding. CR: control vs refeeding. CC: control vs control.

5.3.5.5. CpG distributions and proximal and distal promoters

In addition, the default setting of the promoter definition set by the MethylKit needed some revision. Originally, the promoters were defined as +/-1000 bp from the TSS (termed as 'promoter' on **Figure 34**). However, we hypothesized that the annotation of a

CpG to the intergenic region might include some distal promoters, thus - by modifying the distances -, we introduced the definition of distal and proximal promoters, as well (**Figure 34**). Doing so, from the investigated 200,000 CpGs, 20% of them were localized in proximal and 10% in distal promoters. Moreover, 50% of the CpGs previously defined as intergenic overlapped with the CpGs of distal promoters.

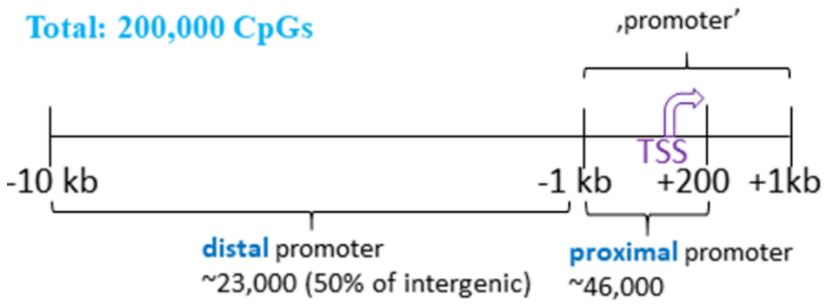


Figure 34. Number of CpGs in proximal and distal promoters. TSS: Transcription start site. Total number of investigated CpGs: 200,000. CpGs in proximal promoter: 20%. CpGs in distal promoter: 10%, 50% of CpGs previously defined as intergenic.

From the approximately 20,000 CpGs in distal promoter, there were altogether almost 600 DMSs (3%) (**Table 7**). Interestingly, more changes were observable for DMSs in distal promoter regions (which can be overlapping with intergenic regions) than for DMSs in distal promoter regions but outside intergenic regions. This suggests that the intergenic regions are indeed important targets of DNA methylation change.

Table 7. Number of DMSs in distal promoters and not intergenic regions undergoing hypo- or hypermethylation upon fasting and refeeding. CF: control vs fasting. FR: fasting vs refeeding. CR: control vs refeeding. CC: control vs control.

		CF	FR	CR	CC
All DMSs in distal promoters		101	339	104	43
Distal promoters	hyper	74	100	30	28
	hypo	27	239	74	15
Not intergenic	hyper	26	46	12	14
	hypo	12	88	32	5

DOI:10.14753/SE.2018.2176

5.3.5.6. CpG distributions of annotated DMSs

Furthermore, we investigated the CpG island distributions of the CpGs with the above mentioned gene annotations. We were interested if there was a category (promoter, exon, intron or intergenic) where the DMSs have significantly lower of higher representation of CpG islands or shores compared to the background.

Firstly, we found low representations for promoter DMSs in CpG islands, as shown on **Figure 35**. Regarding CpGs that underwent hypomethylation upon fasting and were annotated to promoters, only a limited number of them were in CpG islands or CpG shores compared to the background. Concerning hypermethylated CpGs, there were almost as many changes in CpG shores as in the background, but still, CpG-poor regions were affected more, similarly to the case of hypermethylation.

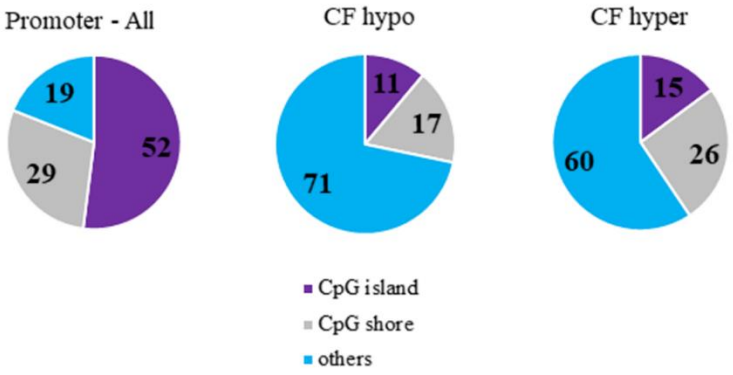


Figure 35. Piecharts for promoter CpGs around CpG islands undergoing hypo- or hypermethylation upon fasting. CF: control versus fasting. All: all the CpGs in the CF comparison. Others: outside CpG islands and shores.

Secondly, we found high representations for intron- and intergenic-annotated DMSs in CpG islands, as shown on **Figure 36**. CpGs that were hypomethylated upon fasting and were located in either introns or intergenic regions were remarkably present in CpG islands or shores, the distribution of which is appreciably different from the background.

DOI:10.14753/SE.2018.2176

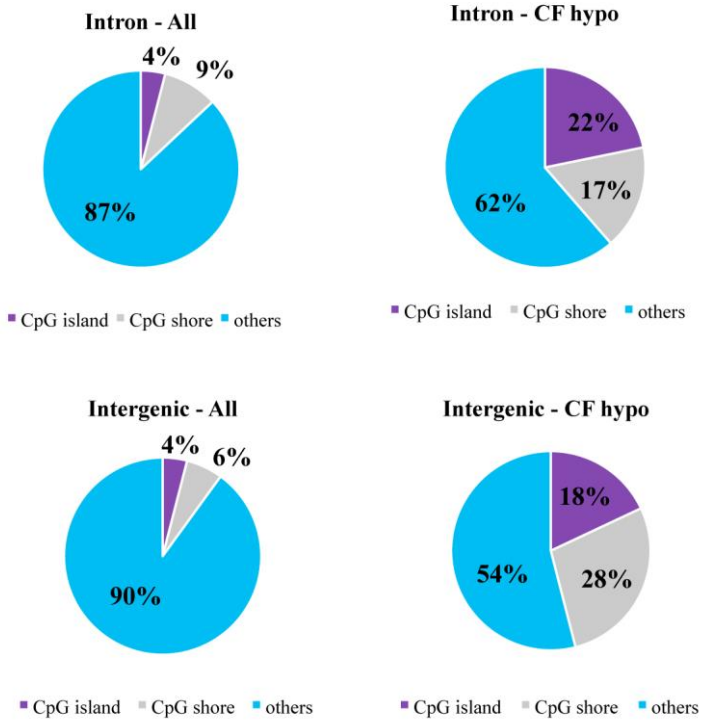


Figure 36. Piecharts for intronic and intergenic CpGs around CpG islands undergoing hypomethylation upon fasting. CF: control versus fasting. All: all the CpGs in the CF comparison. Others: outside CpG islands and shores.

5.3.6. Analysis of differential methylation with the more relaxed analysis

Apart from the stringent analyses, we have tried a more relaxed analysis in order to capture more changes. On the whole, we observed again marked hypermethylation upon fasting and remarkable hypomethylation upon refeeding. With this analysis, we could find much more significant CpGs.

5.3.6.1. 'Reversed' CpGs

Next, we were interested which genes are the ones that have a specific CpG changed in one direction upon fasting and this exact CpG changed in the opposite direction upon refeeding. I call them 'reversed' CpGs. Altogether, we found altogether approximately 2000 CpGs which underwent hypermethylation upon fasting and hypomethylation upon

refeeding. Contrary to this, almost 4 times more CpGs were hypomethylated upon fasting and hypermethylated upon refeeding.

Next, we annotated these CpGs to genes. In order to see which pathways are implicated in these methylation changes, I used the Panther Classification system (www.pantherdb.org). I found several metabolic pathways, for instance cholesterol, fatty acid, phospholipid, amino acid and carbohydrate (gluconeogenic and glycogen metabolic) metabolic processes. The most interesting examples are shown on **Table 8**.

Table 8. Metabolic pathways and genes containing CpGs with 'reversed' methylation status upon fasting and refeeding. CpGs in the C<F>R category are hypermethylated upon fasting and hypomethylated upon refeeding. CpGs in the C>F<R category are hypermethylated upon fasting and hypomethylated upon refeeding.

PATHWAYS	C<F>R	C>F<R
1. Lipid metabolism	**Cholesterol metabolism**	
	Insulin-induced gene 2 protein (Insig2)	
	Fatty acid biosynthesis	
	Fatty acid synthase (Fas)	
	Fatty acid beta-oxidation	
	Peroxisomal acyl-coenzyme A oxidase 1 (Acox1)	
2. Amino acid metabolism	***Glutamine synthetase (GluI)***	***Carnitine O-palmitoyltransferase 1 (Cpt1b)***
	Acetoacetyl-CoA synthetase (Aacs)	
3. Monosaccharide metabolic process	**Gluconeogenesis**	**Gluconeogenesis**
	Fructose-1,6-bisphosphatase 1 (Fbp1)	***Glucose-6-phosphatase3 (G6pc3)***
4. Tricarboxylic cycle	***Isocitrate dehydrogenase (Idh3)***	

DOI:10.14753/SE.2018.2176

5.3.6.2. Analysis of differentially methylated regions (DMR)

We intended to investigate in more detail those CpGs and the genomic environment that are in close proximity to the CpGs significantly changing based on the more stringent analysis. Therefore, we performed DMR analyses.

Firstly, we defined the DMR as +/-1000 bp region from the TSS with a minimum of 3 CpGs in order to gain information about the promoters of the genes which change upon nutritional stress. Applying this method, we have found that *Forkhead box O3* (FoxO3) has a region in its promoter which is hypomethylated upon fasting. The promoter of *Histone deacetylase 1 (Hdac1)* is also hypomethylated upon fasting. In contrast, a promoter region in *Acyl-CoA thioesterase 2* (Acot2) and the epigenetic modifier enzyme *Isocitrate dehydrogenase 3* (Idh3a) is hypermethylated upon fasting.

Moreover, we identified a promoter region in *Acot3* with 20-30% hypomethylation upon refeeding. In addition, a promoter region in *Retinoid X receptor β* and *γ* (RXRβ and γ) and liver-expressed *Cpt1a* was also hypomethylated. (Interestingly, Cpt1a was already hypomethylated upon fasting, but the q value did not reach the threshold of significance.) Concerning hypermethylation, a promoter region of *Hdac10* was identified with more than 10% methylation change. The results are summarized in **Table 9**.

Table 9. Metabolic genes with hypo- or hypermethylated CpGs in their promoter regions of upon fasting or refeeding. DMR analysis was performed with promoter regions +/- 1000 bp from the TSS with at least 3 CpGs changing upon fasting or refeeding. FoxO3: Forkhead box O3, Hdac1: Histone deacetylase 1, Acot: Acyl-CoA thioesterase, Idh3a: Isocitrate dehydrogenase 3a, RXR: Retinoid X receptor and Cpt1: Carnitine O-palmitoyltransferase 1.

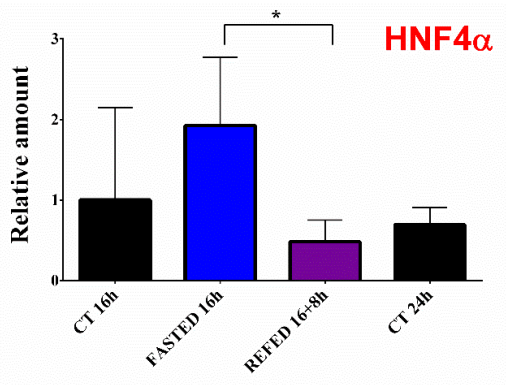
FASTING		**REFEEDING**	
hypo	hyper	hypo	hyper
FoxO3	*Acot2*	*Acot3*	*Hdac10*
Hdac1	*Idh3a*	*RXRβ, γ*	
		Cpt1a	

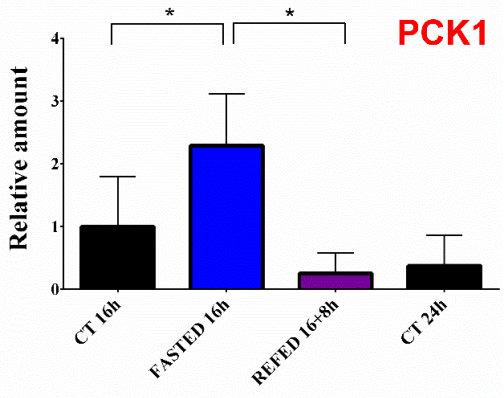
Secondly, we defined the DMR as a region where there are 2 CpGs changing in the same direction within 100 bp in gene promoters. Altogether 16.000 regions were identified, which corresponded to 450 genes. We found that there are DMRs in the metabolic genes *Insig2* and *Acot11* which are hypomethylated upon refeeding.

In conclusion, several metabolic genes were found to be changing their methylation status in their promoters in the direction of either hypo- or hypermethylation. Therefore, we were interested if the methylation changes observed by the RRBS technique are CpG-specific and enriched at CpG-rich regions or they characterize any Cs in the genome. For this, we performed LC-MS/MS, which investigates all the Cs and 5mCs in the genomic DNA.

5.3.6.3. Expression analysis of metabolic genes

I selected four genes for mRNA expression analysis: HNF4α, PCK1, G6P and FAS, since HNF4α is a master regulator in glucose homeostasis and the other enzymes exhibited changes in protein level or methylation upon fasting and refeeding. PCK1 and G6P are key players of gluconeogenesis and FAS has a role in fatty acid synthesis (see the **Introduction** section). The expression of HNF4α is lowered upon refeeding compared to fasting. The expression of PCK1 is elevated upon fasting, but decreased upon refeeding. Moreover, G6P expression is significantly lowered upon refeeding and refeeding cannot restore FAS expression to the value of its respective control group (**Figure 37**) (Student's t-test, $p<0.05$).





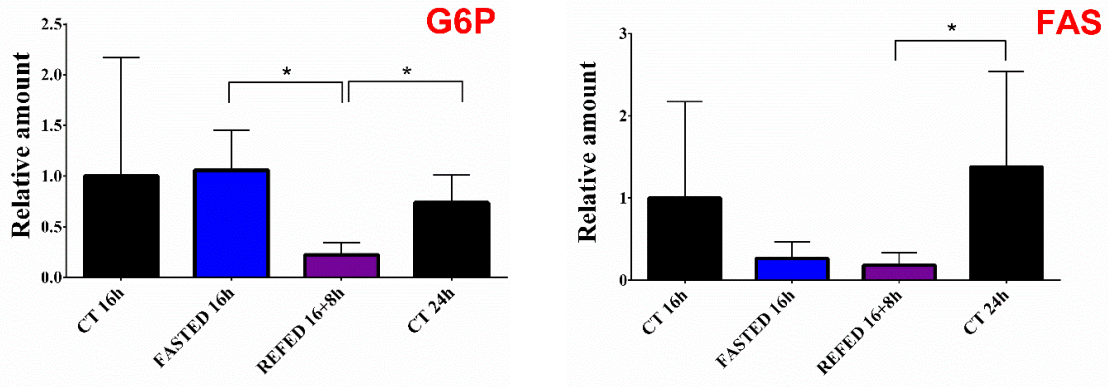


Figure 37. mRNA expression of HNF4α, PCK1, G6P and FAS. Relative amounts are normalized to the housekeeping 18S RNA. CT 16h is normalized to 1. HNF4α: Hepatocyte nuclear factor 4 alpha; PCK1: Phosphoenolpyruvate carboxykinase; G6P: Glucose-6-phosphatase; FAS: Fatty acid synthase. SD. * $p<0.05$. N=6.

5.3.7. Genome-wide methylation measured by LC-MS/MS

After analysing the RRBS methylome data, we also carried out LC-MS/MS measurement (introduced above) on the genomic DNA extracted from the mouse groups undergoing fasting and refeeding and their respective control groups. As shown on **Figure 38**, neither 5mC/C nor 5hmC/C ratios were different among the groups. Therefore, it can be concluded that acute metabolic stress does not cause quantitative changes in the 5mC and 5hmC levels when the cytosines of the whole genome are investigated.

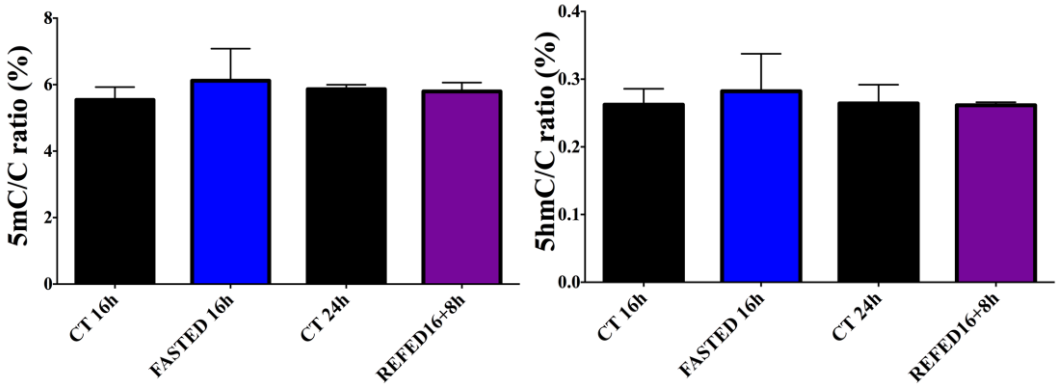


Figure 38. LC-MS/MS measurement of genomic 5mC/C (left panel) and 5hmC/C (right panel) ratios (%) in the groups CT, FASTED and REFED. S.D. N=5.

In the following, the biological validation of the LC-MS/MS method will be described.

5.4. Genome-wide epigenetic changes

As we have seen in the **Objectives**, I intended to find answers to the following scientific question:

IV. How does treatment with DNA methylation inhibitor or ascorbate affect genomic 5mC and 5hmC levels in cell lines?

In answer to this question, my first author publication includes the results described below [89].

5.4.1. Variability of our newly developed LC/MS-MS method

We developed a new liquid chromatography mass spectrometry (LC-MS/MS), which can detect genomic 5mC and 5hmC. For this, formic acid is used to hydrolyse DNA to nucleobases. All the MS measurements were performed by Pál Szabó. The LC-MS/MS was found to be sensitive and selective enough to detect and differentiate 5mC and 5hmC based on their different weights of molecular and fragment ions (not shown). Furthermore, in order to investigate the variability of the MS measurement, I extracted genomic DNA from a C57BL/6 mouse liver and I divided into 21 identical samples. The samples were measured by MS at three different times, each time in 7 parallels to discover intraday and interday variabilities. **Figure 39** displays the intra- and interday variability of the MS measurement of genomic 5mC/C and 5hmC/C ratios from mouse liver. As a result, little intra- and inter-day variabilities were measured and calculated for both 5hmC/C and 5mC/C. I performed the vast majority of genomic DNA extractions. It is important to emphasize that all the parallels of every experiment on a given cell type were always processed together from cell culture through DNA extraction to MS detection.

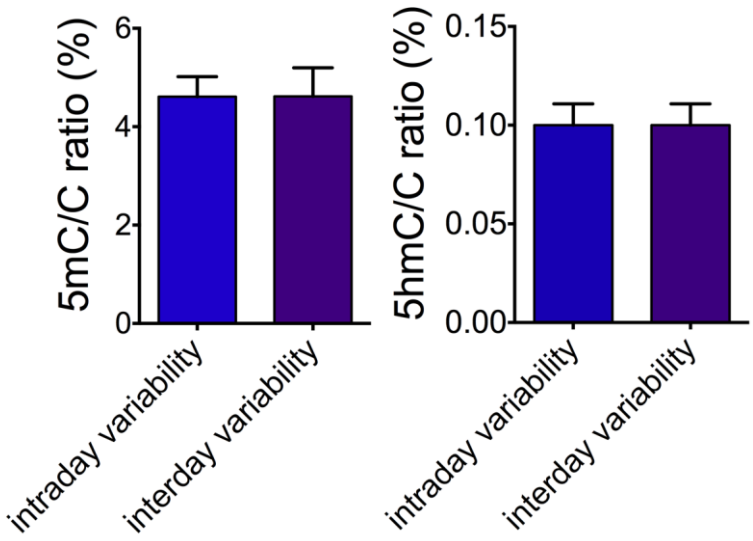


Figure 39. Intra- and inter-day variability of the LC-MS/MS measurement. Error bars indicate S.D. N=8. (modified from [89])

5.4.2. Genomic 5mC and 5hmC levels of different tissues, primary cells and cell lines

To answer this question, we explored genomic 5mC and 5hmC levels in different cell types and tissue types (**Figure 40).** 5mC levels are quite stable among different tissues, but the 5hmC is changeable [87]. Very high 5mC/C ratios were measured in the mouse forebrain. In addition, higher 5mC level was detected in mouse primary hepatocytes compared to mouse liver samples. Furthermore, we detected high genomic 5mC/C ratios in human induced pluripotent stem cells (iPSCs) and in neuronal progenitor cells (NPCs), which are differentiated from them. Chicken DT40 B-cell lymphoma cell lines contained more 5mC than the other cell lines (A375 human melanoma, A2058 human melanoma cells). The 5mC level of the sample with the lowest methylation was approximately three times less than that of the highest level.

DOI:10.14753/SE.2018.2176

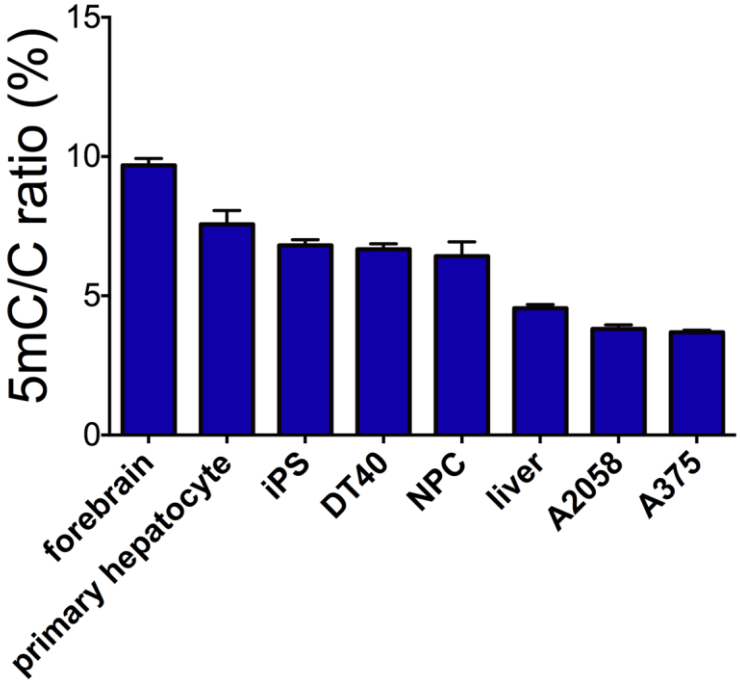


Figure 40A. Genomic 5mC/C ratios (%) in decreasing order in cell lines, primary cells and tissues. Error bars indicate S.D. N=4-6. (modified from [89])

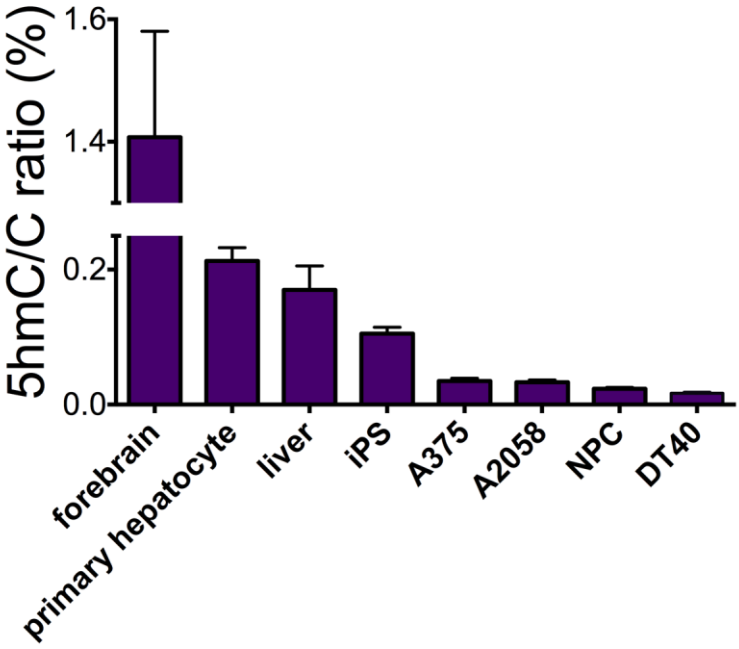


Figure 40B. Genomic 5hmC/C ratios (%) in decreasing order in cell lines, primary cells and tissues. Error bars indicate S.D. N=4-6. (modified from [89])

5hmC levels found in our cell lines and tissue samples were highly variable (**Figure 40**). The 5hmC level of the sample with the lowest methylation was approximately hundred times less than that of the highest level. Mouse forebrain showed high 5hmC level. Mouse liver and primary hepatocytes contain similar 5hmC levels. Lower 5hmC level was detected in iPSCs than brain or liver cells. In addition, conspicuous difference of 5hmC levels was identified between immortalized cell lines and the samples from tissues or primary cells. DT40, A2058 and A375 cells and NPCs, as well, contain low genomic 5hmC/C ratios. In conclusion, the newly developed MS technique is sensitive and simple enough to measure even very low levels of 5hmC.

5.4.3. Effect of DNA methylation inhibition on genomic 5mC and 5hmC levels in cell lines

5azadC, or decitabine is an FDA-approved drug used for treating myelodisplastic syndrome and AML in clinics. It inhibits DNA methylation, and it is widely used in research. Recent reports have shown paradoxical increase of the 5hmC level upon 5azadC treatment in HL60 human promyeloblast cells [137]. It was hypothesized that TET enzymes prefer to bind hemimethylated DNA strands, which appear during replication and stay demethylted because of decitabine treatment. Our aim here was to investigate other cell types, as well. In order to examine whether the above described phenomenon is ubiquitous, the effect of 5azadC on 5hmC level was interrogated in a range of cell lines using our newly developed MS method (**Figure 41**). Cells originated from various tissues (hematopoietic, liver carcinoma, uterine carcinoma, melanoma lung carcinoma and choriocarcinoma). 48-hour treatment was performed because the amount of hemimethylated DNA strands increase during division. As expected, treatment with 5azadC resulted in a significant loss of 5mC levels with approximately 50% (**Figure 41**). Moreover, 5azadC treatment also lead to significant increase of 5hmC level in HL60 cells.

5.4.4. 5hmC changes specific to some cell lines

DOI:10.14753/SE.2018.2176

Interestingly, the surprising effect of 5azadC on 5hmC levels was only found in hematopoietic cells. HL60 human promyeloblast, K562 human myelogenous leukemia and DT40 chicken lymphoblastoma cells exhibited significant increase in 5hmC levels upon treatment ($p<0.05$, *t*-test) (**Figure 41**). The above observed phenomenon suggests a specific mechanism in hematopoietic cells.

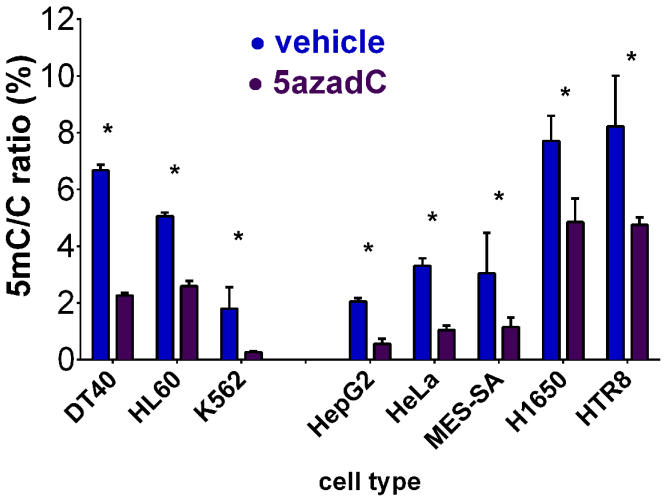


Figure 41A. Genomic 5mC/C levels detected by LC-MS in 5azadC-treated cell lines. Error bars indicate S.D. $p<0.05$. N=3-5. (modified from [89])

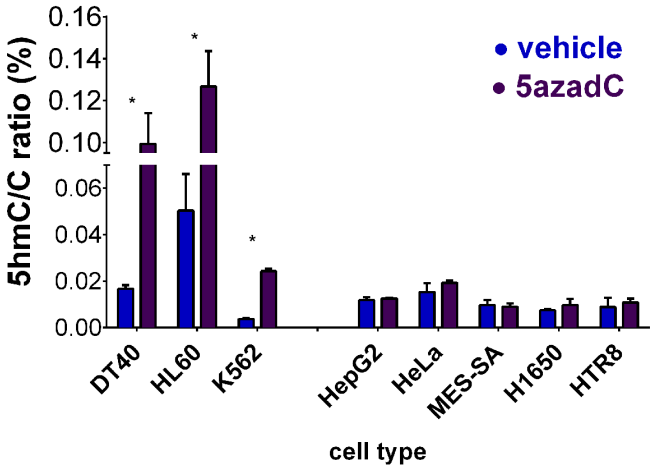


Figure 41B. Genomic 5hmC/C levels detected by LC-MS in 5azadC-treated cell lines. Error bars indicate S.D. $p<0.05$. N=3-5. (modified from [89])

DOI:10.14753/SE.2018.2176

5.4.5. Ascorbate-mediated TET function modulation in cell lines

Based on our observations, we hypothesized that the differences in the effect of 5azadC on 5hmC level in different cells can be attributed to different oxidative states. Accordingly, some cells have low TeT activity of enzymes because they cannot efficiently reducing Fe^{3+} to Fe^{2+} in the reaction. Therefore, we intended to study the effect of ascorbate. Ascorbate is a cofactor regulating DNA methylation and it is required for the full catalytic activity of TET dioxygenases [72, 73]. In our experiments, DNA methylation levels in HL60 and HeLa cells did not change upon ascorbate or decitabine and ascorbate co-treatment. Nevertheless, the 5hmC level were significantly increased upon decitabine and ascorbate treatment in HL60 cells (**Figure 42**). This elevation was higher than that observed upon 5azadC treatment alone. It suggests that the two molecules have synergistic effect and even HL60 cells require some more ascorbate for full TeT enzyme activity. Furthermore, 5hmC levels were also increased upon 5azadC treatment in HeLa cells, but only upon co-treatment with ascorbate. In summary, our results indicate that some cell lines lack functionally fully active TeT enzymes and possibly are deficient in ascorbate, therefore decitabine treatment does not have an effect on their 5hmC levels.

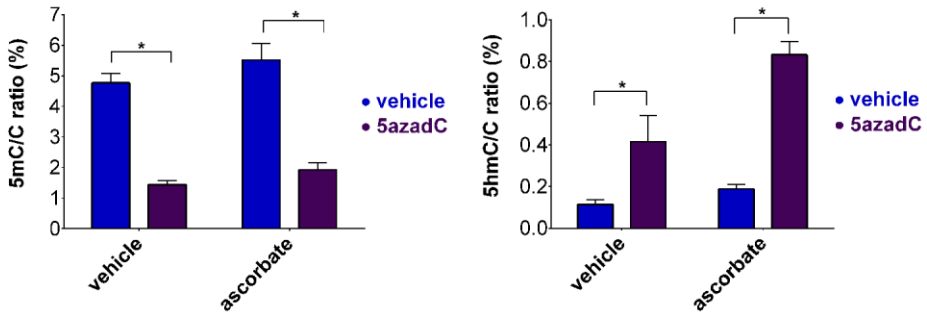


Figure 42A. Genomic 5mC/C and 5hmC/C ratios (%) in vehicle (blue) or 5azadC (purple) and/or ascorbate-treated HL60 cells. Error bars indicate S.D. $p<0.05$. N=5. (modified from [89])

DOI:10.14753/SE.2018.2176

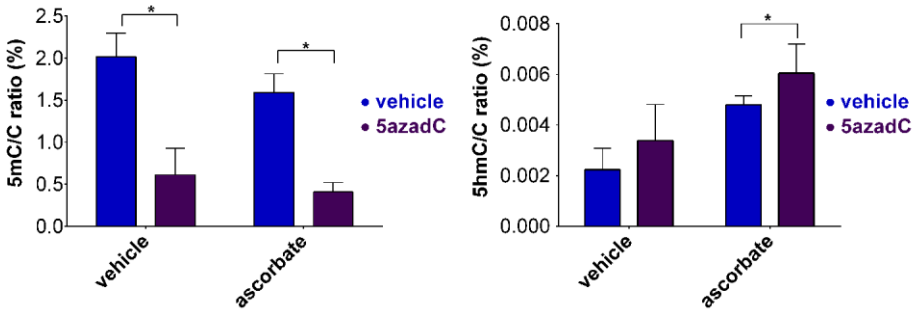


Figure 42B. Genomic 5mC/C and 5hmC/C ratios (%) in vehicle (blue) or 5azadC (purple) and/or ascorbate-treated HeLa cells. Error bars indicate S.D. $p<0.05$. N=5. (modified from [89])

DOI:10.14753/SE.2018.2176

6. DISCUSSION

6.1. Transcriptional regulation by homeobox TFs

During embryonic development, the most important cell fate decisions are made after the morula stage before the formation of the blastocyst. Our initial experiments have revealed that the PRD-class homeobox-containing proteins (Argfx, Dprx, Leutx and Tprx) localise predominantly to the nucleus with some exceptions of cytoplasmic staining.

Since mice lack Argfx, Dprx and Leutx, gain-of-function method was applied. Primary fibroblasts were transfected with the PRD-class genes and samples with ectopic expression were used for RNA-Seq. The proteins in question were also located in the nucleus, although some showed other subcellular staining (**Figure 13**); some proteins (Argfx, Tprx1 and Leutx) showed clear nuclear and one (Dprx) nuclear-cytoplasmic staining. This prominent nuclear localisation is a characteristic of transcription factors. Furthermore, we have shown here that the PRD-class genes are expressed prior to these cell fate decisions. They are strictly and precisely timed showing a sharp increase in expression at the 8-cell stages and a pronounced drop in the blastocyst (**Figure 14**). This observation suggests that they are likely to have roles in driving totipotency to cell fate specifications. It is also intriguing to find a putative downstream effector: the histone H2 variant HIST1H2BD. Therefore, it can be postulated that developmental changes are accompanied by chromatin structure changes and transcriptional changes, as well. Interestingly, the polyadenylated and spliced HIST1H2BD mRNA levels were induced upon cellular differentiation from human mesenchymal stem cells (hMSCs) into osteoblast lineages in a report [138]. These transcripts might be necessary to maintain adequate DNA packing and chromatin structure. Moreover, some histone mRNA transcripts exhibit altered expression during tumorigenesis [139].

The long arm of the human chromosome 19 where Dprx, Leutx and Tprx are located contains an unstable region because of the high frequency of the combination of gene loss and duplications leading to sequence similarity and recombination errors. In addition, this chromosome contains low density of recombination hotspots, but contains high GC content. Therefore, tandem gene duplications and gene loss are predominant in these regions providing diverse roles for these proteins.

In conclusion, we have discovered new roles for the novel PRD-class proteins (Argfx, Dprx, Leutx and Tprx) for early mammalian embryonic development. In addition, we postulate possible epigenetic marks regulating gene expression.

6.2. Transcriptional regulation by HNF4α

Our group has previously shown that HNF4α regulates *ABCC6* gene expression (**Figure 3**). In addition we have found that ERK1/2 activation inhibits *ABCC6* expression in an HNF4α-dependent manner. Here we have revealed that ERK1 is capable of phosphorylating HNF4α at a number of serine and threonine residues. Furthermore, we have found that phosphorylation of HNF4α impedes its *trans*-activational capacity in luciferase reporter assay and its DNA-binding.

HNF4α is a master gene regulator in hepatocytes and it plays an essential role in hepatic development. It regulates a great number of genes playing roles in glucose, lipid and amino acid metabolism, bile acid synthesis, detoxification and inflammation. Our ChIP-seq experiments have revealed approximately 9000 genomic HNF4α binding sites in HepG2 cells, which denote 5500 genes. We have also detected actively transcribed genes regulated by HNF4α; they were also marked by the H3K27ac, a sign of active genes [140] (**Figure 19**). Moreover, KEGG pathway analysis of HNF4α target genes uncovered similar pathways, which are in accordance with literature data. Among the typical genes *ABCA1*, *ABCC6*, *ALDOB (aldolase B), APOA1*, *APOB*, *APOCIII*, *BLVRA* and *B*, *CYP7A1*, *HNF1a* and *4a*, *HPD*, *PKLR* and *SLC2A2 (GLUT2)* were found.

We have selected six - *BLVRA, BLVRB, HPD, PKLR, ABCC6* and *APOA1* - for our further experiments. Pyruvate kinase (PKLR) is an important player in regulating glucose metabolism [16]. The other genes are also involved in metabolism. The enzyme 4-Hydroxyphenyl pyruvate dioxygenase (HPD) takes part in tyrosine metabolism [141]. Apolipoprotein A1 (APOA1) plays a pivotal role in lipid transport [142]. The biliverdin reductase genes *BLVRA* and *BLVRB* participate in heme metabolism and the antioxidant pathway, since they are responsible for catalyzing the synthesis of bilirubin from biliverdin. Finally, the transporter ABCC6 protects against ectopic calcification, furthermore, it is suggested to do a part in ATP homeostasis [32, 33, 143].

Our experiments confirmed that ERK1/2 has posttranslational and suggested transcriptional effects on HNF4α. We have shown that short activation of the ERK1/2 pathway lowered the DNA binding capacity of HNF4α in HepG2 cells implying a post-translational effect. This rapid decline was enhanced after 24h ERK1/2 activation, when already *HNF4a* gene expression is obstructed, as reported earlier [144]. We proved furthermore that short ERK1/2 activation leads to the phosphorylation of HNF4α *in vitro* (**Figure 16**). We have identified several phosphorylation sites mediated by activated ERK1. A number of studies have shown that HNF4α can be phosphorylated by different kinases, for example, PKC, PKA, AMPK and p38. Phosphorylation can alter the DNA-binding, *trans*-activation capacity and intracellular localisation of HNF4 [145]. Phosphorylation of serine 87 by PKC excessively reduced the DNA binding capacity of the protein and also destabilizes it [27]. PKA – induced by cAMP - phosphorylates HNF4α at the serines 142/143 [28] leading to diminished DNA binding activity [28]. Furthermore, it has been reported that PKA – activated by Thyroid-stimulating hormone (TSH) - decreases nuclear localisation of HNF4α in HepG2 cells [146]. AMPK - a metabolic master switch [17] - phosphorylates S313 of HNF4α resulting in its disrupted dimerization [30] and its *trans*-activational capacity. AMPK activation also diminishes the transcription of several hepatic HNF4α target genes. p38 is known to phosphorylate HNF4α at residue T166/S167 leading to its inhibited *trans*-activation [29, 147]. Some suggest indirect effects of p38 [30, 148], whereas some others an inhibitory role [18].

We have examined the functional relevance of several residues phosphorylated by ERK1. We used a construct with *ABCC6* promoter cassette for luciferase reporter gene assays, as previously described [4, 33]. We have reported previously that the *ABCC6* promoter is activated by HNF4α. Here we co-expressed this gene with wild-type (WT) or various phosphomimetic HNF4α mutants.

The mutant S87 for PKC phosphorylation lead to significantly diminished reporter gene activity relative to the WT, as reported previously [27] (**Figure 18**). We have investigated sites firstly identified in our experiments (S451, T457/T459) and the sites to the previously described as p38 target (T166/S167) [29]. These phosphomimetic mutants did not change the *trans*-activational capacity of HNF4α. In our luciferase assay, only the site previously shown as targeted by AMPK (S313) showed diminished reporter gene activity

[30, 148]. In summary, our results imply that ERK1/2 activation leads to HNF4α phosphorylation and reduced DNA binding capacity of the TF.

We have also shown in ChIP-qPCR experiments that ERK1/2 phosphorylates HNF4α, resulting in reduced HNF4α DNA-binding capacity to target sequences (**Figure 21**). ChIP-qPCR experiments require several controls, which we included systematically in our experiments. One of the most important controls is the analysis of immunoprecipitation (IP) of the negative control region performed in the same test tube as the target regions. IgG controls for non-specific IP were also performed. Another negative control experiment is to use only beads and no antibody. Finally, the IP can be expressed as a percentage of input, which serves for further normalization. Furthermore, we performed several independent experiments in order to compensate for the variability of the ChIP experiment. Taking into account these considerations, the ChIP-qPCR technique is capable of detecting quantitative alterations of the chromatin-binding of a TF to target genomic loci.

Activated ERK1/2 pathway is a characteristic of different physiologic and pathologic conditions. For instance, bile salts can act as signalling molecules through the Sphingosine-1-phosphate receptor 2 (S1PR2) G protein coupled receptors (GPCRs), which enhance ERK1/2 action in order to control glucose, lipid and drug metabolism in hepatocytes ([149-151]; [122]). By activating the ERK pathway, HNF4α is quickly downregulated, which leads to diminished expression of *PEPCK* and *G-6-Pase* gluconeogenic genes [150] and *CYP7A1*, the enzyme of which is crucial for bile acid synthesis [152].

Furthermore, bile acids can activate specific nuclear receptors (e.g. FXR and vitamin D receptor (VDR)), as well. VDR can directly activate ERK1/2 [153]. FXR is an important player in bile acid homeostasis [154, 155], which can enhance the expression of the orphan receptor Small heterodimer partner (SHP). SHP binds HNF4α blocking binding of the latter to target *cis*-regulatory elements of the *CYP7A1* promoter [156].

Moreover, oxidative stress (ROS), growth hormones (for example HGF, EGF and FGF15/19 [33, 157]) and cytokines (IL1 and TNFα [18, 24, 29]) can also activate the ERK1/2 [158]. These factors results in reduced HNF4α activity and downregulation of various genes. Our data suggest that the short-term inhibition of HNF4α is due to its

phosphorylation, suggesting that ERK1/2 plays a major role in the complex regulation of a number of hepatic genes through phosphorylating HNF4α.

6.3. Effects of short-term nutritional stress

In this set of experiments, we were interested in the effect of short-term nutritional stress on mouse liver. Firstly, we examined the body weight and blood glucose levels of the mice. Regarding body weight, fasting for 8 hours did not, but fasting for 16 and 24 hours lowered body weight compared to their control group (**Figure 22**). Refeeding for 8 hours could not restore body weight. In conclusion, both fasting and refeeding have a dramatic effect on the body weight of the mice.

Next, we investigated the blood glucose levels of the animals. Fasting (8h, 16h, 24h) drastically lowered blood glucose levels, but the duration of fasting did not have an effect (**Figure 24**). It is interesting to compare the changes in blood glucose levels to those in body weight, because blood glucose was and body weight was not re-established upon refeeding.

Furthermore, we observed that the blood glucose level of the group which was sacrificed at 2 p.m. was significantly higher than that of 2 a.m (**Figure 23**). It can be attributed to the effect of the circadian rhythm which influences blood glucose levels and expression of a great number of genes throughout the day [159]. Zeitgeber (or 'time-giver') is the external timing that synchronizes the animals' biological rhythm to 12 hours' light and 12 hours' dark cycles. In research, ZT0 refers to the starting point of the light period. Although diurnal blood glucose peaks have been reported in the literature at ZT10 [160] (which is 4 p.m. in our experimental setup) or ZT12 [161] (which is 6 p.m. in our case), it is at 2 p.m. (ZT8) in our case. The difference might be attributed to slight differences in circadian rhythm, stressful animal facility conditions or differences in food consumption throughout the day or night.

Secondly, we examined the protein levels of two transcription factors regulating liver homeostasis and gene expression or enzymes playing a role in glucose metabolism (**Figures 25, 26, 27**). We have found that HNF4α protein does not change due to nutrient deprivation. CEBPα was reported to be increased by fasting-related signals from glucagon by binding Early growth response gene-1 (Egr-1) to its promoter and regulating

gluconeogenesis [162] . Moreover, the level of the gluconeogenic enzyme PCK1 is elevated upon fasting, which is in accordance with literature data [162].

Thirdly, we analysed the RRBS data. We have calculated the mean methylation percentages for each group. Altogether, there is no significant difference of average methylation among the groups (**Figure 30**). Furthermore, we have found that the most heterogeneous group is the 'REFED 16+8h' group; some samples have high and others have low methylation. This suggests that the 6 animals respond to refeeding very differently in terms of methylation.

Next, in order to analyse the methylome sequencing data in detail, we have performed a stringent method in order to capture all the possible methylation changes without finding false positives. We have found that fasting resulted in almost 500 CpGs undergoing hypermethylation, whereas less than half undergoing hypomethylation (**Table 5**). In contrast to fasting, refeeding had the opposite effect on methylation change: we found less hypermethylated sites than hypomethylated. It has not been shown before that genome-wide massive methylation changes can occur after 16 hours' fasting and 8 hours' refeeding. Clearly, most of the methylation changes reflect the response to stress in the liver of the mice and a large proportion of the methylation changes are non-specific. However, we consider that methylation changes in the same direction in 6 animals per group analysed by a stringent method are specific. Moreover, this method enables us to analyze at least half of all genes (~10,000), and we can investigate those CpGs in the genome which are important in gene regulation.

One could speculate the biological relevance of our findings. It might be possible that the remarkable hypermethylation triggered by fasting contributes to a stress-induced, saving mode of hepatocytes where the gene expression of a number of hepatic genes is downregulated (for instance, the synthesis of 'luxury' products), at least at certain promoters. In contrast, hypomethylation provoked by refeeding might be associated with increased gene expression allowing hepatocytes to produce any products, as well as try to restore metabolic balance after the shortage of nutrition. According to a recent report, stress induces remodelling of interaction networks, long-term or strong stress switches the cells in a different mode of function with less and weaker connections and hubs (core elements) and more shrinkage [163]. The complex and flexible networks can be re-established after stress, mainly with the help of molecular chaperones. If repeated and

long-term waves of stress affect the cells, it might accelerate aging and disease prevalence. This finding might have implications for nutritional challenges, as well in a way that repeated perturbations in nutrition availability might have long-term, harmful consequences.

When CpG distribution around CpG islands was investigated, we discovered that upon hypermethylation, CpG islands and CpG shores are underrepresented, but the regions beyond them are overrepresented, especially for refeeding (**Figure 32**). In contrast, hypomethylation in refeeding greatly affected CpG shores (**Figure 32**). Fasting and refeeding are characterized with unambiguously distinct CpG distributions. It is well-known that most gene promoters (~70% in mammals) contain unmethylated stretches of DNA with high CpG density, also known as CpG islands [164]. The closer several unmethylated CpGs in a promoter are, the more likely they contribute to gene activation. Therefore, CpG islands and their methylation status play a fundamental role in gene expression.

Furthermore, we discovered that upon hypermethylation, promoters and exons are mildly underrepresented (**Figure 33**). At the same time, introns and intergenic regions are overrepresented. The observed finding might be an instance of gene body methylation change. In the mouse genome, CpG islands are mainly located around the TSS, but approximately 20% are intragenic and 20% are intergenic [49]. It seems that short-term nutritional stress can induce pronounced, CpG-rich intergenic and intragenic methylation changes, as well.

However, refeeding leads to profound changes in the distribution of differential methylation compared to the background. Upon refeeding for hypomethylation DNA methylation changes primarily affect promoters. The explanation for the aforementioned phenomenon might be that a fasted animal endeavours to protect its genome from vast methylation changes, whereas refeeding is such a drastic interference after fasting, that the genome is prone to epigenetic changes. It has not been shown before that genome-wide massive methylation changes occur after short-term fasting and refeeding.

In addition, we have found that CpGs that underwent change upon fasting and were annotated to promoters, only a limited number of them were in CpG islands or CpG shores compared to the background (**Figure 35**). However, CpGs that were hypomethylated upon fasting and were located in either introns or intergenic regions were mainly found

in CpG islands or shores, the distribution of which is appreciably different from the background (**Figure 36**). These changes in intergenic regions might be attributed to enhancers, and promoters are protected from vast methylation changes. However, it is necessary to emphasize that compared to the distribution of CpG islands in the genome (1%) [165], our observed distributions are much more remarkable, even with the stringent analysis. Furthermore, our kit and protocol is optimized in a way that CpG-rich regions (e.g. CpG islands) represent 1/3 of all distributions for all the CpGs [166]. In fact, the regions with almost exclusively CCGG sequences have little chance to be included in the sample pools because of the high efficiency of MspI digestion and several steps of size selection achieved by magnetic beads. The range of the distance between 2 CpGs we examine are theoretically between 200 and 1200 bp. Therefore, we cannot investigate all the changes in CpGs, only a reduced representation of them. However, we can analyze at least half of all genes (~10,000).

Apart from the stringent analyses, we have tried a more relaxed analysis in order to capture more changes. With this analysis, we could find a greater number of significantly changing CpGs. The reason for this was to find genes that have one or more specific CpGs changed in one direction upon fasting and this exact CpG(s) changed in the opposite direction upon refeeding. I identified several metabolic pathways, for instance cholesterol, fatty acid, phospholipid, amino acid and carbohydrate (gluconeogenetic and glycogen metabolic) metabolic processes (**Table 8**). It can be argued that hypermethylation is associated with gene inactivation and hypomethylation with gene activation, therefore the reversal in these directions would contain those genes the expression of which is decreased upon fasting.

We intended to investigate in more detail those CpGs and the genomic environment that are in close proximity to the CpGs significantly changing based on the more stringent analysis. Therefore, we performed DMR analyses. For the DMR analyses, we first defined the DMR as +/-1000 bp region from the TSS with a minimum of 3 CpGs in order to gain information about the promoters of the genes which change upon nutritional stress. Applying this method, we have found that *FoxO3* and *Hdac1* had a region in its promoter which is hypomethylated upon fasting (**Table 9**). Interestingly, it is reported that siRNA knocking down of HDAC1 - but not HDAC2 or HDAC3 - leads to decreased PCK1 and HNF4α expression in HepG2 cells [167]. In contrast to hypomethylation, a promoter

region in *Acot2* and *Idh3a* is hypermethylated upon fasting. Concerning refeeding, we identified a promoter region in *Acot3*, *RXRβ* and *γ* and *Cpt1a* which underwent hypomethylation, whereas a promoter region of *Hdac10* was hypermethylated.

Secondly, we defined the DMR as a region where there are 2 CpGs changing in the same direction within 100 bp in gene promoters. Altogether 16.000 regions were identified, which corresponded to 450 genes. We found that there are DMRs in the metabolic genes *Insig2* and *Acot11* which are hypomethylated upon refeeding. These findings are in accordance with literature data. Indeed, FoxO1 has been reported to be activated upon fasting. Furthermore, the mRNA level of Cpt1, an enzyme playing a role in ketogenesis was also elevated [108]. Refeeding was characterized with elevated RXRβ and γ levels [109].

Furthermore, we investigated the mRNA expression of several metabolic genes. The expression of HNF4α was lowered upon refeeding compared to fasting. The expression of PCK1 was elevated upon fasting, but decreased upon refeeding. Moreover, G6P expression was significantly lowered upon refeeding and refeeding could not restore FAS expression to the value of its respective control group (**Figure 37**). These changes are in accordance with literature data (see **Introduction**) and partly in correlation with protein level and methylation changes. We found that both the protein level and mRNA expression of PCK1 is elevated upon fasting. In addition, the promoter G6P is hypermethylated upon refeeding and its mRNA expression is decreased. Moreover, neither the protein level nor the mRNA expression of HNF4α is changed upon fasting, its expression is decreased upon refeeding. Lastly, the promoter of FAS is hypermethylated upon fasting, but the mRNA expression decrease is not significant.

In conclusion, several metabolic genes were found to be changing their methylation status in their promoters in the direction of either hypo- or hypermethylation. Therefore, we were interested if the methylation changes observed by the RRBS technique are CpG-specific and enriched at CpG-rich regions or they characterize any Cs in the genome.

In order to connect our findings by the RRBS technique, we are planning to follow two directions. First of all, we intend to validate the RRBS technique by performing bisulfite conversion followed by Next-Generation Sequencing on individual samples. Finding the same methylation changes with this technique would prove the reliability and specificity of RRBS. Next, we intend to compare our RRBS data and mRNA expression results with

publicly available ChIP-Seq data. Correlating histone marks of active promoters (H3K4me3) and other regulatory regions (H3K27ac) with observed methylation and mRNA expression changes would also strengthen our findings and shed light on the role of methylation in the regulation of gene expression.

6.4. Genome-wide epigenetic changes

We developed a new liquid chromatography mass spectrometry (LC-MS/MS), which can detect genomic 5mC and 5hmC. This method is selective enough to differentiate cytosine, 5mC and 5hmC in the genomic DNA. Moreover, our method can sensitively detect low levels of genomic 5hmC. 5mC levels are quite stable among different tissues, but the 5hmC is changeable [87]. Very high 5mC/C ratios were detected in the mouse forebrain, possibly correlating with non-CpG methylation, previously reported. The 5mC levels measured in mouse liver were similar to other published reports [168, 169]. Higher 5mC level was detected in mouse primary hepatocytes compared to mouse liver samples (**Figure 40**). Primary hepatocytes are devoid of non-hepatocyte cells and perhaps undergo hypermethylation during cell culture. It is also possible that *de novo* methylation affects non-CpG sites, since CpA dinucleotides can be methylated, as well [170, 171]. iPSCs are characterized by high level of DNA methylation [172]. The 5mC level of the sample with the lowest methylation was approximately three times less than that of the highest level.
5hmC levels found in our cell lines and tissue samples were highly variable (**Figure 40**). The 5hmC level of the sample with the lowest methylation was approximately hundred times less than that of the highest level. Brain-derived cells show very high 5hmC levels, as previously reported [173]. Furthermore, adult mouse liver exhibits low 5hmC level [87, 174, 175]. In cell culture, global 5hmC levels are greatly diminished due to the culturing conditions [175]. Also, NPCs were reported to contain low 5hmC levels [176] In conclusion, the newly developed MS technique is sensitive and simple enough to measure even very low levels of 5hmC.
5azadC, or decitabine is an FDA-approved drug used for research and treatment in the clinics, as well. Recent reports have shown paradoxical increase of the 5hmC level upon 5azadC treatment in HL60 human promyeloblast cells [137]. It was hypothesized that TET enzymes prefer to bind hemimethylated DNA strands, which appear during

replication and stay demethylated because of decitabine treatment. 48-hour treatment was performed because the amount of hemimethylated DNA strands increase during division. Nevertheless, the 5mC loss is greater than the 5hmC gain, since re-methylation is also inhibited due to the lack of newly synthetized methylated strands. Furthermore, 5hmC is not as stable as 5mC, and 5hmC can be further oxidized into carboxylcytosine, formylcytosine and eventually cytosine [87].

In our experiments, DNA methyltransferase inhibitor lead to significant decrease in 5mC levels (**Figure 41**). Moreover, 5azadC treatment also lead to significant increase of 5hmC level in hematopoietic cells. In other cell lines, for instance HepG2 cells we could not see any 5hmC change. However, in other cell lines of hepatic origin (e.g. Huh7) decitabine treatment lead to reorganized nuclear 5hmC by immunostaining, as reported [177].

The above observed phenomenon suggests a specific mechanism in hematopoietic cells. Ascorbate is a cofactor regulating DNA methylation and it is required for the full catalytic activity of TET dioxygenases [72, 73]. It seems that 5azadC and ascorbate have synergistic effect and even HL60 cells require some more ascorbate for full TeT enzyme activity (**Figure 42A**). Our results indicate that some cell lines lack functionally fully active TeT enzymes and possibly are deficient in ascorbate, therefore decitabine treatment does not have an effect on their 5hmC levels.

One can speculate that different cell lines are characterized by very different expression levels of TeT genes. One would expect higher expression of TeT genes in hematopoietic cells or low TeT expression levels would be increased upon treatment. Nevertheless, none of these suppositions are true, based on publicly available data [178]. K562 cells have relatively high TeT1 expression, but lower TeT2 and TeT3 expression levels. In addition, HL60 cells have higher TeT2 and TeT3 expression levels, but lower TeT1 level. Moreover, these expression levels were not effectively different in our assays. Several cells had similar expression to K562, but 5azadC treatment had opposite effect on their 5hmC levels. Furthermore, we have examined the expression levels of other related genes (IDH1 and IDH2, which synthesize α-keto-glutarate or DNMTs), but again no clear correlation was observed between the effect of 5azadC and 5hmC levels in hematopoietic cells.

Our results have revealed that decitabine treatment decreases 5mC and increases 5hmC levels in hematopoietic cell lines. Several reports suggest that DNA hydroxymethylation

and TET2 function are dysregulated in hematologic malignancies (T-ALL, AML, CML) implying that TETs might act as tumour suppressors (reviewed in [179]). Indeed, TeT2 expression level can be positively correlated with tumour-free survival in CLL patients[180]. Decitabine treatment of patients with hematopoietic tumours might result in elevated 5hmC levels, which might have a therapeutic effect. Furthermore, decitabine could be used efficiently in solid tumours or in combination therapy (reviewed in [181]). There, only gene specific DNA methylation changes were investigated, however, global effects of 5hmC level might also be responsible for mechanism of the drug. Our results postulate that co-treatment with dietary ascorbate and decitabine might cause a more pronounced effect and increased overall survival in hematopoietic malignancies.

DOI:10.14753/SE.2018.2176

7. CONCLUSIONS

In the present PhD dissertation, I intended to shed light on four instances where epigenetics plays an essential role. Here, I have shown that transcription and epigenetics are closely related. Moreover, environmental stimuli also influence the transcriptional and/or epigenetic status of the cells, organs or animals.

Firstly, we could see that the homeodomain-containing genes Argfx, Dprx, Leutx and Tprx are transcription factors having essential roles in animal development, which derives from their highly conserved homedomain-containing DNA sequences. Our experiments have revealed that these TFs are responsible for driving totipotency to cell fate decisions. More importantly, a downstream effector has been found, which is the HIST1H2BD histone H2 variant. Therefore, it can be postulated that developmental changes are accompanied by changes in both the chromatin structure and transcription. Indeed, these two processes depend on each other, where most probably epigenetic marks regulate gene expression.

Secondly, we have investigated the transcriptional role of HNF4α - required for hepatocyte differentiation and development in the liver - under physiological conditions and upon extracellular stimuli and its concomitant epigenetic changes. We have revealed that upon ERK1 activation HNF4α is phosphorylated, which results in lowered DNA binding capacity of the TF to target hepatic DNA sequences in HepG2 cells. We have also detected actively transcribed genes regulated by HNF4α; they were also marked by the epigenetic modification H3K27ac. In our ChIP-Seq experiments, 8748 transcription factor binding sites (TFBSs) could be identified for HNF4α. The overlap between loci bound by both HNF4α and H3K27ac was remarkable, suggesting that a great number of HNF4α sites are active regulatory regions. The HNF4α-target regions, which were occupied by expansive H3K27ac signal, showed diminished HNF4α-binding in ChIP-qPCR experiments suggesting that ERK1 activation might not only change TF-binding, but epigenetic mechanisms might also be involved.

Thirdly, we were interested if nutritional stress (short-term fasting and refeeding) influences the methylation status of hepatocyte DNA *in vivo*. We have observed enormous DNA methylation changes, fasting lead to global CpG-hypermethylation, whereas refeeding resulted in global CpG-hypomethylation. These changes affected

several metabolic pathways, for instance cholesterol, fatty acid, phospholipid, amino acid and carbohydrate (gluconeogenetic and glycogen metabolic) metabolic processes. It is clear that CpG methylation influences gene expression. We investigated these methylation changes with several methods, and we could couple them with changes in blood glucose, body weight, protein and mRNA expression levels. It seems that rapid and vast DNA methylation changes occur due to its reversibility and its responsiveness to environmental conditions.

Fourthly, but not lastly, we have also seen that despite DNA methylation changes at CpG islands, genome-wide DNA methylation of cytosines is not affected when animal livers undergo nutritional stress conditions. It seems that the mouse genome is protected from vast methylation changes which affect the whole body and might be too drastic. However, transcription and gene expression seem to be altered and the underlying mechanisms are partly epigenetic. For the genomic 5mC and 5hmC measurements, we have set up a method to measure 5-methylcytosine and 5-hydroxymethylcytosine levels in the genome by LC-MS/MS. We demonstrated in several experiments that DNA methylation inhibition leads to the elevation of 5hmC levels in hematopoietic cell lines. Moreover, adding ascorbate – a cofactor of the TET enzymes – further enhances the formation of 5hmC in hematopoietic cell lines.

In conclusion, all the four projects have clear epigenetic implications and raise very interesting further questions which would be worth investigating.

DOI:10.14753/SE.2018.2176

8. SUMMARY

Eukaryotic gene expression is regulated by transcription *via* the complex network of regulatory proteins, i.e. TFs. Vertebrate homeobox genes are highly conserved in structure and function, moreover, they play fundamental roles in animal development. I showed that the homeodomain-containing Argfx, Dprx, Leutx and Tprx are predominantly located in the nucleus. In addition, their highest expression was found in human embryos at the 4-cell and 8-cell morula stages prior to cell fate restrictions.

The nuclear receptor HNF4α is the master regulator in hepatocytes regulating genes involved in lipid and glucose metabolism. Phosphorylation of the protein can alter its DNA-binding, *trans*-activation capacity and intracellular localisation. We found that ERK1 can phosphorylate HNF4α. I showed that the ERK1 and AMPK- targeted phosphorylation site had an inhibitory effect on target gene transcription. Phosphorylation also resulted in reduced HNF4α DNA-binding capacity to target sequences.

Gene expression is regulated at two interconnected levels: transcriptional and epigenetic. Epigenetics involves various modifications of the chromatin structure in the nuclei of cells. DNA methylation is a reversible modification, thus methylation patterns can be established *de novo* and modified, reflecting environmental conditions. 5hmC is an epigenetic modification of Cs formed by TETs. We developed a new LC-MS/MS method to detect genomic 5mC and 5hmC. We showed that treatment with DNA methylation inhibitor decreased genomic 5mC levels and increased 5hmC levels in hematopoietic cell lines. Ascorbate is a cofactor regulating DNA methylation. Moreover, we found that some cells require additional ascorbate for full TeT enzyme activity.

The highly dynamic epigenetic modifications provide an interface between the genome and the environment. Moreover, there is a connection between cellular metabolic balance and DNA methylation. We observed short-term effects of acute nutritional challenges (fasting and refeeding) *in vivo* in mouse liver. Fasting had a dramatic decrease on the body weight and blood glucose. The protein and the mRNA expression level of the gluconeogenic enzyme PCK1 were elevated upon fasting, whereas G6P expression levels were decreased upon refeeding. Furthermore, fasting resulted in global hypermethylation, while refeeding lead to global hypomethylation. The implicated pathways include cholesterol, fatty acid, phospholipid, amino acid and carbohydrate metabolic processes.

DOI:10.14753/SE.2018.2176

9. ÖSSZEFOGLALÁS

Az eukarióta génexpressziót a transzkripció regulálja szabályozó fehérjék (transzkripciós faktorok) összetett hálózatán keresztül. A gerincesek homeobox génjei konzerváltak a szerkezet és funkció tekintetében, és alapvető szerepet játszanak az állatok fejlődésében. Kimutattam, hogy a homeodomain-tartalmú Argfx, Dprx, Leutx és Tprx fehérjék túlnyomóan a sejtmagban helyezkednek el. Továbbá, humán embriókban 4-sejtes és 8-sejtes morula állapotban – a sejtsors kialakulása előtt - a legmagasabb expressziójuk.

A HNF4α nukleáris receptor a zsír- és cukoranyagcserében részt vevő gének fő szabályozója májsejtekben. A fehérje foszforilációja megváltoztatja a DNS-kötését, a transz-aktivációs képességét és a sejten belüli elhelyezkedését. Kimutattuk, hogy az ERK1 kináz foszforilálja a HNF4α-t. Megmutattam, hogy az ERK1 és AMPK által célzott foszforilációs hely gátló hatású a célgén transzkripciójára nézve. Ezen felül a foszforiláció csökkentette a HNF4α célgének DNS-éhez való kötési képességét is.

A génexpressziót két, egymással szorosan összefüggő szint szabályozza: transzkripciós és epigenetikai. Az epigenetika a sejtmagban lévő kromatinszerkezet módosításait foglalja magába. A DNS metiláció egy visszafordítható módosítás, így a metilációs mintázat *de novo* kialakítható és módosítható, mely a környezeti körülményeket tükrözi. Az 5hmC egy olyan epigenetikai módosítás, melyet a TET enzimek hajtanak végre. Kimutattuk, hogy a DNS metiláció gátlásával csökkent az 5mC szint és nőtt az 5hmC szint hematopoetikus sejtvonalakban. Az aszkorbát egy kofaktor, mely szabályozza a DNS metilációt. Kimutattam, hogy néhány sejtvonal további aszkorbát hozzáadását igényli a teljes TET enzimaktivitás eléréséhez.

A dinamikus epigenetikai módosítások kapcsolódási pontot jelentenek a genom és a környezet között. Továbbá, a sejtek anyagcsere-egyensúlya és DNS metilációja között összefüggés van. Akut táplálkozási stresszhatásásoknak (éhezés és újraetetés) rövidtávú hatásait figyeltük meg egerek májában *in vivo*. Éhezés hatására a testsúly és a vércukorszint drámaian lecsökkent. A glükoneogenezis kulcsenzimének, a PCK1-nek megnőtt a fehérje és mRNS szintje éhezés hatására, míg a G6P expressziója lecsökkent újraetetés hatására. Továbbá éhezésre hipermetiláció, míg újraetetésre hipometiláció következett be, melyek a koleszterin-, zsírsav, foszfolipid, aminosav és szénhidrát-anyagcsere útvonalait érintették.

DOI:10.14753/SE.2018.2176

10. BIBLIOGRAPHY

1. Cooper GM, Hausman RE. The cell: a molecular approach. ASM Press, Sunderland. 2000:251-259.
2. Gilbert SF, Susan RS. Developmental biology. Sinauer Associates, Incorporated Publishers, Sunderland, 2014:96-97.
3. Tamaru T, Isojima Y, Yamada T, Okada M, Nagai K, Takamatsu K. (2000) Light and glutamate-induced degradation of the circadian oscillating protein BMAL1 during the mammalian clock resetting. J Neurosci, 20(20): 7525-30.
4. Ratajewski M, de Boussac H, Sachrajda I, Bacquet C, Kovacs T, Varadi A, Pulaski L, Aranyi T. (2012) ABCC6 expression is regulated by CCAAT/enhancer-binding protein activating a primate-specific sequence located in the first intron of the gene. J Invest Dermatol, 132(12): 2709-17.
5. Holland PW. (2013) Evolution of homeobox genes. Wiley Interdiscip Rev Dev Biol, 2(1): 31-45.
6. Goodman FR. (2002) Limb malformations and the human HOX genes. Am J Med Genet, 112(3): 256-65.
7. Holland PW, Marletaz F, Maeso I, Dunwell TL, Paps J. (2017) New genes from old: asymmetric divergence of gene duplicates and the evolution of development. Philos Trans R Soc Lond B Biol Sci, 372(1713).
8. Zhong YF, Holland PW. (2011) The dynamics of vertebrate homeobox gene evolution: gain and loss of genes in mouse and human lineages. BMC Evol Biol, 11: 169.
9. Evans RM, Mangelsdorf DJ. (2014) Nuclear Receptors, RXR, and the Big Bang. Cell, 157(1): 255-66.
10. Plutzky J. (2011) The PPAR-RXR transcriptional complex in the vasculature: energy in the balance. Circ Res, 108(8): 1002-16.
11. Boergesen M, Pedersen TA, Gross B, van Heeringen SJ, Hagenbeek D, Bindesboll C, Caron S, Lalloyer F, Steffensen KR, Nebb HI, Gustafsson JA, Stunnenberg HG, Staels B, Mandrup S. (2012) Genome-wide profiling of liver X receptor, retinoid X receptor, and peroxisome proliferator-activated receptor

alpha in mouse liver reveals extensive sharing of binding sites. Mol Cell Biol, 32(4): 852-67.

12. Rosen ED, Spiegelman BM. (2001) PPARgamma : a nuclear regulator of metabolism, differentiation, and cell growth. J Biol Chem, 276(41): 37731-4.
13. Hayhurst GP, Lee YH, Lambert G, Ward JM, Gonzalez FJ. (2001) Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. Mol Cell Biol, 21(4): 1393-403.
14. Drewes T, Senkel S, Holewa B, Ryffel GU. (1996) Human hepatocyte nuclear factor 4 isoforms are encoded by distinct and differentially expressed genes. Mol Cell Biol, 16(3): 925-31.
15. Odom DT, Zizlsperger N, Gordon DB, Bell GW, Rinaldi NJ, Murray HL, Volkert TL, Schreiber J, Rolfe PA, Gifford DK, Fraenkel E, Bell GI, Young RA. (2004) Control of pancreas and liver gene expression by HNF transcription factors. Science, 303(5662): 1378-81.
16. Stoffel M, Duncan SA. (1997) The maturity-onset diabetes of the young (MODY1) transcription factor HNF4alpha regulates expression of genes required for glucose transport and metabolism. Proc Natl Acad Sci U S A, 94(24): 13209-14.
17. Leclerc I, Lenzner C, Gourdon L, Vaulont S, Kahn A, Viollet B. (2001) Hepatocyte nuclear factor-4alpha involved in type 1 maturity-onset diabetes of the young is a novel target of AMP-activated protein kinase. Diabetes, 50(7): 1515-21.
18. Mogilenko DA, Dizhe EB, Shavva VS, Lapikov IA, Orlov SV, Perevozchikov AP. (2009) Role of the nuclear receptors HNF4 alpha, PPAR alpha, and LXRs in the TNF alpha-mediated inhibition of human apolipoprotein A-I gene expression in HepG2 cells. Biochemistry, 48(50): 11950-60.
19. Gupta RK, Kaestner KH. (2004) HNF-4alpha: from MODY to late-onset type 2 diabetes. Trends Mol Med, 10(11): 521-4.
20. Bogan AA, Dallas-Yang Q, Ruse MD, Jr., Maeda Y, Jiang G, Nepomuceno L, Scanlan TS, Cohen FE, Sladek FM. (2000) Analysis of protein dimerization and ligand binding of orphan receptor HNF4alpha. J Mol Biol, 302(4): 831-51.

21. Dhe-Paganon S, Duda K, Iwamoto M, Chi YI, Shoelson SE. (2002) Crystal structure of the HNF4 alpha ligand binding domain in complex with endogenous fatty acid ligand. J Biol Chem, 277(41): 37973-6.
22. Wisely GB, Miller AB, Davis RG, Thornquest AD, Jr., Johnson R, Spitzer T, Sefler A, Shearer B, Moore JT, Miller AB, Willson TM, Williams SP. (2002) Hepatocyte nuclear factor 4 is a transcription factor that constitutively binds fatty acids. Structure, 10(9): 1225-34.
23. Yokoyama A, Katsura S, Ito R, Hashiba W, Sekine H, Fujiki R, Kato S. (2011) Multiple post-translational modifications in hepatocyte nuclear factor 4alpha. Biochem Biophys Res Commun, 410(4): 749-53.
24. Wang Z, Salih E, Burke PA. (2011) Quantitative analysis of cytokine-induced hepatocyte nuclear factor-4alpha phosphorylation by mass spectrometry. Biochemistry, 50(23): 5292-300.
25. Caron S, Huaman Samanez C, Dehondt H, Ploton M, Briand O, Lien F, Dorchies E, Dumont J, Postic C, Cariou B, Lefebvre P, Staels B. (2013) Farnesoid X receptor inhibits the transcriptional activity of carbohydrate response element binding protein in human hepatocytes. Mol Cell Biol, 33(11): 2202-11.
26. Ding X, Lichti K, Kim I, Gonzalez FJ, Staudinger JL. (2006) Regulation of constitutive androstane receptor and its target genes by fasting, cAMP, hepatocyte nuclear factor alpha, and the coactivator peroxisome proliferator-activated receptor gamma coactivator-1alpha. J Biol Chem, 281(36): 26540-51.
27. Sun K, Montana V, Chellappa K, Brelivet Y, Moras D, Maeda Y, Parpura V, Paschal BM, Sladek FM. (2007) Phosphorylation of a conserved serine in the deoxyribonucleic acid binding domain of nuclear receptors alters intracellular localization. Mol Endocrinol, 21(6): 1297-311.
28. Viollet B, Kahn A, Raymondjean M. (1997) Protein kinase A-dependent phosphorylation modulates DNA-binding activity of hepatocyte nuclear factor 4. Mol Cell Biol, 17(8): 4208-19.
29. Guo H, Gao C, Mi Z, Zhang J, Kuo PC. (2007) Characterization of the PC4 binding domain and its interactions with HNF4alpha. J Biochem, 141(5): 635-40.

30. Hong YH, Varanasi US, Yang W, Leff T. (2003) AMP-activated protein kinase regulates HNF4alpha transcriptional activity by inhibiting dimer formation and decreasing protein stability. J Biol Chem, 278(30): 27495-501.
31. Chandra V, Huang P, Potluri N, Wu D, Kim Y, Rastinejad F. (2013) Multidomain integration in the structure of the HNF-4alpha nuclear receptor complex. Nature, 495(7441): 394-8.
32. Aranyi T, Bacquet C, de Boussac H, Ratajewski M, Pomozi V, Fulop K, Brampton CN, Pulaski L, Le Saux O, Varadi A. (2013) Transcriptional regulation of the ABCC6 gene and the background of impaired function of missense disease-causing mutations. Front Genet, 4: 27.
33. de Boussac H, Ratajewski M, Sachrajda I, Koblos G, Tordai A, Pulaski L, Buday L, Varadi A, Aranyi T. (2010) The ERK1/2-hepatocyte nuclear factor 4alpha axis regulates human ABCC6 gene expression in hepatocytes. J Biol Chem, 285(30): 22800-8.
34. John RM, Surani MA. (1996) Imprinted genes and regulation of gene expression by epigenetic inheritance. Curr Opin Cell Biol, 8(3): 348-53.
35. Pagliaroli L, Veto B, Aranyi T, Barta C. (2016) From Genetics to Epigenetics: New Perspectives in Tourette Syndrome Research. Front Neurosci, 10: 277.
36. Yamagata Y, Szabo P, Szuts D, Bacquet C, Aranyi T, Paldi A. (2012) Rapid turnover of DNA methylation in human cells. Epigenetics, 7(2): 141-5.
37. Aranyi T, Stockholm D, Yao R, Poinsignon C, Wiart T, Corre G, Touleimat N, Tost J, Galy A, Paldi A. (2016) Systemic epigenetic response to recombinant lentiviral vectors independent of proviral integration. Epigenetics Chromatin, 9: 29.
38. Allfrey VG, Faulkner R, Mirsky AE. (1964) ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. Proc Natl Acad Sci U S A, 51: 786-94.
39. Kouzarides T. (2007) Chromatin Modifications and Their Function. Cell, 128(4): 693-705.
40. Bannister AJ, Kouzarides T. (2011) Regulation of chromatin by histone modifications. Cell Res, 21(3): 381-95.

41. Kuo MH, Allis CD. (1998) Roles of histone acetyltransferases and deacetylases in gene regulation. Bioessays, 20(8): 615-26.
42. Legube G, Trouche D. (2003) Regulating histone acetyltransferases and deacetylases. EMBO reports, 4(10): 944-947.
43. Sakamoto S, Potla R, Larner AC. (2004) Histone deacetylase activity is required to recruit RNA polymerase II to the promoters of selected interferon-stimulated early response genes. J Biol Chem, 279(39): 40362-7.
44. Roelfsema JH, White SJ, Ariyurek Y, Bartholdi D, Niedrist D, Papadia F, Bacino CA, den Dunnen JT, van Ommen GJ, Breuning MH, Hennekam RC, Peters DJ. (2005) Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in both the CBP and EP300 genes cause disease. Am J Hum Genet, 76(4): 572-80.
45. Greenblatt SM, Liu F, Nimer SD. (2016) Arginine methyltransferases in normal and malignant hematopoiesis. Exp Hematol, 44(6): 435-41.
46. Jahan S, Davie JR. (2015) Protein arginine methyltransferases (PRMTs): role in chromatin organization. Adv Biol Regul, 57: 173-84.
47. Ziller MJ, Muller F, Liao J, Zhang Y, Gu H, Bock C, Boyle P, Epstein CB, Bernstein BE, Lengauer T, Gnirke A, Meissner A. (2011) Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. PLoS Genet, 7(12): e1002389.
48. Bird AP. (1993) Functions for DNA methylation in vertebrates. Cold Spring Harb Symp Quant Biol, 58: 281-5.
49. Deaton AM, Bird A. (2011) CpG islands and the regulation of transcription. Genes Dev, 25(10): 1010-22.
50. Zeng J, Nagrajan HK, Yi SV. (2014) Fundamental diversity of human CpG islands at multiple biological levels. Epigenetics, 9(4): 483-91.
51. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature, 405(6785): 486-9.
52. Smith ZD, Meissner A. (2013) DNA methylation: roles in mammalian development. Nat Rev Genet, 14(3): 204-20.

53. Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, Wolffe AP. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat Genet, 19(2): 187-91.
54. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat Genet, 23(2): 185-8.
55. Seisenberger S, Peat JR, Hore TA, Santos F, Dean W, Reik W. (2013) Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers. Philos Trans R Soc Lond B Biol Sci, 368(1609): 20110330.
56. Isagawa T, Nagae G, Shiraki N, Fujita T, Sato N, Ishikawa S, Kume S, Aburatani H. (2011) DNA methylation profiling of embryonic stem cell differentiation into the three germ layers. PLoS One, 6(10): e26052.
57. Yamagata Y, Parietti V, Stockholm D, Corre G, Poinsignon C, Touleimat N, Delafoy D, Besse C, Tost J, Galy A, Paldi A. (2012) Lentiviral transduction of CD34(+) cells induces genome-wide epigenetic modifications. PLoS One, 7(11): e48943.
58. Guo JU, Ma DK, Mo H, Ball MP, Jang MH, Bonaguidi MA, Balazer JA, Eaves HL, Xie B, Ford E, Zhang K, Ming GL, Gao Y, Song H. (2011) Neuronal activity modifies the DNA methylation landscape in the adult brain. Nat Neurosci, 14(10): 1345-51.
59. Ressler KJ, Mercer KB, Bradley B, Jovanovic T, Mahan A, Kerley K, Norrholm SD, Kilaru V, Smith AK, Myers AJ, Ramirez M, Engel A, Hammack SE, Toufexis D, Braas KM, Binder EB, May V. (2011) Post-traumatic stress disorder is associated with PACAP and the PAC1 receptor. Nature, 470(7335): 492-7.
60. Weaver IC, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR, Dymov S, Szyf M, Meaney MJ. (2004) Epigenetic programming by maternal behavior. Nat Neurosci, 7(8): 847-54.
61. Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K, Sun YE, Zhang Y. (2011) Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual

function in transcriptional regulation in mouse embryonic stem cells. Genes Dev, 25(7): 679-84.

62. Kang J, Lienhard M, Pastor WA, Chawla A, Novotny M, Tsagaratou A, Lasken RS, Thompson EC, Surani MA, Koralov SB, Kalantry S, Chavez L, Rao A. (2015) Simultaneous deletion of the methylcytosine oxidases Tet1 and Tet3 increases transcriptome variability in early embryogenesis. Proc Natl Acad Sci U S A, 112(31): E4236-45.
63. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature, 466(7310): 1129-33.
64. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science, 333(6047): 1300-3.
65. Zhang H, Zhang X, Clark E, Mulcahey M, Huang S, Shi YG. (2010) TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. Cell Res, 20(12): 1390-3.
66. Dawlaty MM, Ganz K, Powell BE, Hu YC, Markoulaki S, Cheng AW, Gao Q, Kim J, Choi SW, Page DC, Jaenisch R. (2011) Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. Cell Stem Cell, 9(2): 166-75.
67. Li T, Yang D, Li J, Tang Y, Yang J, Le W. (2015) Critical role of Tet3 in neural progenitor cell maintenance and terminal differentiation. Mol Neurobiol, 51(1): 142-54.
68. Ono R, Taki T, Taketani T, Taniwaki M, Kobayashi H, Hayashi Y. (2002) LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). Cancer Res, 62(14): 4075-80.
69. Ko M, An J, Rao A. (2015) DNA methylation and hydroxymethylation in hematologic differentiation and transformation. Curr Opin Cell Biol, 37: 91-101.
70. Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, Miller CR, Ding L, Golub T, Mesirov JP, Alexe G, Lawrence M, O'Kelly M, Tamayo P, Weir BA, Gabriel S, Winckler W, Gupta S, Jakkula L, Feiler HS, Hodgson

JG, James CD, Sarkaria JN, Brennan C, Kahn A, Spellman PT, Wilson RK, Speed TP, Gray JW, Meyerson M, Getz G, Perou CM, Hayes DN. (2010) Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell, 17(1): 98-110.

71. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, Li Y, Bhagwat N, Vasanthakumar A, Fernandez HF, Tallman MS, Sun Z, Wolniak K, Peeters JK, Liu W, Choe SE, Fantin VR, Paietta E, Lowenberg B, Licht JD, Godley LA, Delwel R, Valk PJ, Thompson CB, Levine RL, Melnick A. (2010) Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell, 18(6): 553-67.
72. Blaschke K, Ebata KT, Karimi MM, Zepeda-Martinez JA, Goyal P, Mahapatra S, Tam A, Laird DJ, Hirst M, Rao A, Lorincz MC, Ramalho-Santos M. (2013) Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature, 500(7461): 222-6.
73. Minor EA, Court BL, Young JI, Wang G. (2013) Ascorbate induces ten-eleven translocation (Tet) methylcytosine dioxygenase-mediated generation of 5-hydroxymethylcytosine. J Biol Chem, 288(19): 13669-74.
74. Baylin SB, Jones PA. (2011) A decade of exploring the cancer epigenome - biological and translational implications. Nat Rev Cancer, 11(10): 726-34.
75. You JS, Jones PA. (2012) Cancer genetics and epigenetics: two sides of the same coin? Cancer Cell, 22(1): 9-20.
76. Rideout WM, 3rd, Coetzee GA, Olumi AF, Jones PA. (1990) 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. Science, 249(4974): 1288-90.
77. Kanai Y, Ushijima S, Nakanishi Y, Sakamoto M, Hirohashi S. (2003) Mutation of the DNA methyltransferase (DNMT) 1 gene in human colorectal cancers. Cancer Lett, 192(1): 75-82.
78. Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, Kandoth C, Payton JE, Baty J, Welch J, Harris CC, Lichti CF, Townsend RR, Fulton RS, Dooling DJ, Koboldt DC, Schmidt H, Zhang Q, Osborne JR, Lin L, O'Laughlin

M, McMichael JF, Delehaunty KD, McGrath SD, Fulton LA, Magrini VJ, Vickery TL, Hundal J, Cook LL, Conyers JJ, Swift GW, Reed JP, Alldredge PA, Wylie T, Walker J, Kalicki J, Watson MA, Heath S, Shannon WD, Varghese N, Nagarajan R, Westervelt P, Tomasson MH, Link DC, Graubert TA, DiPersio JF, Mardis ER, Wilson RK. (2010) DNMT3A mutations in acute myeloid leukemia. N Engl J Med, 363(25): 2424-33.

79. Wu Y, Strawn E, Basir Z, Halverson G, Guo SW. (2007) Aberrant expression of deoxyribonucleic acid methyltransferases DNMT1, DNMT3A, and DNMT3B in women with endometriosis. Fertil Steril, 87(1): 24-32.
80. Shankar S, Srivastava RK. (2008) Histone deacetylase inhibitors: mechanisms and clinical significance in cancer: HDAC inhibitor-induced apoptosis. Adv Exp Med Biol, 615: 261-98.
81. Hennessy BT, Garcia-Manero G, Kantarjian HM, Giles FJ. (2003) DNA methylation in haematological malignancies: the role of decitabine. Expert Opin Investig Drugs, 12(12): 1985-93.
82. Laird PW. (2010) Principles and challenges of genomewide DNA methylation analysis. Nat Rev Genet, 11(3): 191-203.
83. Clark SJ, Harrison J, Paul CL, Frommer M. (1994) High sensitivity mapping of methylated cytosines. Nucleic Acids Research, 22(15): 2990-2997.
84. Jin SG, Kadam S, Pfeifer GP. (2010) Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. Nucleic Acids Res, 38(11): e125.
85. Robertson AB, Dahl JA, Vågbø CB, Tripathi P, Krokan HE, Klungland A. (2011) A novel method for the efficient and selective identification of 5-hydroxymethylcytosine in genomic DNA. Nucleic Acids Research, 39(8): e55-e55.
86. Matarese F, Carrillo-de Santa Pau E, Stunnenberg HG. (2011) 5-Hydroxymethylcytosine: a new kid on the epigenetic block? Mol Syst Biol, 7: 562.
87. Globisch D, Munzel M, Muller M, Michalakis S, Wagner M, Koch S, Bruckl T, Biel M, Carell T. (2010) Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. PLoS One, 5(12): e15367.

88. Le T, Kim KP, Fan G, Faull KF. (2011) A sensitive mass spectrometry method for simultaneous quantification of DNA methylation and hydroxymethylation levels in biological samples. Anal Biochem, 412(2): 203-9.
89. Veto B, Szabo P, Bacquet C, Apro A, Hathy E, Kiss J, Rethelyi JM, Szeri F, Szuts D, Aranyi T. (2018) Inhibition of DNA methyltransferase leads to increased genomic 5-hydroxymethylcytosine levels in hematopoietic cells. FEBS Open Bio, 8(4): 584-592.
90. Etchegaray JP, Mostoslavsky R. (2016) Interplay between Metabolism and Epigenetics: A Nuclear Adaptation to Environmental Changes. Mol Cell, 62(5): 695-711.
91. Mentch SJ, Merhmohamadi M, Huang L, Liu X, Gupta D, Mattocks D, Gomez P, Ables G, Bamman MM, Thalacker-Mercer AE, Nichenametla S, Locasale JW. (2015) Histone Methylation Dynamics and Gene Regulation Occur through the Sensing of One-Carbon Metabolism. Cell metabolism, 22(5): 861-873.
92. Dang L, White DW, Gross S, Bennett BD, Bittinger MA, Driggers EM, Fantin VR, Jang HG, Jin S, Keenan MC, Marks KM, Prins RM, Ward PS, Yen KE, Liau LM, Rabinowitz JD, Cantley LC, Thompson CB, Vander Heiden MG, Su SM. (2009) Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature, 462(7274): 739-44.
93. Xiao M, Yang H, Xu W, Ma S, Lin H, Zhu H, Liu L, Liu Y, Yang C, Xu Y, Zhao S, Ye D, Xiong Y, Guan KL. (2012) Inhibition of alpha-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. Genes Dev, 26(12): 1326-38.
94. Larsson SC, Wolk A. (2007) Overweight, obesity and risk of liver cancer: a meta-analysis of cohort studies. Br J Cancer, 97(7): 1005-8.
95. O'Neill S, O'Driscoll L. (2015) Metabolic syndrome: a closer look at the growing epidemic and its associated pathologies. Obes Rev, 16(1): 1-12.
96. WHO. 06/02/2018]; Available from: http://www.who.int/en/.
97. Kaati G, Bygren LO, Pembrey M, Sjostrom M. (2007) Transgenerational response to nutrition, early life circumstances and longevity. Eur J Hum Genet, 15(7): 784-90.

98. Ng SF, Lin RC, Laybutt DR, Barres R, Owens JA, Morris MJ. (2010) Chronic high-fat diet in fathers programs beta-cell dysfunction in female rat offspring. Nature, 467(7318): 963-6.
99. Eckel-Mahan KL, Patel VR, de Mateo S, Orozco-Solis R, Ceglia NJ, Sahar S, Dilag-Penilla SA, Dyar KA, Baldi P, Sassone-Corsi P. (2013) Reprogramming of the circadian clock by nutritional challenge. Cell, 155(7): 1464-78.
100. McKay JA, Xie L, Manus C, Langie SA, Maxwell RJ, Ford D, Mathers JC. (2014) Metabolic effects of a high-fat diet post-weaning after low maternal dietary folate during pregnancy and lactation. Mol Nutr Food Res, 58(5): 1087-97.
101. Poston L, Igosheva N, Mistry HD, Seed PT, Shennan AH, Rana S, Karumanchi SA, Chappell LC. (2011) Role of oxidative stress and antioxidant supplementation in pregnancy disorders. Am J Clin Nutr, 94(6 Suppl): 1980s-1985s.
102. Sloboda DM, Howie GJ, Pleasants A, Gluckman PD, Vickers MH. (2009) Pre- and postnatal nutritional histories influence reproductive maturation and ovarian function in the rat. PLoS One, 4(8): e6744.
103. Blondeau B, Joly B, Perret C, Prince S, Bruneval P, Lelievre-Pegorier M, Fassot C, Duong Van Huyen JP. (2011) Exposure in utero to maternal diabetes leads to glucose intolerance and high blood pressure with no major effects on lipid metabolism. Diabetes Metab, 37(3): 245-51.
104. Yang A, Sun Y, Mao C, Yang S, Huang M, Deng M, Ding N, Yang X, Zhang M, Jin S, Jiang Y, Huang Y. (2017) Folate Protects Hepatocytes of Hyperhomocysteinemia Mice From Apoptosis via Cystic Fibrosis Transmembrane Conductance Regulator (CFTR)-Activated Endoplasmic Reticulum Stress. J Cell Biochem, 118(9): 2921-2932.
105. Barnett MP, Bermingham EN, Young W, Bassett SA, Hesketh JE, Maciel-Dominguez A, McNabb WC, Roy NC. (2015) Low folate and selenium in the mouse maternal diet alters liver gene expression patterns in the offspring after weaning. Nutrients, 7(5): 3370-86.
106. McKay JA, Mathers JC. (2016) Maternal folate deficiency and metabolic dysfunction in offspring. Proc Nutr Soc, 75(1): 90-95.

107. Hoile SP, Lillycrop KA, Grenfell LR, Hanson MA, Burdge GC. (2012) Increasing the folic acid content of maternal or post-weaning diets induces differential changes in phosphoenolpyruvate carboxykinase mRNA expression and promoter methylation in rats. Br J Nutr, 108(5): 852-7.
108. Geisler CE, Hepler C, Higgins MR, Renquist BJ. (2016) Hepatic adaptations to maintain metabolic homeostasis in response to fasting and refeeding in mice. Nutr Metab (Lond), 13: 62.
109. Kim SJ, Kim JE, Kim YW, Kim JY, Park SY. (2017) Nutritional regulation of renal lipogenic factor expression in mice: comparison to regulation in the liver and skeletal muscle. Am J Physiol Renal Physiol, 313(4): F887-f898.
110. Ou J, Tu H, Shan B, Luk A, DeBose-Boyd RA, Bashmakov Y, Goldstein JL, Brown MS. (2001) Unsaturated fatty acids inhibit transcription of the sterol regulatory element-binding protein-1c (SREBP-1c) gene by antagonizing ligand-dependent activation of the LXR. Proc Natl Acad Sci U S A, 98(11): 6027-32.
111. Yabe D, Komuro R, Liang G, Goldstein JL, Brown MS. (2003) Liver-specific mRNA for Insig-2 down-regulated by insulin: implications for fatty acid synthesis. Proc Natl Acad Sci U S A, 100(6): 3155-60.
112. Oh KJ, Han HS, Kim MJ, Koo SH. (2013) Transcriptional regulators of hepatic gluconeogenesis. Arch Pharm Res, 36(2): 189-200.
113. Marieb EN, Hoehn K, Hutchinson M, *Human anatomy & physiology*. 2013, [San Francisco, Calif.]: Pearson Education/Benjamin Cummings.
114. Goldstein I, Hager GL. (2015) Transcriptional and Chromatin Regulation during Fasting - The Genomic Era. Trends Endocrinol Metab, 26(12): 699-710.
115. Iynedjian PB, Pilot PR, Nouspikel T, Milburn JL, Quaade C, Hughes S, Ucla C, Newgard CB. (1989) Differential expression and regulation of the glucokinase gene in liver and islets of Langerhans. Proc Natl Acad Sci U S A, 86(20): 7838-42.
116. Hirota K, Daitoku H, Matsuzaki H, Araya N, Yamagata K, Asada S, Sugaya T, Fukamizu A. (2003) Hepatocyte nuclear factor-4 is a novel downstream target of insulin via FKHR as a signal-regulated transcriptional inhibitor. J Biol Chem, 278(15): 13056-60.

117. Hirota K, Sakamaki J, Ishida J, Shimamoto Y, Nishihara S, Kodama N, Ohta K, Yamamoto M, Tanimoto K, Fukamizu A. (2008) A combination of HNF-4 and Foxo1 is required for reciprocal transcriptional regulation of glucokinase and glucose-6-phosphatase genes in response to fasting and feeding. J Biol Chem, 283(47): 32432-41.
118. Bergot MO, Diaz-Guerra MJ, Puzenat N, Raymondjean M, Kahn A. (1992) Cis-regulation of the L-type pyruvate kinase gene promoter by glucose, insulin and cyclic AMP. Nucleic Acids Res, 20(8): 1871-7.
119. Rhee J, Inoue Y, Yoon JC, Puigserver P, Fan M, Gonzalez FJ, Spiegelman BM. (2003) Regulation of hepatic fasting response by PPARgamma coactivator-1alpha (PGC-1): requirement for hepatocyte nuclear factor 4alpha in gluconeogenesis. Proc Natl Acad Sci U S A, 100(7): 4012-7.
120. Lee JM, Seo WY, Han HS, Oh KJ, Lee YS, Kim DK, Choi S, Choi BH, Harris RA, Lee CH, Koo SH, Choi HS. (2016) Insulin-Inducible SMILE Inhibits Hepatic Gluconeogenesis. Diabetes, 65(1): 62-73.
121. De Fabiani E, Mitro N, Anzulovich AC, Pinelli A, Galli G, Crestani M. (2001) The negative effects of bile acids and tumor necrosis factor-alpha on the transcription of cholesterol 7alpha-hydroxylase gene (CYP7A1) converge to hepatic nuclear factor-4: a novel mechanism of feedback regulation of bile acid synthesis mediated by nuclear receptors. J Biol Chem, 276(33): 30708-16.
122. De Fabiani E, Mitro N, Gilardi F, Caruso D, Galli G, Crestani M. (2003) Coordinated control of cholesterol catabolism to bile acids and of gluconeogenesis via a novel mechanism of transcription regulation linked to the fasted-to-fed cycle. J Biol Chem, 278(40): 39124-32.
123. Crestani M, Sadeghpour A, Stroup D, Galli G, Chiang JY. (1998) Transcriptional activation of the cholesterol 7alpha-hydroxylase gene (CYP7A) by nuclear hormone receptors. J Lipid Res, 39(11): 2192-200.
124. Song KH, Chiang JY. (2006) Glucagon and cAMP inhibit cholesterol 7alpha-hydroxylase (CYP7A1) gene expression in human hepatocytes: discordant regulation of bile acid synthesis and gluconeogenesis. Hepatology, 43(1): 117-25.

125. Shin DJ, McGrane MM. (1997) Vitamin A regulates genes involved in hepatic gluconeogenesis in mice: phosphoenolpyruvate carboxykinase, fructose-1,6-bisphosphatase and 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase. J Nutr, 127(7): 1274-8.
126. Goldstein I, Baek S, Presman DM, Paakinaho V, Swinstead EE, Hager GL. (2017) Transcription factor assisted loading and enhancer dynamics dictate the hepatic fasting response. Genome Res, 27(3): 427-439.
127. Wang Y, Viscarra J, Kim SJ, Sul HS. (2015) Transcriptional regulation of hepatic lipogenesis. Nat Rev Mol Cell Biol, 16(11): 678-89.
128. Horton JD, Bashmakov Y, Shimomura I, Shimano H. (1998) Regulation of sterol regulatory element binding proteins in livers of fasted and refed mice. Proc Natl Acad Sci U S A, 95(11): 5987-92.
129. Tusnady GE, Simon I, Varadi A, Aranyi T. (2005) BiSearch: primer-design and search tool for PCR on bisulfite-treated genomes. Nucleic Acids Res, 33(1): e9.
130. Smith ZD, Gu H, Bock C, Gnirke A, Meissner A. (2009) High-throughput bisulfite sequencing in mammalian genomes. Methods, 48(3): 226-32.
131. Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res, 33(18): 5868-77.
132. Maeso I, Dunwell TL, Wyatt CD, Marletaz F, Veto B, Bernal JA, Quah S, Irimia M, Holland PW. (2016) Evolutionary origin and functional divergence of totipotent cell homeobox genes in eutherian mammals. BMC Biol, 14: 45.
133. Veto B, Bojcsuk D, Bacquet C, Kiss J, Sipeki S, Martin L, Buday L, Balint BL, Aranyi T. (2017) The transcriptional activity of hepatocyte nuclear factor 4 alpha is inhibited via phosphorylation by ERK1/2. PLoS One, 12(2): e0172020.
134. Lorenz K, Schmitt JP, Schmitteckert EM, Lohse MJ. (2009) A new type of ERK1/2 autophosphorylation causes cardiac hypertrophy. Nat Med, 15(1): 75-83.
135. Bolotin E, Liao H, Ta TC, Yang C, Hwang-Verslues W, Evans JR, Jiang T, Sladek FM. (2010) Integrated approach for the identification of human hepatocyte nuclear factor 4alpha target genes using protein binding microarrays. Hepatology, 51(2): 642-53.

136. Illes A, Enyedi B, Tamas P, Balazs A, Bogel G, Buday L. (2006) Inducible phosphorylation of cortactin is not necessary for cortactin-mediated actin polymerisation. Cell Signal, 18(6): 830-40.
137. Chowdhury B, McGovern A, Cui Y, Choudhury SR, Cho IH, Cooper B, Chevassut T, Lossie AC, Irudayaraj J. (2015) The hypomethylating agent Decitabine causes a paradoxical increase in 5-hydroxymethylcytosine in human leukemia cells. Sci Rep, 5: 9281.
138. Kari V, Karpiuk O, Tieg B, Kriegs M, Dikomey E, Krebber H, Begus-Nahrmann Y, Johnsen SA. (2013) A subset of histone H2B genes produces polyadenylated mRNAs under a variety of cellular conditions. PLoS One, 8(5): e63745.
139. Yan B, Yang X, Lee TL, Friedman J, Tang J, Van Waes C, Chen Z. (2007) Genome-wide identification of novel expression signatures reveal distinct patterns and prevalence of binding motifs for p53, nuclear factor-kappaB and other signal transcription factors in head and neck squamous cell carcinoma. Genome Biol, 8(5): R78.
140. Struhl K. (1998) Histone acetylation and transcriptional regulatory mechanisms. Genes Dev, 12(5): 599-606.
141. Aarenstrup L, Falch AM, Jakobsen KK, Neve S, Henriksen LL, Tommerup N, Leffers H, Kristiansen K. (2002) Expression and post-translational modification of human 4-hydroxy-phenylpyruvate dioxygenase. Cell Biol Int, 26(7): 615-25.
142. Malik S. (2003) Transcriptional regulation of the apolipoprotein AI gene. Front Biosci, 8: d360-8.
143. Jansen RS, Kucukosmanoglu A, de Haas M, Sapthu S, Otero JA, Hegman IE, Bergen AA, Gorgels TG, Borst P, van de Wetering K. (2013) ABCC6 prevents ectopic mineralization seen in pseudoxanthoma elasticum by inducing cellular nucleotide release. Proc Natl Acad Sci U S A, 110(50): 20206-11.
144. Hatzis P, Kyrmizi I, Talianidis I. (2006) Mitogen-activated protein kinase-mediated disruption of enhancer-promoter communication inhibits hepatocyte nuclear factor 4alpha expression. Mol Cell Biol, 26(19): 7017-29.
145. Gonzalez FJ. (2008) Regulation of hepatocyte nuclear factor 4 alpha-mediated transcription. Drug Metab Pharmacokinet, 23(1): 2-7.

146. Song Y, Zheng D, Zhao M, Qin Y, Wang T, Xing W, Gao L, Zhao J. (2015) Thyroid-Stimulating Hormone Increases HNF-4alpha Phosphorylation via cAMP/PKA Pathway in the Liver. Sci Rep, 5: 13409.
147. Xu Z, Tavares-Sanchez OL, Li Q, Fernando J, Rodriguez CM, Studer EJ, Pandak WM, Hylemon PB, Gil G. (2007) Activation of bile acid biosynthesis by the p38 mitogen-activated protein kinase (MAPK): hepatocyte nuclear factor-4alpha phosphorylation by the p38 MAPK is required for cholesterol 7alpha-hydroxylase expression. J Biol Chem, 282(34): 24607-14.
148. Puigserver P, Rhee J, Lin J, Wu Z, Yoon JC, Zhang CY, Krauss S, Mootha VK, Lowell BB, Spiegelman BM. (2001) Cytokine stimulation of energy expenditure through p38 MAP kinase activation of PPARgamma coactivator-1. Mol Cell, 8(5): 971-82.
149. Strub GM, Maceyka M, Hait NC, Milstien S, Spiegel S. (2010) Extracellular and intracellular actions of sphingosine-1-phosphate. Adv Exp Med Biol, 688: 141-55.
150. Cao R, Cronk ZX, Zha W, Sun L, Wang X, Fang Y, Studer E, Zhou H, Pandak WM, Dent P, Gil G, Hylemon PB. (2010) Bile acids regulate hepatic gluconeogenic genes and farnesoid X receptor via G(alpha)i-protein-coupled receptors and the AKT pathway. J Lipid Res, 51(8): 2234-44.
151. Qiao L, Studer E, Leach K, McKinstry R, Gupta S, Decker R, Kukreja R, Valerie K, Nagarkatti P, El Deiry W, Molkentin J, Schmidt-Ullrich R, Fisher PB, Grant S, Hylemon PB, Dent P. (2001) Deoxycholic acid (DCA) causes ligand-independent activation of epidermal growth factor receptor (EGFR) and FAS receptor in primary hepatocytes: inhibition of EGFR/mitogen-activated protein kinase-signaling module enhances DCA-induced apoptosis. Mol Biol Cell, 12(9): 2629-45.
152. Kir S, Zhang Y, Gerard RD, Kliewer SA, Mangelsdorf DJ. (2012) Nuclear receptors HNF4alpha and LRH-1 cooperate in regulating Cyp7a1 in vivo. J Biol Chem, 287(49): 41334-41.
153. Han S, Li T, Ellis E, Strom S, Chiang JY. (2010) A novel bile acid-activated vitamin D receptor signaling in human hepatocytes. Mol Endocrinol, 24(6): 1151-64.

154. Matsubara T, Li F, Gonzalez FJ. (2013) FXR signaling in the enterohepatic system. Mol Cell Endocrinol, 368(1-2): 17-29.
155. Chiang JY. (2004) Regulation of bile acid synthesis: pathways, nuclear receptors, and mechanisms. J Hepatol, 40(3): 539-51.
156. Lee YK, Dell H, Dowhan DH, Hadzopoulou-Cladaras M, Moore DD. (2000) The orphan nuclear receptor SHP inhibits hepatocyte nuclear factor 4 and retinoid X receptor transactivation: two mechanisms for repression. Mol Cell Biol, 20(1): 187-95.
157. Li T, Chiang JYL. (2014) Bile Acid Signaling in Metabolic Disease and Drug Therapy. Pharmacological Reviews, 66(4): 948.
158. Lin CY, Hu CT, Cheng CC, Lee MC, Pan SM, Lin TY, Wu WS. (2016) Oxidation of heat shock protein 60 and protein disulfide isomerase activates ERK and migration of human hepatocellular carcinoma HepG2. Oncotarget, 7(10): 11067-82.
159. Ripperger JA, Jud C, Albrecht U. (2011) The daily rhythm of mice. FEBS Lett, 585(10): 1384-92.
160. Wang C, Xu CX, Krager SL, Bottum KM, Liao DF, Tischkau SA. (2011) Aryl hydrocarbon receptor deficiency enhances insulin sensitivity and reduces PPAR-alpha pathway activity in mice. Environ Health Perspect, 119(12): 1739-44.
161. Dufour CR, Levasseur MP, Pham NH, Eichner LJ, Wilson BJ, Charest-Marcotte A, Duguay D, Poirier-Heon JF, Cermakian N, Giguere V. (2011) Genomic convergence among ERRalpha, PROX1, and BMAL1 in the control of metabolic clock outputs. PLoS Genet, 7(6): e1002143.
162. Shen N, Jiang S, Lu JM, Yu X, Lai SS, Zhang JZ, Zhang JL, Tao WW, Wang XX, Xu N, Xue B, Li CJ. (2015) The constitutive activation of Egr-1/C/EBPa mediates the development of type 2 diabetes mellitus by enhancing hepatic gluconeogenesis. Am J Pathol, 185(2): 513-23.
163. Szalay MS, Kovacs IA, Korcsmaros T, Bode C, Csermely P. (2007) Stress-induced rearrangements of cellular networks: Consequences for protection and drug design. FEBS Lett, 581(19): 3675-80.

164. Saxonov S, Berg P, Brutlag DL. (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci U S A, 103(5): 1412-7.
165. Bird A, Taggart M, Frommer M, Miller OJ, Macleod D. (1985) A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell, 40(1): 91-9.
166. Veillard A-C, Datlinger P, Laczik M, Squazzo S, Bock C. (2016) Diagenode® Premium RRBS technology: cost-effective DNA methylation mapping with superior coverage. Nature Methods, 13.
167. Oiso H, Furukawa N, Suefuji M, Shimoda S, Ito A, Furumai R, Nakagawa J, Yoshida M, Nishino N, Araki E. (2011) The role of class I histone deacetylase (HDAC) on gluconeogenesis in liver. Biochem Biophys Res Commun, 404(1): 166-72.
168. Ivanov M, Kals M, Kacevska M, Barragan I, Kasuga K, Rane A, Metspalu A, Milani L, Ingelman-Sundberg M. (2013) Ontogeny, distribution and potential roles of 5-hydroxymethylcytosine in human liver function. Genome Biology, 14(8): R83-R83.
169. Kinde B, Gabel HW, Gilbert CS, Griffith EC, Greenberg ME. (2015) Reading the unique DNA methylation landscape of the brain: Non-CpG methylation, hydroxymethylation, and MeCP2. Proc Natl Acad Sci U S A, 112(22): 6800-6.
170. Arányi T, Kerjean A, Tóth S, Mallet J, Meloni R, Páldi A. (2002) Paradoxical Methylation of the tyrosine hydroxylase Gene in Mouse Preimplantation Embryos. Genomics, 80(6): 558-563.
171. Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R. (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proceedings of the National Academy of Sciences, 97(10): 5237-5242.
172. Nishino K, Toyoda M, Yamazaki-Inoue M, Fukawatase Y, Chikazawa E, Sakaguchi H, Akutsu H, Umezawa A. (2011) DNA Methylation Dynamics in Human Induced Pluripotent Stem Cells over Time. PLoS Genetics, 7(5): e1002085.

173. Kriaucionis S, Heintz N. (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science, 324(5929): 929-30.

174. Lin IH, Chen Y-F, Hsu M-T. (2017) Correlated 5-Hydroxymethylcytosine (5hmC) and Gene Expression Profiles Underpin Gene and Organ-Specific Epigenetic Regulation in Adult Mouse Brain and Liver. PLoS ONE, 12(1): e0170779.

175. Nestor CE, Ottaviano R, Reddington J, Sproul D, Reinhardt D, Dunican D, Katz E, Dixon JM, Harrison DJ, Meehan RR. (2012) Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. Genome Research, 22(3): 467-477.

176. Wheldon LM, Abakir A, Ferjentsik Z, Dudnakova T, Strohbuecker S, Christie D, Dai N, Guan S, Foster JM, Correa IR, Jr., Loose M, Dixon JE, Sottile V, Johnson AD, Ruzov A. (2014) Transient accumulation of 5-carboxylcytosine indicates involvement of active demethylation in lineage specification of neural stem cells. Cell Rep, 7(5): 1353-1361.

177. Sajadian SO, Tripura C, Samani FS, Ruoss M, Dooley S, Baharvand H, Nussler AK. (2016) Vitamin C enhances epigenetic modifications induced by 5-azacytidine and cell cycle arrest in the hepatocellular carcinoma cell lines HLE and Huh7. Clinical Epigenetics, 8: 46.

178. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson A, Kampf C, Sjostedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J, Ponten F. (2015) Proteomics. Tissue-based map of the human proteome. Science, 347(6220): 1260419.

179. Han JA, An J, Ko M. (2015) Functions of TET Proteins in Hematopoietic Transformation. Mol Cells, 38(11): 925-35.

180. Van Damme M, Crompot E, Meuleman N, Maerevoet M, Mineur P, Bron D, Lagneaux L, Stamatopoulos B. (2016) Characterization of TET and IDH gene

expression in chronic lymphocytic leukemia: comparison with normal B cells and prognostic significance. Clin Epigenetics, 8: 132.

181. Nie J, Liu L, Li X, Han W. (2014) Decitabine, a new star in epigenetic therapy: the clinical application and biological mechanism in solid tumors. Cancer Lett, 354(1): 12-20.

DOI:10.14753/SE.2018.2176

11. BIBLIOGRAPHY OF THE CANDIDATE'S PUBLICATIONS

11.1. Publications related to the thesis

Vető B, Szabo P, Bacquet C, Apro A, Hathy E, Kiss J, Rethelyi JM, Szeri F, Szuts D, Aranyi T. (2018) Inhibition of DNA methyltransferase leads to increased genomic 5-hydroxymethylcytosine levels in hematopoietic cells. FEBS Open Bio, 8(4): 584-592. IF: 2.143

Vető B, Bojcsuk D, Bacquet C, Kiss J, Sipeki S, Martin L, Buday L, Balint BL, Aranyi T. (2017) The transcriptional activity of hepatocyte nuclear factor 4 alpha is inhibited via phosphorylation by ERK1/2. PLoS One, 12(2): e0172020. IF: 2.806

Maeso I, Dunwell TL, Wyatt CD, Marletaz F, Vető B, Bernal JA, Quah S, Irimia M, Holland PW. (2016) Evolutionary origin and functional divergence of totipotent cell homeobox genes in eutherian mammals. BMC Biol, 14: 45. IF: 6.779

11.2. Publications not directly related to the thesis

Pagliaroli L, Vető B, Aranyi T, Barta C. (2016) From Genetics to Epigenetics: New Perspectives in Tourette Syndrome Research. Front Neurosci, 10: 277. IF: 3.556

DOI:10.14753/SE.2018.2176

12. ACKNOWLEDGEMENTS

I would like to pay special thankfulness and appreciation to the persons below who made my research successful and assisted me at every point to reach my goal.

First of all, I would like to thank my supervisor, Tamás Arányi for the continuous support of my PhD study and research, for his patience, motivation and knowledge. Without his guidance and persistent help this dissertation would not have been possible.
Secondly, I would like to thank my group leader, András Váradi who gave access to the laboratory and research facilities and for being encouraging and helpful.
I am utterly grateful to Caroline Bacquet for teaching me a great variety of techniques during the time we worked together and for being enthusiastic and supportive at all times.
My sincere thanks also goes to Peter Holland who gave me the opportunity to join their team and for providing me professional guidance and advice.
I would also like to thank the collaborators for all the expertise and the useful comments I always received from them.
Special thanks go to my colleagues for the constant encouragement and support.
Last but not the least, I would like to thank my family - my parents, my brothers and my sister - and my partner - Gábor - for and supporting me throughout my this difficult period.