
DNA methylation in CHO cells

Characterization of epigenetic phenomena
under process conditions

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat)
der Technischen Fakultät der Universität Bielefeld

vorgelegt von

Anna Wippermann

geboren am 06.06.1984 in Bielefeld

Dezember 2016

Scientific research is one of the most exciting and rewarding of occupations.

Frederick Sanger

Gutachter

1. Gutachter:

Prof. Dr. Thomas Noll

Lehrstuhl für Zellkulturtechnik

Technische Fakultät

Universität Bielefeld

2. Gutachter:

Prof. Dr. Karsten Niehaus

Lehrstuhl für Proteom- und Metabolomforschung

Fakultät für Biologie

Universität Bielefeld

Publications

Parts of this work have already been published in scientific journals and at conferences as listed below.

Peer-reviewed journals

- 11/2016 Anna Wippermann, Oliver Rupp, Karina Brinkrolf, Raimund Hoffrogge, Thomas Noll. Integrative analysis of DNA methylation and gene expression in butyrate-treated CHO cells. *Journal of Biotechnology*. DOI: 10.1016/j.jbiotec.2016.11.020
- 04/2015 Anna Wippermann, Oliver Rupp, Karina Brinkrolf, Raimund Hoffrogge, Thomas Noll. The DNA methylation landscape of Chinese hamster ovary (CHO) DP-12 cells. *Journal of Biotechnology*. DOI: 10.1016/j.jbiotec.2015.02.014
- 10/2013 Anna Wippermann, Sandra Klausning, Oliver Rupp, Stefan P. Albaum, Heino Büntemeyer, Thomas Noll, Raimund Hoffrogge. Establishment of a CpG island microarray for analyses of genome-wide DNA methylation in Chinese hamster ovary cells. *Applied Microbiology and Biotechnology*. DOI:10.1007/s00253-013-5282-2

Oral presentations at scientific conferences

- 06/2015 DNA methylation in CHO DP-12 cells: Landscape and effects on gene expression. 24th ESACT meeting, Barcelona, Spain
- 03/2015 The CHO DP-12 DNA methylation landscape and its effect on gene expression. 10th CeBiTec Symposium on Bioinformatics for Biotechnology and Biomedicine, Bielefeld, Germany
- 06/2013 First CpG island microarray for genome-wide analyses of DNA methylation in Chinese hamster ovary cells: new insights into the epigenetic answer to butyrate treatment. 23rd ESACT meeting, Lille, France

Conference proceedings

- 12/2015 Anica Schmidt, Anna Wippermann, Raimund Hoffrogge, Thomas Noll. RNAi-mediated knockdown of Dnmt3a enhances antibody titer in CHO cells up to 200 percent. BMC Proceedings 2015, 9(Suppl 9):P8. DOI:10.1186/1753-6561-9-S9-P8
- 12/2013 Anna Wippermann, Sandra Klausning, Oliver Rupp, Thomas Noll, Raimund Hoffrogge. First CpG island microarray for genome-wide analyses of DNA methylation in Chinese hamster ovary cells: new insights into the epigenetic answer to butyrate treatment. BMC Proceedings 2013, 7 (Suppl 6):O5. DOI:10.1186/1753-6561-7-S6-O5

Posters presented at scientific conferences

- 06/2015 Anica Schmidt, Anna Wippermann, Raimund Hoffrogge, Thomas Noll. RNAi-mediated knockdown of Dnmt3a enhances antibody titer in CHO cells up to 200 percent. 24th ESACT meeting, Barcelona, Spain
- 05/2014 Thomas Noll, Tobias Jakobi, Karina Brinkrolf, Anna Wippermann, Alfred Pühler. Exploring the promoter landscape of the Chinese hamster by next-generation RNA sequencing. Cell Culture Engineering XIV, Quebec City, Canada
- 04/2014 Anna Wippermann, Oliver Rupp, Karina Brinkrolf, Thomas Noll, Raimund Hoffrogge. Effects of butyrate treatment on the methylome of Chinese hamster ovary (CHO) cells. Keystone Symposium on Epigenetic Programming and Inheritance, Boston, Massachusetts, USA
- 10/2013 Anna Wippermann, Sandra Klausning, Oliver Rupp, Thomas Noll, Raimund Hoffrogge. DNA-methylation dynamics in Chinese hamster ovary cells. The genomics revolution and its impact on biotechnology, Bielefeld, Germany

Contents

Abstract	1
1. Introduction	3
1.1. A brief history of epigenetics	4
1.1.1. The relationship between DNA methylation and gene regulation . .	5
1.1.2. The impact of histone modifications on chromatin structure	6
1.1.3. Small RNA-mediated epigenetic regulation	7
1.1.4. ENCODE and the Roadmap Epigenomics Project	8
1.2. Biological processes involving DNA methylation events	9
1.3. Setting and removing DNA methylation marks	11
1.3.1. DNA methyltransferases	12
1.3.1.1. Maintenance DNA methylation	13
1.3.1.2. <i>De novo</i> DNA methylation	13
1.3.2. DNA demethylation	14
1.4. DNA methylation and transcriptional regulation	16
1.4.1. CpG islands	17
1.4.2. DNA methylation changes at CpG-poor genomic regions	18
1.5. Methods for the analysis of DNA methylation	19
1.5.1. Utilization of methylation-sensitive restriction enzymes	19
1.5.2. Bisulfite-mediated conversion of unmethylated cytosines	20
1.5.3. Enrichment of methylated or unmethylated DNA	21
1.5.4. State-of-the-art DNA methylation analysis and its limitations	21
1.5.4.1. Novel technologies	22
1.6. Epigenetic studies in the production of biopharmaceuticals	24
1.6.1. Analysis of single recombinant genes and promoters	24
1.6.2. Epigenetic engineering	25
1.6.3. Genome-scale epigenomics in long-term cultured cells	26
1.7. Butyrate: Physiological effects and biotechnological application	26
1.7.1. Cells respond differently to butyrate due to the Warburg effect . . .	27

1.7.2.	Effects of butyrate on DNA methylation	27
1.7.3.	The effect of butyrate on CHO cell cultures	29
2.	Research objective	31
3.	Materials and methods	33
3.1.	Cell culture and cultivation process analysis	33
3.1.1.	CHO cell lines	33
3.1.2.	Cultivation conditions	34
3.1.3.	Cryopreservation and thawing	34
3.1.4.	Measurement of cell count and viability	35
3.1.5.	Measurement of glucose and lactate concentrations	35
3.1.6.	Measurement of IgG concentrations	35
3.1.7.	Transient transfection of CHO cell cultures	35
3.2.	Microbiological methods	36
3.2.1.	Bacterial culture	36
3.2.2.	Preparation of competent <i>E. coli</i> cells using calcium chloride	36
3.2.3.	Transformation of competent <i>E. coli</i> cells	36
3.2.4.	Plasmid preparation	37
3.3.	Molecular biology methods	37
3.3.1.	Preparation and analysis of DNA	37
3.3.1.1.	Isolation of DNA from CHO cells using Proteinase K and phenol	37
3.3.1.2.	Agarose gel electrophoresis of DNA	38
3.3.1.3.	DNA extraction from agarose gels	39
3.3.1.4.	Analysis of DNA quantity and quality	39
3.3.1.5.	Sanger sequencing	39
3.3.2.	Polymerase chain reaction (PCR)	39
3.3.2.1.	Amplification from bisulfite-treated DNA	40
3.3.2.2.	Amplification from native DNA	40
3.3.2.3.	Quantitative Real-Time PCR	41
3.3.3.	Preparation and analysis of RNA	42
3.3.3.1.	RNA extraction	42
3.3.3.2.	Quality control on denaturing agarose gels	42
3.3.3.3.	Reverse transcription for generation of cDNA	43
3.3.4.	Gene expression profiling by microarrays	43
3.3.4.1.	Sample preparation	43
3.3.4.2.	Microarray hybridization and wash	44
3.3.4.3.	Scanning and feature extraction	44
3.3.5.	Molecular cloning	45

3.3.5.1.	Subcloning of PCR products amplified from bisulfite-treated DNA	45
3.3.5.2.	Generation of promoter-reporter constructs for dual luciferase assays	45
3.3.5.3.	Cloning of shRNA oligonucleotides into <i>pLVX-shRNA1</i>	46
3.3.6.	ShRNA design for RNA interference experiments	47
3.4.	DNA methylation analysis	47
3.4.1.	Bisulfite Sanger sequencing of individual genomic loci	48
3.4.1.1.	Primer design for bisulfite sequencing	48
3.4.1.2.	Bisulfite conversion of genomic DNA	48
3.4.1.3.	Sequencing of amplicates from bisulfite treated DNA	48
3.4.2.	Combined Bisulfite Restriction Analysis (COBRA)	49
3.4.3.	CpG island microarray analysis	49
3.4.3.1.	Sonication of genomic DNA	50
3.4.3.2.	Immunoprecipitation of methylated DNA	50
3.4.3.3.	Fluorescence labeling of enriched methylated DNA	51
3.4.3.4.	Hybridization	52
3.4.3.5.	Scanning and feature extraction	52
3.4.4.	Whole-genome bisulfite sequencing	52
3.4.4.1.	Bisulfite conversion	52
3.4.4.2.	Library generation and quality control	53
3.4.4.3.	Illumina <i>MiSeq</i> sequencing	53
3.4.4.4.	Illumina <i>HiSeq 1500</i> sequencing	53
3.5.	Protein biochemistry	54
3.5.1.	Westernblot analysis	54
3.5.1.1.	Protein extraction	54
3.5.1.2.	Protein quantification by Bradford assay	55
3.5.1.3.	SDS polyacrylamide gelelectrophoresis (SDS-PAGE)	55
3.5.1.4.	Westernblot analysis	56
3.5.2.	Dual luciferase assay	56
3.6.	Data analysis	57
3.6.1.	Analysis of gene-specific DNA methylation data	57
3.6.2.	Analysis of data generated by CHO CpG island microarray experiments	58
3.6.3.	Analysis of gene expression microarray data	58
3.6.4.	Analysis of next generation sequencing data	59
3.6.5.	Functional gene set analysis	60
3.6.6.	Statistics applied for DNA methylation and gene expression analysis	60
4.	Results and discussion	61
4.1.	Integrative analysis of DNA methylation and gene expression in CHO cells	62

4.1.1.	Generation of samples	62
4.1.2.	Implementation of whole-genome bisulfite sequencing for CHO DP-12 cells	62
4.1.2.1.	Quality control of DNA samples and bisulfite sequencing libraries	64
4.1.2.2.	Illumina <i>HiSeq</i> sequencing of libraries from native and bisulfite-treated CHO DP-12 DNA and generation of the CHO DP-12 reference genome	66
4.1.2.3.	Selection of a mapping algorithm for data analysis	67
4.1.2.4.	Processing of whole-genome bisulfite sequencing data	69
4.1.2.5.	Integration of whole-genome DNA methylation data into <i>GenDBE</i> and inspection of exemplary genomic regions	69
4.1.2.6.	Discussion: Generation of genome-scale DNA methylation data for CHO DP-12 cells	70
4.1.2.7.	Conclusion: Implementation of whole-genome DNA methylation analysis enables characterization of epigenetic phenomena	71
4.1.3.	The CHO DP-12 DNA methylation landscape	72
4.1.3.1.	Global assessment of the CHO DP-12 DNA methylation landscape	72
4.1.3.2.	Evidence for partially methylated domains in the CHO DP-12 genome	74
4.1.3.3.	Functional classification of CHO DP-12 PMD-localized by DNA-microarray analysis	76
4.1.3.4.	Discussion: General characterization of the CHO DP-12 cell DNA methylome	79
4.1.3.5.	Discussion: Partially methylated domains in the CHO DP-12 cell DNA methylome	80
4.1.3.6.	Conclusion: The first map of a CHO cell DNA methylation landscape	82
4.1.4.	DNA methylation dynamics and differential gene expression in response to butyrate	82
4.1.4.1.	Gene expression changes upon butyrate addition	83
4.1.4.2.	Mapping of gene expression changes to cellular pathways	87
4.1.4.3.	Comparison of DNA-microarray results with published gene expression analyses in butyrate-treated CHO cells	98
4.1.4.4.	Genome-wide analysis of DNA methylation changes upon butyrate addition	101
4.1.4.5.	Combination of DNA methylation and gene expression data	106

4.1.4.6.	Discussion: Butyrate-induced changes in gene expression	110
4.1.4.7.	Discussion: Transcription factor binding-site motifs in DMRs	111
4.1.4.8.	Discussion: Dynamics of DNA methylation changes upon butyrate addition	113
4.1.4.9.	Conclusion: Data integration allows for novel perspectives	113
4.2.	Comparison of whole-genome bisulfite sequencing and CpG island microarray analysis	115
4.2.1.	Analysis of the butyrate effect by CpG island microarrays	116
4.2.2.	Overlaps between the results of whole-genome bisulfite sequencing and CpG island microarray analysis	120
4.2.3.	Assessment of CpG microarray performance	120
4.2.3.1.	Assessment of CpG microarray performance by Combined Bisulfite Restriction Analysis (COBRA)	121
4.2.3.2.	Assessment of CpG microarray performance by hybridization experiments	123
4.2.4.	Discussion: Comparable overall results of CpG island microarray and whole-genome methylation analyses	124
4.2.5.	Discussion: Variation in DNA methylation between identical CHO cell cultures	127
4.2.6.	Conclusion: Perspectives of CHO-specific DNA methylation microarrays	128
4.3.	Functional analysis of putative central mediators of the butyrate effect . . .	129
4.3.1.	Discussion: DNMT3A could be involved in butyrate-induced hyperproductivity	131
5.	Summary and outlook	133
5.1.	Transgene integration into partially methylated domains might be associated with production instability or product heterogeneity	134
5.2.	Further questions and novel perspectives	135
5.3.	General considerations	136
6.	Bibliography	139
A.	Abbreviations	173
B.	Supplementary material	177
B.1.	Attempt to identify transgene integration sites in the CHO DP-12 genome .	178
B.2.	Analysis of a putative involvement of microRNA 199a in the response of CHO DP-12 cells to butyrate treatment	180
B.3.	CHO cell promoters	182

B.3.1. Discussion: Bioinformatically identified CHO promoters are functional	185
B.4. Microarray analysis of CHO-XL99 clones with differences in productivity .	186
B.4.1. Confirmation of <i>Actb</i> and <i>Bcat1</i> promoter regions as methylation controls in CHO-XL99	186
B.4.2. Batch cultivation of CHO-XL99 clones	187
B.4.3. CpG island microarray analysis of good and bad producing clones in comparison to the CHO K1 host cell	187
B.4.4. Data analysis	189
B.4.5. Functional annotation of differentially methylated genes in good and bad producing CHO-XL99 clones	191
B.4.6. Differential methylation of apoptotic networks in bad producing clones	193
B.4.7. Discussion: Possible maintenance of detrimental hypoxia-induced epigenetic marks through cell line generation	195
B.5. R code	197
B.6. Supplementary figures	200
B.7. Supplementary tables	224
List of Figures	251
List of Tables	255
Danksagung	257
Erklärung	261

Abstract

The cellular mechanisms responsible for the versatile properties of CHO cells as the major production cell line for biopharmaceutical molecules are not entirely understood yet, although several 'omics' data facilitate the understanding of CHO cells and their reactions to environmental conditions. However, genome-wide studies of epigenetic processes such as DNA methylation as a crucial epigenetic modification and an important element in mammalian genome regulation and development are still limited.

In order to study DNA methylation phenomena in CHO cell cultures, whole-genome bisulfite sequencing was implemented for CHO DP-12 cells that produce a recombinant human anti-IL-8 antibody. Methylome data generated for this cell line allowed for the characterization of the CHO DP-12 DNA methylation landscape as well as for the analysis of differential DNA methylation upon butyrate addition. The combination with gene expression studies by DNA-microarrays enabled an assessment of potential phenotypic effects and functional significances.

As a DNA methylation map of a CHO cell line did not exist before, the analysis of the CHO DP-12 cell DNA methylation landscape was of special importance. CHO DP-12 cells were found to exhibit global hypomethylation compared to a majority of mammalian methylomes and hypermethylation of CpG-dense regions at gene promoters called CpG islands. Also, partially methylated domains were observed that covered 62% of the CHO DP-12 cell genome and contained functional clusters of genes. Gene expression analysis showed these clusters to be either highly or weakly expressed with regard to CHO-specific characteristics. Generally, the results showed that DNA methylation might significantly contribute to CHO phenotypes and allowed for a deeper understanding of CHO cell properties.

Whole-genome bisulfite sequencing and DNA-microarray analyses were next used to analyze the time course of cellular reactions upon butyrate addition in CHO DP-12 cells, as supplementation of butyrate is known to enhance cell specific productivities and leads to alterations of epigenetic silencing events. Gene expression and DNA methylation analyses showed that the central cellular pathways cell cycle, apoptosis and RNA processing, as well as pathways

potentially associated with increased cell specific productivity such as energy metabolism and protein biosynthesis were affected by butyrate. Differentially methylated regions were furthermore found to contain binding-site motifs of specific transcription factors and were hypothesized to represent regulatory regions closely connected to the cellular response to butyrate. The experiment underlined the benefit of integrating DNA methylation and gene expression data, as it provided potential novel candidate genes for rational cell line development and allowed for new insights into the butyrate effect on CHO cells.

Whole-genome bisulfite sequencing was finally compared to the analysis of DNA methylation changes upon butyrate addition by a recently designed CHO-specific microarray. 32 % of the genes detected as differentially methylated by CpG island microarrays were also found by whole-genome bisulfite sequencing. Gene-specific analyses for verification of microarray results did not show significant differences between the methods. Furthermore, hybridization experiments for control of experimental parameters indicated that CpG island microarray analyses were only biased by variations in DNA extraction, which led to false positive calls at 0.15 % of the probes. Therefore, CpG island microarrays were considered suitable tools for the analysis of DNA methylation in CHO cells in a medium-throughput format. As both the whole-genome bisulfite sequencing and CpG island microarray data sets indicated a potential involvement of DNA methyltransferase 3a (DNMT3A) in the butyrate-response, DNMT3A was functionally analyzed in transient knockdown experiments which indicated that reduced levels of the enzyme might be associated with increased cell specific productivity.

1

Introduction

More than 200 biopharmaceutical molecules have been produced by means of mammalian expression systems since the first recombinant protein, Tissue plasminogen activator (tPA), was approved by the Food and Drug Administration (FDA) in 1986 (Walsh 2010). About 70 % of these products are made in Chinese hamster ovary (CHO) cells, which provide proteins with proper foldings and posttranslational modifications and account for the production of six out of the ten top-selling biopharmaceuticals in 2013. These comprised e.g. Humira (AbbVie), Rituxan (Biogen-IDEC) or Avastin (Roche; Walsh 2014).

Even though CHO cells were extensively used over the last three decades, still little is known about the molecular basis for their versatile properties regarding robustness, adaptability and productivity. For this reason several techniques have been applied to unravel the molecular mechanisms in CHO-based production processes and a plethora of 'omics' data sets was generated. These data include the CHO K1 genome sequence which serves as a CHO reference genome and facilitates genome-scale studies on multiple levels (Xu et al. 2011). The CHO K1 genomic data was furthermore assigned to Chinese hamster chromosomes, which enables the analysis of chromosomal rearrangements and genome stability (Brinkrolf et al. 2013). The CHO cell transcriptome was sequenced and published (Becker et al. 2011) and several transcriptomics studies were conducted to analyze the reactions of CHO cells to specific environmental conditions or to compare CHO clones with differences in productivity. For example, differential gene expression in CHO cells was analyzed upon treatment with the productivity-enhancing substance butyrate, which was shown to affect expression of genes related to protein secretion (Kantardjieff et al. 2010). It was furthermore shown that CHO

cell productivity increased due to temperature downshifts and that this was accompanied by a significant upregulation of genes related to energy metabolism (Bedoya-López et al. 2016). Additionally, several proteomic and metabolomic studies of CHO cells were performed, which, e.g., showed that high passage numbers correlated with increased robustness due to differential expression of anti-stress proteins (Beckmann et al. 2012). When extracellular metabolites in CHO cell fed-batch cultures were examined, an accumulation of detrimental metabolites over the cultivation process was found, which was accompanied by depletion of media components despite the addition of feed medium (Chong et al. 2009).

Despite the immense effort to explain CHO cell characteristics by 'omics' experiments, numerous open questions remain and biopharmaceutical production processes are still largely empirical. For example, when the correlation between lactate accumulation and its consumption and the outcomes of production-scale cultures were examined using a large amount of bioprocess data (243 bioreactor runs including inoculum trains and subsequent production processes), it was shown that the rate of lactate formation and the switch to its consumption varied even between cultures of the same strain under the same conditions. The hypothesis was raised that this finding was a hint towards a predetermination of metabolic characteristics in early stages of culturing (Le et al. 2012).

An explanation to such unanswered questions might arise from epigenomics studies in CHO cells, which represent one missing link between the above-mentioned cellular levels and are facilitated by recent developments in analysis techniques. The following sections provide an overview of the history of epigenetic research, definitions and the relevance of epigenomic studies.

1.1. A brief history of epigenetics

Until the middle of the 20th century developmental biologists took little account of genes and gene action to answer the question of how a single fertilized egg can give rise to a complex organism consisting of various specialized cell types. Instead, they focused on the roles of nucleus and cytoplasm in the developmental process and either believed that preformed elements in a cell enlarged during development ('preformationism') or that chemical reactions occurred among cellular components that executed a developmental plan ('epigenesis'; Felsenfeld 2014). As time went on, it became apparent that certain fundamental features of development were not explainable by either of these theories, e.g. the ability of differentiated cells to maintain their phenotypes through cell division. However, a combination of both theories provided a possible explanation, with preformation representing the static nature of the gene and epigenesis representing the dynamic nature of gene expression (Deans and Maggert 2015). Believing in this complementarity, the embryologist Conrad Waddington

coined the term 'epigenetics' in 1942, which he defined as 'the branch of biology that studies the causal interactions between genes and their products which bring the phenotype into being', or, briefly, as changes in phenotype without changes in genotype (Waddington 1942; Holliday 2006; Allis and Jenuwein 2016). At that time the gene was being accepted as the unit of inheritance without knowledge about its biochemical nature (Deans and Maggert 2015).

1.1.1. The relationship between DNA methylation and gene regulation

During the following decades the work of Avery et al., Hershey and Chase and, obviously, Watson and Crick proved that a protein-free DNA molecule could carry and inherit the genetic information (Avery, Macleod, and McCarty 1944; Hershey and Chase 1952; Watson and Crick 1953). As it was furthermore shown that the specialized gene expression potential in differentiated cells was not due to mutations of the germline DNA sequence, mechanisms were proposed in which information was superimposed on DNA sequences (Laskey and Gurdon 1970; Felsenfeld 2014). In 1975 two independent groups proposed that DNA methylation at cytosine residues could act as an epigenetic mark that could be transmitted to the next generation if methylation sites were palindromic and if a maintenance methyltransferase existed that could bind to hemimethylated DNA (Riggs 1975; Holliday and Pugh 1975). This hypothesis was confirmed by experiments using methylation-sensitive restriction enzymes that showed a strong effect of DNA methylation at CpG sites on gene expression (Doerfler 1983; Bird and Southern 1978).

Further evidence for the involvement of DNA methylation in the control of gene expression came from experiments using 5-azacytidine, which is a cytosine analog that is incorporated into DNA and deactivates DNA methyltransferases (Holliday 2006; Jones and Taylor 1980). It was shown that 5-azacytidine reactivated silent genes in many cases, e.g. the *Phosphoglycerate kinase 1 (Pgk1)* gene that is linked to the inactive human X chromosome in hamster-human cell hybrids (Hansen and Gartler 1990) or genes of the herpes simplex virus (HSV) that were silenced in mammalian cells (Clough, Kunkel, and Davidson 1982).

These findings led to a revised definition of epigenetics as 'the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail change in DNA sequence' (Wu and Morris 2001). In the following years several studies focused on patterns of DNA methylation in different organisms or cell types, the impact of DNA methylation on gene expression or the identification of enzymes responsible for *de novo* and maintenance DNA methylation. These studies were immensely accelerated by the advent of next generation sequencing technologies.

1.1.2. The impact of histone modifications on chromatin structure

As transmission of gene activity states during cell division was shown in organisms such as *Drosophila* that exhibit extremely low levels of DNA methylation, it was clear that DNA methylation could not be the only epigenetic mechanism (Felsenfeld 2014). In fact, another epigenetic mechanism was discovered in 1964, when Allfrey and co-workers showed that histones are posttranslationally acetylated and methylated (Allfrey, Faulkner, and Mirsky 1964). Their speculation that histone acetylation might be correlated with chromatin structure and, thereby, gene activation, was confirmed by the findings that DNA was packaged into nucleosomes containing histone octamers (Thomas and Kornberg 1975) and that amino-terminal tails of histones were essential for transcriptional regulation (Durrin et al. 1991). Fig. 1.1 summarizes the basic concept of chromatin organization.

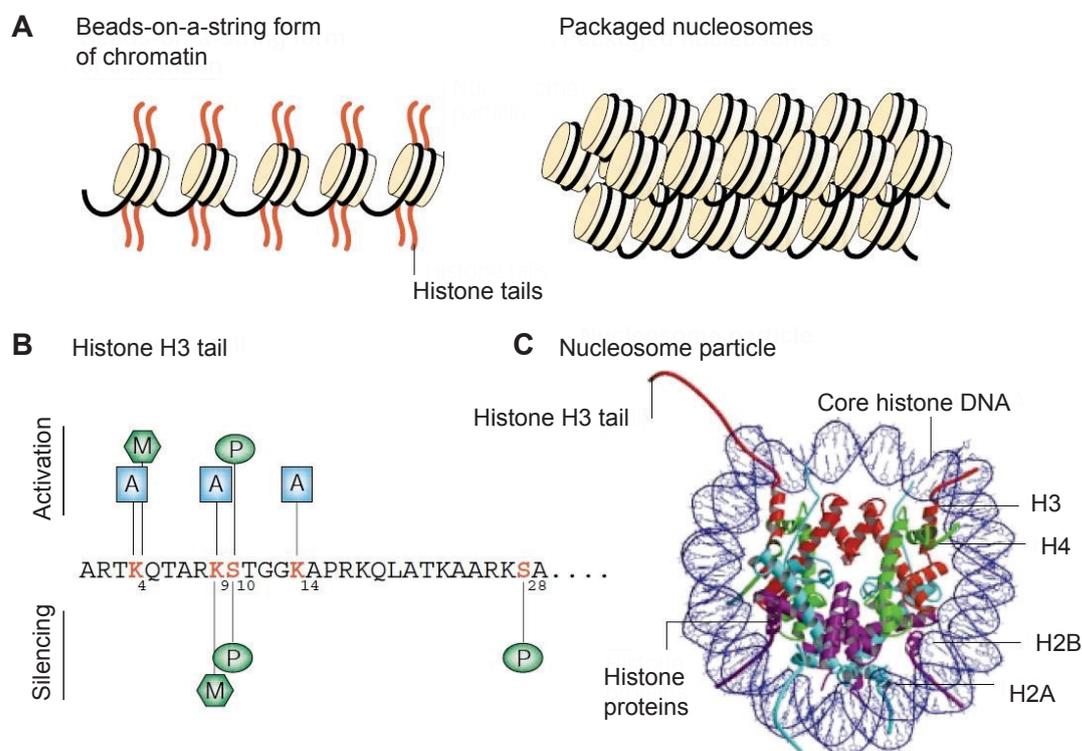


Figure 1.1.: Chromatin organization and histone tail modifications (modified from Marmorstein 2001).

A Histone octamers are depicted as yellow cylinders, histone tails are shown in red and DNA in black. Nucleosomes comprise 145–147 bp of DNA wrapped around a core of each two molecules of histones H2A, H2B, H3 and H4. Each core histone contains a flexible, basic tail region which is the site for posttranslational modifications. As a result of these modifications nucleosomes can either be tightly associated (condensed heterochromatin) or relaxed (permissive euchromatin, 'beads-on-a-string'). **B** Histone H3 tail region with gene regulatory modifications acetylation (A), phosphorylation (P) and methylation (M). Acetylation is associated with transcriptional activation, methylation and phosphorylation are associated with both transcriptional activation and silencing. **C** High resolution structure of a nucleosome core particle. DNA is shown in blue, the histone H2A dimer in aqua, H2B in purple, H3 in red and H4 in green.

The question of how specific chromatin sites are chosen for modification was answered by the discovery of DNA-binding proteins such as the helix-loop-helix myogenic regulatory factor MyoD (which is involved in skeletal muscle development; Berkes and Tapscott 2005) that recruit cofactors like histone modification complexes (Lassar et al. 1989; Felsenfeld 2014). Chromatin Isolation by RNA Purification sequencing (ChIRP-seq) experiments furthermore showed that long non-coding RNAs (lncRNAs) can bind to DNA in a sequence-specific way and are able to recruit histone modifying complexes (Chu et al. 2011). However, it is still not clear how chromatin states can be transmitted through cell divisions. Different kinds of mechanisms for propagation of histone marks have been suggested though, which are rather based on variants of histones and not on histone modifications (Felsenfeld 2014). For example, histone variant H3.3 (exhibiting four different amino acids compared to histone H3) was shown to accumulate in active chromatin, while histone H3 is only incorporated during DNA replication. Histone H3.3 was suspected to maintain gene activity through cell division and to lead to replacement of histone H3, which is incorporated during this process. Thereby, transmission of the active state of chromatin into the next generation would be possible (McKittrick et al. 2004; Felsenfeld 2014).

1.1.3. Small RNA-mediated epigenetic regulation

A third component of the epigenetic system is a class of single-stranded, small non-coding RNAs with a length of about 22 nucleotides, called microRNAs (miRNAs). MiRNAs were discovered in *C. elegans* in 1993 (Lee, Feinbaum, and Ambros 1993) and regulate gene expression on the posttranscriptional level by base-pairing with target mRNAs that leads to translational repression or mRNA decay (Ha and Kim 2014). Although many of the miRNA promoters have not been characterized yet, miRNA genes appear to be regulated similarly to protein-coding genes by transcription factors and epigenetic mechanisms (Vrba et al. 2013). For example, cell-type specific miRNA expression in human mammary epithelial and fibroblast cells was shown to be regulated by DNA methylation and histone 3 lysine 27 trimethylation (H3K27me₃; Vrba et al. 2011). Other miRNAs were shown to regulate the expression of epigenetic mediators such as DNA methyltransferases (DNMTs, e.g. miR-29) and Histone deacetylases (HDACs, e.g. miR-449a) themselves (Garzon et al. 2009; Noonan et al. 2009). In fact, several publications describe a dynamic interplay between DNA methylation events and microRNAs. For example, in fibrotic lung tissue expression of the miR-17-92 cluster targeting DNMT1 was found to be decreased, while DNMT1 expression and methylation of the miR-17-92 promoter itself were increased, thereby suggesting a negative feedback-loop (Sun et al. 2013). Fig. 1.2 summarizes the epigenetics-miRNA regulatory circuit (Sato et al. 2011). Generally, miRNAs are suspected to play a key role within the gene regulatory network and could furthermore be able to transmit epigenetic signals by moving from one cell to another (Holliday 2006).

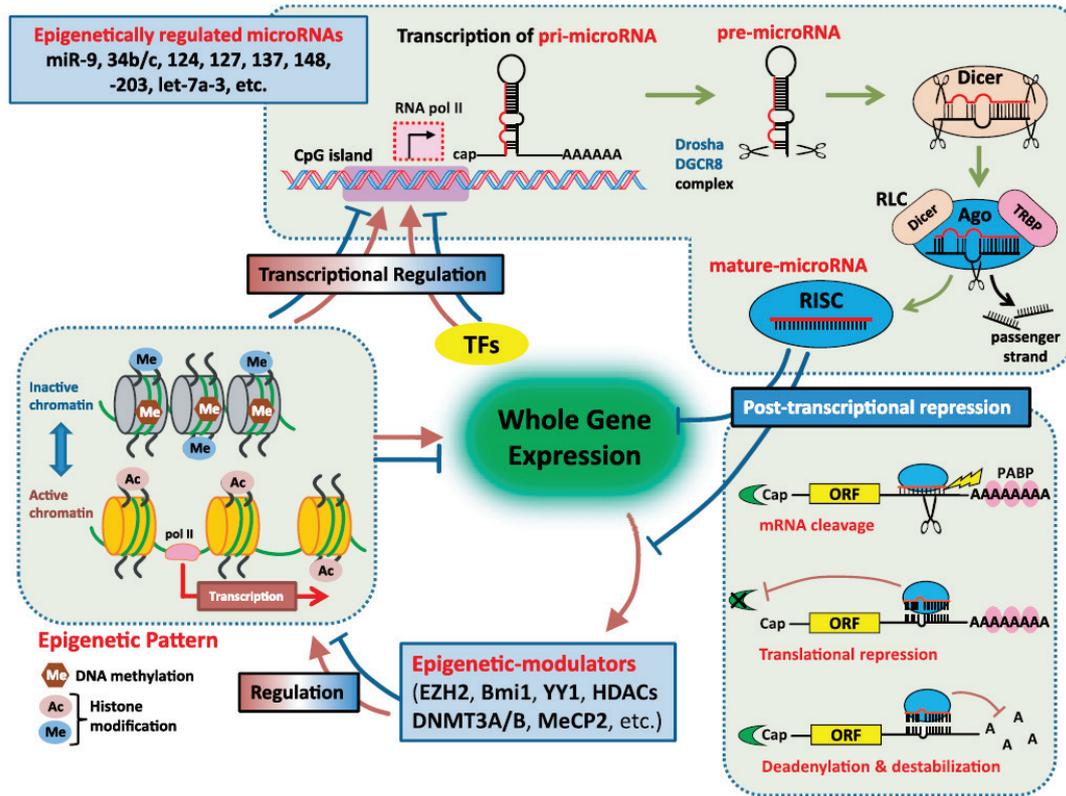


Figure 1.2.: Regulatory epigenetic-miRNA circuit (Sato et al. 2011). MiRNA genes can contain CpG islands in their promoter regions and are transcribed by RNA polymerase II as primary micro RNAs (pri-miRNAs) that are processed to mature miRNAs by the Drosha/DGCR8 complex and Dicer. The mature, single-stranded miRNAs mediate post-transcriptional repression (mRNA cleavage, translational repression or deadenylation and destabilization of target mRNAs) of target genes as part of the RNA induced silencing complex (RISC). Amongst others, epigenetic regulators and transcription factors are targeted by RISC complexes, which can lead to modification of chromatin structures and thereby (feedback) regulation of miRNA genes themselves. RLC = RISC loading complex, PABP = Poly A binding protein, TRBP = Tar RNA binding protein.

1.1.4. ENCODE and the Roadmap Epigenomics Project

It is understood today that gene expression is largely regulated by the epigenome, which provides information about DNA methylation, histone and chromatin modifications as well as other regulatory factors not encoded within the DNA such as non-coding RNAs. The importance of mapping and cataloging epigenomic structures is underlined by the immense efforts of the Encyclopedia of DNA Elements (ENCODE) and Epigenomics Roadmap Projects that followed the unraveling of the human genome. When the first sequencing of the complete human genome was finished in 2001 (Lander et al. 2001), the ENCODE Project, an international consortium of computational and laboratory-based scientists, was launched to identify all functional elements comprised within it. The aim of the consortium was to generate a comprehensive catalog of the structural and functional components, including epigenetic features, encoded in the human genome in order to understand the biology of human health and

disease (The ENCODE Project Consortium 2004). In a pilot phase (2004–2007) 44 genomic target regions, which represented 1 % of the human genome, were manually selected and analyzed regarding the identification of genes, promoters, enhancers, origins of replication, transcription factor binding sites, histone modifications and Desoxyribonuclease I (DNase I) hypersensitive sites (i.e. chromatin accessibility) by tiling arrays and genomic sequencing (Birney et al. 2007). In the next step the ENCODE production phase was launched, which aimed to analyze the entire genome of human cells grown in culture. The realization of this project was eased by the advances in low-cost, rapid DNA-sequencing technologies and the project was supposed to provide information about DNA methylation, histone modifications, chromatin accessibility and interactions, transcription factor binding sites, the location and sequence of gene-regulatory DNA elements as well as sequences and quantities of RNAs transcribed from non-coding and protein-coding regions. The results of an initial analysis of 1.640 data sets were published in a total of 30 papers in 2012, resulting in the conclusion that 80 % of the human genome is transcribed and therefore considered as being functional (Ecker et al. 2012). It has to be noted that this final conclusion is regarded very critically within the scientific community (Graur et al. 2013; Eddy 2012). However, the ENCODE data were used as a basis for the follow-up Roadmap Epigenomics Project which aimed to analyze samples taken directly from human tissues and cells under diverse conditions and generated epigenomic data for 127 tissues and cell types until the first release of a package of papers in 2015 (Romanoski et al. 2015). The data generated in this way are e.g. used to identify fresh entry points for treatment of diseases like Alzheimer's (Gjoneska et al. 2015).

According to the above-mentioned advances in epigenomic research in human cells and tissues that followed the sequencing of the human genome, the relevance of similar studies in biotechnologically important cell lines such as CHO cells is evident. Such studies will enable a deeper understanding of gene regulation in the production of biopharmaceuticals, especially in response to environmental conditions, and will finally enable rational design of better production cell lines. As this thesis focused on the analysis of DNA methylation patterns in CHO cells and the effect on environmental conditions on dynamic DNA methylation, the following sections aim to provide a detailed overview of the current knowledge about this epigenetic mechanism.

1.2. Biological processes involving DNA methylation events

In humans, the greatest changes in overall DNA methylation are observed during early embryonic development, when the fertilized egg undergoes cleavage divisions and develops into a morula and then a blastocyst (Marcho, Cui, and Mager 2015). In this period of time DNA methyl groups in the paternal and maternal pronuclei are largely erased by active demethylation (Fig. 1.3; Reik 2007). DNA methylation marks present at some genes, however,

escape the overall demethylation events. The underlying mechanism is not entirely understood yet, but it is speculated that specific proteins actively protect certain genomic regions from demethylation (Reik 2007; Lane, Robker, and Robertson 2014). At the time of implantation the whole genome is again methylated by the *de novo* DNA methyltransferases DNMT3A and DNMT3B (see section 1.3.1.2), which establish a bimodal pattern of methylation for a majority of mammalian genomes. This means that highly methylated regions alternate with unmethylated regions, which mostly exhibit a specific base composition with a high G+C content (CpG islands; see section 1.4; Cedar and Bergman 2012). These patterns are maintained over cell divisions by the maintenance methyltransferase DNMT1 (see section 1.3.1.1). Following implantation, additional changes in methylation are possible, but mostly affect tissue-specific genes or other specific loci, and overall DNA methylation landscapes remain stable. Accordingly, a study examining 42 whole-genome methylation data sets including diverse human cell and tissue types showed that only 21.8 % of autosomal CpG dinucleotides are dynamically regulated within normal development (Ziller et al. 2013).

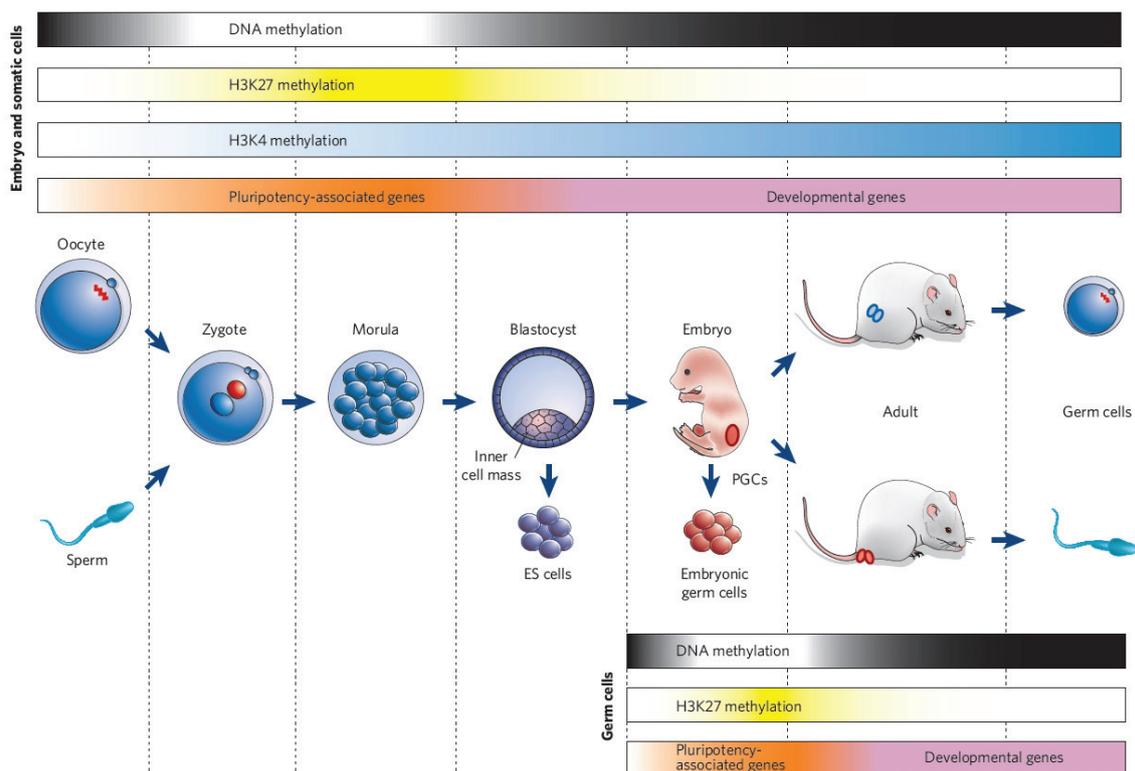


Figure 1.3.: Regulatory epigenetic processes during development (Reik 2007). When DNA methylation is erased early in development, pluripotency-associated genes are expressed and developmental genes are repressed by Histone H3 lysine 27 (H3K27) methylation. When differentiation sets in, pluripotency-associated genes become repressed by increasing DNA methylation which is paralleled by an increase in H3K4 methylation. During germ cell development DNA methylation marks are erased in large parts and pluripotency-related genes are re-expressed. Afterwards developmental genes are expressed again.

DNA methylation has been shown to be responsible for processes such as X chromosome inactivation (Herzing et al. 1997), imprinting (Feil and Berger 2007) or silencing of retrotransposons and repetitive elements (Ndlovu, Denis, and Fuks 2011). Changes in DNA methylation patterns are furthermore associated with a plethora of diseases like cancer (Kulis and Esteller 2010), diabetic nephropathy (Bell et al. 2010), atherosclerosis (Hai and Zuo 2016) or Parkinson's (Wüllner et al. 2016) and it was shown that aberrant changes in epigenetic pathways are often linked to environmental factors such as nutrients or the exposure to chemicals (Ruiz-Hernandez et al. 2015). For example, changes in maternal diets during pregnancy (regarding abnormal concentrations of folate, choline or docosahexaenoic acid) influence the risk of developing metabolic and cardiovascular diseases in adulthood by induction of aberrant DNA methylation, which is associated with differential expression of genes involved in hormonal regulation or inflammation (Lee 2015; Jiang et al. 2012; Lee et al. 2013). Recent evidence proves that it is even possible to transmit acquired DNA methylation marks to the next generation, a process which is described by the term 'transgenerational epigenetic inheritance' (Prokopuk, Western, and Stringer 2015). It was for example shown in mice that preconceptional feeding of a zinc deficient diet for 3 to 5 days disrupts chromatin methylation and preimplantation development in oocytes (Tian and Diaz 2013). DNA methylation levels (and, also, histone H3K4 trimethylation) were found to be decreased and associated with differential expression of repetitive elements and oocyte specific genes, which led to impaired fertilization potential and reduced blastocyst formation. Fertilization potential could in parts be restored by supplementation with a methyl group donor. Another, yet controversial, example was provided by Dias et al. in 2014, who showed that parental olfactory experience influences behavior and neuronal structure of subsequent generations (Dias and Ressler 2014). In this study F0 mice were subjected to odor fear conditioning before conception. It was shown that F1 and F2 generations were similarly sensitive to the F0-conditioned odor without ever being exposed to it. Sperm DNA from conditioned F0 males and F1 offspring showed differential DNA methylation of a gene involved in olfactory pathways.

1.3. Setting and removing DNA methylation marks

The following section provides an overview of the structures and catalytic activities of the above-mentioned enzymes responsible for setting and removing DNA methylation marks, including DNA methyltransferase 1 (DNMT1), which mediates maintenance DNA methylation during cell divisions and DNMT3A and DNMT3B as *de novo* methyltransferases. Furthermore, the methylcytosine dioxygenase family of enzymes (TET1, TET2, TET3) is described, which conveys the removal of methyl groups.

1.3.1. DNA methyltransferases

Three DNA methyltransferases (DNMT1, DNMT3A, DNMT3B) and one related protein without catalytic activity (DNMT3L) exist that consist of an N-terminal part responsible for localization of the enzymes and a C-terminal, catalytically active part exhibiting a common core structure for all DNMTs (see domain structure in Fig. 1.4). In comparison to DNMT1 the members of the DNMT3 family of proteins contain shorter N-terminal domains, which do not possess a nuclear localization signal (NLS), PCNA and DNA replication foci binding domains (PBD and TS) as well as the polybromo-1 homology domain (PBHD). All of these domains mediate DNMT1-association with DNA replication sites. Instead, the DNMT3 family members possess PWWP domains for unspecific DNA binding as well as ATRX-DNMT3-DNMT3L (ADD) domains for interaction with histone tails.

Mechanistically, DNMTs flip their target base out of the double helix and bury it in the active center, thereby forming a covalent bond with it, which increases the negative charge at the C5 atom and leads to addition of a methyl group from the donor molecule S-adenosyl-methionine (SAM or AdoMet; Jurkowska, Jurkowski, and Jeltsch 2011). The following subsections provide a more detailed view on the maintenance methyltransferase, DNMT1, and the *de novo* methyltransferases DNMT3A and DNMT3B.

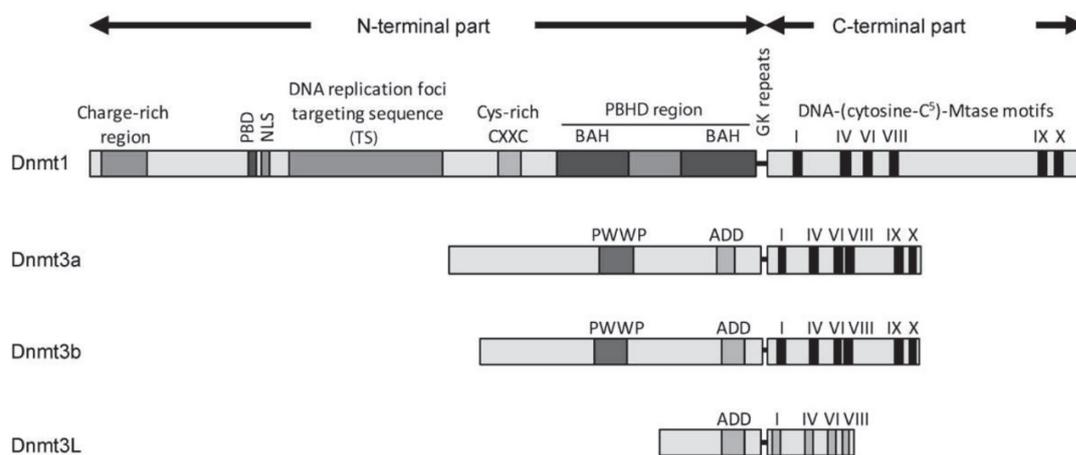


Figure 1.4.: Schematic representation of the domain structure of mammalian DNA methyltransferases (DNMTs; Jurkowska et al. 2011). PBD = PCNA binding domain, NLS = nuclear localization signal, PBHD = polybromo-1 homology domain, CXXC = binding domain for unmethylated CpGs, BAH = bromo-adjacent homology, PWWP = unspecific DNA binding domain domain, ADD = ATRX-DNMT3-DNMT3L domain (histone interaction domain).

1.3.1.1. Maintenance DNA methylation

The hypothesis that maintenance of DNA methylation patterns through cell division could be mediated by a methyltransferase with a strong preference for hemimethylated DNA was confirmed in 1982, when Gruenbaum et al. studied enzymatic DNA methylation activity in crude extracts from somatic cells. They showed that hemimethylated DNA, in contrast to completely unmethylated DNA, was a better substrate for methylation with a 100-fold higher K_m value (Gruenbaum, Cedar, and Razin 1982). It was subsequently shown that the enzyme DNA methyltransferase I (DNMT1) is permanently localized to replication foci and mediates maintenance DNA methylation (Li, Bestor, and Jaenisch 1992; Leonhardt et al. 1992). DNMT1 is recruited to replication foci by a protein complex containing Ubiquitin-like PHD and RING finger domain-containing protein 1 (UHRF1). Deletion of UHRF1 was shown to drastically decrease DNA methylation in mouse embryonic stem cells and UHRF1 was shown to co-localize with DNMT1 at replication foci during S phase (Bostick et al. 2007). Another protein involved in this regulatory complex is Proliferating cell nuclear antigen (PCNA), which dynamically binds to DNMT1 and increases its methylation efficiency by 100 %. However, this interaction was found to be not strictly required for maintaining global methylation (Schermele et al. 2007).

The multimodular protein DNMT1 with its C-terminal methyltransferase domain and the N-terminal regulatory region, which is composed of the replication foci targeting sequence (RFTS) domain, the CXXC zinc finger domain and the pair of bromo adjacent homology (BAH) domains (see Fig. 1.4 and ribbon representation in Fig. 1.5A) is recruited to replication foci by the RFTS domain and interaction with UHRF1 (Leonhardt et al. 1992; Berkuyrek et al. 2014). Furthermore, the RFTS domain interacts with the methyltransferase domain in an autoinhibitory mechanism, which is proposed to be responsible for occlusion of unmethylated CpGs from the active site of the enzyme in order to ensure that only hemimethylated DNA is methylated (Song et al. 2011). The association of the RFTS and methyltransferase domains is strengthened by the helical linker between the CXXC and BAH1 domains. Binding of the CXXC domain to CpG dinucleotides in unmethylated DNA results in positioning of the CXXC-BAH1 domain linker into the catalytic site of the methyltransferase domain, thereby generating a second variant of autoinhibitory conformation. When hemimethylated DNA is bound, DNMT1 inhibition is released (Fig. 1.5B; Zhang et al. 2015). The BAH domains of DNMT1 are suspected to serve as a platform for interaction with other proteins (Song et al. 2011).

1.3.1.2. *De novo* DNA methylation

DNMT3A, DNMT3B and DNMT3L belong to the DNMT3 family of methyltransferases, with DNMT3A and DNMT3B being responsible for the *de novo* establishment of DNA

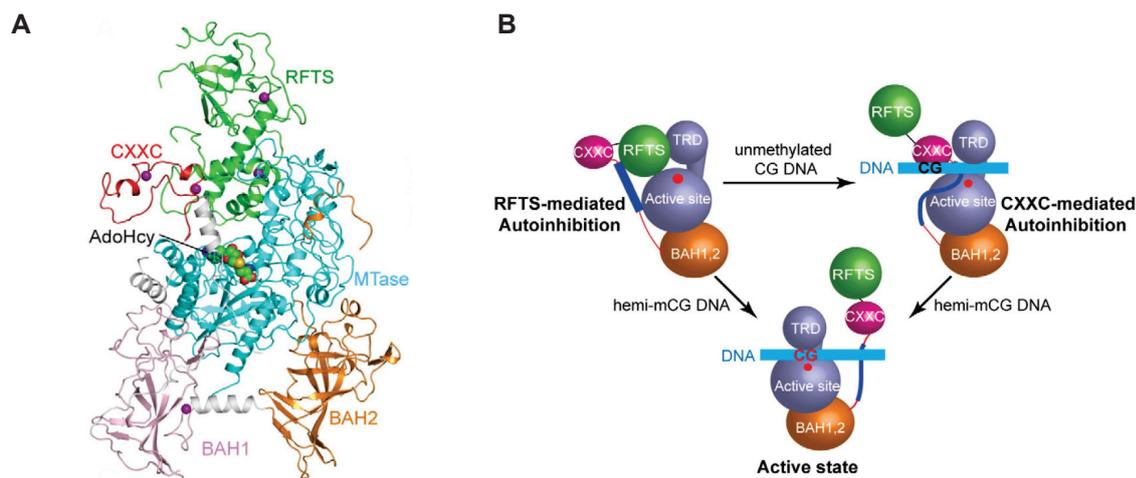


Figure 1.5.: Structural overview of human DNMT1 bound to S-adenosyl homocysteine (AdoHcy) and schematic representation of the autoinhibitory mechanism (Zhang et al. 2015). **A** Ribbon representation with color-coding of RFTS, CXXC, BAH1, BAH2 and methyltransferase domains as well as domain linkers (silver) and zinc ions (purple). **B** Conformational transitions of DNMT1 upon binding to different DNA substrates with a key role of the CXXC-BAH1 linker. TRD = Target Recognition Domain of the methyltransferase.

methylation patterns and DNMT3L as a catalytically inactive regulatory factor for *de novo* methylation (Jurkowska, Jurkowski, and Jeltsch 2011). DNMT3L was found to co-localize and co-immunoprecipitate with DNMT3A and DNMT3B (Hata et al. 2002) and to enhance their *de novo* methylation potential (Chedin, Lieber, and Hsieh 2002). Two monomers of DNMT3A (or DNMT3B) form a 16 nm long, tetrameric complex with two monomers of DNMT3L (Fig. 1.6A). In this complex DNMT3L stabilizes the active site of DNMT3A and thereby promotes methylation activity. The ADD domains present in both proteins mediate association of the complex to nucleosomes (which exhibit a diameter of 11 nm) by binding to unmethylated histone tails, which induces *de novo* methylation of DNA (Fig. 1.6B; Cheng and Blumenthal 2008; Ooi et al. 2007). Accordingly, histone methylation was shown to protect promoters from *de novo* methylation (Appanah et al. 2007). Recently it was shown that the ADD domain of DNMT3A also interacts with the catalytic domain and blocks its activity in an autoinhibitory conformation that is disrupted by histone H3 (and not H3K4me) binding (Guo et al. 2015b).

1.3.2. DNA demethylation

The question of how DNA methylation marks might be removed from DNA strands was answered in 2009, when a methylcytosine dioxygenase family of enzymes (TET1, TET2 and TET3) was discovered and shown to be able to modify 5-methylcytosine through oxidation to 5-hydroxymethylcytosine (5hmC; Tahiliani et al. 2009). In following experiments TET enzymes were furthermore shown to be capable of iterative oxidation of 5hmC that results in

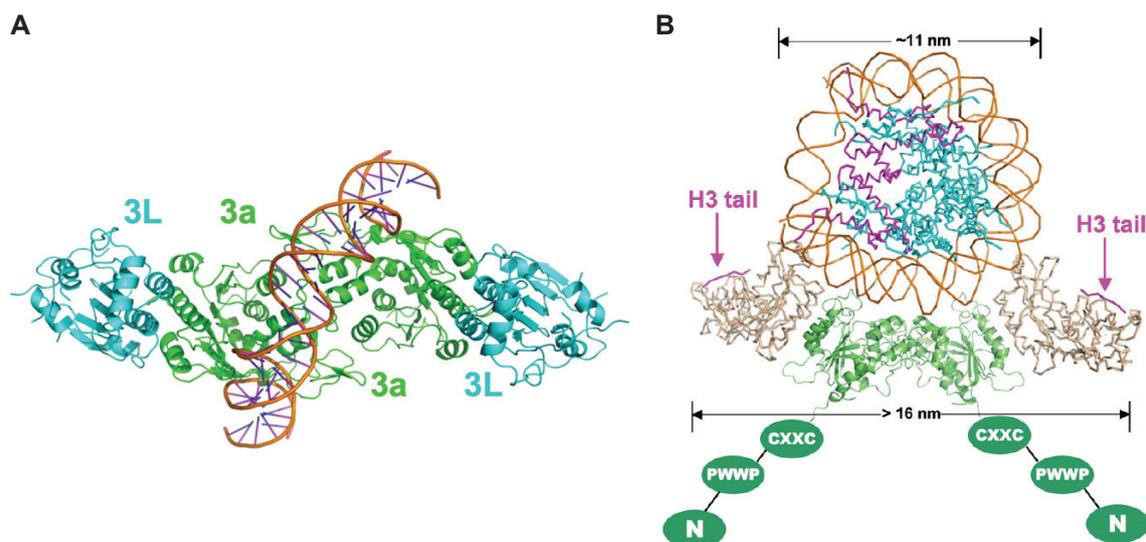


Figure 1.6.: DNMT3A/3L interaction (modified from Cheng and Blumenthal 2008). **A** DNMT3A/3L-tetramer with a DNA molecule (red) that covers two active sites. **B** DNMT3A/3L-tetramer interaction with a nucleosome. DNMT3A is shown in green, DNMT3L in grey. The DNMT3L molecules is capable of binding both histone tails if the DNMT3A/DNMT3L-tetramer is wrapped around the nucleosome. The amino-proximal part of DNMT3A is shown as a cartoon (with N = variable region at amino terminus, PWWP = region involved in unspecific DNA binding and CXXC = unmethylated CpG binding region) and might interact with histone tails from neighboring nucleosomes through the CXXC domain.

the formation of 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC; Fig. 1.7A; Ito et al. 2011). Notably, all three modifications are chemically distinct and could therefore function as substrates for different enzymes.

Three mechanisms of possible roles of these modifications in the regeneration of unmodified cytosine bases were proposed (Kohli and Zhang 2013): In the first scenario, oxidized bases could be diluted in a DNA-replication dependent manner ('AM-PD' in Fig. 1.7A). In a second hypothetical scenario, a direct removal of the C5-position substituent is proposed. Accordingly, it was shown that decarboxylation of 5caC could restore unmodified cytosine - however, a decarboxylase responsible for this process has yet to be identified (Schuesser et al. 2012). Furthermore, DNMTs were shown to function in demethylation events when their methyl-group donor S-adenosyl-methionine (SAM) was missing, suggesting a 'reverse' DNMT reaction that remains to be further explored as well (Kohli and Zhang 2013). The third pathway to restore unmodified cytosine, which is considered most reliably confirmed yet, is based on active cytosine restoration by involvement of enzymes of the base excision repair (BER) pathway, that has been shown to be linked to genome-wide changes in DNA methylation, e.g. in mouse germ cells or zygotes (Hajkova et al. 2010; Wossidlo et al. 2010). In fact, Thymine DNA glycosylase (TDG), that also functions as a DNA repair enzyme capable of removing thymines from T-G-mismatches, was shown to exhibit base excision activity on 5fC and 5caC *in vitro* (Maiti and Drohat 2011). This specificity could be due to

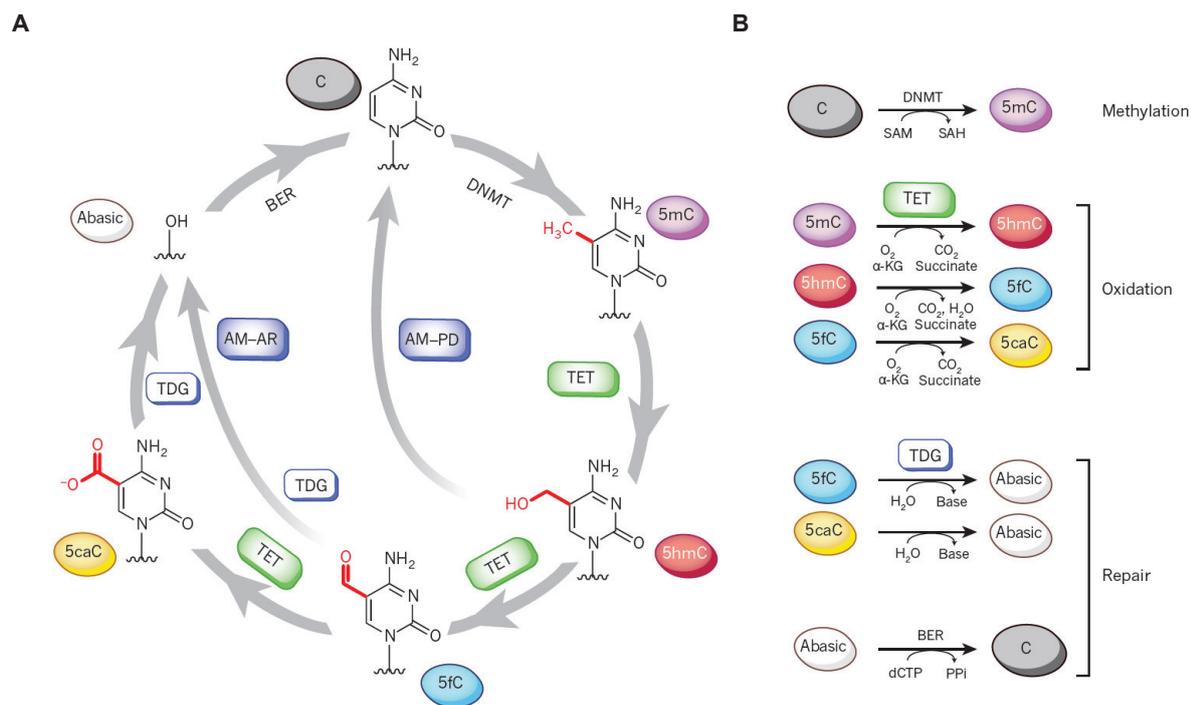


Figure 1.7.: Pathway of dynamic cytosine methylation and proposed functions of 5hmC (modified from Kohli and Zhang 2013). **A** Pathway of dynamic cytosine modifications. Cytosine methylation is mediated by DNMTs. 5-methylcytosine (5mC) can be iteratively oxidized to 5hmC, 5fC and 5caC. Oxidation products can either be passively diluted in a replication-dependent manner ('active modification-passive dilution' (AM-PD) pathway) or actively restored (AM-AR) by TDG-mediated excision during the base excision repair (BER) process. **B** Representation of individual reactions.

the fact that 5fC and 5caC have destabilized N-glycosidic bonds compared to unmethylated cytosine, 5mC and 5hmC (Williams and Wang 2012).

TET-mediated oxidation of 5mC followed by both passive dilution and active removal by TDG/BER plays a key role in the drastic DNA methylation changes during early development (Messerschmidt, Knowles, and Solter 2014). In somatic cells, TET enzymes have furthermore been associated with locus-specific demethylation events, e.g. in hormonal or cytokine-mediated regulation. This suggests a general role of this pathway in the modulation of transcriptional control without the necessity of DNA replication taking place (Kohli and Zhang 2013).

1.4. DNA methylation and transcriptional regulation

The following section describes how DNA methylation influences gene expression. It also underlines the fact that many aspects of this relationship are still unknown. For example, methylation of specific genomic regions with a high density of CpG dinucleotides (so-called

CpG islands) is related to transcriptional repression in some cases of long-term silencing (see e.g. Herzing et al. 1997), but CpG island methylation does not change in many other cases despite of changes in transcriptional activity of linked genes (Schübeler 2015). Furthermore, although DNA methylation at genomic regions outside of CpG islands was shown to be related to gene regulation (Stadler et al. 2011), it is not entirely clear if DNA methylation at these sites is instructive or a result of transcription factor binding.

1.4.1. CpG islands

The largest part of the mammalian genome exhibits comparatively low levels of CpG dinucleotides, as these are prone to spontaneous deamination when they are methylated, resulting in the formation of thymine ('C to T transition'; Lander et al. 2001; Walser, Ponger, and Furano 2008). Due to inefficient base-excision repair processes this led to CpG depletion over evolutionary time (Schübeler 2015). An exception to this finding are genomic regions with a high density of CpG dinucleotides that are called CpG islands (CGIs). CGIs were discovered in experiments using the methylation-sensitive restriction enzyme *HpaII* to digest genomic DNA, which showed that about 1 % of the genomes of all analyzed tissues were detectable as separate fractions with a high number of *HpaII* restriction sites (i.e. unmethylated CpG positions) in a gel (Bird et al. 1985). This discovery led to the annotation of about 26,300 CGIs in the human genome that were defined by a length of at least 200 bp, a G+C content of at least 50 %, and a ratio of observed to expected CpGs of at least 0.6 (Antequera and Bird 1993; Illingworth and Bird 2009; Gardiner-Garden and Frommer 1987). These parameters are still used to predict CGIs within DNA sequences. To reduce the risk of identifying false positive sequences such as Alu elements (the most abundant repetitive element in primate genomes; Zhang, Chen, and Zhao 2015) which resemble CGIs regarding their base composition, Takai and Jones proposed in 2002 to raise the minimal required length to 500 bp, the G+C content to 55 % and the observed-to-expected ratio to 0.65 (Takai and Jones 2002).

70 % of all CGIs in the human genome are associated with promoter regions of annotated genes, including all constitutively expressed genes (Bird 2002). CGI-associated promoters often allow for transcription initiation from different positions, whereas most promoters without CpG islands initiate transcription from a single position ('broad type' and 'sharp type' promoters; Sandelin et al. 2007). However, not all CpG islands co-localize with annotated transcription start sites - so-called 'orphan CGIs' are suspected to either be associated with tissue-specific transcription or to represent sites of unknown transcripts or non-coding RNAs (Ndlovu, Denis, and Fuks 2011).

In most human and murine genomes the majority of CGIs is unmethylated, whereas the remainder of the genome is highly methylated. For this reason most mammalian genomes show a bimodal distribution of DNA methylation. Only a minority of CpG islands (about 10 to

25 %) becomes methylated during embryonic development (Eckhardt et al. 2006; Zeschnigk et al. 2009). CGI methylation can cause stable transcriptional repression of linked genes and has been reported in all cases of long-term silencing such as X chromosome inactivation and genomic imprinting (Schübeler 2015). In some cases CGI methylation events are furthermore associated with malignant transformations. For example, genes of the Melanoma antigen encoding gene (MAGE) family are, in germ cells, associated with unmethylated CGIs which become methylated during differentiation. In some kinds of cancer cells MAGE methylation is lost and the genes are expressed again (Honda et al. 2004). However, in many cases inactive CGI promoters are linked to methylation of lysine 27 of histone H3 (H3K27me3) instead of acquiring DNA methylation, whereas inactive promoters with low CpG-contents are often found to be methylated (Schübeler 2015; Fig. 1.8).

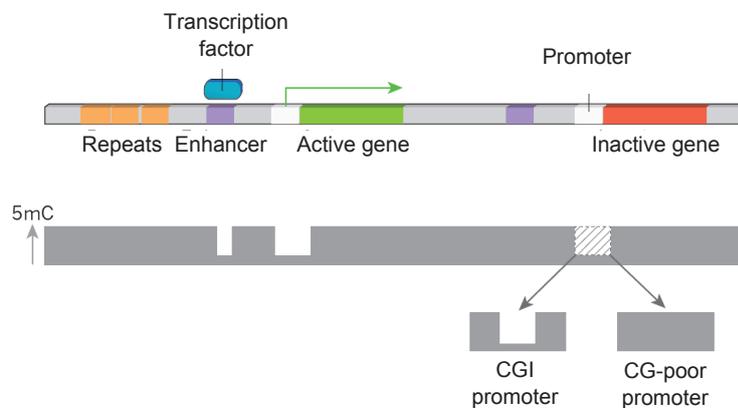


Figure 1.8.: Schematic representation of bimodal DNA methylation in a typical mammalian genome and its correlation with transcription of associated genes (modified from Schübeler 2015). Whereas methylation levels are high at intergenic, repeat and intragenic regions, promoter and enhancer regions often overlap with CGIs which generally remain unmethylated, no matter if the associated gene is active or not. Inactive CpG-poor promoters are generally methylated.

1.4.2. DNA methylation changes at CpG-poor genomic regions

A general decrease of DNA methylation levels associated with binding of transcription factors was described for CpG-poor, regulatory genomic regions, and such variable DNA methylation at CpG-poor regions was proposed to closely reflect on-going gene activity (Schübeler 2015). However, it is still unclear if DNA methylation has a generally repressive effect by facilitation or inhibition of binding of transcription factors, or if DNA methylation changes occur downstream of transcription factor binding. Fig. 1.9 shows proposed scenarios for the interaction between transcription factors and DNA methylation, which include reduction of methylation levels by binding of methylation-insensitive transcription factors, specific binding of transcription factors to methylated genomic regions and repulsion of methylation-sensitive transcription factors by methylation of genomic regions. Furthermore, methyl-CpG-

binding domain (MBD) proteins might inhibit transcription factor binding by interaction with methylated regions that exhibit a high density of CpG dinucleotides (shading in Fig. 1.9D). Finally, methylation insensitive transcription factors might create sites of reduced methylation for methylation-sensitive transcription factors (Schübeler 2015). It has been shown that transcriptional regulators can trigger demethylation events of endogenous genes in cultured cells, which underlines the hypothesis that DNA methylation changes occur downstream of transcription factor binding (Kress, Thomassin, and Grange 2006).

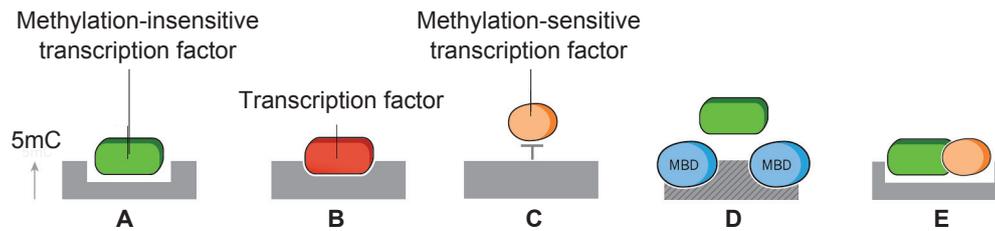


Figure 1.9.: Proposed interactions between transcription factors and DNA methylation at CpG-poor genomic regions (modified from Schübeler 2015). **A** Methylation levels might be reduced by binding of methylation-insensitive transcription factors. **B,C** Methylated genomic regions might either attract or repulse specific transcription factors. **D** Methyl-CpG-binding domain (MBD) proteins might inhibit transcription factor binding by interaction with methylated regions that exhibit a high density of CpG dinucleotides (shading). **E** Methylation-insensitive transcription factors might mediate changes in DNA methylation levels that enable binding of methylation-sensitive factors.

1.5. Methods for the analysis of DNA methylation

Methyl groups are removed from DNA molecules when these are subjected to methods of molecular biology such as PCR and cloning, and are not detectable by hybridization methods as well. Therefore, the detection of methyl-groups at cytosine bases requires one of three general methods, which are based on (i) the digestion of genomic DNA with methylation-sensitive restriction enzymes, (ii) the conversion of unmethylated cytosines to uracil by treatment of DNA with bisulfite or (iii) the enrichment of either unmethylated or methylated DNA by using proteins with specific binding domains. Each of these methods can be combined with gene specific analyses or genome-wide detection of DNA methylation patterns by microarrays or next generation sequencing techniques.

1.5.1. Utilization of methylation-sensitive restriction enzymes

The first group of methods that is based on digestion of native genomic DNA with methylation-sensitive restriction enzymes such as *HpaII*, led to the identification of CpG islands, which appeared as a distinct fraction of DNA fragments in a gel (see section 1.4; Bird et al. 1985).

The technique Methylation-Sensitive Restriction Enzyme Digestion PCR (MSRE-PCR) is also based on digestion of genomic DNA with methylation-sensitive restriction enzymes. DNA digestion is followed by multiplexed PCR to enable determination of the methylation status of several promoter regions in the same sample (Melnikov et al. 2005). Direct digestion of genomic DNA is furthermore used for the microarray-based technique Differential Methylation Hybridization (DMH; Yan et al. 2002). For DMH, genomic DNA is first cut with *MseI* which leaves CpG islands intact and cuts bulk DNA into small fragments. CpG islands are then ligated to linkers and cut with methylation-sensitive restriction enzymes. Only CpG islands that remain intact and keep their linkers can subsequently be amplified and hybridized to a microarray.

1.5.2. Bisulfite-mediated conversion of unmethylated cytosines

The second group of methods is based on bisulfite conversion of DNA. Treatment of genomic DNA with sodium bisulfite results in conversion of unmethylated cytosine bases to uracil (Frommer et al. 1992). Bisulfite-mediated deamination of cytosine takes place by reversible addition of HSO_3^- , followed by liberation of NH_3 by hydrolysis and a subsequent release of HSO_3^- , which results in the formation of uracil (Fig. 1.10; Hayatsu 2008). Several analysis methods can subsequently be performed which employ the fact that the newly generated uracil bases are amplified as thymines when bisulfite-treated DNA is subjected to a PCR reaction.

Most commonly, bisulfite-treated DNA is either subjected to amplification of genomic regions of interest followed by subcloning and Sanger sequencing (Frommer et al. 1992) or generation of sequencing libraries for genome-wide analysis of DNA methylation patterns by next generation sequencing (Bock 2012; Ziller et al. 2013; Lou et al. 2014; Rauch et al. 2012). It is furthermore possible to interrogate the methylation status of specific loci by PCR using primers designed for either unmethylated or methylated DNA (Methylation-Specific PCR (MSP); Herman et al. 1996) or by Real-Time PCR-based analysis of differences in melting temperatures (Methylation-Sensitive High-Resolution Melting (MS-HRM); Wojdacz, Dobrovic, and Hansen 2008). Another method applied to analyze bisulfite-treated genomic DNA is Combined Bisulfite Restriction Analysis (COBRA). COBRA is based on the fact that, due to the methylation-dependent conversion of cytosines, specific restriction enzyme recognition sites are formed or lost in the resulting PCR products (Xiong and Laird 1997; Bilichak and Kovalchuk 2017). Common enzymes used for digestion of such PCR products are *TaqI* (TICGA) or *BstUI* (CG|CG).

Bisulfite treated DNA can also be analyzed by specific microarrays such as Illuminas *Infinium*[®] *HumanMethylation450 BeadChip* which was designed to cover 450,000 methylation sites in the human genome at single-nucleotide resolution. For this assay, genomic DNA is treated with bisulfite followed by whole genome amplification (WGA). Each CpG site is covered by

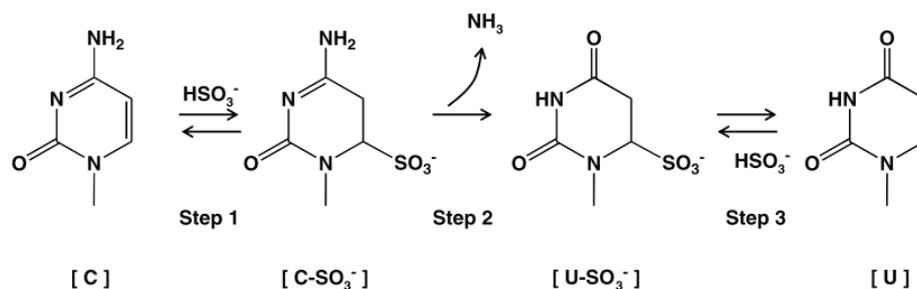


Figure 1.10.: Conversion of cytosine to uracil by bisulfite-treatment (Hayatsu 2008).

two types of beads present on the array, one corresponding to the methylated variant of the respective locus and the other corresponding to the unmethylated variant. Depending on the methylation status DNA fragments anneal to one of these two bead types and fluorescence detection allows for identification of the bead type (Bibikova et al. 2009). The *Infinium*[®] *HumanMethylation450 BeadChip* assay is widely used in epigenome-wide association studies (EWAS; see e.g. Rakyan et al. 2011) and in studies within the scope of the efforts of the Human Epigenome Consortium (Morris and Beck 2015).

1.5.3. Enrichment of methylated or unmethylated DNA

Using the third detection principle for cytosine methylation, methylated DNA fragments can be enriched from a pool of sheared, untreated genomic DNA using an antibody against methylated DNA (Methylated DNA Immunoprecipitation (MeDIP)) or a Methyl-CpG binding domain (MBD) protein (Weber et al. 2005; Rauch et al. 2006a; Bock 2012). When both methods were compared, the 5mC-antibody was shown to exhibit a higher affinity towards DNA sequences with a low CpG density, whereas MBD-based enrichment favored regions with a higher density of CpGs. Accordingly, by using an MBD protein the highest proportion of CpG islands could be enriched (Nair et al. 2011). It is furthermore possible to specifically enrich unmethylated DNA fragments by using zinc finger CXXC-domain containing proteins like DNMT1, which physiologically recruit chromatin-modifying complexes to unmethylated DNA (Blackledge et al. 2011; Long, Blackledge, and Klose 2013). Enriched fractions can subsequently be analyzed by microarray hybridization (Rauch et al. 2006b) or next generation sequencing (Li et al. 2010).

1.5.4. State-of-the-art DNA methylation analysis and its limitations

Due to technological advances and reduced costs of whole-genome sequencing experiments, DNA methylation analyses shifted from gene-specific to genome-scale approaches. Whole-genome bisulfite sequencing is considered the current 'gold standard method' as it enables a

higher genome coverage than microarrays and provides DNA methylation data in the highest possible resolution. However, as this method still requires a sufficient sequencing depth and is not feasible for large numbers of samples (e.g. patient cohorts), methylation microarrays such as the *Infinium*[®] *HumanMethylation450 BeadChip* are popular as well. By such experiments, a variety of DNA methylation maps was generated which revealed several interesting, yet not completely understood, insights into an overall effect of DNA methylation on certain phenotypes. For example, inter-individual variability of drug metabolizing enzymes was shown to be associated with DNA methylation-dependent control of alternative splicing events (Habano et al. 2015). In human placenta tissue, global DNA methylation analyses resulted in the identification of apparently unstructured, partially methylated domains (PMDs) that could be associated with specific functional groups of genes related to developmental processes (Schroeder and LaSalle 2013). DNA methylation marks were furthermore shown to represent suitable biomarkers. For example, promoter methylation of the *Fragile X mental retardation 1 (Fmr1)* gene was shown to be associated with differences in the response of patients to a specific antagonistic drug (Jacquemont et al. 2011).

However, there are still some limitations to each method. Incomplete bisulfite conversions and bisulfite PCR errors are considered to be sources of bias in DNA methylation analyses by next generation sequencing as well as bisulfite-based microarray experiments. These microarrays can also be biased by cross-hybridizations (Laird 2010). As the *Infinium*[®] *HumanMethylation450 BeadChip* assay is a tool which is used frequently (De Meyer et al. 2015; Yue et al. 2016; Smith et al. 2016), several studies exist that analyze its performance in general and also in comparison to sequencing-based methodologies. In these studies it was shown that SNPs present in the analyzed genomes may lead to false interpretations of methylation signal differences (Daca-Roszak et al. 2015) and that cross-reactive probes and polymorphic CpGs exist on this array (Chen et al. 2013). According to this multitude of potential biases that exist for each individual method, a comparison of the *Infinium*[®] *HumanMethylation27 BeadChip* assay (the predecessor of the *HumanMethylation450 BeadChip*) to different genome-wide sequencing-based methods showed that only about one-fifth of the differentially methylated regions identified by whole-genome sequencing approaches were detected by the microarray-based method (Bock et al. 2010). Between sequencing-based methods, differences in the detection of differentially methylated regions (DMRs) were shown in this publication as well, although the results of these methods were generally shown to overlap to a larger extent (Walker et al. 2015).

1.5.4.1. Novel technologies

When the methylcytosine dioxygenase family of TET enzymes was discovered in 2009 and 5mC was shown to be actively modified to 5hmC, 5fC and 5caC, followed by cytosine

restoration by the TDG enzymes (Tahiliani et al. 2009), methods were developed to map these modifications in single-base resolution on a genome-scale. Booth et al. developed the method Oxidative Bisulfite Sequencing (oxBS-Seq) to discriminate 5mC and 5hmC, as both modified bases are protected from bisulfite conversion. OxBS-Seq is based on the conversion of 5hmC to 5fC, which is converted to uracil by bisulfite treatment and is therefore read as thymine in sequencing. By combination of bisulfite sequencing and oxBS-Seq data the identification of 5hmC sites is possible (Fig. 1.11, left; Booth et al. 2012). 5fC maps could be generated using Reduced Bisulfite Sequencing (redBS-Seq; Fig. 1.11 right). This method employs the fact that 5fC is reduced to 5hmC through treatment of DNA with sodium borohydride, and 5hmC, when treated with bisulfite, is resistant to the bisulfite conversion reaction. Alignment of bisulfite reads from reduced and non-reduced DNA allows for single-base mapping of 5fC (Booth et al. 2014).

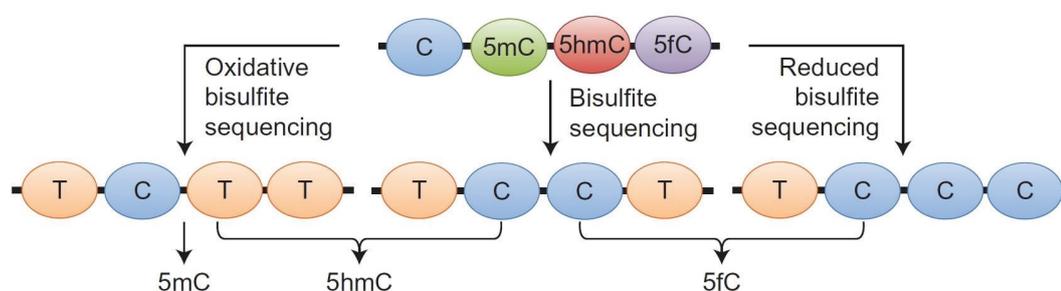


Figure 1.11.: Principle of Oxidative Bisulfite Sequencing (oxBS-Seq) and Reduced Bisulfite Sequencing (redBS-Seq; Booth et al. 2014).

5caC detection was shown to be possible by Chemical Modification-Assisted Bisulfite Sequencing (CAB-Seq), which is based on chemical labeling of 5caC and changed behavior of the modified base during bisulfite conversion (Lu et al. 2013). Thereby, it was e.g. shown that 5fC and 5caC accumulated at hypomethylated promoters in embryonic stem cells (Neri et al. 2015). Bachman et al., as another example, analyzed temporal dynamics of 5fC levels in mouse DNA and found them to be stable and not correlated to the levels of the precursors 5mC and 5hmC (Bachman et al. 2015). As 5fC has a genomic profile distinct from the ones of its precursors and was shown to alter the DNA structure, the authors hypothesized that 5fC is not only a demethylation intermediate, but influences gene regulation distinctly to 5mC.

In parallel to the development of novel technologies for mapping of 5mC intermediates, bisulfite sequencing of single cells is attracting increasing attention, as the current whole-genome approaches require millions of cells and do not allow for studies of cell populations or cell-to-cell heterogeneity. Recently published protocols for single-cell methylomics were already tested in different biological systems (see e.g. Farlik et al. 2015) and will eventually enable the elucidation of phenomena such as the above-mentioned partially methylated domains.

1.6. Epigenetic studies in the production of biopharmaceuticals

While epigenomics in human cells and tissues got into the focus of scientists after completion of the Human Genome Project in 2001, it took about a decade longer for the biotechnological community to address epigenetic phenomena in the production of biopharmaceuticals on a genome-scale. This is in parts due to the fact that the CHO genome and transcriptome sequences are only available since 2011 (Xu et al. 2011; Becker et al. 2011). Therefore, most of the current studies focus on recombinant genes and regulatory elements. Most recently, a first report on epigenetically engineered CHO cell lines was published, underlining the increasing amount of evidence for the importance of epigenetic phenomena in these cells (Tastanova et al. 2016).

1.6.1. Analysis of single recombinant genes and promoters

Several studies were conducted that analyzed recombinant genes or promoters. Amongst them, many were specifically interested in methylation changes of the cytomegalovirus (CMV) promoter. Kim and co-workers, for example, analyzed monoclonal antibody producing CHO cell lines during extended sub-culture in the presence of selective pressure (Kim et al. 2011). A detailed examination of CMV promoter methylation (29 CpG sites) was performed using bisulfite sequencing and revealed that the instable and low generation number cell lines were affected by pCMV methylation that strongly increased during sub-culturing. The correlation between pCMV methylation, instability of productivity and its potential to be used as a biomarker was also addressed by Osterlehner et al., who established a qPCR-based screening method allowing for enrichment of stable producers early in cell line development (Osterlehner, Simmeth, and Göpfert 2011). In a follow-up study that aimed to identify further markers of production stability, histone modifications in the vicinity of the CMV promoter were analyzed and proved to be even more indicative of stability than DNA methylation. Enrichment of stable producer cells was shown to be able by selecting those clones with the highest levels of histone H3 acetylation (Moritz et al. 2016). Adding to this is a study which examined DNA methylation and chromatin modifications along the human Eukaryotic translation elongation factor 1 alpha 1 (EEF1A1) promoter as part of an antibody expression cassette in CHO cells (Veith et al. 2016). Here, also, histone acetylation was associated with high expression levels and was found to be more predictive than DNA methylation. As a consequence of these findings, Saunders et al. assessed different chromatin modifying elements for their potential to negate negative effects of the insertion sight and maintain transgene expression. The elements were integrated either between sequences coding for heavy and light antibody chains or in flanking positions. The A2UCOE (Ubiquitously acting

chromatin-opening element derived from the human HNRPA2B1-CBX3 locus) element was shown to be able to increase antibody yields about 6.5-fold, to mitigate the reduction of antibody production over 120 generations and to reduce transgene promoter methylation (Saunders et al. 2015).

Besides promoter regions, recombinant genes were also analyzed regarding their potential to influence gene expression by induction of epigenetic mechanisms. In 2010 Bauer and colleagues analyzed the impact of intragenic CpG content on gene expression in transiently and stably transfected CHO cells (Bauer et al. 2010). By modulation of the codon usage of green fluorescent protein (GFP), the group created a CpG-depleted version of the reporter and observed a promoter-independent and severe decrease in mRNA and protein abundance upon CpG depletion. The decrease in gene expression was shown to be due to reduced *de novo* transcription in nuclear run-on assays, and it was speculated that this indicated an influence of CpG content on the chromatin structure along the open reading frames. In a similar but extended study, the same comparison between the original GFP gene and a poorly expressed, CpG depleted version, was performed using various promoters (CMV, EF-1alpha and A2UCOE) and different cell systems (CHO and embryonic pluripotent carcinoma cells). CpG depletion was shown to increase local chromatin density, and usage of the CMV promoter led to silencing of recombinant genes irrespective of CpG content. In contrast, the EF-1alpha promoter enabled high expression levels and delayed silencing. Silencing phenomena could be prevented completely by using the A2UCOE promoter, which was hypothesized to be due to its chromatin-opening abilities (Krunner et al. 2015).

1.6.2. Epigenetic engineering

A first report of epigenetic engineering in the context of recombinant protein production was published by Tastanova et al., who screened 29 polycomb proteins for their potential to enhance recombinant gene expression in CHO cells (Tastanova et al. 2016). Upon binding, polycomb proteins are known to recruit chromatin-remodeling proteins and modify the chromatin (both locally and globally) surrounding regulatory elements. Thereby, they regulate key developmental genes (Entrevan, Schuettengruber, and Cavalli 2016). In the screening experiment, the Yin yang 1 (YY1) protein had the greatest impact on the expression of several product genes and CHO cell lines stably expressing YY1 in moderate amounts showed 6-fold higher antibody titers. Interestingly, higher levels of YY1 had the opposite effect on recombinant gene expression. This is in accordance with a previous study that shows a repressive effect of YY1 on the CMV promoter in CHO cells (Brown et al. 2015). The switch between the activating role of moderate YY1 levels and the repressive effect of YY1 over expression was shown in other contexts as well (Deng et al. 2009) and is referred to as a so-called 'squelching effect' (Gill and Ptashne 1988).

1.6.3. Genome-scale epigenomics in long-term cultured cells

In contrast to the multitude of publications about global DNA methylation phenomena in human and mouse, only one estimation of global DNA methylation in long-term cultured cells existed previously to the beginning of this thesis and one study of global epigenetic characteristics of CHO cells was published in 2016, when the major experiments of the study presented here were already performed and published.

This first estimation of global DNA methylation in long-term cultured cells was made in 1990, when Antequera et al. showed that CpG island methylation has a pivotal role in permanent inactivation of genes in cultured cells. They hypothesized that transcriptional repression was often falsely related to mutations (Antequera, Boyes, and Bird 1990) based on experiments using permanent human and murine cell lines which exhibited unexpectedly high numbers of methylated CpG islands. These were mostly associated with genes unnecessary in culture. High methylation levels were e.g. detected for the *Alpha-globin* gene. Housekeeping genes were not associated with methylated CpG islands, nor were genes relevant for the selection process (*Dihydrofolate reductase* gene, *Dhfr*; first shown by Stein et al. 1983). As similar cell lines showed comparable methylation patterns, the authors hypothesized that about 50 % of all CpG islands become methylated in long-term cultured cells.

In 2016, Feichtinger et al. analyzed six related CHO cell lines regarding their genome sequences and DNA methylation patterns as well as global histone modifications under different culture conditions, in different media and after subcloning (Feichtinger et al. 2016). Generally, all CHO cell lines showed a bimodal distribution of either hypo- or hypermethylated CpGs with overall methylation levels of 55–65 %. The numbers of differentially methylated regions between the cell lines were shown to be low in response to passaging and during batch cultivation, and high as a consequence of several rounds of subcloning and selection of a certain phenotype or in response to media changes. In contrast, histone modifications were found to significantly change during a batch culture and growth phases could clearly be distinguished. The authors hypothesized that DNA methylation changes play a major role in defining CHO cell phenotypes, while short term changes in gene expression, like those related to different nutrient availabilities in batch cultures, are rather associated with alterations in histone marks.

1.7. Butyrate: Physiological effects and biotechnological application

Butyrate is a short chain fatty acid (SCFA; Fig. 1.12) which is generated in high levels (mM range) during fermentation of carbohydrates in the colon (Steliou et al. 2012).

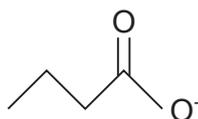


Figure 1.12.: Chemical structure of butyrate.

It was shown to be the primary energy source and an anti-inflammatory agent for normal colorectal epithelium and, simultaneously, to prevent growth of colorectal cancer cells by inhibiting Histone deacetylases (HDACs), thereby altering the expression of genes with functions in proliferation, apoptosis and differentiation. On the contrary, butyrate stimulates growth of normal colon cells. This phenomenon is referred to as 'the butyrate paradox' (Fleming et al. 1991; Hamer et al. 2008; Zhang et al. 2010; Donohoe and Bultman 2012).

1.7.1. Cells respond differently to butyrate due to the Warburg effect

What are putative reasons for the 'butyrate paradox'? In 2012 Donohoe et al. provided proof that the differences in cell metabolism that lead to the contradicting behavior of normal and transformed cells in response to butyrate are due to the Warburg effect (Donohoe and Bultman 2012) This effect describes the phenomenon of aerobic glycolysis being the predominant process of energy generation in cancer cells (and not oxidative metabolism; Vander Heiden, Cantley, and Thompson 2009). Donohoe et al. suggested that in cancer cells butyrate is not metabolized in the mitochondria, but accumulates in the cell and inhibits histone deacetylases, thereby regulating genes that inhibit proliferation and induce apoptosis (Fig. 1.13). Accordingly, Cyclin-dependent kinase inhibitor 1 (p21), Nuclear factor kappa B (NFkB) and components of the Mitogen-activated protein kinase (MAPK) pathway have previously been shown to be affected by butyrate (Archer et al. 2005; Lührs et al. 2002; Scott, Longpre, and Loo 2008; Kolar et al. 2007). In their experiment, Donohoe et al. prevented the Warburg effect from occurring in cancer cells and showed that this was sufficient to induce butyrate-mediated induction of growth. When higher doses of butyrate were used, proliferation was inhibited nevertheless, as the cellular capacities for oxidation of the SCFA were used up. The study also showed that butyrate not only functions as an HDAC inhibitor - a butyrate function that has long been known (Boffa et al. 1978) - but also as a histone acetyltransferase (HAT) activator due to its contribution to acetyl-CoA production when it is oxidatively metabolized (Donohoe and Bultman 2012; Donohoe, Curry, and Bultman 2013).

1.7.2. Effects of butyrate on DNA methylation

Although research on butyrate-induced effects has mainly focused on its role as an HDAC inhibitor yet, several lines of evidence also hint towards a contribution of butyrate-induced

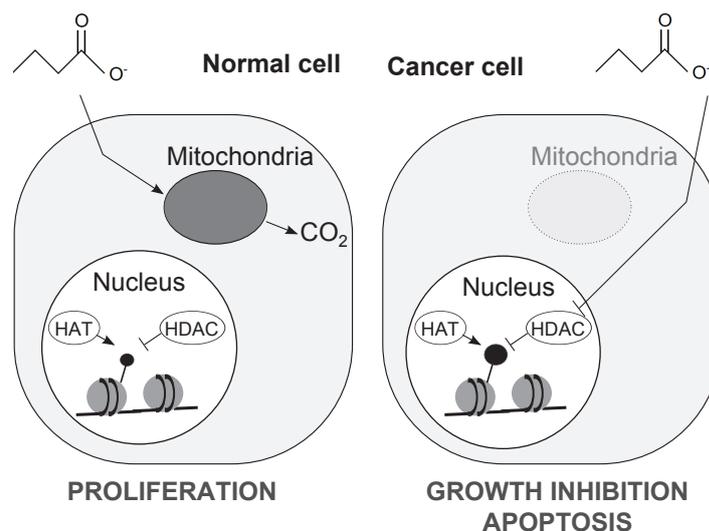


Figure 1.13.: Differences in cellular metabolism between normal and cancer cells result in opposing effects upon butyrate addition (figure based on Donohoe et al. 2013).

changes in DNA methylation to its overall effect. A striking hint on a putative major influence of butyrate on DNA methylation came from research on honey bees. Fertile queens and sterile worker bees have identical genetic backgrounds and queen development is induced by feeding a substance called royal jelly. When the effect of RNA interference-mediated knockdown of DNMT3 was analyzed in honey bee larvae, a majority of these larvae emerged as queens and exhibited fully developed ovaries. This suggested that the major phenotypic difference between queen and worker bees is established by nutritional input (Kucharski et al. 2008). Although it is not fully clear yet what components of royal jelly indeed have such a great impact on the epigenetic system, it is speculated that a substance similar to butyrate, phenylbutyrate, which is also a known HDAC inhibitor, primarily contributes to this effect (Chittka and Chittka 2010; Lyko et al. 2010).

Several other key enzymes responsible for setting and removing DNA methylation marks were shown to be affected by butyrate. In a study examining the relationship between butyrate (which is also used as an antidepressant) and DNA methylation in brain tissue, lower expression levels of TET1 were found that were associated with increased hydroxymethylation of a gene relevant for neurogenesis and synaptic plasticity (Wei et al. 2015). As another major regulator of DNA methylation, DNMT1 was also found to be influenced by butyrate. It was shown that butyrate repressed MAP kinase I (ERK) activation, which was linked to downregulation of DNMT1 protein levels and subsequent changes in DNA methylation (Sarkar et al. 2011). Furthermore, butyrate was shown to influence the expression of Growth arrest and DNA damage inducible protein (GADD45) isoforms which are suspected to contribute to DNA demethylation by DNA damage repair (Chen et al. 2002; Rai et al. 2008). Several studies furthermore report changes of gene-specific DNA methylation levels in response to butyrate treatment. For example, butyrate-induced promoter demethylation of

Secreted frizzled-related protein (SFRP) proved to be related to changes in SFRP gene expression and subsequent inhibition of WNT signaling in gastric cancer (Shin et al. 2012). It was also shown that butyrate exposure resulted in histone acetylation and decreased DNA methylation of *Gamma-globin* genes upon pharmacological, butyrate-dependent induction of fetal hemoglobin (Fathallah et al. 2007).

1.7.3. The effect of butyrate on CHO cell cultures

The first report of an effect of butyrate addition on morphology and growth of Chinese hamster cells was published in 1973 (Wright 1973). Only few years later, this effect was related to cellular energy processes, and in 1980 it was shown that butyrate influences cell cycle progression by induction of G1 phase arrest and histone acetylation in CHO cells (Storrie, Puck, and Wenger 1978; D'Anna, Tobey, and Gurley 1980). Okabayashi et al. were amongst the first to report of an effect of butyrate addition on the production of a recombinant gene in CHO cells. They expressed a *Preprourokinase* gene in CHO-K1 cells under the control of the Simian virus 40 early region (SV40) promoter. Addition of 5 mM butyrate led to an increase in productivities of 2-3-fold (Okabayashi, Kanéda, and Arimura 1989). Since then, many publications proved the productivity enhancing effect of butyrate on CHO cells expressing recombinant proteins. The effect was shown for proteins such as Tissue plasminogen activator, Thrombopoietin, Glycosyltransferases or several monoclonal antibodies, and it was observed for both the CMV and the SV40 promoter (Palermo et al. 1991; Jiang and Sharfstein 2008; Klausning, Krämer, and Noll 2011; Sung et al. 2004; Choi et al. 2005).

The positive effect of butyrate addition on CHO cells are not only accompanied by detrimental effects such as growth inhibition and apoptosis. It was furthermore reported that the quality and bioactivity of glycoproteins can be influenced by butyrate, as butyrate treatment resulted in changed galactosylation of monoclonal antibodies (Hong et al. 2014). Lee et al. also proved a correlation of butyrate addition and glycosylation. They showed that butyrate treatment led to a reduction of the relative proportion of the sialic acid content of a recombinant glycoprotein that was accompanied by altered expression of several genes related to glycosylation. These included UDP-glucose pyrophosphorylase or Sialidase-2 (Lee et al. 2014). However, another study examining an effect of butyrate treatment on protein quality with respect to glycosylation did not show significant changes in sialic acid content or antibody activity (Chen et al. 2011).

As butyrate addition apparently did not change product qualities in some cases, several attempts have been made to reduce the detrimental side-effects regarding growth inhibition and induction of apoptosis and enhance the positive effect of butyrate on productivities. For example, a combination of supplementation of 2 mM butyrate and a cultivation temperature of 30 °C was shown to prolong the culture duration and increase the final antibody concentration

(Chen et al. 2011). Suppression of apoptotic pathways by overexpression of Apoptosis regulator Bcl-2 (BCL2) or X-linked inhibitor of apoptosis (XIAP) was also shown to have a positive effect (Sung and Lee 2005; Kim, Leeds, and Chuang 2009). Further studies showed that apoptosis-inhibitory antioxidants such as N-acetylcystein were capable of maintaining long-term survival and enhance production of recombinant Interferon-beta-1a, or that CHO cells deficient in the pro-apoptotic Apoptosis regulator BAX (BAX) and Bcl-2 homologous antagonist/killer (BAK) proteins were able to withstand butyrate-induced apoptosis (Oh et al. 2005; Misaghi et al. 2013).

As knowledge about the mechanism underlying the butyrate effect promises to enable rational cell line development in terms of employing the positive butyrate effects while omitting the bad ones, several 'omics' studies aimed to analyze the effect of butyrate addition on the CHO cell transcriptome or proteome. De Leon Gatti et al., e.g., could show that butyrate affected histone modifications, lipid metabolism and protein processing as well as the expression of chaperone genes (De Leon Gatti et al. 2007). In a study performed by Kantardjieff et al. butyrate was shown to affect expression of genes related to protein secretion and signaling (Kantardjieff et al. 2010). Klausning et al. could furthermore show that genes related to carbohydrate, lipid, amino acid and glycan metabolism were upregulated upon butyrate addition, whereas genes associated with cell growth and apoptosis were downregulated (Klausning, Krämer, and Noll 2011). Yee et al. could also show that butyrate increased the expression of genes involved in protein secretion and redox activity and decreased the expression of genes related to cell cycle and apoptosis (Yee et al. 2008).

2

Research objective

Chinese hamster ovary (CHO) cells represent the major mammalian production cell line for therapeutic proteins. Several aspects of the molecular mechanisms underlying their unique properties remain poorly understood, although a variety of 'omics' studies was performed that analyzed the CHO cell genome (Xu et al. 2011), transcriptome (Becker et al. 2011; Kantardjieff et al. 2010) or proteome (Beckmann et al. 2012; Nissom et al. 2006). A cellular level that has attracted less attention yet is the CHO cell epigenome. The term 'epigenome' summarizes mechanisms which regulate gene expression without altering the DNA sequence and includes methylation of DNA, posttranslational modifications of histones and posttranscriptional gene silencing by small RNAs. DNA methylation in CHO cells has mostly been studied in experiments focusing on single recombinant genes or promoters before. Here, genome-wide DNA methylation and its effect on gene expression was supposed to be analyzed in an exemplary, antibody-producing CHO cell line.

The detection of methyl-groups at cytosine bases can be performed employing the conversion of unmethylated cytosines to uracil by treatment of DNA with bisulfite and subsequent analysis by next generation sequencing. In comparison to other methods for DNA methylation analysis, this technique allows for the generation of DNA methylation data with the highest genome coverage and resolution. Therefore, whole-genome bisulfite sequencing was supposed to be implemented for CHO DP-12 cells that produce a recombinant human anti-IL-8 antibody. Furthermore, whole-genome DNA methylation data were supposed to be combined with gene expression measurements to allow for an assessment of potential phenotypic effects of DNA methylation and functional significance.

Whole-genome bisulfite sequencing and gene expression studies were supposed to be applied to shed light on the DNA methylation landscape of CHO DP-12 cells. As a DNA methylation map for CHO cells did not exist before, this was considered an important step towards understanding CHO cell characteristics. Data generated in this way for several other mammalian cell types offered for example new explanations regarding causalities of diseases such as cancer or could be used to identify diagnostic biomarkers. In a second application the effects of the productivity enhancing short chain fatty acid butyrate on DNA methylation and gene expression were supposed to be analyzed in a time course approach, as supplementation of butyrate is known to enhance cell specific productivities and leads to alterations of epigenetic silencing events.

As another tool for DNA methylation analysis in CHO cells, a specifically designed CpG island microarray, already existed (Wippermann 2012; Wippermann et al. 2013), CpG island microarray performance was supposed to be compared to whole-genome bisulfite sequencing regarding the butyrate effect on DNA methylation. CpG island microarray experiments were supposed to be verified by gene-specific analysis of DNA methylation and control hybridization experiments. Finally, putative central mediators of the butyrate effect were supposed to be identified based on the comparative data analysis and functionally analyzed in transient knockdown experiments by RNA interference.

3

Materials and methods

3.1. Cell culture and cultivation process analysis

3.1.1. CHO cell lines

For analysis of the CHO cell DNA methylation landscape and differential DNA methylation upon butyrate treatment CHO DP-12 cells were used. The effect of siRNA-mediated knockdown of DNMT3A was analyzed in CHO DP-12 cell cultures, too. The strength of endogenous CHO promoters was measured using both CHO DP-12 and CHO K1 cell lines. The comparative analysis of CHO clones with differences in their production stability was performed using CHO-XL99 cell clones.

The CHO K1 cell line (ATCC CCL-61) was derived as a subclone from the parental CHO cell line that was established by Theodor Puck in 1957 (Puck 1958).

Anti-Interleukine-8-antibody-producing CHO DP-12 cells (clone#1934, ATCC CRL-12445) were derived by co-transfection of CHO DP-12 cells with the vectors p6G4V11N35A.choSD.9 and p6G4V11N35E.cho-SD.10 (U.S. patent 6,025,158) that enable the expression of variable heavy and light regions of two variants of the murine 6G4.2.5 monoclonal antibody. Both vectors were generated based on a dicistronic DHFR expression vector that produces the selectable marker and recombinant cDNAs of the heavy and light chain from a single primary transcript by differential splicing (Lucas et al. 1996). The SV40 promoter was used for recom-

binant gene expression. CHO DP-12 cells were transfected by lipofection and methotrexate was used for gene amplification.

CHO-XL99 cells were kindly provided by the Australian Institute for Bioengineering and Nanotechnology (AIBN). The CHO-XL99 clones used here were CHO K1 derivatives (from the original ATCC-61 vial) that have been adapted step-wise to chemically defined conditions. They produce a therapeutic IgG1 antibody. For this thesis CHO-XL99 clones were used that showed differences in the stability of IgG1 production depending on the presence of selection pressure by puromycin.

3.1.2. Cultivation conditions

CHO K1 and CHO DP-12 cells were cultivated in chemically defined, animal component-free medium *CHOMACS*[®] *CD* (Miltenyi Biotec, Bergisch Gladbach, Germany) supplemented with 6 mM L-glutamine (Fagron GmbH&Co.KG, Barsbüttel, Germany). For cultivation of CHO DP-12 cells 200 nM Methotrexate (MTX; Wyeth, Madison, New Jersey) and 0.1 mg/L Insulin-like growth factor (Long R3 IGF-I; Sigma-Aldrich, Munich, Germany) were added to the media. Cells were cultivated in 250 mL shaker flasks (Corning Inc., Corning, New York) with a culture volume of 70 mL and a seeding cell density of 3×10^5 cells/mL. Culture flasks were kept at 37 °C, 80 % humidity and 5 % CO₂ on a shaking platform at 185 rpm (50 mm diameter rotation). Batch cultivation of CHO-XL99 cells was performed at the Australian Institute for Bioengineering and Nanotechnology.

3.1.3. Cryopreservation and thawing

To preserve CHO cells, 1×10^7 cells from the exponential growth phase were centrifuged for 5 min at 800 rpm (Megafuge 1.0, Thermo Fisher Scientific) and resuspended in ice cold freezing medium (10 % DMSO, 90 % *CHOMACS*[®] *CD*, 6 mM glutamine). 1.5 mL resuspended cells were transferred to cryo vials (Nalgene Cryoware Cryogenic Vials, ThermoFisher Scientific) that were placed into freezing containers at -80 °C (*Mr. Frosty*[™] *Freezing Container*, Thermo Scientific) to achieve a rate of cooling of -1 °C/min. 24 h later the tubes were transferred to the nitrogen tank (*Chronos Biosafe 150*). After thawing cryopreserved cells were washed with ice cold PBS (137 mmol/L NaCl, 2.7 mmol/L KCl, 10 mmol/L Na₂HPO₄, 1.8 mmol/L KH₂PO₄), centrifuged at $200 \times g$ and resuspended in pre-warmed culture medium.

3.1.4. Measurement of cell count and viability

Measurements of cell densities and viabilities were performed on a *Cedex AS20* cellcounter (Innovatis AG, Bielefeld, Germany) in triplicates. Viabilities were determined by utilization of the Trypan Blue exclusion method. The required sample volume was 300 μ L. Samples were diluted 1:2 with phosphate buffered saline (PBS) buffer (155.2 mM NaCl, 1.5 mM KH_2PO_4 , 5.1 mM Na_2HPO_4 , pH 7.2).

3.1.5. Measurement of glucose and lactate concentrations

Glucose and lactate concentrations were measured using the *YSI 2700 Select Analyzer*, which employs immobilized enzymes (Glucose Oxidase and L-Lactate Oxidase) to oxidate glucose and lactate, respectively, and produce hydrogen peroxide that generates a probe signal current (Yellow Springs Instruments Inc., Yellow Springs, Ohio, USA). Cell suspensions were centrifuged for 5 min at $200 \times g$ and the culture supernatants were diluted 1:5 with PBS prior to the measurement.

3.1.6. Measurement of IgG concentrations

IgG concentrations were measured with a HPLC method on protein A columns according to the manufacturer's protocol (Poros A, Applied Biosystems, Foster City, California). Cell suspensions were centrifuged for 5 min at $200 \times g$. Subsequently, 120 μ L supernatant were centrifuged for 5 min at $16,000 \times g$ through a 0.2 μ m filter (VWR, Radnor, Pennsylvania, USA). 100 μ L were used for each measurement. Quantification of peak areas to determine antibody concentrations was performed in comparison to internal standards.

3.1.7. Transient transfection of CHO cell cultures

For RNA interference experiments shRNA-expression plasmids were delivered to CHO DP-12 cell cultures by transient transfection. Also, the strength of endogenous CHO promoters was analyzed in transiently transfected CHO DP-12 and K1 cultures. Transfection of promoter-reporter constructs was performed on a *Nucleofector[®] 2b* (Lonza, Basel, Switzerland) using the *Amaxa[®] Cell Line Nucleofector[®] Kit L* (Lonza, Basel, Switzerland). For each transfection reaction 1×10^6 cells were harvested by centrifugation at $200 \times g$ for 5 min and resuspended in 100 μ L *Nucleofector[®] solution* (supplied supplement added at a ratio of 4.5:1) that was then combined with 6 μ g of plasmid DNA. The mixture was transferred to a *Nucleofector[®]* cuvette and *Nucleofector[®]* program X-001 was applied. Immediately 500 μ L pre-warmed culture medium were added and the cells were transferred into a 12-well plate containing

1.5 mL pre-warmed medium. The transfected cells were incubated for 24 h at 125 rpm, 37 °C and 5 % CO₂ before they were harvested.

3.2. Microbiological methods

3.2.1. Bacterial culture

Microbiological experiments were performed using *OneShot*[®] *TOP10 Chemically Competent E. coli* cells (Invitrogen, Carlsbad, California, USA. Genotype: F- *mcrA* Δ (*mrr-hsdRMS-mcrBC*) Φ 80*lacZ* Δ M15 Δ *lacX74 recA1 araD139* Δ (*araleu*)7697 *galU galK rpsL* (StrR) *endA1 nupG*). Bacteria were cultivated in LB medium (5 g/L yeast extract, 10 g/L tryptone, 10 g/L NaCl; autoclaved) or on LB agar plates (LB medium, 12 g/L agar; autoclaved) at 37 °C on a benchtop shaker at 200 rpm under an incubation hood (*Certomat*[®] *MO II* and *Certomat H*, Sartorius, Göttingen, Germany) and in an incubator (*Function Line*, Heraeus Instruments, Hanau, Germany), respectively. Transformants were selected by antibiotics resistance. All plasmids used in this work contain an ampicillin resistance gene and ampicillin was added to the cultivation media in a concentration of 100 μ g/L.

3.2.2. Preparation of competent *E. coli* cells using calcium chloride

Chemically competent *E. coli* cells were generated using calcium chloride according to Sambrook and Russell (Sambrook and Russell 2006). 50 μ L of an *E. coli* overnight preculture were added to 100 mL LB medium. The culture was incubated at 37 °C and 200 rpm while its growth was monitored by measuring the optical density at 600 nm (OD₆₀₀) every 15-20 min. When the OD₆₀₀ reached 0.35 the culture was harvested by cooling it for 10 min on ice in 50 mL tubes (Corning Inc., Corning, New York, USA) and a following centrifugation at 2,700 \times g for 10 min at 4 °C (Rotina 420R, Hettich Lab Technology, Tuttlingen, Germany). The pellet was resuspended in 30 mL ice-cold MgCl₂-CaCl₂ solution (80 mM MgCl₂, 20 mM CaCl₂; sterile filtered) followed by centrifugation for 10 min at 4 °C. Subsequently, the pellet was again resuspended, this time in 2 mL ice-cold 0.1 M CaCl₂, and dispensed into aliquots of 50 μ L that were either directly used for transformation or stored at -80 °C.

3.2.3. Transformation of competent *E. coli* cells

For transformation DNA was added to the competent cells and the mixture was stored on ice for 30 min followed by a heat shock at 42 °C for 1 min. After the heat shock the cells were rapidly transferred to ice and 800 μ L SOC medium (20 g/L tryptone, 5 g/L yeast extract, 10

mM NaCl, 2,5 mM KCl, 10 mM MgSO₂, 10 mM MgCl₂, 20 mM glucose; autoclaved) were added. The cells were incubated at 37 °C on a shaking platform (*ThermoMixer*[®], Eppendorf, Hamburg, Germany) for 45 min. Cells were then collected by centrifugation for 2 min at room temperature in a benchtop centrifuge (*MiniSpin*, Eppendorf), resuspended in SOC medium and plated on LB agar.

3.2.4. Plasmid preparation

Preparation of up to 20 µg plasmid DNA from culture volumes of 5 mL was performed using the *GeneJet Plasmid Miniprep Kit* (ThermoFisher Scientific Inc., Waltham, Massachusetts, USA) according to the manufacturer's instructions. For minipreparations, *E. coli* clones from selection LB agar plates were inoculated into selective LB medium and grown overnight at 37 °C and 200 rpm. Large scale isolation of plasmid DNA (up to 200 µg) was performed using the *GeneJet Plasmid Midiprep Kit* (ThermoFisher Scientific Inc.) according to the manufacturer's instructions. For midipreparations, a 5 mL preculture of the respective clone was prepared and incubated for 8 h at 37 °C and 200 rpm. The pre-culture was then diluted 1:10 and incubated for 12-16 h.

3.3. Molecular biology methods

3.3.1. Preparation and analysis of DNA

For DNA preparation by phenol chloroform or columns CHO cells were washed twice with PBS and the cell pellets were stored at -80 °C. Phenol chloroform extraction was used for gene specific experiments and CpG island microarrays. For Illumina sequencing and generation of the CHO cell methylome, DNA was isolated from washed cell pellets using the *Qiagen Genomic-tip 100/G system* (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

3.3.1.1. Isolation of DNA from CHO cells using Proteinase K and phenol

Cell pellets were resuspended in 600 µL Tris ethylenediaminetetraacetic acid (EDTA) buffer (TE buffer; 10 mM Tris, 1 mM EDTA, pH 8.0). 2 mL lysis buffer (0.5 % SDS in TE buffer, 100 µg/mL Proteinase K, 20 µg/mL RNase A) were added. The sample was carefully mixed by pipetting and incubated at 37 °C for 1 h, followed by an incubation at 50 °C for 4 h. Subsequently, 1 volume of phenol chloroform isoamyl alcohol (PCI, ratio 25:24:1, pH 7.5-8.0) was added and the sample was mixed for 10 min on a rotating mixer (*VWR*[®] *Tube Rotator*,

VWR, Radnor, Pennsylvania, USA). Phase separation was performed by centrifugation at maximum speed for 10 min (*Rotina 420R*, Hettich Lab Technology, Tuttlingen, Germany). The upper aqueous phase was carefully removed and PCI extraction was repeated until no white precipitate was visible within the interphase any more. The aqueous phase was then extracted with one volume of chloroform isoamyl alcohol (CI, ratio 24:1) followed by centrifugation for 1 min. To precipitate the DNA from the aqueous phase 0.5 volumes NH_4OAc (4.4 mM) and 2.5 volumes of 100 % ethanol were added and the sample was placed into the freezer ($-20\text{ }^\circ\text{C}$) overnight. The precipitate was collected by centrifugation at maximum speed and $4\text{ }^\circ\text{C}$ for 30 min. The supernatant was removed, the pellet was washed with 70 % ethanol and resuspended in 1 mL TE buffer. To achieve complete dissolution of the DNA, the tube was placed on a rocking platform (*ThermoMixer*[®], Eppendorf, Hamburg, Germany) for 2 h at $50\text{ }^\circ\text{C}$. DNA integrity was checked on 0.6 % agarose gels.

3.3.1.2. Agarose gel electrophoresis of DNA

Agarose gels (*SeaKem LE-Agarose*, Lonza, Switzerland) were prepared at concentrations of 0.6 % (w/v), 1 % (w/v) and 2 % (w/v) for separation of high molecular weight DNA, DNA fragments and plasmids ≥ 500 bp, and DNA fragments ≤ 500 bp. Gels were cast and run using Tris acetate EDTA (TAE) electrophoresis buffer (40 mM tris acetate, 1 mM EDTA). DNA samples were mixed with 6-fold gel-loading buffer (0.25 % (w/v) bromophenol blue, 0.25 % (w/v) xylene cyanol FF, 30 % (v/v) glycerol in water). A voltage of 5 V/cm was applied until the loading dye reached the desired position within the gel. Gels were stained with *GelRed* (Biotium, Hayward, California, USA) for 10 min on a shaking platform and detection of fluorescence signals was performed using a *Fusion FX-7* system (Vilber-Lourmat, Eberhardzell, Germany). DNA fragment size was determined by comparison with DNA ladders (Fig. 3.1).

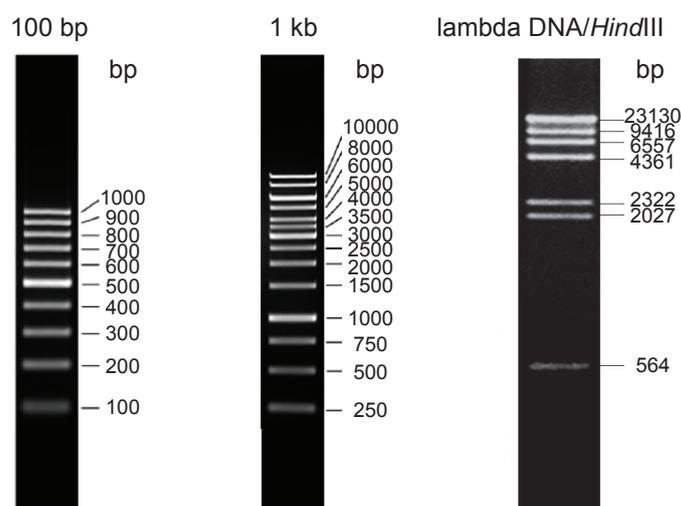


Figure 3.1.: DNA ladders (Thermo Scientific) used for determination of DNA fragment sizes.

3.3.1.3. DNA extraction from agarose gels

DNA was extracted from agarose gels using the *GeneJet Gel Extraction Kit* (ThermoFisher Scientific, Waltham, Massachusetts, USA) according to the manufacturer's instructions. DNA was eluted in varying volumes of elution buffer depending on the subsequent application.

3.3.1.4. Analysis of DNA quantity and quality

For quantitative PCR, gene specific DNA methylation analyses and CpG island microarray experiments, DNA quantity and quality (OD_{260}/OD_{280}) were determined using a *NanoPhotometer* with a *LabelGuard™ Microliter Cell* (Implen, München, Germany). For Illumina sequencing of the CHO cell methylome, DNA concentrations were determined using a *Quant-iT PicoGreen® dsDNA Assay Kit* (Invitrogen, Carlsbad, USA) on a *Microplate Reader Tecan Infinite 200* (Tecan, Männedorf, Switzerland). Sequencing libraries were quantified with a *High-Sensitivity Chip* on a *Bioanalyzer* (Agilent, Böblingen, Germany).

3.3.1.5. Sanger sequencing

To analyze site-specific DNA methylation in single-base resolution, PCR products amplified from bisulfite-treated DNA were cloned into *pJet1.2/blunt* cloning vector in order to enable sequencing of individual molecules (see section 3.4). Also, constructs for siRNA-mediated knockdown of *Dnmt3a* (section 3.3.5.3; vector *pLVX-shRNA1*) and for the analysis of the strength of endogenous CHO promoters (section 3.3.5.3; vector *pMCS-Green Renilla Luc*) were controlled by Sanger sequencing in regard to the potential presence of mutations and the completeness of the cloned oligonucleotides. DNA sequences were determined by the *MPIZ DNA core facility* on *Applied Biosystems Abi Prism 377, 3100 and 3730* sequencers (Weiterstadt, Germany) using *BigDye-terminator v3.1* chemistry. Premixed reagents were also obtained from Applied Biosystems. Per sequencing reaction 10 µL plasmid DNA in elution buffer at a concentration of 250 ng/µL and 1 µL primer at a concentration of 10 pmol/µL were required. Table 3.1 shows the primers used for sequencing. For sequencing reactions with *pJet1.2/blunt* as a template, only the reverse primer was used, as no reliable sequencing results were obtained using a forward primer.

3.3.2. Polymerase chain reaction (PCR)

Polymerase chain reactions (PCRs) were performed to amplify regions of interest from bisulfite treated DNA or to amplify endogenous CHO promoter sequences from genomic CHO K1 DNA. PCR products amplified from bisulfite treated DNA were either analyzed by

Table 3.1.: Primers used for sequencing of DNA fragments inserted into *pJet1.2/blunt* cloning vector, pMCS-Green *Renilla* Luc and pLVX-shRNA1.

Vector	Primer	Sequence (5' → 3')
pJet1.2/blunt	pJet rev	AAGAACATCGATTTTCCATGGCAG
pMCS-Green <i>Renilla</i> Luc	pMCS_LUC_for	GGATAACCGTATTACCGCCATG
	pMCS_LUC_rev	CTGTCCAGCACGTTTCATCTG
pLVX-shRNA1	pLVX-shRNA1-fwd	TTTCTTGGGTAGTTTGCAGTTTT
	pLVX-shRNA1-rev	GCCAGAGGCCACTTGTGTAG

Combined Bisulfite Restriction Analysis (COBRA, see 3.4.2) or subcloned into *pJet1.2/blunt* cloning vector for sequencing (see 3.3.1.5). Endogenous CHO promoter sequences were used to generate reporter-promoter constructs for dual luciferase assays (see 3.5.2).

3.3.2.1. Amplification from bisulfite-treated DNA

An *EpiMark*[®] *Hot Start Taq DNA Polymerase* (NEB) was used to amplify regions of interest from bisulfite treated DNA. Reactions were set up in a volume of 25 μ L according to the manufacturer's recommendations. Generally, 2 μ L purified, bisulfite treated DNA were used as a template. Table 3.2 shows the cycling program used for PCRs with *EpiMark*[®] *Hot Start Taq DNA Polymerase*. Table 3.3 lists the primers that were used.

Table 3.2.: Cycling program used for amplification of regions of interest from bisulfite treated genomic DNA with *EpiMark*[®] *Hot Start Taq DNA Polymerase*. Annealing temperatures specific for each pair of primers are listed in Table 3.3.

Time	Temperature	Cycles	Comment
30 sec	95 °C	1	Initial denaturation
15 sec	95 °C	35-40	Denaturation
30 sec	variable		Annealing
1 min/kb	68 °C		Elongation
5 min	68 °C	1	Final extension
∞	12 °C	1	Hold

3.3.2.2. Amplification from native DNA

Genomic regions that were analyzed regarding their potential to initiate the expression of recombinant genes by dual luciferase reporter assays were amplified from CHO K1 DNA using a *Q5*[®] *High-Fidelity DNA Polymerase* (NEB) that exhibits a more than 100-fold lower error rate than *Taq* polymerases. Reactions were set up in a volume of 25 μ L according to the manufacturer's recommendations without 5 \times *Q5*[®] *GC enhancer*. About 50 ng of genomic DNA were used as a template. Table 3.4 shows the cycling program used for PCRs with *Q5*[®] *High-Fidelity DNA Polymerase*. Table 3.5 lists the primers that were used.

Table 3.3.: Primers used for amplification of regions of interest from bisulfite treated CHO DNA. T_A = annealing temperature.

Name	Sequence (5' → 3')	Gene	Amplificate [bp]	T_A [°C]
pBca1 for	GGGTGTTAGGAATTAATTTTAAATT	<i>Bcat1</i>	373	57
pBca1 rev	CACAATAACTTTCTCTAAAACCTCCC			
p β Actin for	GGTAAGTAGGGATAATAGGTTTAGT	<i>Actb</i>	318	57
p β Actin rev	CCCCAAAATAAACAAATAC			
dvl1_for_BS	TTTTAGTTAGGAGTAAAAGTTGTATAATG	<i>Dvl1</i>	594	56
dvl1_rev_BS	TAAATCAAATCTCAAATTAATAACATCC			
frizzled7_for_BS for	GTAGGGTATTAAGAAGGAGGGTTGTAT	<i>Fzd7</i>	723	56
frizzled7_rev_BS	ATAACTACAAAATCTTACCAAACCAAATC			
myc_for_BS	ATTTGGAAGGATAGTAAAGTATATTGGAAG	<i>Myc</i>	541	56
myc_rev_BS	AAATAAACTCTAACTCACCATATCTCCT			
dnmt3a_for	TTTAGATTAAGAGTGGGAAAATATTGA	<i>Dnmt3a</i>	407	53
dnmt3a_rev	TTTTTCTCCTTAAAACCTCCAAAATAAC			

Table 3.4.: Cycling program used for amplification of promoter regions from CHO genomic DNA with *Q5[®] High-Fidelity DNA Polymerase*. Annealing temperatures specific for each pair of primers are listed in Table 3.5.

Time	Temperature	Cycles	Comment
30 sec	98 °C	1	Initial denaturation
10 sec	98 °C	25-35	Denaturation
15 sec	variable		Annealing
20-30 sec/kb	72 °C		Elongation
2 min	72 °C	1	Final extension
∞	12 °C	1	Hold

3.3.2.3. Quantitative Real-Time PCR

Quantitative Real-Time PCR (qPCR) experiments for analysis of relative gene expression or quality control of methylated DNA enrichment were performed on a *LightCycler[®] 480* (Roche, Basel, Switzerland) using microtiterplates (*White Multiwell Plate 96*, Roche) and *Sealing Foil* (Roche). For all experiments the *Platinum[®] SYBR[®] Green qPCR SuperMix-UDG Kit* (ThermoFisher Scientific) was used. Per reaction 15 μ L *Platinum[®] SYBR[®] Green qPCR SuperMix-UDG*, 0.6 μ L of both the forward and reverse primer (10 μ M), 12.8 μ L dH₂O and 1 μ L sample DNA were used. Sample analysis was performed in triplicates. Table 3.6 shows the cycling conditions used for qPCR. Table 3.7 lists the primers that were used.

Table 3.5.: Primers used for amplification of endogenous CHO promoter regions from CHO K1 genomic DNA. *Eco*RI and *Bam*HI restriction endonuclease recognition sites are shown in lowercase. T_A = annealing temperature. NA = not assigned to any annotated gene.

Name	Sequence (5' → 3')	Gene	Amplificate [bp]	T _A [°C]
P2_Eco_for	GGACgaattcTTGGGGTGTGAAGAAAAGCC	NA	529	56
P2_Bam_rev	GGACggatccCCACAAGCCAGGGATGGAT			
P3_Eco_for	GGACgaattcATACAGTTCCCCATGCTGTG	<i>Hist1h4n</i>	530	56
P3_Bam_rev	GGACggatccTTTTCCAAGACCCTTACCGC			
P5_Eco_for	GGACgaattcATTCTCCAGCCTCCACCTCT	<i>Rps3</i>	339	56
P5_Bam_rev	GGACggatccGTCCGAAATGCTTACCTTCCT			

Table 3.6.: Cycling program for quantitative Real-Time PCR.

Time	Temperature	Cycles	Comment
2 min	50 °C	1	UDG incubation
2 min	95 °C	1	Initial denaturation
15 sec	95 °C	40	Denaturation
30 sec	60 °C		Annealing and elongation
0.11 °C/sec	45 °C to 95 °C	1	Melting curve
∞	40 °C	1	Hold

3.3.3. Preparation and analysis of RNA

Total RNA was extracted for reverse transcriptase quantitative PCR (RT-qPCR) experiments and for gene expression microarray analyses. RNA samples were permanently kept on ice.

3.3.3.1. RNA extraction

RNA extraction was performed using the *Direct-zol*TM *RNA MiniPrep Kit* (Zymo Research, Irvine, California, USA) according to the manufacturer's recommendations. RNA concentrations were measured on a *NanoPhotometer* (Implen) that also enabled determination of RNA purity by calculation of OD₂₆₀/OD₂₈₀ ratios. RNA samples were stored at -80 °C in aliquots of 10 µL.

3.3.3.2. Quality control on denaturing agarose gels

Labeling of cDNA for gene expression microarray experiments requires RNA of high integrity. Therefore, RNA integrity was checked on denaturing agarose gels prior to microarray analyses. 1 % agarose gels were prepared using 1 × MOPS buffer (3-(N-morpholino)propanesulfonic acid, 200 mM) with 6.5 % (v/v) formaldehyde (all work was performed under a fume hood). Gels were pre-run in 1 × MOPS buffer for 20 min at 5 V/cm to reach a temperature beneficial for denaturation of the RNA molecules. 0.5 µg of total RNA were mixed with 6 µL 2 × RNA

Table 3.7.: Primers used for quantitative Real-Time PCR. gDNA = genomic DNA, cDNA = complementary DNA.

Name	Sequence (5' → 3')	Gene	Amplificate [bp]	Template
Bcat1-MIRA-for	GCAGGGACGCTGTTTGGCCT	<i>Bcat1</i>	97	gDNA
Bcat1-MIRA-rev	GGCTTTCAGGGCTCTGCGT			
Actb-MIRA-for	CCGCGGAGCGGACACTTTCA	<i>Actb</i>	100	gDNA
Actb-MIRA-rev	AGCGGGTCCACCGGTGTCTA			
rtPCR_Dnmt3a_for	TGATGAACGCACAAGAGAGC	<i>Dnmt3a</i>	97	cDNA
rtPCR_Dnmt3a_rev	GGTGACATTGAGGCTTCCAC			
rtPCR_Hdac1_for	TTCGGGATATTGGGGCTGG	<i>Hdac1</i>	102	cDNA
rtPCR_Hdac1_rev	TGACTGGCTTGAAAATGGCT			
Aktin-fwd	CACCCTGTGCTGCTCACC	<i>Actb</i>	100	cDNA
Aktin-rev_rev	CGTACATGGCTGGGGTGT			

loading buffer (ThermoFisher Scientific) and heated up for 10 min at 65 °C. Sample were then chilled and loaded onto the gel that was run for about 70 min at 5 V/cm. RNA sample integrity was evaluated regarding the presence of clear 28S and 18S rRNA bands, with the 28S rRNA band supposed to be about twice as intense as the 18S rRNA band.

3.3.3.3. Reverse transcription for generation of cDNA

Synthesis of first strand cDNA from RNA templates was performed using the *RevertAid First Strand cDNA Synthesis Kit* (ThermoFisher Scientific) according to the manufacturer's instructions with 1 µg of template RNA.

3.3.4. Gene expression profiling by microarrays

To analyze the CHO DP-12 cell transcriptome in response to butyrate treatment, gene expression microarray analyses were performed using a CHO cell-specific microarray in a one-color labeling approach (Becker et al. 2014). The microarrays were printed in a 4 × format (4 arrays per slide), which required defined buffer volumes and RNA amounts for hybridization. The volumes and concentrations indicated in the following refer to this format and would need to be adjusted for other array formats. Sample labeling and microarray hybridization was performed according to the protocol *One-Colour Microarray-Based Gene Expression Analysis - Low Input Quick Amp Labeling* (Agilent Technologies). The procedure is summarized in the following sections.

3.3.4.1. Sample preparation

To include positive control transcripts into the gene expression microarray analysis the *Agilent One Colour RNA Spike-In Kit* was used according to the manufacturer's recommendations.

The *RNA Spike Mix Stock Solution* was heated at 37 °C for 5 min and diluted for the total RNA input amount of 50 ng. cRNA generation and labeling reactions were performed according to the protocol and samples were stored at -80 °C overnight. Subsequently, the amplified and labeled cRNA samples were purified using the *RNeasy Mini Kit* (Qiagen) according to the manufacturer's recommendations and eluted in 30 µL RNase-free water. The *NanoPhotometer* was used to determine the yield and specific activity of each reaction by measuring the Cy3 dye concentration (pmol/µL), the RNA absorbance ratio (260/280 nm) and the cRNA concentration (ng/µL). Equations 3.1 (yield) and 3.2 (dye activity) were used for the calculations.

$$\mu\text{g of cRNA} = \frac{\text{Concentration of cRNA} \times 30 \mu\text{L}}{1000} \quad (3.1)$$

$$\text{pmol Cy3 per } \mu\text{g cRNA} = \frac{\text{Concentration of Cy3}}{\text{Concentration of cRNA}} \quad (3.2)$$

The recommended yield for 4 × microarrays is 1.65 µg of cRNA. The recommended specific activity is ≥ 6 pmol Cy3 per µg cRNA.

3.3.4.2. Microarray hybridization and wash

1.65 µg of cRNA were used for hybridization of each microarray according to the manufacturer's recommendations. Slides were hybridized at 10 rpm and 65 °C for 17 h in a *Microarray Hybridization Oven* (Agilent Technologies). The subsequent washing procedure was performed according to the protocol. Slides were put into the slide holders for scanning without an *Agilent Ozone Barrier Slide*, as ozone levels were considered to not exceed 50 ppb.

3.3.4.3. Scanning and feature extraction

After washing, microarray slides were scanned with a *High-Resolution Microarray C Scanner* (Agilent Technologies) using the following settings: Green dye channel, scan resolution 5 µm, Tiff file dynamic range 20 bit, green PMT gain 100 %. Data extraction from scan data was then performed using the *Agilent Feature Extraction Software*. The grid template 'one-color' was assigned to the extraction as well as the protocol 'GE1_107_Sep09'.

3.3.5. Molecular cloning

3.3.5.1. Subcloning of PCR products amplified from bisulfite-treated DNA

To enable sequencing of amplicates from bisulfite treated DNA, PCR products were cloned into *pJet1.2/blunt* cloning vector (see vector map in Fig. 3.2) using the *CloneJet PCR Cloning Kit* (Thermo Scientific) according to the manufacturer's recommendations. As amplification was performed using the *EpiMark®Hot Start Taq DNA Polymerase* (NEB) which produces 3'-dA overhangs, PCR products were blunted prior to ligation.

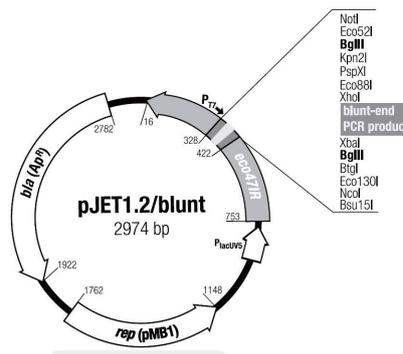


Figure 3.2.: Vector map of *pJet1.2/blunt* cloning vector (Thermo Scientific).

3.3.5.2. Generation of promoter-reporter constructs for dual luciferase assays

Promoter-reporter constructs for dual luciferase assays were generated by cloning of 5 selected endogenous CHO promoter sequences that were provided by Tobias Jakobi (Jakobi et al. 2014) into *pMCS-Green Renilla Luc* cloning vector (see vector map in Fig. 3.3 A) via the *EcoRI* and *BamHI* restriction endonuclease recognition sites contained in its multiple cloning site (MCS).

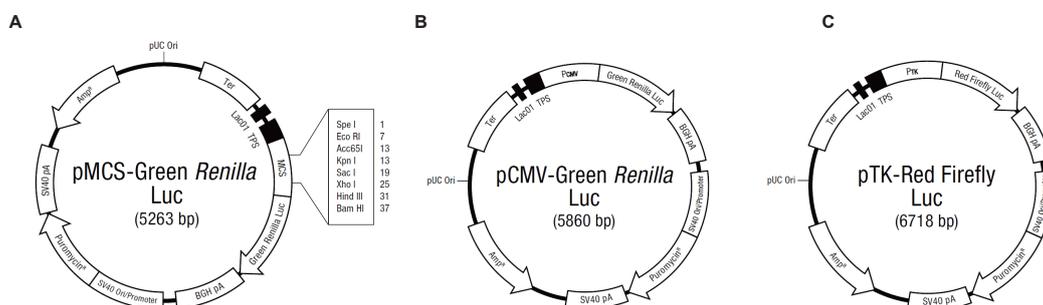


Figure 3.3.: Vector maps of *pMCS-Green Renilla Luc*, *pCMV-Green Renilla Luc* and *pTK-Red Firefly Luc* (ThermoFisher Scientific).

Two promoters (promoters 1 and 5) were synthesized as complementary oligonucleotides by Metabion (Tab. 3.8; Steinkirchen, Germany) and three promoters (promoters 2 to 4) were

amplified from CHO K1 genomic DNA. Primers used for amplification of the promoters 2 to 4 are listed in Tab. 3.5. Restriction endonuclease sites were added to the ordered oligonucleotides and primers to enable cloning into *pMCS-Green Renilla Luc* vector. Annealing of 100 pmol per oligonucleotide was performed in annealing buffer (100 mM potassium acetate, 30 mM HEPES-KOH, 2 mM magnesium acetate, pH 7.4) using a *Thermocycler* (Eppendorf) with the oligonucleotide annealing program shown in Tab. 3.9. PCRs were performed according to section 3.3.2.2.

Table 3.8.: Oligonucleotide sequences used for generation of promoter-reporter constructs by direct cloning into *pMCS-Green Renilla Luc* vector. *EcoRI* and *BamHI* restriction endonuclease recognition sites are shown in lowercase.

No.	Gene	Strand	Sequence (5' → 3')
1	NA	forward	GGACgaattcGCGGACTTCAGGAGAGGAAGCGGAAATGCGTGGCCGCTTTAAA- GGCGTCGGCCCCGCCCGCGCCCTCTTTCCGTGTCCGGCCGCCGTGGGGAC- AGGAGGTGAGTGC GGCGTGC GAGCGGTCCGGGggatccGGAC
		reverse	GTCCggatccCCCCGACCGCTCGCACGCCCGACTCACCTCCTGTCCCCACGG- CGGCCGGACACGGAAGAGGGGCGCCGGGGCGGGCCGACGCCTTTAAAG- CGGCCACGCATTTCCGCTTCTCTCTCTGAAGTCCGCgaattcGTCC
5	<i>Hist1h4n</i>	forward	GGACgaattcCCGCTCCTAGGGAATATAAGCTCTGGCTTGGGCGCACTGCATT- GTAGTTCTCTTTGTTTCTCTGCTGATTGTTTGTTCACCAggatccGGAC
		reverse	GTCCggatccTGGTGAACAAACAATCAGCAGAGAAACAAGAGAGAACTACA- ATGCAGTGCGCCAAGCCAGAGCTTATATCCCTAGGAGCGGgaattcGTCC

Table 3.9.: Program for annealing of oligonucleotides in a *Thermocycler*.

Time	Temperature
3 min	95 °C
2 min	72 °C
30 min	Ramp cool to 37 °C
2 min	25 °C
hold	4 °C

3.3.5.3. Cloning of shRNA oligonucleotides into *pLVX-shRNA1*

ShRNA oligonucleotides for RNAi-mediated knockdown of DNMT3A were cloned into *pLVX-shRNA1* (Clontech, Mountain View, California, USA) which contains a human Pol III-dependent U6 promoter upstream of the MCS for shRNA expression, an ampicillin resistance gene for cloning and a puromycin resistance gene for selection of stably transfected mammalian cells. *pLVX-shRNA1* also contains several elements for packaging into viral particles, which was not performed within this thesis (WPRE, cPPT/CTS and RRE in Fig. 3.4).

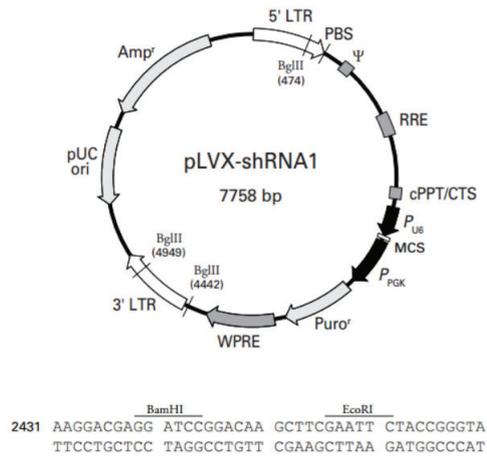


Figure 3.4.: Vector map of *pLVX-shRNA1* (Clontech).

3.3.6. ShRNA design for RNA interference experiments

For RNAi experiments siRNAs were designed using the *Clontech RNAi Sequence Selector* and the following criteria for selection of shRNAs with high efficacies: A maximal melting temperature of 45 °C, an optimal G+C content of 45 %, a higher melting temperature at the 5' end than at the 3' end, a maximal length of homopolymeric sequences of 4 bases, adenine bases at positions 3, 6, 18 or 19 and no guanine at position 13. A scrambled control sequence was generated using the *InVivoGen siRNA Wizard v3.1*. All siRNA sequences were mapped to the CHO genome sequence in *GenDBE* (Rupp et al. 2014) to avoid cross-hybridizations. Subsequently, the *Clontech shRNA Sequence Designer* was used to generate shRNA oligonucleotides that contained loop and terminator sequences as well as restriction enzyme sites (*Bam*HI and *Eco*RI) for cloning of annealed oligos into an expression vector. ShRNA oligonucleotides (Tab. 3.10) were synthesized by Metabion. Annealing of shRNA oligonucleotides was performed in annealing buffer in a *Thermocycler* (see annealing program in Tab. 3.9). Annealed oligos were cloned into *pLVX-shRNA1* (see section 3.3.5.3).

3.4. DNA methylation analysis

The following section describes the techniques that were used to analyze both site-specific and genome-wide DNA methylation in the CHO cell genome. Analysis of single-base methylation patterns of genomic regions with a maximum size of 600 bp was performed by bisulfite Sanger sequencing. Combined Bisulfite Restriction Analysis (COBRA; Xiong and Laird 1997) was applied to confirm bisulfite sequencing results. Genome wide DNA methylation analyses were performed using a CHO specific CpG island microarray and next generation sequencing.

Table 3.10.: shRNA sequences used for RNAi-mediated knockdown of DNMT3A. Indicated are exons of the DNMT3A gene targeted by the three siRNAs as well as the shRNA oligonucleotide sequences (uppercase) including added restriction enzyme sites (lowercase). sc = scrambled control.

siRNA	<i>Dnmt3a</i> - Exon	Strand	Sequence (5' → 3')
2	3	forward	gatccGGTGGAAAGTAGTGACACATTCAAGAGATGTGTCACTACTTTCCACCTTTTTTg
2	3	reverse	aattcAAAAAAGGTGGAAAGTAGTGACACATCTCTTGAATGTGTCACTACTTTCCACCg
3	8	forward	gatccGTCGAGAAACTCATGCCACTTTCAAGAGAAGTGGCATGAGTTTCTCGATTTTTTg
3	8	reverse	aattcAAAAAATCGAGAAACTCATGCCACTTCTTTGAAAGTGGCATGAGTTTCTCGACg
9	20	forward	gatccGTCTGGAACATGGCAGAATATTCAAGAGATATTCTGCCATGTTCCAGATTTTTTg
9	20	reverse	aattcAAAAAATCTGGAACATGGCAGAATATCTCTTGAATATTCTGCCATGTTCCAGACg
sc	-	forward	gatccGGGAGTACATCGTGAGAAATTCAAGAGATTTCTCACGATGTACTCCCtttttg
sc	-	reverse	aattcaaaaaGGGAGTACATCGTGAGAAATCTCTTGAATTTCTCACGATGTACTCCCg

3.4.1. Bisulfite Sanger sequencing of individual genomic loci

3.4.1.1. Primer design for bisulfite sequencing

As bisulfite treated DNA exhibits a low percentage of cytosine bases and many homopolymeric regions, primer design for PCRs with bisulfite DNA as a template requires consideration of certain parameters. Primers were designed with the online software *MethPrimer* (Li and Dahiya 2002) using the following criteria: In order to avoid discrimination between methylated and unmethylated DNA, primers should not contain any CpG dinucleotides. To prevent amplification of unconverted sequences, 3' ends were supposed to contain cytosines from CpH (H = A, T, C) contexts. Primers should furthermore be longer than 25 nt to ensure their specificity. Amplificates should not exceed 500 bp to 600 bp, as bisulfite treated DNA is single stranded and therefore sensitive to shearing forces (Warnecke et al. 2002; Zhang et al. 2009). Primers were synthesized by Metabion.

3.4.1.2. Bisulfite conversion of genomic DNA

Bisulfite conversion of genomic DNA was performed using the *Epimark Bisulfite Conversion Kit* (NEB, Ipswich, Massachusetts, USA) or the *EZ DNA Methylation-Lightning™ Kit* (ZymoResearch, Freiburg, Germany) according to the manufacturer's instructions and with a total amount of 100 ng of DNA per reaction. Bisulfite treated DNA samples were stored at -20° C in aliquots of 10 µL.

3.4.1.3. Sequencing of amplicates from bisulfite treated DNA

Amplification of genomic regions from bisulfite treated DNA was performed using an *Epimark Taq Polymerase* (NEB). PCR products were purified from 2 % agarose gels and cloned into

pJet1.2/blunt cloning vector. 10 μ L per ligation reaction were used for transformation of chemically competent *E. coli* cells. At least 5 clones were randomly picked from selection plates and plasmid DNA was isolated. Sequencing was performed according to 3.3.1.5. Sequence analysis was performed using the online tool *BiQ analyzer* (Bock et al. 2005). Fig. 3.5 shows an exemplary sequencing result of good quality (regarding the conversion rate of cytosines in non-CpG contexts and peak heights) in a comparison to the original sequence.

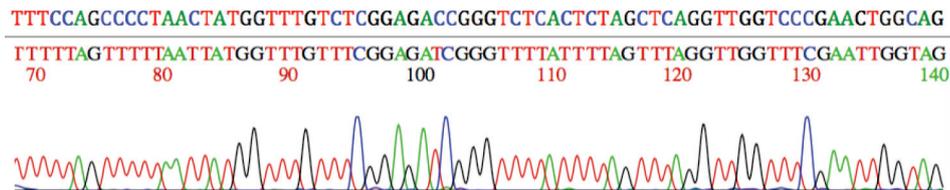


Figure 3.5.: Exemplary Sanger sequencing result. The region of interest was amplified from bisulfite treated DNA, subcloned into *pJet1.2/blunt* cloning vector and Sanger sequenced. Top: Original DNA sequence, bottom: sequence from bisulfite-treated DNA.

3.4.2. Combined Bisulfite Restriction Analysis (COBRA)

To confirm the results of bisulfite sequencing experiments, Combined Bisulfite Restriction Analysis (COBRA) was used (Xiong and Laird 1997). For this experiment purified amplicates from bisulfite treated DNA were digested with *TaqI* (recognition site 5'-TCGA; Promega, Madison, USA) and *BstUI* (recognition site 5'-CGCG; *FastDigest*[®], Thermo Scientific) according to the manufacturer's recommendations. The technique is based on the fact that the presence of restriction enzyme recognition sites depends on the methylation status of cytosine bases within this sequences - i.e., if the cytosine within the *TaqI* site was not methylated, the site reads 5'-TTGA-3' after bisulfite conversion and amplification. Therefore, *TaqI* would only cut fragments from DNA which carried a methyl group at the respective position (see schematic representation of the COBRA principle in Fig. 3.6). DNA fragments were separated on 2 % agarose gels. Undigested DNA fragments served as a control.

3.4.3. CpG island microarray analysis

The analysis of differential CpG island methylation was carried out using customized CHO specific oligonucleotide CpG island microarrays (Wippermann et al. 2013; Wippermann 2012). The process of microarray design included prediction of promoter-associated and intragenic CpG islands in the CHO genome based on the available genomic data from Xu et al. 2011 and transcriptomic CHO data (Becker et al. 2011), and was performed according to the algorithm of Takai and Jones (Takai and Jones 2002). A total of 43,318 CpG islands in the CHO genome were identified, 21,993 of which could be associated with promoter

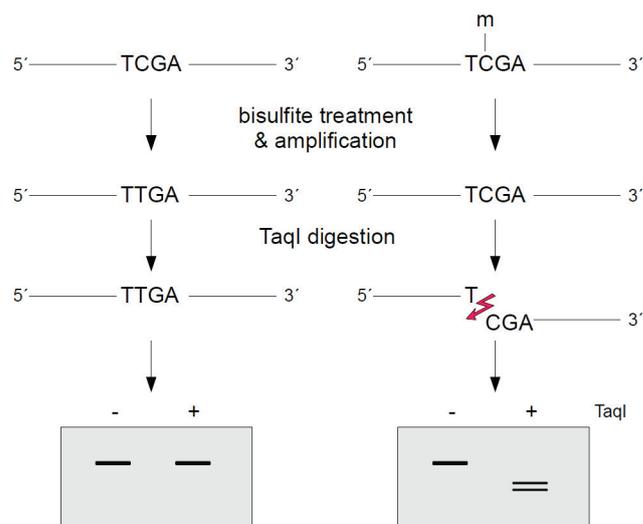


Figure 3.6.: Principle of Combined Bisulfite Restriction Analysis (COBRA).

and intragenic regions. 27,466 individual probes were designed with average distances of 500 bp, in regard to the average DNA fragment lengths after sonication. Customized arrays were printed by Agilent Technologies in an 8 x 60K (62,976 features) high density format. Therefore, all probes are present at least twice per array, so that experiments with two arrays (including a dye swap) result in 4 to 6 replicates per probe. The CHO-specific CpG island microarray contains the ControlGrid IS-62976-8-V2_8by60K_LA_Genomic_20080924 with the negative controls 'DarkCorner', 'Structural Negative' and 'Biological Negative' (indicated as *DarkCorner*, $3 \times SLv1$ and *NC1_00000002* in the .txt file after feature extraction). Negative controls are used to measure the background signal intensity. DarkCorner and Structural Negative controls form a hairpin that prevents hybridization. Biological Negative controls were shown not to hybridize to any DNA sample tested. The microarray was assigned the Agilent ID 040858.

3.4.3.1. Sonication of genomic DNA

Genomic DNA was sonicated to an average fragment size of 500 bp using a *Branson Sonifier 450-A* (Branson Ultrasonics, Danbury, Connecticut). Sonication was carried out 6 times in TE buffer for 30 s with 40 % amplitude while cooling the tubes in icewater. Subsequently, 300 ng DNA per sample were checked on 2 % agarose gels. Sonication was repeated depending on the persistent presence of high molecular DNA.

3.4.3.2. Immunoprecipitation of methylated DNA

35 µg of fragmented genomic DNA per sample were used for enrichment of methylated DNA by magnetic beads coupled MBD2-Fc protein (*Epimark Methylated DNA Enrichment Kit*,

NEB, Ipswich, Massachusetts). Magnetic beads were resuspended by inverted rotation for 30 min. Per 1 µg of sheared DNA 1 µL beads and 0.1 µL MBD protein were used. The required amount of beads was washed twice in $1 \times$ *bind/wash buffer* (3 minutes rotation followed by removal of buffer). Subsequently, 200 µL buffer and the required amount of MBD protein were added and the mixture was incubated by inverted rotation for 1 h. The beads were washed twice again, resuspended in $1 \times$ *bind/wash buffer* and added to the DNA that was diluted in $1 \times$ *bind/wash buffer*. Samples were incubated for 1 h. Eight washing steps were used to wash off unbound DNA. Methylated DNA was eluted by adding 100 µL dH₂O and incubation at 65 °C and 1000 rpm for 20 min on a ThermoMixer (Eppendorf). Removal of MBD protein present in the supernatant was performed using a *GeneJet PCR Purification Kit* (ThermoFisher Scientific). DNA was eluted in 13 µL dH₂O and stored at -80 °C.

3.4.3.3. Fluorescence labeling of enriched methylated DNA

Enriched DNA samples were labeled with Cyanine 3- and 5-dUTP using the *SureTag DNA Labeling Kit* (Agilent Technologies, Santa Clara, California). 2.5 µL random primer were added to 13 µL enriched methylated DNA and the samples were incubated for 5 min at 95 °C in a *Thermocycler* (Eppendorf). After cooling to 4 °C 9.5 µL reaction master mix (5 µL buffer, 2.5 µL $10 \times$ dNTPs, 1.5 µL Cy3-dUTP or Cy5-dUTP and 0.5 µL Exo-Klenow) were added per sample. Incubation was performed at 37 °C for 120 min. The labeling reaction was stopped by incubation at 65 °C for 10 min. For removal of unincorporated dye *Amicon 30 kDa spin filters* (Merck Millipore, Billerica, Massachusetts, USA) were used. Washing was done with TE buffer. After vacuum concentration the samples were eluted in 6.3 µL TE buffer. 1 µL per sample was used to measure the DNA and dye concentrations in order to calculate the yield of fluorescently labeled DNA and the specific dye activity (see equations 3.3 and 3.4). 9 µg to 12 µg of labeled DNA was considered to be optimal yields. Optimal dye activity values for Cy3 were 25–40 pmol/µg and for Cy5 20–25 pmol/µg. Samples were stored overnight at -20 °C and used for microarray hybridization on the next day.

$$\text{Specific activity [pmol dye/}\mu\text{g DNA]} = \frac{\text{pmol per } \mu\text{L of dye}}{\mu\text{g per } \mu\text{L DNA}} \quad (3.3)$$

$$\text{Yield } [\mu\text{g}] = \frac{\text{DNA concentration [ng/}\mu\text{L]} \times \text{Sample volume } [\mu\text{L}]}{1000\text{ng/}\mu\text{g}} \quad (3.4)$$

3.4.3.4. Hybridization

Subsequently the Cy3- and Cy5-labeled samples were combined and 29 μL hybridization mastermix (4.3 μL $10 \times$ blocking agent, 21.5 μL $2 \times$ hi-rpm buffer (Agilent Technologies, Santa Clara, California), 8 % formamide (Carl Roth, Karlsruhe, Germany), 1.2 μL *Mouse Cot-1 DNA* (Life Technologies, Darmstadt, Germany)) were added per tube. The mixture was incubated for 3 min at 95 °C and subsequently for 30 min at 60 °C. 40 μL per sample were used for hybridization to the microarrays by incubation at 67 °C with 20 rpm for 24 h.

3.4.3.5. Scanning and feature extraction

After washing and drying the slides by centrifugation, the arrays were scanned with a *High Resolution Microarray C Scanner* (Agilent Technologies, Santa Clara, California) using the following settings: Scanning profile 'Agilent G3_CGH', channel 'R+G', resolution 3 μm , 16 bit .tiff image. Data extraction from scan data was carried out using the *Feature Extraction Software* (Agilent Technologies). The grid template that was automatically loaded from Agilent's *Grid Template Browser* was assigned to the extraction as well as the protocol 'ChIP_107_Sep09'.

3.4.4. Whole-genome bisulfite sequencing

In order to analyze the CHO DNA methylation landscape and to display the butyrate effect on DNA methylation in single-base resolution, sequencing libraries were generated from bisulfite treated DNA and subjected to next generation Illumina sequencing. As the CHO DP-12 cell line as an exemplary production cell line was used for this experiment, unconverted CHO DP-12 DNA was also sequenced to account for SNPs present between the CHO DP-12 and the reference CHO K1 cell genomes.

3.4.4.1. Bisulfite conversion

50 ng of extracted DNA were spiked with 0.01 % (w/w) of *Unmethylated cl857 Sam7 Lambda DNA* (Promega, Madison, Wisconsin) that served as an unmethylated control and was subjected to bisulfite treatment (*ZymoResearch Methylation Lightning Kit*; ZymoResearch, Irvine, California). Bisulfite converted DNA was stored at -80 °C until library construction.

3.4.4.2. Library generation and quality control

Library generation from bisulfite treated DNA and 10 ng of unconverted genomic DNA was performed using the *EpiGnome Methyl-Seq Kit* (Epicentre, Madison, Wisconsin) according to the manufacturer's instructions. Briefly, bisulfite-treated DNA was randomly primed and DNA molecules with 5' tags were synthesized. Subsequently, the 3' ends of the newly synthesized molecules were tagged as well and dsDNA was amplified by PCR with barcoded primers to allow for multiplexed sequencing of libraries from bisulfite treated and native DNA. For this, *EpiGnome™ Index PCR Primers* (Epicentre) were used. The sequence of each Index PCR primer is 5'-CAAGCAGAAGACGGCATAACGAGAT N1-N6 GTGACTGGAGTTCA-GACGTGTGCTCTTCCGATCT. N1-N6 represents the reverse complement of the index sequence generated during sequencing (see Tab. 3.11). The quality of both libraries was checked on *High Sensitivity DNA Chips* on the *Bioanalyzer* (Agilent Technologies, Santa Clara, California).

Table 3.11.: Barcoding of individual libraries for DNA methylome analysis by next generation sequencing.

Sample	No. of Index PCR Primer	Index sequence
Genome	4	5'-TGACCA-3'
0 h	5	5'-ACAGTG-3'
6 h	2	5'-CGATGT-3'
12 h	10	5'-TAGCTT-3'
24 h	11	5'-GGCTAC-3'
36 h	8	5'-ACTTGA-3'

3.4.4.3. Illumina MiSeq sequencing

In order to estimate the behavior of sequencing libraries generated from bisulfite treated DNA, samples were sequenced on an *Illumina MiSeq desktop sequencer* (Illumina) in a 2×76 bp run according to the manufacturer's instructions. Briefly, sequencing libraries (i.e. 0 h to 36 h) were pooled and denatured. The pool was then diluted to a concentration of 8.4 pM for sequencing. *PhiX Control v3* (Illumina) was added in a concentration of 2.1 pM. The mixture was transferred into the reagent cartridge and sequencing was started after washing and loading the flow cell.

3.4.4.4. Illumina HiSeq 1500 sequencing

Paired-end sequencing was performed on an *Illumina HiSeq 1500* system in a 2×100 bp run with 1 % *PhiX Control v3* (Illumina) in each lane. One entire lane was used as a control to sequence the unconverted CHO DP-12 genomic DNA library and to allow for phasing and

pre-phasing. Lanes containing bisulfite libraries were additionally loaded with 10 % of the unconverted DNA library. Tab. 3.12 contains the theoretical compositions of the sequencing reactions.

Table 3.12.: Composition of Illumina *HiSeq* sequencing reactions for flow cells 1 and 2. Calculations were based on a theoretical maximum output of 280 Gb per flow cell. The theoretical coverages of both sequencing runs were 47-fold for the CHO DP-12 genome and 32-fold for each library. Clusters were generated in a final concentration of 10 pM.

Sample	Lane 1	Lane 2-8
Sequencing run 1		
<i>PhiX</i>	1 % (0.35 Gb)	1 % (0.35 Gb)
CHO DP-12 genome	99 % (34.65 Gb)	10 % (3.5 Gb)
Bisulfite libraries	–	89 % 0, 6, 12, 24 h (31.15 Gb)
Sequencing run 2		
<i>PhiX</i>	1 % (0.35 Gb)	1 % (0.35 Gb)
CHO DP-12 genome	99 % (34.65 Gb)	10 % (3.5 Gb)
Bisulfite libraries	–	12 % 0 h, 16 % 6, 12, 24 h, 40 % 36 h (31.15 Gb)

3.5. Protein biochemistry

Protein biochemical methods were used to determine the expression level of DNMT3A and HDAC1 in butyrate-treated CHO DP-12 cells. For this purpose westernblot analyses were used. Furthermore, dual luciferase assays were performed to analyze the gene expression potential of endogenous CHO regulatory elements in promoter-reportergene assays employing a firefly luciferase.

3.5.1. Westernblot analysis

3.5.1.1. Protein extraction

Cells were washed twice by centrifugation for 5 min at 1000 rpm in PBS buffer. Cell pellets were resuspended in 500 μ L ice cold lysis buffer (50 mM Tris-HCl pH 7.2, 150 mM NaCl, 2 mM EDTA, 0.1 % SDS, 1 % NP40, 1 mM phenylmethylsulfonyl fluoride (PMSF)) and mixed. Samples were incubated for 5 min on ice, followed by 5 min of sonication with 40 % amplitude at RT. After 30 min of incubation on ice, samples were centrifuged for 30 min at 13,000 rpm and 4 °C to remove cell debris from the supernatant that contained the extracted protein.

3.5.1.2. Protein quantification by Bradford assay

Protein quantifications were performed using the *Coo Protein Assay Kit* (Uptima-Interchim, Montlucon, France). 5 μL per sample or standard dilution (BSA) were combined with 250 μL *Coomassie reagent* (Uptima-Interchim) in 96-well plates. Plates were incubated for 1 min on a rotating platform at RT. Absorption was measured within the next 15 min at 570 nm on a plate reader (*PowerWave HT Reader*, BioTek Instruments Inc., USA). Protein concentrations were calculated based on the standard curve.

3.5.1.3. SDS polyacrylamide gelelectrophoresis (SDS-PAGE)

Gels were cast as discontinuous systems consisting of an SDS-polyacrylamide-stacking gel (4 %, pH 6.8) and an SDS-polyacrylamide-resolving gel (10 % for detection of both DNMT3A and HDAC1, pH 8.8). For one stacking gel 130 μL *Rotiphorese[®] Gel 30* (37.5:1, Carl Roth), 610 μL of MilliQ, 250 μL 1.3 \times gel buffer (pH 6.8, 0.5 M Tris-HCl, adjust pH with 6 M HCl) and 10 μL SDS were used. For one resolving gel of 10 % 1650 μL *Rotiphorese[®] Gel 30*, 2050 μL MilliQ, 1250 μL 4 \times gel buffer (pH 8.8, 1.5 M Tris-HCl, adjust pH with 6 M HCl) and 50 μL SDS were used. 5 μL 10 % APS and 1 μL TEMED per mL gel were added to induce polymerization. Resolving gels were overlaid with isopropanol before polymerization for 45 min to assure even edges. Stacking gels were also allowed to polymerize for 45 min. 4 \times SDS sample buffer (8 % SDS, 250 mM TRIS, 40 % glycerol, 0.04 % bromophenol blue) was added to 20 μg of total protein per sample to a final concentration of 1 \times and incubated for 10 min at 70 $^{\circ}\text{C}$ and 400 rpm. Samples were centrifuged and loaded onto the gel that was run at 175 volts in 1 \times SDS electrophoresis buffer (125 mM TRIS, 960 mM glycine, 0.5 % SDS) for 50 to 60 min. A *PageRuler[™] Prestained Protein Ladder* (Fig. 3.7; ThermoFisher Scientific) was used as a size standard.

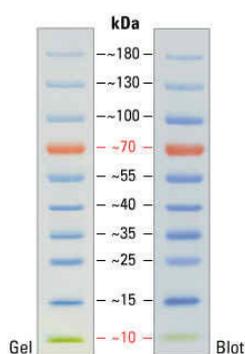


Figure 3.7.: *PageRuler[™] Prestained Protein Ladder* (ThermoFisher Scientific). kDa = kilodalton.

3.5.1.4. Westernblot analysis

Transfer of proteins from the SDS-polyacrylamide gel to a polyvinylidene fluoride (PVDF) membrane that was equilibrated in EtOH for 30 seconds was performed for 60 min at 35 volts in transfer buffer (20 times: 500 mM bicine, 500 mM bis-Tris, 20 mM EDTA. 1 times: Addition of 10 % EtOH and 1 μ L/mL *NuPAGE antioxidant*). The membrane was than incubated overnight at 4 °C in 10 mL blocking solution (PBS buffer, 5 % milk powder, 0.3 % Tween 20) on a shaking platform. After 2 washing steps incubation with the primary antibody (1:1000 for both DNMT3A and HDAC1) was performed for 1 h on a shaking platform. The primary antibody was removed by 3 washing steps (each 15 min; washing solution: PBS buffer, 1 % milk powder, 0.3 % *Tween 20*). Subsequently, incubation with the cyanine dye (Cy) coupled secondary antibody (1:300 for DNMT3A, which was detected with an Cy5-coupled secondary antibody, and 1:500 for HDAC1, which was detected with an Cy3-coupled secondary antibody) was performed for 1 h on a shaking platform followed by three washing steps (each 15 min). As the secondary antibody was Cy-conjugated, gels were scanned using an *Ettan DIGE Imager* with pixel size 100 μ m and variable exposure times (depending on signal intensities).

After scanning and data analysis, PVDF membranes were stained with *Coomassie dye* and scanned to enable quantification of the total amount of protein that has been transferred. This was used to normalize band intensities of the proteins that were analyzed.

3.5.2. Dual luciferase assay

In order to analyze the gene expression potential of endogenous CHO regulatory elements, 5 promoters were analyzed in promoter-reportergene assays employing a firefly luciferase. To enable correction of the results for variability in CHO cell transfection a *Renilla* luciferase was expressed as an internal control. Activities of both enzymes were simultaneously measured using the *Dual-Luciferase[®] Reporter (DLR[™]) Assay System* (Promega) according to the manufacturer's recommendations. Fig. 3.8 shows the bioluminescent reactions the assay system is based on. Central to the experiment is the fact that the amount of relative light units (RLU) produced by Green *Renilla Luciferase* depends on the expression strength of the endogenous CHO promoter, whereas the luminescence signal from the constitutively expressed *Tk* promoter does not change regardless of any intrinsic regulations, but depends on experimental conditions such as the transfection efficiency. For this reason, *Tk* promoter-dependent luciferase expression can be used for normalization in dual luciferase assays, which are capable of measuring both luciferase activities sequentially from a single sample. In the assay used here, *Firefly* luciferase activity was determined first by adding beetle luciferin and coenzyme A, which results in a stable luminescence signal. The signal generated in this way

is subsequently quenched and the substrate for *Renilla* Luciferase, coelenterate-luciferin, is added (Dual Luciferase[®] Reporter Assay System protocol, Promega). Luminescence was measured using a plate-reading luminometer (*FLUOstar Galaxy*; BMG LABTECH GmbH, Ortenberg, Germany).

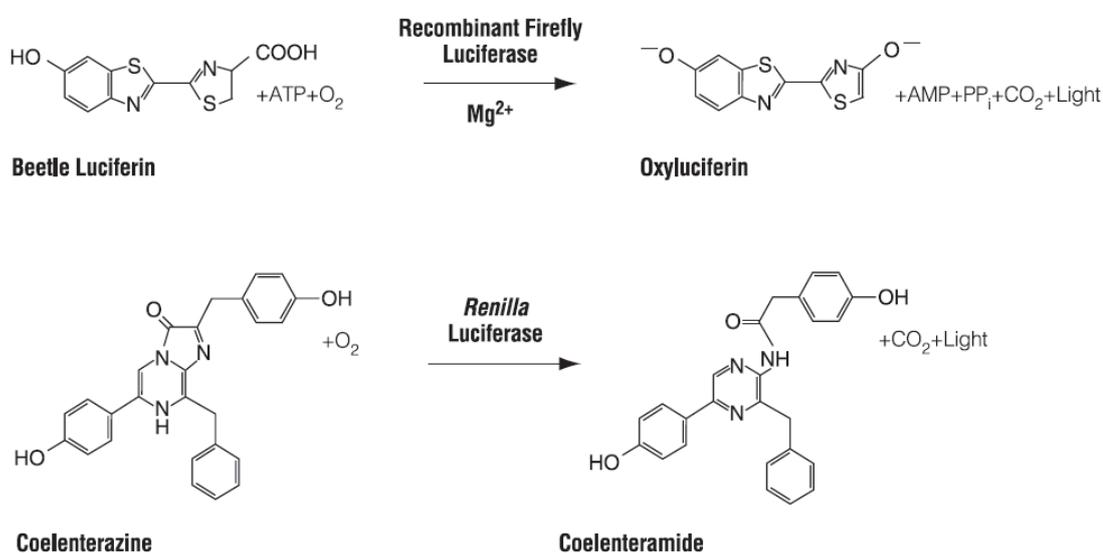


Figure 3.8.: Bioluminescent reactions employed for the Dual Luciferase[®] Reporter Assay System (Promega).

3.6. Data analysis

Data generated by bisulfite sequencing of individual genomic regions were analyzed using the *BiQ Analyzer* software (Bock et al. 2005). For evaluation of microarray data generated by CHO CpG island and CHO gene expression microarrays, analyses were performed using the Bioconductor *limma* package in the *R* programming environment (Ritchie et al. 2015). Further CpG island microarray data evaluation was performed using the *EMMA2* software platform (Dondrup et al. 2009). Analysis of data generated by next generation sequencing (NGS) was done in collaboration with Oliver Rupp from the Bioinformatics and Systems Biology group at Justus Liebig University Giessen.

3.6.1. Analysis of gene-specific DNA methylation data

DNA methylation data generated for genomic regions of interest by Sanger sequencing were analyzed using the *BiQ Analyzer* software tool (Bock et al. 2005). Sequence files and sequences of the genomic regions of interest were imported in FASTA format. The software aligns the *in silico* converted (as if fully unmethylated, i.e. complete C to T conversion)

genomic sequence with the sequencing reads and applies different quality control measures. First the conversion rate, which is defined as the number of correctly converted cytosines outside of CpG contexts divided by all cytosines outside of CpG contexts, is calculated. The software then suggests to exclude sequences that are likely to have originated from only a few individual chromosomes. In a third step putative sequencing errors are analyzed and *BiQ Analyzer* suggests to exclude sequences that exhibit a local sequence identity of less than 80 %. The software finally generates a HTML documentation that includes sequence alignments, statistics and methylation diagrams in 'lollipop' style.

3.6.2. Analysis of data generated by CHO CpG island microarray experiments

Data quality was explored using the Bioconductor *limma* package that enables visual inspection of experimental data to identify potential quality problems (Ritchie et al. 2015). Such problems might arise during sample processing (e.g. due to differences in labeling efficiencies) or microarray hybridization (e.g. due to minor differences in the composition of the hybridization mixtures). For quality assessment of two-color microarray data boxplots of red and green signals and red and green background signals were generated. To get an overview of potential spatial artifacts, intensity plots for all four types of signals were generated. Furthermore, density plots were used to get an idea of the signal distribution across each array. Supplementary listing B.1 shows the *R* script for installation of *limma*, import of microarray data and generation of quality plots.

The microarray data was evaluated using the open source software platform *EMMA2*, which has been designed for consistent storage and efficient analysis of microarray data (Dondrup et al. 2009). Normalization was carried out using the LOWESS method (Yang et al. 2002). Normalized data was subjected to Student's t-tests to determine significantly regulated regions with *p*-values below a threshold of 0.05. All *p*-values were corrected based on the method of Benjamini and Hochberg (Benjamini and Hochberg 1995) to account for the multiple testing situation.

3.6.3. Analysis of gene expression microarray data

Data analysis of gene expression microarray experiments was performed using the *limma* package in *R*. In a first step, the signal ranges before normalization were compared between the arrays. Data were then background corrected and quantile normalized. In order to collect the subset of genes highly expressed in the samples, all genes with expression values above the 75th percentile were defined as highly expressed (according to Schroeder et al. 2011). Furthermore, for each comparison between samples of interest, statistical parameters

for differential expression were calculated. Genes with adjusted p values below 0.01 were considered to be significantly differentially expressed. Supplementary listing B.2 shows the *R* script for import of microarray data, normalization and calculation of statistical parameters for comparisons between data sets of interest.

3.6.4. Analysis of next generation sequencing data

Parts of the analysis described in this section were performed in collaboration with Oliver Rupp, who bioinformatically processed the sequencing data, integrated DNA methylation profiles into the *GenDBE* genome browser and performed *de novo* discovery of DNA sequence motifs in DMRs.

Paired-end reads were quality checked using *FastQC* (Andrews 2010) and trimmed using *Trimmomatic* (Bolger, Lohse, and Usadel 2014). The reference CHO K1 genome (Xu et al. 2011) was corrected for single nucleotide polymorphisms (SNPs) within the CHO DP-12 cell line. First the reads from the sequencing of unconverted CHO DP-12 DNA were mapped to the CHO K1 genome with *Bowtie2* (Langmead and Salzberg 2012). SNPs were then called using the *SAMtools* mpileup algorithm and filtered using in house perl scripts. SNPs with a minimal coverage of 4 reads and at least 25 % of the reads containing the SNP were used for cell line correction using the *vcfutils* *vcf2fq* method (Li et al. 2009). The *Smithlab Methylation Data Analysis Pipeline* (*MethPipe*; Song et al. 2013) was used for mapping of reads from bisulfite treated libraries to the SNP-corrected reference genome, removal of duplicate reads, calculation of single-site methylation levels, estimation of bisulfite conversion rates and calculation of methylation levels in untranslated regions (UTRs), exons, introns and CGIs (annotation of CGIs based on (Wippermann et al. 2013). For comparison of two methylomes, the pipeline used a one-directional version of Fisher's exact test to determine the significance of differential methylation at individual CpGs. Differentially methylated regions (DMRs) were determined by using differential methylation scores of single-CpGs. Methylation data tracks were integrated into the in-house *GenDBE* genome browser (Rupp et al. 2014) to inspect selected regions with known methylation patterns for accordance of whole-genome bisulfite sequencing results. For this, *Branched chain amino acid aminotransferase cytosolic-like* gene, *Bcat1*, served as a methylated control. The *Beta-actin* gene, *Actb*, served as an unmethylated control. Partially methylated domains (PMDs) of more than 10 kb and methylation levels below 70 % were identified using a sliding window approach according to Lister et al. (Lister et al. 2009). For this, SNP-corrected CHO K1 scaffolds containing more than 100 CpGs were included into the analysis (5,874 scaffolds comprising 96 % of the CHO genome). For PMDs and hypermethylated domains (HMDs) with > 100 CpGs, average methylation levels, size and coverage by coding sequence (CDS) were analyzed. Assignment of CHO DP-12 PMDs

to the recently sequenced Chinese hamster (CH) chromosomes (Brinkrolf et al. 2013) was based on read coverages.

3.6.5. Functional gene set analysis

Gene Ontology annotations were performed using the *Database for Annotation, Visualization and Integrated Discovery (DAVID)* tool (Huang et al. 2009). Generally Applicable Gene-set Enrichment analysis of significantly differentially expressed genes was performed using *GAGE* (Luo et al. 2009). The *STRING database* was used to identify potential protein-protein interaction networks (Szklarczyk et al. 2015). The Bioconductor *Pathview* package was applied for pathway-based data integration and visualization of gene expression time courses on *KEGG* pathways (Luo and Brouwer 2013; Kanehisa and Goto 2000). Heatmaps were generated using the *ComplexHeatmap* package (Gu, Eils, and Schlesner 2016). The *de novo* motif discovery for determination of putative transcription factor binding sites within differentially methylated regions was performed using the motif discovery algorithm *DREME*, which was designed to find short core motifs of eukaryotic transcription factors (Bailey 2011). *De novo* discovered motifs were then subjected to a similarity search in transcription factor databases using the *Tomtom* algorithm (Gupta et al. 2007).

3.6.6. Statistics applied for DNA methylation and gene expression analysis

For comparison of two methylomes, a one-directional version of Fisher's exact test according to Altham (1971) was applied by using the default parameters of the *MethPipe* analysis pipeline to determine the significance of differential methylation at individual CpGs. Differentially methylated regions (DMRs) were determined by combining differential methylation scores of single-CpGs and DMRs with ≥ 1 significantly differentially methylated CpG were considered significant (Song et al. 2013). For gene expression analysis by CHO cDNA-microarrays genes with a false discovery rate ≤ 0.01 were considered to be significantly differentially expressed. *GAGE* was also performed with a false discovery rate ≤ 0.01 (Luo et al. 2009). Gene ontology analyses using *DAVID* were performed with Expression Analysis Systematic Explorer (EASE) scores ≤ 0.01 (Hosack et al. 2003; Huang et al. 2009). Significant enrichment of sequence motifs in DMRs by *DREME* (Bailey 2011) was calculated using the Fisher's Exact Test in comparison to a ten-fold bigger set of randomly chosen sequences with an equal length distribution and the default significance threshold of $p \leq 0.05$. *De novo* discovered motifs were subjected to a similarity search in transcription factor databases using the *Tomtom* algorithm and uncorrected p values ≤ 0.0005 (Gupta et al. 2007).

4

Results and discussion

Chinese hamster ovary (CHO) cells represent the most commonly used production cell line for therapeutic proteins. By recent genome and transcriptome sequencing a basis was created for future investigations of genotype-phenotype relationships and for improvement of CHO cell productivity and product quality. In this context information is missing about DNA cytosine methylation as a crucial epigenetic modification and an important element in mammalian genome regulation and development. For this reason this doctoral thesis aimed to analyze DNA methylation in CHO DP-12 cells that produce a recombinant human anti-IL-8 antibody. DNA methylation patterns in CHO cells were analyzed in general as well as changes of DNA methylation in response to the fatty acid butyrate. To be able to assess the impact of both the CHO DNA methylation landscape as well as differential DNA methylation upon butyrate addition on gene expression, CHO-specific DNA-microarrays were applied and data were analyzed in an integrative approach.

As another tool for DNA methylation analysis in CHO cells, a specifically designed CpG island microarray, already existed (Wippermann 2012; Wippermann et al. 2013), CpG island microarray performance was compared to whole-genome bisulfite sequencing regarding the butyrate effect on DNA methylation. Exemplary results of the CpG island microarray experiment were confirmed by gene-specific analysis of DNA methylation. Control hybridization experiments were furthermore performed in order to be able to assess potential sources of bias originating from the experimental procedure. Finally, putative central mediators of the butyrate effect were identified based on the comparative data analysis and functionally analyzed in transient knockdown experiments by RNA interference.

4.1. Integrative analysis of DNA methylation and gene expression in CHO cells

For integrative analysis of DNA methylation and gene expression changes in CHO cells upon butyrate treatment, samples were taken from batch cultivations of the exemplary antibody-producing CHO DP-12 cell line. Gene expression profiling by DNA-microarrays was performed as described previously (Becker et al. 2014), while whole-genome bisulfite sequencing was newly implemented for this cell line. Sample generation and the implementation process are described in the following sections.

4.1.1. Generation of samples

Replicate ($n = 4$) batch cultures of reference CHO DP-12 cells and cultures that were treated with 3 mM butyric acid were monitored regarding viable cell density, viability, cell specific productivity, glucose consumption and lactate formation (Fig. 4.1). The reference cultures showed exponential cell growth with a maximum viable cell density of 1.2×10^7 cells/mL on cultivation day 8 and high viabilities until day 9, while the maximum cell specific productivity of $7.4 \text{ pg/cell} \times \text{d}$ was reached on day 4. When glucose was limited the reference cells started to metabolize lactate.

CHO DP-12 cell cultures that were treated with butyrate reached a maximum viable cell density of about 3×10^6 cells/mL and proliferation was reduced after butyrate addition. One day post treatment, viabilities started to decrease. At the same time, the cultures started to metabolize lactate. However, cell specific productivities of the butyrate-treated cells reached a maximum of $17 \text{ pg/cell} \times \text{d}$. Samples for integrative DNA methylation and gene expression analysis were taken prior to butyrate addition as well as 6, 12, 24 and 36 hours post butyrate treatment. The CHO DP-12 DNA methylation landscape was analyzed using the replicate samples that were taken before butyrate addition. These samples furthermore served as the reference for the analysis of differential DNA methylation and gene expression upon butyrate addition.

4.1.2. Implementation of whole-genome bisulfite sequencing for CHO DP-12 cells

The analysis and interpretation of DNA methylation data by whole-genome bisulfite sequencing generally includes the bisulfite-mediated conversion of unmethylated cytosines, library generation, which leads to the amplification of converted cytosines as thymines, and next generation sequencing. Subsequently, the sequencing data are aligned to a reference genome

4.1. Integrative analysis of DNA methylation and gene expression in CHO cells

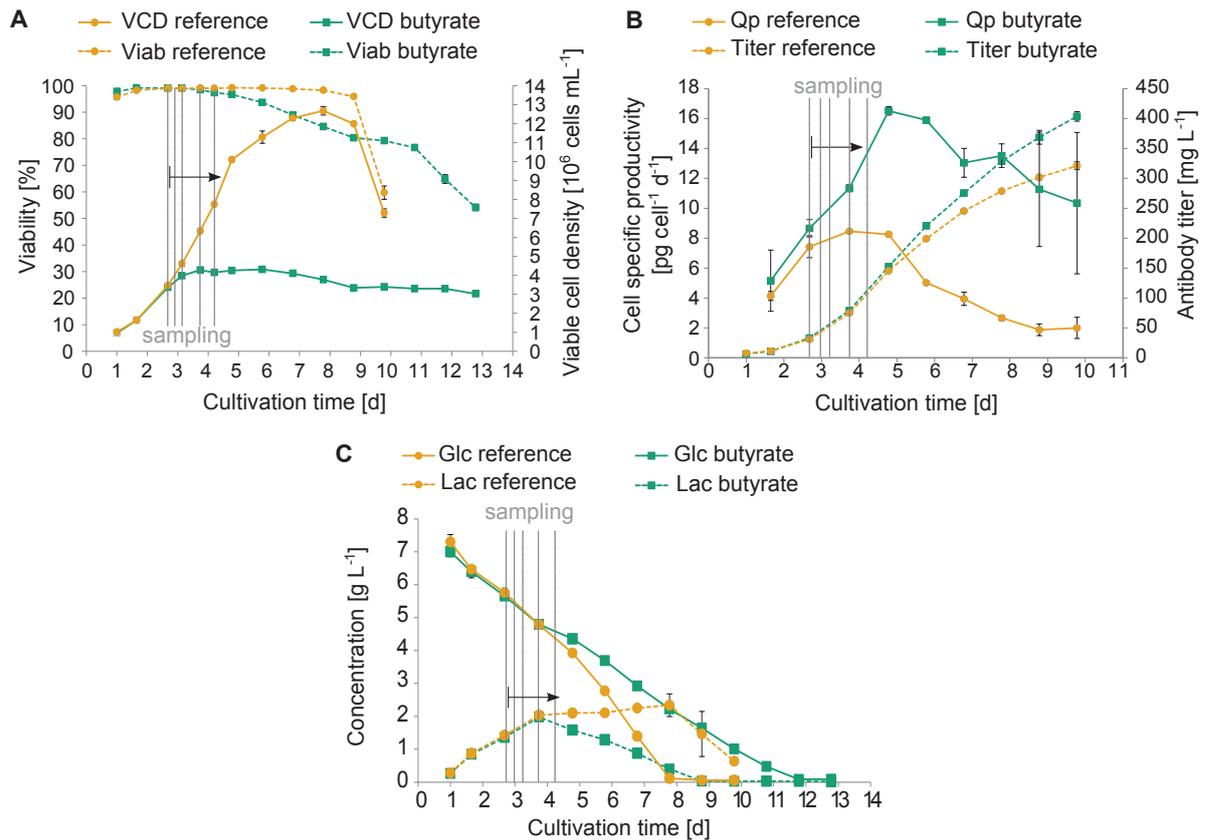


Figure 4.1.: Cell growth, IgG production, glucose and lactate concentrations of CHO DP-12 cell cultures (Wippermann et al. 2016). Error bars represent standard deviations of four replicate cultures. The vertical lines mark the sampling points for whole-genome bisulfite sequencing and gene expression analysis. Butyrate was added immediately after the first sampling. **A** Viable cell density (VCD) and mean viability (Viab) of reference and butyrate-treated cultures. **B** Cell specific productivity (Qp) and mAb titer of reference and butyrate-treated cultures. **C** Glucose (Glc) and lactate (Lac) concentrations of reference and butyrate-treated cultures.

to enable calculation of absolute cytosine methylation levels. After single-base methylation levels are determined, methylation data are visualized in a genome-browser, the distribution of methylation levels across genomes is analyzed and differentially methylated regions (DMRs) can be identified comparing individual methylomes. Combination of data generated in this way with gene expression measurements allows for an assessment of potential phenotypic effects and functional significance (Bock 2012).

The work flow of whole-genome bisulfite sequencing for determination of CHO DP-12 DNA methylation patterns and its combination with gene expression data is summarized in Fig. 4.2. Due to the fact that the CHO DP-12 genome is not yet publicly available, this procedure first of all required correction of the reference genome CHO K1 (Xu et al. 2011) for single nucleotide polymorphisms (SNPs) present between both genomes. Next generation sequencing data from bisulfite-treated reference samples and samples taken after butyrate addition could then be mapped to this customized genome (section 4.1.2). This allowed for the analysis

of the genome-wide distribution of DNA methylation in CHO DP-12 cells and differential DNA methylation could be determined between individual methylomes. Genome-wide DNA methylation data were next combined with gene expression data to analyze the CHO DP-12 DNA methylation landscape (section 4.1.3) and differential DNA methylation upon butyrate treatment in a time-course study including time points 6 h, 12 h, 24 h and 36 h post butyrate addition (section 4.1.4).

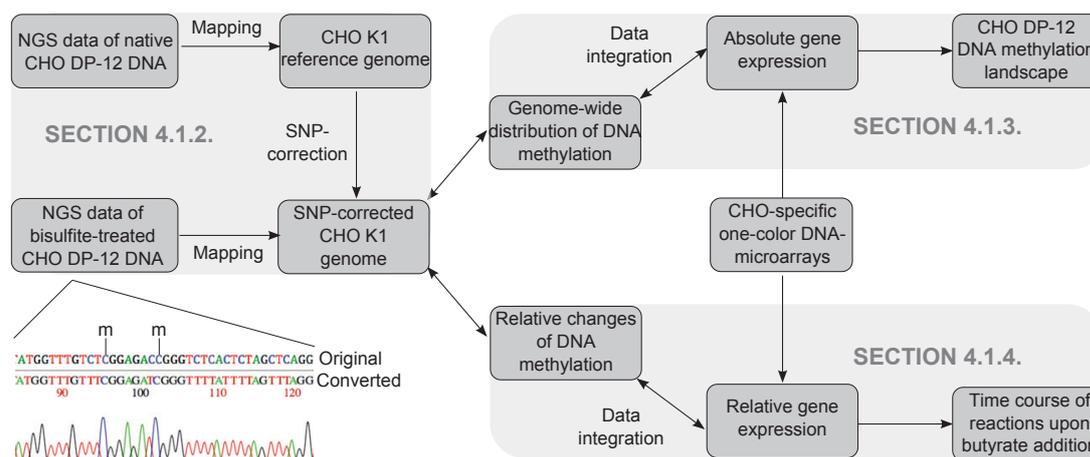


Figure 4.2.: Schematic representation of the work flow of integrating whole-genome bisulfite sequencing and CHO DNA-microarray data.

As the implementation of experimental methods requires utilization of controls, positive and negative methylated and unmethylated control regions were employed that were previously identified in the promoters of *Beta-actin* (*Actb*) and *Branched chain amino acid aminotransferase cytosolic-like* (*Bcat1*; Wippermann 2012). The methylation status of these regions was determined by bisulfite Sanger sequencing of pooled DNA samples from CHO DP-12 cells cultivated with and without butyrate addition. Fig. 4.3 summarizes the results of this analysis and shows a schematic representation of the *Actb* and *Bcat1* promoters with indication of the CpG island regions that were analyzed. The 'lollipop' plots below show that the analyzed region of the *Actb* promoter was found to be completely unmethylated (indicated by white circles), whereas the analyzed region of the *Bcat1* promoter was found to be completely methylated (indicated by black circles). Therefore, these regions served as controls for whole-genome bisulfite sequencing experiments.

4.1.2.1. Quality control of DNA samples and bisulfite sequencing libraries

Quality control of bisulfite sequencing experiments mainly focuses on potential biases that are introduced by the sample material (e.g. DNA quality and integrity) or incomplete bisulfite conversion (Bock 2012). To account for these quality requirements DNA samples extracted from CHO DP-12 cell cultures were checked photometrically for their purity by calculating the

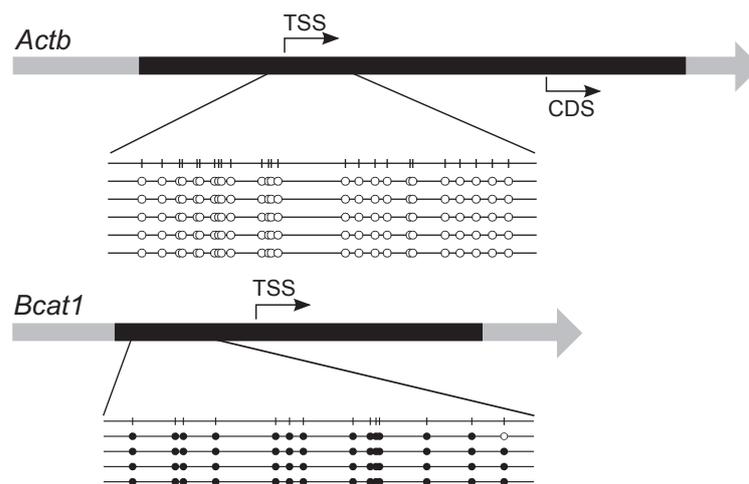


Figure 4.3.: Bisulfite sequencing results of parts of the CGIs of *Beta-actin* (*Actb*) and *Branched chain amino acid aminotransferase cytosolic-like* (*Bcat1*). CGIs are indicated in black. The arrows mark the transcription start sites (TSS) and the beginning of the coding sequence (CDS) for *Actb*. Bisulfite Sanger sequencing results are shown as lollipop plots and black circles represent a methylated CpG position, whereas white circles indicate unmethylated CpG positions.

OD₂₆₀/OD₂₈₀ ratio and for their integrity by application of 0.6 % agarose gels (Supplementary Fig. B.13) before they were subjected to bisulfite treatment.

The efficiency of bisulfite-mediated conversion of unmethylated cytosines was considered to be the next critical step, as it is known to potentially be incomplete. This is due to the fact that bisulfite conversion of unmethylated cytosines depends on complete denaturation of DNA molecules, which can be difficult for intact chromosomal DNA. For this reason the efficiency is controlled by spiking DNA samples with unmethylated lambda phage DNA. It was hypothesized that bisulfite-mediated conversion of DNA might be more efficient when the DNA was sheared before bisulfite treatment. However, sequencing libraries that were generated from sheared, bisulfite-treated DNA turned out to be inhomogeneous regarding their size distribution (Supplementary Fig. B.15). Therefore, five Illumina sequencing libraries were generated from unsheared bisulfite-treated DNA for the reference and sampling points '6 h', '12 h', '24 h' and '36 h' by using barcoded PCR primers to allow for multiplexed sequencing. Library qualities were checked for the correct distribution of fragment sizes on a *Bioanalyzer*[®] (Supplementary Fig. B.14). The sequencing libraries generated from bisulfite-treated DNA exhibited a mean size of about 360 bp.

In order to get an idea of library qualities, potential difficulties that might compromise sequencing, and in order to be able to adjust reaction conditions in the subsequent *HiSeq* sequencing runs if necessary, the libraries were pooled and subjected to a 2×75 bp paired-end run on an Illumina *MiSeq* system with comparably low sequencing output as a pre-test. Also, the conversion rate of the unmethylated spike-in lambda DNA was controlled in this step,

which was above 99 %. *MiSeq* sequencing resulted in 25.14 millions of reads. The pooled libraries exhibited a base composition of 29 % adenine, 22 % guanine, 42 % thymine and 6 % cytosine (Fig. 4.4) up to sequencing cycle 80, as most cytosines outside of CpG contexts are unmethylated in mammalian genomes (Schübeler 2015) and were therefore converted to thymines by bisulfite treatment. After cycle 80 DNA fragments were read from the other end. After this reversal and up to the end of the sequencing reactions guanine and adenine, which are complementary to cytosine and thymine, exhibited the expected percentages of 6 % (as guanine is complementary to cytosine, which was reduced by bisulfite treatment) and 42 % (as adenine is complementary to thymines, which exhibited a higher percentage because unmethylated cytosines were converted). Generally, the libraries were considered suitable for sequencing and no adjustments were necessary in the subsequent *HiSeq* runs.

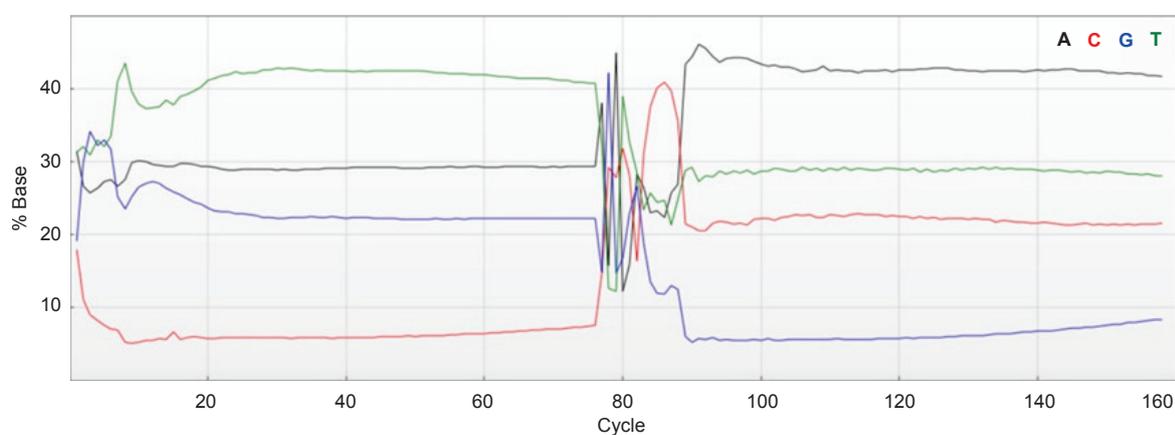


Figure 4.4.: Percent base plot of *MiSeq* sequencing of pooled libraries from whole-genome bisulfite sequencing.

4.1.2.2. Illumina *HiSeq* sequencing of libraries from native and bisulfite-treated CHO DP-12 DNA and generation of the CHO DP-12 reference genome

Multiplexed sequencing was performed on an Illumina *HiSeq 1500* system in two 2×100 bp runs. Next to the libraries from bisulfite-treated DNA, a library generated from a pool of DNAs from both reference and butyrate-treated cells was analyzed that was used for SNP-correction of the CHO K1 genome later on. This native library (mean library size 446 bp) was sequenced with a theoretical sequencing depths of 47-fold. Libraries from bisulfite-treated DNA of reference and butyrate-induced cultures were sequenced with a theoretical coverage of 32-fold per sample. In total, 489 Gb of sequence output were obtained by the two sequencing runs. Tab. 4.1 shows the amount of sequence data per sample and the corresponding calculated coverages. The CHO genome size of 2.5 Gb as determined by Xu et al. 2011 was used for this calculation.

Table 4.1.: Numerical summary of raw sequence output and calculated genome coverage from Illumina *HiSeq* sequencing of libraries from native and bisulfite-treated (BS) CHO DP-12 DNA before further data processing.

Sample	BS treatment	Output [Gb]	Coverage [fold]
Pooled DNAs	–	115.8	46.3
Reference	✓	81.1	32.4
6 h	✓	71.4	28.6
12 h	✓	69.2	27.7
24 h	✓	72.8	29.1
36 h	✓	66.7	26.7

116 Gb of sequence information were obtained for the untreated pool of DNAs, whereas 67-81 Gb of sequence information were obtained for the bisulfite-treated samples. The genomic coverages, which were calculated based on the amount of raw data, corresponded to the theoretical coverages that were used for setting up the multiplexed sequencing reactions only for the pool of DNAs as well as for the reference. For the samples '6 h' to '36 h' a mean coverage of 28-fold instead of 32-fold was achieved.

To generate a CHO DP-12 reference genome by SNP-correction of the CHO K1 genome data, sequencing reads from the library of pooled DNAs were trimmed and 95.8 Gb (575,986,927 reads) of sequence information remained. 84 % of these reads could be mapped to the CHO K1 reference genome and 74 % were properly paired. SNP-detection showed 4,054,041 SNPs with a coverage of ≥ 4 reads and at least 25 % of the reads containing the SNP between both genomes. The mean coverage of the CHO K1 genome by reads from CHO DP-12 DNA after trimming and mapping was 32-fold. In the following sections the customized CHO DP-12 reference genome is referred to as 'the SNP-corrected CHO K1 genome'.

4.1.2.3. Selection of a mapping algorithm for data analysis

The *Smithlab Methylation Data Analysis Pipeline* (*MethPipe*; Song et al. 2013) was chosen for the analysis of data from whole-genome bisulfite sequencing experiments. However, prior to data analysis the mapping tool integrated into the *MethPipe* analysis pipeline was compared to the mapping tool *BS-Seeker2* (Guo et al. 2013) in order to select the algorithm best suitable for data analysis.

Therefore, reads generated by *HiSeq* sequencing from bisulfite-treated libraries of the reference cultures were mapped to exemplary parts of the SNP-corrected CHO K1 genome. Generally, mapping tools for processing of reads from bisulfite-treated libraries employ one of two different strategies. The *MethPipe* mapping tool uses a principle called 'wild-card' alignment, which converts every C in the original sequence to the 'wild-card letter' Y. Fig. 4.5A shows an exemplary alignment with a reference sequence at the top and bisulfite sequencing reads in the center (Fig. modified from Bock 2012). Generally, Cs and Ts present in

the reads are allowed to map to Ys. If a read corresponds to multiple positions in the reference sequence it is removed from the analysis (crossed out). Methylation levels are calculated based on the number of Cs at each Y position. The *BS-Seeker2* mapping tool, in contrast, uses a 'three-letter' alignment and converts each C in the reference sequence to an uppercase T and each C in the reads to a lowercase t (Fig. 4.5B). During alignment, both T and t can map to a T in the reference sequence and ambiguous reads are discarded. Methylation levels are calculated based on the number of lowercase ts.

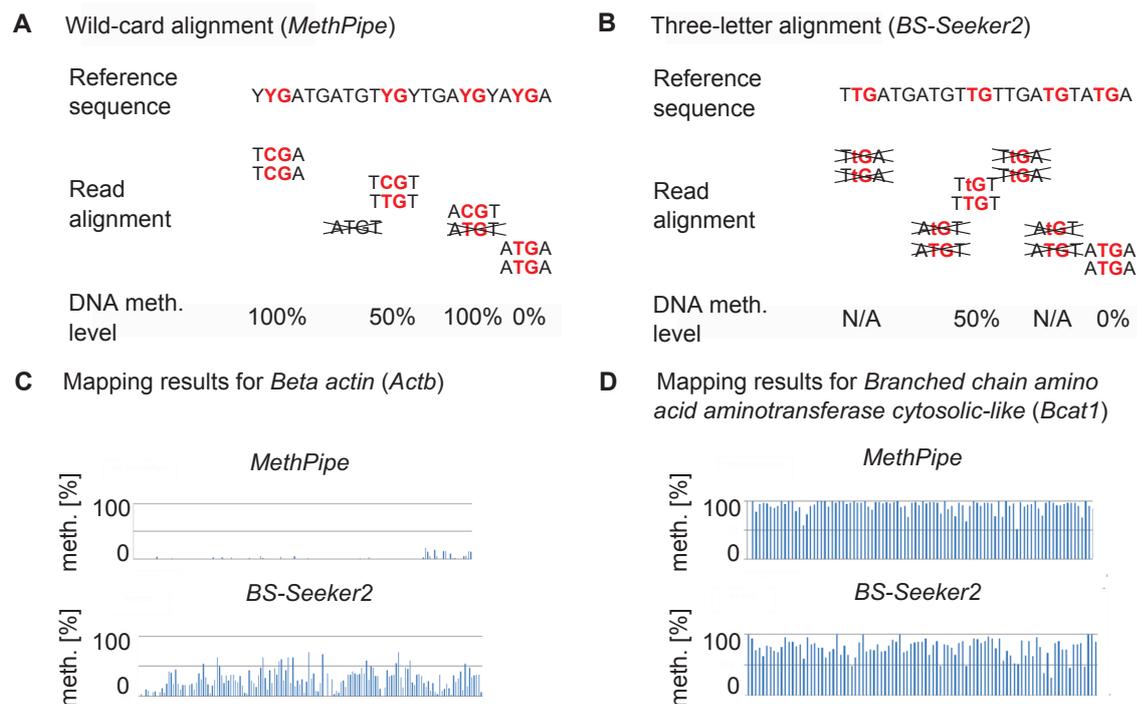


Figure 4.5.: Comparison of mapping tool performance. Two pipelines for DNA methylation data analysis were compared regarding the results for CHO cell genomic regions with known methylation levels (*Actb* and *Bcat1*). **A,B** The *MethPipe* pipeline (Song et al. 2013) uses a wild-card aligner, while *BS-Seeker2* (Guo et al. 2013) employs a three-letter alignment tool (figure based on Bock 2012). **C** Single-base methylation levels for the unmethylated control region *Actb*. **D** Single-base methylation levels for the methylated control region *Bcat1*.

The *MethPipe* mapping tool provided correct results for the internal CHO methylated (the *Bcat1* CpG island) and unmethylated (the 5' UTR-associated CpG island of *Actb*) controls, whereas the *BS-Seeker2* results only gave an impression of the methylation status, as it generated a high background noise (Fig. 4.5C, D). CpG island regions suspected to be unmethylated were inspected as well and identically noisy signals were found to be generated by *BS-Seeker2* (see Supplementary Fig. B.16 for exemplary results). Therefore, the internal *MethPipe* mapping algorithm was used for data processing, which included the removal of duplicate reads as well as calculation of single-site methylation levels, mean methylation of genomic regions, and differential methylation between methylomes.

4.1.2.4. Processing of whole-genome bisulfite sequencing data

In the first step of data analysis, absolute methylation levels for each CpG dinucleotide within the genomes of reference and butyrate-treated samples were determined by calculating the ratio of mapped thymines and cytosines at individual CpG positions of the SNP-corrected CHO K1 genome.

Table 4.2.: Numerical summary of whole-genome bisulfite sequencing results after mapping to the SNP-corrected CHO K1 genome.

Sample	Mean methylation [%]	CpG sites covered [%]	Mean coverage [fold]	Total reads
Reference	61.09	94.04	14.30	398,995,016
6 h	59.89	93.86	13.62	353,379,241
12 h	59.55	93.86	13.24	342,468,374
24 h	60.06	94.03	14.02	364,018,688
36 h	61.23	93.39	11.01	329,546,415

Tab. 4.2 shows that mapping resulted in an average sequencing depth of 11- to 14-fold. This means that, per sample, about 50 % of the reads from bisulfite-treated DNA could be mapped to the SNP-corrected CHO K1 genome. About 94 % of all CpG sites known in the CHO genome were covered by each library (i.e. about 15.3 million CpGs in each genome were analyzed). Furthermore, butyrate treatment was found to induce a maximum decrease in genome-wide mean methylation levels from 61.09 % in the reference samples to 59.55 % at sampling point '12 h'.

4.1.2.5. Integration of whole-genome DNA methylation data into *GenDBE* and inspection of exemplary genomic regions

Next, DNA methylation data of reference and butyrate-treated CHO DP-12 cells were integrated into the genome browser *GenDBE*, which is a web-based tools for browsing eukaryotic genomes (Rupp et al. 2014). First, overall DNA methylation patterns were compared between the sampling points and proved to be very consistent (see exemplary genomic region in Supplementary Fig. B.17). To control the correctness of data integration, the control promoter regions of *Actb* and *Bcat1* were inspected (Fig. 4.6A; data are only shown for the reference, as no differences were observed between the individual samples). The 5' UTR-associated CpG island of *Actb* was found to be completely unmethylated, whereas the *Bcat1* CpG island was found to exhibit methylation levels of nearly 100 %.

Assuming that CHO DP-12 cells exhibited a DNA methylation profile similar to other mammalian cells, promoter-associated CpG islands of CHO housekeeping genes (according to Bahr et al. 2009) were checked for their methylation status as another measure of data quality (Fig. 4.6B). All CpG island regions of the CHO housekeeping genes *Component of*

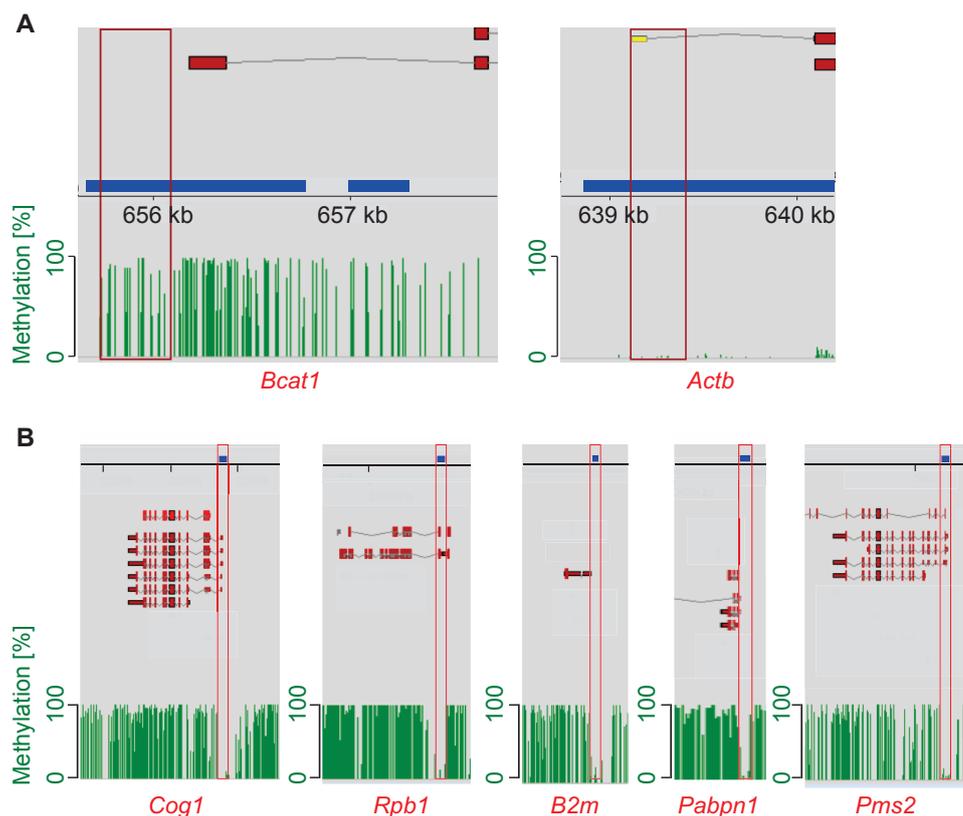


Figure 4.6.: Methylation data tracks were loaded into the *GenDBE* genome browser (Rupp et al. 2014) to visually inspect selected genomic regions. Exons are indicated by red rectangles, CpG islands by blue rectangles and UTRs by yellow rectangles. Methylation levels for each CpG are represented by green bars. **A** Promoter regions of *Branched chain amino acid aminotransferase cytosolic-like* gene (*Bcat1*) and 5' UTR of *Beta-actin* gene (*Actb*). Regions of the CHO DP-12 genome with known methylation levels are marked by a red box. **B** Visual inspection of CpG islands at housekeeping gene promoters. Inspected were the genes *Cog1*, *Rpb1*, *B2m*, *Pabpn1* and *Pms2*.

oligomeric golgi complex 1 (*Cog1*), *DNA-directed RNA polymerase II subunit RPB1* (*Rpb1*), *Beta-2-microglobulin* (*B2m*), *Poly(A) binding protein, nuclear 1* (*Pabpn1*) and *Mismatch repair endonuclease PMS2* (*Pms2*) were found to be unmethylated.

4.1.2.6. Discussion: Generation of genome-scale DNA methylation data for CHO DP-12 cells

In order to analyze genome-wide DNA methylation in CHO DP-12 cells by whole-genome bisulfite sequencing, a customized reference genome was generated for CHO DP-12 cells. Reads from bisulfite-treated CHO DP-12 DNA, which proved to be fully converted based on the complete conversion of spike-in lambda DNA, could be successfully mapped to it using the *MethPipe* data analysis pipeline. The *BS-Seeker2* mapping algorithm was excluded, as the results exhibited a high background noise. The observation that the two mapping algorithms that were tested performed very differently can be explained by the fact that the

three-letter aligner *BS-Seeker2* removes all remaining Cs from the sequencing reads, thereby further reducing the sequence complexity and increasing the number of ambiguous alignment positions (Bock 2012). Absolute methylation levels for each CpG dinucleotide within the CHO DP-12 genome were therefore generated using the internal *MethPipe* mapping algorithm. Single-site methylation data were then integrated into the *GenDBE* genome browser (Rupp et al. 2014) for visual inspection of genomic regions with known methylation levels as well as promoter CpG islands of CHO housekeeping genes. The inspected regions exhibited the expected methylation levels.

Recent publications show similar approaches of mapping genome-scale DNA methylation data from bisulfite sequencing experiments using the *MethPipe* analysis pipeline. For example, Lin et al. compared the methylomes of different breast cancers to normal human breast tissue and analyzed the data using the *MethPipe* suite (Lin et al. 2015). Similarly to the results of the mapping that was performed within the scope of the experiment in this thesis, they could map 79 % of the bisulfite reads to the *Homo sapiens* reference genome and obtained a sequencing depth of about 19-fold, with 91 % of all CpG sites covered and a bisulfite conversion rate of at least 99 %. The authors did not include data about quality control by visual inspection of selected genomic regions. The *MethPipe* analysis pipeline was also used in a study that analyzed the epigenomes of human primordial germ cells. Progressive genome-wide DNA demethylation to basal levels was shown for these cells during early development, with some genomic regions being resistant to overall demethylation (Tang et al. 2015). The authors also reported a bisulfite conversion rate of at least 99 % and mapping efficiencies of a maximum 58 %.

4.1.2.7. Conclusion: Implementation of whole-genome DNA methylation analysis enables characterization of epigenetic phenomena

Taken together, whole-genome bisulfite sequencing was successfully implemented for CHO DP-12 cells and enabled generation of methylome data for reference and butyrate-treated samples. In combination with gene expression studies by DNA-microarrays this promised to enable the characterization and functional analysis of the CHO DP-12 DNA methylation landscape, which was considered necessary for a comprehensive understanding of epigenetic effects linking genotype and phenotype in the production of biopharmaceuticals. The CHO DP-12 DNA methylation landscape and corresponding gene expression data were analyzed both in a comparison to published mammalian methylomes and with regard to CHO phenotypic properties. In a second step, the effects of the productivity enhancing substance butyrate on DNA methylation and gene expression were analyzed regarding affected pathways, putative dynamic effects and time courses of induced changes. Both cellular levels were analyzed

separately as well as in combination to display the cellular reaction to butyrate addition as detailed as possible.

4.1.3. The CHO DP-12 DNA methylation landscape

As a DNA methylation map of a CHO cell line did not exist before, the analysis of the CHO DP-12 cell DNA methylation landscape was of special importance. To be able to assess the impact of the CHO DP-12 DNA methylation landscape on CHO phenotypic properties, this analysis was combined with gene expression data generated by DNA-microarrays. For gene expression microarray analysis four replicate DNA-microarrays were hybridized. Data analysis was performed using the *limma* package in *R* (Smyth 2005) after background correction and quantile normalization across the technical replicates. As the subset of genes highly expressed in the samples was supposed to be correlated with the DNA methylation data, all genes with expression values above the 75th percentile were defined as highly expressed (according to Schroeder et al. 2011).

4.1.3.1. Global assessment of the CHO DP-12 DNA methylation landscape

To assess the levels and patterning of DNA methylation across the CHO DP-12 cell genome, a general analysis was performed first. Mean CpG methylation levels were calculated and CHO DP-12 cells were found to exhibit an average methylation level of 61 % (see section 4.1.2.4). As global methylation levels of > 70 % were reported for a wide variety of cells (e.g. 82 % methylation for HUES64 cells or 78 % methylation for fetal brain cells; Ziller et al. 2013), this finding suggests hypomethylation for the samples analyzed. A global analysis of DNA methylation on the level of single CpG dinucleotides showed a generally bimodal distribution of methylation with 52 % of all CpGs being largely methylated (> 70 % of reads showing methylation) and 30 % of unmethylated CpGs (\leq 10 % of reads showing methylation; Fig. 4.7A). The remainder of CpG dinucleotides showed partial methylation. The DNA methylation data were also searched for methylation in a non-CpG context. Overall, such cytosines exhibited a mean methylation of 3.4 %. When the number of non-CpG cytosines that were covered by more than 4 reads and exhibited a methylation level of \geq 70 % was calculated, only 0.1 % of all cytosines were found to be methylated. This indicated that non-CpG methylation might be irrelevant in CHO DP-12 cells, as high levels of non-CpG methylation were for example found in brain tissue, which was shown to exhibit a mean level of non-CpG methylation of about 25 % (Guo et al. 2014). Due to the very low number of non-CpG methylation in the CHO DP-12 genome, this type of methylation was considered to be irrelevant and potentially due to sequencing bias.

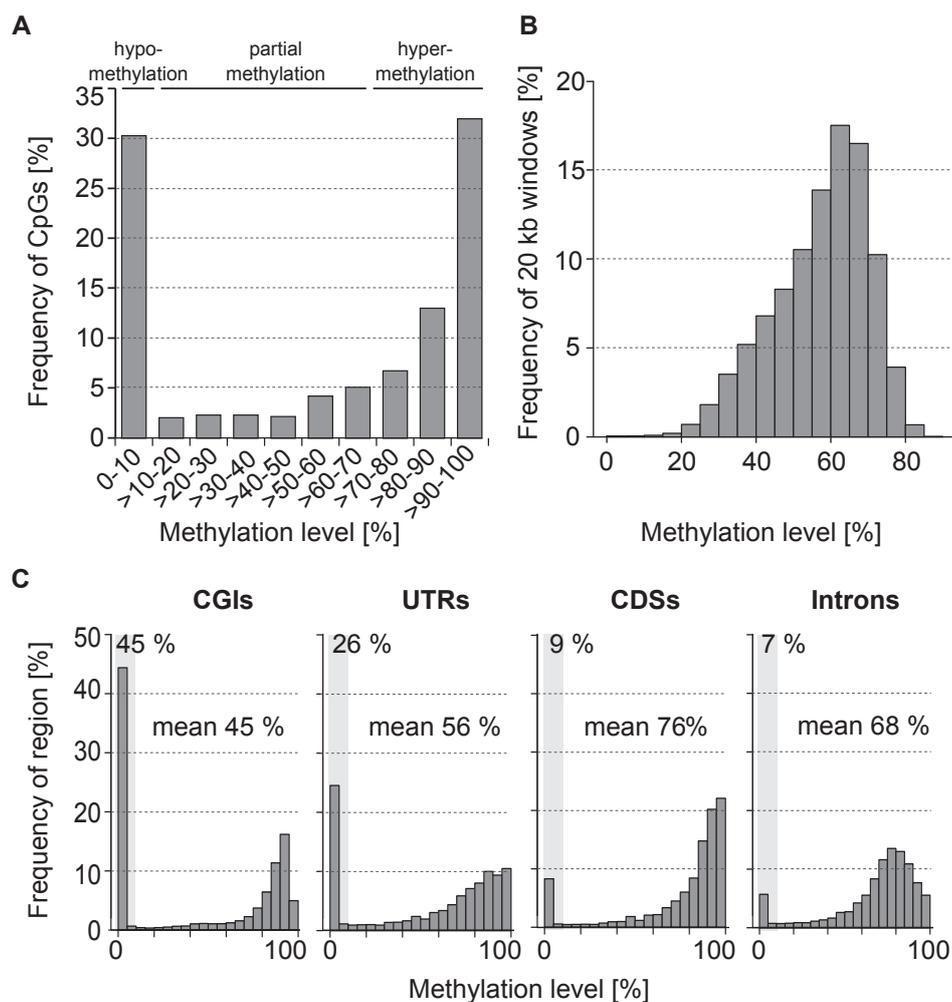


Figure 4.7.: Global distribution of DNA methylation levels in reference CHO DP-12 cultures (Wippermann et al. 2015). **A** Percent methylation of CpG dinucleotides. **B** Percent methylation of 20 kb, non-overlapping windows tiled through the CHO DP-12 genome. **C** Percent methylation of CpG islands (CGIs), untranslated regions (UTRs), coding sequences (CDSs) and introns. Lightgrey bars mark the percentage of regions with methylation levels below 10 %. Also indicated is the average methylation level of each genomic feature.

To get an initial overview of the CHO DP-12 cell DNA methylation landscape, it was examined whether the observed overall hypomethylation was due to a uniform distribution of CpGs with low and high methylation levels or if there were local differences in average methylation between genomic regions. Therefore, the distribution of average methylation levels for non-overlapping 20 kb windows tiled through the genome was analyzed (according to Schroeder et al. 2011; Fig. 4.7B). Windows containing less than 100 CpGs were excluded from the analysis. The purpose of this calculation was to diminish the effect of putatively unmethylated CpG islands. This analysis showed a very broad methylation spectrum between 20 % and 80 % methylation with a peak at 60 %, indicating a high degree of heterogeneity in local CHO DP-12 DNA methylation.

To further characterize the CHO DP-12 cell methylome with regard to the methylation of annotated genomic features, DNA methylation levels for CGIs, untranslated regions (UTRs), coding sequences (CDSs) and introns (CDS and introns also referred to as gene-bodies) were determined (Fig. 4.7C). 45 % of all CGIs in the CHO DP-12 cell genome proved to be unmethylated and, generally, CGIs showed a mean methylation level of 45 %. UTRs showed a mean methylation of 56 %, with 26 % of them being unmethylated. CDSs and intronic regions exhibited mean methylation levels of 76 % and 68 %, respectively, a low percentage of unmethylated regions (9 % and 7 %) and a high frequency of highly methylated regions.

In summary, CHO DP-12 cells showed global hypomethylation and evidence for local differences in average methylation levels. In accordance with previously described DNA methylomes, CGIs proved to have the lowest methylation levels, although a larger fraction of them was methylated as e.g. in comparison to CGIs in human tissues (9 % of hypermethylated CGIs; Eckhardt et al. 2006). Gene-bodies in CHO DP-12 cells exhibited a higher level of DNA methylation than UTRs.

4.1.3.2. Evidence for partially methylated domains in the CHO DP-12 genome

As indications of heterogeneous DNA methylation patterns in the CHO DP-12 cell genome were found, methylation levels were plotted along the SNP-corrected CHO K1 scaffolds to visualize their distribution. In this analysis heterogeneous patterns of DNA methylation with regions showing a high degree of methylation and neighboring regions with partial methylation were observed (an example is shown in Fig. 4.8 and Supplementary Fig. B.23).

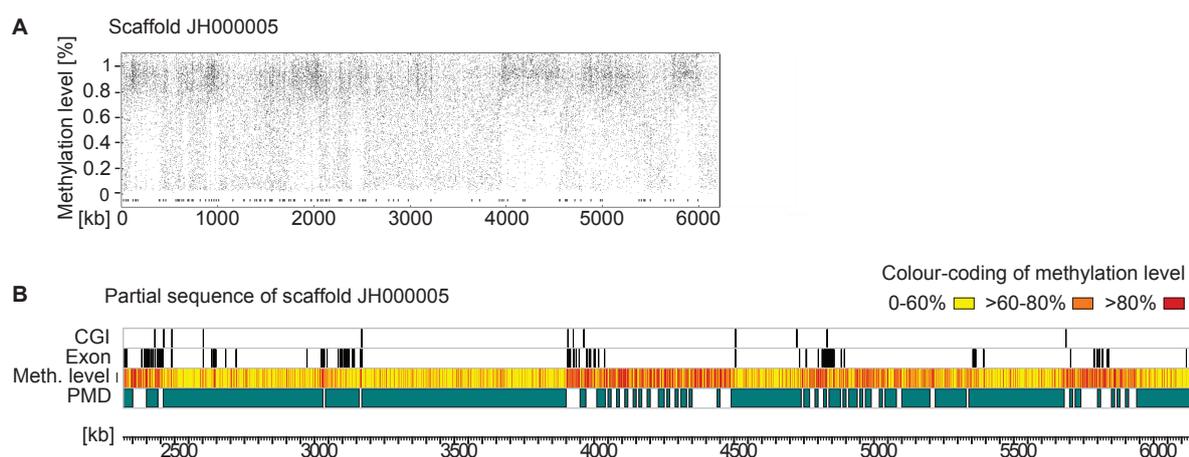


Figure 4.8.: Exemplary display of SNP-corrected CHO K1 scaffold JH000005 (Wippermann et al. 2015). **A** Methylation levels of single CpGs were plotted along the scaffold and jittered by 2 % to prevent overplotting. **B** CHO DP-12 DNA methylation levels for 1 kb windows were plotted along the scaffold and color-coded according to the legend. Also shown are exons and CpG islands (CGI) in black. Regions identified as CHO DP-12 PMDs are marked by green bars. Complete scaffolds JH000001 to JH000009 are shown in Supplementary Fig. B.23.

Similar patterns of alternating methylation domains were e.g. also found in colorectal cancer, breast cancer, fibroblast cultures, SH-SY5Y neuronal cells and placenta. Regions with partial methylation in these cells and tissues were termed partially methylated domains (PMDs) (Schroeder et al. 2011; Schroeder et al. 2013; Gaidatzis et al. 2014; Lister et al. 2009; Hon et al. 2012; Berman et al. 2012). According to the properties of these known PMDs, regions with a minimum size of 10 kb and a methylation level below 70 % were classified as CHO DP-12 PMDs. Calculation of the PMD-coverage of all SNP-corrected CHO K1 scaffolds with more than 100 CpGs (4,874 scaffolds comprising 96 % of the CHO genome) showed that 62 % of the analyzed scaffolds were partially methylated. Accordingly, 38 % of the CHO DP-12 cell genome were covered by highly methylated domains (HMDs).

As it was shown that X chromosome inactivation complicates PMD analyses (Schroeder et al. 2013) and due to the fact that CHO cells originate from a female donor, the next question was whether CHO DP-12 PMDs showed differences in their distribution on the single chromosomes. This possibility was checked by mapping the whole-genome bisulfite sequencing reads to the Chinese hamster (CH) chromosomes sequenced and annotated by Brinkrolf et al. (Brinkrolf et al. 2013). Normalized to chromosome size, 18.4 PMDs/Mb on average and a very significant difference in PMD count per Mb for chromosome X were detected, which was much higher with 26.4 PMDs/Mb (Fig. 4.9). A similar finding was reported by Lister et al. for IMR90 PMDs (Lister et al. 2009). Therefore, PMDs on the X chromosome were omitted from further analyses of PMD and HMD size, methylation level and coding sequence (CDS) coverage.

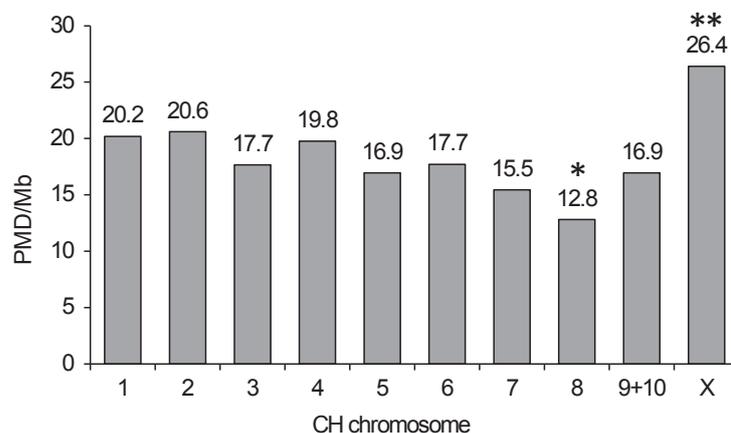


Figure 4.9.: PMD count per Mb for CHO DP-12 PMDs based on Illumina read coverage assigned to Chinese hamster chromosomes (Wippermann et al. 2015). One standard deviation of the mean is indicated by one, two standard deviations by two asterisks.

CHO DP-12 PMDs exhibited 54 % methylation on average, a mean size of 64 kb and a maximum size of 2,990 kb (Fig. 4.10A-B). In contrast, HMDs showed a mean methylation of 77 %, a mean size of 34 kb and a maximum size of 450 kb. Next, the coverage of CHO DP-12 PMDs and HMDs by coding sequence was calculated (Fig. 4.10C). CHO DP-12

PMDs showed a mean CDS coverage of 1.8 % and a maximum coverage of 65 %, whereas HMDs exhibited a mean CDS coverage of 3.4 % and a maximum coverage of 90 %. In order to furthermore obtain the number of CHO DP-12 PMD-localized genes, all annotated CHO genes were categorized according to their genomic position within or outside of a CHO DP-12 PMD. Genes with partial PMD coverage were excluded. 4,265 unique genes co-localized with a CHO DP-12 PMD, representing only 18 % of the total of 24,383 genes predicted for the CHO K1 cell line (Xu et al. 2011). In summary, CHO DP-12 PMDs proved to be larger but also showed clear evidence for containing fewer genes than CHO DP-12 HMDs.

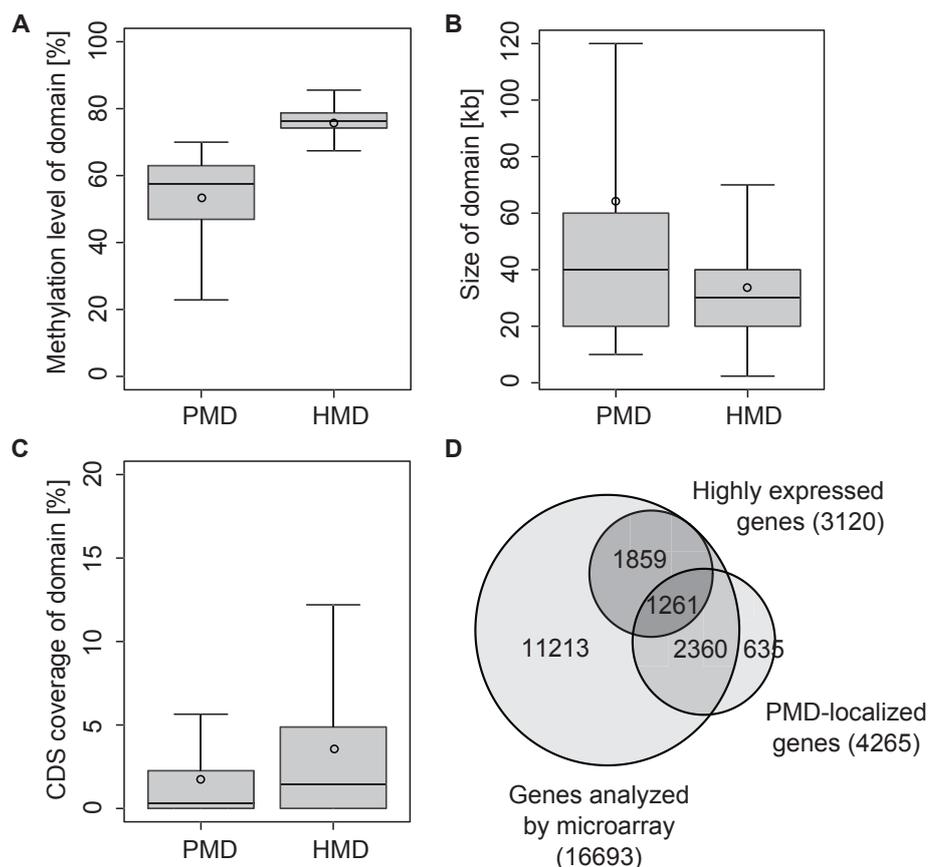


Figure 4.10.: Characterization of CHO DP-12 partially methylated domains (PMDs) and highly methylated domains (HMDs; Wippermann et al. 2015). **A–C** Mean values of data are indicated by black circles. Outliers are not shown. **A** Percent methylation. **B** Size in kb. **C** Coding sequence (CDS) coverage. **D** 4265 genes co-localized with CHO DP-12 PMDs. Microarray analysis enabled assessment of the expression levels of 16,693 genes. Of these, 3120 were highly expressed. 1261 of the highly expressed genes also co-localized with CHO DP-12 PMDs.

4.1.3.3. Functional classification of CHO DP-12 PMD-localized by DNA-microarray analysis

In a next step, PMD-gene expression was analyzed using a CHO-specific DNA-microarray (Becker et al. 2014), as a known characteristic of PMDs is transcriptional repression of

co-localized genes. The DNA-microarray enabled the analysis of 16,693 unique gene clusters and covered 85 % of the identified CHO DP-12 PMD-genes. The remaining (635) PMD-genes were not present on the microarray and were therefore not analyzed (Fig. 4.10D). As a one-color labeling protocol was applied, an estimation of gene expression strengths was possible by this experiment and allowed for the identification of the subset of genes that were highly expressed. By this means, 3120 highly expressed genes were identified. 1261 of them were amongst the CHO DP-12 PMD-localized genes (3621 genes in total). This means that about 65 % of the PMD-localized genes were not highly expressed.

Table 4.3.: Gene ontology (GO) classification of total CHO DP-12 PMD-genes (4265; Wippermann et al. 2015).

GO ID	GO term	Count	<i>p</i> value
GO:0006455	Translational elongation	81	5.59×10^{-29}
GO:0006412	Translation	169	1.96×10^{-22}
GO:0007608	Sensory perception of smell	202	5.95×10^{-21}
GO:0007606	Sensory perception of chemical stimulus	214	3.16×10^{-19}
GO:0050877	Neurological system process	427	1.31×10^{-13}
GO:0050890	Cognition	334	2.90×10^{-13}
GO:0007600	Sensory perception	302	6.29×10^{-13}
GO:0008277	G-protein coupled receptor protein signaling pathway	376	8.68×10^{-9}
GO:0042254	Ribosome biogenesis	60	9.95×10^{-8}
GO:0022613	Ribonucleoprotein complex biogenesis	77	1.67×10^{-6}
GO:0034470	ncRNA processing	78	4.47×10^{-6}
GO:0006364	rRNA processing	45	5.89×10^{-6}
GO:0034660	ncRNA metabolic process	91	8.94×10^{-6}
GO:0016072	rRNA metabolic process	45	2.28×10^{-5}
GO:0000375	RNA splicing, via transesterification reactions	64	3.19×10^{-5}
GO:0000377	RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	64	3.19×10^{-5}
GO:0000398	Nuclear mRNA splicing, via spliceosome	64	3.19×10^{-5}
GO:0006396	RNA processing	182	1.25×10^{-4}
GO:0042274	Ribosomal small subunit biogenesis	10	1.80×10^{-4}
GO:0007166	Cell surface receptor linked signal transduction	549	1.95×10^{-4}
GO:0042273	Ribosomal large subunit biogenesis	9	5.70×10^{-4}
GO:0030005	Cellular di-, tri-valent inorganic cation homeostasis	82	8.27×10^{-4}

Tab. 4.3 shows the gene ontology (GO) analysis of the 4265 genes that were in total associated with CHO DP-12 PMDs and were found to be significantly enriched in two different general categories. These genes were on the one hand related to gene expression and RNA splicing, e.g. represented by the GO terms 'translational elongation' (81 genes), 'translation' (169 genes) or 'ribosome biogenesis' (60 genes) as well as 'RNA processing' (182 genes). On the other hand, genes related to neurological processes were enriched, e.g. represented by the specific GO terms 'sensory perception of smell' (202 genes), 'neurological system process' (427 genes) or 'cognition' (334 genes).

Next, separate gene ontology annotations were performed for the highly expressed PMD-genes (1261) and the weakly or not expressed PMD-genes (2360), respectively. Interestingly,

Table 4.4.: Separate gene ontology (GO) classifications for highly expressed (1261) and weakly or not expressed (2360) CHO DP-12 PMD-genes (Wippermann et al. 2015). Displayed are the top 20 GO terms for both groups according to p values. Complete lists of gene ontology annotations can be found in Supplementary Tabs. B.5 and B.6.

GO ID	GO term	Count	p value
Highly expressed CHO DP-12 PMD-genes			
GO:0006412	Translation	147	4.54×10^{-76}
GO:0006414	Translational elongation	77	7.25×10^{-64}
GO:0006396	RNA processing	117	6.47×10^{-25}
GO:0042254	Ribosome biogenesis	50	1.14×10^{-23}
GO:0022613	Ribonucleoprotein complex biogenesis	60	6.73×10^{-23}
GO:0000398	Nuclear mRNA splicing, via spliceosome	50	9.30×10^{-19}
GO:0000375	RNA splicing, via transesterification reactions	50	9.30×10^{-19}
GO:0000377	RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	50	9.30×10^{-19}
GO:0006364	rRNA processing	37	2.61×10^{-17}
GO:0008380	RNA splicing	68	2.84×10^{-17}
GO:0016072	rRNA metabolic process	37	1.31×10^{-16}
GO:0006397	mRNA processing	70	2.09×10^{-15}
GO:0006119	Oxidative phosphorylation	36	2.32×10^{-15}
GO:0034660	ncRNA metabolic process	56	1.47×10^{-14}
GO:0006091	Generation of precursor metabolites and energy	66	7.95×10^{-14}
GO:0016071	mRNA metabolic process	73	1.12×10^{-13}
GO:0034470	ncRNA processing	47	7.37×10^{-13}
GO:0015985	Energy coupled proton transport, down electrochemical gradient	18	1.61×10^{-09}
GO:0015986	ATP synthesis coupled proton transport	18	1.61×10^{-09}
GO:0042274	Ribosomal small subunit biogenesis	10	4.30×10^{-09}
GO:0000278	Mitotic cell cycle	62	8.82×10^{-09}
Weakly or not expressed CHO DP-12 PMD-genes			
GO:0022610	Biological adhesion	150	3.53×10^{-10}
GO:0007155	Cell adhesion	149	6.01×10^{-10}
GO:0007268	Synaptic transmission	72	2.02×10^{-07}
GO:0019226	Transmission of nerve impulse	81	2.25×10^{-07}
GO:0007610	Behavior	98	1.70×10^{-06}
GO:0007204	Elevation of cytosolic calcium ion concentration	34	2.23×10^{-06}
GO:0055074	Calcium ion homeostasis	49	2.45×10^{-06}
GO:0050877	Neurological system process	212	2.48×10^{-06}
GO:0006874	Cellular calcium ion homeostasis	48	2.61×10^{-06}
GO:0030005	Cellular di-, tri-valent inorganic cation homeostasis	55	5.97×10^{-06}
GO:0006875	Cellular metal ion homeostasis	49	8.43×10^{-06}
GO:0009611	Response to wounding	105	8.63×10^{-06}
GO:0051480	Cytosolic calcium ion homeostasis	34	1.18×10^{-05}
GO:0055082	Cellular chemical homeostasis	80	1.27×10^{-05}
GO:0006873	Cellular ion homeostasis	79	1.28×10^{-05}
GO:0055065	Metal ion homeostasis	50	1.38×10^{-05}
GO:0055066	Di-, tri-valent inorganic cation homeostasis	56	1.39×10^{-05}
GO:0007267	Cell-cell signaling	113	3.90×10^{-05}
GO:0030003	Cellular cation homeostasis	57	4.26×10^{-05}
GO:0044057	Regulation of system process	66	5.03×10^{-05}
GO:0051345	Positive regulation of hydrolase activity	43	8.70×10^{-05}

the highly expressed PMD-genes proved to belong to the same GO classes related to gene expression and RNA processing as generally enriched in CHO DP-12 PMD-localized genes (Tab. 4.4 and Supplementary Tab. B.5). Additionally, genes contributing to energy metabolism ('generation of precursor metabolites and energy', 66 genes) and cell cycle control ('mitotic cell cycle', 62 genes) were found to be significantly enriched in this group. Within the weakly or not expressed PMD-genes those contributing to neurological processes were overrepresented, e.g. including GO terms like 'synaptic transmission' (81 genes) or 'neurological system process' (212 genes; Tab. 4.4 and Supplementary Tab. B.6). Also, genes with functions in cellular adhesion ('biological adhesion', 150 genes) and calcium ion homeostasis ('cellular calcium ion homeostasis', 48 genes) were significantly enriched in this group.

4.1.3.4. Discussion: General characterization of the CHO DP-12 cell DNA methylome

Whole-genome bisulfite sequencing and gene expression microarray analyses showed that the CHO DP-12 cell genome is hypomethylated (61 % CpG methylation) compared to most mammalian cell types like embryonic stem (ES) cells, their derivatives and primary cells that exhibit global methylation levels of >70 % (Ziller et al. 2013). Global hypomethylation has been reported for cancers such as colon tumors and also for long-term cultured cells like IMR90, foreskin fibroblasts and HepG2 cells and proved to be either related to a uniform or a locally concentrated distribution of hypomethylated regions (Ziller et al. 2013; Laurent et al. 2010; Berman et al. 2012; Schroeder and LaSalle 2013).

The analysis of methylation of genomic features showed that CGIs and UTRs in CHO DP-12 cells are less frequently methylated than so-called gene bodies (coding sequences and introns) that are predominantly methylated. This is in accordance with previously described DNA methylomes that show higher levels of DNA methylation in gene-bodies compared to UTRs and was e.g. shown in bovine muscle tissue (Huang et al. 2014). For multiple species a genome-wide negative correlation between promoter methylation and gene expression was described, whereas gene-body methylation has been reported to be positively correlated with gene expression (Deaton and Bird 2011; Jjingo et al. 2012). The meaning of evolutionary conserved gene-body methylation is not completely understood yet, but studies indicate that intragenic DNA methylation modulates alternative RNA splicing (Maunakea et al. 2013). CGIs proved to have the lowest overall methylation levels in CHO DP-12 cells, although a larger fraction of them is methylated compared to human tissues, mouse embryonic stem cells or differentiated mouse cells (Eckhardt et al. 2006; Meissner et al. 2008; Stadler et al. 2011). This finding is in contrast to the concept of CGIs being usually characterized by their lack of DNA methylation, but in accordance with the hypothesis that a great number of CGIs become methylated in permanent cell lines (Antequera, Boyes, and Bird 1990). Cancer cells are also known to exhibit global hypomethylation that accompanies aberrant hypermethylation of CGIs

(Sproul and Meehan 2013). In many of these cases hypermethylation of promoter CGIs can be related to silencing of tumor suppressor genes regulating cell growth, differentiation and apoptosis (Yang and Zheng 2014; Liu et al. 2014). Hu et al., e.g., identified novel candidate tumor suppressor genes in natural killer cell lymphoma (NKCL) by promoter methylation analysis and gene expression profiling, including Bcl-2-like protein 11 (BCL2L11; Hu et al. 2015). Ectopic BCL2L11 expression led to increased apoptosis in natural killer cell lines. A similar approach could be applied to find key regulators of growth and apoptosis in CHO cells useful for cell line engineering.

One study exists that can be used for direct comparison of results from whole-genome bisulfite sequencing studies in CHO cells. In 2016, Feichtinger et al. analyzed six related CHO K1-derived cell lines regarding their genome sequences and DNA methylation landscapes under different culture conditions, in different media and after subcloning (Feichtinger et al. 2016). They showed the numbers of differentially methylated regions between the cell lines to be low in response to passaging and during batch cultivation. High numbers of differentially methylated regions were only found to be high as a consequence of several rounds of subcloning and selection of a certain phenotype. DNA methylation was furthermore found to change significantly in response to media changes. Generally, this study confirmed the results of the one performed here, as all CHO cell lines analyzed by Feichtinger et al. showed a bimodal distribution of either hypo- or hypermethylated CpGs with overall methylation levels of 55 – 65 % and a proportion of CpG dinucleotides with intermediate methylation levels. Differences between CHO K1 and CHO DP-12 cells could be found regarding the proportions of unmethylated and methylated CpGs, as in CHO DP-12 cells 30 % of all CpGs were unmethylated and in CHO K1-derived cells only about 18 % of all CpGs. On the contrary, CHO K1 cells were shown to exhibit higher numbers of methylated CpGs (60 % compared to 30 % methylated CpGs in CHO DP-12). CHO DP-12 cells had, accordingly, a higher proportion of CpG dinucleotides with intermediate methylation levels. Feichtinger et al. did not analyze the overall distribution of DNA methylation across individual genomes in more detail, but rather focused on the comparisons between methylomes. Therefore, no general conclusions regarding the DNA methylation landscapes of the CHO K1-derived cells could be deduced from this study.

4.1.3.5. Discussion: Partially methylated domains in the CHO DP-12 cell DNA methylome

The global characterization of DNA methylation in the CHO DP-12 cell genome indicated a heterogeneous arrangement of hypomethylated regions instead of an even distribution. When methylation levels were plotted along the SNP-corrected CHO K1 scaffolds, large genomic regions with partial methylation were observed. Methylation levels of 54 % on average could

be assigned to these regions which clearly separated them from the neighboring regions with a mean methylation of 77 %. These regions were termed CHO DP-12 PMDs according to the PMDs found in colorectal cancer, breast cancer, fibroblast cultures, SH-SY5Y neuronal cells and placenta (Schroeder et al. 2011; Schroeder et al. 2013; Gaidatzis et al. 2014; Schroeder and LaSalle 2013; Lister et al. 2009; Hon et al. 2012; Berman et al. 2012). PMDs correspond to genomic domains of an average size of 135 kb that cover approximately 40 % of the above-mentioned genomes and have lower levels of DNA methylation than the neighboring genomic regions (Gaidatzis et al. 2014; Schroeder and LaSalle 2013). For CHO DP-12 PMDs a mean size of 64 kb was detected, which is probably due to the current draft status of the CHO K1 genome. A coverage of all analyzed SNP-corrected CHO K1 scaffolds by CHO DP-12 PMDs of 62 % was observed.

Another characteristic of PMDs is transcriptional repression of genes localized within these regions that are generally gene-poor (Schroeder et al. 2013). Accordingly, only 18 % (4,265) of all genes were located within CHO DP-12 PMDs. However, gene expression analysis showed that 30 % (1261) of the CHO DP-12 PMD-localized genes were highly expressed, contradicting the notion of generally repressed gene expression by PMDs. In agreement with the differences in expression, GO analysis of all genes located within CHO DP-12 PMDs showed a significant enrichment in two different functional groups that were related to gene expression on the one hand and to neurological processes on the other hand. When a separate GO analysis was performed for the highly expressed PMD-genes, a significant over-representation was found for those genes with functions in gene expression and, additionally, in energy metabolism and cell cycle, thus representing functions vital for CHO cell growth and productivity. Several reports have shown strong expression of genes related to transcription, translation, chromatin modification and energy metabolism in CHO cell cultures, too (Carlage et al. 2009; Meleady et al. 2011; Nissom et al. 2006).

Weakly or not expressed CHO DP-12 PMD-localized genes were significantly enriched for the functions related to neurological processes such as synaptic transmission and also to cell adhesion and calcium ion homeostasis. It is known that PMDs often contain repressed genes with tissue-specific functions unrelated to the tissue of origin (Schroeder and LaSalle 2013). Neurological processes clearly are unnecessary functions in CHO cells. These processes are strongly related to the cellular calcium homeostasis, as intracellular calcium plays a critical role in neuron development (Gomez and Zheng 2006). It is striking that reduced capacity of cellular adhesion is relevant for adaption of CHO cells to suspension culture, which is a CHO cell property necessary for scalability of industrial processes. It is also strongly related to the conversion of benign tumors to aggressive cancers (Tokuda et al. 2014). Interestingly, PMDs in placenta tissue, SH-SY5Y and IMR90 cells are most significantly enriched in genes contributing to neurological processes, too (Schroeder et al. 2013). This finding is supported

by the fact that methylation patterns in PMDs were shown to be conserved across cell types (Gaidatzis et al. 2014).

4.1.3.6. Conclusion: The first map of a CHO cell DNA methylation landscape

The results presented here represent a first global analysis of the CHO DNA methylation landscape. The CHO DP-12 cell methylome was shown to exhibit general properties that are typical for mammalian cells, although it also turned out to be heterogeneous and characterized by global hypomethylation and hypermethylation of CpG islands. It furthermore shows strong evidence of partially methylated domains. Similar epigenomic features are observed for many cancer types and also for placenta (Schroeder and LaSalle 2013). In fact, placental trophoblastic and malignant cells share features such as high proliferation or their lack of cell-contact inhibition and it was shown that a subset of malignant tumors activate ectopic expression of germline and placental genes (Ferretti et al. 2006; Rousseaux et al. 2013). CHO cells are, similarly to cancer and placental trophoblast cells, characterized by their proliferative properties and deregulated apoptotic pathways. Here, it was additionally shown that CHO DP-12 cells not only exhibit epigenomic features similar to cancer and placenta cells, but that there is evidence of these features being involved in the repression of tissue-specific genes, as it was shown for placenta PMDs, too. Interestingly, the repressive effect of CHO DP-12 PMDs does not affect genes relevant to CHO cell properties.

To fully understand this relationship, it still needs to be elucidated how PMDs evolved, if they are generated in a dynamic manner and what they look like on the cellular level. A comparison of different CHO cell lines with respect to the presence and location of PMDs is also required. However, these findings show that DNA methylation might significantly contribute to CHO phenotypes and allow for a deeper understanding of CHO cell properties. As a consequence, future experiments to identify targets for rational design of CHO cell lines with improved production characteristics should not only address gene expression changes, but also the underlying epigenetic patterns. The presence of PMDs in the CHO DP-12 cell genome for example underlines the necessity of site-specific integration of recombinant genes in order to avoid potentially unfavorable epigenetic environments leading to low expression levels.

4.1.4. DNA methylation dynamics and differential gene expression in response to butyrate

In the second part of the integrative analysis of DNA methylation and gene expression in CHO DP-12 cells, the butyrate effect on DNA methylation dynamics and differential gene expression was analyzed in a time-course study. Butyrate is known to enhance cell specific

productivity and was shown to influence epigenetic processes both on the levels of histones and DNA methylation (Choi et al. 2005; Jiang and Sharfstein 2008; Klausning, Krämer, and Noll 2011; Sung et al. 2004; Kucharski et al. 2008; Donohoe and Bultman 2012). However, the positive effect of butyrate addition on CHO cells is also accompanied by detrimental effects such as growth inhibition and apoptosis. It was furthermore reported that the quality and bioactivity of glycoproteins can be influenced by butyrate addition (Hong et al. 2014; Lee et al. 2014). Knowledge of the mechanism underlying the butyrate effect promises to enable rational cell line development in terms of employing only the positive aspects. For this reason, several studies aimed to analyze the effect of butyrate addition on the CHO cell transcriptome and proteome. It was for example shown that butyrate affected histone modifications, lipid metabolism and protein processing as well as the expression of chaperone genes (De Leon Gatti et al. 2007). Additionally, Kantardjieff et al. reported that butyrate affected expression of genes related to protein secretion and signaling (Kantardjieff et al. 2010). Klausning et al. could show that genes related to carbohydrate, lipid, amino acid and glycan metabolism were upregulated upon butyrate addition, whereas genes associated with cell growth and apoptosis were downregulated (Klausning, Krämer, and Noll 2011). It was furthermore shown that butyrate increased the expression of genes involved in protein secretion and redox activity and decreased the expression of genes related to cell cycle and apoptosis (Yee et al. 2008).

Here, genome-wide butyrate-induced changes of DNA methylation and gene expression were analyzed. DNA methylation and gene expression data were integrated regarding functional annotations of differentially methylated and expressed genes and potential correlations of DNA methylation and gene expression profiles over the time course of sampling. A *de novo* sequence motif discovery was applied furthermore to assess an involvement of specific transcription factors.

4.1.4.1. Gene expression changes upon butyrate addition

Samples for integrative DNA methylation and gene expression analysis were taken from CHO DP-12 batch cultivations prior to butyrate addition as well as 6, 12, 24 and 36 hours post butyrate treatment (please refer to section 4.1.1 for details). The replicate samples taken before butyrate addition served as the reference for the analysis of differential DNA methylation and gene expression upon butyrate addition. RNAs extracted from these samples were checked for integrity (Supplementary Fig. B.18) and pooled before cDNA was generated by reverse transcription. 1.65 µg of each cDNA was used for microarray hybridization. Four DNA-microarrays were hybridized for the reference and each 3 DNA-microarrays for the time points 6, 12, 24 and 36 hours post butyrate treatment (see quality plots in Supplementary Fig. B.19). Data analysis was performed using the *limma* package in *R*. In a first step, the signal ranges before normalization were compared between the arrays, which proved to be very

similar even before further processing. Data were then background corrected and quantile normalized (Supplementary Fig. B.20). For each comparison between samples of interest, statistical parameters for differential expression were calculated. Genes with adjusted p values below 0.01 were considered to be significantly differentially expressed.

To get a first impression of differential gene expression, the top 10 up- and downregulated genes were examined regarding their functions and the extent of differential expression. Table 4.5 shows each of the top 10 up- and downregulated genes 36 hours post butyrate addition with indication of M-values for all four sampling points. M-values represent the \log_2 -ratio of the experimental sample to the reference. Accordingly, M-values of 1.0 indicate a two-fold increase in gene expression in the experimental sample relative to the reference culture (fold change 2). An M-value of -1.0 indicates a decrease in gene expression of 50 % in the experimental sample relative to the reference culture (fold change 0.5). The table also shows key protein functions retrieved from the *UniProt* database (The UniProt Consortium 2014).

Table 4.5.: Expression time courses of the top 10 up- and downregulated genes 36 hours post butyrate addition ($p \leq 0.01$). M-values ≤ -1.0 are associated with downregulated genes. M-values ≥ 1.0 are associated with upregulated genes. Key protein functions were retrieved from the UniProt database (The UniProt Consortium 2014).

Abbr.	Name	Gene expression [M-value]				Function
		6 h	12 h	24 h	36 h	
Downregulated genes						
PRRX1	Paired mesoderm homeobox protein 1	-0.77	-2.07	-2.52	-2.34	Development
EMPL	Epithelial membrane protein 1	-0.58	-1.77	-2.38	-2.24	Apoptosis, growth
IL1RL1	Interleukin-1 receptor-like 1	-0.96	-1.76	-1.99	-2.11	Receptor, growth
PTH LH	Parathyroid hormone-related protein	-1.99	-2.12	-2.11	-1.97	Growth, development
FBXO4	F-box only protein 4	-0.84	-1.95	-2.32	-1.94	Ubiquitination
NUPR1	Nuclear protein 1	0.00	-0.66	-1.53	-1.86	Activator, stress
ACSL5	Long-chain-fatty-acid-CoA ligase 5	-0.48	-1.53	-2.04	-1.79	Lipid metabolism
CFLAR	CASP8 and FADD-like apoptosis regulator	-0.53	-0.65	-1.32	-1.77	Inhibitor of apoptosis
COL4A1	Collagen alpha-1(IV) chain	0.00	-1.33	-2.13	-1.55	Angiogenesis, growth
PLK1	Serine/threonine-protein kinase PLK1	-0.89	-0.84	-1.38	-1.54	Cell cycle
Upregulated genes						
GLYAT	Glycine N-acyltransferase-like protein	0.00	0.50	1.68	2.80	Detoxification, acyl-CoA metabolism
MYL9	Myosin regulatory light polypeptide 9	0.00	1.01	2.09	2.59	Motor protein
ULBP1	NKG2D ligand 1	0.00	0.61	1.62	2.29	Immune signaling
HMGCS2	Hydroxymethylglutaryl-CoA synthase	0.00	0.00	1.31	2.19	Lipid biosynthesis
TXNIP	Thioredoxin-interacting protein	2.47	2.68	2.42	1.84	Repressor, cell cycle
CCPG1	Cell cycle progression protein 1	1.24	1.63	1.81	1.80	Cell cycle
CTNNA2	Catenin alpha-2	0.00	0.60	1.55	1.79	Development
LOXL2	Lysyl oxidase homolog 2	0.00	0.00	0.70	1.77	Chromatin regulator
ATOX1	Copper transport protein ATOX1	0.45	0.95	1.54	1.76	Ion transport
DGAT2	Diacylglycerol O-acyltransferase 2	0.00	0.81	1.27	1.70	Lipid biosynthesis

The strongest downregulation was detected for Paired mesoderm homeobox protein 1 (PRRX1) with an M-value of -2.34 (fold change 0.19 relative to the reference) 36 hours upon butyrate

addition. Six hours post butyrate addition PRRX1 showed downregulation by $M = -0.77$ (fold change 0.59), i.e. PRRX1 was increasingly downregulated over the course of sampling. PRRX1 serves as a co-activator in developmental processes. As another example, Glycine N-acyltransferase-like protein (GLYAT) was found to be the gene with the highest M-value of 2.80 (fold change 6.9 relative to the reference) 36 hours upon butyrate addition. Upregulation of GLYAT was first detected 12 hours post butyrate addition with $M = 0.5$ (fold change 1.4). Strikingly, GLYAT is a mitochondrial acyltransferase and contributes to detoxification processes and acyl-CoA-metabolism. Similar expression profiles were found for the other genes, with maximum M-values of means of -2.0 and 2.0, respectively. Downregulated genes had functions in growth, apoptosis and cell cycle control, whereas upregulated genes had functions in signaling, lipid metabolism, cell cycle and chromatin modification.

In general, 2,724, 5,346, 6,610 and 7,672 genes were detected to be differentially expressed upon butyrate addition at time points 6, 12, 24 and 36 hours (Fig. 4.11A). The Venn diagrams in Fig. 4.11B show that most of the butyrate-responsive genes were expressed at more than one time point with a majority being expressed at the '36 h' sampling point. When expression profiles were analyzed by hierarchical cluster analysis, most genes indeed turned out to be continuously up or downregulated over the time course of sampling (Fig. 4.11C).

Using k-means clustering, expression profiles could be grouped into 4 distinct clusters (Fig. 4.11D). Cluster I contained 929 genes that were upregulated during the first 12 hours of butyrate treatment by a mean of two-fold ($M = 1.0$). In contrast, cluster II included 1014 genes that were downregulated during the first 12 hours by a mean of 0.5-fold ($M = -1.0$). Gene Ontology annotations showed that the upregulated genes in cluster I were most significantly related to energy generation (e.g. 33 genes contributing to 'generation of precursor metabolites and energy'), transport (e.g. 'vesicle-mediated transport', 53 genes) and response reactions (e.g. 'response to endogenous stimulus', 37 genes; see complete GO annotation in Supplementary Tab. B.10). The downregulated genes in cluster II were most significantly related to cell cycle regulation (e.g. 'cell cycle', 106 genes), RNA processing (e.g. 'RNA splicing', 43 genes) and chromatin modification (e.g. 'chromosome organization', 71 genes; Supplementary Tab. B.11). 2861 genes in cluster III and 3039 genes in cluster IV showed expression time courses that continuously, but only slightly, increased or decreased, with minimum and maximum mean values of differential expression of 1.2-fold ($M = 0.25$) and 0.8-fold ($M = -0.25$). The continuously but weakly upregulated genes were most significantly related to protein transport and protein catabolism (e.g. 'protein localization', 227 genes and 'proteolysis involved in cellular protein catabolic process', 128 genes) as well as protein synthesis (e.g. 'translational elongation', 52 genes; Supplementary Tab. B.12). The downregulated genes were most significantly related to RNA metabolism and splicing (e.g. 'RNA processing', 196 genes), cell cycle regulation (e.g. 'mitotic cell cycle', 132 genes) and gene expression (e.g. 'transcription', 425 genes; Supplementary Tab. B.13).

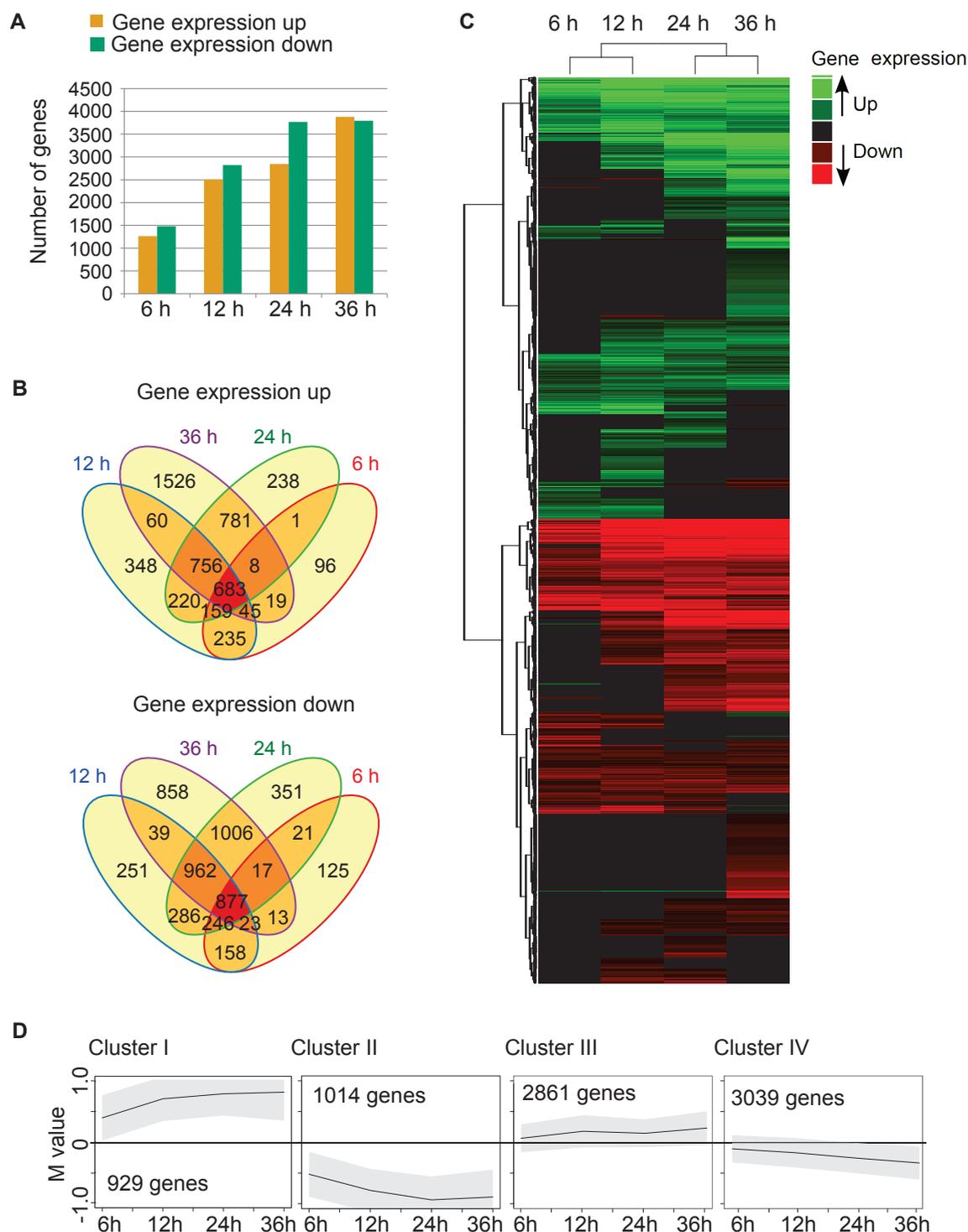


Figure 4.11.: Analysis of differential gene expression data 6 to 36 hours after butyrate addition (modified from Wippermann et al. 2016). **A** Numbers of differentially expressed genes. **B** Un-weighted Venn diagrams of genes showing up- and downregulated expression 6 hours to 36 hours after butyrate addition. **C** Hierarchical clustering of differential gene expression data. A heatmap of gene expression data (fold changes color-coded according to the legend) was generated using the *ComplexHeatmap* package in *R*. Cluster analysis based was performed using euclidean clustering distance and Ward's method for agglomeration (Ward 2012). **D** Time courses of gene expression changes. By k-means clustering 4 distinct general gene expression profiles were identified. Standard deviations are represented in grey.

When the changes in cellular pathways over time were examined by *Generally Applicable Gene-set Enrichment (GAGE; Luo et al. 2009)* analysis (Fig. 4.12) in order to discriminate early response reactions from potential secondary ones, genes related to the processes 'protein amino acid acetylation', 'regulation of transcription', 'regulation of DNA replication' and 'chromatin remodeling' as well as processes related to differentiation (e.g. 'skeletal system development') were found to be specifically downregulated at the first sampling point.

4.1.4.2. Mapping of gene expression changes to cellular pathways

To get a detailed view on the time course of differential gene expression upon butyrate treatment with regard to affected cellular pathways, significantly differentially expressed genes were mapped to *KEGG* pathways (Kanehisa and Goto 2000) based on M-values. Mapping was performed using the Bioconductor *pathview* package (Luo and Brouwer 2013) that allows for visualization of multiple values for a single gene, i.e. enables representation of time course data on *KEGG* pathway maps. According to the GO analysis, the pathways 'cell cycle' and 'apoptosis' were examined in detail. As butyrate was shown to have a major impact on cellular energy generation and, importantly, a butyrate-induced switch from lactate production to consumption could be observed during the experiment, the pathways 'citrate cycle' and 'glycolysis' were also examined. The pathway 'spliceosome' was examined more closely, as butyrate addition affected several genes related to RNA processing. Additionally, the impact of butyrate addition on the 'N-glycan biosynthesis' pathway was analyzed to account for the reports of changes in product glycosylation (see e.g. Hong et al. 2014). The pathways 'RNA transport', 'Aminoacyl tRNA biosynthesis', 'Ribosome biogenesis' and 'Ribosome' were examined in detail in order to assess a potential impact of these processes on recombinant protein production.

Impact of butyrate on cell cycle regulation 36 hours after butyrate addition a large part of the genes involved in cell cycle regulation showed downregulated expression (Fig. 4.13). These included key regulatory elements such as Proliferating cell nuclear antigen (PCNA) or components of the Mini chromosome maintenance (MCM) complex. PCNA regulates DNA replication and cell cycle by interacting with several proteins. It functions e.g. as a platform for DNA polymerase during DNA replication and interaction of PCNA with the p21 protein stops DNA synthesis. Therefore, PCNA expression is correlated with cell proliferation (Strzalka and Ziemienowicz 2011). The structurally related and highly conserved MCM proteins 2 to 7, which form a hexameric complex that is recruited to replication origins during G1 phase and functions as a helicase, are also associated with cell proliferation (Nosedá and Karsan 2006).

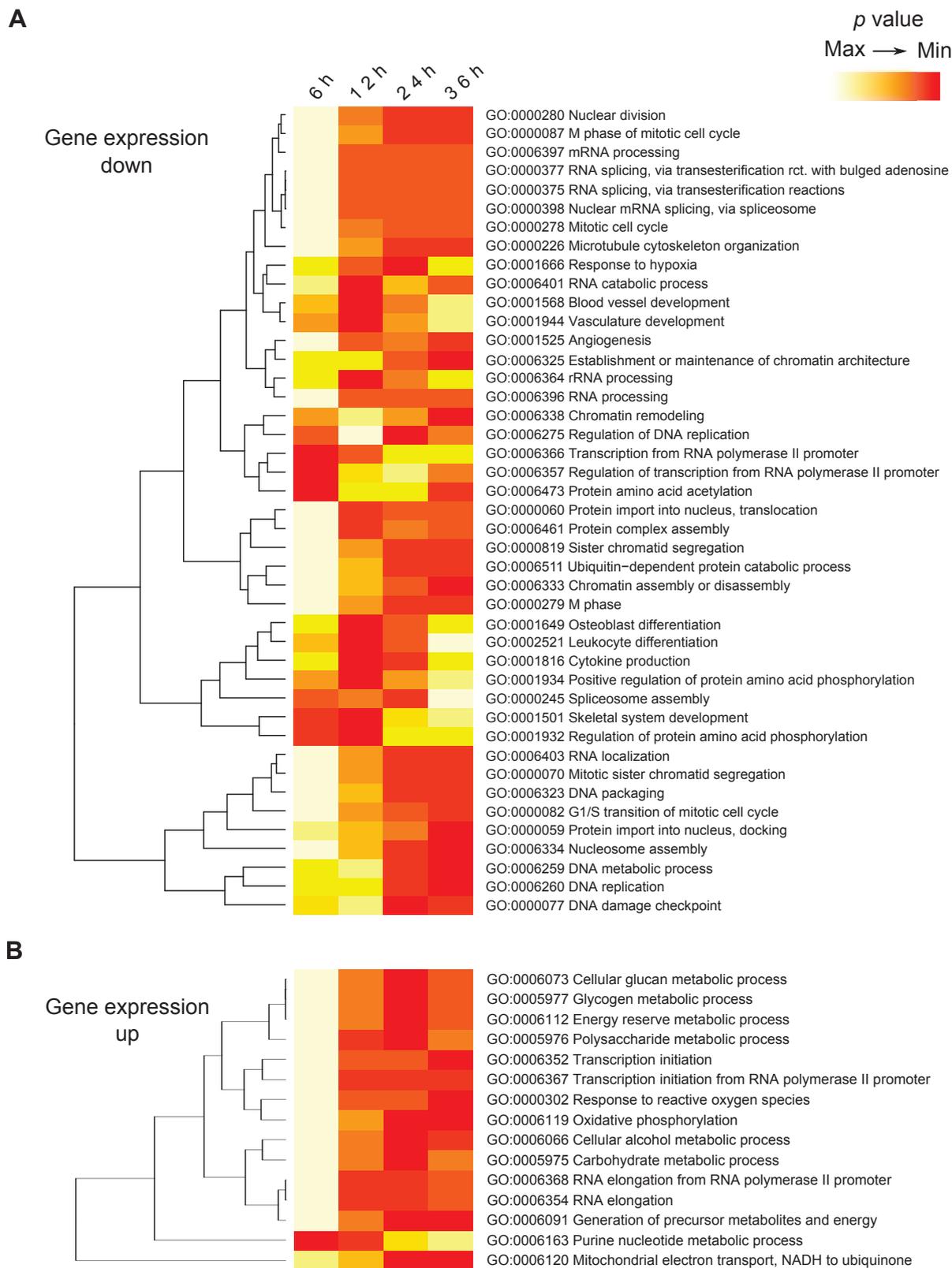


Figure 4.12.: *Generally Applicable Gene-set Enrichment (GAGE, Luo et al. 2009)* analysis of significantly differentially expressed genes after butyrate addition (Wippermann et al. 2016). Heatmaps depict either significantly down- or upregulated cellular pathways (based on Gene Ontology (GO) terms) over the time course of sampling.

Expression time courses differed, as some genes were already downregulated 6 hours post butyrate addition (e.g. PCNA) and for some genes downregulation was detected at later time points (e.g. at the 36 hours sampling point for the MCM complex components). Some central regulators of the pathway were furthermore upregulated, e.g. MYC, which showed upregulated expression 6 to 12 hours post butyrate. MYC is associated with cell proliferation as well, as it induces several positive cell cycle regulators and inhibits the activity of inhibitory factors such as p21 (Bretones, Delgado, and León 2015). Further genes involved in the cell cycle were GADD45A and GADD45B, which were found to be upregulated at all sampling points. GADD45 expression has been shown to be induced by DNA damage and in response to cellular stress associated with apoptosis and growth arrest. GADD45 action is mediated via MAPK signaling (Salvador, Brown-Clay, and Fornace 2013).

For some genes differences in up- and downregulation of subunits or isoforms were detected, e.g. including Anaphase-promoting complex (APC/C). ANAPC4 and ANAPC5 were downregulated after butyrate addition, whereas ANAPC11 and ANAPC13 were upregulated. The APC/C complex regulates cell cycle transitions by controlling proteolysis of cell cycle proteins and is regulated by specific cell cycle checkpoints (Chang and Barford 2014). Similarly, Mitotic spindle assembly checkpoint proteins MAD2A and MAD2B (MAD2L1 and MAD2L2) were down- and upregulated, respectively.

As cell cycle regulation and WNT signaling are closely linked, differential gene expression was analyzed for this pathway as well, which was found to exhibit both up- and downregulated genes (Supplementary Fig. B.31).

Impact of butyrate on the induction of apoptotic pathways Several genes encoding major regulators of apoptosis were upregulated upon butyrate addition, most of them already at the first sampling point or 12 hours after butyrate addition at the latest (Fig. 4.14). These genes included Caspase-3, Caspase-6 and Caspase-9 (CASP3, CASP6, CASP9), Apoptotic protease-activating factor 1 (APAF-1) and Apoptosis regulator Bax (BAX). Caspases (= cysteine-aspartic proteases) are proteases that are involved in cell death mediated by apoptosis and also in non-apoptotic functions such as proliferation and differentiation. Within the intrinsic pathway of apoptosis induction which is activated by cellular stress, the pro-apoptotic effectors BAX and BAK induce the release of cytochrome c. APAF-1 then forms a complex with cytochrome c and activates CASP9, which is a so-called initiator caspase that activates effector caspases such as CASP3 and CASP6 (Shalini et al. 2015).

Further pro-apoptotic genes found to be upregulated upon butyrate addition were Bcl-2-like protein 11 (BIM/BCL2L11) and C-Jun N-terminal kinase (JNK1/MAPK8). BIM was shown to induce apoptosis in response to stimuli such as growth factor deprivation. It induces conformational changes in BAX and thereby leads to its activation. It furthermore uncouples

mitochondrial respiration and leads to the formation of reactive oxygen species. BIM was also shown to affect cytoskeletal integrity (Sionov, Vlahopoulos, and Granot 2015). JNK1 also mediates cellular response reactions to stresses and is involved in apoptosis, but also in processes such as proliferation and differentiation. JNKs in general belong to the Mitogen-activated protein kinase (MAPK) family of proteins and are activated by the MAP2K proteins MKK4 or MKK7, which are in turn activated by several MAP3K proteins that react to specific signals. JNK proteins then upregulate pro-apoptotic genes by influencing transcription factors or by direct phosphorylation events that impact the activity of further proteins (Dhanasekaran and Reddy 2008).

Interestingly, Caspase-7, Caspase-8 and Caspase-12 (CASP7, CASP8 and CASP12) were downregulated at all sampling points following butyrate addition. CASP8 represents the initiator caspase of the extrinsic apoptotic pathway, which depends on ligand binding to so-called death-receptors (McIlwain, Berger, and Mak 2013). While CASP7 belongs, like CASP3 and CASP6, to the group of initiator caspases, CASP12 was shown to be involved in apoptotic events due to ER stress which can be induced by chemicals (Lamkanfi, Kalai, and Vandenabeele 2004).

According to the crosstalk of apoptotic processes with several signaling pathways, differential gene expression was also found for the MAPK signaling pathway (largely upregulated, Supplementary Fig. B.32), the calcium signaling pathway (both up- and downregulated components, Supplementary Fig. B.33) and the p53 signaling pathway (both up- and downregulated components as well, Supplementary Fig. B.34).

Butyrate-dependent changes in central energy metabolism Several genes encoding major regulatory enzymes involved in glycolysis and the TCA cycle were upregulated upon butyrate addition (Supplementary Fig. B.27 and 4.15). Interestingly, upregulation of gene expression occurred earlier than downregulation, with several genes being not differentially upregulated anymore 36 hours after butyrate addition. On the contrary, downregulation of gene expression happened later and most genes were still found to be downregulated 36 hours after butyrate addition.

Genes involved in glycolysis and gluconeogenesis that were upregulated upon butyrate addition comprised e.g. ADP-dependent glucokinase (ADPGK) and Acetyl-CoA synthetase (ACSS2). Also, expression of key players involved in the citric acid cycle was upregulated, including Pyruvate carboxylase (PC), ATP-citrate synthase (ACLY), Isocitrate dehydrogenase (IDH1) and Malate dehydrogenase (MDH1 and MDH2). Genes of the glycolysis pathway that were downregulated at later sampling points included Hexokinase-1 (HK1) and Glyceraldehyde-3-phosphate dehydrogenase (GAPDH). The only gene that was already downregulated 6 hours after butyrate addition was Lactate dehydrogenase (LDHA). Expression of

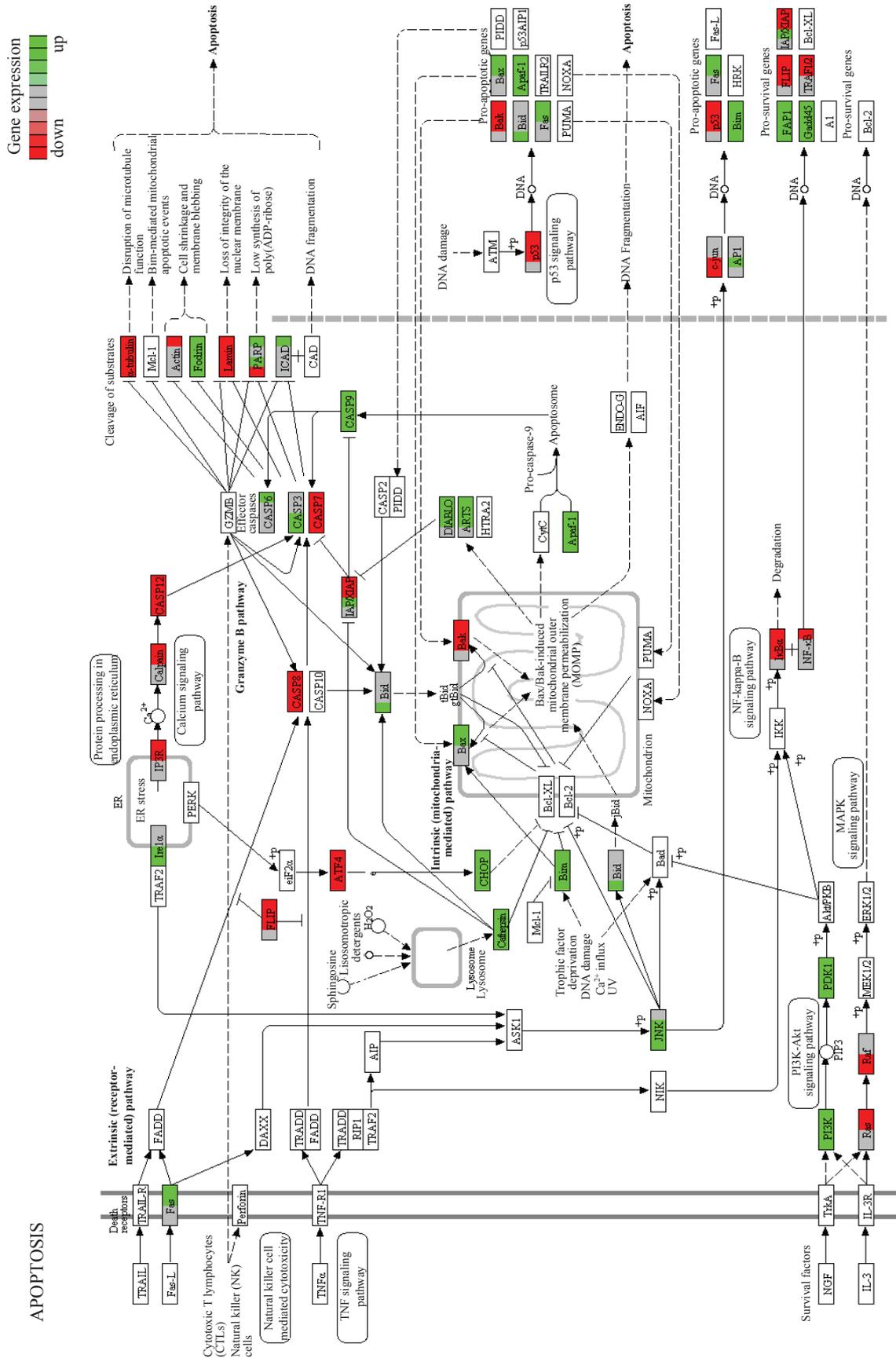


Figure 4.14.: Visualization of gene expression time course after butyrate addition for KEGG pathway Apoptosis (Kanehisa and Goto 2000). Pathway analysis was performed using *pathview* (Luo and Brouwer 2013). Gene expression changes are color-coded according to the legend. Please refer to legend of Fig. 4.13 for a more detailed description.

TCA enzymes including Aconitate hydratase (ACO) and Succinyl-CoA ligase (SUCLG1 and SUCLG2) was found to be downregulated 24 hours and/or 36 hours after butyrate addition as well.

Only minor changes in gene expression within the pentose phosphate pathway were observed. However, both purine and pyrimidine metabolism pathways were strongly affected upon butyrate addition (Supplementary Figs. B.35 and B.36).

The effect of butyrate addition on splicing For a majority of genes involved in splicing (Fig. 4.16) a similar behavior was observed, as most of the genes belonging to the small nuclear ribonucleoprotein particles (snRNPs) U1, U2 and U4 to U6 were downregulated. Differences were observed in the times regulation set in. In contrast, components of the PRP19 complex and PRP19 related genes showed either up- or downregulated expression. Briefly, splicing begins with assembly of the small nuclear RNAs with the pre-mRNA, recognition of the 5' splice and branch sites and formation of the active splicing complex, which involves association of PRP19 and related proteins (the NineTeen Complex, NTC). After several rearrangements the catalytically active complex carries out the splicing reaction, which includes the attack at the 5' splice site by the branch site adenosine and the attack of the 3' splice site by the 5' exon, followed by exon joining and release of an intron-snRNP complex. During this splicing cycle, modulation of the NTC and associated proteins are required for spliceosome confirmations during splicing and disassembly (Almeida and O'Keefe 2015). PRP19 also is involved in mediating cell survival after DNA damage and loss of PRP19 was shown to result in accumulation of DNA double-strand breaks (Mahajan and Mitchell 2003). Interestingly, an interaction of PRP19 complex components with the nucleotide excision repair machinery and, additionally, with PCNA was found (Chanarat and Sträßer 2013; Chu et al. 2006). Strikingly, the base excision repair pathway is linked to the regeneration of unmodified cytosine during TET-mediated demethylation and DNMT1-PCNA-interaction is essential for DNA maintenance methylation. As PCNA was also found to be downregulated, this finding hints towards a complex system of butyrate-responsive cellular pathways.

Butyrate addition leads to changes in expression of genes relevant to glycosylation Butyrate was shown to be capable of influencing the quality and bioactivity of glycoproteins. For example, Hong et al. showed that butyrate treatment resulted in changed galactosylation of monoclonal antibodies (Hong et al. 2014), while Lee et al. showed that butyrate treatment led to a reduction of the relative proportion of the sialic acid content of a recombinant glycoprotein that was accompanied by altered expression of several genes related to glycosylation, such as UDP-glucose pyrophosphorylase or Sialidase-2 (Lee et al. 2014). For this reason the gene expression time course upon butyrate addition in CHO DP-12 cells for genes relevant for N-glycan biosynthesis was examined in more detail (Fig. 4.17).

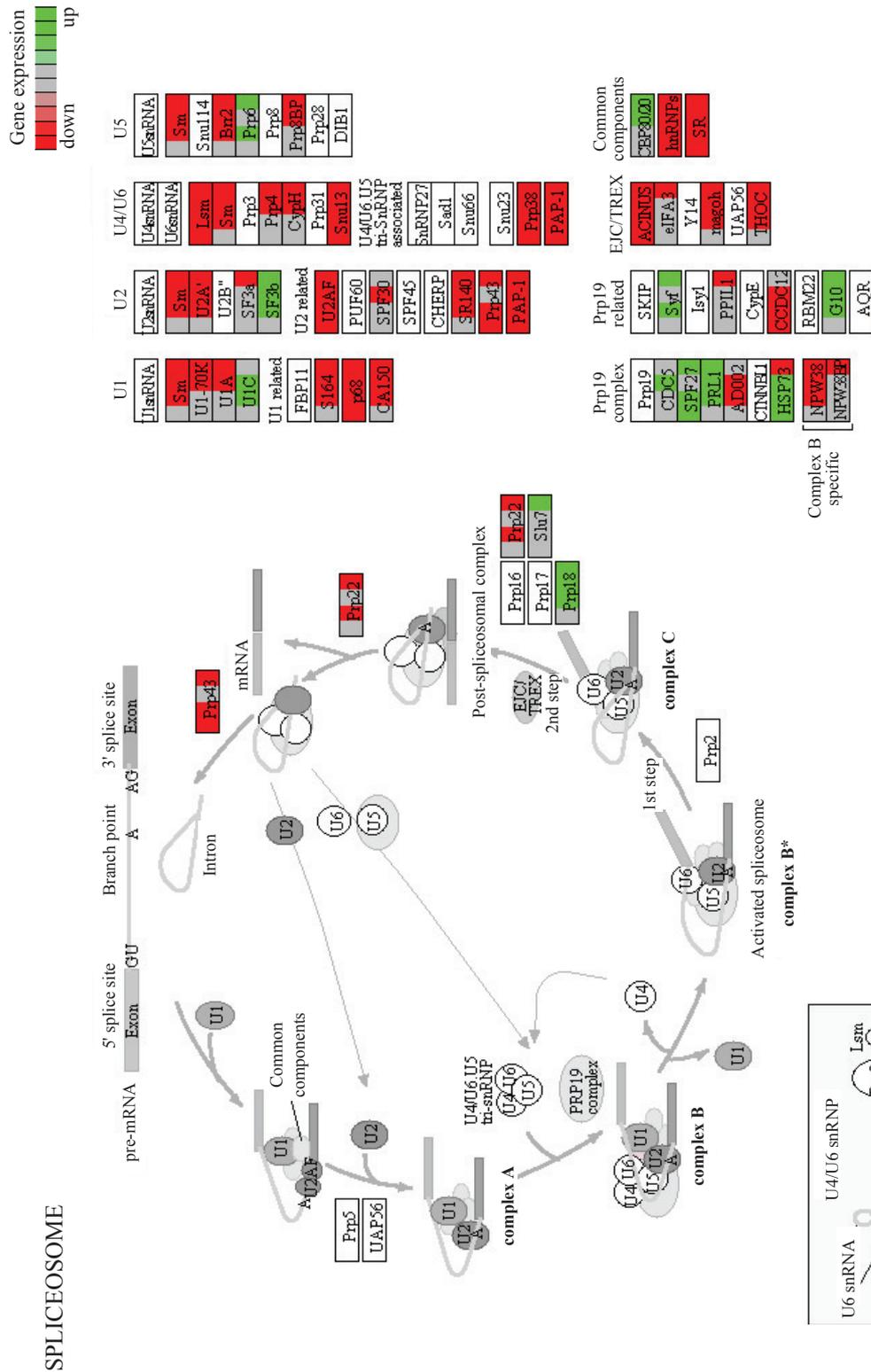


Figure 4.16.: Visualization of gene expression time course after butyrate addition for *KEGG* pathway Spliceosome (Kanehisa and Goto 2000). Pathway analysis and visualization of gene expression changes based on M-values was performed using *pathview* (Luo and Brouwer 2013). Gene expression changes are color-coded according to the legend. Each node is divided into four parts representing differential expression at sampling points '6 h' to '36 h'. For nodes representing multiple proteins (isoforms, subunits etc.) maximal absolute values are used as the summary/representative of a node.

Generally, N-glycosylation describes the attachment of glycans to polypeptides by an amide linkage to asparagine (Asn) side chains (Moremen, Tiemeyer, and Nairn 2012). ALG glycosyltransferases catalyze the addition of monosaccharides to an N-glycan precursor, which is attached to the nascent polypeptide chain during translation by the oligosaccharyltransferase (OST) complex (Kelleher and Gilmore 2006). In the ER, glycoproteins are trimmed by glucosidases. This process is important for ER-retention of such proteins until they are properly folded and part of the quality control process. If folding fails nevertheless, glycoproteins are targeted to ER-associated degradation (ERAD) by trimming of α 1,2-linked mannose residues (Lederkremer 2009). Glycoproteins that pass ER quality control are further processed to a mature state by trafficking through the Golgi apparatus, resulting in distinct glycan profiles (Moremen, Tiemeyer, and Nairn 2012).

Here, five genes of the N-glycosylation pathway were found to be upregulated after butyrate addition, including Alpha-1,3/1,6-mannosyltransferase (ALG2), Beta-1,4-N-acetylglucosaminyltransferase (ALG14), Dolichyl-phosphate mannosyltransferase polypeptide 3 (DPM3), Dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit 1 (OST) and Mannosyl-oligosaccharide 1,2-alpha-mannosidase IA (MAN1). On the contrary, Alpha-1,3-mannosyltransferase (ALG3), Alpha-1,3-glucosyltransferase (ALG8), Alpha-1,2-glucosyltransferase (ALG10), Beta-1,4-N-acetylglucosaminyltransferase (ALG13) and Alpha-mannosidase II (MAN2) were found to be downregulated post butyrate addition. Taken together, butyrate-induced differential expression of genes related to N-glycosylation affected several steps of the pathway, including the initial generation of the glycan precursor by ALGs, its attachment to the nascent polypeptide by the OST complex, as well as the trimming process in the ER and Golgi. Based on this, variations in glycosylation patterns of recombinant proteins as described elsewhere could be expected.

Impact of butyrate on protein synthesis The productivity enhancing effect of butyrate on CHO cells expressing recombinant proteins was shown in several cases (Palermo et al. 1991; Jiang and Sharfstein 2008; Klausning, Krämer, and Noll 2011; Sung et al. 2004; Choi et al. 2005). However, the reason for this phenomenon is not entirely clear. As it was hypothesized that butyrate might lead to increased productivity by eliciting an increase in cellular capacities for translation, the gene expression time course data were analyzed regarding changes in the pathways 'RNA transport' (Supplementary Fig. B.28), 'Aminoacyl tRNA biosynthesis' (data not shown), 'Ribosome biogenesis' (Fig. 4.18) and 'Ribosome' (Supplementary Fig. B.29). No significant changes within the pathway 'Aminoacyl tRNA biosynthesis' were found, but RNA transport, ribosome biogenesis and ribosomal subunit expression were affected by butyrate addition.

Ribosome biogenesis is considered one of the cellular processes that demand the largest amount of energy and involves the action of many transiently associated proteins. Impor-

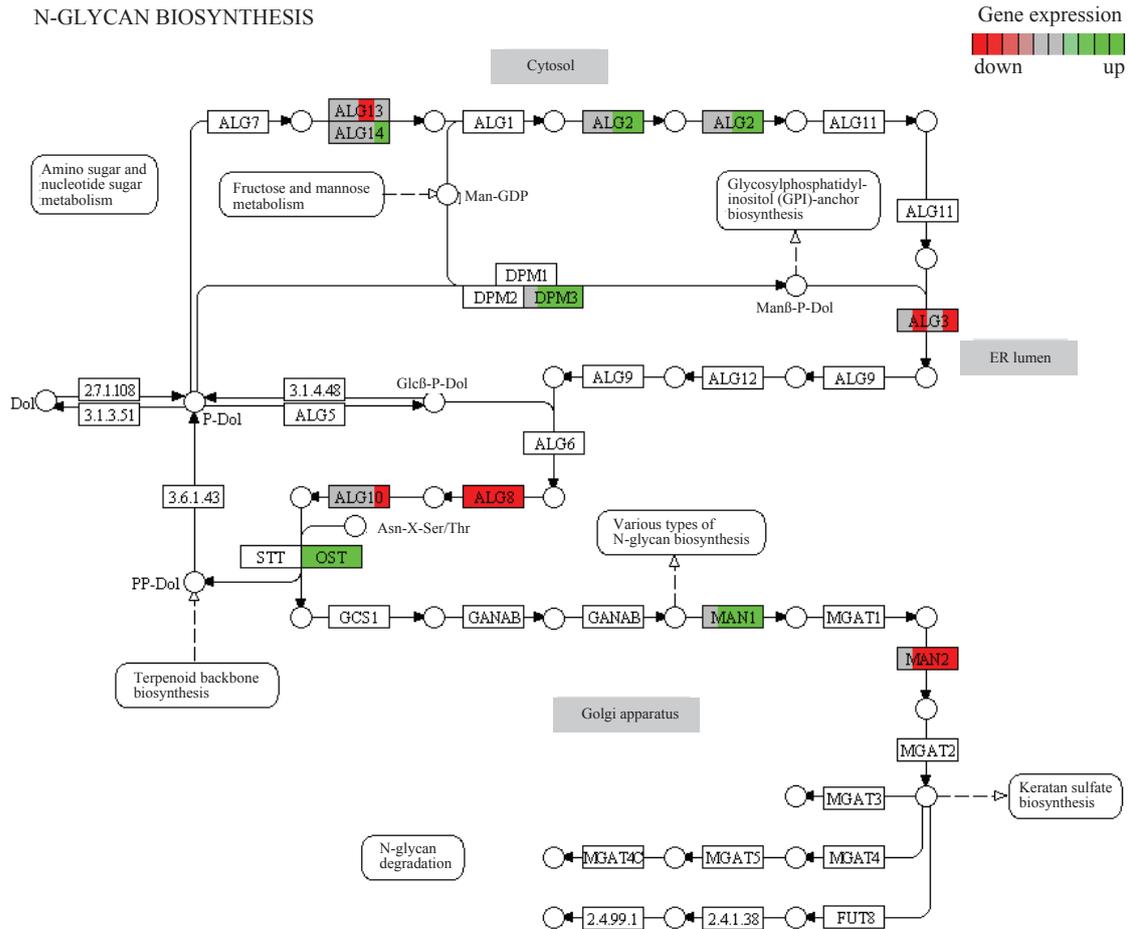


Figure 4.17.: Visualization of gene expression time course after butyrate addition for *KEGG* pathway N-glycan biosynthesis (Kanehisa and Goto 2000). Pathway analysis and visualization of gene expression changes based on M-values was performed using *pathview* (Luo and Brouwer 2013). Gene expression changes are color-coded according to the legend. Each node is divided into four parts representing differential expression at sampling points '6 h' to '36 h'. For nodes representing multiple proteins (isoforms, subunits etc.) maximal absolute values are used as the summary/representative of a node.

tantly, the process of ribosome synthesis is linked to other central processes such as growth and cell division (Thomson, Ferreira-Cerca, and Hurt 2013). It begins in the nucleolus by co-transcription of rRNA genes through RNA polymerase I. Subsequently, several small nucleolar ribonucleoproteins (snoRNPs) mediate co-transcriptional modification of the nascent transcript, which forms ball-like structures at its 5' end that are considered to represent the earliest pre-ribosomes (Grandi et al. 2002). Subsequently, the rRNA transcript is cleaved co-transcriptionally and the 90S pre-ribosome is separated into the pre-40S and pre-60S particles. The pre-40S particles are then exported to the cytoplasm for maturation, whereas maturation of the 60S is more complex and is mediated by several proteins within the nucleolus and nucleus before 60S particles are exported as well (Thomson, Ferreira-Cerca, and Hurt 2013).

Butyrate addition induced downregulation of several components of the pathway, including for example GTP-binding nuclear protein RAN and Exportin-1 (CRM1), which are responsible for the export of ribosomal subunits (Thomas and Kutay 2003). As another example, Nucleolar protein 56 (NOP56), Nucleolar protein 58 (NOP58) and NHP2-like protein 1 (SNU13) were downregulated. These proteins are essential for pre-rRNA processing (Gautier et al. 1997). Upregulation was for example detected for RNA 3'-terminal phosphate cyclase-like protein (RCL1), which was shown to associate with U3 snoRNP and to be required for 18S rRNA biogenesis (Billy et al. 2000).

Taken together, the pathway exhibited mostly downregulated genes. A similar overall downregulation was found for the pathway 'RNA transport'. Several genes encoding ribosomal subunits showed differential expression as well, with similar proportions of up- and down-regulated genes. However, in summary these results suggest that the cellular capacities for protein synthesis were limited instead of amplified upon butyrate addition, which contradicted the hypothesis of an increase in gene expression potential on the level of translation.

Interestingly though, three genes with upregulated expression over the entire time course of sampling were detected that contribute to protein export, namely Heat shock 70kDa protein 5 (BIP), Signal recognition particle subunit SRP54 and Signal peptidase SEC11. The ER luminal chaperone BIP binds newly synthesized proteins to ensure proper folding, mediates export of proteins destined for degradation and BIP expression was shown to increase in response to the accumulation of unfolded proteins in the ER (Gething 1999). Strikingly, BIP is known to be involved in the expression of monoclonal antibodies and increased productivity was shown to be associated with increased BiP levels (Meleady 2007).

4.1.4.3. Comparison of DNA-microarray results with published gene expression analyses in butyrate-treated CHO cells

In order to be able to estimate the reliability of the DNA-microarray experiment it was compared to published data generated under similar conditions. Generally, microarray analyses of gene expression changes depend on several factors such as the microarray type, thresholds set for classification of genes as significantly differentially expressed and the numbers of genes analyzed by the microarray. Other factors that make such comparisons even more complex are, naturally, differences in cell types, cultivation parameters, sampling points and various additional experimental conditions. Furthermore, most publications do not contain complete lists of M-values for the entire group of differentially expressed genes. Therefore, comparisons of microarray experiments need to be carefully assessed. Keeping this in mind, studies that aimed to analyze butyrate-induced gene expression changes by microarrays in cultured cells were compared regarding the extent to which butyrate was found

RIBOSOME BIOGENESIS IN EUKARYOTES

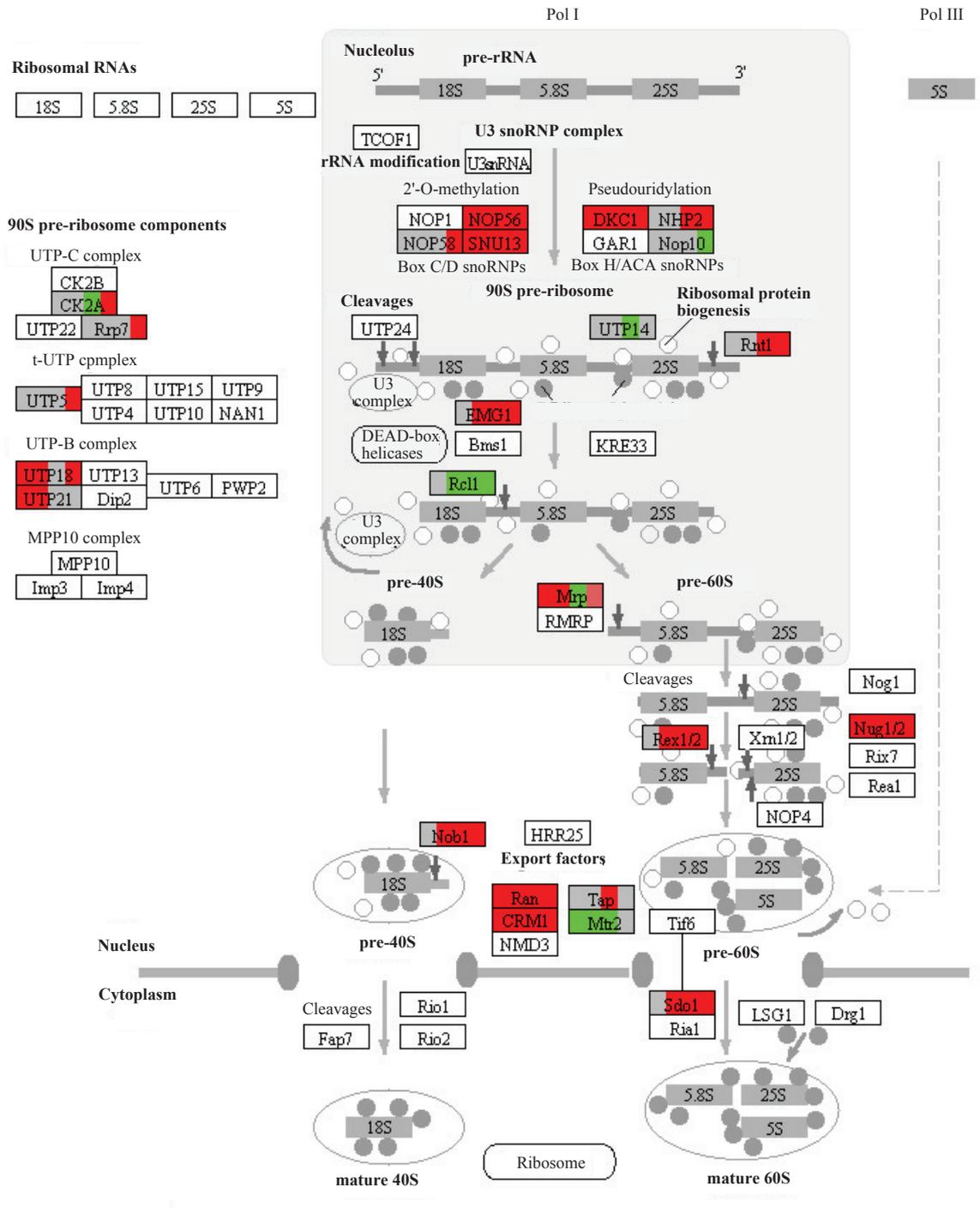


Figure 4.18.: Visualization of gene expression time course after butyrate addition for *KEGG* pathway Ribosome Biogenesis (Kanehisa and Goto 2000). Pathway analysis and visualization of gene expression changes based on M-values was performed using *pathview* (Luo and Brouwer 2013). Gene expression changes are color-coded according to the legend. Each node is divided into four parts representing differential expression at sampling points '6 h' to '36 h'. For nodes representing multiple proteins (isoforms, subunits etc.) maximal absolute values are used as the summary/representative of a node.

to change gene expression both on the global (i.e. numbers of genes) and gene specific (i.e. M-values) scale.

The study closest related to the one performed here was done by Klausning et al., who analyzed the butyrate effect (concentration 2 mM) in CHO DP-12 cells that were cultivated in batch bioreactor experiments. They found a total of 1,461 genes to be differentially expressed upon butyrate treatment (Klausning, Krämer, and Noll 2011). The array used for this analysis covered 38,310 transcripts. It has to be noted that not all of these transcripts were annotated at that time and that the microarray was a preliminary version of the one used in this thesis. For this reason only few expression values could be obtained. The second study used for comparison was performed using a CHO cell line expressing two anti-apoptotic genes (Kantardjieff et al. 2010). Analyses of mRNA and protein levels were conducted for cells cultured with addition of 2 mM butyrate at 33 °C using a microarray that covered 10,118 probe sets. A maximum of 994 differentially expressed genes were found in this experiment. In a third publication used for comparison, gene expression changes after treatment with 10 mM butyrate were analyzed in bovine kidney epithelial cells using a microarray that contained 45,383 unique cattle sequences (Li and Li 2006a). In this experiment, 450 genes were detected to be significantly differentially expressed.

Despite the differences in the experimental conditions or microarray types, all of these studies showed a negative effect of butyrate addition on cell growth, too. Furthermore, common functional groups of genes enriched within the respective datasets were, in accordance to the GO analyses performed in the present study, related to cell cycle control, apoptosis, metabolism and signaling. Li and Li as well as Kantardjieff et al. furthermore reported an enrichment of genes related to the extracellular matrix such as Matrix metalloproteinases. A similar enrichment was not observed in the present study.

Next, fold changes of gene expression for 99 cell cycle-related genes were compared between the studies. Generally, 76 of these were also found to be differentially expressed in the present study. Tab. 4.6 shows the M-values for selected examples (see Supplementary Tab. B.17 for a complete list).

Cyclin A2 (CCNA2) was found to be downregulated by Li and Li (fold change 0.17), Kantardjieff et al. (fold change 0.02) and also by the DNA-microarray experiments performed in this thesis (fold change 0.39). Origin recognition complex subunit 1 (ORC1L) as well as Minichromosome maintenance deficient 3 and 5 (MCM3 and MCM5) expression was shown to be downregulated in all studies. Max interacting protein 1 (MXI1) was found to be strongly upregulated by Li and Li (fold change 12.55) and was also found to be upregulated in the present study (fold change 2.26). MYC expression was detected to be downregulated only in CHO DP-12 cells by Klausning et al. (fold change 0.29) and by the present study (fold change 0.82). Taken together, gene-specific fold changes were generally comparable, especially

Table 4.6.: Comparison of studies that analyzed gene expression in cultured cells under butyrate treatment by microarrays (Kantardjieff et al. 2010; Klausning, Krämer, and Noll 2011; Li and Li 2006a) with the results from the present study ('36 h' sampling point). As fold changes were indicated differently in these studies they were converted for comparability. A complete list of fold changes for 99 cell cycle-related genes can be found in Supplementary Tab. B.17.

Transcript	Butyrate-induced gene expression fold changes measured by			
	Li and Li 2006	Klausning et al. 2011	Kantardjieff et al. 2010	Present study
CCNA2	0.17	n.d.	0.02	0.39
MCM3	0.10	0.86	0.05	0.41
MCM5	0.17	0.27	0.01	0.37
MXI1	12.55	n.d.	n.d.	2.26
MYC	n.d.	0.29	n.d.	0.82
ORC1L	0.23	0.27	0.07	0.52

those obtained by Klausning et al. and the present study. As the same cell line and microarray type were used in both experiments this underlines the reliability of this study. The highest changes in gene expression were measured by Kantardjieff et al., which may be due to the combination of butyrate treatment with lower cultivation temperatures. For this reason the highest discrepancy was found between these values and the ones found in the present study.

4.1.4.4. Genome-wide analysis of DNA methylation changes upon butyrate addition

In parallel to gene expression analyses, genomic regions showing butyrate-induced changes in DNA methylation were identified and analyzed regarding their association with genes. It has to be noted that, for this analysis, no discrimination was made between HMDs and PMDs, as it is not entirely clear if gene regulation is indeed directly influenced by PMDs and, importantly, PMD localization did not change in response to butyrate.

In the first analysis step, mean CpG methylation levels were compared (see Tab. 4.2 in section 4.1.2.4). Butyrate treatment was found to induce a maximum decrease in genome-wide mean methylation levels from 61.09 % in the reference samples to 59.55 % at sampling point '12 h'. 24 hours post butyrate the cells exhibited a mean methylation of 60.06 % and, at the '36 h' sampling point, of 61.23 %. Fig. 4.19A shows that butyrate addition furthermore did not affect the general bimodal distribution of unmethylated and highly methylated CpGs. Also, similarly small changes in mean methylation levels of CpG islands, microRNA (miRNA) genes, coding sequences (CDSs), introns, untranslated regions (UTRs) and non-codingRNA (ncRNA) genes were found (Fig. 4.19B–F). This indicated that differential DNA methylation in response to butyrate was a site-specific and not an overall effect.

Next, differentially methylated genomic regions (DMRs) were identified based on comparisons between the methylomes of reference and butyrate-treated cultures. Accordingly, the term 'hypomethylation' refers to lower methylation levels of a genomic region in a given

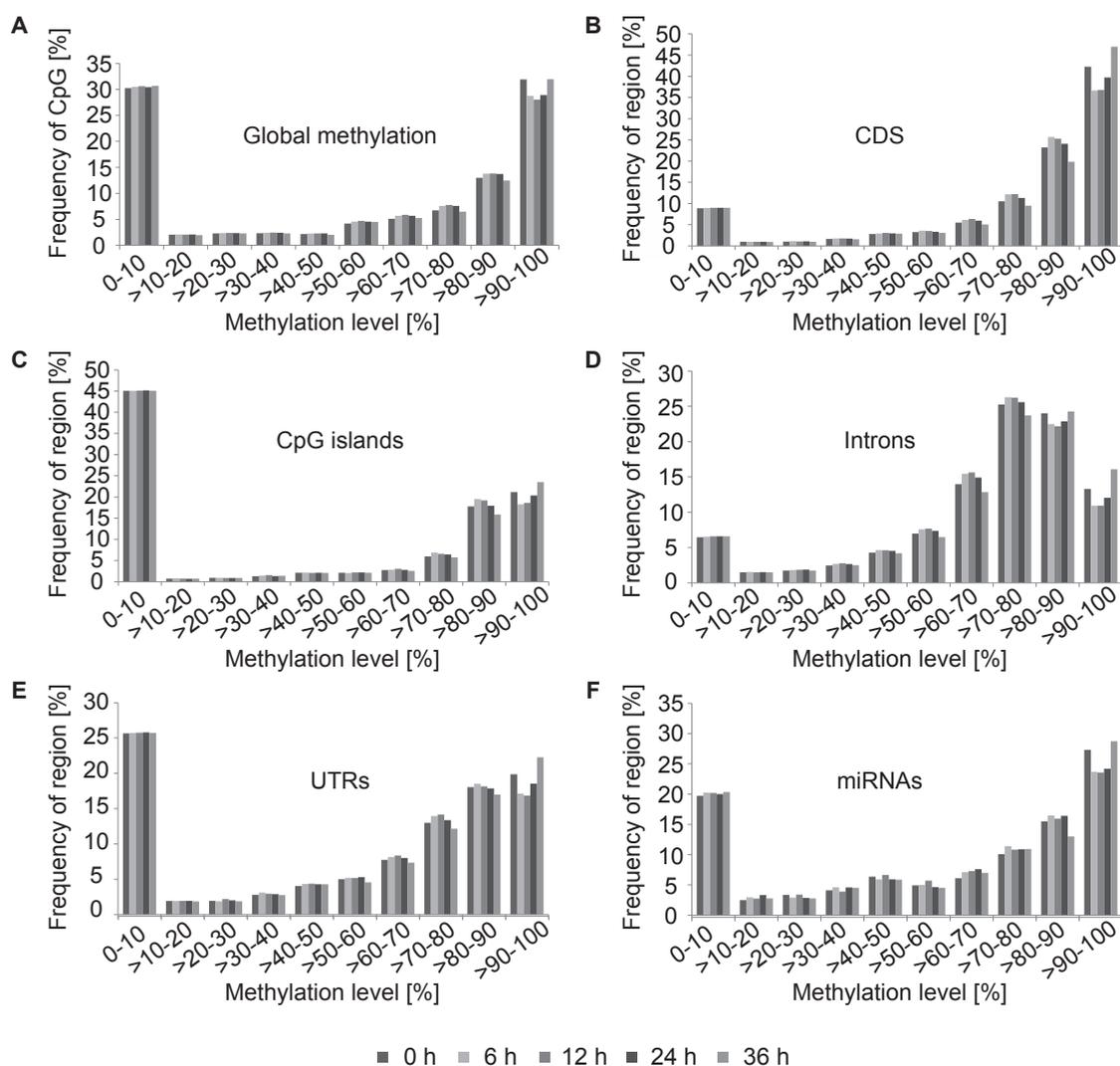


Figure 4.19.: Percent methylation of CpG dinucleotides and frequencies of mean methylation levels of genomic features. **A** Frequency of single CpGs within methylation level intervals (Wippermann et al. 2016). **B–F** Frequencies of mean methylation levels of the genomic regions coding sequence (CDS, **B**) CpG island (CGI, **C**), intron (**D**), untranslated region (UTR, **E**) and micro RNA (miRNA, **F**).

sample compared to the reference and the term 'hypermethylation' refers to relatively higher methylation levels. DMRs were subsequently grouped according to their position relative to genes, if at least one base pair of an individual region overlapped with the DMR. For this analysis isoforms of a gene were grouped and overlapping exons were merged. Four individual regions were defined for each gene: (i) the 5' regions '1 kb upstream of the start codon up to first intron' (as regulatory DNA methylation at 5' ends often affects CGIs, which are mostly found in promoter regions and first exons; Brenet et al. 2011; Jones and Takai 2001), (ii) the intragenic regions 'second exon to last intron', (iii) the 3' regions 'last exon to 1 kb downstream' and, to get an idea about potential methylation events further upstream of genes, (iv) the 'upstream region' (1001 bp to 20 kb upstream). If an upstream region

overlapped with another gene, the region was shortened. DMRs that could not be assigned to any of these regions were grouped into the category 'intergenic'.

The total numbers of DMRs varied slightly over time with a maximum difference in the numbers of hypomethylated regions between '6 h' (8113 DMRs) and '12 h' (8617 DMRs; Tab. 4.7). For the hypermethylated regions sampling point '36 h' (5589 DMRs) exhibited a maximum difference compared to the number of DMRs at the '6 h' sampling point (6524 DMRs). DMR sizes were found to range from about 2 to 8.5 kb for both hypomethylated and hypermethylated DMRs at all sampling points.

Table 4.7.: Numbers of relatively hypomethylated and relatively hypermethylated DMRs with ≥ 1 significantly differentially methylated CpG 6, 12, 24 and 36 hours after butyrate addition in comparison to the reference (Wippermann et al. 2016).

Sample	Hypomethylated DMRs	Hypermethylated DMRs
6 h	8,113	6,524
12 h	8,617	6,256
24 h	8,505	6,374
36 h	8,175	5,589

Fig. 4.20A shows the percentages of relatively hypomethylated and relatively hypermethylated regions that were associated with genes or could not be associated with any genes for the '6 h' time point, as no significant differences in DMR-count per feature were found between the sampling points (see pie charts for all sampling points in Supplementary Fig. B.24). 54 % of the hypomethylated regions at the '6 h' sampling point were assigned to intergenic regions and 46 % were associated with genes (23 % were associated with upstream regions, 10 % with 5', 11 % intragenic and only 2 % with 3' regions). Similarly, 57 % of the hypermethylated regions were associated with intergenic regions and 43 % could be associated with genes.

As potential correlation of DNA methylation changes and differential gene expression were supposed to be analyzed, DMRs that were not associated with genes were removed from further analyses at this point and 6536, 6739, 6929 and 6327 DMRs, corresponding to 4739, 4721, 4967 and 4459 unique genes, remained for the sampling points '6 h', '12 h', '24 h' and '36 h', respectively. This indicated that several DMRs were associated with single genes. The Venn diagrams in Fig. 4.20B furthermore show that differential methylation of individual genes could be present at one or multiple time points. Additionally, both hypo- and hypermethylation was detected for 30 % (2890) of the differentially methylated genes (Fig. 4.20C). For these genes associated with multiple DMRs, differential methylation occurred both at the same region or at different regions and DMRs could also be present at either the same or different time points.

Visual inspection of differentially methylated genomic regions in a genome browser illustrated various dynamics of butyrate-induced DNA methylation changes. Fig. 4.21 shows 5'-associated DMRs within CpG islands for the genes *DNA-binding protein SATB2* (*Satb2*, e.g.

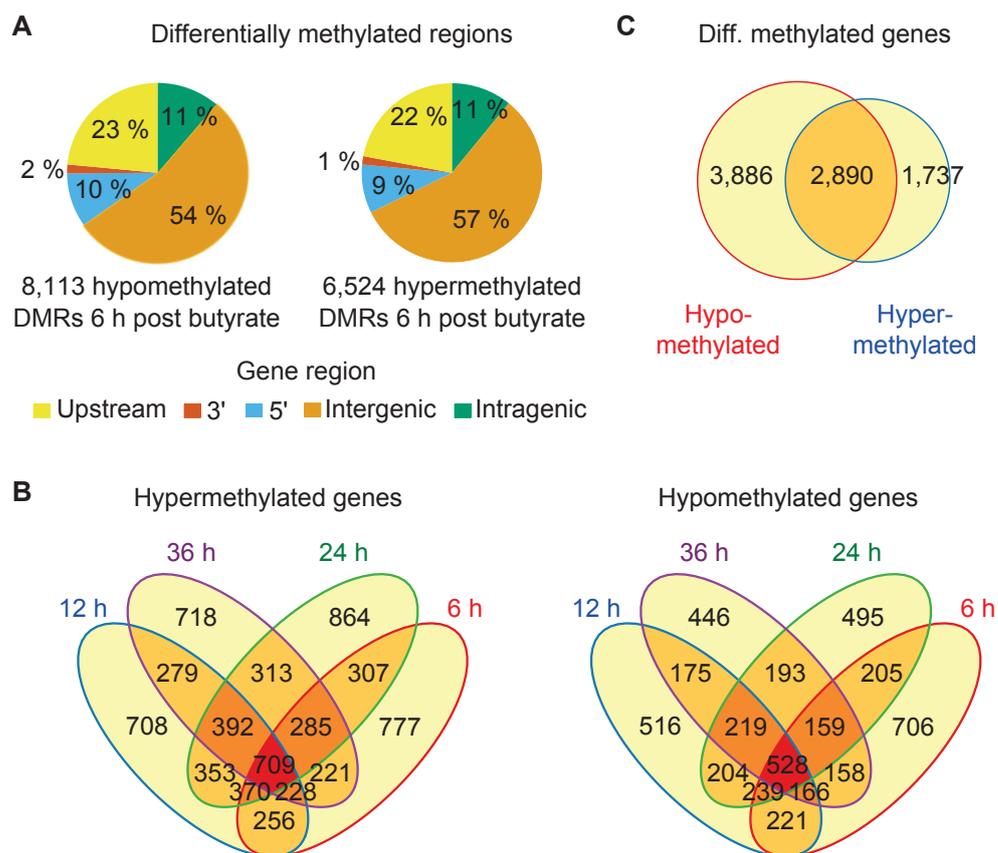


Figure 4.20.: DNA methylation changes in CHO DP-12 cultures upon butyrate addition (Wippermann et al. 2016). **A** Percentages of relatively hypomethylated and hypermethylated regions associated with genes at the '6 h' sampling point. **B,C** Venn diagrams of differentially methylated genes.

involved in chromatin remodeling; Gyorgy et al. 2008), *Microtubule-associated protein 10* (*Kiaa1338*, e.g. involved in mitotic spindle assembly; Fong et al. 2013) and *Oral-facial-digital syndrome 1 protein* (*Ofd1*, e.g. involved in cell cycle regulation; Singla et al. 2010). *Satb2* represents an example of an exclusively hypomethylated gene and *Kiaa1338* represents an example of exclusive hypermethylation. *Ofd1* was found to be associated with both kinds of differential methylation which occurred next to each other.

DMRs were also visually examined in a higher resolution on the single-nucleotide scale (Fig. 4.22). Fig. 4.22A shows selected examples of genomic regions detected to be hypomethylated (a set of each five exemplary regions per time point can be found in Supplementary Fig. B.25). In several cases (16 of 20 inspected regions) differential hypomethylation was detected exclusively for one time point. At the following time points methylation patterns were found to be re-established and no significant differential methylation was detected anymore (notice numbers of significantly differentially methylated CpGs).

The examination of relatively hypermethylated regions showed different dynamics of methylation changes over time (Fig. 4.22B and Supplementary Fig. B.26). For all of the hypermethy-

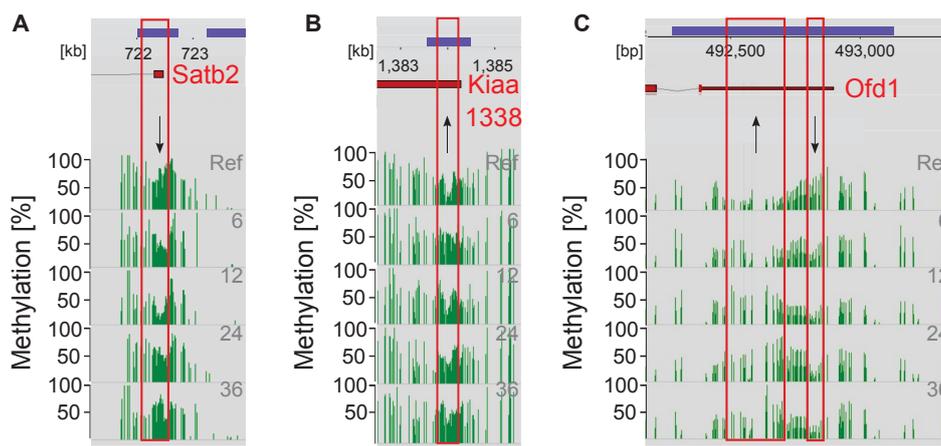


Figure 4.21.: Exemplary display of 5'-associated differentially methylated regions within CpG islands (Wippermann et al. 2016). Methylation data tracks were loaded into the *GenDBE* genome browser (Rupp et al. 2014) to visually inspect selected genomic regions. Exons are indicated by red rectangles and CpG islands by blue rectangles. Methylation levels for each CpG are represented by green bars. Regions that were identified as significantly hypo- or hypermethylated compared to the reference are marked by a red box and arrows indicate decreasing or increasing methylation levels. All three genes are located on the lagging strand. **A** Display of 3 kb of CHO K1 scaffold JH000295 including the 5' region of *DNA-binding protein SATB2* (*Satb2*). **B** Display of 3 kb of CHO K1 scaffold JH000057 including the 5' region of *Microtubule-associated protein 10* (*Kiaa1338*). **C** Display of 1 kb of scaffold JH001054 including the 5' region of *Oral-facial-digital syndrome 1 protein* (*Ofd1*).

lated regions examined (20 of 20), significant differential methylation was detected at more than one time point and methylation patterns generally showed some variability. For example, the hypermethylated DMR found to be present on scaffold JH000953 (Fig. 4.22B) showed 15 and 13 significantly differentially methylated CpGs 6 hours and 12 hours post butyrate addition, only 4 significant CpGs 24 hours post butyrate addition and again 14 significant CpGs 36 hours post butyrate.

To get an overview of the functional significance of genes that were affected by differential methylation upon butyrate addition, gene ontology analyses were performed separately for genes associated with both hypo- and hypermethylated DMRs as well as for genes exclusively associated with either hypo- or hypermethylated DMRs. Generally, enriched functional groups were very similar to those significantly enriched for the differentially expressed genes. Genes that were both hypo- and hypermethylated were functionally related to protein biosynthesis (e.g. 'translational elongation', 36 genes), differentiation processes (e.g. 'skeletal system development', 62 genes) and RNA metabolism (e.g. 'negative regulation of RNA metabolic process', 67 genes; Supplementary Tab. B.14). Genes associated with exclusively hypomethylated DMRs were most significantly enriched in functions related to cell cycle regulation (e.g. 'cell cycle phase', 101 genes), transcription (e.g. 'regulation of transcription from RNA polymerase II promoter', 157 genes), signaling (e.g. 'intracellular signaling

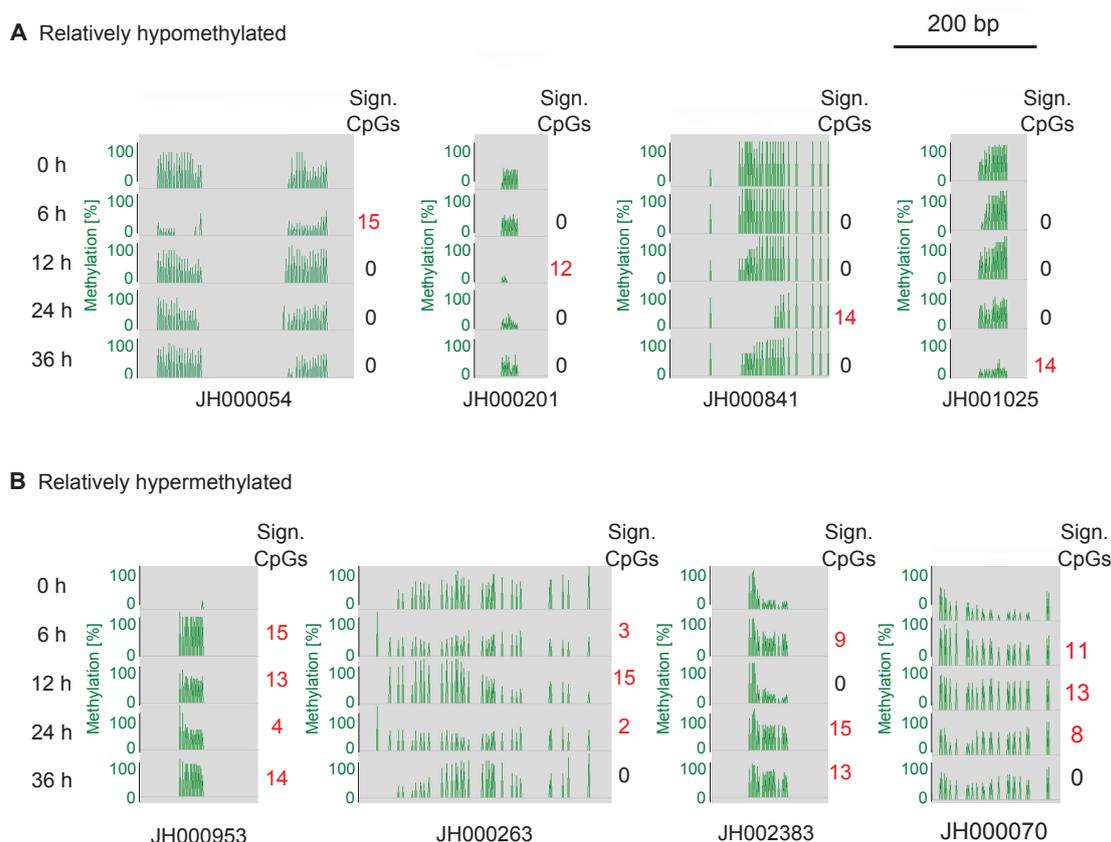


Figure 4.22.: Screenshots of relatively hypomethylated (**A**) and hypermethylated (**B**) DMRs as visualized in the *GenDBE* genome browser (Rupp et al. 2014). Methylation levels for each CpG are represented by green bars. All regions are represented in the same scale (see scale bar). Indicated are scaffold numbers as well as the numbers of CpGs detected as significantly differentially methylated.

cascade’, 250 genes) and apoptosis (e.g. ’regulation of cell death’, 170 genes; Supplementary Tab. B.15). Those genes that were exclusively associated with hypermethylated DMRs (Supplementary Tab. B.16) were most significantly related to protein transport (e.g. ’protein localization’, 110 genes) and RNA processing (e.g. ’RNA processing’, 66 genes).

4.1.4.5. Combination of DNA methylation and gene expression data

In order to combine the data generated by whole-genome bisulfite sequencing and gene expression microarray analyses, potential correlations between DNA methylation and gene expression profiles were analyzed. First, a general analysis was performed in order to assess if up- or downregulation of gene expression was associated with specific methylation types (simultaneous hypo- and hypermethylation or exclusive hypo- or hypermethylation). For this purpose density plots of gene expression M-values were generated (Fig. 4.23). The majority of genes associated with both hypo- and hypermethylated DMRs at different regions was not differentially expressed (about 81 %, represented by the peak at $M = 0$ in the density plot).

For genes associated with hypo- and hypermethylated DMRs at the same genomic region as well as exclusively hypo- or hypermethylated DMRs the density curves showed significantly lower percentages of genes that were not differentially expressed with about 38 %, 47 % and 48 %, respectively.

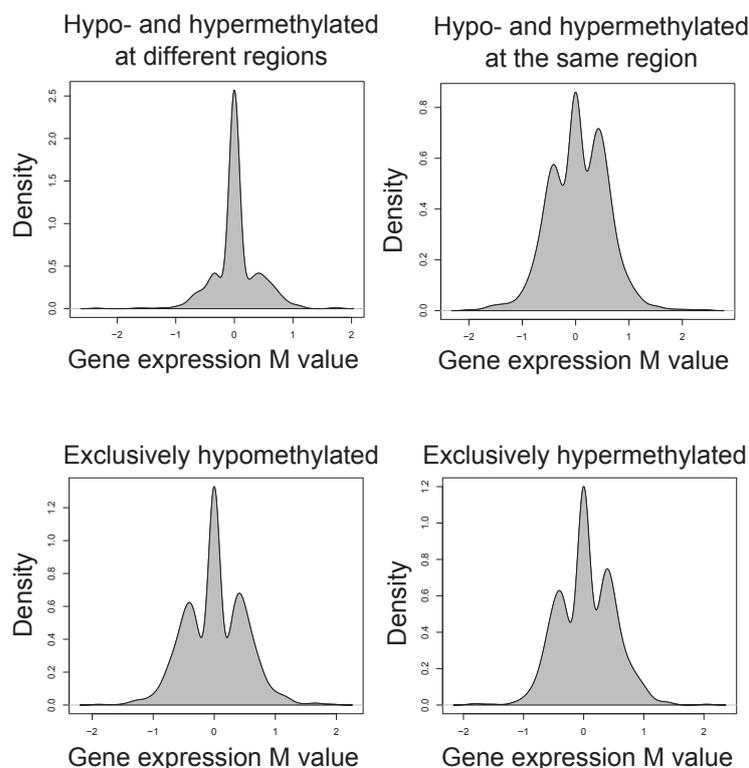


Figure 4.23.: Density plots of gene expression M-values associated with different types of differentially methylated genes.

In total, 973 genes (11 % of the differentially methylated genes) were found to be both differentially methylated and expressed. A hierarchical cluster analysis was performed for these genes to assess potential correlations between gene expression profiles and DNA methylation changes. Five distinct clusters of DNA methylation and gene expression patterns were detected (Fig. 4.24A). Inspection of the two main clusters (clusters 1 and 5) showed that the majority of up- or downregulated genes was not associated with specific DNA methylation profiles. However, three smaller clusters differed from the majority of genes. Cluster 4 contained a group of genes associated with both hypo- and hypermethylation in intragenic regions which was associated with both up- and downregulated gene expression, but no significant enrichment of any gene functions was found for these genes. Cluster 2 showed hypomethylation of upstream regions, whereas cluster 3 exhibited hypermethylation of upstream regions at multiple time points. Both clusters were associated with both up- and downregulated genes as well. The genes of these clusters were significantly related to 'translational elongation' ($p = 3.7 \times 10^{-9}$), 'ribosome biogenesis' ($p = 1.2 \times 10^{-3}$) and 'ribonucleoprotein complex biogenesis' ($p = 8.9 \times 10^{-3}$). A search for protein associations between the corresponding gene products using the STRING database revealed a network of

strongly interacting ribosomal proteins and associated factors such as Heat shock cognate 71 kDa protein (HSPA8), mRNA turnover protein 4 (MRTO4) or Bystin (BYSL; Fig. 4.24B). Most of the genes within this network were not differentially expressed at the '6 h' sampling point, but were found to be differentially expressed 12 hours after butyrate addition or later (Fig. 4.24C). A majority of the associated factors were downregulated, whereas most of the ribosomal components were upregulated. Interestingly, HSPA8 showed a very different expression profile as it was already 2.1-fold upregulated 6 hours post butyrate addition and showed a maximum of 2.6-fold upregulation 12 hours post butyrate followed by a constant decrease in differential expression.

For further exploration of mechanisms underlying the effects of butyrate addition on DNA methylation and gene expression *de novo* motif discovery was performed using the motif discovery algorithm *DREME*, which was designed to find short core motifs of eukaryotic transcription factors (Bailey 2011). Significant enrichment of sequence motifs in DMRs was calculated using the Fisher's Exact Test in comparison to a ten-fold bigger set of randomly chosen sequences with an equal length distribution (i.e. ten 'negative' sequences per 'positive' sequence). Very low *p* values with a minimum of 2.3×10^{-438} were calculated for the motif 5'-CGD-3' (D = AGT). As DMRs naturally contain CpG dinucleotides and a lot of them in high densities, the significant overrepresentation of CG-containing motifs was expected and proofed the reliability of the sequence discovery method. *De novo* discovered motifs were subsequently compared to human and mouse transcription factor databases using the *Tomtom* algorithm (Gupta et al. 2007) in order to determine putative transcription factor binding sites. Eight DNA sequence motifs were discovered in upstream regions which significantly corresponded to 42 putative transcription factor binding sites. At 5' ends two motifs were found which corresponded to eight putative transcription factor binding sites (see Supplementary Tab. B.9 for a list of DMR motifs and putatively corresponding transcription factors). Strikingly, ten factors with putative binding sites in upstream regions and five transcription factors with putative binding sites in 5' regions were indeed found to be differentially expressed when the gene expression microarray data were searched for differential expression of transcription factors corresponding to these binding sites (Fig. 4.25). These included factors with upregulated expression such as Zinc finger and SCAN domain containing protein 4 (ZSCAN4) and Forkhead box protein 1 (FOXO1) as well as several factors showing downregulated expression such as Transcription factor AP-4 (TFAP4), Transcription factor AP-2 gamma (TFAP2C), Ras-responsive element-binding protein 1 (RREB1), Transcription factor SP1 (SP1), Krueppel-like factor 13 (KLF13), Proto-oncogene c-Rel (REL), Nuclear factor NF-kappa-B p65 subunit (RELA) and Nuclear factor NF-kappa-B p105 subunit (NFKB1).

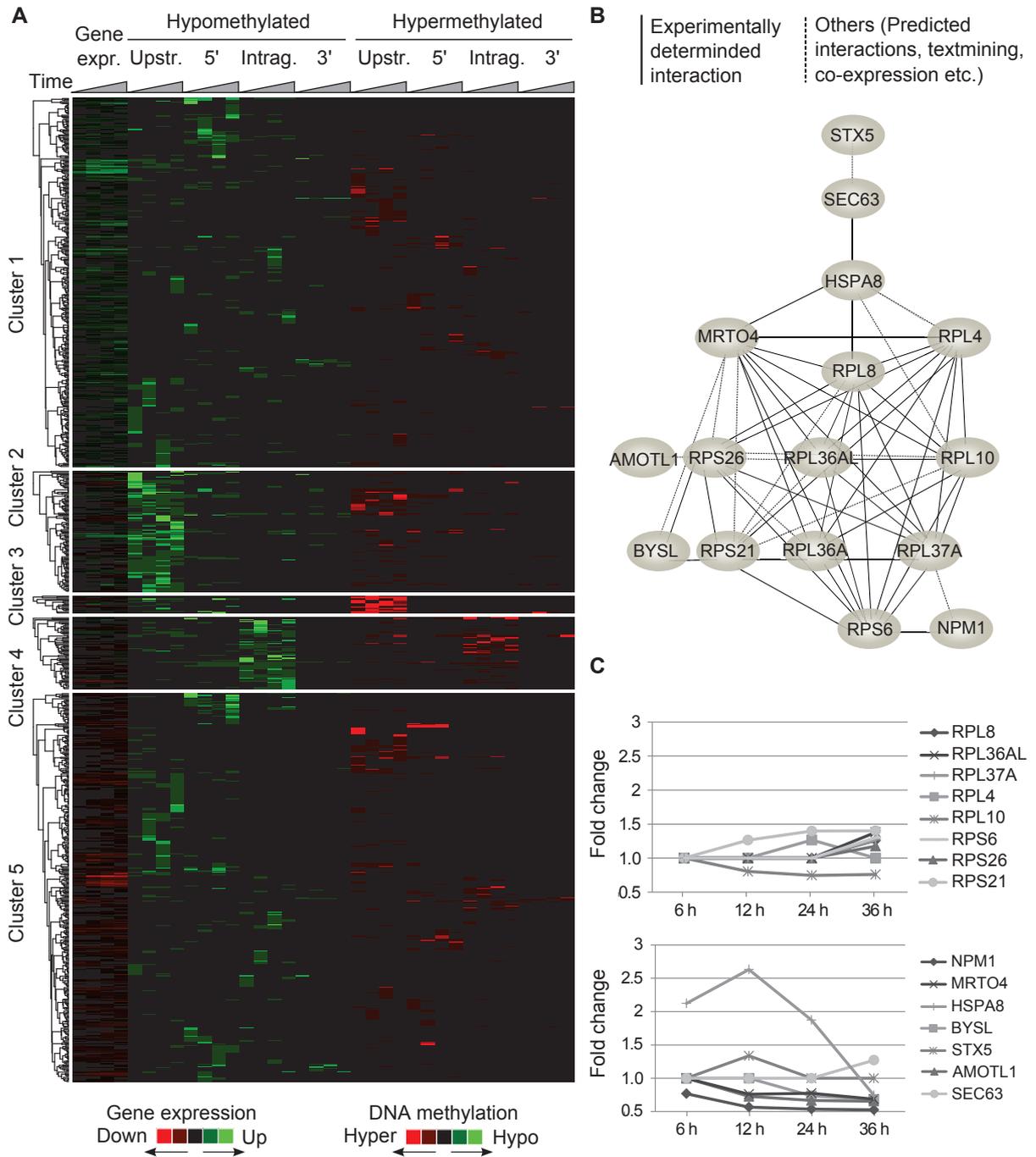


Figure 4.24.: Integration of DNA methylation and gene expression data (Wippermann et al. 2016). **A** Hierarchical cluster analysis of gene expression and DNA methylation profiles was performed using Ward's method for agglomeration (Ward 2012). Changes in gene expression based on fold-changes and numbers of differentially methylated CpGs are color-coded according to the legend. **B** Analysis of the genes comprised within clusters 2 and 3 using the *STRING* database (Szklarczyk et al. 2015) revealed a network of ribosomal proteins and associated factors. **C** Gene expression fold-changes are represented for ribosomal proteins (top diagram) and associated factors (bottom diagram).

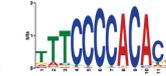
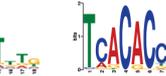
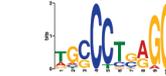
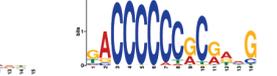
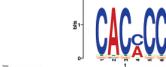
TF ID	TFAP4	ZSCAN4	TFAP2C	FOXO1	RREB1
Eukar. TF site					
DMR motif					
P value	4.2×10^{-5}	6.5×10^{-5}	1.7×10^{-5}	4.2×10^{-4}	9.6×10^{-4}
Position	upstream	upstream	upstream	5'	5'
Expression	0.7-fold	1.2-fold	0.8-fold	1.4-fold	0.7-fold
TF ID	SP1	KLF13	MGA	TFAP2A	GLIS2
Eukar. TF site					
DMR motif					
P value	1.6×10^{-4}	4.8×10^{-4}	1.9×10^{-4}	3.7×10^{-4}	1.1×10^{-4}
Position	upstream	upstream	upstream	upstream	upstream
Expression	0.7-fold	0.8-fold	0.6-fold	0.6-fold	0.6-fold
TF ID	REL	RELA	NFKB1	GLIS3	GLIS1
Eukar. TF site					
DMR motif					
P value	7.6×10^{-5}	1.3×10^{-4}	2.5×10^{-4}	1.3×10^{-4}	2.3×10^{-4}
Position	5'	5'	5'	upstream	upstream
Expression	0.7-fold	0.7-fold	0.8-fold	1.4-fold	1.4-fold

Figure 4.25.: Putative transcription factor binding-site motifs *de novo* discovered in DMRs upon butyrate addition (Wippermann et al. 2016). Motif comparisons were performed using *Tomtom* algorithm (Gupta et al. 2007). Shown are sequence logos and *p* values. It is furthermore indicated whether the DMR was located within an upstream region or a 5' region and if putative transcription factor expression was up- or downregulated.

4.1.4.6. Discussion: Butyrate-induced changes in gene expression

Functional analysis of differentially expressed genes upon butyrate addition showed that cellular pathways such as cell cycle, apoptosis, signaling, protein transport, central energy metabolism and RNA processing were significantly affected by butyrate. These processes have been described to be affected by butyrate addition in CHO cells and other cell types before (Blouin et al. 2011; Bordonaro, Lazarova, and Sartorelli 2008; Kantardjieff et al. 2010; Klausning, Krämer, and Noll 2011; Li and Li 2006b; Scott, Longpre, and Loo 2008; Sun et al. 2012; Wu et al. 2012). A closer examination of specific pathways furthermore indicated that butyrate induced downregulation of most cell cycle genes over the time course of 36 hours, potentially activated the intrinsic pathway of apoptosis and led to changes in glycolysis and TCA cycle, which might indicate an increase in energetic resources. Furthermore, several genes contributing to RNA splicing processes were found to be downregulated and

some were hypothesized to be functionally linked to DNA methylation events through the base excision repair pathway. Genes involved in protein biosynthesis were found to be differentially expressed as well, although an overall downregulation of RNA transport and ribosome biogenesis rather suggested reduced translational capacities than the opposite. As ribosome biogenesis requires large amounts of energy (Thomson, Ferreira-Cerca, and Hurt 2013), downregulation of this process might contribute to an increase in energy resources available for mAb production downstream of this process, for example during translation itself. Strikingly, upregulated genes were found to be enriched for those contributing to translational elongations and, additionally, a central factor relevant for protein export (BIP) was found to be upregulated as well. BIP was shown to be related to high productivities in CHO cells before (see e.g. Carlage et al. 2009). Finally, butyrate was shown to affect several genes related to N-glycan biosynthesis, suggesting that its addition might very well contribute to changes in the quality of recombinant proteins, as it has been reported before, too (Hong et al. 2014; Lee et al. 2014).

The gene expression analysis furthermore showed that genes related to 'chromatin remodeling' and 'protein amino acid acetylation' as well as differentiation events were strongly affected specifically at the first sampling point, indicating a potential role for these processes within the early response to butyrate. As butyrate is known for its function as an inhibitor of histone deacetylation (Boffa et al. 1978; Corfe 2012; Donohoe and Bultman 2012), an early onset of processes related to chromatin remodeling could be expected. An involvement of processes related to differentiation was less expected in the context of CHO-based production of biopharmaceuticals, although butyrate has generally been shown to induce differentiation of several cell types, including colonic T cells, adipocytes and medulloblastoma cells (Furusawa et al. 2013; Li, Yao, and Jiang 2014; Nör et al. 2013).

4.1.4.7. Discussion: Transcription factor binding-site motifs in DMRs

Differentially methylated regions (DMRs) were found to be associated with genes related to cellular processes similar to the ones significantly enriched within the group of differentially expressed genes. Furthermore, DMRs were found to contain binding-site motifs of specific transcription factors which were also differentially expressed and were therefore hypothesized to represent regulatory sites closely connected to the cellular response to butyrate. The fact that 973 genes were both differentially methylated and differentially expressed underlined the hypothesis of these regions being active regulatory sites, although this number represented only 11 % of the total differentially methylated genes. However, similarly small overlaps between differential methylation and gene expression have been found in other studies as well, even when RNA-sequencing was applied instead of gene expression analyses by microarrays (see e.g. Kubo et al. 2015; Sadler et al. 2016). The presence of transcription factor binding-

site motifs within DMRs furthermore underlined a potential involvement of differentiation processes in the butyrate response, as many transcription factors with decreased expression levels were related to cell growth, the regulation of apoptotic events and differentiation. These included SP1, KLF13, TFAP2 and RREB1 as well as REL, RELA and NFKB1, which represent three of the REL-like domain containing subunits of NFKB (Bragança et al. 2002; Cruz-Topete et al. 2016; Guo et al. 2015a; Jiang et al. 2015; Perkins 2000; Vizcaíno, Mansilla, and Portugal 2015).

On the contrary, transcription factors involved in pluripotency-related processes as well as cell cycle arrest, cell death and energy homeostasis were upregulated. These included ZSCAN4, which is involved in processes related to pluripotency by regulation of telomere elongation (Lee and Gollahon 2015) and FOXO1, which is related to cell cycle arrest and cell death (Martins, Lithgow, and Link 2016). FOXO1 was furthermore shown to be involved in the regulation of energy metabolism by promotion of the expression of gluconeogenic enzymes (Gross, Heuvel, and Birnbaum 2008). Strikingly, our results showed an increased expression of pyruvate carboxylase (PC), which might indeed indicate an increase in gluconeogenesis.

The finding that processes related to central metabolism were potentially affected by butyrate addition was confirmed on the phenotypic level by a switch from lactate production to consumption that was visible 24 hours post butyrate addition. Interestingly, several studies link upregulated expression of genes related to energy metabolism to increased productivities in CHO cells, including Müller et al. who analyzed the effect of butyrate on the CHO cell proteome by label-free MS quantification and found (amongst others) several proteins with altered expression that were involved in glycolysis, TCA cycle and the pentose phosphate pathway (Müller et al. 2016). These included proteins with increased expression levels such as IDH1 and MDH1 as well as proteins showing decreased expression such as GAPDH, which were found to be differentially expressed in this study, too. Bedoya-Lopez et al. furthermore showed that CHO cell productivity increased due to temperature downshifts and that this was accompanied by a significant upregulation of genes related to energy metabolism (Bedoya-López et al. 2016). Meleady et al. could also identify changes in expression levels of genes related to energy metabolism that were significantly associated with productivities (Meleady et al. 2011). In the latter study HSPA8, which was found here within a network of genes exhibiting constant differential methylation of upstream regions, was also amongst the genes related to sustained productivity and has been found in similar contexts as well (Lee et al. 2009; Seth et al. 2007; Smales et al. 2004). Interestingly, HSPA8 is a member of the HSP70 family of chaperones with additional functions in the regulation of pluripotency (Geng et al. 2015; Stricher et al. 2013) and was significantly upregulated by butyrate addition already at the first sampling point. This finding hints towards a specific, yet to be further elucidated role of HSPA8 within the butyrate-responsive regulatory network.

4.1.4.8. Discussion: Dynamics of DNA methylation changes upon butyrate addition

The genome-wide detection of differential DNA methylation in CHO DP-12 cells after butyrate addition showed that the overall CHO DP-12 DNA methylation landscape did not change significantly over the time course of 36 hours, as the general bimodal distribution of either unmethylated and fully methylated CpGs did not change. Also, the location of partially methylated domains (PMDs) was not affected by butyrate addition and mean methylation levels of genomic features did not change. For this reason differential methylation was analyzed site-specifically. When differentially methylated regions (DMRs) were detected, no significant changes in their numbers or association with genomic features were found. Nevertheless, DMRs showed a great variability regarding (i) different kinetics of the removal of DNA methylation marks and (ii) multiple DNA methylation events affecting individual genes at multiple time points. The differences in methylation kinetics observed between hypo- and hypermethylated DMRs on the single-nucleotide level, which showed that hypomethylated DMRs could be quickly established whereas hypermethylated DMRs were only slowly removed, could be due to differences within the TET-based pathway for regeneration of unmodified cytosine bases (see section 1.3.2 for a detailed description). Briefly, TET enzymes are capable of oxidizing 5-methylcytosine and oxidation products can be removed passively by being diluted in a DNA-replication dependent manner or by active removal mediated by Thymine DNA glycosylase (TDG), which functions as a DNA repair enzyme capable of removing thymines from T-G-mismatches, within the base excision repair (BER) pathway (Kohli and Zhang 2013). The hypothesis that hypermethylated DMRs were passively removed was underlined by the finding that DNA methylation patterns in the respective regions were unstructured and variable, which possibly represented an uncoordinated process of 'dilution-by-replication' due to the butyrate-induced cell cycle arrest. On the contrary, the establishment of hypomethylated DMRs was found to occur precisely and quick, thereby emphasizing a possible involvement of an active enzymatic process. The fact that, in both cases, the establishment of hypermethylated states happened comparably quick adds to this hypothesis, as the establishment of DNA methylation marks is only possible through an active process. Variable DNA methylation at individual genes has been found in other biological systems as well. For example, Wan et al. analyzed genome-wide DNA methylation and gene expression patterns in different tissues and found that multiple DMRs can be associated with one gene. Among 247 genes that were found in this study to be associated with more than one tissue-specific DMR, 79 exhibited incoherent DMR sets (Wan et al. 2015).

4.1.4.9. Conclusion: Data integration allows for novel perspectives

The integrative analysis of time course DNA methylation and gene expression data from butyrate-treated CHO cells proved the utility of integrating DNA methylation and gene

expression data in a biotechnological context. Gene expression analyses proved that butyrate addition affected pathways such as cell cycle or apoptosis, and regions that exhibited variable DNA methylation in response to butyrate were associated with genes related to similar functional groups. As a subset of genes furthermore exhibited both differential methylation and expression and differentially methylated regions were found to contain potential binding-site motifs of transcription factors putatively related to the observed phenotypic effects of butyrate addition, these regions were hypothesized to represent regulatory sites closely connected to the cellular response to butyrate addition. Generally, DNA methylation was considered to reflect on-going gene activity. Obviously, questions remain regarding the relevance of DNA methylation changes at genes that were not found to be differentially expressed. As it is not clear how DNA methylation affects gene expression and if it occurs up- or downstream of gene regulation (Schübeler 2015), answering these questions will require further experiments. Further studies are required to assess the involvement of the transcription factors putatively involved in the butyrate response as well. However, these results generally underline the fact that integrating DNA methylation and gene expression data allows for a novel perspective regarding target selection for rational cell line development. Taken together, this experiment showed that integration of epigenetic and gene expression data allowed for new insights into the butyrate effect on CHO cells and generally prove the benefit of this kind of studies in the production of biopharmaceuticals.

4.2. Comparison of whole-genome bisulfite sequencing and CpG island microarray analysis

Although technological advances and reduced costs of whole-genome sequencing experiments shifted DNA methylation analyses from gene-specific to genome-scale approaches, these methods still require a sufficient sequencing depth and are not feasible for large numbers of samples, which, for example, could be necessary to monitor DNA methylation changes over a cultivation process. A more economical method is the analysis of DNA methylation by microarrays, which employ the fact that methylated DNA fragments can be enriched from sheared genomic DNA using a Methyl-CpG binding domain (MBD) protein. Enriched fractions can subsequently be analyzed by hybridization to microarrays which contain probes corresponding to pre-selected genomic loci. In the context of CHO-based production of recombinant proteins, CpG islands were considered to be of special interest, as CGI methylation can cause stable transcriptional repression of linked genes and has been described to occur during malignant transformation of several types of cancer (Park et al. 2011; Rauch et al. 2012). According to this fact it was hypothesized that CpG island methylation in CHO cells might correlate with specific phenotypes or reactions to environmental conditions as well.

Generally, CpG island microarrays for human and mouse are commercially available. For example, Agilent's human CGI microarray covers 27,800 CGIs with 237,220 probes. As a CGI microarray specifically designed for CHO cells did not exist, a CHO-specific array was developed in a master thesis. This array covers 19,598 promoter-associated and intragenic CGIs in the CHO genome (89 % of the total CGIs that were bioinformatically identified) with 27,446 unique probes in a 60K format (62,976 probes per array; Wippermann 2012; Wippermann et al. 2013). Fig. 4.26A shows a schematic representation of the location of probes relative to genomic features. Due to this layout, the CHO-specific CpG island microarray contains each probe at least twice, which results in at least four replicates when microarray analyses are performed in dye swap pairs.

Fig. 4.27 outlines the experimental procedure of CpG island microarray analysis. Central to the experiment was the enrichment of methylated DNA from a pool of DNA fragments, which exhibited an average size of 500 bp and were generated by repeated cycles of sonication. Methylated DNA fragments were then enriched using a methyl-DNA-binding domain (MBD) protein coupled to magnetic beads and the enrichment efficiency was controlled by determining the ratio between the two previously identified control regions (*Actb* and *Bcat1*). Enriched fractions of methylated DNA were subsequently labeled with fluorescent dyes and comparatively hybridized to the microarray.

The CpG island microarray was applied to analyze the butyrate effect on CHO DP-12 cells (Wippermann 2012) and the data generated by this means were functionally analyzed in

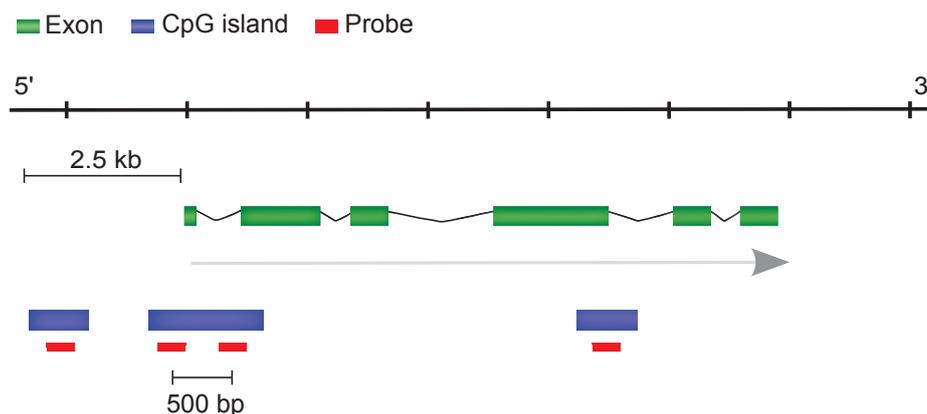


Figure 4.26.: Schematic representation of CHO CpG island microarray probe distribution relative to genomic features. Genomic features are color-coded according to the legend. Promoter regions were defined as '2.5 kb upstream to end of first exon' as CHO transcription start sites were not known at that point. For long CGIs probes were designed with average distances of about 500 bp (please refer to Wippermann 2012 for details).

detail during this doctoral thesis. The results from this analysis were compared to the whole-genome bisulfite sequencing results in order to assess microarray performance and identify potential sources of bias. Additionally, exemplary results were confirmed by gene-specific analysis of DNA methylation and CpG island microarray performance was controlled in further hybridization experiments. Finally, putative central mediators of the butyrate effect were identified based on the comparative data analysis and functionally analyzed in transient RNAi experiments.

4.2.1. Analysis of the butyrate effect by CpG island microarrays

Cultivations ($n = 4$) were performed in the same way as for the analysis of the effect of butyrate treatment by whole-genome bisulfite sequencing (Supplementary Fig. B.37) and a very similar effect of butyrate on growth and productivity was observed. Briefly, reduced cell growth could already be detected 24 hours after butyrate addition in treated cultures and the viability started to decrease. The butyrate-treated cells reached a maximum of 31×10^5 cells/mL 48 hours after butyrate addition and enhanced cell specific productivities were observed on day 5 with a maximum value of 17 pg/cell/day. The control cultures continued with normal cell growth until a maximum viable cell density of 110×10^5 cells/mL was achieved. They reached a maximum cell specific productivities of 9 pg/cell/day. Here, also, a butyrate-induced switch to lactate consumption could be observed. Samples for DNA methylation analysis by CpG island microarrays were taken before butyrate addition as well as 24 hours and 48 hours after butyrate addition from both the reference and butyrate-treated cultures. DNA was extracted and sheared (see images of control agarose gels in Supplementary Fig. B.21) and methylated DNA fragments were enriched. Per sampling point two CpG island microarrays were hybridized as dye swap pairs. It has to be noted that the comparisons at each time point

4.2. Comparison of whole-genome bisulfite sequencing and CpG island microarray analysis

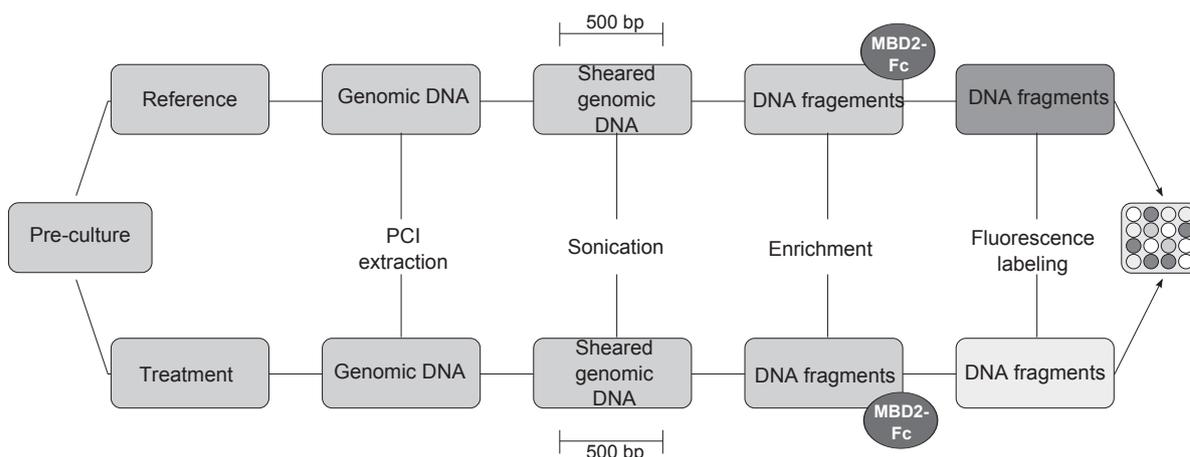


Figure 4.27.: Schematic representation of the experimental procedure of CpG island microarray analysis. After extraction, DNA is sonicated to an average fragment size of about 500 bp. Methylated DNA fragments are enriched using a methyl-DNA-binding domain (MBD) protein coupled to magnetic beads. The enriched methylated fractions are labeled with Cyanine 3- and 5-dUTP and the samples are subsequently hybridized to the microarray.

were made directly between the reference cultures and the cultures that were treated with butyrate, which is in contrast to the whole-genome bisulfite sequencing experiment. This means that for the '0 h' time point the reference cultures were compared to the cultures that were not yet treated with butyrate. Therefore, the comparison between the cultures at the '0 h' time point served as a control, as differences between identical cultures were not expected. Data analysis was performed as described previously. Corresponding MA plots can be found in Supplementary Fig. B.38.

The Venn diagrams in Fig. 4.28 illustrate the numbers of genes associated with differentially methylated CpG islands that were found in this analysis and show that the effect of butyrate on CpG island methylation was dynamic and reversible. Surprisingly, 42 genes were relatively hypermethylated and 141 genes were relatively hypomethylated in a comparison between the reference cultures and the cultures that were treated with butyrate later on ('0 h' sampling point; see discussion in section 4.2.5). These numbers and also the ratios between the respective sampling points remained constant even after applying the most stringent statistical parameters. Therefore, these genes were excluded from the '24 h' and '48 h' datasets before further data analysis. Small numbers of differentially methylated genes were also found at the end of the experiment (32 genes affected by relative hypermethylation and 110 genes affected by hypomethylation). A strongly increased number of 1221 genes associated with differentially methylated CpG islands was found 24 hours after butyrate supplementation, with 342 genes showing relative hypermethylation of corresponding CGIs and 879 genes being associated with relatively hypomethylated genomic regions. The greatest overlap in the sampling points regarding affected genes existed between 24 hours and 48 hours (139 genes in total). Furthermore, 30 genes showed differential methylation both before butyrate

treatment and 24 hours later. Only 10 genes were found being significantly differentially methylated amongst all three time points and 5 differentially methylated genes were both present at 0 and 48 hours.

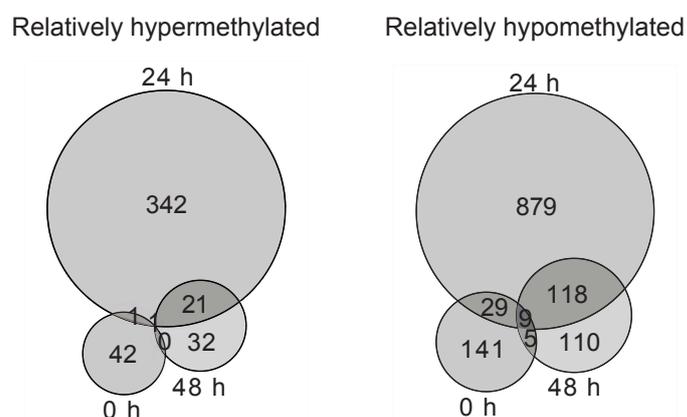


Figure 4.28.: Significantly differentially methylated genes upon butyrate treatment (Wippermann et al. 2013). Venn diagrams of genes associated with hypermethylated or hypomethylated CpG islands prior to (0 h), 24 hours (24 h) and 48 hours (48 h) upon treatment of CHO DP-12 cells with 3 mM butyrate.

Gene ontology analyses were carried out in order to gain an overview of affected cellular processes and components (Fig. 4.29). In the dataset recorded before butyrate treatment it was found that the protein group 'positive regulation of developmental process' was significantly overrepresented. The data of sampling point '24 h' suggested that the genes associated with hypermethylated CGIs were related to the processes 'response to stimulus', 'phosphorylation' and 'biogenesis', e.g. represented by the more specific functions 'DNA repair', 'response to organic substance' or 'programmed cell death'. Within the group of genes associated with hypomethylated CGIs the categories 'transcription', 'chromatin', 'signaling', 'immune response', 'developmental process' and 'phosphate metabolism' were frequently affected, e.g. comprising functions such as 'regulation of gene-specific transcription', 'chromatin modification', 'regulation of MAP kinase activity' or 'immune system development' (also see complete list of GO annotations for the '24 h' sampling point in Supplementary Tab. B.18 and B.18). Another 24 hours later (sampling point '48 h'), genes related to 'carbohydrate biosynthetic process' were abundantly associated with hypermethylated CGIs. Hypomethylated genes at the 48 hours sampling point could be clustered according to the functional categories 'chromatin modification', 'dendrite morphogenesis', 'post-embryonic development' or 'protein modification by small protein conjugation or removal'.

4.2. Comparison of whole-genome bisulfite sequencing and CpG island microarray analysis

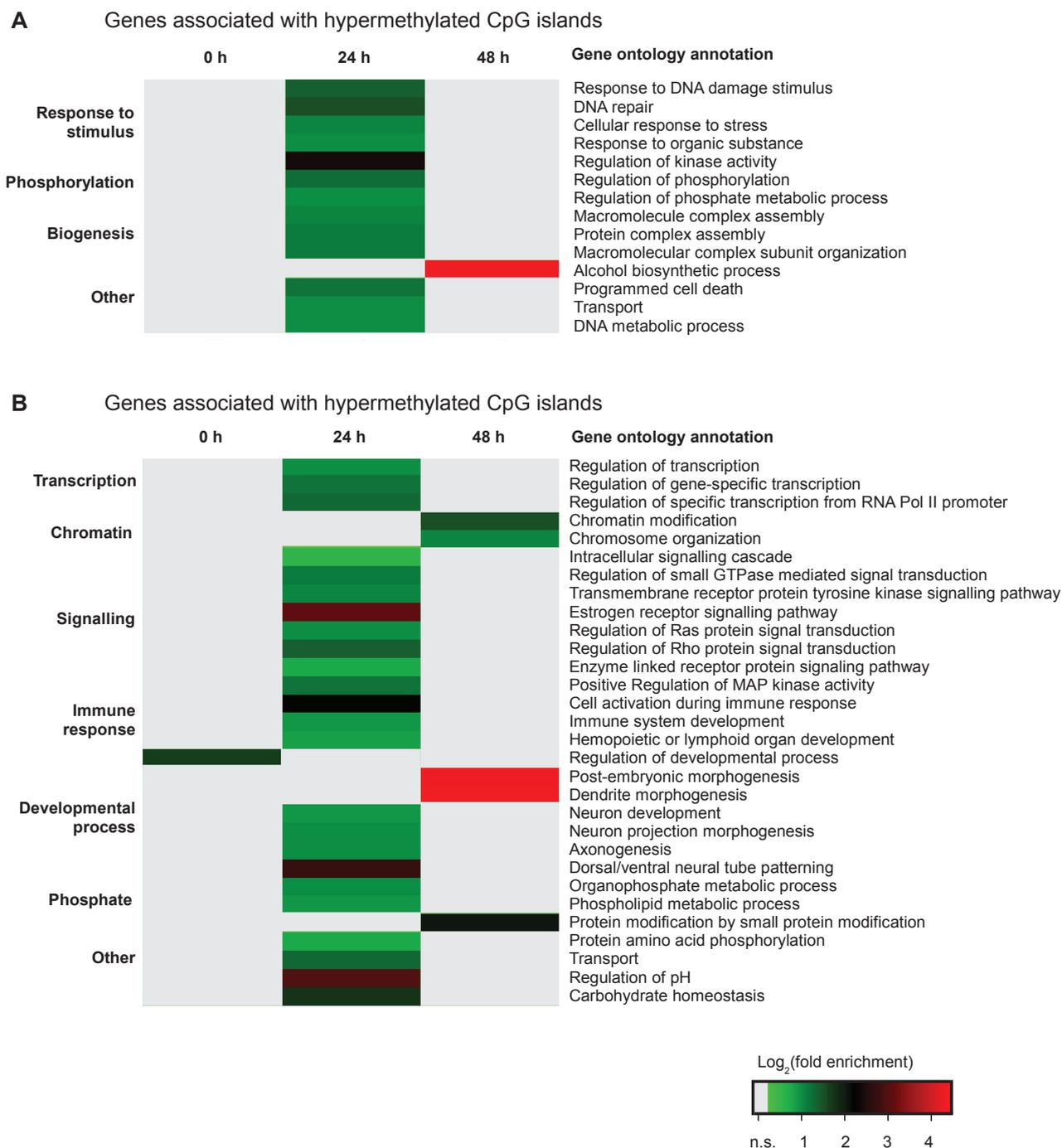


Figure 4.29.: Gene ontology analysis of genes affected by butyrate-mediated differential methylation of associated CpG islands (Wippermann et al. 2013). Log₂ values of fold enrichment are color-coded according to the legend, n.s.: not significantly enriched.

4.2.2. Overlaps between the results of whole-genome bisulfite sequencing and CpG island microarray analysis

The CpG island microarray results were compared to the whole-genome bisulfite sequencing analysis in order to get an impression of the reliability of the custom-designed microarray. Fig. 4.30 shows the comparison of genes detected to be differentially methylated by CpG island microarray analysis with genes detected to be exclusively hypo- *or* hypermethylated, genes that were found to be both hypo- and hypermethylated and, furthermore, genes that were differentially expressed but not differentially methylated. The distinction between the types of differential methylation was made because the presence of both hypo- and hypermethylated cytosines within the same DNA fragment was suspected to potentially lead to biased enrichment reactions. Therefore, genes both hypo- and hypermethylated were hypothesized to be less frequently detected by the CpG island microarray.

448 genes (270 exclusively hypo- *or* hypermethylated genes and 178 genes both hypo- and hypermethylated) that were detected to be differentially methylated by whole-genome bisulfite sequencing were also detected by the CpG island microarray experiment. Obviously, the hypothesis that genes both hypo- and hypermethylated would escape microarray detection because of an indifferent behavior during methylated DNA enrichment did not prove to be true. In summary about 32 % of the microarray result was confirmed by whole-genome bisulfite sequencing. Although this number was apparently low, gene ontology analyses showed enrichment of similar functional categories in both datasets, which included differentiation processes, transcription, signaling and apoptosis.

Interestingly, 356 genes that were detected to be differentially expressed by DNA-microarray analysis but were not found to be differentially methylated by whole-genome bisulfite sequencing were also found to be differentially methylated by the CpG island microarray. This finding might be due to the fact that DNA methylation changes in response to butyrate potentially reflected gene activity. It is obviously possible that gene activity-induced DNA methylation changes were detected by the CpG island microarray but not the whole-genome approach.

4.2.3. Assessment of CpG microarray performance

In order to assess the reliability of differential DNA methylation detection by the CHO-specific CpG island microarray in more detail, exemplary results from the analysis were verified by Combined Bisulfite Restriction Analysis (COBRA; Xiong and Laird 1997). Furthermore, control hybridization experiments were performed in order to be able to assess potential sources of bias originating from the experimental procedure.

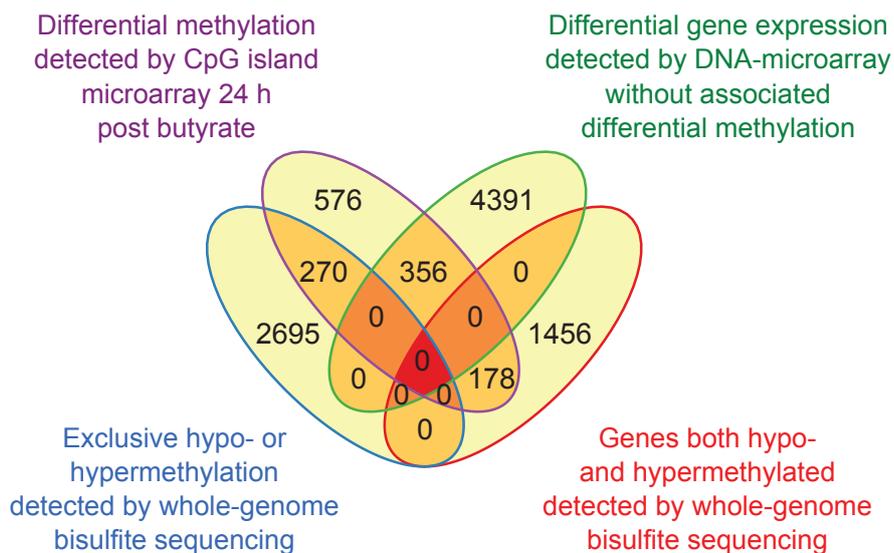


Figure 4.30.: Comparison of genes detected to be differentially methylated and/or differentially expressed by different techniques. Purple: Genes detected to be differentially methylated upon butyrate addition by CpG island microarray analysis. Red and blue: Genes detected to be either exclusively hypo- or hypermethylated or both by whole-genome bisulfite sequencing (including genes that were differentially expressed as well as those only differentially methylated but not differentially expressed). Green: Genes detected to be differentially expressed by DNA-microarrays but not associated with differential methylation.

4.2.3.1. Assessment of CpG microarray performance by Combined Bisulfite Restriction Analysis (COBRA)

Differentially methylated genes detected by CpG island microarrays were analyzed by COBRA (Xiong and Laird 1997). Components of the WNT signaling pathway were selected as examples for this analysis, because several genes that were found to be differentially methylated upon butyrate addition by both CpG island microarray analysis and whole-genome bisulfite sequencing were related to the 'intracellular signaling cascade'.

Butyrate-mediated CpG island methylation was detected within the canonical WNT/Beta-catenin pathway ranging from members of WNT biogenesis and signal transduction in the cytoplasm to transcriptional regulation in the nucleus. Briefly, Porcupine, an ER localized membrane-spanning protein, mediates secretion and palmitoylation of WNT proteins (Hausmann, Bänziger, and Basler 2007; Janda et al. 2012). This modification is essential for their interaction with the Frizzled seven-pass transmembrane protein family and the LDL receptor-related proteins LRP5 and LRP6 (Janda et al. 2012). The formation of a WNT-FZD-LRP complex leads to recruitment of the scaffolding protein Dishevelled, subsequent interaction with the Axin complex and thereby stabilization of Beta-catenin (Logan and Nusse 2004). Beta-catenin moves into the nucleus and regulates gene expression of WNT target genes alongside the transcription factors TCF/LEF (Hoppler and Kavanagh 2007). The subsequent

recruitment of several co-activators regulates a plethora of transcriptional events (MacDonald, Tamai, and He 2009). For example, expression of the WNT responsive gene MYC is activated by Beta-catenin (He et al. 1998). MYC has decisive influence on growth, proliferation and differentiation (Massagué 2004). Its expression was previously shown to be influenced by butyrate (Krupitza et al. 1995; Mariani et al. 2003) and a general influence of butyrate on WNT signaling has been stated before on other cellular levels, e.g. by measuring WNT activity or WNT target gene expression upon butyrate treatment (Bordonaro, Lazarova, and Sartorelli 2008; Shin et al. 2012).

COBRA experiments depend on the presence of specific restriction endonuclease recognition sites within the DNA region to be analyzed. These sites contain CpG dinucleotides that either vanish after bisulfite conversion if the DNA was unmethylated at that specific position, or remain within the DNA sequence if the respective base was methylated. Suitable restriction enzymes are *TaqI* (recognition site 5'-TCGA-3') and *BstUI* (recognition site 5'-CGCG-3'). Due to this requirement not every gene can be analyzed by this technique. However, analysis was possible for the genes *Segment polarity dishevelled homolog DVL-1 (Dvl1)*, *Frizzled-7 (Fzd7)* and *Myc proto-oncogene (Myc)*. *Fzd7* and *Dvl1* were only detected to be differentially methylated by CpG island microarray analysis, whereas differential methylation of *Myc* was detected by both techniques.

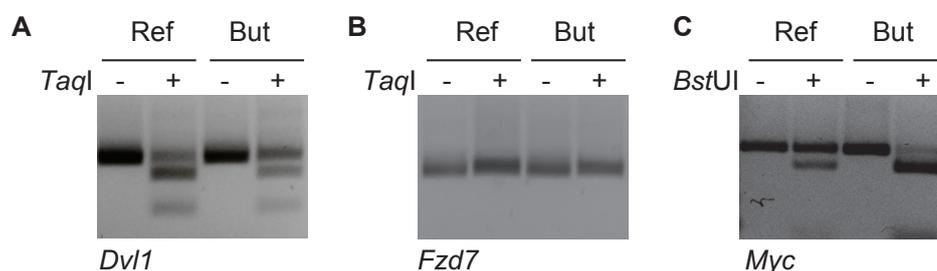


Figure 4.31.: Combined Bisulfite Restriction Analysis of *Segment polarity dishevelled homolog DVL-1 (Dvl1)*, **A**), *Frizzled-7 (Fzd7)*, **B**) and *Myc proto-oncogene (Myc)*, **C**). Promoter regions were amplified from bisulfite-treated DNA of reference and butyrate-treated CHO DP-12 cell cultures (24 h post butyrate). PCR products were digested with *TaqI* (recognition site 5'-TCGA-3') or *BstUI* (recognition site 5'-CGCG-3'). Undigested DNA was used to control the signal strength.

Fig. 4.31 shows the COBRA results as a comparison of reference and butyrate-treated samples at time point 24 hours. Undigested PCR products were used to enable a better estimation of signal strength. For *Fzd7*, both the amplicates from reference and DNA from butyrate-treated cultures were not cut by *TaqI*, which indicates that both DNA regions were not methylated at the cytosine position that was analyzed. This was in accordance with the whole-genome bisulfite sequencing analysis, but not with the CpG island microarray result. Differential DNA methylation could be detected for *Dvl1* and *Myc*, though, which indicated less methylation upon butyrate treatment for *Dvl1*. For *Myc* the experiment indicated a higher methylation level. Both results were in conformity with the changes in CpG island methylation that were

detected by the microarray. On the contrary, differential *Dvll* methylation apparently escaped the whole-genome bisulfite sequencing approach, while differential methylation of *Myc* was also measured by this technique. Taken together, the gene specific analysis showed that two of the three differentially methylated CpG islands that were examined indeed exhibited differential methylation, which confirmed the CpG island microarray result in parts. The results of whole-genome bisulfite sequencing, too, could be verified in parts by this analysis, as two of three results were confirmed as well. Interestingly, this analysis did not allow for a statement regarding a higher reliability of either of the two methods.

4.2.3.2. Assessment of CpG microarray performance by hybridization experiments

In order to be able to assess potential sources of bias originating from the experimental procedure of CpG island microarray analysis, three hybridization experiments were performed in dye swap pairs. The potential sources of bias that were identified in the experimental procedure and which were addressed by this experiment were (i) incomplete purification of DNA during the extraction by phenol-chloroform, (ii) inhomogeneities already present within the input fractions of sheared DNA before enrichment (this experiment was performed during the master studies (Wippermann 2012) but is included here for the sake of completeness) and (iii) variations in the procedure of methylated DNA enrichment.

The first experiment addressed a potential variability introduced by DNA extraction using phenol-chloroform. This step was suspected to be affected by several experimental factors such as incomplete proteolysis by Proteinase K or impurities in the DNA samples due to carryovers during phase separation. This could in turn lead to maintenance of histones or other DNA-binding proteins at some DNA regions and affect the subsequent methylated DNA enrichment reaction. According to this assumption, this experimental step was hypothesized to be the one with the possibly strongest impact on the detection of differential DNA methylation. In order to control the impact of DNA extraction on CpG island microarray results, two separate DNA extractions were performed from the same sample of CHO DP-12 cells and analyzed by the CHO-specific CpG island microarray. The MA plot in Fig. 4.32A shows the result of this comparison. 40 probes indicated significant differences in signal intensities and therefore falsely detected differential methylation. These probes exhibited A-values ≥ 7.0 and M-values ≥ 1.0 or ≤ -1.0 . A-values represent the logarithmic product of red and green signal intensities (\log_2RG), whereas M-values represent the signal ratio (\log_2R/G). That is, the M-value represents the difference in enrichment of a specific region from two samples. The threshold for M-values to be considered significant was generally set to $M = 1.0$ (representing a two-fold difference in enrichment).

The second experiment addressed potential variations within input fraction of sheared DNA from CHO DP-12 cells that were cultivated under different conditions. For this experiment

DNA samples (see images of control agarose gels in Supplementary Fig. B.21) from reference and butyrate-treated CHO DP-12 cell cultures were compared. No significant differences in microarray signals were expected for this comparison, as each sample contained the entire CHO DP-12 genome. Indeed, no significant differences between the two input fractions from untreated and butyrate-treated CHO DP-12 cell cultures were detected (Tab. 4.8). The MA plot in Fig. 4.32B shows that all values centered around a mean intensity of $A = 10$ and M -values range from 0.5 to -0.5.

In the third experiment, two distinct enrichment reactions were performed from the same sample of sheared DNA to control a potential bias introduced by methylated DNA enrichment. As the sheared DNA samples contained 500 bp fragments of the genomes of millions of cells, they were supposed to be very homogeneous regarding the distribution of fragments. Therefore, it was expected that two individual enrichment reactions from the same sample would lead to very similar fractions of enriched methylated DNA. According to this expectation, no significantly differentially enriched CpG islands were found. The corresponding MA plot in Fig. 4.32C shows only few signals above the threshold of $A = 7.0$ (represented in green or red). According to the Student's t-test these were not statistically significant.

Table 4.8.: Significant differences (p adjusted ≤ 0.05) between technical control hybridizations.

Sample 1	Sample 2	$M \geq 1.0$	$M \leq -1.0$
Input (untreated)	Input (butyrate-treated)	0	0
Extraction 1	Extraction 2	40	0
Enrichment 1	Enrichment 2	0	0

In summary, the control experiments to assess the three variables suspected to be able to have a strong impact on reliable CpG island microarray performance showed that only the DNA extraction procedure led to false positive results. However, only a small number of probes was falsely detected as differentially methylated (40 probes representing only 0.15 % of the total number of probes present on the array). DNA extraction most probably has an effect on binding of the MBD protein to specific DNA sites and not on the general composition of the extracted DNA, as even sheared fractions from different samples were found to be completely homogeneous. A general bias introduced by the enrichment procedure itself could be excluded as well.

4.2.4. Discussion: Comparable overall results of CpG island microarray and whole-genome methylation analyses

The analysis of differentially methylated regions by CpG island microarrays showed 1,221 genes to be affected 24 hours upon butyrate addition. When gene ontology analyses were carried out to get an overview of involved cellular processes and components, the functions

4.2. Comparison of whole-genome bisulfite sequencing and CpG island microarray analysis

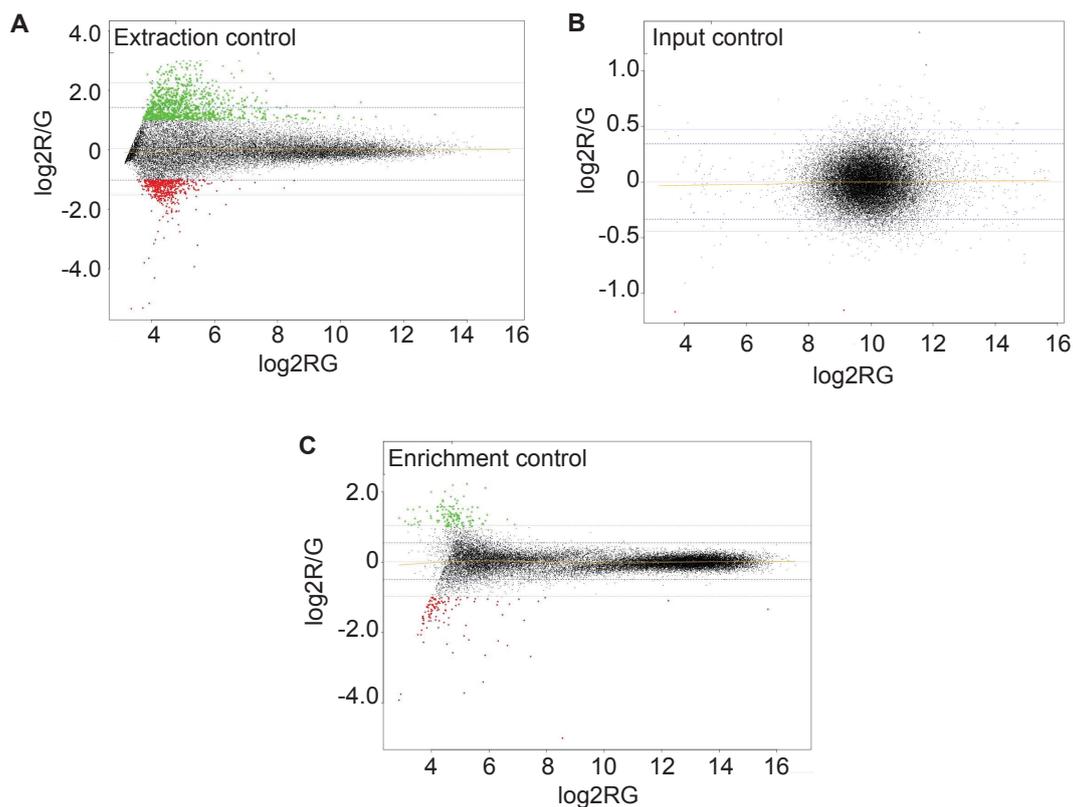


Figure 4.32.: MA plots of quality control hybridization experiments. Data analysis was performed using *EMMA2*. Data were LOWESS normalized and Student's t-tests were calculated. MA plots were generated from these tests. M-values below the threshold of -1.0 (red) represent *de novo* methylated regions. M-values above the threshold of 1.0 (green) represent hypomethylated regions. The blue lines mark for each time point the 95 % and 99 % confidence intervals of all observed M-values. The yellow line represents the LOWESS fit. **A** Extraction control. Comparison of enriched methylated DNA fractions from separately extracted DNA from the same cell sample. **B** Input control (Wippermann 2012). Comparison of input fractions from reference and butyrate-treated CHO DP-12 cell cultures without enrichment of methylated DNA. **C** Enrichment control. Comparison of enriched methylated DNA fractions from the same sample of sheared DNA.

of genes found to be associated with differentially methylated CpG islands were similar to those detected by whole-genome bisulfite sequencing and DNA-microarray analysis, including response reactions, protein biosynthesis, cell death, DNA metabolism, transcription, developmental processes, phosphate metabolism, signaling and transport. More than 50 % of the genes found to be associated with differentially methylated CpG islands were also detected to be either differentially methylated or differentially expressed or both by whole-genome bisulfite sequencing and DNA-microarray analysis.

The presence of components of 'neuron development', 'synaptic transmission' or 'dendrite morphogenesis' within the genes found to be differentially methylated by CpG island microarray analysis was interesting, as similar functional groups were not found amongst the genes detected to be differentially methylated by the whole-genome approach or the gene

expression microarray. Though butyrate is able to induce differentiation of stem cells into neural cells (Yao et al. 2010), an involvement of such specific factors in butyrate-treated CHO cells has not been reported yet. On the one hand, this finding might be related to the fact that genes involved in developmental processes were found to be both differentially methylated and differentially expressed in the integrative, sequencing-based approach. On the other hand, genes contributing to neuron development and related processes were found to be located within partially methylated domains in CHO DP-12 cells. As methylation patterns are more unstructured within these regions, it is possible that detection by enrichment and microarray hybridization leads to some kind of bias for PMD-genes.

To the best of my knowledge, no publications can be found that show a direct comparison of next generation bisulfite sequencing with enrichment based CpG island microarray hybridization, which complicates an assessment of the CHO CGI-microarray performance in comparison to the whole-genome approach. This lack of literature is probably due to the plethora of factors that complicate such comparisons. However, some publications compare bisulfite sequencing with the *Illumina Infinium HumanMethylation27* assay, which is based on bisulfite conversion of DNA followed by hybridization to a microarray containing 27,000 CpG sites. In a comparison of the *Infinium* assay with different genome-wide, sequencing-based methods Bock et al. showed that only about one-fifth of the differentially methylated regions identified by whole-genome sequencing were detected by the microarray-based method. Between different sequencing methods differences in DMR detection were also shown in this publication (Bock et al. 2010). Interestingly, comparisons between sequencing- and hybridization-based methods show only partial overlaps even when no enrichment procedure is included into the experiment, which is the case for RNA sequencing (RNA-Seq) and transcriptome analysis by microarrays. Zhao et al. analyzed gene expression profiles of human T cells after activation and showed that, although a high correlation between gene expression profiles was present, several genes were found to be differentially expressed only by one method. For example, for the comparison '0 h' and '2 h' 602 genes were detected as regulated by both methods, whereas 911 and 801 genes were uniquely detected by either RNA-Seq or microarray hybridization (Zhao et al. 2014). Black et al. also compared these techniques and applied them to the analysis of dose-response reactions in rats. In this study it was also shown that the overlap between genes detected as differentially expressed was only 30–40 % (Black et al. 2014).

Another important finding of both experiments was the dynamic change of DNA methylation in response to butyrate. CpG island microarray analysis showed that 48 hours upon butyrate supplementation 91 % of the observed differential CpG island methylation that was present at sampling point '24 hours' was not detectable anymore. This is obviously in contrast to the whole-genome bisulfite sequencing analysis which, on the one hand, showed that DMRs were variable and DNA methylation was highly dynamic as well, but, on the other hand, also

showed that DMR numbers were still high 36 hours after butyrate addition. Although the sampling point '48 hours' was used in the microarray experiment, which makes a comparison obviously unprecise, the whole-genome approach allows for the guess that the number of DMRs did not rapidly decrease between 36 and 48 hours post butyrate. Interestingly though, in the whole-genome bisulfite sequencing experiment both genome-wide and region-specific mean methylation levels were found to be decreased 12 hours post butyrate and increased again at later time points as well. This effect might be caused by the hypothesized differences in the kinetics of setting and removing DNA methylation marks. Importantly, the phenomenon was also detected for CpG islands and might therefore be the reason for the apparent dynamics detected by the CpG island microarray experiment.

4.2.5. Discussion: Variation in DNA methylation between identical CHO cell cultures

One result of CpG island microarray analysis could not be compared to the whole-genome approach, namely the occurrence of differentially methylated genes in a comparison between identical cultures that were not treated with butyrate. There are several possible explanations for this interesting finding. First, it could be possible that a statistically relevant occurrence of (false positive) differentially methylated genes might be caused by a biased enrichment reaction. Accordingly, the hybridization experiments for control of experimental parameters showed that DNA extraction led to a small fraction of false positives. However, it is furthermore known that within cell populations or even between individual cells a certain degree of noise can occur within regulatory networks (Chalancon et al. 2012). It was shown that switching between states of DNA methylation can contribute to this noise (Lim and Oudenaarden 2007). The differentially methylated genes detected before butyrate treatment may represent this effect, which could be further magnified by asynchronous cell growth or minimal chronological differences in sampling. Interestingly, it is known that cultured cells can behave very differently over the course of a production process even if they were treated in the same way (Le et al. 2012). The reason for this differential behavior has not been found yet.

The reason for the detection of differential methylation between identical cultures could not be identified based on the data at hand. However, the genes found by the '0 h' comparison were removed from the sets of differentially methylated genes at the '24 h' and '48 h' sampling points, as they were suspected to be potentially unrelated to the butyrate-induced changes in CpG island methylation.

4.2.6. Conclusion: Perspectives of CHO-specific DNA methylation microarrays

Taken together, the comparative analysis of DNA methylation changes upon butyrate addition by CpG island microarrays and whole-genome bisulfite sequencing showed that both techniques were suitable for DNA methylation detection. Although the sequencing-based method was expected to be more reliable, COBRA experiments and control hybridizations proved the reliability of the CpG island microarray as well. For this reason CpG island microarray analyses might indeed be performed for economical reasons when sample numbers would be too high for the whole-genome approach. In such experiments it could be advisable to change the genomic sites covered by the array, as CpG islands are obviously not exclusively targeted by DNA methylation changes, but rather keep their methylation state under many conditions (Schübeler 2015). Besides the possibility of broadening the coverage to the whole CHO genome or at least to regions with annotated genes, it would furthermore be possible to create arrays for specific questions that could include genes and regulatory regions belonging to certain pathways. An advantage of such an array would obviously be a higher coverage of individual genes and thereby a higher statistical power. An exemplary application would be the analysis of DNA methylation in regard to product properties like glycosylation, which was shown here to be potentially influenced by butyrate addition. It would furthermore be possible to add those promoters and enhancer regions to the array which are not associated with CpG islands to achieve a more comprehensive view on gene regulation. Additionally, sequences of recombinant genes and viral promoters should be added to analyze potential changes in DNA methylation directly related to changes in productivity. As an alternative it would also be possible to combine bisulfite-mediated conversion of DNA with microarrays instead of using methylated DNA enrichment for sample preparation. The *Illumina Infinium HumanMethylation BeadChip Array* represents an example of such a combination and is one of the most popular DNA methylation microarray platforms today (see e.g. Daca-Roszak et al. 2015).

4.3. Functional analysis of putative central mediators of the butyrate effect

The analysis of the effect of butyrate on CHO DP-12 cells by both whole-genome bisulfite sequencing and CpG island microarrays showed that DNA methylation patterns changed dynamically. As hypothesized previously (please refer to section 4.1.4.7), these observed dynamics of DNA methylation upon butyrate addition were potentially due to transcription factor binding. However, an involvement of central components of the DNA methylation machinery was considered to be necessary for these changes. Therefore, the CpG island microarray and whole-genome bisulfite sequencing data were subjected to an analysis regarding regulatory components of this machinery. The genes *Dnmt3a* and *Uhrfl*, encoding major mediators of *de novo* and maintenance methylation (Bostick et al. 2007; Cedar and Bergman 2009), were found to be relatively hypomethylated by CpG island microarray analysis at the 24 hours sampling point. Furthermore, the genes *Gadd45g* and *Apobec2*, encoding mediators of DNA demethylation (Barreto et al. 2007; Rai et al. 2008) showed changes in their promoter-methylation. *Dnmt3a* and *Gadd45a* were also found to be differentially methylated in the whole-genome bisulfite sequencing experiment. The DNA-microarray gene expression analysis furthermore showed that DNMT3A expression was downregulated, whereas GADD45A expression was upregulated. As the hypothesis that DNMT3A could have a key role in the cellular response to butyrate addition was strengthened by the fact that silencing of DNMT3A in honey bee larvae led to the development into honey bee queens and, strikingly, this developmental process is induced in nature by feeding a diet called royal jelly (which contains phenyl butyrate; Kucharski et al. 2008), a putative role of DNMT3A in the butyrate-response was analyzed more closely by transient knockdown in RNA interference experiments.

In order to first confirm differential methylation of the *Dnmt3a* gene by a gene specific analysis method, Combined Bisulfite Restriction Analysis (COBRA) was applied. Butyrate-induced downregulation of DNMT3A expression was confirmed by Real-Time PCR and westernblot analysis. Fig. 4.33A shows the COBRA result, which indicated a clear difference in DNA methylation between reference and butyrate-treated cultures, although a higher level of methylation was detected. As only single cytosine sites are analyzed by this method and differential DNA methylation can be very variable anyway, this finding does not necessarily contradict the results of whole-genome bisulfite sequencing and CpG island microarray analysis, which indicated lower methylation of *Dnmt3a* upon butyrate treatment. The expression analysis (Fig. 4.33B and C) showed that DNMT3A abundance was decreased 24 hours as well as 48 hours upon butyrate addition on both the mRNA and protein level by about 60 %. In the reference cultures mRNA and protein levels did not change significantly over this period of time.

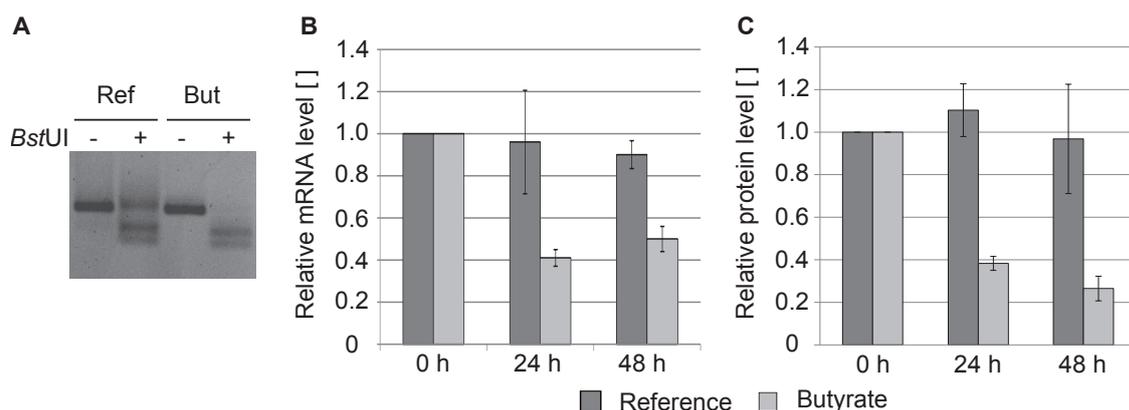


Figure 4.33.: Confirmatory analysis of *Dnmt3a* methylation and expression changes upon butyrate addition. **A** Combined bisulfite restriction analysis (COBRA). The promoter region was amplified from bisulfite-treated DNA of reference and butyrate-treated CHO DP-12 cell cultures (24 h post butyrate). PCR products were digested with *Bst*UI (recognition site 5'-CGCG-3'). Undigested DNA was used to control the signal strength. **B,C** Gene expression analysis of DNMT3A. **B** Real-Time PCR result. Error bars represent standard deviations of triplicate measurements. **C** Westernblot analysis. Error bars represent standard deviations of three separate westernblots from each replicate CHO DP-12 culture. An exemplary westernblot can be found in Supplementary Fig. B.39.

As differential methylation and downregulation of DNMT3A gene expression was confirmed, DNMT3A was considered to be a promising target for RNAi-mediated knockdown. siRNAs targeting 3 different exons (exon 3, 8 and 20, see Fig. 4.34A) of the *Dnmt3a* gene were designed and used to generate 60 nt shRNA-oligonucleotides which formed the typical hairpin structure that is necessary for proper processing and subsequent silencing potential of the released siRNA. Also, a scrambled control, i.e. an siRNA sequence similar in base composition to the actual siRNAs but unspecific for the CHO genome, was used to assess potential effects resulting from the general presence of siRNAs in the cells. The shRNA-oligonucleotides were cloned into the *pLVX-shRNA1* vector, which contains a human Pol III-dependent U6 promoter upstream of the MCS for shRNA expression, and used for transient transfection experiments.

In order to test the knockdown efficiency of the siRNAs designed to target DNMT3A, CHO DP-12 cell cultures were transiently transfected by electroporation and grown for two days. A reference culture was transfected with an empty vector. All cultures exhibited high viabilities until the experiment was terminated. Samples were taken on both days to enable determination of cell specific productivities compared to the reference culture and DNMT3A expression levels were measured by qPCR on the second day. The qPCR analysis (Fig. 4.34B) showed that shRNA2 had a very strong knockdown effect that resulted in the reduction of DNMT3A mRNA levels by 90 %. Transfection of the pLVX-shRNA constructs 3 and 9 led to a reduction in DNMT3A mRNA levels of about 20 %, although the measurements exhibited a large standard deviation and may therefore be not significant. In the cultures transfected with the

scrambled control a significant relative decrease of DNMT3A mRNA abundance by 35 % was observed. Interestingly, the changes in cell specific productivities did not correlate with knockdown efficiencies, as the cultures transfected with shRNA2 only showed an increase in productivities of 2 pg mAB per cell and day compared to the reference, whereas the cultures transfected with shRNAs 3 and 9 and the scrambled control showed a similar increase of about 8 pg mAB per cell and day compared to the reference.

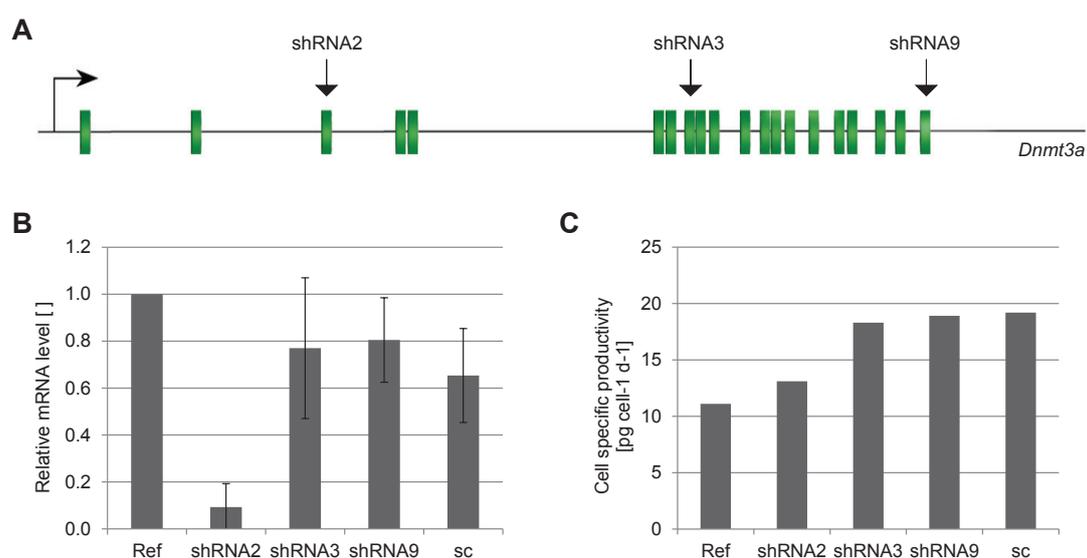


Figure 4.34.: Knockdown of DNMT3A by RNA interference. **A** Schematic representation of *Dnmt3a* gene and indication of exons that were targeted by shRNAs. **B** Relative DNMT3A mRNA abundance in CHO DP12 cells transiently transfected with pU6-shRNA constructs. Reference cultures were transfected with an empty pLVX-shRNA1 vector. Realtime PCR measurements were performed 48 hours post transfection. Values were normalized to the ACTB expression. Error bars represent standard deviations of triplicate measurements. **C** Cell specific productivities were calculated based on antibody titers that were measured by HPLC 24 hours and 48 hours post transfection.

4.3.1. Discussion: DNMT3A could be involved in butyrate-induced hyperproductivity

Transient transfection of CHO DP-12 cell cultures with siRNA constructs targeting DNMT3A led to an apparently successful knockdown of the gene. However, transfection experiments with a scrambled control also had an effect on DNMT3A knockdown. This effect has already been described previously for transient transfection experiments using the same cell line but different target genes (see e.g. Klausung 2013). An explanation for this behavior has not been found yet. However, it seems possible that DNMT3A as a protein involved in many key regulatory processes is generally affected by the process of electroporation and/or the introduction of siRNAs into the cell. Strikingly, a relationship of DNMT3A and small

RNAs was reported in several publications that link the process of antisense RNA-mediated transcriptional gene silencing to the epigenetic machinery (Weinberg and Morris 2016).

Most relevant to this experiment was the finding that decreased expression of DNMT3A, no matter if it was induced by an actual siRNA or by the scrambled control, correlated with an increase in cell specific productivities. Interestingly, the strong knockdown achieved by shRNA2 was associated with the lowest increase in productivities, which might be due to the fact that a drastic change in DNMT3A abundance is likely to affect several important pathways and thereby probably leads to cell damage. A slight decrease in DNMT3A gene expression might only cause a slight impairment of cellular processes and thereby allow for the cells to remain functional. Therefore, shRNAs3 and 9 might be used in further experiments for stable knockdown of DNMT3A. shRNA2 could be tested in an inducible expression system, as it is possible that permanent expression of shRNA2 has a detrimental effect but that a short pulse or controlled expression in specific growth phases might be suitable to enhance cell specific productivities a lot more than observed in this experiment.

5

Summary and outlook

Chinese hamster ovary (CHO) cells represent the major mammalian production cell line for therapeutic proteins and improvement of CHO-based bioprocess performance and product characteristics of biopharmaceuticals, e.g. by molecular engineering approaches, is of great interest. An essential basis for this purpose is the CHO K1 genomic sequence (Xu et al. 2011) and transcriptome data (Becker et al. 2011) that became available in 2011. These data furthermore enable investigations of genotype-phenotype relationships in CHO cells. In this regard the analysis of DNA methylation is of special importance, as this epigenetic process is linked to gene regulation and, importantly, changes in DNA methylation are often associated with environmental factors such as nutrients or the exposure to chemicals (Ruiz-Hernandez et al. 2015).

In order to study DNA methylation phenomena in CHO cell batch cultures, whole-genome bisulfite sequencing was implemented here for the antibody-producing CHO DP-12 cell line. This was achieved by customization of the publicly available CHO K1 genome, evaluation of library qualities and selection of a suitable mapping algorithm. Methylome data generated in this way were combined with gene expression measurements. This allowed for the first global analysis of the CHO DNA methylation landscape as well as for the analysis of differential DNA methylation upon butyrate addition. CHO DP-12 cells were found to exhibit global hypomethylation compared to a majority of mammalian methylomes and hypermethylation of CpG-dense regions at gene promoters called CpG islands. Also, partially methylated domains were observed that covered 62% of the CHO DP-12 cell genome and contained functional clusters of genes. Gene expression analysis showed these clusters to be either

highly or weakly expressed with regard to CHO-specific characteristics. When whole-genome bisulfite sequencing and DNA-microarray analyses were applied to analyze the time course of cellular reactions upon butyrate addition, known butyrate-responsive pathways were found to be affected, including cell cycle, splicing and apoptosis. Also, pathways potentially involved in butyrate-induced hyperproductivity such as central energy metabolism and protein biosynthesis were affected. Interestingly, the data suggested reduced biosynthetic capacities but an increase in protein export. Differentially methylated regions were furthermore found to contain binding-site motifs of specific transcription factors with a potentially central role in the butyrate response. Similar effects of butyrate addition on DNA methylation were obtained when data generated by a recently designed, CHO-specific microarray was analyzed.

5.1. Transgene integration into partially methylated domains might be associated with production instability or product heterogeneity

The partially methylated domains (PMDs) that were found in the CHO DP-12 cell methylome represent epigenomic features similar to the methylomes of cancer and placenta cells (see e.g. Schroeder and LaSalle 2013). The gene expression analysis strongly suggested these features to be involved in the repression of tissue-specific genes, as it was e.g. shown for placenta PMDs, too. Interestingly, the repressive effect of CHO DP-12 PMDs was not found to affect genes relevant to CHO cell properties. To fully understand this relationship, it needs to be elucidated how PMDs evolved, if they are generated in a dynamic manner and what they look like on the cellular level. Also, potential functions of PMDs need to be analyzed and the impact that they might have on cellular behavior. In the context of recombinant gene expression it would also be interesting to elucidate if the integration of transgenes into PMDs correlates with production stability or gene expression strength. To answer these questions, single-cell epigenomic studies are necessary. Also, comparative site-directed integration of recombinant genes into both highly methylated domains and partially methylated domains could provide answers regarding a potential influence of the integration site on production stability and possibly also on product heterogeneity.

Single-cell epigenomics are not as established as single-cell genomics studies, which are e.g. employed for the analysis of genomes from unculturable microorganism or intra-tumor genetic heterogeneity (Gawad, Koh, and Quake 2016; Marcy et al. 2007), or single-cell transcriptomics, which e.g. enable the analysis of cell sub-populations in tumors or heterogeneity in gene expression in dependency on circadian rhythms (Kanter and Kalisky 2015; Dalerba et al. 2011; Shalek et al. 2013; Zopf et al. 2013). Nevertheless, single-cell epigenomics methods are currently being implemented to elucidate DNA methylation dynamics during

development or cell-to-cell heterogeneity in cell populations. The most critical factor in this regard is the bisulfite conversion, which affects DNA integrity and thereby leads to low genomic coverages. In 2015 first single-cell methylomes were produced by Farlik et al., who implemented a whole-genome bisulfite sequencing-based protocol and established a bioinformatical data analysis pipeline for the analysis of low-coverage methylome data (Farlik et al. 2015). It was also shown that single-cell analyses of DNA methylation can be combined with whole-genome amplification to achieve higher signals in sequencing, but this combination obviously comes at the cost of introducing another source of bias into the sequencing results (see e.g. Gravina, Ganapathi, and Vijg 2015).

5.2. Integrative analysis of butyrate-induced cellular reactions raised further questions and allowed for novel perspectives

Integration of DNA methylation and gene expression data in order to analyze the effect of butyrate addition on epigenetic phenomena showed that 11% of the differentially methylated genes were also found to be differentially expressed. Questions remain regarding the relevance of DNA methylation changes at genes that were not found to be differentially expressed. As it is not clear how DNA methylation affects gene expression and if it occurs up- or downstream of gene regulation (Schübeler 2015), answering these questions will require further experiments as well. Such experiments would include genome-wide profiling studies of the epigenetic processes of DNA methylation, histone modifications and small RNA-mediated regulation and their impact on gene expression changes. Furthermore, gene-specific approaches would need to be applied to determine functional relationships between epigenetic phenomena and gene regulation. It is also necessary to refine bioinformatics analysis pipelines in order to be able to cope with the increasing amount of whole-genome data and to efficiently integrate these data to create maps of multiple regulatory layers that might provide further clues.

Generally, the integrative approach for data analysis allowed for a novel perspective regarding the selection of potential target genes for rational cell line engineering, as it enabled identification of transcription factors with potentially central roles within the butyrate-responsive network. Follow-up experiments should address the roles of these transcription factors, e.g. by mutagenization of binding sites (see e.g. Whitfield et al. 2012) or over-expression (see e.g. Deniaud et al. 2009). Furthermore, DNA patterns at putative transcription factor binding sites could be modified by the very recently developed technique of targeted DNA methylation. Targeted DNA methylation repurposes the CRISPR/Cas9 system by using a deactivated Cas9 nuclease fused to the catalytic domain of DNMT3A, which is targeted to specific genomic

loci through a co-expressed guide RNA (Vojta et al. 2016). Vojta and co-authors showed that the system was capable of specific methylation of larger genomic regions by targeting multiple adjacent sites, that the newly generated DNA methylation pattern was heritable through mitosis and that it could repress gene activity. This system would be very suitable for verification of direct regulatory functions of DNA methylation marks at specific genomic sites.

Regarding a potential role of DNMT3A in the butyrate response and the inverse correlation of DNMT3A expression and cell specific productivity, further experiments would be necessary. These should, on the one hand, elucidate if DNMT3A is directly or indirectly affected by butyrate addition and how its interaction network looks like. On the other hand, the relationship between DNMT3A expression and cell specific productivity should be closer examined by employment of inducible expression systems, stable integration of RNAi-expression vectors for DNMT3A knockdown and characterization of cell behavior in fed-batch or perfusion processes.

5.3. General considerations regarding CHO cell metabolism, cultivation parameters and potential epigenetic effects

The results of this thesis are suitable to serve as a basis for future studies of DNA methylation in CHO cells. Results obtained by epigenomic studies might be useful for improvement of production processes of biopharmaceuticals, as they could provide novel targets for rational cell line development. DNA methylation marks might also serve as biomarkers useful for the enrichment of stable, high producing clones during cell line development or to predict the outcomes of production-scale processes (see e.g. Le et al. 2012). Lastly, epigenome-wide studies might offer explanations for molecular mechanisms underlying the unique properties of CHO cells, which are still poorly understood. This was underlined by the insights gained from the establishment of the first DNA methylation map of a CHO cell line that was performed here. The results obtained from the integrative analysis of DNA methylation and gene expression in CHO cells furthermore underlined that environmental conditions might be directly linked to epigenome-wide CHO cell characteristics and, thereby, influence CHO cell metabolism and performance. In the following, general aspects regarding epigenetics, metabolism and the influence of environmental conditions are considered more generally in the context of the production of biopharmaceuticals.

An influence of metabolism on gene expression is an accepted phenomenon and epigenetic processes were shown to be related to it several times. For example, patients suffering

from type 2 diabetes exhibited several changes of DNA methylation in genomic regions associated with genes relevant for carbohydrate metabolism and glycan degradation (Nilsson et al. 2014). Prenatal malnutrition, as another example, was shown to be associated with differential DNA methylation of genes related to growth and metabolism (Tobi et al. 2014). In CHO cells metabolic capacities are obviously related to growth and productivity. In this context it is interesting to have a closer look at lactate, a metabolite that has long been known to affect the height of cell densities and final titers in production processes and is generated from pyruvate during aerobic glycolysis (Lunt and Vander Heiden 2011). In this thesis, a butyrate-induced switch from lactate formation to its consumption correlated with changes in cell specific productivity, too. Variations in the rate of lactate formation and the switch to its consumption between cultures of the same strain under the same conditions were furthermore shown by Le et al., who applied multivariate analysis to a large amount of bioprocess data from 243 bioreactor runs (Le et al. 2012). It was shown that the rate of lactate formation and the switch to its consumption varied even between cultures of the same strain under the same conditions. The hypothesis was raised that this finding was a hint towards a predetermination of metabolic characteristics in early stages of culturing. Adding to the hypothesis that epigenetic phenomena might be responsible for metabolic changes in CHO cells is the fact that acetyl-CoA and other energy metabolites including S-adenosyl methionine (SAM), nicotinamide adenine dinucleotide (NAD⁺) and adenosine triphosphate (ATP) serve as co-factors for most epigenetic enzymes (Donohoe and Bultman 2012). ATP-citrate lyase (ACL)-dependent production of acetyl-CoA was e.g. shown to contribute to increased histone acetylation in cells responding to growth factor stimulation and during differentiation processes (Wellen et al. 2009). This effect is suspected to be due to acetyl-CoAs regulatory effect on histone acetyl transferases (HATs), which play a key role in cellular function. NAD⁺ is an essential co-factor for class III HDACs (Sauve and Youn 2012). Glucose abundance leading to a high NADH/ NAD⁺ ratio therefore results in decreased HDAC activity, enabling the epigenetic system to 'sense' the nutritional cellular state and respond to changing environmental conditions (Donohoe and Bultman 2012).

Taken together, studies linking epigenetic phenomena to central metabolism might be advisable in order to deepen the understanding of mammalian cultivation systems. Knowledge gained in this way might then be employed for modulation of metabolic processes. In this regard it appears possible to improve the production of biopharmaceuticals by using media compositions or alternative supplements that alter methylation patterns in a way that leads to increased growth and productivity, less apoptotic activity or a more efficient metabolism. Underlining this hypothesis is the fact that addition of 5-azacytidine was shown to be responsible for activation of the CMV promoter in several cell types including CHO cells (Radhakrishnan et al. 2008; Choi et al. 2005). Interestingly, several studies address the question of an epigenetic effect of media components in the context of cultured human cell

lines and reprogramming of stem cells. For example, an influence of supra-physiological folic acid concentrations on *in vitro* cultured human cells was reported (Charles, Johnson, and Belshaw 2012). The micronutrient folate is required for the generation of SAM, which is the primary methyl group donor not only for DNA, but also for histone, RNA and phospholipid methylation (Donohoe and Bultman 2012). At physiological doses (nanomolar range), an effect of folate on global or gene-specific DNA methylation was not shown in non-transformed cell lines. Strikingly, supra-physiological concentrations of folic acid (micromolar range), as present in commercially available culture media, induced LINE-1 hypomethylation, CpG island hypermethylation and decreased the intracellular SAM:SAH ratio, that is, caused aberrant DNA methylation (Charles, Johnson, and Belshaw 2012). Interestingly, folate is processed to the essential co-factor tetrahydrofolate (THF) by DHFR, which is used for gene amplification in CHO cells (Jiang and Sharfstein 2009). Other studies addressing epigenetic effects of media supplements focus on vitamin C (L-ascorbic acid), which harbors general antioxidant properties but is rarely included in common cell culture media and supplements (Monfort and Wutz 2013). Vitamin C was shown to facilitate the generation of induced pluripotent stem cells with an improved differentiation potential by attenuating epigenetic silencing of an imprinted gene cluster (Stadtfield et al. 2012). Based on the current knowledge it is speculated that this effect might be explained by a vitamin C-based enhancement of key enzymes changing histone and DNA methylation patterns (Monfort and Wutz 2013). It remains to be elucidated if absence of vitamin C in common cell culture media has an effect on enzyme activities in production cell lines.

The cultivation temperature is another critical parameter in biopharmaceutical production processes. Several lines of evidence establish mild hypothermia as a factor positively affecting the production of recombinant proteins (Kou et al. 2011; Vergara et al. 2014). Interestingly, it has been shown that an inverse correlation exists between global DNA methylation levels and body temperature in fish (Varriale and Bernardi 2006). In plants an effect of temperature on DNA methylation has been shown several times (Stępiński 2012; Zhang et al. 2012; Chen et al. 2014) - however, this might obviously be irrelevant for mammalian cells. In many reptilian species, as another example, temperature is the major determinant of sex (Mork, Czerwinski, and Capel 2014). In developing gonads derived from embryos of the *American alligator* that were incubated at different temperatures, it was e.g. shown that genes playing a critical role in sex determination were differentially methylated depending on the respective incubation temperature (Parrott et al. 2014). Proof of temperature-dependent changes of DNA methylation in mammalian cells is, to the best of my knowledge, still lacking. However, analyzing the impact of temperature changes on epigenetic processes would be interesting, especially in regard to pre-culture handling and potential effects of the culture history on subsequent production processes.

6

Bibliography

- ALLFREY, V. G., FAULKNER, R., and MIRSKY, A. E. (1964). "Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis." In: *Proceedings of the National Academy of Sciences of the United States of America* 51, pp. 786–94.
- ALLIS, C. D. and JENUWEIN, T. (2016). "The molecular hallmarks of epigenetic control." In: *Nature reviews. Genetics* 17.8, pp. 487–500.
- ALMEIDA, R. A. de and O'KEEFE, R. T. (2015). "The NineTeen Complex (NTC) and NTC-associated proteins as targets for spliceosomal ATPase action during pre-mRNA splicing." In: *RNA biology* 12.2, pp. 109–14.
- AMELIO, I., INOUE, S., MARKERT, E. K., LEVINE, A. J., KNIGHT, R. A., MAK, T. W., and MELINO, G. (2015). "TAp73 opposes tumor angiogenesis by promoting hypoxia-inducible factor 1 α degradation." In: *Proceedings of the National Academy of Sciences of the United States of America* 112.1, pp. 226–31.
- ANDREWS, S (2010). "FastQC: A Quality Control tool for High Throughput Sequence Data". In: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- ANTEQUERA, F and BIRD, A (1993). "Number of CpG islands and genes in human and mouse." In: *Proceedings of the National Academy of Sciences of the United States of America* 90.24, pp. 11995–9.
- ANTEQUERA, F, BOYES, J, and BIRD, A (1990). "High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines." In: *Cell* 62.3, pp. 503–14.

- APPANAH, R., DICKERSON, D. R., GOYAL, P., GROUDINE, M., and LORINCZ, M. C. (2007). “An unmethylated 3’ promoter-proximal region is required for efficient transcription initiation.” In: *PLoS genetics* 3.2, e27.
- ARCHER, S. Y., JOHNSON, J., KIM, H.-J., MA, Q., MOU, H., DAESETY, V., MENG, S., and HODIN, R. A. (2005). “The histone deacetylase inhibitor butyrate downregulates cyclin B1 gene expression via a p21/WAF-1-dependent mechanism in human colon cancer cells.” In: *American journal of physiology. Gastrointestinal and liver physiology* 289.4, G696–703.
- AVERY, O. T., MACLEOD, C. M., and MCCARTY, M (1944). “Studies on the chemical nature of the substance inducing transformation of pneumococcal types: Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III.” In: *The Journal of experimental medicine* 79.2, pp. 137–58.
- BACHMAN, M., URIBE-LEWIS, S., YANG, X., BURGESS, H. E., IURLARO, M., REIK, W., MURRELL, A., and BALASUBRAMANIAN, S. (2015). “5-Formylcytosine can be a stable DNA modification in mammals.” In: *Nature chemical biology* 11.8, pp. 555–7.
- BAHR, S. M., BORGSCHULTE, T., KAYSER, K. J., and LIN, N. (2009). “Using microarray technology to select housekeeping genes in Chinese hamster ovary cells.” In: *Biotechnology and bioengineering* 104.5, pp. 1041–6.
- BAILEY, T. L. (2011). “DREME: motif discovery in transcription factor ChIP-seq data.” In: *Bioinformatics (Oxford, England)* 27.12, pp. 1653–9.
- BARRETO, G., SCHÄFER, A., MARHOLD, J., STACH, D., SWAMINATHAN, S. K., HANDA, V., DÖDERLEIN, G., MALTRY, N., WU, W., LYKO, F., and NIEHRS, C. (2007). “Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation.” In: *Nature* 445.7128, pp. 671–675.
- BAUER, A. P., LEIKAM, D., KRINNER, S., NOTKA, F., LUDWIG, C., LÄNGST, G., and WAGNER, R. (2010). “The impact of intragenic CpG content on gene expression.” In: *Nucleic acids research* 38.12, pp. 3891–908.
- BECKER, J., HACKL, M., RUPP, O., JAKOBI, T., SCHNEIDER, J., SZCZEPANOWSKI, R., BEKEL, T., BORTH, N., GOESMANN, A., GRILLARI, J., KALTSCHMIDT, C., NOLL, T., PÜHLER, A., TAUCH, A., and BRINKROLF, K. (2011). “Unraveling the Chinese hamster ovary cell line transcriptome by next-generation sequencing”. In: *Journal of Biotechnol* 156.3, pp. 227–235.
- BECKER, J., TIMMERMAN, C., RUPP, O., ALBAUM, S. P., BRINKROLF, K., GOESMANN, A., PÜHLER, A., TAUCH, A., and NOLL, T. (2014). “Transcriptome analyses of CHO cells with the next-generation microarray CHO41K: development and validation by analysing the influence of the growth stimulating substance IGF-1 substitute LongR(3.).” In: *Journal of biotechnology* 178, pp. 23–31.

-
- BECKMANN, T. F., KRÄMER, O., KLAUSING, S., HEINRICH, C., THÜTE, T., BÜNTEMEYER, H., HOFFROGGE, R., and NOLL, T. (2012). “Effects of high passage cultivation on CHO cells: a global analysis.” In: *Applied Microbiology and Biotechnology* 94.3, pp. 659–671.
- BEDOYA-LÓPEZ, A., ESTRADA, K., SANCHEZ-FLORES, A., RAMÍREZ, O. T., ALTAMIRANO, C., SEGOVIA, L., MIRANDA-RÍOS, J., TRUJILLO-ROLDÁN, M. A., and VALDEZ-CRUZ, N. A. (2016). “Effect of Temperature Downshift on the Transcriptomic Responses of Chinese Hamster Ovary Cells Using Recombinant Human Tissue Plasminogen Activator Production Culture.” In: *PloS one* 11.3, e0151529.
- BELL, C. G., TESCHENDORFF, A. E., RAKYAN, V. K., MAXWELL, A. P., BECK, S., and SAVAGE, D. A. (2010). “Genome-wide DNA methylation analysis for diabetic nephropathy in type 1 diabetes mellitus.” In: *BMC medical genomics* 3.1, p. 33.
- BENJAMINI, Y. and HOCHBERG, Y. (1995). “Controlling the false discovery rate: A practical and powerful approach to multiple testing”. In: *J R Stat Soc Series B Stat Methodol* 57.1, pp. 289–300.
- BERKES, C. A. and TAPSCOTT, S. J. (2005). “MyoD and the transcriptional control of myogenesis.” In: *Seminars in cell & developmental biology* 16.4-5, pp. 585–95.
- BERKYUREK, A. C., SUETAKE, I., ARITA, K., TAKESHITA, K., NAKAGAWA, A., SHIRAKAWA, M., and TAJIMA, S. (2014). “The DNA methyltransferase Dnmt1 directly interacts with the SET and RING finger-associated (SRA) domain of the multifunctional protein Uhrf1 to facilitate accession of the catalytic center to hemi-methylated DNA.” In: *The Journal of biological chemistry* 289.1, pp. 379–86.
- BERMAN, B. P., WEISENBERGER, D. J., AMAN, J. F., HINOUE, T., RAMJAN, Z., LIU, Y., NOUSH-MEHR, H., LANGE, C. P. E., DIJK, C. M. van, TOLLENAAR, R. A. E. M., VAN DEN BERG, D., and LAIRD, P. W. (2012). “Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains.” In: *Nature genetics* 44.1, pp. 40–6.
- BIBIKOVA, M., LE, J., BARNES, B., SAEDINIA-MELNYK, S., ZHOU, L., SHEN, R., and GUNDERSON, K. L. (2009). “Genome-wide DNA methylation profiling using Infinium® assay.” EN. In: *Epigenomics* 1.1, pp. 177–200.
- BILICHAK, A. and KOVALCHUK, I. (2017). “The Combined Bisulfite Restriction Analysis (COBRA) Assay for the Analysis of Locus-Specific Changes in Methylation Patterns.” In: *Methods in molecular biology (Clifton, N.J.)* 1456, pp. 63–71.
- BILLY, E., WEGIERSKI, T., NASR, F., and FILIPOWICZ, W. (2000). “Rcl1p, the yeast protein similar to the RNA 3'-phosphate cyclase, associates with U3 snoRNP and is required for 18S rRNA biogenesis.” In: *The EMBO journal* 19.9, pp. 2115–26.

- BIRD, A. P. and SOUTHERN, E. M. (1978). "Use of restriction enzymes to study eukaryotic DNA methylation: I. The methylation pattern in ribosomal DNA from *Xenopus laevis*." In: *Journal of molecular biology* 118.1, pp. 27–47.
- BIRD, A. (2002). "DNA methylation patterns and epigenetic memory." In: *Genes & development* 16.1, pp. 6–21.
- BIRD, A., TAGGART, M., FROMMER, M., MILLER, O. J., and MACLEOD, D. (1985). "A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA". In: *Cell* 40.1, pp. 91–99.
- BIRNEY, E. et al. (2007). "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project." In: *Nature* 447.7146, pp. 799–816.
- BLACK, M. B., PARKS, B. B., PLUTA, L., CHU, T.-M., ALLEN, B. C., WOLFINGER, R. D., and THOMAS, R. S. (2014). "Comparison of microarrays and RNA-seq for gene expression analyses of dose-response experiments." In: *Toxicological sciences : an official journal of the Society of Toxicology* 137.2, pp. 385–403.
- BLACKLEDGE, N. P., LONG, H. K., ZHOU, J. C., KRIAUCIONIS, S., PATIENT, R., and KLOSE, R. J. (2011). "Bio-CAP : a versatile and highly sensitive technique to purify and characterise regions of non-methylated DNA". In: *Analysis* 12, pp. 1–14.
- BLOUIN, J.-M., PENOT, G., COLLINET, M., NACFER, M., FOREST, C., LAURENT-PUIG, P., COUMOUL, X., BAROUKI, R., BENELLI, C., and BORTOLI, S. (2011). "Butyrate elicits a metabolic switch in human colon cancer cells by targeting the pyruvate dehydrogenase complex." In: *International journal of cancer. Journal international du cancer* 128.11, pp. 2591–601.
- BOCK, C. (2012). "Analysing and interpreting DNA methylation data". In: *Nature Reviews Genetics* 13.10, pp. 705–719.
- BOCK, C., REITHER, S., MIKESKA, T., PAULSEN, M., WALTER, J., and LENGAUER, T. (2005). "BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing." In: *Bioinformatics (Oxford, England)* 21.21, pp. 4067–8.
- BOCK, C., TOMAZOU, E. M., BRINKMAN, A. B., MÜLLER, F., SIMMER, F., GU, H., JÄGER, N., GNIRKE, A., STUNNENBERG, H. G., and MEISSNER, A. (2010). "Quantitative comparison of genome-wide DNA methylation mapping technologies." In: *Nature Biotechnology* 28.10, pp. 1106–1114.
- BOFFA, L. C., VIDALI, G, MANN, R. S., and ALLFREY, V. G. (1978). "Suppression of histone deacetylation in vivo and in vitro by sodium butyrate." In: *Journal of Biological Chemistry* 253.10, pp. 3364–3366.

-
- BOLGER, A. M., LOHSE, M., and USADEL, B. (2014). “Trimmomatic: a flexible trimmer for Illumina sequence data.” In: *Bioinformatics (Oxford, England)* 30.15, pp. 2114–20.
- BOOTH, M. J., BRANCO, M. R., FICZ, G., OXLEY, D., KRUEGER, F., REIK, W., and BALASUBRAMANIAN, S. (2012). “Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution.” en. In: *Science (New York, N.Y.)* 336.6083, pp. 934–7.
- BOOTH, M. J., MARSICO, G., BACHMAN, M., BERALDI, D., and BALASUBRAMANIAN, S. (2014). “Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution.” In: *Nature chemistry* 6.5, pp. 435–40.
- BORDONARO, M., LAZAROVA, D. L., and SARTORELLI, A. C. (2008). “Butyrate and Wnt signaling: a possible solution to the puzzle of dietary fiber and colon cancer risk?” In: *Cell Cycle* 7.9, pp. 1178–1183.
- BOSTICK, M., KIM, J. K., ESTÈVE, P.-O., CLARK, A., PRADHAN, S., and JACOBSEN, S. E. (2007). “UHRF1 plays a role in maintaining DNA methylation in mammalian cells.” In: *Science* 317.5845, pp. 1760–1764.
- BRAGANÇA, J., SWINGLER, T., MARQUES, F. I. R., JONES, T., ELORANTA, J. J., HURST, H. C., SHIODA, T., and BHATTACHARYA, S. (2002). “Human CREB-binding protein/p300-interacting transactivator with ED-rich tail (CITED) 4, a new member of the CITED family, functions as a co-activator for transcription factor AP-2.” In: *The Journal of biological chemistry* 277.10, pp. 8559–65.
- BRENET, F., MOH, M., FUNK, P., FEIERSTEIN, E., VIALE, A. J., SOCCI, N. D., and SCANDURA, J. M. (2011). “DNA methylation of the first exon is tightly linked to transcriptional silencing.” In: *PloS one* 6.1, e14524.
- BRETONES, G., DELGADO, M. D., and LEÓN, J. (2015). “Myc and cell cycle control.” In: *Biochimica et biophysica acta* 1849.5, pp. 506–16.
- BRINKROLF, K. et al. (2013). “Chinese hamster genome sequenced from sorted chromosomes.” en. In: *Nature biotechnology* 31.8, pp. 694–5.
- BROWN, A. J., SWEENEY, B., MAINWARING, D. O., and JAMES, D. C. (2015). “NF- κ B, CRE and YY1 elements are key functional regulators of CMV promoter-driven transient gene expression in CHO cells.” In: *Biotechnology journal* 10.7, pp. 1019–28.
- CARLAGE, T., HINCAPIE, M., ZANG, L., LYUBARSKAYA, Y., MADDEN, H., MHATRE, R., and HANCOCK, W. S. (2009). “Proteomic profiling of a high-producing Chinese hamster ovary cell culture.” In: *Analytical chemistry* 81.17, pp. 7357–62.
- CEDAR, H and BERGMAN, Y. (2012). “Programming of DNA Methylation Patterns”. In: *Annual Review of Biochemistry* February, pp. 97–117.

- CEDAR, H. and BERGMAN, Y. (2009). "Linking DNA methylation and histone modification: patterns and paradigms." In: *Nature Reviews Genetics* 10.5, pp. 295–304.
- CHALANCON, G., RAVARANI, C. N. J., BALAJI, S, MARTINEZ-ARIAS, A., ARAVIND, L, JOTHI, R., and BABU, M. M. (2012). "Interplay between gene expression noise and regulatory network architecture." In: *Trends in Genetics* 28.5, pp. 221–232.
- CHANARAT, S. and STRÄSSER, K. (2013). "Splicing and beyond: the many faces of the Prp19 complex." In: *Biochimica et biophysica acta* 1833.10, pp. 2126–34.
- CHANG, L. and BARFORD, D. (2014). "Insights into the anaphase-promoting complex: a molecular machine that regulates mitosis." In: *Current opinion in structural biology* 29, pp. 1–9.
- CHARLES, M. A., JOHNSON, I. T., and BELSHAW, N. J. (2012). "Supra-physiological folic acid concentrations induce aberrant DNA methylation in normal human cells in vitro." In: *Epigenetics : official journal of the DNA Methylation Society* 7.7, pp. 689–94.
- CHEDIN, F., LIEBER, M. R., and HSIEH, C.-L. (2002). "The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a." In: *Proceedings of the National Academy of Sciences of the United States of America* 99.26, pp. 16916–21.
- CHEN, F., KOU, T., FAN, L., ZHOU, Y., YE, Z., ZHAO, L., and TAN, W.-S. (2011). "The combined effect of sodium butyrate and low culture temperature on the production, sialylation, and biological activity of an antibody produced in CHO cells". In: *Biotechnology and Bioprocess Engineering* 16.6, pp. 1157–1165.
- CHEN, X., HU, J., ZHANG, H., and DING, Y. (2014). "DNA methylation changes in photoperiod-thermo-sensitive male sterile rice PA64S under two different conditions." In: *Gene* 537.1, pp. 143–8.
- CHEN, Y.-A., LEMIRE, M., CHOUFANI, S., BUTCHER, D. T., GRAFODATSKAYA, D., ZANKE, B. W., GALLINGER, S., HUDSON, T. J., and WEKSBERG, R. (2013). "Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray." In: *Epigenetics* 8.2, pp. 203–9.
- CHEN, Z., CLARK, S., BIRKELAND, M., SUNG, C. M., LAGO, A., LIU, R., KIRKPATRICK, R., JOHANSON, K., WINKLER, J. D., and HU, E. (2002). "Induction and superinduction of growth arrest and DNA damage gene 45 (GADD45) α and β messenger RNAs by histone deacetylase inhibitors trichostatin A (TSA) and butyrate in SW620 human colon carcinoma cells." In: *Cancer Letters* 188.1-2, pp. 127–140.
- CHENG, X. and BLUMENTHAL, R. M. (2008). "Mammalian DNA methyltransferases: a structural perspective." In: *Structure (London, England : 1993)* 16.3, pp. 341–50.

-
- CHI, J.-T. et al. (2006). “Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers.” In: *PLoS medicine* 3.3, e47.
- CHITTKA, A. and CHITTKA, L. (2010). “Epigenetics of royalty.” In: *PLoS biology* 8.11, e1000532.
- CHOI, K. H., BASMA, H., SINGH, J., and CHENG, P.-W. (2005). “Activation of CMV promoter-controlled glycosyltransferase and beta -galactosidase glycogenes by butyrate, trichostatin A, and 5-aza-2'-deoxycytidine.” In: *Glycoconjugate journal* 22.1-2, pp. 63–9.
- CHONG, W. P. K., GOH, L. T., REDDY, S. G., YUSUFI, F. N. K., LEE, D. Y., WONG, N. S. C., HENG, C. K., YAP, M. G. S., and HO, Y. S. (2009). “Metabolomics profiling of extracellular metabolites in recombinant Chinese Hamster Ovary fed-batch culture.” In: *Rapid communications in mass spectrometry : RCM* 23.23, pp. 3763–71.
- CHU, C., QU, K., ZHONG, F. L., ARTANDI, S. E., and CHANG, H. Y. (2011). “Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions.” In: *Molecular cell* 44.4, pp. 667–78.
- CHU, P.-C., YANG, Y.-C., LU, Y.-T., CHEN, H.-T., YU, L.-C., and CHANG, M.-S. (2006). “Silencing of p29 affects DNA damage responses with UV irradiation.” In: *Cancer research* 66.17, pp. 8484–91.
- CLOUGH, D., KUNKEL, L., and DAVIDSON, R. (1982). “5-Azacytidine-induced reactivation of a herpes simplex thymidine kinase gene”. en. In: *Science* 216.4541, pp. 70–73.
- CORFE, B. M. (2012). “Hypothesis: butyrate is not an HDAC inhibitor, but a product inhibitor of deacetylation.” en. In: *Molecular bioSystems* 8.6, pp. 1609–12.
- CRUZ-TOPETE, D., HE, B., XU, X., and CIDLOWSKI, J. A. (2016). “Krüppel-like Factor 13 Is a Major Mediator of Glucocorticoid Receptor Signaling in Cardiomyocytes and Protects These Cells from DNA Damage and Death.” In: *The Journal of biological chemistry* 291.37, pp. 19374–86.
- DACA-ROSZAK, P., PFEIFER, A., ŻEBRACKA-GALA, J., RUSINEK, D., SZYBIŃSKA, A., JARZĄB, B., WITT, M., and ZIĘTKIEWICZ, E. (2015). “Impact of SNPs on methylation readouts by Illumina Infinium HumanMethylation450 BeadChip Array: implications for comparative population studies.” En. In: *BMC genomics* 16.1, p. 1003.
- DALERBA, P. et al. (2011). “Single-cell dissection of transcriptional heterogeneity in human colon tumors.” In: *Nature biotechnology* 29.12, pp. 1120–7.
- DANG, C. V., KIM, J.-w., GAO, P., and YUSTEIN, J. (2008). “The interplay between MYC and HIF in cancer”. In: *Nature Reviews Cancer* 8.1, pp. 51–56.
- D'ANNA, J. A., TOBEY, R. A., and GURLEY, L. R. (1980). “Concentration-dependent effects of sodium butyrate in Chinese hamster cells: cell-cycle progression, inner-histone acetylation, histone H1 dephosphorylation, and induction of an H1-like protein.” In: *Biochemistry* 19.12, pp. 2656–71.

- DE LEON GATTI, M., WLASCHIN, K. F., NISSOM, P. M., YAP, M., and HU, W.-S. (2007). “Comparative transcriptional analysis of mouse hybridoma and recombinant Chinese hamster ovary cells undergoing butyrate treatment.” In: *Journal of bioscience and bioengineering* 103.1, pp. 82–91.
- DE MEYER, T., BADY, P., TROOSKENS, G., KURSCHEID, S., BLOCH, J., KROS, J. M., HAINFELLNER, J. A., STUPP, R., DELORENZI, M., HEGI, M. E., and VAN CRIEKINGE, W. (2015). “Genome-wide DNA methylation detection by MethylCap-seq and Infinium HumanMethylation450 BeadChips: an independent large-scale comparison.” In: *Scientific reports* 5, p. 15375.
- DEANS, C. and MAGGERT, K. A. (2015). “What do you mean, "epigenetic"?” In: *Genetics* 199.4, pp. 887–96.
- DEATON, A. M. and BIRD, A. (2011). “CpG islands and the regulation of transcription.” In: *Genes & development* 25.10, pp. 1010–22.
- DENG, Z., WAN, M., CAO, P., RAO, A., CRAMER, S. D., and SUI, G (2009). “Yin Yang 1 regulates the transcriptional activity of androgen receptor.” In: *Oncogene* 28.42, pp. 3746–57.
- DENIAUD, E., BAGUET, J., CHALARD, R., BLANQUIER, B., BRINZA, L., MEUNIER, J., MICHALLET, M.-C., LAUGRAUD, A., AH-SOON, C., WIERINCKX, A., CASTELLAZZI, M., LACHUER, J., GAUTIER, C., MARVEL, J., and LEVERRIER, Y. (2009). “Overexpression of transcription factor Sp1 leads to gene expression perturbations and cell cycle inhibition.” In: *PloS one* 4.9, e7035.
- DHANASEKARAN, D. N. and REDDY, E. P. (2008). “JNK signaling in apoptosis.” In: *Oncogene* 27.48, pp. 6245–51.
- DIAS, B. G. and RESSLER, K. J. (2014). “Parental olfactory experience influences behavior and neural structure in subsequent generations.” In: *Nature neuroscience* 17.1, pp. 89–96.
- DOERFLER, W (1983). “DNA methylation and gene activity.” en. In: *Annual review of biochemistry* 52, pp. 93–124.
- DONDRUP, M. et al. (2009). “EMMA 2 - a MAGE-compliant system for the collaborative analysis and integration of microarray data.” In: *BMC Bioinformatics* 10, p. 50.
- DONOHOE, D. R. and BULTMAN, S. J. (2012). “Metaboloepigenetics: interrelationships between energy metabolism and epigenetic control of gene expression.” In: *Journal of cellular physiology* 227.9, pp. 3169–77.
- DONOHOE, D. R., CURRY, K. P., and BULTMAN, S. J. (2013). *Microbial Oncotarget: Bacterial-Produced Butyrate, Chemoprevention and Warburg Effect*.
- DURRIN, L. K., MANN, R. K., KAYNE, P. S., and GRUNSTEIN, M (1991). “Yeast histone H4 N-terminal sequence is required for promoter activation in vivo.” In: *Cell* 65.6, pp. 1023–31.

-
- ECKER, J. R., BICKMORE, W. A., BARROSO, I., PRITCHARD, J. K., GILAD, Y., and SEGAL, E. (2012). “Genomics: ENCODE explained.” In: *Nature* 489.7414, pp. 52–5.
- ECKHARDT, F. et al. (2006). “DNA methylation profiling of human chromosomes 6, 20 and 22.” In: *Nature genetics* 38.12, pp. 1378–85.
- EDDY, S. R. (2012). “The C-value paradox, junk DNA and ENCODE.” In: *Current biology : CB* 22.21, R898–9.
- ENTREVAN, M., SCHUETTENGRUBER, B., and CAVALLI, G. (2016). “Regulation of Genome Architecture and Function by Polycomb Proteins.” In: *Trends in cell biology* 26.7, pp. 511–25.
- FARLIK, M., SHEFFIELD, N. C., NUZZO, A., DATLINGER, P., SCHÖNEGGER, A., KLUGHAMMER, J., and BOCK, C. (2015). “Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics.” English. In: *Cell reports* 10.8, pp. 1386–97.
- FATHALLAH, H., WEINBERG, R. S., GALPERIN, Y., SUTTON, M., and ATWEH, G. F. (2007). “Role of epigenetic modifications in normal globin gene regulation and butyrate-mediated induction of fetal hemoglobin.” en. In: *Blood* 110.9, pp. 3391–7.
- FEICHTINGER, J. et al. (2016). “Comprehensive genome and epigenome characterization of CHO cells in response to evolutionary pressures and over time.” In: *Biotechnology and bioengineering* 113, pp. 2241–53.
- FEIL, R. and BERGER, F. (2007). “Convergent evolution of genomic imprinting in plants and mammals.” In: *Trends in genetics : TIG* 23.4, pp. 192–9.
- FELSENFELD, G. (2014). “A brief history of epigenetics.” In: *Cold Spring Harbor perspectives in biology* 6.1.
- FERRETTI, C, BRUNI, L, DANGLES-MARIE, V, PECKING, A. P., and BELLET, D (2006). “Molecular circuits shared by placental and cancer cells, and their implications in the proliferative, invasive and migratory capacities of trophoblasts.” In: *Human reproduction update* 13.2, pp. 121–41.
- FLEMING, S. E., FITCH, M. D., DEVRIES, S, LIU, M. L., and KIGHT, C (1991). “Nutrient utilization by cells isolated from rat jejunum, cecum and colon.” In: *The Journal of nutrition* 121.6, pp. 869–78.
- FONG, K.-w., LEUNG, J. W.-c., LI, Y., WANG, W., FENG, L., MA, W., LIU, D., SONGYANG, Z., and CHEN, J. (2013). “MTR120/KIAA1383, a novel microtubule-associated protein, promotes microtubule stability and ensures cytokinesis.” In: *Journal of cell science* 126.Pt 3, pp. 825–37.
- FROMMER, M, McDONALD, L. E., MILLAR, D. S., COLLIS, C. M., WATT, F, GRIGG, G. W., MOLLOY, P. L., and PAUL, C. L. (1992). “A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands.” In: *Proceedings of the National Academy of Sciences of the United States of America* 89.5, pp. 1827–31.

- FURUSAWA, Y. et al. (2013). “Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells.” In: *Nature* 504.7480, pp. 446–50.
- GAIDATZIS, D., BURGER, L., MURR, R., LERCH, A., DESSUS-BABUS, S., SCHÜBELER, D., and STADLER, M. B. (2014). “DNA sequence explains seemingly disordered methylation levels in partially methylated domains of Mammalian genomes.” In: *PLoS genetics* 10.2, e1004143.
- GARDINER-GARDEN, M and FROMMER, M (1987). “CpG islands in vertebrate genomes.” In: *Journal of molecular biology* 196.2, pp. 261–82.
- GARZON, R. et al. (2009). “MicroRNA-29b induces global DNA hypomethylation and tumor suppressor gene reexpression in acute myeloid leukemia by targeting directly DNMT3A and 3B and indirectly DNMT1.” In: *Blood* 113.25, pp. 6411–8.
- GAUTIER, T, BERGÈS, T, TOLLERVEY, D, and HURT, E (1997). “Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis.” In: *Molecular and cellular biology* 17.12, pp. 7088–98.
- GAWAD, C., KOH, W., and QUAKE, S. R. (2016). “Single-cell genome sequencing: current state of the science.” In: *Nature reviews. Genetics* 17.3, pp. 175–88.
- GENG, Y., ZHAO, Y., SCHUSTER, L. C., FENG, B., LYNN, D. A., AUSTIN, K. M., STOKLOSA, J. D., and MORRISON, J. D. (2015). “A Chemical Biology Study of Human Pluripotent Stem Cells Unveils HSPA8 as a Key Regulator of Pluripotency.” In: *Stem cell reports* 5.6, pp. 1143–54.
- GETHING, M. J. (1999). “Role and regulation of the ER chaperone BiP.” In: *Seminars in cell & developmental biology* 10.5, pp. 465–72.
- GILL, G and PTASHNE, M (1988). “Negative effect of the transcriptional activator GAL4.” In: *Nature* 334.6184, pp. 721–4.
- GJONESKA, E., PFENNING, A. R., MATHYS, H., QUON, G., KUNDAJE, A., TSAI, L.-H., and KELLIS, M. (2015). “Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer’s disease”. In: *Nature* 518.7539, pp. 365–369.
- GODA, N. and KANAI, M. (2012). “Hypoxia-inducible factors and their roles in energy metabolism.” In: *International journal of hematology* 95.5, pp. 457–63.
- GOMEZ, T. M. and ZHENG, J. Q. (2006). “The molecular basis for calcium-dependent axon pathfinding.” In: *Nature reviews. Neuroscience* 7.2, pp. 115–25.
- GRANDI, P., RYBIN, V., BASSLER, J., PETFALSKI, E., STRAUSS, D., MARZIOCH, M., SCHÄFER, T., KUSTER, B., TSCHOCHNER, H., TOLLERVEY, D., GAVIN, A. C., and HURT, E. (2002). “90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors.” In: *Molecular cell* 10.1, pp. 105–15.

-
- GRAUR, D., ZHENG, Y., PRICE, N., AZEVEDO, R. B. R., ZUFALL, R. A., and ELHAIK, E. (2013). "On the immortality of television sets: "function" in the human genome according to the evolution-free gospel of ENCODE." In: *Genome biology and evolution* 5.3, pp. 578–90.
- GRAVINA, S., GANAPATHI, S., and VIJG, J. (2015). "Single-cell, locus-specific bisulfite sequencing (SLBS) for direct detection of epimutations in DNA methylation patterns." In: *Nucleic acids research* 43.14, e93.
- GROSS, D. N., HEUVEL, A. P. J. van den, and BIRNBAUM, M. J. (2008). "The role of FoxO in the regulation of metabolism." In: *Oncogene* 27.16, pp. 2320–36.
- GRUENBAUM, Y., CEDAR, H., and RAZIN, A. (1982). "Substrate and sequence specificity of a eukaryotic DNA methylase". In: *Nature* 295.5850, pp. 620–622.
- GU, Z., EILS, R., and SCHLESNER, M. (2016). "Complex heatmaps reveal patterns and correlations in multidimensional genomic data." In: *Bioinformatics (Oxford, England)*, btw313–.
- GUO, J. U., SU, Y., SHIN, J. H., SHIN, J., LI, H., XIE, B., ZHONG, C., HU, S., LE, T., FAN, G., ZHU, H., CHANG, Q., GAO, Y., MING, G.-L., and SONG, H. (2014). "Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain." In: *Nature neuroscience* 17.2, pp. 215–22.
- GUO, T., CHUNG, J. H., WANG, T., MCDONALD-MCGINN, D. M., KATES, W. R., HAWUŁA, W., COLEMAN, K., ZACKAI, E., EMANUEL, B. S., and MORROW, B. E. (2015a). "Histone Modifier Genes Alter Conotruncal Heart Phenotypes in 22q11.2 Deletion Syndrome." In: *American journal of human genetics* 97.6, pp. 869–77.
- GUO, W., FIZIEV, P., YAN, W., COKUS, S., SUN, X., ZHANG, M. Q., CHEN, P.-Y., and PELLEGRINI, M. (2013). "BS-Seeker2 : a versatile aligning pipeline for bisulfite sequencing data". In: *BMC Genomics* 14, p. 774.
- GUO, X., WANG, L., LI, J., DING, Z., XIAO, J., YIN, X., HE, S., SHI, P., DONG, L., LI, G., TIAN, C., WANG, J., CONG, Y., and XU, Y. (2015b). "Structural insight into autoinhibition and histone H3-induced activation of DNMT3A." In: *Nature* 517.7536, pp. 640–4.
- GUPTA, S., STAMATOYANNOPOULOS, J. A., BAILEY, T. L., and NOBLE, W. S. (2007). "Quantifying similarity between motifs." En. In: *Genome biology* 8.2, R24.
- GYORGY, A. B., SZEMES, M., DE JUAN ROMERO, C., TARABYKIN, V., and AGOSTON, D. V. (2008). "SATB2 interacts with chromatin-remodeling molecules in differentiating cortical neurons." In: *The European journal of neuroscience* 27.4, pp. 865–73.
- HA, M. and KIM, V. N. (2014). "Regulation of microRNA biogenesis." In: *Nature reviews. Molecular cell biology* 15.8, pp. 509–524.

- HABANO, W., KAWAMURA, K., IIZUKA, N., TERASHIMA, J., SUGAI, T., and OZAWA, S. (2015). “Analysis of DNA methylation landscape reveals the roles of DNA methylation in the regulation of drug metabolizing enzymes.” En. In: *Clinical epigenetics* 7.1, p. 105.
- HAI, Z. and ZUO, W. (2016). “Aberrant DNA methylation in the pathogenesis of atherosclerosis.” In: *Clinica chimica acta; international journal of clinical chemistry* 456, pp. 69–74.
- HAIKOVA, P., JEFFRIES, S. J., LEE, C., MILLER, N., JACKSON, S. P., and SURANI, M. A. (2010). “Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway.” In: *Science (New York, N.Y.)* 329.5987, pp. 78–82.
- HAMER, H. M., JONKERS, D., VENEMA, K., VANHOUTVIN, S, TROOST, F. J., and BRUMMER, R.-J. (2008). “Review article: the role of butyrate on colonic function.” In: *Alimentary pharmacology & therapeutics* 27.2, pp. 104–19.
- HANSEN, R. S. and GARTLER, S. M. (1990). “5-Azacytidine-induced reactivation of the human X chromosome-linked PGK1 gene is associated with a large region of cytosine demethylation in the 5' CpG island.” In: *Proceedings of the National Academy of Sciences of the United States of America* 87.11, pp. 4174–8.
- HATA, K., OKANO, M., LEI, H., and LI, E. (2002). “Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice.” In: *Development (Cambridge, England)* 129.8, pp. 1983–93.
- HAUSMANN, G., BÄNZIGER, C., and BASLER, K. (2007). “Helping Wingless take flight: how WNT proteins are secreted.” en. In: *Nature Reviews Molecular Cell Biology* 8.4, pp. 331–336.
- HAYATSU, H. (2008). “Discovery of bisulfite-mediated cytosine conversion to uracil , the key reaction for DNA methylation analysis - A personal account”. In: *Proceedings of the Japan Academy. Series B, Physical and Biological Sciences* 84.8, pp. 2–11.
- HE, T. C., SPARKS, A. B., RAGO, C, HERMEKING, H, ZAWEL, L, COSTA, L. T. da, MORIN, P. J., VOGELSTEIN, B, and KINZLER, K. W. (1998). “Identification of c-MYC as a target of the APC pathway.” In: *Science* 281.5382, pp. 1509–1512.
- HERMAN, J. G., GRAFF, J. R., MYÖHÄNEN, S, NELKIN, B. D., and BAYLIN, S. B. (1996). “Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands.” In: *Proceedings of the National Academy of Sciences of the United States of America* 93.18, pp. 9821–6.
- HERSHEY, A. D. and CHASE, M (1952). “Independent functions of viral protein and nucleic acid in growth of bacteriophage.” In: *The Journal of general physiology* 36.1, pp. 39–56.
- HERZING, L. B., ROMER, J. T., HORN, J. M., and ASHWORTH, A (1997). “Xist has properties of the X-chromosome inactivation centre.” In: *Nature* 386.6622, pp. 272–5.

-
- HOLLIDAY, R and PUGH, J. E. (1975). “DNA modification mechanisms and gene activity during development.” In: *Science (New York, N.Y.)* 187.4173, pp. 226–32.
- HOLLIDAY, R. (2006). “Epigenetics: A Historical Overview”. In: *Epigenetics* 1.2, pp. 76–80.
- HON, G. C. et al. (2012). “Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer.” In: *Genome research* 22.2, pp. 246–58.
- HONDA, T, TAMURA, G, WAKI, T, KAWATA, S, TERASHIMA, M, NISHIZUKA, S, and MOTOYAMA, T (2004). “Demethylation of MAGE promoters during gastric cancer progression.” In: *British journal of cancer* 90.4, pp. 838–43.
- HONG, J. K., LEE, S. M., KIM, K.-Y., and LEE, G. M. (2014). “Effect of sodium butyrate on the assembly, charge variants, and galactosylation of antibody produced in recombinant Chinese hamster ovary cells.” In: *Applied microbiology and biotechnology* 98.12, pp. 5417–25.
- HOPPLER, S. and KAVANAGH, C. L. (2007). “Wnt signalling: variety at the core.” In: *Journal of Cell Science* 120.Pt 3, pp. 385–393.
- HOSACK, D. A., DENNIS, G., SHERMAN, B. T., LANE, H. C., and LEMPICKI, R. A. (2003). “Identifying biological themes within lists of genes with EASE.” In: *Genome biology* 4.10, R70.
- HOU, H. and YU, H. (2010). “Structural insights into histone lysine demethylation.” In: *Current opinion in structural biology* 20.6, pp. 739–48.
- HU, S., DONG, T. S., DALAL, S. R., WU, F., BISSONNETTE, M., KWON, J. H., and CHANG, E. B. (2011). “The microbe-derived short chain fatty acid butyrate targets miRNA-dependent p21 gene expression in human colon cancer.” In: *PloS one* 6.1, e16221.
- HU, X., KUCUK, C., JIANG, B., KLINKEBIEL, D., GENG, H., GONG, Q., BOUSKA, A., IQBAL, J., GAULARD, P., MCKEITHAN, T. W., and CHAN, W. C. (2015). “Global promoter methylation analysis reveals novel candidate tumor suppressor genes in natural killer cell lymphoma.” In: *Clin Cancer Res* 21, pp. 1699–711.
- HUANG, D. W., SHERMAN, B. T., ZHENG, X., YANG, J., IMAMICHI, T., STEPHENS, R., and LEMPICKI, R. A. (2009). “Extracting biological meaning from large gene lists with DAVID.” In: *Current Protocols in Bioinformatics* Chapter 13, Unit 13.11.
- HUANG, Y.-Z., SUN, J.-J., ZHANG, L.-Z., LI, C.-J., WOMACK, J. E., LI, Z.-J., LAN, X.-Y., LEI, C.-Z., ZHANG, C.-L., ZHAO, X., and CHEN, H. (2014). “Genome-wide DNA Methylation Profiles and Their Relationships with mRNA and the microRNA Transcriptome in Bovine Muscle Tissue (Bos taurine).” en. In: *Scientific reports* 4, p. 6546.
- ILLINGWORTH, R. S. and BIRD, A. P. (2009). “CpG islands -’a rough guide’.” In: *FEBS Letters* 583.11, pp. 1713–1720.

- ITO, S., SHEN, L., DAI, Q., WU, S. C., COLLINS, L. B., SWENBERG, J. A., HE, C., and ZHANG, Y. (2011). "Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine." In: *Science (New York, N.Y.)* 333.6047, pp. 1300–3.
- JACQUEMONT, S. et al. (2011). "Epigenetic modification of the FMR1 gene in fragile X syndrome is associated with differential response to the mGluR5 antagonist AFQ056." In: *Science translational medicine* 3.64, 64ra1.
- JAKOBI, T., BRINKROLF, K., TAUCH, A., NOLL, T., STOYE, J., PÜHLER, A., and GOESMANN, A. (2014). "Discovery of transcription start sites in the Chinese hamster genome by next-generation RNA sequencing." In: *Journal of biotechnology* 190, pp. 64–75.
- JANDA, C. Y., WAGHRAY, D., LEVIN, A. M., THOMAS, C., and GARCIA, K. C. (2012). "Structural basis of Wnt recognition by Frizzled." In: *Science* 337.6090, pp. 59–64.
- JELTSCH, A. (2013). "Oxygen, epigenetic signaling, and the evolution of early life." In: *Trends in biochemical sciences* 38.4, pp. 172–6.
- JIANG, S., WEI, H., SONG, T., YANG, Y., ZHANG, F., ZHOU, Y., PENG, J., and JIANG, S. (2015). "KLF13 promotes porcine adipocyte differentiation through PPAR γ activation." In: *Cell & bioscience* 5, p. 28.
- JIANG, X., YAN, J., WEST, A. A., PERRY, C. A., MALYSHEVA, O. V., DEVAPATLA, S., PRESSMAN, E., VERMEYLEN, F., and CAUDILL, M. A. (2012). "Maternal choline intake alters the epigenetic state of fetal cortisol-regulating genes in humans." In: *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 26.8, pp. 3563–74.
- JIANG, Z and SHARFSTEIN, S. T. (2009). "Characterization of gene localization and accessibility in DHFR-amplified CHO cells." In: *Biotechnology progress* 25.1, pp. 296–300.
- JIANG, Z. and SHARFSTEIN, S. T. (2008). "Sodium butyrate stimulates monoclonal antibody over-expression in CHO cells by improving gene accessibility." In: *Biotechnology and Bioengineering* 100.1, pp. 189–194.
- JJINGO, D., CONLEY, A. B., YI, S. V., LUNYAK, V. V., and JORDAN, I. K. (2012). "On the presence and role of human gene-body DNA methylation." In: *Oncotarget* 3.4, pp. 462–74.
- JONES, P. A. and TAKAI, D (2001). "The role of DNA methylation in mammalian epigenetics." en. In: *Science (New York, N.Y.)* 293.5532, pp. 1068–70.
- JONES, P. A. and TAYLOR, S. M. (1980). "Cellular differentiation, cytidine analogs and DNA methylation." In: *Cell* 20.1, pp. 85–93.
- JURKOWSKA, R. Z., JURKOWSKI, T. P., and JELTSCH, A. (2011). "Structure and function of mammalian DNA methyltransferases." In: *ChemBiochem : a European journal of chemical biology* 12.2, pp. 206–22.

-
- KANEHISA, M and GOTO, S (2000). "KEGG: kyoto encyclopedia of genes and genomes." In: *Nucleic Acids Research* 28.1, pp. 27–30.
- KANTARDJIEFF, A., JACOB, N. M., YEE, J. C., EPSTEIN, E., KOK, Y.-J., PHILP, R., BETENBAUGH, M., and HU, W.-S. (2010). "Transcriptome and proteome analysis of Chinese hamster ovary cells under low temperature and butyrate treatment." In: *Journal of biotechnology* 145.2, pp. 143–59.
- KANTER, I. and KALISKY, T. (2015). "Single cell transcriptomics: methods and applications." English. In: *Frontiers in oncology* 5, p. 53.
- KELLEHER, D. J. and GILMORE, R. (2006). "An evolving view of the eukaryotic oligosaccharyltransferase." In: *Glycobiology* 16.4, 47R–62R.
- KIM, H. J., LEEDS, P., and CHUANG, D.-M. (2009). "The HDAC inhibitor, sodium butyrate, stimulates neurogenesis in the ischemic brain." In: *Journal of neurochemistry* 110.4, pp. 1226–40.
- KIM, M., O'CALLAGHAN, P. M., DROMS, K. A., and JAMES, D. C. (2011). "A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies." In: *Biotechnology and Bioengineering* 108.10, pp. 2434–2446.
- KIM, S., LEE, U. J., KIM, M. N., LEE, E.-J., KIM, J. Y., LEE, M. Y., CHOUNG, S., KIM, Y. J., and CHOI, Y.-C. (2008). "MicroRNA miR-199a* regulates the MET proto-oncogene and the downstream extracellular signal-regulated kinase 2 (ERK2)." In: *The Journal of biological chemistry* 283.26, pp. 18158–66.
- KLAUSING, S. (2013). *Optimierung von CHO Produktionszelllinien: RNAi-vermittelter Gen-knockdown und Untersuchungen zur Klonstabilität.*
- KLAUSING, S., KRÄMER, O., and NOLL, T. (2011). "Bioreactor cultivation of CHO DP-12 cells under sodium butyrate treatment - comparative transcriptome analysis with CHO cDNA microarrays." In: *BMC proceedings* 5 Suppl 8.Suppl 8, P98.
- KOH, M. Y. and POWIS, G. (2012). "Passing the baton: the HIF switch." In: *Trends in biochemical sciences* 37.9, pp. 364–72.
- KOHLI, R. M. and ZHANG, Y. (2013). "TET enzymes, TDG and the dynamics of DNA demethylation." In: *Nature* 502.7472, pp. 472–9.
- KOLAR, S. S. N., BARHOUMI, R., LUPTON, J. R., and CHAPKIN, R. S. (2007). "Docosahexaenoic acid and butyrate synergistically induce colonocyte apoptosis by enhancing mitochondrial Ca²⁺ accumulation." In: *Cancer research* 67.11, pp. 5561–8.
- KOU, T.-C., FAN, L., ZHOU, Y., YE, Z.-Y., LIU, X.-P., ZHAO, L., and TAN, W.-S. (2011). "Detailed understanding of enhanced specific productivity in Chinese hamster ovary cells at low culture temperature." In: *Journal of bioscience and bioengineering* 111.3, pp. 365–9.

- KRESS, C., THOMASSIN, H., and GRANGE, T. (2006). “Active cytosine demethylation triggered by a nuclear receptor involves DNA strand breaks.” In: *Proceedings of the National Academy of Sciences of the United States of America* 103.30, pp. 11112–7.
- KRINNER, S., HEITZER, A., ASBACH, B., and WAGNER, R. (2015). “Interplay of Promoter Usage and Intragenic CpG Content: Impact on GFP Reporter Gene Expression.” In: *Human gene therapy* 26.12, pp. 826–40.
- KRUPITZA, G, HARANT, H, DITTRICH, E, SZEKERES, T, HUBER, H, and DITTRICH, C (1995). “Sodium butyrate inhibits c-myc splicing and interferes with signal transduction in ovarian carcinoma cells.” In: *Carcinogenesis* 16.5, pp. 1199–1205.
- KUBO, N. et al. (2015). “DNA methylation and gene expression dynamics during spermatogonial stem cell differentiation in the early postnatal mouse testis.” En. In: *BMC genomics* 16.1, p. 624.
- KUCHARSKI, R, MALESZKA, J, FORET, S, and MALESZKA, R (2008). “Nutritional control of reproductive status in honeybees via DNA methylation.” In: *Science (New York, N.Y.)* 319.5871, pp. 1827–30.
- KULIS, M. and ESTELLER, M. (2010). “DNA methylation and cancer.” In: *Advances in genetics* 70, pp. 27–56.
- LAIRD, P. W. (2010). “Principles and challenges of genomewide DNA methylation analysis.” In: *Nature Reviews Genetics* 11.3, pp. 191–203.
- LAMKANFI, M, KALAI, M, and VANDENABEELE, P (2004). “Caspase-12: an overview.” In: *Cell death and differentiation* 11.4, pp. 365–8.
- LANDER, E. S. et al. (2001). “Initial sequencing and analysis of the human genome.” In: *Nature* 409.6822, pp. 860–921.
- LANE, M., ROBKER, R. L., and ROBERTSON, S. A. (2014). “Parenting from before conception”. In: *Science* 345.6198, pp. 756–760.
- LANGMEAD, B. and SALZBERG, S. L. (2012). “Fast gapped-read alignment with Bowtie 2.” In: *Nature methods* 9.4, pp. 357–9.
- LASKEY, R. A. and GURDON, J. B. (1970). “Genetic content of adult somatic cells tested by nuclear transplantation from cultured cells.” In: *Nature* 228.5278, pp. 1332–4.
- LASSAR, A. B., BUSKIN, J. N., LOCKSHON, D, DAVIS, R. L., APONE, S, HAUSCHKA, S. D., and WEINTRAUB, H (1989). “MyoD is a sequence-specific DNA binding protein requiring a region of myc homology to bind to the muscle creatine kinase enhancer.” In: *Cell* 58.5, pp. 823–31.

-
- LAURENT, L., WONG, E., LI, G., HUYNH, T., TSIRIGOS, A., ONG, C. T., LOW, H. M., KIN SUNG, K. W., RIGOUTSOS, I., LORING, J., and WEI, C.-L. (2010). "Dynamic changes in the human methylome during differentiation." In: *Genome research* 20.3, pp. 320–31.
- LE, H., KABBUR, S., POLLASTRINI, L., SUN, Z., MILLS, K., JOHNSON, K., KARYPIS, G., and HU, W.-S. (2012). "Multivariate analysis of cell culture bioprocess data-lactate consumption as process indicator." In: *Journal of Biotechnology* 162.2-3, pp. 210–223.
- LEDERKREMER, G. Z. (2009). "Glycoprotein folding, quality control and ER-associated degradation." In: *Current opinion in structural biology* 19.5, pp. 515–23.
- LEE, H.-S. (2015). "Impact of Maternal Diet on the Epigenome during In Utero Life and the Developmental Programming of Diseases in Childhood and Adulthood." In: *Nutrients* 7.11, pp. 9492–507.
- LEE, H.-S., BARRAZA-VILLARREAL, A., HERNANDEZ-VARGAS, H., SLY, P. D., BIESSY, C., RAMAKRISHNAN, U., ROMIEU, I., and HERCEG, Z. (2013). "Modulation of DNA methylation states and infant immune system by dietary supplementation with ω -3 PUFA during pregnancy in an intervention study." In: *The American journal of clinical nutrition* 98.2, pp. 480–7.
- LEE, K. and GOLLAHON, L. S. (2015). "ZSCAN4 and TRF1: A functionally indirect interaction in cancer cells independent of telomerase activity." In: *Biochemical and biophysical research communications* 466.4, pp. 644–9.
- LEE, R. C., FEINBAUM, R. L., and AMBROS, V. (1993). "The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*". In: *Cell* 75.5, pp. 843–854.
- LEE, S. M., KIM, Y.-G., LEE, E. G., and LEE, G. M. (2014). "Digital mRNA profiling of N-glycosylation gene expression in recombinant Chinese hamster ovary cells treated with sodium butyrate." In: *Journal of biotechnology* 171, pp. 56–60.
- LEE, Y. Y., WONG, K. T. K., TAN, J., TOH, P. C., MAO, Y., BRUSIC, V., and YAP, M. G. S. (2009). "Overexpression of heat shock proteins (HSPs) in CHO cells for extended culture viability and improved recombinant protein production." In: *Journal of biotechnology* 143.1, pp. 34–43.
- LEONHARDT, H, PAGE, A. W., WEIER, H. U., and BESTOR, T. H. (1992). "A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei." In: *Cell* 71.5, pp. 865–73.
- LI, E, BESTOR, T. H., and JAENISCH, R (1992). "Targeted mutation of the DNA methyltransferase gene results in embryonic lethality." In: *Cell* 69.6, pp. 915–26.
- LI, G., YAO, W., and JIANG, H. (2014). "Short-chain fatty acids enhance adipocyte differentiation in the stromal vascular fraction of porcine adipose tissue." In: *The Journal of nutrition* 144.12, pp. 1887–95.

- LI, H., HANDSAKER, B., WYSOKER, A., FENNEL, T., RUAN, J., HOMER, N., MARTH, G., ABECASIS, G., and DURBIN, R. (2009). “The Sequence Alignment/Map format and SAMtools.” In: *Bioinformatics (Oxford, England)* 25.16, pp. 2078–9.
- LI, L.-C. and DAHIYA, R. (2002). “MethPrimer: designing primers for methylation PCRs.” In: *Bioinformatics* 18.11, pp. 1427–1431.
- LI, N., YE, M., LI, Y., YAN, Z., BUTCHER, L. M., SUN, J., HAN, X., CHEN, Q., ZHANG, X., and WANG, J. (2010). “Whole genome DNA methylation analysis based on high throughput sequencing technology.” In: *Methods (San Diego, Calif.)* 52.3, pp. 203–12.
- LI, R. W. and LI, C. (2006a). “Butyrate induces profound changes in gene expression related to multiple signal pathways in bovine kidney epithelial cells.” En. In: *BMC genomics* 7.1, p. 234.
- (2006b). “Butyrate induces profound changes in gene expression related to multiple signal pathways in bovine kidney epithelial cells.” In: *BMC Genomics* 7, p. 234.
- LIM, H. N. and OUDENAARDEN, A. van (2007). “A multistep epigenetic switch enables the stable inheritance of DNA methylation states.” In: *Nature Genetics* 39.2, pp. 269–275.
- LIN, I.-H. et al. (2015). “Hierarchical clustering of breast cancer methylomes revealed differentially methylated and expressed breast cancer genes.” In: *PloS one* 10.2, e0118453.
- LISTER, R. et al. (2009). “Human DNA methylomes at base resolution show widespread epigenomic differences.” In: *Nature* 462.7271, pp. 315–22.
- LIU, G., LIU, Y.-J., LIAN, W.-J., ZHAO, Z.-W., YI, T., and ZHOU, H.-Y. (2014). “Reduced BMP6 expression by DNA methylation contributes to EMT and drug resistance in breast cancer cells.” In: *Oncology reports* 32.2, pp. 581–8.
- LOGAN, C. Y. and NUSSE, R. (2004). “The Wnt signaling pathway in development and disease.” In: *Annual Review of Cell and Developmental Biology* 20, pp. 781–810.
- LONG, H. K., BLACKLEDGE, N. P., and KLOSE, R. J. (2013). “ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection.” In: *Biochemical Society transactions* 41.3, pp. 727–40.
- LOU, S. et al. (2014). “Whole-genome bisulfite sequencing of multiple individuals reveals complementary roles of promoter and gene body methylation in transcriptional regulation.” In: *Genome biology* 15.7, p. 408.
- LU, X., SONG, C.-X., SZULWACH, K., WANG, Z., WEIDENBACHER, P., JIN, P., and HE, C. (2013). “Chemical modification-assisted bisulfite sequencing (CAB-Seq) for 5-carboxylecytosine detection in DNA.” In: *Journal of the American Chemical Society* 135.25, pp. 9315–7.

-
- LUCAS, B. K., GIERE, L. M., DEMARCO, R. A., SHEN, A, CHISHOLM, V, and CROWLEY, C. W. (1996). "High-level production of recombinant proteins in CHO cells using a dicistronic DHFR intron expression vector." In: *Nucleic acids research* 24.9, pp. 1774–9.
- LÜHRS, H, GERKE, T, MÜLLER, J. G., MELCHER, R, SCHAUBER, J, BOXBERGE, F, SCHEPPACH, W, and MENZEL, T (2002). "Butyrate inhibits NF-kappaB activation in lamina propria macrophages of patients with ulcerative colitis." In: *Scandinavian journal of gastroenterology* 37.4, pp. 458–66.
- LUNT, S. Y. and VANDER HEIDEN, M. G. (2011). "Aerobic glycolysis: meeting the metabolic requirements of cell proliferation." In: *Annual review of cell and developmental biology* 27, pp. 441–64.
- LUO, W. and BROUWER, C. (2013). "Pathview: an R/Bioconductor package for pathway-based data integration and visualization." In: *Bioinformatics (Oxford, England)* 29.14, pp. 1830–1.
- LUO, W., FRIEDMAN, M. S., SHEDDEN, K., HANKENSON, K. D., and WOOLF, P. J. (2009). "GAGE: generally applicable gene set enrichment for pathway analysis." En. In: *BMC bioinformatics* 10.1, p. 161.
- LYKO, F., FORET, S., KUCHARSKI, R., WOLF, S., FALCKENHAYN, C., and MALESZKA, R. (2010). "The honey bee epigenomes: differential methylation of brain DNA in queens and workers." In: *PLoS biology* 8.11, e1000506.
- MACDONALD, B. T., TAMAI, K., and HE, X. (2009). "Wnt/ β -catenin signaling: components, mechanisms, and diseases." In: *Developmental Cell* 17.1, pp. 9–26.
- MAHAJAN, K. N. and MITCHELL, B. S. (2003). "Role of human Pso4 in mammalian DNA repair and association with terminal deoxynucleotidyl transferase." In: *Proceedings of the National Academy of Sciences of the United States of America* 100.19, pp. 10746–51.
- MAITI, A. and DROHAT, A. C. (2011). "Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites." In: *The Journal of biological chemistry* 286.41, pp. 35334–8.
- MARCHO, C., CUI, W., and MAGER, J. (2015). "Epigenetic dynamics during preimplantation development." In: *Reproduction (Cambridge, England)* 150.3, R109–20.
- MARCY, Y., OUVERNEY, C., BIK, E. M., LÖSEKANN, T., IVANOVA, N., MARTIN, H. G., SZETO, E., PLATT, D., HUGENHOLTZ, P., RELMAN, D. A., and QUAKE, S. R. (2007). "Dissecting biological "dark matter" with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth." In: *Proceedings of the National Academy of Sciences of the United States of America* 104.29, pp. 11889–94.

- MARIANI, M. R., CARPANETO, E. M., ULIVI, M., ALLFREY, V. G., and BOFFA, L. C. (2003). “Correlation between butyrate-induced histone hyperacetylation turn-over and c-myc expression”. In: *The Journal of Steroid Biochemistry and Molecular Biology* 86.2, pp. 167–171.
- MARMORSTEIN, R (2001). “Protein modules that manipulate histone tails for chromatin regulation.” In: *Nature reviews. Molecular cell biology* 2.6, pp. 422–32.
- MARTINS, R., LITHGOW, G. J., and LINK, W. (2016). “Long live FOXO: unraveling the role of FOXO proteins in aging and longevity.” In: *Aging cell* 15.2, pp. 196–207.
- MASSAGUÉ, J. (2004). “G1 cell-cycle control and cancer.” In: *Nature* 432.7015, pp. 298–306.
- MAUNAKEA, A. K., CHEPELEV, I., CUI, K., and ZHAO, K. (2013). “Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition.” In: *Cell research* 23.11, pp. 1256–69.
- MCILWAIN, D. R., BERGER, T., and MAK, T. W. (2013). “Caspase functions in cell death and disease.” In: *Cold Spring Harbor perspectives in biology* 5.4, a008656.
- MCKITTRICK, E., GAFKEN, P. R., AHMAD, K., and HENIKOFF, S. (2004). “Histone H3.3 is enriched in covalent modifications associated with active chromatin.” In: *Proceedings of the National Academy of Sciences of the United States of America* 101.6, pp. 1525–30.
- MEISSNER, A., MIKKELSEN, T. S., GU, H., WERNIG, M., HANNA, J., SIVACHENKO, A., ZHANG, X., BERNSTEIN, B. E., NUSBAUM, C., JAFFE, D. B., GNIRKE, A., JAENISCH, R., and LANDER, E. S. (2008). “Genome-scale DNA methylation maps of pluripotent and differentiated cells.” In: *Nature* 454.7205, pp. 766–70.
- MELEADY, P. (2007). “Proteomic profiling of recombinant cells from large-scale mammalian cell culture processes.” In: *Cytotechnology* 53.1-3, pp. 23–31.
- MELEADY, P., DOOLAN, P., HENRY, M., BARRON, N., KEENAN, J., O’SULLIVAN, F., CLARKE, C., GAMMELL, P., MELVILLE, M. W., LEONARD, M., and CLYNES, M. (2011). “Sustained productivity in recombinant Chinese hamster ovary (CHO) cell lines: proteome analysis of the molecular basis for a process-related phenotype.” In: *BMC biotechnology* 11, p. 78.
- MELNIKOV, A. A., GARTENHAUS, R. B., LEVENSON, A. S., MOTCHOULSKAIA, N. A., and LEVENSON CHERNOKHVOSTOV, V. V. (2005). “MSRE-PCR for analysis of gene-specific DNA methylation.” In: *Nucleic acids research* 33.10, e93.
- MESSERSCHMIDT, D. M., KNOWLES, B. B., and SOLTER, D. (2014). “DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos.” In: *Genes & development* 28.8, pp. 812–28.

-
- MILLER, W. M., WILKE, C. R., and BLANCH, H. W. (1987). “Effects of dissolved oxygen concentration on hybridoma growth and metabolism in continuous culture.” In: *Journal of cellular physiology* 132.3, pp. 524–30.
- MISAGHI, S., QU, Y., SNOWDEN, A., CHANG, J., and SNEDECOR, B. (2013). “Resilient immortals, characterizing and utilizing Bax/Bak deficient Chinese hamster ovary (CHO) cells for high titer antibody production.” In: *Biotechnology progress* 29.3, pp. 727–37.
- MONFORT, A. and WUTZ, A. (2013). “Breathing-in epigenetic change with vitamin C.” In: *EMBO reports* 14.4, pp. 337–46.
- MOREMEN, K. W., TIEMEYER, M., and NAIRN, A. V. (2012). “Vertebrate protein glycosylation: diversity, synthesis and function.” In: *Nature reviews. Molecular cell biology* 13.7, pp. 448–62.
- MORITZ, B., WOLTERING, L., BECKER, P. B., and GÖPFERT, U. (2016). “High levels of histone H3 acetylation at the CMV promoter are predictive of stable expression in Chinese hamster ovary cells.” In: *Biotechnology progress* 32.3, pp. 776–86.
- MORK, L., CZERWINSKI, M., and CAPEL, B. (2014). “Predetermination of sexual fate in a turtle with temperature-dependent sex determination.” In: *Developmental biology* 386.1, pp. 264–71.
- MORRIS, T. J. and BECK, S. (2015). “Analysis pipelines and packages for Infinium HumanMethylation450 BeadChip (450k) data”. In: *Methods* 72, pp. 3–8.
- MÜLLER, B., HEINRICH, C., JABS, W., KASPAR, S., SCHMIDT, A., RODRIGUES DE CARVALHO, N., ALBAUM, S. P., BAESSMANN, C., NOLL, T., and HOFFFROGGE, R. (2016). “Label-free protein quantification of sodium butyrate treated CHO cells by ESI-UHR-TOF-MS”. In: *Submitted to Journal of Biotechnology*.
- NAIR, S. S., COOLEN, M. W., STIRZAKER, C., SONG, J. Z., STATHAM, A. L., STRBENAC, D., ROBINSON, M. D., and CLARK, S. J. (2011). “Comparison of methyl-DNA immunoprecipitation (MeDIP) and methyl-CpG binding domain (MBD) protein capture for genome-wide DNA methylation analysis reveal CpG sequence coverage bias.” In: *Epigenetics : official journal of the DNA Methylation Society* 6.1, pp. 34–44.
- NALLAMSHETTY, S., CHAN, S. Y., and LOSCALZO, J. (2013). “Hypoxia: a master regulator of microRNA biogenesis and activity.” In: *Free radical biology & medicine* 64, pp. 20–30.
- NDLOVU, M. N., DENIS, H., and FUKS, F. (2011). “Exposing the DNA methylome iceberg.” In: *Trends in Biochemical Sciences* 36.7, pp. 381–387.
- NERI, F., INCARNATO, D., KREPELOVA, A., RAPELLI, S., ANSELMINI, F., PARLATO, C., MEDANA, C., DAL BELLO, F., and OLIVIERO, S. (2015). “Single-Base Resolution Analysis of 5-Formyl and 5-Carboxyl Cytosine Reveals Promoter DNA Methylation Dynamics.” In: *Cell reports* 10.5, pp. 674–683.

- NILSSON, E. et al. (2014). “Altered DNA methylation and differential expression of genes influencing metabolism and inflammation in adipose tissue from subjects with type 2 diabetes.” In: *Diabetes* 63.9, pp. 2962–76.
- NISSOM, P. M., SANNY, A., KOK, Y. J., HIANG, Y. T., CHUAH, S. H., SHING, T. K., LEE, Y. Y., WONG, K. T. K., HU, W.-S., SIM, M. Y. G., and PHILP, R. (2006). “Transcriptome and proteome profiling to understanding the biology of high productivity CHO cells.” In: *Molecular biotechnology* 34.2, pp. 125–40.
- NOONAN, E. J., PLACE, R. F., POOKOT, D, BASAK, S, WHITSON, J. M., HIRATA, H, GIARDINA, C, and DAHIYA, R (2009). “miR-449a targets HDAC-1 and induces growth arrest in prostate cancer.” In: *Oncogene* 28.14, pp. 1714–24.
- NÖR, C., SASSI, F. A., FARIAS, C. B. de, SCHWARTSMANN, G., ABUJAMRA, A. L., LENZ, G., BRUNETTO, A. L., and ROESLER, R. (2013). “The histone deacetylase inhibitor sodium butyrate promotes cell death and differentiation and reduces neurosphere formation in human medulloblastoma cells.” In: *Molecular neurobiology* 48.3, pp. 533–43.
- NOSEDA, M. and KARSAN, A. (2006). “Notch and minichromosome maintenance (MCM) proteins: integration of two ancestral pathways in cell cycle control.” en. In: *Cell cycle (Georgetown, Tex.)* 5.23, pp. 2704–9.
- OH, H. K., SO, M. K., YANG, J., YOON, H. C., AHN, J. S., LEE, J. M., KIM, J. T., YOO, J. U., and BYUN, T. H. (2005). “Effect of N-acetylcystein on butyrate-treated Chinese hamster ovary cells to improve the production of recombinant human interferon- β -1a.” In: *Biotechnology Progress* 21.4, pp. 1154–1164.
- OKABAYASHI, K, KANÉDA, T, and ARIMURA, H (1989). “Effect of butyrate on the expression of the human preprourokinase gene introduced into Chinese hamster ovary cells.” In: *Cell structure and function* 14.5, pp. 579–86.
- OOI, S. K. T., QIU, C., BERNSTEIN, E., LI, K., JIA, D., YANG, Z., ERDJUMENT-BROMAGE, H., TEMPST, P., LIN, S.-P., ALLIS, C. D., CHENG, X., and BESTOR, T. H. (2007). “DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA.” In: *Nature* 448.7154, pp. 714–7.
- OSTERLEHNER, A., SIMMETH, S., and GÖPFERT, U. (2011). “Promoter methylation and transgene copy numbers predict unstable protein production in recombinant Chinese hamster ovary cell lines.” In: *Biotechnology and Bioengineering* 108.11, pp. 2670–2681.
- PALERMO, D., DEGRAAF, M., MAROTTI, K., REHBERG, E., and POST, L. (1991). “Production of analytical quantities of recombinant proteins in Chinese hamster ovary cells using sodium butyrate to elevate gene expression”. In: *Journal of Biotechnology* 19.1, pp. 35–47.

-
- PARK, S. Y., KWON, H. J., LEE, H. E., RYU, H. S., KIM, S.-W., KIM, J. H., KIM, I. A., JUNG, N., CHO, N.-Y., and KANG, G. H. (2011). “Promoter CpG island hypermethylation during breast cancer progression.” In: *Virchows Archiv : an international journal of pathology* 458.1, pp. 73–84.
- PARROTT, B. B., KOHNO, S., CLOY-MCCOY, J. A., and GUILLETTE, L. J. (2014). “Differential incubation temperatures result in dimorphic DNA methylation patterning of the SOX9 and aromatase promoters in gonads of alligator (*Alligator mississippiensis*) embryos.” In: *Biology of reproduction* 90.1, p. 2.
- PERKINS, N. D. (2000). “The Rel/NF- κ B family: friend and foe”. In: *Trends in Biochemical Sciences* 25.9, pp. 434–440.
- PROKOPIUK, L., WESTERN, P. S., and STRINGER, J. M. (2015). “Transgenerational epigenetic inheritance: adaptation through the germline epigenome?” In: *Epigenomics* 7.5, pp. 829–46.
- PUCK, T. (1958). “The genetics of somatic mammalian cells.” In: *Advances in Biological and Medical Physics* 5, pp. 75–101.
- RADHAKRISHNAN, P., BASMA, H., KLINKEBIEL, D., CHRISTMAN, J., and CHENG, P.-W. (2008). “Cell type-specific activation of the cytomegalovirus promoter by dimethylsulfoxide and 5-Aza-2'-deoxycytidine”. In: *The International Journal of Biochemistry & Cell Biology* 40.9, pp. 1944–1955.
- RAI, K., HUGGINS, I. J., JAMES, S. R., KARPF, A. R., JONES, D. A., and CAIRNS, B. R. (2008). “DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45.” In: *Cell* 135.7, pp. 1201–1212.
- RAIMONDI, A., FERGUSON, S. M., LOU, X., ARMBRUSTER, M., PARADISE, S., GIOVEDI, S., MESSA, M., KONO, N., TAKASAKI, J., CAPPELLO, V., O'TOOLE, E., RYAN, T. A., and DE CAMILLI, P. (2011). “Overlapping role of dynamin isoforms in synaptic vesicle endocytosis.” In: *Neuron* 70.6, pp. 1100–14.
- RAKYAN, V. K., DOWN, T. A., BALDING, D. J., and BECK, S. (2011). “Epigenome-wide association studies for common human diseases.” In: *Nature reviews. Genetics* 12.8, pp. 529–41.
- RAUCH, T., LI, H., WU, X., and PFEIFER, G. P. (2006a). “MIRA-assisted microarray analysis, a new technology for the determination of DNA methylation patterns, identifies frequent methylation of homeodomain-containing genes in lung cancer cells.” In: *Cancer research* 66.16, pp. 7939–47.
- (2006b). “MIRA-assisted microarray analysis, a new technology for the determination of DNA methylation patterns, identifies frequent methylation of homeodomain-containing genes in lung cancer cells.” In: *Cancer research* 66.16, pp. 7939–47.

- RAUCH, T. A., WANG, Z., WU, X., KERNSTINE, K. H., RIGGS, A. D., and PFEIFER, G. P. (2012). “DNA methylation biomarkers for lung cancer.” In: *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* 33.2, pp. 287–96.
- REIK, W. (2007). “Stability and flexibility of epigenetic gene regulation in mammalian development.” In: *Nature* 447.7143, pp. 425–32.
- RIGGS, A. (1975). “X inactivation, differentiation, and DNA methylation”. en. In: *Cytogenetic and Genome Research* 14.1, pp. 9–25.
- RITCHIE, M. E., PHIPSON, B., WU, D., HU, Y., LAW, C. W., SHI, W., and SMYTH, G. K. (2015). “limma powers differential expression analyses for RNA-sequencing and microarray studies”. In: *Nucleic Acids Research* 43.7, e47.
- ROMANOSKI, C. E., GLASS, C. K., STUNNENBERG, H. G., WILSON, L., and ALMOUZNI, G. (2015). “Epigenomics: Roadmap for regulation.” In: *Nature* 518.7539, pp. 314–6.
- ROUSSEAU, S. et al. (2013). “Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers.” In: *Science translational medicine* 5.186, 186ra66.
- RUIZ-HERNANDEZ, A., KUO, C.-C., RENTERO-GARRIDO, P., TANG, W.-Y., REDON, J., ORDOVAS, J. M., NAVAS-ACIEN, A., and TELLEZ-PLAZA, M. (2015). “Environmental chemicals and DNA methylation in adults: a systematic review of the epidemiologic evidence.” In: *Clinical epigenetics* 7.1, p. 55.
- RUPP, O., BECKER, J., BRINKROLF, K., TIMMERMANN, C., BORTH, N., PÜHLER, A., NOLL, T., and GOESMANN, A. (2014). “Construction of a public CHO cell line transcript database using versatile bioinformatics analysis pipelines.” In: *PloS one* 9.1, e85568.
- SADLER, T., BHASIN, J. M., XU, Y., BARNHOLZ-SLOAN, J., CHEN, Y., TING, A. H., and STYLIANOU, E. (2016). “Genome-wide analysis of DNA methylation and gene expression defines molecular characteristics of Crohn’s disease-associated fibrosis.” In: *Clinical epigenetics* 8, p. 30.
- SALVADOR, J. M., BROWN-CLAY, J. D., and FORNACE, A. J. (2013). “Gadd45 in stress signaling, cell cycle control, and apoptosis.” In: *Advances in experimental medicine and biology* 793, pp. 1–19.
- SAMBROOK, J. and RUSSELL, D. W. (2006). “Purification of nucleic acids by extraction with phenol:chloroform.” In: *CSH Protoc* 2006.1, pdb.prot4455.
- SANDELIN, A., CARNINCI, P., LENHARD, B., PONJAVIC, J., HAYASHIZAKI, Y., and HUME, D. A. (2007). “Mammalian RNA polymerase II core promoters: insights from genome-wide studies.” en. In: *Nature reviews. Genetics* 8.6, pp. 424–36.
- SARKAR, S., ABUJAMRA, A. L., LOEW, J. E., FORMAN, L. W., PERRINE, S. P., and FALLER, D. V. (2011). “Histone deacetylase inhibitors reverse CpG methylation by regulating DNMT1 through ERK signaling.” In: *Anticancer research* 31.9, pp. 2723–32.

-
- SATO, F., TSUCHIYA, S., MELTZER, S. J., and SHIMIZU, K. (2011). “MicroRNAs and epigenetics.” In: *FEBS Journal* 278.10, pp. 1598–1609.
- SAUNDERS, F., SWEENEY, B., ANTONIOU, M. N., STEPHENS, P., and CAIN, K. (2015). “Chromatin function modifying elements in an industrial antibody production platform—comparison of UCOE, MAR, STAR and cHS4 elements.” In: *PloS one* 10.4, e0120096.
- SAUVE, A. A. and YOUN, D. Y. (2012). “Sirtuins: NAD(+)-dependent deacetylase mechanism and regulation.” In: *Current opinion in chemical biology* 16.5-6, pp. 535–43.
- SCHERMELLEH, L., HAEMMER, A., SPADA, F., RÖSING, N., MEILINGER, D., ROTHBAUER, U., CARDOSO, M. C., and LEONHARDT, H. (2007). “Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation.” In: *Nucleic acids research* 35.13, pp. 4301–12.
- SCHIESSER, S., HACKNER, B., PFAFFENEDER, T., MÜLLER, M., HAGEMEI, C., TRUSS, M., and CARELL, T. (2012). “Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing.” In: *Angewandte Chemie (International ed. in English)* 51.26, pp. 6516–20.
- SCHROEDER, D. I. and LASALLE, J. M. (2013). “How has the study of the human placenta aided our understanding of partially methylated genes?” In: *Epigenomics* 5.6, pp. 645–54.
- SCHROEDER, D. I., LOTT, P., KORF, I., and LASALLE, J. M. (2011). “Large-scale methylation domains mark a functional subset of neuronally expressed genes.” In: *Genome research* 21.10, pp. 1583–91.
- SCHROEDER, D. I., BLAIR, J. D., LOTT, P., YU, H. O. K., HONG, D., CRARY, F., ASHWOOD, P., WALKER, C., KORF, I., ROBINSON, W. P., and LASALLE, J. M. (2013). “The human placenta methylome.” In: *Proceedings of the National Academy of Sciences of the United States of America* 110.15, pp. 6037–42.
- SCHÜBELER, D. (2015). “Function and information content of DNA methylation”. In: *Nature* 517.7534, pp. 321–326.
- SCOTT, D. W., LONGPRE, J. M., and LOO, G. (2008). “Upregulation of GADD153 by butyrate: involvement of MAPK.” In: *DNA and cell biology* 27.11, pp. 607–14.
- SETH, G., PHILP, R. J., LAU, A., JIUN, K. Y., YAP, M., and HU, W.-S. (2007). “Molecular portrait of high productivity in recombinant NS0 cells.” In: *Biotechnology and bioengineering* 97.4, pp. 933–51.
- SHALEK, A. K. et al. (2013). “Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells.” In: *Nature* 498.7453, pp. 236–40.

- SHALINI, S., DORSTYN, L., DAWAR, S., and KUMAR, S (2015). “Old, new and emerging functions of caspases.” In: *Cell death and differentiation* 22.4, pp. 526–39.
- SHI, Y., LAN, F., MATSON, C., MULLIGAN, P., WHETSTINE, J. R., COLE, P. A., CASERO, R. A., and SHI, Y. (2004). “Histone demethylation mediated by the nuclear amine oxidase homolog LSD1.” In: *Cell* 119.7, pp. 941–53.
- SHIN, H., KIM, J.-H., LEE, Y. S., and LEE, Y. C. (2012). “Change in gene expression profiles of secreted frizzled-related proteins (SFRPs) by sodium butyrate in gastric cancers: induction of promoter demethylation and histone modification causing inhibition of Wnt signaling.” In: *International Journal of Oncology* 40.5, pp. 1533–1542.
- SHUKUROV, R. R., KAZACHENKO, K. Y., KOZLOV, D. G., NURBAKOV, A. A., SAUTKINA, E. N., KHAMITOV, R. A., and SERYOGIN, Y. A. (2014). “Optimization of genetic constructs for high-level expression of the darbepoetin gene in mammalian cells”. In: *Applied Biochemistry and Microbiology* 50.9, pp. 802–811.
- SINGLA, V., ROMAGUERA-ROS, M., GARCIA-VERDUGO, J. M., and REITER, J. F. (2010). “Ofd1, a human disease gene, regulates the length and distal structure of centrioles.” In: *Developmental cell* 18.3, pp. 410–24.
- SIONOV, R. V., VLAHOPOULOS, S. A., and GRANOT, Z. (2015). “Regulation of Bim in Health and Disease.” In: *Oncotarget* 6.27, pp. 23058–134.
- SMALES, C. M., DINNIS, D. M., STANSFIELD, S. H., ALETE, D, SAGE, E. A., BIRCH, J. R., RACHER, A. J., MARSHALL, C. T., and JAMES, D. C. (2004). “Comparative proteomic analysis of GS-NS0 murine myeloma cell lines with varying recombinant monoclonal antibody production rate.” In: *Biotechnology and bioengineering* 88.4, pp. 474–88.
- SMITH, A. R., SMITH, R. G., CONDLIFFE, D., HANNON, E., SCHALKWYK, L., MILL, J., and LUNNON, K. (2016). “Increased DNA methylation near TREM2 is consistently seen in the superior temporal gyrus in Alzheimer’s disease brain.” In: *Neurobiology of aging* 47, pp. 35–40.
- SMYTH, G. K. (2005). *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Ed. by R. GENTLEMAN, V. J. CAREY, W. HUBER, R. A. IRIZARRY, and S. DUDOIT. Statistics for Biology and Health. New York: Springer-Verlag.
- SONG, J., RECHKOBLIT, O., BESTOR, T. H., and PATEL, D. J. (2011). “Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation.” In: *Science (New York, N.Y.)* 331.6020, pp. 1036–40.
- SONG, Q., DECATO, B., HONG, E. E., ZHOU, M., FANG, F., QU, J., GARVIN, T., KESSLER, M., ZHOU, J., and SMITH, A. D. (2013). “A reference methylome database and analysis pipeline to facilitate integrative and comparative epigenomics.” In: *PLoS one* 8.12. Ed. by O. EL-MAARRI, e81148.

-
- SPROUL, D. and MEEHAN, R. R. (2013). “Genomic insights into cancer-associated aberrant CpG island hypermethylation.” In: *Briefings in functional genomics* 12.3, pp. 174–90.
- SROUGI, M. C. and BURRIDGE, K. (2011). “The nuclear guanine nucleotide exchange factors Ect2 and Net1 regulate RhoB-mediated cell death after DNA damage.” In: *PloS one* 6.2, e17108.
- STADLER, M. B., MURR, R., BURGER, L., IVANEK, R., LIENERT, F., SCHÖLER, A., NIMWEGEN, E. van, WIRBELAUER, C., OAKELEY, E. J., GAIDATZIS, D., TIWARI, V. K., and SCHÜBELER, D. (2011). “DNA-binding factors shape the mouse methylome at distal regulatory regions.” In: *Nature* 480.7378, pp. 490–5.
- STADTFELD, M., APOSTOLOU, E., FERRARI, F., CHOI, J., WALSH, R. M., CHEN, T., OOI, S. S. K., KIM, S. Y., BESTOR, T. H., SHIODA, T., PARK, P. J., and HOCHEDLINGER, K. (2012). “Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all-iPS cell mice from terminally differentiated B cells.” In: *Nature genetics* 44.4, 398–405, S1–2.
- STEIN, R., SCIAKY-GALLILI, N., RAZIN, A., and CEDAR, H (1983). “Pattern of methylation of two genes coding for housekeeping functions.” In: *Proceedings of the National Academy of Sciences of the United States of America* 80.9, pp. 2422–6.
- STELIOU, K., BOOSALIS, M. S., PERRINE, S. P., SANGERMAN, J., and FALLER, D. V. (2012). “Butyrate Histone Deacetylase Inhibitors”. en. In: *Biores Open Access* 1, pp. 192–8.
- STĘPIŃSKI, D. (2012). “Levels of DNA methylation and histone methylation and acetylation change in root tip cells of soybean seedlings grown at different temperatures.” In: *Plant physiology and biochemistry : PPB / Société française de physiologie végétale* 61, pp. 9–17.
- STORRIE, B, PUCK, T. T., and WENGER, L (1978). “The role of butyrate in the reverse transformation reaction in mammalian cells.” In: *Journal of cellular physiology* 94.1, pp. 69–75.
- STRICHER, F., MACRI, C., RUFF, M., and MULLER, S. (2013). “HSPA8/HSC70 chaperone protein”. en. In: *Autophagy*.
- STRZALKA, W. and ZIEMIENOWICZ, A. (2011). “Proliferating cell nuclear antigen (PCNA): a key factor in DNA replication and cell cycle regulation.” In: *Annals of botany* 107.7, pp. 1127–40.
- SUN, S., LI, W., ZHANG, H., ZHA, L., XUE, Y., WU, X., and ZOU, F. (2012). “Requirement for store-operated calcium entry in sodium butyrate-induced apoptosis in human colon cancer cells.” In: *Bioscience reports* 32.1, pp. 83–90.
- SUN, X., HE, Y., HUANG, C., MA, T.-T., and LI, J. (2013). “The epigenetic feedback loop between DNA methylation and microRNAs in fibrotic disease with an emphasis on DNA methyltransferases.” In: *Cellular signalling* 25.9, pp. 1870–6.

- SUNG, Y. H. and LEE, G. M. (2005). “Enhanced human thrombopoietin production by sodium butyrate addition to serum-free suspension culture of bcl-2-overexpressing CHO cells.” In: *Biotechnology Progress* 21.1, pp. 50–57.
- SUNG, Y. H., SONG, Y. J., LIM, S. W., CHUNG, J. Y., and LEE, G. M. (2004). “Effect of sodium butyrate on the production, heterogeneity and biological activity of human thrombopoietin by recombinant Chinese hamster ovary cells.” In: *Journal of biotechnology* 112.3, pp. 323–35.
- SZKLARCZYK, D., FRANCESCHINI, A., WYDER, S., FORSLUND, K., HELLER, D., HUERTA-CEPAS, J., SIMONOVIC, M., ROTH, A., SANTOS, A., TSAFOU, K. P., KUHN, M., BORK, P., JENSEN, L. J., and MERING, C. von (2015). “STRING v10: protein-protein interaction networks, integrated over the tree of life.” In: *Nucleic acids research* 43.Database issue, pp. D447–52.
- TAHILIANI, M., KOH, K. P., SHEN, Y., PASTOR, W. A., BANDUKWALA, H., BRUDNO, Y., AGARWAL, S., IYER, L. M., LIU, D. R., ARAVIND, L., and RAO, A. (2009). “Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1.” In: *Science (New York, N.Y.)* 324.5929, pp. 930–5.
- TAKAI, D. and JONES, P. A. (2002). “Comprehensive analysis of CpG islands in human chromosomes 21 and 22.” In: *Proceedings of the National Academy of Sciences* 99.6, pp. 3740–3745.
- TANG, W. W. C., DIETMANN, S., IRIE, N., LEITCH, H. G., FLOROS, V. I., BRADSHAW, C. R., HACKETT, J. A., CHINNERY, P. F., and SURANI, M. A. (2015). “A Unique Gene Regulatory Network Resets the Human Germline Epigenome for Development.” In: *Cell* 161.6, pp. 1453–67.
- TASTANOVA, A., SCHULZ, A., FOLCHER, M., TOLSTRUP, A., PUKLOWSKI, A., KAUFMANN, H., and FUSSENEGGER, M. (2016). “Overexpression of YY1 increases the protein production in mammalian cells.” In: *Journal of biotechnology* 219, pp. 72–85.
- THE ENCODE PROJECT CONSORTIUM, A. A. (2004). “The ENCODE (ENCyclopedia Of DNA Elements) Project.” en. In: *Science (New York, N.Y.)* 306.5696, pp. 636–40.
- THE UNIPROT CONSORTIUM (2014). “UniProt: a hub for protein information”. In: *Nucleic Acids Research* 43.D1, pp. D204–212.
- THOMAS, F. and KUTAY, U. (2003). “Biogenesis and nuclear export of ribosomal subunits in higher eukaryotes depend on the CRM1 export pathway.” In: *Journal of cell science* 116.Pt 12, pp. 2409–19.
- THOMAS, J. O. and KORNBERG, R. D. (1975). “An octamer of histones in chromatin and free in solution.” In: *Proceedings of the National Academy of Sciences of the United States of America* 72.7, pp. 2626–30.
- THOMSON, E., FERREIRA-CERCA, S., and HURT, E. (2013). “Eukaryotic ribosome biogenesis at a glance.” In: *Journal of cell science* 126.Pt 21, pp. 4815–21.

-
- TIAN, X and DIAZ, F. J. (2013). “Acute dietary zinc deficiency before conception compromises oocyte epigenetic programming and disrupts embryonic development.” In: *Developmental biology* 376.1, pp. 51–61.
- TOBI, E. W. et al. (2014). “DNA methylation signatures link prenatal famine exposure to growth and metabolism.” en. In: *Nature communications* 5, p. 5592.
- TOKUDA, E., ITOH, T., HASEGAWA, J., IJUIN, T., TAKEUCHI, Y., IRINO, Y., FUKUMOTO, M., and TAKENAWA, T. (2014). “Phosphatidylinositol 4-phosphate in the Golgi apparatus regulates cell-cell adhesion and invasive cell migration in human breast cancer.” In: *Cancer research* 74.11, pp. 3054–66.
- TSUKADA, Y.-I., FANG, J., ERDJUMENT-BROMAGE, H., WARREN, M. E., BORCHERS, C. H., TEMPST, P., and ZHANG, Y. (2006). “Histone demethylation by a family of JmjC domain-containing proteins.” In: *Nature* 439.7078, pp. 811–6.
- TURCHI, L, ABERDAM, E, MAZURE, N, POUYSSÉGUR, J, DECKERT, M, KITAJIMA, S, ABERDAM, D, and VIROLLE, T (2008). “Hif-2alpha mediates UV-induced apoptosis through a novel ATF3-dependent death pathway.” In: *Cell death and differentiation* 15.9, pp. 1472–80.
- VANDER HEIDEN, M. G., CANTLEY, L. C., and THOMPSON, C. B. (2009). “Understanding the Warburg effect: the metabolic requirements of cell proliferation.” In: *Science (New York, N.Y.)* 324.5930, pp. 1029–33.
- VARRIALE, A. and BERNARDI, G. (2006). “DNA methylation and body temperature in fishes.” In: *Gene* 385, pp. 111–21.
- VEITH, N., ZIEHR, H., MACLEOD, R. A. F., and REAMON-BUETTNER, S. M. (2016). “Mechanisms underlying epigenetic and transcriptional heterogeneity in Chinese hamster ovary (CHO) cell lines.” In: *BMC biotechnology* 16, p. 6.
- VERGARA, M., BECERRA, S., BERRIOS, J., OSSES, N., REYES, J., RODRÍGUEZ-MOYÁ, M., GONZALEZ, R., and ALTAMIRANO, C. (2014). “Differential Effect of Culture Temperature and Specific Growth Rate on CHO Cell Behavior in Chemostat Culture.” In: *PloS one* 9.4, e93865.
- VIZCAÍNO, C., MANSILLA, S., and PORTUGAL, J. (2015). “Sp1 transcription factor: A long-standing target in cancer chemotherapy.” In: *Pharmacology & therapeutics* 152, pp. 111–24.
- VOJTA, A., DOBRINIĆ, P., TADIĆ, V., BOČKOR, L., KORAC, P., JULG, B., KLASIĆ, M., and ZOLDOŠ, V. (2016). “Repurposing the CRISPR-Cas9 system for targeted DNA methylation.” In: *Nucleic acids research* 44.12, pp. 5615–28.
- VRBA, L., GARBE, J. C., STAMPFER, M. R., and FUTSCHER, B. W. (2011). “Epigenetic regulation of normal human mammary cell type-specific miRNAs.” In: *Genome research* 21.12, pp. 2026–37.

- VRBA, L., MUÑOZ-RODRÍGUEZ, J. L., STAMPFER, M. R., and FUTSCHER, B. W. (2013). “miRNA gene promoters are frequent targets of aberrant DNA methylation in human breast cancer.” In: *PloS one* 8.1, e54398.
- WADDINGTON, C. H. (1942). “The epigenotype.” In: *International journal of epidemiology* 41.1, pp. 10–3.
- WALKER, D. L., BHAGWATE, A. V., BAHETI, S., SMALLEY, R. L., HILKER, C. A., SUN, Z., and CUNNINGHAM, J. M. (2015). “DNA methylation profiling: comparison of genome-wide sequencing methods and the Infinium Human Methylation 450 Bead Chip.” EN. In: *Epigenomics* 7.8, pp. 1287–302.
- WALSER, J.-C., PONGER, L., and FURANO, A. V. (2008). “CpG dinucleotides and the mutation rate of non-CpG DNA.” In: *Genome research* 18.9, pp. 1403–14.
- WALSH, G (2014). “Biopharmaceutical benchmarks 2014.” In: *Nature biotechnology* 32.10, pp. 992–1000.
- WALSH, G. (2010). “Post-translational modifications of protein biopharmaceuticals.” In: *Drug Discovery Today* 15.17-18, pp. 773–780.
- WAN, J., OLIVER, V. F., WANG, G., ZHU, H., ZACK, D. J., MERBS, S. L., and QIAN, J. (2015). “Characterization of tissue-specific differential DNA methylation suggests distinct modes of positive and negative gene expression regulation”. En. In: *BMC Genomics* 16.1, p. 49.
- WARD, J. H. (2012). “Hierarchical Grouping to Optimize an Objective Function”. en. In: *Journal of the American Statistical Association*, pp. 236–244.
- WARNECKE, P. M., STIRZAKER, C., SONG, J., GRUNAU, C., MELKI, J. R., and CLARK, S. J. (2002). “Identification and resolution of artifacts in bisulfite sequencing.” In: *Methods (San Diego, Calif.)* 27.2, pp. 101–7.
- WATSON, J. D. and CRICK, F. H. (1953). “Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid.” In: *Nature* 171.4356, pp. 737–8.
- WEBER, M., DAVIES, J. J., WITTIG, D., OAKELEY, E. J., HAASE, M., LAM, W. L., and SCHÜBELER, D. (2005). “Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells.” In: *Nature genetics* 37.8, pp. 853–62.
- WEI, Y., MELAS, P. A., WEGENER, G., MATHÉ, A. A., and LAVEBRATT, C. (2015). “Antidepressant-like effect of sodium butyrate is associated with an increase in TET1 and in 5-hydroxymethylation levels in the Bdnf gene.” en. In: *The international journal of neuropsychopharmacology / official scientific journal of the Collegium Internationale Neuropsychopharmacologicum (CINP)* 18.2, pyu032.

-
- WEINBERG, M. S. and MORRIS, K. V. (2016). “Transcriptional gene silencing in humans.” In: *Nucleic acids research*, gkw139–.
- WELLEN, K. E., HATZIVASSILIOU, G., SACHDEVA, U. M., BUI, T. V., CROSS, J. R., and THOMPSON, C. B. (2009). “ATP-citrate lyase links cellular metabolism to histone acetylation.” In: *Science (New York, N.Y.)* 324.5930, pp. 1076–80.
- WHITFIELD, T. W., WANG, J., COLLINS, P. J., PARTRIDGE, E. C., ALDRED, S. F., TRINKLEIN, N. D., MYERS, R. M., and WENG, Z. (2012). “Functional analysis of transcription factor binding sites in human promoters.” En. In: *Genome biology* 13.9, R50.
- WILLIAMS, R. T. and WANG, Y. (2012). “A density functional theory study on the kinetics and thermodynamics of N-glycosidic bond cleavage in 5-substituted 2'-deoxycytidines.” In: *Biochemistry* 51.32, pp. 6458–62.
- WIPPERMANN, A., RUPP, O., BRINKROLF, K., HOFFFROGGE, R., and NOLL, T. (2016). “Integrative analysis of DNA methylation and gene expression in butyrate-treated CHO cells”. In: *Journal of Biotechnology*.
- WIPPERMANN, A. (2012). “Untersuchung differentieller DNA-Methylierung in CHO-Zellen nach Butyratbehandlung”. In: *Master thesis*.
- WIPPERMANN, A., KLAUSING, S., RUPP, O., ALBAUM, S. P., BÜNTEMEYER, H., NOLL, T., and HOFFFROGGE, R. (2013). “Establishment of a CpG island microarray for analyses of genome-wide DNA methylation in Chinese hamster ovary cells.” In: *Applied microbiology and biotechnology*.
- WIPPERMANN, A., RUPP, O., BRINKROLF, K., HOFFFROGGE, R., and NOLL, T. (2015). “The DNA methylation landscape of Chinese hamster ovary (CHO) DP-12 cells.” In: *Journal of biotechnology* 199, pp. 38–46.
- WOJDACZ, T. K., DOBROVIC, A., and HANSEN, L. L. (2008). “Methylation-sensitive high-resolution melting.” In: *Nature protocols* 3.12, pp. 1903–8.
- WOSSIDLO, M., ARAND, J., SEBASTIANO, V., LEPIKHOV, K., BOIANI, M., REINHARDT, R., SCHÖLER, H., and WALTER, J. (2010). “Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes.” In: *The EMBO journal* 29.11, pp. 1877–88.
- WRIGHT, J. (1973). “Morphology and growth rate changes in Chinese hamster cells cultured in presence of sodium butyrate”. In: *Experimental Cell Research* 78.2, pp. 456–460.
- WU, C. T. and MORRIS, J. R. (2001). “Genes, genetics, and epigenetics: a correspondence.” In: *Science (New York, N.Y.)* 293.5532, pp. 1103–5.
- WU, S., LI, C., HUANG, W., LI, W., and LI, R. W. (2012). “Alternative splicing regulated by butyrate in bovine epithelial cells.” In: *PloS one* 7.6, e39182.

- WÜLLNER, U, KAUT, O, DEBONI, L, PISTON, D, and SCHMITT, I (2016). “DNA methylation in Parkinson’s disease.” In: *J Neurochem* 139 Suppl 1, pp. 108–120.
- XIONG, Z. and LAIRD, P. W. (1997). “COBRA : a sensitive and quantitative DNA methylation assay”. In: *System* 25.12, pp. 2532–2534.
- XU, X. et al. (2011). “The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line.” In: *Nat Biotechnol* 29.8, pp. 735–741.
- YAN, P. S., CHEN, C.-M., SHI, H., RAHMATPANAHI, F., WEI, S. H., and HUANG, T. H.-M. (2002). “Applications of CpG Island Microarrays for High-Throughput Analysis of DNA Methylation”. In: *Journal of Nutrition* 132.8, 2430S–2434.
- YANG, W.-T. and ZHENG, P.-S. (2014). “Promoter hypermethylation of KLF4 inactivates its tumor suppressor function in cervical carcinogenesis.” In: *PloS one* 9.2, e88827.
- YANG, Y. H., DUDOIT, S., LUU, P., LIN, D. M., PENG, V., NGAI, J., and SPEED, T. P. (2002). “Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation.” In: *Nucleic Acids Res* 30.4, e15.
- YAO, X., ZHANG, J.-R., HUANG, H.-R., DAI, L.-C., LIU, Q.-J., and ZHANG, M. (2010). “Histone deacetylase inhibitor promotes differentiation of embryonic stem cells into neural cells in adherent monoculture.” In: *Chinese Medical Journal* 123.6, pp. 734–738.
- YEE, J. C., DE LEON GATTI, M., PHILP, R. J., YAP, M., and HU, W.-S. (2008). “Genomic and proteomic exploration of CHO and hybridoma cells under sodium butyrate treatment.” In: *Biotechnology and Bioengineering* 99.5, pp. 1186–1204.
- YUE, W., CHENG, W., LIU, Z., TANG, Y., LU, T., ZHANG, D., TANG, M., and HUANG, Y. (2016). “Genome-wide DNA methylation analysis in obsessive-compulsive disorder patients.” In: *Scientific reports* 6, p. 31333.
- ZESCHNIGK, M., MARTIN, M., BETZL, G., KALBE, A., SIRSCH, C., BUITING, K., GROSS, S., FRITZILAS, E., FREY, B., RAHMANN, S., and HORSTHEMKE, B. (2009). “Massive parallel bisulfite sequencing of CG-rich DNA fragments reveals that methylation of many X-chromosomal CpG islands in female blood DNA is incomplete.” In: *Human molecular genetics* 18.8, pp. 1439–48.
- ZHANG, L., CHEN, J.-G., and ZHAO, Q. (2015). “Regulatory roles of Alu transcript on gene expression.” In: *Experimental cell research* 338.1, pp. 113–8.
- ZHANG, L., WANG, Y., ZHANG, X., ZHANG, M., HAN, D., QIU, C., and HAN, Z. (2012). “Dynamics of phytohormone and DNA methylation patterns changes during dormancy induction in strawberry (*Fragaria Å— ananassa* Duch.)” In: *Plant cell reports* 31.1, pp. 155–65.

-
- ZHANG, Y., ROHDE, C., TIERLING, S., JELTSCH, A., and REINHARDT, R. (2009). “Chapter 14 DNA Methylation Analysis by Bisulfite Conversion , Cloning , and Sequencing of Individual Clones”. In: *Methods* 507, pp. 177–187.
- ZHANG, Y., ZHOU, L., BAO, Y. L., WU, Y., YU, C. L., HUANG, Y. X., SUN, Y., ZHENG, L. H., and LI, Y. X. (2010). “Butyrate induces cell apoptosis through activation of JNK MAP kinase pathway in human colon cancer RKO cells.” In: *Chemico-Biological Interactions* 185.3, pp. 174–181.
- ZHANG, Z.-M., LIU, S., LIN, K., LUO, Y., PERRY, J. J., WANG, Y., and SONG, J. (2015). “Crystal Structure of Human DNA Methyltransferase 1.” In: *Journal of molecular biology* 427.15, pp. 2520–31.
- ZHAO, S., FUNG-LEUNG, W.-P., BITTNER, A., NGO, K., and LIU, X. (2014). “Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells.” In: *PloS one* 9.1, e78644.
- ZILLER, M. J., GU, H., MÜLLER, F., DONAGHEY, J., TSAI, L. T.-Y., KOHLBACHER, O., DE JAGER, P. L., ROSEN, E. D., BENNETT, D. A., BERNSTEIN, B. E., GNIRKE, A., and MEISSNER, A. (2013). “Charting a dynamic DNA methylation landscape of the human genome.” en. In: *Nature* 500.7463, pp. 477–81.
- ZOPF, C. J., QUINN, K., ZEIDMAN, J., and MAHESHRI, N. (2013). “Cell-cycle dependence of transcription dominates noise in gene expression.” In: *PLoS computational biology* 9.7, e1003161.

A

Abbreviations

%	Percent
°C	Degree Celsius
<i>g</i>	Gravitational acceleration (9,81 m/s ²)
μL	Microliter
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methyl-cytosine
5mC	5-methylcytosine
A	Log2RxG
A2UCOE	Ubiquitously acting chromatin-opening element
AdoMet	S-adenosyl-methionine
ATP	Adenosine triphosphate
BER	Base excision repair
bp	Basepairs
Ca.	Circa
CAB-Seq	Chemical Modification-Assisted Bisulfite Sequencing
CDS	Coding sequence
CGI	CpG island
CH	Chinese hamster
ChIRP-seq	Chromatin Isolation by RNA Purification sequencing

CHO	Chinese hamster ovary
COBRA	Combined Bisulfite Restriction Analysis
CpG	Cytosine-Phosphate-Guanine (Dinucleotide)
Cy3-dUTP	Cyanine3-5-(3-aminoallyl)-2'-deoxy-uridine-5'-triphosphate
Cy5-dUTP	Cyanine5-5-(3-aminoallyl)-2'-deoxy-uridine-5'-triphosphate
d	Day
DMH	Differential methylation hybridization
DMH	Differential Methylation Hybridization
DMR	Differentially methylated region
DNA	Desoxyribonucleic acid
DNase I	Desoxyribonuclease I
DNMT	DNA Methyltransferase
dNTP	Desoxynucleotide triphosphat
EDTA	Ethylendiamintetraacetate
EF-1a	Elongation factor 1 alpha
ENCODE	Encyclopedia of DNA Elements
ER	Endoplasmic reticulum
ERAD	ER-associated degradation
EtOH	Ethanol
EWAS	Epigenome-wide association studies
FDA	Food and Drug Administration
Fig.	Figure
g	Gramm
GAGE	Generally applicable gene-set enrichment
gDNA	Genomic DNA
Glc	Glucose
GO	Gene ontology
h	Hour
H3K27	Histone H3 lysine 27
HAT	Histone acetyl transferase
HDAC	Histone deacetylase
HepG2	Human liver cancer cell line HepG2
HMD	Highly methylated domain
IL8	Interleukin 8
IMR90	Human foetal lung (IMR-90) cells
kb	Kilobases
KEGG	Kyoto Encyclopedia of Genes and Genomes
Lac	Lactate
lncRNA	long noncoding RNA

M	Log ₂ R/G
mAB	Monoclonal antibody
MBD	Methyl-CpG-binding domain
MDR	Methylation determining region
MeDIP	Methylated DNA Immunoprecipitation
min	Minutes
miRNA	MicroRNA
miRNA	microRNA
mL	Milliliter
mRNA	messenger RNA
MS-HRM	Methylation-sensitive high-resolution melting
MSP	Methylation-specific PCR
MSRE-PCR	Methylation-Sensitive Restriction Enzyme Digestion PCR
NaBu	Natriumbutyrate
ncRNA	Non-coding RNA
ncRNA	Non-coding RNA
NKCL	Natural killer cell lymphoma
nt	Nucleotides
o/e	Observed/expected ratio
oxBS-Seq	Oxidative Bisulfite Sequencing
p	Probability
PCR	Polymerase chain reaction
PMD	Partially methylated domain
pri-miRNA	Primary microRNA
Qp	Cell specific productivity
qPCR	Quantitative realtime PCR
redBS-Seq	Reduced Bisulfite Sequencing
RISC	RNA-induced silencing complex
RLU	Relative light units
RNA	Ribonucleic acid
RNAi	RNA interference
RNAPII	RNA polymerase II
RNase	Ribonuclease
rpm	Rounds per minute
rRNA	Ribosomal RNA
RT	Room temperature
SCFA	Short chain fatty acid
SH-SY5Y	Neuroblastoma cell line SH-SY5Y
siRNA	Small interfering RNA

snoRNA	small nucleolar RNA
SNP	Single nucleotide polymorphism
SV40	Simian virus 40 early region promoter
Tab.	Table
TDG	Thymine DNA glycosylase
TF	Transcription factor
tRNA	transfer RNA
TSS	Transcription start site
TTS	Transcription termination site
UTR	Untranslated region
VCD	Viable cell density
Viab	Viability
WGA	Whole genome amplification

B

Supplementary material

In the following four sections experimental results are summarized that did not lead to a satisfying final conclusion. First, the attempt to identify the transgene integration site within the CHO DP-12 genome is described that was performed on the basis of whole-genome sequencing of the CHO DP-12 cell genome. No reliable results could be obtained in this case because the paired-end reads were too short for the stringent mapping procedure that would have been necessary.

Second, the analysis of a putative involvement of microRNA 199a in the response of CHO DP-12 cells to butyrate treatment is presented, which did not allow for a final answer as well due to the complexity of the topic. However, a broader experimental setup was not possible within the scope of this thesis.

In a third experiment, endogenous CHO cell promoters were tested regarding their expression potential and capability to potentially replace the methylation prone CMV promoter. Generally, the endogenous promoter sequences were found to be functional, but further studies need to examine if these promoters are targeted by DNA methylation or histone modification events as well.

Lastly, the CHO-specific CpG island microarray was applied to analyze and compare two CHO-XL99 clones with differences in production characteristics that were kindly provided by the Australian Institute of Bioengineering and Nanotechnology (AIBN). Unfortunately, some microarrays were affected by hybridization artifacts that did not allow for a comparison between some relevant samples.

B.1. Attempt to identify transgene integration sites in the CHO DP-12 genome

In order to assess the effect of butyrate treatment on DNA methylation of the IgG gene integration site within the CHO DP-12 genome, the location of the recombinant gene needed to be investigated. Therefore, the unmapped reads from the first sequencing run were aligned to the sequence of the vector that was used to generate the CHO DP-12 cell line (Fig. B.1). Surprisingly, in contrast to an average coverage of the genome (post trimming and mapping) of about 34-fold, a coverage of the vector sequence of 350-fold was detected. This finding hints towards multiple copies of the vector in the CHO DP-12 genome, which further complicated identification of the integration site. The plot of coverages against the vector sequence also shows that the coverage is significantly higher for the first 4000 bases, which represent the coding region of the recombinant gene. In some cases the non-coding part of the vector seemed to be removed from the genome.

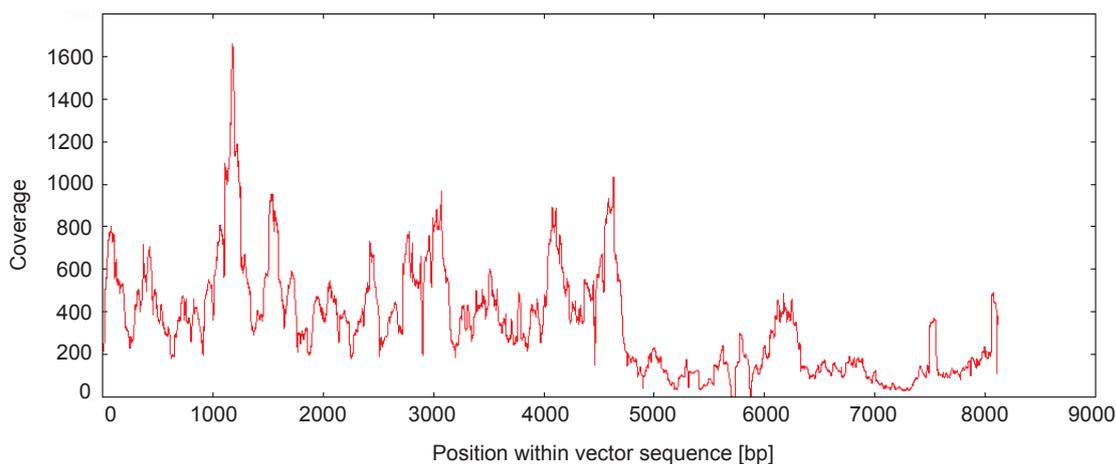


Figure B.1.: Coverage of the recombinant CHO DP-12 vector DNA by reads that could not be mapped to the SNP-corrected CHO K1 genome. The region from 40 bp to 4300 bp represents the sequence coding for heavy and light chain of the anti-IL8-antibody.

In an attempt to identify the integration site, the sequencing data generated from native CHO DP-12 DNA was searched for paired-end reads that mapped to the vector sequence with one read and to the CHO K1 genome with the corresponding second read. Hits were supposed to accumulate around the putative integration site. Fig. B.2 shows an exemplary mapping result for the recombinant vector sequence on the x-axis and a section of CHO scaffold JH002371.1 on the y-axis. Positions where paired-end reads could be mapped to both sequences are marked by a red cross. Interestingly, seven positions within this scaffold, which co-localized with genes encoding IgG heavy chain regions, showed accumulation of mapping hits at the beginning and end of the expression cassette. Similar results were found for several other scaffolds as well with more or less clear accumulation patterns. Based on

this mapping result it was speculated (i) that the short paired-end sequencing reads could lead to a biased mapping process, which led to several false-positive hits within genes very similar to the recombinant IgG gene or (ii) that the integration of the recombinant DNA, which was performed randomly for CHO DP-12 cell line generation, happened by a mechanism similar to homologous recombination. However, only repetition of sequencing with longer reads that would allow for more stringent mapping procedures would have enabled us to reliably exclude one or the other option.

As interpretation of the results for transgene integration site identification was not possible by bioinformatics measures, a putative integration site was supposed to be analyzed by PCR. For this, a *Long PCR Enzyme Mix* (ThermoFisher) was used, which is supposed to be able to amplify extremely long DNA fragments of up to 46 kb. In this experiment CHO DP-12 DNA was compared to DNA from the Chinese hamster. Primers were designed that flanked a putative integration region of 6504 bp, as the exact integration position could not be predicted. Fig. B.3 shows a 0.6 % agarose gel used for separation of long-range PCR reactions that were performed at 53 °C and 56 °C. At both temperatures amplicates of a size of about 6500 bp were generated from Chinese hamster DNA. This showed that the primers worked and that, of course, no putative transgene was present in the host DNA. For CHO DP-12 DNA no amplicate could be detected.

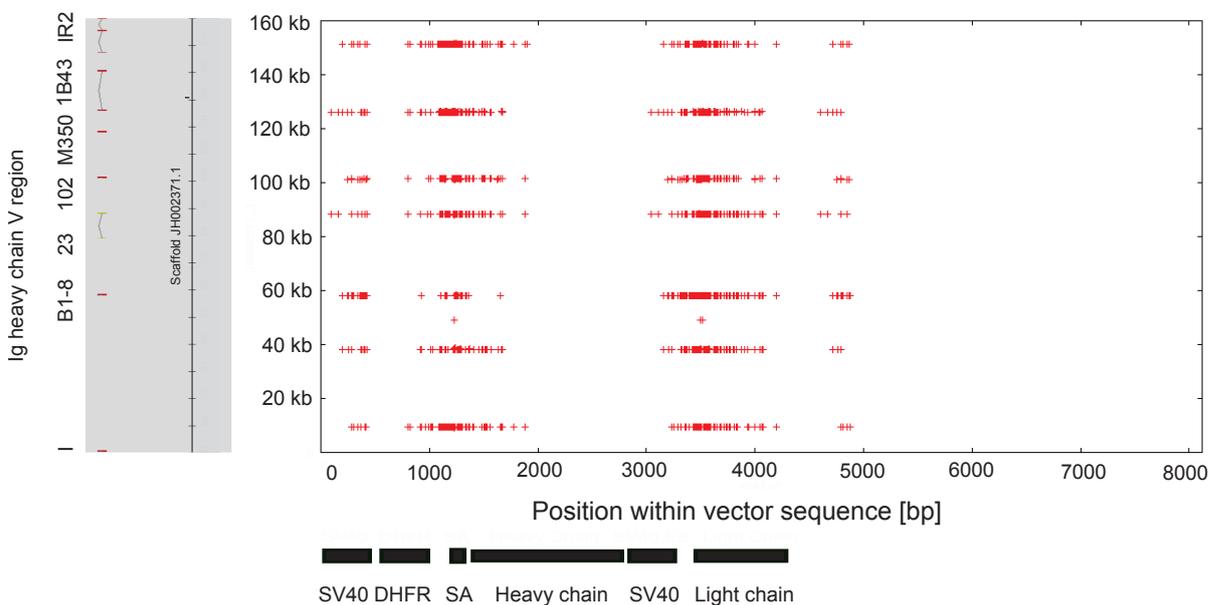


Figure B.2.: Mapping of reads from libraries of native CHO DP-12 DNA to both the recombinant vector sequence (x axis) and the SNP corrected CHO K1 genome (y axis). Each read that could be mapped to both sequences is indicated by a red cross. The figure shows a cluster of IgG V genes on scaffold JH002371.1 with 7 putative integration sites as an example.

There may be various reasons for this result. On the one hand it is possible that the prediction of at least this specific integration site was correct, but that the recombinant vector integrated multiple times. In this scenario the resulting genomic region would have been too long for

amplification. On the other hand it is possible that the primers simply did not work, as a lot of SNPs were detected between CHO K1 and CHO DP-12 cells and it is highly likely that even more SNPs are present between DNA from a Chinese hamster and an antibody-producing CHO cell line.

However, the attempt to identify transgene integration sites in the CHO DP-12 cell genome was stopped at this point. A reliable identification would have been possible e.g. by a 2×300 bp Illumina *MiSeq* sequencing experiment which could not be performed within the scope of this thesis.

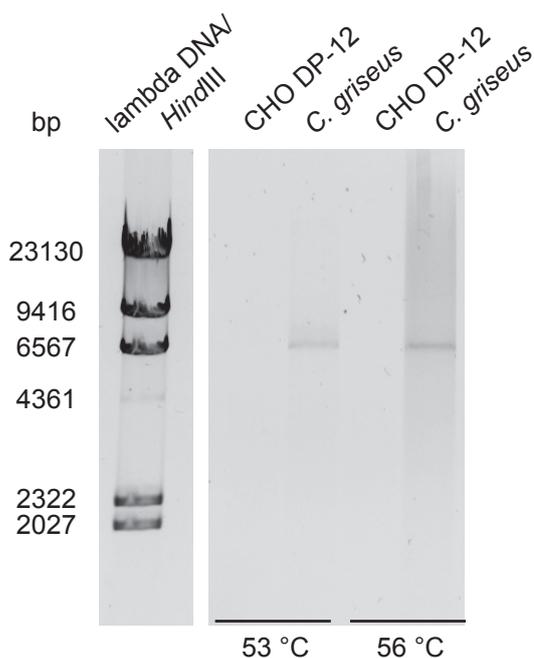


Figure B.3.: Attempt to verify a predicted transgene integration site in the CHO DP-12 genome by long-range PCR. Primers flanked a putative integration site of a size of 6504 bp (i.e. without integrated DNA). CHO DP-12 DNA was compared with *C. griseus* (Chinese hamster) DNA. PCR was performed at annealing temperatures 53 °C and 56 °C.

B.2. Analysis of a putative involvement of microRNA 199a in the response of CHO DP-12 cells to butyrate treatment

As it has been shown that the expression of microRNAs can be regulated by changes in the methylation status of their host genes and, additionally, butyrate has been shown to induce *p21* protein expression by targeting the miR-106b family (Hu et al. 2011), a potential involvement of the miRNA system in the CHO DP-12 cell response to butyrate addition was analyzed. For this reasons, the DNA methylation data generated by the CpG island microarray experiment

B.2. Analysis of a putative involvement of microRNA 199a in the response of CHO DP-12 cells to butyrate treatment

were searched for hints towards an effect of butyrate treatment on the microRNA system and, in fact, several miRNA-containing genes were found to be differentially methylated upon butyrate treatment. These included *Coatomer subunit zeta-2 (Copz2)*, *Ena/VASP-like protein (Evl)* and *Dynammin-2 (Dnm2)*. Human *Dnm2*, a microtubule-associated GTPase involved in clathrin-mediated endocytosis (Raimondi et al. 2011) contains the *miR-199a* gene in intron 15 (Kim et al. 2008) and was selected to be analyzed regarding a potential effect of butyrate addition on miRNA expression. For this purpose, DNMT2 and primary microRNA (pri-miRNA) 199a expression as well as the expression of two miR199a target genes, Prostaglandin G/H synthase 2 (PTGS2) and NAD-dependent protein deacetylase sirtuin-1 (SIRT1) was analyzed by qPCR 24 and 48 hours after butyrate addition (see schematic representation of the regulatory system in Fig. B.4A).

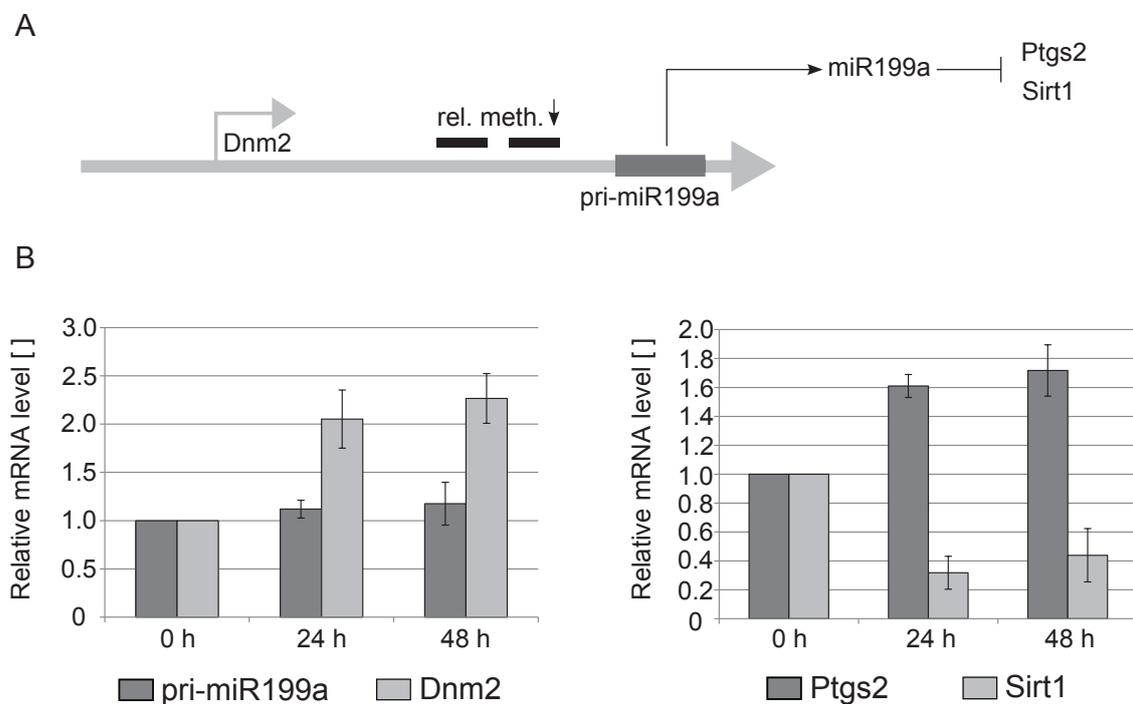


Figure B.4.: Analysis of putative effects of butyrate treatment on the expression of pri-miR199a and miR-199a target genes. **A** Schematic representation of *Dnm2* gene with CpG island positions (black) and the *pri-miR199a* gene located at the 3' terminus. Mature miR-199a leads to repression of target genes SIRT1 and PTGS2. **B** Analysis of relative mRNA levels of pri-miR199a, DNMT2, PTGS2 and SIRT1 in butyrate treated cells by qPCR. Error bars represent standard deviations of triplicate measurements.

Fig. B.4B shows the qPCR results. Although decreased methylation of an intragenic *Dnm2* region located before the miRNA gene was detected, no significant changes in pri-miRNA expression upon butyrate treatment were measured by qPCR. DNMT2 expression, however, was found to be increased by more than 100 %. mRNA levels of the miR1992 targets PTGS2 and SIRT1 were found to be increased by 60 % and decreased by 80 %, respectively. Based on these results, no direct effect of butyrate on miRNA199a-mediated regulation of the analyzed target genes could be shown. As microRNAs regulated a plethora of target transcripts and

several transcripts can be regulated by multiple microRNAs, this result does not allow for a final answer. Further experiments would be necessary, which could for example combine next generation sequencing of gene and microRNA expression after butyrate addition with DNA methylation analyses. An experiment such big was not possible within the scope of this thesis.

B.3. Analysing the expression strength of CHO cell promoters by dual luciferase assays

Several publications show a DNA methylation-dependent silencing of viral promoters used for recombinant gene expression in CHO cells. For example, when Kim et al. analyzed CMV promoter methylation (29 CpG sites) by bisulfite sequencing in mAb producing CHO cell lines during extended sub-culture in the presence of selective pressure, they revealed that the instable, low generation number cell lines were affected by pCMV methylation that strongly increased during sub-culturing (Kim et al. 2011). A correlation between pCMV methylation, instability of productivity and its potential to be used as a biomarker was also addressed by Osterlehner et al., who established a qPCR-based screening method which allows for the enrichment of stable producers early in cell line development (Osterlehner, Simmeth, and Göpfert 2011).

As it was hypothesized that endogenous CHO cell promoters might represent an alternative to viral promoters in terms of either not being silenced as quickly or in response to different conditions, the gene expression potential of 5 CHO promoters was analyzed by luciferase reporter assays. This first estimation of promoter strength was possible because transcription start sites were discovered in CHO cells by Jakobi et al. (Jakobi et al. 2014). The researchers applied a dual-library RNA sequencing approach, which is based on the fact that intact mRNA possess a 5' cap that inhibits ligation of sequencing adapters. The actual peak library was therefore treated with Tobacco acid pyrophosphatase (TAP) to remove 5' caps prior to adapter ligation, whereas a second library was not TAP-treated. In this second library only fragmented mRNAs were sequenced and therefore represented the background noise that could be subtracted from the actual signals. By this approach 6,547 CHO cell transcription starts sites were identified that were used to annotate putative promoter sequences and allowed for a characterization of the CHO cell promoter landscape that was found to be, accordingly to other mammalian cell types, characterized by the presence of broad and sharp type transcription start sites. Furthermore, promoter regions were screened for common regulatory elements such as TATA boxes.

For the analysis of promoter expression strength, 5 promoter sequences associated with high expression levels (i.e. high abundance of reads) in the RNA sequencing experiment performed

by Jakobi et al. were examined in dual-luciferase reporter assays regarding their recombinant gene expression potential (Tab. B.1). Two of them could not be assigned to any gene, two were associated with isoforms of *Histone H4-like protein (Hist1h4n)* and one was assigned to *Ribosomal protein S3 (Rps3)*. Besides the fact that all of the selected promoters were associated with high expression levels, they varied regarding the types of expression initiation (sharp or broad) and the presence of regulatory elements. These elements included the core promoter motifs TATA box, downstream processing element (DPE), initiator region (INR) and TFIIB recognition element (BRE), which represent sequence elements that can allow for basal transcription from mammalian core promoters (Smales et al. 2004; see schematic representation in Fig. B.5).

Table B.1.: Key features of 5 endogenous CHO promoters examined by dual luciferase assays regarding the potential of recombinant gene expression (Jakobi et al. 2014). Indicated are CHO scaffold numbers, assigned genes (if possible), heights of the peaks generated from mapping of RNA sequencing reads to the reference genome, types of peaks (sharp or broad) and identified regulatory elements. DPE = downstream processing element, INR = initiator region, BRE = TFIIB recognition element. NA = not assigned, ND = not detected.

No.	Source	Gene	Peak height	Peak type	Elements
1	Scaffold 157	NA	921,977	Sharp	ND
2	Scaffold 115	NA	376,649	Broad	TATA box, DPE
3	Scaffold 69	<i>Hist1h4n</i>	849,029	Sharp	TATA box, INR
4	Scaffold 5	<i>Rps3</i>	159,343	Broad	BRE
5	Scaffold 37	<i>Hist1h4n</i>	218,538	Sharp	TATA box

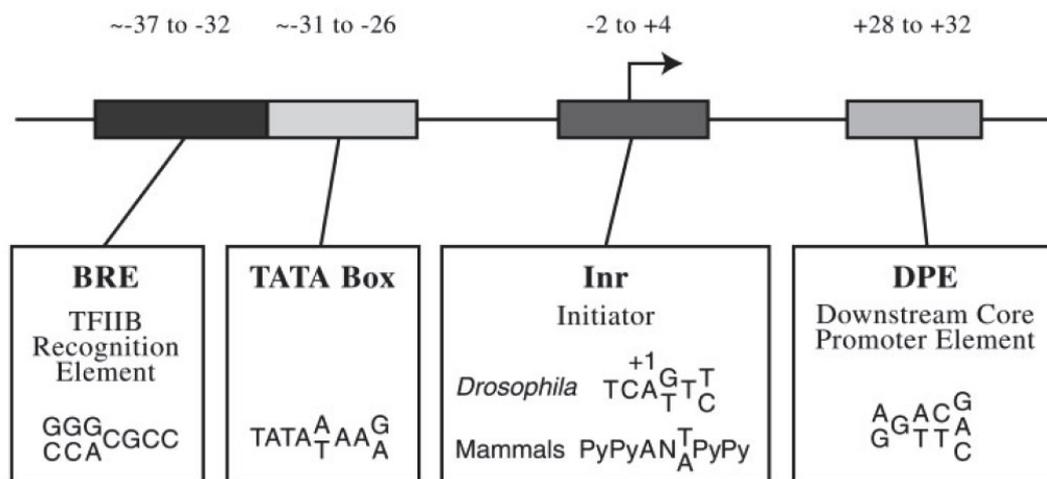


Figure B.5.: Core promoter sequence motifs (Smales et al. 2004).

Reporter assays were performed by cloning of the selected promoter sequences into the multiple cloning site of a *Green Renilla Luciferase* vector. Transient co-transfections of the experimental promoters within the *Renilla Luciferase* plasmid and a normalization plasmid that contained the *Thymidine kinase (Tk)* promoter within a *Firefly Luciferase* vector was

performed using electroporation. Fig. B.6 outlines the experimental procedure. Central to the experiment is the fact that the amount of relative light units (RLU) produced by *Green Renilla Luciferase* depends on the expression strength of the endogenous CHO promoter, whereas the luminescence signal from the constitutively expressed *Tk* promoter does not change regardless of any intrinsic regulations, but depends on experimental conditions such as the transfection efficiency. For this reason, *Tk* promoter-dependent luciferase expression can be used for normalization in dual luciferase assays, which are capable of measuring both luciferase activities sequentially from a single sample. In the assay used here, *Firefly Luciferase* activity was determined first by adding beetle luciferin and coenzyme A, which results in a stable luminescence signal. The signal generated in this way is subsequently quenched and the substrate for *Renilla* luciferase, coelenterate-luciferin, is added (*Dual Luciferase*[®] Reporter Assay System protocol, Promega).

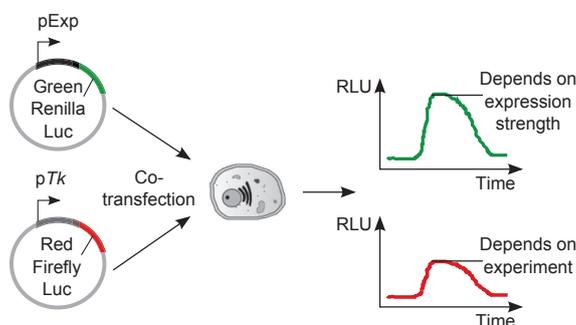


Figure B.6.: Principle of dual luciferase reporter assays. Endogenous CHO promoter sequences ('pExp') were cloned into a plasmid encoding the *Green Renilla Luciferase* gene and co-transfected into CHO K1 and CHO DP-12 cells simultaneously with a normalization plasmid which contained the constitutive *Tk* promoter ('pTk') and a *Red Firefly Luciferase*. The luminescent signal produced by *Renilla Luciferase* abundance was normalized to the *Red Firefly* signal in regard to varying experimental conditions such as the transfection efficiency.

Experiments were performed using both CHO K1 and CHO DP-12 cells. The results of luminescence measurements are shown in Fig. B.7. For both cell lines, untransfected control cells (NTC) and cells transfected with the empty vector (MCS) showed reporter signals close to the detection limit. Similarly low signals were observed for the endogenous CHO promoters 2 and 5, which showed expression potentials of less than 1 % of the CMV promoter strength. The promoters 1 and 3 were found to possess an expression strength of 11 % relative to the CMV promoter in CHO DP-12, but not CHO K1 cells (0.5 % and 1.5 %, respectively). Promoter 4 exhibited a comparable expression potential in both cell lines with 5.5 % and 7.5 % of the CMV promoter expression capacity. Expression potential could not be related to certain promoter types or specific functional elements, as both broad and sharp type promoters were amongst the three functional promoters. Also, they were associated with either no regulatory elements (promoter 1), with a TATA box and INR (promoter 3) or with a BRE box (promoter 4).

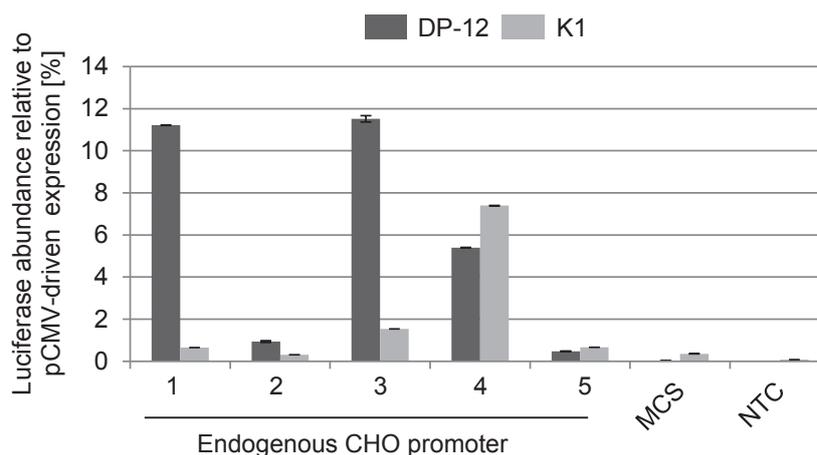


Figure B.7.: Reporter assay for measurement of the expression potential of endogenous CHO promoters in comparison to the CMV promoter. 'MCS' = CHO cultures transfected with the empty *Green Renilla* vector, 'NTC' = untransfected control cells. Error bars represent standard deviations of triplicate measurements.

B.3.1. Discussion: Bioinformatically identified CHO promoters are functional

The evaluation of expression potentials of newly discovered CHO cell promoters proved that the annotated transcription start sites were indeed functional and could be employed to mediate recombinant gene expression. Three of the five promoter sequences that were tested showed an expression potential of about 1/10 of the CMV promoter strength in CHO DP-12 cells and one of them in CHO K1 cells. Obviously, expression from endogenous CHO promoters depends on intrinsic cellular pathways that are active or inactive in a specific CHO cell line under given culturing conditions. This might be an explanation for the differences found between CHO K1 and DP-12 cells, as CHO DP-12 cells were selected for a high producer phenotype while CHO K1 cells do not produce any recombinant gene and were not subjected to a selection process. Therefore, a general statement on the functionality of endogenous CHO promoters might be impossible and promoters should be tested in the specific CHO cell line they are supposed to be used in.

A more reliable screening for functional sequences with even higher expression potentials than those found in this experiment would furthermore require examination of a larger number of promoter sequences. Such an analysis would also possibly enable an estimation about preferences regarding certain functional elements, which was not possible based on the small number of promoters used here. However, this experiment shows that endogenous CHO promoters could indeed be used to replace viral promoters, although optimization of culturing conditions or the promoters themselves would possibly be required. Strong enhancement of expression potentials was for example shown to be possible by artificial combination of regulatory elements such as TATA box and Initiator Region (INR) in metazoans (??). In

CHO cells expressing darbepoetin it was shown that a synthetic promoter in combination with codon optimization of the recombinant gene was capable of achieving high expression levels (Shukurov et al. 2014).

Endogenous CHO promoter elements might be capable of replacing methylation-prone viral promoters, but further studies need to examine if these promoters are targeted by DNA methylation or histone modification events as well. However, in general this experiment provided a basis for further experiments and underlines the usefulness of a closer examination of endogenous CHO cell promoters.

B.4. Comparative CpG island microarray analysis of CHO clones with differences in productivity

The CHO-specific CpG island microarray was supposed to be applied in a second experiment to analyze and compare two CHO-XL99 clones with differences in production characteristics that were kindly provided by the Australian Institute of Bioengineering and Nanotechnology (AIBN). The clones were derived from the original CHO K1 cell line by transfection with a plasmid encoding the gene of a therapeutic IgG1 antibody, a puromycin resistance gene and a glutamine synthetase gene. Antibody expressing cell pools (in the following referred to as 'clones' - however, it has to be noted that no single clones were isolated and analyses were performed on polyclonal populations) were selected by puromycin and methionine sulfoximine (MSX) and positive clones were either kept without selection pressure or with supplementation of puromycin. Samples for CpG island microarray analysis of one clone with poor production behavior and one clone with good production characteristics were taken from a 7 day batch cultivation at day 4 or 5 and shipped to Bielefeld. A parallel batch cultivation with identical sampling was performed for the host CHO K1 cell line in Bielefeld.

B.4.1. Confirmation of *Actb* and *Bcat1* promoter regions as methylation controls in CHO-XL99

The promoter regions of *Actb* and *Bcat1* were previously identified as completely methylated and unmethylated in CHO DP-12 cells, respectively. No changes in their DNA methylation status could be detected upon butyrate addition. Therefore they were used as unmethylated and methylated controls. To assess the potential of these promoter regions as methylation controls in CHO-XL99 cells as well, their methylation status was determined by bisulfite sequencing. The *Actb* promoter region exhibited completely unmethylated CpG dinucleotides, whereas the *Bcat1* promoter region was found to be completely methylated in CHO-XL99

(Fig. B.8). Therefore, these regions could be used as methylation controls for the following CpG island microarray experiments.

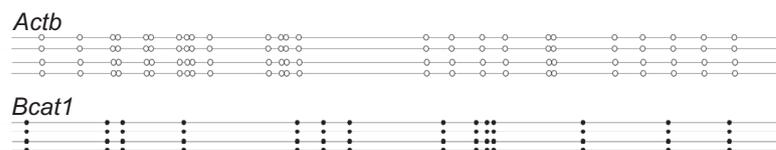


Figure B.8.: Confirmation of *Actb* and *Bcat1* promoter regions as methylation controls in CHO-XL99 by bisulfite sequencing. Black circles represent a methylated CpG position, white circles indicate unmethylated CpG positions.

B.4.2. Batch cultivation of CHO-XL99 clones

After transfection, mAb producing CHO-XL99 clones 'C' and 'G' were cultivated with selection pressure by puromycin ('Early cell line development' in Fig. B.9). Gene amplification was performed in a second round of selection with both puromycin addition and MSX ('Mid cell line development') and the clones were subsequently subjected to prolonged passaging without selection pressure ('Late cell line development'). In all stages of cell line development, the 'G' clone showed lower maximum viable cell densities compared to the 'C' clone, but higher mAb expression levels during early and mid cell line development (i.e. when selection pressure was added; 'Combined expression levels' in Fig. B.9). Without selection pressure, mAb expression in the 'G' clone was low. On the contrary, the 'C' clone showed comparatively high mAb expression only during early cell line development but mAb production could not be increased by MSX selection and was also low without any selection pressure. Therefore, the 'G' clone is in the following referred to as 'the good producer' and the 'C' clone is referred to as 'the bad producer'. Sample generation for CpG island microarray analysis was done by performing seven day batch cultivation with addition of puromycin only, puromycin and MSX and without adding selection pressure. Also, the parental cells were cultivated and samples from each were taken in triplicates in the middle of the exponential growth phase.

B.4.3. CpG island microarray analysis of good and bad producing clones in comparison to the CHO K1 host cell

DNA was extracted from three replicate samples of each the CHO K1 host cell samples and the good and bad producer clones that were kept without selection pressure and with selection pressure, respectively. DNA integrity was checked on agarose gels after DNA extraction and no degradation was observed. DNAs were then sonicated (see images of control agarose gels in Supplementary Fig. B.21) and methylated DNA fragments were

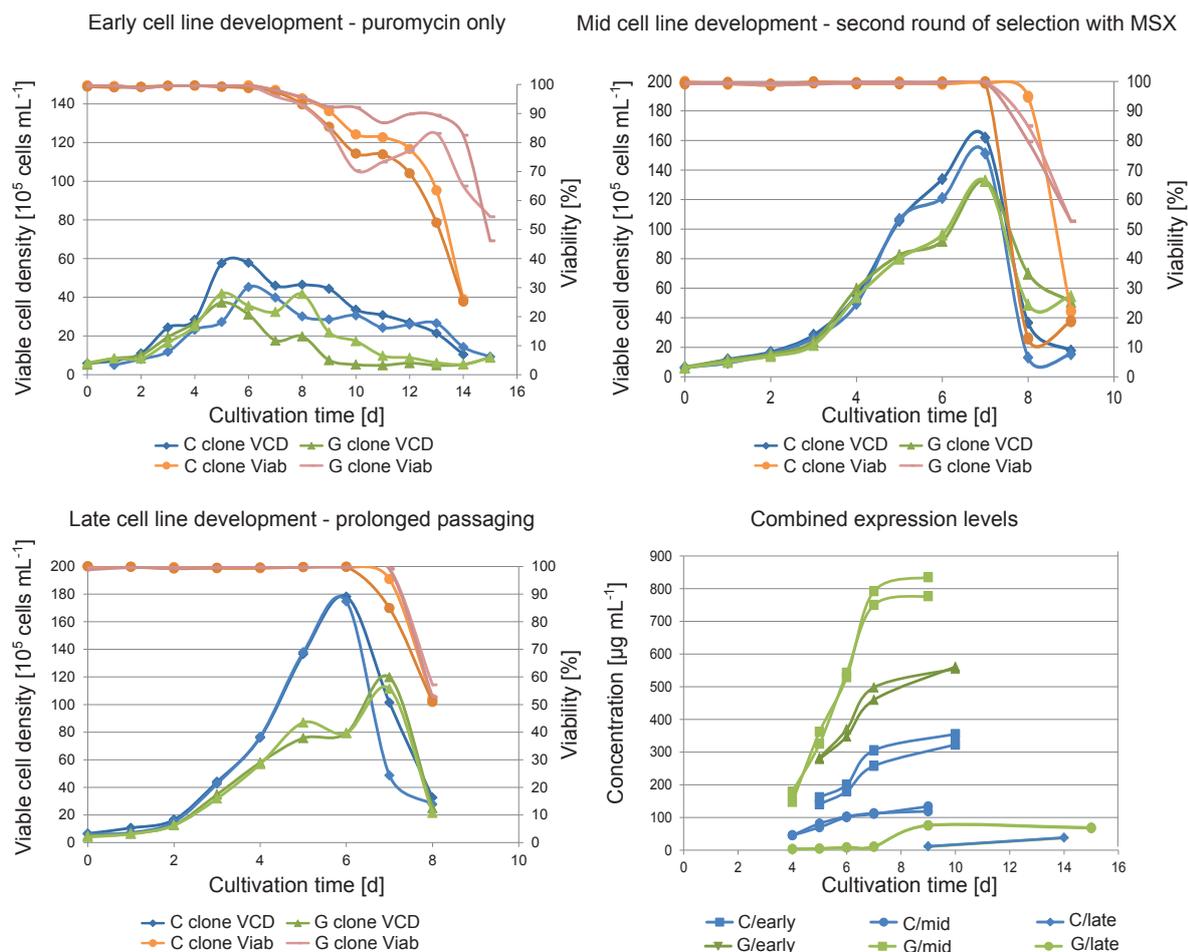


Figure B.9.: VCDs, viabilities and mAb expression levels of CHO-XL99 clone cultivations during early, mid and late cell line development.

enriched. Enrichment efficiencies were controlled by qPCR with primers specific for *Actb* and *Bcat1*, respectively. The enriched methylated DNA fractions of each three replicates were pooled and compared to the CHO K1 host DNA by CpG island microarray analysis, which was performed and controlled as described previously. Unfortunately, several microarrays exhibited hybridization artifacts. Examples are shown in Fig. B.10. Hybridization artifacts can occur due to difficulties in handling of microarrays in a $8 \times$ format because the volume of sample plus hybridization buffer is very small. Several arrays exhibited regions where air bubbles stuck during hybridization or regions with low signal intensities in the middle part of the array when too little hybridization mixture was present inside the chamber. These artifacts were obtained regardless of the person setting up the array sandwich, even when a very well trained person performed the experiments. Arrays exhibiting such artifacts were excluded from the analysis and, consequently, also arrays from dye swap pairs when the statistical power would have been too low for reliable calculations of probabilities. For this reason, all samples generated from cells under both puromycin and MSX selection were excluded from the analysis.

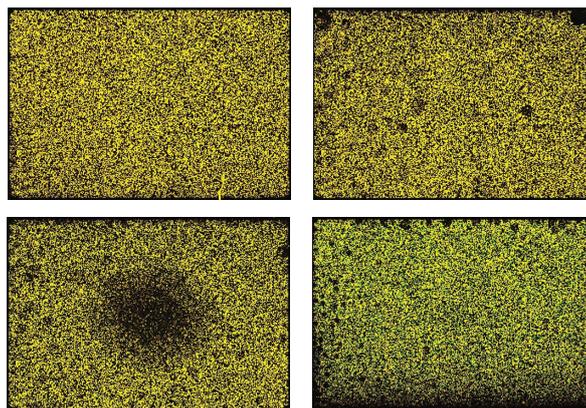


Figure B.10.: Exemplary images of artifacts in hybridized CpG island microarrays after scanning and applying Agilent's *Feature Extraction* software. Top left: Array without artifacts. Top right: Small air bubbles. Bottom left: Too little hybridization mixture to achieve good signal intensities in the middle of the array. Bottom right: Weak signals at one side.

B.4.4. Data analysis

In a next step data were normalized and t-tests were calculated for each comparison (i.e. both clones with and without selection pressure were compared to the parental cells). To get an initial overview, MA plots were generated for these tests (Fig. B.11). Probes with M-values between -1.0 and 1.0 (i.e. not significantly differentially methylated) are represented in black, probes with M-values below -1.0 are represented in red (i.e. hypomethylation compared to the host cells) and probes with M-values above 1.0 (i.e. hypermethylation compared to the host cells) are represented in green. The probes represented in black, i.e. those without a significant change in DNA methylation, showed the characteristic distribution that was observed for other CHO cell lines as well: The majority was either unmethylated (low A values/intensities) or methylated (high A values/intensities), a minority showed intermediate methylation levels. Furthermore, the MA plots indicated that less genes were significantly differentially methylated when both clones were cultivated without selection pressure. Under selection with puromycin both clones exhibited a higher degree of differential methylation in comparison to the host cells. According to the expectations, puromycin treatment resulted in an increase in the numbers of differentially methylated CpG islands. For subsequent analyses and in accordance with the established CpG island microarray analysis procedure, probes with M-values of ≤ -1.0 and ≥ 1.0 were considered to be significant when they exhibited A values of more than 7.0 and p adjusted values of less than 0.05.

Fig. B.12 shows the numbers of differentially methylated CpG islands present in the genomes of the good and the bad producer when each of them was compared to the CHO K1 host cell. Both clones had similar numbers of differentially methylated genes when they were kept without selection pressure. The bad producer exclusively exhibited 377 hypomethylated and 108 hypermethylated genes, whereas 291 genes were exclusively hypomethylated and 152

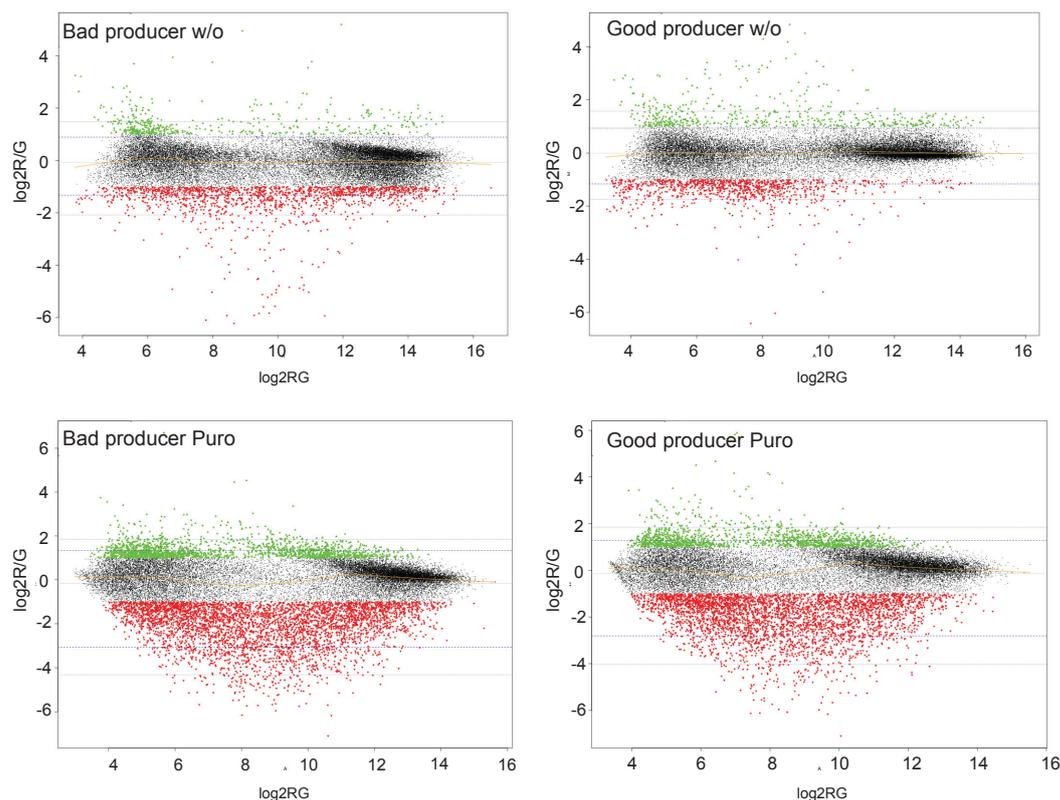


Figure B.11.: MA plots for the comparison of good and bad producing CHO-XL99 clones to the CHO K1 host cell line. M-values below the threshold of -1.0 (red) represent hypomethylated regions compared to the CHO K1 host cell line. M-values above the threshold of 1.0 (green) represent hypermethylated regions compared to the CHO K1 host cell line. The blue lines mark for each time point the 95 % and 99 % confidence intervals of all observed M-values. The yellow line represents the LOWESS fit.

genes were exclusively hypermethylated in the good producer in comparison to the CHO K1 host cell. 83 genes were differentially hypomethylated and 17 genes were differentially hypermethylated in both clones compared to the host cell. These differences represent the effect of transfection, selection and gene amplification during cell line generation and are of greatest relevance regarding the hypothesis that epigenetic states are established during this procedure that can lead to unfavorable outcomes of production processes.

When both clones were subjected to puromycin selection, differential methylation increased especially regarding the numbers of differentially hypomethylated genes. 1,140 genes were hypomethylated in both clones, 1,128 were exclusively hypomethylated in the bad producer and 331 were exclusively hypomethylated in the bad producer. The effect of puromycin selection on differential hypermethylation was comparably small, with 182 hypermethylated genes in the bad producer, 29 genes in the good producer and 6 overlapping genes.

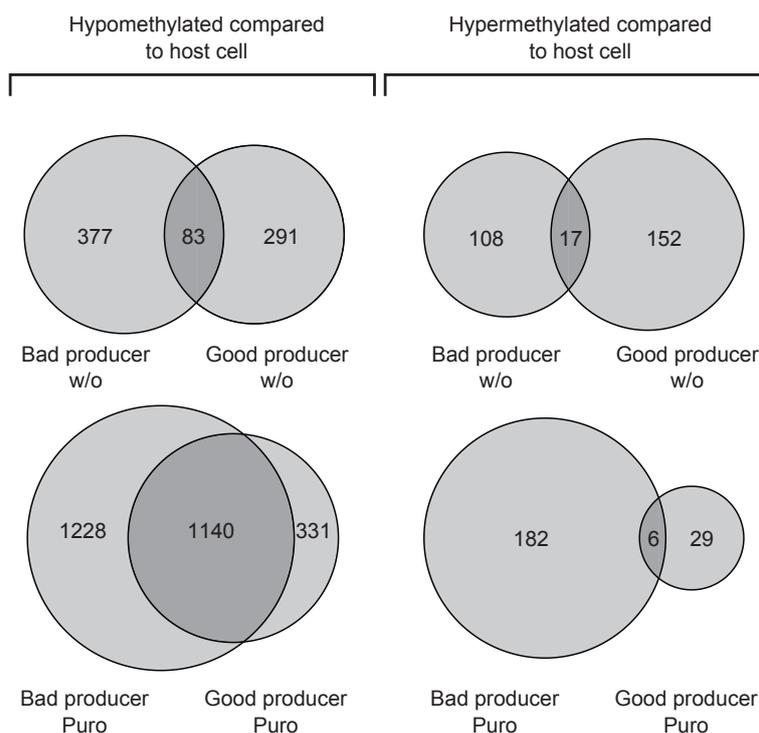


Figure B.12.: Significantly differentially methylated genes in good and bad producing CHO-XL99 clones with ('Puro') and without ('w/o') selection pressure compared to the CHO K1 host cell line.

B.4.5. Functional annotation of differentially methylated genes in good and bad producing CHO-XL99 clones

To assess if specific functional groups of genes were affected by differential DNA methylation Gene Ontology analyses were performed with EASE scores below 0.05. The genes that were identically differentially methylated in both clones were related to transport and gene expression. Besides this, the clones differed significantly regarding the functions of differentially methylated genes, as most of the functional groups enriched in the genes that were detected to be exclusively differentially methylated in the good producer (Tab. B.2) were related to cellular response reactions, including e.g. response reactions to hormones, vitamins or nutrients. This finding raises the possibility that in this case selection resulted in growth of a cell pool more responsive to components of the culture medium than other clones, i.e. a beneficial phenotype that could possibly be related to the good performance of the good producer.

In the bad producer (Tab B.3), genes related to cytoskeleton organization, biosynthetic and metabolic processes were found to be differentially methylated. Also, genes contributing to the response to hypoxia were overrepresented. This is interesting as it was shown in human cancer cells that hypoxia-related genes encode proteins that contribute to processes such as glucose transport, proliferation and apoptosis (Chi et al. 2006). Strikingly, cell death-related

Table B.2.: Gene Ontology analysis of genes associated with differentially methylated CpG islands in replicate cultures of the good producing CHO-XL99 cell line without selection pressure.

GO term (category 'biological process')	%	Fold enrichment	<i>p</i> value	Benjamini <i>p</i> value
Response to estrogen stimulus	2.42	4.07	0.004	0.997
Response to nutrient levels	3.32	2.99	0.004	0.960
Response to vitamin	1.81	4.86	0.008	0.986
Response to extracellular stimulus	3.32	2.67	0.008	0.969
Response to hormone stimulus	4.53	2.19	0.009	0.951
Regulation of wnt receptor signaling pathway	1.51	5.81	0.010	0.945
Translational elongation	2.11	3.71	0.012	0.910
RNA processing	5.74	1.86	0.014	0.927
Response to nutrient	2.42	3.06	0.016	0.931
Response to endogenous stimulus	4.53	1.98	0.020	0.922
Response to retinoic acid	1.21	6.48	0.023	0.938
cAMP biosynthetic process	0.91	11.46	0.027	0.943
Response to steroid hormone stimulus	2.72	2.51	0.028	0.935
Response to organic substance	6.65	1.63	0.029	0.932
Sulfur amino acid biosynthetic process	0.91	10.69	0.031	0.937
Cyclic nucleotide metabolic process	1.21	5.63	0.033	0.941
Detection of mechanical stimulus	0.91	10.03	0.035	0.941
Negative regulation of catalytic activity	3.32	2.12	0.035	0.933
Cellular response to hormone stimulus	2.11	2.81	0.038	0.941
Regulation of gene-specific transcription	2.11	2.79	0.039	0.939
RNA splicing	3.32	2.07	0.040	0.936
Negative regulation of specific transcription from RNA polymerase II promoter	1.21	5.22	0.041	0.930
Response to vitamin A	1.21	5.09	0.043	0.934

genes were significantly overrepresented, including some major cellular regulators such as *Myc* or *Death-associated protein kinase 1 (Dapk1)*. As an epigenetic 'predisposition' toward apoptotic processes that results from the procedure of cell line generation could be a reason for the variable performance of CHO production cell lines, the network of differentially methylated apoptotic genes was analyzed more closely (Section B.4.6).

Upon puromycin treatment 1,140 genes became hypomethylated in both the good and bad producing CHO-XL99 clones. These genes were related to the functional groups of 'phosphate metabolism/phosphorylation', 'calcium ion transport', 'chromatin modification', 'signaling' and 'cell death' (Tab. B.4). As puromycin treatment induces cell death in those cells that do not carry the resistance gene (anymore) or maybe integrated the gene at a silent locus within the genome, apoptotic processes could be expected. The occurrence of genes related to the processes phosphorylation, calcium signaling and chromatin modification might be an indication for growth and protein production within those cells that are resistant to puromycin.

Fewer genes were found to be uniquely differentially methylated in the good producer upon puromycin treatment in comparison to the bad producing clone. However, both groups comprised the functional categories 'gene expression' and 'signaling'. Differentially methylated

genes in the bad producer furthermore belonged to the categories 'metabolism', 'macro-molecule biosynthesis', 'transport' and 'exocytosis' (data not shown).

Table B.3.: Gene Ontology analysis of genes associated with differentially methylated CpG islands in replicate cultures of the bad producing CHO-XL99 cell line without selection pressure.

GO term (category 'biological process')	%	Fold enrichment	<i>p</i> value	Benjamini <i>p</i> value
Actin filament-based process	3.61	2.74	0.003	0.988
Cell death	7.22	1.84	0.004	0.951
Death	7.22	1.83	0.004	0.885
Actin cytoskeleton organization	3.33	2.70	0.005	0.858
Response to oxygen levels	2.50	3.25	0.007	0.819
Auditory receptor cell development	0.83	21.80	0.008	0.813
Ubiquitin-dependent protein catabolic process	3.33	2.52	0.008	0.800
Programmed cell death	6.11	1.83	0.009	0.789
Triglyceride biosynthetic process	0.83	16.95	0.013	0.805
Apoptosis	5.83	1.77	0.015	0.812
Phospholipid biosynthetic process	1.94	3.49	0.015	0.792
Regulation of RNA metabolic process	13.61	1.37	0.017	0.796
Regulation of transcription, DNA-dependent	13.33	1.38	0.018	0.775
Response to protein stimulus	1.94	3.33	0.019	0.768
Neutral lipid biosynthetic process	0.83	12.71	0.022	0.795
Acylglycerol biosynthetic process	0.83	12.71	0.022	0.795
Cytoskeleton organization	4.44	1.87	0.025	0.819
Glycerol ether biosynthetic process	0.83	11.74	0.026	0.817
Negative regulation of molecular function	3.61	1.98	0.032	0.846
Phospholipid metabolic process	2.50	2.41	0.034	0.850
Organophosphate metabolic process	2.50	2.29	0.043	0.875
Response to hypoxia	1.94	2.66	0.048	0.882
Regulation of transcription from rna polymerase ii promoter	6.11	1.54	0.048	0.875

B.4.6. Differential methylation of apoptotic networks in bad producing clones

Gene Ontology analyses showed a significant overrepresentation of cell death-related genes in the bad producer without application of selection pressure. As an epigenetic 'predisposition' toward unfavorable cellular processes resulting from the procedure of cell line generation or due to differences in culture handling was previously suspected to be responsible for the variable performance of CHO production cell lines (Le et al. 2012), the significant presence of differential methylation of genes relevant for apoptotic processes was analyzed more closely. An analysis using the STRING database (Szklarczyk et al. 2015) showed two clusters of associated gene-products with known or predicted interactions. The first cluster comprised the proteins Inositol 1,4,5-trisphosphate receptor type 1 (ITPR1) and Rho guanine nucleotide exchange factor 16 and 17 (ARHGEF16 and ARHGEF17) that interact with Rho-related GTP-binding protein RhoB (RHOB) to mediate apoptosis in response to DNA damage (Srougi and Burrige 2011). The second cluster of associated proteins encoded by differentially

Table B.4.: Gene Ontology analysis of genes affected by puromycin induced differential methylation of associated CpG islands.

GO term (category 'biological process')	%	Fold enrichment	<i>p</i> value	Benjamini <i>p</i> value
Di-, tri-valent inorganic cation transport	2.41	2.43	0.000	0.425
Phosphate metabolic process	8.26	1.51	0.000	0.265
Phosphorylation	6.69	1.49	0.001	0.623
Calcium ion transport	1.88	2.36	0.002	0.606
Chromatin modification	2.93	1.90	0.002	0.568
Regulation of small GTPase mediated signal transduction	2.72	1.92	0.002	0.616
Protein amino acid phosphorylation	5.65	1.51	0.003	0.616
Cell death	5.96	1.48	0.003	0.589
Death	5.96	1.46	0.004	0.612
Embryonic morphogenesis	3.03	1.76	0.004	0.652
Ephrin receptor signaling pathway	0.42	10.63	0.005	0.649
Proline biosynthetic process	0.42	10.63	0.005	0.649
Positive regulation of apoptosis	3.87	1.60	0.005	0.683
Cellular amino acid biosynthetic process	0.94	3.28	0.005	0.659
Positive regulation of programmed cell death	3.87	1.59	0.006	0.651
Positive regulation of cell death	3.87	1.58	0.007	0.657
Programmed cell death	5.02	1.46	0.008	0.709
Nitrogen compound biosynthetic process	3.03	1.66	0.009	0.721
Apoptosis	4.92	1.45	0.010	0.729

methylated genes comprised cellular key regulators such as MYC, Tumor protein p73 (TP73) and Caspase-7 (CASP7) as well as 5 additional proteins. Strikingly, all of these proteins interact with hypoxia inducible factors within the cellular response to low oxygen levels. In hypoxic conditions Hypoxia-inducible factor 1 (HIF1) inhibits MYC to induce cell-cycle arrest (Dang et al. 2008). P73 not only controls proapoptotic genes, but was shown to be involved in the response to hypoxic conditions by promoting ubiquitin-dependent degradation of Hypoxia-inducible factor 1-alpha (Amelio et al. 2015). CASP7 was shown to be induced by Hypoxia-inducible factor 2-alpha in response to UV-induced apoptosis (Turchi et al. 2008).

In the bad producer (without selection pressure) all genes encoding these proteins (except for *Proteasome activator complex subunit 3 (Psmc3)* and *Myc*) were hypomethylated compared to the XL99 host. For most of them, differential methylation was even detected at consecutive loci (e.g CpG islands 1, 2 and 3 of *Casp7*). When selection pressure by puromycin supplementation was applied, differential methylation of *Myc*, *Tp73*, *Ddx41*, *Ube4b* and *Ddit4* increased in the bad producer. For example, *Ube4b* became hypomethylated at one additional CpG island (two differentially methylated CGIs without selection pressure, three CGIs under puromycin selection) and the relative methylation of the *Tp73* CGI decreased from 0.34 to 0.02. Interestingly, differential methylation of *Myc* reversed from a mean relative methylation value of 4 to a mean value of 0.2.

None of the genes encoding the cell death related proteins comprised within the potentially hypoxia-responsive protein cluster were significantly differentially methylated in the good

producer without selection pressure. When selection pressure was applied, this clone also exhibited hypomethylation of *Myc*, *Tp73*, *Ube4b* and *Ddit4* with fold methylation changes comparable to the bad producer (e.g. a fold methylation change of 0.02 was detected for the *Tp73* CGI in both clones.)

B.4.7. Discussion: Possible maintenance of detrimental hypoxia-induced epigenetic marks through cell line generation

This experiment showed that the bad producing CHO-XL99 clone possessed several differentially methylated CpG islands associated with hypoxia-inducible, apoptosis-related genes even when kept without selection pressure and, most importantly, while maintaining high viabilities. This finding supports the hypothesis that the process of cell line generation could have an effect on DNA methylation without immediately resulting in a specific phenotype and, thereby, bearing the chance of selecting unfavorable clonal populations. In this scenario only the addition of additional cellular stress would cause the cells to switch into unwanted phenotypes.

The cluster of apoptotic genes found to be associated with differentially methylated CpG islands was closely related to the cellular reaction to hypoxic conditions. Strikingly, the dissolved oxygen concentration represents a critical parameter in production processes. Disturbances of this parameter can lead to significantly impaired growth and productivity. Therefore, oxygen transport is a major limitation in large-scale mammalian cell culture (Miller, Wilke, and Blanch 1987). As it is well accepted that environmental factors interact with the human epigenome to adjust gene expression to specific conditions, an influence of oxygen and oxidative processes on epigenetic events is, based on the current knowledge, reasonable to assume. Demethylation of DNA in mammals is linked to oxidative processes by the involvement of the ten eleven translocation (TET) dioxygenases, which convert 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) (Tahiliani et al. 2009). Furthermore, histone lysine demethylation is mediated by oxygen-dependent enzymes, too. Lysine Specific Demethylase 1 (LSD1) catalyzes oxidation of the methylamine group to the imine state, thereby allowing hydrolysis and release of the former methyl carbon atom as formaldehyde (Shi et al. 2004; Jeltsch 2013). Another group of protein lysine demethylases harbors a Jumonji C (JmjC) domain and uses, similar to TET enzymes, alpha-ketoglutarate as a co-substrate in an oxidation reaction to remove all methylation states (Donohoe and Bultman 2012; Tsukada et al. 2006; Hou and Yu 2010). The removal of epigenetic marks therefore requires molecular oxygen as a highly reactive reaction partner and the oxygenases involved in this reactions need to be optimized for a certain oxygen pressure (Jeltsch 2013). Thus, setting up reversible epigenetic patterns might depend on proper environmental oxygen levels. Adding to this hypothesis is recent research that shows a major effect of hypoxia on transcription, maturation

and function of a specific group of miRNAs (termed “hypoxamirs”™; Nallamshetty, Chan, and Loscalzo 2013). Furthermore, hypoxia inducible factors (HIFs) are known to regulate energy metabolism in hypoxic conditions by triggering a switch from oxidative phosphorylation to anaerobic glycolysis and inhibit the conversion of pyruvate to acetyl-coA, a process that has already been described in the context of the butyrate experiments in this thesis and that is linked to epigenetic processes on several levels (Koh and Powis 2012; Goda and Kanai 2012).

However, it has to be kept in mind that this experiment was performed using a polyclonal cell population and differential DNA methylation as detected by the CHO-specific CpG island microarray is therefore difficult to assess, especially regarding the variability of DNA methylation that was reported in the main part of this thesis.

B.5. R code

Listing B.1: R script for import of CHO CpG island microarray data into *limma* and generation of quality plots.

```
# Installation
source("http://www.bioconductor.org/biocLite.R")
biocLite()
biocLite("limma")
biocLite("statmod")

# Set working directory
setwd("C:\\Users\\...")

# Load package
library(limma)

# Import target txt-file (Columns: 'FileName', 'Cy3', 'Cy5',
'Name' (of sample))
targets <- readTargets("Targets.txt", row.names="Name")

# Import microarray data files
RG <- read.maimages(targets, source="agilent", columns = list(G =
"gMedianSignal", Gb = "gBGMedianSignal", R = "rMedianSignal", Rb =
"rBGMedianSignal"), annotation = c("Row", "Col", "FeatureNum",
"ControlType", "ProbeName", "genes"), names = targets$Name)

# Generation of boxplots that are saved as PDF files in working
directory
pdf("GBackground.pdf")
boxplot(data.frame(log2(RG$Gb)), main="Green background")

pdf("RBackground.pdf")
boxplot(data.frame(log2(RG$Rb)), main="Red background")

pdf("GSignal.pdf")
boxplot(data.frame(log2(RG$G)), main="Green signal")

pdf("RSignal.pdf")
boxplot(data.frame(log2(RG$R)), main="Red signal")
```

```
# Generation of intensity plots for each array
for (i in 1:8){
  name.Gb <- paste(i,"imageplotGb.pdf", sep = "")
  pdf(name.Gb)
  imageplot(log2(RG$Gb[,i]),RG$printer, main =
  paste("Imageplot Gb Array", i))

  name.G <- paste(i,"imageplotG.pdf", sep = "")
  pdf(name.G)
  imageplot(log2(RG$G[,i]),RG$printer, main =
  paste("Imageplot G Array", i))

  name.Rb <- paste(i,"imageplotRb.pdf", sep = "")
  pdf(name.Rb)
  imageplot(log2(RG$Rb[,i]),RG$printer, main =
  paste("Imageplot Gb Array", i))

  name.R <- paste(i,"imageplotR.pdf", sep = "")
  pdf(name.R)
  imageplot(log2(RG$R[,i]),RG$printer, main =
  paste("Imageplot R Array", i))
}

# Generation of one density plot for all arrays
MA <- normalizeWithinArrays(RG, method="none", bc.method="none")
pdf("PlotDensitiesMA.pdf")
plotDensities(MA)
```

Listing B.2: R script for import of CHO gene expression microarray data into *limma* as well as normalization and calculation of statistics for the comparison of data sets of interest.

```
# Installation
source("http://www.bioconductor.org/biocLite.R")
biocLite()
biocLite("limma")
biocLite("statmod")

# Set working directory
setwd("C:\\Users\\...")

# Load package
library(limma)
```

```
# Import target txt-file (Columns: 'SampleNumber', 'FileName',  
'Condition')  
targets <- readTargets("targets.txt")  
x <- read.maimages(targets, source="agilent", green.only=TRUE)  
  
# Normalize within and between arrays  
y <- backgroundCorrect(x, method="normexp", offset=16)  
y <- normalizeBetweenArrays(y, method="quantile")  
y.ave <- avereps(y, ID=y$genes$ProbeName)  
  
# Build a design matrix, perform linear modeling  
f <- factor(targets$Condition, levels = unique(targets$Condition))  
design <- model.matrix(~0 + f)  
colnames(design) <- levels(f)  
fit <- lmFit(y.ave, design)  
  
# Compare data sets of interest, e.g. data of timepoints '0h' and '6h'  
contrast.matrix <- makeContrasts("zero-six", levels=design)  
fit2 <- contrasts.fit(fit, contrast.matrix)  
fit2 <- eBayes(fit2)  
  
# Generate txt file containing statistics  
topTable(fit2, adjust="BH", coef="zero-six", genelist=y.ave$genes)
```

B.6. Supplementary figures

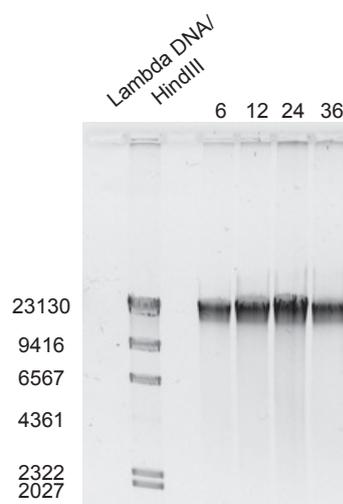


Figure B.13.: Agarose gel for integrity control of genomic DNA extracted from reference and butyrate treated CHO DP-12 cells for whole-genome bisulfite sequencing.

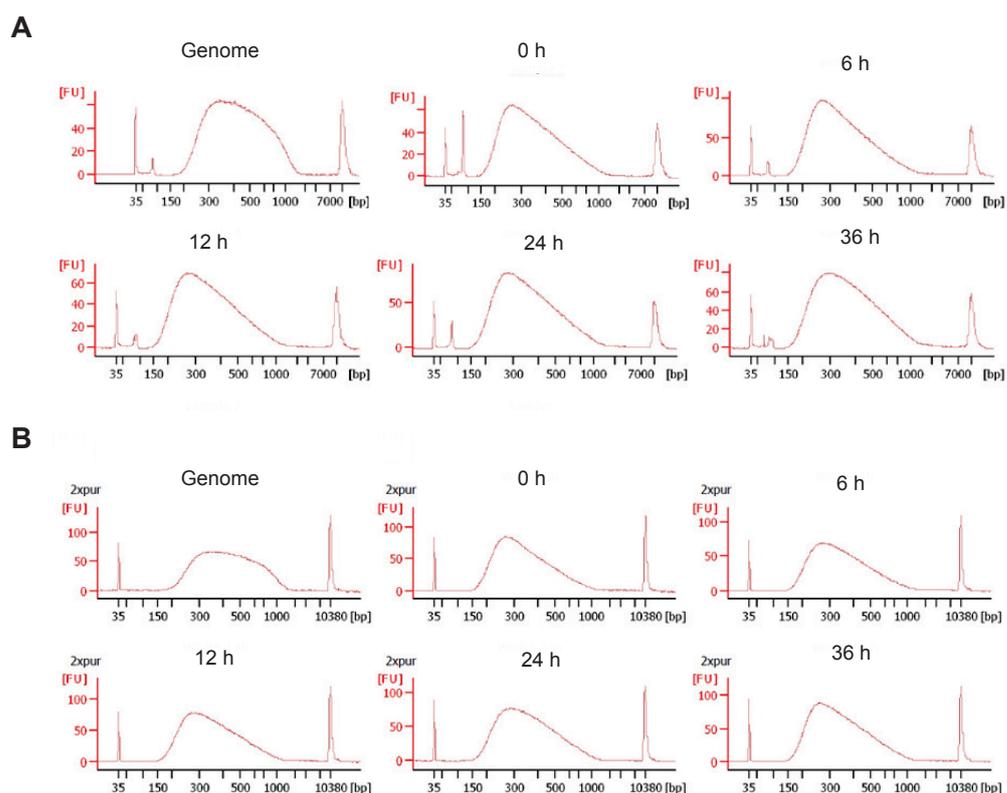


Figure B.14.: Quality of libraries for whole genome bisulfite sequencing was checked using *High Sensitivity DNA Chips* on a *Bioanalyzer*. At first (A), all samples showed an additional peak at about 85 bp that disappeared after a second cleanup procedure with *AMPURE* beads (B).

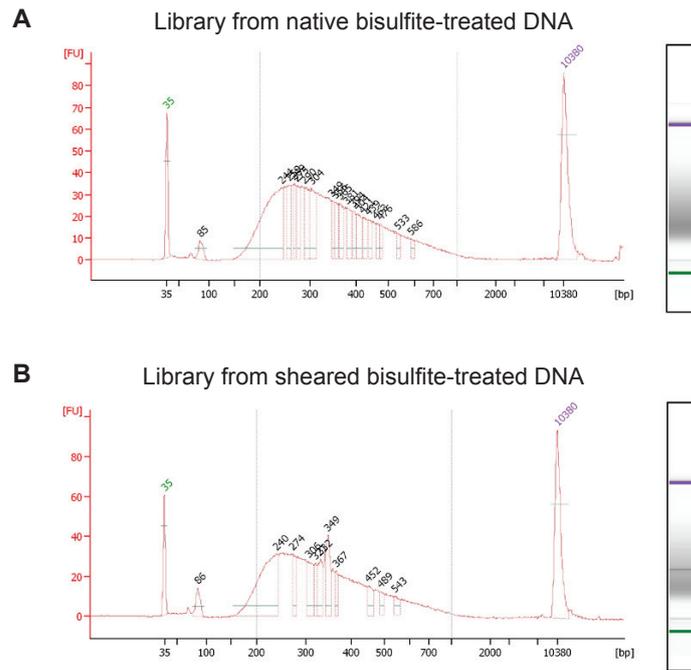


Figure B.15.: Evaluation of an effect of shearing on library generation from bisulfite treated DNA. **A** Library generated from native bisulfite-treated DNA. **B** Library generated from sheared bisulfite-treated DNA.

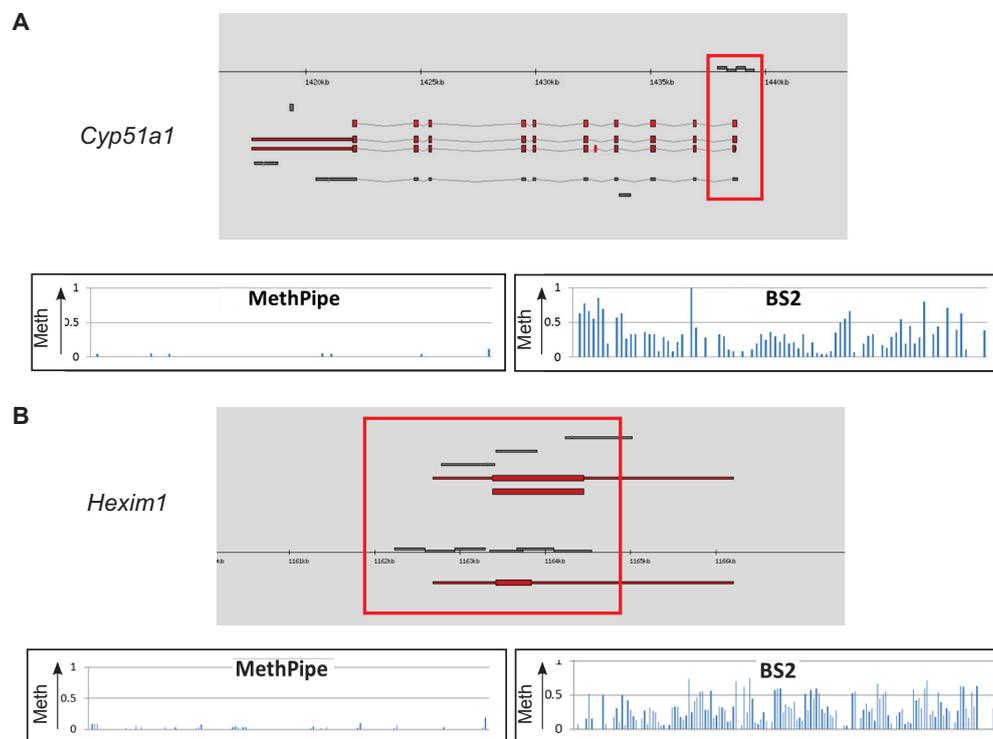


Figure B.16.: Comparison of mapping results for exemplary CpG islands obtained with the *MethPipe* mapping tool (Song et al. 2013) and *BS-Seeker2* (Guo et al. 2013). Red boxes mark the CpG island regions that were analyzed regarding their methylation level. **A** CpG island of *Lanosterol 14-alpha demethylase (Cyp51a1)* gene. **B** CpG island of *Protein HEXIM1 (Hexim1)* gene.

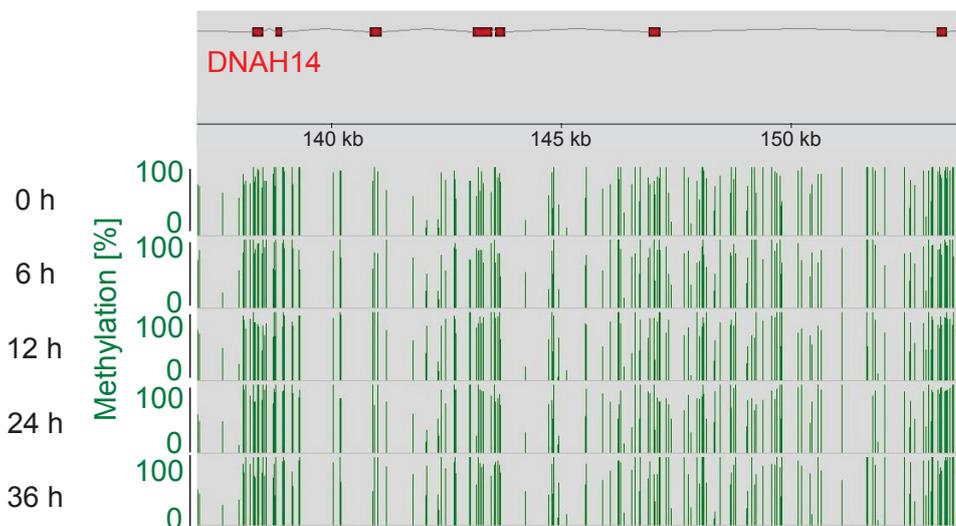


Figure B.17.: Methylation data tracks were loaded into the *GenDBE* genome browser (Rupp et al. 2014) to check for consistency of whole-genome bisulfite sequencing results between the sampling points. Exons of the exemplary *Dnah14* gene are indicated by red rectangles. Methylation levels for each CpG are represented by green bars.

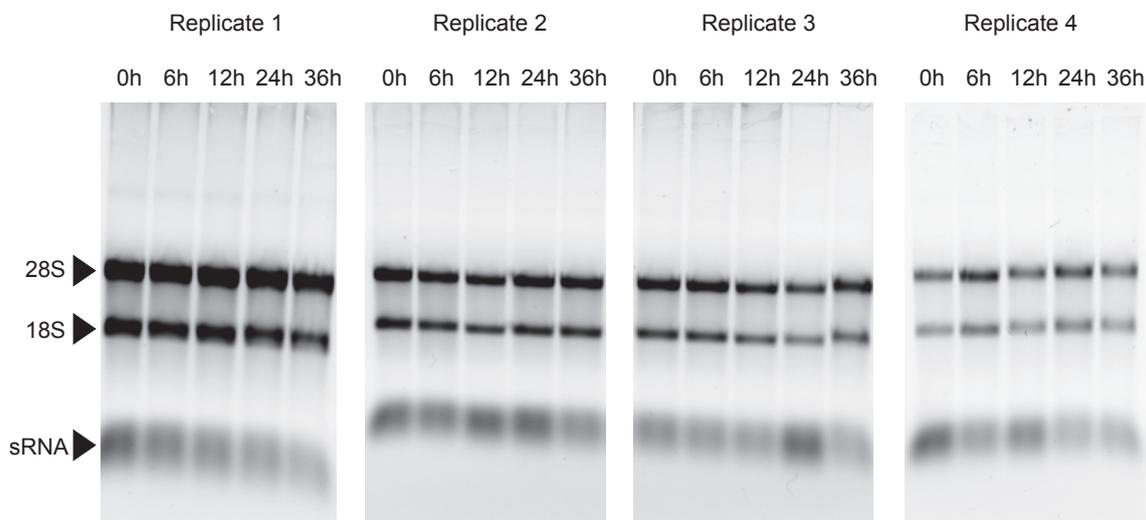


Figure B.18.: Denaturing agarose gels for control of integrity of RNAs used for gene expression microarray experiments. RNA extraction was performed separately for replicate cultures at time points 0h (reference) and 6h to 36h post butyrate addition. Indicated are the 28S and 18S rRNA bands and also the position of RNAs smaller than 200 bp ('sRNA') that are additionally captured by the preparation columns that were used.

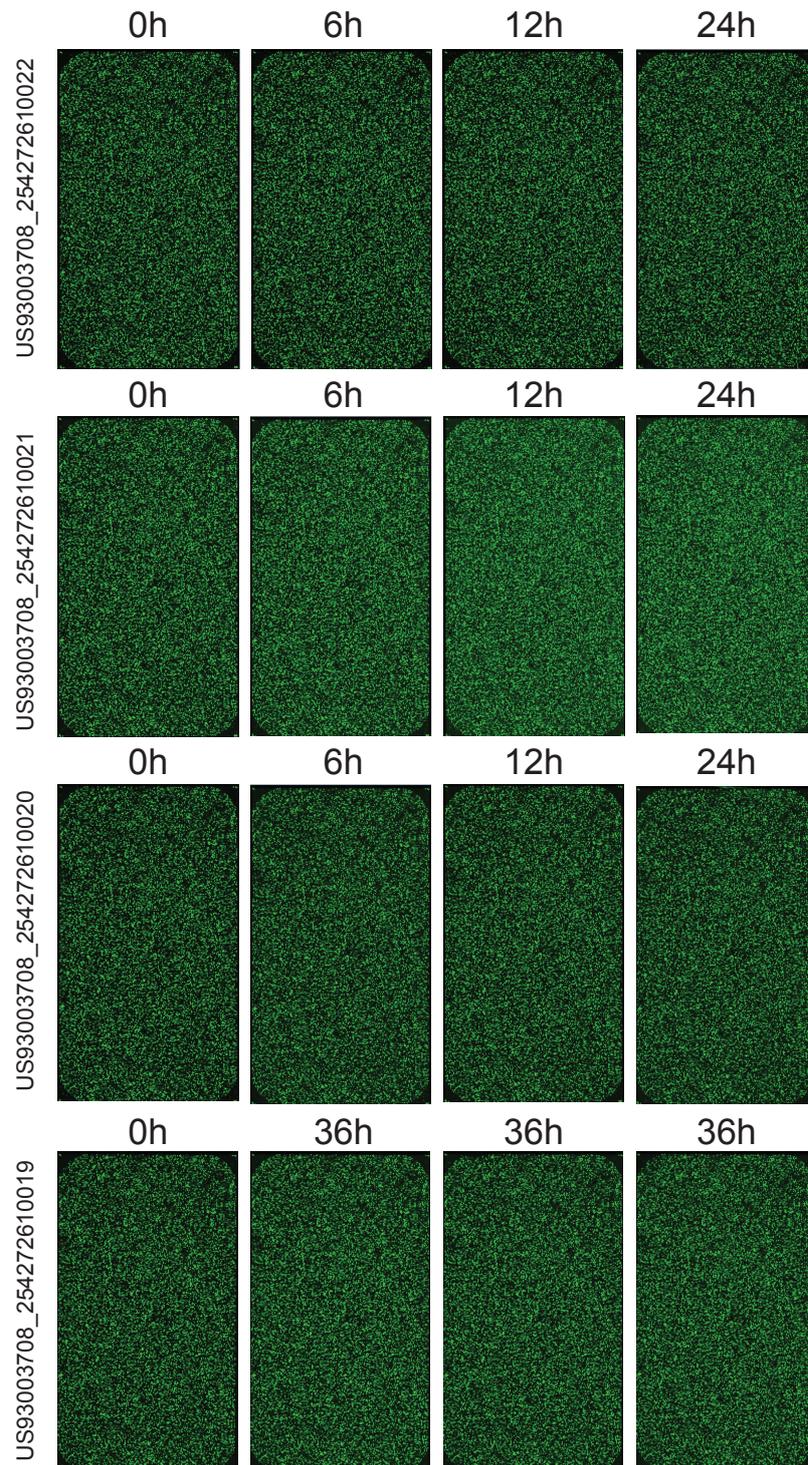


Figure B.19.: Artificial images of spatial signal distributions of gene expression microarrays. Images were generated by the *Feature Extraction* software. Sampling points of hybridized RNAs are indicated as well as internal slide numbers.

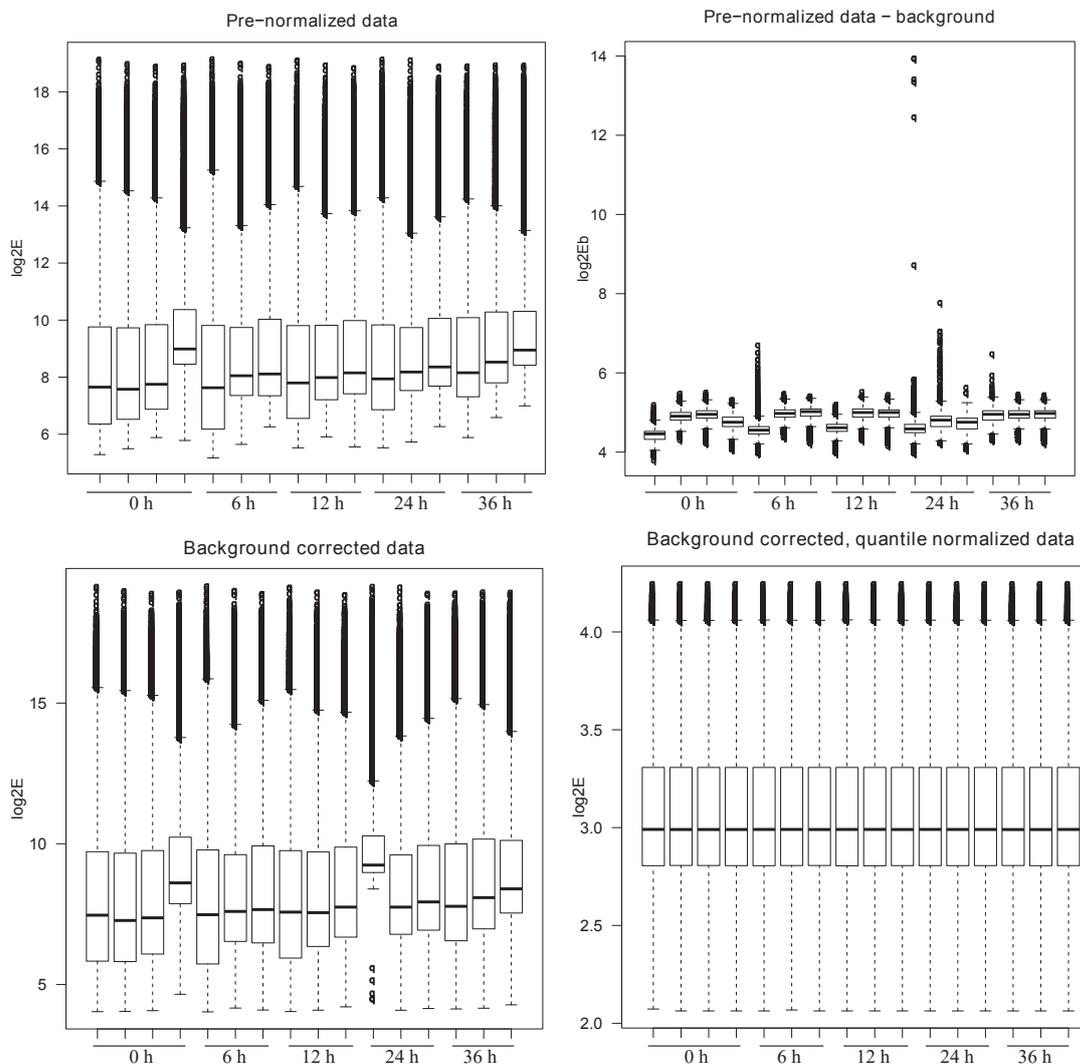


Figure B.20.: Boxplots of signal intensities for gene expression microarrays before and after background correction and quantile normalization for 4 reference ('0 h') and each 3 microarrays for time points 6, 12, 24 and 36 hours. Signal intensities ($\log_2 E$) were corrected for the background signals ($\log_2 E_b$) and normalized (see *R* script in B.2).

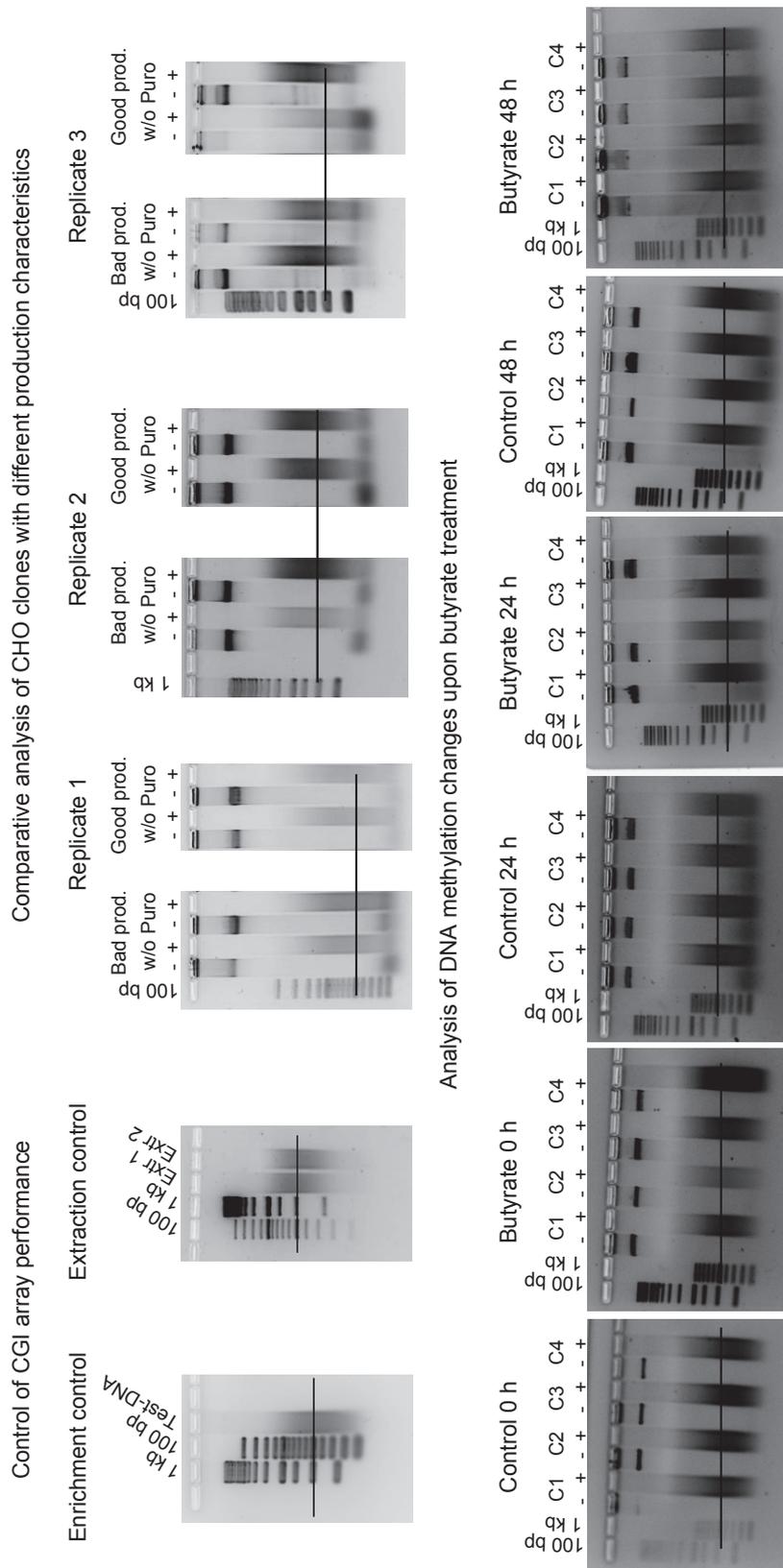


Figure B.21.: Agarose gels for the control of sheared DNA samples for CpG island microarray analysis. Per lane, 300 ng of DNA were separated to control if the samples exhibited the desired fragment size of about 500 bp.

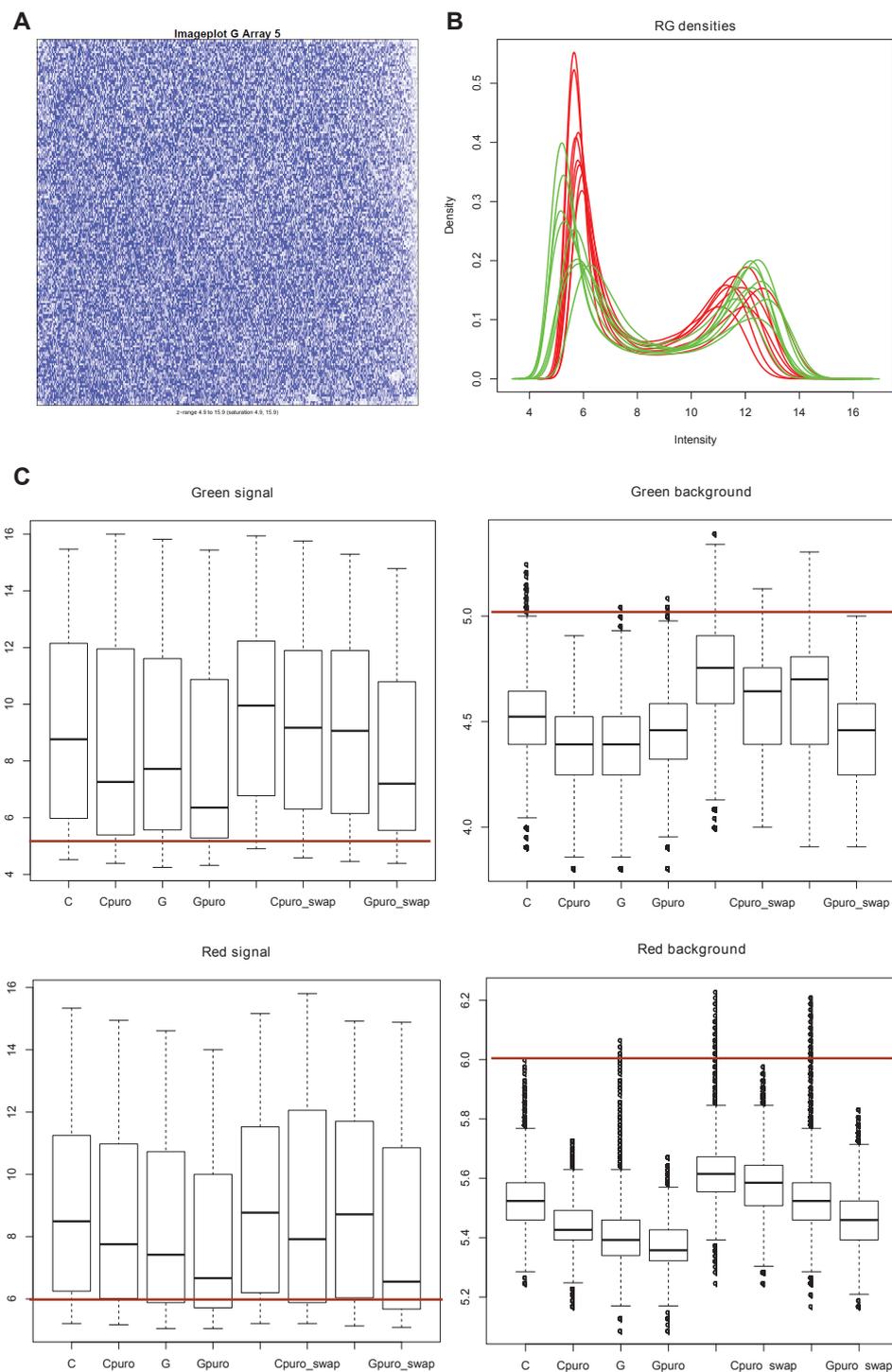


Figure B.22.: Exemplary display of quality plots from CHO CpG island microarray data. **A** Imageplots were generated to control the spatial distribution of signals on the slide. **B** Densities of red and green signals were plotted and expected to show a generally bimodal distribution, which corresponds to either unmethylated or methylated CpG islands. **C** Red and green signals and background signals were controlled in order to exclude main signals that were too low or background signals that were too high to allow for a discrimination between foreground and background.

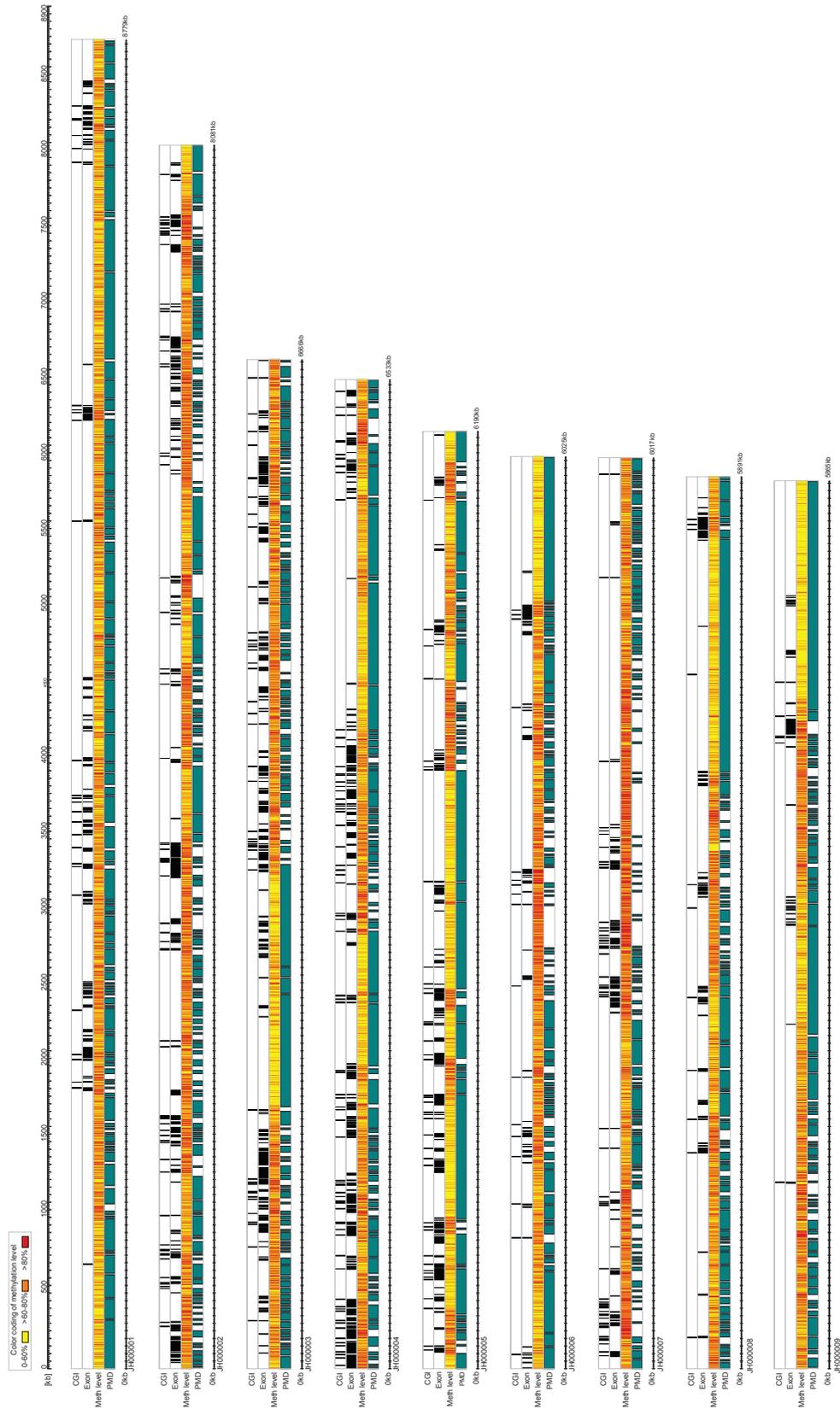


Figure B.23.: Display of SNP-corrected CHO K1 scaffolds JH000001 to JH000009 (Wippermann et al. 2015). CHO DP-12 DNA methylation levels for 1 kb windows were plotted along the scaffolds and color-coded according to the legend. Also shown are exons and CpG islands (CGI) in black. Regions identified as CHO DP-12 PMDs are marked by green bars.

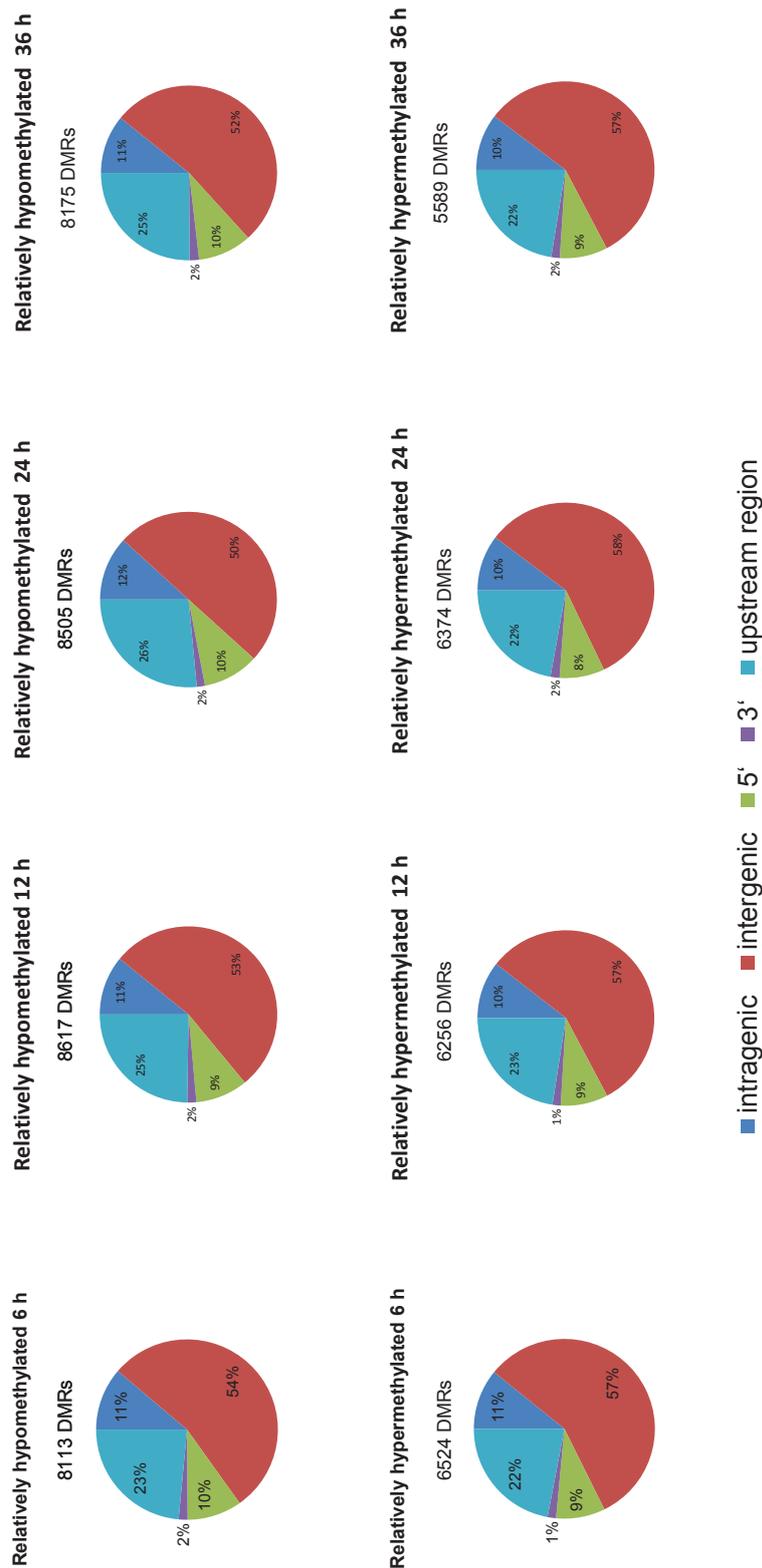


Figure B.24.: Percentages of relatively hypomethylated and relatively hypermethylated DMRs with ≥ 1 significantly differentially methylated CpG 6, 12, 24 and 36 hours after butyrate addition, which were assigned to the genomic features upstream region, 5' and 3' end, intergenic and intragenic region (Wippermann et al. 2016).

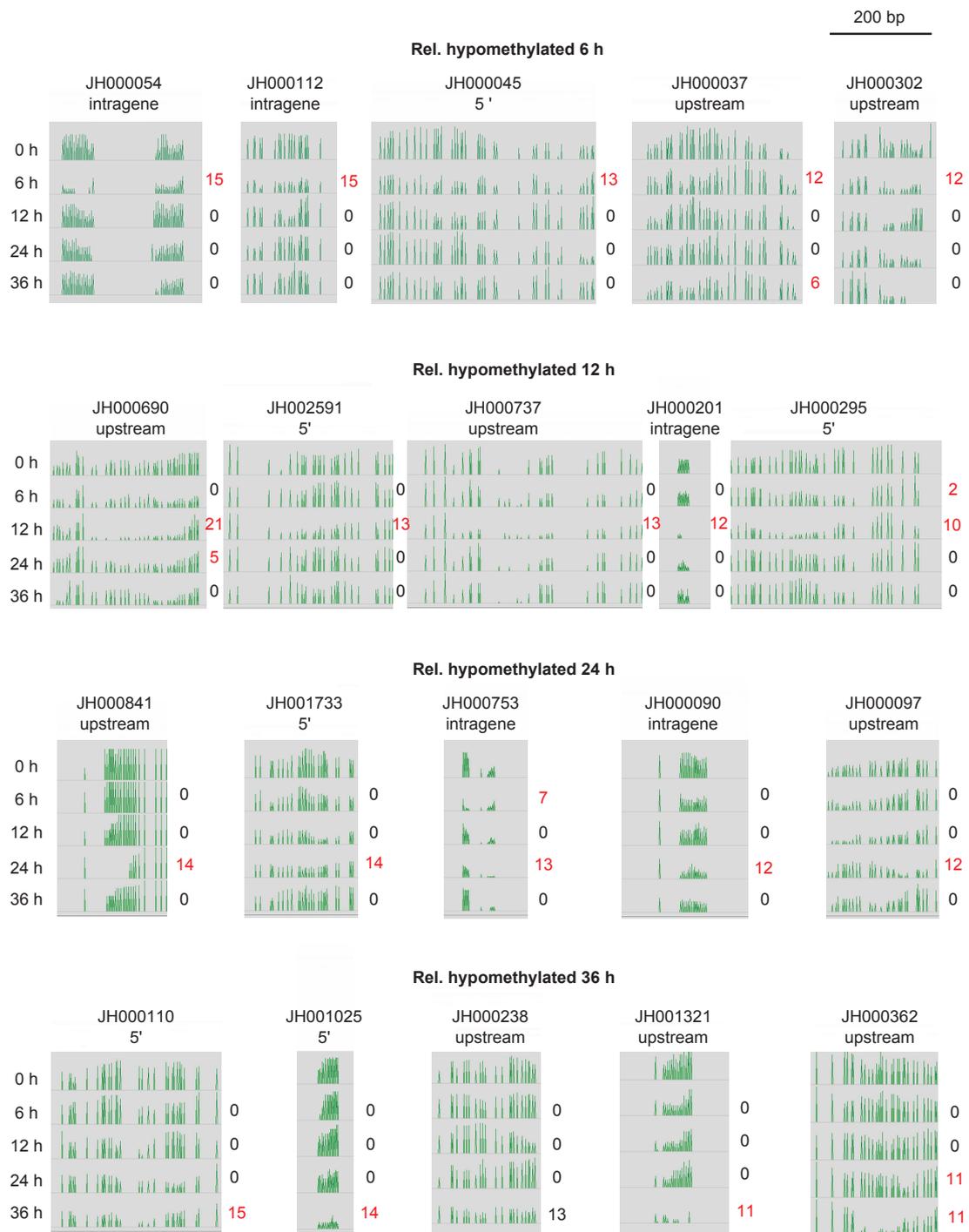


Figure B.25.: Screenshots of relatively hypomethylated DMRs as visualized in the *GenDBE* genome browser (Rupp et al. 2014). All regions are represented in the same scale.

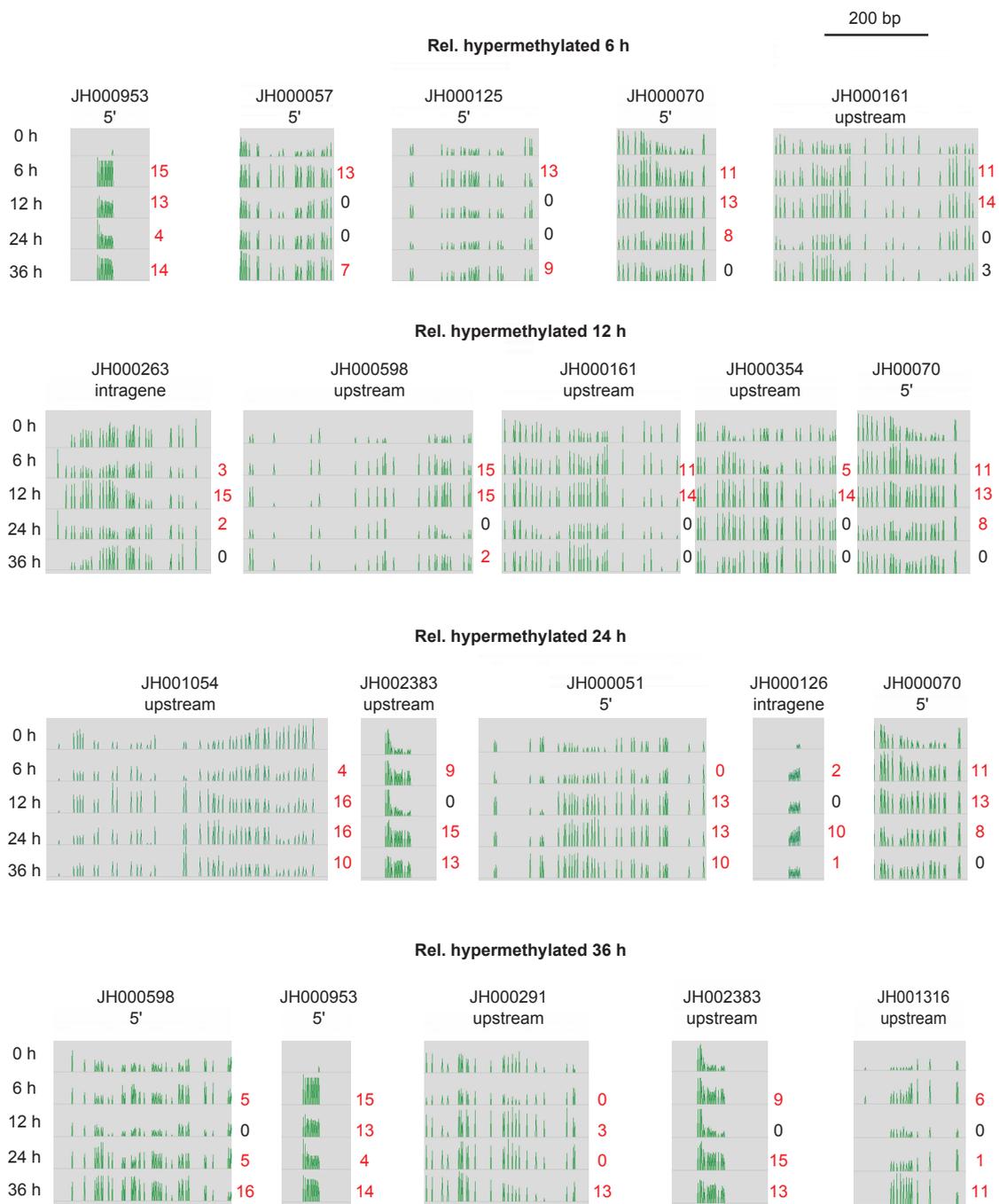


Figure B.26.: Screenshots of relatively hypermethylated DMRs as visualized in the *GenDBE* genome browser (Rupp et al. 2014). All regions are represented in the same scale.

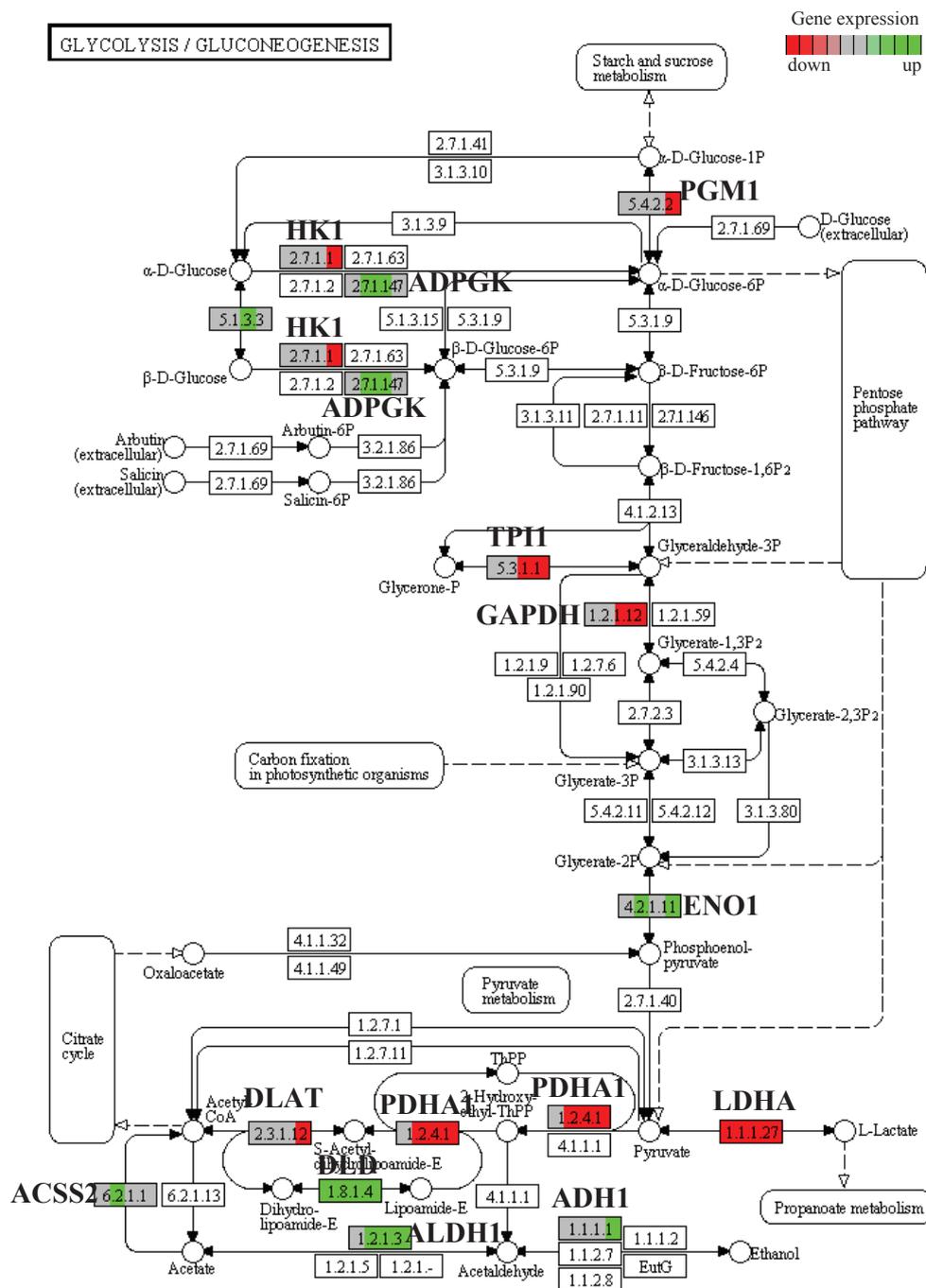


Figure B.27.: Visualisation of gene expression time course after butyrate addition for *KEGG* pathway Glycolysis (Kanehisa and Goto 2000). Pathway analysis and visualization of gene expression changes based on M-values was performed using *pathview* (Luo and Brouwer 2013). Gene expression changes are color-coded according to the legend.

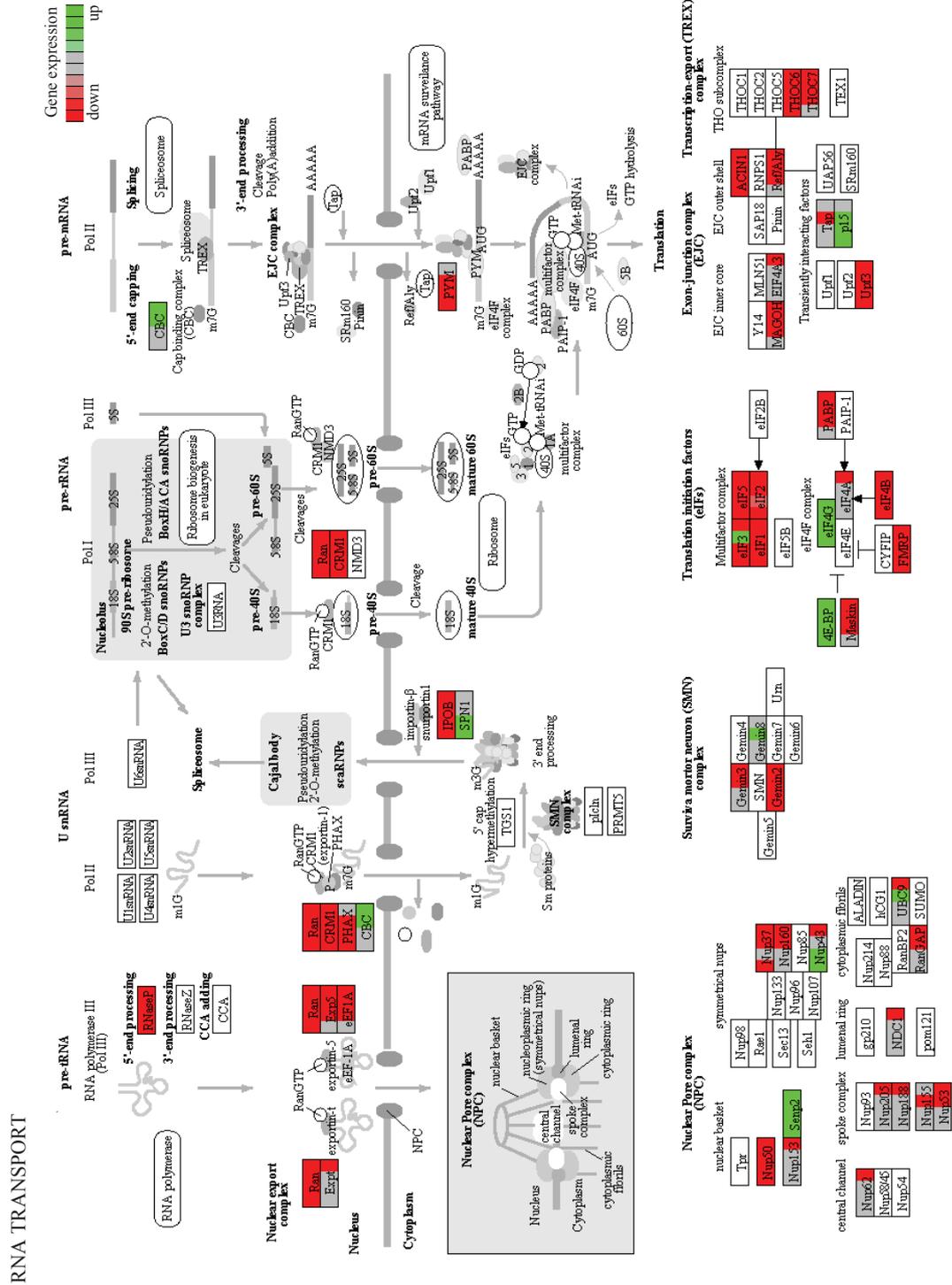


Figure B.28.: Visualisation of gene expression time course after butyrate addition for KEGG pathway RNA Transport (Kanehisa and Goto 2000). Pathway analysis and visualization of gene expression changes based on M-values was performed using *pathview* (Luo and Brouwer 2013). Gene expression changes are color-coded according to the legend.

RIBOSOME

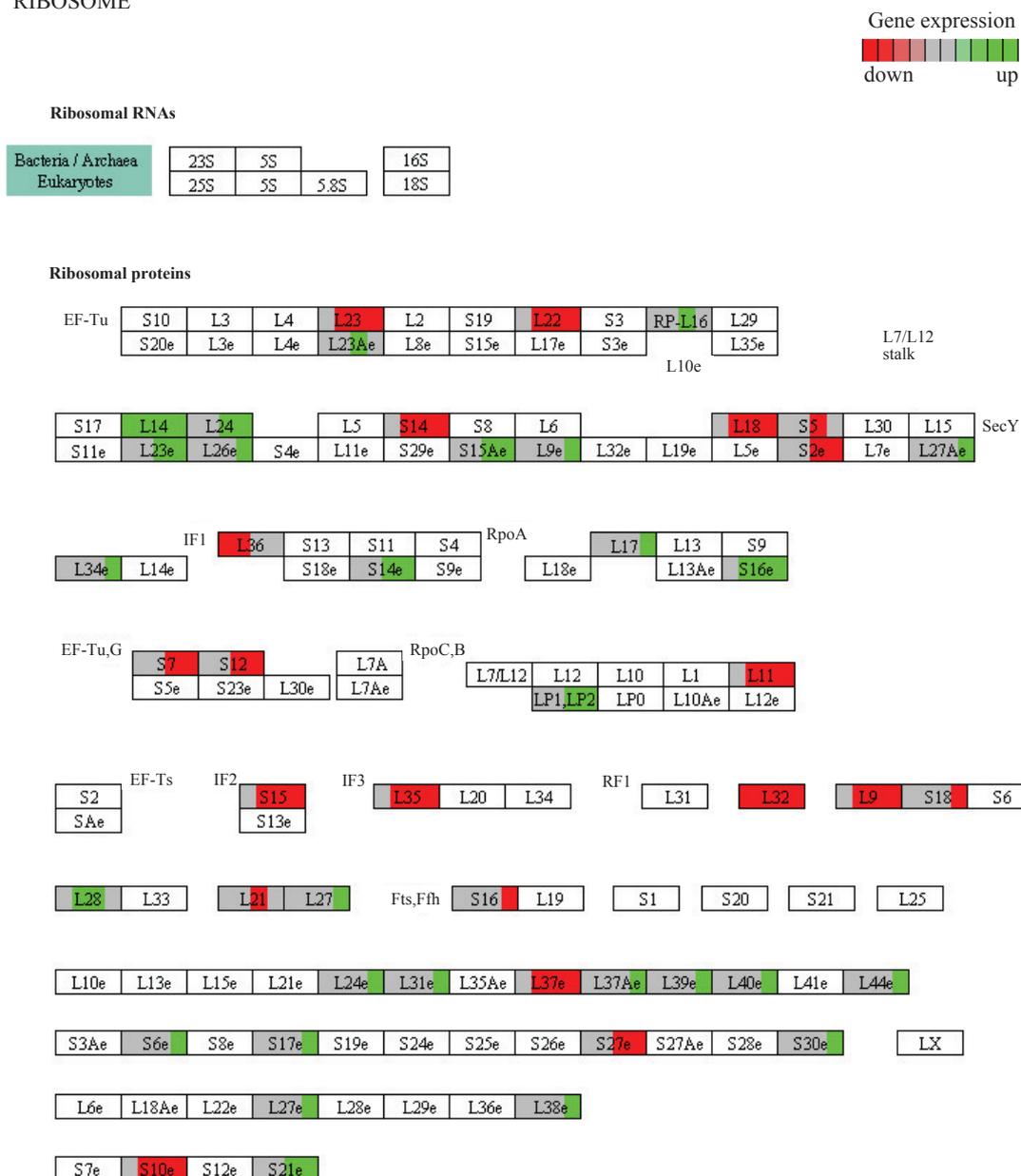


Figure B.29.: Visualisation of gene expression time course after butyrate addition for *KEGG* pathway Ribosome (Kanehisa and Goto 2000). Pathway analysis and visualization of gene expression changes based on M-values was performed using *pathview* (Luo and Brouwer 2013). Gene expression changes are color-coded according to the legend.

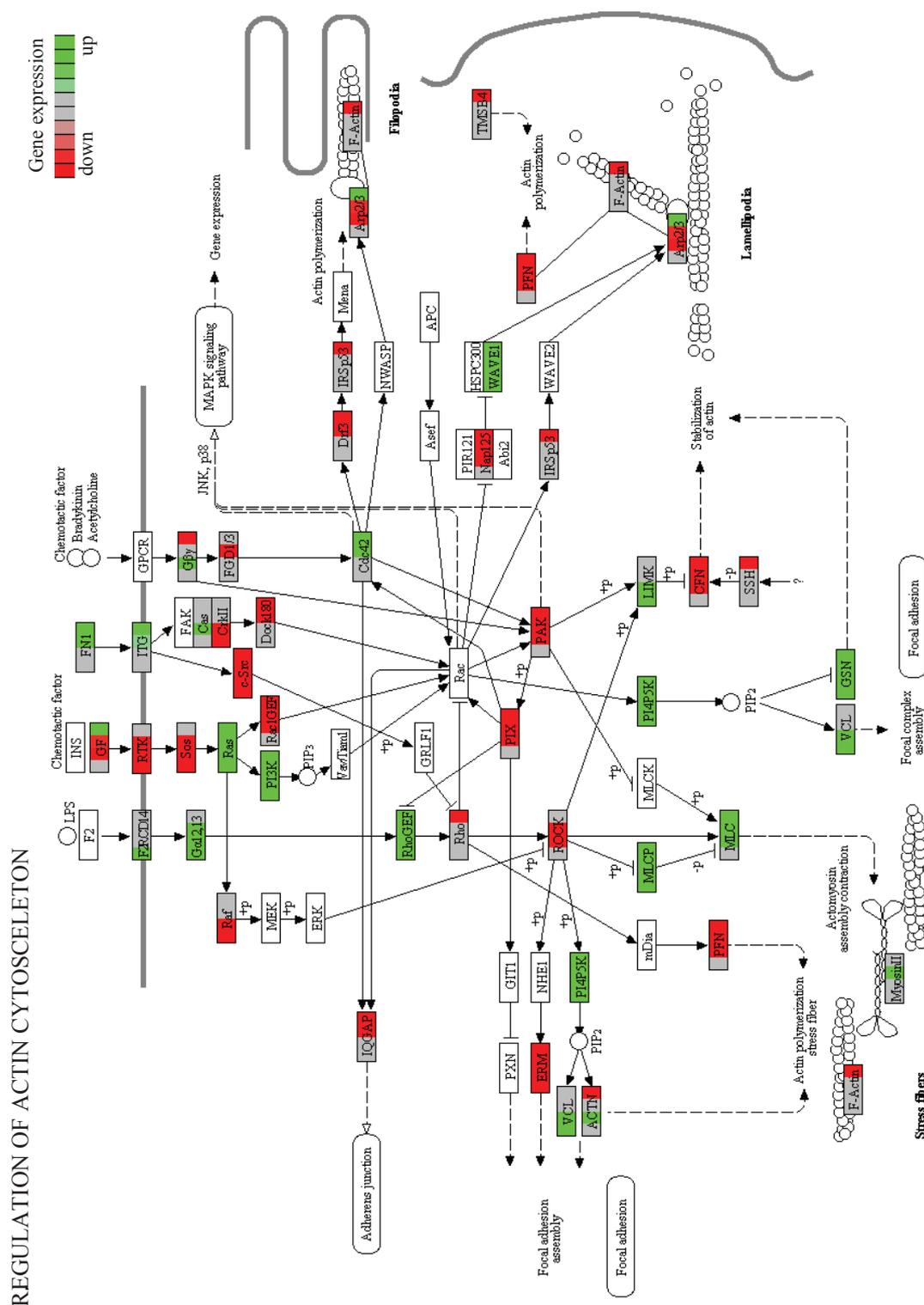


Figure B.30.: Visualisation of gene expression time course after butyrate addition for *KEGG* pathway Regulation of Actin Cytoskeleton (Kanehisa and Goto 2000). Pathway analysis and visualization of gene expression changes based on M-values was performed using *pathview* (Luo and Brouwer 2013). Gene expression changes are color-coded according to the legend.

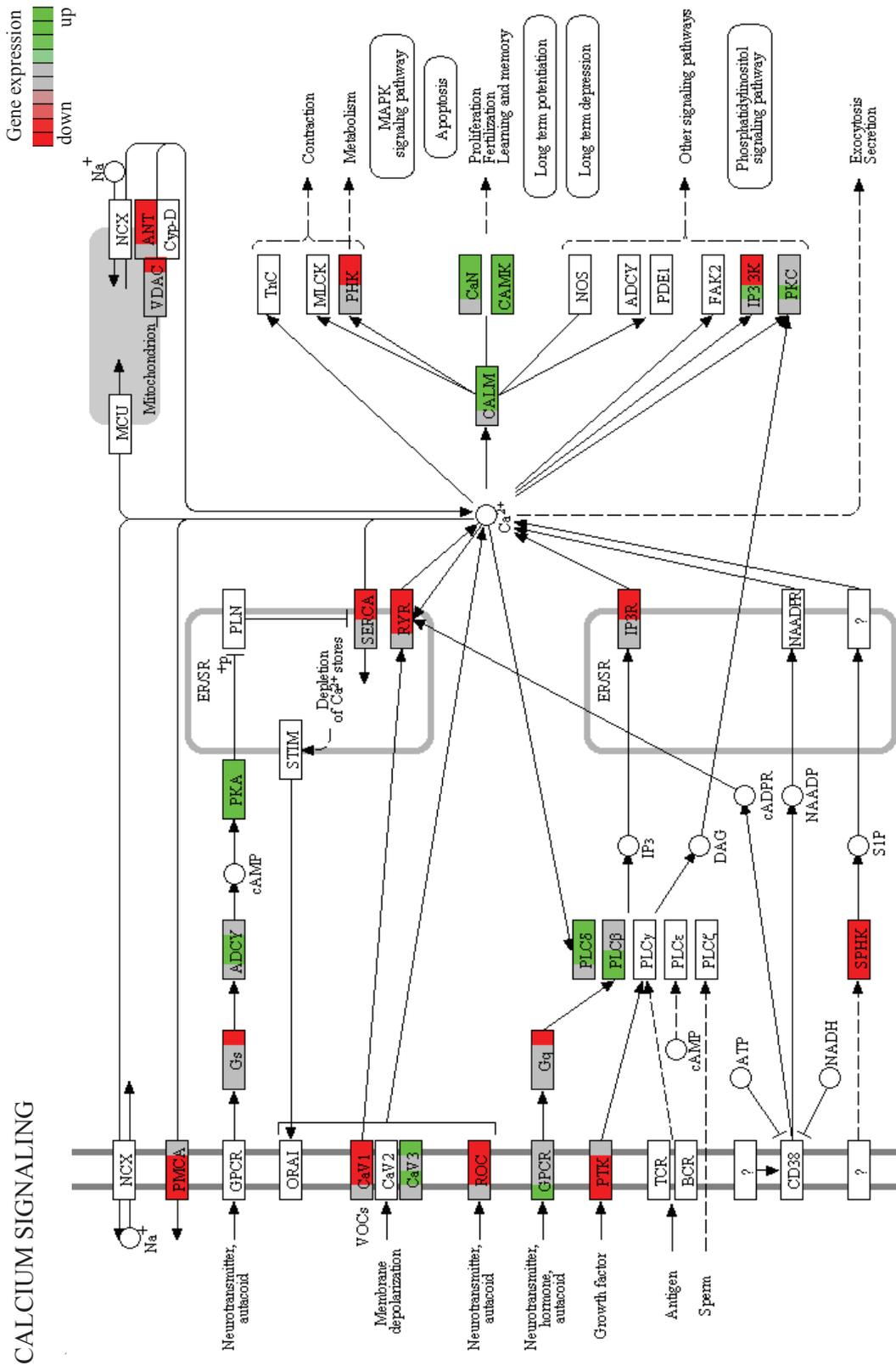
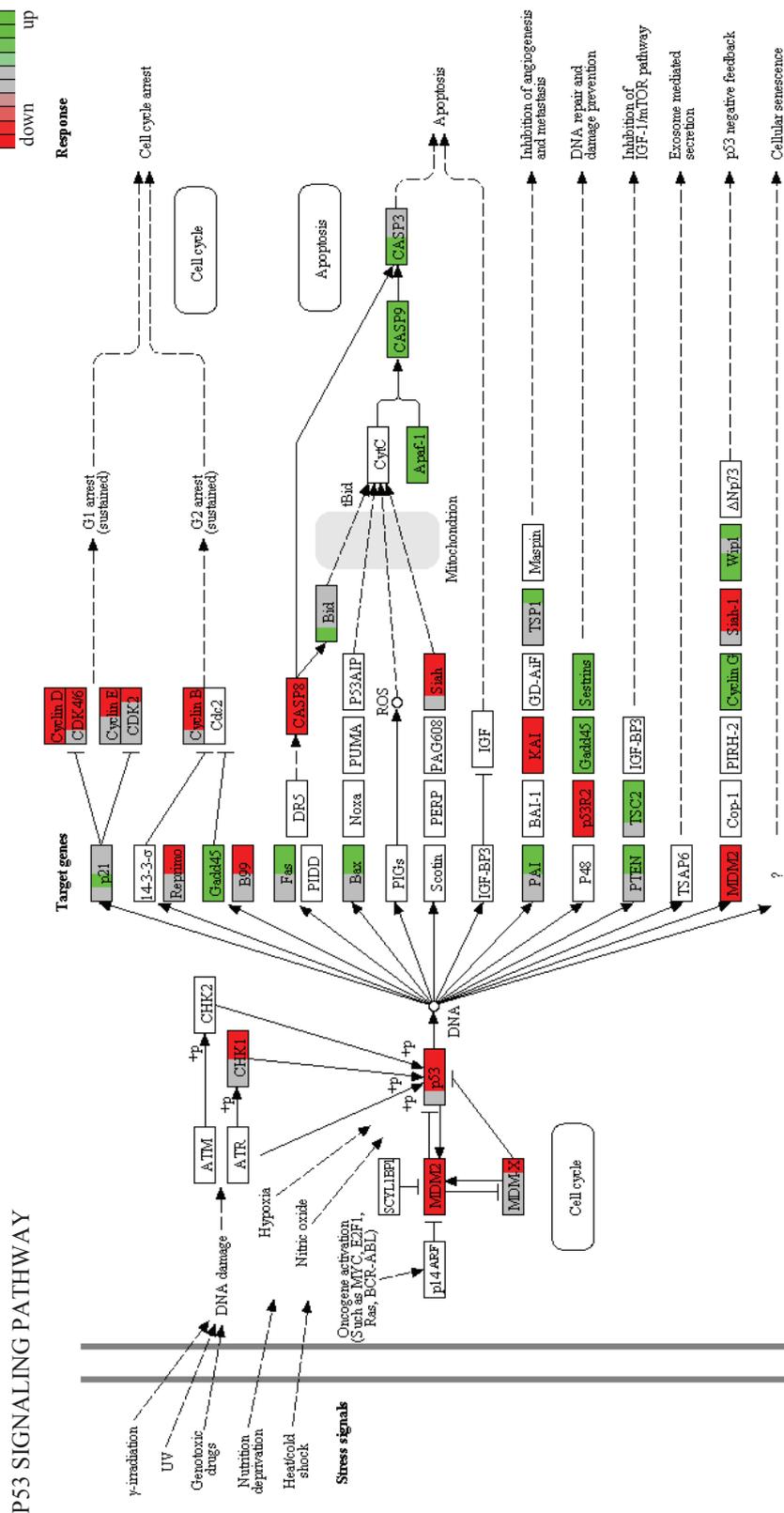


Figure B.33.: Visualisation of gene expression time course after butyrate addition for *KEGG* pathway Regulation of Calcium Signaling (Kanehisa and Goto 2000). Pathway analysis and visualization of gene expression changes based on M-values was performed using *pathview* (Luo and Brouwer 2013). Gene expression changes are color-coded according to the legend.



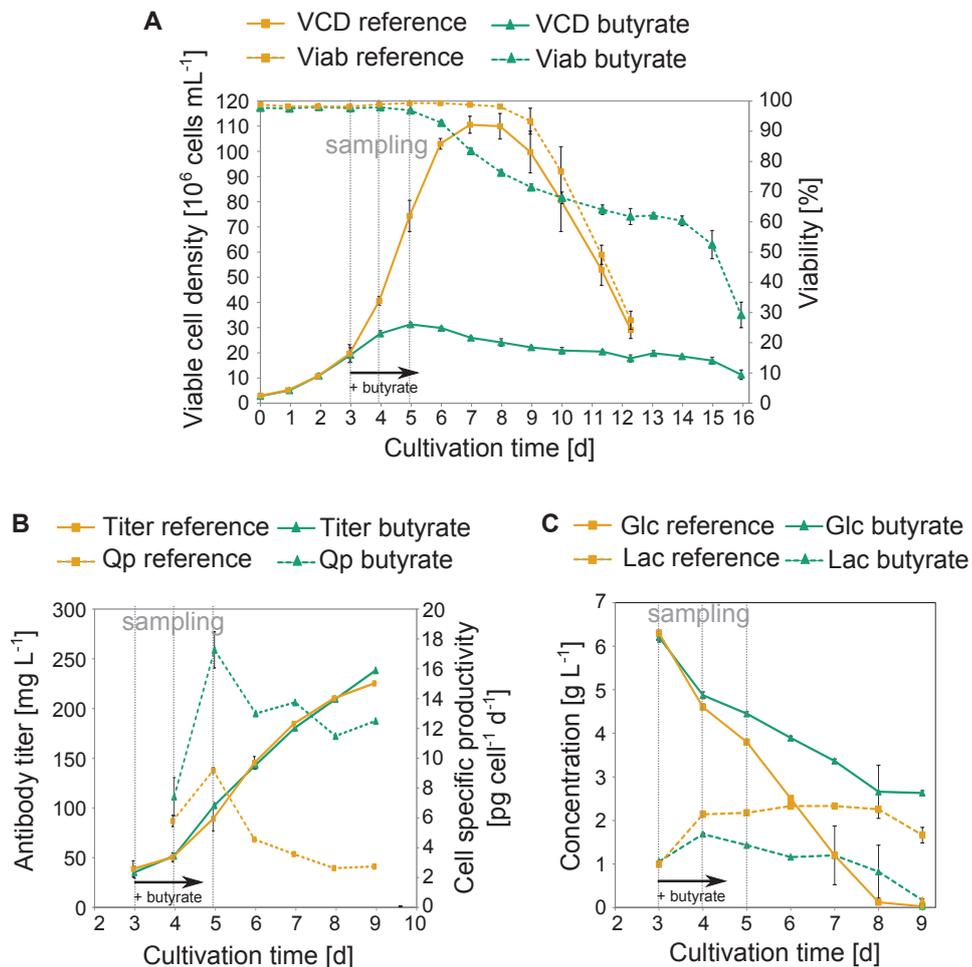


Figure B.37.: Cell growth and IgG production as well as glucose and lactate concentrations of butyrate treated and control CHO DP-12 cell cultures (Wippermann 2012, Wippermann et al. 2013). Shown are the mean values of each four replicate cultures, error bars represent standard deviations. **A** Mean viable cell density (continuous lines) and viability (dashed lines). Addition of 3 mM butyrate is marked by a red vertical line, sampling points for microarray analyses are marked by grey arrows. **B** Mean IgG titer (continuous lines) and cell specific productivity (dashed lines) of butyrate and reference cultures. **C** Glucose (Glc) and lactate (Lac) concentrations.

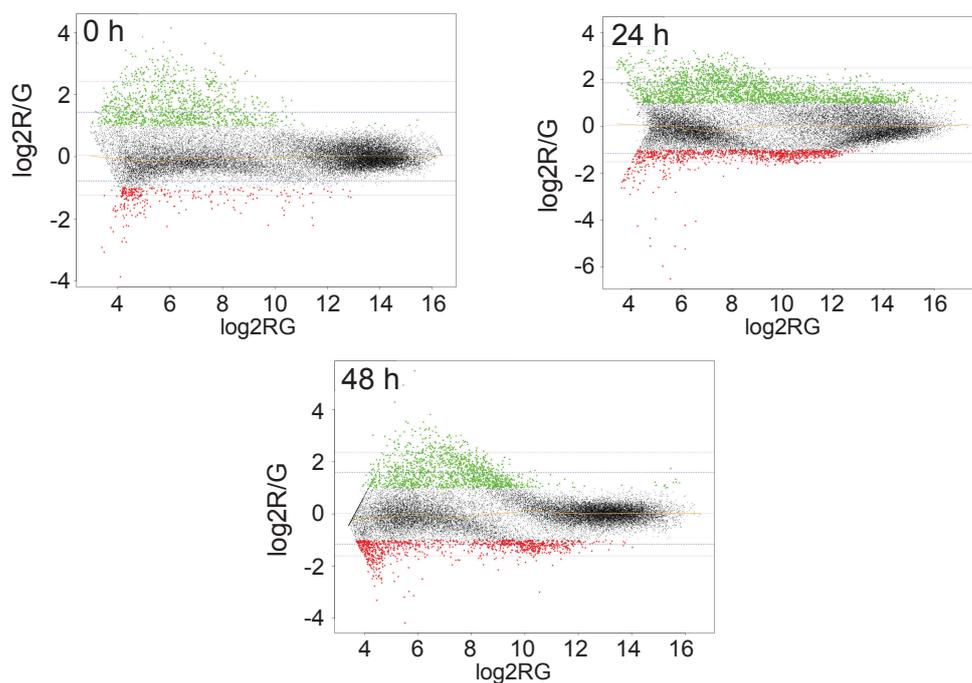


Figure B.38.: MA plots for the indicated sampling points (Wippermann 2012, Wippermann et al. 2013). M-values below the threshold of -1.0 (red) represent *de novo* methylated regions upon butyrate treatment. M-values above the threshold of 1.0 (green) represent hypomethylated regions upon butyrate treatment. The blue lines mark for each time point the 95 % and 99 % confidence intervals of all observed M-values. The yellow line represents the LOWESS fit.

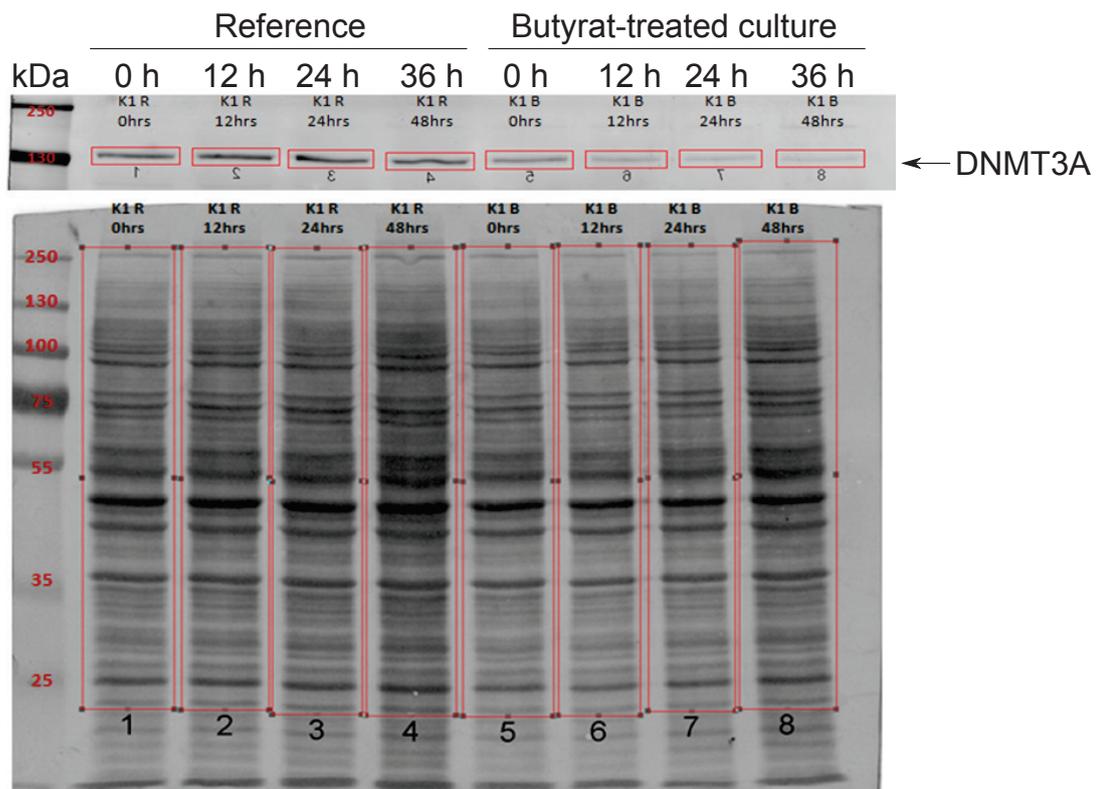


Figure B.39.: Exemplary westernblot for measurement of DNMT3A expression in reference and butyrate treated CHO DP-12 cell cultures. Expression strength was normalized to whole-protein abundance as determined by coomassie staining of the membranes. A *PageRuler™ Prestained Protein Ladder* was used as a size standard.

B.7. Supplementary tables

Table B.5.: Complete GO annotation for highly expressed CHO DP-12 PMD genes (1261) with Benjamini p values ≤ 0.05 (Wippermann et al. 2015).

GO term (category 'biological process')	%	Fold enrichment	p value	Benjamini p value
Translation	11.7	5.78	4.54E-76	1.41E-72
Translational elongation	6.1	9.92	7.25E-64	1.13E-60
RNA processing	9.3	2.78	6.47E-25	6.72E-22
Ribosome biogenesis	4	5.33	1.14E-23	8.88E-21
Ribonucleoprotein complex biogenesis	4.8	4.34	6.73E-23	4.19E-20
Nuclear mRNA splicing, via spliceosome	4	4.25	9.30E-19	4.83E-16
RNA splicing, via transesterification reactions	4	4.25	9.30E-19	4.83E-16
RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	4	4.25	9.30E-19	4.83E-16
rRNA processing	2.9	5.23	2.61E-17	1.16E-14
RNA splicing	5.4	3.11	2.84E-17	1.11E-14
rRNA metabolic process	2.9	5.01	1.31E-16	3.84E-14
mRNA processing	5.6	2.84	2.09E-15	6.57E-13
Oxidative phosphorylation	2.9	4.78	2.32E-15	6.60E-13
ncRNA metabolic process	4.4	3.17	1.47E-14	3.83E-12
Generation of precursor metabolites and energy	5.2	2.74	7.95E-14	1.91E-11
mRNA metabolic process	5.8	2.57	1.12E-13	2.50E-11
ncRNA processing	3.7	3.27	7.37E-13	1.53E-10
Energy coupled proton transport, down electrochemical gradient	1.4	5.85	1.61E-09	3.14E-07
ATP synthesis coupled proton transport	1.4	5.85	1.61E-09	3.14E-07
Ribosomal small subunit biogenesis	0.8	11.83	4.30E-09	7.88E-07
Mitotic cell cycle	4.9	2.18	8.82E-09	1.53E-06
Ribosomal large subunit biogenesis	0.7	11.71	4.62E-08	7.57E-06
Cellular respiration	2.1	3.49	4.84E-08	7.55E-06
Ion transmembrane transport	1.4	4.78	6.09E-08	9.04E-06
Negative regulation of protein modification process	2.3	3.17	6.51E-08	9.22E-06
Translational initiation	1.3	4.91	1.00E-07	1.36E-05
Mitochondrial electron transport, NADH to ubiquinone	1.3	4.96	2.36E-07	3.06E-05
Intracellular transport	7	1.74	2.92E-07	3.63E-05
Electron transport chain	2.1	3.08	3.68E-07	4.41E-05
Proton transport	1.5	4.05	4.10E-07	4.74E-05
Cell cycle process	6.2	1.8	4.69E-07	5.22E-05
Positive regulation of ubiquitin-protein ligase activity during mitotic cell cycle	1.6	3.83	5.05E-07	5.42E-05
Protein folding	2.8	2.57	5.14E-07	5.34E-05
ATP synthesis coupled electron transport	1.4	4.18	5.49E-07	5.51E-05
Mitochondrial ATP synthesis coupled electron transport	1.4	4.18	5.49E-07	5.51E-05
Proteasomal protein catabolic process	2	3.19	5.61E-07	5.46E-05
Proteasomal ubiquitin-dependent protein catabolic process	2	3.19	5.61E-07	5.46E-05
Hydrogen transport	1.5	3.92	6.99E-07	6.60E-05
Negative regulation of cellular protein metabolic process	2.8	2.53	7.73E-07	7.08E-05

Positive regulation of ubiquitin-protein ligase activity	1.6	3.72	8.25E-07	7.34E-05
Cell cycle	7.8	1.64	9.03E-07	7.81E-05
Respiratory electron transport chain	1.5	3.86	9.03E-07	7.61E-05
Macromolecular complex assembly	6.9	1.7	9.64E-07	7.90E-05
Positive regulation of protein ubiquitination	1.7	3.41	9.79E-07	7.82E-05
Regulation of ubiquitin-protein ligase activity during mitotic cell cycle	1.6	3.66	1.05E-06	8.15E-05
Anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process	1.5	3.8	1.16E-06	8.82E-05
Negative regulation of ubiquitin-protein ligase activity during mitotic cell cycle	1.5	3.8	1.16E-06	8.82E-05
Positive regulation of ligase activity	1.6	3.56	1.66E-06	1.23E-04
Negative regulation of ubiquitin-protein ligase activity	1.5	3.69	1.88E-06	1.36E-04
Negative regulation of ligase activity	1.5	3.69	1.88E-06	1.36E-04
Negative regulation of protein metabolic process	2.8	2.43	1.91E-06	1.35E-04
Cellular macromolecular complex assembly	4	2.05	1.99E-06	1.38E-04
Energy derivation by oxidation of organic compounds	2.3	2.62	4.00E-06	2.71E-04
Regulation of ubiquitin-protein ligase activity	1.6	3.34	4.83E-06	3.20E-04
Purine ribonucleotide biosynthetic process	2	2.78	7.44E-06	4.83E-04
Negative regulation of protein ubiquitination	1.5	3.34	8.73E-06	5.55E-04
Regulation of ligase activity	1.6	3.21	8.74E-06	5.45E-04
Macromolecular complex subunit organization	6.9	1.59	1.35E-05	8.22E-04
Purine ribonucleoside triphosphate biosynthetic process	1.7	2.92	1.35E-05	8.09E-04
Purine nucleoside triphosphate biosynthetic process	1.7	2.89	1.59E-05	9.36E-04
Ribonucleoside triphosphate biosynthetic process	1.7	2.89	1.59E-05	9.36E-04
Protein targeting	2.9	2.18	1.74E-05	0.001
Regulation of protein ubiquitination	1.7	2.86	1.87E-05	0.001
Ribonucleotide biosynthetic process	2	2.62	2.08E-05	0.001
Cellular macromolecular complex subunit organization	4	1.86	2.36E-05	0.001
Nucleoside triphosphate biosynthetic process	1.7	2.81	2.57E-05	0.001
Oxidation reduction	6.3	1.61	2.59E-05	0.001
Protein targeting to mitochondrion	1	4.59	2.86E-05	0.001
Protein localization in mitochondrion	1	4.59	2.86E-05	0.001
ATP biosynthetic process	1.6	2.92	3.63E-05	0.002
Purine ribonucleotide metabolic process	2.1	2.45	4.53E-05	0.002
DNA metabolic process	5.2	1.67	4.76E-05	0.002
Mitochondrial transport	1.3	3.2	5.14E-05	0.002
Purine ribonucleoside triphosphate metabolic process	1.8	2.56	7.16E-05	0.003
Ribonucleoside triphosphate metabolic process	1.8	2.54	8.18E-05	0.004
NADH metabolic process	0.5	9.76	1.22E-04	0.006
ATP metabolic process	1.7	2.6	1.25E-04	0.006
Ribonucleotide metabolic process	2.1	2.3	1.31E-04	0.006
Purine nucleoside triphosphate metabolic process	1.8	2.45	1.36E-04	0.006
Cell cycle phase	4.3	1.7	1.54E-04	0.007
Intracellular protein transport	4	1.74	1.55E-04	0.007
Protein transport	6.9	1.49	1.70E-04	0.007
Ribonucleoprotein complex assembly	1.3	3.02	1.90E-04	0.008
Cofactor metabolic process	2.5	2.07	1.93E-04	0.008
Establishment of protein localization	6.9	1.47	2.33E-04	0.009
Cellular amide metabolic process	1.1	3.25	2.52E-04	0.01
Regulation of cellular protein metabolic process	4.7	1.62	2.57E-04	0.01

Response to unfolded protein	1.3	2.93	2.65E-04	0.01
Purine nucleotide biosynthetic process	2	2.2	3.71E-04	0.014
Nucleoside triphosphate metabolic process	1.8	2.28	3.90E-04	0.015
Cellular protein localization	4.1	1.65	4.34E-04	0.016
Protein complex assembly	4.8	1.57	4.44E-04	0.017
Protein complex biogenesis	4.8	1.57	4.44E-04	0.017
M phase of mitotic cell cycle	2.6	1.92	4.70E-04	0.017
Ribosome assembly	0.5	7.8	4.81E-04	0.017
Cellular macromolecule localization	4.1	1.63	5.13E-04	0.018
Nitrogen compound biosynthetic process	3.4	1.72	5.99E-04	0.021
Nicotinamide metabolic process	0.9	3.58	6.90E-04	0.024
Nicotinamide nucleotide metabolic process	0.9	3.58	6.90E-04	0.024
protein localization in organelle	1.9	2.14	7.39E-04	0.026
M phase	3.4	1.7	7.65E-04	0.026
Nucleobase, nucleoside and nucleotide biosynthetic process	2.3	1.95	8.11E-04	0.027
Nucleobase, nucleoside, nucleotide and nucleic acid biosynthetic process	2.3	1.95	8.11E-04	0.027
Negative regulation of macromolecule metabolic process	6.4	1.44	8.27E-04	0.028
Alkaloid metabolic process	0.9	3.49	8.51E-04	0.028
Protein localization	7.5	1.39	9.47E-04	0.031
Purine nucleotide metabolic process	2.2	1.96	9.87E-04	0.032

Table B.6.: Complete GO annotation for weakly or not expressed CHO DP-12 PMD-genes (2360) with Benjamini p values ≤ 0.05 (Wippermann et al. 2015).

GO term (category 'biological process')	%	Fold enrichment	p value	Benjamini p value
Cell adhesion	6.30	1.63	6.01E-10	1.23E-06
Synaptic transmission	3.10	1.85	2.02E-07	2.77E-04
Transmission of nerve impulse	3.40	1.78	2.25E-07	2.31E-04
Behavior	4.20	1.60	1.70E-06	0.001
Elevation of cytosolic calcium ion concentration	1.40	2.37	2.23E-06	0.002
Calcium ion homeostasis	2.10	2.00	2.45E-06	0.001
Neurological system process	9.00	1.34	2.48E-06	0.001
Cellular calcium ion homeostasis	2.00	2.01	2.61E-06	0.001
Cellular di-, tri-valent inorganic cation homeostasis	2.30	1.86	5.97E-06	0.002
Cellular metal ion homeostasis	2.10	1.92	8.43E-06	0.003
Response to wounding	4.50	1.52	8.63E-06	0.003
Cytosolic calcium ion homeostasis	1.40	2.21	1.18E-05	0.004
Cellular chemical homeostasis	3.40	1.62	1.27E-05	0.004
Cellular ion homeostasis	3.40	1.62	1.28E-05	0.003
Metal ion homeostasis	2.10	1.87	1.38E-05	0.004
Di-, tri-valent inorganic cation homeostasis	2.40	1.80	1.39E-05	0.003
Cell-cell signaling	4.80	1.45	3.90E-05	0.009
Cellular cation homeostasis	2.40	1.72	4.26E-05	0.009
Regulation of system process	2.80	1.64	5.03E-05	0.01
Positive regulation of hydrolase activity	1.80	1.84	8.70E-05	0.017
Ion homeostasis	3.40	1.52	1.03E-04	0.019
Cell surface receptor linked signal transduction	12.50	1.22	1.13E-04	0.02

Regulation of G-protein coupled receptor protein signaling pathway	0.80	2.71	1.53E-04	0.026
Chemical homeostasis	4.10	1.44	1.88E-04	0.03
Regulation of phospholipase activity	0.90	2.38	1.88E-04	0.029
Activation of phospholipase C activity by G-protein coupled receptor protein signaling pathway coupled to IP3 second messenger	0.80	2.66	2.00E-04	0.03
Regulation of lipase activity	1.10	2.20	2.22E-04	0.032
Embryonic skeletal system development	1.00	2.29	2.32E-04	0.032
G-protein coupled receptor protein signaling pathway	7.90	1.27	2.88E-04	0.039
Locomotory behavior	2.40	1.60	3.52E-04	0.046
Positive regulation of phospholipase activity	0.90	2.34	3.56E-04	0.045
Metal ion transport	3.70	1.44	4.10E-04	0.05

Table B.7.: Motif enrichment in DMRs 6, 12, 24 and 36 hours after butyrate addition identified by *DREME* (Bailey 2011). A = Adenine, T = Thymine, C = Cytosine, G = Guanine, N "Any base" = ACGT, V "Not T" = ACG, H "Not G" = ACT, D "Not C" = AGT, B "Not A" = CGT, M "Amino" = AC, R "Purine" = AG, W "Weak" = AT, S "Strong" = CG, Y "Pyrimidine" = CT, K "Keto" = GT, U = T. Significantly enriched DNA motifs are displayed separately for upstream, 5', intragene and 3' regions. Also indicated are the corresponding *p* values.

Sample	Upstream region		5'		Intragene		3'	
	Motif (5'→3')	<i>p</i> value	Motif (5'→3')	<i>p</i> value	Motif (5'→3')	<i>p</i> value	Motif (5'→3')	<i>p</i> value
6 h	CGD	3.7E-338	CGD	6.60E-48	CGD	5.70E-107	CGB	2.10E-10
	CGC	6.80E-29	CGC	1.20E-16	CGC	2.40E-12		
	CCWSCC	1.90E-23	GGSC	9.00E-07	CCCWGSAG	4.80E-11		
	GRGGGS	1.40E-15	GGTCTAYC	1.00E-06	CCTSCAG	5.00E-10		
	CAGGGB	1.60E-11			GGGRCC	5.30E-09		
	GGACAS	8.20E-11			CCTCMYC	6.00E-10		
	GAGKTCCC	8.70E-11			CACASMC	9.20E-09		
	GAGGGAGM	1.60E-09			GRGTCAC	5.90E-08		
	BCCAG	1.50E-09			GGCTKCC	6.90E-08		
	GRGSC	2.10E-08			CATGGGGA	1.10E-07		
	ATWCCATC	6.30E-08			CTVCTGTC	1.40E-08		
	AKGAWGGG	2.20E-07			CAGGSAC	9.30E-08		
	ACWCCTG	2.30E-07			AGWSCCC	3.10E-07		
	CACAGASA	3.40E-07			GGTGSC	5.00E-07		
	GGAGATSA	7.40E-07						
	AKCCTGG	7.50E-07						
	ACCYACAG	1.10E-06						
12 h	CGD	1.6E-328	CGD	1.30E-139	CGW	1.40E-50	CGB	7.70E-10
	CGC	9.50E-28	CGC	3.00E-17	CGS	5.10E-36	ATTCCWGK	1.90E-08
	CHCCHCC	1.70E-16			CTGNCCC	6.10E-12		
	SCCAGS	1.90E-14			CMCASCC	1.80E-11		
	GCCYCC	1.00E-10			CCCTSCTS	1.90E-11		
	CCAGGR	9.80E-09			GGKGACY	9.00E-11		
	CTCTGRG	8.50E-08			CCWGAGAS	3.90E-10		
	GAGGGAYA	5.30E-07			AGGGSWC	1.50E-09		
	ARACCAG	1.00E-06			CCCCARG	6.20E-09		
	STGTCAGA	9.70E-07			GCCMCWC	1.20E-08		
	CAGGGK	1.10E-06			GASTCAC	4.00E-08		
	ATCCCTKC	1.10E-06			GCCASC	6.80E-08		
					BCTGCTGC	7.70E-08		

					CCTSCAG	1.50E-07		
					CTCMYCC	4.90E-07		
					STGGGGA	5.80E-07		
					CCWGGAY	9.20E-07		
					AKCTSTGG	1.50E-06		
24 h	CGD	1.9E-312	CGD	5.10E-126	CGD	1.30E-100	CGB	1.40E-10
	CGC	6.90E-26	CGC	4.00E-15	CCWGSAG	1.10E-11		
	GRGGGS	2.40E-16	GGAAMTCC	1.10E-06	CGCGCGC	2.40E-11		
	CCWGSC	4.00E-14			CGC	1.40E-09		
	GGAAVTCC	2.20E-09			CCCAGRGR	6.50E-09		
	CCCAGS	3.90E-09			RGGGMCA	9.50E-09		
	CTAKGAAC	7.20E-08			CCAGSYC	2.20E-08		
	CCHCC	7.40E-08			CMCACHC	1.80E-08		
	GGWGATCA	4.60E-08			GGKGAC	3.40E-08		
	GGASWCC	2.50E-07			GASTCAC	1.00E-07		
	AMTGATCC	4.20E-07			CTGCMC	6.50E-07		
	ACAGSTGA	4.90E-07			CTACCTAG	1.10E-06		
	GTACYCCA	7.30E-07			GAGCCYTA	1.50E-06		ACCCARTC
	CCAAKAYC	8.10E-07						
	CTCTCAYC	1.40E-06						
	ACCCARTC	1.50E-06						
36 h	CGD	1.80E-298	CGD	5.10E-137	CGD	3.20E-103	CGY	8.40E-08
	CGC	1.60E-24	CGC	3.90E-18	CGC	2.80E-14	GATAGASA	3.60E-07
	GRGGS	5.30E-16			AGGGSYC	2.00E-12		
	CCCWGS	2.00E-11			CAGCTCAK	2.40E-09		
	CCAGGM	1.70E-10			CCTGCWG	2.00E-09		
	GTACTCCA	6.90E-08			CMCASAC	3.00E-09		
	CCTSAACA	8.30E-08			CCYASAGC	6.10E-09		
	CTGGATCW	9.60E-08			CCTGAGAS	1.80E-08		
	GGTMCAGA	1.70E-07			CCCCTYC	2.60E-08		
	AGATCWG	1.20E-06			SCCTGGR	6.10E-08		
	CTCCTSAA	1.70E-06			GRGTCAC	1.60E-07		
					GAGGCTSC	2.10E-07		
					GTRAGGGA	3.10E-07		
					ACCCWGC	3.20E-07		
					CCAGDGAC	3.20E-07		
					CCTGTSYC	6.50E-07		
					CCTGGRC	1.10E-06		
					GCCAAGCC	1.40E-06		
					CTGTGGYC	1.40E-06		

Table B.8.: Motifs enriched in hypomethylated DMRs 6, 12, 24 and 36 hours after butyrate addition identified by *DREME* (Bailey 2011). A = Adenine, T = Thymine, C = Cytosine, G = Guanine, N "Any base" = ACGT, V "Not T" = ACG, H "Not G" = ACT, D "Not C" = AGT, B "Not A" = CGT, M "Amino" = AC, R "Purine" = AG, W "Weak" = AT, S "Strong" = CG, Y "Pyrimidine" = CT, K "Keto" = GT, U = T. Significantly enriched DNA motifs are displayed separately for upstream, 5', intragene and 3' regions. Also indicated are the corresponding *p* values.

Sample	Upstream region		5'		Intragene		3'	
	Motif (5'→3')	<i>p</i> value	Motif (5'→3')	<i>p</i> value	Motif (5'→3')	<i>p</i> value	Motif (5'→3')	<i>p</i> value
6 h	CGD	2.9e-426	CGH	1.60E-153	CGD	4.00E-132	CGH	6.50E-14
	CGC	3.60E-39	CCG	8.00E-26	CGC	2.10E-12		
	CCWGGDC	8.80E-26	BCCAGC	2.30E-11	GGGCWSC	9.10E-12		
	GSTGGS	1.00E-21	CCHCC	3.40E-09	CACWGGG	2.40E-11		
	GCCYCY	3.60E-18	ARCCCTG	6.40E-07	CCMCMCCC	8.70E-11		
	CCCTGB	3.40E-14	CTGGGBC	7.40E-07	CWCCTGS	2.60E-10		

	CCCCMC	1.20E-12			GCWGTGSC	2.90E-11		
	CCWGSC	2.10E-11			CTGARGCC	8.80E-10		
	TTCCTGGR	2.30E-09			CYCCAGC	2.80E-09		
	GCHC	2.70E-09			CAGRGYC	4.40E-09		
	AGGGGWC	1.00E-07			CAKCTGCA	9.10E-09		
	CAGACWGG	1.20E-07			CCTGRKG	4.70E-08		
	GATCCTR	4.70E-07			GGGACAS	2.80E-08		
	CCAGGW	7.70E-07			CCWGGCY	9.00E-08		
	GGTCYAC	8.40E-07			CAGGGS	3.80E-08		
	GKGTGTS	2.00E-06			CCTTGRCC	3.30E-07		
					RGCACCY	3.20E-07		
					CAGMCACA	5.10E-07		
					GGTGGMCW	4.90E-08		
					AGGAYCCC	1.20E-06		
					CHCTTGGA	1.10E-06		
					CWGCTCTC	1.70E-06		
12 h	CGD	5.9e-459	CGD	6.30E-184	CGH	1.40E-126	CGB	2.20E-13
	CGC	1.60E-49	CGC	5.50E-17	CCG	2.40E-18	CCTCTTAG	1.10E-06
	BCCAGS	2.10E-26	CTGSCY	6.20E-13	CHCAGCC	1.10E-12	GAAACTTA	1.50E-06
	GGDGGB	4.00E-22	CCAGS	1.20E-08	CTGBCCC	3.10E-12		
	CCTGYC	1.80E-13	CCCMCMC	4.20E-07	AGCYCAGR	5.10E-12		
	SAGGS	1.30E-12	GGTATACA	1.20E-06	CWCAGRGC	3.40E-11		
	CACMCC	4.50E-11			CYCCWGC	1.10E-10		
	CHCAGAG	1.10E-09			GGWGCY	3.00E-09		
	CAGGWGTY	4.50E-09			CCWKGKAC	3.00E-10		
	GWTCWGGA	3.90E-08			GAGHCCC	5.80E-09		
	GGGAYACC	6.10E-08			GCAGSY	1.10E-08		
	BAGACC	2.00E-07			CARGYCC	2.60E-09		
	ACWCCC	3.20E-07			CTBAGGGA	2.70E-08		
					CCAGGS	3.60E-08		
					GAGRGGMC	4.20E-09		
					CTGSCC	4.40E-08		
					CACWGGAG	2.10E-07		
					AMCCCCAG	2.80E-07		
					CWGWGCC	3.40E-07		
					GATGAGWC	3.80E-07		
					CCCCWTCC	5.00E-07		
					CACCYCC	9.50E-07		
					CYCAATGC	1.40E-06		
24 h	CGD	3.7e-446	CGD	2.00E-190	CGD	7.40E-153	CGB	4.10E-15
	CGC	4.60E-47	CGC	1.30E-24	CGC	2.20E-16	CAAKATCC	7.30E-07
	CCWGGGS	1.10E-28	SCAGGS	6.00E-11	CTGGGSY	8.00E-13		
	CCASCMC	1.30E-23	CCHCC	1.40E-09	CMCMCMC	1.10E-11		
	GCCYCY	4.40E-20	GCTGGS	1.80E-07	CTSAGMC	1.10E-11		
	CAGGGTKK	8.80E-15	ARCCC	1.70E-06	GCASCC	3.50E-11		
	CYCMC	1.00E-13			HCAGGGY	4.00E-11		
	GGSCY	5.30E-12			CAVCTCC	2.60E-10		
	ASACCWG	7.90E-11			CCMCAGM	3.00E-10		
	TCCAGGA	1.50E-10			CASWGGG	3.60E-09		
	CTGCCY	4.30E-08			CTGCCCY	2.70E-08		
	GGSTS	5.90E-08			GCCTGS	4.70E-09		
	GAGATRCC	4.00E-07			ASCTGCAS	3.20E-09		
	GGTGS	9.80E-07			CCCTSCWG	5.00E-08		
	GGTGTCCM	9.20E-07			CCCCTR	4.30E-07		
					GGACACA	7.70E-07		
					CCAGAWGR	8.00E-07		
					AGWGAGC	1.30E-06		

					AGGRGCC	1.40E-06		
36 h	CGD	2.3e-483	CGB	3.00E-196	CGD	4.30E-156	CGD	4.60E-14
	CGC	1.40E-49	CGA	3.60E-21	CGC	1.60E-13	ATTCYACC	1.40E-07
	CCWGSCY	1.20E-33	CCWGCC	4.10E-12	CCAGGSH	4.00E-12	ACAGGAAY	4.80E-07
	GGDGGGB	4.60E-23	CCCMMCC	3.00E-10	KCTGGGS	1.10E-11	GTTGAGGK	8.30E-07
	CCWGGRS	1.80E-16	CTCTGRG	8.00E-08	CCTGWGRC	4.40E-11		
	AVWCCC	4.40E-14	CTGTAGMC	1.20E-07	GCYCTGA	9.40E-11		
	GGBCY	1.30E-10	AGGGTTB	1.30E-07	CTGKCCC	3.00E-10		
	GGCAS	1.10E-08	CCAGGM	1.20E-06	CTSAGCC	3.20E-10		
	CTCTGWG	2.10E-08	AGTVACCC	4.20E-07	CTGCTCWS	1.20E-09		
	GGGAYAC	8.00E-08			CWCTGGS	1.10E-09		
	GACTYCTC	2.60E-07			CTGAGGHC	6.60E-09		
	CAGGATS	3.60E-07			ABTGTGGG	6.40E-09		
	BTGGGG	1.30E-06			GRGKAC	6.10E-09		
	RCACCTCA	1.80E-06			GCCCCY	1.00E-08		
					CWCCYCC	7.80E-09		
					CAGGWS	1.20E-08		
					AGGACYC	1.60E-07		
					CAGTGTGS	1.70E-07		
					AGKCCATG	3.90E-07		
					CCACCHAA	5.90E-07		
					GGCTTSA	1.00E-06		
					GCTGCAS	1.30E-06		
					GCAGTGTY	1.70E-06		

Table B.9.: DNA motifs *de novo* discovered in DMRs upon butyrate addition were compared to human and mouse transcription factor databases (Wippermann et al. 2016). Motif comparisons were performed using *Tomtom* (Gupta et al. 2007). *P* values are indicated for both the motif discovery and motif comparison. Seed and wobble bases: A = adenine, T = thymine, C = cytosine, G = guanine, M "amino" = AC, R "purine" = AG, W "weak" = AT, S "strong" = CG, Y "pyrimidine" = CT. Also shown are gene expression values as determined by CHO-specific DNA-microarrays as fold-changes. n.a. = gene not present on the microarray; n.d. = gene present on microarray but not significantly differentially expressed.

DMR motif (5'→3')	Genomic position	<i>p</i> value DREME	Putative TF ID	<i>p</i> value Tomtom	TF expression (CHO DNA-microarray)
ACCCARTC	Upstream	1.50E-06	DUXL	6.27E-05	n.a.
CCWGGRS	Upstream	1.80E-16	EBF1	2.04E-05	n.d.
CCTGYC	Upstream	1.80E-13	ID4	4.22E-05	n.a.
CCTGYC	Upstream	1.80E-13	SNAI2	4.69E-05	n.d.
GCCYCY	Upstream	4.40E-20	SP2	2.79E-05	n.d.
GCCYCY	Upstream	4.40E-20	SP1	1.58E-04	0.7-fold
GCCYCY	Upstream	4.40E-20	KLF13	4.83E-04	0.8-fold
CACMCC	Upstream	4.50E-11	TBX15	6.74E-05	n.d.
CACMCC	Upstream	4.50E-11	TBX1	9.65E-05	n.d.
CACMCC	Upstream	4.50E-11	TBX20	1.35E-04	n.d.
CACMCC	Upstream	4.50E-11	TBX21	1.61E-04	n.d.
CACMCC	Upstream	4.50E-11	TBR1	1.93E-04	n.d.
CACMCC	Upstream	4.50E-11	MGA	1.93E-04	0.6-fold
CACMCC	Upstream	4.50E-11	TBX5	1.93E-04	n.d.

CACMCC	Upstream	4.50E-11	HIC1	2.14E-04	n.d.
CACMCC	Upstream	4.50E-11	SREBF2	2.14E-04	n.d.
CACMCC	Upstream	4.50E-11	TBX4	2.27E-04	n.d.
CACMCC	Upstream	4.50E-11	EOMES	3.42E-04	n.d.
ACAGSTGA	Upstream	4.90E-07	TFAP4	4.23E-05	0.7-fold
ACAGSTGA	Upstream	4.90E-07	MEIS3	2.84E-04	n.d.
ACAGSTGA	Upstream	4.90E-07	MEIS2	2.84E-04	n.d.
ACAGSTGA	Upstream	4.90E-07	PKNOX1	3.34E-04	n.d.
ACAGSTGA	Upstream	4.90E-07	PKNOX2	3.54E-04	n.d.
ACAGSTGA	Upstream	4.90E-07	TGIF2	3.97E-04	n.d.
GRGGS	Upstream	5.30E-16	TFAP2C	1.69E-05	0.8-fold
GRGGS	Upstream	5.30E-16	TFAP2A	3.73E-04	0.6-fold
CCCCMC	Upstream	1.20E-12	ZNF740	3.66E-05	n.d.
CCCCMC	Upstream	1.20E-12	ZFP281	4.37E-05	n.a.
CCCCMC	Upstream	1.20E-12	ZFP740	7.21E-05	n.d.
CCCCMC	Upstream	1.20E-12	FOXO4	1.01E-04	n.d.
CCCCMC	Upstream	1.20E-12	GLIS2	1.07E-04	0.6-fold
CCCCMC	Upstream	1.20E-12	ZIC1	1.07E-04	n.d.
CCCCMC	Upstream	1.20E-12	ZIC3	1.19E-04	n.d.
CCCCMC	Upstream	1.20E-12	FOXO3	1.23E-04	n.d.
CCCCMC	Upstream	1.20E-12	GLIS3	1.30E-04	1.4-fold
CCCCMC	Upstream	1.20E-12	FOXO1	1.43E-04	n.d.
CCCCMC	Upstream	1.20E-12	PLAGL1	1.73E-04	n.d.
CCCCMC	Upstream	1.20E-12	FOXO6	1.84E-04	n.d.
CCCCMC	Upstream	1.20E-12	GLIS1	2.25E-04	1.4-fold
CCCCMC	Upstream	1.20E-12	ZIC4	4.84E-04	n.d.
GKGTGTS	Upstream	2.00E-06	ZSCAN4	6.47E-05	1.2-fold
GKGTGTS	Upstream	2.00E-06	KLF1	2.20E-04	n.d.
GGAAMTCC	5'	1.10E-06	FLI1	3.05E-05	n.d.
GGAAMTCC	5'	1.10E-06	ERG	4.06E-05	n.d.
GGAAMTCC	5'	1.10E-06	REL	7.60E-05	0.7-fold
GGAAMTCC	5'	1.10E-06	RELA	1.29E-04	0.7-fold
GGAAMTCC	5'	1.10E-06	NFKB1	2.51E-04	0.8-fold
CCMCMC	5'	4.20E-07	GLI2	5.74E-05	n.d.
CCMCMC	5'	4.20E-07	RREB1	9.56E-05	0.7-fold
CCMCMC	5'	4.20E-07	FOXO1	4.16E-04	1.4-fold

Table B.10.: Gene ontology (GO) classification of differentially expressed genes (929) with expression profiles belonging to cluster I (Wippermann et al. 2016).

GO ID	GO term	Count	<i>p</i> value
GO:0055114	oxidation reduction	62	9.61E-08
GO:0016192	vesicle-mediated transport	53	4.67E-06
GO:0006091	generation of precursor metabolites and energy	33	3.16E-05
GO:0048193	Golgi vesicle transport	19	4.14E-05
GO:0046907	intracellular transport	54	8.53E-05

Appendix B. Supplementary material

GO:0009719	response to endogenous stimulus	37	1.89E-04
GO:0006119	oxidative phosphorylation	15	1.91E-04
GO:0010035	response to inorganic substance	23	2.67E-04
GO:0006892	post-Golgi vesicle-mediated transport	11	3.36E-04
GO:0008610	lipid biosynthetic process	30	6.62E-04
GO:0019220	regulation of phosphate metabolic process	40	7.72E-04
GO:0051174	regulation of phosphorus metabolic process	40	7.72E-04
GO:0010033	response to organic substance	54	8.27E-04
GO:0045859	regulation of protein kinase activity	31	9.07E-04
GO:0031667	response to nutrient levels	21	0.001
GO:0010038	response to metal ion	16	0.001
GO:0009725	response to hormone stimulus	32	0.001
GO:0051338	regulation of transferase activity	32	0.002
GO:0043549	regulation of kinase activity	31	0.002
GO:0006886	intracellular protein transport	32	0.002
GO:0006368	RNA elongation from RNA polymerase II promoter	9	0.002
GO:0009991	response to extracellular stimulus	22	0.002
GO:0051348	negative regulation of transferase activity	13	0.002
GO:0019318	hexose metabolic process	20	0.002
GO:0006979	response to oxidative stress	18	0.002
GO:0030384	phosphoinositide metabolic process	11	0.002
GO:0015986	ATP synthesis coupled proton transport	8	0.002
GO:0015985	energy coupled proton transport, down electrochemical gradient	8	0.002
GO:0042325	regulation of phosphorylation	37	0.002
GO:0006354	RNA elongation	9	0.002
GO:0006469	negative regulation of protein kinase activity	12	0.002
GO:0022900	electron transport chain	14	0.003
GO:0015980	energy derivation by oxidation of organic compounds	16	0.003
GO:0033673	negative regulation of kinase activity	12	0.003
GO:0008219	cell death	51	0.004
GO:0006650	glycerophospholipid metabolic process	14	0.004
GO:0034613	cellular protein localization	33	0.004
GO:0032870	cellular response to hormone stimulus	15	0.004
GO:0016265	death	51	0.004
GO:0046486	glycerolipid metabolic process	17	0.004
GO:0070727	cellular macromolecule localization	33	0.004
GO:0005996	monosaccharide metabolic process	21	0.004
GO:0006796	phosphate metabolic process	64	0.006
GO:0006793	phosphorus metabolic process	64	0.006
GO:0007584	response to nutrient	15	0.006
GO:0009206	purine ribonucleoside triphosphate biosynthetic process	12	0.006
GO:0033554	cellular response to stress	41	0.007
GO:0009201	ribonucleoside triphosphate biosynthetic process	12	0.007
GO:0009145	purine nucleoside triphosphate biosynthetic process	12	0.007
GO:0008202	steroid metabolic process	19	0.007
GO:0034220	ion transmembrane transport	8	0.008

GO:0019217	regulation of fatty acid metabolic process	8	0.008
GO:0009142	nucleoside triphosphate biosynthetic process	12	0.008
GO:0009141	nucleoside triphosphate metabolic process	14	0.009
GO:0006754	ATP biosynthetic process	11	0.009
GO:0009205	purine ribonucleoside triphosphate metabolic process	13	0.009
GO:0010243	response to organic nitrogen	9	0.009
GO:0045454	cell redox homeostasis	9	0.009
GO:0006818	hydrogen transport	9	0.009
GO:0009199	ribonucleoside triphosphate metabolic process	13	0.01
GO:0022904	respiratory electron transport chain	9	0.01

Table B.11.: Gene ontology (GO) classification of differentially expressed genes (1014) with expression profiles belonging to cluster II (Wippermann et al. 2016).

GO ID	GO term	Count	<i>p</i> value
GO:0007049	cell cycle	106	8.21E-19
GO:0000278	mitotic cell cycle	68	1.07E-18
GO:0022402	cell cycle process	86	3.68E-18
GO:0022403	cell cycle phase	68	5.17E-16
GO:0000087	M phase of mitotic cell cycle	48	6.53E-16
GO:0000280	nuclear division	47	1.61E-15
GO:0007067	mitosis	47	1.61E-15
GO:0048285	organelle fission	47	7.86E-15
GO:0000279	M phase	57	1.74E-14
GO:0051276	chromosome organization	71	4.75E-14
GO:0051301	cell division	50	2.05E-12
GO:0006325	chromatin organization	57	6.45E-12
GO:0006396	RNA processing	71	1.64E-11
GO:0008380	RNA splicing	43	3.18E-09
GO:0000375	RNA splicing, via transesterification reactions	30	3.23E-09
GO:0000377	RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	30	3.23E-09
GO:0000398	nuclear mRNA splicing, via spliceosome	30	3.23E-09
GO:0034622	cellular macromolecular complex assembly	45	1.07E-08
GO:0006323	DNA packaging	25	1.47E-08
GO:0065003	macromolecular complex assembly	73	1.49E-08
GO:0034621	cellular macromolecular complex subunit organization	48	1.66E-08
GO:0006333	chromatin assembly or disassembly	26	1.78E-08
GO:0043933	macromolecular complex subunit organization	76	2.13E-08
GO:0031497	chromatin assembly	21	3.14E-08
GO:0065004	protein-DNA complex assembly	21	7.02E-08
GO:0007017	microtubule-based process	37	1.16E-07
GO:0008283	cell proliferation	51	4.84E-07
GO:0034728	nucleosome organization	20	5.01E-07
GO:0051726	regulation of cell cycle	42	7.23E-07
GO:0006397	mRNA processing	41	8.45E-07

GO:0016071	mRNA metabolic process	45	8.83E-07
GO:0032886	regulation of microtubule-based process	14	1.44E-06
GO:0048145	regulation of fibroblast proliferation	12	1.47E-06
GO:0010605	negative regulation of macromolecule metabolic process	72	1.52E-06
GO:0006334	nucleosome assembly	18	2.28E-06
GO:0016568	chromatin modification	36	2.33E-06
GO:0007346	regulation of mitotic cell cycle	25	2.38E-06
GO:0000226	microtubule cytoskeleton organization	24	4.51E-06
GO:0031399	regulation of protein modification process	37	4.85E-06
GO:0010604	positive regulation of macromolecule metabolic process	78	8.71E-06
GO:0070507	regulation of microtubule cytoskeleton organization	12	1.06E-05
GO:0032268	regulation of cellular protein metabolic process	50	1.19E-05
GO:0051172	negative regulation of nitrogen compound metabolic process	53	1.54E-05
GO:0016481	negative regulation of transcription	48	2.30E-05
GO:0007059	chromosome segregation	16	2.74E-05
GO:0010558	negative regulation of macromolecule biosynthetic process	54	3.22E-05
GO:0031327	negative regulation of cellular biosynthetic process	55	3.28E-05
GO:0010564	regulation of cell cycle process	19	4.27E-05
GO:0045934	negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	51	4.45E-05
GO:0048146	positive regulation of fibroblast proliferation	9	4.90E-05
GO:0045449	regulation of transcription	185	5.24E-05
GO:0009890	negative regulation of biosynthetic process	55	5.79E-05
GO:0051247	positive regulation of protein metabolic process	30	5.99E-05
GO:0032270	positive regulation of cellular protein metabolic process	29	7.12E-05
GO:0031400	negative regulation of protein modification process	19	7.63E-05
GO:0008284	positive regulation of cell proliferation	43	7.66E-05
GO:0007093	mitotic cell cycle checkpoint	11	8.15E-05
GO:0031401	positive regulation of protein modification process	25	8.21E-05
GO:0022613	ribonucleoprotein complex biogenesis	24	1.23E-04
GO:0007051	spindle organization	11	1.23E-04
GO:0042516	regulation of tyrosine phosphorylation of Stat3 protein	7	1.90E-04
GO:0010629	negative regulation of gene expression	48	2.23E-04
GO:0050658	RNA transport	16	2.31E-04
GO:0051236	establishment of RNA localization	16	2.31E-04
GO:0050657	nucleic acid transport	16	2.31E-04
GO:0006260	DNA replication	24	2.76E-04
GO:0042127	regulation of cell proliferation	67	2.97E-04
GO:0045596	negative regulation of cell differentiation	26	3.07E-04
GO:0006403	RNA localization	16	3.25E-04
GO:0051173	positive regulation of nitrogen compound metabolic process	57	3.57E-04
GO:0051253	negative regulation of RNA metabolic process	37	3.58E-04
GO:0015931	nucleobase, nucleoside, nucleotide and nucleic acid transport	17	4.07E-04
GO:0070271	protein complex biogenesis	47	4.39E-04

GO:0006461	protein complex assembly	47	4.39E-04
GO:0045935	positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	55	5.04E-04
GO:0040001	establishment of mitotic spindle localization	5	5.15E-04
GO:0045892	negative regulation of transcription, DNA-dependent	36	5.32E-04
GO:0012501	programmed cell death	54	5.37E-04
GO:0051603	proteolysis involved in cellular protein catabolic process	53	6.17E-04
GO:0006915	apoptosis	53	6.67E-04
GO:0044257	cellular protein catabolic process	53	6.89E-04
GO:0033043	regulation of organelle organization	25	7.77E-04
GO:0009891	positive regulation of biosynthetic process	59	7.80E-04
GO:0006259	DNA metabolic process	46	8.17E-04
GO:0045787	positive regulation of cell cycle	11	9.20E-04
GO:0042254	ribosome biogenesis	17	9.58E-04
GO:0051781	positive regulation of cell division	9	0.001
GO:0042509	regulation of tyrosine phosphorylation of STAT protein	8	0.001
GO:0001932	regulation of protein amino acid phosphorylation	21	0.001
GO:0007018	microtubule-based movement	16	0.001
GO:0019941	modification-dependent protein catabolic process	50	0.001
GO:0043632	modification-dependent macromolecule catabolic process	50	0.001
GO:0016569	covalent chromatin modification	17	0.001
GO:0051329	interphase of mitotic cell cycle	15	0.001
GO:0030163	protein catabolic process	53	0.001
GO:0051340	regulation of ligase activity	13	0.001
GO:0051653	spindle localization	5	0.001
GO:0051293	establishment of spindle localization	5	0.001
GO:0001933	negative regulation of protein amino acid phosphorylation	8	0.001
GO:0045941	positive regulation of transcription	49	0.001
GO:0010557	positive regulation of macromolecule biosynthetic process	55	0.001
GO:0051439	regulation of ubiquitin-protein ligase activity during mitotic cell cycle	12	0.002
GO:0031328	positive regulation of cellular biosynthetic process	57	0.002
GO:0006511	ubiquitin-dependent protein catabolic process	26	0.002
GO:0051098	regulation of binding	19	0.002
GO:0008219	cell death	59	0.002
GO:0051254	positive regulation of RNA metabolic process	43	0.002
GO:0022415	viral reproductive process	11	0.002
GO:0051325	interphase	15	0.002
GO:0051351	positive regulation of ligase activity	12	0.002
GO:0032269	negative regulation of cellular protein metabolic process	21	0.002
GO:0016265	death	59	0.002
GO:0044265	cellular macromolecule catabolic process	59	0.002
GO:0006913	nucleocytoplasmic transport	19	0.002
GO:0010627	regulation of protein kinase cascade	26	0.002
GO:0051169	nuclear transport	19	0.002
GO:0045893	positive regulation of transcription, DNA-dependent	42	0.003
GO:0051028	mRNA transport	13	0.003

Appendix B. Supplementary material

GO:0032259	methylation	12	0.003
GO:0010628	positive regulation of gene expression	49	0.003
GO:0006351	transcription, DNA-dependent	29	0.003
GO:0009792	embryonic development ending in birth or egg hatching	32	0.003
GO:0070647	protein modification by small protein conjugation or removal	19	0.003
GO:0051294	establishment of spindle orientation	4	0.003
GO:0034508	centromere complex assembly	4	0.003
GO:0000132	establishment of mitotic spindle orientation	4	0.003
GO:0042532	negative regulation of tyrosine phosphorylation of STAT protein	4	0.003
GO:0046425	regulation of JAK-STAT cascade	8	0.003
GO:0051248	negative regulation of protein metabolic process	21	0.003
GO:0006338	chromatin remodeling	10	0.003
GO:0006350	transcription	143	0.003
GO:0009967	positive regulation of signal transduction	29	0.003
GO:0032774	RNA biosynthetic process	29	0.003
GO:0051438	regulation of ubiquitin-protein ligase activity	12	0.003
GO:0051302	regulation of cell division	9	0.004
GO:0051437	positive regulation of ubiquitin-protein ligase activity during mitotic cell cycle	11	0.004
GO:0034660	ncRNA metabolic process	24	0.004
GO:0000075	cell cycle checkpoint	13	0.004
GO:0043414	biopolymer methylation	11	0.004
GO:0010563	negative regulation of phosphorus metabolic process	9	0.004
GO:0045936	negative regulation of phosphate metabolic process	9	0.004
GO:0051640	organelle localization	13	0.004
GO:0006364	rRNA processing	13	0.004
GO:0051100	negative regulation of binding	10	0.004
GO:0043009	chordate embryonic development	31	0.004
GO:0051443	positive regulation of ubiquitin-protein ligase activity	11	0.004
GO:0019221	cytokine-mediated signaling pathway	11	0.004
GO:0051129	negative regulation of cellular component organization	17	0.005
GO:0010941	regulation of cell death	63	0.005
GO:0060216	definitive hemopoiesis	4	0.005
GO:0006474	N-terminal protein amino acid acetylation	4	0.005
GO:0046426	negative regulation of JAK-STAT cascade	4	0.005
GO:0010639	negative regulation of organelle organization	12	0.005
GO:0016032	viral reproduction	11	0.005
GO:0044092	negative regulation of molecular function	31	0.005
GO:0006917	induction of apoptosis	30	0.005
GO:0012502	induction of programmed cell death	30	0.005
GO:0033135	regulation of peptidyl-serine phosphorylation	5	0.006
GO:0031365	N-terminal protein amino acid modification	5	0.006
GO:0031398	positive regulation of protein ubiquitination	12	0.006
GO:0016072	rRNA metabolic process	13	0.006

GO:0031110	regulation of microtubule polymerization or depolymerization	7	0.006
GO:0043392	negative regulation of DNA binding	9	0.006
GO:0006275	regulation of DNA replication	10	0.006
GO:0051101	regulation of DNA binding	15	0.006
GO:0009057	macromolecule catabolic process	60	0.007
GO:0016570	histone modification	15	0.007
GO:0043067	regulation of programmed cell death	62	0.007
GO:0034470	ncRNA processing	20	0.007
GO:0006473	protein amino acid acetylation	9	0.007
GO:0051252	regulation of RNA metabolic process	123	0.007
GO:0030307	positive regulation of cell growth	8	0.007
GO:0043623	cellular protein complex assembly	18	0.007
GO:0046605	regulation of centrosome cycle	4	0.007
GO:0046784	intronless viral mRNA export from host nucleus	4	0.007
GO:0007052	mitotic spindle organization	5	0.007
GO:0010647	positive regulation of cell communication	30	0.007
GO:0043193	positive regulation of gene-specific transcription	12	0.008
GO:0006730	one-carbon metabolic process	14	0.008
GO:0042981	regulation of apoptosis	61	0.008
GO:0007569	cell aging	7	0.008
GO:0045937	positive regulation of phosphate metabolic process	13	0.008
GO:0031396	regulation of protein ubiquitination	13	0.008
GO:0010562	positive regulation of phosphorus metabolic process	13	0.008
GO:0031145	anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process	10	0.008
GO:0031055	chromatin remodeling at centromere	3	0.009
GO:0006366	transcription from RNA polymerase II promoter	23	0.009
GO:0051052	regulation of DNA metabolic process	14	0.009
GO:0007026	negative regulation of microtubule depolymerization	5	0.009
GO:0031114	regulation of microtubule depolymerization	5	0.009

Table B.12.: Gene ontology (GO) classification of differentially expressed genes (2861) with expression profiles belonging to cluster III (Wippermann et al. 2016).

GO ID	GO term	Count	<i>p</i> value
GO:0008104	protein localization	227	1.61E-20
GO:0015031	protein transport	203	3.43E-20
GO:0045184	establishment of protein localization	203	1.05E-19
GO:0006414	translational elongation	52	4.95E-18
GO:0046907	intracellular transport	171	7.89E-16
GO:0070727	cellular macromolecule localization	115	7.64E-13
GO:0034613	cellular protein localization	114	1.08E-12
GO:0006886	intracellular protein transport	105	4.16E-12
GO:0006412	translation	90	9.97E-10
GO:0006605	protein targeting	64	8.39E-09

GO:0009057	macromolecule catabolic process	162	4.50E-07
GO:0044265	cellular macromolecule catabolic process	152	5.69E-07
GO:0043632	modification-dependent macromolecule catabolic process	124	1.46E-06
GO:0019941	modification-dependent protein catabolic process	124	1.46E-06
GO:0051603	proteolysis involved in cellular protein catabolic process	128	1.91E-06
GO:0044257	cellular protein catabolic process	128	2.49E-06
GO:0006793	phosphorus metabolic process	189	4.51E-06
GO:0006796	phosphate metabolic process	189	4.51E-06
GO:0030163	protein catabolic process	130	4.73E-06
GO:0016192	vesicle-mediated transport	122	4.93E-06
GO:0070647	protein modification by small protein conjugation or removal	44	2.10E-05
GO:0016310	phosphorylation	156	2.84E-05
GO:0033554	cellular response to stress	116	3.98E-05
GO:0032446	protein modification by small protein conjugation	36	1.59E-04
GO:0007264	small GTPase mediated signal transduction	68	1.68E-04
GO:0042274	ribosomal small subunit biogenesis	8	2.35E-04
GO:0048193	Golgi vesicle transport	35	3.03E-04
GO:0006612	protein targeting to membrane	14	3.48E-04
GO:0016567	protein ubiquitination	32	5.11E-04
GO:0016044	membrane organization	79	5.22E-04
GO:0033365	protein localization in organelle	37	5.91E-04
GO:0017038	protein import	34	6.56E-04
GO:0009144	purine nucleoside triphosphate metabolic process	32	8.06E-04
GO:0032869	cellular response to insulin stimulus	21	9.74E-04
GO:0006119	oxidative phosphorylation	27	0.001
GO:0019637	organophosphate metabolic process	46	0.001
GO:0006974	response to DNA damage stimulus	76	0.001
GO:0009116	nucleoside metabolic process	20	0.001
GO:0008286	insulin receptor signaling pathway	14	0.001
GO:0006644	phospholipid metabolic process	44	0.001
GO:0009141	nucleoside triphosphate metabolic process	33	0.001
GO:0022613	ribonucleoprotein complex biogenesis	42	0.001
GO:0032870	cellular response to hormone stimulus	33	0.002
GO:0008654	phospholipid biosynthetic process	27	0.002
GO:0045333	cellular respiration	26	0.002
GO:0016197	endosome transport	18	0.002
GO:0042981	regulation of apoptosis	144	0.003
GO:0043067	regulation of programmed cell death	145	0.003
GO:0051656	establishment of organelle localization	20	0.003
GO:0010941	regulation of cell death	145	0.003
GO:0007243	protein kinase cascade	73	0.003
GO:0006468	protein amino acid phosphorylation	121	0.004
GO:0043068	positive regulation of programmed cell death	83	0.004
GO:0006281	DNA repair	58	0.004
GO:0010942	positive regulation of cell death	83	0.005
GO:0051650	establishment of vesicle localization	12	0.005

GO:0032483	regulation of Rab protein signal transduction	15	0.005
GO:0032313	regulation of Rab GTPase activity	15	0.005
GO:0043065	positive regulation of apoptosis	82	0.005
GO:0018105	peptidyl-serine phosphorylation	11	0.005
GO:0034976	response to endoplasmic reticulum stress	12	0.006
GO:0006916	anti-apoptosis	44	0.006
GO:0009150	purine ribonucleotide metabolic process	32	0.006
GO:0007242	intracellular signaling cascade	211	0.006
GO:0032868	response to insulin stimulus	25	0.006
GO:0007259	JAK-STAT cascade	13	0.007
GO:0018209	peptidyl-serine modification	13	0.007
GO:0009205	purine ribonucleoside triphosphate metabolic process	28	0.007
GO:0009060	aerobic respiration	12	0.008
GO:0000209	protein polyubiquitination	10	0.008
GO:0000737	DNA catabolic process, endonucleolytic	10	0.008
GO:0016072	rRNA metabolic process	24	0.008
GO:0006508	proteolysis	179	0.008
GO:0006650	glycerophospholipid metabolic process	28	0.008
GO:0009199	ribonucleoside triphosphate metabolic process	28	0.008
GO:0016558	protein import into peroxisome matrix	6	0.008
GO:0006163	purine nucleotide metabolic process	40	0.008
GO:0006308	DNA catabolic process	17	0.009
GO:0051187	cofactor catabolic process	11	0.009
GO:0009259	ribonucleotide metabolic process	33	0.009
GO:0006364	rRNA processing	23	0.009

Table B.13.: Gene ontology (GO) classification of differentially expressed genes (3039) with expression profiles belonging to cluster IV (Wippermann et al. 2016).

GO ID	GO term	Count	<i>p</i> value
GO:0006396	RNA processing	196	7.05E-32
GO:0000278	mitotic cell cycle	132	2.16E-21
GO:0016071	mRNA metabolic process	132	2.16E-21
GO:0008380	RNA splicing	108	5.47E-20
GO:0007049	cell cycle	219	5.51E-20
GO:0006397	mRNA processing	117	6.97E-20
GO:0048285	organelle fission	90	7.57E-18
GO:0006259	DNA metabolic process	154	2.31E-17
GO:0022402	cell cycle process	165	2.20E-16
GO:0022403	cell cycle phase	131	3.48E-16
GO:0007067	mitosis	84	1.07E-15
GO:0000280	nuclear division	84	1.07E-15
GO:0000087	M phase of mitotic cell cycle	85	1.08E-15
GO:0000279	M phase	106	8.98E-14
GO:0051276	chromosome organization	138	6.31E-13
GO:0034470	ncRNA processing	68	9.61E-12

Appendix B. Supplementary material

GO:0051301	cell division	92	3.36E-11
GO:0006260	DNA replication	67	6.56E-11
GO:0007059	chromosome segregation	38	1.63E-10
GO:0000377	RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	57	1.88E-10
GO:0000398	nuclear mRNA splicing, via spliceosome	57	1.88E-10
GO:0000375	RNA splicing, via transesterification reactions	57	1.88E-10
GO:0022613	ribonucleoprotein complex biogenesis	62	1.06E-9
GO:0006350	transcription	425	1.29E-9
GO:0016568	chromatin modification	83	1.60E-9
GO:0006281	DNA repair	85	1.89E-9
GO:0043933	macromolecular complex subunit organization	171	2.74E-9
GO:0006357	regulation of transcription from RNA polymerase II promoter	174	3.22E-9
GO:0006974	response to DNA damage stimulus	103	4.19E-9
GO:0006913	nucleocytoplasmic transport	54	1.18E-8
GO:0034621	cellular macromolecular complex subunit organization	98	1.44E-8
GO:0051169	nuclear transport	54	1.93E-8
GO:0034660	ncRNA metabolic process	70	2.81E-8
GO:0010605	negative regulation of macromolecule metabolic process	171	3.42E-8
GO:0044265	cellular macromolecule catabolic process	169	4.01E-8
GO:0065003	macromolecular complex assembly	157	5.43E-8
GO:0051236	establishment of RNA localization	38	6.44E-8
GO:0050658	RNA transport	38	6.44E-8
GO:0050657	nucleic acid transport	38	6.44E-8
GO:0000819	sister chromatid segregation	21	7.98E-8
GO:0051028	mRNA transport	35	1.09E-7
GO:0006403	RNA localization	38	1.61E-7
GO:0015931	nucleobase, nucleoside, nucleotide and nucleic acid transport	41	2.06E-7
GO:0009057	macromolecule catabolic process	176	2.44E-7
GO:0000070	mitotic sister chromatid segregation	20	2.79E-7
GO:0010604	positive regulation of macromolecule metabolic process	188	6.25E-7
GO:0042254	ribosome biogenesis	42	7.20E-7
GO:0045449	regulation of transcription	491	8.39E-7
GO:0070271	protein complex biogenesis	121	9.73E-7
GO:0006461	protein complex assembly	121	9.73E-7
GO:0006325	chromatin organization	96	1.04E-6
GO:0051254	positive regulation of RNA metabolic process	116	1.19E-6
GO:0045893	positive regulation of transcription, DNA-dependent	115	1.35E-6
GO:0016569	covalent chromatin modification	42	1.86E-6
GO:0016570	histone modification	41	1.99E-6
GO:0051329	interphase of mitotic cell cycle	36	3.44E-6
GO:0010628	positive regulation of gene expression	133	3.61E-6
GO:0016072	rRNA metabolic process	34	4.88E-6
GO:0045941	positive regulation of transcription	129	5.37E-6
GO:0008033	tRNA processing	29	5.54E-6

GO:0034622	cellular macromolecular complex assembly	81	6.99E-6
GO:0051325	interphase	36	7.15E-6
GO:0010629	negative regulation of gene expression	117	7.49E-6
GO:0030163	protein catabolic process	139	8.11E-6
GO:0051173	positive regulation of nitrogen compound metabolic process	143	8.45E-6
GO:0006412	translation	83	1.00E-5
GO:0033554	cellular response to stress	128	1.05E-5
GO:0031328	positive regulation of cellular biosynthetic process	150	1.09E-5
GO:0051603	proteolysis involved in cellular protein catabolic process	134	1.24E-5
GO:0019941	modification-dependent protein catabolic process	129	1.33E-5
GO:0043632	modification-dependent macromolecule catabolic process	129	1.33E-5
GO:0017038	protein import	41	1.40E-5
GO:0006364	rRNA processing	32	1.48E-5
GO:0045935	positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	138	1.52E-5
GO:0044257	cellular protein catabolic process	134	1.60E-5
GO:0010557	positive regulation of macromolecule biosynthetic process	143	1.92E-5
GO:0006511	ubiquitin-dependent protein catabolic process	64	2.10E-5
GO:0009891	positive regulation of biosynthetic process	150	2.40E-5
GO:0006605	protein targeting	58	2.91E-5
GO:0010558	negative regulation of macromolecule biosynthetic process	122	3.37E-5
GO:0032774	RNA biosynthetic process	74	3.55E-5
GO:0006351	transcription, DNA-dependent	73	4.03E-5
GO:0045944	positive regulation of transcription from RNA polymerase II promoter	88	5.00E-5
GO:0051168	nuclear export	23	5.96E-5
GO:0016481	negative regulation of transcription	104	7.09E-5
GO:0022618	ribonucleoprotein complex assembly	25	7.49E-5
GO:0051436	negative regulation of ubiquitin-protein ligase activity during mitotic cell cycle	24	7.76E-5
GO:0033365	protein localization in organelle	42	9.42E-5
GO:0045934	negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	113	1.06E-4
GO:0006261	DNA-dependent DNA replication	22	1.08E-4
GO:0051170	nuclear import	29	1.19E-4
GO:0051172	negative regulation of nitrogen compound metabolic process	114	1.20E-4
GO:0043123	positive regulation of I-kappaB kinase/NF-kappaB cascade	31	1.23E-4
GO:0051444	negative regulation of ubiquitin-protein ligase activity	24	1.32E-4
GO:0051352	negative regulation of ligase activity	24	1.32E-4
GO:0010498	proteasomal protein catabolic process	32	1.37E-4
GO:0043161	proteasomal ubiquitin-dependent protein catabolic process	32	1.37E-4
GO:0030261	chromosome condensation	13	1.51E-4
GO:0031327	negative regulation of cellular biosynthetic process	121	1.63E-4
GO:0009890	negative regulation of biosynthetic process	123	1.78E-4

Appendix B. Supplementary material

GO:0043543	protein amino acid acylation	22	1.87E-4
GO:0031145	anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process	23	2.32E-4
GO:0006405	RNA export from nucleus	17	2.60E-4
GO:0022411	cellular component disassembly	20	2.63E-4
GO:0051248	negative regulation of protein metabolic process	49	2.75E-4
GO:0051439	regulation of ubiquitin-protein ligase activity during mitotic cell cycle	24	3.50E-4
GO:0000059	protein import into nucleus, docking	10	4.14E-4
GO:0006399	tRNA metabolic process	34	4.72E-4
GO:0040029	regulation of gene expression, epigenetic	25	4.99E-4
GO:0006606	protein import into nucleus	27	5.04E-4
GO:0006338	chromatin remodeling	20	5.85E-4
GO:0046907	intracellular transport	135	5.97E-4
GO:0051438	regulation of ubiquitin-protein ligase activity	25	6.15E-4
GO:0006473	protein amino acid acetylation	19	6.16E-4
GO:0016573	histone acetylation	18	6.41E-4
GO:0031397	negative regulation of protein ubiquitination	24	6.77E-4
GO:0051443	positive regulation of ubiquitin-protein ligase activity	23	7.43E-4
GO:0043122	regulation of I-kappaB kinase/NF-kappaB cascade	31	7.99E-4
GO:0006406	mRNA export from nucleus	14	8.74E-4
GO:0034504	protein localization in nucleus	28	9.57E-4
GO:0043623	cellular protein complex assembly	42	9.98E-4
GO:0000079	regulation of cyclin-dependent protein kinase activity	19	0.001
GO:0006366	transcription from RNA polymerase II promoter	56	0.001
GO:0045892	negative regulation of transcription, DNA-dependent	79	0.001
GO:0032984	macromolecular complex disassembly	12	0.001
GO:0034623	cellular macromolecular complex disassembly	12	0.001
GO:0051340	regulation of ligase activity	25	0.001
GO:0051253	negative regulation of RNA metabolic process	80	0.001
GO:0000122	negative regulation of transcription from RNA polymerase II promoter	62	0.001
GO:0051437	positive regulation of ubiquitin-protein ligase activity during mitotic cell cycle	22	0.001
GO:0051351	positive regulation of ligase activity	23	0.001
GO:0032269	negative regulation of cellular protein metabolic process	45	0.001
GO:0007010	cytoskeleton organization	93	0.001
GO:0045184	establishment of protein localization	152	0.001
GO:0000082	G1/S transition of mitotic cell cycle	19	0.002
GO:0007076	mitotic chromosome condensation	8	0.002
GO:0006401	RNA catabolic process	21	0.002
GO:0042176	regulation of protein catabolic process	18	0.002
GO:0009451	RNA modification	17	0.002
GO:0031398	positive regulation of protein ubiquitination	25	0.002
GO:0042177	negative regulation of protein catabolic process	9	0.002
GO:0015031	protein transport	149	0.003
GO:0022614	membrane to membrane docking	5	0.003

GO:0051101	regulation of DNA binding	32	0.003
GO:0006289	nucleotide-excision repair	18	0.003
GO:0000226	microtubule cytoskeleton organization	37	0.004
GO:0006730	one-carbon metabolic process	30	0.004
GO:0051252	regulation of RNA metabolic process	325	0.004
GO:0010608	posttranscriptional regulation of gene expression	49	0.004
GO:0016458	gene silencing	18	0.004
GO:0032479	regulation of type I interferon production	8	0.005
GO:0031396	regulation of protein ubiquitination	27	0.005
GO:0006302	double-strand break repair	19	0.006
GO:0009156	ribonucleoside monophosphate biosynthetic process	10	0.006
GO:0006809	nitric oxide biosynthetic process	7	0.006
GO:0007017	microtubule-based process	56	0.006
GO:0006417	regulation of translation	34	0.007
GO:0022616	DNA strand elongation	5	0.007
GO:0051098	regulation of binding	37	0.007
GO:0006270	DNA replication initiation	8	0.007
GO:0006413	translational initiation	15	0.007
GO:0051099	positive regulation of binding	22	0.008
GO:0009411	response to UV	18	0.008
GO:0051090	regulation of transcription factor activity	27	0.008
GO:0045089	positive regulation of innate immune response	15	0.009
GO:0031400	negative regulation of protein modification process	30	0.009
GO:0006308	DNA catabolic process	18	0.009
GO:0006886	intracellular protein transport	77	0.009
GO:0046209	nitric oxide metabolic process	7	0.009
GO:0031050	dsRNA fragmentation	7	0.009
GO:0031047	gene silencing by RNA	12	0.009

Table B.14.: Gene ontology (GO) classification of genes both hypomethylated and hypermethylated upon butyrate addition (2890; Wippermann et al. 2016).

GO ID	GO term	Count	<i>p</i> value
GO:0006414	translational elongation	36	1.55E10-9
GO:0009952	anterior/posterior pattern formation	38	1.49E10-6
GO:0007389	pattern specification process	58	5.48E10-6
GO:0048706	embryonic skeletal system development	24	1.67E10-5
GO:0003002	regionalization	45	1.96E10-5
GO:0001501	skeletal system development	62	8.77E10-5
GO:0051253	negative regulation of RNA metabolic process	67	1.93E10-4
GO:0006364	rRNA processing	24	3.36E10-4
GO:0045892	negative regulation of transcription, DNA-dependent	65	3.59E10-4
GO:0042254	ribosome biogenesis	29	3.77E10-4
GO:0016072	rRNA metabolic process	24	6.44E10-4
GO:0048704	embryonic skeletal system morphogenesis	17	6.79E10-4
GO:0045926	negative regulation of growth	26	8.83E10-4

GO:0007155	cell adhesion	111	9.24E10-4
GO:0022610	biological adhesion	111	9.75E10-4
GO:0022613	ribonucleoprotein complex biogenesis	37	0.001
GO:0051270	regulation of cell motion	39	0.001
GO:0015031	protein transport	118	0.002
GO:0010605	negative regulation of macromolecule metabolic process	114	0.002
GO:0008361	regulation of cell size	40	0.002
GO:0034470	ncRNA processing	37	0.002
GO:0045184	establishment of protein localization	118	0.002
GO:0030308	negative regulation of cell growth	22	0.002
GO:0001558	regulation of cell growth	38	0.002
GO:0031056	regulation of histone modification	7	0.002
GO:0045792	negative regulation of cell size	23	0.002
GO:0030334	regulation of cell migration	34	0.002
GO:0032535	regulation of cellular component size	49	0.003
GO:0048705	skeletal system morphogenesis	25	0.003
GO:0016481	negative regulation of transcription	75	0.003
GO:0010629	negative regulation of gene expression	81	0.003
GO:0034101	erythrocyte homeostasis	14	0.004
GO:0035051	cardiac cell differentiation	10	0.004
GO:0043433	negative regulation of transcription factor activity	13	0.005
GO:0045934	negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	81	0.005
GO:0031057	negative regulation of histone modification	5	0.005
GO:0031058	positive regulation of histone modification	5	0.005
GO:0008104	protein localization	130	0.005
GO:0042274	ribosomal small subunit biogenesis	6	0.006
GO:0030097	hemopoiesis	42	0.007
GO:0051172	negative regulation of nitrogen compound metabolic process	81	0.007
GO:0051130	positive regulation of cellular component organization	34	0.007
GO:0034660	ncRNA metabolic process	41	0.008
GO:0006986	response to unfolded protein	17	0.008
GO:0051318	G1 phase	8	0.008
GO:0051090	regulation of transcription factor activity	22	0.008
GO:0006412	translation	55	0.008
GO:0048534	hemopoietic or lymphoid organ development	45	0.009
GO:0001667	ameboidal cell migration	11	0.009

Table B.15.: Gene ontology (GO) classification of hypomethylated genes upon butyrate addition (3886; Wippermann et al. 2016).

GO ID	GO term	Count	<i>p</i> value
GO:0022403	cell cycle phase	101	1.09E-05
GO:0000279	M phase	84	1.12E-05
GO:0022402	cell cycle process	129	2.01E-05

GO:0000087	M phase of mitotic cell cycle	61	2.79E-05
GO:0006357	regulation of transcription from RNA polymerase II promoter	157	5.59E-05
GO:0007067	mitosis	59	6.15E-05
GO:0000280	nuclear division	59	6.15E-05
GO:0048285	organelle fission	60	1.07E-04
GO:0000278	mitotic cell cycle	88	1.12E-04
GO:0007242	intracellular signaling cascade	250	1.29E-04
GO:0006396	RNA processing	121	1.55E-04
GO:0060627	regulation of vesicle-mediated transport	31	1.58E-04
GO:0010941	regulation of cell death	170	1.78E-04
GO:0043067	regulation of programmed cell death	169	2.08E-04
GO:0032891	negative regulation of organic acid transport	8	2.11E-04
GO:0006397	mRNA processing	77	2.36E-04
GO:0007049	cell cycle	162	2.51E-04
GO:0042981	regulation of apoptosis	167	2.56E-04
GO:0016071	mRNA metabolic process	86	3.03E-04
GO:0044265	cellular macromolecule catabolic process	152	3.19E-04
GO:0008380	RNA splicing	69	3.51E-04
GO:0043632	modification-dependent macromolecule catabolic process	124	3.56E-04
GO:0019941	modification-dependent protein catabolic process	124	3.56E-04
GO:0060191	regulation of lipase activity	28	3.75E-04
GO:0060193	positive regulation of lipase activity	25	4.72E-04
GO:0009057	macromolecule catabolic process	161	4.76E-04
GO:0051051	negative regulation of transport	38	5.62E-04
GO:0016192	vesicle-mediated transport	122	8.80E-04
GO:0051603	proteolysis involved in cellular protein catabolic process	126	9.97E-04

Table B.16.: Gene ontology (GO) classification of hypermethylated genes upon butyrate treatment (1737; Wippermann et al. 2016).

GO ID	GO term	Count	<i>p</i> value
GO:0008104	protein localization	110	2.89E-8
GO:0045184	establishment of protein localization	97	1.27E-7
GO:0015031	protein transport	95	3.04E-7
GO:0008380	RNA splicing	43	9.63E-6
GO:0006397	mRNA processing	45	4.03E-5
GO:0046907	intracellular transport	77	4.25E-5
GO:0006396	RNA processing	66	6.89E-5
GO:0016071	mRNA metabolic process	49	7.58E-5
GO:0051028	mRNA transport	18	1.65E-4
GO:0050657	nucleic acid transport	19	2.12E-4
GO:0051236	establishment of RNA localization	19	2.12E-4
GO:0050658	RNA transport	19	2.12E-4
GO:0006403	RNA localization	19	3.15E-4
GO:0015931	nucleobase, nucleoside, nucleotide and nucleic acid transport	20	5.32E-4

GO:0048193	Golgi vesicle transport	22	5.39E-4
GO:0070727	cellular macromolecule localization	50	5.76E-4
GO:0000398	nuclear mRNA splicing, via spliceosome	24	7.61E-4
GO:0000377	RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	24	7.61E-4
GO:0000375	RNA splicing, via transesterification reactions	24	7.61E-4
GO:0034613	cellular protein localization	49	8.70E-4

Table B.17.: Comparison of studies that aimed to analyze butyrate-induced gene expression changes by microarrays in cultured cells (Klausing, Krämer, and Noll 2011; Kantardjieff et al. 2010; Li and Li 2006a). Fold changes of expression of cell cycle-related genes were, if necessary, converted to enable comparability. n.d. = significant change in gene expression not detected in the respective microarray experiment.

Transcript	Butyrate-induced gene expression fold changes measured by			
	Li and Li 2006	Klausing et al. 2011	Kantardjieff et al. 2010	Present study
ATR	n.d.	0.53	0.06	0.77
AURKA	0.17	n.d.	n.d.	0.46
AURKAIP1	0.19	n.d.	n.d.	n.d.
AURKB	0.18	n.d.	n.d.	0.50
BUB1	0.10	n.d.	0.18	n.d.
BUB1B	0.06	n.d.	n.d.	0.50
CBX5	0.10	n.d.	n.d.	0.50
CCAR1	0.17	n.d.	n.d.	0.70
CCNA2	0.17	n.d.	0.02	0.39
CCNA2	0.17	n.d.	n.d.	0.39
CCNB1	n.d.	n.d.	0.01	0.43
CCNB1	0.06	n.d.	n.d.	0.43
CCNB2	0.22	n.d.	0.02	0.48
CCNB2	0.22	n.d.	n.d.	0.48
CCNG1	0.07	n.d.	n.d.	n.d.
CDC20	n.d.	n.d.	0.01	0.35
CDC20	0.10	n.d.	n.d.	0.35
CDC25A	0.11	n.d.	n.d.	0.59
CDC2A	n.d.	n.d.	0.05	n.d.
CDC451	0.09	n.d.	n.d.	0.57
CDCA2	0.08	n.d.	n.d.	0.58
CDCA8	0.17	n.d.	n.d.	0.38
CDK1	0.12	n.d.	n.d.	n.d.
CDK11B	0.04	n.d.	n.d.	n.d.
CDK4	0.15	n.d.	n.d.	n.d.
CDKN2C	0.22	n.d.	n.d.	n.d.
CENPA	0.09	n.d.	n.d.	0.32
CENPF	0.06	n.d.	n.d.	0.69
CHAF1A	0.20	n.d.	n.d.	0.81
CHD3	0.14	n.d.	n.d.	0.54
CHEK1	0.16	n.d.	n.d.	0.79

CHEK2	n.d.	1.43	0.04	n.d.
CKS1B	0.09	n.d.	n.d.	0.65
CKS2	0.02	n.d.	n.d.	0.72
CYCD	n.d.	2.49	n.d.	n.d.
DKC1	0.11	n.d.	n.d.	0.40
DLGAP5	0.22	n.d.	n.d.	0.52
DSCC1	0.06	n.d.	n.d.	0.55
DTYMK	0.15	n.d.	n.d.	0.68
E2F4	0.11	n.d.	n.d.	0.66
E2F6	n.d.	n.d.	2.60	0.63
ESCO2	0.13	n.d.	n.d.	n.d.
ESPL1	0.08	n.d.	n.d.	0.63
FEN1	0.18	n.d.	n.d.	0.63
GTF3C2	0.24	n.d.	n.d.	0.71
GTSE1	0.23	n.d.	n.d.	0.57
H2AFZ	0.08	n.d.	n.d.	0.54
HDAC7A	n.d.	0.71	0.18	0.78
HDAC8	0.15	n.d.	n.d.	n.d.
HELLS	0.13	n.d.	n.d.	0.60
KIF11	0.23	n.d.	n.d.	n.d.
KIF14	0.20	n.d.	n.d.	0.66
KIF20A	0.12	n.d.	n.d.	0.52
KIF20B	0.13	n.d.	n.d.	0.51
KIF22	0.09	n.d.	n.d.	0.39
KIF23	0.07	n.d.	n.d.	0.51
KIF2c	0.19	n.d.	n.d.	0.65
LMNB1	0.05	n.d.	n.d.	0.67
MAD2L1	0.07	n.d.	n.d.	0.72
MAP126	0.04	n.d.	n.d.	n.d.
MCM3	0.10	0.86	0.05	0.41
MCM4	0.14	n.d.	n.d.	0.50
MCM5	0.17	0.27	0.01	0.37
MCM6	0.09	n.d.	n.d.	0.65
MCM7	n.d.	n.d.	0.11	0.51
MND1	0.07	n.d.	n.d.	n.d.
MSH3	0.24	n.d.	n.d.	n.d.
MSH6	0.21	n.d.	n.d.	0.64
MXI1	12.55	n.d.	n.d.	2.26
MYC	n.d.	0.29	n.d.	0.82
NCAPG	0.06	n.d.	n.d.	0.52
NCL	0.10	n.d.	n.d.	0.66
NDC80	0.15	n.d.	n.d.	0.63
NEK2	0.13	n.d.	n.d.	0.57
NUCKS1	0.19	n.d.	n.d.	0.53
NUSAP1	0.08	n.d.	n.d.	0.41
ORC1L	0.23	0.27	0.07	0.52
PBK	0.15	n.d.	n.d.	0.50

PCNA	0.09	n.d.	n.d.	0.44
PES1	0.17	n.d.	n.d.	n.d.
PKMYT1	n.d.	0.41	0.01	0.58
PLK1	n.d.	n.d.	0.05	0.34
PRC1	0.07	n.d.	n.d.	0.54
PRKDC	n.d.	n.d.	1.10	n.d.
PSMB8	0.14	n.d.	n.d.	0.35
RAD51	0.11	n.d.	n.d.	n.d.
RAD51AP1	0.21	n.d.	n.d.	0.70
REXO4	0.21	n.d.	n.d.	0.62
SAFB2	0.20	n.d.	n.d.	n.d.
SMC4	0.13	n.d.	n.d.	0.46
SPC25	0.14	n.d.	n.d.	0.59
TCERG1	0.09	n.d.	n.d.	0.61
UEB2C	0.14	n.d.	n.d.	0.39
UBE2K	0.25	n.d.	n.d.	0.59
WEE1	0.16	n.d.	n.d.	n.d.
YWHAG	n.d.	n.d.	0.18	n.d.

Table B.18.: Gene Ontology analysis (p values ≤ 0.01 .) of genes affected by butyrate-mediated differential hypomethylation of associated CpG islands 24 hours upon butyrate addition (Wippermann et al. 2013). Genes detected without treatment with 3 mM butyrate ('0 h' sampling point) were removed from the dataset before ontology analysis.

GO ID	GO term	Count	p value
GO:0010605	negative regulation of macromolecule metabolic process	70	4.98E-06
GO:0010629	negative regulation of gene expression	53	6.06E-06
GO:0045934	negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	53	9.53E-06
GO:0016481	negative regulation of transcription	49	9.66E-06
GO:0051172	negative regulation of nitrogen compound metabolic process	53	1.38E-05
GO:0016310	phosphorylation	73	1.40E-05
GO:0006468	protein amino acid phosphorylation	63	2.12E-05
GO:0010558	negative regulation of macromolecule biosynthetic process	54	2.83E-05
GO:0031327	negative regulation of cellular biosynthetic process	55	2.92E-05
GO:0006793	phosphorus metabolic process	83	3.87E-05
GO:0006796	phosphate metabolic process	83	3.87E-05
GO:0009890	negative regulation of biosynthetic process	55	5.17E-05
GO:0045892	negative regulation of transcription, DNA-dependent	39	5.20E-05
GO:0051253	negative regulation of RNA metabolic process	39	7.45E-05
GO:0007242	intracellular signaling cascade	100	8.44E-05
GO:0051056	regulation of small GTPase mediated signal transduction	30	1.07E-04
GO:0000904	cell morphogenesis involved in differentiation	29	1.47E-04
GO:0048667	cell morphogenesis involved in neuron differentiation	26	1.73E-04

GO:0030030	cell projection organization	38	2.22E-04
GO:0048666	neuron development	35	4.12E-04
GO:0030182	neuron differentiation	42	4.68E-04
GO:0002366	leukocyte activation during immune response	9	5.69E-04
GO:0002263	cell activation during immune response	9	5.69E-04
GO:0000122	negative regulation of transcription from RNA polymerase II promoter	29	6.15E-04
GO:0051051	negative regulation of transport	18	0.001
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	25	0.001
GO:0030520	estrogen receptor signaling pathway	5	0.002
GO:0002285	lymphocyte activation during immune response	6	0.002
GO:0046578	regulation of Ras protein signal transduction	23	0.002
GO:0032583	regulation of gene-specific transcription	17	0.002
GO:0035023	regulation of Rho protein signal transduction	14	0.003
GO:0048812	neuron projection morphogenesis	23	0.003
GO:0045851	pH reduction	5	0.003
GO:0032989	cellular component morphogenesis	36	0.003
GO:0031175	neuron projection development	26	0.003
GO:0007167	enzyme linked receptor protein signaling pathway	32	0.004
GO:0032582	negative regulation of gene-specific transcription	9	0.004
GO:0007409	axonogenesis	21	0.004
GO:0002520	immune system development	27	0.004
GO:0033500	carbohydrate homeostasis	9	0.005
GO:0042593	glucose homeostasis	9	0.005
GO:0010551	regulation of specific transcription from RNA polymerase II promoter	13	0.005
GO:0006357	regulation of transcription from RNA polymerase II promoter	57	0.005
GO:0045860	positive regulation of protein kinase activity	23	0.005
GO:0021904	dorsal/ventral neural tube patterning	5	0.006
GO:0007389	pattern specification process	26	0.006
GO:0010553	negative regulation of specific transcription from RNA polymerase II promoter	8	0.006
GO:0019637	organophosphate metabolic process	21	0.006
GO:0007268	synaptic transmission	28	0.006
GO:0000902	cell morphogenesis	32	0.006
GO:0030004	cellular monovalent inorganic cation homeostasis	6	0.007
GO:0060122	inner ear receptor stereocilium organization	4	0.007
GO:0033674	positive regulation of kinase activity	23	0.007
GO:0006644	phospholipid metabolic process	20	0.008
GO:0048534	hemopoietic or lymphoid organ development	25	0.008
GO:0035295	tube development	22	0.009
GO:0032147	activation of protein kinase activity	14	0.009
GO:0019226	transmission of nerve impulse	31	0.009
GO:0043406	positive regulation of MAP kinase activity	13	0.009
GO:0042592	homeostatic process	57	0.010

Table B.19.: Gene Ontology analysis (p values ≤ 0.01 .) of genes affected by butyrate-mediated differential hypermethylation of associated CpG islands 24 hours upon butyrate addition (Wippermann et al. 2013). Genes detected without treatment with 3 mM butyrate ('0 h' sampling point) were removed from the dataset before ontology analysis.

GO ID	GO term	Count	p value
GO:0051348	negative regulation of transferase activity	10	9.97E-05
GO:0006469	negative regulation of protein kinase activity	9	2.77E-04
GO:0033673	negative regulation of kinase activity	9	3.50E-04
GO:0006974	response to DNA damage stimulus	19	3.53E-04
GO:0006281	DNA repair	16	4.33E-04
GO:0065003	macromolecular complex assembly	27	5.25E-04
GO:0043933	macromolecular complex subunit organization	28	6.24E-04
GO:0010033	response to organic substance	28	7.91E-04
GO:0006461	protein complex assembly	22	8.43E-04
GO:0070271	protein complex biogenesis	22	8.43E-04
GO:0051338	regulation of transferase activity	18	9.45E-04
GO:0045859	regulation of protein kinase activity	17	0.001
GO:0033554	cellular response to stress	23	0.002
GO:0043549	regulation of kinase activity	17	0.002
GO:0010605	negative regulation of macromolecule metabolic process	27	0.002
GO:0043086	negative regulation of catalytic activity	14	0.003
GO:0016192	vesicle-mediated transport	22	0.004
GO:0042325	regulation of phosphorylation	19	0.004
GO:0043069	negative regulation of programmed cell death	16	0.004
GO:0060548	negative regulation of cell death	16	0.004
GO:0006259	DNA metabolic process	20	0.005
GO:0043407	negative regulation of MAP kinase activity	5	0.005
GO:0044092	negative regulation of molecular function	15	0.006
GO:0051172	negative regulation of nitrogen compound metabolic process	20	0.006
GO:0019220	regulation of phosphate metabolic process	19	0.006
GO:0051174	regulation of phosphorus metabolic process	19	0.006
GO:0051235	maintenance of location	6	0.008
GO:0009890	negative regulation of biosynthetic process	21	0.008
GO:0043066	negative regulation of apoptosis	15	0.009
GO:0010629	negative regulation of gene expression	19	0.009

List of Figures

1.1. Chromatin organization and histone tail modifications.	6
1.2. Regulatory epigenetic-miRNA circuit.	8
1.3. Regulatory epigenetic processes during development (Reik 2007).	10
1.4. Schematic representation of the domain structure of mammalian DNA methyltransferases.	12
1.5. Structural overview of human DNMT1 bound to S-adenosyl homocysteine (AdoHcy) and schematic representation of the autoinhibitory mechanism of DNMT1.	14
1.6. DNMT3A/3L interaction.	15
1.7. Pathway of dynamic cytosine methylation and proposed functions of 5hmC (modified from Kohli and Zhang 2013).	16
1.8. Mammalian DNA methylation landscape (modified from Schübeler 2015).	18
1.9. Proposed interactions between transcription factors and DNA methylation at CpG-poor genomic regions (modified from Schübeler 2015).	19
1.10. Conversion of cytosine to uracil by bisulfite-treatment (Hayatsu 2008).	21
1.11. Principle of Oxidative Bisulfite Sequencing (oxBS-Seq) and Reduced Bisulfite Sequencing (redBS-Seq; Booth et al. 2014).	23
1.12. Chemical structure of butyrate.	27
1.13. Differences in cellular metabolism between normal and cancer cells result in opposing effects upon butyrate addition (figure based on Donohoe et al. 2013).	28
3.1. DNA ladders (Thermo Scientific) used for determination of DNA fragment sizes.	38
3.2. Vector map of <i>pJet1.2/blunt</i> cloning vector (Thermo Scientific).	45
3.3. Vector maps of <i>pMCS-Green Renilla Luc</i> , <i>pCMV-Green Renilla Luc</i> and <i>pTK-Red Firefly Luc</i> (ThermoFisher Scientific).	45
3.4. Vector map of <i>pLVX-shRNA1</i> (Clontech).	47
3.5. Exemplary Sanger sequencing result.	49
3.6. Principle of Combined Bisulfite Restriction Analysis (COBRA).	50
3.7. <i>PageRuler™ Prestained Protein Ladder</i> (ThermoFisher Scientific). kDa = kilodalton.	55
3.8. Bioluminescent reactions employed for the Dual Luciferase® Reporter Assay System (Promega).	57
4.1. Cell growth, IgG production, glucose and lactate concentrations of CHO DP-12 cell cultures (Wippermann et al. 2016).	63
4.2. Schematic representation of the work flow of integrating whole-genome bisulfite sequencing and CHO DNA-microarray data.	64
4.3. Bisulfite sequencing results of the CGIs of <i>Beta-actin (Actb)</i> and <i>Branched chain amino acid aminotransferase cytosolic-like (Bcat1)</i>	65

4.4. Percent base plot of <i>MiSeq</i> sequencing of pooled libraries from whole-genome bisulfite sequencing.	66
4.5. Comparison of mapping tool performance.	68
4.6. Control of methylation data tracks of methylation controls <i>Actb</i> and <i>Bcat1</i> as well as CHO housekeeping genes in the <i>GenDBE</i> genome browser.	70
4.7. Global distribution of DNA methylation levels in reference CHO DP-12 cultures.	73
4.8. Exemplary display of SNP-corrected CHO K1 scaffold JH000005 (Wippermann et al. 2015).	74
4.9. PMD count per Mb for CHO DP-12 PMDs based on Illumina read coverage assigned to Chinese hamster chromosomes (Wippermann et al. 2015).	75
4.10. Characterization of CHO DP-12 partially methylated domains (PMDs) and highly methylated domains (HMDs; Wippermann et al. 2015).	76
4.11. Analysis of differential gene expression data 6 to 36 hours after butyrate addition (modified from Wippermann et al. 2016).	86
4.12. <i>Generally Applicable Gene-set Enrichment (GAGE)</i> analysis of significantly differentially expressed genes after butyrate addition (Wippermann et al. 2016).	88
4.13. Visualization of gene expression time course after butyrate addition for <i>KEGG</i> pathway Cell Cycle.	90
4.14. Visualization of gene expression time course after butyrate addition for <i>KEGG</i> pathway Apoptosis.	92
4.15. Visualization of gene expression time course after butyrate addition for <i>KEGG</i> pathway Citrate Cycle (Wippermann et al. 2016).	94
4.16. Visualization of gene expression time course after butyrate addition for <i>KEGG</i> pathway Spliceosome.	95
4.17. Visualization of gene expression time course after butyrate addition for <i>KEGG</i> pathway N-Glycan Biosynthesis.	97
4.18. Visualization of gene expression time course after butyrate addition for <i>KEGG</i> pathway Ribosome Biogenesis.	99
4.19. Percent methylation of CpG dinucleotides and frequencies of mean methylation levels of genomic features.	102
4.20. DNA methylation changes in CHO DP-12 cultures upon butyrate addition (Wippermann et al. 2016).	104
4.21. Exemplary display of 5â€™-associated differentially methylated regions within CpG islands (Wippermann et al. 2016).	105
4.22. Screenshots of relatively hypomethylated (A) and hypermethylated (B) DMRs as visualized in the <i>GenDBE</i> genome browser.	106
4.23. Association of differential DNA methylation types and gene expression	107
4.24. Integration of DNA methylation and gene expression data.	109
4.25. Putative transcription factor binding-site motifs <i>de novo</i> discovered in hypermethylated and hypomethylated DMRs were compared to human and mouse transcription factor databases (Wippermann et al. 2016).	110
4.26. Schematic representation of CHO CpG island microarray probe distribution relative to genomic features.	116
4.27. Schematic representation of the experimental procedure of CpG island microarray analysis.	117
4.28. Significantly differentially methylated genes upon butyrate treatment.	118
4.29. Gene ontology analysis of differentially methylated genes upon butyrate treatment.	119
4.30. Comparison of genes detected to be differentially methylated and/or differentially expressed by different techniques.	121
4.31. Combined Bisulfite Restriction Analysis of the genes <i>Segment polarity dishevelled homolog DVL-1 (Dvl1)</i> , <i>Frizzled-7 (Fzd7)</i> and <i>Myc proto-oncogene (Myc)</i> .	122

4.32. MA plots of quality control hybridization experiments for CpG island microarray analysis.	125
4.33. Confirmatory analysis of <i>Dnmt3a</i> methylation and expression changes upon butyrate addition.	130
4.34. Knockdown of DNMT3A by RNA interference.	131
B.1. Coverage of recombinant CHO DP-12 vector DNA by reads that could not be mapped to the SNP-corrected CHO K1 genome.	178
B.2. Mapping of reads from libraries of native CHO DP-12 DNA to both the recombinant vector sequence and the SNP corrected CHO K1 genome.	179
B.3. Attempt to verify a predicted transgene integration site in the CHO DP-12 genome by long-range PCR.	180
B.4. qPCR analysis of putative effects of butyrate treatment on the expression of pri-miR199a and miR-199a target genes.	181
B.5. Core promoter sequence motifs (Smales et al. 2004).	183
B.6. Principle of dual luciferase reporter assays.	184
B.7. Analysis of expression strength of endogenous CHO promoters.	185
B.8. Confirmation of <i>Actb</i> and <i>Bcat1</i> promoter regions as methylation controls in CHO-XL99 by bisulfite sequencing	187
B.9. VCDs, viabilities and mAB expression levels of CHO-XL99 clones.	188
B.10. Exemplary images of artifacts in hybridized CpG island microarrays after scanning and applying Agilent's <i>Feature Extraction</i> software.	189
B.11. MA plots for the comparison of good and bad producing CHO-XL99 clones to the CHO K1 host cell line.	190
B.12. Significantly differentially methylated genes in good and bad producing CHO-XL99 clones with and without selection pressure compared to the CHO K1 host cell line.	191
B.13. Agarose gel for integrity control of genomic DNA extracted from reference and butyrate treated CHO DP-12 cells for whole-genome bisulfite sequencing.	200
B.14. Quality of libraries for whole genome bisulfite sequencing was checked using <i>High Sensitivity DNA Chips</i> on a <i>Bioanalyzer</i>	200
B.15. Evaluation of an effect of shearing on library generation from bisulfite treated DNA.	201
B.16. Comparison of mapping results for exemplary CpG islands obtained with the <i>MethPipe</i> mapping tool and <i>BSseeker2</i>	201
B.17. Methylation data tracks were loaded into the <i>GenDBE</i> genome browser (Rupp et al. 2014) to check for consistency of whole-genome bisulfite sequencing results between the sampling points.	202
B.18. Denaturing agarose gels for control of integrity of RNAs used for gene expression microarray experiments.	202
B.19. Artificial images of spatial signal distributions of gene expression microarrays.	203
B.20. Boxplots of signal intensities for gene expression microarrays before and after background correction and normalization.	204
B.21. Agarose gels for the control of sheared DNA samples for CpG island microarray analysis.	205
B.22. Exemplary display of quality plots from CHO CpG island microarray data.	206
B.23. Display of PMD distribution of SNP-corrected CHO K1 scaffolds JH000001 to JH000009 (Wippermann et al. 2015).	207
B.24. Percentages of relatively hypomethylated and relatively hypermethylated DMRs for all sampling points assigned to the genomic features upstream region, 5' end, 3', intergenic and intragenic region (Wippermann et al. 2016).	208
B.25. Screenshots of relatively hypomethylated DMRs as visualized in the <i>GenDBE</i> genome browser (Rupp et al. 2014). All regions are represented in the same scale.	209

B.26. Screenshots of relatively hypermethylated DMRs as visualized in the <i>GenDBE</i> genome browser (Rupp et al. 2014). All regions are represented in the same scale.	210
B.27. Visualisation of gene expression time course after butyrate addition for <i>KEGG</i> pathway Glycolysis.	211
B.28. Visualisation of gene expression time course after butyrate addition for <i>KEGG</i> pathway RNA Transport.	212
B.29. Visualisation of gene expression time course after butyrate addition for <i>KEGG</i> pathway Ribosome.	213
B.30. Visualisation of gene expression time course after butyrate addition for <i>KEGG</i> pathway Regulation of Actin Cytoskeleton.	214
B.31. Visualisation of gene expression time course after butyrate addition for <i>KEGG</i> pathway Regulation of Wnt Signaling.	215
B.32. Visualisation of gene expression time course after butyrate addition for <i>KEGG</i> pathway MAPK Signaling.	216
B.33. Visualisation of gene expression time course after butyrate addition for <i>KEGG</i> pathway Regulation of Calcium Signaling.	217
B.34. Visualisation of gene expression time course after butyrate addition for <i>KEGG</i> pathway Regulation of p53 Signaling.	218
B.35. Visualisation of gene expression time course after butyrate addition for <i>KEGG</i> pathway Purine Metabolism.	219
B.36. Visualisation of gene expression time course after butyrate addition for <i>KEGG</i> pathway Pyrimidine Metabolism.	220
B.37. Cell growth, IgG production as well as glucose and lactate concentrations of butyrate treated and control CHO DP-12 cell cultures (Wippermann 2012, Wippermann et al. 2013).	221
B.38. MA plots for samples taken 0 h, 24 h and 48 h upon butyrate treatment (Wippermann 2012, Wippermann et al. 2013).	222
B.39. Exemplary westernblot for measurement of DNMT3A expression in reference and butyrate treated CHO DP-12 cell cultures.	223

List of Tables

3.1. Primers used for sequencing of DNA fragments inserted into <i>pJet1.2/blunt</i> cloning vector, pMCS-Green <i>Renilla</i> Luc and pLVX-shRNA1.	40
3.2. Cyling program used for amplification of regions of interest from bisulfite treated genomic DNA with <i>EpiMark[®] Hot Start Taq DNA Polymerase</i> . Annealing temperatures specific for each pair of primers are listed in Table 3.3.	40
3.3. Primers used for amplification of regions of interest from bisulfite treated CHO DNA. T _A = annealing temperature.	41
3.4. Cyling program used for amplification of promoter regions from CHO genomic DNA with <i>Q5[®] High-Fidelity DNA Polymerase</i> . Annealing temperatures specific for each pair of primers are listed in Table 3.5.	41
3.5. Primers used for amplification of endogenous CHO promoter regions from CHO K1 genomic DNA. <i>EcoRI</i> and <i>BamHI</i> restriction endonuclease recognition sites are shown in lowercase. T _A = annealing temperature. NA = not assigned to any annotated gene.	42
3.6. Cyling program for quantitative Real-Time PCR.	42
3.7. Primers used for quantitative Real-Time PCR. gDNA = genomic DNA, cDNA = complementary DNA.	43
3.8. Oligonucleotide sequences used for generation of promoter-reporter constructs by direct cloning into <i>pMCS-Green Renilla Luc</i> vector.	46
3.9. Program for annealing of oligonucleotides in a <i>Thermocycler</i>	46
3.10. shRNA sequences used for RNAi-mediated knockdown of DNMT3A. Indicated are exons of the DNMT3A gene targeted by the three siRNAs as well as the shRNA oligonucleotide sequences (uppercase) including added restriction enzyme sites (lowercase). sc = scrambled control.	48
3.11. Barcoding of individual libraries for DNA methylome analysis by next generation sequencing.	53
3.12. Composition of Illumina <i>HiSeq</i> sequencing reactions for flow cells 1 and 2.	54
4.1. Numerical summary of raw sequence output and calculated genome coverage from Illumina <i>HiSeq</i> sequencing of libraries from native and bisulfite-treated (BS) CHO DP-12 DNA before further data processing.	67
4.2. Numerical summary of whole-genome bisulfite sequencing results after mapping to the SNP-corrected CHO K1 genome.	69
4.3. Gene ontology (GO) classification of total CHO DP-12 PMD-genes (4265; Wippermann et al. 2015).	77
4.4. Separate gene ontology (GO) classifications for highly expressed and weakly or not expressed CHO DP-12 PMD-genes (Wippermann et al. 2015).	78
4.5. Top 10 up- and downregulated genes 36 hours post butyrate addition.	84

4.6. Comparison of studies that analyzed gene expression in cultured cells under butyrate treatment by microarrays.	101
4.7. Numbers of relatively hypomethylated and relatively hypermethylated DMRs with ≥ 1 significantly differentially methylated CpG 6, 12, 24 and 36 hours after butyrate addition in comparison to the reference (Wippermann et al. 2016).	103
4.8. Significant differences (p adjusted ≤ 0.05) between technical control hybridizations.	124
B.1. Key features of 5 endogenous CHO promoters examined by dual luciferase assays regarding the potential of recombinant gene expression.	183
B.2. Gene Ontology analysis of genes associated with differentially methylated CpG islands in replicate cultures of the good producing CHO-XL99 cell line without selection pressure. . . .	192
B.3. Gene Ontology analysis of genes associated with differentially methylated CpG islands in replicate cultures of the bad producing CHO-XL99 cell line without selection pressure.	193
B.4. Gene Ontology analysis of genes affected by puromycin induced differential methylation of associated CpG islands.	194
B.5. Complete GO annotation for highly expressed CHO DP-12 PMD genes (1261) with Benjamini p values ≤ 0.05 (Wippermann et al. 2015).	224
B.6. Complete GO annotation for weakly or not expressed CHO DP-12 PMD-genes (2360) with Benjamini p values ≤ 0.05 (Wippermann et al. 2015).	226
B.7. Motifs enriched in hypermethylated DMRs 6, 12, 24 and 36 hours after butyrate addition by <i>DREME</i>	227
B.8. Motifs enriched in hypomethylated DMRs 6, 12, 24 and 36 hours after butyrate addition identified by <i>DREME</i>	228
B.9. DNA motifs <i>de novo</i> discovered in hypermethylated and hypomethylated DMRs were compared to human and mouse transcription factor databases (Wippermann et al. 2016).	230
B.10. Gene ontology (GO) classification of differentially expressed genes (929) with expression profiles belonging to cluster I (Wippermann et al. 2016).	231
B.11. Gene ontology (GO) classification of differentially expressed genes (1014) with expression profiles belonging to cluster II (Wippermann et al. 2016).	233
B.12. Gene ontology (GO) classification of differentially expressed genes (2861) with expression profiles belonging to cluster III (Wippermann et al. 2016).	237
B.13. Gene ontology (GO) classification of differentially expressed genes (3039) with expression profiles belonging to cluster IV (Wippermann et al. 2016).	239
B.14. Gene ontology (GO) classification of genes both hypomethylated and hypermethylated upon butyrate addition (2890; Wippermann et al. 2016).	243
B.15. Gene ontology (GO) classification of hypomethylated genes upon butyrate addition (3886; Wippermann et al. 2016).	244
B.16. Gene ontology (GO) classification of hypermethylated genes upon butyrate treatment (1737; Wippermann et al. 2016).	245
B.17. Comparison of studies that aimed to analyze butyrate-induced gene expression changes by microarrays in cultured cells (Klausing, Krämer, and Noll 2011; Kantardjieff et al. 2010; Li and Li 2006a).	246
B.18. Gene Ontology analysis of genes affected by butyrate-mediated differential hypomethylation of associated CpG islands 24 hours upon butyrate addition (Wippermann et al. 2013).	248
B.19. Gene Ontology analysis of genes affected by butyrate-mediated differential hypermethylation of associated CpG islands 24 hours upon butyrate addition (Wippermann et al. 2013).	250

Danksagung

An erster Stelle möchte ich Herrn Prof. Dr. Thomas Noll für die Möglichkeit danken, diese Arbeit am Lehrstuhl Zellkulturtechnik anfertigen zu dürfen. Besonders bedanken möchte ich mich dabei für die Bereitstellung des spannenden Themas und das Ermöglichen von Publikationen und Konferenzteilnahmen. Weiterhin möchte ich mich grundsätzlich dafür bedanken, dass Du stets ein offenes Ohr für alle Anliegen, ob fachlich oder nicht, hattest und mir mit gezielten Ratschlägen immer weiterhelfen konntest.

Herrn Prof. Dr. Karsten Niehaus möchte ich für die Bereitschaft danken, als Zweitbetreuer zu fungieren und diese Arbeit zu begutachten. Weiterhin möchte ich mich für die guten Ratschläge zur Konzeption der Arbeit bedanken und ebenfalls für die Möglichkeit, das Luminometer für die Promotorstudien zu testen.

Herrn Dr. Raimund Hoffrogge möchte ich nicht nur für die gute Zusammenarbeit, unzählige fachliche Diskussionen, das Betreuen meiner Masterarbeit, die Begleitung meiner Doktorarbeit und das Korrigieren von Publikationen, Abschlussarbeiten, Postern und Vorträgen danken - darüber hinaus danke ich Dir auch für die schöne Zeit und netten Gespräche in der Mensa, den nachmittäglichen Kaffeeservice und viele lustige Momente. Nie vergessen werde ich zum Beispiel den Abend in Lille, an dem wir über Deine Tante sprachen.

Besonders bedanken möchte ich mich bei Herrn Oliver Rupp für die kompetente bioinformatische Unterstützung - ohne Dich wäre diese Arbeit nicht in dieser Form möglich gewesen. Und obwohl ich bestimmt manchmal sehr penetrant war, hast Du immer eine Engelsgeduld bewiesen. Auch dafür möchte ich mich herzlich bedanken.

Bei Frau Dr. Karina Brinkrolf möchte ich für die gute Zusammenarbeit bedanken und für viele fruchtbare Diskussionen. Danke für die vielen Dinge, die ich von Dir hinsichtlich des Next Generation Sequencing lernen konnte. Auch möchte ich mich für das detaillierte Korrekturlesen von Papern und Abschlussarbeiten bedanken, sowie für nette Abende und mehr (!) oder weniger anstrengende Wandertouren mit guten Gesprächen. Dank Dir habe ich außerdem gelernt, Reiseinformationen immer gründlich zu checken!

Frau Anica Schmidt und Frau Klaudia Grunwald (Tic und Trick? Man weiß nie wer wer ist) danke ich für die gute Labor- und Büronachbarschaft, lustige Kaffee- und Sushi-Runden und die unzähligen Gespräche nach denen es hieß: "Und heute ist erst Montag!". Auch wenn die Zeit stressig war, haben wir wirklich viel gelacht. Besonders herzlich bedanken möchte ich mich auch für die medizinische Betreuung während einer speziellen Konferenzsession und nachdem die Zentrifuge versuchte, mir den Finger abzubeißen. Ich danke Euch ebenfalls für die schöne Zeit Lille und Barcelona. Ich werde bestimmt noch meinen Enkeln die Geschichte davon erzählen, wie wir zwischen den Glastüren feststeckten.

Bei Frau Nadine Rodrigues de Carvalho bedanke ich mich nicht nur für die hervorragende technische Assistenz, sondern auch für die kompetente Beratung in technischen Belangen ("Hast du schon überprüft, ob der Stecker drin ist? – Ja guck, jetzt geht es wieder."), ständige Hilfsbereitschaft und viele nette Gespräche über Katzen und das Leben im Allgemeinen.

Den CLIB-Kollegen danke ich für gute Diskussionen, spannende Vorträge, lustige und informative Retreats und nette Abende. Bei Frau Dr. Iris Brune möchte ich mich in diesem Zusammenhang für hervorragende Koordination, großes Engagement und die gute Zusammenarbeit bei der Organisation von Weihnachtsfeiern und Barbecues bedanken.

Bei Frau Dr. Sandra Klausning möchte ich mich noch einmal für die gute Betreuung und Zusammenarbeit während meiner Masterarbeitszeit und danach bedanken und auch für die gute Büronachbarschaft.

Der gesamten Arbeitsgruppe Zellkulturtechnik danke ich für die schöne Zeit, nette Gespräche in der Kaffeepause, im Labor und bei Ausflügen und Hilfe in allen Lebenslagen.

Frau Anica Winkler und Herrn Andreas Albersmeier danke ich für die kompetente Hilfe bei der Vorbereitung der Sequenzierlibraries und der kompetenten Durchführung des Sequenzierexperimentes.

Bei Frau Eva Schulte-Berndt möchte ich mich besonders herzlich für die Hilfe bei der kniffligen Hybridisierung der Methylierungsarrays bedanken.

Meinen Bacheloranden Daniela Dey, Greta Bischof, Maximilian Jung und Johannes Britz danke ich für ihre Mitarbeit am Projekt, die zu einigen interessanten Ergebnissen geführt hat.

Meiner Familie danke ich für alles. Ohne euch wäre ich nicht, wer ich bin und wo ich bin. Danke für die Unterstützung während dieser langen Ausbildung, dafür, dass Ihr immer da seid wenn es darauf ankommt und dass wir immer zusammenhalten. Ich könnte mir keine bessere und lustigere Familie vorstellen oder wünschen und ich bin so froh, dass es Euch gibt.

Bei Robert möchte ich mich für die große Unterstützung während des letzten Jahres bedanken. Danke, dass Du immer hinter mir stehst, mir Kraft gibst und mein bester Freund bist.

"Last but not least" danke ich Freya dafür, dass sie alle Abschlussarbeiten bis hierher begleitet hat und sogar meistens Verständnis hatte, wenn ich zu viel Zeit am Schreibtisch verbringen musste (obwohl das bestimmt nicht immer leicht war). Danke, dass es Dich gibt! Ich hab Dich lieb bis zum Mond und zurück.

Erklärung

Hiermit erkläre ich, dass mir die geltende Promotionsordnung der Technischen Fakultät der Universität Bielefeld bekannt ist und ich die vorliegende Dissertation selbstständig und ohne Hilfe Dritter angefertigt habe. Es wurden ausschließlich die genannten Quellen und Hilfsmittel verwendet. Wörtliche und sinngemäße Zitate, sowie anderen Publikationen entnommene Abbildungen sind als solche gekennzeichnet. Die Arbeit wurde weder in dieser noch einer anderen Fassung einer Prüfungsbehörde vorgelegt. Ich habe keine erfolglosen Promotionsversuche unternommen.

Enger, den 6.12.2016