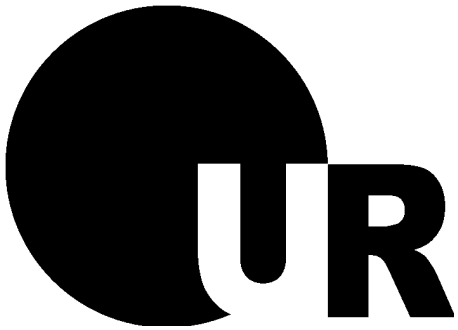
# Analysis of tissue-specific & allele-specific DNA methylation

Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.) der Naturwissenschaftlichen Fakultät IV – Chemie und Pharmazie der Universität Regensburg



vorgelegt von

**Elmar Schilling**

aus

Schwenningen

2009

The present work was carried out in the Department of Hematology and Oncology at the University Hospital Regensburg from June 2005 to June 2009 and was supervised by PD. Dr. Michael Rehli.

Die vorliegende Arbeit entstand in der Zeit von Juni 2005 bis Juni 2009 in der Abteilung für Hämatologie und internistische Onkologie des Klinikums der Universität Regensburg unter der Anleitung von PD. Dr. Michael Rehli.

Promotionsgesuch eingereicht am: 30. Juli 2009
Die Arbeit wurde angeleitet von PD. Dr. Michael Rehli.

Prüfungsausschuss:

Vorsitzender: Prof. Dr. Sigurd Elz
1. Gutachter: Prof. Dr. Roland Seifert
2. Gutachter: PD. Dr. Michael Rehli
3. Prüfer: Prof. Dr. Gernot Längst

Lob und Tadel bringen den Weisen
nicht aus dem Gleichgewicht.

(Budha)

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# Acknowledgement

I would like to thank Prof. Dr. Reinhard Andreesen for accepting me as a PhD student in his department and for his honest and generous support during all those years.

I would like to thank Prof. Dr. Roland Seifert for accepting to supervise and review my thesis.

I am very grateful to PD Dr. Michael Rehli for his enthusiasm, his continuous support and advice and for the great atmosphere in our lab. He was always willing to listen to my problems. He kept patience and calmness all the time and during his leadership I had the chance to learn many new things.

I would like to thank everybody inside our laboratory for the wonderful atmosphere, for your patience, technical and moral support. In detail: Many thanks to: Dr Carol El-Chartouni, PhD Hang Thu Pham, Claudia Gebhard, Maja Klug, Lucia Schwarzfischer-Pfeilschifter, Dagmar Glatz, Christian Schmidl, Katrin Dietl, Michael Kastenberger, Julia Wegner, Prof. Dr. Marina Kreutz, Dr. Eva Gottfried, Monika Wehrstein, Alice Peuker, Gabi Hartmannsgruber, Martina Köhler, Ireen Ritter and Sandra Vogl and all those I probably forgot to mention. In addition, I would like to take the opportunity to thank our closely connected co-workinggroup – the Kreutz-Lab. We not only shared room and working material, I also appreciated our combined seminars where I learned much about tumor and monocyte metabolism. I also would like to thank the whole Forschungsbau H1 where I could find all I needed for both: the last minute experiments and coffee breaks, climbing or grilling events.

Thanks to Carol for the help in the preparation of BMM's. Special thanks to Luci for her precious organization skills and her technical and moral support. I will also like to thank Dagmar and Ireen for their technical support and (not to forget) funny coffee breaks and spare time events.

I would also like to thank all former crewmembers, who already left the ‘ship’ sailing to various different harbors: Dr. Monika Lichtinger, Dr. Tobias Weil, Dr. Mike Aigner, Monika Eigner, Sabine Pape and Alex Müller.

I am glad that I found some people having the same life of suffering each day: pangs of hunger. This group consisted of Kaste, Chris, Tobi, Dagmar, mostly Michael and was sometimes complemented by Katrin and Sandra, at least for the obligatory “after lunch” coffee (At this point it is indicated to thank Nespresso…)
I was also a great pleasure for me to live in our flat-sharing community consisting of my beloved brother Reini and Martin Lachinger, who successfully substituted my good old friend Dr. Werner Müsch after three eventful years. I’ll also like to thank all other friends not personally mentioned yet who helped to make my stay in Regensburg not boring at all.

Last but not least, I would like to thank my parents. They really always supported me in a very unselfish way. Without their help I would never had the change to start and finish my study, to work in and explore the United States for over half a year and to perform the present thesis. I really appreciate their never-ending help and patience.

# 1 Introduction

## 1.1 The concept of Epigenetics

The human genome project resulted in enormous amounts of genomic raw data. It is now known, that there are two meters of tightly compacted DNA, containing roughly 30,000 genes, coiled into the nucleus of each cell. However, identity and developmental potential of individual cells within organisms is not only defined by its genetic component. It is important to understand changes that occur in the function of genes without a change in the genome sequence. This branch of research is known as “epigenetics”, and concerns itself with chemical interactions that modify either DNA itself or its packaging proteins, the histones. The term "epi" means "on top of", and the epigenetic code compromises a second layer of information on top of the genetic code. **Figure 1.1** illustrates basic epigenetic modifications, observed on the level of DNA and histones. Epigenetics, in a broad sense, is a connection between genotype and phenotype that alters the final outcome of genomes without changing the underlying DNA sequences. For example, even though the great majority of cells in a multicellular organism share an identical genotype, organisms contain a variety of cell types with diverse, yet stable, profiles of gene expression and distinct cellular functions. More specifically, epigenetics can be defined as the study of any stable heritable change in gene expression or cellular phenotype that occurs without changes in the sequence of DNA. Epigenetic research nowadays includes the study of covalent and non-covalent modifications of DNA and histone proteins and the mechanisms by which such modifications influence overall chromatin structure. The following chapters particularly focus on DNA methylation, which provides a stable, heritable and critical component of epigenetic regulation.

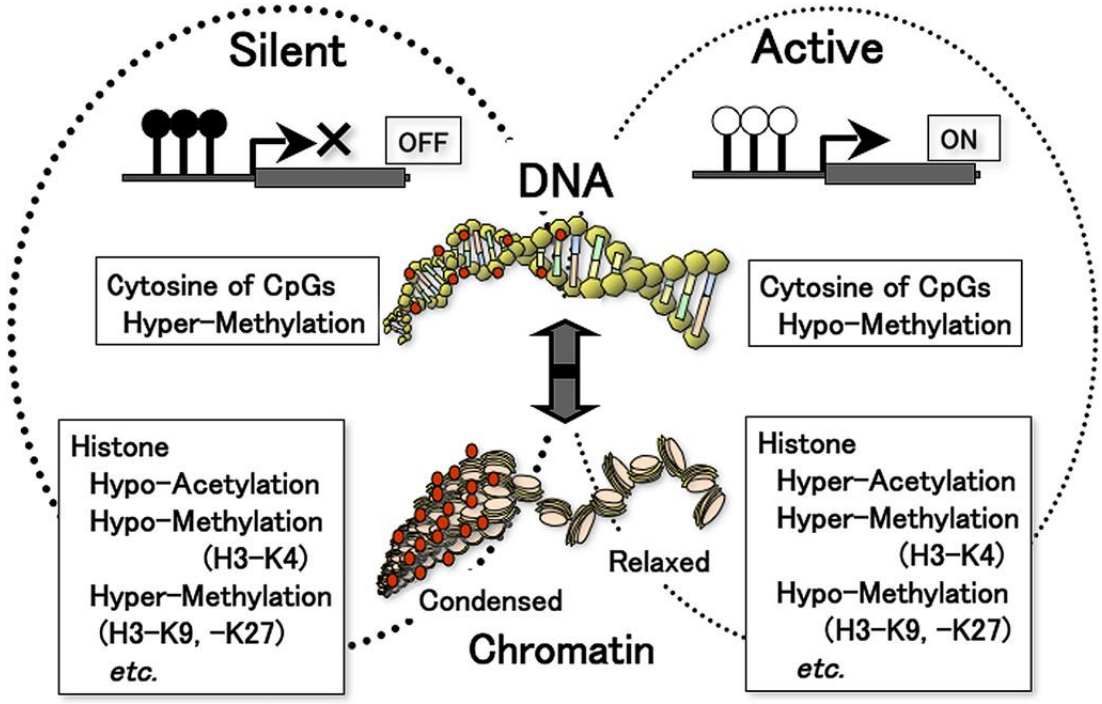



**Figure 1.1 Epigenetic regulatory mechanisms: DNA methylation and chromatin structure.** Methylation of CpGs occurring in the regulatory region of genes causes gene silencing. Methylation of DNA and chromatin structure is coordinated through modification of histones, including by acetylation and methylation. In general, DNA methylation and chromatin condensation are associated with gene silencing. The epigenetic state consisting of DNA methylation and chromatin configuration results in changes in the differentiation of cells. Once established, the epigenetic state is maintained during the proliferation of cells (Figure taken from reference[1]).

## 1.2 DNA methylation

The longest known epigenetic modifications is the covalent addition of a methyl group to the 5' carbon atom of the base cytosine (C). In mammals, this modification is mainly assigned to cytosines followed by guanine (G), so called CpG dinucleotides (CpGs). DNA methylation is associated with gene silencing and was reported to be essential for embryonic development[2], genomic imprinting[3], X-inactivation in mammals[4] and silencing of potential harmful DNA elements like transposons or endogenous retroviruses[5]. In addition, aberrant DNA methylation has been linked to abnormal developmental processes and cancer formation[6]. Interestingly, CpG dinucleotides are distributed unequally over the genome: Most CpGs in mammals are methylated, distributed randomly and appear rarer than statistically expected[7]. One possible explanation for this distribution is the observation that 5mC hydrolytically deaminates to thymine (T). Thymine - a naturally occurring genomic base - is not recognized as misplaced, resulting in a C to T

transition and a decrease of CpGs over time in evolution. However, there are regions with higher CpG density, so called CpG islands (CGIs). They were initially defined as regions with a GC content of 50 % or higher and a ratio of observed versus expected CpG frequency of 0.6 in a region of 200 base pairs minimum[8]. CGIs are often associated with promoter regions. More than half of the genes in the human genome are associated with one or more CGIs which are preferentially unmethylated, raising the question how the maintenance of hypomethylation in development and differentiation is achieved[7].
Two classes of proteins are involved in creating and recognizing methylated CpG's (DNA methyltransferases and Methyl-CpG-binding proteins) and are introduced in the next two sections.

### 1.2.1 DNA methyltransferases

DNA methyltransferase activity was first detected in crude cellular extracts and subsequently purified as a 200 kDa protein named DNA methyltransferase 1 (Dnmt1)[9]. In eukaryotes, three different families of DNA methyltransferases (Dnmts) have been identified (Dnmt1, Dnmt2 and Dnmt3)[10,11]. So far four different proteins belong to these families (1, 2, 3a and 3b). The different family members have similar molecular structures with most of them containing ten specific sequence motifs within the C-terminal catalytic domain, six of which are highly conserved[12,13]. However, the regulatory domains located at the N-terminal part of the proteins show little sequence homologies[14] and are completely absent in Dnmt2. It was proposed, that DNA methyltransferases could act the same way than other enzymes catalyzing electrophilic substitutions at carbon-5 of pyrimidins[15]. The contact between DNA and enzyme is established by extreme distortion of the DNA helix. The reaction is started when cytosine is flipped out of the DNA and inserted into the binding pocket of a methyltransferase enzyme. In the active site, the catalytic cysteine thiolate forms a transition state intermediate with the carbon-6 of the cysteine's ring. This creates a reactive, negative charged 4-5 enamine, which attacks the methyl group provided by the AdoMet cofactor (S-adenosyl methionine, SAM) acting as methyl donor[16]. After transfer of the methyl group to the C-5 position within the cytosine ring the proton is removed from this position leading to reconstitution of the 5,6 double bond and the release of the enzyme by ß-elimination[17]. Although the mechanism of transfer is similar between the different family members, the biological roles are different. Dnmt1 enzyme exhibits a 5-10 fold preference for hemi-methylated DNA[18], which suggests that this enzyme is responsible for the maintenance of DNA methylation patterns after daughter strand formation during the S-phase of cell cycle. Functional experiments in mammalian cells additionally proposed a role in the X-chromosome inactivation, allelic silencing of imprinted genes and silencing of incorporated transposons[19]. Silencing of

human Dnmt1 is lethal for normal cells[20]. The catalytic significance of Dnmt2 has not been identified. It was shown, that Dnmt2 is not essential for global de novo or maintenance methylation in embryonic stem cells[21]. It may be involved in maintenance of centromeric structure[10]. Dnmt3a and Dnmt3b are generally considered to be de novo DNA methyltransferases. During early development, when most of the de novo methylation occurs, both enzymes are highly expressed and essential for the establishment of new methylation patterns and following correct development[2]. Another nuclear protein not mentioned yet, DNA methyltransferase-like protein 3L (Dnmt3L), shows similarity to DNA methyltransferases. This protein is not thought to function as a DNA methyltransferase as it does not contain the amino acid residues necessary for methyltransferase activity. However, this protein does stimulate de novo methylation by Dnmt3a[22] and it is thought to be required for the establishment of maternal genomic imprints[23]. This protein also plays an important role differentiation of spermatocytes[24].

### 1.2.2 Methyl-CpG-binding proteins

Another class of proteins binds to methylated CpG's and comprises the so called methyl-CpG-binding proteins (MBD). In 1993, Bird and colleagues defined the MBD domain by molecular analysis of the prototype MBD protein, MeCP2[25]. This domain was both, necessary and sufficient, to bind mCpG DNA and was found in all MBD family members. The family contains 4 other proteins, named MBD1-4 (see **Figure 1.2**). MBD2 is the DNA-binding component of MeCP1 complex, which was observed to be specific for methylated DNA in a variety of mammalian cell types[26]. Three of these proteins (MBD1, MBD2 and MeCP2) are involved in methylation-dependent repression of transcription[27]. An unrelated protein named KAISO, which lacks the MBD domain, has also been shown to bind methylated DNA (via zinc fingers) leading to repression in model systems[28,29]. Each of the four methyl-CpG-binding proteins has been shown to associate with a different repressor complex[30-32]. A prerequisite for understanding the biological roles of the different MBD proteins is the knowledge of their target sites. One might expect that the various proteins compete with one another for access to methylated sites, due to their overlapping DNA sequence specificity. However, Klose *et al*. studied MBD binding sites within a primary human cell lines and demonstrated, that they were largely nonoverlapping[33], suggesting that each protein is targeted independently. In this study MeCP2 strongly prefers mCpG sites flanked by a run of AT-rich DNA. MBD1 binds to mCpG via its MBD domain, but a major splice form is additionally able to bind unmethylated CpG via a CxxC domain. A pair of adjacent mCpG motifs is recognized by Kaiso[28]. Only MBD2 so far appears to have an exclusive affinity for methylated CpG dinucleotides.

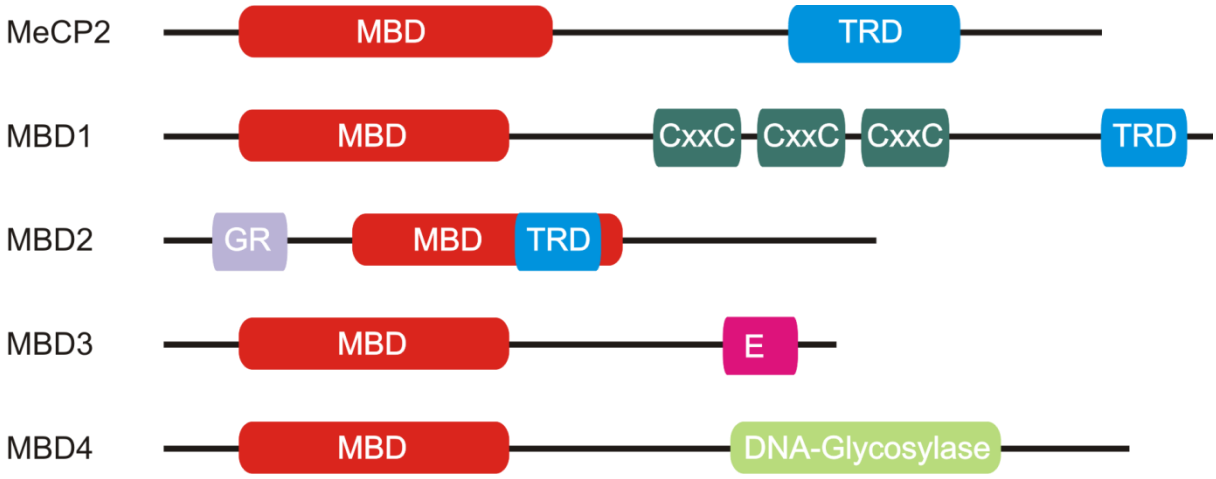



**Figure 1.2 Characteristic domains of the MBD protein family**. Five members of the MBD protein family are shown. MBD domains are marked in red. Other domains are labelled as well and include transcriptional repression domains (TRD, colored in blue) CxxC domains (colored in green) GR repeats of unknown function (GR, labelled purple) and an E-repeat (E, labelled in rose). A DNA N-glycosylase domains in, which is involved in the T:G mismatch repair after 5'-methylcytosine deamination (shown in light green) (Figure adopted from reference[34]).

## 1.3 Functions and molecular consequences of DNA methylation in healthy cells

Since every cell of a multicellular organism contains essentially the same genome, it is crucial to understand why genetically homogenous cells are different from each other and how this is controlled. As mentioned above, DNA methylation does not alter the base composition itself. Nevertheless, a fifth base is introduced to the genomic code, leading to an altered surficial area, which may trigger various reactions. Several observations indicate that DNA methylation has an essential regulatory function in mammalian development, which is to establish the correct pattern of gene expression, and that distinct DNA methylation patterns are tightly correlated to specific chromatin structures. Various physiological processes are controlled by specific DNA methylation patterns including genomic imprinting[3], inactivation of one X chromosome in females[4], regulation of tissue-specific gene expression[35,36] and silencing of transposons[5]. Moreover, aberrant methylation can confer a selective advantage, leading to cancerous growth[37]. The following part focuses on the epigenetic molecular mechanisms during normal development.

### 1.3.1 Global methylation landscapes

In mammals DNA methylation is found throughout the genome in mammalian studies. Interestingly short regions - so called CpG islands (CGIs) - remained unmethylated[38,39] and are believed to be protected from methylation through *cis*-acting elements[40]. CpG islands typically occur at or near the transcription start site of genes, particularly housekeeping genes, in vertebrates[41,42]. CGIs are usually defined as regions with at least 200 bp and with a GC percentage that is greater than 50% and with an observed/expected CpG ratio that is greater than 60%[8]. Based on an extensive search on the complete sequences of human chromosomes 21 and 22, DNA regions greater than 500 bp with a GC content of more than 55% and observed/expected CpG ratio of 65% were more likely to be the true CpG islands associated with the 5' regions of genes[43]. However, the global DNA methylation pattern seen in vertebrates is not ubiquitous among eukaryotes. Several well-studied model systems have no recognizable *Dnmt*-like genes and are devoid of DNA methylation (for example, the yeast *Saccharomyces cerevisiae* and the nematode worm *Caenorhabditis elegans*). In fungi that have genomic 5-methylcytosine (5mC), only repetitive DNA sequences are methylated[44]. Invertebrates mostly exhibit a ‘mosaic methylation’, comprising domains of heavily methylated DNA interspersed with domains free of methylation[45]. In plants up to 50% of cytosine’s have been reported to be methylated in some species, which represents the highest level of DNA methylation among eukaryotes[46]. Large numbers of methylated transposons, seem to be responsible for this high levels in maize[47,48].

An accumulation of 5-methylcytosines is also found in tandem and interspersed repeats, constituting a second group of CpG-rich regions beside CpG islands. Among these repeats, *Alu* repeats with more than one million copies per haploid genome are considered to be the most dominant family[49]. They belong to the 'short interspersed nucleotide elements' (SINEs) and alone constitute for more than 10% of the human genome[49]. They are not randomly distributed within the human genome, but tend to accumulate in gene-rich regions[50]. Most *Alu r*epeats have been reported to be highly methylated in somatic tissues[51,52]. DNA methylation in general represents the major mechanism to repress transposable elements, as detailed below.

### 1.3.2 Genomic immunity: *De novo* methylation of integrated foreign DNA

DNA methylation is an ancient process found in all domains of life. In eukaryotes, which include diverse organisms such as plants as well as humans, DNA methylation is found exclusively at cytosine residues. The observation, that mice lacking Dnmts die very early

during embryogenesis, suggests that this modification has important roles and is essential for mammalian embryonic development[2,53]. *Dnmt*-null mice show reduced DNA methylation levels but the specific reasons for death during development remain unclear. Defects in repression of the inactivated X chromosome in female cells and in the establishment and maintenance of allele-specific expression of imprinted genes have been observed[54-56]. Thus, lethality might result from aberrant gene dosage. In addition, increased expression of transposon RNA was reported in these embryos[57].

In *Dnmt*-knockout mice, global demethylation as a consequence of Dnmt1 absence likely triggers mutations through the activation of cryptic transposons, which might contribute to early lethality. Normally these transposons are methylated and thereby repressed. Consequently, DNA methylation masks the effects of transposon insertion by mechanisms that do not directly depend on regulation of transcription or transposition. Instead of that, methylated transposons are hidden from the genome mediated amongst others by DNA methylation[5]. This property may have allowed transposons to accumulate to high copy numbers without major scrambling effects on the host genome. Amongst others, these observations resulted in the conclusion that cytosine DNA methylation functions to maintain the repressed chromatin state leading to silencing of (alternative) promoter, pseudo gene or transposon activity[27]. It is therefore conceivable that intragenic methylation helps to weaken transcriptional noise[58].

However, DNA methylation has been adapted for a variety of uses throughout evolution, beside defense against transposable elements, including control of gene expression, discussed in the next paragraph.

### 1.3.3 Development: Tissue specific DNA methylation

As mentioned before, every cell of a multicellular organism contains essentially the same genome. However, entire organisms incorporate a huge variety of differentiated and specialized cells, all of them using only a small proportion of available genes.

These patterns of differential gene activity are clonally inherited through cell division. Because specific methyl-CpG dinucleotides are maintained through DNA replication, DNA methylation states provide an attractive mechanism (epigenetic mark) to maintain a particular state of gene activity through cell division and, thus, to contribute to the maintenance of the differentiated state. However, the precise extent of tissue-specific hyper- or hypomethylation throughout the genome is largely unknown and has been controversially discussed[35,36].

## 1.3.4 Imprinting

Imprinting describes the rare but remarkable situation when alleles of the same gene are expressed unequal in a parent-of-origin dependent manner. This means that some genes within the diploid genome are only expressed from alleles inherited from the father, whereas some other are only expressed from alleles derived from the mother. This phenomenon was first reported in 1984, when it was discovered, that parthogenic (maternal only) and androgenic (paternal only) embryos fail to develop after murine nuclear transfer experiments[59,60]. Shortly after, the responsible chromosomes carrying the imprinted loci where discovered via studies of UPD (uniparental contribution of one chromosome pair) in mice[61]. UPD for some chromosomes was lethal or with severe developmental defects contrary to other chromosomes, where according phenotypes were unaltered. Imprinted genes must carry some distinguishing mark on one parental allele in order to assure sex-specific gene expression. *Igf2* and H19, nearby on murine chromosome 7, were the first mammalian imprinted genes to be identified[62-64]. In addition, a role for DNA methylation in imprinting was first demonstrated by defective imprinting observed in *Dnmt-/-* mice, where imprinting of the H19, *Igf2* and *Igf2r* genes was disturbed[3]. This study revealed that maintenance of DNA methylation was essential for imprinted expression. This region is under the control of a differentially methylated region (DMR) which can bind an enhancer blocking protein (CTCF) in a methylation sensitive way[6]. The unmethylated DMR on the maternal allele binds the CCCTC-binding-factor *CTCF* and blocks a downstream enhancer of the *Igf2* gene. In contrast H19, which is located upstream of the “blockade”, is expressed. The paternal allele is methylated at the DMR preventing the binding of CTCF, what enables the enhancer to stimulate transcription of *Igf2*[65].

Differentially methylated regions are typically found within the vicinity of imprinted genes. Some of which are called imprinting control regions (ICR) and by definition are already established in germ cells and are resistant to early embryonic epigenetic reprogramming (see **Figure 1.3**). It is noteworthy that all ICRs/DMRs would not formally be defined as CGIs and differences in GC and CpG content between maternal and paternal have been reported[66].

## 1.3.5 Resetting of imprints

During development two major waves of genome-wide demethylation take place affecting a substantial part of the genome. The first wave is in the germ cell stage and the second later after fertilization. Both are important for the inheritance of imprinting, though in a different mode. The first genome-wide demethylation is thought to erase methylation in

order for the new parent-of-origin methylation to be set. During the second wave after fertilization the methylation pattern of imprinted loci is believed to be protected, which is necessary for the proper dissemination of imprinting to somatic cells[67]. Interestingly, the mechanism of epigenetic reprogramming seems to differ between male and female germ cells, at least regarding establishment dynamics[68]. *Snrpn* (small nuclear riboprotein N) and *Mest* (mesoderm specific transcript homolog) are two examples of maternally expressed genes that were unmethylated in sperm but fully methylated in mature oocytes[69]. The latter two examples represent genes of particular interest, because of their important biological functions and characteristic phenotypes associated with the absence of their gene products. *Snrpn* is reported to be involved in the development of Angelmann or Prader-Willi syndromes[70], whereas paternal transmission of *Mest* gene results in growth-retarded embryos and increased postnatal death. In addition abnormal adult maternal behavior has been noted for *Mest*-deficient females[71,72]. Reprogramming during development is summarized in **Figure 1.3**.

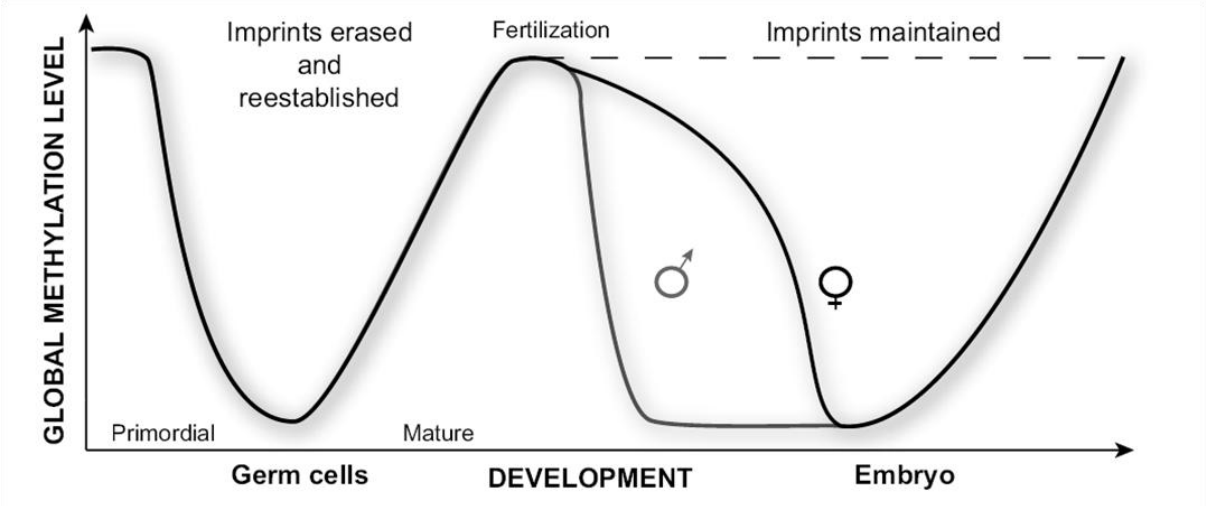



**Figure 1.3 Imprinting during developmental reprogramming**. During germ cell formation DNA methylation is erased and re-established in a gender-dependent manner. After fertilization, the paternal genome (♂) is demethylated by an active mechanism whereas the maternal genome (♀) is demethylated by a passive mechanism that depends on DNA replication. After fertilization, imprints are resistant to global waves of de- and remethylation. (Adopted from reference[67]).

## 1.3.6 X-chromosome inactivation

Besides imprinting, another system of monoallelic gene exclusion in mammals is known. In mammals, sex is determined by differential inheritance of a pair of dimorphic chromosomes: the gene-rich X chromosome and the gene-poor Y chromosome. To balance the unequal X-chromosome dosage between the XX female and XY male, mammals have adopted a unique form of dosage compensation in which one of the two X

chromosomes is inactivated in the female. This mechanism involves a complex, highly coordinated sequence of events.

X-chromosome inactivation in eutherian mammals was first discovered in 1961 by Mary Lyon in studies of coat-colour genes in mice[73] and is a random process, commonly divided into the steps of counting, choice, initiation, establishment and maintenance (see **Table 1.1**). These steps are genetically separable and appear to be controlled by the X inactivation center (Xic) except for maintenance[73-75]. Xic contains *Xist*, a noncoding RNA gene first described 1991[76,77]. During the counting step, the number of X-chromosomes is measured, in relation to haploid autosome sets. A region at the 3' end of the *Xist* gene is thought to be responsible for this initial step[78]. During choice all but one X-chromosome is committed to inactivation and it has been shown that the genes *Xist* and *Tsix* participate in regulating this step[79,80]. *Tsix* is the second non-translated RNA, transcribed antisense to *Xist*, encoded by the Xic region[81]. Xist RNA accumulates along the X chromosome containing the active *Xist* gene and proceeds to inactivate almost all of the other hundreds of genes on that chromosome[82].

*Tsix* mRNA has been shown to be a negative regulator of *Xist* and prevents its abundant accumulation on the active X chromosome[81]. Hence, the inition of silencing relies on *Xist* expression. However, once silencing is established, maintenance of the inactive X is apparently independent of further Xic and *Xist* function[83,84]. Following this physical deactivation, large parts of the chromosome are silenced by DNA methylation. For random X-inactivation maintenance, DNA methylation is the key stabilizing factor, as deletion of maintenance DNA methyltransferase Dnmt1 results in reactivation of the silenced X in the embryo[85]. In addition, DNA methylation is also required to stably repress the *Xist* gene on the active X chromosome[82,86,87].

It is worth mentioning, that X-inactivation exists in an imprinted form as well, which is believed to be the ancestral mechanism. Marsupial mammals undergo nonrandom X-chromosome inactivation and preferentially shut off the paternally derived X-chromosome[88].

**Table 1.1** Steps and active components of mammalian X-inactivation

| ***Step*** | ***Component*** |
|---|---|
| Counting | *Xist* 3' region |
| Choice | *Xist* and *Tsix* |
| Initiation of silencing | *Xist* |
| Establishment | *Xist*? |
| Maintenance | i.a. DNA methylation |

### 1.3.7 Interindividual phenotypical differences and inheritance of DNA methylation

In the past it has been considered that epigenetic differences participate in characteristic individual phenotypes too. Although phenotypic variation between individuals is mainly driven by genetic traits, there is also evidence that epigenetic mechanisms may contribute to phenotypic differences in mammals[89]. Examples for epigenetic differences between individuals are rare and mostly, but not exclusively, confined to the level of DNA methylation. It has been demonstrated that supposedly genetically identical, monozygotic twins can show differences at the level of DNA methylation[90,91]. These differences, however, are likely not inherited but acquired during the lifetime of each individual[90]. In addition, there are several reports demonstrating that DNA sequence variants associate with specific epigenetic states[92-94]. A recent study in humans identified several cases of allele-specific DNA-methylation at non-imprinted gene loci[95], where the methylation status of each allele was likely controlled in *cis* by the local DNA sequence. This raises the question how epigenetic states can be inherited to the next generation. The underlying DNA sequence itself seems to play an important factor (regulated in *cis*), as highlighted before. In contrary, in mice there is well-documented evidence that epigenetic states can be inherited across generations, e.g. at the Agouti viable yellow (*Avy*) allele[96,97], and that allelic variation at certain epigenetic modifier genes in mice, like DNA methyl-transferases or chromatin remodelling factors may influence the inheritance of CpG methylation patterns in *trans*[98,99].

Taken together, this suggests that three types of inheritance may exist *in vivo*: methylation patterns at non-imprinted loci may be inherited based on genetic mechanisms (in *cis* and in *trans*) or based on epigenetic mechanisms. The extent of contribution of each type of mechanism in shaping individual epigenetic or phenotypic differences is currently unknown.

## 1.4 Mapping DNA methylation

As detailed above, DNA methylation has various regulatory roles and functions in both healthy and malignant cells. On the basis of these discoveries, the investigation of DNA methylation has drawn much attention and was subjected to continuous further developments. However, the direct examination of DNA methylation has been hindered by the fact, that DNA methylation cannot be analyzed by standard DNA sequencing methods, since they are unable to distinguish 5-methylcytosine from unmodified cytosine. Thus, alternative techniques must be used to access normal or aberrant methylation. A number

of methodologies used for detection of DNA methylation have so far been developed, each with its own advantages and inherent disadvantages, but no single technique can provide a complete assessment, in terms of resolution (ideally single-base) and feasibility of high throughput and global covering.

Early studies relied on methylation-sensitive restriction enzymes such as NotI, HpaII, or SmaI in conjunction with Southern hybridization[100]. While this provides information on the specific restriction enzyme site, southern hybridization is time-consuming and requires large quantities of DNA and it is not useful for determining global methylation patterns. A popular isoschizomer pair for methylation analysis is HpaII/MspI. Both enzymes cleave CCGG sites, but HpaII is unable to cleave if the internal cytosine is methylated. However, methods based on methylation specific restriction enzymes suffer from the limitation that they provide information only about CpGs within the cleavage site.

Treatment of genomic DNA with sodium bisulfite ($NaHSO_3$) overcomes this limitation and allows the analysis of virtually any CpG position within the genome. As stated previously, standard sequencing methods cannot distinguish 5-methylcytosine from unmodified cytosine, thus hindering the analysis of CpG site methylation. In 1992, Frommer and colleagues[101] described genomic sequencing of bisulfite-treated DNA as a way to distinguish methylated from unmethylated cytosines at single base resolution level. When single stranded DNA is treated with sodium bisulfite, unmethylated cytosines are converted to uracil, whereas 5-methycytosines are unreactive[102] (see **Figure 1.4**).

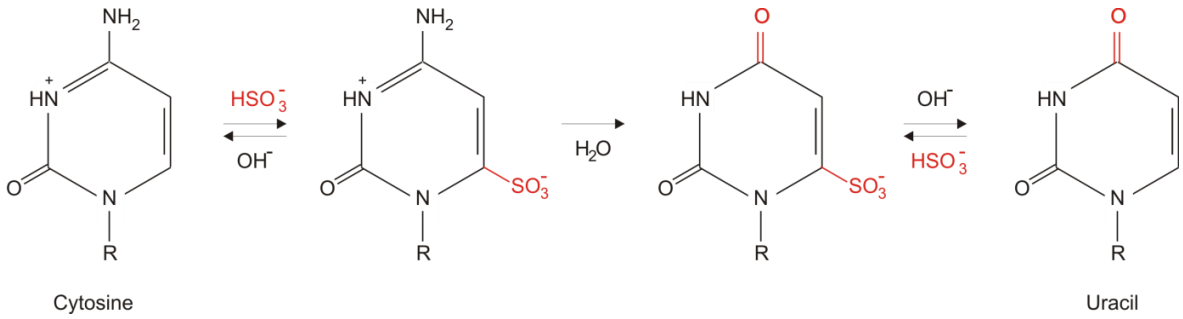



**Figure 1.4 Chemical reaction during bisulfite treatment.** Cytosine-derivates undergo reversible reactions with bisulphite yielding a 5,6-Dihydro-6-sulfonate, which deaminates spontaneously. After that the sulphate is eliminated under alkaline conditions, leaving Uracil. Methylated cytosine is not affected by this reaction, due to sterical hindrance by the methyl-group in 5'-position.

Using primers that surround a CpG-containing area of interest, but do not contain CpGs themselves, the modified DNA strands are then amplified by PCR. The PCR-product can then be sequenced, directly or after subcloning of the amplified fragment. Direct sequencing gives information about the average methylation of a CpG site in a sample, while sequencing of cloned DNA allows the analysis of individual CpG sites on

independent half strand DNA molecules. However, sequencing of bisulfite-treated DNA is very laborious and time consuming and therefore not practicable on a global scale. This approach has become very important in the investigation of methylation of cancer-related genes and has been used in the analysis of novel methylation targets identified by restriction landmark genomic scanning (RLGS)[103,104].

Specific adapted PCR techniques have been developed in several laboratories to map DNA methylation without the need of sequencing. These methods rely on single nucleotide polymorphisms (SNPs) introduced by the bisulfite-treatment at the CpG site. Specific PCR primers for the methylated and unmethylated states (i.e. MSP)[105] or methylation specific probes are used for real time PCR (i.e. MethyLight)[106].

Other methods to map DNA methylation utilize specific proteins or antibodies, which are able to bind methylated DNA. Methylated DNA can be immunoprecipitated using an antibody directed against 5-methyl Cytidine (5-meC). Immunoprecipitation of methylated DNA (Methyl DNA IP, also called MeDIP) was first described in 2004[107,108]. In brief, genomic DNA is sheared, either using restriction enzymes or via sonication, and denatured. Then, immunoselection and immunoprecipitation can take place using the antibody directed against 5-methyl Cytidine and antibody binding beads. Another strategy for *in vitro* methylation detection uses column- or bead-immobilized recombinant methylated-CpG binding domain (MBD) proteins, particularly MeCP2 and MBD2, to enrich for methylated DNA fragments for subsequent detection by PCR or microarray hybridization[109,110]. In 2006 a robust technique, called methyl-CpG-immunoprecipitation (MCIp) was established in our laboratory. The approach is based on a recombinant, antibody-like protein that efficiently binds native CpG-methylated DNA. In combination with CpG island microarrays, the technique was used to identify genes with aberrantly methylated CpG islands in three myeloid leukemia cell lines[110]. A great advantage of this method is that DNA is not only enriched for highly methylated DNA (whereas unmethylated DNA is lost, as in MeDIP), but also recovers mainly unmethylated DNA, since the DNA is fractionated depending on the methylation density and each individual fraction is kept. This methodology is intensively used in the present thesis and explained in more detail in subsequent sections.

Most techniques used to map DNA methylation described in the literature depend more or less on the principles explained above, but may contain certain modifications or specific adjustments. Besides different pretreatment techniques much effort was done in establishing new downstream high-throughput readout principles. Thus, microarrays, mass spectrometry and next generation sequencing were intensively used to globally mine DNA methylation in various approaches.

Hence, new readout approaches arose, by combination of bisulfite treatment of DNA and next-generation sequencing (NGS)[111,112]. NGS platforms share a common technological feature: massively parallel sequencing of clonally amplified or single DNA molecules that are spatially separated in a flow cell. As a massively parallel process, NGS generates hundreds of mega bases to giga bases of nucleotide-sequence output in a single instrument run, depending on the platform. Ehrich and colleagues recently advanced the treatment of genomic DNA with sodium bisulfite in an alternative direction by combination with mass spectrometry (MS)[113]. This method employs a PCR step after bisulfite treatment during which a T7 promoter primer is incorporated into the PCR product. *In vitro* transcription and subsequent T-specific cleavage follows the PCR. The method then utilizes matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS)[114] in order to detect the G to A sequence difference in the generated RNA fragments. This permits high-throughput identification of methylation sites and their semi quantitative measurement at single or multiple CpG positions.

**Table 1.2** (page 15) summarizes several methodologies used to explore DNA methylation, ether globally or gene specifically, and points out underlying sample pretreatment techniques. Current high throughput readout methods together with key references are summarized in **Table 1.3** (page 16).

**Table 1.2** Current methods for DNA methylation analysis: sample pretreatment

| ***Method (+reference)*** | ***General basis*** |
|---|---|
| Bisulphite conversion[101] | Sodium bisulphite converts unmethylated cytosine to uracil, whereas methylated cytosines are protected from conversion<br>**Resolution: high, single base** |
| RLGS; HELP assay[115,116] | DNA is differentially fragmented with a methylation-sensitive restriction enzyme. Following size fractionation, this method enriches methylated DNA<br>**Resolution: moderate** |
| McrBC digestion[117-119] | DNA digestion with a methylation-specific restriction enzyme, McrBC. Following size fractionation, this method enriches unmethylated DNA<br>**Resolution: moderate** |
| Methylated DNA immunoprecipitation (MeDIP)[107,120-123] | Immunoprecipitates DNA containing methylated cytosines using a monoclonal antibody; requires single strand DNA<br>**Resolution: moderate** |
| MBD affinity purification (e.g. MAP; MCIp; MB-PCR)[109,110,120,124-126] | Immunoprecipitates DNA containing methylated CpG using an MBD column<br>**Resolution:** moderate |
| CXXC affinity purification (CAP)[126] | Immunoprecipitates DNA containing unmethylated CpG using a CXXC-domain column<br>**Resolution: moderate** |

**Table 1.3** Current methods for DNA methylation analysis: high-throughput readout

| ***Readout method (+reference)*** | ***General basis*** | ***Uses*** |
|---|---|---|
| Oligonucleotide arrays[110,120,121,127-130] | Short (25-mer) or long (60-mer) oligonucleotide array | Tiling genomic arrays, promoter arrays, CpG-island arrays and custom arrays<br>**Resolution: moderate** |
| SNP arrays[131] | SNP selective probe array | Detection of allele-specific DNA methylation<br>**Resolution: moderate** |
| BeadArray[132] | Ratio of the methylated and unmethylated PCR products is determined at single CpG sites | Detection of methylation polymorphisms. However, a large set of primers needs to be designed<br>**Resolution: high, quantitative** |
| Standard sequencing | Sanger sequencing | Expensive and labor intensive for genome-wide analysis<br>**Resolution: high, quantitative** |
| Direct large-scale sequencing (NGS)[111,112] | Short-read sequencing (Solexa sequencing: 40 millon reads of 25-30 bases; 454 sequencing: 400,000 reads of >100 bases) | Fast and relatively inexpensive. Genotype information can be obtained simultaneously. However, high-quality reference sequence is required<br>**Resolution: high, quantitative** |
| EpiTyper (MALDI-TOF)[113] | detection of polymorphisms via mass spectrometry | Fast and relatively inexpensive.<br>**Resolution: high, quantitative** |

# 2 Aims

A global picture of epigenetic variation between cell lineages or individuals will be key to a better understanding of DNA methylation and its contribution to phenotypes. The main goals of this thesis were to develop methods to address epigenetic variation, in particular on the level of DNA methylation, and to apply them to study tissue- and mouse strain-specific DNA methylation. Based on the previously developed methyl-CpG immunoprecipitation (MCIp) approach, this method should be modified and adapted to enable the comparative analysis of differential DNA methylation in regions with low or intermediate CpG content, which is difficult to study with existing technologies. One pilot study should compare the methylation profiles of three human tissues and correlate methylation differences with transcriptional activities of the corresponding genes. The second, major study should compare the methylation profiles of a defined cell type (bone marrow-derived macrophages) from two different inbred mouse strains. This study should aim at identifying interindividual differences in DNA methylation and provide details on how these differences are established in individuals.

# 3 Material

## 3.1 Equipment

| | |
|---|---|
| 8-Channel Pipettor MATRIX Impact2 Equalizer 384 | Thermo Fisher Scientific, Hudson, USA |
| Autoclave | Technomara, Fernwald, Germany |
| Biofuge fresco | Heraeus, Osterode, Germany |
| BioPhotometer | Eppendorf, Hamburg, Germany |
| Camera | Polaroid, Cambridge, USA |
| Densitometer | Molecular Dynamics, Krefeld, Germany |
| Electrophoresis equipment | Biometra, Göttingen, Germany |
| Heat sealer Fermant 400 | Josten & Kettenbaum, Bensberg, Germany |
| Heatblock | Stuart Scientific, Staffordshire, UK |
| Incubators | Heraeus, Hanau, Germany |
| Laminar air flow cabinet Lamin Air HA 2472 | Heraeus, Osterode, Germany |
| Luminometer Sirius | Berthold Detec. Systems, Pforzheim, Germany |
| MassARRAY Compact System | Sequenom, Hamburg, Germany |
| MassARRAY MATRIX Liquid Handler | Sequenom, Hamburg, Germany |
| MassARRAY Phusio chip module | Sequenom, Hamburg, Germany |
| Megafuge 3,0 R | Heraeus, Osterode, Germany |
| Microarray hybridisation chambers SureHyb | Agilent Technologies, Böblingen, Germany |
| Microarray scanner; 5 micron resolution | Agilent Technologies, Böblingen, Germany |
| Microarray slide holder | Agilent Technologies, Böblingen, Germany |
| Microscopes | Zeiss, Jena, Germany |
| Multifuge 3S-R | Heraeus, Osterode, Germany |
| Multipipettor Multipette plus | Eppendorf, Hamburg, Germany |
| NanoDrop | PeqLab, Erlangen, Germany |
| PCR-Thermocycler PTC-200 | MJ-Research/Biometra, Oldendorf, Germany |
| PCR-Thermocycler Veriti 384 well | Applied Biosystems, Foster City, USA |
| pH-Meter | Knick, Berlin, Germany |
| Picofuge | Heraeus, Osterode, Germany |
| Power supplies | Biometra, Göttingen, Germany |

| | |
|---|---|
| Realplex Mastercycler epGradient S | Eppendorf, Hamburg, Germany |
| Sigma 2 - Sartorius | Sartorius, Göttingen, Germany |
| Sonifier 250 | Branson, Danbury, USA |
| Sorvall RC 6 plus | Thermo Fisher Scientific, Hudson, USA |
| Spectra Fluor Plus | Tecan, Salzburg, Austria |
| Spectrophotometer | Perkin Elmer, Überlingen, Germany |
| Speed Vac | Christ, Osterode, Germany |
| Thermomixer | Eppendorf, Hamburg, Germany |
| Typhoon 9200 | Molecular Dynamics, Krefeld, Germany |
| Water purification system | Millipore, Eschborn, Germany |
| Waterbath | Julabo, Seelstadt, Germany |

## 3.2 Consumables

| | |
|---|---|
| 384-well PCR plates | Thermo Fisher Scientific, Hudson, USA |
| 8-channel pipettor tips Impact 384 | Thermo Fisher Scientific, Hudson, USA |
| adhesive PCR sealing film | Thermo Fisher Scientific, Hudson, USA |
| Cell culture flasks and pipettes | Costar, Cambridge, USA |
| CLEAN resin | Sequenom, Hamburg, Germany |
| Cryo tubes | Nunc, Wiesbaden, Germany |
| Eppendorf reaction tubes | Eppendorf, Hamburg, Germany |
| filter tubes: Millipore Ultrafree-MC | Millipore, Eschborn, Germany |
| Heat sealing film | Eppendorf, Hamburg, Germany |
| Luminometer vials | Falcon, Heidelberg, Germany |
| MATRIX Liquid Handler D.A.R.Ts tips | Thermo Fisher Scientific, Hudson, USA |
| Micro test tubes (0.5, 1.5, 2 ml) | Eppendorf, Hamburg, Germany |
| Microarray gasket slides | Agilent Technologies, Santa Clara, USA |
| Multiwell cell culture plates and tubes | Falcon, Heidelberg, Germany |
| nProteinA Sepharose 4 FastFlow | GE Healthcare, Munic, Germany |
| PCR plate Twin.tec 96 well | Eppendorf, Hamburg, Germany |
| rProteinA Sepharose 4 FastFlow | GE Healthcare, Munic, Germany |
| Sepharose Cl-4 beads | Sigma-Aldrich, Munic, Germany |
| SpectroCHIP bead array | Sequenom, Hamburg, Germany |
| Sterile combitips for Eppendorf multipette | Eppendorf, Hamburg, Germany |
| Sterile micropore filters | Millipore, Eschborn, Germany |

| | |
|---|---|
| Sterile plastic pipettes | Costar, Cambridge, USA |
| Syringes and needles | Becton Dickinson, Heidelberg, Germany |

# 3.3 Chemicals

Unless noted otherwise, chemicals used in this thesis were purchased either from Sigma Chemicals (Deisenhofen, Germany) or from Merck (Darmstadt, Germany). Readymade buffers and cell culture media were obtained from Biochrom (Berlin, Germany) or from Invitrogen (Karlsruhe, Germany). Water was generally of Millipore-purified/distilled quality. If denoted, ultra-pure, DEPC-treated $H_2O_{USB}$ purchased from USB Corporation through Amersham (Braunschweig, Germany) was used.

# 3.4 DNA oligonucleotides

Utilized DNA oligonucleotide primers are listed according to species and application. All standard primers were purchased from Metabion (Martinsried, Germany) or Operon (Cologne, Germany) at a synthesis scale of usually 0.02 µmol and HPLC purified. Primers for the MassARRAY application were ordered from Sigma (Deisenhofen, Germany) in 96-deep well plates.

## 3.4.1 Human

### 3.4.1.1 Real-time primer for MCIp

| Gene | Primer sequence (sense & antisense) |
|---|---|
| *AKAP3* | 5`-AAATGAATGGTCAACAGGACAAAGAGG-3` |
| | 5`-CCCAGTCCTTCTAAACTAAACATCAGCC-3` |
| *APBA2* | 5`-ATTTACGCGTTGCTGAATCTCCTG-3` |
| | 5`-CGGGTTCTACTCACTCAGAATCTCCA-3` |
| *AURKC* | 5`-AATATTTGGCGAGTGGGAGGAGAG-3` |
| | 5`-GCTTCAAGGTCACGTCCGATTCAG-3` |
| *BACH* | 5`-GGCTCAGCTTCATTACCCTCC-3` |
| | 5`-AAAGCCCGGAGCTAACGAC-3` |
| *CORT* | 5`-GGCACAATTTCACATTCAAGCTCCT-3` |
| | 5`-AGTTCTGCTTCCTCGTAATTATGCC-3` |
| *DAZ* | 5`-CAGAGGCAGTGTTTCACCCACC-3` |
| | 5`-TCCTTTGACCACTCGAAGCCC-3` |

| Gene | Primer sequence (sense & antisense) |
|---|---|
| *DAZL* | 5`-CAGGAAAGCCGAGGATGACTTCAC-3` |
| | 5`-ATCATGGTGAGTTGAGGGAGAGG-3` |
| *DCX* | 5`-TCTGAGCTCCAAGCAAGAAATTCCTG-3` |
| | 5`-GATTCACTAAGCCTGAGTTACATGGGAG-3` |
| *DDX3Y* | 5`-AATCTGGACTAGAGGGAGGACTGAG-3` |
| | 5`-CTGAAGGAACCACTACACCAAAGTTGTC-3` |
| *DMRT1* | 5`-CCCAGGCCGCTCTGTATCTC-3` |
| | 5`-GACTCCAGCCACCGCTCTC-3` |
| *EIF1AY* | 5`-AGCTGGGAATGCCCTATTTGGTC-3` |
| | 5`-GTTAAGGTCCCTCAACCGTCCC-3` |
| *FGF6* | 5`-AAAGCCTCCATCGGGCACTC-3` |
| | 5`-ATCATCAACTAGGGACGGAGCA-3` |
| *H1FNT* | 5`-ACTAAAGAAAGCGGTGGAGATCGAG-3` |
| | 5`-TCCTTCTAGGTCTAAACCAATCCAAGTC-3` |
| *LMO2* | 5`-AGACGCAATGCAGTTCCATGG-3` |
| | 5`-ATCTCCCTACACCGTGAACCC-3` |
| *MAGEB10* | 5`-ATGTCTACATTTCGCCAGTGAAGAG-3` |
| | 5`-TTCCAGGCTTGATGGGAGAAGAG-3` |
| *MGP* | 5`-TTTCCCAAACAGTCATTCCTTCCAG-3` |
| | 5`-GTAACACAGTCAGGTCCAGCAG-3` |
| *MIRN127* | 5`-TTTGTGAACCACTCGGCTCC-3` |
| | 5`-TGACCAACAGACAGGACTACATACAG-3` |
| *MIRN142* | 5`-GCCAGACTTGCCTCCTCTACCT-3` |
| | 5`-GAGGCCCTAGTCTCTACCTGAGTG-3` |
| *MIRN338* | 5`-GATATTGTTGGAGAGGACAGCCGT-3` |
| | 5`-CTGGAAGAAGTGGCGAAGGAC-3` |
| *MIRN363* | 5`-AAATTGCATCGTGATCCACCCGAC-3` |
| | 5`-CCGGCCTGTGGAAGAAAGGA-3` |
| *MORC1* | 5`-AAGCTCCCACTGCGTCGTC-3` |
| | 5`-AGATACGTCATTCTCAAATTGTGGCTC-3` |
| *NLGN4Y* | 5`-CAGATGACGAGTGGGTTGGG-3` |
| | 5`-GAGACCGCCTGAAGGATGAGAG-3` |
| *PRAME* | 5`-GAAGTACTCCGCCTCCACAAACTC-3` |
| | 5`-TGTTGCTCTTAGCCACCATGCC-3` |
| *PRM1* | 5`-GTGAGCAGGTGGAACTCTGTGG-3` |
| | 5`-GGTTTGTGAGGTCCCAGCCC-3` |
| *PRY* | 5`-ACACCTGCATGTCTTGGCCC-3` |
| | 5`-CTGTTGCTTTGTTTGTGCTGTGG-3` |
| *RBMX* | 5`-GCAACGAGGGCGAACAAAGG-3` |
| | 5`-GCGTTTCAAATTCTAAGCTCCGTTTCC-3` |
| *RBMY1A1* | 5`-CGCGTGTGCCTTAATCCACC-3` |
| | 5`-AGTTACTGCAAAGGCTGTGGG-3` |
| *SKD3* | 5`-TCCACAGCCCTAAGCCTTTCC-3` |

| Gene | Primer sequence (sense & antisense) |
|---|---|
| | 5`-CCCGAACTCTATTTCCCAATATGCC-3` |
| *SNRPN* | 5`-TACATCAGGGTGATTGCAGTTCC-3` |
| | 5`-TACCGATCACTTCACGTACCTTCG-3` |
| *SPTBN2* | 5`-GACAGCAGTGTGCAGCCC-3` |
| | 5`-TGCCACTTTCTAGTCATTATTGCCTG-3` |
| *SYBL1* | 5`-GCTCTTCATTTATGCCGCCTACTC-3` |
| | 5`-ACAGCTCTTGTGGTAAACAGCGTC-3` |
| *TAF7L* | 5`-ATCACCCTAGTGGCTCTTCTGC-3` |
| | 5`-CTTTTATGGGAACTGAGCTATGTGTC-3` |
| *TEKT3* | 5`-CCGCCTTTCCTATCTGTGACGA-3` |
| | 5`-GCTCCCGCCTCTCCTATAACTG-3` |
| *TGIF2L* | 5`-CAGCGTGGAGAAGAAAGGATGTG-3` |
| | 5`-CTCACCTGTTGTTTCCGAGAAAGAC-3` |
| *TLR4* | 5`-CGCTATCACCGTCTGACCGAG-3` |
| | 5`-CTTTCACTTCCTCTCACCCTTTAGCC-3` |
| *TMSB4Y* | 5`-TGCTCTCAAAGGTATTGTGCGTTGTC-3` |
| | 5`-TCGGGAGATGAGTGAGCCTCAG-3` |
| *TSPY* | 5`-CGGAGCCTCTATTTGCCTCGAC-3` |
| | 5`-CGGTGCTGAGGTGGAATTGATCTG-3` |
| *TSPYL5* | 5`-AACGCTAAGAGCCATCACCACC-3` |
| | 5`-CTTTCTTCATGACAGATGCCACAAGAG-3` |
| *TTTY14* | 5`-CTGAGCAAGTTATGCGGTCGG-3` |
| | 5`-AATGCGACGTCCTCCTACTCC-3` |
| *TYROBP* | 5`-GGGTGGTGGCAGAAGGGA-3` |
| | 5`-CCTTCACGTCTAGTTGCTGGG-3` |
| *UTY* | 5`-CCGTGAGGTCTCAGTAAGAGTAGG-3` |
| | 5`-TACTGCATCTTGAAAGTGTGATGAAGG-3` |
| *VCX* | 5`-TGTAGCCCTCCATCCATCCC-3` |
| | 5`-CCACACCTGCCATCATTACCC-3` |
| *VCY* | 5`-GTCTTCGTACAGCCCTCCATCC-3` |
| | 5`-GTTCCTCCACAACTACCATCATTACCC-3` |

### 3.4.1.2 Nested amplification of bisulfite-treated primer

| Gene | | Primer sequence (sense & antisense) |
|---|---|---|
| *ZSCAN5 (down)* | outer | 5`-ATTTTTATTGGTTTAGTAGGTTGATG-3` |
| | | 5`-CTTTACACCCCATAATTAACTCT-3` |
| | inner | 5`-TTAAAAAGGGTGTGTGGTAGGT-3` |
| | | 5`-AAAAATAATACTAACTTTCTCATCAAACT-3` |
| *ZSCAN5 (up)* | outer | 5`-TTAGGAATTTTAGGTTAGTTTGGA-3` |
| | | 5`-CAAATCTATATTTTCTAAAACCACC-3` |

| Gene | | Primer sequence (sense & antisense) |
|---|---|---|
| | inner | 5`-AATTTAAGTTTGATGAGAAAGTTAGT-3` |
| | | 5`-CAAATACTTCAACAATATAACAACTC-3` |
| *MAGEB10* | outer | 5`-TTAGTGAAGAGATTTTATAGTTATAATGT-3` |
| | | 5`-TCTCTAATAACTTAATACCAACACCT-3` |
| | inner | 5`-TTTGTTTTGGGTATTTTATTTTTGTG-3` |
| | | 5`-TACTATCACTTCCCTAACTCTCCA-3` |
| *TGIF2LY* | outer | 5`-ATAGAGAAGAAAAATTTAAGTAGTAAGG-3` |
| | | 5`-TTACTATATTAAAAATCACAAACAACT-3` |
| | inner | 5`-TTTGAATATGGAAGAATGGAAGAGG-3` |
| | | 5`-AACAAATCTATAAAAAAAATTCTCACCT-3` |
| *BACH* | outer | 5`-TTTTAGTAATGTAGGTTTTATAGGGTGG-3` |
| | | 5`-CTATAAACCAAAACCTCTCTACCCA-3` |
| | inner | 5`-TAGGGTGGGAGATTAGTTTTGGG-3` |
| | | 5`-AACTCATTCTTCTACTTACCTAAAAACCA-3` |
| *TLR4* | outer | 5`-TAGGTAGAGGTTAGATGATTAATTGGG-3` |
| | | 5`-ACAAAAATAAACAAACCAAAACACAC-3` |
| | inner | 5`-TATTAGAGAAATAAAAGTTTAGAATGTTAAGG-3` |
| | | 5`-AAAAAAACTAACTCCAACCACATACCT-3` |

## 3.4.2 Murine

### 3.4.2.1 MassARRAY Primer

| Gene Symbol | Primer Ids | | Primer Sequences |
|---|---|---|---|
| *1600021P15Rik* | Epi01 | 10F | 5'-aggaagagagGGATTTATAATAGGTTATGGAATGATGAAG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAAAAAATTTCCCCAAATATAAAAAACTC-3' |
| | Epi02 | 10F | 5'-aggaagagagTTTGTTAATGGGAATAAAGTTGGTTTG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAAAAAACCCAAACACACAACTCTCC-3' |
| *Asb4* | Epi03 | 10F | 5'-aggaagagagGATATGAAAGGTAGATGGTTTGGGTATTAG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctTTTTCTCCTTCCACCAATTAAAAAC-3' |
| | Epi04 | 10F | 5'-aggaagagagAGGAAAGTTTGTTTTGGTATTATGGG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctTCTCTACTAAAATTTTACCATAAATTTACC-3' |
| | Epi05 | 10F | 5'-aggaagagagGGAGGTGGGTTTTGTATTTTTTTAAAAAG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctATCCCAAAATACCCTTTAACCCTCA-3' |
| | Epi06 | 10F | 5'-aggaagagagAGTGGGGGTATTAGTTGTGTGTAGG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctTCCCACATAAAAAACAATTTTCTTCC-3' |
| | Epi07 | 10F | 5'-aggaagagagGTTTTTATAAATGATTGAAATGGTTGTTTG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCATTATACCAAAAACCAAAAACCTCC-3' |
| | Epi08 | 10F | 5'-aggaagagagGAGGATGAATTTTTTTGTTGAAATTTTATT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCAAACTTATCTTACTTTCCCACAACCC-3' |
| *Coro2a* | Epi09 | 10F | 5'-aggaagagagTTGGTTTTGGTGGATGAATTTTTTATT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCATCTCATCCCAAACCCTTATTCTAAA-3' |
| | Epi10 | 10F | 5'-aggaagagagTTTAGAATAAGGGTTTGGGATGAGATG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCCCTATAACCCCAACCTAAAACTTC-3' |

| Gene Symbol | Primer Ids | | Primer Sequences |
|---|---|---|---|
| *Coro2a* | Epi11 | 10F | 5'-aggaagagagTTAGGTTGGGGTTATAGGGTAAGGTTTT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCCCCACTAAAAAACAAACTCAACAC-3' |
| | Epi12 | 10F | 5'-aggaagagagTAGAGGGTTGGGTGGGTATTATAGG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctTTCAAATCAAAAAATCCACAACTCC-3' |
| *Eps8l1* | Epi13 | 10F | 5'-aggaagagagTTTGTTGTTTTTTTAGGATTGTGGAG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctACCTCCAACTCTAATACTTTTTCTCTATCC-3' |
| | Epi14 | 10F | 5'-aggaagagagTTTTGTGTTAATGAGGAGTTGTAGT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAAACCCTTAATCTACAACCAAACAC-3' |
| | Epi15 | 10F | 5'-aggaagagagTTTGTGGTTTTTAATTAGAGTTTGTTTTTG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctTCCCTAACCCCAACACCTAAAATAC-3' |
| *Frap1* | Epi16 | 10F | 5'-aggaagagagAGGTTTTGGTTTATTTGGGTTGAGAT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAAAATCAAACCCACTCACTACTACCC-3' |
| | Epi17 | 10F | 5'-aggaagagagAGAATTTGATAGAAGTGTGGATGATTAATT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCAAACAAAATAAAAAACTTTCCCAACTC-3' |
| | Epi18 | 10F | 5'-aggaagagagTTGAGGGATTAGGGTAGGTAAAGTTGTAG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctACCCAAAACTCAAACTCAAAAAAAA-3' |
| *Isoc2b* | Epi19 | 10F | 5'-aggaagagagGGGGATATAGATGGGAGAGGATTTA-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAACTATTTACAAACCCACCATCCC-3' |
| | Epi20 | 10F | 5'-aggaagagagGTTTAATTTTGGTTAGTGTAGGGTTTATGG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCACCCAAAAAAACAAATCCTAAATTC-3' |
| *Isoc2b* | Epi21 | 10F | 5'-aggaagagagGGTGGGTTTGTAAATAGTTAGTAAGTAA-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAACTCTCAACCAAAACCCTCCCT-3' |
| | Epi22 | 10F | 5'-aggaagagagTTGGTAAGTAGTGGGGGAAAATGTTT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAATTCATAACACCATCCACTCCATC-3' |
| | Epi23 | 10F | 5'-aggaagagagTTATAAAGGGTAAAAAGGAGGTTTAGTA-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctTTTTTCTCCAAAATTAACTCACTAAC-3' |
| | Epi24 | 10F | 5'-aggaagagagGTGGATGGTGTTATGAATTTTATATTAGTT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctACCTCCTTTTTACCCTTTATAAACAC-3' |
| *Pdgfrb* | Epi25 | 10F | 5'-aggaagagagTTGTTGGAGATATTGGGAGATGTAGAA-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCCCCCACTAAAAAAAACAAAACAAC-3' |
| | Epi26 | 10F | 5'-aggaagagagGGTGATTGAAGGTTTTAGGTTGTTTG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctTCCAAAAACCCTCCACTAAAAATAAAC-3' |
| *Pik2b* | Epi27 | 10F | 5'-aggaagagagTTTAGATGATTTGGAATTTTTGGAATTATA-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctTTCCTTCCCAATAAAAAAATTAACAC-3' |
| | Epi28 | 10F | 5'-aggaagagagGGTTGTTGGGATTTGAATTTAGGATT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAAACACCTCCTATAAAACCTTTTCCC-3' |
| *Pop4* | Epi29 | 10F | 5'-aggaagagagGGTGAGGTAGAGAGATTTTGAGTTTTTAGA-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCCTATAACCATAACCCATACTCCCAAC-3' |
| *Ppp1r14d* | Epi30 | 10F | 5'-aggaagagagTTTTGAGAAGAGGAAGAAGGAGTTTTT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCACCCAAACACAAAACAAAAATAACC-3' |
| | Epi31 | 10F | 5'-aggaagagagTTTTATTTATTGGTTGAGATTGTTTT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAATATAAAACTCCCACCTATCTTCC-3' |
| *Rab6b* | Epi32 | 10F | 5'-aggaagagagTTAGTAAGTGGTTGGAGAGGAAGAGG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAACAACTCCACTAACCCAAAAACAATAC-3' |
| | Epi33 | 10F | 5'-aggaagagagGGTATATGGTTTGTGGATAGTGGGT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCCCTCACAAAAAACTCCATAAAAACA-3' |

| Gene Symbol | Primer Ids | | Primer Sequences |
|---|---|---|---|
| *Sfi1* | Epi34 | 10F | 5'-aggaagagagTTAGGGGGAGAAGGATGTATAGATTTTG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCTAAAAAATCACCTCCCTCAAATAC-3' |
| | Epi35 | 10F | 5'-aggaagagagTTGAGGGAGGTGATTTTTTAGTAGA-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctATCATATCTCCTACCCCTTAATTTTC-3' |
| | Epi36 | 10F | 5'-aggaagagagGAGAAGTTTTTGTTTTTTGGAAGATAGAG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctATAACAAAAAACCCCAACCCAAAAT-3' |
| | Epi37 | 10F | 5'-aggaagagagATTTTGGGTTGGGGTTTTTTGTTAT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCTCAACTCCTTACTTAATCACTTAAACTCA-3' |
| *Slc13a3* | Epi38 | 10F | 5'-aggaagagagTAGAGTTTAAATTTTAGAAGAGTGGGAGGA-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAAAACCCCTTCCCTACTCCTTATCC-3' |
| | Epi39 | 10F | 5'-aggaagagagTTTAAATTTTAGAAGAGTGGGAGGAT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctATCCCACCTAAACTTTACCCCATTT-3' |
| | Epi40 | 10F | 5'-aggaagagagTTTTAAGTTTGGAAGAGGTTGTAATTA-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctACACCAACAACAACACCAAAAAAC-3' |
| *Slc27a6* | Epi41 | 10F | 5'-aggaagagagGGAAAGTTTTTATTATTTATGAGGGGGA-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCCAATTTAACCAAACCAAACCACA-3' |
| | Epi42 | 10F | 5'-aggaagagagAAGGTTGGAGTTGAGGAAGGTTATTA-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCTAAACCACACCAAAAAACAACCAC-3' |
| | Epi43 | 10F | 5'-aggaagagagATGTTGTGGGTTTATTAGGTGTTTA-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctTTACAAAATTACAACAAATCAATTATTAAC-3' |
| | Epi44 | 10F | 5'-aggaagagagTTAAGTGTTGTTATATTTAAAGTAGGGGAG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAAAAAAAATCTCAAACAAAACAAATAAC-3' |
| *Spint1* | Epi45 | 10F | 5'-aggaagagagTGTGGAGTGTAGGGAGTAGGATATATAGG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctATTCCAAAAAAATCCCCATTCTAAC-3' |
| | Epi46 | 10F | 5'-aggaagagagGGGGGATGTTGTGAAGTAGGTATTTTAT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAAAAACCCAAACCACCCATTCTAA-3' |
| | Epi47 | 10F | 5'-aggaagagagGGGTAGGGTTTAGGTGTTGTGTGTATT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCCAAACCACCCATTCTAAAAAAAATT-3' |
| | Epi48 | 10F | 5'-aggaagagagTATTGGGTTGGAGTTTTTTTGTTTG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctATCCTAAAACCAACCTTCCTCTTTC-3' |
| *Zfp568* | Epi49 | 10F | 5'-aggaagagagTTTTGGGAAATGTAGTTTATAAAGGGATAG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAACACAAATACTCCCAAAAAAACCC-3' |
| | Epi50 | 10F | 5'-aggaagagagATGTTTTAGAGTATAGGTGTTTTTAGAAGG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctATCCACAAAAAAACAACCACCTAAC-3' |
| *3110007F17Rik* | Epi51 | 10F | 5'-aggaagagagGGTTTAGTTTTTTGGTGAATGTTTTGT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctTACTCTCCCCATATCCTACCTACAC-3' |
| | Epi52 | 10F | 5'-aggaagagagGTGTAGGTAGGATATGGGGAGAGTA-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCCCACCATAACAATACACCAAAAC-3' |
| | Epi53 | 10F | 5'-aggaagagagAATTTTAGGGTTTAGAAAGTTAGATGTTTG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctACCCACTAATCTAACTCCTTAATAAATACC-3' |
| *Pop4* | Epi54 | 10F | 5'-aggaagagagGTTTTTTAAGTGTTGGGATTAAAGG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCAAAAATTCAAAAAAACAAAAATTAC-3' |
| *Zfp568* | Epi55 | 10F | 5'-aggaagagagGAGAATGGAGATTAGGAGAGAATTGGTATT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctTCCCAAAAAAACTAATAACCTATCCAAC-3' |
| *3110007F17Rik* | Epi56 | 10F | 5'-aggaagagagTTTTAGGGTTTAGAAAGTTAGATGTTTGT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctACCCACTAATCTAACTCCTTAATAAATACC-3' |

| Gene Symbol | Primer Ids | | Primer Sequences |
|---|---|---|---|
| *3110007F17Rik* | Epi57 | 10F | 5'-aggaagagagGTTGGGGATTGTTTTTTTTGGTTAG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAAACATCTAACTTTCTAAACCCTAAAATTT-3' |
| *Zfp568* | Epi58 | 10F | 5'-aggaagagagTTTTGGAATTTATTTTGTAGATTAGGTTGG-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctCATTTCCCAAAAAAACTAATAACCTATCC-3' |
| | Epi59 | 10F | 5'-aggaagagagGAAAGGTTAGTGGTTTGTTTAGTTTTT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctACCCTAACTATCCTAAAACTCACTCTATAA-3' |
| *Sfi1* | Epi60 | 10F | 5'-aggaagagagATTTTGGGTTGGGGTTTTTTGTTAT-3' |
| | | T7R | 5'-cagtaatacgactcactatagggagaaggctAAC**R**CAAC**R**TAAACTCAATATACCAAC-3' |

### 3.4.2.2 Real-time PCR primer for MCIp

| Gene | Primer sequence (sense & antisense) |
|---|---|
| "empty" | 5'-GGTGAGTTGTATGACCTTGTTCAATTCC-3' |
| | 5'-AGTGCAGTGAGTCAGACATAACCC-3' |
| *Mest* | 5'-CAGACGCCACCTCCGATCC-3' |
| | 5'- GGCCGCATTATCCCATGCC-3' |
| *Snrpn* | 5'- ACATTCCGGTCAGAGGGACGAAG-3' |
| | 5'- CCGCAATGGCTCAGGTTTGTC-3' |

### 3.4.2.3 Real-time PCR primer for validation of CNV

| Gene | Primer sequence (sense & antisense) |
|---|---|
| *Camk2b* | 5'-CTTCACTTTCTCTTTCTGTCTTTGCCTC-3' |
| | 5'-TAAATAATCCATCTCGCCATAGTGCCAG-3' |
| *Chst10* | 5'-CAGTTGAACCAGGTTTGTGAGGGAG-3' |
| | 5'-ATGAGGTGGGAAGCTAAATGGTATGTG-3' |
| *Marco* | 5'-GAATGCAAACCAGGATGCCCAC-3' |
| | 5'-GGGTAGGTGAGTAGTGGATGGAG-3' |
| *Chia* | 5'-TTTACTGCCACTGCTTCATCAAGCC-3' |
| | 5'-CACTTCCTGGAGGAGTAGTCACTG-3' |
| *Chi3l3* | 5'-TCAGTTTATGTCCAAAGTGCACATACC-3' |
| | 5'-GAACACAGCACACTGTCCTTTCAC-3' |
| *Chi3l4* | 5'-TTTCAGTAAGGAGGCACAGGAAGAG-3' |
| | 5'-CCACAGGCTTAGCACTGATACC-3' |
| *Glp1r* | 5'-GCTCTCGGAAGGACACTCATCTC-3' |
| | 5'-CCTATGACATCACCAGAGAACAACAC-3' |
| *Dnahc8 (1)* | 5'-CTTCATCCTCCCTTCATCGCTTCAC-3' |
| | 5'-AGACAATCGCCTGCCCATCC-3' |
| *Dnahc8 (2)* | 5'-AATTCAGTTGGCAGCAGGGAG-3' |
| | 5'-ATTGTAGCTCTGTCAATCAACCTGTC-3' |
| *Glo1* | 5'-CATGAAATATTAGGACAGTGGGCAAAGG-3' |
| | 5'-GGTCATTGCTACTGTTCTCATCCC-3' |
| *Btbd9* | 5'-CTCTCCTTTCACAACCTGAGCC-3' |

| Gene | Primer sequence (sense & antisense) |
|---|---|
| *Gpb1* | 5'-TCTCCTTTCCCTCGTTACCTAGTTACC-3'<br>5'-TTTCCAGTATTCTACCTACCACACCAG-3'<br>5'-GAGGCACAATGACAGTCCCTAGAG-3' |

### 3.4.2.4 Real-time PCR primer for RT-PCR

| Gene | Primer sequence (sense & antisense) |
|---|---|
| *Abcb10* | 5'-TGCCTATGGAGCTGACAACCT-3'<br>5'-TGCAATCCTCTGCTTCTGCC-3' |
| *Adamts1* | 5'-GCTGTGATCGCATCATAGACTCCA-3'<br>5'-GCTGCCATTGTTTCTGGACCC-3' |
| *Asb4* | 5'-GCACAGCCTGAGATCTGCTACC-3'<br>5'-AATGGCTTTCGGGCAAGAGTG-3' |
| *Ccl24* | 5'-CTTGCTGCACGTCCTTTATTTCCA-3'<br>5'-GGTCAGTACAGATCTTATGGCCCT-3' |
| *Ctse* | 5'-CACCCAGTATTCCATCCATCGCA-3'<br>5'-CCATCCACAGTCAACCCTTCC-3' |
| *Gbp1* | 5'-AGCAGGAACGGAAAGAGTTAATGG-3'<br>5'-TGGTGCATGATCGAGGTGGAG-3' |
| *Ifi202b* | 5'-AAAGAAAGGCTGGTTGATGGAGAG-3'<br>5'-ATACCACCACTTTCATTGCTCCTG-3' |
| *Isoc2b* | 5'-TGGCGTCTTTCTGTCCACGAG-3'<br>5'-TTGCCTTGGAAGAAGCCTAACAG-3' |
| *Itga9* | 5'-GCATGACAACACTCTCACACTGAC-3'<br>5'-CGTCCACAGACTCGCCATACAC-3' |
| *Ltb* | 5'-AGCTACTTCCCTGGTGACCCTG-3'<br>5'-ACGACGGTTTGCTGTCATCCA-3' |
| *Marco* | 5'-GGGCAGCACAGAAGACAGAG-3'<br>5'-CGTGAGCAGGATCAGGTGGA-3' |
| *Msh5* | 5'-AGTTGTCCCACACCAAGGTCAG-3'<br>5'-GCAAGACTTTCCTCAAAGTCCACC-3' |
| *Pdgfrb* | 5'-CCAACAAGCATTGTCCGCCC-3'<br>5'-TCCATGTAGCCACCGTCACTC-3' |
| *Phemx* | 5'-AGGCAGATACTCCCTGACACCA-3'<br>5'-GCACTTCAGCAGGGCATTGAG-3' |
| *Plau* | 5'-TTCTCATGAACAGTGTATGCAGCCC-3'<br>5'-GCCTTCGATGTTACAGATAAGCGG-3' |
| *Sfi1* | 5'-TGGCACAGAGACTCAGTCGG-3'<br>5'-CTTTGCTTGTAGTGAGAAGGCCC-3' |
| *SpiC* | 5'-ACCAATCCGTACAGAACATAGCTG-3'<br>5'-AAGGTATTCAAACAGCCGAAGC-3' |
| *Trim12* | 5'-CTCTCATTGAAGAGGTGGCCCAG-3'<br>5'-AAACTCTGACCGAACATTCTGCAC-3' |
| *Vill* | 5'-CCGACTGTTCGAGTGCTCCA-3'<br>5'-ATTGCAGATCCCGGTCCCAG-3' |

### 3.4.2.5 Primer for amplification of genomic DNA for sequencing

| Gene | Primer sequence (sense & antisense) |
|---|---|
| *md_1* | 5'-GCTGAGCTGGCTAGGTAGG-3' |
| | 5'-ACGATCCTTTGTCATTAACGG-3' |
| *md_3* | 5'-CCTCATGTTTCCACACATTGAC-3' |
| | 5'-GTGCTCTGTCAATTTTGCTCAG-3' |
| *md_4* | 5'-TGCCTCCCAGTAAGTTCTATCC-3' |
| | 5'-ATTCAGTTACCTGATCATCATGG-3' |
| *md_5* | 5'-ACAGCGTAACTATGTTCCATGTC-3' |
| | 5'-ACACATACCCCATACATATAAAACAAC-3' |
| *Tbc1d2* | 5'-ATTTGGGGTTGCTACAACTG-3' |
| | 5'-CATGTTTTATGTAGCTGAGTGGC-3' |
| *Cd48* | 5'-AGGGCCATGTGAACGG-3' |
| | 5'-CAGCCCAGGAGCATGC-3' |
| *Tmtc2* | 5'-AGTACATTTACTGACACAACGTTCAAG-3' |
| | 5'-CTGGACAGAAGAGATGCTATTTGC-3' |
| *Pgbd5* | 5'-GCTGAGAACCTTGTAGAAAACG-3' |
| | 5'-GGCTGGTATGGATCCTGG-3' |
| *Tmtc2* | 5'-GATGGTGTTGAAAATTAATAAGGATG-3' |
| | 5'-GTGTGGGCAGCTATGTGAG-3' |
| *1600021P15Rik* | 5'-GCAGGAGGAGTTTCTATCCTGTGTAGTC-3' |
| | 5'-CACAGGTCCTAGCTAGCCTGGT-3' |
| *2310009E04Rik (fggy) (1)* | 5'-TGAGCTAGCACTGGCAATGAG-3' |
| | 5'-CACAAGCAGCAAACACCAGC-3' |
| *2310009E04Rik (fggy) (2)* | 5'-ATGCATAAGGACAGAACACAGG -3' |
| | 5'-ACCATGATTTGCCATCTGTG-3' |
| *3110007F17Rik* | 5'-TGCACAGTGGCTGTTTACTCTC-3' |
| | 5'-AAAGGCGCACGTCCAACTG-3' |
| *Arhgdib* | 5'-TGCTTCACCTCTGACTCTCACAC-3' |
| | 5'-CACTCAAGAAGAACCCAGAACCCA-3' |
| *Asb4* | 5'-ACCTGAAGGAACTGAGATCTGG -3' |
| | 5'-ACTAACAAGTTGATCCTGATGCC-3' |
| *Cd48* | 5'-CATCCATCAACTTTGTGCTTTGTCC-3' |
| | 5'-GTGTGCCAATACCTGAAACCTGAG-3' |
| *Coro2a* | 5'-CTGGCTTTGGTGGATGAATTCCC-3' |
| | 5'-CTTTAAAGACCTCTGGTGTAGGGCTG-3' |
| *Eps8l1* | 5'-TAGAGGCTCCCGAATCCCTG-3' |
| | 5'-CTTCTCCATCACGGCCTCCA-3' |
| *Frap1* | 5'-GGGCAGTGTTTGTTTCCTCGT-3' |
| | 5'-AGTGCCATCACCCAGTCAGATACTC-3' |
| *Isoc2b* | 5'-CATGACAGCCAGGCTCAG -3' |
| | 5'-AGACCAACTTTGAACCTCTGC-3' |
| *Pdgfrb* | 5'-CCCAATGAAGTATAGCAGTCCCAC-3' |
| | 5'-GGCCAGGGTGTATACCCATCTC-3' |
| *Pi4k2b* | 5'-AGGTAGGTGGACTCAGAGATAACAAAGG-3' |
| | 5'-ACTTGCATTTATGTCACATGCTCCTC-3' |
| *Pop4* | 5'-TTCCCACTTTCCTGTTCTCTAAGACCA-3' |
| | 5'-GGAAATCAGGCAAGAGTAGAGGGA-3' |

| Gene | Primer sequence (sense & antisense) |
|---|---|
| *Ppp1r14d* | 5'-CCTGGGCAACTTAGAAATGTCC-3' |
| | 5'-GGGTTACATCTGGAGAAGGGAG-3' |
| *Rab6b* | 5'-GGCTCTGCTAAGGTACCGTG -3' |
| | 5'-CCCTTCTTCCGTCTCAGACCCA-3' |
| *Sfi1* | 5'-GATACACCTTCAGGAACTAAACCAGA-3' |
| | 5'-CGCGACTTGGGAGTCACGG-3' |
| *Sfi1 (ps)* | 5'-TCATGGCACCATCATAATTAAGGC -3' |
| | 5'-CCACAAATGACTGTTACCAATAATTGC-3' |
| *Slc13a3* | 5'-AAAGCAGTTAGCACCCATTTAAGACAC-3' |
| | 5'-CGTCCATAATGCCAACCCTCAG-3' |
| *Slc27a6 (1)* | 5'-CTCAGCTCCAAAGTAGATCCC -3' |
| | 5'-TATTGCATACGCATATCTTACATTCC-3' |
| *Slc27a6 (2)* | 5'-GGCACACGTGATGATGTTAGC -3' |
| | 5'-ACAAATTAGCATCCTTTCTACAAGCC-3' |
| *Spint1* | 5'-ACATAGGACCACCTCCTTAGTAGAACAG-3' |
| | 5'-CTAAGTGAAAGTGGCAGGCAGG-3' |
| *Tiam2* | 5'-GAGGTCATGAAGCATCTCAAGTG -3' |
| | 5'-TGGAACGCTATGAGGCAGAC-3' |
| *Zfp568* | 5'-GGTGTGGCAGGTAGGAGG -3' |
| | 5'-CTCTCCGCAAAGCTGTGAC-3' |

### 3.4.2.6 Additional internal sequencing primer

| Gene | Primer sequence |
|---|---|
| *Asb4* | 5'-CTGGAGCTATCAAAGGCCAC-3' |
| | 5'-TCTGCGTTTGATTTCATCTAGACC-3' |
| | 5'-AAGGACTGAAATTTGACTCTCC-3' |
| | 5'-GGGTCAGTGTTCTGATTGAG-3' |
| | 5'-TCCAAATAGAGGACCTTGAGTC-3' |
| | 5'-GCATCGAGACCAGAATCATCC-3' |
| | 5'-GTCCCACTCTTAAACTTTACTGC-3' |
| | 5'-GTTGTTGCAGAGGTCCACAC-3' |
| | 5'-CATGCAAATGAAGCTGCAGGAC-3' |
| | 5'-ATGTGGGAAAGAAATATGGCAGGC-3' |
| *Isoc2b* | 5'-CTCAGTTCGCCAGATTTGATCG-3' |
| | 5'-TTCGCCTCACTCTGTTCTACTC-3' |
| | 5'-AAAGAACAAGCCTCGTAAGAC-3' |
| | 5'-CCTGGCATGAATAAAGCCT-3' |
| *Ppp1r14d* | 5'-CCACCTGTCTTCCGGGATAGCT-3' |
| | 5'-TTCTCCCTTCTCCAGATGTAACCC-3' |
| *Rab6b* | 5'-GGCTCTGCTAAGGTACCG-3' |
| | 5'-TATCTATCACATGGGAAATGCC-3' |
| | 5'-GCATATGGCCTGTGGACAG-3' |
| *Sfi1 (ps)* | 5'-CATGGAATTGGTGATATGAGCT-3' |
| | 5'-GTGGCTCACAATCACCTG-3' |
| | 5'-ATTAAGACAGCATTTACAGCACGG-3' |
| | 5'-CAGAAGAAAGCAGAAGCCC-3' |
| | 5'-AAACTAGTGGGTGGGTGTC-3' |

# 3.5 Enzymes, kits and reagents

| | |
|---|---|
| BioPrime Purification Module | Invitrogen, Karlsruhe, Germany |
| BioPrime Total Genomic Labelling System | Invitrogen, Karlsruhe, Germany |
| Blood & Cell Culture DNA Midi Kit | Qiagen, Hilden, Germany |
| BSA | Sigma, Deisenhofen, Germany |
| DNA Ladder 1 kb plus | Invitrogen, Karlsruhe, Germany |
| DNA molecular weight standard | Invitrogen, Karlsruhe, Germany |
| dNTPs | NEB, Frankfurt, Germany |
| Dual-Luciferase Reporter Assay System | Promega, Madison, USA |
| EZ DNA methylation kit | Zymo Research, Orange, USA |
| Genomic DNA isolation kit | Qiagen, Hilden, Germany |
| Glycogen | Ambion, Austin, USA |
| Human Cot-1 DNA | Invitrogen, Karlsruhe, Germany |
| In-Fusion cloning kit | Clonetech, Saint-Germain-en-Laye, France |
| Klenow Enzyme | NEB, Frankfurt, Germany |
| Klenow exo- (3'-5' exo minus) | NEB, Frankfurt, Germany |
| MinElute PCR Purification Kit | Qiagen, Hilden, Germany |
| NucleoSpin Plasmid Quick Pure | Macherey-Nagel, Düren, Germany |
| Phusion DNA Polymerase | NEB, Frankfurt, Germany |
| Phusion DNA Polymerase hot-start | NEB, Frankfurt, Germany |
| Plasmid Midi Kit | Qiagen, Hilden, Germany |
| Proteinase K | Roche, Mannheim |
| QIAEX II gel extraction kit | Qiagen, Hilden, Germany |
| QIAquick PCR Purification Kit | Qiagen, Hilden, Germany |
| QuantiTect SYBR green | Qiagen, Hilden, Germany |
| Quick Ligation Kit | NEB, Frankfurt, Germany |
| Repli-G Midi Kit | Qiagen, Hilden, Germany |
| Restriction endonucleases | NEB, Roche |
| Reverse Transkriptase SuperSkript II | Promega, Madison, USA |
| RNAse A | Qiagen, Hilden, Germany |
| S-Adenosyl-Methionin (SAM) | NEB, Frankfurt, Germany |
| Shrimp Alkaline Phosphatase (SAP) | Sequenom, Hamburg, Germany |
| SssI CpG methylase | NEB, Frankfurt, Germany |
| T-Cleavage MassCleave Reagent kit | Sequenom, Hamburg, Germany |
| T4 DNA Ligase | Promega, Madison, USA |

| | |
|---|---|
| T4 DNA Ligase buffer | NEB, Frankfurt, Germany |
| T4 DNA Polymerase | NEB, Frankfurt, Germany |
| T4 Poly-Nucleotide-Kinase | NEB, Frankfurt, Germany |

## 3.6 Molecular weight standards

DNA Ladder 1 kb Plus was purchased from Invitrogen (Karlsruhe, Germany). The Kaleidoscope pre-stained standard protein marker was purchased from BioRad Laboratories (Munich, Germany)

## 3.7 Bacterial strains and plasmids

The following bacterial strains were used:

| | |
|---|---|
| TOP10 | F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 deoR araD139 Δ(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG |
| DH10ß | F-mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 deoR araD139 Δ(ara-leu)7697 galU galK rpsL endA1 nupG |

Following plasmid was used:

| | |
|---|---|
| pCR®2.1-TOPO | Invitrogen, Karlsruhe, Germany |

## 3.8 Databases and software

| | |
|---|---|
| Adobe Illustrator CS4 | |
| Agilent feature extraction version 9.5.1 | |
| BioEdit version 7.0.9.0 | |
| BioGPS (formerly SymATLAS) | http://biogps.gnf.org/#goto=welcome |
| Corel Draw10 | |
| EpiTyper 1.0.5 | |
| GeneRunner version 3.05 | http://www.generunner.net/ |
| GEO (Gene Expression Omnibus) | http://www.ncbi.nlm.nih.gov/geo/ |
| ImageQuant 5.0 | |
| Microsoft Office 2003 and 2007 | |
| Mouse Ancestry Mapper | http://mouse.perlegen.com/mouse/index.html |
| Mouse Haplotype Block Viewer | http://mouse.perlegen.com/mouse/index.html |
| PerlPrimer version 1.1.14 | http://perlprimer.sourceforge.net |
| PubMed | www.ncbi.nlm.nih.gov/entrez |
| Reference Manager 10 | |
| RefExA | http://157.82.78.238/refexa/main_search.jsp |
| Spotfire decision site | |
| UCSC Genome Browser | http://genome.ucsc.edu/ |

# 4 Methods

Unless otherwise mentioned, all methods were based on protocols described in „Current protocols of Molecular Biology“[133] and in the “Molecular cloning laboratory manual”[134].

## 4.1 General molecular biology

### 4.1.1 Bacterial culture

#### 4.1.1.1 Bacterial growth medium

LB-medium for bacterial growth was prepared using the following protocol:

- LB broth
  - 10 g NaCL
  - 10 g Bacto Tryptone (Difco)
  - 5 g Yeast extract
    - Add $H_2O$ to 1 l
    - Adjust to pH 7.5 with NaOH, autoclave

- LB-agar plates
  - 15 g Agar
  - 10 g NaCL
  - 10 g Bacto Tryptone (Difco)
  - 5 g Yeast extract
    - Add $H_2O$ to 1 L,
    - Adjust pH to 7.5 with NaOH, autoclave
    - Cool to 50°C and add desired antibiotic
    - Pour into 10 cm petri dishes and store inverted at 4°C

*E. coli* strains (DH10ß, Top10 or Pir1) were streaked out on solid LB-agar plates with specific antibiotics and grown over night at 37°C. Then single colonies were picked and used to inoculate liquid LB broth, which were incubated with shaking at 200 rpm at 37°C over night. Different antibiotics were used for selective growth of *E. coli* strains based on the resistance provided on the plasmids used, namely: Ampicillin or Kanamycin at concentration of 50µg/ml.

#### 4.1.1.2 Preparation of chemically competent *E. coli*

The desired bacterial strain was streaked out on solid LB agar without antibiotic, grown overnight and a single bacterial colony was picked into 5 ml Ψ Broth. Bacteria were grown at 37°C with shaking to an OD 550 of 0.3 and the 5 ml culture was used to inoculate 100 ml Ψ Broth prewarmed to 37°C. After growing to an OD 550 of 0.48, cells were chilled on ice and pelleted at 1300 g and 4°C for 15 min. The pellet was loosened up by vortexing and resuspended in 30 ml ice-cold TFBI. After incubation on ice for 5 min, cells were collected via centrifugation (5 min, 4000 g, 4°C) and resuspended in 4 ml ice-cold TFBII. The suspension was dispensed into 50 µl aliquots and frozen at -80°C. Required broth and buffers:

- Ψ Broth

| | | |
|---|---|---|
| 2 % | 20 g | Bacto Tryptone (Difco) |
| 0.5 % | 5 g | Bacto Yeast extract (Difco) |
| 0.4 % | 8.18 g | $MgSO_4*7H_2O$ |
| 10 mM | 0.745 g | KCL |

  - Adjust to pH 7.6 with KOH and add $H_2O$ to 1 l

- TFBI

| | | |
|---|---|---|
| 100 mM | 6.045 g | $RbCL_2$ |
| 50 mM | 4.5 g | $MnCl_2$ |
| 30 mM | 1.472 g | KOAc |
| 10 mM | 0.735 g | $CaCL_2*2H_2O$ |
| 15 % | 75 ml | Glycerol |

  - Adjust pH to 5.8 with 0.2 M HOAc
  - Add $H_2O$ to500 ml, filter sterile and store at 4°C

- TFBII

| | | |
|---|---|---|
| 10 mM | 1.047 g | MOPS/NaOH, pH 7.0 |
| 10 mM | 0.605 g | $RbCL_2$ |
| 75 mM | 5.513 g | $CaCL_2$ |
| 15 % | 75 ml | Glycerol |

  - Add $H_2O$ to 500 ml, filter sterile and store at 4°C

### 4.1.1.3 Transformation of chemically competent *E. coli*

Chemically competent *E. coli* were thawed on ice, 2 µl of the certain ligation reaction (e.g. from TOPO cloning) was added to the bacterial suspension, was mixed gently and incubated on ice for 30 min. Cells were then heat-shocked in a water bath at 42°C for 30 s or 90 s, depending on the bacterial strain and immediately cooled on ice for 2 min. On ice 250 µl SOC medium was added and the bacteria and subsequently incubated at 37°C for one hour with shaking in order to express the resistance. For blue/white screening of insert-containing clones after transformation (in case of TOP10 cells), 40 µl of X-gal was dispersed on a prewarmed LB-plate prior to use and incubated at 37°C for an additional 30 min. Afterwards 50-150 µl of the transformation reaction were plated and incubated at 37°C on LB-agar containing the antibiotic necessary for selection of transformed cells over night.

- SOC medium

| | | |
|---|---|---|
| 2% | 20 g | Bacto Trypton |
| 0.5% | 5 g | Bacto YeastExtract |
| 10 mM | 0.6 g | NaCl |
| 3 mM | 0.2 g | KCl |

  - Add $H_2O$ to 1 L, autoclave
  - Add to the to the cooled solution:

| | | |
|---|---|---|
| 10 mM | 10 ml | $MgCl_2$ 1 M sterile filtered |
| 10 mM | 10 ml | $MgSO_4$ 1 M sterile filtered |
| 20 mM | 10 ml | Glucose 2 M, sterile filtered |

- X-gal: 40 mg X-gal (5-bromo-4-chloro-3-indolyl-ß-D-galactoside)
  - In 1 ml DMF, store at -20°C protected from light

### 4.1.1.4 Glycerol stocks

For long-term storage, bacteria were stored in autoclaved LB-Glycerol containing 40 % LB-medium and 60 % glycerol. 250 µl of bacterial suspension was added to 750 µl of LB-glycerol, mixed and frozen at -20°C followed by storage at -80°C.

### 4.1.1.5 Plasmid isolation from *E. coli*

For plasmid isolation from single *E. coli* colonies, DNA mini-prep was carried out using NucleoSpin® Plasmid Quick Pure Kit from Macherey-Nagel following the supplied instructions. To isolate larger amounts of ultra pure Plasmid-DNA (ca. 100 µg) for

transfection experiments, plasmids were isolated using QIAGEN Plasmid Midi-Kit for endotoxin-free midipreps, according to the manufacturers' instructions.

# 4.2 Molecular technologies

## 4.2.1 Polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) allows *in-vitro* synthesis of large amounts of DNA by primed, sequence specific polymerization catalysed by DNA polymerase[135,136]. A thermo stable DNA polymerase synthesizes the sister strand of a heat denatured single stranded DNA-fragment when deoxynucleotide triphosphates are added under appropriate conditions. The polymerisation reaction is "primed" with small oligonucleotides that anneal to the template DNA strand through base pairing, giving the reaction its specificity by defining the borders of the segment to be amplified. Standard applications of PCR reactions are explained in the following and are used unless otherwise mentioned. More specialized applications are explained in more detail within the specific method.

### 4.2.1.1 Primer design

Unless otherwise mentioned sequences for generating primers were extracted using the UCSC Genome Browser. In general primers were designed using PerlPrimer Software and controlled using PCR and BLAT functions of the UCSC Genome Browser and GeneRunner Software. Following settings were used to design primers

| | |
|---|---|
| Primer Tm: | 65-68°C |
| Primer Length: | 18-28 bp |
| Amplicon size: | 80-150 bp |

Primers for cloning were designed according to the utilized cloning technique. For in-Fusion cloning, the PCR product was inserted into the target vector by recombination. Therefore the primers for amplification are tagged with sequences overlapping the target vector sequence at the desired recombination site, whereas recombination of linearised vector and PCR product is enhanced by the dry-down reaction from the In-Fusion cloning kit.

For direct cloning into the vector by restriction digest and fragment combination through ligation, PCR products were PEG-precipitated (see section 4.2.7) and the precipitate as well as the cloning vector were digested with the adequate restriction endonucleases. Restriction sites were introduced by adding the appropriate recognition sites to the primer

sequences. To prevent re-assembly of the cut vector ends, the digested vector was CIAP treated (see section 4.2.5). The cut fragment and vector were gel-purified (see section 4.2.8) and combined in a ligation reaction. 1- 5 µl of the reaction was used to transform chemically competent *E. coli* cells (see section 4.1.1.3).

#### 4.2.1.2 Standard PCR for cloning or sequencing of gDNA

Standard PCRs were generally performed in PCR tubes in 20-100 µl reaction volume in a MJ Research PTC 200 thermocycler (BioZym). The "calculated temperature" feature was used to decrease temperature hold times and additionally the lid was heated to 105°C to prevent vaporisation. The nucleotide sequences of the utilized primers are given in section 3.4. The primer annealing temperatures varied between 57 and 66°C. This method was used to generate PCR products for sequencing or cloning reactions. Slightly different protocols were used for the MassARRAY application and for real-time PCR. For a typical reaction the Phusion™ Hot Start High-Fidelity DNA Polymerase (Finnzymes; Espoo, Finnland) was used with the following basic reaction conditions:

| *Component* | *Volume* | *Final concentration* |
|---|---|---|
| $H_2O$ | Add 50 µl | |
| 5 x Phusion HF buffer | 10 µl | 1 x |
| 10 mM dNTPs | 1 µl | 200 µM each |
| Primer S | 1 µl | 0.2 µM |
| Primer AS | 1 µl | 0.2 µM |
| Template DNA | X µl | |
| Phusion Polymerase (2 U/ml) | 0.5 µl | 0.02 U/µl |

Typical thermocycler programme:

| ***Cycle step*** | ***Temp.*** | ***Time*** | ***Number of cycles*** |
|---|---|---|---|
| Initial denaturation | 98°C | 30 s | 1 |
| Denaturation | 98°C | 10 s | |
| Annealing | 60-66°C | 20 s | 35 |
| Extension | 72°C | 30 s / 1 kb | |
| Final extension | 72°C | 5 min | 1 |
| | 4°C | hold | |

PCR-reactions were either gel purified using agarose gel electrophoresis followed by purification with a PCR-Purification kit (Quiagen) or directly used for PEG precipitation.

#### 4.2.1.3 Reverse transcription (RT-PCR)

RT-PCR was used to quantitatively measure the mRNA level of certain transcript within cells total RNA preparations. RNA was first transcribed into cDNA using 1 µg RNA and the M-MLV Reverse Transcriptase Kit (Promega) together with *random decamers* (Ambion) using following schema:

| | | | |
|---|---|---|---|
| **RT-PCR reaction:** | 1 µg | Total RNA | |
| | 1 µl | Random Decamers | |
| | 1 µl | dNTP's (10 µM each) | |
| | | • Ad $H_2O$ 15 µl | → 5 min 65°C |
| | | • Cool on ice and spin down | |
| | 4 µl | M-MLV Buffer (5x) | → 2 min, 42°C |
| | 1 µl | M-MLV reverse transcriptase | → 50 min, 42°C |
| | | | 15 min, 70°C |

#### 4.2.1.4 Real-time quantitative PCR analysis

Real-time polymerase chain reaction, also called quantitative real time polymerase chain reaction (Q-PCR/qPCR) is laboratory technique based on the polymerase chain reaction, which is used to amplify and simultaneously quantify a targeted DNA molecule. It enables both detection and quantification (as absolute number of copies or relative amount when normalized to DNA input or additional normalizing genes) of a specific sequence in a DNA sample. The procedure follows the general principle of polymerase chain reaction; its key feature is that the amplified DNA is quantified as it accumulates in the reaction in *real time* after each amplification cycle. On the one hand, this method was used to quantify cDNA after reverse transcription (see section 4.2.1.3).On the other hand to quantitatively analyse genomic DNA in order to detect copy number variations or after fragmentation by Methyl-CpG-immunoprecipitation (MCIp, see section 4.3).

Real time-PCR's were performed in 96-well optical plates (Eppendorf), in a volume of 10 µl per well, using the QuantiFast SYBR Green PCR Kit (Qiagen). cDNA was used 1:5 diluted with $H_2O$. Reaction conditions were as follows:

| *Component* | *Volume* |
|---|---|
| SYBR Green Mix (2x) | 5 µl |
| Primer S (10 µM) | 0.5 µl |
| Primer AS (10 µM) | 0.5 µl |
| $H_2O$ | 2 µl |
| Template DNA | 2 µl |

Typical Two step real-time programme:

| *Cycle step* | *Temp.* | *Time* | *Number of cycles* |
|---|---|---|---|
| Initial denaturation | 95°C | 5 min | 1 |
| Denaturation | 95°C | 8 s | 50 |
| Annealing/Extension | 60 | 20 s | |
| Melting curve | 95°C | 15 sec | 1 |
| | 65°C | 15 sec | |
| | 65-95°C | 10 min | |

For all primer pairs a standard curve was performed by diluting an appropriate sample (1:1000; 1:100; 1:50; 1:10, 1) to measure the efficiency of amplification. SYBR green is a fluorescent dye which is able to intercalate into double stranded DNA. After each cycle the fluorescence is measured in all wells. For each sample a “ct-value” is created which refers to the reflection point of the sigmoidal curve after plotting the fluorescence signal of the exponential PCR-amplification against its cycle number. The points of the standard curve were used to generate a linear fitting. The slope of this straight line and its intercept (calculated value for the undiluted standard) was used to calculate the particular DNA amounts of the certain samples relative to the undiluted sample using following formula:

$$(10^{\frac{-1}{s}})^{(Y-ct)}$$

S=slope; Y=intercept; ct=ct-value

All measurements were performed in duplicates and repeated two to three times. The specificity of the PCR was controlled by melting curve analysis. Expression values were normalized against the expression of the housekeeper *HPRT*. PCR reactions were run and analyzed using the Mastercycler EP Realplex (Eppendorf).

Nested PCR for quantitative methylation analysis

Methylation analysis of specific DNA fragments was performed using a nested PCR after bisulfite treatment of genomic DNA (see section 4.5.3). Nested PCR means that the PCR-product of initial PCR reaction generated with an outer primer pair is used as a template for a second PCR using a different set of primers inside of the first PCR-product (inner or nested primer). A common reaction was prepared as follows:

| *Components* | *1°PCR* | *2°PCR* |
|---|---|---|
| Bisulfite DNA | 10 µl | |
| DNA from 1°PCR | | 0.5 µl |
| 10 x Taq-Buffer | 5 µl | 5 µl |
| dNTPs (10 mM each) | 1 µl | 1 µl |
| Out S (10 µM) | 2 µl | |
| Out AS (10 µM) | 2 µl | |
| In S (10 µM) | | 2 µl |
| In AS (10 µM) | | 2 µl |
| Taq | 0.5 µl | |
| FastStart-Taq | | 0.5 µl |
| $H_2O$ | Ad 50 µl | Ad 50 µl |

| *step* | *1°PCR* | | | *2°PCR* | | |
|---|---|---|---|---|---|---|
| | **Temp** | **Time** | **cycle** | **Temp** | **Time** | **cycle** |
| Initial denaturation | 93°C | 5 sec | 1 | 94°C | 3 min | 1 |
| Denaturation | 93°C | 15 sec | | 94°C | 15 sec | |
| Annealing | 55°C | 15 sec | 30 | 55°C | 15 sec | 30-35 |
| Extension | 72°C | 70 sec | | 72°C | 80 sec | |
| Final extension | 72°C | 5 min | 1 | 72°C | 5 min | 1 |

### 4.2.2 Creation of 0% to 100% methylated DNA as a control

In order to generate fully unmethylated genomic DNA *in vitro* as a control for methylation analysis (e.g. for section 4.5), genomic DNA was amplified using the REPLI-g Mini/Midi kit (Qiagen) according to the manufacturer's instructions. Purification of amplification products was performed using QIAamp DNA Micro Kit (Qiagen) as indicated in the manufacturer's manual. Secondary, to generate fully methylated DNA as a control for methylation analysis genomic DNA was methylated using SssI methyltransferase (New England Biolabs). S-Adenosylmethionin (SAM) was used as a methyl donor. In general following reaction conditions were used:

| | |
|---|---|
| DNA | 20 µg |
| NEB buffer 2 | 10 µl |
| SAM | 0.5 µl |
| SssI | 5 µl |
| H2O | Ad 100 µl |

After 2 hours incubation at 37°C the reaction was supplied with additional 0.75 µl SAM and 1 µl SssI followed by incubation of another hour.
A desired percentage of methylation was generated by mixing an appropriate amount of unmethylated (0%) and fully methylated (100%) DNA.

### 4.2.3 Molecular cloning

Direct cloning of PCR-products was done using the TOPO-TA Cloning kit (Invitrogen) according to the manufacturers' instructions. Alternatively DNA fragments were PEG-precipitated and the precipitates as well as the cloning vector were digested with the appropriate restriction endonucleases (New England Biolabs or Roche). For directional cloning, adapter sequences were added to the 5'-end of the PCR-primers containing the appropriate recognition sequences. The cut fragments and the vector were gel-purified and combined in a 10 µl ligation reaction with a 3- to 5-fold molar excess of insert to vector, using 25-50 ng of vector. Ligation was carried out over night at 16°C with 1U T4 DNA Ligase or alternatively 5 min using a rapid ligation system. Two µl of the reaction was then used to transform chemically competent *E. coli*.

### 4.2.4 Restriction digest

To verify the presence and orientation of plasmid-inserts, and to clone PCR-products into plasmids, DNA was digested with the appropriate restriction enzymes. The digestion of plasmid DNA or PCR products was carried out at 37°C for 2 hours using 5 U enzyme for every µg DNA in a volume of 20 µl.

### 4.2.5 CIAP treatment

To prevent re-ligation of restriction digested vectors, vector-ends were treated with CIAP (calf intestine alkaline phosphatise, Roche) at 37°C for 30 min prior to gel purification and cloning.

### 4.2.6 Ligation reaction

Restriction enzyme treated vectors and PCR products were ligated in a 10 µl reaction at 3- to 5-fold molar excess of insert to vector, using 25-50 ng of vector. Ligation was carried out over night at 16°C with 1 U T4 DNA ligase and 1 µl T4 DNA ligase buffer.

### 4.2.7 PEG-precipitation

To precipitate DNA from small volumes, e.g. PCR reactions or endonuclease digestion, one volume of PEG-mix was added to the DNA-containing solution. This mix was vortexed and incubated for 10 min at room temperature. After centrifugation (10 min, 13000 g, RT), the supernatant was discarded and the precipitated DNA was washed by carefully adding 200 µl 100 % EtOH to the tube wall without disturbing the pellet, followed by a centrifugation step (10 min, 13000 g, RT). After centrifugation the supernatant was carefully removed and the remaining pellet was air-dried for 5 to 10 min. The pellet was resuspended in 20 to 30 µl $H_2O$.

| | | | |
|---|---|---|---|
| **PEG-mix** | 26.2 % | 26.2 g | PEG 8000 |
| | 0.67 M | 20 ml | NaOAc (3 M) pH 5.2 |
| | 0.67 mM | 660 µl | $MgCl_2$ (1 M) |
| | Add $H_2O$ to 250 ml | | |

## 4.2.8 Agarose gel electrophoresis

Agarose gel electrophoresis was used to determine the length of linearized DNA-fragments and to separate them for further applications. It was also used to check the integrity of genomic DNA. The required amount of agarose was determined according to Table 4.1:

**Table 4.1** Agarose concentration for different separation ranges

| ***Agarose concentration in gel in %*** | ***Efficient range of separation in kb*** |
|---|---|
| 0.3 | 5-60 |
| 0.6 | 1-20 |
| 1.2 | 0.4-6 |
| 1.5 | 0.2-3 |
| 2.0 | 0.1-2 |

Agarose powder was added to TAE (1x) and the slurry was heatedin a microwave oven until the agarose was completely dissolved. Ethidium bromide was added after cooling the solution to 50-60°C. The gel was cast after hardening and mounted in the electrophoresis tank and covered with TAE (1x). DNA-containing samples were diluted 4:1 with DNA loading Dye (5x), mixed and loaded into the slots of the submerged gel. One slot was used for a 1kb size standard (invitrogen). Depending on the size and the desired resolution, gels were run at 40-100 V for 30 min to 3 hours. Required buffers:

| | | | |
|---|---|---|---|
| ○ TAE (50x) | 2 M | 252.3 g | Tris |
| | 250 mM | 20.5 g | NaOAc/HOAc pH 7.8 |
| | 50 mM | 18.5 g | EDTA |
| | Add $H_2O$ to 1 l | | |
| ○ EDTA (0.5 M) | 0.5 M | 18.6 g | EDTA/NaOH, pH 8.0 |
| | Add $H_2O$ to 100 ml | | |
| ○ DNA loading dye (5x) | 50 mM | 500 µl | Tris/HCl, pH 7.8 |
| | 1 % | 500 µl | SDS (20 %) |
| | 50 mM | 1 ml | EDTA (0.5 M), pH 8.0 |
| | 40 % | 4 ml | Glycerol |
| | 1 % | 10 mg | Bromophenol blue |
| | Add $H_2O$ to 10 ml and store at 4°C | | |

| | | | |
|---|---|---|---|
| ○ 1.0 % Agarose (w/v) | 1 % | 1 g | Agarose (Biozym) |
| | Add 1 x TAE to 100 ml and heat till agarose dissolves | | |
| | Cool to 50°C and add 1 µl EtBr per 10 ml gel | | |
| | (EtBr: 10 mg/ml; Sigma) | | |

### 4.2.9 Purification of DNA fragment by gel extraction

PCR fragments were purified by running on an ethidium bromide containing agarose gel. Specific bands were excised under UV illumination and subsequent gel extracted using QIAEX II Gel Extraction Kit (Qiagen) or NucleoSpin® Extract II following the manufacterer's instructions.

### 4.2.10 Sequencing of genomic DNA

Primers for amplifying specific genomic region were designed and controlled as described above. PCR-reactions were carried out using Phusion Hot Start High Fidelity DNA Polymerase (Finnzymes, Espoo, Finland) according to the manufacturer's instructions and 100 ng gDNA as a template. The cycling parameters were as follows: initial denaturation 98°C for 30 s, amplification at 98°C for 10 s, 62°C for 20 s, 72°C for 30 s / 1 kb, for 35 cycles. PCR-reactions were either gel purified using agarose gel electrophoresis followed by purification with a PCR-Purification kit (Qiagen) or directly used for PEG precipitation. Purified PCR-products were sequenced using the PCR primers and additional internal sequencing primers in corporation with GENART (Regensburg, Germany) and analysed using the UCSC Browser and GeneRunner and BioEdit Software, respectively.

# 4.3 Methyl-CpG-immunoprecipitation (MCIp)

In general, the MCIp procedure allows rapid enrichment of CpG methylated DNA. DNA was bound to the MBD2-Fc fusion protein produced and purified in our lab[110] (The entire procedure is also available from the lab's webpage: www.ag-rehli.de). The affinity to DNA is increased with the density of methylated CpGs and lowered with higher salt concentrations in the buffer, allowing enrichment of methylated DNA on a matrix-protein complex with increasing NaCl concentration in the wash buffer. Washing with buffers containing different NaCl concentrations and collection of according flow through's leads to the fragmentation of DNA depending on the methylation status of CpG dinucleotides.

**Required buffers:**

| | | |
|---|---|---|
| ○ TME (10x) | 200 mM | Tris-HCl (1 M) pH 8.0 |
| | 20 mM | $MgCl_2$ (1 M) |
| | 5 mM | EDTA (500 mM) |
| ○ Buffer A (300 mM NaCl) | 1x | TME (10x) |
| | 300 mM | NaCL (5M) |
| | 0.1% | NP-40 (10%) |
| ○ Buffer X (300 mM NaCl) | 1x | TME (10x) |
| | 300 mM | NaCL (5M) |
| | 0.1% | NP-40 (10%) |
| ○ Buffer B-G | 1x | TME (10x) |
| | 0.1% | NP-40 (10%) |
| | 350 (B), 400 (C), 450 (D), 500 (E), 600 (F), 1000 mM (G), | |

## 4.3.1 Preparation of MBD2-Fc Fusionprotein

### 4.3.1.1 Protein production using MBD2-Fc expressing *Drosophila* S2 cells

MBD2-Fc expressing *Drosophila* S2 cells were previously produced by our lab[110]. For large scale protein production the cells are in seeded in 2000 ml roller bottles (Corning, USA) and incubated at 21°C during permanent rolling. The protein production is separated in two phases: selection and production. For protein production the cells are spun down (without loosening the adherent cells) at 300xg for 8 minutes at 4°C. Subsequently the cells are transferred back to the same roller bottle at a density of 10-15 million cells/ml in 400 ml Insect-Xpress medium (Lonza, Walkersville, MD USA) and supplemented with 0.5 mM CuSO4. The MBD2-Fc containing culture medium is harvested after 4 days (production phase) via centrifugation at 300xg at 4°C for 8 minutes. After production phase the cells are re-cultured in 400 ml Insect-Xpress medium at a density of 3-4 million cells/ml supplemented with 400 µg/ml hygromycin B (Clontech Laboratories, USA) and incubated for 3 days in the same bottle (selection phase). A continuing protein production was maintained by alternating cycles of production and selection phases always using the same bottle.

### 4.3.1.2 MBD2-Fc protein purification

Cell culture supernatant of production phase were combined (from different bottles or production phases) and centrifuged at 2000xg for 20min at 4°C and subsequently at 15000xg, 40 60 min, 4°C in order to remove remaining cells and debris. In the following supernatants are dialyzed against 1xTBS (pH 7.4) for 72h at 4°C (whereas TBS was replaced 3 times) using Dialysis Tubing VISKING (diameter 32mm; Serva, Heidelberg). The MBD-Fc protein was afterwards affinity-purified using rProtein A Sepharose Fast Flow (GE Healthcare, Uppsala, Sweden) together with a Pharmacia column (Pharmacia, ). Constant flow rate of dialysed cell culture supernatant was achieved using a peristaltic pump (Heidolph, Schwabach, Germany). Elution was performed in 1 ml fractions with elution buffer into 1.5 ml tubes each containing 50 µl neutralization buffer. The different fractions were measured at a Biophotometer (Eppendorf, Hamburg, Germany) and protein containing fractions were combined and dialyzed against 1xTBS using Slide-A-Lyzer Dialysis Cassettes (Pierce, Rockford, USA). Dialysis was carried out for 48h and TBs was replaced once. For long term storage gelatine and sodium azid was added to final concentrations of 0.2% and 0.05%, respectively. The preserved MBD2-Fc solution was stored at 4°C. After each protein preparation the rProtein A Sepharose containing column was washed with 3 M KCL and 1xTBS, in order to remove all remaining proteins and to recover its binding capacity.

**Required buffers and solutions:**

| | | | |
|---|---|---|---|
| ○ 10x TBS | 500 mM | 302.85 g | Tris/HCl pH 7.4 |
| | 1.5 M | 438.3 g | NaCl |
| | 10 mM | 18.6 g | EDTA |
| | 0.05 % | 2.5 g | $NaN_3$ |
| | Adjust to pH 7.4 with HCl and add $H_2O$ to 5 L | | |
| ○ Elution buffer | 0.1 M | 2.9 g | Citric acid |
| | Adjust pH to 3.0 and add $H_2O$ to 100 ml | | |
| ○ Neutralisation buffer | 1.5 M | 18 g | Tris/HCl pH 8.8 |
| | Adjust to pH 8.8 with HCl and add $H_2O$ to 100 ml | | |
| ○ KCl | 3M | 223.65 g | KCl |
| | Add $H_2O$ to 1 L | | |

#### 4.3.1.3 MBD2-Fc quality and quantity assessment

Each MBD-Fc protein preparation was quality controlled and quantified using sodium dodecyl sulfate (SDS) polyacrylamide (AA) gel electrophoresis (SDS-PAGE). This technique separates proteins according to their electrophoretic mobility, which is, besides other characteristics, a function of the polypeptide chain length. A polyacrylamide gel is composed of stacking and separating gel layers that differ in acrylamide concentration. For a SDS-PAGE the protein preparation was diluted 1:5 with $H_2O$ in a volume of 10 µ and supplemented with 10 µl SDS-sample buffer. Accordingly, a bovine serum albumin (BSA) standard curve was prepared containing 4 different dilutions comprising 1, 0.5, 0.25 and 0.125 mg/ml. All samples were incubated to 95°C for 5 minutes and subsequently loaded into a SDS-PAGE assembly together with a pre-stained protein size standard (Bio-Rad Laboratories, Munich, Germany).

SDS-PAGE stock solutions:

| *Stock solution* | *Separating gel stock (13.5%)* | *Stacking gel stock (5%)* |
|---|---|---|
| Stacking buffer | - | 25 ml |
| Separating buffer | 25 ml | - |
| SDS (10%) | 1 ml | 1 ml |
| Rotiphorese Gel (30%) | 45 ml | 16.65 ml |
| $H_2O$ | Add 100 ml | Add 100 ml |

SDS-PAGE gel mixture:

| *Stock solution* | *Separating gel* | *Stacking gel* |
|---|---|---|
| Stock solution | 10 ml | 5 ml |
| TEMED | 10 µl | 5 µl |
| APS (10%) | 50 µl | 40 µl |

Prior to loading, a 13.5 % polyacrylamide SDS gel was prepared as follows: The separating gel was cast and overlaid with water-saturated isobutanol until polymerized. Isobutanol was exchanged for separating gel buffer diluted 1:3 with water. The stacking gel was poured on top of the separating gel, and the comb inserted immediately. After polymerization, the stacking gel was mounted in the electrophoresis tank, which was filled with 1x Laemmli buffer. Meanwhile, the protein and BSA solutions (prepared as described above) were loaded in the wells. The gel was run with 25 mA/110 volts till the sample buffer bands reached the surface of the stacking gel, then the current was increased to 200 volts and the gel was run for 2-3 h. Afterwards the bands were stained using Coomassie BLUE solution (Biorad) for one hour with gentle shaking. De-staining was carried out over night in desalted Millipore water. For documentation purposes the ready stained gel was scanned using a personal Densitometer SI (Molecular Dynamics). The gel image was loaded into the ImageQuant 5.0 software and protein bands were quantified using the BSA-standard curve as a reference.

**Required buffers and solutions:**

| | | | |
|---|---|---|---|
| ○ Separating buffer | 1.5 M<br>Add $H_2O$ to 500 ml | 90.83 g | Tris/HCl pH 8.8 |
| ○ Stacking buffer | 0.5 M<br>Add $H_2O$ to 500 ml | 30 g | Tris/HCl pH 8.8 |
| ○ SDS (10%) | 10% (w/v)<br>Add $H_2O$ to 100 ml | 10 g | SDS |
| ○ AP (fresh prepared) | 10%<br>Add $H_2O$ to 1 ml | 100 mg | Ammonium persulfate |

| | | | |
|---|---|---|---|
| ○ Laemmli (5x) | 40 mM | 15 g | Tris |
| | 0.95 M | 21 g | Glycin |
| | 0.5 % | 15 g | SDS |
| | Add $H_2O$ to 3 L | | |

## 4.3.2 Binding MBD2-Fc to beads

For single-gene analysis, typically 18 µg purified MBD2–Fc protein per 40 µl nProtein A–Sepharose 4 Fast Flow beads (Amersham Biosciences) were rotated in 2 ml TBS overnight at 4°C in order to bind the Fc-part of the protein to the beads. On the next day, the MBD2–Fc-bead complexes (40µl/assay) were transferred and dispersed equally into 0.5-ml Ultrafree-MC centrifugal filter devices (Millipore) and spinwashed twice with buffer A.

## 4.3.3 DNA fragmentation

Genomic DNA was fragmented as follows: To reduce viscosity, gDNA was initially sheared using a 20 gauge needle attached to a 2 ml syringe (BD) before quantification using the NanoDrop ND 1000 spectrophotometer (Peqlab). Sonication to a mean fragment size of 400–500 bp was carried out with the Branson Sonifier 250 (Danbury) using settings shown below. After sonication the sample was immediately cooled on ice. Fragment range was controlled using agarose gel electrophoresis.

| | | |
|---|---|---|
| For 5 µg DNA in 500 µl TE | Duty cycle | 30 % |
| | output | 3 |
| | Sonication time | 60 sec |
| For 2.5 µg DNA in 500 µl TE | Duty cycle | 80 % |
| | output | 0.5 |
| | Sonication time | 2*30 sec |

### 4.3.4 Enrichment of highly methylated DNA

Sonicated DNA (300 ng) was added to the washed MBD2–Fc beads in 350 µl buffer X and rotated for 3 h at 4°C. Beads were centrifuged to recover unbound DNA fragments (300 mM fraction) and subsequently washed twice with 200 µl and 150 µl of buffers containing increasing NaCl concentrations (350-1000 mM, see buffers B-G). The flow-through of each washing step was collected in separate tubes and desalted using a QIAquick PCR Purification Kit (Qiagen). In parallel, 300 ng sonicated input DNA was resuspended in 350 µl buffer X and desalted using a QIAquick PCR Purification Kit (Qiagen) as a control. This MCIp protocol was scaled up to generate DNA fragments for direct microarray hybridisation. Here, for each sample, 84 µg purified MBD2–Fc protein was added to 200 µl Protein A–Sepharose beads (Amersham Biosciences) in 15 ml TBS and rotated overnight at 4°C. For the precipitation, 2-ml Ultrafree-MC centrifugal filter devices (Millipore) were used and 2 or 4 µg of sonicated DNA was added to the washed MBD2–Fc bead complexes in 2000 µl buffer X. Unbound DNA fragments (300 mM fraction) were collected and desalted using a QIAquick PCR Purification Kit (Qiagen) Flow-throughs were analysed for successful enrichment of highly methylated DNA with the real-time PCR (see section 4.2.1.4). Based on the distribution of fragments analysed with RT-PCR primers, a conclusion could be made about the methylation level of the sample DNA in this region. For the “mirror image approach” a threshold was defined and flow-throughs were combined to a hypomethylated and a hypermethylated fraction for subsequent labeling and microarray analysis.

# 4.4 Microarray handling and analysis

To generate fluorescently labelled DNA for microarray hybridization, the hypo- and hypermethylated pools of the MCIp immunoprecipitation of C57BL/6 were directly labelled with Alexa Fluor 5. The corresponding pools from BALB/c were labelled with Alexa Fluor 3. In both cases labelling was performed with the BioPrime Total Genomic Labeling System (Invitrogen, Carlsbad, CA, USA). The labelling reaction was performed using the manufacturer's instructions. The differently labelled fractions of the two mice strains were combined to a final volume of 80 µl, supplemented with 50 µg mouse Cot-1 DNA (Invitrogen), 52 µl of Agilent blocking agent (10-fold) (Agilent Technologies, Böblingen, Germany), 78 µl Deionized formamide (Sigma) and 250 µl Agilent hybridization buffer (2-fold) as supplied in the Agilent oligo aCGH Hybridization kit. The samples are heated to 95° for 3 min, mixed and subsequently incubated at 37°C for 30 min and spun down afterwards for 1 min. Hybridization on a custom designed Mouse array (Agilent) was then carried out 65°C for 40 h using an SureHyb chamber and an Agilent hybridization oven. Slides were washed in Wash I (6xSSPE, 0.005% N-lauroylsarcosine) at room temperature for 5 min and in Wash II (0.06xSSPE) for an additional 5 min. Afterwards slides were dried and incubated using acetonitrile within an ozone free facility. Images were scanned immediately and analysed using a DNA Microarray scanner (Agilent) and processed using Feature Extraction Software 9.5.1 (Agilent).

## 4.4.1 Gene Expression analysis

Total RNA preparations from BMM cultures from two independent experiments were used to globally analyse gene expression patterns on Whole Mouse Genome Oligo Microarrays (Agilent).

### 4.4.1.1 Labelling reaction

Labelling was performed using the Agilent Gene Expression system according to the manufacturer's instructions. In brief, 200 ng to 1000 ng of high-quality RNA were amplified and Cyanine 3-CTP labelled with the One Color Low RNA Input Linear Amplification Kit (Agilent). Labelling efficiency was controlled using the NanoDrop ND-1000 Spectrophotometer (PeqLab, Erlangen, Germany).

### 4.4.1.2 Microarray hybridization

Of each sample 1.65 µg labelled cRNA was fragmented and hybridized on the Whole Mouse Genome Expression Array (4x44K, Agilent). Fragmantation mix was prepared as follows:

| *Component* | *Volume/Mass* |
|---|---|
| Labelled, linearly amplified cRNA | 1.65 µg |
| Agilent Blocking Agent (10x) | 11 µl |
| Nuclease free water | Ad 52.8 µl |
| Fragmentation Buffer (25x) | 2.2 µl |

The sample was incubated at 60°C for exactly 30 minutes in order to fragment RNA. Afterwards the fragmentation was stopped by adding of 2x Hybridisation buffer. The final hybridization mixture for the 4x44k (4 array/slide; total volume 110 µl each) Whole Mouse Genome microarrays was prepared as follows:

| *Component* | *Volume* |
|---|---|
| cRNA from Fragmentation Mix | 55 µl |
| Agilent Hybridization Buffer (2x) (2xGE, HI-RPM) | 55 µl |
| Agilent Blocking Agent (10x) | 50 µl |

The sample was spun down afterward for 1 min and kept on ice until loading onto array, which was performed as soon as possible. Hybridization on microarrays slides (Agilent) was then carried out at 65°C for 17 h using an Agilent SureHyb chamber and an Agilent hybridization oven. Slides were washed in Gene Expression Wash Buffer I (Agilent) at room temperature for one minute and in Gene Expression Wash Buffer II (Agilent, prewarmed to 37°C) for an additional minute. Afterwards slides were dried and incubated using acetonitrile for 30 s. Images were scanned immediately using a DNA microarray scanner (Agilent), and processed with Feature Extraction Software 9.5.1 (Agilent) using default parameters (protocol GE1-v5_95_Feb 97) to obtain background subtracted and spatially detrended, processed signal intensities. Extracted data were further processed with GeneSpring GX 7.3.1 as follows: Data Values below 10.0 were set to 10.0. Each measurement was divided by the 50,0th percentile of all measurements in that sample. The percentile was calculated using only genes marked present. Each gene was divided by the median of its measurements in all samples. If the median of the raw values was below 0 then each measurement for that gene was divided by 0 if the numerator was above 0, otherwise the measurement was thrown out. Median polishing was done, where each chip was normalized to its median and each gene was normalized to its median. These normalizations were repeated until the medians converged.

## 4.4.2 CGH microarrays

Samples for DNA methylation mapping were analysed using comparative genome hybridization microarrays (CGH) either using human proximal promoter microarrays or murine custom arrays, both purchased from Agilent.

### 4.4.2.1 Labelling reaction

Samples from MCIp (section 4.3) were labelled directly with Alexa Fluor® 3–aha–dCTP and Alexa Fluor® 5–aha–dCTP, respectively, using the BioPrime Plus Array CGH Genomic Labelling System (Invitrogen). The labelling reaction was carried out according to the manufacturer's Manual. Labelling efficiency was controlled using the NanoDrop ND-1000 Spectrophotometer (PeqLab, Erlangen, Germany).

### 4.4.2.2 Microarray hybridization

The differently labelled DNA fragments or pools of two samples were combined and supplemented with human or mouse Cot-1 DNA, depending on the samples used, (Invitrogen), Agilent blocking agent (10-fold) (Agilent Technologies, Böblingen, Germany), Agilent hybridization buffer (2-fold) as supplied in the Agilent oligo aCGH Hybridization Kit. For more stringent hybridisation conditions deionised formamide was additionally added in order to prevent cross-hybridisation of GC-rich DNA sequences. The correct amount differed according to the arrays format used:

**44k** (1 array/slide; total volume 500 µl) human proximal promoter microarrays

| *Component* | *1x Mix* | *Final concentration* |
|---|---|---|
| DNA samples combined in 150 µl TE | 80 µl | |
| Cot-1 DNA (1.0 mg/ml) | 50 µl | 0.1 mg/ml |
| Agilent Blocking Agent (10x) | 50 µl | 1x |
| Agilent Hybridisation Buffer (2x) | 250 µl | 1x |

**244k** (1 array/slide; total volume 500 µl) custom arrays

| *Component* | *1x Mix* | *Final concentration* |
|---|---|---|
| DNA samples combined in 80 µl TE | 80 µl | |
| Cot-1 DNA (1.0 mg/ml) | 50 µl | 0.1 mg/ml |
| Agilent Blocking Agent (10x) | 52 µl | 1x |
| Deionized formamide | 78 µl | 15 % |
| Agilent Hybridisation Buffer (2x) | 250 µl | 1x |

The sample was heated to 95°C for 3 min, mixed, and subsequently incubated at 37°C for 30 min and spun down afterward for 1 min. Hybridization on microarrays slides (Agilent) was then carried out at 65°C (or 67°C for the stringent protocol) for 40 h using an Agilent SureHyb chamber and an Agilent hybridization oven. Slides were washed in Wash I (6× SSPE, 0.005% *N*-lauroylsarcosine) at room temperature for 5 min and in Wash II (0.06× SSPE; prewarmed to 37°C for stringent protocol) for additional 5 min. Afterwards slides were dried and incubated using acetonitrile for 30 s. Images were scanned immediately and analyzed using a DNA microarray scanner (Agilent). Microarray images were processed using Feature Extraction Software 9.5.1 (Agilent) using the standard CGH protocol for samples from MCIp. Processed data was imported into Microsoft Office Excel for further analysis. Graphical presentations of datasets were obtained using Spotfire Decision Site Software 7.0 (Spotfire).

Calculation of CpG index

At a mean fragment size of 400-500 bp, a single oligonucleotide probe should detect methylation events within a radius of at least 350 bp. During the MCIp procedure the fragments of the 350 mM fraction were allowed to be methylated at one or two CpG dinucleotides. A given DNA fragment should contain at least four or five CpG residues within the above radius, in order to allow the reproducible separation of unmethylated and methylated fragments. To calculate the CpG index for each individual microarray probe, the relative chromosomal location and in a next step the genomic DNA sequence surrounding each individual probe was extracted using the UCSC Genome Browser (http://genome.ucsc.edu/). The number of CpG residues in a region 250, 350 and 850 bp upstream and downstream of each probe (CpG index) was calculated using Microsoft Exel 2007. Following analyses was restricted to probes with a CpG index above 4.

# 4.5 Quantitative DNA methylation analysis with the MassARRAY Compact System

## 4.5.1 General overview

All reagents were purchased from Sequenom unless otherwise mentioned. All centrifugation steps were carried out at 3000rpm unless otherwise mentioned. Quantitative methylation analysis of the regions of interest was performed with the MassARRAY Compact System (Sequenom). This method is based on matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) measurement of bisulphite converted DNA. Bisulphite treatment generates methylation dependent sequence variations: Cytosine (C) is deaminated to uracil (U) which generates Thymine (T) after PCR-amplification, whereas 5'-methyl-Cytosine remains unchanged. This C/T variation appears as a G/A (guanine/adenine) variation after *in vitro* RNA transcription resulting in a mass difference of 16 kDa (the mass difference between G and A) in the fragments which can be quantified with the MassARRAY Compact System. In the mass spectrum, the relative amount of methylation can be calculated by comparing the signal intensity between the mass signals of methylated and non-methylated DNA-template. Primers were designed using Methprimer software (www.urogene.org/methprimer/), which is available online. Genomic DNA (gDNA) was extracted from cell cultures of murine bone-marrow derived macrophages (BMM), spleen and testis with the DNeasy Blood and Tissue Kit (Qiagen). At the beginning, the genomic DNA was treated with sodium bisulphite (Zymo Research), where unmethylated cytosine is converted to uracil whereas 5'-methyl-cytosine is not affected. The regions of interest were then amplified by PCR with the reverse primer harbouring a T7-Polymerase promoter for *in vitro* transcription. Unincorporated deoxynucleotide triphosphates were inactivated by dephosphorylation with shrimp alkaline phosphatase (SAP, Sequenom).

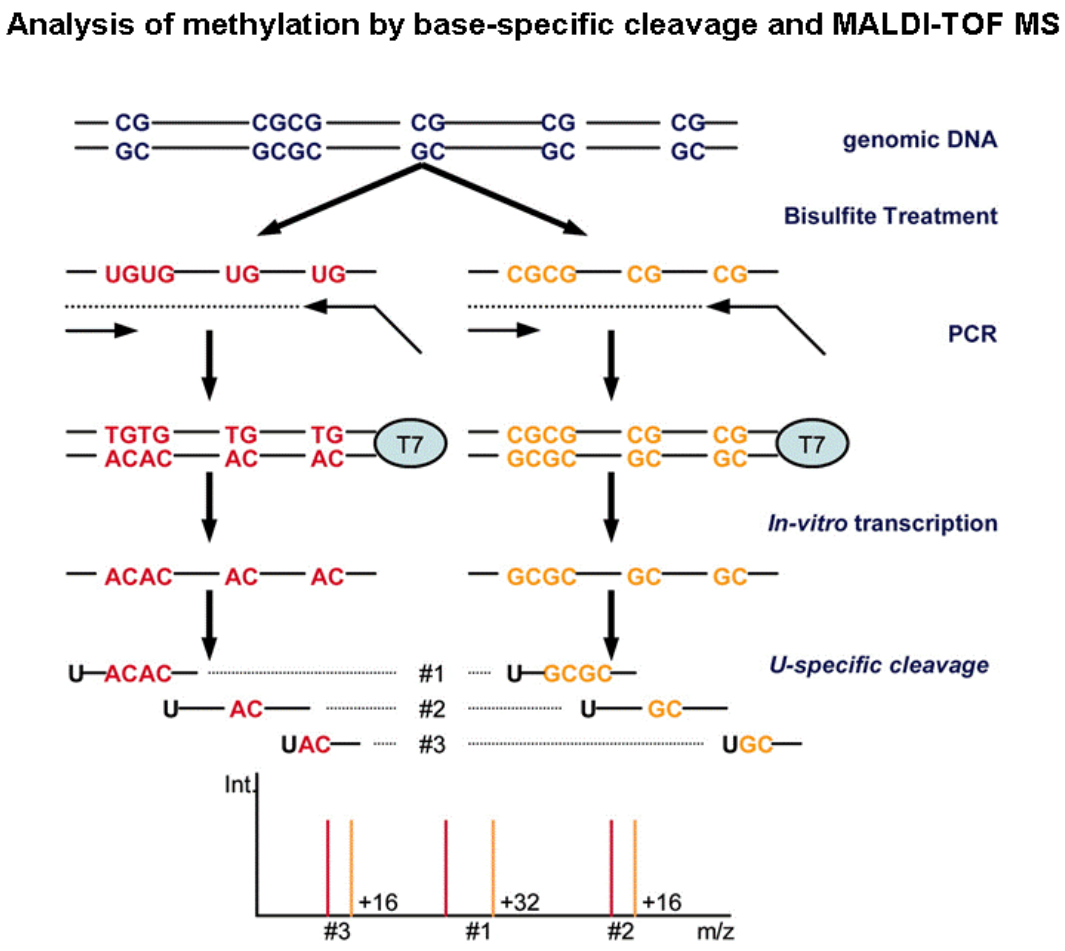



**Figure 4.1 MassARRAY workflow overview**: Genomic DNA is isolated and bisulphite treated to generate methylation specific mass differences. Regions of interest are amplified by PCR following in vitro transcription and base specific cleavage (in the figure U-specific, in the assay used for this work T-specific). Mass differences are then analysed with MALDI-TOF MS (Figure taken from www.sequenom.com).

After SAP treatment the amplicons were transcribed *in vitro* into RNA using a T7-RNA Polymerase and simultaneously “T”-specific cleavage was achieved with RNaseA and modified cytosine triphosphate nucleotides to avoid “C” cleavage within the same reaction. Water was added and reactions were desalted with 6 mg of cation exchange resin (Sequenom). Finally reactions were spotted on a SpectroCHIP with the Phusio Chip Module and measured with the MassARRAY Compact System (all Sequenom)

### 4.5.2 Primer Design

The primers were designed to amplify fragments with the limited size of 200 to 600 base pairs, because bisulphite conversion destabilises the integrity of gDNA and may lead to fragmentation of template DNA and therefore limits PCR-amplification. Unlike methylation-specific primers these primers here bind to both methylated and unmethylated template because the software excludes regions with CpG dinucleotides for potential primer binding sites. The reverse primer is tagged with a T7-promoter sequence and a short eight base pair spacer sequence while the forward primer has a ten mere tag to balance primer lengths. The primers were designed using Methprimer software (www.urogene.org/methprimer/) and purchased from Sigma in a 96 deep well system.

### 4.5.3 Bisulfite treatment of genomic DNA

A common method for analyzing cytosine methylation is bisulphite conversion of DNA followed by sequencing. Cytosine-derivates undergo reversible reactions with bisulphite yielding a 5,6-Dihydro-6-sulfonate, which deaminates spontaneously. After that the sulphate is eliminated under alkaline conditions, leaving Uracil. 5'-methyl Cytosine is not affected by this reaction and so unmethylated Cytosine appears as a Uracil in the sequencing reaction whereas 5'-methyl Cytosine remains cytosine. As already mentioned, in the following procedure this methylation specific difference is not used for sequencing but for generating methylation depending mass differences to be analysed by mass spectrometry. All reagents in this section were obtained from the EZ DNA methylation kit (Zymo). 1 µg of gDNA was brought to a volume of 45 µl and was diluted with 5 µl M-Dilution Buffer, mixed and incubated at 37°C for 15 minutes. After incubation 100 µl of CT Conversion Reagent was added, lightly vortexed and incubated in the dark with the following protocol:

| | | |
|---|---|---|
| Step 1: | 95°C | 30 sec |
| Step 2: | 50°C | 15 min |
| Step 3: | Repeat steps 1-2 for 20 cycles | |
| Step 4: | 4°C | hold |

Afterwards the samples were incubated on ice for 10 minutes, 400 µl of M-Binding Buffer was added and the sample was loaded on a Zymo-Spin I Column placed in a 2 ml collection tube. DNA was bound by centrifuging at full speed for 15-30 seconds, washed with 200 µl M-Wash Buffer, centrifuged again for 15-30 seconds and then treated with 200 µl M-Desulphonation Buffer for 15 minutes at room temperature. After incubation the

column was centrifuged for 15-30 seconds, washed twice with 200 µl M-Wash Buffer centrifuging 30 seconds and 1 minute respectively at full speed to remove wash buffer residues. To elute the DNA 100 µl water was added directly to the centre of the column and centrifuged 30 seconds at 3000 rpm. The procedure yields 100 µl of bisulphite converted DNA with a concentration of 7-8 ng/µl.

### 4.5.4 PCR-amplification

Polymerase Chain Reaction (PCR) allows the specific amplification of DNA segments (see 4.2.1). The PCR-reactions were prepared in 384 well plates (ABgene) with the following reagents according to the manufacturer:

| *Component* | *Volume for single reaction* | *Final concentration* |
|---|---|---|
| $ddH_2O$ | 1.42 µl | N/A |
| 10x HotStarBuffer | 0.5 µl | 1x |
| dNTP mix 25 mM each | 0.04 µl | 200 µM |
| 5 U/µl Hot Star Taq | 0.04 ml | 0.2 U |
| DNA Template | 1 µl | 5-10 ng |

To each reaction 2 µl primer mix was added, giving a final reaction volume of 5 µl, with the concentration of 500 pM of the forward and reverse primer. Then the plate was sealed with AB-0558 spun down, centrifuged and incubated in a Veriti 384 well thermal cycler (Applied Biosystems) with the following programme:

| *Step* | *Temperature* | *Time* | *Cycle* |
|---|---|---|---|
| Initial denaturation | 94°C | 4 min | 1 |
| Denaturation | 94°C | 20 sec | |
| Annealing | 59°C | 30 sec | 45 |
| Elongation | 72°C | 1 min | |
| Final elongation | 72°C | 3 min | 1 |
| Cooling | 4°C | hold | 1 |

### 4.5.5 Shrimp Alkaline Phosphatase (SAP) Treatment

Unincorporated nucleotides can disturb downstream applications and are therefore enzymatically inactivated. Under alkaline conditions SAP removes phosphate groups from

several substrates including deoxynucleotide triphosphates, rendering it unavailable for further polymerase catalyzed reactions. The SAP solution was prepared as follows:

| *Component* | *Volume for single reaction* |
|---|---|
| RNAse free water | 1,7 µl |
| SAP | 0.3 µl |

2 µl of the SAP solution was added to each PCR-reaction with the 96 channel pippeting robot MassARRAY Liquid Handler and Fusio$^{TM}$ Chip Module (Matrix). The plate was sealed with AB-0558, centrifuged and incubated as follows on a Veriti 384 well thermal cycler (Applied Biosystems):

| | | |
|---|---|---|
| Step 1: | 37°C | 20 min |
| Step 2: | 85°C | 5 min |
| Step 3: | 4°C | hold |

## 4.5.6 In vitro transcription and RNaseA treatment

The PCR reaction is transcribed into RNA *in vitro* with the T7 RNA polymerase, which is guided to the amplified PCR-products by the introduced T7 promoter tag in the reverse primer. The transcribed RNA is in the same reaction enzymatically cleaved by RNaseA, cleaving specifically after cytosine and thymine. T-specific cleavage is achieved by using modified cytosine triphosphate nucleotides which protect from RNaseA digestion when incorporated in an RNA polymer. The RNase and T-cleavage mix was prepared according to the manufacturer's instruction:

| *Component* | *Volume for single reaction* |
|---|---|
| RNase free water | 3.21 µl |
| 5x T7 Polymerase buffer | 0.89 µl |
| Cleavage Mix (**T mix**) | 0.22 µl |
| DTT (100 mM) | 0.22 µl |
| T7 R&DNA Polymerase (50 U/µl) | 0.4 µl |
| RNaseA | 0.06 µl |
| Total volmume | 5 µl |

5 µl of the mix and 2µl of the SAP treated PCR reaction were transferred into a new 386-well plate with the 96 channel pippeting robot MassARRAY Liquid Handler and Fusio™ Chip Module (Matrix), sealed with AB-0558, centrifuged and incubated on a Veriti 384 well thermocycler C (Applied Biosystems) for three hours at 37°C.

### 4.5.7 Desalting of Cleavage Reaction: resin treatment

Because salt ions are co-vaporised when acquired during MALDI-TOF analysis they are therefore visible in the mass-spectra. This would irritate the analysis of the mass-spectra and therefore the reactions need to be desalted. For desalting of the transcription/cleavage mix 20 µl water was added to each reaction with the MassARRAY Liquid Handler (Matrix) followed by the addition of 6 mg CLEAN resin per reaction. This mix was rotated for slowly for 10 minutes and spun down to collect the resin at the bottom of the wells.

### 4.5.8 Transfer on SpectroCHIP and acquisition

The SpectroCHIP holds the matrix on which the sample probes are spotted and consists of a crystallized acidic compound. When the analyte is spotted on the matrix its solvent dissolves the matrix, and when the solvent evaporates the matrix recrystallizes with analyte-molecules spread enclosed in the crystals. The DNA samples are transferred on a SpectroCHIP with the Phusio Chip Module and analysed with the MassARRAY Compact System MALDI-TOF MS (all Sequenom). The co-crystallized analyte is acquired with a laser while the matrix is predominantly ionized, protecting the DNA from the disruptive laser beam. Eventually the charge is transferred to the sample and charged ions are created which are accelerated in a vacuum towards a detector that measures the particle's time of flight.

### 4.5.9 Data processing

Acquired data was processed with the EpiTyper Analyzer software (version 1.0.5, Sequenom). The MS is calibrated with a four point calibrant (Sequenom) containing 1479, 3004, 5044.4 and 8486.6 kDa particles. Relative to this calibration the accelerated analytes generate signal intensity (y-axis) versus mass (kDa, x-axis) plots. With the sequence of every amplicon known, the software can virtually process the sequence and predict the fragments from the in vitro transcription/RNase digestion and relocate CpG units. If fragments contain a single CpG this is called a CpG-site. More CpG-sites within one fragment are summarized to a CpG-unit get a sum methylation value since the software averages the methylation of the individual CpG-sites. If expected and incoming information match, the signal intensities of the methylated and unmethylated DNA

templates are compared and quantified. A normal calibrated system is able to measure fragments between a range of 1500 and 7000 Dalton. Fragments outside of this range and fragments whose mass peaks are overlapping with multiple other fragments cannot be analysed.

# 4.6 Laboratory animals

Wild-type inbred mice were purchased from Charles River (Sulzfeld, Germany).

## 4.6.1 Wild-type inbred mice

Wild-type inbred C57BL/6JCrl and BALB/cAnNCrl mice were purchased from Charles River (Sulzfeld, Germany). Bone marrow derived macrophages (BMM) of different mice were produced as described in section 4.7.6. Total cellular RNA was isolated using the RNeasy Kit (Qiagen). Genomic DNA was prepared using the Qiagen Blood & Cell Culture DNA Kit. DNA and RNA concentrations were determined using the NanoDrop ND-1000 Spectrophotometer (PeqLab, Erlangen, Germany) and quality was assessed by agarose gel electrophoresis. In addition, spleen and testis DNA were prepared using a TissueLyser (Qiagen) and the DNeasy Blood&Tissue kit (Qiagen).

## 4.6.2 F1 Hybrids

F1 hybrid offspring were produced by natural matings within our animal facility. Two female mice from one strain and one male mouse from the other strain were kept together to produce hybrid mice separately for both possible sires. Different gender of F1 hybrids were separated at time of fertility. Genomic DNA and RNA from bone marrow derived macrophages, spleen and testis (RNA and DNA), were produced as described for wild-type mice.

# 4.7 Cell culturing methods

## 4.7.1 Cell culture

RAW264.7 cells were grown in 90 % 1640 RPMI plus 10 % fetal calf serum (FCS) supplied with 1 mM sodium pyruvate (Gibco), 2 mM L-Glutamine (Biochrome), 2 ml MEM Non-essential amino acids (Gibco), MEM Vitamines (Gibco), 50U/ml Penicillin/Streptomycin (Gibco), 50 µM 2-Mercaptoethanol (Gibco) in an humified incubator at 37°C and 5 % $CO_2$. Cells were passaged every two to three days using a Cell scraper (Sarstedt).

## 4.7.2 Analysing cell vitality

The numbers of viable and dead cells were determined by trypan blue exclusion. Therefore the cell suspension was diluted with trypan blue solution and counted in a Neubauer haemocytometer. Dead cells appear blue since the blue stain is able to enter the cytoplasm. The concentration of viable cells was then calculated using following formula:

Number of viable cells per ml = $Z * D * 10^4$

Z=average of unstained cells per corner square (1 mm containing 16 sub-squares); D=dilution factor

**Required solution:** 0.2 % (w/v) trypan blue in 0.9 % NaCl solution

## 4.7.3 Freezing and thawing cells

Cells were harvested and about $6*10^6$ cells per ml were suspended in 800 µl ice-cold medium, including 10 % FCS. After inverting the mix and transferring into cryo-vials(Corning), 160 µl DMSO (Sigma) and 640 µ FCS were added and the tubes were rapidly inverted to mix the cells properly. In order to allow gradual freezing at a rate of ~1°C/min, the cryo-vials were placed in isopropanol-filled cryo-containers (Nalgene) and frozen at -20°C. After 2 hours the vials were transferred to -80°C for 48 hours. Long-term storage was accomplished in liquid nitrogen (-196°C).

## 4.7.4 Mycoplasma assay

Cells were routinely checked for mycoplasma contamination by ELISA with a Mycoplasma Detection Kit (Roche, Germany) or by a biochemical test with a MycoAlert® Mycoplasma detection assay (Cambrex, Rocheland, USA) according to the manufacturer's intructions.

## 4.7.5 Isolation of human monocytes

Leukapheresis, which is a procedure to separate white blood cells from peripheral blood, was used to extract a leukocyte-enriched concentrate of blood from healthy volunteers (Graw et al, 1971). Mononuclear cells (MNC) were isolated by density-gradient centrifugation using the Ficoll-Paque technique (Pharmacia, Freiburg) and were washed three times with phosphate buffered saline (PBS) (Johnson et al, 1977). Subsequently monocytes were isolated from the MNC mixture by counterflow centrifugation (Elutriation)

(Sanderson et al, 1977). Elutriation was performed using a Beckman centrifuge (J6-MC) within a 50 ml chamber. Prior to use the assembly was sterilised by incubation with 6% $H_2O_2$-solution and washed twice for 20 minutes with PBS and HANKS (balanced salt solution), respectively. The peristaltic pump was calibrated with HANKS prior to use at a constant centrifugation force of 1500 rpm at 4°C. HANKS used in the following, was supplemented with autologous plasma, which was collected as a by-product during the leukaphoresis procedure, to a final concentration of 6%. **Table 4.2** shows an exemplary flow rate pattern which varied marginally between individual preparations.

**Table 4.2** Elutriation steps, settings and expected cells

| *Fraction* | *Flowrate (ml/min)* | *Volume (ml)* | *Cell in fraction* |
|---|---|---|---|
| 1a | 52 | 500 | Thrombocytes |
| 1b | 57 | 500 | Lymphocytes (B-and T-cells) |
| 2a | 64 | 500 | Lymphocytes (B-and T-cells) |
| 2b | 74 | 400 | Lymphocytes (B-and T-cells) |
| 2c | 92 | 400 | Lymphocytes (B-and T-cells) |
| 3 | 111 | 4*200 | Monocytes |

In this example the MNC were loaded at a flow-rate of 52 ml/min. In the following the inflow of HANKS (including 6% plasma) was continued and the flow-rate was stepwise increased. Different fractions were collected as depicted in **Table 4.2**. The biggest cells within the MNC preparation, the monocytes, eluted in the last fraction (fraction 3) using the highest flow-rate. The monocytes were subsequently pelleted at 300 g at 4°C for 8 minutes and resuspended in RPMI medium prior to counting. Normally the MNC preparation contained 10-30% monocytes. In addition, flow-cytometry was used to determine the proportion of CD14-positive cells (monocytes). Usually the preparation yielded 85-95% monocytes. Yield and purity varied slightly, depending on the donor.

## 4.7.6 Mouse bone marrow macrophage preparation

Mouse bone marrow progenitor cells were extracted from the femurs and tibia of 8-12 weeks old mice and differentiated in bone marrow derived macrophages (BMM) by adding rCSF to the culturing media. All Mice were ordered from Charles River (Sulzfeld) and F1 hybrid offspring were generated by natural matings. Cells from several individual were ether pooled or treated separately depending on the experiment. Cells were flushed out of the bones with TBS using 27 gauge needle attached to a 5 ml syringe (BD) from both directions, centrifuged at 300 g, for 8 min at 4°C counted and seeded at a concentration of $5*10^5$ cells/ml culture medium in bacteriological square plates (Sterilin; Barloworld scientific, UK) After 5 days the medium was changed and cells were harvested on day 6 and re-plated at a density of $10*10^6$ cells/10 ml media in 10 cm tissue culture dish (Falcon). On day 7 the cells were harvested and DNA and RNA were prepared as described.

**Required cell culture medium:**

| | | |
|---|---|---|
| BMM growth medium | 500 ml | RPMI with supplements (see 4.7.1) |
| | 200 ng/ml | rCSF (Cetus) |
| | 10 % | 50 ml FCS (PAN Biotech) |

# 5 Results

Due to space limitations, this section only contains parts or summaries of the sequence, expression, MCIp-on-chip and MassARRAY data. Complete figures, tables, lists and UCSC tracks are or will be available online within the supplementary information of the corresponding publications.

## 5.1 Global, comparative analysis of tissue specific promoter CpG methylation

The data from this section has been published in the journal *Genomics*[A]. Microarray data has been deposited with the GEO Data Library under Series Entry GSE5548.

### 5.1.1 Adaptation of the MCIp-on-chip approach

Methyl-CpG-immunoprecipitation (MCIp) is based on the recombinant antibody-like MBD2–Fc fusion protein that was originally designed by our lab to globally detect disease-related hypermethylation in CpG islands[110]. In this work the existing MCIp procedure was modified and established for a novel application, the genome-wide detection of tissue-specific methylation differences. In general, cell type-specific methylation differences are thought to mainly occur in CpG-depleted regulatory regions, whereas most CpG islands, in particular promoters, remain unmethylated in all tissues[137]. We therefore hypothesized, that those tissue-specific methylation differences may preferentially be detected in the nonbinding fraction of a MCIp experiment which contains mainly unmethylated CpG dinucleotides. The newly developed approach takes advantage of the fact that MCIp not only enriches hypermethylated DNA but separates DNA depending on its methylation grade and therefore also allows the collection of hypomethylated DNA. All fractions, hypomethylated as well as hypermethylated can be collected within one experiment. In this pilot study the original method was for the first time adapted to comparatively analyze the hypomethylated genomic DNA of different human tissues using microarray technology (MCIp-on-chip).

[A] Schilling,E. & Rehli,M. Global, comparative analysis of tissue-specific promoter CpG methylation. *Genomics*. **90**, 314-323 (2007).

#### 5.1.1.1 Separation of differentially methylated DNA by MCIp

The previously described procedure[110] was adapted to separate and isolate unmethylated DNA fragments as follows. In contrast to the original approach genomic DNA was fragmented to a mean length of 400-500 base pairs by ultrasonication rather than endonuclease digestion to avoid sequence bias and to generate fragments of a defined mean length. To allow the direct labeling of DNA fragments for subsequent microarray analysis without the need of prior amplification, which again can produce a bias, the whole procedure was scaled up (see methods section for details). All genomic DNA samples were derived from different human individuals. Genomic DNA from monocytes was obtained from male volunteers after their isolation from mononuclear cells via counter current centrifugation (Elutriation), whereas all other genomic DNA samples (brain, liver and two testis samples) were purchased from BioChain Institute (Hayward, CA, USA). Testis samples were chosen as a reference, because of their general low global methylation content. I used 4 µg of sonicated genomic DNA fragments together with large spin columns to get sufficient amounts of DNA fragments for the direct labeling of DNA for microarray hybridization experiments. The salt concentration of the binding buffer was adjusted to allow fragments methylated at only a few sites (one or two mCpG's) flow through into the 350 mM fraction. The affinity of the MBD2-Fc fusion protein to DNA is decreased with higher salt concentrations in the buffer, allowing separation of methylated DNA using increasing NaCl concentration in the wash buffer. Additional fractions were obtained at 400, 500, 570 and 1000 mM NaCl.

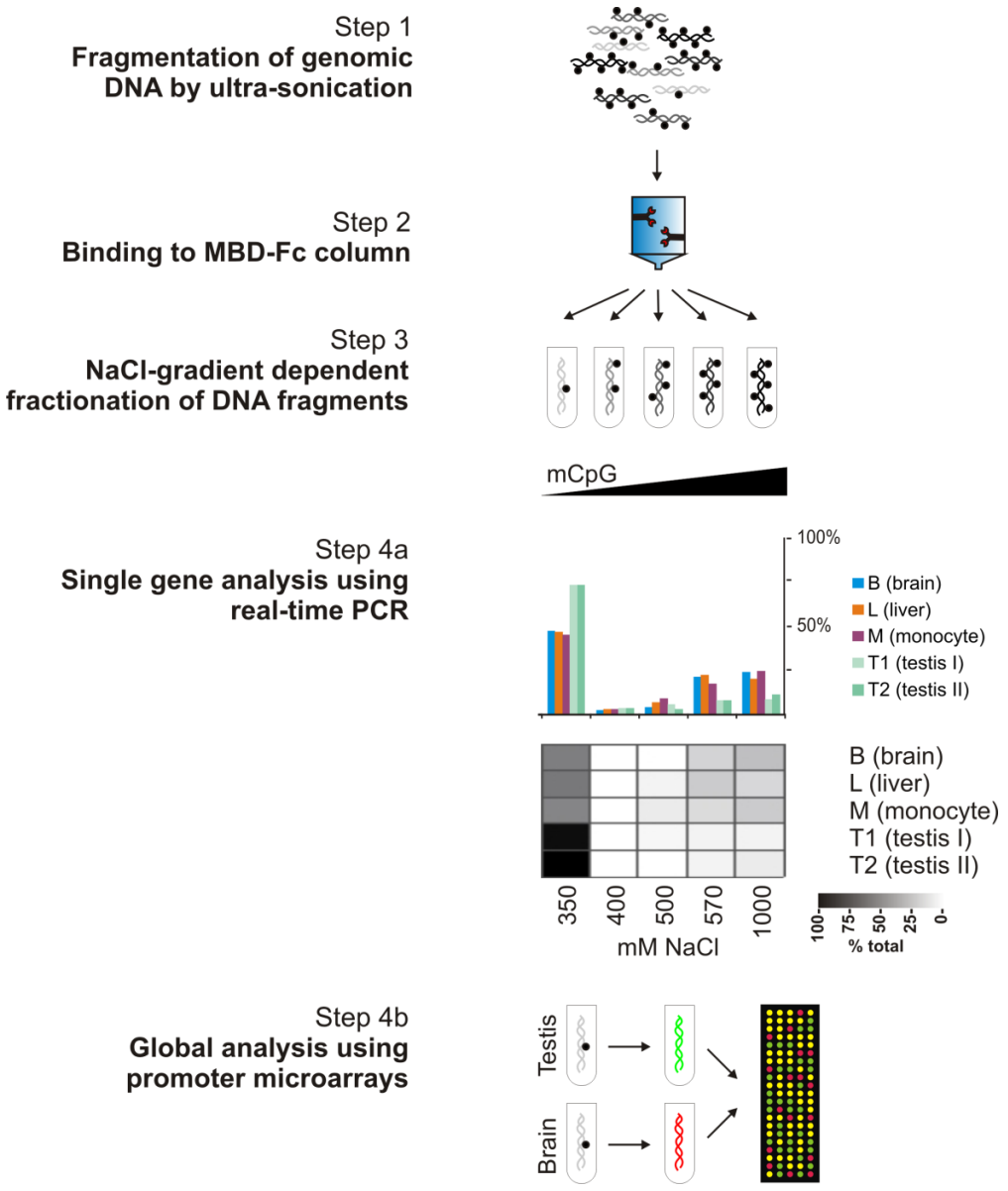


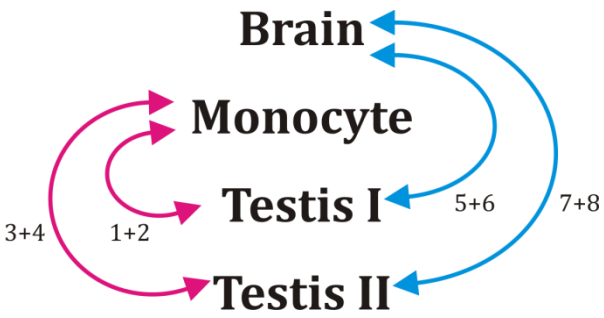

**Figure 5.1 MCIp-on-Chip approach to detect promoter regions with tissue-specific demethylation.** Genomic DNA of tissues or purified cell populations is fragmented to a mean size of 400-500 bp using ultrasonication (Step1). The material is bound to MBD-Fc sepharose beads at 350 mM NaCl for 3 h (Step 2) and fractionated using wash/elution buffers containing increasing concentrations of NaCl (400-1000 mM, Step 3). For single gene analysis, the amount of a single gene fragment can be quantified by real-time PCR in each fraction (Step 4a). The data of an exemplary analysis of the *SNRPN* promoter is graphically shown as a bar chart and heat map. The percentage of amplified material in a given fraction is represented by a gray scale. Since the *SNRPN* promoter is subject to maternal imprinting, approximately half of the amplified material in somatic tissues is present in the 350 mM fraction (unmethylated) while the second half elutes mainly above 570 mM NaCl representing the methylated alleles. As expected, germ-line cells in testis lack the maternal imprint and approximately 75% of the SNRPN fragments are found in the unmethylated fraction. Alternative to single gene analysis, individual fractions (here the unmethylated 350 mM NaCl fraction) of two tissues may be used for dual color microarray analysis (Step 4b).

The strategy for purification and subsequent detection of the fractionated material is shown in **Figure 5.1**. Prior to labeling the separation of differentially methylated DNA during the MCIp procedure was controlled by qPCR using primers covering a maternally imprinted region of *SNRPN*, which gives rise to two elution peaks in somatic samples. In addition, the utilized MBD2-Fc lot was continuously tested for functionality prior to each experiment using a small scale MCIp followed by qPCR for *SNRPN*.

The methylated and unmethylated alleles of *SNRPN* were separated in all three somatic tissues (brain, liver and monocytes) and were eluted in the low salt and high salt fractions, respectively. In germ-line cells (testis samples) both alleles were unmethylated and were collected mainly in the low salt fraction, as demonstrated by the heat map in **Figure 5.1** (Step 4a). Hence, it could be shown that the modified MCIp approach can be used to efficiently separate the methylated and the unmethylated alleles in somatic tissues, whereas the majority of this locus, as expected, appears unmethylated in germ line cells (testis).

#### 5.1.1.2 Detection of tissue specific hypomethylation by microarray readout

To demonstrate the feasibility of this method to detect tissue-specific differences on a global scale, the microarray technology was chosen. I used the hypomethylated DNA from the MCIp separation (350 mM fractions) of the three different tissues (testis, brain and monocytes) as described above. Fractionated DNA fragments were directly labeled with different fluorescent dyes. All testis samples were labeled with Alexa-Fluor-555 (shown in green in **Figure 5.1** Step 4b), whereas all brain and monocyte samples were labeled with Alexa-Fluor-567 (shown in red in **Figure 5.1** Step 4b). Testis DNA fragments served as a reference and were co-hybridized with either monocyte or brain DNA fragments onto commercial available 44k human promoter arrays (Agilent). Each comparison uses a set of two microarray slides (Slide 1/2 and slide 2/2), each containing 44,000 oligonucleotide probes of 45 to 60 base pairs in length. In total these 88,000 probes per “array” corresponded to nonrepetitive sequences and represented approximately 17,000 annotated genes and 20,000 promoters. Slide 1/2 contained probes for promoters on human chromosomes 1-10, whereas all other chromosomes including X- and Y-chromosomes were represented on slide 2/2. In order to limit interindividual differences and to increase the significance of these analysis, two hybridizations per comparison were performed using two independent testis samples (biological replicates). Hence, every set included two hybridizations with independent testis samples for each comparison. Efficiency of labeling was controlled by UV-spectroscopy and samples were hybridized on microarrays as follows:




| | | | | |
|---|---|---|---|---|
| 1) | Monocyte | versus | Testis I | Slide 1/2; repl. 1 |
| 2) | Monocyte | versus | Testis I | Slide 2/2; repl. 1 |
| 3) | Monocyte | versus | Testis II | Slide 1/2; repl. 2 |
| 4) | Monocyte | versus | Testis II | Slide 2/2; repl. 2 |
| 5) | Brain | versus | Testis I | Slide 1/2; repl. 1 |
| 6) | Brain | versus | Testis I | Slide 2/2; repl. 1 |
| 7) | Brain | versus | Testis II | Slide 1/2; repl. 2 |
| 8) | Brain | versus | Testis II | Slide 2/2; repl. 2 |

Representative scatter plots of MCIp microarray analyses are shown in **Figure 5.2**, where the Log2 signal ratios (testis/somatic tissue) were plotted against the Log2 processed signal of testis (reference). Each probe on the array is represented by a single dot. On a global scale, both testis MCIp preparations appeared to be less methylated than brain or monocytic DNA. In addition, brain was usually less different from testis than monocytes. Assuming that DNA methylation in proximal promoter regions would be most influential in terms of gene regulation the further analysis was restricted to oligonucleotide probes located at proximal promoter regions (-600 to +200 relative to the transcription start site). A probe was defined as differentially enriched when the fluorescent signal was significantly present ($p<0.01$ for processed signal) and more than 2.4-fold different between the samples (**Figure 5.2**; red spots).

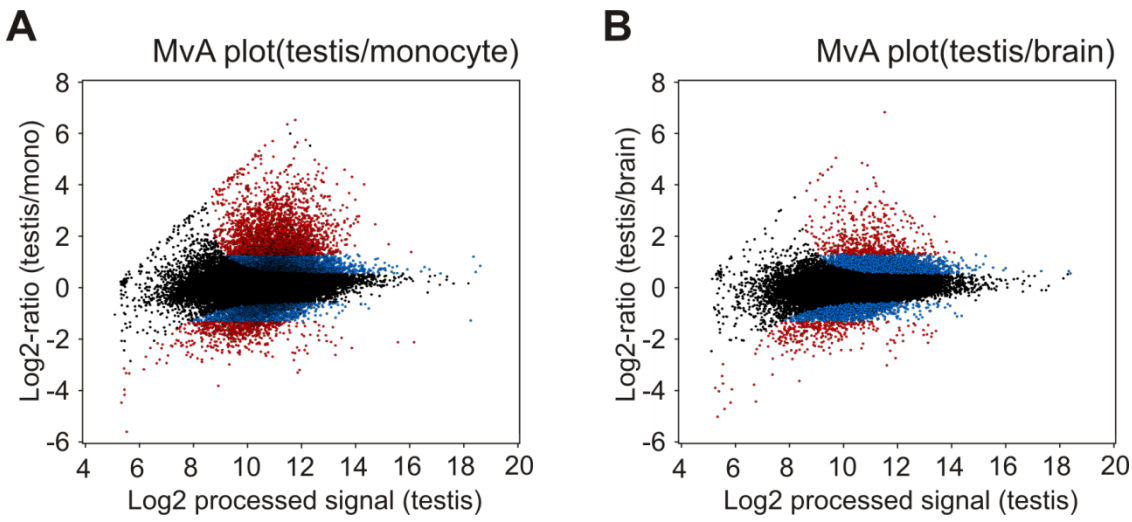



**Figure 5.2 Global profiling of tissue DNA methylation using human promoter microarrays.** MvA plots (log2 signal ratio versus log2 processed signal testis) are shown for representative hybridization experiments of human testis versus monocytes (A) or brain (B). Black dots represent probes that were not significantly changed (p>0.01). Blue and red spots were significantly enriched (ratio < 2.4 fold and > 2.4 fold, respectively; p<0.01).

## 5.1.2 Identification of differentially methylated promoters

To identify differentially methylated promoters, the analysis was resticted to probes with a CpG index above 4. Assuming, that a single oligonucleotide probe should detect methylation events within a radius of at least 350 bp (see **Figure 5.3A**), a given DNA fragment should contain at least four or five CpG residues within this radius, in order to allow the reproducible separation of unmethylated and methylated fragment (see methods section 0). In total, less than 10% of all probes in proximal promoter regions and 4% of all proximal promoters were excluded.

A proximal gene promoter was considered comparatively less methylated (hypomethylated) if at least one probe within this promoter was concordantly different in both hybridization experiments using the criteria explained above. A promoter was grouped as hypomethylated in testis if it was less methylated compared to both somatic tissues (brain and monocytes). Promoters were grouped as hypomethylated in brain if they appeared less methylated in the brain/testis comparisons but not significantly different in the monocytes/testis hybridizations. The latter was also used to identify promoters defined as hypomethylated in monocytes. In total, 283 gene promoters that were specifically hypomethylated in testis were identified. In addition, 123 gene promoters in brain as well as 190 gene promoters in monocytes were less methylated compared to testis (complete list is shown in **Table 12.1** in the appendix section).

### 5.1.3 Correlation of promoters and nearby genetic elements (CGI and *Alu* repeats)

As mentioned above, a single oligonucleotide probe should detect methylation events within a radius of at least 350 bp. Therefore it is possible that strongly methylated sequences (such as methylated CpG islands or *Alu* repeats) in the vicinity of a probe may affect its signal within a 250 to 350 bp distance. To address this question, the presence of classical CpG islands[8] or *Alu* repeats[138,139] near individual probes was determined at radii of 250, 350, and 850 bp around each probe. The relative locations of microarray probe, CpG islands, and *Alu* repeats were determined using the UCSC Genome Browser. As demonstrated in **Figure 5.3B** almost 50% of hypomethylated promoters were associated with nearby CpG islands in testis at all distances ($p<0.0001$; $\chi^2$-test). In contrast, less than 5% of hypomethylated promoters found in the somatic tissues (brain and monocytes) showed the same association. When comparing the percentage of hypomethylated promoters that are associated with *Alu* repeats, a significantly higher incidence of *Alu* repeats (50%) was noticed in the vicinity of promoter sequences in testis (**Figure 5.3C**). Whereas in the brain sample only 30% of demethylated promoters were found close to *Alu*-repeats and even less in the monocytes sample (22%). The difference was obvious at all distances, but was most significant at 250 bp around individual hypomethylated probes ($p<0.0001$; $\chi^2$-test). Genes with tissue-specific demethylated promoters were generally distributed over all chromosomes (data not shown). In addition, the higher incidence of CpG islands or *Alu*-repeats in the vicinity of hypomethylated promoters in testis also resulted in the observation that these promoter generally showed higher median CpG indices (CpGindex=27) as corresponding promoters in brain (CpGindex=13) or monocytes (CpGindex=8).

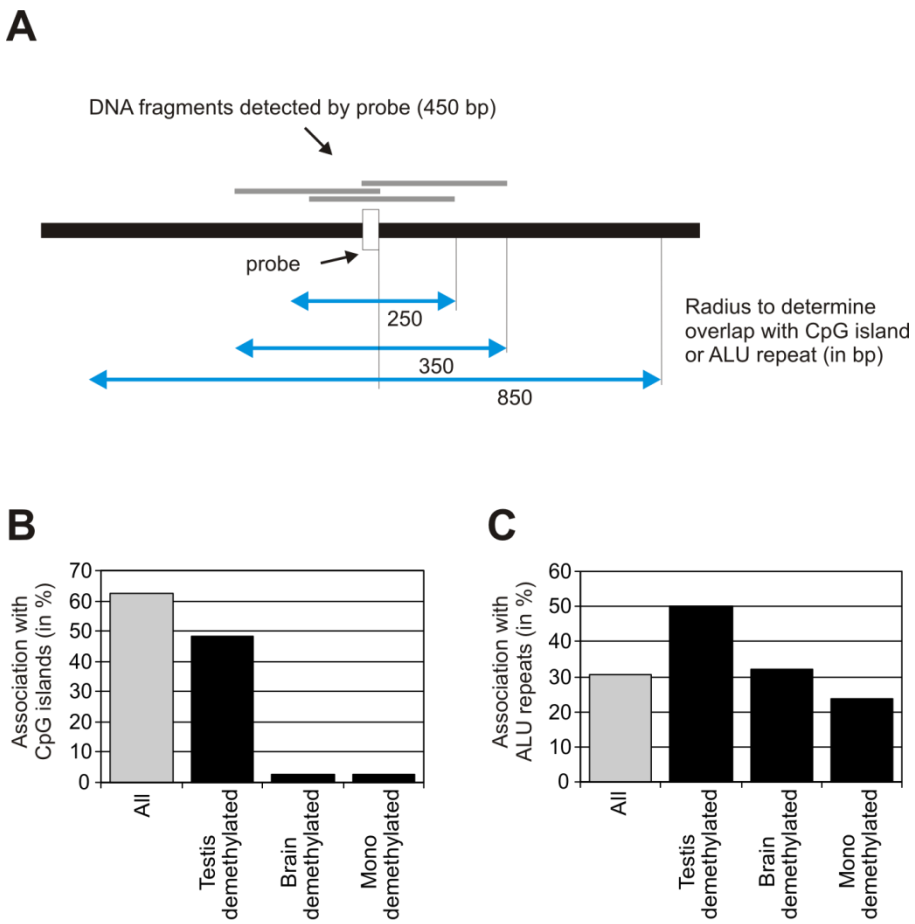



**Figure 5.3 Association of hypomethylated promoters with CpG islands and *Alu* repeats.** (A) Graphical presentation of a probe (50 bp), detected fragments (mean of 450 bp) and distances used to determine overlaps with CpG islands and *Alu* repeats (shown in blue). The presence of CpG islands or *Alu* repeats in the vicinity of proximal promoter probes was calculated for all individual promoter regions with a CpG index above 4 (All, 19145 promoters in total) or within each hypomethylated group (testis, brain and monocytes with 283, 123, 190 promoters, respectively). The percentage of association of hypomethylated probes within a given radius (250 bp) with either CpG island (B) or *Alu* repeats (C) is shown relative to all promoters in a group.

Interestingly, a previously unobserved accumulation of testis-specific hypomethylation was detected on the Y-chromosome. Several of the annotated Y-chromosomal genes appeared hypomethylated compared to both somatic tissues. These genes included *VCY*, *DAZ*, *TSPY* and *RBMY*, which reside in areas on the Y chromosome that are known as ampliconic areas[140] (see **Figure 5.4**). These genes were chosen amongst others for the initial validation set.

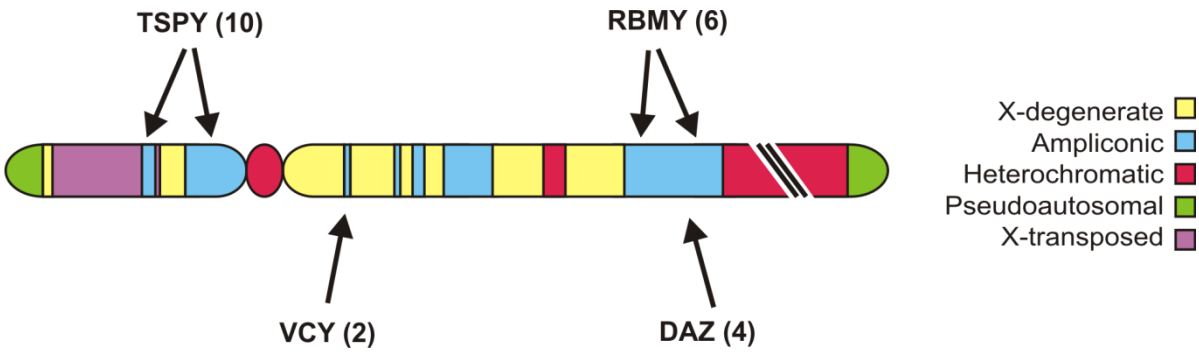



**Figure 5.4 Different areas on the human Y chromosome.** The various areas on the human Y chromosome as defined by Skaletsky[140] are shown color coded. The ampliconic areas include large regions (about 30% of the male-specific euchromatin on the Y-chromosome) where sequence pairs show greater than 99.9% identity, which is maintained by frequent gene conversion (non-reciprocal transfer). The most prominent features here are eight massive palindromes, at least six of which contain testis genes. Arrows mark the position of 4 genes located within ampliconic areas, which are chosen among others for the initial single-gene MCIp validation set. Numbers in brackets indicate the number of amplicons detected by the corresponding primer pairs.

## 5.1.4 Validation of MCIp microarray results by real-time PCR

### 5.1.4.1 Y-chromosomal genes

To validate the MCIp microarray data I initially focused on the Y chromosome and analyzed all CpG-rich promoters of testis-specific genes in ampliconic areas[140] and also number of additionally randomly chosen genes of the Y chromosome. Some genes within this ampliconic area are present in multiple copies with highly similar promoter sequences for each gene cluster. Hence, primers for *TSPY*, *VCY*, *PRY*, *RBMY1A1* and *DAZ* detect 10, 2, 4, 6 and 4 copies, respectively. Furthermore, with the primer set for the *TGIF2L* gene, two copies were amplified, one of which is located on the Y chromosome and the other on the X chromosome. I designed primers for 12 different promoters within this region. In order to accomplish the validation I performed small scale MCIp's for single-gene detection, which used 300 ng of DNA fragments per sample. In this experiment I also included genomic DNA derived from liver to increase the number of somatic tissues used. Each fraction of a MCIp was measured by subsequent real-time PCR independently for two times (technical replicates) and each MCIp was carried out twice (biological replicates).

The MCIp data correlated well with the microarray data (see **Figure 5.5**). Most testis-specific genes within the ampliconic regions of the Y-chromosome showed hypermethylation in somatic cells. In particular following genes, *VCY*, *DAZ*, *TSPY, RBMY*

and *TGIF2LY*, which are highly methylated according to the microarray analysis, showed strong hypermethylation in somatic tissues in comparison to the two testis samples. Furthermore, in randomly chosen control genes such as *NLGN4Y* and *EIF1AY*, methylation could be neither detected in the microarray experiment nor in the single-gene MCIp.

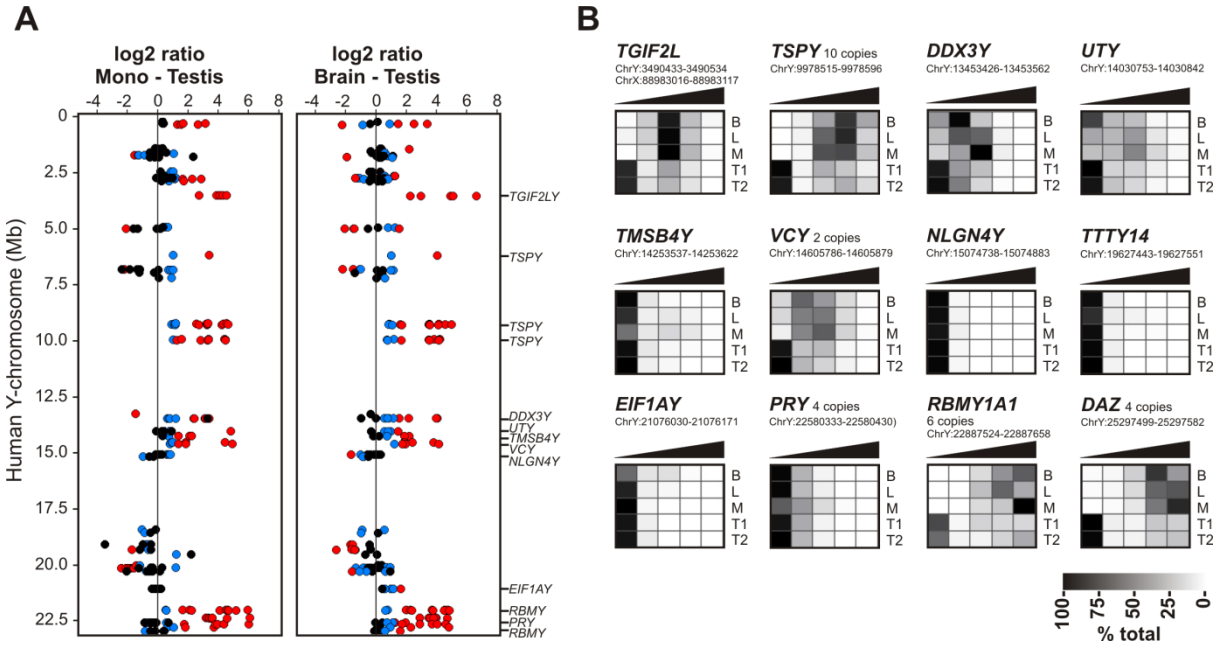



**Figure 5.5 Tissue-specific CpG methylation profiles of the human Y chromosome.** (A) Schematic representation of microarray results from one of two hybridizations for the human Y chromosome. Log2 signal ratios for each probe are represented relative to their absolute chromosomal position (based on NCBI build 35 of the human genome sequence). Black spots represent probes that were not significantly changed ($p>0.01$), blue spots were significantly enriched but less than 2.4-fold different, and red spots were significantly enriched ($p<0.01$) and over 2.4-fold different. The positions of genes shown in (B) are indicated. (B) Graphical representation of single-gene MCIp real-time PCR data for several tested Y-chromosome genes. Gene symbols, copy numbers, and representative chromosomal localizations of the amplified product (based on NCBI build 35 of the human genome sequence) are given above the heat maps, which represent the percentages of amplified material in a given NaCl fraction by gray colouring for DNA samples of brain (B), liver (L), monocytes (M), testis sample 1 (T1), and testis sample 2 (T2) as described for **Figure 5.1**.

### 5.1.4.2 Autosomal Genes

The validation of MCIp results was further extended to autosomes. In total, I designed real-time PCR primers for 25 coding and 4 non-coding gene promoters and performed single-gene MCIp's as described above. The real-time PCR results of these 25 coding genes are presented as heat maps in **Figure 5.6A**. A number of previous studies already noted the testis-specific hypomethylation of selected genes such as *MAGE-A1*[140,141]. Also in non-coding micro-RNAs a variation of tissue-specific promoter methylation could be detected. Therefore, 4 micro-RNA promoters (*MIRN127*, *142*, *338* and *363*) were included in the validation analysis. As illustrated in **Figure 5.6B,** tissue-specific hypomethylation as seen on the microarray was validated using single-gene MCIp analysis. The promoter region of *MIRN142* was hypomethylated in monocytes, whereas promoters of *MIRN338* and *363* were hypomethylated in the brain sample compared to the other tissues. According to the literature, testis-specific hypomethylation was demonstrated for the *MIRN127* promoter. A recent publication also described the testis-specific demethylation of *MIRN127* and its activation in somatic tissues by chromatin-modifying drugs[142], confirming the MCIp result.

In total, 24 (out of 29) chosen gene promoters showed essentially the same tissue-specific methylation patterns as seen on the microarray (see **Table 5.1** on page 80). Only *DAZL*, *MGP* and *DCX* gene promoters exhibited a different pattern. *DAZL* appears to be false negative on the microarray, whereas *MGP* and *DCX* seem to be false positive, because microarray data and validation results differed in these cases. Two examples (*TYROBP* and *MIRN363)*, although correctly validated, were significantly different in only one of the two microarray hybridizations, but showed the same trend in each case.

Besides the large number of testis-specific gene promoters that demonstrated hypermethylation in somatic tissues, a large group of testis-specific genes was also noticed, that did not exhibit a differential methylation status in any of the examined tissues. For instance *AKAP3*, *DMRT1*, *TEKT3* and *TTTY14* gene promoters were generally unmethylated in all tissues examined. Additionally it was observed, that copies of Y-linked genes on other chromosomes like *DAZL* (Chr. 3) or *VCX* (Chr. X), that share almost identical promoter regions with their homologue on chromosome Y, show somatic hypermethylation. Whereas homologous genes that are regulated by different promoter sequences like *RBMX* (Chr. X) or *TSPYL5* (Chr. 8) show a different methylation profile as compared to their Y-chromosome homologues (data not shown). Only, the *SYBL1* genes on both X and Y chromosomes appear to represent an exception in having almost identical promoters but being solely methylated on the Y chromosome[143].

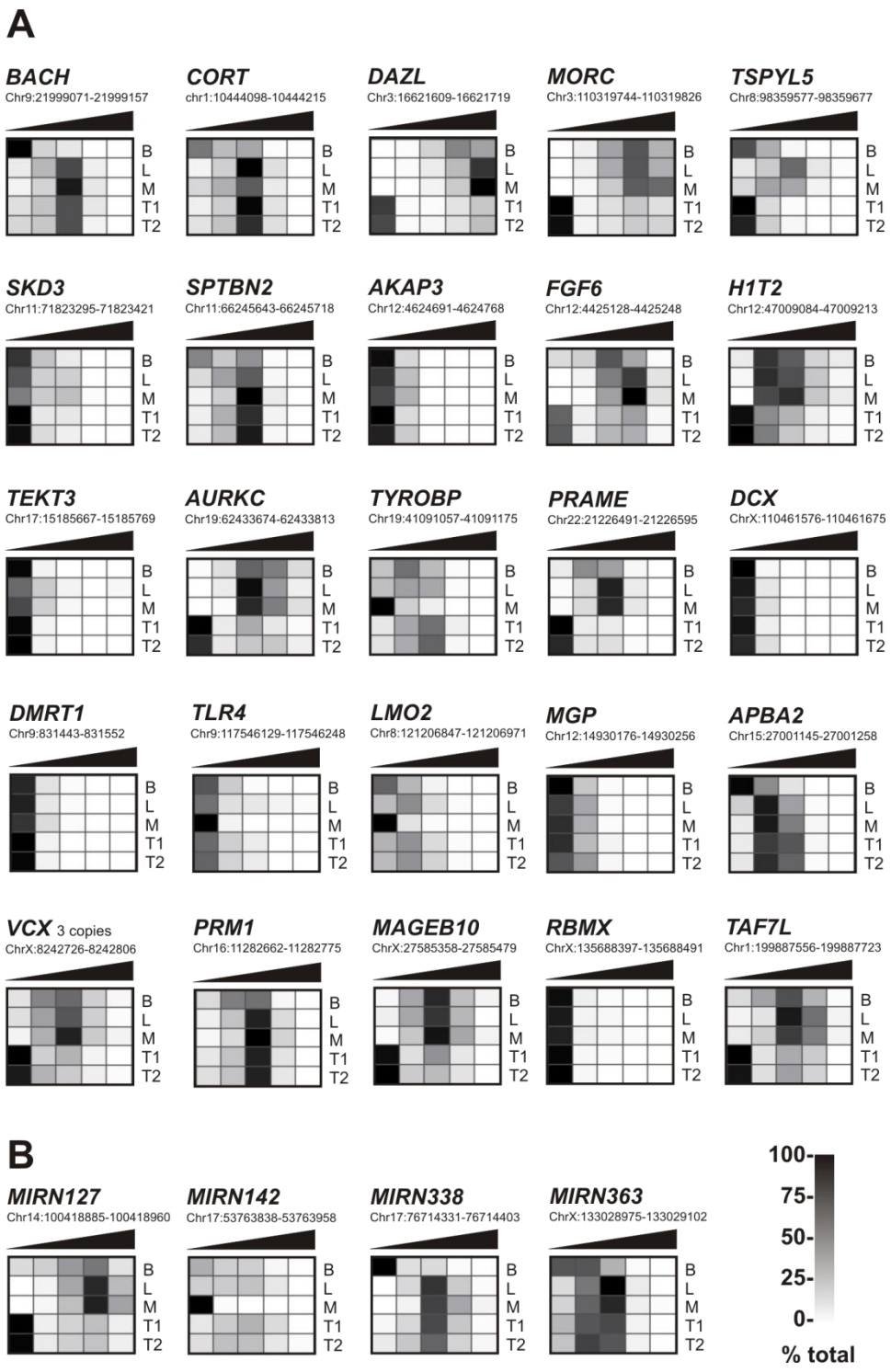



**Figure 5.6 Heat maps of MCIp real-time PCR results.** Graphical presentation of single gene MCIp real-time PCR data for several coding (A) and non-coding (B) genes listed in **Table 5.1**. Gene symbols, copy numbers and representative chromosomal localizations of the amplified product (based on the NCBI Build 35 of the human genome sequence) are given above the heat maps which are presented as described in **Figure 5.1**.

In total, 35 out of 40 (87.5%) validated gene promoters demonstrated the same methylation profile in global and single-gene analysis. Taken together, the MCIp microarray approach demonstrated a high reproducibility. For further confirmation I used another MCIp independent approach to validate the microarray results.

### 5.1.4.3 Microarray-validation using bisulfite-treatment of genomic DNA

As a second, MCIp-independent validation approach I directly analyzed cytosine methylation by bisulfite-treatment of genomic DNA followed by sequencing. Promoters of five genes were selected and primer pairs specific for bisulfite-treated DNA were designed. Three of them exhibited hypomethylation specifically in testis samples, whereas one was hypomethylated in the brain and one in the monocytes sample. For each promoter two pair of primers were used comprising of an 'outer' pair and an 'inner' pair. The 'outer' primer pair was used in the initial amplification step (1$^{st}$ PCR), whose product was used as a template for the 'inner' pair of primers during the 2$^{nd}$ PCR, in order to increase specificity of PCR amplification (nested PCR). Each PCR steps was monitored by agarose gel electrophoresis. The purified PCR products of the 2$^{nd}$ PCR were cloned into the pCR2.1-TOPO vector and transfected into chemically competent *E. coli* TOP10. Insert-containing plasmids were isolated from single colonies and sequenced. The results are shown in **Figure 5.7**. The hypomethylation patterns seen for *BACH* and *TLR4* were clearly brain and monocyte-specific, respectively. *BACH* was hypomethylated to a stronger extend in the brain compared to the other tissues. *TLR4* was completely unmethylated in monocytes only, whereas all other samples contained both fully and partially methylated sequences. The other two examples, *MAGEB10* and *TGIF2LY,* exhibited almost complete methylation of all analyzed sequences only in somatic tissues (brain, monocytes and liver). In contrast, completely unmethylated fragments occurred only in both testis samples. Taken together these four gene promoters showed the same tissue-specific methylation pattern in the single-gene MCIp validation experiment (see **Figure 5.5** and **Figure 5.6**), as well as in the microarray analysis.

The fifth example gene shown in **Figure 5.8** *ZSCAN5* appeared to be strongly unmethylated in both testis samples. Interestingly, the promoter of this gene contained an *Alu* repeat, whose methylation status was also examined. Although it was not analysed by single gene-MCIp, the bisulfite sequencing result correlated well with the microarray result: *ZSCAN5* exhibited hypomethylation only in the two testis samples. Interestingly the *Alu* repeat was mainly unmethylated as well.

In total, all five bisulfite sequencing examples validated the microarray results, indicating that the initial global experiment correctly identified tissue-specific hypomethylation with a low false positive rate.


**Figure 5.7 Bisulfite sequencing results of selected gene promoters.** (A-D) Genomic DNA from the indicated tissues was analyzed for CpG methylation at *BACH* (A), *TLR4* (B), *MAGEB10* (C) and *TGIF2LY* (D) loci by bisulfite sequencing. Bisulfite-treated DNA was amplified in individual PCR reactions. PCR-products (representing the sense strand) were sub-cloned and several individual clones were sequenced. The analyzed fragment is schematically represented above each figure where a CpG dinucleotide is shown as a vertical line. The positions of exons are displayed as gray boxes and transcription start sites by arrows. For each tissue, the position of CpG dinucleotides is shown as a square (empty: unmethylated; filled: methylated).

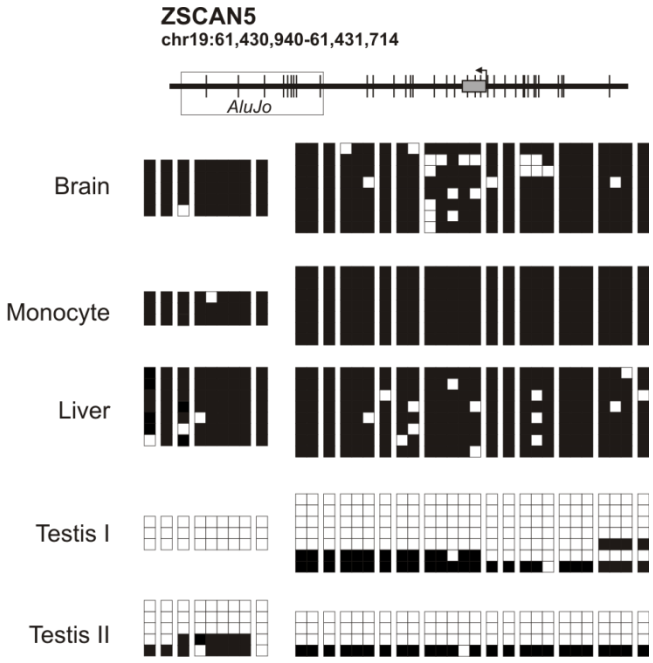



**Figure 5.8 Bisulfite sequencing of *ZSCAN5*.** Additional example of bisulfite sequenced promoter region presented as described in **Figure 5.7**.

### 5.1.4.4 Summary of validation

In total, I validated 44 genes (summarized in **Table 5.1**) using single-gene MCIp and sequencing of bisulfite-treated genomic DNA as independent DNA methylation mapping techniques. Within **Table 5.1** the MCIp-Array column indicates the tissues that were found to be less methylated (hypomethylated) compared to the two other tissues used in the microarray experiment. The promoters of *TYROBP* and *MIRN363* were significantly different in only one of the two microarray hybridizations, but showed the same trend in each case. The data obtained from the liver DNA sample in single-gene MCIp were not included, since this sample was not analyzed by microarray. Single-gene MCIp validation data obtained from the brain, testis and monocyte samples is shown (MCIp-PCR), as well as data derived from sequencing of bisulfite-treated genomic DNA (Bisulfite). Primers for *AURKC* and *DDX3Y* were located more than 600 bp upstream ("up") of the putative transcription start site in a region that was differentially methylated as indicated, whereas the proximal region was not different in both cases. In addition, **Table 5.1** shows information obtained from public available databases. The presence or absence of a CpG island in the 5'-proximal promoter region (CGI promoter) was determined using the UCSC Genome Browser (http://genome.ucsc.edu/). Expression data (mRNA expression) were downloaded from SymATLAS (http://symatlas.gnf.org/SymAtlas/) and RefEXA databases (http://www.lsbm.org/site_e/database/) and indicate the primary source of expression.

**Table 5.1** Profiles of promoter hypomethylation and mRNA expression of selected genes

| *Gene symbol* | *CGI promoter* | *Chr.* | *MCIp-Array* | *MCIp-PCR* | *Bisulfite* | *mRNA Expression* |
|---|---|---|---|---|---|---|
| *BACH (ACOT7)* | no | 1 | brain | brain | brain | brain |
| *CORT* | no | 1 | brain | brain | - | *ubiq.** |
| *DAZL* | yes | 3 | *not diff.** | testis | - | testis |
| *MORC* | yes | 3 | testis | testis | - | testis |
| *TSPYL5* | yes | 8 | testis, brain | testis, brain | - | testis, brain, blood |
| *DMRT1* | yes | 9 | *not diff.* | *not diff.* | - | testis |
| *TLR4* | no | 9 | mono | mono | mono | blood |
| *LMO2* | no | 11 | mono | mono | - | blood, endothelium |
| *SKD3* | yes | 11 | testis, brain | testis, brain | - | testis |
| *SPTBN2* | no | 11 | brain | brain | - | testis, brain, other |
| *AKAP3* | no | 12 | *not diff.* | *not diff.* | - | testis |
| *FGF6* | yes | 12 | testis | testis | - | testis, muscle |
| *H1T2* | no | 12 | testis | testis | - | testis |
| *MGP* | no | 12 | mono | *not diff.* | - | *ubiq. (not blood)* |
| *MIRN127* | yes | 14 | testis | testis | - | testis |
| *APBA2* | no | 15 | brain | brain | - | brain, blood |
| *SNRPN* | yes | 15 | testis | testis | - | *ubiq. (impr.)** |
| *PRM1* | no | 16 | *not diff.* | *not diff.* | - | testis |
| *MIRN142* | no | 17 | mono | mono | - | *n.a.** |
| *MIRN338* | no | 17 | brain | brain | - | *n.a.* |
| *TEKT3* | yes | 17 | *not diff.* | *not diff.* | - | testis |
| *AURKC (up)*[6] | yes | 19 | testis | testis | - | testis |
| *TYROBP* | no | 19 | *not diff. (mono)* | mono | - | blood |
| *ZSCAN5* | yes | 19 | testis | - | testis | testis |
| *PRAME* | yes | 22 | testis | testis | - | testis |
| *DCX* | no | X | mono | *not diff.* | - | fetal brain |
| *MAGEB10* | no | X | testis | testis | testis | *n.a.* |
| *MIRN221* | no | X | *not diff.* | - | *not diff.* | *n.a.* |
| *MIRN363* | no | X | *not diff. (brain)* | brain | - | *n.a.* |
| *RBMX* | yes | X | *not diff.* | *not diff.* | - | *ubiq.* |
| *TAF7L* | no | X | testis | testis | - | testis |
| *VCX* | no | X | *n.a.* | testis | - | testis |
| *DAZ* | yes | Y | *n.a.* | testis | - | testis |
| *DDX3Y (up)* | no | Y | testis | testis | - | testis, blood |
| *EIF1AY* | no | Y | *not diff.* | *not diff.* | - | blood |
| *NLGN4Y* | yes | Y | *not diff.* | *not diff.* | - | *ubiq.* |
| *PRY* | no | Y | *not diff.* | *not diff.* | - | testis |
| *RBMY1A1* | yes | Y | testis | testis | - | testis |
| *SYBL1* | yes | Y | *n.a.* | testis | - | testis |
| *TGIF2LY* | no | Y | testis | testis | testis | testis |
| *TMSB4Y* | yes | Y | *not diff.* | *not diff.* | - | *ubiq.* |
| *TSPY* | yes | Y | testis | testis | - | testis |
| *TTTY14* | yes | Y | *n.a.* | *not diff.* | - | testis |
| *UTY* | no | Y | testis, brain | testis, brain | - | *ubiq.* |
| *VCY* | no | Y | testis | testis | - | testis |

*ubig.: ubiquitous; not diff.: not different; impr.: imprinted; n.a.: not analyzed

The present microarray analysis confirmed the already published testis-specific methylation status of several genes, which were not independently validated by single-gene MCIp. Those genes comprise *DDX4*[144], *ACTL7B*[145], as well as *TAF7L* and *LUZP4*[146]. *MAGEA1* and *TKTL1* that had been analyzed previously by bisulfite sequencing[146,147] were detected as hypomethylated in only one of the two testis samples tested.

### 5.1.5 Relationship between DNA methylation and gene expression

Next the relationship between CpG DNA methylation of gene promoters and the expression patterns of the corresponding genes was correlated. For this RNA expression data was extracted from publicly available reference databases for gene expression analysis for all corresponding differentially methylated fragments. In particular BioGPS (formerly SymATLAS) and RefExA databases were used. Genes with unknown transcription status were removed and the analysis was restricted to genes, for which expression data were available. Hence, in total 426 genes were analyzed composed of 201 testis-, 88 brain- and 137 monocyte-specific hypomethylated genes. The distribution of expression patterns for tissue-specific hypomethylated fragments is summarized in **Figure 5.9**. The largest group of testis-hypomethylated genes (75; 37%) was specifically expressed exclusively in testis. Another 17% was expressed in testis but also in other tissues excluding (14; 7%) or including (19; 10%) monocytes or brain. Thus, 55% of hypomethylated genes were expressed in testis, even though not solely. Expression of genes that were hypomethylated in monocytes or brain was also often specific for the respective gene (19 or 21%, respectively). Similar to testis, 57% of hypomethylated genes were expressed in brain, but also in other tissues excluding or including testis or monocytes (18% each).

Interestingly, most genes in the monocyte-hypomethylated group were transcribed specifically in the hematopoietic system (in total 104; 76%), rather than in the specialized cell type itself (monocyte). Only 4% of these genes were also expressed in brain or testis. The coexistence of DNA methylation and transcription in a certain tissue was also detected.

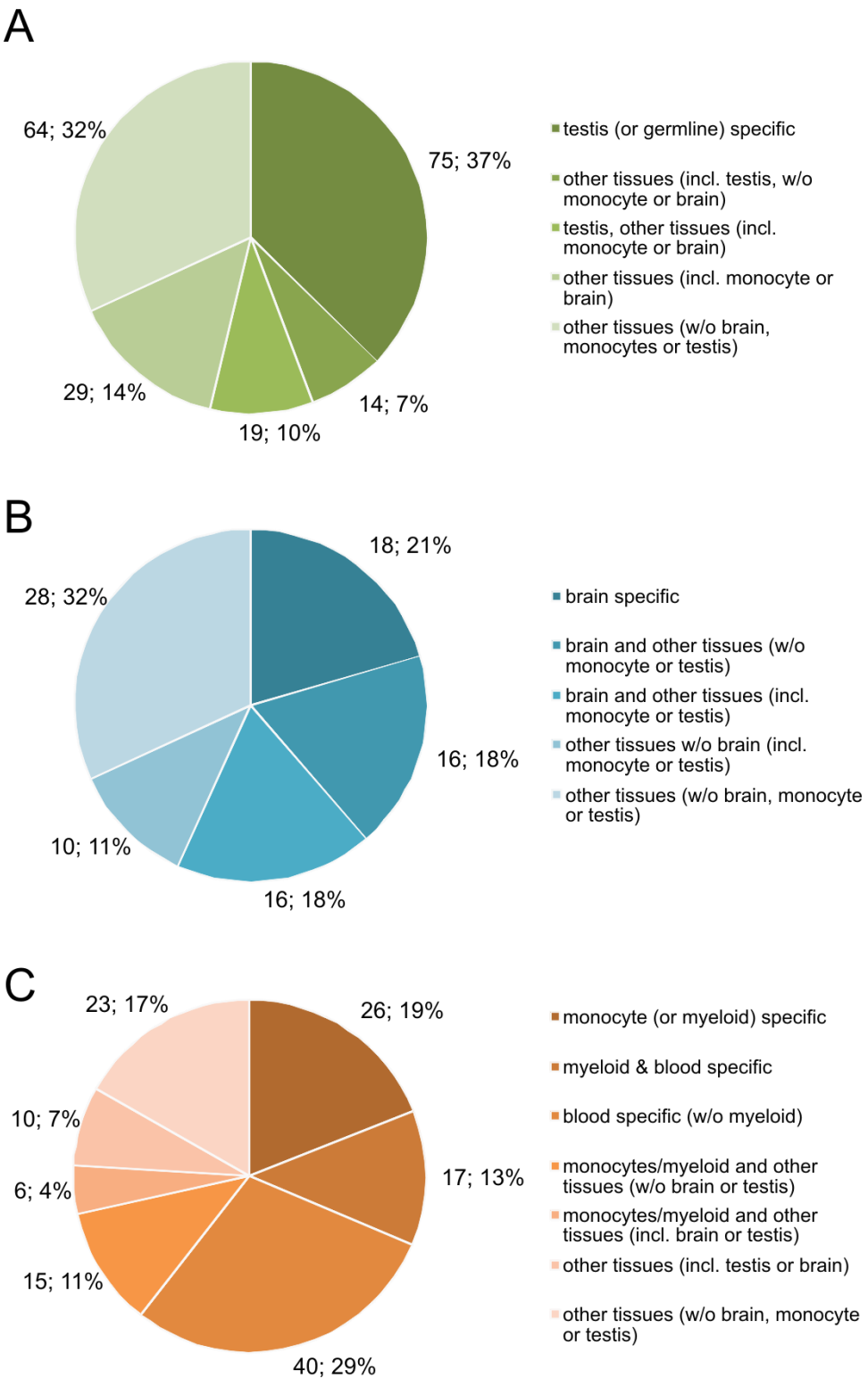



**Figure 5.9 Comparison of methylation and expression profiles.** Distribution of expression patterns for genes with hypomethylated promoters in (A) testis, (B) brain or (C) monocytes. The methylation data derived from microarray experiments while the expression data was obtained from publicly available reference databases for gene expression analysis (RefExa and BioGPS, formerly SymATLAS).

We also determined the number of genes showing a restricted expression pattern in testis, monocytes or brain within each group of hypomethylated genes. 95% of genes which were exclusively demethylated in testis were only expressed testis. Although slightly reduced, a similar degree of correlation appeared to be true for brain and monocytes, ranging from 90% to 79%, respectively. As a summary **Figure 5.10** illustrates that the number of genes expressed in a certain tissue was significantly enriched ($p<0.0001$) in the hypomethylated group of the same tissue. This observation strongly suggests that tissue-specific hypomethylation significantly correlates with tissue-specific transcription.

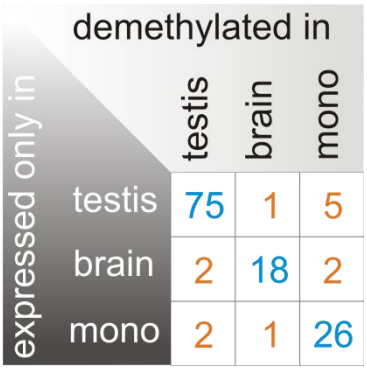



**Figure 5.10 Correlation of tissue-specific expression and demethylation.** Distribution of testis-, brain-, or monocytes-specific genes in each demethylated group.

## 5.2 Allele-specific DNA methylation in mouse strains is mainly determined by *cis*-acting sequences

The data obtained from this project has already been submitted[B], but the review process was still in progress at time of writing this thesis. Microarray gene expression data was deposited with the NCBI GEO database and accession code GSE14644 and will be available upon publication. Results from locus-wide comparative hybridizations will be available under the accession code GSE14463. DNA sequencings were deposited with GenBank (Accession numbers: FJ751937-FJ752004).

### 5.2.1 Inbred mice as model organisms & Preliminary work

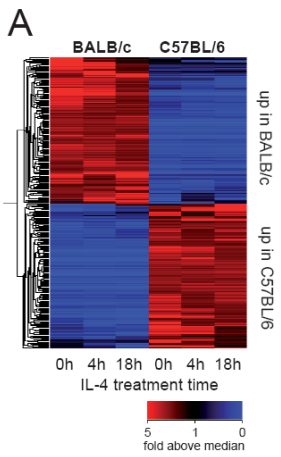
The true extent of epigenetic variation between individuals is still unknown. The development of assays to detect such variation will facilitate more systematic analyses and will contribute to a better understanding of the epigenetic influence on individual phenotypes. In addition, it is currently unknown as well, if such individual variations are inherited based on genetic mechanisms (in *cis* and in *trans*) or based on epigenetic mechanisms. To map interindividual DNA methylation differences we chose to analyze bone marrow-derived macrophages of two distinct inbred mouse strains C57BL/6 and BALB/c. Inbred mice are ideally suited to study the inheritance of an epigenetic mark like DNA methylation, because they have a distinct genetic background, which allows for reproducible mating conditions between two inbred mouse strains. To study the inheritance of an epigenetic mark in the F1 generation, one needs to define regions that are epigenetically different in cells of both strains. However, systematic screens for epigenetic differences in mice have not been performed and only a handful of epigenetically variable regions have been studied so far[89]. In the present study, we performed a combined genetic and epigenetic profiling of strain- specific differences in inbred mice. Therefore the MCIp procedure was further advanced to compare methylation profiles of a defined cell type (bone marrow-derived macrophages) from two inbred mouse strains (C57BL/6 & BALB/c) that represent prototypic models for Th1- or Th2- dominated immune responses[148,149]. We chose to analyze macrophages, because they may actually contribute to some immune-related phenotypic differences in these strains and they can

[B] submitted publication: Schilling,E., El-Chartouni,C. & Rehli,M. Allele-specific DNA methylation in mouse strains is mainly determined by *cis*-acting sequences

be grown under identical conditions from bone-marrow progenitors. It was planned to use tiling arrays for comparative methylation profiling. However, although possible, the coverage of the whole genome is very expensive due to the high amounts of arrays needed. To circumvent this problem, the analysis was restricted to regions where genes showed differences in their mRNA expression profile in bone marrow-derived macrophages (BMMs) from the two inbred mouse strains, because epigenetic differences are likely to occur in such regions. Gene expression analyses using microarrays of untreated BMMs as well as IL-4 treated BMMs were previously performed in our lab. Microarray hybridizations were performed in two independent biological replicates. The results are summarized in **Figure 5.11A.** With few exceptions, differential gene expression was already detected in untreated BMM. However, some genes were shown to be differentially regulated by IL-4 and are therefore included in the list of differentially expressed genes. In total, 311 hits (probes), representing 311 genes or transcripts were identified, which showed at least five-fold differences between strains. The array results were controlled on a single gene level by quantitative reverse transcriptase-(RT)-PCR for a number of representative genes. Quantitative PCR results are shown in **Figure 5.11B**. All chosen examples validated the microarray results, indicating that the generated list contains few, if any, false positives.

## 5.2.2 Design of custom tiling array

Based on the expression analysis, the set of 311 differential probes was further reduced by removing those which are redundant or covering invalidated transcripts, resulting in a final set of 165 known genes showing significantly different gene expression between C57BL/6 and BALB/c BMM. Based on the transcriptome analysis, a custom tiling microarray was designed using the eArray software from Agilent. Assuming that epigenetically differentially regulated regions were not only located within the promoter sequence itself, the tiling array was designed to cover whole regions around differentially expressed transcripts. Depending on the genetic neighborhood about 50 kb downstream and upstream of a specific gene were covered with partially overlapping microarray probes. Regions containing repeats were excluded by the eArray software in order to minimize cross-hybridization effects. In total, the custom designed 240k Array covered 28 mega bases of the mouse genome and contained 181 genomic segments, with a median size of 110 kb. These genomic regions contained more than 500 proximal promoter regions and at least 800 genes, which included a number of well known and immunologically relevant genes like *Plau*, *Marco*, *Ltb* and *Il1b* (see a complete list in **Table 12.2,** page 173). The tiling array also contained 16 randomly chosen control regions that were either equally expressed or not expressed at all in both cell types.




| Term | GO ID | P-value | |
|---|---|---|---|
| C-type_lectin | IPR001304 | 2.461e-04 | (99/5/14438/82) |
| Immune response | GO:0006955 | 5.173e-04 | (339/8/12558/69) |
| Carbohydrate binding | GO:0030246 | 5.386e-04 | (254/7/13943/78) |

| Term | GO ID | P-value | |
|---|---|---|---|
| Cellular polysaccharide metabolic process | GO:0044264 | 8.525e-05 | (46/4/12558/64) |
| Carbohydrate metabolic process | GO:0005975 | 1.127e-03 | (318/7/12558/64) |

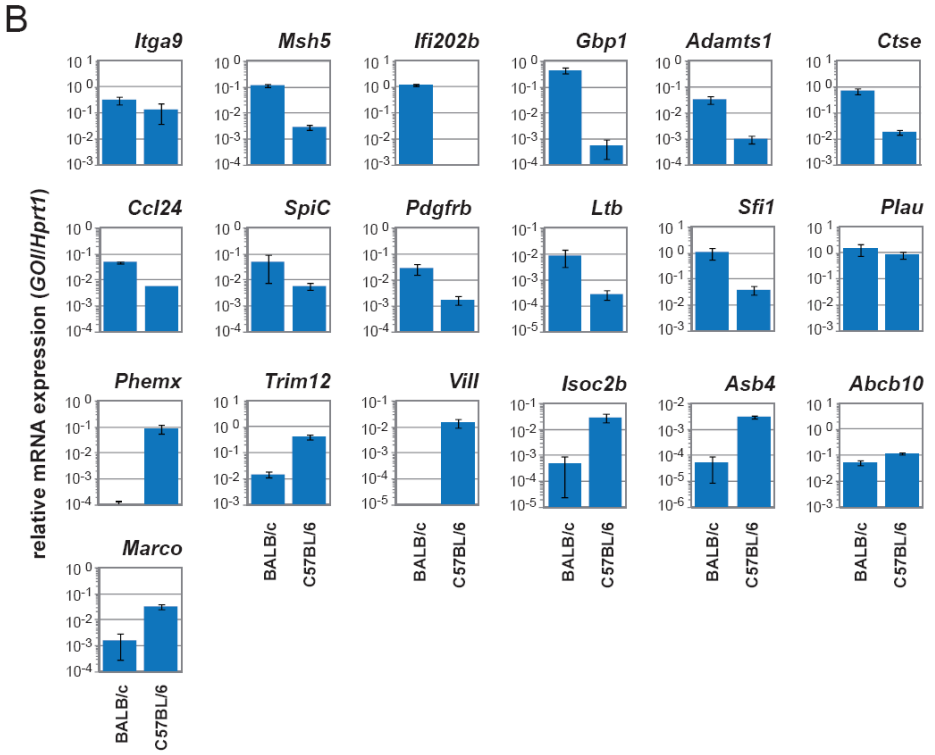



**Figure 5.11 Expression profiling of BMM from two inbred mouse strains.** (A) BMM from BALB/c and C57BL/6 strains were left untreated or treated with IL-4 for 4 h or 18 h. Expression analysis was performed using Mouse Whole Genome arrays (Agilent) to identify genes that are differentially expressed in either strain. Shown is a hierarchical clustering of 311 probes that show at least five-fold differences in median normalized signals between untreated (0 h) or IL-4 treated (4 h, 18 h) BMM. Gene ontology (GO) terms enriched in either of the two gene clusters are shown on the right. Numbers in brackets indicate: genes in cluster/genes in cluster with specific GO term/total genes/total genes with specific GO term. (B) Validation of differential gene expression in untreated macrophages using qRT-PCR. Results for all genes (GOI, genes of interest) were normalized for Hprt1 expression. Data represent mean values ± SD of at least four independent qPCR analyses from at least two independent experiments. The top two rows contain genes indentified as higher expressed in BALB/c, while the lower two rows show genes which are higher expressed in C57BL/6 mice.

### 5.2.3 MCIp “mirror-image” procedure

The general MCIp procedure is described in the methods section and was further improved for a novel application in section 5.1. Here I separated genomic DNA from macrophages from both mice strains into unmethylated (CpG) and methylated (mCpG) pools and co-hybridized the two unmethylated or the two methylated DNA pools onto locus-wide tiling array. As enriched DNA from one strain in the methylated fraction should be depleted in the unmethylated fraction, the signal intensities from CpG and mCpG pool hybridizations should complement themselves and, in addition, serve as an internal control. This “mirror-image” approach thereby allows the identification of differentially methylated regions (DMR), which is supported by reducing false positive hits. In addition, the stringency of hybridization was increased by a combination of a higher incubation temperature (67°C) and by adding formamide (15%) to the hybridization reaction mix. These modified reaction conditions were expected to additionally reduce crosshybridization events (nonspecific binding), which possibly give rise to misleading results. I used large scale MCIp (4 µg of each DNA) in order to get sufficient amount of gDNA for subsequent labeling and microarray hybridization. Prior to labeling, the separation of CpG methylation densities of individual MCIp fractions was controlled by qPCR using primers covering an imprinted region of Mest and a genomic region lacking CpG’s within a range of 1000 bp (empty_mm). The methylated and unmethylated alleles of Mest are separated in BMM's from each mouse strain and elute in the low salt fractions and high salt fractions, respectively. The region defined by the empty_mm primers (negative control) eluted in the low salt fraction of the fractionation due to the complete absence of CpGs (data not shown). After defining a threshold at a salt concentration of 500 mM, the fractions containing mainly unmethylated DNA (CpG pool) or methylated DNA (mCpG pool) were pooled. Subsequently the samples were directly labeled for microarray hybridization. C57BL/6 samples were labeled with Alexa-Fluor-5 and BALB/c samples using Alexa-Fluor-3. Efficiency of labeling was controlled with UV-spectroscopy and samples were hybridized on our custom design Agilent 240k microarrays as follows:

| | | | |
|---|---|---|---|
| 1) | C57BL/6 (CpG pool) repl. 1 | versus | BALB/c (CpG pool) repl. 1 |
| 2) | C57BL/6 (mCpG pool) repl. 1 | versus | BALB/c (mCpG pool) repl. 1 |
| 3) | C57BL/6 (CpG pool) repl. 2 | versus | BALB/c (CpG pool) repl. 2 |
| 4) | C57BL/6 (mCpG pool) repl. 2 | versus | BALB/c (mCpG pool) repl. 2 |

Image-data was extracted with Agilent feature extraction software and imported to Microsoft Excel for further analysis. Because the signal intensities were biased in correlation to their GC content (higher GC content lowered the average signals), the

probes were GC normalized. In general, the replicates for each pool correlated well within individual hybridizations ($r^2$=0.84 CpG pools and $r^2$=0.85 for mCpG pools). However, the analysis is additionally complicated by the fact that the sequence of microarray probes corresponds to only one strain's genome (the reference genome of C57BL/6) and sequence variants between mouse strains can additionally influence the hybridization efficiency of microarray probes. Since the extent of genetic variation between both strains at the investigated loci was largely unknown, it was necessary to analyze genetic variations between both mouse strains at first (see next paragraph). Utilizing the fact that essentially the entire genome of both strains was analyzed in two halves (CpG pool and mCpG pool), the two independent microarray results were re-combined to obtain a virtual comparative hybridization (vCGH) of both strains. **Figure 5.12** illustrates the whole procedure and explains the workflow.

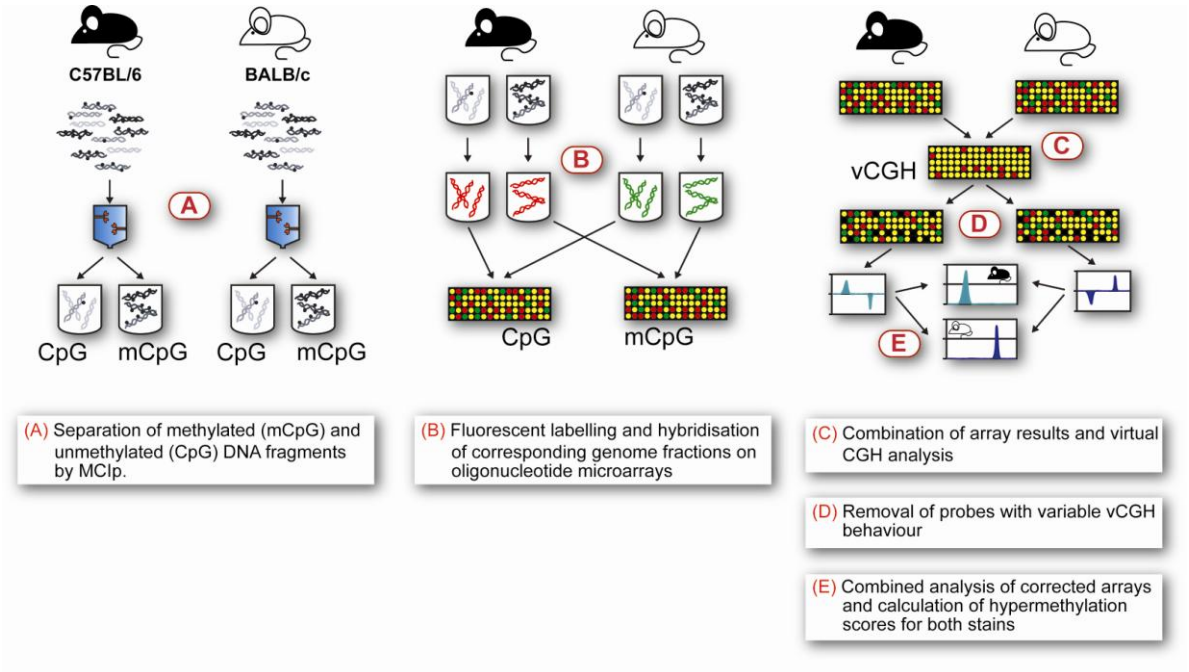



**Figure 5.12 Simultaneous detection of epigenetic and genetic differences using MCIp.** The experimental workflow is presented schematically. (A) Fragmented genomic DNA from bone marrow-derived macrophages of either mouse strain is fractionated using a MBD2-Fc column and separated into methylated (mCpG) and unmethylated (CpG) DNA pools. (B) Both DNA pools are fluorescently labeled and compared between mouse strains by co-hybridization on a locus-specific microarray using stringent conditions. (C) Array results are combined in a virtual CGH analysis to detect copy number variations (CNVs) and sequence polymorphic regions that are removed (D) from further analysis. (E) Differentially methylated regions are detected by analyzing remaining array probes for diametrically opposed enrichment behaviour between both hybridizations (e.g. a region that is relatively enriched in the unmethylated pool of BALB/c and shows reverse enrichment behaviour in the methylated pool is considered hypomethylated in BALB/c).

## 5.2.4 Virtual CGH to detect genetic differences

As mentioned above, the two genomes of both mouse strains were virtually compared. For this, the signal intensities of both independent experiments and both genome pools (from CpG pool and mCpG pool) at each individual probe were added up for each strain. Next, the $\log_{10}$ ratios of cumulative signal intensities (C57BL/6 / BALB/c) were plotted against their chromosomal localization. Probes hybridizing to genomic DNA of both strains to the same extent should exhibit a $\log_{10}$ ratio of cumulative signal intensities close to zero. Unbalanced hybridization behavior in the vCGH with a $\log_{10}$ ratio of cumulative signal intensities >0.15 as compared to the surrounding probes was used as a threshold for detecting genetic differences between strains. **Figure 5.13A** illustrates that 15.1% of all microarray probes (34096) demonstrated preferential hybridization with gDNA from C57BL/6 macrophages. In addition, the majority of known SNPs (13535) overlapped with probe sequences showing unequal hybridization behavior, whereas all other probes associated with relatively few SNPs (2801). A representative set of regions (9 regions) that were indicative of sequence variations were sequenced for both strains and it was found that every affected probe contained at least one, but more often two or three sequence variations including SNP and microdeletions or –insertions: An example of sequence alignment is shown in **Figure 5.13B** (the complete list of sequence alignments can be found in the appendix section). In addition, it was also found, that the sequence variation was generally higher in genes that showed differential expression in macrophages of both strains (**Figure 5.13A**, gray dots) than in control genes (either equally expressed or not expressed; **Figure 5.13A**, black dots). Furthermore a number of regions where vCGH signals indicated copy number variations (CNV) were also indentified. Two regions (*Nlrp1b* and A430090E18Rik) appear to be deleted in BALB/c, two regions (*Chia*/*Chi3l3*/*Chi3l4* and 2610305D13Rik) seem to be duplicated in C57BL/6 and five regions (*Btbd9*/*Glo1*/*Dnahc8*/*Glpr1*, *CD244*, *Ifi202b*, *Tmem14a* and *Gbp1*/*Gbp2*) appear at least duplicated in BALB/c. I validated gene duplications for three locations using quantitative real-time PCR (**Figure 5.13C**). Some of the larger CNV's were previously identified using standard CGH approaches[150] suggesting that vCGH analysis correctly identified genetic alterations between both mouse strains. All probes with underlying genetic differences were removed before microarray results were further analyzed for differential DNA methylation.

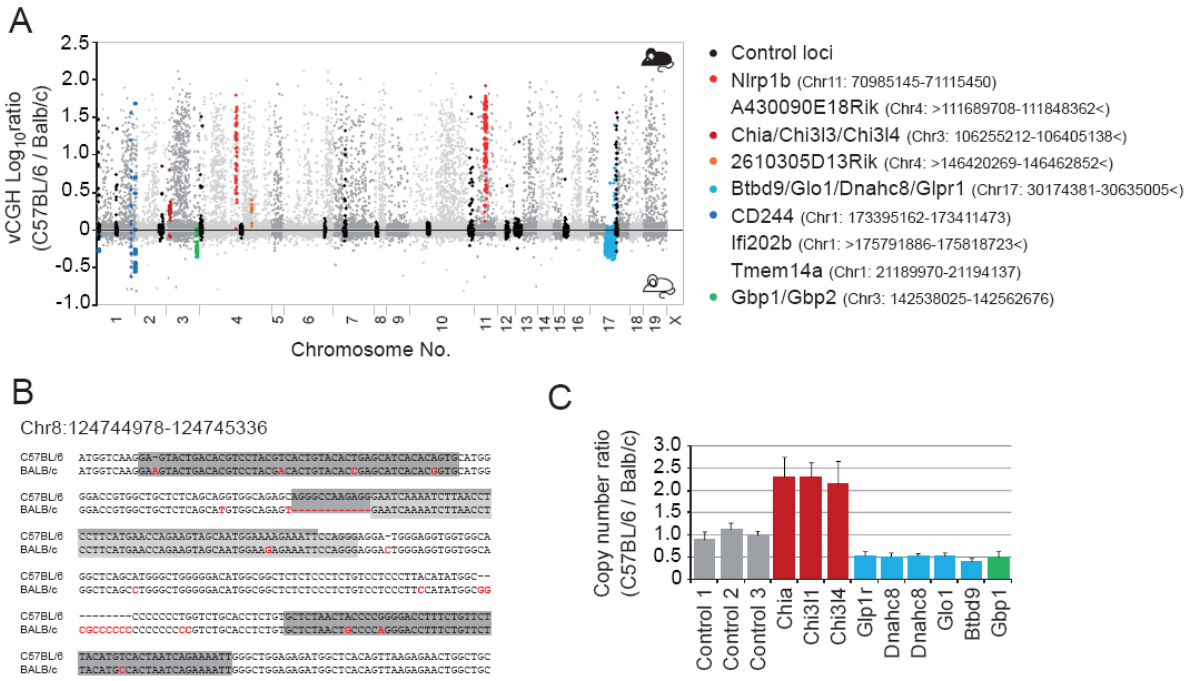



**Figure 5.13 Detection of sequence variation using virtual CGH.** (A) The histogram shows a CGH-like representation of combined signal intensities from separate hybridizations of methylated and unmethylated DNA pools. Control loci (in black) show relatively few C57BL/6-enriched signals. Two regions (*Nlrp1b* and A430090E18Rik, in red) are deleted in BALB/c, two regions (*Chia*/*Chi3l3*/*Chi3l4* and 2610305D13Rik; in dark red and orange, respectively) are duplicated in C57BL/6 and five regions (*Btbd9*/*Glo1*/*Dnahc8*/*Glpr1*, *CD244*, *Ifi202b*, *Tmem14a* and *Gbp1*/*Gbp2*; in blue or green) appear at least duplicated in BALB/c. Genomic locations are provided for individual regions. The signs > and < indicate that the affected regions extended over the analyzed area. (B) Sequences of an exemplary region where several probes (boxed in gray) showed C57BL/6-enriched signals. Compared to the reference strain, BALB/c contains several nucleotide exchanges (highlighted in red). (C) QPCR-validation of three copy number variations (*Chia*/*Chi3l3*/*Chi3l4*, three amplicons; *Btbd9*/*Glo1*/*Dnahc8*/*Glpr1*, five amplicons, *Gbp1*/*Gbp2*, one amplicon) detected by vCGH. Regions where no sequence variations were seen have been chosen as controls (control 1-3; *Camk2b*, *Chst10*, *Marco*).

## 5.2.5 Detection of DMR by microarray

The vCGH-corrected datasets were further analyzed to identify differentially methylated regions. All probes indicating sequence variations were excluded (34096). A representative scatter plot of vCGH-corrected comparative microarray hybridizations from unmethylated (CpG) and methylated pools (mCpG) is shown in **Figure 5.14**, where microarray probes showing the expected complementary behavior are colored in red (hypomethylated in C57BL/6 BMM) and blue (hypomethylated in BALB/c BMM). Next, hypomethylation scores were calculated by subtracting the $\log_{10}$ signal intensity ratios of the CpG pool hybridization from those of the mCpG pool hybridization of the same experiment. Hypomethylation scores were then analyzed using a sliding window approach. The window included five probes with a maximal distance of 500 bp between two neighboring probes. A cumulative hypomethylation score of the five probes >1.5 (in both replicate experiments) or >1.3 (if the value for each of the five probes was >0.15 in both replicate experiments) was used as a threshold for detecting hypomethylation in C57BL/6. A cumulative hypomethylation score of the five probes <-1.5 (in both replicate experiments) or <-1.3 (if the value for each of the five probes was <-0.15 in both replicate experiments) was used as a threshold for detecting hypomethylation in BALB/c.

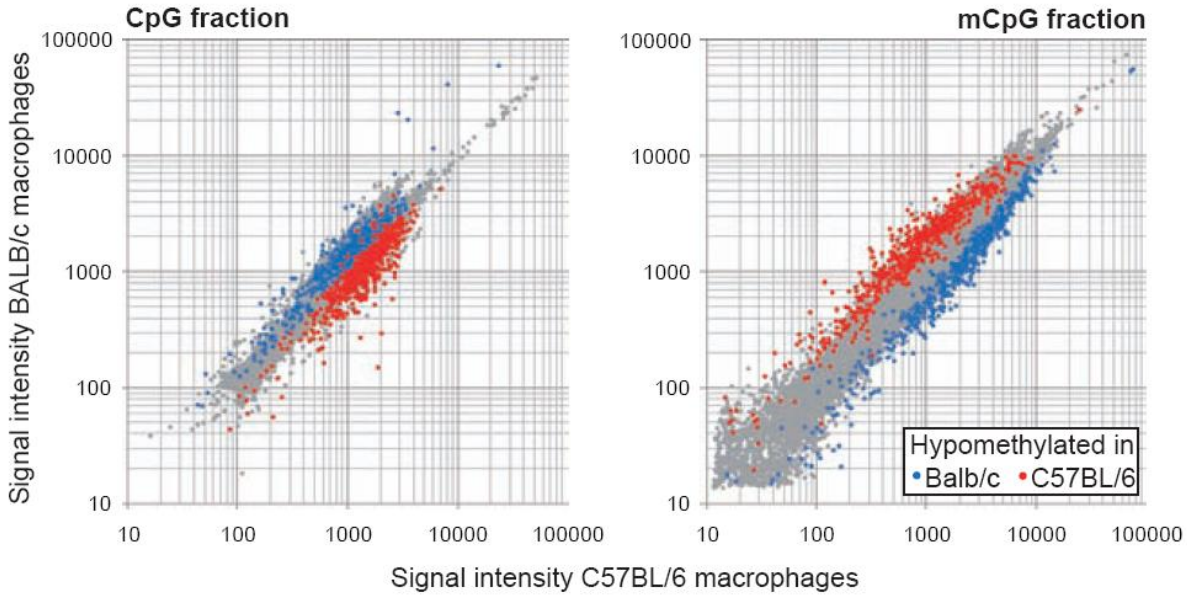



**Figure 5.14 Sequence regions with strain-specific CpG methylation.** (A) Scatter plots of normalized signal intensities from independent hybridizations of methylated (mCpG) and unmethylated (CpG) DNA pools. Probes in differentially methylated regions (colored in red and blue) show the expected intensity distribution (enriched in one pool and depleted in the other one).

In total, 435 regions with lineage-specific CpG methylation that were associated with 171 genes were identified. **Table 5.2** contains a list of 32 selected mouse strain-specific DMR. The genomic location (Build 36 assembly by NCBI, mm8) of the 'core' region of the DMR is given along with information on the associated (neighboring) gene, its relative position to this gene, its methylation status and BMM expression data (if available) for the associated gene. Microarray expression data was median-normalized to the untreated C57BL/6 BMM (0 h) sample (data for 4h and 18h treatments is not shown). The call status of individual probes in each sample is given in brackets behind the relative expression values (P, present; A, absent). As expected, DNA methylation at proximal promoter regions correlated with lower gene expression levels in all but one cases (except *Slc13a3*), whereas no significant correlation between methylation and transcription state was evident at promoter distal sites. In addition, a remarkable number (about 90%) of the identified DMR contained probes that showed unbalanced hybridization behavior in the vCGH, suggesting that the large majority of DMR are associated with genetic variations.

**Table 5.2** Annotated list of selected mouse strain-specific DMR

| ***Center of DMR (Mouse assembly mm8)*** | ***Gene symbol*** | ***Position relative to gene*** | ***De-methylated in*** | ***BMM BALB/c (0h)*** | ***BMM C57BL/6 (0h)*** |
|---|---|---|---|---|---|
| chrX:119220336-119220381 | ***3110007F17Rik*** | prom | C57BL/6 | 0.0515 (A) | 0.992 (P) |
| chr6:5333111-5334506 | ***Asb4*** | distal | C57BL/6 | 0.0642 (P.A) | 0.997 (P) |
| chr4:46594578-46596131 | ***Coro2a*** | intra | C57BL/6 | 0.209 (P) | 0.997 (P) |
| chr4:46610429-46610801 | ***Coro2a*** | intra | C57BL/6 | 0.209 (P) | 0.997 (P) |
| chr4:46620908-46622454 | ***Coro2a*** | intra | C57BL/6 | 0.209 (P) | 0.997 (P) |
| chr7:4064078-4064123 | ***Eps8l1*** | intra | BALB/c | 0.0861 (A) | 0.979 (P) |
| chr7:4073211-4073587 | ***Eps8l1*** | intra | BALB/c | 0.0861 (A) | 0.979 (P) |
| chr7:4073898-4074764 | ***Eps8l1*** | intra | BALB/c | 0.0861 (A) | 0.979 (P) |
| chr7:4081211-4081948 | ***Eps8l1*** | intra | BALB/c | 0.0861 (A) | 0.979 (P) |
| chr4:147309756-147309900 | ***Frap1*** | intra | C57BL/6 | 1.22 (P) | 0.998 (P) |
| chr7:4468706-4470122 | ***Isoc2b*** | prom | C57BL/6 | 0.188 (P) | 0.987 (P) |
| chr7:4476223-4476274 | ***Isoc2b*** | distal | C57BL/6 | 0.188 (P) | 0.987 (P) |
| chr5:53010351-53010409 | ***Pi4k2b*** | distal | C57BL/6 | 7.234 (P) | 0.999 (P) |
| chr5:53012858-53013017 | ***Pi4k2b*** | distal | C57BL/6 | 7.234 (P) | 0.999 (P) |
| chr5:53013951-53014431 | ***Pi4k2b*** | distal | C57BL/6 | 7.234 (P) | 0.999 (P) |
| chr5:53026418-53026767 | ***Pi4k2b*** | distal | C57BL/6 | 7.234 (P) | 0.999 (P) |
| chr5:53049276-53049634 | ***Pi4k2b*** | intra | BALB/c | 7.234 (P) | 0.999 (P) |
| chr5:53058948-53058993 | ***Pi4k2b*** | distal | BALB/c | 7.234 (P) | 0.999 (P) |
| chr5:53061668-53062228 | ***Pi4k2b*** | distal | C57BL/6 | 7.234 (P) | 0.999 (P) |
| chr5:53063596-53064069 | ***Pi4k2b*** | distal | BALB/c | 7.234 (P) | 0.999 (P) |
| chr7:37961808-37962602 | ***Pop4*** | distal | BALB/c | 0.1 (P) | 0.999 (P) |
| chr2:118921263-118922015 | ***Ppp1r14d*** | prom | C57BL/6 | 1.414 (A) | 0.996 (A) |
| chr9:102967337-102967700 | ***Rab6b*** | distal | C57BL/6 | 0.0133 (A) | 0.988 (P) |

| *Center of DMR (Mouse assembly mm8)* | *Gene symbol* | *Position relative to gene* | *De-methylated in* | *BMM BALB/c (0h)* | *BMM C57BL/6 (0h)* |
|---|---|---|---|---|---|
| chr9:102971565-102971612 | ***Rab6b*** | intra | BALB/c | 0.0133 (A) | 0.988 (P) |
| chr11:3092678-3093978 | ***Sfi1*** | prom | BALB/c | 16.38 (P) | 0.974 (P) |
| chr2:165163847-165163893 | ***Slc13a3*** | intra | BALB/c | 0.0413 (P.A) | 0.999 (P) |
| chr2:165165074-165165182 | ***Slc13a3*** | prom | BALB/c | 0.0413 (P.A) | 0.999 (P) |
| chr18:58681880-58682279 | ***Slc27a6*** | intra | C57BL/6 | 8.848 (P) | 0.997 (A) |
| chr18:58686133-58686695 | ***Slc27a6*** | intra | C57BL/6 | 8.848 (P) | 0.997 (A) |
| chr2:118931167-118931543 | ***Spint1*** | intra | C57BL/6 | 0.0282 (P) | 0.839 (P) |
| chr2:118933687-118934010 | ***Spint1*** | intra | C57BL/6 | 0.0282 (P) | 0.839 (P) |
| chr7:29692014-29692149 | ***Zfp568*** | prom | C57BL/6 | | |

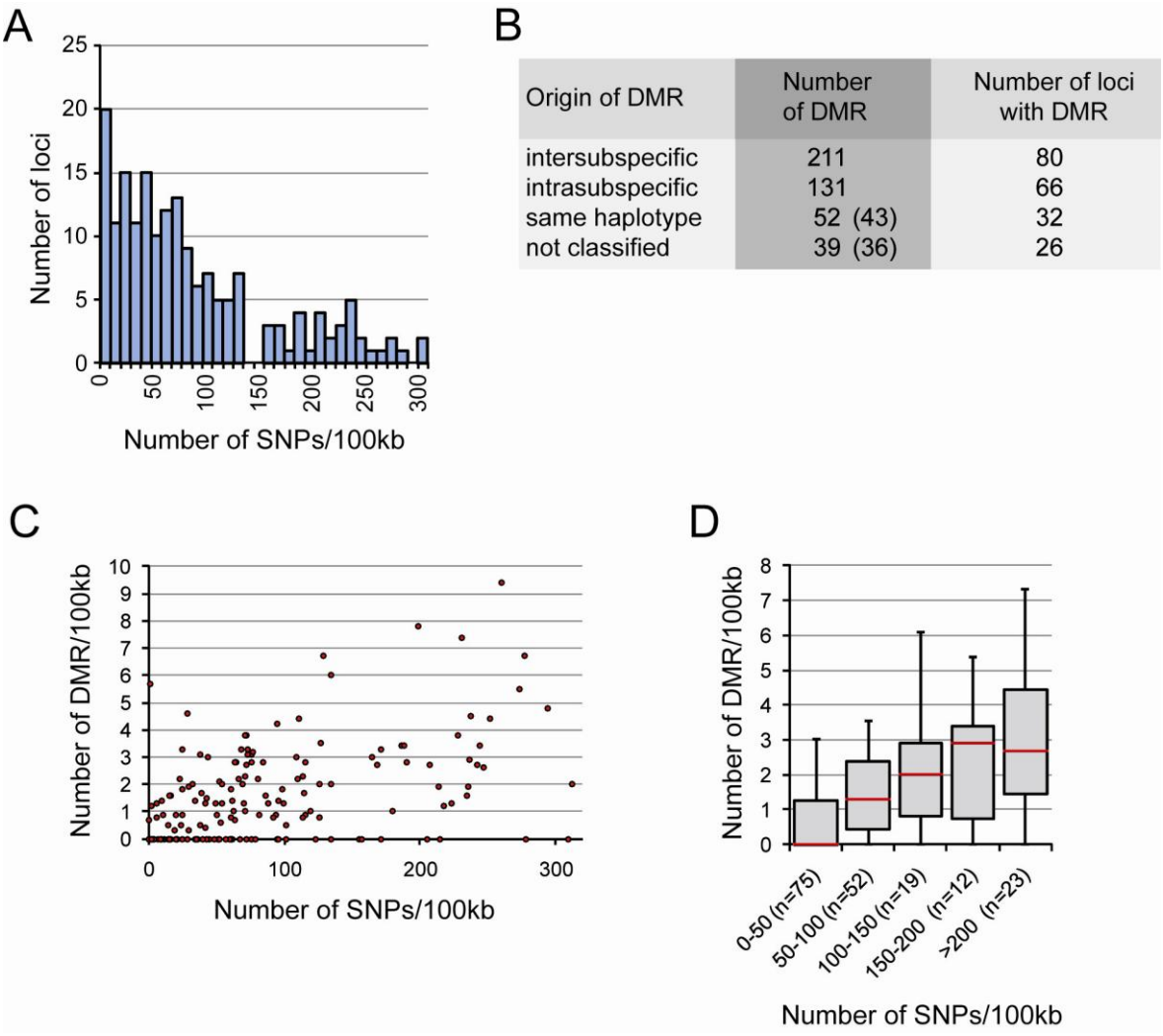
Complete list is shown in the appendix section 12.3.

### 5.2.6 Correlation between genetic variability and differential methylation status

As described earlier, the number of probes demonstrating unbalanced hybridization behavior (34096) was substantially larger that the number of known variations. This suggested that sequences of both strains are far more variable at the studied genomic regions than previously published SNP data[151] suggests. This is in line with an earlier study, suggesting that the published SNP data only captures a fraction of the variation found in the laboratory mouse[152]. To analyze if genomic intervals investigated in this study were enriched for genetic variation as compared to the whole genome, which could produce a bias towards *cis*-dependent DMR, we calculated the distribution of known SNP across all genomic loci on the microarray.

The distribution of known SNP markedly varied across all interrogated regions with 15-20% of all regions having comparably low or high SNP densities (less than 2/10000 or more than 15/10000, respectively; a diagram is provided in **Figure 5.15A**). The average SNP density (1 SNP every 1200 bp) on our array was approximately two times higher as compared to the whole genome (1 SNP every 2500 bp), suggesting that differentially expressed genes tend to contain more polymorphisms. Furthermore, we studied correlations between genetic variability and differential methylation status. According to the classifications provided by the Perlegen-study[151], eighty of the genomic intervals contained DMR of intersubspecific origin and sixty six intervals contained DMR of intrasubspecific origin. Thirty two regions were predicted to be of the same haplotype. However, the majority of DMR in the latter class were also associated with probes showing unbalanced hybridization behavior in the vCGH, suggesting that most of them actually are of intrasubspecific origin (**Figure 5.15B**). In total, about 90% of the identified DMR contained probes that showed unbalanced hybridization behaviour in the vCGH.

DMR and SNP count per region correlated ($r^2$=0.59) suggesting a link between DMR occurrence and genetic variation between both strains. A corresponding diagram and a box plot on DMR/SNP correlation are provided in **Figure 5.15C** and **Figure 5.15D**.




| Origin of DMR | Number of DMR | Number of loci with DMR |
|---|---|---|
| intersubspecific | 211 | 80 |
| intrasubspecific | 131 | 66 |
| same haplotype | 52 (43) | 32 |
| not classified | 39 (36) | 26 |

**Figure 5.15 Correlation of DMR and genetic variation.** (A) The bar graph shows the distribution of SNP-densities across all genomic loci on the microarray (bin size 10 SNP/100000 bp). (B) The origin of each DMR was determined using the Mouse Ancestry Mapper and Mouse Haplotype Block Viewer provided by Perlegen Sciences (mouse.perlegen.com). The first column indicates classifications for individual DMR. For regions that were either not classified or predicted to be of the same haplotype the number of DMR showing unbalanced hybridisation behaviour are given in brackets. The second column indicates how many loci were associated with either type of classification. (C) Numbers of DMR/100kb were plotted against SNP counts per 100 kb for all loci present on the array. (D) The box plot shows the distribution of DMR counts conditional on the SNP counts at corresponding loci. The red lines denote medians, boxes the interquartile ranges, and whiskers the 5th and 95th percentiles. The total numbers of loci within each SNP-density group are given in brackets.

## 5.2.7 MALDI-TOF mass spectrometry: Validation of microarray data

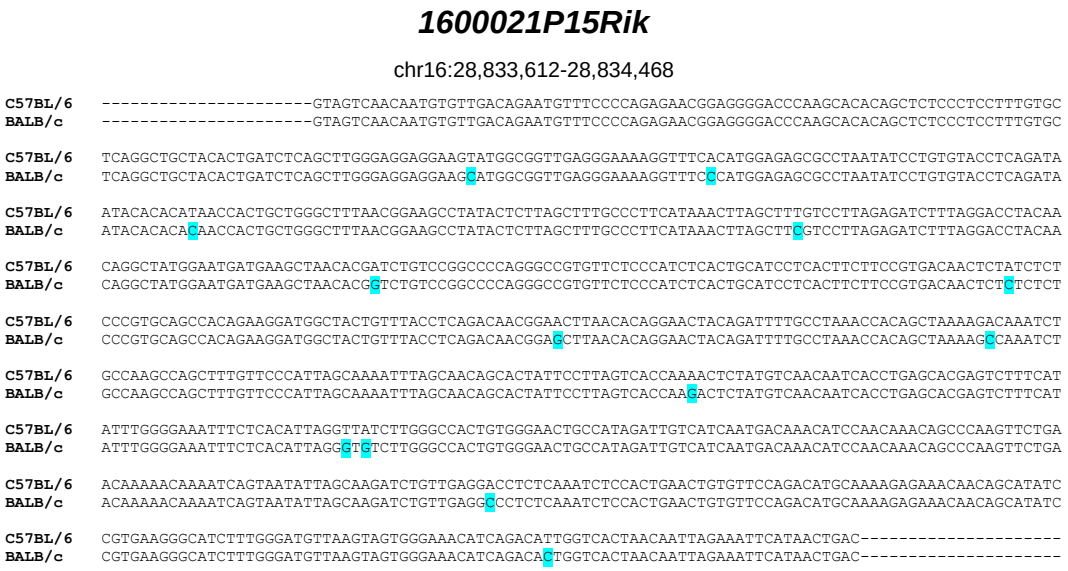
### 5.2.7.1 Method characteristics

To validate and quantify methylation differences at the resolution of single CpGs, a representative set of DMR (from **Table 5.2**, page 92) was chosen to be analyzed by MALDI-TOF mass spectrometry. This method is based on matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) measurement of bisulfite converted DNA using the EpiTYPER platform (Sequenom, San Diego, USA). Bisulfite treatment generates methylation dependent sequence variations, which can be measured using MALDI-TOF mass spectrometry. Moreover this procedure enables one to analyze multiple CpGs in one amplicon and to compare their methylation status between individual samples. Unlike classic methods for assessing methylation events at individual sites in the genome, EpiTYPER quantitatively assesses methylation ratios simultaneously across multiple CpG sites over multiple samples (see methods section for more details).

### 5.2.7.2 Genomic DNA sequences in C57BL/6 and BALB/c

Since many sequence variations were detected between both mouse strains, it was necessary to obtain the precise genomic DNA sequences at the desired loci prior to primer design for MALDI-TOF analysis. Initially a set of 22 regions was chosen based on the MCIp microarray results and underlying genomic DNA was sequenced from two male individuals of both strains. DNA sequencing indicated that DMRs were often associated with strain-specific insertions of CpG-containing repetitive sequences, not overlapping with microarray probes, which were likely methylated and responsible for the observed methylation pattern. Thus, four regions were excluded from further analysis, mainly because of large indels (insertion-deletion polymorphisms). The complete lists of sequence alignments between C57BL/6 and BALB/c are attached in the appendix section (see section 12.4, page 187). The residual 18 regions (see **Table 5.3**, page 100) were analyzed on the EpiTYPER platform. With one exception (*Eps8l1*), all of the analyzed regions contained sequence variations between mice strains. Four representative sequence alignments, including 1600021P15Rik (A), *Coro2a* (B), 3110007F17Rik (C) and *Zfp568* (D) genes, are presented **Figure 5.16** on the following two pages. All four examples exhibited single nucleotide exchanges within BALB/c sequences as compared to the C57BL/6 reference strain (13 in (A), 3 in (B), 21 in (C) and 12 in (D)). The latter two regions additionally contained small deletions (two in 3110007F17Rik) or insertions (one in *Zfp568*) within the sequenced BALB/c regions.

**A)**




**B)**

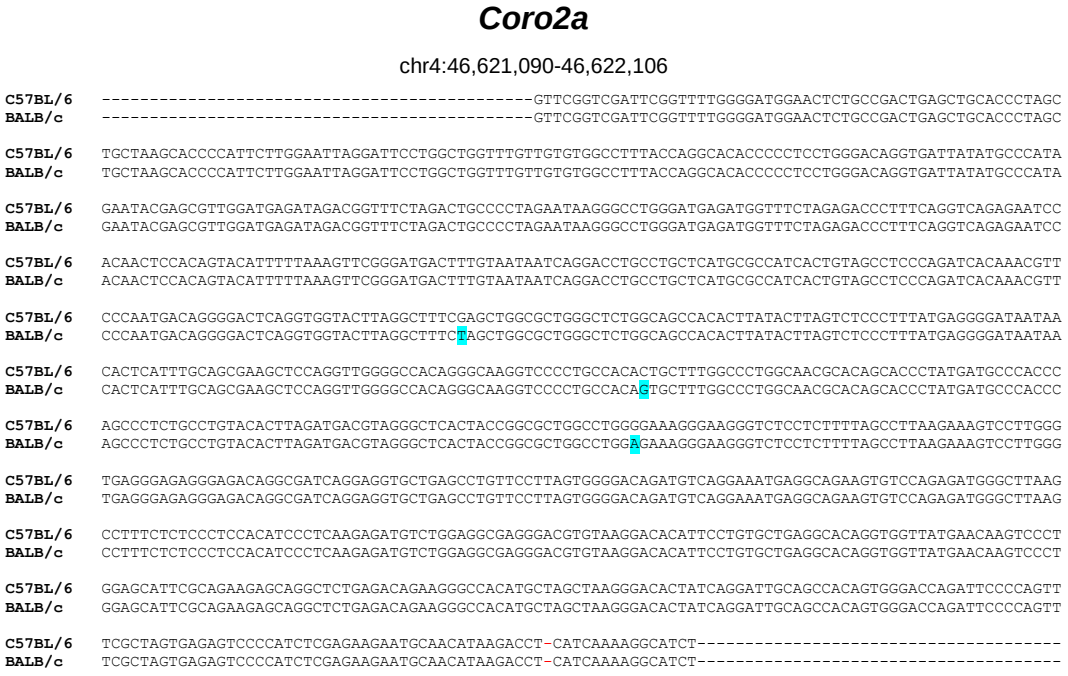



**Figure 5.16 DNA sequences of 4 selected DMRs in C57BL/6 and BALB/c.** Regions that were selected as possible targets for methylation analysis using MALDI-TOF MS. Genomic DNA was amplified by PCR and directly sequenced. Differences between BALB/c and the reference strain C57BL/6 are marked as follows: deletions or insertions (relative to C57BL/6) are in red lettering; single nucleotide polymorphisms are boxed in blue (BALB/c). The Figure is continued on the next page showing two more examples. The complete list of sequence alignments is shown in the appendix section (see section 12.4, page 187).

**C)**

***3110007F17Rik***

chrX:119,220,163-119,220,873

```
C57BL/6   ------------------TCTCGTTCAGCTGTTCTACCCACTGGTCTAGCTCCTTGGTGAATGCCTTGTCGTCTATGGCGGTGAGCTGCTCATCTCCGCC
BALB/c    ------------------TCTCGTTCAGCTGTTCTACCCACTGGTCTAGCTCCTTGGTGAATGCCTTGTCGTCTCTGGCGGTGAGCTGCTCATCTCCGCC

C57BL/6   GCCTCCCCCTCCGCTGATCCCCTGGCGGGACCGCTAAGCTGCAGAGAGCGCAGCCCGCGCCCCGCCCAAGCCCAGCCATTATCTTCGGGTCCCTGCTGTT
BALB/c    GCCTCCCCCTCCGCTGATCCCCTGGCGGGACCGCTAAGCTGCAGAGAGCGCAGCCCGCGCCCTGCCCAAGCCCAGCCATTATCTTCGGGTCCCTGCTGTT

C57BL/6   TGCCTCTAAGTCTGTGGCGGTTCCGTAGCTGTAGTGCCTGGCGGCCCCAGCCCCGCCCACGCGCTTCTGTCGCACCGCCCCCTGCGCTGTTGTCAGAGCG
BALB/c    TGCCTCTAAGTCTGTGGCGGTTCCGTAGCTGTAGTGCCTGGCGGCCCCAGCCCCGCCCACGAGCTTCTGTCGCACCGCCCCCTGCGCTGTTGTCAGAGCG

C57BL/6   CGGGACACAGTGAGGGAGGGGGGAATGGGCGAGCGGTCGCGCCCGTTGGGAGGAGCAGTCCGGACCCATGCCCGTGCAGGCAGGACATGGGGAGAGCACGC
BALB/c    CGGGACACAGTGAGGGAGGGGGGAATGGGCGAGCGGTCGCGCCCATTGGGAGGAGCAGTCCGGACCCCTGCCCGTGCAGGCGGGACGTGGGGAGAGCACGC

C57BL/6   AGCGACAGGCATCTGGCTTTCTGAGCCCTGGGATTTGGTGAGCGAACACTCGGAGTGGGATGGTTCGGCGCGCGCTGAGACCCGCTATCCCGGCCCCCGC
BALB/c    AGCGACAGGCATCTGGCTTTCTGAGCCCTGGGATTTGGCGAGCGAACACTCGGAGTGGGA----------------AGACCCGCTATCCCGGCCCCCGC

C57BL/6   CCACCTTTCCCAGCGGGCCTTTCCAGCGGTCCCGGGGCGGTCTTGGTGTACTGCCATGGTGGGTGGAGGTGCCCTTGGGGCCTGGTCGCTGGCCAGGGGG
BALB/c    CCACCTTTCCCAGCGGGCCTTTCCAGCCGTCCCAGGGCGGTCGAGGTGTACCGCCATGGTGG----AGGTGCGCTTGGGGCCTGGTCGCTGGCCAGGGGG

C57BL/6   GGCAGTCCCCCGCCGACCCCAGCCTGGCCCTGCCCGCGAGCCCGCCACCAGAGCACGGGACGCAGGAAAAGGGGGCGGGGAATGACGCGCGGGTGGAGGA
BALB/c    GGCAGTCCCCAGCCGACCCCAGCCTGGCCCTGCCCGCGCGCCGCCAGAGCACTGGACGCAGGAAAAGGGGACGGGGAATGACGCGCGGGTGGAGGA

C57BL/6   GTTTTCGCTGTGCGCCGCCAGCACAGTTG----------------------------------------------------------------------
BALB/c    GTTTTCACTGTGCGCCGCCAGCACAGTAG----------------------------------------------------------------------
```

**D)**

***Zfp568***

chr7:29,691,813-29,692,926

```
C57BL/6   ---------------------------------------GCTATTCTAAGA-CAGAGGTCTTATCTCAGCCTTTAGGGAGCTCGCCCAGGCCTGGTGGTTTT
BALB/c    ---------------------------------------GCTATTCTAAGAACGGAGGTCTTATCTCAGCCTTTAGGGAGCTCGCCCAGGCCTGGTGGTTTT

C57BL/6   ACTGCAGCTTGAAAGAAAGTAAGGCTTGCGGGCATGCAGTAAGATGGGGGCGGCCGTCCAATACCTGGGATCACTCTGGGCTCGGCACAAAGGAAAGCCC
BALB/c    ACTGCAGCTTGAAAGAAAGTAAGGCTTGCGGGCATGCAGTAAGATGGGGGCGGCTGTCCAATACCTGGGATCACTCTGGGCTCGGCACAAAGGAAAGCCC

C57BL/6   GGCTCTTTGTGGTTTGTAGTGCATTCTTTGCCTCTTTTGGGTAACTGACCTGGCAATGGGTAAAACTGGGAGTAGACATAAAAAAATTGTGACCCTCCAG
BALB/c    GGCTCTTTGTGGTTTGTAGTGCATTCTTTGCCTCTTTTGGGTAACTGACCTGGCAATGGGTAAAACTGGGAGTAGACATAAAAAAATTGTGACCCTCCAG

C57BL/6   GTCTTTCTTGGCCAGGCCAGAGATCATTACCAGCCTTGTCCCCATTATCCATCCTGCGTAGAGAATGGAGATTAGGAGAGAACTGGCACCCTGACCAATT
BALB/c    GTCTTTCTTGGCCAGGCCAGAGATCATTACCAGCCTTGTCCCCATTATCCATCCTGCGTAGAGAATGGAGATTAGGAGAGAACTGGCACCCTGACCAATT

C57BL/6   TGCGAATCTAAACCTCACCTCAGAGATTGTTCCCAGGAGCAGAGTTCTGAAGTCACAACCACTCGGGTGACAGGAGCCCCATGGTTACATGCAAGATGCC
BALB/c    TGCGAATCTAAACCTCACCTCAGAGATTGTTCCCAGGAGCAGAGTTCTGAAGTCACAACCACTCGGGTGACAGGAGCCCCATGGTTACATGCAAGATGCC

C57BL/6   ACATAGAGGTAGGCACAGTCTT----------------------------------------------------------------------------
BALB/c    ACATAGAGGTAGGCACAGTCTTTTTTTTGTTTGTTTGTTTGTTTTTTGTTTTTTTTTGTTTGTTTTTGTTTTTTTGAGACAGGGTTTCTCTGTATAGCCC

C57BL/6   ----------------------------------------------------------------------------------------------
BALB/c    TGGCTGTCCTGGAACTCACTCTGTAGACCAGGCTGGCCTCGAACTCAGAAATCCGCCTGCCTCTGCCTCCCGAGTGCTGAGATTAAAGGCGTGCGCCACC

C57BL/6   ---------------------------ACACTTAAGCCACAACTCCACAACGACCAGACATTTCATGTCCTCATAGAAAAGCTATCAGAATGTATTACAAAC
BALB/c    ACGCCCGGCTGGTAGGCACAGTCTTACACTTAAGCCACAACTCCACAACGACCAGACATTTCATGTCCTCATAGAAAAGCTATCAGAATGTATTACAAAC

C57BL/6   AACCACAAAAACTATTCCAGGGCACTGCGTCCGCCCAGTCGCAAGGCTCCTGCCGAGAGGCTGGACAGGCCACTAGCCTTTCTGGGAAATGTAGTCCACA
BALB/c    AACCACAAAAACTATTCCAGGGCACTGCGTCCGCCCAGTCGCAAGGCTCCTGCCGAGAGGCTGGACAGGCCACTAGCCTTTCTGGGAAATGTAGTCCACA

C57BL/6   AAGGGACAGCCACCTAGCGCCCCCAGCGCCTCCTCTGGTTGTTCCGGAAAAGACCGTCCCCGCAGCTGCTCACGTCCAGCGCTGCTTCGCCGGCGGCGTC
BALB/c    AAGGGACAGCCACCTAGCGCCCCCAGCGCCTCCTCTGGTTGTTCCGGAAAAGACCGTCCCCACAGCTGCTCACGTCCAGCGCTGCTTCGCCGCCGGCGTC

C57BL/6   ACCGTCTCGCCGCGCTAATGGGCTCAGCCACCTCGGCGCCTTCGGCCATGACTCCTGTCCCTAGCACAGGAAACGCCAGCCGAGCTCTGACCGGAAGTTG
BALB/c    ACCGTCTCGCGGCGCTAATGGGCTCAGCCACCTCGCGCCTTCGGCCATGACTCCTGTCCCTAGTACAGGAAACGCCAGCGGAGCTCTGACCGGAAGTTG

C57BL/6   GGTCGGGCCGCGGGTCCTTCTGGGAGCACCTGTGCTCTGGGACATAACGGGCGCGCGCTGCCGCCCGAGTGGAGGGGCTGGCCGTGCGTGGCCCGCATGC
BALB/c    GGTCGGGCCGCGGGTCCTTCTGGGAGCACCTGTGCTCTGGGACATAACGGGCGCGCGCTGCTGCCCGAGTGGAGGGGCTGGCCGTGCGTGGCCCGCATGC

C57BL/6   CGCGTCTCTGAGGTTGCGTCTCGGGAGGTTTTGGCTGGTGGCGCTGCGGGGACCTCTCCACGGGAGAGGCACGGGTGGTCCGGTTTGCGCCGGCCGTGGG
BALB/c    CGCGTCTCTGAGGTTGAGTCTCGGGAGGTTTTGGCTGGTGACGCTGCGGGGACCTCTCCACGGGAGAGGCACGGGTGGTCCGGTTTGCGCCGGCCGTGGG

C57BL/6   CTTCGGGCTGGGCGCGGGCTCTCGGGGGGCTCGGAGTATCCCGGGTCACAGCTTT-----------------------------------------------
BALB/c    CTTCGGGCTGGGCGCGGGCTCTCGGGGGGCTCGGAGTATCCCGG-TCACAGTTTT-----------------------------------------------
```

**Figure 5.16** continued from previous page.

### 5.2.7.3 Validation of DMR by combination of bisulfite and MALDI-TOF MS

The validation panel primarily focused on sequences showing no repeat variations. First, an initial round of MALDI-TOF MS analysis for the 18 selected DMRs was performed using gDNA from two male mice of each strain (C57BL/6 and BALB/c). Six of these 18 DMRs resided within the promoter region of the associated gene (3110007F17Rik, *Isoc2b*, *Ppp1r14d*, *Rab6b*, *Slc13a3* and *Zfp568*), *Asb4* and *Pi4k2b* reside in regions distal to the promoter and the remaining 10 DMRs are associated with intragenic sequences. Genomic sequencing of the *Sfi1* promoter region indicated the presence of an additional pseudogene (ps) of unknown chromosomal localization. This pseudogene differed slightly in sequence and was present in both strains, according to the equal PCR sequencing results from both strains. Specific primers for both possible *Sfi1* regions (endogenous copy and pseudogene) were designed in order to discriminate between them.

In total I designed 60 primer pairs for the amplification of bisulfite-treated DNA. Five primer pairs were specific only for the underlying sequence of C57BL/6 and four primer pairs were specific for BALB/c sequences. The remaining 51 primer pairs resided in regions where no sequence polymorphism was detected within the primer binding sites between the two strains. This ensures an unbiased amplification of both alleles (C57BL/6 or BALB/c derived) independent of their origin. Data for the validation set is summarized in **Table 5.3** and four examples are shown and described in **Figure 5.17**. (The complete MALDI-TOF MS data will be available online within the supplementary information of the corresponding publication).

All four examples shown, exhibited a similar degree of methylation difference between the strains as detected by the microarray experiment. Thus, the MALDI-TOF analysis proved that the microarray experiment correctly identified DMRs. Thirteen out of eighteen regions showed ‘true’ methylation differences. In the remaining five cases, the differential MCIp enrichment behavior resulted from an increased number of methylated CpGs in one strain, either due to the presence of SNPs (*Pi4k2b*, *Ppp1r14d*, Spint1) or due to the insertion of methylated sequences in one strain (*Pop4*, *Zfp568*). Alleles with higher numbers of methylated CpG lead to a stronger binding to the MBD-Fc column during the MCIp procedure and therefore appear as stronger methylated in the microarray data.

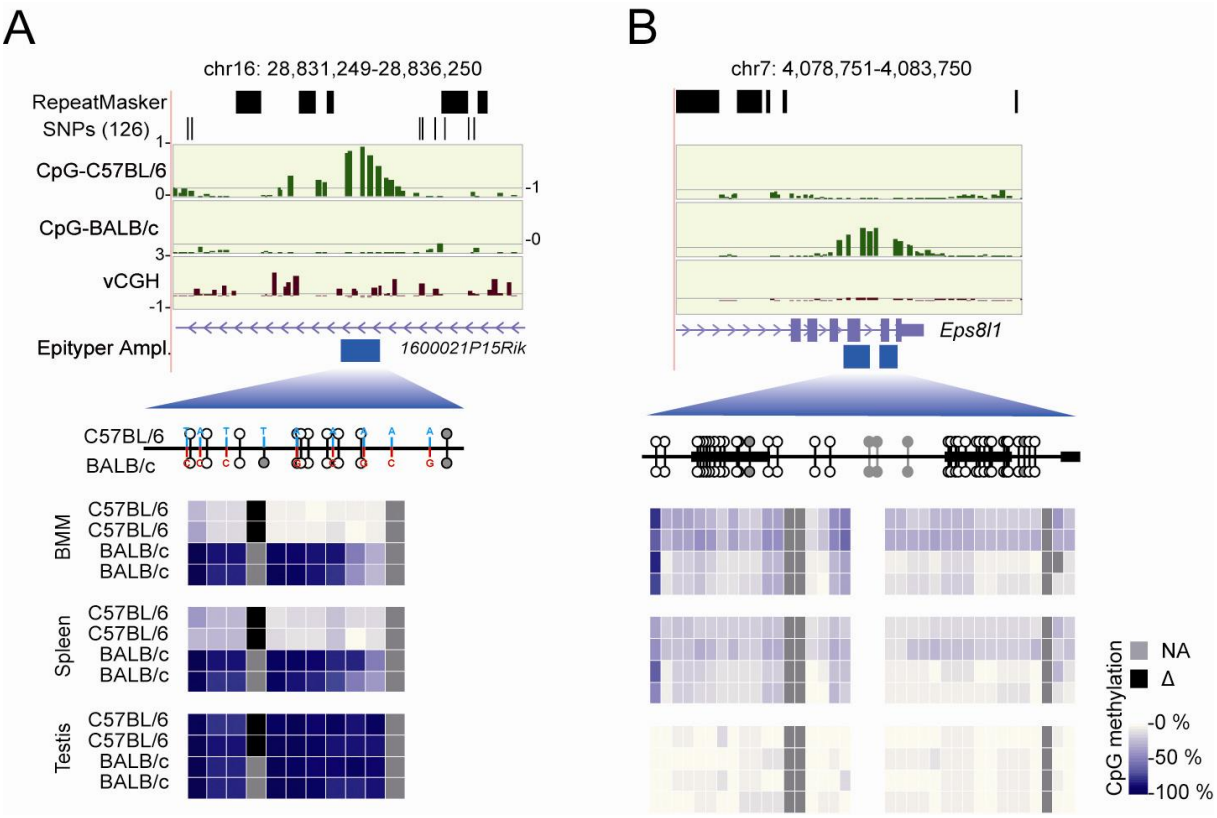



**Figure 5.17 Validation of strain-specific CpG methylation by MALDI-TOF MS of bisulfite treated DNA.** Four examples of DMRs detected by the MCIp-microarray approach and validation using MALDI-TOF MS of bisulfite treated DNA: (A) 1600021P15Rik, (B) Eps8L1, (C) Pdgfrb and (D) Coro2a (C and D are shown on next page). MCIp results are presented in the upper panels. Shown are the following tracks (from top to bottom) that were generated using the UCSC Genome Browser: repetitive regions as identified by the RepeatMasker program, single nucleotide polymorphisms from the dbSNP (NCBI database for genomic variation) build 126 (both in black), hypomethylation scores for BMM of both mouse strains (defined as the difference product of log10 signal intensity ratios of both hybridizations; shown in green), vCGH signals indicating the presence of genetic variation at probe level (in brown) as well as gene structures (in light blue) and the position of amplicons (Epityper Ampl.; in blue) that were designed for MALDI-TOF MS analysis of bisulfite treated DNA. The relative position of CpGs within amplicons is indicated below by small lollipops (with the upward orientation representing C57BL/6, and the downward orientation representing BALB/c). Sequence variations are highlighted in red and blue, black bars mark the position of exons and gray lollipops are not analyzed by the MS due to system restrictions (explained in the methods section). Methylation levels of individual CpGs in the indicated cell types (two individuals for each strain) are shown color-coded. The scale ranges from pale yellow (0% methylation) to dark blue (100% methylation), strain-specifically absent CpGs are coloured black, non-detectable CpGs are marked in gray.

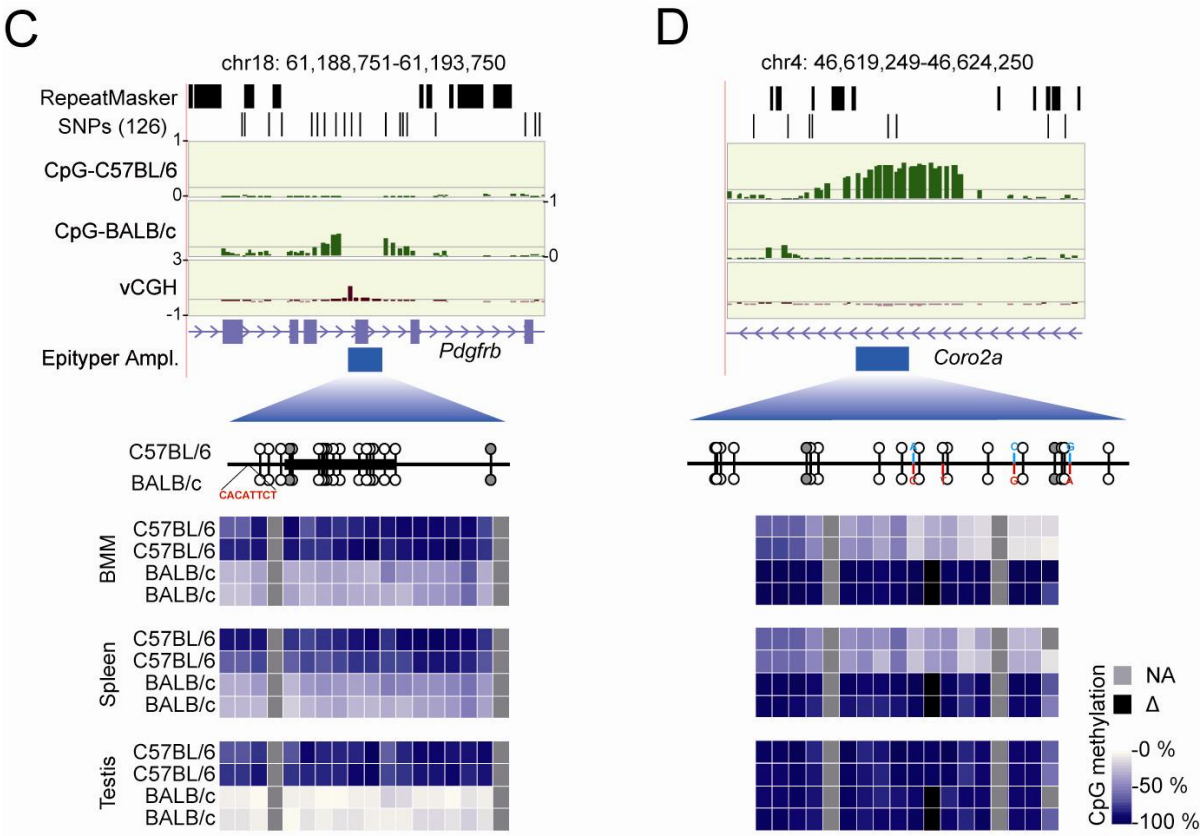



**Figure 5.17** continued from previous page.

**Table 5.3** MCIp-detected DMR validation set and germ line methylation

| Genomic Region (Mouse Assembly mm8) | Associated Gene | Distance from TSS (in kB) | MCIp indicated Demethylation in | Bisulfite-MS analysis (Nos. of amplicons) | Methylation status in sperm (C57 – BALB) |
|---|---|---|---|---|---|
| chr16: 28833304-28834357 | *1600021P15Ri* | 15.6 | C57BL/6 | confirmed DMR (2) | mCpG - mCpG |
| chrX: 119220336-119220381 | *3110007F17Rik* | -0.1 | C57BL/6 | confirmed DMR (5) | CpG - CpG |
| chr6: 5333111-5334506 | *Asb4* | -6.6 | C57BL/6 | confirmed DMR (6) | CpG - CpG |
| chr4: 46620908-46622454 | *Coro2a* | 1.3 | C57BL/6 | confirmed DMR (4) | mCpG - mCpG |
| chr7: 4064078-4064123 | *Eps8l1* | 0.2 | BALB/c | confirmed DMR (3) | CpG - CpG |
| chr4: 147309756-147309900 | *Frap1* | 17.4 | C57BL/6 | confirmed DMR (3) | mCpG - mCpG |
| chr7: 4468706-4470122 | *Isoc2b* | -0.2 | C57BL/6 | confirmed DMR (6) | CpG - CpG |
| chr18: 61190864-61191597 | *Pdgfrb* | 20.7 | BALB/c | confirmed DMR (2) | mCpG - CpG |
| chr5: 53026418-53026767 | *Pi4k2b* | -3.2 | C57BL/6 | one additional mCpG in BALB/c (2) | mCpG - mCpG |
| chr7: 37961808-37962602 | *Pop4* | 18.2 | BALB/c | methylated insertion in C57BL/6 (2) | m-CpG - ND |
| chr2: 118921263-118922015 | *Ppp1r14d* | -0.3 | C57BL/6 | two additional mCpGs in BALB/c (2) | mCpG - mCpG |
| chr9: 102968600-102969350 | *Rab6b* | -0.9 | BALB/c | confirmed DMR (2) | CpG - CpG |
| chr11: 3092678-3093978 | *Sfi1* | 0.1 | BALB/c | confirmed DMR* (5) | mCpG - CpG |
| chr2: 165163847-165165182 | *Slc13a3* | -0.1 | BALB/c | confirmed DMR (3) | CpG - CpG |
| chr18: 58681880-58682279 | *Slc27a6* | 0.5 | C57BL/6 | confirmed DMR (2) | CpG - CpG |
| chr18: 58686133-58686695 | *Slc27a6* | 4.8 | C57BL/6 | confirmed DMR (2) | CpG - CpG |
| chr2: 118931167-118931543 | *Spint1* | 2.5 | C57BL/6 | two additional mCpGs in BALB/c (4) | mCpG - mCpG |
| chr7: 29692014-29692149 | *Zfp568* | -0.6 | C57BL/6 | methylated insertion in BALB/c (4) | CpG - CpG |

## 5.2.8 Analysis of gender differences

Since the initial screen included only male mice, it was also of interest to look at gender specific methylation differences. To address this issue BMM of female individuals of each strain were examined as well. In addition, DNAs with different grades of methylation (0%, 33%, 66% and 100% methylation, generated by mixing SssI-methylated DNA and unmethylated DNA produced in the appropriate ratios; see methods section for more details) were analyzed to test if the procedure yields reliable and accurate results in the data analysis. To accomplish this and for all following MassARRAY experiments the initial set was reduced to 24 amplicons. Since some primer pairs amplified the same region, but on different DNA strands, it was possible to exclude one or several pairs without losing information. In addition, other regions, originally covered by several primer pairs, were restricted to fewer primer pairs, especially when removed primer pairs were redundant in the methylation level as investigated with the remaining primer pairs.

We analyzed genomic DNA from BMM of three different female individuals of each strain and also included a third male per strain in order to limit possible individual differences. For correlation analysis the complete dataset was reduced to all autosomal CpG units that were measured in all samples. The comparison of Bisulfite/MALDI-TOF MS analysis of methylation patterns in males and females suggested that the DMRs were equally detected on autosomes of both genders. A correlation matrix for this calculation is illustrated in **Figure 5.18**. All three biological replicates of each strain and gender were highly similar for the analyzed regions. (median $r^2$>0.97). Moreover, both genders also correlated well within one strain (median $r^2$>0.97).

| | C57BL/6 male BMM (1) | C57BL/6 male BMM (2) | C57BL/6 male BMM (3) | C57BL/6 female BMM (1) | C57BL/6 female BMM (2) | C57BL/6 female BMM (3) | BALB/c male BMM (1) | BALB/c male BMM (2) | BALB/c male BMM 3) | BALB/c female BMM (1) | BALB/c female BMM (2) | BALB/c female BMM (3) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| C57BL/6 male BMM (1) | 1.00 | 0.97 | 0.97 | 0.97 | 0.97 | 0.97 | 0.05 | 0.05 | 0.04 | 0.05 | 0.05 | 0.08 |
| C57BL/6 male BMM (2) | | 1.00 | 0.97 | 0.98 | 0.95 | 0.95 | 0.04 | 0.04 | 0.03 | 0.04 | 0.04 | 0.07 |
| C57BL/6 male BMM (3) | | | 1.00 | 0.98 | 0.99 | 0.99 | 0.07 | 0.07 | 0.05 | 0.07 | 0.06 | 0.11 |
| C57BL/6 female BMM (1) | | | | 1.00 | 0.98 | 0.97 | 0.06 | 0.06 | 0.04 | 0.07 | 0.06 | 0.10 |
| C57BL/6 female BMM (2) | | | | | 1.00 | 0.99 | 0.07 | 0.07 | 0.06 | 0.08 | 0.07 | 0.11 |
| C57BL/6 female BMM (3) | | | | | | 1.00 | 0.08 | 0.08 | 0.06 | 0.08 | 0.07 | 0.12 |
| BALB/c male BMM (1) | | | | | | | 1.00 | 0.99 | 0.97 | 0.94 | 0.97 | 0.97 |
| BALB/c male BMM (2) | | | | | | | | 1.00 | 0.96 | 0.94 | 0.97 | 0.96 |
| BALB/c male BMM (3) | | | | | | | | | 1.00 | 0.98 | 0.98 | 0.94 |
| BALB/c female BMM (1) | | | | | | | | | | 1.00 | 0.98 | 0.91 |
| BALB/c female BMM (2) | | | | | | | | | | | 1.00 | 0.95 |
| BALB/c female BMM (3) | | | | | | | | | | | | 1.00 |

**Figure 5.18 Correlation matrix of MALDI-TOF MS detected methylation ratios in BMM.** The complete dataset was reduced to all autosomal CpG units that were measured in all BMM samples of three males and females of each strain. Correlation was calculated using the 'coefficient of determination' function in Excel 2007 and colored according to a three-color-scale ranging from the lowest value (red) to the highest value (dark blue).

## 5.2.9 Analysis of the influence of the BMM culture system

To exclude the possibility that the observed differences in methylation are due to an influence of the artificial *in vitro* BMM culture system, spleen samples of each mouse were additionally prepared and analyzed. Concordantly, the genomic DNA derived from the spleen samples from three individuals per strain and gender was evaluated. In the following, the methylation data for the reduced set of 24 amplicons (female samples were only measured in the reduced set) were used to calculate correlation matrixes. Of note, the correlation analysis involved only autosomal CpG units, which were measured in all samples. This correlation showed, that the three biological replicates of each cell or tissue type were highly similar (see **Figure 5.18** for BMM; data not shown for spleen) for the analyzed regions. Moreover, as shown in **Figure 5.19** both spleen and BMM also correlated well within the same strain (median $r^2$>0.93), suggesting that the observed patterns are not a cell culture-induced artifact.

| | C57BL/6 male BMM | C57BL/6 female BMM | C57BL/6 male spleen | C57BL/6 female spleen | BALB/c male BMM | BALB/c female BMM | BALB/c male spleen | BALB/c female spleen |
|---|---|---|---|---|---|---|---|---|
| C57BL/6 male BMM | 1.00 | 0.99 | 0.96 | 0.95 | 0.05 | 0.07 | 0.06 | 0.08 |
| C57BL/6 female BMM | | 1.00 | 0.94 | 0.95 | 0.07 | 0.09 | 0.08 | 0.10 |
| C57BL/6 male spleen | | | 1.00 | 0.97 | 0.02 | 0.04 | 0.03 | 0.04 |
| C57BL/6 female spleen | | | | 1.00 | 0.02 | 0.03 | 0.03 | 0.04 |
| BALB/c male BMM | | | | | 1.00 | 0.99 | 0.98 | 0.97 |
| BALB/c female BMM | | | | | | 1.00 | 0.98 | 0.98 |
| BALB/c male spleen | | | | | | | 1.00 | 0.99 |
| BALB/c female spleen | | | | | | | | 1.00 |

**Figure 5.19 Correlation matrix of MALDI-TOF MS detected methylated ratios in BMM and spleen.** The complete dataset was reduced to all autosomal CpG units that were measured in all samples. Individual male or female methylation data was averaged (n=3 in each case) and correlation was calculated using Excel 2007 and colored according to a three-color-scale ranging from the lowest value (red) to the highest value (dark blue).

### 5.2.10 Comparison of somatic cells with germ line cells

It is still unclear how epigenetic states are inherited to the next generation. The inheritance of 5-methylcytosine marks itself could provide a conceivable possibility. To investigate this it was important to access the level of methylation in reproductive cells as well, because a prerequisite for this option is, that reproductive cells show similar differential patterns like non-reproductive cells. Accordingly, I assessed how methylation patterns differed between somatic and germ line cells. Germ line cells were prepared from testis and analyzed by mass spectrometry using the full initial amplicon set. Assuming that testis mainly contains germ line cells or cells committed to the germ line lineage, the amount of germ line cells within the testis preparations was determined. Therefore, a single-gene MCIp was performed followed a by real-time PCR with two independent genomic DNA preparations for each tissue (BMM, spleen and testis) and strain. But in this case the MCIp process differed in that the genomic DNA was digested with MseI restriction endonuclease instead of shearing by ultrasonication. This resulted in a more stable DNA fragment length, which generally leads to a more distinct distribution pattern during column fractionation. As judged by the predominantly unmethylated status (95-85%) of the maternally imprinted *Snrpn* alleles in these samples (see **Figure 5.20**), the testis samples were prepared to mainly contain germ line cells (70-90%).

When comparing germ line with somatic tissues, it appeared that the majority (16/18) of the validated DMR only acquired differential DNA methylation in somatic cells (BMM, spleen) but showed similar methylation states in the testis (containing mostly germ line cells) of both strains. **Table 5.3** (page 100) shows that six of the analyzed regions were methylated in testis samples of both strains and nine regions were unmethylated. Only two DMRs (*Pdgfrb* and *Sfi1*) also appeared in germ line. The methylation status for *Pop4* in BALB/c was not measurable, since the DMRs resulted from a methylated insertion in the other strain. Thus, the differential patterns are more likely established after fertilization and DNA methylation itself is unlikely to be the inherited mark.

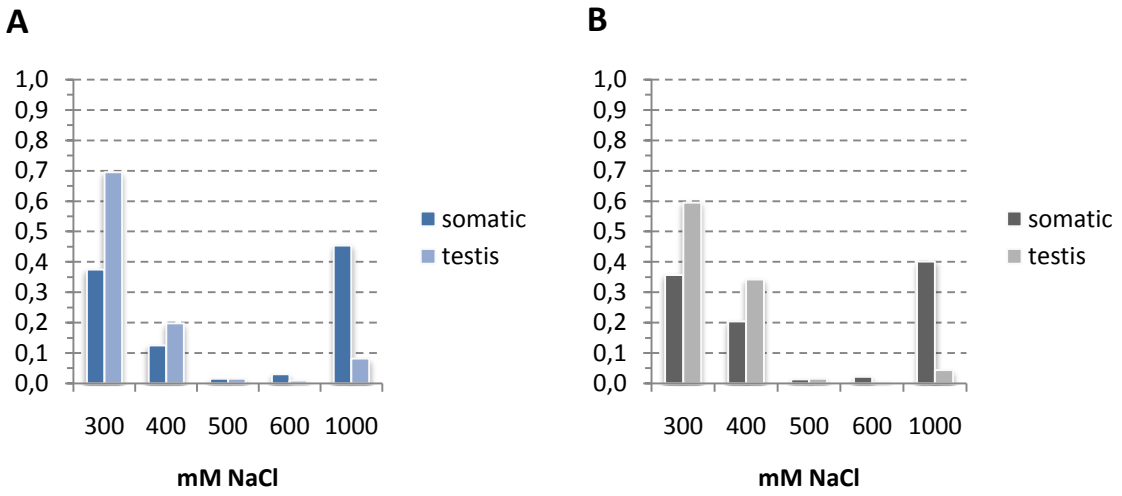



**Figure 5.20 Real-time PCR results for a maternally imprinted gene.** MCIp elution profiles are schematically represented as bar charts. During the small scale MCIp procedure increasing concentrations of NaCl were used to elute the DNA into different fractions, which were analyzed by real-time PCR for the distribution of both alleles of the maternal imprinted gene *Snrpn*. The results are shown relative to the amount of input-DNA (1.0; 100%). Normalized results of somatic tissues (BMM and spleen) were averaged for both parental strains shown separately in (A) BALB/c and (B) C57BL/6. In germ line cells (testis) mainly the unmethylated allele was found, whereas in somatic tissues both the methylated and unmethylated versions of the allele were present.

## 5.2.11 Inheritance of DNA methylation patterns in F1-hybrid animals

Since germ line cells lack the differential methylation pattern seen in somatic cells, the question arose, how these methylation differences are inherited to F1 hybrids. In order to investigate this question, the degree of CpG methylation at selected loci was compared between of F1 hybrids and their wild-type parents. I produced F1 hybrid offspring in both possible mating combinations (C57BL/6 sires or BALB/c sires). For each mating combination at least two litters, containing two individuals of both genders minimum, were produced, grown to the age of 6 to 8 weeks and independently analyzed. Genomic DNA from bone marrow-derived macrophages (BMM), spleen and germ line (testis) was prepared as described for both parental strains and the methylation status was measured by MALDI-TOF mass spectrometry using the reduced amplicon set (24 amplicons).

### 5.2.11.1 Methylation pattern of DMRs in F1-hybrid mice

As mentioned above most primer pairs used to amplify bisulfite-treated DNA were not specific for one of the parental strains. However some primer pairs were indeed specifically designed for one parental strain (for example, primer pairs for 3110007F17Rik

and *Zfp568*). In addition, some amplicons contained single nucleotide polymorphisms (SNP) located between the primer binding sites within the remaining amplicon sequence. These SNP's had no impact on PCR reactions, but sometimes resulted in altered fragment mixtures after fragmentation of cRNA during the MassARRAY application. Therefore I analyzed F1-hybrid animals with both possible genomic DNA sequences, derived from C57BL/6 and BALB/c parental mice, and averaged the methylation data at the level single CpG units, in order to eliminate a possible influence of fragment composition on the methylation data. If SNPs resulted in additional CpG sites, the methylation data for this CpG site was obtained only from the analysis with the underlying parental sequence. Three examples of DMR, corresponding to two genes (*Slc27a6* and *Zfp568*) are shown in **Figure 5.21**.

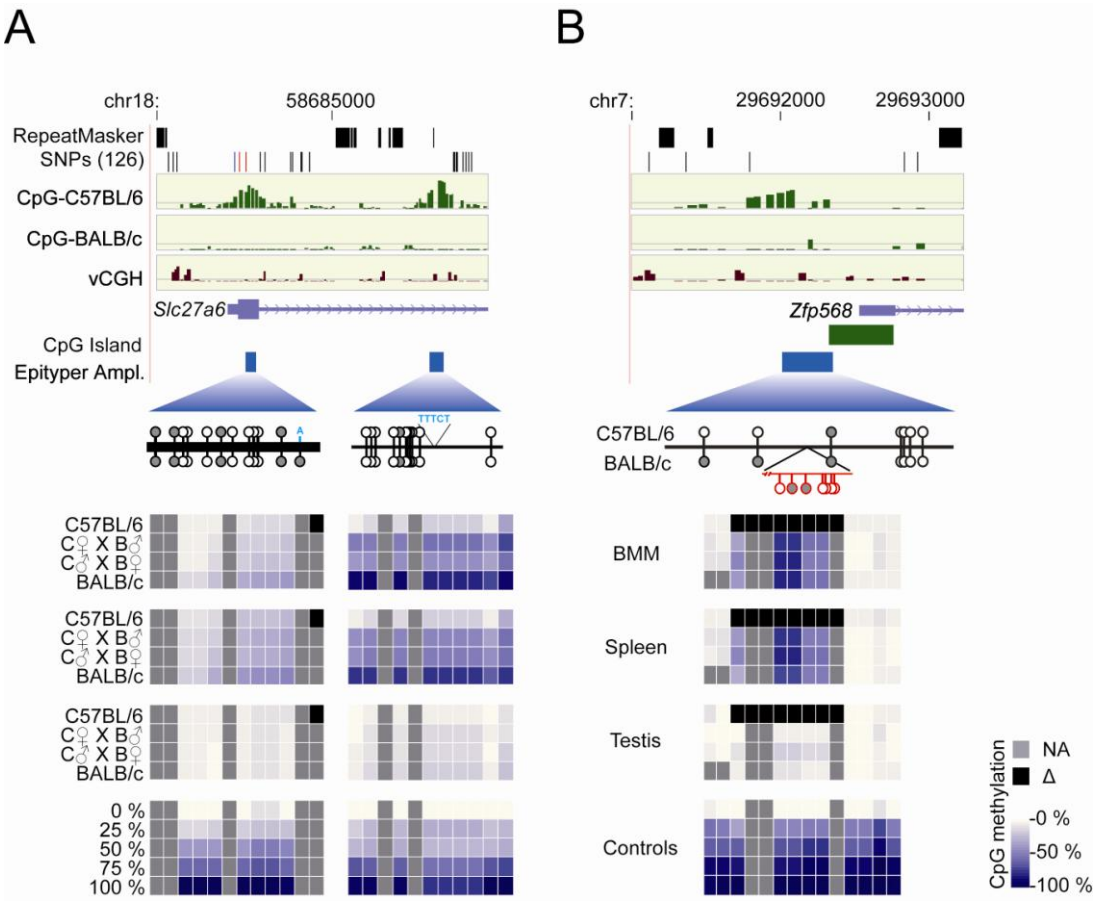



**Figure 5.21 Inheritance of strain-specific methylation patterns in F1 hybrids.** This figure illustrates two examples for allelic inheritance of strain-specific methylation patterns in (A) *Slc27a6* and (B) *Zfp568*. In the top panels, hypomethylation scores for BMM of both strains are displayed as described in **Figure 5.17**. Averaged methylation levels of individual CpGs were determined by MALDI-TOF analysis at the indicated DMRs in BMM, spleen, and testis and are shown color coded for parental strains and F1 hybrids. Data are mean values of 2-4 individual samples.

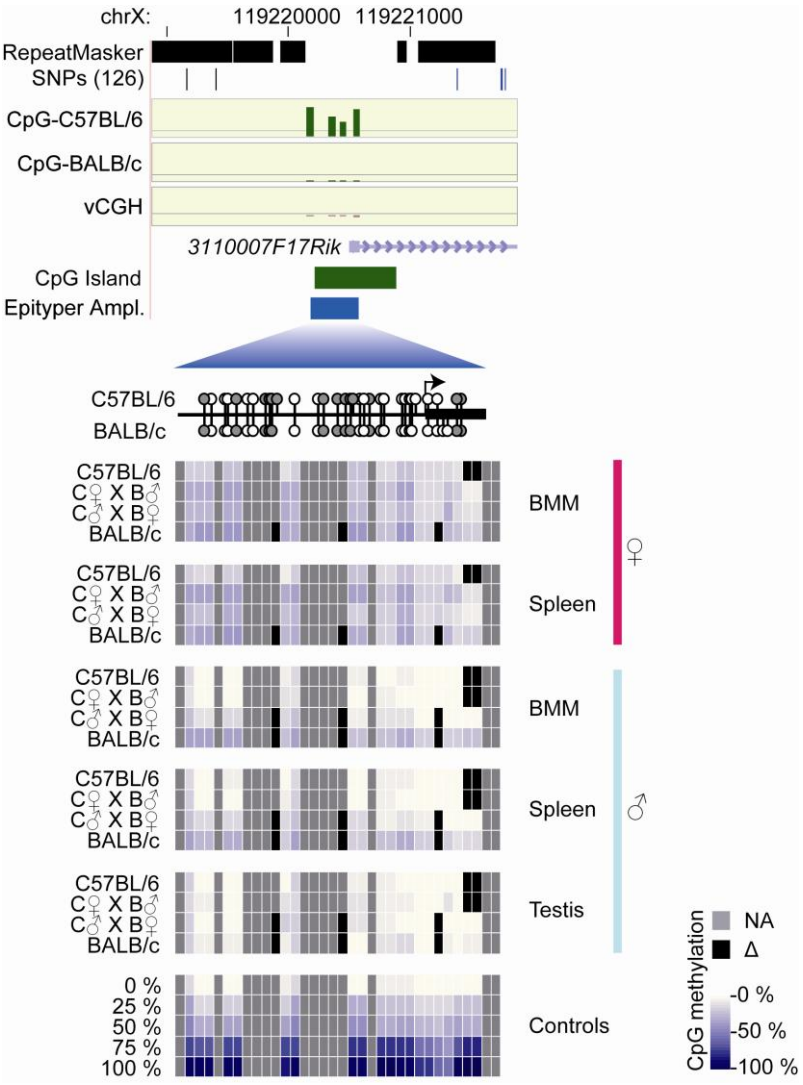



**Figure 5.22 Inheritance of a X-linked DMR in F1-hybrid mice.** Example of allelic inheritance of strain-specific methylation patterns in the *3110007F17Rik* gene on the X-chromosome. In the top panels, hypomethylation scores for BMM of both strains are displayed as described in **Figure 5.17**. Gender-specific averaged methylation levels of individual CpGs were determined by MALDI-TOF analysis at the indicated DMRs in BMM, spleen, and testis and are shown color coded for parental strains and F1 hybrids. Data are mean values of 2-4 individual samples. Methylation in this region is only seen in male BALB/c animals and male F1-hybrids derived from a BALB/c mother.

The first example (*Slc27a6*) given in **Figure 5.21A** exhibited two DMRs where the methylation status in F1-hybrid mice ranged between the levels seen in the parental strains. The differentially methylated region in the promoter of *Zfp568* (**Figure 5.21B**) resided within a BALB/c-specific insertion. F1-hybrid mice, although harboring both parental alleles, contained only methylated CpG's within the parental BALB/c-specific insertion and not on the C57BL/6-derived allele.

Another DMR example is located at the 3110007F17Rik gene, which is located at the X-chromosome. In parental strains this region is methylated in BALB/c and unmethylated in C57BL/6 males. **Figure 5.22** illustrates, that methylation of the X-chromosomal gene in F1-hybrid males is only seen if the corresponding mother is of BALB/c origin (transmitting the methylated allele). However this phenomenon was only evident in males. In females the methylation was more similar between all F1 individuals analyzed, which is likely due to the presence of the second X-chromosome. Random X-inactivation, occurring in female cells, leads to an altered DNA methylation creating a different methylation pattern as compared to males.

#### 5.2.11.2 Impact of mating combination or offspring gender

The next question to be addressed was how the various F1-hybrid animals differed, first in terms of mating combination (C57BL/6 or BALB/c sire) and in gender. Initially, I compared the DNA methylation profiles of the different F1 individuals within a mating combination and gender and found that all F1 samples were methylated to a highly similar extent as measured for single CpG units (complete MassARRAY data is not shown due to space limitations but will be available online in the supplementary part of the corresponding publication). However, an exception was observed at the *Isoc2b* promoter were the methylation levels in F1 hybrids were considerably more variable (ranging from 15-45% in individual F1 mice). In a next step, the methylation levels for each gender within a mating combination were averaged at single CpG units and correlation ratios were calculated as described above (section 5.2.8). The corresponding complete correlation matrix is given in **Figure 5.23**. The matrix shows that averaged methylation levels in F1 animals were highly similar between samples of the somatic tissues with only a minor, if any, difference between mating combination (median $r^2$>0.95 for BMM and median $r^2$>0.96 for spleen). The same was apparent when comparing the two possible offspring gender within or between mating combinations. Hence, the methylation levels for individual CpG units of individual F1-hybrid mice were averaged for following comparisons to parental strains.

| | C57BL/6 BMM male (n=3) | C57BL/6 BMM female (n=3) | F1 (Cmother) BMM male (n=1) | F1 (Bmother) BMM male (n=3) | F1 (Cmother) BMM female (n=2) | F1 (Bmother) BMM female (n=2) | BALB/c BMM male (n=3) | BALB/c BMM female (n=3) | C57BL/6 spleen male (n=3) | C57BL/6 spleen female (n=3) | F1 (Cmother) spleen male (n=2) | F1 (Bmother) spleen male (n=3) | F1 (Cmother) spleen female (n=2) | F1 (Bmother) spleen female (n=2) | BALB/c spleen male (n=3) | BALB/c spleen female (n=3) | C57BL/6 testis male (n=3) | F1 (Cmother) testis male (n=2) | F1 (Bmother) testis male (n=2) | BALB/c testis male (n=3) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| C57BL/6 BMM male | 1.00 | 0.99 | 0.66 | 0.71 | 0.60 | 0.63 | 0.00 | 0.01 | 0.95 | 0.94 | 0.65 | 0.60 | 0.55 | 0.63 | 0.01 | 0.01 | 0.43 | 0.37 | 0.34 | 0.05 |
| C57BL/6 BMM female | | 1.00 | 0.60 | 0.66 | 0.56 | 0.58 | 0.01 | 0.01 | 0.94 | 0.95 | 0.60 | 0.55 | 0.51 | 0.59 | 0.01 | 0.02 | 0.41 | 0.34 | 0.31 | 0.03 |
| F1 (Cmother) BMM male | | | 1.00 | 0.97 | 0.96 | 0.97 | 0.15 | 0.13 | 0.67 | 0.66 | 0.95 | 0.94 | 0.93 | 0.95 | 0.13 | 0.12 | 0.32 | 0.43 | 0.39 | 0.23 |
| F1 (Bmother) BMM male | | | | 1.00 | 0.95 | 0.98 | 0.16 | 0.15 | 0.74 | 0.72 | 0.97 | 0.95 | 0.92 | 0.96 | 0.14 | 0.13 | 0.42 | 0.51 | 0.48 | 0.30 |
| F1 (Cmother) BMM female | | | | | 1.00 | 0.97 | 0.23 | 0.22 | 0.64 | 0.63 | 0.95 | 0.93 | 0.97 | 0.94 | 0.22 | 0.20 | 0.35 | 0.44 | 0.42 | 0.28 |
| F1 (Bmother) BMM female | | | | | | 1.00 | 0.21 | 0.20 | 0.65 | 0.65 | 0.95 | 0.93 | 0.94 | 0.95 | 0.19 | 0.18 | 0.35 | 0.46 | 0.43 | 0.30 |
| BALB/c BMM male | | | | | | | 1.00 | 0.99 | 0.00 | 0.00 | 0.13 | 0.17 | 0.24 | 0.16 | 0.98 | 0.97 | 0.00 | 0.01 | 0.01 | 0.24 |
| BALB/c BMM female | | | | | | | | 1.00 | 0.00 | 0.00 | 0.12 | 0.16 | 0.22 | 0.16 | 0.98 | 0.98 | 0.00 | 0.01 | 0.01 | 0.27 |
| C57BL/6 spleen male | | | | | | | | | 1.00 | 0.97 | 0.72 | 0.68 | 0.63 | 0.70 | 0.00 | 0.01 | 0.53 | 0.46 | 0.43 | 0.08 |
| C57BL/6 spleen female | | | | | | | | | | 1.00 | 0.70 | 0.66 | 0.61 | 0.68 | 0.00 | 0.01 | 0.52 | 0.48 | 0.44 | 0.09 |
| F1 (Cmother) spleen male | | | | | | | | | | | 1.00 | 0.96 | 0.96 | 0.98 | 0.12 | 0.11 | 0.43 | 0.54 | 0.51 | 0.32 |
| F1 (Bmother) spleen male | | | | | | | | | | | | 1.00 | 0.94 | 0.97 | 0.17 | 0.16 | 0.43 | 0.55 | 0.51 | 0.33 |
| F1 (Cmother) spleen female | | | | | | | | | | | | | 1.00 | 0.94 | 0.23 | 0.21 | 0.35 | 0.46 | 0.44 | 0.29 |
| F1 (Bmother) spleen female | | | | | | | | | | | | | | 1.00 | 0.16 | 0.15 | 0.44 | 0.54 | 0.50 | 0.33 |
| BALB/c spleen male | | | | | | | | | | | | | | | 1.00 | 0.98 | 0.00 | 0.01 | 0.01 | 0.26 |
| BALB/c spleen female | | | | | | | | | | | | | | | | 1.00 | 0.00 | 0.01 | 0.01 | 0.28 |
| C57BL/6 testis male | | | | | | | | | | | | | | | | | 1.00 | 0.88 | 0.87 | 0.39 |
| F1 (Cmother) testis male | | | | | | | | | | | | | | | | | | 1.00 | 0.99 | 0.61 |
| F1 (Bmother) testis male | | | | | | | | | | | | | | | | | | | 1.00 | 0.62 |
| BALB/c testis male | | | | | | | | | | | | | | | | | | | | 1.00 |

**Figure 5.23 Correlation matrix of MALDI-TOF MS detected methylation ratios in parental strains and F1-hybrid mice.** The complete dataset was reduced to all autosomal CpG units that were measured in all samples. Individual male or female methylation data was averaged and the number of individuals used is indicated in brackets (top lane). Correlation was calculated using Excel 2007 and colored according to a three-color-scale ranging from the lowest value (red) to the highest value (dark blue).

### 5.2.11.3 Comparison of mean methylation levels between parents and offspring

Most DMRs exhibited a methylation pattern in F1 mice as seen at the *Slc27a6* gene (see **Figure 5.21A**), where offspring methylation levels were not comparable to one of the parental strains, but ranged in between the parental levels indicating an intermediate methylation status. To explore if this observation is a common phenomenon, the averaged methylation levels of parental BMM and spleen data were compared by plotting against averaged CpG methylation ratios of F1 hybrids derived from both mating combinations (C57BL/6 or BALB/c sires). In eight out of eleven DMRs analyzed by MALDI-TOF (*1600021P15Rik*, *Asb4*, *Coro2a*, *Pdgfrb*, endogenous *Sfi*, *Slc13a3*, *Slc27a6*, and *Zfp568*), methylation patterns in F1 hybrids were almost identical to the average methylation levels

seen in parental strains ($r^2$>0.97) (see **Figure 5.24**). The remaining three DMRs (*Sfi1* pseudogene, *Isoc2b*, and *Eps8l1*) were slightly different and either acquired (*Sfi1* pseudogene) or lost methylation (*Isoc2b*, *Eps8l1*) in F1 hybrids relative to their parental strains. However, it was obvious that at the investigated loci neither one of the parental methylation levels was completely reconstituted in F1 animals, which would be the case if one parental methylation pattern is inherited dominantly. The simplest and most plausible explanation for this pattern is that the methylation level of each allele is inherited independently. During the MassARRAY application both parental alleles were not analyzed separately, since primer pairs did not discriminate between parental alleles. Hence, the methylation status of parental alleles in F1-mice was averaged. Given this situation I wondered if both parental alleles exhibited the same level of methylation in F1-hybrid animals or if, as a second possible alternative, individual methylation status of parental alleles was maintained.

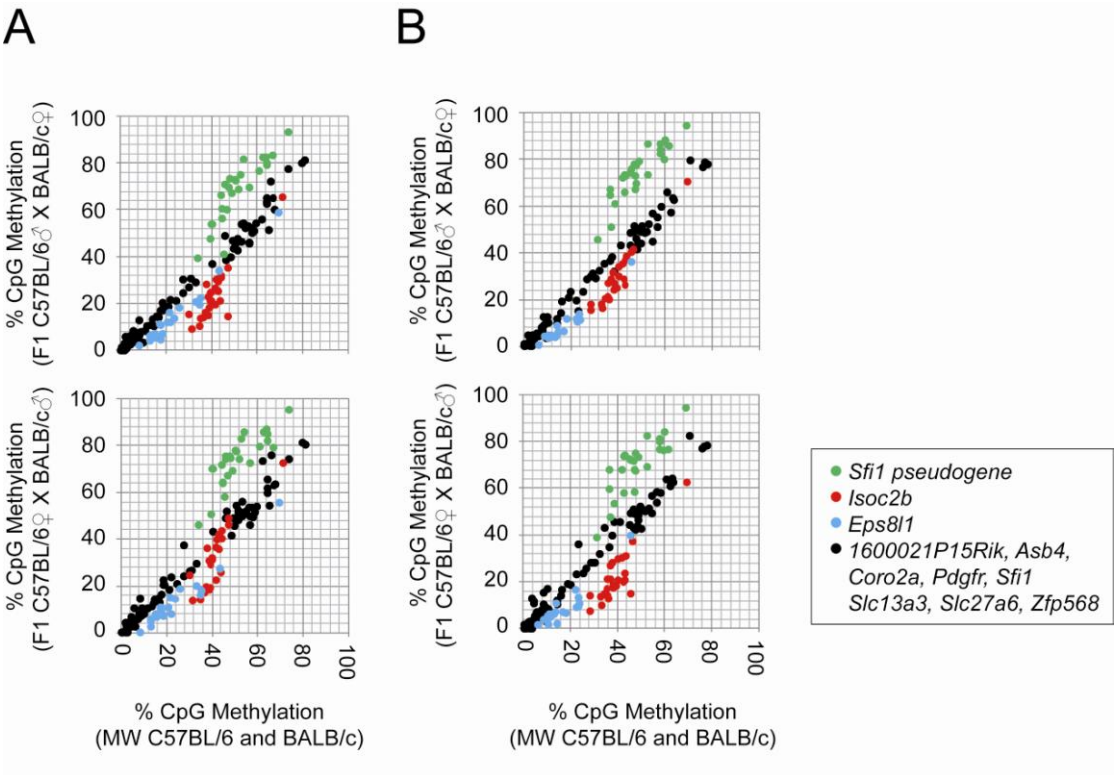



**Figure 5.24 Methylation ratios of wildtype mice versus F1 hybrids.** Averaged CpG methylation ratios of parental (A) BMM or (B) spleen (n=3 for each strain and tissue) are plotted against averaged CpG methylation ratios of F1 hybrids derived from C57BL/6 (top, n=5) or BALB/c (bottom, n=4) sires. In eight out of eleven DMR analyzed by MALDI-TOF, methylation patterns in F1 hybrids are almost identical to the average methylation level in parental strains ($r^2$>0.97; marked in black). Three DMR (*Sfi1* pseudogene, *Isoc2b*, and *Eps8l1*, marked in red, green and blue, respectively) either acquired (*Sfi1* pseudogene) or lost methylation (*Isoc2b*, *Eps8l1*) in F1 hybrids relative to parental strains.

#### 5.2.11.4 Allele-specific bisulfite sequencing of DMRs

To address if parental alleles maintained their own methylation pattern in F1-hybird animals I performed traditional bisulfite sequencing, since the MS analysis did not discriminate between the two different parental alleles. I chose DMR regions where a characteristic sequence difference remained after bisulfite-treatment of genomic DNA and analyzed the tissue where the greatest methylation difference between parental strains was observed. In total, six regions were analyzed; four containing single nucleotide polymorphisms (*Coro2a*, *Isoc2b, 1600021P15Rik* and *Asb4*) and two regions harboring small insertions in parental C57BL/6 (*Slc27a6*) or BALB/c (*Pdgfrb*) sequences. Besides two male F1-hybrid animals (one from each mating combination) I also included two male individuals from each parental strain. Genomic DNA, derived from BMMs (or testis in case of *Pdgfrb*) was treated with sodium-bisulfite, similar to the MassARRAY experiments (see methods section 4.5.3, page 57). PCR products of these five regions were generated using the same primer pairs, specific to bisulfite-treated DNA as well as the same PCR conditions which have been used in the MassARRAY experiments. PCR-products were then cloned into the pCR®2.1TOPO-vector and transfected in *E. coli*. Insert-containing plasmids, derived from single colonies were further purified and sequenced. Sequencing results for five amplicons are shown in **Figure 5.25**. The figure illustrates that in the parental strains the genes *Coro2a, Isoc2b, Asb4, 1600021P15Rik* and *Slc27a6* were methylated in BALB/c, whereas *Pdgfrb* was methylated in C57BL/6. This observation was in line with the microarray and MassARRAY results. In F1-hybrid animals the methylation pattern within one animal was mixed, with some sequences being methylated and some unmethylated. Both F1-mice used showed the same methylation distribution pattern between individually cloned sequences and were, therefore, combined to one pool. Furthermore, since I restricted this analysis to amplicons, which can be traced back to their strain origin, I was able to determine the strain parentage of a specific amplicon within the F1-mice. Hence, alleles of *Coro2a, Pdgfrb, Asb4, 1600021P15Rik* and *Slc27a6* that derived from C57BL/6 exhibited a very similar methylation pattern as seen in the parental strain. The same was observed for alleles, which can be traced back to the BALB/c origin. However, the *Isoc2b* gene uniquely showed a different pattern, where both alleles within F1-hybrid animals contained unmethylated as well as methylated sequences. This pattern was obvious in both F1-mice examined in this experiment.

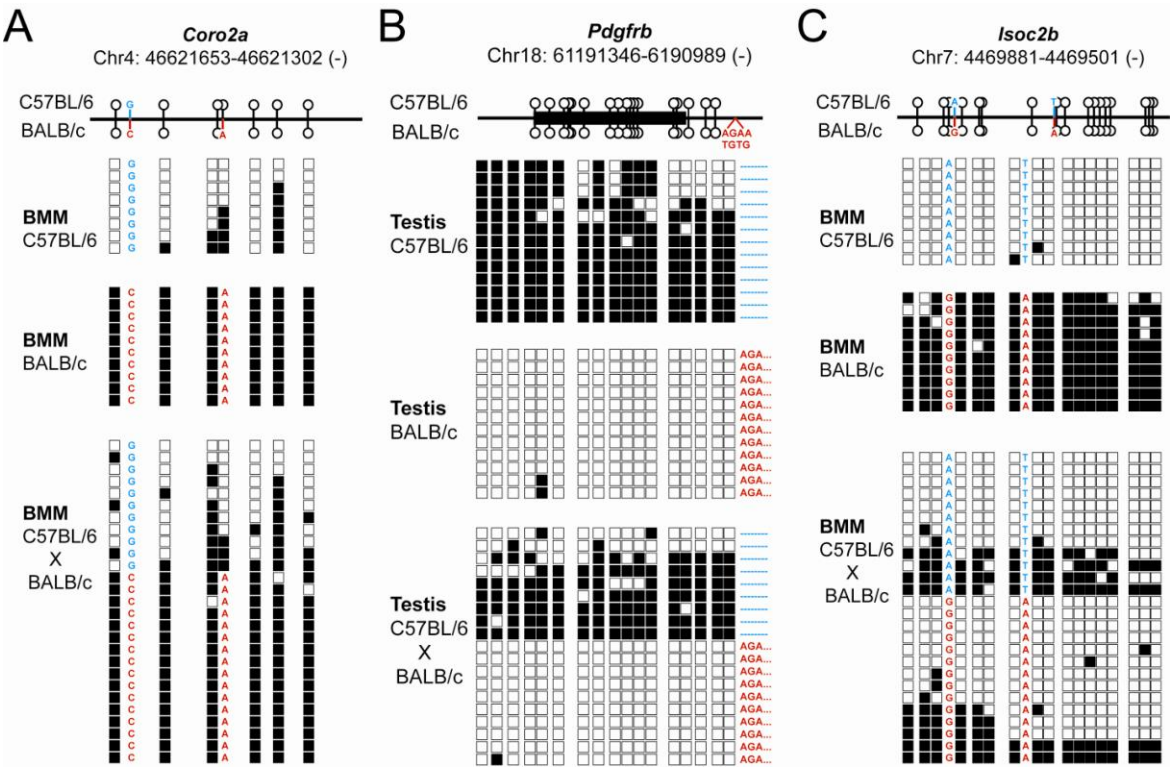



**Figure 5.25 Strain-specific methylation patterns are mainly controlled in *cis*.** Allele-specific bisulfite sequencing of DMRs in *Coro2a* (A, controlled in *cis*), *Pdgfrb* (B, controlled in *cis*), *Isoc2b* (C, controlled in *trans*), *1600021P15Rik* (D, controlled in *cis*), *Asb4* (E, controlled in *cis*) and *Slc27a6* (F, controlled in *cis*) (D-F are shown on the next page). The genomic position of CpGs within the amplicons is shown at the top. Sequence variations used to distinguish the different parental alleles are marked in blue for C57BL/6 and in red for BALB/c. Individual CpGs are represented by either white (unmethylated) or black (methylated) squares. Lines of squares represent independently sequenced clones derived from two independent sample preparations, derived from reciprocal crosses.

Taken together, in 5 (out of 6) regions the allelic methylation patterns in F1 hybrids were established essentially as observed on the parental alleles, suggesting that DNA-methylation at these sites is largely controlled in *cis* by the local genetic sequence. Amongst the studied regions, only at the *Isoc2b* promoter, a mixed DNA methylation pattern was observed with some sequences being unmethylated and others methylated independent of their origin. Therefore the differential pattern must be controlled in *trans*, for example through a yet unknown strain-specific epigenetic modifier.

In summary it became obvious, that allele-specific DNA methylation in mouse strains is mainly determined by local *cis*-acting sequences.

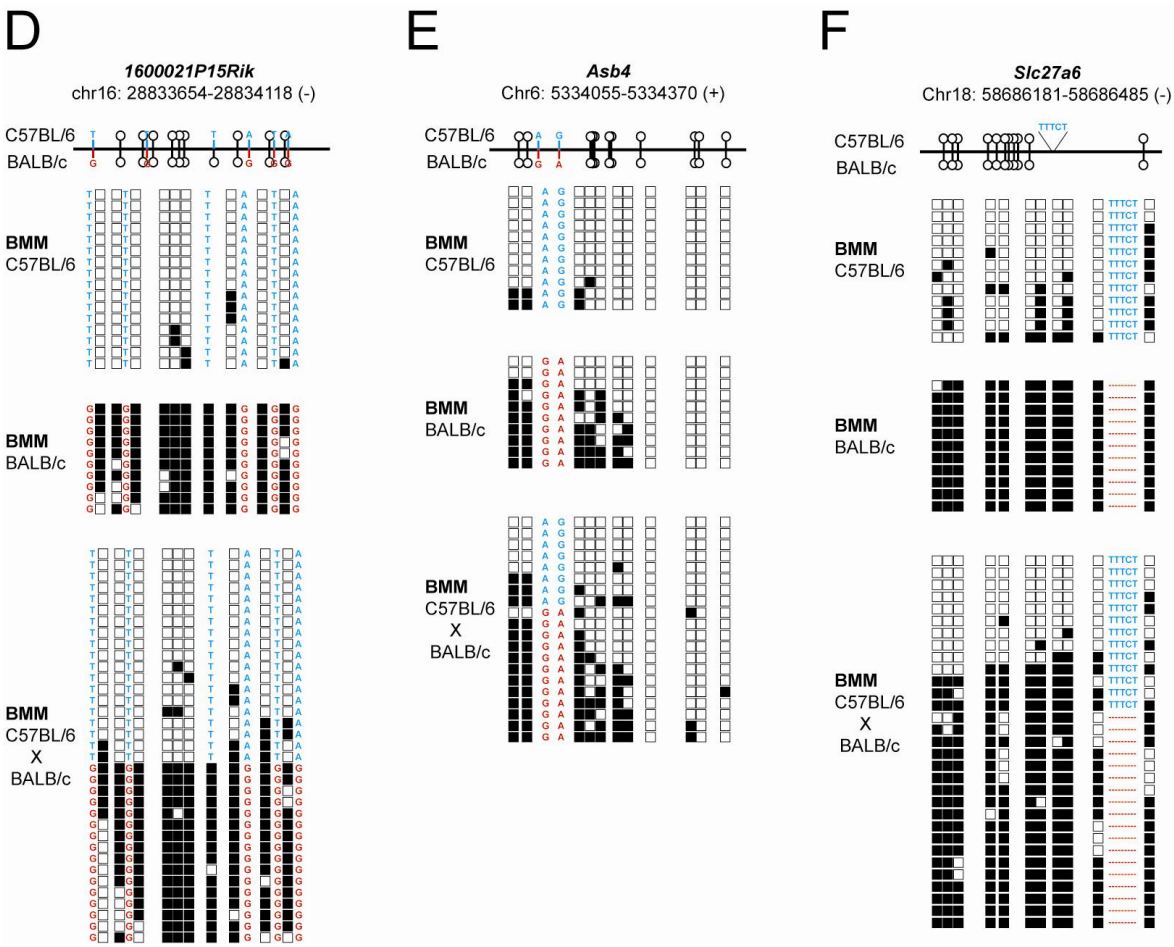



**Figure 5.25** continued from previous page.

# 6 Discussion

Epigenetic phenomena in general and DNA methylation in particular play a major role in the regulation of gene expression. The latter mechanism contributes to both the shaping of cellular phenotypes within an organism and the establishing of interindividual differences between organisms. In addition, differences in DNA methylation can not only be observed in transformed cells or other non-cancer diseases, but also at very different levels in various tissue types and developmental stages. In this work, I have established methods to study the contribution of differential DNA methylation to cell type-specific gene expression as well as the occurrence of differentially methylated regions between individuals (inbred mice strains).

## 6.1 Mapping DNA methylation

In the past years epigenetic modifications such as DNA methylation have been intensively studied. Several methodologies have been developed to map DNA methylation, each of which having its own advantages and disadvantages as highlighted in the introduction. As part of this thesis, methyl-CpG immunoprecipitation (MCIp)[110] was modified and utilized to map DNA methylation both on single-gene level and on a global scale by combination with quantitative real-time PCR or microarrays, respectively. MCIp is based on the recombinant antibody-like MBD2–Fc fusion protein that was originally designed by our lab to globally detect disease-related hypermethylation in CpG islands on a global level[110]. A crucial step for accurate DNA methylation mapping is the sample pretreatment. Sanger sequencing of bisulfite treated genomic DNA, which is the standard method, provides the best resolution, because it maps methylated cytosines at single base-pair resolution. Large-scale bisulfite-DNA sequencing has been successfully initiated[153,154], but is very laborious and resource-intensive. Alternative, less direct methods have been developed to map DNA methylation on a global scale. Approaches based on methylation-sensitive restriction enzymes enrich fragment due to digestion of methylated[118,119] or unmethylated DNA[116,155] followed by size fractionation. However, the restriction enzyme-based enrichment methods lack sensitivity and do not cover all CpGs due to their dependence of restriction sites[7]. Other methods have been developed which use protein affinity to enrich methylated or unmethylated sequences. Methylated DNA fragments are affinity purified with either an anti-m5C antibody (methylated DNA immunoprecipitation; MeDIP) or by

using MBD affinity purification (MAP, using the MBD domain for binding of methylated DNA)[107,123]. Both methods gave comparable results[120], but require a relatively high density of DNA methylation such as methylated CGIs[7]. A method to enrich unmethylated DNA using CxxC affinity purification (CAP, with x being any residue) was also recently published[126] but is also insensitive to regions with low CpG content[7]. In contrast, MCIp can divide the bulk of genomic DNA fragments into separate fractions of increasing methylation density. Therefore, during the MCIp procedure, not only the methylated DNA can be enriched, but also unmethylated DNA is recovered without sample loss. This allows the simultaneous analysis of the whole range of DNA methylation density, including both hyper- and hypomethylated DNA.

Samples enriched by either of these techniques can be globally analyzed using DNA microarrays or by direct large-scale sequencing techniques. The latter, such as 454 sequencing (Roche) or Solexa bisulfite sequencing (Illumina), are relatively fast and cheap and have been used in a number of publications[156-158]. However, these readout techniques are still resource intensive and analysis cannot be restricted to distinct regions. Microarrays are cheaper and more flexible, since one can either use custom-designed or pre-designed commercially available tiling arrays, depending on the genomic regions of interest. However, compared to large-scale sequencing approaches, microarrays receive data with only moderate resolution.

During this thesis, the MCIp-on-chip application (MCIp combined with microarray readout) originally designed by our lab to globally detect disease-related hypermethylation in CpG islands[110] was further developed to a novel application. This work was designed as a pilot study to demonstrate the applicability of this method for mapping of cell type-specific hypomethylation. To overcome resolution restrictions and to validate the microarray analysis, the methylation levels were additionally quantified for some regions using single-gene MCIp and the standard bisulfite sequencing approach as a MCIp-independent approach. In summary, our data showed a good correlation between tissue-specific gene expression and promoter hypomethylation. Hence, the MCIp-on-chip was successfully applied to a novel application and seems to be a reliable method to detect cell type-specific hypomethylation.

Furthermore, by using MCIp I successfully fractionated genomic DNA into methylated and unmethylated pools, which has recently been performed by our group in another study as well[159]. In their study, Schmidl and colleagues were able to identify and functionally characterize differentially methylated regions between two different human T-cell populations. In the present study the performance of this “mirror image” approach was improved by using more stringent hybridization conditions to identify differentially methylated regions (DMRs) in specific, selected genomic intervals. Since the "mirror

image" data correlated well with the MassARRAY results, this elaborate approach seems to be a reliable method for locus-wide comparative methylation analysis. With the pre-selection of differentially expressed genes and the downstream high-throughput analysis of selected CpGs by mass spectrometry, this work systematically identified DMRs and yielded new insights in individual (allele-specific) DNA methylation. Moreover, the vCGH analysis allowed the identification of genomic alternations, such as copy number variations (CNV) or deletions. Some of the larger CNV's, which have been validated in this thesis by qRT-PCR, were previously identified using standard comparative genomic hybridization approaches (CGH)[150]. Standard CGH approaches utilize genomic DNA without pre-treatment steps (like MCIp), to screen for genetic changes (gains or losses) on a regular basis[160,161]. The fact that these copy number variations between the two mice strains were correctly identified strongly suggests that the MCIp technique fractionated genomic DNA without sample loss like in MAP or MeDIP.

Taken together, the MCIp technique seems to be a reliable methodology to map differential DNA methylation at the level of both hypo- and hypermethylation.

## 6.2 Cell type-specific promoter methylation

Entire organisms incorporate a huge variety of differentiated and specialized cells. Every cell of a multicellular organism contains essentially the same genome but only a small proportion of all available genes are actively transcribed, whereas all remaining genes are silent. It is widely accepted that methylation of the promoter region is related to gene silencing. However, experimental evidence on this topic tends to be unclear. Initially, theoretical models were proposed by Riggs or Holliday and Pugh, based on DNA methylation to explain X chromosome inactivation[162] and other events in cell differentiation during development[163]. Although those early studies proposed that DNA methylation could have a role in regulating development, following studies of DNA methylation in known tissue-specific genes during development did not support a major role for DNA methylation. Studies by Warnecke and Clark[36] found that the tissue-specific expression of the skeletal α-actin gene in the adult mouse did not correlate with the methylation state of the promoter. Walsh and Bestor[35] investigated the methylation status of seven tissue-specific genes and found no correlation with tissue-specific expression. However, more recent publications systematically comparing differential methylation patterns on a genome-wide level came to different conclusions. De Bustos and colleagues for example received very different methylation profiles from diverse organs, suggesting the existence of tissue-specific epigenetic modification patterns across human chromosome 1[164]. The results of another study utilizing MeDIP to analyze 16 different human tissues, including

sperm, indicates that tissue specific DNA methylation, including CpG islands, is relatively common[165]. These studies are in line with several others by Song, associating tissue-specific differentially methylated regions with differential gene expression[166,167]. In addition, a study by Eckhardt identified cell type specific methylation using bisulfite sequencing across chromosomes 6, 20 and 22 across various tissues, which inversely correlated with transcription[153]. Similar results were also shown by a methylation sequencing study of 13 transcription factors suggesting that epigenetic remodelling of transcription factor genes is a frequent mechanism during hematopoietic development[168]. Recently Straussman and colleagues investigated the methylation of human CGIs in various tissues and found that some CpG-rich regions, which are constitutively unmethylated during early embryogenesis, gain tissue-specific methylation at later developmental steps[169].

Concordant to these publications, we found in our analysis a correlation between tissue-specific hypomethylation and tissue-specific gene expression. This was most obvious in monocytes, with 76% of hypomethylated genes belonging to the hematopoietic system. Among the different tissue types used in this study, the monocyte sample was the most homogenous cell population with purity of up to 90%. Tissues are generally composed of different, specialized cell types, likely resulting in mixed gene expression and methylation patterns. This was also demonstrated by bisulfite sequencing, where fully methylated sequences were found among, usually unmethylated, germ line-derived sequences (testis). These sequences are likely derived from contaminating somatic cells, since testis contains various levels of cells from spermatocytes to sperm as well as the surrounding mesenchymal cell population. The same is true for the brain sample, which is also derived from a more heterogeneous cell population than purified monocytes. It is therefore difficult, if not impossible, to correlate tissue-specific methylation or expression patterns for all genes, because both signatures may be determined by independent cell populations within the tissue. However, the amount of genes that showed a restricted expression pattern in testis, monocytes or brain within each group of hypomethylated genes, led to the result that the number of genes expressed in a certain tissue was significantly enriched in the hypomethylated group of the same tissue. This observation is in line with recent studies[164,166,167] and suggests that tissue-specific DNA hypomethylation correlates significantly with tissue-specific transcription.

It is currently believed that tissue-specific variations in DNA methylation largely occur in CpG-depleted promoters, whereas most CpG island promoters remain unmethylated in all tissues[137]. Recently it was shown, that CpG island promoters are preferentially unmethylated independent of their expression state and that low density CpG promoters

(LCP) are mainly methylated without precluding gene expression[122,170]. However, an exception of this general rule was recently provided by Shen and colleagues who identified a class of normally methylated CGI promoters in normal blood comprising 4% of known CGI promoters[171]. Moreover, Straussman and colleagues showed that CGI-like, constitutively unmethylated regions, which seem to be formed during early embryogenesis, undergo *de novo* methylation in a tissue-specific manner and have a dynamic role in development[169]. Our study contrasts the suggestion that especially LCP associated genes are not influenced in their expression by DNA methylation, because we showed that tissue-specific DNA hypomethylation correlates significantly with tissue-specific transcription, not only at CGI genes. Our results were recently supported by a study, which demonstrated that 'weak' CpG islands associated with specific developmentally regulated genes undergo aberrant hypermethylation in vitro[172]. It has also been reported that CpG-poor promoters are more prone to demethylation occurring in cancerous cells leading to the activation of oncogenes[155]. The greater flexibility in methylation states of regions with low CpG density may therefore contribute to the diversity of methylation states that underlies heterogeneity of cell types.

The present study also showed that hypomethylated promoters were associated with classical CpG islands in every cell type analyzed. However, strikingly more hypomethylated promoters were associated with CpG islands in testis than in somatic tissues. This observation is in line with the previously postulated general absence of tissue-specific variation in somatic CpG island methylation[137]. Notably, a study by Weber suggests that de novo methylation of CGIs in somatic cells preferentially occurs at germ line-specific genes as well as promoters with an intermediate CpG frequency[122]. The marked association of testis-specific hypomethylation with CpG island promoters in our analysis extends earlier studies reporting the somatic methylation of individual, so-called cancer-testis antigens, which are generally controlled by CpG-rich promoters[147,173-175]. These studies favor a model, in which DNA methylation is the primary silencing mechanism for a unique subset of germ line- and tumor-specific genes with a CpG-rich promoter. In line with this a recent study demonstrated tissue-specific *de novo* methylation of CpG-rich regions, which are normally unmethylated during early embryogenesis and gain methylation upon development[169].

Interestingly, when comparing the percentage of hypomethylated promoters that are associated with *Alu* repeats[138,139], we found a significant higher incidence of *Alu* repeats near hypomethylated promoters in testis as compared to the other two tissues. Oligonucleotide probes on the microarray do not directly correspond to repetitive

sequences. However, they may be located next to repeats and detect their methylation status. This observation is in line with previous studies indicating that *Alu* repeats in testis may be in general less methylated than those in somatic cells[52,176,177]. It was also reported that a significant number of all human genes may be controlled by *Alu*-associated CpG islands[178] and that *Alu* elements exhibit tissue-specific methylation patterns, which was recently shown for different brain regions[179]. However, little information on the positions of hypomethylated *Alu* repeats is available. Our results represent a further indication for a role of *Alu* repeats in testis-specific gene regulation. It was reported that the chromosomal density of *Alu* repeats is around three times higher on chromosome Y than on chromosome X and over two times higher than the average density for all human autosomes[180]. In line with this it was demonstrated, that treatment of lymphoid tumor cells with a specific anti-cancer drug (clofarabine) induces, as a side effect, both hypomethylation of *Alu* repeats and expression of cancer-testis antigens[181]. Recently it was shown by two publications that demethylated regions are associated with repetitive sequences such as LTR repeats by analyzing purified spermatogenic cells[182] or testis samples[167]. Moreover, it was published that some repetitive elements are activated and transcribed during spermatogenesis and therefore primarily passed through the paternal germ line[176,180]. During activation, these Y-linked *Alu* repeats have been shown to be hypomethylated due to the activity of a specific *Alu* repeat-binding protein (SABP) which protects element CpG dinucleotides from methyl-transferase enzyme activity[183]. This demethylation process may additionally affect nearby genes and activate their expression. For example, it was demonstrated in *Arabidopsis thaliana* that the expression of a flowering time gene correlates with the methylation status of a nearby SINE-like transposable element[184].

In the present study genes with cell type-specifically hypomethylated promoters were generally found on all chromosomes. However, an accumulation of testis-specific hypomethylation that has not been observed previously was detected on the Y chromosome, on which several of the annotated genes appeared hypomethylated compared to both somatic tissues and, as described earlier, on the X chromosome[146].
Besides the large number of testis-specific gene promoters that demonstrated hypermethylation in somatic tissues, a large group of testis-specific genes was also detected that did not exhibit a differential methylation status in any of the studied tissues. This suggests that only a subgroup of testis-specific genes is affected by somatic hypermethylation. In addition, we also noted that copies of Y-linked genes like *DAZL* (chromosome 3) or *VCX* (chromosome X) that share almost identical promoter regions with their gene homologues on chromosome Y (*DAZ* and *VCY*, respectively) also show

somatic hypermethylation, whereas homologous genes that are regulated by different promoter sequences, like *RBMX* (chromosome X) or *TSPYL5* (chromosome 8), show a different methylation profile compared to their Y-chromosome homologues *RBMY1A1* and *TSPY1*. This indicates that a sequence-dependent (promoter-specific) mechanism establishes the hypermethylation in somatic tissues. The *SYBL1* genes on both X and Y chromosomes appear to represent an exception in having almost identical promoters but being solely methylated on the Y chromosome[143,185]. Based on the methylation characteristics of *SYBL1*, the existence of Y-chromosome inactivation resembling the well-known X-chromosome inactivation in females was previously proposed[143]. However, according to the methylation pattern of all other Y-chromosome genes that were analyzed in the present study, this kind of phenomenon appears to be restricted to the short pseudo-autosomal region at the outmost end of the long arm of the Y-chromosome, which is flanked by a large heterochromatic region.

Taken together, these observations suggest the presence of sequence-specific regulatory mechanisms controlling either the specific demethylation of some promoters in testis/germ-line cells or their methylation in somatic tissues. Common *cis*-acting sequences responsible for this phenomenon are unknown so far and the attempt to associate specific DNA motifs with either group of promoters failed. The methylation of CpG islands in lymphocytes was recently shown to correlate with DNA sequence, repeats, and predicted DNA structure using bioinformatics scoring and prediction methods, suggesting that multiple factors may predispose CpG islands for DNA methylation[186]. A number of other bioinformatic attempts, however, indentified specific nucleotide sequences[187,188] that correlated with methylated or unmethylated CGI in cancer samples but failed to identify consensus sites for known transcription factors. Only one recent study identified SP1 to be associated with unmethylated CGI[169]. However, an independent study, done by our lab, recently obtained evidence that the methylation states of CGI in normal and cancer cells correlates with combinatorial binding of specific *cis*-acting factors (unpublished data).

Insulator proteins such as CTCF and Boris/CTCFL may also be involved in regulatory mechanisms controlling gene expression. An insulator is a DNA sequence that can act as a barrier to the influences of neighboring *cis*-acting elements and for example prevents gene activation when located between an enhancer and a promoter[189]. Normally, CTCF and BORIS are expressed in a mutually exclusive pattern that correlates with re-setting of methylation marks during male germ cell differentiation[190]. Recent studies have shown, that the testis-specific nuclear protein and the CTCF paralog BORIS/CTCFL[190] can initiate the demethylation of the cancer-testis antigens MAGE-A1[191] and NY-ESO-1[192] in normal somatic cells, suggesting that this transcription factor may be involved in the testis-specific

demethylation of at least some of the somatically methylated promoters. Moreover, De Smet and colleagues gained evidence, that the site-specific hypomethylation of MAGE-A1 in tumor cells relies on a transient process of demethylation followed by a persistent local inhibition of remethylation due to the presence of transcription factors[141]. Another report indicates that E2F6 may be essential for the long-term somatic silencing of certain male-germ-cell-specific genes[193]. In addition, Storre et al. suggests a molecular mechanism for the stable transcriptional silencing of meiotic genes in somatic cells by E2F6[194]. In line with this, a recent work indicates that E2F6 possesses a broad ability to bind to and regulate the meiosis-specific gene population[195]. Taken together, the above observations favor a model in which specific *cis*-acting sequences control promoter methylation at the level of somatic hypermethylation rather than by testis-specific demethylation.

The coexistence of DNA methylation and transcription in a certain tissue was also detected in this study. Transcription may indeed proceed in the presence of DNA methylation at those sites. However, there are a number of possibilities that may also explain these observations, including alternative promoter usage, tissue heterogeneity or simply false-positive microarray signals. It is clear that the presented pilot study provides only an initial and limited view of the human methylation landscape, although it is clearly in line with several other recent reports. Interindividual methylation variability, which might in addition be necessary to understand the role of DNA methylation in gene regulation, has not been addressed in this study. Moreover, only differences in dominant cell types within the tissue samples analyzed may be detected in these experiments, because tissues are composed of different cell types.

## 6.3 Allele-specific DNA methylation in mice

As detailed previously, epigenetic processes are fundamental for the differentiation and development of multicellular organisms by controlling chromatin accessibility and transcription[38]. Although phenotypic variation between individuals is mainly driven by genetic traits, there is also evidence that epigenetic mechanisms may contribute to inter-individual phenotypic differences in mammals[89]. Examples for epigenetic differences between individuals are comparatively rare and mostly, but not exclusively, confined to the level of DNA methylation, which represents a relatively stable epigenetic modification that can often be measured in an allelic context. There is well-documented evidence that epigenetic states can be inherited across generations. For example, agouti viable yellow (*Avy*) mice display inheritance of yellow fur as a result of incomplete erasure of the methylation signal associated with a retrotransposon insertion[96,97], and kinked-tail mice

transmit phenotype through multiple generations due to the loss of the silent epigenetic state at the *Axin* gene[196] as well as a heritable white-tail phenotype associated with Kit-specific microRNAs[197]. It was also reported, that allelic variation at certain epigenetic modifier genes in mice, like DNA methyl-transferases or chromatin remodeling factors may influence the inheritance of CpG methylation patterns in *trans*[98,99]. Differences at the level of DNA methylation have also been demonstrated in supposedly genetically identical, monozygotic twins that are acquired during the lifetime of each individual through largely unknown mechanisms[90,91].

In addition to *trans*-regulation, there are several reports demonstrating that DNA sequence variants associate with specific epigenetic states. Differential methylation states of individual alleles were first described for the in *c-Ha-ras-1* gene in human cells[198,199] More recently, a study by Murrell for example showed that loss of imprinting of the IGF2 gene, which is frequently observed in Beckwith-Wiedemann syndrome (BWS), correlated with loss of maternal allele-specific methylation (LOM) for certain haplotypes[94]. This observation was supported by another study, showing that variation in DNA methylation of the *IGF2/H19* locus is amongst others mainly determined by single nucleotide polymorphisms (SNPs) in *cis*[93]. In addition, Flanagan demonstrated allele-specific methylation patterns in the *CDH13* gene[92], and Yamada reported a CGI, which is methylated in an allele-specific but parental-origin-independent (non-imprinted) manner[200]. Along this line, a recent study in humans identified several cases of allele-specific DNA-methylation at non-imprinted gene loci[95], where the methylation status of each allele was likely controlled in *cis* by the local DNA sequence.

Taken together, the published data suggests that three types of inheritance of DNA methylation patters may exist in vivo: methylation patterns at non-imprinted loci may be inherited across generations based on genetic mechanisms (in *cis* and in *trans*) or based on epigenetic mechanisms (e.g. inheritance of DNA methylation, or other epigenetic marks).

As part of this thesis a broad analysis of differential DNA methylation in two inbred mouse strains was performed utilizing a modified protocol of DNA methylation-dependent genome fractionation (MCIp) in combination with locus-wide tiling arrays. Stringent hybridization conditions for microarray analyses of CpG-methylated and unmethylated genome fractions allowed the simultaneous detection of several hundred differentially methylated regions (DMRs). In addition, previously unrecognized strain-specific sequence variations that track with the large majority of DMRs were detected. The microarray results were validated by DNA sequencing (to confirm sequence variations) and a combination of bisulfite treatment of genomic DNA and mass spectrometry (to confirm CpG methylation

differences). The results suggest that the chosen approach is highly reproducible and correctly identified both DMRs and DNA sequence variations.

Why were inbred mice chosen as model organisms to investigate allele-specific DNA methylation? The house mouse is one of the most successful mammals and the premier research animal in mammalian biology. A strain is defined as inbred when it has been mated brother versus sister for 20 or more consecutive generations. Such inbred strains are therefore isogenic (genetically identical) allowing offspring to posses both genetic and phenotypic uniformity. Each inbred strain has a unique genotype and consequently a unique phenotype. The normal pigmented C57BL strain was originally derived by Little in 1921, whereas the strain BALB/c, which is albino and small in size, originated in 1923 by McDowell. C57BL is probably the most widely used of all inbred strains. Its sub strain C57BL/6, which developed before 1937, alone accounts for more than 14% of occasions on which an inbred strain is used.

Concerning immunology various differences between these two strains are known. The C57BL/6-BALB/c model has been intensively used to study the immunopathogenesis of several intracellular infections such as listeriosis[201], leihsmaniasis[202,203], yersiniosis[204,205] and mycobacterial infections[206,207]. For example, the BALB/c strain is highly susceptible to the infection of *Leishmania major*, in contrast to resistant C57BL/6[208,209]. Susceptible mice as BALB/c fail to control parasite proliferation and develop progressive lesions and systemic disease, which is associated with Th2 response manifested by high levels of IL-4, IL-13, IL-10 and antibody production. In contrast, resistant C57BL/6 mice develop IL-12-driven, interferon-gamma (INF-gamma) dominated Th1 response that promotes healing and parasite clearance[210,211]. It is likely that epigenetical differences are found in immune-cells derived from the two inbred stains especially in regions containing genes which are differentially expressed without or upon stimulation. Macrophages are widely distributed immune system cells and beyond that involved in all stages of the immune response. They were chosen because they actually contribute to the clearance of an intracellular pathogen. The leishmanial parasites for example live inside the macrophage, but the macrophage is also the host's major effector cell in the elimination of this pathogen[212]. Macrophages can be phenotypically polarized by the local cytokine pattern to trigger specific functional programs. Polarized macrophages are broadly classified in two main groups: classically activated macrophages (or M1), and alternatively activated macrophages (or M2). M1 exhibit potent microbicidal properties and promote strong Th1 responses, whilst M2 support Th2-associated effector functions. An activation of inappropriate effector mechanisms can lead to enhanced pathogen growth, rather than effective pathogen clearance. For example, classically activated macrophages protect

from *L. major* (Th1 response) in C57BL/6 mice, because they contribute to the clearance of the bacterium. On the contrary, in susceptible BALB/c mice, macrophages with a wound-healing phenotype (Th2 response) allow for its persistence[210,211]. It is also important, that macrophages can be grown *in vitro*, without possible environmental impacts.

Both mice strains used in this study (C57BL/6 and BALB/c) belong to the most widely studied inbred mice. Especially the underlying genetic background of the C57BL/6 strain has been sequenced in detail and serves as the reference sequence for murine genetic research. It was conceivable that differences between inbred mice strains are present. In particular, since it was proposed that the genomes of inbred strains are a mosaic of regions of different sub specific origins. The fine structure of such mosaic variation was described by Wade, reporting extremely high variation spanning one-third of the genome and low variation in the remaining part[213]. This bimodal distribution was thought to represent regions having different, or the same, sub specific origins, respectively. Following studies raised questions about the haplotype structure[214,215], the effect of ascertainment biases in subspecific assignment[216,217] and the contributions of intersubspecific versus intrasubspecific variation[218]. In addition, the presence of substantial intrasubspecific variation, ancestral polymorphisms and secondary introgression after the divergence of the subspecies was reported[218,219]. Frazer and colleagues recently identified 8.3 million SNPs by resequencing of 15 different mouse strains indicating that modern inbred strains are composed of at least four *mus musculus* substrains (*domesticus*, *molossinus*, *musculus* and *castaneus*)[151]. In contrast to the mosaic model, another recent work found that most of the genome has intermediate levels of variation of intrasubspecific origin[152]. However the true extent of genetic differences between these two specific mice strains still remains elusive.

In order to account for both, the reported genetic differences between strains and the unbalanced hybridization behavior seen in our experiments, we applied the virtual CGH analysis. The amount of microarray probes showing preferential hybridization only with the reference strain (C57BL/6) was unexpectedly high (>15%) in the present experiments. Although the analysis focused on regions where differential gene expression was observed, a huge number of array probes were affected. Subchromosomal alterations such as copy number variations (CNV) were seen and validated during the virtual CGH analysis. These findings were in line with a recently published study, which identified some of the larger CNV's as well using standard CGH approaches[150]. In addition to these large scale variations a huge number of smaller alterations were observed including

insertions and deletions within one strain and single nucleotide polymorphisms (SNPs). The accumulation of these sequence variations correlated well with probes exhibiting unequal hybridization behavior. Since probes were designed according to the genomic sequence of the reference strain C57BL/6, the affected ones were only conspicuous for BALB/c. Most previously known SNPs, as identified by Frazer[151], correlated with affected probes. In addition, a large number of previously unknown SNPs were found in exemplary sequenced regions showing unequal hybridization behavior. The genomic sequencing data yielded also that the position of a number of repetitive elements was altered in BALB/c as compared to C57BL/6. However, it is unclear if the BALB/c strain contains more or less repetitive elements in general.
Taken together, it is therefore conceivable that the genetic backgrounds of these two mouse strains vary to a greater extent than previously known and expected. These variations may in part contribute to the differential immune answer in both strains.

After correction by vCGH our experiments defined a large number of allele-specific methylation events. In the following a representative set of these genomic regions were traced in the germ line of parental animals. Interestingly, most DMRs that were studied in detail, were only differentially methylated within the somatic tissues studied (BMM and spleen) and showed equal levels of methylation in the male germ line, suggesting that the observed differences are established post-fertilization and DNA methylation itself is not inherited across generations.
Furthermore, the exemplary mass spectrometry validation of DMRs and allele specific examination of DNA methylation in F1 Hybrids using traditional sequencing of bisulfite treated genomic DNA suggests, that the propensity for acquiring DNA methylation during development mainly depended on the local sequence context in *cis*. It is formally possible that another, yet uncharacterized epigenetic mark at these sites determines the allelic methylation status in a heritable fashion. However, since a common association of DMRs with sequence variations was found, it is much more likely that the establishment of differential methylation patterns during embryonic development may involve *cis*-acting regulators, including sequence-specific DNA-binding proteins or other sequence-dependent regulatory mechanisms. Several publications accentuate the contribution of *cis*-acting regulators as observed in this thesis. The findings of this work are in line with a number of previous studies linking allelic DNA-methylation and genetic variation which were already discussed in the beginning of this chapter[92-95,200] and share an obvious corollary with the recent observation that the genetic sequence is largely responsible for directing species-specific transcription in mice carrying the human chromosome 21[220]. Moreover, investigating the same human chromosome, Zhang and colleagues very

recently indentified three cases of allele-specific DNA methylation showing strong sequence dependence similar to our results[221]. However, our analysis contrasts a lately published study that did not find significant variation in DNA methylation between inbred or outbred mice and concluded that the impact of DNA polymorphisms on DNA methylation would not seem to be common[91]. The reasons for the deviant findings are not clear but may relate to the different technologies and readout systems used in each study. For example, the above study used a CpG island microarray covering 2176 unique CpG island regions of the mouse genome that are likely more conserved, frequently contain gene promoters and are often unmethylated. Based on our findings that differential DNA methylation is associated with genetic variation and promoter-distal sites, we would predict that the microarray platform utilized by Kaminski et al. is less likely to detect allele-specific DNA methylation than the considerably larger and locus-wide microarray used in our study.

It is conceivable that *trans*-acting mechanisms contribute to the epigenetic status of a specific region as well. It has recently been published that the absence or functional inactivation of the few known epigenetic modifiers that mediate DNA methylation in *trans* show profound developmental abnormalities[222]. However, our data suggests, that only a relatively small proportion of strain-specific DMRs (like the ones identified in *Isoc2b*, *Sfi1* pseudogene, or *Eps8l1*) depend dominantly on *trans*-acting mechanisms.

Taken together, the presented findings suggest that genetic variability has a major impact on epigenetic variability, likely due to its influence on the propensity of individual alleles to become methylated (or demethylated) during development. This relationship is likely not restricted to early embryonic development in mice, but may also apply to other developmental (including disease- and age-related) processes in mammals. For example Zhang et al. demonstrated allele-specific DNA methylation in human samples as well[221]. In essence, the results imply that any given DNA sequence (at a non-imprinted locus) may carry intrinsic information that recruits *cis*-acting factors and determines its DNA methylation status during mammalian development.

What is the relevance of the present findings? The systematic screening of the human genome for genetic variants that affect gene regulation was thought to advance the fundamental understanding of phenotypic diversity and lead to the identification of alleles that modify disease risk. It becomes more and more obvious, that the allele-specific expression of genes is, for instance, controlled by *cis*-acting factors (such as DNA polymorphisms and methylation) in the flanking DNA sequence and *trans*-acting modulators (such as transcription factors) that are themselves regulated by other genetic (developmental or environmental) factors. Recently a study showed that genetics may be an important factor that influences global chromatin states by using allele-specific ChiP

(chromatin immunoprecipitation)[223]. These allelic differences in chromatin may provide an explanation for the observed allelic variation in gene expression[224]. Evidence for the medical importance of *cis*-acting polymorphisms has been provided by positional cloning of susceptibility genes for common diseases such as stroke[225] and type 2 diabetes[226]. The polymorphisms in these examples were not associated with protein coding or splice-site variations. However, diverse *cis*-acting mechanisms might also be important for malignant diseases. Successful epigenetic control of gene expression is necessary and aberrant DNA methylation, which has been recognized as an important feature in cancer formation, leads to the altered gene expression seen in almost all human cancers. Originally only genome-wide DNA hypomethylation was linked with cancer and was the focus of tumor research[37,227,228]. But during the mid 1980s regional hypermethylation of specific cancer-associated gene promoter sites attracted greater attention and the focus relocated to the participation of hypermethylation to cancer formation[229-231]. The present study strongly suggests, that individual alleles can exhibit different probabilities to become methylated (regulated in *cis*). Interestingly, a recent study investigated the regulation of allele-specific expression (ASE) by *cis*-acting elements in acute lymphoblastic leukemia (ALL)[232]. They found that allele-specific CpG site methylation is one of the factors that regulate allele-specific gene expression in ALL cells. Besides, an independent study, done by our lab, recently obtained evidence of allele-specific methylation of certain CGIs in acute myeloid leukemia (AML) samples (unpublished data). However, further studies are certainly required, in order to shed more light on a possible connection between *cis*-acting mechanisms and cancer formation.

Taken together, this study describes the first comprehensive identification of allele-specific DNA-methylation in mice, and provides a resource for further mechanistic studies. The findings presented in this thesis imply that the genetic context of a specific region is strongly regulating the epigenetic status in terms of DNA methylation. It seems that genetic variability has a major impact on epigenetic variability and that this variability contributes to the shaping of different individual phenotypes. However, this might not only have an impact for early developmental steps, but also for disease-related processes in mammals.

## 6.4 Outlook

This pilot study enabled the detection of cell type-specific differences in promoter DNA methylation with a reasonable resolution. Hence, the study presents a novel application of the MCIp methodology that may fill previous methodical gaps allowing the reliable detection of hypomethylated DNA fragments on a genome-wide level. It is also clear that this study provides only an initial and limited view of the human methylation landscape. Since tissues are composed of different cell types, we likely detected differences in dominant cell types. Further analysis would certainly benefit when purified cell populations are used. In addition, matched methylation and expression profiling data will be necessary to completely understand the relationship between DNA methylation and gene expression. Our studies concentrated on promoters so far. In the literature, there is rising data on distribution and dynamics of DNA methylation off-side promoter regions. For example, Eckhardt et al. reported that tissue-specific DNA methylation patterns are not only found in promoters but also in conserved intergenic elements, leading the authors to propose that DNA methylation regulates the activity of distant enhancers[153]. In line with this a recent study done by our lab showed that cell type-specific methylation differences mainly occur in promoter off-side regions of low CpG content (no CGIs)[159]. This study concentrated on specific T-cell subsets, but the conclusions very likely apply to other cell types as well and would be interesting to study on more diverse cell types or tissues. In addition, a combined analysis of DNA methylation and chromatin marks may lead to the identification of promoter distal regions with enhancer/silencer activity.

On the other hand the present study demonstrated that genetic sequence variations contribute to epigenetic regulation in a *cis*-acting manner. The relevance of these findings for human disease-related research is stated above. For years, much effort has been made to understand epigenetic reasons for malignant growth as well as other non-cancerous human diseases, but the interpretation of disease-related DNA methylation pattern was and is certainly very difficult and additionally complicated by the occurrence of normal cell developmental processes as well as age-related processes. Moreover, there are other determinants as well, such as the local genetic context regulating the methylation pattern in *cis*. Other determinates might regulate DNA methylation in a *trans*-acting manner. In general the true extent of contribution from genetic determinants such as sequence polymorphisms to epigenetic patterns is currently unknown. Hence, comparative DNA methylation profiling is interesting for other model systems as well, which could be easily achieved with this approach. In principal, all disease processes could be analyzed regarding allele-specific DNA methylation differences. Allele-specific DNA methylation may also have an impact on psychological and behavioral issues as well

as disease- and cancer-related issues. For instance, individual phenotypical differences between twins could be addressed using our methodology. An in-depth investigation of human allele-specific methylation will likely contribute to the identification of distinct epialleles, which are not only involved in individual phenotypical characteristics but may also affect different individual disease susceptibility and probability to develop cancer. As a first step one could investigate the occurrence of allele-specific DNA methylation in human cell lines. However, CGH-microarray-based detection has limited ability to measure allelic DNA methylation and bisulfite sequencing, though less limited in this regard, requires extensive resources when applied genome-wide. MCIp in combination with SNP arrays for example may provide an opportunity to identify specific alleles, which are more prone to DNA methylation. Data on allelic methylation asymmetry will likely be useful for fine mapping and interpreting associations of (noncoding) SNPs to complex genetic diseases.

In our analysis we also gained evidence that the DNA methylation of some regions might be regulated by a certain strain-specific epigenetic modifier (in *trans*). The identification of the factor being responsible for the epigenetic control of for example the *Isoc2b* gene promoter in parental strains can provide interesting information about the epigenetic regulation of such particular regions. A useful tool in to search for such a factor might be a methodology called marker-assisted backcrossing. Backcrossing leads to offspring with a genetic identity which is closer to that of one parent. By screening for F2-offspring individuals with either F1-like or parental-like patterns one might be able to clarify the location of genomic regions harboring the modifier possibly leading to its identification.

# 7 Summary

Epigenetic mechanisms such as histone modifications and DNA methylation extend the information of the underlying genomic DNA sequence. Understanding cell type-specific epigenetic codes on a global level is a major challenge after the sequencing of the human genome has been completed. In addition, differences in epigenetic patterns between individuals may contribute to phenotypic variation and disease susceptibility. However, little is known about the extent of such variation or how different epigenetic patterns are established. In the present thesis, I applied and modified the methyl-CpG immunoprecipitation (MCIp) method to globally map DNA methylation. This method is based on a recombinant, methyl-CpG-DNA binding protein (MBD2-Fc), which is used to fractionate genomic DNA depending on its methylation status.

During this thesis the MCIp-method was on the one hand modified and established for a novel application. This study was designed as a pilot project to establish the MCIp approach for comparative hypomethylation analysis. To study cell-type-specific DNA methylation differences, the MCIp procedure was used in combination with oligonucleotide promoter microarrays (MCIp-on-chip) to identify differences in promoter hypomethylation of coding and non-coding genes in human tissues. Forty-four identified tissue-specifically methylated promoters were independently validated using bisulfite sequencing or single gene MCIp, confirming the results obtained by the MCIp microarray approach. A comparison of the obtained tissue methylation profiles with corresponding gene expression data indicated a significant association between tissue-specific promoter methylation and gene expression, not only in CpG-rich promoters. The data also highlighted the exceptional epigenetic status of germ line cells in testis, where many testis-specific promoters were demethylated as compared to somatic tissues. Somatic hypermethylation particularly affected many CpG-rich promoters of testis-specific genes e.g. in ampliconic areas of the Y-chromosome.

In order to analyze individual (allele-specific) DNA methylation the MCIp-on-chip method was on the other hand further improved to the “mirror image” procedure. The “mirror image” approach was used to analyze the methylation profiles of immune cells from two inbred mice strains (C57BL/6 and BALB/c) that represent prototypic models for Th1- or Th2-dominated immune responses. Using MCIp the genomes of bone marrow-derived macrophages were fractionated into methylated and unmethylated genome pools and separately analyzed by microarray at 181 genomic regions (covering 28 Mb of mouse

genome) that were selected based on differential gene expression between both mouse strains. The combined analysis of hypo- and hypermethylated profiles allowed the identification of several hundred differentially methylated regions, but also uncovered several copy number variations as well as many single nucleotide polymorphisms (SNPs) or micro- and macro-deletions/insertions. By using DNA sequencing and mass spectrometry it was shown for a subset of regions that specific allelic methylation patterns in somatic cells were maintained in F1- hybrid animals, regardless of their status in germ line cells. A common association with sequence polymorphisms suggests that the genomic context at these differentially methylated regions determines their developmentally regulated methylation in *cis*.

# 8 Deutsche Zusammenfassung

Epigenetik im Allgemeinen und DNA Methylierung im Speziellen, erweitern den Informationsgehalt der zugrundeliegenden genomischen DNA-Sequenz. Nachdem das menschliche Genom weitgehend sequenziert wurde, ist insbesondere das Verständnis von zelltypspezifischen epigenetischen Kodierungen in genomweiter Hinsicht die nächste zu bewältigende Herausforderung. Darüber hinaus sind unterschiedliche epigenetische Muster auch im Verdacht, sowohl zu unterschiedlichen individuellen phänotypischen Ausprägungen, als auch zu der individuellen Anfälligkeit für bestimmte Krankheiten beizutragen. Allerdings ist über das Ausmaß solcher Unterschiede und deren Etablierung sehr wenig bekannt. In dieser Arbeit wurde die Methyl-CpG Immunpräzipitations (MCIp)-Methode verwendet, um die DNA Methylierung unter zwei Aspekten zu untersuchen. Die Methode basiert auf einem rekombinanten, methyl-CpG-DNA bindenden Protein (MBD2-Fc), mit dessen Hilfe genomische DNA hinsichtlich ihres Methylierungsgrades fraktioniert werden kann.

Zunächst wurde die MCIp Methode während dieser Arbeit modifiziert und für eine neue Anwendung etabliert. Diese Studie war als Pilotprojekt angelegt um die Verwendung der MCIp Methode zur vergleichenden Analyse der DNA Hypomethylierung zu etablieren. Um zelltypspezifische DNA Methylierungsunterschiede untersuchen zu können, wurde die MCIp Methode mit Promotor Microarrays kombiniert, um gewebespezifische Unterschiede in der Hypomethylierung von Promotoren codierender und nicht codierender Gene zu analysieren. Daraufhin wurden 44 als gewebespezifisch methyliert identifizierte Promotoren mit Microarray-unabhängigen Methoden validiert. Hierzu wurde sowohl die MCIp-Methode zum Nachweis der Methylierung an Einzelgenen als auch die Sequenzierung von Natriumbisulfit-behandelter genomischer DNA angewendet. Nach Validierung der Arrayergebnisse wurden die resultierenden gewebsspezifischen Methylierungsmuster mit der Expression der entsprechenden Gene verglichen. Es stellte sich heraus, dass nicht nur in stark CpG-haltigen Promotoren die gewebespezifische Methylierung und die Expression signifikant miteinander korrelierten. Sondern auch, dass Keimbahnzellen aus dem Hodengewebe, ein im Vergleich zu somatischen Zellen, außergewöhnliches epigenetisches Muster aufwiesen. So waren im Besonderen, stark CpG-haltige Promotoren hodenspezifischer Gene von somatischer Hypermethylierung betroffen. Diese Gene waren vor allem in ampliconischen Bereichen des Y-Chromosoms zu finden. Die kombinierte Anwendung von MCIp, zur Probenvorbehandlung, und

Microarrays, als globales Auslesesystem, ermöglichten in diesem Projekt die unverfälschte und unvoreingenommene Analyse gewebespezifischer Promotormethylierung.

Um individuelle (allele-spezifische) DNA Methylierungsunteschiede untersuchen zu können, wurde die MCIp-on-chip zur „mirror image“ Anwendung weiterentwickelt. Damit wurden in einem zweiten Projekt Immunzellen zweier Maus-Inzuchtstämme (C57BL/6 und BALB/c), die ein prototypisches Model für die Th1- oder Th2-dominierte Immunantwort darstellen, analysiert. Hierbei wurde die MCIp-Methode verwendet, um die Genome von murinen Knochenmarksmakrophagen in einen methylierten und einen unmethylierten Teil aufzutrennen. Die jeweiligen korrespondierenden DNA-Pools beider Mausstämme wurden auf einen spezifisch hergestellten Microarray hybridisiert. Da epigenetische Unterschiede vermehrt in Bereichen differentieller Genexpression zwischen Knochenmarks-makrophagen beider Mausstämme zu erwarten sind, wurden zunächst 181 dieser Regionen ausgewählt und für das Arraydesign herangezogen. Insgesamt wurden damit 28 Mb des murinen Genoms abgedeckt. Die kombinierte Analyse der hypo- und hypermethylierten Teilbereiche ermöglichte die Identifizierung von mehreren hundert unterschiedlich methylierten Regionen. Ferner wurden auch verstärkt genetische Variationen, wie Unterschiede in der Genkopiezahl, Insertionen und Deletionen im Mikro- und Makromaßstab, sowie Einzelnukleotid-Polymorphismen zwischen beiden Mausstämmen festgestellt. Anschließend wurde eine ausgewählte Anzahl von Regionen mittels MALDI-TOF Massenspektrometrie und Sequenzierung von Natriumbisulfit-behandelter DNA eingehender untersucht. Dabei zeigte sich, dass die Methylierungsmuster elterlicher Allele im Wesentlichen auch in F1-Tieren, unabhängig vom Methylierungsgrad dieser Regionen in Keimbahnzellen, beibehalten wurden. Dieses Phänomen war häufig mit Sequenz-unterschieden zwischen den Allelen assoziiert. Dies führte zu der Annahme, dass die entwicklungsspezifische Methylierung dieser Regionen maßgeblich durch die lokale genetische Sequenz in *cis* reguliert wird.

# 9 Reference List


1. Ohgane,J., Yagi,S. & Shiota,K. Epigenetics: the DNA methylation profile of tissue-dependent and differentially methylated regions in cells. *Placenta.* **29 Suppl A:S29-35. Epub;%2007 Nov 26.**, S29-S35 (2008).

2. Okano,M., Bell,D.W., Haber,D.A. & Li,E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell.* **99**, 247-257 (1999).

3. Li,E., Beard,C. & Jaenisch,R. Role for DNA methylation in genomic imprinting. *Nature.* **366**, 362-365 (1993).

4. Goto,T. & Monk,M. Regulation of X-chromosome inactivation in development in mice and humans. *Microbiol. Mol. Biol. Rev.* **62**, 362-378 (1998).

5. Kass,S.U., Pruss,D. & Wolffe,A.P. How does DNA methylation repress transcription? *Trends Genet.* **13**, 444-449 (1997).

6. Plass,C. & Soloway,P.D. DNA methylation, imprinting and cancer. *Eur. J. Hum. Genet.* **10**, 6-16 (2002).

7. Suzuki,M.M. & Bird,A. DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* **9**, 465-476 (2008).

8. Gardiner-Garden,M. & Frommer,M. CpG islands in vertebrate genomes. *J. Mol. Biol.* **%20;196**, 261-282 (1987).

9. Bestor,T.H. & Ingram,V.M. Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 5559-5563 (1983).

10. Bestor,T.H. The DNA methyltransferases of mammals. *Hum. Mol. Genet.* **9**, 2395-2402 (2000).

11. Colot,V. & Rossignol,J.L. Eukaryotic DNA methylation as an evolutionary device. *Bioessays.* **21**, 402-411 (1999).

12. Lauster,R., Trautner,T.A. & Noyer-Weidner,M. Cytosine-specific type II DNA methyltransferases. A conserved enzyme core with variable target-recognizing domains. *J. Mol. Biol.* **%20;206**, 305-312 (1989).

13. Posfai,J., Bhagwat,A.S., Posfai,G. & Roberts,R.J. Predictive motifs derived from cytosine methyltransferases. *Nucleic Acids Res.* **17**, 2421-2435 (1989).

14. Margot,J.B., Ehrenhofer-Murray,A.E. & Leonhardt,H. Interactions within the mammalian DNA methyltransferase family. *BMC. Mol. Biol.* **4:7. Epub;%2003 May 30.**, 7 (2003).

15. Santi,D.V., Garrett,C.E. & Barr,P.J. On the mechanism of inhibition of DNA-cytosine methyltransferases by cytosine analogs. *Cell.* **33**, 9-10 (1983).

16. Villar-Garea,A. & Esteller,M. DNA demethylating agents and chromatin-remodelling drugs: which, how and why? *Curr. Drug Metab.* **4**, 11-31 (2003).

17. Erlanson,D., Dhen,L. & Verdine,G.L. Enzymatic DNA methylation through a locally unpaired intermediate. *J. Am. Chem. Soc.* 12583-12584 (1993).

18. Yoder,J.A., Soman,N.S., Verdine,G.L. & Bestor,T.H. DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. *J. Mol. Biol.* **270**, 385-395 (1997).

19. Siedlecki,P. & Zielenkiewicz,P. Mammalian DNA methyltransferases.

20. Lei,H. *et al.* De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development.* **122**, 3195-3205 (1996).

21. Okano,M., Xie,S. & Li,E. Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res.* **26**, 2536-2540 (1998).

22. Chedin,F., Lieber,M.R. & Hsieh,C.L. The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 16916-16921 (2002).

23. Bourc'his,D., Xu,G.L., Lin,C.S., Bollman,B. & Bestor,T.H. Dnmt3L and the establishment of maternal genomic imprints. *Science.* **294**, 2536-2539 (2001).

24. Hata,K., Okano,M., Lei,H. & Li,E. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development.* **129**, 1983-1993 (2002).

25. Nan,X., Meehan,R.R. & Bird,A. Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res.* **21**, 4886-4892 (1993).

26. Meehan,R.R., Lewis,J.D. & Bird,A.P. Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic Acids Res.* **20**, 5085-5092 (1992).

27. Bird,A.P. & Wolffe,A.P. Methylation-induced repression--belts, braces, and chromatin. *Cell.* **99**, 451-454 (1999).

28. Prokhortchouk,A. *et al.* The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes Dev.* **15**, 1613-1618 (2001).

29. Yoon,H.G., Chan,D.W., Reynolds,A.B., Qin,J. & Wong,J. N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso. *Mol. Cell.* **12**, 723-734 (2003).

30. Jones,P.L. *et al.* Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.* **19**, 187-191 (1998).

31. Nan,X. *et al.* Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature.* **393**, 386-389 (1998).

32. Sarraf,S.A. & Stancheva,I. Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. *Mol. Cell.* **15**, 595-605 (2004).

33. Klose,R.J. *et al.* DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. *Mol. Cell.* **19**, 667-678 (2005).

34. Dhasarathy,A. & Wade,P.A. The MBD protein family-reading an epigenetic mark? *Mutat. Res.* **647**, 39-43 (2008).

35. Walsh,C.P. & Bestor,T.H. Cytosine methylation and mammalian development. *Genes Dev.* **13**, 26-34 (1999).

36. Warnecke,P.M. & Clark,S.J. DNA methylation profile of the mouse skeletal alpha-actin promoter during development and differentiation. *Mol. Cell Biol.* **19**, 164-172 (1999).

37. Feinberg,A.P. & Tycko,B. The history of cancer epigenetics. *Nat. Rev. Cancer.* **4**, 143-153 (2004).

38. Bird,A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6-21 (2002).

39. Bird,A.P. CpG-rich islands and the function of DNA methylation. *Nature.* **321**, 209-213 (1986).

40. Gidekel,S. & Bergman,Y. A unique developmental pattern of Oct-3/4 DNA methylation is controlled by a cis-demodification element. *J. Biol. Chem.* **277**, 34521-34530 (2002).

41. Ponger,L., Duret,L. & Mouchiroud,D. Determinants of CpG islands: expression in early embryo and isochore structure. *Genome Res.* **11**, 1854-1860 (2001).

42. Robinson,P.N., Bohme,U., Lopez,R., Mundlos,S. & Nurnberg,P. Gene-Ontology analysis reveals association of tissue-specific 5' CpG-island genes with development and embryogenesis. *Hum. Mol. Genet.* **13**, 1969-1978 (2004).

43. Takai,D. & Jones,P.A. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl. Acad. Sci. U. S. A.* **%19;99**, 3740-3745 (2002).

44. Selker,E.U. *et al.* The methylated component of the Neurospora crassa genome. *Nature.* **422**, 893-897 (2003).

45. Tweedie,S., Charlton,J., Clark,V. & Bird,A. Methylation of genomes and genes at the invertebrate-vertebrate boundary. *Mol. Cell Biol.* **17**, 1469-1475 (1997).

46. Montero,L.M. *et al.* The distribution of 5-methylcytosine in the nuclear genome of plants. *Nucleic Acids Res.* **20**, 3207-3210 (1992).

47. Palmer,L.E. *et al.* Maize genome sequencing by methylation filtration. *Science.* **%19;302**, 2115-2117 (2003).

48. SanMiguel,P. *et al.* Nested retrotransposons in the intergenic regions of the maize genome. *Science.* **274**, 765-768 (1996).

49. Batzer,M.A. & Deininger,P.L. Alu repeats and human genomic diversity. *Nat. Rev. Genet.* **3**, 370-379 (2002).

50. Chen,C., Gentles,A.J., Jurka,J. & Karlin,S. Genes, pseudogenes, and Alu sequence organization across human chromosomes 21 and 22. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 2930-2935 (2002).

51. Gama-Sosa,M.A. *et al.* The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Res.* **11**, 6883-6894 (1983).

52. Kochanek,S., Renz,D. & Doerfler,W. DNA methylation in the Alu sequences of diploid and haploid primary human cells. *EMBO J.* **12**, 1141-1151 (1993).

53. Li,E., Bestor,T.H. & Jaenisch,R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell.* **69**, 915-926 (1992).

54. Chen,T. & Li,E. Structure and function of eukaryotic DNA methyltransferases. *Curr. Top. Dev. Biol.* **60:55-89.**, 55-89 (2004).

55. Dodge,J.E. *et al.* Inactivation of Dnmt3b in mouse embryonic fibroblasts results in DNA hypomethylation, chromosomal instability, and spontaneous immortalization. *J. Biol. Chem.* **280**, 17986-17991 (2005).

56. Karpf,A.R. & Matsui,S. Genetic disruption of cytosine DNA methyltransferase enzymes induces chromosomal instability in human cancer cells. *Cancer Res.* **65**, 8635-8639 (2005).

57. Walsh,C.P., Chaillet,J.R. & Bestor,T.H. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* **20**, 116-117 (1998).

58. Bird,A.P. Gene number, noise reduction and biological complexity. *Trends Genet.* **11**, 94-100 (1995).

59. McGrath,J. & Solter,D. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell.* **37**, 179-183 (1984).

60. Surani,M.A., Barton,S.C. & Norris,M.L. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature.* **308**, 548-550 (1984).

61. Cattanach,B.M. & Kirk,M. Differential activity of maternally and paternally derived chromosome regions in mice. *Nature.* **315**, 496-498 (1985).

62. Barlow,D.P., Stoger,R., Herrmann,B.G., Saito,K. & Schweifer,N. The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. *Nature.* **349**, 84-87 (1991).

63. Bartolomei,M.S., Zemel,S. & Tilghman,S.M. Parental imprinting of the mouse H19 gene. *Nature.* **351**, 153-155 (1991).

64. DeChiara,T.M., Robertson,E.J. & Efstratiadis,A. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell.* **64**, 849-859 (1991).

65. Wendt,K.S. *et al.* Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature.* **451**, 796-801 (2008).

66. Kobayashi,H. *et al.* Bisulfite sequencing and dinucleotide content analysis of 15 imprinted mouse differentially methylated regions (DMRs): paternally methylated

DMRs contain less CpGs than maternally methylated DMRs. *Cytogenet. Genome Res.* **113**, 130-137 (2006).

67. Reik,W., Dean,W. & Walter,J. Epigenetic reprogramming in mammalian development. *Science.* **293**, 1089-1093 (2001).

68. Davis,T.L., Yang,G.J., McCarrey,J.R. & Bartolomei,M.S. The H19 methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. *Hum. Mol. Genet.* **9**, 2885-2894 (2000).

69. Lucifero,D., Mertineit,C., Clarke,H.J., Bestor,T.H. & Trasler,J.M. Methylation dynamics of imprinted genes in mouse germ cells. *Genomics.* **79**, 530-538 (2002).

70. Bielinska,B. *et al.* De novo deletions of SNRPN exon 1 in early human and mouse embryos result in a paternal to maternal imprint switch. *Nat. Genet.* **25**, 74-78 (2000).

71. Lefebvre,L. *et al.* Abnormal maternal behaviour and growth retardation associated with loss of the imprinted gene Mest. *Nat. Genet.* **20**, 163-169 (1998).

72. Li,L. *et al.* Regulation of maternal behavior and offspring growth by paternally expressed Peg3. *Science.* **284**, 330-333 (1999).

73. LYON,M.F. Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature.* **190:372-3.**, 372-373 (1961).

74. Lee,J.T., Strauss,W.M., Dausman,J.A. & Jaenisch,R. A 450 kb transgene displays properties of the mammalian X-inactivation center. *Cell.* **86**, 83-94 (1996).

75. Rastan,S. & Robertson,E.J. X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. *J. Embryol. Exp. Morphol.* **90:379-88.**, 379-388 (1985).

76. Brockdorff,N. *et al.* Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature.* **351**, 329-331 (1991).

77. Brown,C.J. *et al.* A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature.* **349**, 38-44 (1991).

78. Clerc,P. & Avner,P. Role of the region 3' to Xist exon 6 in the counting process of X-chromosome inactivation. *Nat. Genet.* **19**, 249-253 (1998).

79. Lee,J.T. & Lu,N. Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell.* **99**, 47-57 (1999).

80. Marahrens,Y., Loring,J. & Jaenisch,R. Role of the Xist gene in X chromosome choosing. *Cell.* **92**, 657-664 (1998).

81. Lee,J.T., Davidow,L.S. & Warshawsky,D. Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat. Genet.* **21**, 400-404 (1999).

82. Wutz,A. & Jaenisch,R. A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol. Cell.* **5**, 695-705 (2000).

83. Brown,C.J. & Willard,H.F. The human X-inactivation centre is not required for maintenance of X-chromosome inactivation. *Nature.* **368**, 154-156 (1994).

84. Csankovszki,G., Panning,B., Bates,B., Pehrson,J.R. & Jaenisch,R. Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. *Nat. Genet.* **22**, 323-324 (1999).

85. Sado,T. *et al.* X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. *Dev. Biol.* **225**, 294-303 (2000).

86. Beard,C., Li,E. & Jaenisch,R. Loss of methylation activates Xist in somatic but not in embryonic cells. *Genes Dev.* **9**, 2325-2334 (1995).

87. Panning,B. & Jaenisch,R. DNA hypomethylation can activate Xist expression and silence X-linked genes. *Genes Dev.* **10**, 1991-2002 (1996).

88. Cooper,D.W. Directed genetic change model for X chromosome inactivation in eutherian mammals. *Nature.* **230**, 292-294 (1971).

89. Rakyan,V.K., Blewitt,M.E., Druker,R., Preis,J.I. & Whitelaw,E. Metastable epialleles in mammals. *Trends Genet.* **18**, 348-351 (2002).

90. Fraga,M.F. *et al.* Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 10604-10609 (2005).

91. Kaminsky,Z.A. *et al.* DNA methylation profiles in monozygotic and dizygotic twins. *Nat. Genet.* **41**, 240-245 (2009).

92. Flanagan,J.M. *et al.* Intra- and interindividual epigenetic variation in human germ cells. *Am. J. Hum. Genet.* **79**, 67-84 (2006).

93. Heijmans,B.T., Kremer,D., Tobi,E.W., Boomsma,D.I. & Slagboom,P.E. Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus. *Hum. Mol. Genet.* **16**, 547-554 (2007).

94. Murrell,A. *et al.* An association between variants in the IGF2 gene and Beckwith-Wiedemann syndrome: interaction between genotype and epigenotype. *Hum. Mol. Genet.* **13**, 247-255 (2004).

95. Kerkel,K. *et al.* Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. *Nat. Genet.* **40**, 904-908 (2008).

96. Blewitt,M.E., Vickaryous,N.K., Paldi,A., Koseki,H. & Whitelaw,E. Dynamic reprogramming of DNA methylation at an epigenetically sensitive allele in mice. *PLoS. Genet.* **2**, e49 (2006).

97. Morgan,H.D., Sutherland,H.G., Martin,D.I. & Whitelaw,E. Epigenetic inheritance at the agouti locus in the mouse. *Nat. Genet.* **23**, 314-318 (1999).

98. Blewitt,M.E. *et al.* SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat. Genet.* **40**, 663-669 (2008).

99. Chong,S. *et al.* Modifiers of epigenetic reprogramming show paternal effects in the mouse. *Nat. Genet.* **39**, 614-622 (2007).

100. Reik,W. *et al.* Imprinting mutations in the Beckwith-Wiedemann syndrome suggested by altered imprinting pattern in the IGF2-H19 domain. *Hum. Mol. Genet.* **4**, 2379-2385 (1995).

101. Frommer,M. *et al.* A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 1827-1831 (1992).

102. Hayatsu,H., Wataya,Y., Kai,K. & Iida,S. Reaction of sodium bisulfite with uracil, cytosine, and their derivatives. *Biochemistry.* **9**, 2858-2865 (1970).

103. Dai,Z. *et al.* Global methylation profiling of lung cancer identifies novel methylated genes. *Neoplasia.* **3**, 314-323 (2001).

104. Yoshikawa,H. *et al.* SOCS-1, a negative regulator of the JAK/STAT pathway, is silenced by methylation in human hepatocellular carcinoma and shows growth-suppression activity. *Nat. Genet.* **28**, 29-35 (2001).

105. Yang,H.J. *et al.* Detection of hypermethylated genes in tumor and plasma of cervical cancer patients. *Gynecol. Oncol.* **93**, 435-440 (2004).

106. Eads,C.A. *et al.* MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Res.* **28**, E32 (2000).

107. Weber,M. *et al.* Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* **37**, 853-862 (2005).

108. Mukhopadhyay,R. *et al.* The binding sites for the chromatin insulator protein CTCF map to DNA methylation-free domains genome-wide. *Genome Res.* **14**, 1594-1602 (2004).

109. Cross,S.H., Charlton,J.A., Nan,X. & Bird,A.P. Purification of CpG islands using a methylated DNA binding column. *Nat. Genet.* **6**, 236-244 (1994).

110. Gebhard,C. *et al.* Genome-wide profiling of CpG methylation identifies novel targets of aberrant hypermethylation in myeloid leukemia. *Cancer Res.* **66**, 6118-6128 (2006).

111. Bentley,D.R. Whole-genome re-sequencing. *Curr. Opin. Genet. Dev.* **16**, 545-552 (2006).

112. Cokus,S.J. *et al.* Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature.* **452**, 215-219 (2008).

113. Ehrich,M. *et al.* Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15785-15790 (2005).

114. Stanssens,P. *et al.* High-throughput MALDI-TOF discovery of genomic sequence polymorphisms. *Genome Res.* **14**, 126-133 (2004).

115. Hatada,I., Hayashizaki,Y., Hirotsune,S., Komatsubara,H. & Mukai,T. A genomic scanning method for higher organisms using restriction sites as landmarks. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 9523-9527 (1991).

116. Khulan,B. *et al.* Comparative isoschizomer profiling of cytosine methylation: the HELP assay. *Genome Res.* **16**, 1046-1055 (2006).

117. Vaughn,M.W. *et al.* Epigenetic Natural Variation in Arabidopsis thaliana. *PLoS. Biol.* **%19;5**, e174 (2007).

118. Irizarry,R.A. *et al.* Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome Res.* **18**, 780-790 (2008).

119. Lippman,Z., Gendrel,A.V., Colot,V. & Martienssen,R. Profiling DNA methylation patterns using genomic tiling microarrays. *Nat. Methods.* **2**, 219-224 (2005).

120. Zhang,X. *et al.* Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell.* **126**, 1189-1201 (2006).

121. Zilberman,D., Gehring,M., Tran,R.K., Ballinger,T. & Henikoff,S. Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nat. Genet.* **39**, 61-69 (2007).

122. Weber,M. *et al.* Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* **39**, 457-466 (2007).

123. Keshet,I. *et al.* Evidence for an instructive mechanism of de novo methylation in cancer cells. *Nat. Genet.* **38**, 149-153 (2006).

124. Gebhard,C. *et al.* Rapid and sensitive detection of CpG-methylation using methyl-binding (MB)-PCR. *Nucleic Acids Res.* **34**, e82 (2006).

125. Jorgensen,H.F., Adie,K., Chaubert,P. & Bird,A.P. Engineering a high-affinity methyl-CpG-binding protein. *Nucleic Acids Res.* **34**, e96 (2006).

126. Illingworth,R. *et al.* A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS. Biol.* **6**, e22 (2008).

127. Dalma-Weiszhausz,D.D., Warrington,J., Tanimoto,E.Y. & Miyada,C.G. The affymetrix GeneChip platform: an overview. *Methods Enzymol.* **410:3-28.**, 3-28 (2006).

128. Nuwaysir,E.F. *et al.* Gene expression analysis using oligonucleotide arrays produced by maskless photolithography. *Genome Res.* **12**, 1749-1755 (2002).

129. Reinders,J. *et al.* Genome-wide, high-resolution DNA methylation profiling using bisulfite-mediated cytosine conversion. *Genome Res.* **18**, 469-476 (2008).

130. Rauch,T.A. *et al.* High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 252-257 (2008).

131. Yuan,E. *et al.* A single nucleotide polymorphism chip-based method for combined genetic and epigenetic profiling: validation in decitabine therapy and tumor/normal comparisons. *Cancer Res.* **66**, 3443-3451 (2006).

132. Fan,J.B. *et al.* Illumina universal bead arrays. *Methods Enzymol.* **410:57-73.**, 57-73 (2006).

133. Ausubel F.M., Brent R. & Kington RE. Current Protocols in Molecular Biology. John Wiley & Sons, Inc., (2006).

134. Sambrook J. & Russel DW. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, (2001).

135. Mullis,K. *et al.* Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* **51 Pt 1:263-73.**, 263-273 (1986).

136. Saiki,R.K. *et al.* Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science.* **239**, 487-491 (1988).

137. Jaenisch,R. & Bird,A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33 Suppl:245-54.**, 245-254 (2003).

138. Jurka,J. Repbase update: a database and an electronic journal of repetitive elements. *Trends Genet.* **16**, 418-420 (2000).

139. Jurka,J. *et al.* Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* **110**, 462-467 (2005).

140. Skaletsky,H. *et al.* The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **423**, 825-837 (2003).

141. De Smet,C., Loriot,A. & Boon,T. Promoter-dependent mechanism leading to selective hypomethylation within the 5' region of gene MAGE-A1 in tumor cells. *Mol. Cell Biol.* **24**, 4781-4790 (2004).

142. Saito,Y. *et al.* Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell* **9**, 435-443 (2006).

143. Matarazzo,M.R. *et al.* Allelic inactivation of the pseudoautosomal gene SYBL1 is controlled by epigenetic mechanisms common to the X and Y chromosomes. *Hum. Mol. Genet.* **11**, 3191-3198 (2002).

144. Song,F. *et al.* Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proc. Natl. Acad. Sci. U. S. A* **102**, 3336-3341 (2005).

145. Hisano,M., Ohta,H., Nishimune,Y. & Nozaki,M. Methylation of CpG dinucleotides in the open reading frame of a testicular germ cell-specific intronless gene, Tact1/Actl7b, represses its expression in somatic cells. *Nucleic Acids Res.* **31**, 4797-4804 (2003).

146. Koslowski,M., Sahin,U., Huber,C. & Tureci,O. The human X chromosome is enriched for germline genes expressed in premeiotic germ cells of both sexes. *Hum. Mol. Genet.* **15**, 2392-2399 (2006).

147. De Smet,C., Lurquin,C., Lethe,B., Martelange,V. & Boon,T. DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. *Mol. Cell Biol.* **19**, 7327-7335 (1999).

148. Gessner,A., Blum,H. & Rollinghoff,M. Differential regulation of IL-9-expression after infection with Leishmania major in susceptible and resistant mice. *Immunobiology.* **189**, 419-435 (1993).

149. Locksley,R.M., Heinzel,F.P., Sadick,M.D., Holaday,B.J. & Gardner,K.D., Jr. Murine cutaneous leishmaniasis: susceptibility correlates with differential expansion of helper T-cell subsets. *Ann. Inst. Pasteur Immunol.* **138**, 744-749 (1987).

150. She,X., Cheng,Z., Zollner,S., Church,D.M. & Eichler,E.E. Mouse segmental duplication and copy number variation. *Nat. Genet.* **40**, 909-914 (2008).

151. Frazer,K.A. *et al.* A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. *Nature.* **448**, 1050-1053 (2007).

152. Yang,H., Bell,T.A., Churchill,G.A. & Pardo-Manuel,d., V. On the subspecific origin of the laboratory mouse. *Nat. Genet.* **39**, 1100-1107 (2007).

153. Eckhardt,F. *et al.* DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat. Genet.* **38**, 1378-1385 (2006).

154. Rollins,R.A. *et al.* Large-scale structure of genomic methylation patterns. *Genome Res.* **16**, 157-163 (2006).

155. Hatada,I. *et al.* Genome-wide profiling of promoter methylation in human. *Oncogene.* **25**, 3059-3064 (2006).

156. Taylor,K.H. *et al.* Ultradeep bisulfite sequencing analysis of DNA methylation patterns in multiple gene promoters by 454 sequencing. *Cancer Res.* **67**, 8511-8518 (2007).

157. Zilberman,D. & Henikoff,S. Genome-wide analysis of DNA methylation patterns. *Development.* **134**, 3959-3965 (2007).

158. Korshunova,Y. *et al.* Massively parallel bisulphite pyrosequencing reveals the molecular complexity of breast cancer-associated cytosine-methylation patterns obtained from tissue and serum DNA. *Genome Res.* **18**, 19-29 (2008).

159. Schmidl,C. *et al.* Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity. *Genome Res.* **19**, 1165-1174 (2009).

160. Oostlander,A.E., Meijer,G.A. & Ylstra,B. Microarray-based comparative genomic hybridization and its applications in human genetics. *Clin. Genet.* **66**, 488-495 (2004).

161. Albertson,D.G. & Pinkel,D. Genomic microarrays in human genetic disease and cancer. *Hum. Mol. Genet.* **12 Spec No 2:R145-52. Epub;%2003 Aug 5.**, R145-R152 (2003).

162. Riggs,A.D. X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet.* **14**, 9-25 (1975).

163. Holliday,R. & Pugh,J.E. DNA modification mechanisms and gene activity during development. *Science.* **187**, 226-232 (1975).

164. De Bustos,C. *et al.* Tissue-specific variation in DNA methylation levels along human chromosome 1. *Epigenetics. Chromatin.* **2**, 7 (2009).

165. Rakyan,V.K. *et al.* An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Res.* **18**, 1518-1529 (2008).

166. Song,F. *et al.* Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 3336-3341 (2005).

167. Song,F. *et al.* Tissue specific differentially methylated regions (TDMR): Changes in DNA methylation during development. *Genomics.* **93**, 130-139 (2009).

168. Ivascu,C. *et al.* DNA methylation profiling of transcription factor genes in normal lymphocyte development and lymphomas. *Int. J. Biochem. Cell Biol.* **39**, 1523-1538 (2007).

169. Straussman,R. *et al.* Developmental programming of CpG island methylation profiles in the human genome. *Nat. Struct. Mol. Biol.* **16**, 564-571 (2009).

170. Bernstein,B.E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell.* **125**, 315-326 (2006).

171. Shen,L. *et al.* Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. *PLoS. Genet.* **3**, 2023-2036 (2007).

172. Meissner,A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature.* **454**, 766-770 (2008).

173. De Smet,C. *et al.* The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 7149-7153 (1996).

174. Ortmann,C.A. *et al.* Aberrant hypomethylation of the cancer-testis antigen PRAME correlates with PRAME expression in acute myeloid leukemia. *Ann. Hematol.* **87**, 809-818 (2008).

175. Koslowski,M. *et al.* Frequent nonrandom activation of germ-line genes in human cancer. *Cancer Res.* **64**, 5988-5993 (2004).

176. Hellmann-Blumberg,U., Hintz,M.F., Gatewood,J.M. & Schmid,C.W. Developmental differences in methylation of human Alu repeats. *Mol. Cell Biol.* **13**, 4523-4530 (1993).

177. Rubin,C.M., VandeVoort,C.A., Teplitz,R.L. & Schmid,C.W. Alu repeated DNAs are differentially methylated in primate germ cells. *Nucleic Acids Res.* **22**, 5121-5127 (1994).

178. Oei,S.L. *et al.* Clusters of regulatory signals for RNA polymerase II transcription associated with Alu family repeats and CpG islands in human promoters. *Genomics.* **83**, 873-882 (2004).

179. Xie,H. *et al.* High-throughput sequence-based epigenomic analysis of Alu repeats in human cerebellum. *Nucleic Acids Res.* **%20.**, (2009).

180. Jurka,J., Krnjajic,M., Kapitonov,V.V., Stenger,J.E. & Kokhanyy,O. Active Alu elements are passed primarily through paternal germlines. *Theor. Popul. Biol.* **61**, 519-530 (2002).

181. Zhang,Y., Shahriar,M., Zhang,J., Ahmed,S.U. & Lim,S.H. Clofarabine induces hypomethylation of DNA and expression of Cancer-Testis antigens. *Leuk. Res.* (2009).

182. Oakes,C.C., La Salle,S., Smiraglia,D.J., Robaire,B. & Trasler,J.M. Developmental acquisition of genome-wide DNA methylation occurs prior to meiosis in male germ cells. *Dev. Biol.* **307**, 368-379 (2007).

183. Chesnokov,I.N. & Schmid,C.W. Specific Alu binding protein from human sperm chromatin prevents DNA methylation. *J. Biol. Chem.* **270**, 18539-18542 (1995).

184. Lippman,Z. *et al.* Role of transposable elements in heterochromatin and epigenetic control. *Nature.* **430**, 471-476 (2004).

185. Huber,R. *et al.* DNA methylation in transcriptional repression of two differentially expressed X-linked genes, GPC3 and SYBL1. *Proc. Natl. Acad. Sci. U. S. A.* **%19;96**, 616-621 (1999).

186. Bock,C. *et al.* CpG island methylation in human lymphocytes is highly correlated with DNA sequence, repeats, and predicted DNA structure. *PLoS. Genet.* **2**, e26 (2006).

187. Feltus,F.A., Lee,E.K., Costello,J.F., Plass,C. & Vertino,P.M. Predicting aberrant CpG island methylation. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 12253-12258 (2003).

188. Feltus,F.A., Lee,E.K., Costello,J.F., Plass,C. & Vertino,P.M. DNA motifs associated with aberrant CpG island methylation. *Genomics.* **87**, 572-579 (2006).

189. Bell,A.C., West,A.G. & Felsenfeld,G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell.* **98**, 387-396 (1999).

190. Loukinov,D.I. *et al.* BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. *Proc. Natl. Acad. Sci. U. S. A* **99**, 6806-6811 (2002).

191. Vatolin,S. *et al.* Conditional expression of the CTCF-paralogous transcriptional factor BORIS in normal cells results in demethylation and derepression of MAGE-A1 and reactivation of other cancer-testis genes. *Cancer Res.* **65**, 7751-7762 (2005).

192. Hong,J.A. *et al.* Reciprocal binding of CTCF and BORIS to the NY-ESO-1 promoter coincides with derepression of this cancer-testis gene in lung cancer cells. *Cancer Res.* **65**, 7763-7774 (2005).

193. Pohlers,M. *et al.* A role for E2F6 in the restriction of male-germ-cell-specific gene expression. *Curr. Biol.* **15**, 1051-1057 (2005).

194. Storre,J. *et al.* Silencing of the meiotic genes SMC1beta and STAG3 in somatic cells by E2F6. *J. Biol. Chem.* **280**, 41380-41386 (2005).

195. Kehoe,S.M. *et al.* A conserved E2F6-binding element in murine meiosis-specific gene promoters. *Biol. Reprod.* **79**, 921-930 (2008).

196. Rakyan,V.K. *et al.* Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 2538-2543 (2003).

197. Rassoulzadegan,M. *et al.* RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature.* **441**, 469-474 (2006).

198. Chandler,L.A., Ghazi,H., Jones,P.A., Boukamp,P. & Fusenig,N.E. Allele-specific methylation of the human c-Ha-ras-1 gene. *Cell.* **50**, 711-717 (1987).

199. Silva,A.J. & White,R. Inheritance of allelic blueprints for methylation patterns. *Cell.* **54**, 145-152 (1988).

200. Yamada,Y. *et al.* A comprehensive analysis of allelic methylation status of CpG islands on human chromosome 21q. *Genome Res.* **14**, 247-266 (2004).

201. Cheers,C. *et al.* Production of colony-stimulating factors (CSFs) during infection: separate determinations of macrophage-, granulocyte-, granulocyte-macrophage-, and multi-CSFs. *Infect. Immun.* **56**, 247-251 (1988).

202. Heinzel,F.P., Rerko,R.M. & Hujer,A.M. Underproduction of interleukin-12 in susceptible mice during progressive leishmaniasis is due to decreased CD40 activity. *Cell Immunol.* **184**, 129-142 (1998).

203. Scharton,T.M. & Scott,P. Natural killer cells are a source of interferon gamma that drives differentiation of CD4+ T cell subsets and induces early resistance to Leishmania major in mice. *J. Exp. Med.* **178**, 567-577 (1993).

204. Autenrieth,I.B., Beer,M., Bohn,E., Kaufmann,S.H. & Heesemann,J. Immune responses to Yersinia enterocolitica in susceptible BALB/c and resistant C57BL/6 mice: an essential role for gamma interferon. *Infect. Immun.* **62**, 2590-2599 (1994).

205. Bohn,E., Heesemann,J., Ehlers,S. & Autenrieth,I.B. Early gamma interferon mRNA expression is associated with resistance of mice against Yersinia enterocolitica. *Infect. Immun.* **62**, 3027-3032 (1994).

206. Appelberg,R. *et al.* Role of gamma interferon and tumor necrosis factor alpha during T-cell-independent and -dependent phases of Mycobacterium avium infection. *Infect. Immun.* **62**, 3962-3971 (1994).

207. Brett,S.J. & Butler,R. Resistance to Mycobacterium lepraemurium is correlated with the capacity to generate macrophage activating factor(s) in response to mycobacterial antigens in vitro. *Immunology.* **59**, 339-345 (1986).

208. Laskay,T., Diefenbach,A., Rollinghoff,M. & Solbach,W. Early parasite containment is decisive for resistance to Leishmania major infection. *Eur. J. Immunol.* **25**, 2220-2227 (1995).

209. Scott,P., Eaton,A., Gause,W.C., di,Z., X & Hondowicz,B. Early IL-4 production does not predict susceptibility to Leishmania major. *Exp. Parasitol.* **84**, 178-187 (1996).

210. Sacks,D. & Noben-Trauth,N. The immunology of susceptibility and resistance to Leishmania major in mice. *Nat. Rev. Immunol.* **2**, 845-858 (2002).

211. von Stebut,E. & Udey,M.C. Requirements for Th1-dependent immunity against infection with Leishmania major. *Microbes. Infect.* **6**, 1102-1109 (2004).

212. Foote,S.J. & Handman,E. Genetics of murine leishmaniasis. *Brief. Funct. Genomic. Proteomic.* **4**, 270-276 (2005).

213. Wade,C.M. *et al.* The mosaic structure of variation in the laboratory mouse genome. *Nature.* **420**, 574-578 (2002).

214. Frazer,K.A. *et al.* Segmental phylogenetic relationships of inbred mouse strains revealed by fine-scale analysis of sequence variation across 4.6 mb of mouse genome. *Genome Res.* **14**, 1493-1500 (2004).

215. Yalcin,B. *et al.* Unexpected complexity in the haplotypes of commonly used inbred strains of laboratory mice. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9734-9739 (2004).

216. Harr,B. Genomic islands of differentiation between house mouse subspecies. *Genome Res.* **16**, 730-737 (2006).

217. Boursot,P. & Belkhir,K. Mouse SNPs for evolutionary biology: beware of ascertainment biases. *Genome Res.* **16**, 1191-1192 (2006).

218. Zhang,J. *et al.* A high-resolution multistrain haplotype analysis of laboratory mouse genome reveals three distinctive genetic variation patterns. *Genome Res.* **15**, 241-249 (2005).

219. Ideraabdullah,F.Y. *et al.* Genetic and haplotype diversity among wild-derived mouse inbred strains. *Genome Res.* **14**, 1880-1887 (2004).

220. Wilson,M.D. *et al.* Species-specific transcription in mice carrying human chromosome 21. *Science.* **322**, 434-438 (2008).

221. Zhang,Y. *et al.* DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. *PLoS. Genet.* **5**, e1000438 (2009).

222. Ashe,A. *et al.* A genome-wide screen for modifiers of transgene variegation identifies genes with critical roles in development. *Genome Biol.* **9**, R182 (2008).

223. Kadota,M. *et al.* Allele-specific chromatin immunoprecipitation studies show genetic influence on chromatin state in human genome. *PLoS. Genet.* **3**, e81 (2007).

224. Yan,H., Yuan,W., Velculescu,V.E., Vogelstein,B. & Kinzler,K.W. Allelic variation in human gene expression. *Science.* **297**, 1143 (2002).

225. Gretarsdottir,S. *et al.* The gene encoding phosphodiesterase 4D confers risk of ischemic stroke. *Nat. Genet.* **35**, 131-138 (2003).

226. Horikawa,Y. *et al.* Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nat. Genet.* **26**, 163-175 (2000).

227. Feinberg,A.P. & Vogelstein,B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature.* **301**, 89-92 (1983).

228. Feinberg,A.P. & Vogelstein,B. Hypomethylation of ras oncogenes in primary human cancers. *Biochem. Biophys. Res. Commun.* **111**, 47-54 (1983).

229. Baylin,S.B. *et al.* DNA methylation patterns of the calcitonin gene in human lung cancers and lymphomas. *Cancer Res.* **46**, 2917-2922 (1986).

230. Greger,V., Passarge,E., Hopping,W., Messmer,E. & Horsthemke,B. Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. *Hum. Genet.* **83**, 155-158 (1989).

231. Sakai,T. *et al.* Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. *Am. J. Hum. Genet.* **48**, 880-888 (1991).

232. Milani,L. *et al.* Allele-specific gene expression patterns in primary leukemic cells reveal regulation of gene expression by CpG site methylation. *Genome Res.* **19**, 1-11 (2009).

# 10 Abbreviations

| | |
|---|---|
| °C | Grad Celsius |
| µg | Microgram |
| µl | Microliter |
| 5mC | 5-methylcytosine |
| A | Adenine |
| AA | Acrylamide |
| ALL | Acute lymphoblastic leukemia |
| APS | Ammonium persulfate |
| AS | Antisense |
| ASE | Allele-specific expression |
| BMM | Bone-marrow derived macrophages |
| bp | Base pair |
| BSA | Bovine serum albumine |
| C | Cytosine |
| CAP | CxxC affinity purification |
| cDNA | Copy DNA |
| CGH | Comparative genome hybridisation |
| CGI | CpG-island |
| CIAP | Calf intestinal alkaline phosphatase |
| CNV | Copy number variation |
| CpG | Cytosine/Guanine dinucleotide |
| cRNA | Copy RNA |
| CTCF | CCCTC-binding factor |
| C-terminal | Carboxyterminal |
| (d) CTP | (deoxy) Cytosine triphosphate |
| DEPC | Diethylpyrocarbonate |
| DMF | Dimethylformamide |
| DMR | Differentially methylated region |
| DNA | Deoxyribonucleic acid |
| Dnmt | DNA-methyltransferase |
| dNTP | Deoxyribonucleotide |
| EDTA | Ethylenediaminetetraacetic acid |
| EtBr | Ethydiumbromide |
| F1 | F1 generation |
| G | Guanine |
| g | Rotational force or gram |
| GC | Guanine/Cytosine |

| | |
|---|---|
| gDNA | Genomic DNA |
| GE | Gene expression |
| GO | Gene ontology |
| h | Hour |
| HELP | HpaII-tiny fragment Enrichment by Ligation-mediated PCR |
| hg | Human genome |
| i.a. | Inter alia |
| ICR | Inprinting control region |
| IP | Immunopreciptitation |
| k | Kilo |
| kb | Kilo base |
| kDa | Kilo Dalton |
| l | Liter |
| LB | Lysogeny broth |
| M | Molar |
| M1/M2 | classically/alternatively activated macrophage |
| mA | Milli Ampere |
| MALDI_TOF MS | Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry |
| MBD | Methyl-binding domain |
| MB-PCR | Methyl-binding PCR |
| MCIp | Methyl-CpG-immunoprecipitation |
| mCpG | Methylated CpG dinucleotide |
| MeCP2 | Methyl CpG binding protein 2 |
| MeDIP | Methylated DNA immunoprecipitation |
| mg | Milligram |
| min | Minute |
| ml | Milliliter |
| mM | Millimolar |
| mm | Mus musculus |
| MNC | Mononuclear cells |
| mRNA | Messenger RNA |
| MS | Mass spectrometry |
| MSP | Methyl-specific PCR |
| n | Number |
| NaCL | Sodium chloride |
| $NaHSO_3$ | Sodium bisulfite |
| NCBI | National Center for Biotechnology Information |
| ng | Nanogram |
| NGS | Next generation sequencing |
| nos | Numbers |
| N-ternimal | Aminoternimal |
| OD | Optical density (Absorbance) |

| | |
|---|---|
| PAGE | Polyacrylamide gel electrophoresis |
| PBS | Phosphate buffered saline |
| PCR | Polymerase chain reaction |
| PEG | Polyethylene glycol |
| ps | Pseudogene |
| qPCR | Quantitative PCR |
| rCSF | Recombinant colony stimulating factor |
| RLGS | Restriction landmark genomic scanning |
| RNA | Ribonucleic acid |
| rpm | Rotations per minute |
| RT | Room temperature |
| RT-PCR | Reverse transcription PCR |
| s | Second |
| S | Sense |
| SAM | S-Adenosyl-Methionine |
| SAP | Shrimp Alkaline Phosphatase |
| SD | Standard diviation |
| SDS | Sodium dodecyl sulfate |
| SINE | Short interspersed element |
| SNP | Single nucleotide polymorphism |
| SOC | Rich bacterial growth media |
| SSPE | Sodium chloride, sodium phosphate, EDTA buffer |
| T | Thymine |
| TAE | Tris-acetate buffer with EDTA |
| TBS | Tris buffered saline |
| TE | Tris EDTA buffer |
| TEMED | Tetramethylethylenediamine |
| TFBI/II | Media for competent cells |
| Th1/Th2 | Type 1/Type 2 helper T cell |
| Tm | Melting temperature |
| TRD | Transcriptional repression domain |
| U | Unit or Uracil |
| UCSC | University of California, Santa Cruz |
| UPD | Uniparental contribution of one chromosome pair |
| vCGH | Virtual comparative genome hybridisation |
| w/v | Weigth per volume |
| Xic | X-inactivation center |

# 11 List of publications

**Accepted publications:**

- Silversmith,R.E., Levin,M.D., **Schilling,E.**, & Bourret,R.B. Kinetic characterization of catalysis by the chemotaxis phosphatase CheZ. Modulation of activity by the phosphorylated CheY substrate. *J. Biol. Chem*. **283**, 756-765 (2008).

- **Schilling,E**. & Rehli,M. Global, comparative analysis of tissue-specific promoter CpG methylation. *Genomics*. **90**, 314-323 (2007).

- Gebhard,C., Schwarzfischer,L., Pham,T.H., **Schilling,E.**, Klug,M., Andreesen,R., & Rehli,M. Genome-wide profiling of CpG methylation identifies novel targets of aberrant hypermethylation in myeloid leukemia. *Cancer Res*. **66**, 6118-6128 (2006).

**Submitted publications:**

- **Schilling,E.**, El-Chartouni,C, Rehli,M. Allele-specific DNA methylation in mouse strains is mainly determined by *cis*-acting sequences.

- Gebhard,C. Benner,C. Ehrich,M. Schwarzfischer,L. **Schilling,E.** Klug,M. Dietmeier,W. Thiede,C. Holler,E. Andreesen,R. Rehli,M. DNA methylation states of CpG islands in normal and in cancer cells correlate with defined *cis*-acting sequences and combinatorial transcription factor binding.

# 12 Appendix

**Appendix I – Tissue-specific methylated promoters (section 12.1; page 154)**

**Table 12.1** contains all tissue-specific differentially methylated promoters. Genomic locations are based on the Build 35 assembly by NCBI and the Human Genome Sequencing Consortium (hg17). The distance of probes to transcription start sites (TSS) of corresponding genes is indicated as well as the CpG index (+/- 350 bp). The tissue where this promoter is demethylated is shown in the last column.

**Appendix II – Differentially expressed loci (section 12.2; page 173)**

**Table 12.2** lists all gene loci that were selected for the locus-wide tiling array. Genomic locations are based on the Build 36 assembly by NCBI and the Mouse Genome Sequencing Consortium (mm8). Randomly chosen control loci are boxed in light green.

**Appendix III – Annotated list of mouse strain-specific DMR (section 12.3; page 177)**

**Table 12.3** contains the complete list of mouse strain-specific DMR discovered in this study. The genomic location (Build 36 assembly by NCBI, mm8) of the 'core' region of the DMR is given along with information on the associated (neighboring) gene, its position relative to this gene, its methylation status and BMM expression data (if available) for the associated gene. Microarray expression data was median-normalized to the C57BL/6 BMM sample (data for 4h and 18h treatments is not shown). The call status of individual probes in each sample is given in brackets behind the relative expression values (P, present; A, absent; M, marginal).

**Appendix IV – sequence alignments (section 12.4; page 187)**

This supplement contains DNA sequence alignments of 34 selected regions in C57BL/6 and BALB/c. Regions were randomly selected from the vCGH data (A-I), or as possible targets for methylation analysis using MALDI-TOF MS (J-AH). Genomic DNA was PCR amplified and directly sequenced. All sequences were deposited with GenBank (accession numbers: FJ751937-FJ752004). Differences between BALB/c and the reference strain C57BL/6 are marked as follows: deletions or insertions (relative to C57BL/6) are in red lettering; single nucleotide polymorphisms are boxed in blue (BALB/c). Variations between the published reference sequence and the C57BL/6 sequence used in this study were only observed in a fragment that represents a pseudogene or duplication of the Sfi1 locus (in AC). Genomic locations are based on the Build 36 assembly by NCBI and the Mouse Genome Sequencing Consortium (mm8).

# 12.1 Appendix I – Tissue-specific methylated promoters

**Table 12.1** Microarray-detected differentially methylated gene promoter

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr13:113156292-113156352 | ADPRHL1 | -482 | 29 | brain |
| chr15:027001096-027001156 | APBA2 | -18 | 17 | brain |
| chr15:027001215-027001265 | APBA2 | 96 | 16 | brain |
| chr18:053550053-053550112 | ATP8B1 | -45 | 11 | brain |
| chr1:006380184-006380236 | BACH | -71 | 23 | brain |
| chr19:010471818-010471874 | BC002417 | -107 | 16 | brain |
| chr12:049926071-049926129 | BC003379 | 61 | 12 | brain |
| chrX:141018179-141018224 | BC005891 | -521 | 21 | brain |
| chr20:018070821-018070881 | BC007537 | -112 | 11 | brain |
| chr1:001204306-001204358 | BC007625 | -158 | 32 | brain |
| chr12:127824516-127824576 | BC009492 | -371 | 22 | brain |
| chr16:083581862-083581922 | BC017702 | -453 | 25 | brain |
| chr9:094602320-094602379 | BC020194 | 75 | 12 | brain |
| chr18:013672133-013672193 | BC022410 | -57 | 9 | brain |
| chr10:050401989-050402049 | BC028954 | 130 | 11 | brain |
| chr7:112018428-112018488 | BC030804 | -262 | 8 | brain |
| chr7:112018654-112018714 | BC030804 | -488 | 9 | brain |
| chr21:041670735-041670795 | BC035293 | 128 | 15 | brain |
| chr17:007226637-007226691 | BC035782 | -211 | 14 | brain |
| chr14:090821721-090821781 | BC035914 | 8 | 24 | brain |
| chr12:039507992-039508052 | BC036569 | -258 | 13 | brain |
| chr8:001794402-001794449 | BC040474 | 98 | 27 | brain |
| chr8:021966477-021966533 | BC052805 | -596 | 15 | brain |
| chr17:024426164-024426224 | BC058075 | -477 | 21 | brain |
| chr16:002319879-002319939 | BC062779 | -211 | 12 | brain |
| chr2:135822033-135822093 | BC064616 | -123 | 11 | brain |
| chr20:041626880-041626940 | BC065511 | -349 | 17 | brain |
| chr1:200471739-200471799 | BC069650 | 102 | 13 | brain |
| chr14:089112094-089112145 | BC069658 | -245 | 20 | brain |
| chr14:089112268-089112328 | BC069658 | -66 | 16 | brain |
| chr1:032054426-032054486 | BC070182 | -369 | 14 | brain |
| chr14:020580502-020580555 | BC074960 | -434 | 13 | brain |
| chr1:034990667-034990712 | BC093838 | -252 | 18 | brain |
| chr17:019126986-019127046 | BC093972 | 64 | 26 | brain |
| chr17:040238543-040238600 | BC096213 | 140 | 19 | brain |
| chr17:041395305-041395363 | BC098281 | -185 | 15 | brain |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr10:011401728-011401788 | C10orf31 | 95 | 13 | brain |
| chr11:030558663-030558723 | C11orf8 | -77 | 9 | brain |
| chr11:030558830-030558890 | C11orf8 | -244 | 6 | brain |
| chr11:030559065-030559125 | C11orf8 | -479 | 8 | brain |
| chr14:066725463-066725523 | C14orf54 | -405 | 11 | brain |
| chr5:034079089-034079144 | C1QTNF3 | -42 | 12 | brain |
| chr22:042144979-042145035 | C22orf1 | -401 | 14 | brain |
| chr8:011025219-011025279 | C8orf16 | -94 | 13 | brain |
| chr1:019491575-019491635 | CAPZB | -54 | 15 | brain |
| chr22:048538819-048538879 | CCDS14080.1 | -23 | 26 | brain |
| chr1:010443963-010444008 | CORT | -251 | 20 | brain |
| chr1:010444106-010444166 | CORT | -100 | 18 | brain |
| chr1:010444380-010444438 | CORT | 173 | 22 | brain |
| chr22:023919940-023919992 | CRYBB3 | -412 | 14 | brain |
| chr10:126706348-126706408 | CTBP2 | 65 | 25 | brain |
| chr10:035934088-035934148 | CX40.1 | -225 | 14 | brain |
| chrX:019744390-019744450 | CXorf23 | -339 | 9 | brain |
| chr2:228323317-228323377 | DKFZp547H025 | 46 | 9 | brain |
| chr8:001436591-001436651 | DLGAP2 | -354 | 14 | brain |
| chr8:001436801-001436854 | DLGAP2 | -148 | 23 | brain |
| chr20:001224333-001224385 | ENST00000217276 | -58 | 24 | brain |
| chr16:012516793-012516846 | ENST00000219090 | -68 | 14 | brain |
| chr2:105383785-105383845 | ENST00000258449 | -519 | 13 | brain |
| chr16:000337079-000337137 | ENST00000262320 | 2 | 19 | brain |
| chr16:000337208-000337268 | ENST00000262320 | -128 | 18 | brain |
| chrX:110171873-110171933 | ENST00000262836 | -573 | 6 | brain |
| chrX:110172089-110172149 | ENST00000262836 | -357 | 6 | brain |
| chr3:102987196-102987254 | ENST00000273347 | 95 | 11 | brain |
| chr5:138238070-138238126 | ENST00000274711 | 47 | 9 | brain |
| chr5:138238643-138238699 | ENST00000274711 | -526 | 16 | brain |
| chr12:001007477-001007535 | ENST00000299183 | 176 | 16 | brain |
| chr16:015640655-015640715 | ENST00000300014 | -32 | 14 | brain |
| chr18:070475055-070475115 | ENST00000309902 | -246 | 12 | brain |
| chr12:016644843-016644903 | ENST00000320122 | 188 | 8 | brain |
| chr16:011677271-011677316 | ENST00000329565 | -37 | 31 | brain |
| chr15:089569526-089569586 | ENST00000330276 | -547 | 11 | brain |
| chr15:089569658-089569718 | ENST00000330276 | -415 | 10 | brain |
| chr8:043077295-043077355 | ENST00000331373 | -523 | 11 | brain |
| chr6:012826568-012826628 | ENST00000332995 | 71 | 12 | brain |
| chr10:123770911-123770971 | ENST00000334433 | -499 | 19 | brain |
| chr10:123771372-123771432 | ENST00000334433 | -38 | 20 | brain |
| chr4:154549026-154549086 | ENST00000338700 | -86 | 15 | brain |
| chr2:198773879-198773939 | ENST00000341860 | -132 | 10 | brain |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr2:198774045-198774098 | ENST00000341860 | 30 | 11 | brain |
| chr2:198774190-198774249 | ENST00000341860 | 178 | 12 | brain |
| chr10:123343655-123343715 | ENST00000346997 | -364 | 21 | brain |
| chr6:118994211-118994271 | ENST00000359516 | -1 | 15 | brain |
| chr11:131035896-131035956 | ENST00000360508 | -157 | 13 | brain |
| chr11:131036065-131036125 | ENST00000360508 | -326 | 10 | brain |
| chr1:040591855-040591915 | ENST00000361584 | 83 | 8 | brain |
| chr17:026673154-026673214 | EVI2A | -341 | 8 | brain |
| chr17:077910023-077910083 | FLJ35767 | -359 | 18 | brain |
| chr17:032806792-032806852 | FLJ39647 | -275 | 13 | brain |
| chr17:032806919-032806979 | FLJ39647 | -148 | 8 | brain |
| chr20:014266631-014266691 | FLRT3 | -391 | 12 | brain |
| chr15:089212954-089213014 | FURIN | 96 | 9 | brain |
| chr6:010693856-010693916 | GCNT2 | -94 | 16 | brain |
| chr6:024597559-024597619 | GPLD1 | 189 | 12 | brain |
| chrX:013594808-013594864 | GPM6B | 135 | 14 | brain |
| chrX:013595126-013595186 | GPM6B | -185 | 18 | brain |
| chrX:013595243-013595303 | GPM6B | -302 | 15 | brain |
| chrX:013595385-013595445 | GPM6B | -444 | 11 | brain |
| chr19:052113811-052113857 | GRLF1 | 62 | 17 | brain |
| chr6:114490978-114491038 | HS3ST5 | -306 | 8 | brain |
| chr3:035760481-035760539 | hsa-mir-128b | -462 | 9 | brain |
| chr3:035760665-035760725 | hsa-mir-128b | -277 | 13 | brain |
| chr3:035760825-035760885 | hsa-mir-128b | -117 | 12 | brain |
| chr3:035761037-035761087 | hsa-mir-128b | 90 | 14 | brain |
| chr9:094927591-094927651 | hsa-mir-24-1 | -237 | 35 | brain |
| chr7:130019996-130020056 | hsa-mir-29b-1 | -473 | 12 | brain |
| chr17:076714638-076714698 | hsa-mir-338 | -324 | 34 | brain |
| chr3:054937590-054937641 | HT017 | -503 | 17 | brain |
| chr11:128242463-128242523 | KCNJ1 | -15 | 11 | brain |
| chr17:065612388-065612448 | KCNJ16 | -171 | 8 | brain |
| chr17:065612548-065612600 | KCNJ16 | -15 | 10 | brain |
| chr19:048977667-048977725 | KCNN4 | -447 | 15 | brain |
| chr13:044666706-044666766 | KCTD4 | 116 | 13 | brain |
| chr19:039483115-039483175 | KIAA0354 | -73 | 15 | brain |
| chrX:116888523-116888583 | KLHL13 | 182 | 8 | brain |
| chr4:038886645-038886705 | KLHL5 | 50 | 10 | brain |
| chr21:044817499-044817551 | KRTAP10-4 | -508 | 16 | brain |
| chr21:044817830-044817890 | KRTAP10-4 | -173 | 23 | brain |
| chr12:016652579-016652639 | LMO3 | -318 | 9 | brain |
| chr4:054265227-054265282 | LNX | 41 | 14 | brain |
| chr12:123298434-123298494 | LOC144347 | -125 | 18 | brain |
| chr16:006367728-006367788 | LOC440337 | -6 | 12 | brain |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr1:024473419-024473474 | LOC90529 | -191 | 10 | brain |
| chr5:138238877-138238935 | LRRTM2 | 2 | 16 | brain |
| chr19:051850046-051850102 | MGC15476 | -320 | 22 | brain |
| chr7:077621227-077621287 | MGC34774 | 48 | 7 | brain |
| chr19:014913135-014913195 | OR7C2 | -135 | 12 | brain |
| chr11:123816664-123816719 | OR8B8 | -500 | 14 | brain |
| chr11:123686382-123686442 | OR8D1 | -540 | 7 | brain |
| chr11:057703393-057703453 | OR9Q1 | -69 | 10 | brain |
| chr11:057703574-057703634 | OR9Q1 | 112 | 13 | brain |
| chr11:066478359-066478419 | PC | 67 | 14 | brain |
| chr10:050402588-050402648 | PGBD3 | -286 | 12 | brain |
| chr10:050402894-050402954 | PGBD3 | -592 | 10 | brain |
| chr6:012825817-012825877 | PHACTR1 | 29 | 10 | brain |
| chr20:057729122-057729182 | PHACTR3 | -507 | 15 | brain |
| chr20:057729238-057729298 | PHACTR3 | -391 | 13 | brain |
| chr20:057729743-057729791 | PHACTR3 | 108 | 11 | brain |
| chr5:067557608-067557666 | PIK3R1 | -580 | 18 | brain |
| chr5:067557770-067557830 | PIK3R1 | -417 | 17 | brain |
| chr5:067557898-067557949 | PIK3R1 | -294 | 16 | brain |
| chr5:067558297-067558357 | PIK3R1 | 110 | 13 | brain |
| chr5:146415567-146415627 | PPP2R2B | 74 | 10 | brain |
| chr12:010972575-010972635 | PRH2 | -495 | 7 | brain |
| chr6:150304060-150304115 | RAET1E | 124 | 8 | brain |
| chr16:024765142-024765187 | SLC5A11 | 112 | 17 | brain |
| chr11:066245697-066245757 | SPTBN2 | -281 | 18 | brain |
| chr12:011031222-011031282 | TAS2R50 | -526 | 8 | brain |
| chr15:055299101-055299161 | TCF12 | 185 | 19 | brain |
| chr10:114033350-114033410 | TECTB | -102 | 17 | brain |
| chr20:002224396-002224456 | TGM3 | -248 | 13 | brain |
| chr20:030184552-030184603 | TM9SF4 | 102 | 10 | brain |
| chr9:097366024-097366079 | TMOD1 | 60 | 9 | brain |
| chr2:233559949-233560002 | UNQ830 | -523 | 12 | brain |
| chr9:134034396-134034456 | WDR5 | -41 | 21 | brain |
| chr7:020030099-020030159 | 7A5 | 124 | 5 | monocyte |
| chr7:020030479-020030539 | 7A5 | -256 | 6 | monocyte |
| chr12:021981417-021981477 | ABCC9 | -572 | 7 | monocyte |
| chr11:100367624-100367684 | AD031 | -366 | 6 | monocyte |
| chr1:111758594-111758651 | ADORA3 | -470 | 9 | monocyte |
| chr6:031691579-031691630 | AIF1 | -163 | 12 | monocyte |
| chr6:031691863-031691918 | AIF1 | 123 | 12 | monocyte |
| chr9:072797308-072797368 | ALDH1A1 | 185 | 12 | monocyte |
| chrY:006785869-006785929 | AMELY | -470 | 6 | monocyte |
| chr10:090601872-090601932 | ANKRD22 | -406 | 9 | monocyte |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr7:036537529-036537589 | AOAH | -294 | 7 | monocyte |
| chr11:043289816-043289876 | API5 | -262 | 36 | monocyte |
| chr2:143720458-143720518 | ARHGAP15 | -206 | 7 | monocyte |
| chr2:143720760-143720820 | ARHGAP15 | 96 | 6 | monocyte |
| chrX:011043676-011043733 | ARHGAP6 | 48 | 9 | monocyte |
| chrX:011044130-011044190 | ARHGAP6 | -408 | 7 | monocyte |
| chrX:135589041-135589096 | ARHGEF6 | -45 | 13 | monocyte |
| chrX:135589230-135589279 | ARHGEF6 | -231 | 8 | monocyte |
| chrX:135589371-135589431 | ARHGEF6 | -378 | 6 | monocyte |
| chrX:135589508-135589568 | ARHGEF6 | -515 | 6 | monocyte |
| chr14:075058388-075058448 | BATF | -118 | 15 | monocyte |
| chr18:030427081-030427137 | BC005300 | -210 | 8 | monocyte |
| chr1:150176015-150176069 | BC005928 | 64 | 16 | monocyte |
| chr4:185725651-185725698 | BC015803 | -300 | 21 | monocyte |
| chr17:004284158-004284210 | BC023646 | 169 | 17 | monocyte |
| chr3:100102558-100102603 | BC029658 | 144 | 51 | monocyte |
| chr12:069317856-069317916 | BC051329 | -425 | 5 | monocyte |
| chr12:010953076-010953136 | BC069376 | 58 | 5 | monocyte |
| chr18:065765813-065765860 | BC074787 | -185 | 7 | monocyte |
| chr18:065766196-065766256 | BC074787 | -575 | 5 | monocyte |
| chr6:136214703-136214763 | BC075082 | -70 | 12 | monocyte |
| chr22:021478281-021478341 | BC089414 | -515 | 9 | monocyte |
| chr15:031147428-031147482 | BC103692 | 70 | 23 | monocyte |
| chr15:031148053-031148113 | BC103692 | -558 | 13 | monocyte |
| chr10:098021115-098021175 | BLNK | 171 | 6 | monocyte |
| chr4:068253333-068253393 | BRDG1 | 152 | 6 | monocyte |
| chr20:011846702-011846762 | BTBD3 | 168 | 6 | monocyte |
| chr16:020610453-020610511 | BUCS1 | -403 | 5 | monocyte |
| chr14:057030226-057030286 | C14orf105 | 67 | 6 | monocyte |
| chr14:057030420-057030480 | C14orf105 | -127 | 10 | monocyte |
| chr1:244038040-244038100 | C1orf150 | -421 | 14 | monocyte |
| chr1:111728804-111728864 | C1orf162 | 189 | 16 | monocyte |
| chr20:012937751-012937810 | C20orf38 | -66 | 8 | monocyte |
| chr20:012937957-012938011 | C20orf38 | 138 | 8 | monocyte |
| chr21:033107724-033107784 | C21orf62 | 114 | 8 | monocyte |
| chr13:048873355-048873415 | CAB39L | 115 | 7 | monocyte |
| chr13:048873633-048873685 | CAB39L | -159 | 7 | monocyte |
| chr16:054158083-054158143 | CAPNS2 | 29 | 6 | monocyte |
| chr11:104411424-104411484 | CASP1 | -387 | 5 | monocyte |
| chr2:201873017-201873077 | CASP10 | -313 | 14 | monocyte |
| chr11:104385097-104385157 | CASP5 | -218 | 5 | monocyte |
| chr16:046823116-046823176 | CCDS10732.1 | 187 | 9 | monocyte |
| chr17:031441416-031441463 | CCL3 | 161 | 9 | monocyte |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr17:029669941-029670001 | CCL8 | -207 | 5 | monocyte |
| chr3:032967875-032967935 | CCR4 | -164 | 11 | monocyte |
| chr3:114176584-114176640 | CD200R1 | 15 | 7 | monocyte |
| chr19:040511691-040511744 | CD22 | -234 | 18 | monocyte |
| chr1:157646127-157646187 | CD244 | -439 | 11 | monocyte |
| chr2:204396677-204396737 | CD28 | 4 | 14 | monocyte |
| chr17:070131477-070131537 | CD300LE | -111 | 11 | monocyte |
| chr17:070221099-070221148 | CD300LF | -420 | 18 | monocyte |
| chr1:157494697-157494757 | CD48 | -69 | 9 | monocyte |
| chr1:111127601-111127659 | CD53 | -184 | 10 | monocyte |
| chr1:111127777-111127837 | CD53 | -7 | 10 | monocyte |
| chr12:009804656-009804716 | CD69 | 78 | 5 | monocyte |
| chr12:009805014-009805074 | CD69 | -280 | 7 | monocyte |
| chr3:123279253-123279306 | CD86 | -160 | 8 | monocyte |
| chr3:123279385-123279441 | CD86 | -26 | 9 | monocyte |
| chr1:199930587-199930642 | CHIT1 | -181 | 7 | monocyte |
| chr1:243907508-243907568 | CIAS1 | 147 | 12 | monocyte |
| chr12:009913629-009913689 | CLEC2B | 66 | 11 | monocyte |
| chr12:009913853-009913913 | CLEC2B | -158 | 8 | monocyte |
| chr12:009914031-009914091 | CLEC2B | -336 | 8 | monocyte |
| chr12:008167453-008167513 | CLEC4A | -11 | 6 | monocyte |
| chr12:008557410-008557470 | CLEC4D | 38 | 8 | monocyte |
| chr13:045577337-045577397 | CPB2 | -198 | 5 | monocyte |
| chr14:024115709-024115769 | CTSG | -433 | 7 | monocyte |
| chrX:128639251-128639301 | CXorf9 | -218 | 19 | monocyte |
| chr4:101094606-101094666 | DAPP1 | -531 | 8 | monocyte |
| chrX:110461445-110461503 | DCX | -23 | 6 | monocyte |
| chrX:110461689-110461749 | DCX | -168 | 5 | monocyte |
| chr20:000016400-000016455 | DEFB125 | 65 | 5 | monocyte |
| chrX:031933072-031933121 | DMD | 147 | 6 | monocyte |
| chrX:031933269-031933320 | DMD | -51 | 7 | monocyte |
| chr18:030544384-030544444 | DTNA | 169 | 7 | monocyte |
| chr13:040454715-040454775 | ELF1 | -327 | 8 | monocyte |
| chr4:101796439-101796499 | EMCN | -112 | 5 | monocyte |
| chr19:006838564-006838624 | EMR1 | 13 | 6 | monocyte |
| chrX:030637639-030637699 | ENST00000288422 | -215 | 7 | monocyte |
| chr3:065999103-065999154 | ENST00000295884 | -105 | 60 | monocyte |
| chr12:025096774-025096830 | ENST00000312137 | 38 | 11 | monocyte |
| chr18:064656299-064656359 | ENST00000319445 | -589 | 7 | monocyte |
| chr12:054149628-054149688 | ENST00000327335 | 168 | 5 | monocyte |
| chr3:174804847-174804907 | ENST00000360119 | -213 | 6 | monocyte |
| chr19:060077475-060077534 | FCAR | 144 | 11 | monocyte |
| chr1:158364610-158364658 | FCGR2C | -66 | 6 | monocyte |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr2:150541171-150541231 | FLJ32955 | 12 | 7 | monocyte |
| chr15:036776061-036776113 | FLJ35695 | -3 | 6 | monocyte |
| chr20:014266338-014266394 | FLRT3 | -96 | 9 | monocyte |
| chr7:113648765-113648825 | FOXP2 | -431 | 7 | monocyte |
| chr19:056947246-056947293 | FPR1 | -307 | 5 | monocyte |
| chr5:161426705-161426765 | GABRG2 | -559 | 9 | monocyte |
| chr11:022652630-022652690 | GAS2 | -270 | 5 | monocyte |
| chr3:113334829-113334875 | GCET2 | -45 | 7 | monocyte |
| chr3:113334974-113335034 | GCET2 | -197 | 7 | monocyte |
| chr7:149702231-149702290 | GIMAP4 | 155 | 13 | monocyte |
| chr12:074160760-074160820 | GLIPR1 | -10 | 14 | monocyte |
| chr14:104602606-104602660 | GPR132 | 166 | 15 | monocyte |
| chr14:104602823-104602876 | GPR132 | -50 | 17 | monocyte |
| chrX:078232777-078232837 | GPR174 | 194 | 8 | monocyte |
| chr13:098708452-098708511 | GPR18 | 148 | 7 | monocyte |
| chr19:040631908-040631954 | GPR43 | -525 | 21 | monocyte |
| chr5:054433851-054433911 | GZMA | -348 | 7 | monocyte |
| chr3:122862799-122862859 | HCLS1 | -424 | 6 | monocyte |
| chr12:053177919-053177979 | HEM1 | 170 | 6 | monocyte |
| chr17:053763603-053763663 | hsa-mir-142 | 45 | 25 | monocyte |
| chr21:025867811-025867871 | hsa-mir-155 | -322 | 7 | monocyte |
| chr21:025868007-025868067 | hsa-mir-155 | -126 | 6 | monocyte |
| chrX:065021531-065021588 | hsa-mir-223 | -174 | 6 | monocyte |
| chr7:136044799-136044859 | hsa-mir-490 | -340 | 6 | monocyte |
| chrX:138731834-138731886 | hsa-mir-505 | 50 | 5 | monocyte |
| chrX:138731949-138732009 | hsa-mir-505 | -69 | 5 | monocyte |
| chr19:010311120-010311167 | ICAM3 | 157 | 19 | monocyte |
| chr16:021571336-021571392 | IGSF6 | 100 | 8 | monocyte |
| chr19:018058713-018058773 | IL12RB1 | -46 | 15 | monocyte |
| chr2:113601343-113601394 | IL1RN | 0 | 10 | monocyte |
| chr16:028425706-028425755 | IL27 | -74 | 17 | monocyte |
| chrX:070114605-070114665 | IL2RG | -211 | 10 | monocyte |
| chrX:070114819-070114873 | IL2RG | -422 | 9 | monocyte |
| chr2:218815979-218816033 | IL8RB | -245 | 10 | monocyte |
| chr21:034806400-034806460 | KCNE1 | 13 | 14 | monocyte |
| chrX:086578760-086578815 | KLHL4 | -174 | 9 | monocyte |
| chrX:086578936-086578995 | KLHL4 | 4 | 10 | monocyte |
| chr3:184755993-184756047 | KLHL6 | 90 | 16 | monocyte |
| chr19:059568434-059568489 | LAIR1 | -181 | 18 | monocyte |
| chr19:059705722-059705776 | LAIR2 | -75 | 17 | monocyte |
| chr13:045654201-045654261 | LCP1 | 164 | 11 | monocyte |
| chr4:159800755-159800815 | LGR7 | 52 | 10 | monocyte |
| chr19:059777016-059777064 | LILRA2 | -30 | 13 | monocyte |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr11:033870368-033870419 | LMO2 | 19 | 8 | monocyte |
| chr12:053534187-053534247 | LOC118430 | -370 | 6 | monocyte |
| chr19:007647597-007647647 | LOC199675 | 109 | 17 | monocyte |
| chr7:092493734-092493794 | LOC253012 | -243 | 5 | monocyte |
| chr12:090007918-090007978 | LUM | 62 | 7 | monocyte |
| chr6:006533842-006533894 | LY86 | -64 | 14 | monocyte |
| chrX:138450602-138450662 | MCF2 | -260 | 6 | monocyte |
| chr12:045896156-045896216 | MGC16044 | -132 | 14 | monocyte |
| chr12:045896307-045896367 | MGC16044 | 19 | 14 | monocyte |
| chr4:109171932-109171992 | MGC26963 | -71 | 6 | monocyte |
| chr2:232651796-232651841 | MGC42174 | -80 | 70 | monocyte |
| chr12:014929893-014929953 | MGP | 172 | 7 | monocyte |
| chr12:014930158-014930218 | MGP | -93 | 7 | monocyte |
| chr12:014930491-014930551 | MGP | -426 | 8 | monocyte |
| chr4:010362701-010362752 | MIST | 32 | 9 | monocyte |
| chr1:046771745-046771805 | MKNK1 | 63 | 14 | monocyte |
| chr11:102156587-102156647 | MMP10 | -63 | 5 | monocyte |
| chr4:091173153-091173212 | MMRN1 | -47 | 7 | monocyte |
| chr11:019038698-019038758 | MRGPRX2 | 76 | 7 | monocyte |
| chr11:059580353-059580411 | MS4A3 | -294 | 9 | monocyte |
| chr11:059580782-059580842 | MS4A3 | 136 | 10 | monocyte |
| chr11:059804592-059804648 | MS4A4A | -94 | 6 | monocyte |
| chr11:059707157-059707217 | MS4A6A | 63 | 5 | monocyte |
| chr8:016094406-016094466 | MSR1 | 159 | 5 | monocyte |
| chr17:005428634-005428680 | NALP1 | -107 | 13 | monocyte |
| chr19:059019379-059019433 | NALP12 | -1 | 22 | monocyte |
| chr16:003567563-003567608 | NOD3 | -295 | 23 | monocyte |
| chr15:074091620-074091680 | NRG4 | 190 | 10 | monocyte |
| chr15:074092059-074092119 | NRG4 | -249 | 7 | monocyte |
| chr11:059920331-059920391 | NYD-SP21 | -194 | 6 | monocyte |
| chr11:007917489-007917547 | OR10A3 | 125 | 6 | monocyte |
| chr11:006870334-006870394 | OR2D2 | -57 | 7 | monocyte |
| chr11:055096329-055096389 | OR4C16 | 180 | 6 | monocyte |
| chr15:100163532-100163587 | OR4F6 | 114 | 6 | monocyte |
| chr14:019598158-019598212 | OR4L1 | 142 | 5 | monocyte |
| chr11:004748193-004748253 | OR51F1 | -500 | 5 | monocyte |
| chr11:056137660-056137720 | OR5M1 | -136 | 5 | monocyte |
| chr11:123695081-123695141 | OR8D2 | 192 | 5 | monocyte |
| chr11:007684730-007684790 | OVCH2 | -243 | 6 | monocyte |
| chrY:001699759-001699804 | P2RY8 | 46 | 35 | monocyte |
| chr6:136214016-136214076 | PDE7B | -480 | 5 | monocyte |
| chr7:106099744-106099804 | PIK3CG | -100 | 12 | monocyte |
| chr7:106099890-106099950 | PIK3CG | 46 | 13 | monocyte |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr2:068503807-068503867 | PLEK | -237 | 8 | monocyte |
| chr13:037070784-037070844 | POSTN | 60 | 6 | monocyte |
| chr7:113153060-113153120 | PPP1R3A | -75 | 5 | monocyte |
| chr20:004669765-004669825 | PRNT | -481 | 11 | monocyte |
| chr19:000646594-000646650 | PRSSL1 | -161 | 39 | monocyte |
| chr22:036002457-036002517 | PSCD4 | -436 | 12 | monocyte |
| chr2:158126089-158126148 | PSCDBP | -6 | 11 | monocyte |
| chr2:158126252-158126312 | PSCDBP | -170 | 10 | monocyte |
| chr1:195339670-195339730 | PTPRC | -181 | 6 | monocyte |
| chr1:195339927-195339987 | PTPRC | 76 | 6 | monocyte |
| chr1:189276226-189276286 | RGS1 | -257 | 8 | monocyte |
| chr1:189276641-189276701 | RGS1 | 158 | 7 | monocyte |
| chr14:020429033-020429091 | RNASE3 | -339 | 7 | monocyte |
| chrX:105742642-105742702 | RNF128 | -540 | 5 | monocyte |
| chr21:035343526-035343586 | RUNX1 | -91 | 12 | monocyte |
| chr5:076181489-076181549 | S100Z | -62 | 18 | monocyte |
| chr16:020682896-020682956 | SAH | 114 | 6 | monocyte |
| chr21:014840611-014840671 | SAMSN1 | -106 | 5 | monocyte |
| chr2:165886060-165886120 | SCN3A | -30 | 6 | monocyte |
| chr2:165886326-165886386 | SCN3A | -296 | 7 | monocyte |
| chr1:166412892-166412952 | SELL | -537 | 6 | monocyte |
| chr20:001548843-001548895 | SIRPB1 | -239 | 8 | monocyte |
| chr1:157430268-157430328 | SLAMF1 | -356 | 9 | monocyte |
| chr2:219072069-219072118 | SLC11A1 | -328 | 14 | monocyte |
| chr2:219072566-219072612 | SLC11A1 | 53 | 15 | monocyte |
| chr7:122434066-122434126 | SLC13A1 | -120 | 5 | monocyte |
| chr13:085271470-085271530 | SLITRK6 | -16 | 6 | monocyte |
| chr13:085271744-085271804 | SLITRK6 | -290 | 5 | monocyte |
| chr16:029581461-029581521 | SPN | -309 | 17 | monocyte |
| chr13:032678432-032678492 | STARD13 | -319 | 7 | monocyte |
| chr11:059390799-059390859 | TCN1 | -235 | 5 | monocyte |
| chr9:117546051-117546111 | TLR4 | -56 | 13 | monocyte |
| chr1:169751629-169751682 | TNFSF18 | 58 | 5 | monocyte |
| chr1:169751799-169751859 | TNFSF18 | -116 | 5 | monocyte |
| chr1:169752037-169752097 | TNFSF18 | -354 | 6 | monocyte |
| chr1:074412639-074412699 | TNNI3K | -436 | 5 | monocyte |
| chr1:074412830-074412890 | TNNI3K | -245 | 5 | monocyte |
| chr3:190831478-190831538 | TP73L | -409 | 9 | monocyte |
| chr6:041277331-041277391 | TREML2 | -459 | 11 | monocyte |
| chr21:044597502-044597562 | TRPM2 | -379 | 19 | monocyte |
| chr11:002279235-002279289 | TSPAN32 | -556 | 30 | monocyte |
| chr11:002279412-002279472 | TSPAN32 | -376 | 27 | monocyte |
| chr11:063730212-063730262 | URP2 | -544 | 19 | monocyte |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr11:063730360-063730409 | URP2 | -397 | 17 | monocyte |
| chr6:133076944-133077004 | VNN1 | -93 | 10 | monocyte |
| chrX:065042712-065042769 | VSIG4 | 166 | 5 | monocyte |
| chrX:065043147-065043205 | VSIG4 | -270 | 5 | monocyte |
| chrX:048298024-048298072 | WAS | -410 | 11 | monocyte |
| chrX:048298366-048298413 | WAS | -69 | 18 | monocyte |
| chrX:048298583-048298629 | WAS | 148 | 17 | monocyte |
| chr20:043691414-043691474 | WFDC10A | -354 | 7 | monocyte |
| chr10:102212516-102212567 | WNT8B | -246 | 6 | monocyte |
| chr20:055628978-055629038 | ZBP1 | -73 | 16 | monocyte |
| chr1:226001750-226001795 | ABCB10 | -595 | 61 | testis |
| chrX:070581435-070581495 | ACRC | 184 | 42 | testis |
| chr9:108697985-108698034 | ACTL7B | -244 | 14 | testis |
| chr8:039815105-039815165 | ADAM2 | -199 | 30 | testis |
| chr8:039815434-039815494 | ADAM2 | -528 | 15 | testis |
| chr15:083160423-083160483 | ALPK3 | -461 | 24 | testis |
| chr15:083160861-083160921 | ALPK3 | -23 | 33 | testis |
| chr17:057023731-057023791 | ANAC | -416 | 12 | testis |
| chr8:041774538-041774583 | ANK1 | -263 | 53 | testis |
| chr12:050567855-050567909 | ANKRD33 | -128 | 35 | testis |
| chr1:062497498-062497558 | ANKRD38 | -424 | 29 | testis |
| chr4:165476253-165476311 | ANP32C | 186 | 27 | testis |
| chr14:030564711-030564756 | AP4S1 | 90 | 57 | testis |
| chr2:021179003-021179051 | APOB | -430 | 20 | testis |
| chr2:131507983-131508043 | ARHGEF4 | 58 | 38 | testis |
| chr7:116661739-116661799 | ASZ1 | -275 | 44 | testis |
| chr7:116662038-116662098 | ASZ1 | -574 | 17 | testis |
| chr3:184453274-184453334 | B3GNT5 | -442 | 27 | testis |
| chr10:013427520-013427580 | BC000941 | -526 | 25 | testis |
| chr2:010813651-010813711 | BC012142 | 128 | 7 | testis |
| chr10:065052808-065052868 | BC018658 | -335 | 18 | testis |
| chr6:007930830-007930890 | BC028580 | -476 | 17 | testis |
| chr1:150561140-150561185 | BC029792 | 190 | 77 | testis |
| chr3:040404320-040404380 | BC029869 | -190 | 19 | testis |
| chr4:156019738-156019798 | BC031053 | 171 | 29 | testis |
| chr11:093940419-093940473 | BC031060 | 117 | 30 | testis |
| chr20:005399468-005399528 | BC031676 | -354 | 18 | testis |
| chr17:069781533-069781582 | BC033013 | -432 | 33 | testis |
| chr10:102737912-102737961 | BC033762 | -364 | 21 | testis |
| chr19:047039558-047039603 | BC034629 | 54 | 11 | testis |
| chr19:047039753-047039807 | BC034629 | -146 | 16 | testis |
| chr19:047039869-047039920 | BC034629 | -260 | 21 | testis |
| chr19:047039984-047040031 | BC034629 | -373 | 22 | testis |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr8:028230883-028230931 | BC034758 | 185 | 26 | testis |
| chr1:061831723-061831783 | BC034822 | -185 | 43 | testis |
| chrY:013453945-013454005 | BC034942 | -174 | 20 | testis |
| chr2:175409928-175409988 | BC038989 | -165 | 30 | testis |
| chr10:051774291-051774351 | BC042899 | -226 | 24 | testis |
| chr19:061039769-061039823 | BC050326 | -285 | 21 | testis |
| chr19:061039875-061039927 | BC050326 | -180 | 21 | testis |
| chr19:061039998-061040043 | BC050326 | -61 | 17 | testis |
| chr2:024309223-024309283 | BC050462 | -369 | 48 | testis |
| chrX:042394092-042394138 | BC056673 | -449 | 30 | testis |
| chr11:032772794-032772839 | BC062749 | -39 | 25 | testis |
| chr11:032773245-032773305 | BC062749 | -498 | 17 | testis |
| chr5:156209712-156209772 | BC066922 | -384 | 38 | testis |
| chr2:147178392-147178452 | BC068242 | -570 | 34 | testis |
| chr14:090771094-090771141 | BC069592 | 30 | 29 | testis |
| chr19:058486091-058486151 | BC071665 | 114 | 43 | testis |
| chr4:123211791-123211850 | BC093682 | 53 | 39 | testis |
| chr20:044519568-044519613 | BC100775 | 99 | 22 | testis |
| chr10:005397209-005397263 | BC100867 | 182 | 29 | testis |
| chrX:027586080-027586140 | BC101308 | -179 | 28 | testis |
| chrX:027586324-027586375 | BC101308 | 60 | 21 | testis |
| chr12:050004544-050004596 | BIN2 | -365 | 46 | testis |
| chr5:000944836-000944882 | BRD9 | -456 | 34 | testis |
| chr1:092126698-092126758 | BRDT | -424 | 33 | testis |
| chr1:092127055-092127115 | BRDT | -67 | 37 | testis |
| chr4:015380379-015380435 | BST1 | -502 | 17 | testis |
| chr10:014856735-014856789 | C10orf45 | 140 | 21 | testis |
| chr10:014857023-014857081 | C10orf45 | -150 | 16 | testis |
| chr10:014857215-014857274 | C10orf45 | -342 | 13 | testis |
| chr10:050176694-050176739 | C10orf71 | -476 | 14 | testis |
| chr10:050177017-050177077 | C10orf71 | -145 | 24 | testis |
| chr10:050177136-050177196 | C10orf71 | -26 | 27 | testis |
| chr10:050177316-050177361 | C10orf71 | 146 | 25 | testis |
| chr11:064619729-064619778 | C11orf2 | -505 | 36 | testis |
| chr12:089851642-089851694 | C12orf12 | -247 | 26 | testis |
| chr2:119633051-119633111 | C1QL2 | -386 | 43 | testis |
| chr20:016502637-016502697 | C20orf23 | -589 | 32 | testis |
| chr20:006052697-006052757 | C20orf42 | -536 | 18 | testis |
| chr5:134236976-134237036 | C5orf14 | -319 | 7 | testis |
| chr6:028334561-028334615 | C6orf194 | -488 | 18 | testis |
| chr6:168216888-168216935 | C6orf54 | -598 | 22 | testis |
| chr8:090982710-090982770 | C8orf1 | -528 | 29 | testis |
| chr8:050146829-050146889 | C8orf22 | -600 | 22 | testis |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr9:083466971-083467024 | C9orf103 | -522 | 33 | testis |
| chr9:111636628-111636676 | C9orf84 | 126 | 30 | testis |
| chr19:051609001-051609054 | CCDC8 | -346 | 39 | testis |
| chr17:017421192-017421252 | CCDS11187.1 | -171 | 20 | testis |
| chr20:055532659-055532710 | CCDS13459.1 | -17 | 16 | testis |
| chr20:055532860-055532918 | CCDS13459.1 | -222 | 22 | testis |
| chr20:044179812-044179867 | CD40 | -473 | 35 | testis |
| chr16:054424364-054424416 | CES1 | 186 | 38 | testis |
| chr14:073775463-073775511 | CHX10 | -440 | 13 | testis |
| chr13:094883687-094883747 | CLDN10 | -141 | 25 | testis |
| chr13:094883826-094883886 | CLDN10 | -2 | 26 | testis |
| chr3:139211273-139211331 | CLDN18 | -401 | 23 | testis |
| chr21:034962950-034963003 | CLIC6 | -581 | 30 | testis |
| chr7:093667928-093667975 | COL1A2 | -572 | 26 | testis |
| chr10:101831460-101831506 | CPN1 | 101 | 22 | testis |
| chr13:036578196-036578256 | CSNK1A1L | -425 | 15 | testis |
| chr13:036578320-036578380 | CSNK1A1L | -549 | 16 | testis |
| chrX:119491377-119491437 | CUL4B | 124 | 18 | testis |
| chrX:017998934-017998994 | CXorf20 | -304 | 38 | testis |
| chr10:135229354-135229406 | CYP2E1 | -367 | 13 | testis |
| chr5:055069906-055069963 | DDX4 | -598 | 38 | testis |
| chr5:055070131-055070191 | DDX4 | -371 | 19 | testis |
| chr14:100262518-100262576 | DLK1 | -458 | 62 | testis |
| chr1:053636513-053636573 | DMRTB1 | -549 | 13 | testis |
| chr1:053636683-053636743 | DMRTB1 | -379 | 17 | testis |
| chr1:053636834-053636894 | DMRTB1 | -228 | 23 | testis |
| chr19:047040866-047040916 | DMRTC2 | 28 | 51 | testis |
| chr6:038798487-038798547 | DNAH8 | -358 | 23 | testis |
| chr6:038798726-038798779 | DNAH8 | -123 | 28 | testis |
| chr6:038798990-038799050 | DNAH8 | 145 | 20 | testis |
| chr17:069789351-069789407 | DNAI2 | 134 | 27 | testis |
| chr2:027409437-027409497 | DNAJC5G | -472 | 15 | testis |
| chr2:027409633-027409682 | DNAJC5G | -282 | 28 | testis |
| chr3:110517996-110518051 | DPPA2 | 31 | 33 | testis |
| chr3:110518270-110518330 | DPPA2 | -246 | 33 | testis |
| chr12:007755184-007755244 | DPPA3 | -141 | 15 | testis |
| chr12:007755506-007755557 | DPPA3 | 176 | 20 | testis |
| chrX:044459214-044459274 | DUSP21 | -259 | 16 | testis |
| chr13:049164091-049164151 | EBPL | -509 | 37 | testis |
| chr1:062372292-062372341 | ECAT11 | -240 | 32 | testis |
| chr10:011823894-011823954 | ECHDC3 | -464 | 44 | testis |
| chr11:034492005-034492057 | ELF5 | -125 | 22 | testis |
| chr7:072887099-072887144 | ELN | -47 | 39 | testis |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr19:012335763-012335819 | ENST00000242804 | -346 | 18 | testis |
| chr19:048602268-048602315 | ENST00000253435 | 78 | 27 | testis |
| chr3:115438395-115438445 | ENST00000308095 | 191 | 22 | testis |
| chr3:115438598-115438643 | ENST00000308095 | -9 | 33 | testis |
| chr8:088955150-088955195 | ENST00000319675 | 143 | 46 | testis |
| chr14:088086735-088086791 | ENST00000328736 | -44 | 25 | testis |
| chr1:024486962-024487022 | ENST00000337248 | -99 | 33 | testis |
| chr22:016607778-016607838 | ENST00000342111 | -313 | 34 | testis |
| chr9:096879787-096879842 | ENST00000343150 | 66 | 12 | testis |
| chr10:014412406-014412461 | ENST00000358621 | -326 | 17 | testis |
| chr5:145297403-145297463 | ENST00000359120 | -199 | 12 | testis |
| chr16:023632406-023632466 | ERN2 | -114 | 31 | testis |
| chr3:033293599-033293659 | FBXL2 | -308 | 39 | testis |
| chr12:004425224-004425284 | FGF6 | -213 | 25 | testis |
| chr2:070929759-070929819 | FIGLA | -359 | 41 | testis |
| chr2:070929932-070929992 | FIGLA | -532 | 22 | testis |
| chr6:035765221-035765274 | FKBP5 | -555 | 24 | testis |
| chr19:004197644-004197696 | FLJ10374 | -437 | 24 | testis |
| chr5:014634305-014634365 | FLJ11127 | -596 | 27 | testis |
| chr2:127679908-127679968 | FLJ16008 | -365 | 18 | testis |
| chr13:035687166-035687226 | FLJ20449 | -520 | 37 | testis |
| chr12:014612434-014612494 | FLJ22662 | -406 | 29 | testis |
| chr12:014612584-014612644 | FLJ22662 | -556 | 33 | testis |
| chr4:189401781-189401829 | FLJ25801 | -248 | 13 | testis |
| chr4:189401950-189402010 | FLJ25801 | -423 | 15 | testis |
| chr12:053153513-053153572 | FLJ32942 | 72 | 27 | testis |
| chr22:027492745-027492805 | FLJ33814 | -485 | 26 | testis |
| chr11:006474809-006474866 | FLJ35709 | -269 | 33 | testis |
| chr11:006475069-006475115 | FLJ35709 | -14 | 33 | testis |
| chr16:065393596-065393648 | FLJ35894 | -598 | 21 | testis |
| chr12:038305834-038305894 | FLJ40126 | -387 | 25 | testis |
| chr12:038306039-038306099 | FLJ40126 | -182 | 27 | testis |
| chr3:128394059-128394110 | FLJ40141 | -587 | 20 | testis |
| chr3:128394789-128394849 | FLJ40141 | 148 | 32 | testis |
| chr7:055676401-055676461 | FLJ44060 | 20 | 19 | testis |
| chrX:146767996-146768055 | FMR1NB | -369 | 26 | testis |
| chr7:151090966-151091017 | GALNTL5 | -167 | 19 | testis |
| chr9:076303158-076303218 | GCNT1 | -511 | 31 | testis |
| chr13:019615417-019615468 | GJA3 | -15 | 32 | testis |
| chr12:102946927-102946987 | GLT8D2 | -575 | 14 | testis |
| chrX:009543990-009544046 | GPR143 | -365 | 20 | testis |
| chr19:050787372-050787432 | GPR4 | 155 | 11 | testis |
| chr2:105316635-105316685 | GPR45 | -57 | 22 | testis |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr12:121740032-121740085 | GPR81 | -49 | 30 | testis |
| chr12:121740456-121740516 | GPR81 | -477 | 22 | testis |
| chr3:142979453-142979513 | GRK7 | -257 | 16 | testis |
| chr3:142979636-142979688 | GRK7 | -78 | 28 | testis |
| chr1:042299954-042299999 | GUCA2A | -488 | 17 | testis |
| chr12:047009081-047009135 | H1T2 | 79 | 18 | testis |
| chrX:103074934-103074994 | H2BFWT | -563 | 13 | testis |
| chr2:172603958-172604018 | HAT1 | -493 | 37 | testis |
| chr11:057180646-057180706 | HEAB | -529 | 23 | testis |
| chr6:026126370-026126430 | HIST1H1A | -405 | 30 | testis |
| chr6:026216289-026216349 | HIST1H1T | 24 | 28 | testis |
| chr6:026216516-026216576 | HIST1H1T | -203 | 23 | testis |
| chr6:026216688-026216748 | HIST1H1T | -375 | 16 | testis |
| chr6:025834843-025834893 | HIST1H2AA | -99 | 35 | testis |
| chr6:025835017-025835072 | HIST1H2AA | -275 | 33 | testis |
| chr6:026355415-026355475 | HIST1H4G | -261 | 35 | testis |
| chr1:147506201-147506261 | HORMAD1 | 194 | 43 | testis |
| chr14:100418478-100418527 | hsa-mir-127 | -567 | 42 | testis |
| chr14:100418684-100418741 | hsa-mir-127 | -357 | 35 | testis |
| chr14:100418906-100418957 | hsa-mir-127 | -138 | 33 | testis |
| chr19:058982557-058982617 | hsa-mir-371 | -154 | 25 | testis |
| chr19:058983199-058983249 | hsa-mir-373 | -547 | 41 | testis |
| chr19:058983592-058983652 | hsa-mir-373 | -149 | 34 | testis |
| chr19:058916941-058917001 | hsa-mir-520g | -261 | 10 | testis |
| chrX:069064911-069064958 | HSHIN6 | -427 | 28 | testis |
| chr3:185300293-185300347 | HTR3E | -348 | 30 | testis |
| chr5:158690481-158690530 | IL12B | -446 | 12 | testis |
| chr6:137407996-137408056 | IL20RA | -35 | 36 | testis |
| chr9:136610397-136610457 | INPP5E | -334 | 54 | testis |
| chr9:005175919-005175973 | INSL6 | -328 | 42 | testis |
| chr1:223232075-223232122 | ITPKB | -168 | 76 | testis |
| chr14:062638206-062638266 | KCNH5 | 101 | 24 | testis |
| chr3:180467656-180467707 | KCNMB3 | -141 | 20 | testis |
| chr16:076380089-076380134 | KIAA1576 | 128 | 70 | testis |
| chr17:049254827-049254872 | KIF2B | -411 | 14 | testis |
| chr17:049255007-049255067 | KIF2B | -223 | 22 | testis |
| chr17:036725541-036725601 | KRTAP17-1 | -98 | 29 | testis |
| chr12:050981310-050981369 | KRTHB6 | -576 | 14 | testis |
| chr12:050981695-050981751 | KRTHB6 | -192 | 33 | testis |
| chr12:050981838-050981890 | KRTHB6 | -51 | 46 | testis |
| chr18:019706520-019706573 | LAMA3 | -457 | 16 | testis |
| chr15:098902835-098902895 | LASS3 | -417 | 23 | testis |
| chr2:074642598-074642647 | LBX2 | -524 | 46 | testis |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr11:018433626-018433686 | LDHAL6A | -350 | 31 | testis |
| chr4:041824665-041824715 | LOC285429 | -13 | 29 | testis |
| chr5:118993002-118993062 | LOC340069 | -120 | 12 | testis |
| chr5:118993224-118993284 | LOC340069 | 102 | 23 | testis |
| chr3:140222658-140222718 | LOC389151 | -222 | 38 | testis |
| chr3:140222811-140222871 | LOC389151 | -375 | 24 | testis |
| chr12:031836914-031836974 | LOC440093 | -502 | 14 | testis |
| chr6:112775022-112775082 | LOC442247 | -172 | 13 | testis |
| chr6:112775289-112775349 | LOC442247 | 95 | 17 | testis |
| chr14:106330981-106331033 | LOC90925 | -170 | 8 | testis |
| chrX:114346898-114346958 | LUZP4 | -372 | 24 | testis |
| chrX:114347357-114347414 | LUZP4 | 85 | 30 | testis |
| chr1:163689893-163689953 | MAEL | -253 | 39 | testis |
| chrX:027585180-027585240 | MAGEB10 | -571 | 10 | testis |
| chrX:027585295-027585355 | MAGEB10 | -456 | 15 | testis |
| chrX:027585574-027585634 | MAGEB10 | -177 | 27 | testis |
| chrX:027585816-027585867 | MAGEB10 | 60 | 31 | testis |
| chr5:034020679-034020735 | MATP | -170 | 14 | testis |
| chr4:156059682-156059742 | MGC27016 | -392 | 31 | testis |
| chr19:047040294-047040351 | MGC42718 | 26 | 39 | testis |
| chr13:023781149-023781205 | MGC48915 | -538 | 23 | testis |
| chr16:054070000-054070055 | MMP2 | -576 | 22 | testis |
| chr10:127454430-127454477 | MMP21 | -73 | 31 | testis |
| chr11:004965548-004965608 | MMP26 | -421 | 20 | testis |
| chr3:110319810-110319870 | MORC1 | -182 | 44 | testis |
| chr3:110320025-110320081 | MORC1 | -395 | 27 | testis |
| chr3:110320179-110320239 | MORC1 | -551 | 12 | testis |
| chr22:024462570-024462625 | MYO18B | -76 | 32 | testis |
| chr22:024462774-024462822 | MYO18B | 125 | 32 | testis |
| chr2:105926618-105926665 | NCK2 | -181 | 29 | testis |
| chr2:105926928-105926988 | NCK2 | 136 | 22 | testis |
| chr19:060109208-060109254 | NCR1 | -88 | 22 | testis |
| chr22:028323597-028323657 | NF2 | -491 | 42 | testis |
| chr14:076807728-076807782 | NGB | -347 | 31 | testis |
| chr20:035582479-035582532 | NNAT | -515 | 39 | testis |
| chr1:202023046-202023106 | NUAK2 | -536 | 17 | testis |
| chrX:108585183-108585230 | NXT2 | 52 | 14 | testis |
| chr9:006317882-006317942 | NYD-SP25 | -462 | 14 | testis |
| chr9:006318029-006318089 | NYD-SP25 | -315 | 19 | testis |
| chr9:006318143-006318200 | NYD-SP25 | -203 | 21 | testis |
| chr9:006318527-006318587 | NYD-SP25 | 183 | 23 | testis |
| chr19:015779866-015779917 | OR10H1 | -44 | 21 | testis |
| chr11:057752890-057752942 | OR10Q1 | 7 | 36 | testis |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr11:057753340-057753400 | OR10Q1 | -447 | 29 | testis |
| chr9:104411064-104411120 | OR13C8 | 89 | 9 | testis |
| chr9:122591290-122591350 | OR1L6 | -252 | 14 | testis |
| chr9:122591559-122591619 | OR1L6 | 17 | 15 | testis |
| chr1:243941631-243941691 | OR2B11 | -336 | 16 | testis |
| chr1:244425915-244425966 | OR2L13 | -593 | 31 | testis |
| chr11:057714609-057714666 | OR9Q2 | 99 | 23 | testis |
| chr4:140559484-140559544 | OSAP | -530 | 28 | testis |
| chrX:119032158-119032209 | OTEX | -454 | 12 | testis |
| chrX:119032301-119032354 | OTEX | -598 | 13 | testis |
| chr20:002999795-002999852 | OXT | -442 | 19 | testis |
| chr1:017320543-017320597 | PADI3 | -328 | 18 | testis |
| chrX:049163752-049163805 | PAGE1 | -175 | 22 | testis |
| chrX:055129478-055129538 | PAGE5 | -303 | 28 | testis |
| chr7:139216973-139217033 | PARP12 | -298 | 20 | testis |
| chr5:140481779-140481834 | PCDHB4 | 42 | 27 | testis |
| chr9:122670999-122671059 | PDCL | -565 | 45 | testis |
| chr2:170377235-170377295 | PHOSPHO2 | 34 | 27 | testis |
| chr8:022188432-022188485 | PIWIL2 | -342 | 27 | testis |
| chr15:019336639-019336686 | POTE15 | 5 | 42 | testis |
| chrY:000272912-000272964 | PPP2R3B | -348 | 58 | testis |
| chr9:101437406-101437463 | PPP3R2 | -596 | 14 | testis |
| chr22:021226443-021226499 | PRAME | -221 | 28 | testis |
| chr22:021226692-021226751 | PRAME | -471 | 20 | testis |
| chr18:021967526-021967586 | PSMA8 | -257 | 36 | testis |
| chr14:051803708-051803768 | PTGDR | -442 | 26 | testis |
| chrY:022010829-022010879 | RBMY1A1 | -519 | 27 | testis |
| chrY:022011049-022011104 | RBMY1A1 | -297 | 41 | testis |
| chrY:022011196-022011241 | RBMY1A1 | -155 | 40 | testis |
| chrY:022034356-022034405 | RBMY1A1 | -534 | 26 | testis |
| chrY:022034590-022034645 | RBMY1A1 | -297 | 41 | testis |
| chrY:022379013-022379063 | RBMY1A1 | -265 | 40 | testis |
| chrY:022379305-022379352 | RBMY1A1 | -555 | 23 | testis |
| chrY:022402554-022402604 | RBMY1A1 | -265 | 40 | testis |
| chrY:022402846-022402893 | RBMY1A1 | -555 | 23 | testis |
| chrY:022035074-022035134 | RBMY1D | 175 | 38 | testis |
| chrY:022378632-022378685 | RBMY1E | 100 | 38 | testis |
| chrY:022402173-022402226 | RBMY1E | 100 | 38 | testis |
| chrY:022667088-022667141 | RBMY1F | 100 | 39 | testis |
| chrY:022667499-022667554 | RBMY1F | -312 | 42 | testis |
| chrY:022667706-022667751 | RBMY1F | -514 | 30 | testis |
| chrY:022792588-022792640 | RBMY1F | -520 | 30 | testis |
| chrY:022792785-022792845 | RBMY1F | -319 | 37 | testis |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chr17:009749337-009749382 | RCV1 | 50 | 27 | testis |
| chr13:097627971-097628018 | RNF113B | -474 | 21 | testis |
| chr13:024235857-024235913 | RNF17 | -453 | 25 | testis |
| chr13:024236016-024236076 | RNF17 | -292 | 31 | testis |
| chr17:058239951-058239996 | RNF190 | -546 | 20 | testis |
| chr20:055399371-055399431 | RNPC1 | -468 | 60 | testis |
| chrX:095944828-095944886 | RPA4 | -194 | 22 | testis |
| chrX:095944984-095945044 | RPA4 | -37 | 25 | testis |
| chr20:049852955-049853015 | SALL4 | -564 | 13 | testis |
| chr2:119998637-119998697 | SCTR | -531 | 34 | testis |
| chr15:041879150-041879209 | SERINC4 | -553 | 16 | testis |
| chr18:059295290-059295350 | SERPINB5 | 122 | 23 | testis |
| chr7:134870010-134870061 | SLC13A4 | 153 | 18 | testis |
| chr7:134870181-134870240 | SLC13A4 | -22 | 18 | testis |
| chr7:134870420-134870476 | SLC13A4 | -260 | 16 | testis |
| chr5:140664281-140664334 | SLC25A2 | -511 | 39 | testis |
| chr4:129008782-129008842 | SLC25A31 | -366 | 33 | testis |
| chr15:043331195-043331251 | SLC28A2 | -502 | 26 | testis |
| chr19:054520595-054520655 | SLC6A16 | -339 | 34 | testis |
| chr19:054520701-054520761 | SLC6A16 | -445 | 28 | testis |
| chr3:150287495-150287555 | SMARCA3 | -510 | 36 | testis |
| chr5:157012185-157012243 | SOX30 | -208 | 27 | testis |
| chr17:028342640-028342695 | SPACA3 | -327 | 8 | testis |
| chr20:055337931-055337991 | SPO11 | -276 | 36 | testis |
| chrX:047812424-047812478 | SSX5 | 2 | 18 | testis |
| chr7:023522820-023522873 | STK31 | -286 | 37 | testis |
| chr7:023523231-023523281 | STK31 | 74 | 48 | testis |
| chr15:072282580-072282640 | STRA6 | -365 | 17 | testis |
| chr1:115109205-115109265 | SYCP1 | -261 | 25 | testis |
| chr2:047507506-047507566 | TACSTD1 | -581 | 31 | testis |
| chr9:032625452-032625500 | TAF1L | 191 | 30 | testis |
| chrX:100354018-100354063 | TAF7L | 150 | 32 | testis |
| chrX:100354217-100354269 | TAF7L | -53 | 25 | testis |
| chr14:054977244-054977293 | TBPL2 | -252 | 37 | testis |
| chr18:042816199-042816259 | TCEB3B | -243 | 18 | testis |
| chr6:035217605-035217665 | TCP11 | -531 | 36 | testis |
| chrX:069911419-069911479 | TEX11 | 139 | 28 | testis |
| chrX:069911757-069911817 | TEX11 | -199 | 17 | testis |
| chr11:111542733-111542787 | TEX12 | -544 | 14 | testis |
| chr11:111543010-111543067 | TEX12 | -266 | 31 | testis |
| chr11:111543401-111543446 | TEX12 | 119 | 35 | testis |
| chr17:054124266-054124322 | TEX14 | 121 | 27 | testis |
| chr17:054124532-054124591 | TEX14 | -146 | 36 | testis |

| *Probe Location (NCBI Built 35)* | *Gene Name* | *TSS-Distance of Probe* | *CpG-Index (+/- 350 bp)* | *demethylated in* |
|---|---|---|---|---|
| chrX:132077792-132077852 | TFDP3 | 74 | 52 | testis |
| chr7:093165110-093165170 | TFPI2 | -424 | 35 | testis |
| chrY:003490121-003490181 | TGIF2LY | -335 | 18 | testis |
| chrY:003490590-003490650 | TGIF2LY | 134 | 28 | testis |
| chr3:197553427-197553477 | TM4SF19 | 102 | 21 | testis |
| chr21:010013262-010013309 | TPTE | -494 | 57 | testis |
| chr6:030238601-030238661 | TRIM15 | -340 | 15 | testis |
| chr11:005573297-005573357 | TRIM6 | -595 | 15 | testis |
| chr4:166256899-166256947 | TRIM61 | -500 | 20 | testis |
| chr4:123212030-123212090 | TRPC3 | 156 | 17 | testis |
| chr5:135721032-135721092 | TRPC7 | -90 | 23 | testis |
| chrY:009217882-009217927 | TSPY1 | -529 | 38 | testis |
| chrY:009218125-009218180 | TSPY1 | -281 | 43 | testis |
| chrY:009258493-009258538 | TSPY1 | -530 | 36 | testis |
| chrY:009258675-009258726 | TSPY1 | -345 | 40 | testis |
| chrY:009278839-009278884 | TSPY1 | -529 | 38 | testis |
| chrY:009279082-009279137 | TSPY1 | -281 | 43 | testis |
| chrY:009917685-009917730 | TSPY1 | -529 | 38 | testis |
| chrY:009917928-009917983 | TSPY1 | -281 | 43 | testis |
| chrY:009937968-009938013 | TSPY1 | -529 | 38 | testis |
| chrY:009938211-009938266 | TSPY1 | -281 | 43 | testis |
| chrY:009958298-009958343 | TSPY1 | -529 | 38 | testis |
| chrY:009958541-009958596 | TSPY1 | -281 | 43 | testis |
| chrY:006157316-006157372 | TSPY2 | -284 | 41 | testis |
| chrY:009238215-009238260 | TSPY2 | -532 | 38 | testis |
| chrY:009238458-009238513 | TSPY2 | -284 | 42 | testis |
| chr2:054394979-054395026 | TSPYL6 | -7 | 37 | testis |
| chr13:018654238-018654291 | TUBA2 | -328 | 37 | testis |
| chr17:016224925-016224984 | UBB | -137 | 49 | testis |
| chr7:129186948-129187008 | UBE2H | -238 | 36 | testis |
| chr7:129187262-129187322 | UBE2H | -552 | 22 | testis |
| chr1:064380995-064381055 | UBE2U | -485 | 22 | testis |
| chr1:064381614-064381674 | UBE2U | 134 | 26 | testis |
| chr10:005396663-005396723 | UCN3 | -282 | 21 | testis |
| chr19:059259113-059259173 | UNQ3033 | -124 | 32 | testis |
| chr3:120374655-120374715 | UPK1B | -429 | 11 | testis |
| chr11:017522705-017522756 | USH1C | -191 | 53 | testis |
| chr15:048626028-048626088 | USP50 | -44 | 10 | testis |
| chr15:048626269-048626329 | USP50 | -285 | 15 | testis |
| chrY:014536716-014536766 | VCY | -218 | 23 | testis |
| chrY:014605985-014606035 | VCY | -218 | 23 | testis |
| chr14:096332954-096333014 | VRK1 | -452 | 32 | testis |
| chr8:088955364-088955411 | WDR21C | 25 | 43 | testis |

| ***Probe Location (NCBI Built 35)*** | ***Gene Name*** | ***TSS-Distance of Probe*** | ***CpG-Index (+/- 350 bp)*** | ***demethylated in*** |
|---|---|---|---|---|
| chr8:088955576-088955636 | WDR21C | -194 | 43 | testis |
| chrX:027759211-027759256 | WDR42B | -10 | 24 | testis |
| chrX:027759401-027759461 | WDR42B | -208 | 20 | testis |
| chr17:046267562-046267618 | WFIKKN2 | -248 | 35 | testis |
| chrX:052724088-052724138 | XAGE5 | -135 | 22 | testis |
| chrX:073306853-073306913 | ZCCHC13 | -162 | 32 | testis |
| chr16:070080829-070080881 | ZNF19 | -113 | 32 | testis |
| chr16:070081172-070081226 | ZNF19 | -457 | 13 | testis |
| chr19:057182391-057182449 | ZNF350 | -529 | 24 | testis |
| chr2:027717601-027717646 | ZNF512 | 80 | 40 | testis |
| chr19:009470707-009470767 | ZNF560 | -458 | 27 | testis |
| chrX:022050692-022050752 | ZNF645 | 6 | 22 | testis |
| chr17:035277376-035277431 | ZPBP2 | -591 | 29 | testis |
| chr17:035277492-035277552 | ZPBP2 | -472 | 42 | testis |
| chr19:061431564-061431617 | ZSCAN5 | -119 | 27 | testis |

# 12.2 Appendix II – Differentially expressed loci

**Table 12.2** Genomic regions selected for C57BL/6-BALB/c DMR screening

| ***Gene Symbol*** | ***Genomic Location (mm8 Assembly)*** | ***Region Size*** |
|---|---|---|
| *Tmem14a/Gsta3* | chr1:21175070-21265070 | 90000 |
| *Prim2* | chr1:33353488-33657701 | 304213 |
| *Chst10* | chr1:38765658-38877701 | 112043 |
| 2810022L02Rik/*Kctd18* | chr1:57667722-57937906 | 270184 |
| *Serpinb10* | chr1:109300970-109410970 | 110000 |
| *Marco* | chr1:122250969-122380969 | 130000 |
| *Ctse*/5430435G22Ri*k* | chr1:133440969-133560969 | 120000 |
| *Qscn6* | chr1:157505088-157615088 | 110000 |
| *Rfwd2* | chr1:161011546-161236546 | 225000 |
| *Cd244/Ly9/SlamF7* | chr1:173372619-173530408 | 157789 |
| *Gm1313* | chr1:175248822-175352711 | 103889 |
| *Ifi202b* | chr1:175779250-175898466 | 119216 |
| *Usp6nl* | chr2:6194226-6411092 | 216866 |
| *Ssb* | chr2:69633467-69703643 | 70176 |
| *Gchfr* | chr2:118831229-118890830 | 59601 |
| *Spint1* | chr2:118890918-118975954 | 85036 |
| *Tmem87a* | chr2:120020705-120135705 | 115000 |
| *Sord* | chr2:121881280-121980778 | 99498 |
| *Il1b* | chr2:129006032-129112561 | 106529 |
| *Slc13a3* | chr2:165061308-165201402 | 140094 |
| *Col20a1* | chr2:180896703-180996948 | 100245 |
| *Postn* | chr3:54409054-54528961 | 119907 |
| 5830417I10Rik | chr3:88835975-88948905 | 112930 |
| *Chi3l3/Chi3l4* | chr3:106245610-106405138 | 159528 |
| *D3Bwg0562e* | chr3:117261149-117402880 | 141731 |
| *Ank2* | chr3:126874104-127524104 | 650000 |
| *Dkk2* | chr3:131974607-132167610 | 193003 |
| *Gbp1* | chr3:142392233-142606562 | 214329 |
| *Ifi44/H28* | chr3:151618310-151734422 | 116112 |
| *Coro2a* | chr4:46525228-46668228 | 143000 |
| *Ptprd* | chr4:75330613-76115327 | 784714 |
| 2310009E04Rik | chr4:95005807-95450953 | 445146 |
| A430090E18Rik | chr4:111680177-111848400 | 168223 |

| ***Gene Symbol*** | ***Genomic Location (mm8 Assembly)*** | ***Region Size*** |
|---|---|---|
| *Tie1* | chr4:117945964-118024781 | 78817 |
| *Zfp69* | chr4:120394772-120485053 | 90281 |
| 4732473B16Rik | chr4:132702209-132797015 | 94806 |
| *C1qb* | chr4:136111222-136201253 | 90031 |
| *Padi4* | chr4:139981584-140156279 | 174695 |
| *Fblim1* | chr4:140812020-140903161 | 91141 |
| 2610305D13Rik | chr4:146416268-146536308 | 120040 |
| *Angptl7* | chr4:147289080-147394262 | 105182 |
| *Prdm16* | chr4:153079925-153530673 | 450748 |
| *Mmel1* | chr4:153676058-153769691 | 93633 |
| *Abhd1* | chr5:31199647-31249676 | 50029 |
| A930005I04Rik | chr5:35197772-35279696 | 81924 |
| *Pi4k2b* | chr5:52979822-53082588 | 102766 |
| *Uchl1* | chr5:66906028-67016369 | 110341 |
| 9430027B09Rik*(AK020441)* | chr5:67638314-67688271 | 49957 |
| *Pkd2* | chr5:104676856-104787118 | 110262 |
| *Ccl24* | chr5:135819207-135907670 | 88463 |
| 4921520G13Rik | chr5:144115000-144185000 | 70000 |
| *Sgce* | chr6:4578612-4747100 | 168488 |
| *Ppp1r9a* | chr6:4803428-5166532 | 363104 |
| *Asb4* | chr6:5290427-5433021 | 142594 |
| *Smo* | chr6:29650513-29741369 | 90856 |
| 3321401G04Rik | chr6:42564846-42689678 | 124832 |
| *Atp6v0e2* | chr6:48440201-48490384 | 50183 |
| *Magi1* | chr6:93590731-94299367 | 708636 |
| *Dppa3* | chr6:122556043-122635890 | 79847 |
| *Klrk1* | chr6:129547265-129609534 | 62269 |
| *Klra21* | chr6:130010920-130173550 | 162630 |
| *Dusp16* | chr6:134629170-134807675 | 178505 |
| *Pde6h* | chr6:136890719-136977679 | 86960 |
| *Eps8l1* | chr7:4031936-4102329 | 70393 |
| *Isoc2b(*0610042E07Rik*)* | chr7:4398047-4511485 | 113438 |
| *Gpr77* | chr7:15363248-15438676 | 75428 |
| *Ryr1* | chr7:28673943-28873741 | 199798 |
| *Zfp74* | chr7:29604551-29703741 | 99190 |
| *Dmkn* | chr7:30449165-30523818 | 74653 |
| *Pop4* | chr7:37931908-38030436 | 98528 |
| *Rcn3* | chr7:44922955-44993196 | 70241 |
| *Kcnj14* | chr7:45664509-45712789 | 48280 |

| ***Gene Symbol*** | ***Genomic Location (mm8 Assembly)*** | ***Region Size*** |
|---|---|---|
| *Snrpn* | chr7:59811452-60379127 | 567675 |
| *Hddc3* | chr7:80189650-80259610 | 69960 |
| *Mesdc2* | chr7:83718442-83817717 | 99275 |
| *Dgat2* | chr7:98998855-99094905 | 96050 |
| *Trim12*/A530023O14Rik | chr7:104137416-104293705 | 156289 |
| *Thumpd1* | chr7:119488373-119552914 | 64541 |
| *Crym* | chr7:119938287-120028137 | 89850 |
| *Il21r* | chr7:125344625-125452718 | 108093 |
| *H19* | chr7:142330097-142530245 | 200148 |
| *Tspan32* | chr7:142765081-142859034 | 93953 |
| *Cd209a* | chr8:3703397-3798930 | 95533 |
| *Gdf15* | chr8:73530382-73585445 | 55063 |
| *Cyp4f18* | chr8:74880471-74985606 | 105135 |
| *Arl2bp* | chr8:97536235-97583551 | 47316 |
| *Slc7a5* | chr8:124717240-124823762 | 106522 |
| *Pgbd5* | chr8:127225139-127370026 | 144887 |
| *Thy1* | chr9:43744379-43829574 | 85195 |
| *Amica1* | chr9:44804941-44888846 | 83905 |
| *Ttc12* | chr9:49144191-49278455 | 134264 |
| *Snf1lk* | chr9:50599125-50804125 | 205000 |
| 9830163H01Rik | chr9:51776244-51870405 | 94161 |
| *Mns1* | chr9:72187100-72307141 | 120041 |
| *Rab6b* | chr9:102930224-103076420 | 146196 |
| *Camp* | chr9:109684793-109736865 | 52072 |
| *Vill* | chr9:118883109-118950024 | 66915 |
| *Cck* | chr9:121308523-121394360 | 85837 |
| *Ppp1r14c* | chr10:6820000-7032000 | 212000 |
| *Lats1/Zc3h12d* | chr10:7350392-7600392 | 250000 |
| *Ltb* | chr17:34786107-34829107 | 43000 |
| *Dscr1l1* | chr17:43214106-43541106 | 327000 |
| *Guca1a* | chr17:46836242-46881242 | 45000 |
| *Fzd8* | chr18:9162918-9264975 | 102057 |
| *Colec12* | chr18:9657675-9927990 | 270315 |
| *Slc27a6* | chr18:58631609-58780715 | 149106 |
| *Pdgfrb* | chr18:61145715-61235715 | 90000 |
| *Kcnk7* | chr19:5680000-5725000 | 45000 |
| *Best1(Vmd2)* | chr19:10002219-10098259 | 96040 |
| *Gcnt1* | chr19:17343162-17489863 | 146701 |
| *Tjp2* | chr19:24112607-24341607 | 229000 |

| ***Gene Symbol*** | ***Genomic Location (mm8 Assembly)*** | ***Region Size*** |
|---|---|---|
| *Ankrd15* | chr19:25257607-25550607 | 293000 |
| *Ppp1r2* | chr19:30554873-30641507 | 86634 |
| *D19Ertd652e* | chr19:47803821-48003821 | 200000 |
| 3110007F17Rik | chrX:119170503-119310237 | 139734 |
| *Casp8** | chr1:58,690,040-58,841,559 | 151519 |
| *Ly96* | chr1:16,633,670-16,740,819 | 107149 |
| *Cebpb* | chr2:167,330,172-167,421,592 | 91420 |
| *Psma7* | chr2:179,920,784-180,011,810 | 91026 |
| *Creb3* | chr4:43,546,778-43,630,964 | 84186 |
| *S100pbp* | chr4:128,620,129-128,730,786 | 110657 |
| *Clec4a2* | chr6:123,039,011-123,150,618 | 111607 |
| *Cebpa* | chr7:34,778,063-34,880,686 | 102623 |
| *Tmem34* | chr8:80,452,004-80,556,693 | 104689 |
| *Egr2* | chr10:66,883,235-66,987,541 | 104306 |
| *Lrp1* | chr10:126,895,143-127,065,000 | 169857 |
| *Nef2* | chr12:57,562,909-57,680,019 | 117110 |
| *Jarid2* | chr13:44,691,875-44,981,947 | 290072 |
| *Cebpd* | chr16:15,755,864-15,853,123 | 97259 |
| *Mapk1* | chr16:16,850,945-17,023,015 | 172070 |
| *Psmb9* | chr17:33,752,387-33,837,618 | 85231 |

* Control loci are boxed in light green.

# 12.3 Appendix III – Annotated list of mouse strain-specific DMR

**Table 12.3** Annotated list of all mouse strain-specific DMR

| *Center of DMR (Mouse assembly mm8)* | *Gene Symbol* | *Position relative to Gene* | *Demethylated in* | *BMM BALB/c (0h)* | *BMM C57BL/6 (0h)* |
|---|---|---|---|---|---|
| chr1:33428882-33429004 | ***Prim2*** | intra | C57BL/6 | 0.0513 (P) | 0.998 (P) |
| chr1:33476141-33476727 | ***Prim2*** | intra | C57BL/6 | 0.0513 (P) | 0.998 (P) |
| chr1:33499231-33499713 | ***Prim2*** | intra | C57BL/6 | 0.0513 (P) | 0.998 (P) |
| chr1:33585109-33585169 | ***Prim2*** | intra | BALB/c | 0.0513 (P) | 0.998 (P) |
| chr1:33600588-33600898 | ***Prim2*** | intra | C57BL/6 | 0.0513 (P) | 0.998 (P) |
| chr1:38834245-38834666 | ***Chst10*** | intra | C57BL/6 | 35.28 (P) | 1 (P) |
| chr1:57749504-57749558 | ***2810022L02Rik*** | intra | BALB/c | 10.77 (P) | 0.999 (A) |
| chr1:58749011-58749071 | ***Casp8*** | intra | BALB/c | 0.996 (P) | 0.998 (P) |
| chr1:58777847-58778072 | ***Casp8*** | intra | C57BL/6 | 0.996 (P) | 0.998 (P) |
| chr1:109401592-109402066 | ***Serpinb8*** | distal | C57BL/6 | 1.48 (P) | 0.995 (P) |
| chr1:122290855-122291258 | ***Marco*** | distal | BALB/c | 0.0779 (P) | 1 (P) |
| chr1:122308958-122309018 | ***Marco*** | intra | C57BL/6 | 0.0779 (P) | 1 (P) |
| chr1:122371709-122371770 | ***Marco*** | distal | C57BL/6 | 0.0779 (P) | 1 (P) |
| chr1:133539104-133539235 | ***5430435G22Rik*** | intra | BALB/c | 0.0695 (P) | 0.995 (P) |
| chr1:133540665-133540799 | ***5430435G22Rik*** | intra | C57BL/6 | 0.0695 (P) | 0.995 (P) |
| chr1:133553009-133553065 | ***Slc26a9*** | distal | BALB/c | 0.561 (A) | 0.991 (P) |
| chr1:133559814-133559910 | ***Slc26a9*** | distal | BALB/c | 0.561 (A) | 0.991 (P) |
| chr1:157549961-157550607 | ***Qsox1*** | intra | C57BL/6 | 9.433 (P) | 0.992 (P) |
| chr1:157554929-157555080 | ***Qsox1*** | intra | C57BL/6 | 9.433 (P) | 0.992 (P) |
| chr1:157577435-157577492 | ***Qsox1*** | distal | BALB/c | 9.433 (P) | 0.992 (P) |
| chr1:161072415-161072474 | ***Rfwd2*** | intra | BALB/c | 0.0874 (P) | 0.987 (P) |
| chr1:161132832-161133097 | ***Rfwd2*** | intra | C57BL/6 | 0.0874 (P) | 0.987 (P) |
| chr1:161154201-161154255 | ***Rfwd2*** | intra | BALB/c | 0.0874 (P) | 0.987 (P) |
| chr1:161173713-161173898 | ***Rfwd2*** | intra | C57BL/6 | 0.0874 (P) | 0.987 (P) |
| chr1:161181010-161181534 | ***Rfwd2*** | intra | C57BL/6 | 0.0874 (P) | 0.987 (P) |
| chr1:161220409-161220798 | ***Rfwd2*** | distal | BALB/c | 0.0874 (P) | 0.987 (P) |
| chr1:173418157-173418737 | ***Cd244*** | intra | BALB/c | 0.122 (P) | 0.987 (P) |
| chr1:173424681-173424758 | ***Ly9*** | distal | BALB/c | 0.851 (P) | 0.995 (P) |
| chr1:173434392-173434437 | ***Ly9*** | intra | C57BL/6 | 0.851 (P) | 0.995 (P) |
| chr1:173472612-173472668 | ***Slamf7*** | intra | BALB/c | 0.17 (P) | 0.999 (P) |
| chr1:173480500-173480622 | ***Slamf7*** | intra | BALB/c | 0.17 (P) | 0.999 (P) |
| chr1:173509929-173510843 | ***Cd48*** | distal | C57BL/6 | 0.953 (P) | 0.999 (P) |
| chr1:173521917-173521962 | ***Cd48*** | intra | BALB/c | 0.953 (P) | 0.999 (P) |
| chr10:6828900-6829688 | ***Iyd*** | distal | C57BL/6 | 1.207 (A) | 0.996 (A) |
| chr10:7018786-7019486 | ***Ppp1r14c*** | distal | C57BL/6 | 0.0184 (A) | 1 (P) |
| chr10:7456811-7457022 | ***BC013529*** | distal | C57BL/6 | 0.781 (P) | 0.993 (P) |
| chr10:7465903-7466014 | ***BC013529*** | intra | BALB/c | 0.781 (P) | 0.993 (P) |
| chr10:7554662-7554765 | ***Zc3h12d*** | intra | C57BL/6 | 0.0925 (P) | 0.999 (P) |

| *Center of DMR (Mouse assembly mm8)* | *Gene Symbol* | *Position relative to Gene* | *Demethylated in* | *BMM BALB/c (0h)* | *BMM C57BL/6 (0h)* |
|---|---|---|---|---|---|
| chr10:7556446-7556491 | ***Zc3h12d*** | intra | BALB/c | 0.0925 (P) | 0.999 (P) |
| chr10:8453154-8453214 | ***Sash1*** | intra | BALB/c | 1.302 (P) | 0.994 (P) |
| chr10:8499529-8500702 | ***Sash1*** | intra | BALB/c | 1.302 (P) | 0.994 (P) |
| chr10:8516559-8516781 | ***Sash1*** | intra | BALB/c | 1.302 (P) | 0.994 (P) |
| chr10:8533747-8533799 | ***Sash1*** | intra | C57BL/6 | 1.302 (P) | 0.994 (P) |
| chr10:8565556-8566318 | ***Sash1*** | intra | BALB/c | 1.302 (P) | 0.994 (P) |
| chr10:8569257-8569687 | ***Sash1*** | intra | C57BL/6 | 1.302 (P) | 0.994 (P) |
| chr10:8600425-8600804 | ***Sash1*** | distal | BALB/c | 1.302 (P) | 0.994 (P) |
| chr10:18209707-18210091 | ***Nhsl1*** | distal | C57BL/6 | 1.503 (P) | 0.966 (P) |
| chr10:18262232-18262290 | ***Hebp2*** | distal | C57BL/6 | 21.81 (P) | 1 (P) |
| chr10:18264292-18264337 | ***Hebp2*** | distal | C57BL/6 | 21.81 (P) | 1 (P) |
| chr10:84777151-84777599 | ***Btbd11*** | distal | BALB/c | 0.151 (A) | 0.96 (P) |
| chr10:84813063-84813163 | ***Btbd11*** | distal | C57BL/6 | 0.151 (A) | 0.96 (P) |
| chr10:85015114-85015305 | ***Btbd11*** | intra | C57BL/6 | 0.151 (A) | 0.96 (P) |
| chr10:85038563-85039091 | ***Btbd11*** | intra | BALB/c | 0.151 (A) | 0.96 (P) |
| chr10:85073454-85073867 | ***Btbd11*** | intra | BALB/c | 0.151 (A) | 0.96 (P) |
| chr10:85077536-85077897 | ***Btbd11*** | intra | C57BL/6 | 0.151 (A) | 0.96 (P) |
| chr10:88105615-88105792 | ***Spic*** | intra | BALB/c | 6.522 (P) | 0.991 (P) |
| chr10:92874018-92874212 | ***Lta4h*** | distal | C57BL/6 | 1.032 (P) | 1 (P) |
| chr10:92886703-92886763 | ***Lta4h*** | intra | BALB/c | 1.032 (P) | 1 (P) |
| chr10:92893443-92893580 | ***Lta4h*** | intra | BALB/c | 1.032 (P) | 1 (P) |
| chr10:92894786-92895269 | ***Lta4h*** | intra | C57BL/6 | 1.032 (P) | 1 (P) |
| chr10:92900457-92900512 | ***Lta4h*** | intra | BALB/c | 1.032 (P) | 1 (P) |
| chr10:92912661-92912967 | ***Lta4h*** | intra | C57BL/6 | 1.032 (P) | 1 (P) |
| chr10:92916395-92916703 | ***Hal*** | distal | C57BL/6 | 1293 (P) | 0.997 (P.A) |
| chr10:92931309-92931538 | ***Hal*** | intra | C57BL/6 | 1293 (P) | 0.997 (P.A) |
| chr10:92947751-92947804 | ***Hal*** | distal | C57BL/6 | 1293 (P) | 0.997 (P.A) |
| chr10:93014816-93014876 | ***Ccdc38*** | intra | C57BL/6 | 1.207 (A) | 0.996 (A) |
| chr10:104598765-104598953 | ***Tmtc2*** | intra | BALB/c | 0.0226 (P.A) | 0.993 (P) |
| chr10:104626764-104627103 | ***Tmtc2*** | intra | BALB/c | 0.0226 (P.A) | 0.993 (P) |
| chr10:104634539-104634758 | ***Tmtc2*** | intra | BALB/c | 0.0226 (P.A) | 0.993 (P) |
| chr10:104635908-104635968 | ***Tmtc2*** | intra | C57BL/6 | 0.0226 (P.A) | 0.993 (P) |
| chr10:104671761-104672161 | ***Tmtc2*** | intra | BALB/c | 0.0226 (P.A) | 0.993 (P) |
| chr10:104706006-104706340 | ***Tmtc2*** | intra | C57BL/6 | 0.0226 (P.A) | 0.993 (P) |
| chr10:104747765-104747995 | ***Tmtc2*** | intra | C57BL/6 | 0.0226 (P.A) | 0.993 (P) |
| chr10:104759614-104759674 | ***Tmtc2*** | intra | BALB/c | 0.0226 (P.A) | 0.993 (P) |
| chr10:104789057-104789197 | ***Tmtc2*** | intra | C57BL/6 | 0.0226 (P.A) | 0.993 (P) |
| chr10:104839643-104839694 | ***Tmtc2*** | intra | BALB/c | 0.0226 (P.A) | 0.993 (P) |
| chr10:104840547-104840607 | ***Tmtc2*** | intra | C57BL/6 | 0.0226 (P.A) | 0.993 (P) |
| chr10:104840635-104840749 | ***Tmtc2*** | intra | C57BL/6 | 0.0226 (P.A) | 0.993 (P) |
| chr10:104842784-104842934 | ***Tmtc2*** | intra | BALB/c | 0.0226 (P.A) | 0.993 (P) |
| chr10:104883498-104883558 | ***Tmtc2*** | intra | C57BL/6 | 0.0226 (P.A) | 0.993 (P) |
| chr10:104927384-104927444 | ***Tmtc2*** | intra | C57BL/6 | 0.0226 (P.A) | 0.993 (P) |
| chr10:104951824-104951878 | ***Tmtc2*** | intra | C57BL/6 | 0.0226 (P.A) | 0.993 (P) |
| chr10:104968996-104969374 | ***Tmtc2*** | intra | C57BL/6 | 0.0226 (P.A) | 0.993 (P) |
| chr10:104976330-104976386 | ***Tmtc2*** | intra | BALB/c | 0.0226 (P.A) | 0.993 (P) |

| *Center of DMR (Mouse assembly mm8)* | *Gene Symbol* | *Position relative to Gene* | *Demethylated in* | *BMM BALB/c (0h)* | *BMM C57BL/6 (0h)* |
|---|---|---|---|---|---|
| chr10:105001660-105002240 | ***Tmtc2*** | distal | C57BL/6 | 0.0226 (P.A) | 0.993 (P) |
| chr10:105014390-105014549 | ***Tmtc2*** | distal | C57BL/6 | 0.0226 (P.A) | 0.993 (P) |
| chr10:122672434-122672816 | ***Fam19a2*** | intra | C57BL/6 | 69.92 (P) | 0.996 (A) |
| chr10:122708405-122708511 | ***Fam19a2*** | intra | C57BL/6 | 69.92 (P) | 0.996 (A) |
| chr10:122710469-122710529 | ***Fam19a2*** | intra | C57BL/6 | 69.92 (P) | 0.996 (A) |
| chr10:122822800-122823015 | ***Fam19a2*** | intra | BALB/c | 69.92 (P) | 0.996 (A) |
| chr10:122827614-122827671 | ***Fam19a2*** | intra | BALB/c | 69.92 (P) | 0.996 (A) |
| chr10:122849206-122849328 | ***Fam19a2*** | intra | C57BL/6 | 69.92 (P) | 0.996 (A) |
| chr10:122858308-122858831 | ***Fam19a2*** | intra | C57BL/6 | 69.92 (P) | 0.996 (A) |
| chr10:122859610-122860125 | ***Fam19a2*** | intra | C57BL/6 | 69.92 (P) | 0.996 (A) |
| chr10:122868108-122868477 | ***Fam19a2*** | intra | BALB/c | 69.92 (P) | 0.996 (A) |
| chr10:122904978-122905047 | ***Fam19a2*** | intra | BALB/c | 69.92 (P) | 0.996 (A) |
| chr10:122957032-122957347 | ***Fam19a2*** | intra | C57BL/6 | 69.92 (P) | 0.996 (A) |
| chr10:123131977-123132288 | ***Fam19a2*** | intra | BALB/c | 69.92 (P) | 0.996 (A) |
| chr10:123135244-123135391 | ***Fam19a2*** | intra | C57BL/6 | 69.92 (P) | 0.996 (A) |
| chr11:3092678-3093978 | ***Sfi1*** | prom | BALB/c | 16.38 (P) | 0.974 (P) |
| chr11:3128842-3128985 | ***Eif4enif1*** | intra | C57BL/6 | 1.48 (P) | 0.957 (P) |
| chr11:5857646-5857691 | ***Ykt6*** | intra | C57BL/6 | 1.141 (P) | 0.973 (P) |
| chr11:5911639-5911781 | ***Camk2b*** | intra | BALB/c | 0.0639 (A) | 0.992 (P) |
| chr11:5932196-5932242 | ***Camk2b*** | intra | C57BL/6 | 0.0639 (A) | 0.992 (P) |
| chr11:46600295-46600496 | ***Havcr1*** | intra | C57BL/6 | 0.19 (A) | 0.806 (P) |
| chr11:46707379-46707893 | ***Timd4*** | distal | BALB/c | 8.529 (P) | 0.967 (P.A) |
| chr11:88799550-88799596 | ***Coil*** | intra | BALB/c | 0.682 (P) | 0.999 (P) |
| chr11:88801073-88801240 | ***Coil*** | intra | BALB/c | 0.682 (P) | 0.999 (P) |
| chr11:88804568-88804628 | ***Coil*** | intra | BALB/c | 0.682 (P) | 0.999 (P) |
| chr11:90494121-90494462 | ***Tom1l1*** | intra | BALB/c | 6.307 (P) | 0.998 (P) |
| chr11:90523839-90524032 | ***Tom1l1*** | distal | C57BL/6 | 6.307 (P) | 0.998 (P) |
| chr11:90527633-90527930 | ***Tom1l1*** | distal | BALB/c | 6.307 (P) | 0.998 (P) |
| chr11:90538382-90538517 | ***Tom1l1*** | distal | BALB/c | 6.307 (P) | 0.998 (P) |
| chr11:90543530-90543948 | ***Tom1l1*** | distal | C57BL/6 | 6.307 (P) | 0.998 (P) |
| chr11:90546847-90546916 | ***Tom1l1*** | distal | BALB/c | 6.307 (P) | 0.998 (P) |
| chr12:15770109-15770577 | ***Trib2*** | distal | C57BL/6 | 0.0942 (A) | 0.891 (P) |
| chr12:15784239-15784299 | ***Trib2*** | distal | BALB/c | 0.0942 (A) | 0.891 (P) |
| chr12:15823593-15823638 | ***Trib2*** | intra | BALB/c | 0.0942 (A) | 0.891 (P) |
| chr12:15826824-15826870 | ***Trib2*** | intra | BALB/c | 0.0942 (A) | 0.891 (P) |
| chr12:45187434-45187920 | ***Nrcam*** | distal | C57BL/6 | 42.4 (P) | 0.942 (P) |
| chr12:45213241-45213388 | ***Nrcam*** | intra | C57BL/6 | 42.4 (P) | 0.942 (P) |
| chr12:79919494-79919852 | ***6330442E10Rik*** | intra | C57BL/6 | 0.944 (P) | 0.998 (P) |
| chr12:79968891-79969672 | ***Plekhh1*** | intra | C57BL/6 | 5.628 (P) | 0.998 (P) |
| chr12:79994386-79994739 | ***Plekhh1*** | intra | C57BL/6 | 5.628 (P) | 0.998 (P) |
| chr12:79995677-79995722 | ***Plekhh1*** | intra | BALB/c | 5.628 (P) | 0.998 (P) |
| chr12:80002205-80002717 | ***Pigh*** | intra | BALB/c | 1.281 (P) | 0.998 (P) |
| chr13:56269324-56269539 | ***Neurog1*** | distal | C57BL/6 | 0.756 (P) | 0.979 (P) |
| chr13:56313994-56314118 | ***Cxcl14*** | distal | BALB/c | 17.13 (P) | 0.952 (P) |
| chr13:56347924-56348198 | ***Cxcl14*** | distal | C57BL/6 | 17.13 (P) | 0.952 (P) |
| chr13:62858842-62859176 | ***Fbp2*** | intra | C57BL/6 | 1.207 (A) | 0.996 (A) |

| *Center of DMR (Mouse assembly mm8)* | *Gene Symbol* | *Position relative to Gene* | *Demethylated in* | *BMM BALB/c (0h)* | *BMM C57BL/6 (0h)* |
|---|---|---|---|---|---|
| chr13:67930408-67930468 | ***Zfp459*** | distal | BALB/c | 7.655 (P) | 0.937 (P.A) |
| chr13:76579670-76579730 | ***Ttc37*** | intra | C57BL/6 | 1.925 (P) | 0.939 (P) |
| chr13:109314064-109314489 | ***Ercc8*** | intra | BALB/c | 1.302 (P) | 0.998 (P) |
| chr13:109326652-109326712 | ***Elovl7*** | distal | BALB/c | 0.0586 (A) | 0.996 (P) |
| chr13:109416087-109416140 | ***Elovl7*** | distal | C57BL/6 | 0.0586 (A) | 0.996 (P) |
| chr14:19624572-19624712 | ***Plau*** | prom | C57BL/6 | 20.23 (P) | 0.974 (P) |
| chr14:50788103-50788154 | ***Ear6*** | distal | C57BL/6 | 0.863 (A) | 0.996 (A) |
| chr14:50809903-50810041 | ***Mett11d1*** | intra | C57BL/6 | 1.097 (P) | 1 (P) |
| chr14:50884020-50884348 | ***E130112L23Rik*** | distal | BALB/c | 1.953 (P) | 0.99 (P) |
| chr14:51053374-51053423 | ***Rpgrip1*** | intra | BALB/c | 0.21 (P) | 1 (P) |
| chr14:51060437-51060883 | ***Rpgrip1*** | intra | C57BL/6 | 0.21 (P) | 1 (P) |
| chr14:51081511-51082058 | ***Rpgrip1*** | intra | BALB/c | 0.21 (P) | 1 (P) |
| chr14:51104658-51105062 | ***Supt16h*** | intra | C57BL/6 | 0.652 (P) | 0.999 (P) |
| chr14:65636120-65636251 | ***Adra1a*** | intra | BALB/c | 5.283 (P) | 0.999 (P) |
| chr14:65651727-65652060 | ***Adra1a*** | intra | BALB/c | 5.283 (P) | 0.999 (P) |
| chr14:65710648-65711641 | ***Adra1a*** | distal | C57BL/6 | 5.283 (P) | 0.999 (P) |
| chr14:78033582-78034116 | ***1190002H23Rik*** | intra | C57BL/6 | 0.0526 (P) | 0.994 (P) |
| chr14:78037674-78038068 | ***1190002H23Rik*** | distal | BALB/c | 0.0526 (P) | 0.994 (P) |
| chr14:78042149-78042203 | ***1190002H23Rik*** | distal | BALB/c | 0.0526 (P) | 0.994 (P) |
| chr15:74776186-74776280 | ***Ly6e*** | distal | C57BL/6 | 1.38 (P) | 0.982 (P) |
| chr15:74783393-74783628 | ***Ly6e*** | intra | C57BL/6 | 1.38 (P) | 0.982 (P) |
| chr15:74790167-74790537 | ***Ly6e*** | distal | C57BL/6 | 1.38 (P) | 0.982 (P) |
| chr15:90288150-90288496 | ***Cpne8*** | distal | C57BL/6 | 39.53 (P) | 0.924 (P) |
| chr15:90314666-90314836 | ***Cpne8*** | distal | C57BL/6 | 39.53 (P) | 0.924 (P) |
| chr15:90335420-90335501 | ***Cpne8*** | intra | C57BL/6 | 39.53 (P) | 0.924 (P) |
| chr15:90478475-90478535 | ***Cpne8*** | intra | C57BL/6 | 39.53 (P) | 0.924 (P) |
| chr16:28833304-28834357 | ***1600021P15Rik*** | intra | C57BL/6 | 8.062 (P) | 0.998 (P) |
| chr16:38842237-38842369 | ***Igsf11*** | intra | BALB/c | 0.314 (P.A) | 0.984 (P) |
| chr16:38873534-38873736 | ***Igsf11*** | intra | BALB/c | 0.314 (P.A) | 0.984 (P) |
| chr16:38944175-38944451 | ***Igsf11*** | intra | C57BL/6 | 0.314 (P.A) | 0.984 (P) |
| chr16:85665872-85665932 | ***Adamts1*** | distal | C57BL/6 | 30.23 (P) | 0.987 (P.A) |
| chr17:3293172-3293498 | ***Tiam2*** | intra | C57BL/6 | 0.0958 (P) | 0.979 (P) |
| chr17:3334336-3334651 | ***Tiam2*** | intra | C57BL/6 | 0.0958 (P) | 0.979 (P) |
| chr17:3353108-3353259 | ***Tiam2*** | intra | C57BL/6 | 0.0958 (P) | 0.979 (P) |
| chr17:3355433-3355486 | ***Tiam2*** | intra | BALB/c | 0.0958 (P) | 0.979 (P) |
| chr17:3359059-3359104 | ***Tiam2*** | intra | BALB/c | 0.0958 (P) | 0.979 (P) |
| chr17:3360576-3360633 | ***Tiam2*** | intra | C57BL/6 | 0.0958 (P) | 0.979 (P) |
| chr17:3486560-3486832 | ***Tfb1m*** | intra | C57BL/6 | 1.298 (A) | 0.984 (A) |
| chr17:3489017-3489250 | ***Tfb1m*** | intra | C57BL/6 | 1.298 (A) | 0.984 (A) |
| chr17:3505060-3505778 | ***Tfb1m*** | intra | C57BL/6 | 1.298 (A) | 0.984 (A) |
| chr17:30301717-30302316 | ***Btbd9*** | intra | C57BL/6 | 11.81 (P) | 0.995 (P) |
| chr17:30462176-30462337 | ***Dnahc8*** | intra | C57BL/6 | 4.678 (P) | 0.996 (P) |
| chr17:33787792-33788060 | ***Psmb9*** | distal | C57BL/6 | 0.992 (P) | 0.994 (P) |
| chr17:33914667-33915150 | ***H2-Eb1*** | intra | C57BL/6 | 1.578 (P) | 0.997 (P) |
| chr17:33917083-33917555 | ***H2-Eb1*** | intra | C57BL/6 | 1.578 (P) | 0.997 (P) |
| chr17:33948318-33948367 | ***H2-Ea*** | distal | C57BL/6 | 319.1 (P) | 0.996 (A) |

| *Center of DMR (Mouse assembly mm8)* | *Gene Symbol* | *Position relative to Gene* | *Demethylated in* | *BMM BALB/c (0h)* | *BMM C57BL/6 (0h)* |
|---|---|---|---|---|---|
| chr17:34637963-34638443 | ***Msh5*** | intra | C57BL/6 | 56.31 (P) | 0.994 (P) |
| chr17:43287601-43287829 | ***Rcan2*** | intra | C57BL/6 | 17.43 (P) | 0.996 (A) |
| chr17:43292477-43292537 | ***Rcan2*** | intra | C57BL/6 | 17.43 (P) | 0.996 (A) |
| chr17:43295716-43295900 | ***Rcan2*** | intra | C57BL/6 | 17.43 (P) | 0.996 (A) |
| chr17:43343826-43344046 | ***Rcan2*** | intra | BALB/c | 17.43 (P) | 0.996 (A) |
| chr17:43395146-43395209 | ***Rcan2*** | intra | BALB/c | 17.43 (P) | 0.996 (A) |
| chr17:43417114-43417163 | ***Rcan2*** | intra | C57BL/6 | 17.43 (P) | 0.996 (A) |
| chr17:43475365-43475425 | ***Rcan2*** | intra | BALB/c | 17.43 (P) | 0.996 (A) |
| chr17:43485699-43486569 | ***Rcan2*** | intra | BALB/c | 17.43 (P) | 0.996 (A) |
| chr17:43490290-43490887 | ***Rcan2*** | intra | BALB/c | 17.43 (P) | 0.996 (A) |
| chr18:9209850-9210313 | ***Fzd8*** | distal | C57BL/6 | 5.553 (P) | 0.937 (P) |
| chr18:9214393-9214977 | ***Fzd8*** | intra | C57BL/6 | 5.553 (P) | 0.937 (P) |
| chr18:58681880-58682279 | ***Slc27a6*** | intra | C57BL/6 | 8.848 (P) | 0.997 (A) |
| chr18:58686133-58686695 | ***Slc27a6*** | intra | C57BL/6 | 8.848 (P) | 0.997 (A) |
| chr19:5687110-5687363 | ***Pcnxl3*** | intra | BALB/c | 1.089 (P) | 0.995 (P) |
| chr19:10054676-10054898 | ***Best1*** | intra | BALB/c | 0.111 (P) | 0.997 (P) |
| chr19:10059640-10059975 | ***Best1*** | intra | C57BL/6 | 0.111 (P) | 0.997 (P) |
| chr19:17402472-17402791 | ***Gcnt1*** | intra | C57BL/6 | 0.0663 (P) | 0.994 (P) |
| chr19:17411445-17411968 | ***Gcnt1*** | intra | BALB/c | 0.0663 (P) | 0.994 (P) |
| chr19:17417594-17417742 | ***Gcnt1*** | intra | BALB/c | 0.0663 (P) | 0.994 (P) |
| chr19:17424366-17424594 | ***Gcnt1*** | prom | C57BL/6 | 0.0663 (P) | 0.994 (P) |
| chr19:24146132-24146183 | ***Tjp2*** | distal | BALB/c | 0.0519 (P) | 0.974 (P) |
| chr19:24221222-24221543 | ***Tjp2*** | intra | C57BL/6 | 0.0519 (P) | 0.974 (P) |
| chr19:24229262-24229401 | ***Tjp2*** | intra | BALB/c | 0.0519 (P) | 0.974 (P) |
| chr19:24310167-24310227 | ***Tjp2*** | distal | C57BL/6 | 0.0519 (P) | 0.974 (P) |
| chr19:24331447-24331788 | ***Fxn*** | intra | BALB/c | 0.85 (P) | 0.999 (P) |
| chr19:24337430-24337560 | ***Fxn*** | intra | BALB/c | 0.85 (P) | 0.999 (P) |
| chr19:25277140-25278032 | ***Ankrd15*** | distal | BALB/c | 20.66 (P) | 0.912 (P.A) |
| chr19:25346036-25346164 | ***Ankrd15*** | intra | C57BL/6 | 20.66 (P) | 0.912 (P.A) |
| chr19:25390590-25390638 | ***Ankrd15*** | intra | C57BL/6 | 20.66 (P) | 0.912 (P.A) |
| chr19:25499075-25499142 | ***Ankrd15*** | intra | BALB/c | 20.66 (P) | 0.912 (P.A) |
| chr19:25499563-25499613 | ***Ankrd15*** | intra | BALB/c | 20.66 (P) | 0.912 (P.A) |
| chr19:47809044-47809476 | ***6330577E15Rik*** | intra | BALB/c | 1.013 (P) | 1 (P) |
| chr19:47919990-47920050 | ***Gsto2*** | intra | BALB/c | 0.946 (P) | 0.996 (P) |
| chr19:47932361-47932413 | ***Gsto2*** | intra | BALB/c | 0.946 (P) | 0.996 (P) |
| chr19:47956327-47956551 | ***Itprip*** | intra | C57BL/6 | | |
| chr19:47958144-47958529 | ***Itprip*** | intra | C57BL/6 | | |
| chr19:47961787-47962003 | ***Itprip*** | intra | BALB/c | | |
| chr2:118831692-118831737 | ***1200015F23Rik*** | intra | C57BL/6 | 1.389 (P) | 0.968 (P) |
| chr2:118870105-118870259 | ***Dnajc17*** | intra | C57BL/6 | 1.771 (P) | 0.982 (P) |
| chr2:118921263-118922015 | ***Ppp1r14d*** | prom | C57BL/6 | 1.414 (A) | 0.996 (A) |
| chr2:118931167-118931543 | ***Spint1*** | intra | C57BL/6 | 0.0282 (P) | 0.839 (P) |
| chr2:118933687-118934010 | ***Spint1*** | intra | C57BL/6 | 0.0282 (P) | 0.839 (P) |
| chr2:120043669-120043809 | ***Vps39*** | intra | BALB/c | 1.379 (P) | 1 (P) |
| chr2:121917801-121918164 | ***Sord*** | distal | C57BL/6 | 0.147 (P) | 0.971 (P) |
| chr2:129008854-129009320 | ***Il1a*** | distal | C57BL/6 | 1.584 (A) | 0.996 (A) |

| *Center of DMR (Mouse assembly mm8)* | *Gene Symbol* | *Position relative to Gene* | *Demethylated in* | *BMM BALB/c (0h)* | *BMM C57BL/6 (0h)* |
|---|---|---|---|---|---|
| chr2:129060442-129060527 | ***Il1b*** | intra | BALB/c | 10.84 (P) | 0.997 (P) |
| chr2:129084712-129085352 | ***Il1b*** | distal | C57BL/6 | 10.84 (P) | 0.997 (P) |
| chr2:129095204-129095388 | ***Il1b*** | distal | C57BL/6 | 10.84 (P) | 0.997 (P) |
| chr2:165163847-165163893 | ***Slc13a3*** | intra | BALB/c | 0.0413 (P.A) | 0.999 (P) |
| chr2:165165074-165165182 | ***Slc13a3*** | prom | BALB/c | 0.0413 (P.A) | 0.999 (P) |
| chr2:180916370-180916541 | ***9230112E08Rik*** | distal | C57BL/6 | 0.995 (P) | 0.992 (P) |
| chr2:180924755-180925097 | ***9230112E08Rik*** | distal | C57BL/6 | 0.995 (P) | 0.992 (P) |
| chr3:54505131-54505191 | ***Postn*** | distal | BALB/c | 2.843 (P) | 1 (P.A) |
| chr3:117362502-117362562 | ***D3Bwg0562e*** | distal | BALB/c | 20.58 (P) | 0.996 (A) |
| chr3:117363368-117363427 | ***D3Bwg0562e*** | distal | BALB/c | 20.58 (P) | 0.996 (A) |
| chr3:117378454-117378604 | ***D3Bwg0562e*** | distal | BALB/c | 20.58 (P) | 0.996 (A) |
| chr3:117390058-117390183 | ***D3Bwg0562e*** | distal | BALB/c | 20.58 (P) | 0.996 (A) |
| chr3:126918409-126918466 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:126932154-126932803 | ***Ank2*** | intra | BALB/c | 6.139 (P) | 0.999 (P.A) |
| chr3:126961500-126961737 | ***Ank2*** | intra | BALB/c | 6.139 (P) | 0.999 (P.A) |
| chr3:126964886-126964946 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:126990100-126990231 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:126997578-126997623 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127002910-127003059 | ***Ank2*** | intra | BALB/c | 6.139 (P) | 0.999 (P.A) |
| chr3:127058579-127058783 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127064288-127064344 | ***Ank2*** | intra | BALB/c | 6.139 (P) | 0.999 (P.A) |
| chr3:127115515-127116082 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127144143-127144203 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127144544-127144604 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127152673-127152718 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127157369-127157422 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127161553-127161690 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127181064-127181574 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127183561-127183610 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127204193-127204468 | ***Ank2*** | intra | BALB/c | 6.139 (P) | 0.999 (P.A) |
| chr3:127220202-127220609 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127228341-127228401 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127234728-127235174 | ***Ank2*** | intra | BALB/c | 6.139 (P) | 0.999 (P.A) |
| chr3:127242061-127242747 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127260837-127261164 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127261581-127261716 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127308653-127309047 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127347726-127347863 | ***Ank2*** | intra | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127384454-127384514 | ***Ank2*** | intra | BALB/c | 6.139 (P) | 0.999 (P.A) |
| chr3:127401141-127401603 | ***Ank2*** | prom | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127443926-127443972 | ***Ank2*** | distal | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:127474127-127474259 | ***Ank2*** | distal | BALB/c | 6.139 (P) | 0.999 (P.A) |
| chr3:127492535-127492881 | ***Ank2*** | distal | C57BL/6 | 6.139 (P) | 0.999 (P.A) |
| chr3:131987012-131987072 | ***Dkk2*** | distal | BALB/c | 8.107 (P) | 0.99 (A) |
| chr3:142400177-142400232 | ***Gbp5*** | distal | BALB/c | 1.205 (P) | 1 (P) |
| chr3:142401448-142401499 | ***Gbp5*** | distal | BALB/c | 1.205 (P) | 1 (P) |

| *Center of DMR (Mouse assembly mm8)* | *Gene Symbol* | *Position relative to Gene* | *Demethylated in* | *BMM BALB/c (0h)* | *BMM C57BL/6 (0h)* |
|---|---|---|---|---|---|
| chr3:142548523-142548645 | ***Gbp2*** | distal | C57BL/6 | 0.743 (P) | 1 (P) |
| chr4:46555277-46556225 | ***Trim14*** | intra | C57BL/6 | 1.145 (P) | 0.995 (P) |
| chr4:46594578-46596131 | ***Coro2a*** | intra | C57BL/6 | 0.209 (P) | 0.997 (P) |
| chr4:46610429-46610801 | ***Coro2a*** | intra | C57BL/6 | 0.209 (P) | 0.997 (P) |
| chr4:46620908-46622454 | ***Coro2a*** | intra | C57BL/6 | 0.209 (P) | 0.997 (P) |
| chr4:46645788-46645910 | ***Tbc1d2*** | intra | C57BL/6 | 1.5 (P) | 0.993 (P) |
| chr4:46661251-46661373 | ***Tbc1d2*** | intra | C57BL/6 | 1.5 (P) | 0.993 (P) |
| chr4:66324635-66324996 | ***Tlr4*** | intra | C57BL/6 | 1.542 (P) | 0.994 (P) |
| chr4:75614171-75614231 | ***Ptprd*** | intra | BALB/c | 44.94 (P) | 0.996 (A) |
| chr4:75643880-75643940 | ***Ptprd*** | intra | BALB/c | 44.94 (P) | 0.996 (A) |
| chr4:75666515-75667215 | ***Ptprd*** | intra | C57BL/6 | 44.94 (P) | 0.996 (A) |
| chr4:75709451-75709627 | ***Ptprd*** | intra | C57BL/6 | 44.94 (P) | 0.996 (A) |
| chr4:75749082-75749313 | ***Ptprd*** | intra | C57BL/6 | 44.94 (P) | 0.996 (A) |
| chr4:75762858-75763384 | ***Ptprd*** | intra | C57BL/6 | 44.94 (P) | 0.996 (A) |
| chr4:75781594-75781862 | ***Ptprd*** | intra | BALB/c | 44.94 (P) | 0.996 (A) |
| chr4:75852063-75852194 | ***Ptprd*** | intra | C57BL/6 | 44.94 (P) | 0.996 (A) |
| chr4:75857675-75858015 | ***Ptprd*** | intra | C57BL/6 | 44.94 (P) | 0.996 (A) |
| chr4:75903639-75904077 | ***Ptprd*** | intra | BALB/c | 44.94 (P) | 0.996 (A) |
| chr4:75909398-75909566 | ***Ptprd*** | intra | BALB/c | 44.94 (P) | 0.996 (A) |
| chr4:75970799-75971683 | ***Ptprd*** | intra | C57BL/6 | 44.94 (P) | 0.996 (A) |
| chr4:75998019-75998274 | ***Ptprd*** | intra | BALB/c | 44.94 (P) | 0.996 (A) |
| chr4:76002246-76002407 | ***Ptprd*** | intra | C57BL/6 | 44.94 (P) | 0.996 (A) |
| chr4:76014633-76014693 | ***Ptprd*** | intra | BALB/c | 44.94 (P) | 0.996 (A) |
| chr4:76042627-76042753 | ***Ptprd*** | intra | BALB/c | 44.94 (P) | 0.996 (A) |
| chr4:95020698-95021255 | ***Fggy*** | distal | C57BL/6 | 0.15 (P) | 0.98 (P) |
| chr4:95049158-95049994 | ***Fggy*** | prom | C57BL/6 | 0.15 (P) | 0.98 (P) |
| chr4:95071653-95071921 | ***Fggy*** | intra | BALB/c | 0.15 (P) | 0.98 (P) |
| chr4:95074879-95075609 | ***Fggy*** | intra | C57BL/6 | 0.15 (P) | 0.98 (P) |
| chr4:95255640-95255700 | ***Fggy*** | intra | C57BL/6 | 0.15 (P) | 0.98 (P) |
| chr4:95281652-95281871 | ***Fggy*** | intra | BALB/c | 0.15 (P) | 0.98 (P) |
| chr4:117954421-117954479 | ***Mpl*** | intra | BALB/c | 0.664 (A) | 0.939 (P) |
| chr4:117998531-117998591 | ***Tie1*** | distal | BALB/c | 13.28 (P) | 0.958 (A) |
| chr4:118000679-118000739 | ***Tie1*** | distal | C57BL/6 | 13.28 (P) | 0.958 (A) |
| chr4:136116934-136117007 | ***Ephb2*** | distal | BALB/c | 0.583 (A) | 0.983 (P.A) |
| chr4:136120484-136120813 | ***Ephb2*** | distal | BALB/c | 0.583 (A) | 0.983 (P.A) |
| chr4:136127607-136127655 | ***Ephb2*** | distal | BALB/c | 0.583 (A) | 0.983 (P.A) |
| chr4:136157517-136157754 | ***C1qb*** | prom | C57BL/6 | 0.146 (P) | 0.995 (P) |
| chr4:136160301-136161004 | ***C1qb*** | distal | BALB/c | 0.146 (P) | 0.995 (P) |
| chr4:136188412-136188847 | ***C1qa*** | distal | BALB/c | 0.383 (P) | 0.999 (P) |
| chr4:139992658-139992703 | ***Rcc2*** | intra | C57BL/6 | 1.033 (P) | 0.999 (P) |
| chr4:140001888-140002184 | ***Padi6*** | intra | BALB/c | 0.781 (A) | 0.992 (P.A) |
| chr4:140037365-140037761 | ***Padi4*** | intra | BALB/c | 0.0265 (P) | 0.989 (P) |
| chr4:140053810-140053855 | ***Padi4*** | distal | C57BL/6 | 0.0265 (P) | 0.989 (P) |
| chr4:140066734-140066816 | ***Padi3*** | intra | BALB/c | 0.443 (P.A) | 0.999 (P) |
| chr4:140075656-140075701 | ***Padi3*** | intra | C57BL/6 | 0.443 (P.A) | 0.999 (P) |
| chr4:140086347-140087247 | ***Padi1*** | intra | C57BL/6 | 0.756 (P) | 0.999 (P) |

| *Center of DMR (Mouse assembly mm8)* | *Gene Symbol* | *Position relative to Gene* | *Demethylated in* | *BMM BALB/c (0h)* | *BMM C57BL/6 (0h)* |
|---|---|---|---|---|---|
| chr4:140088960-140089479 | ***Padi1*** | intra | BALB/c | 0.756 (P) | 0.999 (P) |
| chr4:140098544-140098705 | ***Padi1*** | intra | C57BL/6 | 0.756 (P) | 0.999 (P) |
| chr4:140100052-140100287 | ***Padi1*** | intra | C57BL/6 | 0.756 (P) | 0.999 (P) |
| chr4:140102179-140102439 | ***Padi1*** | intra | C57BL/6 | 0.756 (P) | 0.999 (P) |
| chr4:140117495-140117765 | ***Padi1*** | prom | BALB/c | 0.756 (P) | 0.999 (P) |
| chr4:140150659-140150706 | ***Padi1*** | distal | BALB/c | 0.756 (P) | 0.999 (P) |
| chr4:140834500-140834615 | ***Fblim1*** | distal | BALB/c | 0.126 (P) | 0.991 (P) |
| chr4:140848488-140848640 | ***Fblim1*** | distal | BALB/c | 0.126 (P) | 0.991 (P) |
| chr4:140887217-140887262 | ***Tmem82*** | intra | C57BL/6 | 1.56 (P) | 1 (P) |
| chr4:140902197-140902509 | ***Slc25a34*** | intra | BALB/c | 0.779 (P) | 0.99 (P) |
| chr4:146455519-146455683 | ***2610305D13Rik*** | distal | C57BL/6 | 0.0885 (A) | 0.989 (P) |
| chr4:147309756-147309900 | ***Frap1*** | intra | C57BL/6 | 1.22 (P) | 0.998 (P) |
| chr4:147344719-147345290 | ***Angptl7*** | prom | BALB/c | 71.66 (P) | 0.998 (A) |
| chr4:153165776-153166060 | ***Prdm16*** | intra | BALB/c | 0.0185 (A) | 0.988 (P) |
| chr4:153374889-153374938 | ***Prdm16*** | intra | C57BL/6 | 0.0185 (A) | 0.988 (P) |
| chr4:153377867-153377925 | ***Prdm16*** | intra | C57BL/6 | 0.0185 (A) | 0.988 (P) |
| chr4:153397287-153397825 | ***Prdm16*** | intra | C57BL/6 | 0.0185 (A) | 0.988 (P) |
| chr4:153416659-153417244 | ***Prdm16*** | intra | BALB/c | 0.0185 (A) | 0.988 (P) |
| chr4:153521279-153522015 | ***Actrt2*** | distal | C57BL/6 | 0.524 (A) | 0.84 (P.A) |
| chr5:31235102-31235151 | ***Preb*** | intra | BALB/c | 0.987 (P) | 0.996 (P) |
| chr5:35229612-35229843 | ***A930005I04Rik*** | distal | C57BL/6 | 0.143 (P) | 0.999 (P) |
| chr5:53010351-53010409 | ***Pi4k2b*** | distal | C57BL/6 | 7.234 (P) | 0.999 (P) |
| chr5:53012858-53013017 | ***Pi4k2b*** | distal | C57BL/6 | 7.234 (P) | 0.999 (P) |
| chr5:53013951-53014431 | ***Pi4k2b*** | distal | C57BL/6 | 7.234 (P) | 0.999 (P) |
| chr5:53026418-53026767 | ***Pi4k2b*** | distal | C57BL/6 | 7.234 (P) | 0.999 (P) |
| chr5:53049276-53049634 | ***Pi4k2b*** | intra | BALB/c | 7.234 (P) | 0.999 (P) |
| chr5:53058948-53058993 | ***Pi4k2b*** | distal | BALB/c | 7.234 (P) | 0.999 (P) |
| chr5:53061668-53062228 | ***Pi4k2b*** | distal | C57BL/6 | 7.234 (P) | 0.999 (P) |
| chr5:53063596-53064069 | ***Pi4k2b*** | distal | BALB/c | 7.234 (P) | 0.999 (P) |
| chr5:66959934-66959982 | ***Uchl1*** | intra | C57BL/6 | 0.0211 (P) | 0.998 (P) |
| chr5:67640239-67640299 | ***Slc30a9*** | distal | BALB/c | 1.548 (P) | 0.99 (P) |
| chr5:67676687-67677026 | ***Slc30a9*** | distal | BALB/c | 1.548 (P) | 0.99 (P) |
| chr5:67680207-67680368 | ***Slc30a9*** | distal | C57BL/6 | 1.548 (P) | 0.99 (P) |
| chr5:104720508-104721242 | ***Pkd2*** | intra | C57BL/6 | 25.9 (P) | 1 (P) |
| chr5:104725504-104725800 | ***Pkd2*** | intra | C57BL/6 | 25.9 (P) | 1 (P) |
| chr5:104733382-104733438 | ***Pkd2*** | intra | C57BL/6 | 25.9 (P) | 1 (P) |
| chr6:5104955-5105188 | ***Ppp1r9a*** | intra | C57BL/6 | 10.8 (P) | 0.988 (P) |
| chr6:5333111-5334506 | ***Asb4*** | distal | C57BL/6 | 0.0642 (P.A) | 0.997 (P) |
| chr6:29688651-29688794 | ***Smo*** | intra | C57BL/6 | 7.301 (P) | 0.984 (P) |
| chr6:29716478-29716561 | ***4631427C17Rik*** | distal | BALB/c | 1.345 (P) | 0.827 (P) |
| chr6:93630855-93630915 | ***Magi1*** | distal | C57BL/6 | 9.72 (P) | 1 (P.A) |
| chr6:93691281-93691444 | ***Magi1*** | intra | BALB/c | 9.72 (P) | 1 (P.A) |
| chr6:93694981-93695577 | ***Magi1*** | intra | C57BL/6 | 9.72 (P) | 1 (P.A) |
| chr6:93704332-93704392 | ***Magi1*** | intra | C57BL/6 | 9.72 (P) | 1 (P.A) |
| chr6:93713202-93713457 | ***Magi1*** | intra | C57BL/6 | 9.72 (P) | 1 (P.A) |
| chr6:93735273-93735333 | ***Magi1*** | intra | BALB/c | 9.72 (P) | 1 (P.A) |

| *Center of DMR (Mouse assembly mm8)* | *Gene Symbol* | *Position relative to Gene* | *Demethylated in* | *BMM BALB/c (0h)* | *BMM C57BL/6 (0h)* |
|---|---|---|---|---|---|
| chr6:93755446-93756328 | ***Magi1*** | intra | C57BL/6 | 9.72 (P) | 1 (P.A) |
| chr6:93768673-93769578 | ***Magi1*** | intra | C57BL/6 | 9.72 (P) | 1 (P.A) |
| chr6:93814158-93814213 | ***Magi1*** | intra | C57BL/6 | 9.72 (P) | 1 (P.A) |
| chr6:93823430-93823772 | ***Magi1*** | intra | C57BL/6 | 9.72 (P) | 1 (P.A) |
| chr6:94029111-94029381 | ***Magi1*** | intra | C57BL/6 | 9.72 (P) | 1 (P.A) |
| chr6:94085475-94085534 | ***Magi1*** | intra | BALB/c | 9.72 (P) | 1 (P.A) |
| chr6:94115136-94115260 | ***Magi1*** | intra | BALB/c | 9.72 (P) | 1 (P.A) |
| chr6:122566409-122566463 | ***Apobec1*** | intra | C57BL/6 | 2.923 (P) | 1 (P) |
| chr6:122593442-122593497 | ***Dppa3*** | intra | C57BL/6 | 0.0522 (P) | 0.998 (P) |
| chr6:122599007-122599338 | ***Dppa3*** | distal | BALB/c | 0.0522 (P) | 0.998 (P) |
| chr6:129559721-129559770 | ***Klrd1*** | intra | BALB/c | 0.785 (P) | 0.99 (P) |
| chr6:134647920-134648332 | ***Loh12cr1*** | intra | BALB/c | 0.551 (P) | 1 (P) |
| chr6:134652859-134653242 | ***Loh12cr1*** | intra | C57BL/6 | 0.551 (P) | 1 (P) |
| chr6:134666768-134667057 | ***Loh12cr1*** | intra | C57BL/6 | 0.551 (P) | 1 (P) |
| chr6:134684992-134685037 | ***Dusp16*** | intra | BALB/c | 6.056 (P) | 0.989 (P) |
| chr6:134711215-134711398 | ***Dusp16*** | intra | BALB/c | 6.056 (P) | 0.989 (P) |
| chr6:134723586-134723979 | ***Dusp16*** | intra | C57BL/6 | 6.056 (P) | 0.989 (P) |
| chr6:136898012-136898210 | ***Arhgdib*** | intra | C57BL/6 | 0.933 (P) | 0.997 (P) |
| chr6:136920044-136920089 | ***Pde6h*** | intra | BALB/c | 0.125 (A) | 0.999 (P) |
| chr6:136962558-136962606 | ***Pde6h*** | distal | C57BL/6 | 0.125 (A) | 0.999 (P) |
| chr6:136920044-136920089 | ***Pde6h*** | intra | BALB/c | 0.125 (A) | 0.999 (P) |
| chr6:136962558-136962606 | ***Pde6h*** | distal | C57BL/6 | 0.125 (A) | 0.999 (P) |
| chr7:4064078-4064123 | ***Eps8l1*** | intra | BALB/c | 0.0861 (A) | 0.979 (P) |
| chr7:4073211-4073587 | ***Eps8l1*** | intra | BALB/c | 0.0861 (A) | 0.979 (P) |
| chr7:4073898-4074764 | ***Eps8l1*** | intra | BALB/c | 0.0861 (A) | 0.979 (P) |
| chr7:4081211-4081948 | ***Eps8l1*** | intra | BALB/c | 0.0861 (A) | 0.979 (P) |
| chr7:4468706-4470122 | ***Isoc2b*** | prom | C57BL/6 | 0.188 (P) | 0.987 (P) |
| chr7:4476223-4476274 | ***Isoc2b*** | distal | C57BL/6 | 0.188 (P) | 0.987 (P) |
| chr7:15421071-15421190 | ***C5ar1*** | distal | BALB/c | 0.37 (P) | 0.998 (P) |
| chr7:28818237-28818282 | ***Ryr1*** | intra | BALB/c | 13.91 (P) | 1 (P) |
| chr7:28863980-28864027 | ***1110006G06Rik*** | intra | BALB/c | 1.008 (P) | 1 (P) |
| chr7:29665315-29665376 | ***Zfp74*** | intra | BALB/c | 1.336 (P) | 0.999 (P) |
| chr7:29692014-29692149 | ***Zfp568*** | prom | C57BL/6 | | |
| chr7:37961808-37962602 | ***Pop4*** | distal | BALB/c | 0.1 (P) | 0.999 (P) |
| chr7:80210572-80210730 | ***Unc45a*** | intra | C57BL/6 | 2.327 (P) | 0.999 (P) |
| chr7:99033917-99034043 | ***Dgat2*** | intra | C57BL/6 | 11.39 (P) | 0.99 (P) |
| chr7:99037518-99037643 | ***Dgat2*** | intra | BALB/c | 11.39 (P) | 0.99 (P) |
| chr7:99064892-99065884 | ***Dgat2*** | distal | C57BL/6 | 11.39 (P) | 0.99 (P) |
| chr7:119504588-119504967 | ***Thumpd1*** | distal | C57BL/6 | 0.01 (A) | 0.996 (P) |
| chr7:119506866-119506918 | ***Thumpd1*** | distal | BALB/c | 0.01 (A) | 0.996 (P) |
| chr7:119994863-119995241 | ***Abca14*** | prom | C57BL/6 | 1.207 (A) | 0.996 (A) |
| chr7:120010482-120010849 | ***Abca14*** | intra | BALB/c | 1.207 (A) | 0.996 (A) |
| chr7:125348874-125348919 | ***Il4ra*** | intra | BALB/c | 1.04 (P) | 0.999 (P) |
| chr7:125398313-125398366 | ***Il21r*** | intra | C57BL/6 | 0.724 (P) | 0.99 (P) |
| chr7:142430532-142430685 | ***Igf2*** | distal | BALB/c | 1.098 (M.A) | 0.923 (P) |
| chr8:73561622-73562051 | ***Gdf15*** | distal | C57BL/6 | 0.0791 (P) | 0.988 (P) |

| *Center of DMR (Mouse assembly mm8)* | *Gene Symbol* | *Position relative to Gene* | *Demethylated in* | *BMM BALB/c (0h)* | *BMM C57BL/6 (0h)* |
|---|---|---|---|---|---|
| chr8:124746984-124747033 | ***Slc7a5*** | distal | BALB/c | 3.167 (P) | 0.998 (P) |
| chr8:124747942-124748185 | ***Slc7a5*** | distal | BALB/c | 3.167 (P) | 0.998 (P) |
| chr8:124755205-124755348 | ***Slc7a5*** | distal | BALB/c | 3.167 (P) | 0.998 (P) |
| chr8:124757601-124757653 | ***Slc7a5*** | distal | C57BL/6 | 3.167 (P) | 0.998 (P) |
| chr8:124758800-124759357 | ***Slc7a5*** | distal | BALB/c | 3.167 (P) | 0.998 (P) |
| chr8:124760437-124760974 | ***Slc7a5*** | distal | BALB/c | 3.167 (P) | 0.998 (P) |
| chr8:124768035-124768493 | ***Slc7a5*** | intra | C57BL/6 | 3.167 (P) | 0.998 (P) |
| chr8:124774258-124775049 | ***Slc7a5*** | intra | C57BL/6 | 3.167 (P) | 0.998 (P) |
| chr8:124798597-124798956 | ***BC048644*** | intra | C57BL/6 | | |
| chr8:124823004-124823197 | ***Car5a*** | intra | C57BL/6 | 0.693 (A) | 0.973 (P.A) |
| chr8:127227407-127227452 | ***Galnt2*** | intra | BALB/c | 1.505 (P) | 0.999 (P) |
| chr8:127276982-127277478 | ***Pgbd5*** | intra | BALB/c | 0.0115 (A) | 1 (P) |
| chr8:127283682-127283870 | ***Pgbd5*** | intra | C57BL/6 | 0.0115 (A) | 1 (P) |
| chr8:127301989-127302034 | ***Pgbd5*** | intra | C57BL/6 | 0.0115 (A) | 1 (P) |
| chr8:127317898-127318260 | ***Pgbd5*** | intra | BALB/c | 0.0115 (A) | 1 (P) |
| chr8:127359130-127359175 | ***Pgbd5*** | distal | C57BL/6 | 0.0115 (A) | 1 (P) |
| chr8:127361009-127361074 | ***Pgbd5*** | distal | C57BL/6 | 0.0115 (A) | 1 (P) |
| chr8:127364656-127364799 | ***Pgbd5*** | distal | BALB/c | 0.0115 (A) | 1 (P) |
| chr9:49187501-49188250 | ***Ttc12*** | distal | C57BL/6 | 7.06 (P) | 0.998 (P) |
| chr9:49192161-49192634 | ***Ttc12*** | intra | C57BL/6 | 7.06 (P) | 0.998 (P) |
| chr9:49237091-49237345 | ***Ttc12*** | intra | BALB/c | 7.06 (P) | 0.998 (P) |
| chr9:50769119-50769184 | ***Snf1lk2*** | distal | C57BL/6 | 1.375 (P) | 0.966 (P) |
| chr9:51783122-51783240 | ***Rdx*** | distal | C57BL/6 | 0.971 (P) | 1 (P) |
| chr9:51850201-51850259 | ***Rdx*** | distal | C57BL/6 | 0.971 (P) | 1 (P) |
| chr9:51851802-51851859 | ***Rdx*** | distal | C57BL/6 | 0.971 (P) | 1 (P) |
| chr9:72207341-72207401 | ***BC065403*** | intra | BALB/c | | |
| chr9:102967337-102967700 | ***Rab6b*** | distal | C57BL/6 | 0.0133 (A) | 0.988 (P) |
| chr9:102971565-102971612 | ***Rab6b*** | intra | BALB/c | 0.0133 (A) | 0.988 (P) |
| chrX:119220336-119220381 | ***3110007F17Rik*** | prom | C57BL/6 | 0.0515 (A) | 0.992 (P) |

# 12.4 Appendix IV – Sequence alignments

**A**

***md_1***

chr8: 124,744,953-124,745,822

```
C57BL/6  ----------------------------------GAATGCCTAGGAGGGAA-CGCCCAGATGGTCAAGGAGTACTGACACGTCCTACGTCACTGTACACTG
BALB/c   ----------------------------------GAGATGCTAGGAGGGAA-CGCCCAGATGGTCAAGGAGTACTGACACGTCCTACGACACTGTACACCG

C57BL/6  AGCATCACACAGTGCATGGGGACCGTGGCTGCTCTCAGCAGGTGGCAGAGCAGGGCCAAGAGGGAATCAAAATCTTAACCTCCTTCATGAACCAGAAGTA
BALB/c   AGCATCACACGGTGCATGGGGACCGTGGCTGCTCTCAGCATGTGGCAGAGT----------GAATCAAAATCTTAACCTCCTTCATGAACCAGAAGTA

C57BL/6  GCAATGGAAAAGAAATTCCAGGGAGGA-TGGGAGGTGGTGGCAGGCTCAGCATGGGCTGGGGGACATGGCGGCTCTCTCCCTCTGTCCTCCCTTACATAT
BALB/c   GCAATGGAAGAGAAATTCCAGGGAGGACTGGGAGGTGGTGGCAGGCTCAGCCTGGGCTGGGGGACATGGCGGCTCTCTCCCTCTGTCCTCCCTTCCATAT

C57BL/6  GGC----------CCCCCCTGGTCTGCACCTCTGTGCTCTAACTACCCCGGGGACCTTTCTGTTCTTACATGTCACTAATCAGAAAATTGGGCTGGAGAG
BALB/c   GGCGGCGCCCCCCCCCCCCCGTCTGCACCTCTGTGCTCTAACTGCCCCAGGGACCTTTCTGTTCTTACATGCCACTAATCAGAAAATTGGGCTGGAGAG

C57BL/6  ATGGCTCACAGTTAAGAGAACTGGCTGCTCTTCCACAGGACCCAGTGTTAATTCCCAGCGCCCACAGGAAGCTCACAACCATCTGTAACTCTAGCACTAG
BALB/c   ATGGCTCACAGTTAAGAGAACTGGCTGCTCTTCCACAGGACCCAGTGTTAATTCCCAGCGCCCACAGGAAGCTCACAACCATCTGTAACTCTAGCACTAG

C57BL/6  GGGAATCCAATGCCCTTAGTGCCCGGCGAACGCATGATGCACAGGCATATATGCAGACAAATCCCCCATGCACATAAAGAATGGGGAGGAAGAGGAGGGG
BALB/c   GGGA-TCCAATGCCCTTAGTGCCCGGCATACGCATGATGCACAGGCATATATGTAGACAAATCCCCCATGCACATAAAGAATGGGGAGGAAGAGGAGGGG

C57BL/6  GGGCAAGGGGA-GGAGAAGGGGAGGGGGATAGGGAGGAGGAAGAGAAAGAAGAAGGGGAAGGAAGAAGAGGAGGGGGAGAGAAAAAGGAGGGGGAAGAGG
BALB/c   GTGCAAGGGGAAGGAGAAGGGGAGGGGGATAGGGAGGAGGAAGAGAAAGAAGAAGGGGAAGGAAGAAGAGGAGGGGGAGAGAAAAAGGAGGGGGAAGAGG

C57BL/6  AAGAGAAGGAAGAGAAGAGAAAGAAGAAGAGGAGGAGGAGG------GGAAGGAGGAGGAAATGGAAAAAAAAAAAACTCTTTGGGTCCCAAAGTCAGAT
BALB/c   AAGAGAAGGAAGAGAAGAGAAAGAAGAAGAGAAGGAGGAGGAGGAGGAAGGAGGAGGAAATGGAAAA--CAAAACCTCTTTGGGTCCCAAAGTCAGAT

C57BL/6  ACTTTCCATAGACCCCAACTGTATAACAAGACAGAGCCAGGTGTCAAGAAGAGCAGAGCCCAGGGCCACAGCACCAATGACGAGAGCTTGTGAGACACAC
BALB/c   ACTTTCCAGAGACCCCAACTATATAACAAGACAGAGCCAGGTGTCAAGAAGAGCAGAGCCCAGGGCCACAGCACCAATGACGAGAGCTTGTGAGACACAC

C57BL/6  TCACCTGCCACAGACTCACCA
BALB/c   TCATCTGCCACAG-CTCACCA
```

**B**

***md_3***

chr7:59,714,748-59,715,117

```
C57BL/6  ----------------------------------CTGAATT-ACTCTGTACACTTACTTTCTTAATGCCACATTGGAATCCCCTTTCCTGTATTATAAGTA
BALB/c   ----------------------------------CTGAATT-ACTCTGTACACTTACTTTCTTAATGCCACATTGGAATCCCCTTTCCTGTATTATAAGTA

C57BL/6  TCAAAGTGTCAAAACACGTATTTACCGGTTCCTCCTATCCCTTTAGAAACTGGGAGGTGCAAATCACCATTATCTGTTGTGATTTCCCTCTGAAAAAAAT
BALB/c   TCAAAGTGTCAAAACACATATTTACCTGTTCCTCCTATCCCTTTAGAAACTGGGAGCTGCAAATCACCATTATCTGTTATGATTTTCCTCTGAAAATGAT

C57BL/6  CTGCATGTTAACTCTCAAGTGCTTTTATCCCTTCCCATCATTAATTATATTCACAGCAAGGCACAATATAATTTTAATTCCTTTAGGAATATTATTATTG
BALB/c   CTGCATGTTAACTCTCAAGTGCTTTTATCCCTTCC-ATCAGTAATTATATTCACAGCAAGGCACAATATAATTTCAATTCCTTTAGGAATATTATTTTTA

C57BL/6  TGCAACACCCATATACCAAAATAAATGTTCTTTTTTTCTGTCTGCCTTTAATTTCCACACACCTTTAATCCACAATACTGTCTGTTCTTTGAAATATTAA
BALB/c   TGCAACACCCATGTACCAAAATAAATGTTCTTTTTTTCTGTCTGCCTTTAATTTCCACACACGTTTAATCCACAATACTGTCTGTTCTTTGAAATATTAA

C57BL/6  TGT
BALB/c   TGT
```

**C**

***md_4***

chr7:104,152,267-104,153,014

```
C57BL/6  ----------------------------------ATCTTTCAG--TTTCAAATTCTTTCTTATCTGCCATCAGCCTTTGCAGAACTACCTGCAGCTGCTCC
BALB/c   ----------------------------------ACCTTTCGA--TTTC-AATTCTTTCTTGTCTGCCATCAGCCTTTGCAGAACTACCTGCAGCTGCTCC

C57BL/6  TGTAAGAAAGGAATTGTAGCAGAGTCAAGCTGTGTGCTTAGTAGGACTGTTGTTCTATCCAAAGGCCAAAACCTTCTGGAAAGATAGAAAGCTAAGGATT
BALB/c   TGTAAGAAAGGAATTGTAGCAGAGTCAAGCTGTGTGCTTAGTAGGACTGTTGTTCTATCCAAAGGCCAAAACCTTCTGGAAAGATAGAAAGTTAAGGATT

C57BL/6  TCTACCTTTCTTTTAAAGGATTGTGATATGAATTCTTTACATAACAATTATATACACATATAGTTTTAATCTATTATGAATAAGTAGATAAGAAACTTTT
BALB/c   TCTACCTTTCTTTTAAAGGATTGTGATATGAATTCTTTACATAACAATTATATACACATATAGTTTTAATCTATTATGAATAAGTAGATAAGACACTTTT

C57BL/6  GGAAACAAGGCAAACAGGGACAAAATCAACTTGGTCTGGAAAGTTTACAGAAACCCAGGGAATGGCAACTAGAGTCTAGGAGTAGGAAGAAAATAATAAG
BALB/c   GAAATGAAGGCAAACAGGGACAAAATCAACTTGGTCTGGGAAGTTTACAGAAACCCAGGGAATGGCAACTAGAGTCTAGGAGTAGGAAGAAAATAATAAG

C57BL/6  GGATAGCCATGTACTTGGAAAAAGATTCCCCATCTTCTTTCAAAAAGGTGGTCTCAAGCCATGCCAGGTATGTAGATATCAGAGAATAATTTGCAACTTG
BALB/c   GAATAACCATCTACTTGGAAAAGGATTCCCCATCTTCTATCCAAATGGTGGTCTCAAGCCATGCCAGGTATGTAGATATCAGAGAATAATTTGCAACTTG

C57BL/6  CAGCCATCAGTTTTCTCCTTCAGTGGTCTGGTTTCTGAGTTCAGACTCAGGCCGATATACTTGGCAACAAGCCATTGAGGCATTGTTCCCAACACTCATC
BALB/c   CAGCCATCAGTTCTCTCCTTCAGTGGTCTGGTTTCTGATTTCAGACTCAGGCCGTTATACTTGGCAACAAGCCATTAAGGCATTGCTCCCAACCCTCATC

C57BL/6  TTACTTACTTTGACAACTAAAGAATAATCAGAGCTTTTCTTGGACATAAAACCTCAACAGTATACAACAAACTGTAAAAACACTTGCTCTTGTTCTAGAG
BALB/c   TTATTTACTTTGACACGTAAAGTATAATCAGAGCTTTTCGTGGACATAAAATCTCAACAGTATACAACAAACTGTAAAAACACTTGCTATTGTTCTAGAG

C57BL/6  TTTGTCGCCCCTGGTAGATTCTCTTAGGATTAAGCTCAATGGATAGTCCCATTCTTTCAGGCCTCAGAAGGTAACAGGATG
BALB/c   TTTGTCACCCTTGGTAGATTATCTTAGGATTAAGCTCAATGGATAGTCCCATTCTTTCAAGCCTCAGAAGGTAACAGGATG
```

**D**

***md_5***

chr7:29,621,365-29,622,278

```
C57BL/6  ----------------------------------------------------------AGCCTCAGGTGATCTCTTTATTGGAGCAAGGCAAAGAGCCCTGGATG
BALB/c   ----------------------------------------------------------AGCCTCAGGTGATCTCTTTATTGGAGCAAGGCAAAGAGCCCTGGATG

C57BL/6  ATTGGTAAAGAGCTTACAAGAGGCCTGTGCTCGGGTAAGTGAGAAGCATTGCGCAGAGCGGTGCCCACGGGGACAACTCGCTGGTAATCCAGGAAACTGC
BALB/c   ATTGGTAAAGAGCTTACAAGAGGCCTGTGCTCGGGTAAGTGAGAAGCATTGCGCAGAGCGGTGCCCACGGGGACAACTCGCTGGTAATCCAGGAAACTGC

C57BL/6  ACCTTTCCCATGTTATTTTGTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTAAAAATTTTTTTTTTTTTTATTTGAAAGTACACTGTAGCTGTCTTCA
BALB/c   ACCTTTCCCATGTTATTTTGTTTT----------------------------------------------------------------------------

C57BL/6  GACACTCCAGAAGAGAGAGTCAGATCTCCTTACGGATGGTTGTGAGCCACCATGTGGTTGCTGGGATTTGAACTCCTGACCTTCGGAAGAGCAGTCGGGT
BALB/c   ----------------------------------------------------------------------------------------------------

C57BL/6  GCTCTTACCCACTGAGCCATCTCACCAGCCCGTTATTTTGTTTTAAAGAGTATTTCCTAATGTTATGTAGACGAGTGCTCTGTCTGCATGTACACCTGCG
BALB/c   ---------------------------------------------AAAGAGTATTTCCTAATGTTATGTAGACGAGTGCTCTGTCTGCATGCACACCTGCG

C57BL/6  TGCCAGAAGAGAACATCAACTCCCATTATAGATGGCTGTGAGCCACCATGTGGTTGCTGGAAAGTGACCTCTTAACAGCTGAGCCATCTCTCCTGCCCAT
BALB/c   TGCCAGAAGAGAACATCAACTCCCACTATAGATGGCTGTGAGCCACCATGTGGTTGCTGGAAAGTGACCTCTTAACAGCTGAGCCATCTCTCCTGCCCAT

C57BL/6  TGGTTGTTATTTTGTAAGCTTCTGCCAAAAACTTGGGTCTAAGTGAGAAAAGCAAGTTCTTTTTTTAAAAAAAG-AGCATTATTATTATTATTATTTTGT
BALB/c   TGGTTGTTATTTTGTAAGCTTCTGCCAAAAACTTAGATCTAAGTGAGAAAAGCAAGTTCTTTTTTTAAAAAGAGCATTATTATTATTATTATTATTTTGT

C57BL/6  TTCATGTATTGATTGATTGATTGATTGATTGCATGGGGCGAATTGGGGGGGGTCAAAGGACAACTGTGAGAGTTCGTTTTCTCTGTGCATGGTGTTGGTC
BALB/c   TTCATGTATTGATTGATTGATTGATTGATTGCATGGGGCGAGTTGGGGGGG-TCAAAGGACAACTGTGAGAGTTCGTTTTCTCTGTGCATGGTGTTGGCC

C57BL/6  TGGGGATCAGACTCAGGTCATCATTGGTTTTGTTTTCACATAGGAACGGTTTTATAACTGGCTTATCAGTCTTTACCTTCTCTTGTCCATCCAGAGGTCG
BALB/c   TGGGGATCAGACTCAGGTCATCATTGGTTTTGTTTTCACATAGGAACGGTTTTATAGCTGGCTTCTCAGTCTTTACCTTCTCTTGTCCATCCAGAGGTCG

C57BL/6  GATGCCTAAAATCCCAACCTCCTAACAACCTGAGTTACCCTCTGGTGACCAGGCTCAATACTAGAGCT
BALB/c   GATGCCTAAAATCCCAACCTCCTAACAACCTGAGTTACCTTCTGGTGACCGGGCTCAATACTAGAGCT
```

**E**

***Tbc1d2***

chr4:46,658,967-46,659,728

```
C57BL/6  ------------------------------------TTGA-GCAGGTTAACAACAGCAAAGTGGCTGCAAAGTGATCGGTTTTGCTGCTGAAGCCTTCTTGG
BALB/c   ------------------------------------TTGATGCAGGGTTACAACAGCAAAGTGGCTGCAAAGTGATCGGTTTTGCTGCTGAAGCCTTCTTGG

C57BL/6  AGACTCTTCACGAACACGCAGGCCACAGGTGACTCTGTATCTGATAAGCCATTTTAAAAAAAGAGTCTGAAATGTACCAAAAACTTGATGTGCCAAGGTG
BALB/c   AGACTCTTCACGAACACGCAGGCCACAGGTGACTCTGTATCTGATAAGCCATTT-AAAAAAAGAGTCTGAAATGTACCAAAAACTTGATGTGCCAAGGTG

C57BL/6  TCTATCCTAGGGTCATGTATAGTACCTTAAAGTGGAAATAACTTTTCCTAGGTTTCTGTTGCTGTTGTTGTTTTATTTTTATCATGTGTATGTATGTATT
BALB/c   TCTATCCTAGGGTCATGTACAGTACCTTAAAGTGGAAATAACTTTCCCCAGGTTTCTGTTGCTGTTGTTGTTTTATTTTTATCATGTGTATGTATGTATT

C57BL/6  GGAGGGAGCATATATAATGTTTGTTCACATGTGTATGAACACACATGCACATGCATGTGCTATGTGTGGTGCCCAGAGGTTGATGTCAGATTTCTTCTTG
BALB/c   GGAGGGAGCATATATAATGTTTGTTCACATGTGTATGAACACACATGCACATGCATGTGCTATGTGTGGTGCCCAGAGGTTGATGTCAGATATCTTCTTG

C57BL/6  ACCACTCTGCTTTACTTACTGACACAGGTGCTCTAGCTGAGCCCAGAGCTCGCTGATTAGGACAGTCTGTTTATCCAGCTTGCTCTGCATCTCCTGTCTC
BALB/c   ACCACTCTGCTTTACTTACTGACACAGGTTCTCTAGCTGAGCCCAGAGCTCGCTGATTAGGACAGTTTGTTTATCCAGCTTGCTCTGCATCTCCTATCTC

C57BL/6  TGCCTCCCCAGGAA-----GAGCACAAGGTACCATGCCCACCTGGCTTTCCCATGCACTCTGGGGAAATCTGCATGCTGTTTCTCAGGCTTGCATGGCAA
BALB/c   TGCCTCCCCAGGAATACCTGAGCACAAGGTACCACGCCCACCTGGCTTTCCCATGCACTCTGGGGAAATCTGCATGCTGTTTCTCAGGCTTGCATGGCAA

C57BL/6  TCCTTTATCTATGGAGCCATCTCTCTAGACATCCTCCTTTTTTGGGGGGGGGGGGTAGAGTCTCACTAAATTACTCAGACCTGGCCTGGATCTTGAAAT
BALB/c   TCCTTTATCTATGGAGCCATCTCTCTAGACATCCCCCTTTTTTGGGGGGG------CAGAGTCTCACTAAATTACTCATACCTGGCCTGGATCTTGAAAT

C57BL/6  CCATATTCCAGCGCTCAAGAACCTAGGATTACAGGTGCACACCACCAAGCCCGGCCAGAAATGAAGACTTGACAGCCCAGGAGGAGGCTGGGTAAAGCAA
BALB/c   CCATATTCCAATGCTCAAGAACCTAGGATTACAAGTGCACACCACAAAGCCCGGCTGGAAATGAAGACTTGACAGCCCAGGAGGAGGCTGGGTAAAGCAA

C57BL/6  GA
BALB/c   GA
```

**F**

***Cd48***

chr1:173,518,415-173,518,932

```
C57BL/6  ----------------------------------------------------------CTCCCCTCTGCGACGCTGGTGGTCTCTGCCCAGTCACAGCCTCTCC
BALB/c   ----------------------------------------------------------CTCCCCTCTGCGACGCTGGTGGTCTCTGCCCAGTCACAGCCTCTCC

C57BL/6  TTACTCCCCACAGGCAACAGCAGATTCGGCCTCAGTGCTCAAAGTGCACTGCTGAAAAATCCATTCTTTTAAGCCAAGCTATTTTTCTTGTTGATCTCTA
BALB/c   TTACTCCCCACAGGCGACAGCAGATTCGGCCTCAGTGCTCAAAGTGCACTGCTGAAAAATCCATTCTTTTAAGCCAAGCTATTTTTCTTGTTGATCTCTA

C57BL/6  GTGGGTAGGAGAATTTGAAATCTCATTCCATGTGGTGTGATGGTAGTGGTCTGCTTTCCTGTTGTGTATGGTAATTCGTGAAATAGAGAAAAAACAGAAA
BALB/c   GTGGGTAGGAGAATTTTAAATCTCATTCCATGTGGTGTGATGGTAGTG-----CTTTCCTGTTGTGTATGGTAATTCGTAAGATAGAGAAAAAACAGAAA

C57BL/6  AACCTCATACGACTTCCGGTTTTGGGTTTTGCTTCCTGATTGAAGGGCAGGCGCCCTGACTTCTCTTACAGTTGTCTCCAGTGTTCTGGGGAAGCTTCTC
BALB/c   AACCTCATACAACTTCCGGGTTTGGGTTTTGCTTCCTGATTGAAGGGCAGGCGCCCTGCCTTCTCTTACAGTTGTCTCAAGTGTTTCGGGGAAGCTTCTC

C57BL/6  TAAGTATTATGTGCTTCATAAAACAGGGATGGTGTCTGGTCCTGGAACTGCTACTGCTGCCCTTGGGAACTGGATTTCAAGGTAGCTTCCTCCAAGTGTC
BALB/c   TAAGTATTATGTGCTTCAGAAAACAGGGATGGTGTCTGGTCCTGGAACTGCTACTGCTGCCCTTGGGAACTGGATTTCAAGGTAGCTTCCACCAAGTGTC

C57BL/6  CGGCTGTGCACTGAAGGCATGGCCAGCTGGGAAAGGGGAGGTGGGATGTAGCTCAACCCTGGAAATGAGCAT
BALB/c   CGGCTGTGCACTGAAGGCATGGCCAGCTGGGAAAGGGGAGGTGTGATGTAGCTCAACCCTGGAAATGAGCAT
```

## G

### *Tmtc2*

chr10:104,969,448-104,970,451

```
C57BL/6  --------------------TTTGTTGTTCTGTTAAATGTGCTGGCCTTCAAACTCCTGAGGACTTTGCTGTTTCTGCCTCCCATCTTCTCAGGGAAGTAG
BALB/c   --------------------TTTGTTGT-CTGTTAA-TGTGTGG--CTTCAAATTCCTGAGGACTTGGT--GTTCTGCCTCCCATCTTCTCAGGGAAGTAG

C57BL/6  TCCTTCATTGAGGATTCAAGAGGATTTACAGACATTCATGTTATGCTT-TTTTTTTTTTTCTGTAGGTCTGGAAGATGCAAATTTAGGTCCTCAGGCTTG
BALB/c   TCCTTCAT-GAGGATTCAAGAGGATTTACAGACATTCATGTTATACTTCTTTTTTTTTTTCTGTAGTTCTGGAAGATGCAAATTTAGGTC-TCAGGCTTG

C57BL/6  GGTACCAAGTGCTTCTATCCACTGAACCCTCCCTCCACCCAGCTTCAAGGAAATATTTTTAAGGTTTCCTAACAAAATAAAGCAAAATTGAAAACTATGT
BALB/c   GGTACCAAGTGCTTCTATCCACTGAACCTTCCTTCCACCAGCTTCAAGGAAATATTTTTAAGGTTTCCTAACAAAATAAAGCAAAATTAAAAACTATGT

C57BL/6  TGATTTTTTTTTTGCTAGGGTTATGTAAGAAGAGAGAATATCCATTGAGACAACACTTCCATCAAATTGGCCTGTAAATAAGTCTGTGGGAGGCATTTTC
BALB/c   TGATTTTTTTT--GCTAGGGTTATGTAAGAAGAAAGAATATCCATTGAGACAACACTTCCATCAAATTGGCCTGTAAATAAGTCTGTGGGAGGCATTTTC

C57BL/6  TTGGTTAAAAGTTGATGTGGGAGGGCTCAGCCCATTGCAGGTGCTATCACACATAGGCTAGTGGTCCTGGATGCTATAAGAGGAGGTTATGCAAGCCATC
BALB/c   TTGGTAAAAAGTTGATGTGGGAGGGTCCAGCCCATTGTAGGGGCTATCACACATAGGCTAGTGGTCCTGGATGCTATAAGAGTAGGTTACACAAGCCATC

C57BL/6  AAGAGCTTGCCTTGTAAGCAGCACTCTCCCAGGTCTTCTACAGAAGTTCTGTCTCCAGGTACCTGCCTTGAATTCCTGTCCTGACTTCCTTCAGTGATGG
BALB/c   AAGAGCTTGCCTTGTAAGCAGCACTCTTCCAGGTCTTCTACAGAAGTTCTGTCTCCAGGTCCCTGCCTTGAATTCCTGTCCTGACTTCCTTCAGTGATGG

C57BL/6  ACTGTGACGTGAAAATATAAACCAAATAAACCCTCTCCAAGTTGCTTTTGATCAGAAGGCTTTACCATAGCAAGTGAAAATGAACAAGGACATGTACCTA
BALB/c   ACAGTGACGTGAAAATATAAACCAAATAAACCCTCTCCAAGTTGCTTTTGGTCAGAAGGCTTCACCATAGCAAGTGAAAATGAACAAGGACATGGACCTA

C57BL/6  CATTGATCCCATTTTTATAAACTTGTTGCTTCTAAGAAAAATTGCAAATGTATTGTGAAGATCTCCCCAGAAAGGGCGATTAGTCATCAAAAGCAGAGAT
BALB/c   CATTGATCCCATTTTTATAAACGTGTTGCTTCTAAGAAAAATTGCAAATGTATTGTGAAAATCTCCCCAGAAAGGGCGATTAGTCATCAAAAGCAGAGAT

C57BL/6  TTGACAAAAACAACAACAACAAACAAACAAACAAAAACAACAACAACAACAAACTTTTTT------AACTGATCACACCACCTAGCACAGAGGCACAAGT
BALB/c   TTGACAAAAACAACAACAACAAACAAACAAACAAAAACAACAACAAAAAAACCCTTTTTTCCTTTAACTGATCACACCACCTAGCACAGAGGCACACGT

C57BL/6  TCATAGGC-AGGACCATGTTTCATTTGACAAACCAGGAATTCAGACTTGAGAGCAGTCTAGGCCACACCTCCTACCTTACCTCTCACAGATACCACAAAG
BALB/c   TCATAGGCGAGGACCGTGTTTCATTTGACAA-CCAGGAATTCAGACTTGAGAGCAGTCTAGGCCACGCCCCTACCCTACCTCTCACAGATACCACAAAG

C57BL/6  TGACCTACTGGTTGAATCAAGAAACAAAAGC
BALB/c   TGACCTACTGGT-GAATCGAGAAC--GAAGC
```

## H

### *Pgbd5*

chr8:127,300,507-127,300,973

```
C57BL/6  ---------------------------------------CACTCTACCCACAGCTT-CTCAACAAATGCCCCTCACAAAACAAGCTTTCTATCTTTCAATTT
BALB/c   ---------------------------------------CATCTCCAGCACAGTTTTCTCAACAAACACCCCTT-CAAAACAAGCTTTCTACCTTTCAATTT

C57BL/6  CCCTGTCTTCATTTTGACCAAGCC-TTGGTTAGGGCCAC-TCCACAC-TCAGGCATGGAAGGCATGCCTTACACTGACCAAGACCAAGCCATGCCCAATC
BALB/c   CCCTGCCTTCATTTTGACCAAGCCATTGGTTAGGGCCACCTCCACACATCAGGCATGGAAGGCATGCCTTACACTGACCAAGACCAAGCCATGCCCAGTC

C57BL/6  GCCCATTAAGCGCCCACAGGAATGAACTCTAGCTTTGACCCTACACTCTCAGACTGACCTGAGAATTAAAGCTGGTCACTGCCCTCTCAGAGCCCGGTCT
BALB/c   GCCCATTAAGCGCCCACAGGAATGAACTCAAGCTTTAACCCTACGCTCTCAGACTGACCTGAGAATTAAAGCCGGTCACTGCCCTCTCAGAGCCCGGTCT

C57BL/6  CCCTGCAGAAAACACCCAAAGAAGG------ACCCTGAAGGTGTCGGTCGGATCAGGTTGAAGAGTGGCACGAGCCAAGTGGTCCCTAATGTTCTAACGT
BALB/c   CCCTGCAGAAAACACCCAAAGGAGAGAGCCTAATCTGAAGATGTCGGTCGGATCAGGTTGAAGAGTGTCACGAGCCAAGTGGTCCCTAATGTTCTAACAT

C57BL/6  CTTTCTGTAAACCCAGTGGTTCCCTGCAACTTCACGTACCCGTGAGCCTCCTTAAAAGTCTAGAGACCAGTTGGTTGTGGTCCCAGGATCCAGGAATCAC
BALB/c   CTTTCTGTAAACCCAGTGGTTCCCTGCAACTTCACGTACCCATGACCCTCCTTAAAAGTCTAGAGGCCAGTTGGTTGTGGTCCCAGGATCCAGGAATCAT

C57BL/6  ACTCTACCAGGATC
BALB/c   ACTCTAACAGGATC
```

## I

### *Tmtc2*

chr9:49,224,882-49,225,570

```
C57BL/6  ----------------------------------AAAGGAGAAGAAGGTACTAGGAAAGACCAGTGAATCGCTGGGACTAATGTGAGACAATTAGGAAGA
BALB/c   ----------------------------------AAAGGAGAAGAAGGTACTAGGAA-GACCAGTGAATCGCTGGGACTAATGTGAGTCAATTAGGAAGA

C57BL/6  AAAAGAAATAGCAAGAGAATTTAATACTGGGCTTCCAGGAACAGAGCTTTAAGTCACTAGCCAACTCTAGTAATTTATGGACGGCTTATCAGGAGTTGTA
BALB/c   AAAAGAAATAGCAAGAGAATTTAATACCGGGCTTCCAGGAACAGAGCTTTAAGTCACTAGCCAACTCTAGTAATTTATGGACGGCTTATCAGGAGTTGTA

C57BL/6  TGCTTTGCTGAGTATCAGGGAAAGTGGTGAATAAAAAACAAAATCCTTTTGTTTTGTTTGAAAAGCTTGTGATCTTGTGTGGGAGACAAACGTTAACACA
BALB/c   TGCTTTGCTGAGTATCAGGGAAAGTGGTGAATAAAAAACAAAATCCTTTTGTTTTGTTTGAAAAGCTTGTGATCTTGTGTGGGAGACAAACGTTAACACA

C57BL/6  AGCAGCTCTGAGATTCCTAAAACAGAAGCCTTCAGCTGTCTCCCTGCCTCTCTTCACTGTGTTGTCTACGCAGATTGGCATTCATAGACTGTCCTTTAGA
BALB/c   AGCAGCTCTGAGATTCCTAAAACAGAAGCCTTCAGCTGTCTCCCTGCCTCTCTTCACTGTGTTGTCTACCCAGATTGGCAATCATAGACTGTCCTTTAGA

C57BL/6  GGGATCACTCTACCCTCACACCCCACCTAAACAGATAAATTTGTTCACCTCACTCCAATTTCAGATCAATCTAATTATCCATATCTAGCAGCCTTGACTA
BALB/c   GGGATCACTCTACCCTCACACCCCACCTAAACAGATAAATTTGTTCACCTCACTCCAATTTCAGATCAATCTAATTATCCATATCTAGCAGCCTTGACTA

C57BL/6  GGGAGGCTAAACATTGATGAAGAAAAATGGATAAATATCTGGGGGGGGGGGGACTTACTGGTCTGTCAGTGACCCTCACTCTCCCAGATCCTTTACTATG
BALB/c   GGGAGGCTAAACATTGATGAAGAAAACTGGATAAATATCTGTGGGGGG---AACTTACTGGTCTGTCAGTGACCCTCACTCTCCCAGATCCTCTACTATG

C57BL/6  TCATCCTAGGTAGTCTATTATCCTCTCAAACACTTTCAAGTTTTGGACAATCTAAACCTGACGATTCCAAATAGAGGATTCTGCTTATTCATTCGGTGAG
BALB/c   TCATCCTAGGTAGTCTATTATCCTCTCAAACACTTTCAAGTTTTGGACAATCTAAACCTGACGATTCCAAATAGAGGATTCTGCTTATTCATTCGGTGAG

C57BL/6  GAAAGGGATGGTGTTAAGAAAGAG
BALB/c   GAAAGGGATGGTGTTAAGAAAGAG
```

**J**

## *1600021P15Rik*

chr16:28,833,612-28,834,468

```
C57BL/6  ----------------------GTAGTCAACAATGTGTTGACAGAATGTTTCCCCAGAGAACGGAGGGGACCCAAGCACACAGCTCTCCCTCCTTTGTGC
BALB/c   ----------------------GTAGTCAACAATGTGTTGACAGAATGTTTCCCCAGAGAACGGAGGGGACCCAAGCACACAGCTCTCCCTCCTTTGTGC

C57BL/6  TCAGGCTGCTACACTGATCTCAGCTTGGGAGGAGGAAGTATGGCGGTTGAGGGAAAAGGTTTCACATGGAGAGCGCCTAATATCCTGTGTACCTCAGATA
BALB/c   TCAGGCTGCTACACTGATCTCAGCTTGGGAGGAGGAAGCATGGCGGTTGAGGGAAAAGGTTTCCATGGAGAGCGCCTAATATCCTGTGTACCTCAGATA

C57BL/6  ATACACACATAACCACTGCTGGGCTTTAACGGAAGCCTATACTCTTAGCTTTGCCCTTCATAAACTTAGCTTTGTCCTTAGAGATCTTTAGGACCTACAA
BALB/c   ATACACACACAACCACTGCTGGGCTTTAACGGAAGCCTATACTCTTAGCTTTGCCCTTCATAAACTTAGCTTCGTCCTTAGAGATCTTTAGGACCTACAA

C57BL/6  CAGGCTATGGAATGATGAAGCTAACACGATCTGTCCGGCCCCAGGGCCGTGTTCTCCCATCTCACTGCATCCTCACTTCTTCCGTGACAACTCTATCTCT
BALB/c   CAGGCTATGGAATGATGAAGCTAACACGGTCTGTCCGGCCCCAGGGCCGTGTTCTCCCATCTCACTGCATCCTCACTTCTTCCGTGACAACTCTCTCTCT

C57BL/6  CCCGTGCAGCCACAGAAGGATGGCTACTGTTTACCTCAGACAACGGAACTTAACACAGGAACTACAGATTTTGCCTAAACCACAGCTAAAAGACAAATCT
BALB/c   CCCGTGCAGCCACAGAAGGATGGCTACTGTTTACCTCAGACAACGGAGCTTAACACAGGAACTACAGATTTTGCCTAAACCACAGCTAAAAGCCAAATCT

C57BL/6  GCCAAGCCAGCTTTGTTCCCATTAGCAAAATTTAGCAACAGCACTATTCCTTAGTCACCAAAACTCTATGTCAACAATCACCTGAGCACGAGTCTTTCAT
BALB/c   GCCAAGCCAGCTTTGTTCCCATTAGCAAAATTTAGCAACAGCACTATTCCTTAGTCACCAAGACTCTATGTCAACAATCACCTGAGCACGAGTCTTTCAT

C57BL/6  ATTTGGGGAAATTTCTCACATTAGGTTATCTTGGGCCACTGTGGGAACTGCCATAGATTGTCATCAATGACAAACATCCAACAAACAGCCCAAGTTCTGA
BALB/c   ATTTGGGGAAATTTCTCACATTAGGGTGTCTTGGGCCACTGTGGGAACTGCCATAGATTGTCATCAATGACAAACATCCAACAAACAGCCCAAGTTCTGA

C57BL/6  ACAAAAACAAAATCAGTAATATTAGCAAGATCTGTTGAGGACCTCTCAAATCTCCACTGAACTGTGTTCCAGACATGCAAAAGAGAAACAACAGCATATC
BALB/c   ACAAAAACAAAATCAGTAATATTAGCAAGATCTGTTGAGGCCCTCTCAAATCTCCACTGAACTGTGTTCCAGACATGCAAAAGAGAAACAACAGCATATC

C57BL/6  CGTGAAGGGCATCTTTGGGATGTTAAGTAGTGGGAAACATCAGACATTGGTCACTAACAATTAGAAATTCATAACTGAC
BALB/c   CGTGAAGGGCATCTTTGGGATGTTAAGTAGTGGGAAACATCAGACACTGGTCACTAACAATTAGAAATTCATAACTGAC
```

**K**

## *2310009E04Rik (Fggy)*

chr4:95,071,096-95,072,198

```
C57BL/6  ---------------------------------------------CCAGGCTACAAAGCTGAGGAGGAGTGGGCTGTCATAGAAGACAGCCTTTTACTTTAAG
BALB/c   ---------------------------------------------CCAGGCTACAAAGCTGAGGAGGAGTGGGCTGTCATAGAAGACAGCCTTTTATTTTAAG

C57BL/6  AGGAAGAAAGATCCTGAATATGCTGAGAGAAGCCATAGAGAATCAGGTATTATTGTCATGTTGACTCCTTCCTTCCTGTACAAAATTTTCTCACCATCTT
BALB/c   AGGAAGAAAGATCCTGAATATGCTGAGAGAAGCCATAGAGAATCAGGTATTATTGTCATGTTGACTCCTTCCTTCCTGTACAAAATTTTCTCACCATCTT

C57BL/6  AGAGCATGTGGGAGACTTGCAGGACTATTCACTCTTATAGGAGCGGCGAATCTTTTCTGTTAATAAGACTGAAGTGCGCAGTACAGGCCTACTTGTTGCT
BALB/c   AGAGCATGTGGGAGACTTGCAGGACTACTCACTCTTATAGGAGCGGCGAATCTTTTCTGTTAATAAGACCGAAGTGAGCAGTACAGGCCTACTTGTTGCT

C57BL/6  TGTCCTCATCTCCCCTGCTTACCTAAACCGGATTACAATGTAACCAAAAACTAAATTGCATGAGGCTGGAACACATGAGCTCCCAGTAGAGTTCAAAGAT
BALB/c   TGTCCTCATCTCCCCTTCTTACCTAAACCGGATTACAATGTAACCAAAAACTAAATTGCATGAGGCTGGAACACATGAGCTCCCAGTAGAGTTCAAAGAT

C57BL/6  CACTGTTAAAAGAGGACACTGAGTTGAGTTTTCTCCCCTGGCCTGGGACTTACTTTCTGTAGGCTGTCACCAACGAAACTGAACCTGGAGGAGTGGTTCT
BALB/c   CACTGTTAAAAGAGGACACTGAGTTGAGTTTTCTCCCCTGGCCTGGGACTTAATTTCTGTAGGCTGTCACCAACGAAACTGAACCTGGAGGAGTGGTTCT

C57BL/6  TTAAGAAGCAGAATCTAAGGTAAGCCCCTCCATTTTCCAGGGATGTGGTGCTGTGAAGTAGTCATCACTGTGACTAAGTTATTTGAGGGGAAAG------
BALB/c   TTAAGAAGCAGAATCTAAGGTAAGCCCCTCCATTTTCCAGGGATGTGGTGCTGTGAAGTAGTCATCACTGTGACTAGGTTATTTGAGGGGAAAGAAAGGG

C57BL/6  ----------------------------------------------------------------------------------------------------
BALB/c   AATACAGGGACAGCACAATTCCCTAGGGTTGTTTTTGAGCAGAGAGTTTGTAGCATTTACCCCCTTCTAGGGATGGCTGAAATAGTTTTGCAGAGAGCAG

C57BL/6  ----------------------------------------------------------------------------------------------------
BALB/c   TCCCAGTTTCCAGGTCTCCCGCCTATCCTCATATGGCATTTGAATAAGCTTGTACATTTGCTAAACAATTAGCTGTGAAATCTATGACTTCCTGTGGTAC

C57BL/6  ----------------------------------------------------------------------------------------------------
BALB/c   ACTTCTCACAATATTTATACAGGCCATCATTTAGTGAACTGTATAAATATTTATGGCATGGAACAGCCTTTTGATCATAAATAGGGGCCAGGAGTCAGTG

C57BL/6  --------------------GCTTGGTATTACCTGAGTTTTAAGGCATCTTGTAGCCTGAGATATTAGGTATGTCTCAGGACTTGTGCACAGCTTGAATTT
BALB/c   TCGCCCATGCTGGGGACTGGCTTGGTATTACCTGAGTTTTAAGGCATCTTGTAGCCTGAGATATTAGGTATGTCTCAGGACTTGTGCACAGCTTGAATTT

C57BL/6  CTCTTCATTATTTTTCTATCTCTTTGCCATTTATTTTTCTATCTCTTTGCCATCGAGAGGCCATGAGCTTATTTTTCCTGAATGCCCTCCTCTACAACTG
BALB/c   CTCTTCATTATTTTTCTATCTCTTTGCCAT----------CGAGAGGCCATGGGCTTATTTTTCCTGAATGCTCTCCTCTAAAACTG

C57BL/6  TTTTTCTTTTTAGAGTATTCCTCTTTGCCCCAAAATAGTGAGTTGAAACCCAAGTTCCATTCCTCATTATTCATAAAGCAGGCTGTGGAATTTCCGTGAC
BALB/c   TTTTTCTTTTTAGAGTATTCCTCTTTGCCCCAAAATAGTGAGTTGAAACCCAAGTTCCATTCCTCATTATTCATAAAGCAGGCTGTGGAATTTCCGTGAC

C57BL/6  CTGACCTGGAATTATAGCCCCTCTGTTGCTGCTATGCTTTACCTAGTGCTGTCTTTGAATACCTCAGCTCATTTTCCTGAGAAATGGGGCTTCCACACTG
BALB/c   CTGACCTGGAATTATAGCCCCTCTGTTGCTGCTATGCTTTACCTAGTGCTGTCTTTGAATACCTCAGCTCATTTTCCTGAGAAATGGGGCTTCCACACTG

C57BL/6  CCTACTTGCTCAGCTAAAGATGAAATGAGATCATAGAGTCAGGTGGGCAGTTTGCTCAAGAAAGACCAGCTTTTGTGGCATTTGTTGTTTGTTATTGAAG
BALB/c   CCTACTTGCTCAGCTAAAGATGAAATGAGATCATAGAGTCAGGTGGGCAGTTTGCTCAAGAAAGACCAGCTTTTGTGGCATTTGTTG----TTATTGAAG

C57BL/6  TTGCAAAGGCTTTGATTTTCAGTTTTAGCTGAATGTTGTGGAAAGTGTCCGAAACTCACTACTGTATGCT
BALB/c   TTGCAAAGGCTTTGATTTTCAGTTTTAGCTGAATGTTGTGGAAAGTGTCCGAAACTCACTACTGTATGCT
```

## L

### *2310009E04Rik (Fggy)*

chr4:95,074,498-95,075,968

```
C57BL/6  ----------------------------------------CATGCCTCTGGGAATGCCCGTGTTCTCCCTTCTAGCTTTCTGTAAGCACAGCTTATGATGCA
BALB/c   ----------------------------------------CATGCCTCTGGGA-TGCCCGTGTTCTCCCTTCTAGCTTTCTGTAAGCACAGCTTATGATGCA

C57BL/6  TTCTCAAAGCGGACATCATCTACTGCATTTCTCAGAAAGGAACCCTGTTAAGTGACACATGATTGTACAAGCACAGACTGAGAAAGTCAAATAACTAAGG
BALB/c   TTCTCAAAGCGGACATCATCTACTGCATTTCTCAGAAAGGAACCCTGTTAAGTGACACATGATTGTACAAGCACAGACTGAGAAAGTCAAATAACTAAGG

C57BL/6  GATGGCTCTTTTCGCTTTTTTTT-CCTATTGGTTCACTGATAAAGGAGTTTACTTTAGAAAACCACCGCTGGAAACCAATTATATTGCAAGCACTGGAAG
BALB/c   GATGGCTCTTTTCGCTTTTTTTTTCCTATGGGTTCACTGATAAAGGAGTTTACTTTAGAAAACCACCGCTGGAAACCAACTATATTGCAAGCACTGGAAG

C57BL/6  GGAGTGGGGATGCAGCAATCAGTAACATCCCCTTTTCTCCAGGTTCAGTCCAATGAGGAGATGGACATACCACCAAAAACATGGTCCAGAGCGAACATTT
BALB/c   GGAGTGGGGATGCAGCAATCAGTAACATCCCCTTTTCTCCAGGTTCCGTCCAATGAGGAGATGGACATACCACCAAAAACATGGTCCAGAGCGAACATTT

C57BL/6  GAAATGAAAATACAGACACTGTACATATTGACAGCCTTTGTACATGTGCAAGAAAAGTGATAAGGTCCCCCTTAGAGAAGTAGAGGATGGAATTGGGAAG
BALB/c   GAAATGAAAATACAGACACTGTACATATTGACAGCCTTTGTACATGTGCAAGAAAAGTGATAAGGTCCCCCTTAGAGAAGTAGAGGATGGAATTGGGAAG

C57BL/6  GGAAGAGCAAGTGATGGGGGAGGACCTCCATCAGTGCAAAGCACAAGTATACTTCCAGGACTGTGGGAATATCTAGGATCCCAGTGGGCAGGGCCAGGGT
BALB/c   GGAAGAGCAAGTGATGGGGGAGGACCTCCATCAGTGCAAAACACAAGTATACTTCCAGGACTGTGGGAATATCTAGGACCCAGTGGGCAGGGCCAGGGT

C57BL/6  AGCATTCATTCGGGTTAAGAAGCTCACAGAAAGTCTCACTTGAGGAAGACCTTAGGGTGTCTTATTCTATCTATTTCTATCTATTGTGATGATATACTGC
BALB/c   AGCGTTCATTCGGGTTAAGAAGCTCACAGAAAGTCTCACTTGAGGAAGACCTTAGGCTGTCTTATTCTATCTATTTCTATCTATTGTGATGATATTCTGC

C57BL/6  TGCCCCAAGAAAGAGTCTAGCACAGTAATCACTATGGTGTTCATCTGGCTTGACCCAGGATATCTGACTTGGGGTACTAGGACACAAAAGGACAGACAAA
BALB/c   TGCCCCGAGAAAGAGCCTAGCACAGTAATCACTATGGTGTTCATCTGGCTTGACCCAGGATATCTGACATGGGGTACTAGGACACAAAAGGACAGACACA

C57BL/6  CAGATACAAGTACAGGAAAGCCAGGACTGCATGGACCACACACACTCTGATAAAGATGCGCCATCAGAGTCTGCACTGTGCAGGTTGCAGTTATATATAG
BALB/c   CAGATGCAAGTACAGGAAAGCCAGGACTGCATGGACCACACACACTCTGATAAAGATGCGCCATCAGAGTCGGCACTGTGCAGGTTGCAGTTATATATAG

C57BL/6  TTGCAGGAGGAGCTGGGATTACTGTATAAATTTGCAAGAGATGGTGAGGTTATTCTATACACTTGTAGTGAGAGGCAGGGTTATTGTATACAGTTGTCCA
BALB/c   TTGCAGGAGGAGCTGGGATTATTGTATAAAGTTGCAGAAGATGGTGAGGTTATTCTATACACTTGTAGTGAGAGGCAGGGTTATTGTATACAGTTGTCCA

C57BL/6  AGAAGGCAGGACTATTATATATGTAGCTGAGCAAGGAGGCAGGCTTAGCTAATCTTGGTAGGGGAGTCTCTATAGGGGAACCGGCTCAGACTGTAAGCAT
BALB/c   AGAAGGCAGGGCTATTATACATGTAGCTGAGCAAGGAGGCAAGCTTAGCTAATCTTGGTAGGGGAGTCTCTATAGGGGAACCGGCTCAGACTGTAAGCAT

C57BL/6  CTTGGATGAAGTTATTGCACTTGACCTATTTTTAAGGACCAGGGTAAGGTCTTGACTTTCCCATGGGTCTGAGGCATGTGGACATTAACATGGTCATACT
BALB/c   CTTGGATGAAGTTATTGCACTTGACCTATTTTTAAGGACCAGGGTAAGGTTTTGACTTTCCCATGGGTCTGAGGCATGTGGACATTGACATGGTCATACT

C57BL/6  CATGTCAAACAATCCATGTTCACTCAGGACTTCCTCTACTCCCTGTGAACCACTAAGAGCAAACAGGGAGGATGGGAATGACAACTCAACAGGTGTCTGG
BALB/c   CATGTCAAACAATACATGTTCACTCAGGACTTCCTCTACTCCCTGTGAACCACTAAGAGCAAACAGGGAGGATGGGAATGACAACTCAACAGGTGTCTGG

C57BL/6  GTTATTTAAAAATCTGCTTTTATTGGTTACTGGCTTAAAGAGAAGGTATCACTCCCAAGTGAATGAAAGAATGAAAGCTGGTACTCATGTTTATGTGAGC
BALB/c   GTTATTTAAAAATCTGCTTTTATTGGTTACTGGCTTAAAGAGAAGGTATCACTCCCAAGTGAATGAAAGAATGAAAGCTGGTACTCATGTTTATGTGAGC

C57BL/6  TCTCGAGGGGAAAAAAGTCAGGGACCCAGGGAGCCAGGAAGACCCACCTGCTGAGCCTGGGAGCCCTGTCATGCTCAGGGGAGGCTGTGAGCTTTATTAT
BALB/c   TCTCGAGGGGAAAAAAGTCAGGGACCCAGGGAGCCAGGAAGACCCACCTGCTGAGCCTGGGAGCCCTGTCATGCTCAGGGGAGGCTGTGAGCTTTATTAT

C57BL/6  TCTTGCTTCT
BALB/c   TCTTGCTTCT
```

## M

### *3110007F17Rik*

chrX:119,220,163-119,220,873

```
C57BL/6  ------------------TCTCGTTCAGCTGTTCTACCCACTGGTCTAGCTCCTTGGTGAATGCCTTGTCGTCTATGGCGGTGAGCTGCTCATCTCCGCC
BALB/c   ------------------TCTCGTTCAGCTGTTCTACCCACTGGTCTAGCTCCTTGGTGAATGCCTTGTCGTCTGTGGCGGTGAGCTGCTCATCTCCGCC

C57BL/6  GCCTCCCCCTCCGCTGATCCCCTGGCGGGACCGCTAAGCTGCAGAGAGCGCAGCCCGCGCCCCGCCCAAGCCCAGCCATTATCTTCGGGTCCCTGCTGTT
BALB/c   GCCTCCCCCTCCGCTGATCCCCTGGCGGGACCGCTAAGCTGCAGAGAGCGCAGCCCGCGCCCTGCCCAAGCCCAGCCATTATCTTCGGGTCCCTGCTGTT

C57BL/6  TGCCTCTAAGTCTGTGGCGGTTCCGTAGCTGTAGTGCCTGGCGGCCCCAGCCCCGCCCACGCGCTTCTGTCGCACCGCCCCCTGCGCTGTTGTCAGAGCG
BALB/c   TGCCTCTAAGTCTGTGGCGGTTCCGTAGCTGTAGTGCCTGGCGGCCCCAGCCCCGCCCACGAGCTTCTGTCGCACCGCCCCCTGCGCTGTTGTCAGAGCG

C57BL/6  CGGGACACAGTGAGGGAGGGGGGAATGGGCGAGCGGTCGCGCCGTTGGGAGGAGCAGTCCGGACCCATGCCCGTGCAGGCAGGACATGGGGAGAGCACGC
BALB/c   CGGGACACAGTGAGGGAGGGGGGAATGGGCGAGCGGTCGCGCCATTGGGAGGAGCAGTCCGGACCCCTGCCCGTGCAGGCGGGACGTGGGGAGAGCACGC

C57BL/6  AGCGACAGGCATCTGGCTTTCTGAGCCCTGGGATTTGGTGAGCGAACACTCGGAGTGGGATGGTTCGGCGCGCGCTGAGACCCGCTATCCCGGCCCCCGC
BALB/c   AGCGACAGGCATCTGGCTTTCTGAGCCCTGGGATTTGGCGAGCGAACACTCGGAGTGGGA----------------AGACCCGCTATCCCGGCCCCCGC

C57BL/6  CCACCTTTCCCAGCGGGCCTTTCCAGCGGTCCCGGGGCGGTCTTGGTGTACTGCCATGGTGGGTGGAGGTGCGCTTGGGGCCTGGTCGCTGGCCAGGGGG
BALB/c   CCACCTTTCCCAGCGGGCCTTTCCAGCCGTCCCAGGGCGGTCGAGGTGTACCGCCATGGTGG----AGGTGCGCTTGGGGCCTGGTCGCTGGCCAGGGGG

C57BL/6  GGCAGTCCCCCGCCGACCCCAGCCTGGCCCTGCCCGCGAGCCCGCCACCAGAGCACGGGACGCAGGAAAAGGGGGCGGGGAATGACGCGCGGGTGGAGGA
BALB/c   GGCAGTCCCCAGCCGACCCCAGCCTGGCCCTGCCCGCGCGCCGGCCCCAGAGCACTGGACGCAGGAAAAGGGGACGGGGAATGACGCGCGGGTGGAGGA

C57BL/6  GTTTTCGCTGTGCGCCGCCAGCACAGTTG
BALB/c   GTTTTCACTGTGCGCCGCCAGCACAGTAG
```

**N**

***Arhgdib***

chr6:136,897,634-136,898,449

```
C57BL/6  ----------------------------------------TAATTATACAATCCATGACCTTAGGCAGCAGGGTACAGATGGTGCCTCACCCACCTAACC
BALB/c   ----------------------------------------TAATTATACAATCCATGACCTTAGGCAGCAGG-TACAGATGGTGCCTCACCCACCTAACC

C57BL/6  AGGAAGCCCTTTGCTCCTGACTGTGGGTGACAGCAATTGCATGTCTCCGTGGAGGCCTTGCCCTGACCCATCCCTACACGCACACACCTGCTACCACAGG
BALB/c   AGGAAGCCCTTTGCTCCTGACTGTGGGTGACAGCAATTGCATGTCTCCGTGGAGGCCTTGCCCTGACCCAACCCTACACGCACACACCTGCTACCACAGG

C57BL/6  GACATCTCCCAGCAGTGTTTTCTTGTACTTGGTTAGACTCTCGTCATCCTTGTCCATCTCCTGCAGCTCCTTCAAGGACTTCTGAGGGGGTGGCTTATAA
BALB/c   GACATCTCCCAGCAGTGTTTTCTTGTACTTGGTTAGACTCTCGTCATCCTTGTCCATCTCCTGCAGCTCCTTCAAGGACTTCTGAGGGGGTGGCTTATAA

C57BL/6  TTGAGCTTGCTGTCCAGGTCGTCGTCCGCCTCTTCCAGCTGTGGCTGTGCATCCTTCTCCGTCATCTTGATTGCTCTCTATTTCAGAGAGAGAATGATGG
BALB/c   TTGAGCTTGCTGTCCAGGTCGTCATCCGCCTCTTCCAGCTGTGGCTGTGCATCCTTCTCCGTCATCTTGATTGTTCTCTATTTCAGAGAGAGAATGATGG

C57BL/6  TTCATTAGGCGCTGGGGTTGCTACTGGTACCACCTTAGTGGAGATCTGTATTCATCGGTGATTTTCAAATTGTTCCTCAACACAGATATTCCATGAAATA
BALB/c   TTCATTAGGCGCTGGGGTTGCTACTGGTACCACCTTAGTGGAGATCTGTATTCATCGGTGATTTTCAAATTGTTCCTCAACACATATATTCCATGAAATA

C57BL/6  GATAGTATTTCCAGTCATTTCTCTGACTTGTCTTAATTTTTGCTTATAGAATAAAGCTATAAGAATGAACAGCTTTTCTGTGAGCCATGCACTACAGTGA
BALB/c   GATAGTATTTCCAGTCATTTCTCTGACTTGTCTTAATTTTTGCTTATAGAATAAAGCTGTAAGAATGAACAGCTTTTCTGTGAGCCATGCACTACAGTGA

C57BL/6  AGAAAGGTAACATAAACAATCCCAGCGATCGACCATTCTAGCTTAGCCC----------------------------------------------------
BALB/c   AGAAAGGTAACATAAACAATCCCAGCGATCGACCATTCTAGCTTAGCCCATTGCTATTCCCAGCGATCGACCATTCTAGCTTAGCCCATTGCTATTCCCA

C57BL/6  ----------------------------------------------------------------------------------------------------
BALB/c   GCGATCGACCATTCTAGCTTAGCCCCATTGCTATTCCCAGCGATCGACCATTCTAGCTTAGCCCATTGCTATTCCCAGCGATCGACCATTCTAGCTTAGC

C57BL/6  ----------------------------------------CATTGCTATTCCCAGTATGCGCAATCTGAGTTTTGCTGACTTCTGGACACTGGTCAATT
BALB/c   CCCATTGCTATTCCCAGCGATCGACCATTCTAGCTTAGCCCCATTGCTATTCCCAGTATGCGCAATCTGAGTTTTGCTGACTTCTGGACACTGGTCAATT

C57BL/6  ATCTGTGGTCAAAAGGCTGGTTTCTTCTTATTAATGTATTCTTTCGTTGAATACTTATCAGACCTTAAAGGGGAACCCTGGGGAAAACAGACCTCTAATG
BALB/c   ATCTGTGGTCAAAAGGCTGGTTTCTTCTTATTAACGTATTCTTTCCTTGAATACTTATCAGACCTTAAAGGGGAACCCTGGGGAAAACAGACCTCTAATG

C57BL/6  CCTGCCTGGCGGGTATGAACGGTTCTGCTCCATGCTGTTAATCATAGCT
BALB/c   CCTGCCTGGCGGGTATGAACGGTTCTGCTCCATGCTGTTAATCATAGCT
```

**O**

***Asb4***

chr6:5,331,766-5,333,016

```
C57BL/6  ----------------------------------CTGTCTCTTGGCTTCAGCCTGGTGGATCCCAGTAGCATACAGCAATAACTCAAGTCCATAGCCATTG
BALB/c   ----------------------------------CTGTCTCTTGGCT-CAGCCTGGTGGATCCCAGTAGCATACAGCAATAACTCAAGTCCATAGCCATTG

C57BL/6  AGTTATGGTATGTGGAGGCAGGGCTTCCAGCCAAGCCAGAAATGTATCAGATAGCTTTCAGCCAAGGCAATTTATTGTAGCATGTTGGTGTTTGTTCCTA
BALB/c   AGTTATGGTATGTGGAGGCAGGGCTTCCAGCCAAGCCAGAAATGTATCAGATAGCTTTCAGCCAAGGCAATTTATTGTAGCATGTTGGTGTTTGTTCCTA

C57BL/6  AGTCGCCTTTGCTGGTGACTGATGTACCATCCTTCTAACACTCCTGAAAACCTCATAGTCATGCTGAAGAAAGAAATTCACCCAGGAGAAACGCCAGACG
BALB/c   AGTCGCCTTTGCTGGTGACTGATGTACCATCCTTCTAACACTCCTGAAAACCTCATAGTCATGCTGAAGAAAGAAATTCACCCAGGAGAAACGCCAGACG

C57BL/6  GATCAGGAGTACAAGCTCCCCGCCCTTTTATGCCCTGGAGCTATCAAAGGCCACCTAGCAACATTATCAGACAGGGAAAGGAGGGCATATCTTTGATTTC
BALB/c   GATCAGGAGTACAAGCTCCCCGCCCTTTTATGCCCTGGAGCTATCAAAGGCCACCTAGCAAGATTATCAGACAGGGAAAGGAGGGCATATCTTTGATTTC

C57BL/6  TCAGTCTTGCCTGGGGCCTTGACTGACAGTAGCTGCTGTTTTAACCTCCATCTGTGTCTTCGTGCCTTTTTAGAAGGGACTCAGGAAAGCTTGCTCTGGC
BALB/c   TCAGTCTTGCCTGGGGCCTTGACTGACAGTAGCTGCTGTTTTAACCTCCATCTGTGTCTTCATGCCTTTTTAGAAGGGACTCAGGAAAGCTTGCTCTGGC

C57BL/6  ATTATGGGGAATATCATCAACCTAGGAGGAGAATTCTGAACATGAGATGCAAGCTTGTCTTGCTTTCCCACAGCCCAAAACAGACGTATATGACAGTGAG
BALB/c   ATTATGGGGAATATCATCAACCTAGGAGGAGAATTCTGAACATGAGATGCAAGCTTGTCTTGCTTTCCCACAGCCCAAAACAGACGTATATGACAGTGAG

C57BL/6  TATGGGAAATAAACATATAGAAGAAAACACGTCTAGTTCGTTTACAGAACGCTTCTTAATACCAGTGACGTAGGAGCAAAGGATCGTGGGTGGGCGGCTA
BALB/c   TATGGGAAATAAACATATAGAAGAAAACACGTCTAGTTCGTTTACAGAACGCTTCTTAATACCAGTGACGTAGGAGCAAAGGATCGTGGGTGGGCGGCTA

C57BL/6  CTCCTCCCTGACAAGCCTTAATTACTTTTATATTTATTAGACATACATCTTGAAATGATTCTATAGAATCATAGGCAAACCTATGGTAAAATTTCAGCAG
BALB/c   CTCCTCCCTGACAAGCCTTAATTACTTTTATATTTATTAGACATACATCTTGAAATGATTCTATAGAATCATAGGCAAACCTATGGTAAAATTTCAGCAG

C57BL/6  AGAAATTCATCCTCGAAGCAAAATTTAATCAGCAAGGATAATGGTTTATATGGATTAAAACACATTTTTAAAATCTGCGTTTGATTTCATCTAGACCTCG
BALB/c   AGAAATTCATCCTCGAAGCAAAATTTAATCAGCAAGGATAATGGTTTATATGGATTAAAACACATTTTTAAAATCTGCGTTTGATTTCATCTAGACCTCG

C57BL/6  TTTT-ATCTGCATTGGCTCATTTCTCCCATTTTTGAGAGATTTTGCAGCTGGTCTTTCAGGTTCATTTTTCAAGGAATACTGCCCTCATGCGGCCGGGTG
BALB/c   TTTTTATCTGCATTGGCTCATTTCTCCCATTTTTGAGAGATTTTGCAGCTGGTCTTTCAGGTTCATTTTTCAAGGAATACTGCCCTCATGCGGCCGGGTG

C57BL/6  TGCCAAGACCGCTCACTCTCTGATTCCCTCCAGAGGACGGTCTGTGTGGCTCTGTGAGAGAGAGCTTTTGCTGCTTGAGGCTGTGTCTGTGGGTTCCTTT
BALB/c   TGCCAAGACCGCTCACTCTCTGATTCCCTCCAGAGGACGGTCTGTGTGGCTCTGTGAGAG----CTTTTGCTGCTTGAGGTTGTGTCTGTGGGTTCCTTT

C57BL/6  CGAAAACATGAATCCTTAAAATGCTTGCCAGGGGATGCACTCTGGTGCAGCTTTTGACTCAATTATCTGCTTACTGCATCAATAACCAACACATGCAAAA
BALB/c   CGAAAACATGAATCCTTAAAATGCTTGCCAGGGGATGCACTCTGGTGCAGCTTTTGACTCAATTATCTGCTTACTGCATCAATAACCAACACATGCAAAA

C57BL/6  ATAATCAGTCGCTGATCATTTTACACACTCGGAAAGTCCTCACAAACCCGACTCCTTTAAAGTCAGTAGAACGCGTTAGCCTTG
BALB/c   ATAATCAGTCGCTGATCATTTTACACACTCGGAAAGTCCTCACAAACCCGACTCCTTTAAAGTCAGTAGAACGCGTTAGCCTTG
```

**P**

***Asb4***

chr6:5,333,153-5,335,375

```
C57BL/6   ----------AGTGGAACAAACAACGGAGGTGGGCTCTGTACTTTCTTAAAAAGACAAGGTTTGACAAGATACTGTTTCTCAGCTCCAAT
BALB/c    ----------AGTGGAACAAACAACGGAGGTGGGCTCTGTACTTTCTTAAAAAGACAAGGTTTGACAAGATACTGTTTCTCAGCTCCAAT

C57BL/6   TCAGCCGTTCTATAGCAAGAATATCTGTAAGCAGACAGAAAATATGCAACCCTTCTCTCCACTCTTCCCGGGTCCTGCAGCTTCATTTGCATGATACACT
BALB/c    TCAGCCGTTCTATAGCAAGAATATCTGTAAGCAGACAGAAAATATGCAACCCTTCTCTCCACTCTTCCCGGGTCCTGCAGCTTCATTTGCATGATACACT

C57BL/6   CCAAACAAATTCTCTGCAAGAATGTCGATTTTACATTATATCCAGGGAAATGCGCAAAGTGCTACCCCAAAAGGTAAAATCAAATCGATGAATCTTTGCA
BALB/c    CCAAACAAATTCTCTGCAAGAATGTCGATTTTACATTATATCCAGGGAAATGCGCAAAGTGCTACCCCAAAAGGTAAAATCAAATCGATGAATCTTTGCA

C57BL/6   ACTGGTAAACTGCTTCTTGGTTTCTGTGAACTCTGAGGGCTAAAGGGCATCCTGGGATCTTAGGAAAGAAGGCTTAGAAGCCTGCCATATTTCTTTCCCA
BALB/c    ACTGGTAAACTGCTTCTTGGTTTCTGTGAACTCTGAGGGCTAAAGGGCATCCTGGGATCTTAGGAAAGAAGGCTTAGAAGCCTGCCATATTTCTTTCCCA

C57BL/6   CATGAAGGGCAATTTTCTTCCCTAGCTATACTTTGCCTTCAAAAATTTGTCAGAGGTAAAGAAAGAATGCAAACGCTGCCTGGAGTACAATTGCGGGGTC
BALB/c    CATGAAGGGCAATTTTCTTCCCTAGCTATACTTTGCCTTCAAAAATTTGTCAGAGGTAAAGAAAGAATGCAAACGCTGCCTGGAGTACAATTGCGGGGTC

C57BL/6   AGTGTTCTGATTGAGACAGAATTATGGCGAATTATTCCCATGAGTTGATTTTGTGAAGTGTGTGCTGGTG-TGAGGGTGGAAATACAAGCCGGCGCTAGT
BALB/c    AGTGTTCTGATTGAGACAGAATTATGGCGAATTATTCCCATGAGTTGATTTTGTGAAGTGTGTGCTGGTG-TGAGGGTGGAAATACAAGCCGGCGCTAGT

C57BL/6   GTTTAAACAAGGCTGCAAGGGCAGTTGCTAGGAGACCCTGCACACAGCTGATGCCCCCACTGCGGGCTCAGACGTGCTTTGTCTGAAGAACCCGGCCTTT
BALB/c    GTTTAAACAAGGCTGCAAGGGCAGTTGCTAGGAGACCCTGCACACAGCTGATGCCCCCACTGCGGGCTCAGACGTGCTTTGTCTGAAGAACCCGGCCTTT

C57BL/6   CGCAGTCAAAGATGACATTAAATGTATAGTCTGTCAAAATTTAGGTGGAGGAAAAATTCAGGTGGACCAGGGCTCGGTTATTGCTTGTGTCAATCTAGCT
BALB/c    CGCAGTCAAAGATGACATTAAATGTATAGTCTGTCAAAATTTAGGTGGAGGAAAAATTCAGGTGGACCAGGGCTCGGTTATTGCTTGTGTCAATCTAGCT

C57BL/6   ATTCTGACGTCCATCACCCATTAAGGATGGGTTTTGAAAGGGAAAATATAGCATTGTGCCAGGAACCAAGAGCCTCCGGAAACGCTCAAGGGAATTTAGT
BALB/c    ATTCTGACGTCCATCACCCATTAAGGATGGGTTTTGAAAGGGAAAATATAGCATTGTGCCAGGAACCAAGAGCCTCCGGAAACGCTCAAGGGAATTTAGT

C57BL/6   TACAAATGCAGTTTGAATACCCCGATATGAAAGGCAGATGGTCTGGGTACCAGTTCCGGGAAGCAGGCGGCCCCAGCTAAGTAATCACTTCAATGTTCCA
BALB/c    TACAAATGCAGTTTGAATACCCCGATATGAAAGGCAGATGGTCTGGGTACCAGTTCCGGGAAGCAGGCGGCCCCAGCTAAGTGATCACTTCAATGTTCCA

C57BL/6   AATAGAGGACCTTGAGTCCACAAACAGAGACTTCTAAGAGGAAAGGCGCGCGAGTGTGATTTTAGATCCCGCCGTTTACAAACAACCATTTCAGTCATTT
BALB/c    AATAGAAGACCTTGAGTCCACAAACAGAGACTTCTAAGAGGAAAGGCGCGCGAGTGTGATTTTAGATCCCGCCGTTTACAAACAACCATTTCAGTCATTT

C57BL/6   GTGGGAACCGTTTAATAGACTTTTATGAGAAACACATAGGCTTCCAGTTTGTAATTTAGACCTTAGAGATGCATCGAACTCGAGAGCTGAAAGTGGCTTA
BALB/c    GTGGGAACCGTTTAATAGACTTTTATGAGAAACACATAGGCTTCCAGTTTGTAATTTAGACCTTAGAGATGCATCGAACTCGAGAGCTGAAAGTGGCTTA

C57BL/6   GAAGTTCCTCAAGCGCCCCCAACTGGTGGAAGGAGAAAACTGACCCGCTTTGTGGATAAGTGCAGTCTAAAAAAAACGTTGCTTTCTTGGTTTTCAGAAA
BALB/c    GAAGTTCCTCAAGCGCCCCCAACTGGTGGAAGGAGAAAACTGACCCGCTTTGTGGATAAGTGCAGTCTAAAAAAAACGTTGCTTTCTTGGTTTTCAGAAA

C57BL/6   CTCAGGCAAGACTACTTGCCTTTGTAAAATAGTATCACCCACCCTCTGCCAATGCTCAAGATATGATACCCGGTGAGGAGTAAAGTAGAGGAAAGAGGAA
BALB/c    CTCAGGCAAGACTACTTGCCTTTGTAAAATAGTATCACCCACCCTCTGCCAATGCTCAAGATATGATACCCGGTGAGGAGTAAAGTAGAGGAAAGAGGAA

C57BL/6   CATTTTGAATTTTGTGTTTTGAATATCAGTAGCACGGGAGCCAGAGACACAGTCAGTTAAGATTGCCTGTCTTCTCCAGGAGAGTCATGTTGTTCTGTGA
BALB/c    CATTTTGAATTTTGTGTTTTGAATATCAGTAGCACGGGAGCCAGAGACACAGTCAGTTAAGATTGCCTGTCTTCTCCAGGAGAGTCATGTTGTTCTGTGA

C57BL/6   GCATCGAGACCAGAATCATCCACAACTCTTTGTAGTGCTTTTCCGTGTACAAGCACATTCTGTGTATTAATTTTTTTCATTTGTTTCATAAAGTCTGACA
BALB/c    GCATCGAGACCAGAATCATCCACAACTCTTTGTAGTGCTTTTCCGTGTACAAGCACATTCTGTGTATTAATTTTTTTCATTTGTTTCATAAAGTCTGACA

C57BL/6   ATTGCTCTTACACGCTGGTGGTGAAACTAATAAGGATTTGATACAAGGTTTAGTTGTAACTCCACTACCCTACCTCCAAATAAGGAATAACTTCCTTGGC
BALB/c    ATTGCTCTTACACGCTGGTGGTGAAACTAATAAGGATTTGATACAAGGTTTAGTTGTAACTCCACTACCCTACCTCCAAATAAGGAATAACTTCCTTGGC

C57BL/6   CAGTGATCTCACAGAATAGCTAATTGAAATCTCAAAGTGGTTTGAATTTTTAATTTTGTTTGTTTTGTATTTATACTTCATAGCCAGGAAAAACAAAAAC
BALB/c    CAGTGATCTCACAGAATAGCTAATTGAAATCTCAAAGTGGTTTGAATTTTTAATTTTGTTTGTTTTGTATTTATACTTCATAGCCAGGAAAAACAAAAAC

C57BL/6   AAAACAAAAAAAACCTCTGTACTTTTTTTGCTTTTTTATGAAAAAAAAAAAAAAAGAAACCCCATCAGCACTTCCCATGTTGGAGAAATTAGTATTACTAA
BALB/c    AAAACAAAAAAACCTCTGTACTTTTTTTGCTTTTTTATGAAAAAAAAAAAAA-GAAACCCCATCAGCACTTCCCATGTTGGAGAAATTAGTATTACTAA

C57BL/6   CCCTGGTACATAGTAACTTTCGGAATGTTGTTCCAGGCAGCCTTGAGAGAACTGAGCTGGTTGGGTTCAAGAAACAGAAGGAAATAAAAATTACCCATCA
BALB/c    CCCTGGTACATAGTAACTTTCGGAATGTTGTTCCAGGCAGCCTTGAGAGAACTGAGCTGGTTGGGTTCAAGAAACAGAAGGAAATAAAAATTACCCATCA

C57BL/6   GTCCCACTCTTAAACTTTACTGCAAAGAGATGTCTTTAGAGAGATGGAACCCAGGGGCTGATATGGCTTCTTTGCATCCTCTTGGCAGAGCTTAGGGGCC
BALB/c    GTCCCACTCTTAAACTTTACTGCAAAGAGATGTCTTTAGAGAGATGGAACCCAGGGGCTGACATGGCTTCTTTGCATCCTCTTGGCAGAGCTTAGGGGCC

C57BL/6   CAGGACTTGTTGCTTTAGCCCACAAAGCTCTGTCCTTACAGAGAAAACCAGTCTCGATGTTCCCTCCCCTCCCCTTTTCCTTCGTTCTAGGAGATGCCAA
BALB/c    CAGGACTTGTTGCTTTAGCCCACAAAGCTCTGTCCTTACAGAGAAAACCAGTCTCGATGTTCCCTCCCCTCCCCTTTTCCTTCGTTCTAGGAGATGCCAA

C57BL/6   CCACCATGCGGGGAGGGTATTCAGACAGGCTCCTGGAGAGACTG
BALB/c    CCACCATGCGGGGAGGGTATTCAGACAGGCTCCTGGAGAGACTG
```

**Q**

***Cd48***

chr1:173,509,829-173,510,806

```
C57BL/6   ----------------------------------------------------GACATGAGCATTAAGGCCTCCCTCTGCACTGGTCCTCATAAGCCCTGCTT
BALB/c    ----------------------------------------------------GACATGAGCATTAAGGCCTCCCTCTGCACTGGTCCTCATAAGCCCTGCTT

C57BL/6   GCCTGCAGGTGCTGCTGGCTTGCAGGTCAGCTGGATCTCAGCTGCTCTGAGCCTTCTGTCACCTTCTGCTCATTGGTTTGCTAGGTTGCTAGGAAACCGT
BALB/c    GCCTGCAGGTGCTGCTGGCTTGCAGGTCAGCTGGATCTCAGCTGCTCTGAGCCTTCTGTCACCTTCTGCTCATTGGTTTGCTAGGTTGCTAGGAAACCGT

C57BL/6   GTCCCAGGCTTCCATGCTGGGCTTGGCCTAAAGCCTAGTCTGCCTAAGTTTAGATTTCTAGCCCTACATTTGGTGAGCCGAATACTGTTAGACCACCTCA
BALB/c    GTCCCAGGCTTCCGTGCTGGGCTTGGCCTAAAGCCTAGTCTGCCTAAGTTTAGATTTCTAGCCCTACATTTGGTGAGCTGAATACTGTTAGACCGCCTCA

C57BL/6   ACTCGCCAGCTCCCAATAATGACTTTTTACTAGTTATAAAAGCTGAGGCCTTTTATCTGTTTTTT----------------------------------
BALB/c    ACTCGCCAGCTCCCAATAATGACTTTTTACTAGTTATAAAAGCTGAGGCCTTTTATCTGTTTTTTTTTTCTCTTTGTCTTGTCTTGTTTTCCTGTATAAA

C57BL/6   ----------------------------------------------------------------------------------------------------
BALB/c    AAAGATCCCAGTATAAATCAGGGAGGGAAGAAGACAGATACACCTGCCCTTTAGTGTAGAGCTAGGGGAGGGTAATGGGGGTGAACCTCGTAAGGAGAGG

C57BL/6   ----------------------------------------------------------------------------------------------------
BALB/c    TGGGGACTTGTTATTTAAGGCTTAAAATATTCTCTGTAAAAAACGAGGATTCTCCACAGCCTTGGGGGAGACAGAGGCTGCCGCTTCAGTGAGGGTTTCT

C57BL/6   ----------------------------------------------------------------------------------------------------
BALB/c    CCACCAAGACTGGGAAAGGCAGCCCCTCTTCTCACTCTGCGTCGCTTTCCTCCTTGTGCAGGTAGGAGTGCTGCAGGGCTCGGAAGGCAGAGATTCGCTT

C57BL/6   ----------------------------------------------------------------------------------------------------
BALB/c    ATGTGGGTTAAAGGTCAGCATTTCCAGTAGCAGCTGCGCTCCAGACTCCTCCATCTCTGGCACCACTGACTGCACTGGCCGAGGCCCTCTGGGGGCAAAG

C57BL/6   ----------------------------------------------------------------------------------------------------
BALB/c    GCTCCTCGAGGTAGAGATACCTCTCGAGGCCAGTCGTCTTCTGGAGGCAATCCAATGAGATCAAAGATTTTCCCCAACTGGTGGGCTTCAGAGTTTCCAC

C57BL/6   ----------------------------------------------------------------------------------------------------
BALB/c    AGAAGAGAGGCTTCCGACGGAACATCTCTGCAAAGATACAGCCAACGCTCCACATGTCCACGGGTGTTGCGTATGTAGACTGCAGAAGAACTTCAGGAGC

C57BL/6   ----------------------------------------------   ----------------------------------------------------
BALB/c    TCGGTACCAAAGCGTACCACCACAGGCGTGAGGCCATCTGG //GGAGTACACGACTCCTCACCGTCACATTGGAGTCAAAGGCTCCAGCTCCTAACAAG

C57BL/6   ----------------------------------------------------------------------------------------------------
BALB/c    CCACTCACGAACTGTGCT-GATGGAAGGCCCCTCCAGCTGCTCCTCCATTAGGACTCTCACACTTCTTGAGGCCACAAAGTGGCCACTGTGGGGATCTCG

C57BL/6   ----------------------------------------------------------------------------------------------------
BALB/c    GGTTTGTACACCGTCCCATAGGCACCGACACCAATTTCAGCCACGGGTTCATATCGAGTGGCAGCCATTCTCGAAGCAGGGGATCTTACGCTCGGCTAAG

C57BL/6   ----------------------------------------------------------------------------------------------------
BALB/c    AGCTCTGGAGGTGGCCCTTTATCTGGGCCCGGGAGCCGGTTCCTGCGGCGCCATACACCGGGGATCGGTCCCGGGCAGCTGGACGCTTCGGGCCAGACCA

C57BL/6   --------------------------ATGCAATGCAAATATTTACTCTTAATTCACATTGTCAAAATTTTTCAAGAAACAATCACAAAGGTTTATTTTAAA
BALB/c    --TAGACACAGTTATCTGTTTTTTTATGCAATGCAAATATTTACTCTTAATTCACATTGTCAAAAATTTTCAAGAAACAATCACAAAGGTTTATTTTAAA

C57BL/6   TGTGCTTTGTATCTTTGGCTTGTTCTCGACTAGCTTGTATAACTTAGTAACTCCACAGGTAGTCACTATGTCTGACTTTCTCCATGTCTGGCTGGAGCAT
BALB/c    TGTGCTTTGTATCTTTGGCTTGTTCTCGACTAGCTTGTATAACTTAGTAACTCCACAGGTAGTCACTATGTCTGACTTTCTCCATGTCTGGCTGGAGCAT

C57BL/6   TTCCCAGGCCTGATTCTTTCCCAGAGTTCCTATCTATACTCTGAAGTTTCACTTTTCTTCTCCTGTCCTGTTATAGACCATCAGATCTTTATTAAACCAA
BALB/c    TTCCCAGGCCTGATTCTTTCCCAGAGTTCCTATCTATATTCTGAAGTTCCGCTTTTCTTCTCCTGTCCTGTTATAGACCATCAGATCTTTATTAAACCAA

C57BL/6   TCAGAAGATGATGGAGAAAATGTTTACGAAATCTTTAGGCAGGTGTGATGATCCAATAAAAATAACAATACCAAGGTCAGGCCTGTAAACAGCTCTGCGG
BALB/c    TCAGAAGATGATGGAGAAAATGTTTACGAAATCTTTAGGCAGGTGTGATGATCCAATAAAAGTAACAATACCAAGGTCAGGCCTGTAAACAGCTCTGCGG

C57BL/6   GTAGGTAAACCAATATTTGAATAATACAAGGGTAACCTTTACACAGTGCACAAAAAGATTAACCCAACTGAGGAGGAGAGAGGAGGAAGGGAAGAAGGGT
BALB/c    GTAGATAAACCAATATTTGAATAATACAAGGGTAACCTTTACACAGTGCACAAAAAGATTAACCCAACTGAGGAGGAGAGAGGAGGAAGGGAAGAAGGGT

C57BL/6   AGAGACAAGAGCAGAGTGGGACAAAGGTGGCCTGCTAAATTGACTTTGGTATTAAAGGAGGAATCTGTGTAAATGAAAGAGCTGGATGAAAATCCAGTAC
BALB/c    AGAGACAAGAGCGGAGTGGGACAAAGGTGGCCTGCTAAATTGACTTTGGTATTAAAGGAGGAATCTGTGTAAATGAAAGAGCTGGATGAGAATCCAGTAC

C57BL/6   ATGAACACTGTCTGGTCTGTTTGCTTGGTTTAGTTTTGAGCCTCTCAGCATCCCTTCCTAGGTGTT-AGAAGTCCATCTCATCCCTGG
BALB/c    ATGAACGCTGTCTGGTCTGTTTGCTTGGTTTAGTTTTGAGCCTCTCAGCATCTCTTCCTAGGTGTT-AGAAGTCCATCTCATCCCTGG
```

**R**

***Coro2a***

chr4:46,621,090-46,622,106

```
C57BL/6  --------------------------------------------------GTTCGGTCGATTCGGTTTTGGGGATGGAACTCTGCCGACTGAGCTGCACCCTAGC
BALB/c   --------------------------------------------------GTTCGGTCGATTCGGTTTTGGGGATGGAACTCTGCCGACTGAGCTGCACCCTAGC

C57BL/6  TGCTAAGCACCCCATTCTTGGAATTAGGATTCCTGGCTGGTTTGTTGTGTGGCCTTTACCAGGCACACCCCCTCCTGGGACAGGTGATTATATGCCCATA
BALB/c   TGCTAAGCACCCCATTCTTGGAATTAGGATTCCTGGCTGGTTTGTTGTGTGGCCTTTACCAGGCACACCCCCTCCTGGGACAGGTGATTATATGCCCATA

C57BL/6  GAATACGAGCGTTGGATGAGATAGACGGTTTCTAGACTGCCCCTAGAATAAGGGCCTGGGATGAGATGGTTTCTAGAGACCCTTTCAGGTCAGAGAATCC
BALB/c   GAATACGAGCGTTGGATGAGATAGACGGTTTCTAGACTGCCCCTAGAATAAGGGCCTGGGATGAGATGGTTTCTAGAGACCCTTTCAGGTCAGAGAATCC

C57BL/6  ACAACTCCACAGTACATTTTTAAAGTTCGGGATGACTTTGTAATAATCAGGACCTGCCTGCTCATGCGCCATCACTGTAGCCTCCCAGATCACAAACGTT
BALB/c   ACAACTCCACAGTACATTTTTAAAGTTCGGGATGACTTTGTAATAATCAGGACCTGCCTGCTCATGCGCCATCACTGTAGCCTCCCAGATCACAAACGTT

C57BL/6  CCCAATGACAGGGGACTCAGGTGGTACTTAGGCTTTCGAGCTGGCGCTGGGCTCTGGCAGCCACACTTATACTTAGTCTCCCTTTATGAGGGGATAATAA
BALB/c   CCCAATGACAGGGGACTCAGGTGGTACTTAGGCTTTCTAGCTGGCGCTGGGCTCTGGCAGCCACACTTATACTTAGTCTCCCTTTATGAGGGGATAATAA

C57BL/6  CACTCATTTGCAGCGAAGCTCCAGGTTGGGGCCACAGGGCAAGGTCCCCTGCCACACTGCTTTGGCCCTGGCAACGCACAGCACCCTATGATGCCCACCC
BALB/c   CACTCATTTGCAGCGAAGCTCCAGGTTGGGGCCACAGGGCAAGGTCCCCTGCCACAGTGCTTTGGCCCTGGCAACGCACAGCACCCTATGATGCCCACCC

C57BL/6  AGCCCTCTGCCTGTACACTTAGATGACGTAGGGCTCACTACCGGCGCTGGCCTGGGGAAAGGGAAGGGTCTCCTCTTTTAGCCTTAAGAAAGTCCTTGGG
BALB/c   AGCCCTCTGCCTGTACACTTAGATGACGTAGGGCTCACTACCGGCGCTGGCCTGGAGAAAGGGAAGGGTCTCCTCTTTTAGCCTTAAGAAAGTCCTTGGG

C57BL/6  TGAGGGAGAGGGAGACAGGCGATCAGGAGGTGCTGAGCCTGTTCCTTAGTGGGGACAGATGTCAGGAAATGAGGCAGAAGTGTCCAGAGATGGGCTTAAG
BALB/c   TGAGGGAGAGGGAGACAGGCGATCAGGAGGTGCTGAGCCTGTTCCTTAGTGGGGACAGATGTCAGGAAATGAGGCAGAAGTGTCCAGAGATGGGCTTAAG

C57BL/6  CCTTTCTCTCCCTCCACATCCCTCAAGAGATGTCTGGAGGCGAGGGACGTGTAAGGACACATTCCTGTGCTGAGGCACAGGTGGTTATGAACAAGTCCCT
BALB/c   CCTTTCTCTCCCTCCACATCCCTCAAGAGATGTCTGGAGGCGAGGGACGTGTAAGGACACATTCCTGTGCTGAGGCACAGGTGGTTATGAACAAGTCCCT

C57BL/6  GGAGCATTCGCAGAAGAGCAGGCTCTGAGACAGAAGGGCCACATGCTAGCTAAGGGACACTATCAGGATTGCAGCCACAGTGGGACCAGATTCCCCAGTT
BALB/c   GGAGCATTCGCAGAAGAGCAGGCTCTGAGACAGAAGGGCCACATGCTAGCTAAGGGACACTATCAGGATTGCAGCCACAGTGGGACCAGATTCCCCAGTT

C57BL/6  TCGCTAGTGAGAGTCCCCATCTCGAGAAGAATGCAACATAAGACCT-CATCAAAAGGCATCT
BALB/c   TCGCTAGTGAGAGTCCCCATCTCGAGAAGAATGCAACATAAGACCT-CATCAAAAGGCATCT
```

**S**

***Eps8l1***

chr7:4,081,175-4,081,945

```
C57BL/6  ---------------------GCCCCAGCACCTAGAGTGCGCACTCCCTCCCTTCCCGGTTTCCCTCCTGCAGAGAAATTCTCCCAGATGCTTTGTGTCAA
BALB/c   ---------------------GCCCCAGCACCTAGAGTGCGCACTCCCTCCCTTCCCGGTTTCCCTCCTGCAGAGAAATTCTCCCAGATGCTTTGTGTCAA

C57BL/6  TGAGGAGCTGCAGTCGCGCCTGGCGCAGGGCCGTTCGGGTCCCAGCCGGGTAACCCCGGGGCCCCGAGCCCAAGAGCCCCAGCTCAGCCCGCGCTCTGAG
BALB/c   TGAGGAGCTGCAGTCGCGCCTGGCGCAGGGCCGTTCGGGTCCCAGCCGGGTAACCCCGGGGCCCCGAGCCCAAGAGCCCCAGCTCAGCCCGCGCTCTGAG

C57BL/6  GCCTCGGTGGTCCGTGCCTGGCTGCAGACCAAGGGCTTTAGCTCCGGGTGAGTGGGGTCCCGCTGCTACTTTGCAGAGACAGATAACAGACCTGGAAGTC
BALB/c   GCCTCGGTGGTCCGTGCCTGGCTGCAGACCAAGGGCTTTAGCTCCGGGTGAGTGGGGTCCCGCTGCTACTTTGCAGAGACAGATAACAGACCTGGAAGTC

C57BL/6  AAGGTTTTAGTTGGTCTGGGGGTGGGAGCGGGGAAAACTGTAGGGCAATTTGCTACGCCCAAAGACAGGCTCTGATTGGAGACCACAGGGAAGAGCATTC
BALB/c   AAGGTTTTAGTTGGTCTGGGGGTGGGAGCGGGGAAAACTGTAGGGCAATTTGCTACGCCCAAAGACAGGCTCTGATTGGAGACCACAGGGAAGAGCATTC

C57BL/6  CTTTTCTGAAATCTGACACCTCCGGAGCCTGAGCCTCTCCGTCCCAACTTCCTTGCTTGAGAGCTCTGATTGGACAAGCAGCTAGGTGGGTGCGATGTGA
BALB/c   CTTTTCTGAAATCTGACACCTCCGGAGCCTGAGCCTCTCCGTCCCAACTTCCTTGCTTGAGAGCTCTGATTGGACAAGCAGCTAGGTGGGTGCGATGTGA

C57BL/6  TTGGTGAAGCTGAAGGGGCTTGAGCTGTGATGGCCCCTGCTGCCTTCTTAGGACTGTGGAGGCGCTCGGTGTGCTGACCGGCGCACAGCTCTTCTCGCTG
BALB/c   TTGGTGAAGCTGAAGGGGCTTGAGCTGTGATGGCCCCTGCTGCCTTCTTAGGACTGTGGAGGCGCTCGGTGTGCTGACCGGCGCACAGCTCTTCTCGCTG

C57BL/6  CAAAAGGAAGAGTTGCGGGCGGTGTGCCCCGAGGAAGGGGCGCGAGTGTACAGCCAAGTCACTGTGCAGCGCGCGCTGCTGGAGGTGAGAGAATCGCCCA
BALB/c   CAAAAGGAAGAGTTGCGGGCGGTGTGCCCCGAGGAAGGGGCGCGAGTGTACAGCCAAGTCACTGTGCAGCGCGCGCTGCTGGAGGTGAGAGAATCGCCCA

C57BL/6  GGCCCTGACGAGGCGAGGGGTGGGGCGCCCAGGACTCACATACTAAGGTCTTGATTTCTACCCCACCCTCAAGGACAGAGAAAAAGTATCA
BALB/c   GGCCCTGACGAGGCGAGGGGTGGGGCGCCCAGGACTCACATACTAAGGTCTTGATTTCTACCCCACCCTCAAGGACAGAGAAAAAGTATCA
```

**T**

***Frap1***

chr4:147,344,467-147,345,448

```
C57BL/6  ----------------------------------------------AGGAGCTGTGAACTTGTTTCCAGTTGCGCCTATTAGATTATGGCCTGCCAGCCACT
BALB/c   ----------------------------------------------AGGAGCTGTGAACTTGTTCCAGTTGCGCCTATTAGATTATGGCCTGCCAGCCACT

C57BL/6  GGCTCAGCTCTATCAGCAGACAGGCAGAGTGGAAGACTTTCCCAACTCCTGTCTGAAGGTTATATAACACAGATTGAGGGTGAGGTAGCAAGAAAGAGTT
BALB/c   GGCTCAGCTCTATCAGCAGACAGGCAGAGTGGAAGACTTTCCCAACTCCTGTCTGAAGGTTATATAACACAGACTGAGGGTGAGGTAGCAAGAAAGAGTT

C57BL/6  GGACTGGAGACTGCTGAGCCAGGGAGGCTGGCTAAGGTAATGCATGCTGTTTACTCCAGGTCCTGGTTTATCTGGGCTGAGATGAGTCTTAACAGGGCTT
BALB/c   GGACTGGAGACTGCTGAGCCAGGGAGGCTGGCTAAGGTAATGCATGCTGTTTACTCCAGGTCCTGGTTTATCTGGGCTGAGATGAGTCTTAACAGGGCTT

C57BL/6  TGTCTGGCCATCATTCGGTGTTCTTCAAAATCCTGTTGAACCTGATCCTTGGGCGTCAGGTGGTTTGCATCTGTGGGGCTGTCTTCCTCTTTACACATTC
BALB/c   TGTCTGGCCATCATTCGGTGTTCTTCAAAATCCTGTTGAACCTGATCCTTGGGCGTCAGGTGGTTTACATCTGTGGGGCTGTCTTCCTCTTTACACATTC

C57BL/6  TCCTTCTAAAACGATCTTTCAACATTTACAAGGAGTTGTCATAGTTTGGGGCAGAGAGAAAGTCTCAGCCATGCCAGGCGTGGTCTGGGGTTGGTCATCC
BALB/c   TCCTTCTAAAACGATCTTTCAACATTTACAAGGAGTTGTCATAGTTTGGGGCAGAGAGGAAGTCTCAGCCATGCCAGGCGTGGTCTGGGGTTGGTCATCC

C57BL/6  ACACTTCTGTCAGGTTCTGTCTGGGCAGCAGTGAGTGGGCCTGACCTTTACCCAAGGCTCAGACTCAGAGAGAAGATTCCCTAGTACATGCAGATCAGAA
BALB/c   ACACTTCTGTCAGGTTCTGTCTGGGCAGCAGTGAGTGGGCCTGACCTTTACCCAAGGCTCAGACTCAGAGAGAAGACTCCCTAGTACATGCAGATCAGAA

C57BL/6  GGCTCGATTCCCTCTGTGCCTGCAGCGTGGTTTGCAGCTTGATGCTGGCCCCTATACCTACGTACACAGTGTCTGAGAAAACGGGGTGATGGCAGCTTTC
BALB/c   GGCTCGATTCCCTCTGTGCCTGCAGCGTGGTTTGCAGCTTGATGCTGGCCCCTATAGCTACGTACACAGTGTCTGAGAAAACGGGGTGATGGCAGCTTTC

C57BL/6  ATACCTTTAATTGAAGGGCACTGCTCTTCTGAAATCATGTAAAAGATGGAACTTTCCCTATAACTGTACAAAGGTCCCTCTGTCCTTATGGCAGGAGCTC
BALB/c   ATACCTTTAAGTGAAGGGCACTGCTCTTCTGAAATCATGGAAAAGATGGAACTTTCCCTATAACTGTACAAAGGTCCCTCTGTCCTTATGGCAGGAGCTC

C57BL/6  TGTCTGACTTTATAGTGTGGGTCCTACAGCTTTGCCTACCCTGATCCCTCAGAGTCTTTCCCAGATGGGTTTCTGGCTCCTCGGAAAGTGATTTCTGCTC
BALB/c   TGTCTGACTTTATAGTGTGGGTCCTACAGCTTTGCCTACCCTGATCCCTCAGAGTCTTTCCCAGATGGGTTTCTGGCTCCTCGGAAAGTGATTTCTGCTC

C57BL/6  AGACAGCTGGAGTTACTTAGCAGGGACCCATTTCTGAATATAAGCATTGGGTGCTCAGACAGACGACTGTCTGTTTCCTCTCACTCCTGTCTTTTTTGCC
BALB/c   AGACAGCTGGAGTTACTTAGCAGGGACCCATTTCTGAATATAAGCATTGGGTGCTCAGACAGACGACTGTCTGTTTCCTCTCACTCCTGTCTTTTTTGCC

C57BL/6  TCCTTGGGCACCAGGTTGTTACTAAC
BALB/c   TCCTTGGGCACCAGGTTGTTACTAAC
```

**U**

***Pdgfrb***

chr18:61,190,675-61,191,616

```
C57BL/6  ---------------------------------------------ACCTAGACTTCCCACCGAGCCTAGTTCATTCTCAACCCCTCAGTGCAGACCTTGACAGC
BALB/c   ---------------------------------------------ACCTAGACTTCCCACCGAGCCTAGTTCATTCTCAACCCCTCAGTGCAGACCTTGACAGC

C57BL/6  CCCGTTGCCAGAGTCAACCTCTGCATTAGTTTGAAGTTCCCTTGCACCCTGTATATAATATAGACCCAGAATAGAAAACTTCAACTCCACTCCCAACCTC
BALB/c   CCCGTTGCCAGAGTCAACCTCTGCATTAGTTTGAAGTTCCCTTGCACCCTGTATATAATATAGACCCAGAATAGAAAACTTCAACTCCACTCCCAACCTC

C57BL/6  AGCTCTGACCCTAACACTATCAGTTAGGGTCAAATTCCATCCCTGTTTCCGTTCCAGTCAAAGCTTGCACTTCAGCTCTGTTGCCTAAAGAAACCGTGCC
BALB/c   AGCTCTGACCCTAACACTATCAGTTAGGGTCAAATTCCATCCCTGTTTCCGTTCCAGTCAAAGCTTGCACTTCAGCTCTGTTGCCTAAAGAAACCGTGCC

C57BL/6  CTAGTTTTTAGAAGCCCCCAGAAGGTGTTTATCTTGCATTTAGTCCGCGCTGGCCTCCAGGAACCCTCCACTAGAGATGGGCAGGATG--------CATT
BALB/c   CTAGTTTTTAGAAGCCCCCAGAAGGTGTTTATCTTGCATTTAGTCCGCGCTGGCCTCCAGGAACCCTCCACTAGAGATGGGCAGGATGCACATTCTCATT

C57BL/6  GGTGTCTGGGCATAACTCCGCCCACAGGGCCCGCCCCTAACTGTTGTACCTTCCGCAGAGAATGGCTACGTGCGGCTGCTGGAGACACTGGGAGATGTAG
BALB/c   GGTGTCTGGGCATAACTCCGCCCACAGGGCCCGCCCCTAACTGTTGTACCTTCCGCAGAGAATGGCTACGTGCGGCTGCTGGAGACACTGGGAGATGTAG

C57BL/6  AAATTGCTGAGCTGCACCGGAGTCGGACGCTGCGGGTGGTGTTCGAGGCTTATCCGATGCCTTCTGTCCTGTGGCTCAAGGACAACCGTACCTTGGGTGA
BALB/c   AAATTGCTGAGCTGCACCGGAGTCGGACGCTGCGGGTGGTGTTCGAGGCTTATCCGATGCCTTCTGTCCTGTGGCTCAAGGACAACCGTACCTTGGGTGA

C57BL/6  CTCCGGCGCTGGCGAGTTAGTTTTGTCTACTCGCAACATGTCTGAGACCCGGTGAGCACCAGTTAGTCCACTTATGTTGGAAGCTTCAAAGCTGCCAGGC
BALB/c   CTCCGGCGCTGGCGAGTTAGTTTTGTCTACTCGCAACATGTCTGAGACCCGGTGAGCACCAGTTAGTCCACTTATGTTGGAAGCTTCAAAGCTGCCAGGC

C57BL/6  AGCCTGGGGCCTTCAGTCACCAACAAACCTAGGACCTTCTTGTCTGAGAGCCTGTTGTGTGATGGACACATGCCAGGTACTAGGTATGATGCTCCTGCAG
BALB/c   AGCCTGGGGCCTTCAGTCACCAACAAACCTAGGACCTTCTTGTCTGAGAGCCTGTTGTGTGATGGACACATGCCAGGTACTAGGTATGATGCTCCTGCAG

C57BL/6  GAGCAACGCAGGCTGCTCTGCTCCTCCTAGTGGGGGTAGCAGACACTGACGACTTGGCCACTCAGAGGTTTAGTTAACCTTGTGATGAACAGAACCAGCA
BALB/c   GAGCAACGCAGGCTGCTCTGCTCCTCCTAGTGGGGGTAGCAGACACTGACGACTTGGCCACTCAGAGGTTTAGTTAACCTTGTGATGAACAGAACCAGCA

C57BL/6  GGTAATGGCCAAAAGCAAAAA-CACCGGATAGGAAGAGGATTTCCCCAAGGCTAGAGATCTGCAGACAGGCAAGAATGTGCCCCTCTGATGG
BALB/c   GGTAATGGCCAAAAGCAAAAAACACCGGATAGGAAGAGGATTTCCCCAAGGCTAGAGATCTGCAGACAGGCAAGAATGTGCCCCTCTGATGG
```

**V**

***Isoc2b***

***chr7:4,468,816-4,470,431***

```
C57BL/6  ----------------------------------------------------------------------------ATGCTTTAGGGAGAGCGATACAGAGGA
BALB/c   ----------------------------------------------------------------------------ATGCTTTAGGGAGAGCGATACAGAGGA

C57BL/6  GCACGTAATTGTGAGAGAAGCGGGGACCACAACTCCTGGAACCAGGAATTTCAGGCTGTGTGACCCTGGGTGAGTTTCCCTTCCTCTCTGGGTCTCTTTTT
BALB/c   GCACGTAATTGTGAGAGAAGCGGGGACCACAACTCCTGGAACCAGGAATTTCAGGCTGTGTGACCCTGGGTGAGTTTCCCTTCCTCTCTGGGTCTCTTTTT

C57BL/6  TTTCTCCAGGGTTAGCTCACTGGCGGGGTTGATGCAAAATTTGTCTCTAGAGAGGGTTGGGGACTCTCCCCGCGGGTCTATCCTGAGACCAAATGGAATG
BALB/c   TTTCTCCAGGGTTAGCTCACTGGCGGGGTTGATGCAAAATTTGTCTCTAGAGAGGGTTGGGGACTCTCCCCGCGGGTCTATCCTGAGACCAAATGGAATG

C57BL/6  CCATTCCAACCCCCGAACCCGACACCCTGCAAAGCTTCCGTTTCCCGGTCTGAATGACGGGATGGTGGGTTTGCAAACAGCCAGCAAGTAACGAACCGGA
BALB/c   CCATTCCAACCCCCGAACCCGACACCCTGCAAAGCTTCCGTTTCCCGGTCTGAATGACGGGATGGTGGGTTTGCAAACAGCCAGCAAGTAACGAACCGGA

C57BL/6  CTTACTCGAAGGTCCCAGGTTCGAGTTCGGGCCCCTTTTCCAAGAAGTCCCGTGACCTTTTCCCCGCTGGATTCCCACTGACCCCGTACTGGGCCTCCTT
BALB/c   CTTACTCGAAGGTCCCAGGTTCGAGTTCGGGCCCCTTTTCCAAGAAGTCCCGTGACCTTTTCCCCGCTGGATTCCCACTGACCCCGTACTGGGCCTCCTT

C57BL/6  TTTGCCCTTTGTAGACACGTAGTCAGAGCCGCGGGTCTACACCCCCGCGCCGAGCCAACCAATCGGCGTTCGCCTCACTCTGTTCTACTCCATCCTGTTA
BALB/c   TTTGCCCTTTGTAGACACGTAGTCAGAGCCGCGGGTCTACACCCCCGCGCCGAGCCAACCAATCGGCGTTCGCCTCACTCTGTTCTACTCCATCCTGTTA

C57BL/6  ATTCGATTGGTCTGAAGAAACTGGAGCAGGCGGAGAGACAGAGGGTAGGACCAATGAAAATCCACGGAGGGAGGGCCTTGGCTGAGAGCTCGGCTTCTAA
BALB/c   ATTCGATTGGTCTGAAGAAACTGGAGCAGGCGGAGAGACAGAGGGTAGGACCAATGAAAATCCACGGAGGGAGGGCCTTGGCTGAGAGCTCGGCTTCTAA

C57BL/6  CCCGGAAGGAGGGTGAAGAACGACCAGAGAGCTAGAGGCTCAATCTTGGCTAGTGTAGGGTTCATGGCACCATCCACTCCATCCGGGGCGACCGGGGAGA
BALB/c   CCCGGAAGGAGGGTGAAGAACGACCAGAGAGCTAGAGGCTCAATCTTGGCTAGTGTAGGGTTCATGGCACCATCCACTCCATCCGGGGCGACCGGGGAGA

C57BL/6  GGTGATGCTTTCTGTAACCTAGCTCTAGGAAGGCTGAAGCACGAGTATCGCCAGTTCGAGACTAGCCGGGACTGCGTGGTGAGACCCTGTCAACCAAAAC
BALB/c   GGTGATGCTTTCTGTAACCTAGCTCTAGGAAGGCTGAAGCACGAGTATCGCCAGTTCGAGACTAGCCGGGACTGCGTGGTGAGACCCTGTCAACCAAAAC

C57BL/6  AACAGCCGGGCATGGTGGCGCAAATGTTTAGTTCCAGTAATCATCATAAACCGATGATTTAGTTAATGGTTTAGTTCCAGTACTCAGGAGGCAGAGGCAG
BALB/c   AACAGCCGGGCATGGTGGCGCTAATGTTTAGTTCCAGTAATCATCATAAACCGATGATTTAGTTAATGGTTTAGTTCCAGTACTCAGGAGGCAGAGGCAG

C57BL/6  GCATAATTCTGACTCTACAACGCGAATTCCAGAATAGCCAGTACAACACGGAAGTAAGTCCTGTCTCGAAAACCCGAAAGTCAAACTAAGAAAACAAAAG
BALB/c   GCATAATTCTGACTCTACAACGCGAATTCCAGAATAGCCAGTACAACACGGAAGTAAGCCCTGTCTCGAAAACCCGAAAGTCAAACTAAGAAAACAAAAG

C57BL/6  AACAAGCCTCGTAAGACATTTTCCCCC-ACTGCTTGCCAGTTAGCTTTTTCTCCAGGAATTCAGGACTTGTCTCCTTGGGTGGTCTATATAGAAAACTGA
BALB/c   AACAAGCCTCGTAAGACATTTTCCCCCCACTGCTTGCCAGTTAGCTCTTTCTCCAGGAATTCAGGACTTGTCTCCTTGGGTGGTCTATATAGAAAACTGA

C57BL/6  CTCTGTACTTAGAAGTGGGGGTTAAGAGAGAGAGGATTACTTCATCCAGATATATAAACCTGGTATGGTGGCTCAGG-CCGGTAATCCCAGCGTTCTAGA
BALB/c   CTCTGTACTTAGAAGTGGGGGTTAAGAGAGAGAGGATTACTTCATCCAGATATATAAACCTGGTATGGTGGCTCAGG-CCGGTAATCCCAGCGTTCTAGA

C57BL/6  AGCTGAGGCCTTGAGTTCAATGAATAAGAAAACAAAAGCTGATTACGGTGGCTTAATACCCATGACCCCAAGTACTCAGAAGTGGAGACAGGAGGATTTC
BALB/c   AGCTGAGGCCTTGAGTTCAATGAATAAGAAAACAAAAGCTGATTACGGTGGCTTAATACCCATGACCCCAAGTACTCAGAAGTGGAGACAGGAGGATTTC

C57BL/6  TTCACCTTGAAGTTATGAGAAAAGGGCTAGGCATGGTGGCTAACGAGGACTGTAAACCCAGCACTTGGGAGGCTGAGTGTGGGTGAGGATCTTTTGGGGG
BALB/c   TTCACCTTGAAGTTATGAGAAAAGGGCTAGGCATGGTGGCTAACGAGGACTGTAAACCCAGCACTTGGGAGGCTGAGTGTGGGTGAGGATCTTTTGGGGG

C57BL/6  AGAGATATATATATACACAGAGTTCCAGGTTAGCCTGGGCCACAGAGTGTGACACTGTCGAACAAAGAAAAAGAAATAAGGCTGGAACTGTAGCTCAGGT
BALB/c   AGAGATATATATATACATAGAGTTCCAGGTTAGCCTGGGCCACAGAGTGTGACACTGTCGAACAAAGAAAAAGAAATAAGGCTGGAACTGTAGCTCAGGT

C57BL/6  GGCAGAGTCCTTGCCTGGCATGAATAAAGCCTTCGCTTTAGTGCCCAGCACTGTATGCATGGGGGTGGGGTGAGGGTGG-AGAGATGGGG
BALB/c   GGCAGAGTCCTTGCCTGGCATGAATAAAGCCTTCGCTTTAGTGCCCAGCACTGTATGCATGGGGGTGGGGTGAGGGTGGGAGAGATGGGG
```

**W**

***Pi4k2b***

chr5:53,026,310-53,027,151

```
C57BL/6  ------------------------AAGGCCACACAAGATATTTTAGAACTGAGCTTGGATAAGGGCAGTAACAATAGCAGTTGTGGTTGGCATCTGGAAT
BALB/c   ------------------------AAGGCCACACAAGATATTTTAGA-CTGAGCTTGGATAAGGGCAGTAACAATAGCAGTTGTGGTTGGCATCTGGAAT

C57BL/6  GCCCTCCTCCGCCGGCTCACGCTTTGTTCCCAGGCTGCTGATGTGAGAGGTGGTACAGCCTTTAGGAAGATGGGTCTGGTGGGCAGAAGTATGTGACTCT
BALB/c   GCCCTCCTCCGCCGGCTCACGCTTTGTTCCCAGGCTGCTGATGTGAGAGGTGGTACAGCCTTTAGGAAGATGGGTCTGGTGGGCAGAAGTATGTGACTCT

C57BL/6  GGAGGTGACAGTGCCTTTGATATGGGACATGGCTTTTGCTCCATGTCTTTCCTAATGGGGGGGGGGAGGGGGGTGTCTCTGCCTCAGGCACCTCCTGTAA
BALB/c   GGAGGTGACAGTGCCTTTGATATGGGACATGGCTTTTGCTCCATGTCTTTCCTAATGGGGGGGGG-AGGGGGGTGTCTCTGCCTCAGGCACCTCCTGTAA

C57BL/6  GACCTTTTCCCACTCAGATGACTTGGAATTTCTGGAATCATAATTGAAAATGATCTTTCCTCTCCTGGGTTGTTTCCATCAGATGCTTTGGTCATGGTGG
BALB/c   GACCTTTTCCCACTCAGATGACTTGGAATTTCTGGAATCATAATTGAAAATGATCTTTCCTCTCCTGGGTTGTTTCCATCAGATGCTTTGGTCATGATGG

C57BL/6  CAAGAAAGGCGACCGACGTGCGATCCACGACCATTTGAAAATGCCGGTGAAATAAACCACCCCAAGTAAATAAAACCATAGACATCATCCCCAGATCTGT
BALB/c   CAAGAAAGGCGAGCGACGTGCGATCCACGACCATTTGAAAATGCCGGTGAAATAAACCACCCCAAGTAAATAAAACCATAGACATCATCCCCAGATCTGT

C57BL/6  GTTATAACTAACCGTGCTAATTTCTTCACTGGGAAGGAAAATGGTACTAGGTAGAAAATCAGACTTCTAAGAGGGAATAAAACAGGGGCTGGAGAGATGG
BALB/c   GTTATAACTAACCGTGCTAATTTCTTCACTGGGAAGGAAAACGGTACTAGGTAGAAAATCAGACTTCTAAGAGGGAATAAAACAGGGGCTGGAGAGATGG

C57BL/6  CTCAGCGAGTTAAGAGCACTGACTGCTCTTCAGAAGGTCCTGAGTTCAAATCCCAGCAACCACATGGTGGCTCACAACCAGATATGACGCCCTCTTCTGG
BALB/c   CTCAGCGAGTTAAGAGCACTGACTGCTCTTCAGAAGGTCCTGAGTTCAAATCCCAGCAACCACATGGTGGCTCACAACCAGATATGACGCCCTCTTCTGG

C57BL/6  TCTGTCTGAAGACAGCTACAGTGTACTTACATATAATGATAATGATGATGATAATAATAATAATAATTTAATAAATAAAAGAGGTAATAAAACAGTACATT
BALB/c   TCTGTCTGAAGACAGCTACAGTGTACTTACATATAATGATAATGATGATGATAATAATAATAATAATTTAATAAATAAAAGAGGTAATAAAACAGTACATT

C57BL/6  AGTAATTAGAGCCATCGTGTAATATTTATTTTAGGAAGAAGCTTCACAGGGAAAGGGAGGAGCATG
BALB/c   AGTAATTAGAGCCATCGTGTAATATTTATTTTAGGAAGAAGCTTCACAGGGAAAGGGAGGAACATG
```

**X**

**Pop4**

chr7:37,961,843-37,962,567

```
C57BL/6  ----------------------------------------CTGTTTCTATCTCTTTGGGGCTGCTGCCAGCAGGCCTAGGCTTCCCTATGCACTGGTCTAGA
BALB/c   ----------------------------------------CTGTTTCTATCTCTTTGGGGCTGCTGCCAGCAGGCCTAGGCTTCCCTATGCACTGGTCTAGA

C57BL/6  ATGCATGTCTCTCCTAGTAAGCCTTTCTGCCTTTAGGCCCTGGTTGTGGGTACACAAAGGTTCAAAGGGACAGAGATTGCGTTTCTGACACCGTGCTCAT
BALB/c   ATGCATGTCTCTCCTAGTAAGCCTTTCTGCCTTTAGGCCCTGGTTGTGGGTACACAAAGGTTCAAAGGGACAGAGATTGCGTTTCTGACACCGTGCTCAT

C57BL/6  GAGTACTGTTCCCATGCCTCATAACAATATCTGAGATGGCTGTGGTGGAATATAATATGCCTACAATCCTAGCATTTGGAAGGGTGAGGCAGAGAGATCC
BALB/c   GAGTACTGTTCCCATGCCTCATAACAATATCTGAGATGGCTGTGGTGGAATATAATATGCCTACAATCCTAGCATTTGGAAGGGTGAGGCAGAGAGATCC

C57BL/6  TGAGTTCCCAGATAGCCTGGTATATAGGGTAAGTTTCAGGCTAGCCTGTGCTGTAAAGCTAGAGACCCTATTTTAAAAAAAGAATCAAGCCGGGCGTGGT
BALB/c   TGAGTTCCCAGATAGCCTGGTATATAGGGTAAGTTTCAGGCTAGCCTGTGCTGTAAAGCTAGAGACCCTATTTTAAAAAAAGAATCAA----------

C57BL/6  GGCGCACGCCTTTAATCCCAGCACTTGGGAGGCAGAGACAGGCGGATTTCTGAGTTCGAGGCCAGCCTGGTCTACAAAGAGAGTTCCGGGACAGCCAAGA
BALB/c   ----------------------------------------------------------------------------------------------------

C57BL/6  CTATACAGAGAAACCCTGTCTCGAAAAACCAAAAAAAAAAAAAAAAAAAA----AAAGAATCAAAACTAATCAACTAATGAAAGATCAAAAAAGGCTGGG
BALB/c   ------------------------------------------------------------AACTAATCAACTAATGAAAGATCAAAAAAGGCTGGG

C57BL/6  AGTATGGGTCATGGTCACAGGCTCAATGGTGTTCGTTTCCTTCCCGGCATTCAAACAAACAACAAAGAAAACCTTTCTTCTGTGCATCTTCAATGCTGAC
BALB/c   AGTATGGGTCATGGTCACAGGCTCAATGGTGTTCGTTTCCTTCCCGGCATTCAAACAAACAACAAAGAAAACCTTTCTTCTGTGCATCTTCAATGCTGAC

C57BL/6  CCTGGTCTAGGATGCAGATATTCCTCACAGGCCAGAGCCACCTCCTCACTGGCATATGTTCTT
BALB/c   CCTGGTCTAGGATGCAGATATTCCTCACAGGCCAGAGCCACCTCCTCACTGGCATATGTTCTT
```

**Y**

***Ppp1r14d***

chr2:118,921,131-118,922,245

```
C57BL/6  ----------------------------------------------TCTGTGGACGATGCCCTCCTTCTCTTCTCAGAAGCCCAGTGGACCTTCTTGTCTGGG
BALB/c   ----------------------------------------------TCTGTGGACGATGCCCTCCTTCTCTTCTCAGAAGCCCAGTGGACCTTCTTGTCTGGG

C57BL/6  TTGTCTGTGCCAGGGCTGGGAGATGTACAGGAAGCAGGGCTTGAAGACAGCATGGAGAGATCAGACTGTGCAAGGTGCTGGGAGAAACTGCCAGTCTAGG
BALB/c   TTGTCTGTGCCAGGGCTGGGAGATGTACAGGAAGCAGGGCTTGAAGACAGCATGGAGAGATCGGACTGTGCAAGGTGCTGGGAGAAACTGCCAGTCTAGG

C57BL/6  CAAGAGAATAGAGGGAAATGAGGGGACAGAGAATGGAGGGAGAGGCCCACAAGGAAGAAAGGAGGTGGGAAGAGGGCCAAGGCTCTTGGGAGGGGCAGGG
BALB/c   CAAGAGAATAGAGGGAAATGAGGGGACCGAGAATGGAGGGAGAGGCCCACAAGGAAGAAAGGAGGTGGGAAGAGGGCCAAGGCTCTTGGGCGGGGCAGGG

C57BL/6  GTCAAAGATCAATAGGTCTCTAGTTAGTGCCGGGAATGAACTTCCTAGCCTCTCTCTGCCAGCCAGTGGCATTCAGAGCCATATCCCAGCCTCCATTCTG
BALB/c   GTCAAAGATCAATAGGTCTCTAGTTAGTGCCGGGAATGAACTTCCTAGCCTCTCTCTGCCAGCCAGTGGCATTCAGAGCCGTATCCCAGCCTCCATTCTG

C57BL/6  CAGCTGCTGGAGCCCACAGTAACATAAACCTCAAACAAGCAAATGTCAAGAGGCTAGCAAGCCTTGAGAAGAGGAAGAAGGAGCCTTTGAGTGCTGAAGG
BALB/c   CAGCTGCTGGAGCCCACGGTAACATAAACCTCACACAAGCAAATGTCAAGAGGCTAGCAAGCCTTGAGAAGAGGAAGAAGGAGCCTTTGAGTGCTGAAGG

C57BL/6  GTGTGGGGCTCCCACCTGTCTTCCGGGATAGCTCTAGACATACGGTGACTAGGCTCAATCAACTGAAGTCCCTCAGCATCCAATCTTCTAGGAACTAGAT
BALB/c   GTGTGGGGCTCCCACCTGTTTTCCAGGATAGCTCTAGACATACGGTGACTAGGCTCAATCAACTGAAGTCCCTCAGCATCCGATCTTCTAGGAACTAGAT

C57BL/6  AGTCTATTGCAGGCCTGAGCAGCTGTGGTCGGATAATTCTGGTTCCTGCAGGGGTAGGATTCCTCCTCAAGGTGAGGAGCCACCGAGGGCAATCTCAGCC
BALB/c   AGTCTATTGCAGGCCTGAGCAGCTGTGGTCGGATAATTCCGGTTCCTGCGGTGGTAGGATTCCTCCTCACAGTGAGGAGCCACCGAGGGCAATCTCAGCC

C57BL/6  AGTGGATAGAGTTCACAGATGGCTACTTCTGTCCTGTGTCTGGGTGTGTCTGGATATTGGCTTTGCTGCTGTGATTGTTAAAGGCAGATTTATGCAGCCT
BALB/c   AGTGGATAGAGTTCAAAGATGGCTACTTCTGTCCTGTGTCTGGGTGTGTCTGGATATTGGCTTTGCTGCTGTGATTGTTAAAGGCAGATTTATGCAGCCT

C57BL/6  GGCCTACAAGACCTGTTGTCATTCTTAGAAGCCAAGTTTATGGTCATGTGACCCTCCCCTGATTCAGCTTCTTCCTGATAAAAGACACCTTCAGACCTGA
BALB/c   GGCCTACAAGACCTGTTGTCATTCTTAGAAGCCAAGTTTATGGTCATGTGACCCTCCCCTGATTCAGCTTCTTCCTGATAAAAGACACCTTCAGACCTGA

C57BL/6  CTTCACATAGGAATCCTGGACTGGTAAACAAGCTTGCCATTTAAAAGTACTTTCATTATTCAGTCCTTTATAAGAAAAAATTTAAATAACTTAATGTGGT
BALB/c   CTTCACATAGGAATCCTGGACTGGTAAACAAGCTTGCCATTTAAAAGTACTTTCATTATTCAGTCCTTTATAAGAAAAAATTTAAATAACTTAATGTGGT

C57BL/6  TCATTTTGTCTTGTTTTTCGTTTTAATGAACAGCTCAGAAAAACAGATAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA
BALB/c   TCATTTTGTCTTGTTTTTCGTTTTAATGAACAGCTCAGAAAAAAAGATAGAA-GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAG

C57BL/6  AGAAAGGAAGGAAGG--AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG
BALB/c   AGAGAGAGAGAGAGGGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG
```

**Z**

***Rab6b***

chr9:102,967,555-102,969,590

```
C57BL/6   ------------------------------------TATCGTGGGGGCTGTTGGATGAAGAGACACAGAAACTTCCTGTACTATTTTTTGAAACCTCTATGT
BALB/c    ------------------------------------TATCGTGGGGGCTGTTGGATGAAGAGACACAGAAACTTCCTGTACTATTTTTTGAAACCTCTATGT

C57BL/6   AGTCTAAAACTATTTCAGGCTTTAAAAACTGGACACATTTTTACTGAAAATGGTTTCAAAATTGGGGCAGACGATGGAATGCTAAGCAGACACGGGAACC
BALB/c    AGTCTAAAACTATTTCAGGCTTTAAAAACTGGACACATTTTTACTGAAAATGGTTTCAAAATTGGGGCAGACGATGGAATGCTAAGCAGACACGGGAACC

C57BL/6   AAAGGCTCTGCTCTATGCATCAACAGGAATCTTGGGAGTACAGCGAGAAGGGAAGAGCAAGTCAGAAGACAGGGTTTGAAATATGAATTCAAAAGACTCA
BALB/c    AAAGGCTCTGCTCTATGCATCAACAGGAATCTTGGGAGTACAGCGAGAAGGGAAGAGCAAGTCAGAAGACAGGGTTTGAAATATGAATTCAAAAGACTCA

C57BL/6   AGAGAAATGAAAGTGTTGAGAGTTTGGGTTTTCTGCCTTGGGGAAGGAGGAAAAGGGGGAGGGGATAAAAGAAGAATTAGAAGACAGGTGTGTGTTCACT
BALB/c    AGAGAAATGAAAGTGTTGAGAGTTTGGGTTTTCTGCCTTGGGGAAGGAGGAAAAGGGGGAGGGGATAAAAGAAGAATTAGAAGACAGGTGTGTGTTCACT

C57BL/6   GCGATATGATGTCTTTCCATGTTCCTCATCCAGAGAACACACTCATCAGGAAGGGCAGAAAGATGCCATGCAGAAGAAACTCTTTCTAGTGAGACTACCA
BALB/c    GCGATATGATGTCTTTCCATGTTCCTCATCCAGAGAACACACTCATCAGGAAGGGCAGAAAGATGCCATGCAGAAGAAACTCTTTCTAGTGAGACTACCA

C57BL/6   AAAAGAGGCTTATAGTGGCTCTTTTTATCTATCACATGGGAAATGCCTGATATCCCAGGACAGAATAAAGATGATCCCTTCCCCTAAGAGACAAAGGAAT
BALB/c    AAAAGAGGCTTATAGTGGCTCTTTTTATCTATCACATGGGAAATGCCTGATATCCCAGGACAGAATAAAGATGATCCCTTCCCCTAAGAGACAAAGGAAT

C57BL/6   GGCAAGGGTAAGGGAGAGACTTGGAAGAGAGCAGAAGCATTCTAAGGTTAGGATTCACATACAGCATGAAGACCATGCTGTGTAGCTCTGCCCAGACTGC
BALB/c    GGCAAGGGTAAGGGAGAGACTTGGAAGAGAGCAGAAGCATTCTAAGGTTAGGATTCACATACAGCATGAAGACCATGCTGTGTAGCTCTGCCCAGACTGC

C57BL/6   AGGAGCCAACCCAGCCGACTGAACCAGTACAGATGAGGAAATGCTGGGTGGAAGGTGTGACACCCCCCAACCCCCTGAATGCCTTGACCTCCACAGGCAG
BALB/c    AGGAGCCAACCCAGCCGACTGAACCAGTACAGATGAGGAAATGCTGGGTGGAAGGTGTGACACCCCCCAACCCCCTGAATGCCTTGACCTCCACAGGCAG

C57BL/6   GGAGGGGTGTGGCTTAGGAGTAGTTCTTTCACTGCTGGAGTGTCCAACTTCTAGCTACGGAGACACAACATTGCAATCTACAAACTTCACAATGTAAAAC
BALB/c    GGAGGGGTGTGGCTTAGGAGTAGTTCTTTCACTGCTGGAGTGTCCAACTTCTAGCTACGGAGACACAACATTGCAATCTACAAACTTCACAATGTAAAAC

C57BL/6   CAGGCAATGGGATCAACAACAATAAACAAAAACTTGCCCCAAAAAATGTAGAAACCCATCCAGTATCTCATAATGTTCTACATAAGATTACAATATTGTG
BALB/c    CAGGCAATGGGATCAACAACAATAAACAAAAACTTGCCCCAAAAAATGTAGAAACCCATCCAGTATCTCATAATGTTCTACATAAGATTACAATATTGTG

C57BL/6   TTGTGCTGAATTCAAAGCTATCCTGCAAGGCATATGGCCTGTGGACAGTGGGTCGGACACACCTGCCAAAACTTTCCTCAGCCCATGCTGGATCCCGAGA
BALB/c    TTGTGCTGAATTCAAAGCTATCCTGCAAGGCATATGGCCTGTGGACAGTGGGTCGGACACACCTGCCAAAACTTTCCTCAGCCCATGCTGGATCCCGAGA

C57BL/6   CAGCAACACACTCGGGGCAAAAGTAGGGTGGAGTCTGAGTTATCTTAGTCTGGTGCCCAATCTGGCATGCTTATATGCTCAGAGGGATGACCTCCGACTC
BALB/c    CAGCAACACACTCGGGGCAAAAGTAGGGTGGAGTCTGAGTTATCTTAGTCTGGTGCCCAATCTGGCATGCTTATATGCTCAGAGGGATGACCTCCGACTC

C57BL/6   CAGGGACTTCAAACCGAAGAAGACCACAAGAATAGAACAGTTACTCTGATTTGAGAGTAGTTTTAGAGGTTAATTCGAGCATTCAGCAACTGTTTCCTAA
BALB/c    CAGGGACTTCAAACCGAAGAAGACCACAAGAATAGAACAGTTACTCTGATTTGAGAGTAGTTTTAGAGGTTAATTCGAGCATTCAGCAACTGTTTCCTAA

C57BL/6   ACTCTAGGTGAAGTGGAACCTCAGCTGCTTCCATGGAGTTTCCTGTGAGGGCTCAATGTGCGGTGGGAGACTCTAGTGGGTACATTCGGGGTAAATTCAT
BALB/c    ACTCTAGGTGAAGTGGAACCTCAGCTGCTTCCATGGAGTTTCCTGTGAGGGCTCAATGTGCGGTGGGAGACTCTAGTGGGTACATTCGGGGTAAATTCAT

C57BL/6   AAGTGTCCCCTGGGAAATCCGTAGGCACGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGCATTTCACTCCTGGCTTGCTTTGTGTAACATACACTACACA
BALB/c    AAGTGTCCC-TGGGAAATCCGTAAGCACGGGTGTGTGTGTGTGTGTGTGTGTGTGTGCATTTCACTCCTGGCTTGCTTTGTGTAACATACACTACACA

C57BL/6   TTCCAAACAGTCCCAAACAAACATTCCCTGAATGGCCAGAGAAGGTCAAGACCAGCAAGTGGCTGGAGAGGAAGAGGGGAGGGGCTTTGTAAAGCCTCAA
BALB/c    TTCCAAACAGTCCCAAACAAACATTCCCTGAATGGCCAGAGAAGGTCAAGACCAGCAAGTGGCTGGAGAGGAAGAGGGGAGGGGCTTTGTAAAGCCTCAA

C57BL/6   TTAACCCGATGCCTTAGAGTGTACTGCTTGGGGACTGCGTTTAACCTTAGTCTGGCGTCTACTCCTTGCTCTCACGCTACTGTGTGTGCCCTGCAAACTT
BALB/c    TTAACCCGATGCCTTAGAGTGTACTGCTTGGGGACTGCGTTTAACCTTAGTCTGGCGTCTACTCCTTGCTCTCACGCTACTGTGTGTGCCCTGCAAACTT

C57BL/6   AGAGGTCGCTCTCTTATACTCTTGTAGCTCATTACCCGAATAAACCTAACTAGCAGGGCGCAGGAGTGTGGTTGGGCTGACGGTCTCAGGATCAAAAGAG
BALB/c    AGAGGTCGCTCTCTTATACTCTTGTAGCTCATTACCCGAATAAACCTAACTAGCAGGGCGCAGGAGTGTGGTTGGGCTGACGGTCTCAGGATCAAAAGAG

C57BL/6   GCAGGGCATTGTCTTTGGGTTAGTGGAGCTGTCTTCTACCAGGCAGGTTAGGTTTCAAACCCTAGTAGGCATATCCTGTAAGGGTGTCTTCTCTCTCACT
BALB/c    GCAGGGCATTGTCTTTGGGTTAGTGGAGCTGTCTTCTACCAGGCAGGTTAGGTTTCAAACCCTAGTAGGCATATCCTGTAAGGGTGTCTTCTCTCTCACT

C57BL/6   GGCCAGCTCCTCTGGGGACAACCACAGATCCCCTCCGCATGAAGAGTCTGAGCCTAGCTAGATAGGAGGTAAGCTAGAGGCCAGCAGGGGCCACCGAGGA
BALB/c    GGCCAGCTCCTCTGGGGACAACCACAGATCCCCTCCGCATGAAGAGTCTGAGCCTAGCTAGATAGGAGGTAAGCTAGAGGCCAGCAGGGGCCACCGAGGA

C57BL/6   GCCATACCCCAGGTCTGAGCCAGGTATCCGCACATCTTCCTACTCACCACCACACCACACACTGACCCTT
BALB/c    GCCATACCCCAGGTCTGAGCCAGGTATCCGCACATCTTCCTACTCACCACCACACCACACACTGACCCTT
```

## AA

### *Slc13a3*

chr2:165,163,937-165,164,881

```
C57BL/6  --------------------TAGAA---CGACATAGACTGATTCCTTTAT-TTCACCGTGAGCCTATGATGCGTTATCATATTCCCTCTTTCTGGGTAGGG
BALB/c   --------------------TAGCA---CGACATAGACTGATTCCTTTATCTTCACCGTGAGCCTATGATGCGTTATCATATTCCCTCTTTCTGGGTAGGG

C57BL/6  GGTACGCAAACCAAAGTTCAGAGAGGTTAAGTGACTGGCGACAGGTCACACAGTTGTTCTTCGGAAAGACTGTGAAGGGCTATATACCCCAGTAGCACCA
BALB/c   GGTAGCAAACCAAAGTTCAGAGAGGTTAAGTGACTGGCGACAGGTCACACAGTTGTTCTTCGGAAAGACTGTGAAGGGCTATATACCCCAGTAGCACCA

C57BL/6  AGCAAGCCGTCAGAGCTCAAACTTTAGAAGAGTGGGAGGACCGAGAGAGGCAGACCAGGGAGAGGCAAGGAAGGGAGAAGGGGAGCGGGTGCTCCTAGGG
BALB/c   AGCAAGCCGTCAGAGCTCAAACTTTAGAAGAGTGGGAGGACCGAGAGAGGCAGACCAGGGAGAGGCAAGGAAGGGAGAAGGGGAGCGGGTGCTCCTAGGA

C57BL/6  GTAGGCAGGCTCAAGGCGGGAGGGTCTCGACCCGGGACCACCGTACCTTGGGCGGCAGGGCGAAGAGAATGGGCAGCAGAGCCAGCGGCACCAGCAGCAG
BALB/c   GTAGGCAGGCTCAAGGCGGGAGGGTCTCGACCCGGGACCACCGTACCTTGGGCGGCAGGGCGAAGAGAATGGGCAGCAGAGCCAGCGGCACCAGCAGCAG

C57BL/6  CACCAGGAGGCGCCGCGCGCTCCACACCTTCTTGGCCAGCGCCGCCAGCGCCGCCATCCGCGCGATCGCTGGAGACCAGGCGGCCCGACTGCCCCTCAGC
BALB/c   CACCAGGAGGCGCCGCGCGCTCCACACCTTTTTGGCCAGCGCCGCCAGCGCCGCCATCCGCGCGATCGCTGGAGACCAGGCGGCCCGACTGCCCCTCAGC

C57BL/6  GGCCGCCGGCAGGGACCTAAAGCCCCAAGGGACGGGGAGGAGCCATGCGGAACCGCCCCGGGCCGCGCCTCCTTGGTCTGGACAAGGAGCAGGGAAGGGG
BALB/c   GGCCGCCGGCAGGGACCTAAAGCCCCAAGGGACGGGGAGGAGCCATGCGGAACCGCCCCGGGCCGCGCCTCCTTGGTCTGGACAAGGAGCAGGGAAGGGG

C57BL/6  CCCCAGAGCACCCCACCCGCCCCCGCGCCCCACTCTGGACTCACGTGGGAAATGGGGCAAAGCCCAGGTGGGACTAAAGGCCTCAGCACCGTCATCCCAC
BALB/c   CCCCAGAGCACCCCACCCGCCCCCGCGCCCCACTCTGGACTCACGTGGGAAATGGGGCAAAGCCCAGGTGGGACTAAAGGCCTCAGCACTTGTCATCCCAC

C57BL/6  TGTCCAACCCTGTAGGGACAGACCAAGCTACTGAGACCTCTAGTGACCGTGGTTGCAGCCTCTTCCAGACTTGGAGGTTGGTGGGGGGG----TTTAAGT
BALB/c   TGTCCAACCCTGTAGGGACAGACCACGCTACTGAGACCTCTAGTGACCGTGGTTGCAGCCTCTTCCAGACTTGGAGGTTGGTGGGGGGGGGGGTTTAAGT

C57BL/6  GAGATAGTGCTTCGTGCAAAGGGCCGGGTCACCTGTGGGTAGGAGCTTTCCCTTCTCAGTGACAGATGCCACAACCTGCTTGATATTGTGTGTGTTGGGA
BALB/c   GAGATAGTGCTTCGTGCAAAGGGCCCGGTCACCTGTGGGTAGGAGCTTTCCCTTCTCAGTGACAGATGCCACAACCTGCTTGATATTGTGTGTGTTGGGA

C57BL/6  CAGGCTCTTATTTAGCCCAAGCTGGCCTGCATCTCCCTATGCTGCTGAGGACAACCTTGAACTCCCATCCTCCTGC
BALB/c   CAGGCTCTTATTTAGCCCAAGCTGGCCTGCATCTCCCTATGCTGCTGAGGACAACCTT-AACTCCC-TCCTCCTGC
```

## AB

### *Sfi1*

chr11:3,092,948-3,093,861

```
C57BL/6  ---------------------------------------------AAACGCAACGTGAGCTCAGCTCCTTGCTTGGTCACTTGAGCTCAGTGTACCAACGAGC
BALB/c   ---------------------------------------------AAACGCAACGTGAGCTCAGCTCCTTGCTTGGTCACTTGAGCTCAGTGTACCAACGAGC

C57BL/6  CCTACGCCCCGCCACACCCCTGAATCGAACGTGGATCTGCGCAGACTAGTAAGCGAAGCACGAGCTGGAGAATCACTCAGCGGAGCTCTCCACCAGGGGG
BALB/c   CCTACGCCCCGCCACACCCCTGAATCGAACGTGGATCTGCGCAGACTAGTAAGCGAAGCACGAGCTGGAGAATCACTCAGCGGAGCTCTCCACCAGGGGG

C57BL/6  AGAAGGATGCATAGACCCTGCAGAAACAGCGCGAGTGCAGCCGTGACTGCGAGCGACCTTACGCACGCCCTGCTCGCACTTCCACGCATGGACCAGAGCC
BALB/c   AGAAGGATGCATAGACCCTGCAGAAACAGCGCGAGTGCAGCCGTGACTGCGAGCGACCTTACGCACGCCCTGCTCGCACTTCCACGCATGGACCAGAGCC

C57BL/6  TTGTGGGTAGGCAGCATCTCCTGCACTTCTACAGAAGAAAGCAGAAGCCCCGACCAAGCCGTTCCAGGACAAAGTTCCTAGGAACGCGCGCTCCGCCAAC
BALB/c   TTGTGGGAAGGCAGCATCTCCTGCACTTCTACAGAAGAAAGCAGAAGCCCCGACCAAGCCGTTCCAGGACAAAGTTCCTAGGAACGCGCGCTCCGCCAAC

C57BL/6  TTCAAAACGCAAGGGCGAGCCAGTCAAAAGCTGGAGGCGCTAGGCCGGGCCTCGTACCTGAGGGAGGTGACTCCTCAGCAGACGCTAGACAGAGGTGGTA
BALB/c   TTCAAAACGCAAGGGCGAGCCAGTCAAAAGCTGGAGGCGCTAGGCCGGGCCTCGTACCTGAGGGAGGTGACTCCTCAGCAGACGCTAGACAGAGGTGGTA

C57BL/6  GCAAAGGGCCCCAGCCCAGAGTTTCCGGCCCTGCGCCGACGGCGGCTCCTCCCCCGTCTCCACGGCAACGCGACCCGCGCCGGAAGCTGCTCTGTCCAGC
BALB/c   GCAAAGGGCCCCAGCCCAGAGTTTCCGGCCCTGCGCCGACGGCGGCTCCTCCCCCGTCTCCACGGCAACGCGACCCGCGCCGGAAGCTGCTCTGTCCAGC

C57BL/6  TTAGTGCGCTCTGTCTTCCAGGAAACAGAGACTTCTCTCTAAGCTCGGCTTCCTACGGTTTTCTTGCCCTGAAGTTTCTGCGCTTCTGAGAAGCGGGTCA
BALB/c   TTAGTGCGCTCTGTCTTCCAGGAAACAGAGACTTCTCTCTAAGCTCGGCTTCCTACGGTTTTCTTGCCCTGAAGTTTCTGCGCTTCTGAGAAGCGGGTCA

C57BL/6  CTAGTTACAACTCCTGCATGGGACCGGTGTGAGACGGAAAAGTGTCTCTTTTCATTGCGTGGTGGCTCACGTCGGTAATCTCCGCATCAAACAGCCGAAT
BALB/c   CTAGTTACAACTCCTGCATGGGACCGGTGTGAGACGGAAAAGTGTCTCTTTTCATTGCGTGGTGGCTCACGTCGGTAATCTCCGCATCAAACAGCCGAAT

C57BL/6  CTTTAAAGAGAACTTTCAGAACAAAGTATGAAGTTTCATACTAAGAACAGTGAAAACTAGTGGGTGGGTGTCTGAAGACGAAAACCAAGGGGCAGGAGAT
BALB/c   CTTTAAAGAGAACTTTCAGAACAAAGTATGAAGTTTCATACTAAGAACAGTGAAAACTAGTGGGTGGGTGTCTGAAGACGAAAACCAAGGGGCAGGAGAT

C57BL/6  ATGACTCACCAGTTAAATGCACTTTTCATTCATCCAGACTCCCAATACCCACACGG
BALB/c   ATGACTCACCAGTTAAATGCACTTTTCATTCATCCAGACTCCCAATACCCACACGG
```

## AC

***Sfi1*** (pseudogene or duplication)
(location unknown)

```
C57BL/6  ----------------------------------------------------CAAATTAACCATGAAAAATCTCATCCTTGGGGCTGGAGAGATGGCTCAGTG
BALB/c   ----------------------------------------------------CAAATTAACCATGAAAAATCTCATCCTTGGGGCTGGAGAGATGGCTCAGTG

C57BL/6  GTTAAGAGCACTGACTGCTCTTCCAGAGTTCCTGAGTTCAATTCCCAGCAACCACATGGTGGCTCACAACCATCTGTAATGGGATCTGATGCCCTCTTCT
BALB/c   GTTAAGAGCACTGACTGCTCTTCCAGAGTTCCTGAGTTCAATTCCCAGCAACCACATGGTGGCTCACAACCATCTGTAATGGGATCTGATGCCCTCTTCT

C57BL/6  GGTGTGTCTGAAGACAGCGACAGTGTACTCATATACATAAAATAAACAAATTGTTAAAAAAAAAAAAGAAAAGAAAAATCTCATCCTTGCAAGGCTAAGT
BALB/c   GGTGTGTCTGAAGACAGCGACAGTGTACTCATATACATAAAATAAACAAATTGTTAAAAAAAAAAAAGAAAAGAAAAATCTCATCCTTGCAAGGCTAAGT

C57BL/6  TTTTACTGATATATAATCTGCATTTAAGGCTGGAGATGGTTTAATTGTGCTTTTCCACCACAAAAAGAGCTCTAGGTTCAGTCATCAGCCCTGAATAGAA
BALB/c   TTTTACTGATATATAATCTGCATTTAAGGCTGGAGATGGTTTAATTGTGCTTTTCCACCACAAAAAGAGCTCTAGGTTCAGTCATCAGCCCTGAATAGAA

C57BL/6  CGGGGCATGGTGCCACACACCAGTAATCCCAGAGACAGGAGGATCTCTGCAAGCTTTTCTCTTTCTTTAAAATTTATGTATTCTGGAGCTGGAGAGATGG
BALB/c   CGGGGCATGGTGCCACACACCAGTAATCCCAGAGACAGGAGGATCTCTGCAAGCTTTTCTCTTTCTTTAAAATTTATGTATTCTGGAGCTGGAGAGATGG

C57BL/6  CTCAACAGTTAAAACCACTGATTGTTCCTTCAGAGATTCTGAGTTCAGTTCCCAGCAACCACATGGTGGCTCACAACCATCTATAATGGGATTCAATGCC
BALB/c   CTCAACAGTTAAGACCACTGATTGTTCCTTCAGAGATTCTGAGTTCAGTTCCCAGCAACCACATGGGGGCTCACAACCATCTATAAGGGGATTCAATGCC

C57BL/6  CTCTTCTGGTGTGTCTGGACAGCAACAGCGTACTCACATATATAAAATAAATCTTTTTTTAAAATTTTAAGAATTATGTATTCTGTGTATGTATAGAAGTG
BALB/c   CTCTTCTGGTGTGTCTGGACAGCAACAGCGTACTCACATATATAAAATAAATCTTTTTTAAAATTTTAAAATTTATGTATTCTGTGTATGTATAGAAGTG

C57BL/6  CTTTGCCTACATGTGTGTACGAGTACCACATGAGTACATAGTTTCCTTGGAGGTCAAAAGAAGGTATCAGATTCCATGGAATTGGTGATATGAGCTTGAT
BALB/c   CTTTGCCTACATGTGTGTACGAGTACCACATGAGTACATAGTTTCCTTGGAGGTCAAAAGAAGGTATCAGATTCCATGGAATTGGTGATATGAGCTTGAT

C57BL/6  TTTTTTAATTCATTTTGGCCATATGTTTAGAAGCCTTTATACTCTTGTCAAAGTTTTCAAGACTCATCCCATATGGAGTCCTGACCTACAGTTTGAAAAG
BALB/c   TTTTTTAATTCATTTTGGCCATATGTTTAGAAGCCTTTATACTCTTGTCAAAGTTTTCAAGACTCATCCCATATGGAGTCCTGACCTACAGTTTGAAAAG

C57BL/6  CTGAGATCAGGGCTGGAGAAATACCCCAGCAGTTAACAACAGTACATGATATTCTAGCAGAGAAATGAAGTTTGGTTCCTACCACCCACAACCTGTAGTT
BALB/c   CTGAGATCAGGGCTGGAGAAATACCCCAGCAGTTAACAACAGTACATGATATTCTAGCAGAGAAATGAAGTTTGGTTCCTACCACCCACAACCTGTAGTT

C57BL/6  TCCAACCTCCTTTAACTACAGTTCCAAGGAATCCAATAAGAGATCGATCCAACACTCGTGGCTCCAAGAGAGCCTGTACAAATGCACATGTACCCCCACA
BALB/c   TCCAACCTCCTTTAACTACAGTTCCAAGGAATCCAATAAGAGATCGATCCAACACTCGTGGCTCCAAGAGAGCCTGTACAAATGCACATGTACCCCCACA

C57BL/6  AACACATAACTAAAAACAATAAAAGTAAATCTGTGTTGGGGGGGAGAGGGGCTGAAGTGATGGCTCAGCTGT----------------------------
BALB/c   AACACATAACTAAAAACAATAAAAGTAAATCTGTGTTGGGGGGGAGAGGGGCTGAAGTGATGGCTCAGCTGT----------------------------

C57BL/6  ----------TGAGGACGAATCAGATTCAGCTGACCGCTTTTCCAAAGTACCCAGATTAAGTTCCCATAACCCAGGGGTGGCTCACAATCACCTGGAACT
BALB/c   ----------TGAGGACGAATCAGATTCAGCTGACCGCTTTTCCAAAGTACCCAGATTAAGTTCCCATAACCCAGGGGTGGCTCACAATCACCTGGAACT

C57BL/6  CCAGTTCTGAGGTTCCAATGCCCTCGTCTGACCTCTCCATTTGGGCTTCTGAGGGCAACAGGCAAGCAAAAGGTACCCAGACATCCATGCACGCAAAACA
BALB/c   CCAGTTCTGAGGTTCCAATGCCCTCGTCTGACCTCTCCATTTGGGCTTCTGAGGGCAACAGGCAAGCAAAAGGTACCCAGACATCCATGCACGCAAAACA

C57BL/6  GCCACACACATAAATCTTTTTTTTTA-AATGTTTTTAAATAATCATGAATTGGAGGTTAGCCTGGATTCTGAGGCCAGTGTGAGCTAGCTTGTGTTATAG
BALB/c   GCCACACACATAAATCTTTTTTTTTA-AATGTTTTTAAATAATCATGAATTGGAGGTTAGCCTGGATTCTGAGGCCAGTGTGAGCTAGCTTGTGTTATAG

C57BL/6  AGACGCCAGAAATGGAGCCCAAAGCCTTCAGGACACCAGACAATCAGTCTACCTATGAGCTGGCATTTCAGGTGTGTTGTGAGCCACCAGTGGTGGACGC
BALB/c   AGACGCCAGAAATGGAGCCCAAAGCCTTCAGGACACCAGACAATCAGTCTACCTATGAGCTGGCATTTCAGGTGTGTTGTGAGCCACCAGTGGTGGACGC

C57BL/6  TGGGAGTCCAACTGGTATTGTCTCAAAAAGGAGTAAGGGTTTCTTAAGCGCTGAGCCATCTCTCTGGCTCCTTTAATAAATCATCTTTCTTTTTAAAAAT
BALB/c   TGGGAGTCCAACTGGTATTGTCTCAAAAAGGAGTAAGGGTTTCTTAAGCGCTGAGCCATCTCTCTGGCTCCTTTAATAAATCATCTTTCTTTTTAAAAAT

C57BL/6  ATTTAATATTAAATTTTAAACCATGATAACTTCAAATTACTTGATAATTACCAGGCAATCTAAAGCCCAGGGACTACAATTAAGACAGCATTTACAGCAC
BALB/c   ATTTAATATTAAATTTTAAACCATGATAACTTCAAATTACTTGATAATTACCAGGCAATCTAAAGCCCAGGGACTACAATTAAGACAGCATTTACAGCAC

C57BL/6  GGGTTGCACACGTTTTTCTGAAGCACTTTTCATGAACGACCCACTCTAAGAACAGGCACCTCAGCTGGTGCTTATTGTCATTTTCCAGGCTGATGCTCAA
BALB/c   GGGTTGCACACGTTTTTCTGAAGCACTTTTCATGAACGACCCACTCTAAGAACAGGCACCTCAGCTGGTGCTTATTGTCATTTTCCAGGCTGATGCTCAA

C57BL/6  GACCATCAGCTAACAGG-TGGCAAAGGTGCCACTCTCATTCAAACCTCGATACACCTTCAGGAACTACACCAGCAGAGGTTCCATGGTACAAACGCAACG
BALB/c   GACCATCAGCTAACAGGCTGGCAAAGGTGCCACTCTCATTCAAACCTCGATACACCTTCAGGAACTACACCAGGAGAGGTTCCATGGTACAAACGCAACG

C57BL/6  TGAGCTCAG----------------------------TGTACCAACGAGCCCTACGCCCAGCCACACCCCTGAATCGAACGTGGATCTGCGCAGACTAGTAA
BALB/c   TGAGCTCAG----------------------------TGTACCAACGAGCCCTACGCCCAGCCACACCCCTGAATCGAACGTGGATCTGCGCAGACTAGTAA

C57BL/6  GCGAAGCACGAGCTGGAGAATCACTCAGCGGAGCTCTCCACCAGGGGGAGAAGGATGCATAGACCCTGCAGAAACAGCGCGAGTGCAGCCGTGACTGCGA
BALB/c   GCGAAGCACGAGCTGGAGAATCACTCAGCGGAGCTCTCCACCAGGGGGAGAAGGATGCATAGACCCTGCAGAAACAGCGCGAGTGCAGCCGTGACTGCGA

C57BL/6  GCGACCTTACGCACGCCCTGCTCGCACTTCCACGCATGGACCAGAGCCTTGTGGGTAGGCAGCATCTCCTGCACTTCTACAGAAGAAAGCGGAAGCCCCG
BALB/c   GCGACCTTACGCACGCCCTGCTCGCACTTCCACGCATGGACCAGAGCCTTGTGGGTAGGCAGCATCTCCTGCACTTCTACAGAAGAAAGCGGAAGCCCCG

C57BL/6  ACCAAGCCGTTCCAGGACAAAGTTCCTAGGAACGCGAGCTCCGCCAACTTCAAAACACAAGGGCGAGCCAGTCAAAAGCTGGAGGCGCTAGGCCGGGCCT
BALB/c   ACCAAGCCGTTCCAGGACAAAGTTCCTAGGAACGCGAGCTCCGCCAACTTCAAAACACAAGGGCGAGCCAGTCAAAAGCTGGAGGCGCTAGGCCGGGCCT

C57BL/6  CGTACCTGAGGGAGGTGACTCCTCAGCAGACGCTAGACAGAGGTGGTAGCAAAGGGCCCCAGCCCAGAGTTTCCGGCCCTGCGCCGACGGCGGCTCCTCC
BALB/c   CGTACCTGAGGGAGGTGACTCCTCAGCAGACGCTAGACAGAGGTGGTAGCAAAGGGCCCCAGCCCAGAGTTTCCGGCCCTGCGCCGACGGCGGCTCCTCC

C57BL/6  CCCGTCTCCACGGCAACGCGGCCCGCGCCGGAAGCTGCTCTGTCTTGCTTAGTGCGCTCTGTCTTCCAGGAAACCGAGACTTCTCTCTAAGCTCGGCTTC
BALB/c   CCCGTCTCCACGGCAACGCGGCCCGCGCCGGAAGCTGCTCTGTCTTGCTTAGTGCGCTCTGTCTTCCAGGAAACCGAGACTTCTCTCTAAGCTCGGCTTC

C57BL/6  CTACAGTTTTCTTGCCCTGAAGTTTCTGCGCTTCTGAGAAGCGGGTCACTAGTTACAACTCCTGCATGGGACCGGTGTGAGACGGAAAAGTGTCTCTTTT
BALB/c   CTACAGTTTTCTTGCCCTGAAGTTTCTGCGCTTCTGAGAAGCGGGTCACTAGTTACAACTCCTGCATGGGACCGGTGTGAGACGGAAAAGTGTCTCTTTT

C57BL/6  CATTGCGTGGTGGCTCACGTCGGTAATCTCCGCATCAAACAGCCGAATCTTTAAAGAGAACTTTCAGAACAGAGTATGAAGTTTCAAACTAAGAACAGTG
BALB/c   CATTGCGTGGTGGCTCACGTCGGTAATCTCCGCATCAAACAGCCGAATCTTTAAAGAGAACTTTCAGAACAGAGTATGAAGTTTCAAACTAAGAACAGTG

C57BL/6  AAAACTAGTGGGTGGGTGTCTGAAGACGAAAACCAAGGGGCAGGAGATATGACTCACCAGTTAAATGCACTTTTCATTCATCCAGACTCCCAATACCCAC
BALB/c   AAAACTAGTGGGTGGGTGTCTGAAGACGAAAACCAAGGGGCAGGAGATATGACTCACCAGTTAAATGCACTTTTCATTCATCCAGACTCCCAATACCCAC

C57BL/6  ACGGTGGCTCTTAAAACAGTTGTGACTCCCAAGTCGCGCTCCCGTTTGACCTATATGCAAGCAAAACATTCATACTCAAAATTTTTAAAAAGATTAAACC
BALB/c   ACGGTGGCTCTTAAAACAGTTGTGACTCCCAAGTCGCGCTCCCGTTTGACCTATATGCAAGCAAAACATTCATACTCAAAATTTTTAAAAAGATTAAACC

C57BL/6  CAAGAAATATGGGTTTTGCCCTTTTACAAATGGTAATTTTGTAATCCAAAGATTTA
BALB/c   CAAGAAATATGGGTTTTGCCCTTTTACAAATGGTAATTTTGTAATCCAAAGATTTA
```

## AD

### *Slc27a6*

chr18:58,681,397-58,682,391

```
C57BL/6  ----------------------------------------TTAGGCGGGGTTTTCAGAGGCAGCAAGAGGTGCTAAAGACATTCCCAGCTTGAAGCCCAGGG
BALB/c   ----------------------------------------TTAGGCGGGGTTTTCAGAGGCAGCAAGAGGTGCTAAAGACATTCCCAGCTTGAAGCCCAGGG

C57BL/6  AGCTGCAAGGAGCCTGTTGGACAGCCGTGGAGGGTGGGTGCACTGAGCTGTGTTGGGCTCATTTCCAGCGCTGTGAGTGTTGTGAGTACAATAGCCCCCC
BALB/c   AGCTGCAAGGAGCCTGTTGGACAGCCGTGGAGGGTGGGTGCACTGTGCTGTGTTGGGCTCATTTCCAGCGCTGTGAGTGTTGTGAGTACAATAGCTCCCC

C57BL/6  CCCCCCCTCTTTTTTTAATCTAAGAAGACTCAGGGCTCTGCTCCACAACTGTCTAGCCCAATGCCCCAGGGCGTATCACCAGCGTTCTGTAGCTTTCTGG
BALB/c   CCCCCCCTCTTTTTTTAATCTAAGAAGACTCAGGGCTCTGCTCCACAACTGTCTAGCCCAATGCCCCAGGGCTTATCACCAGCATTCTGAAGCTTTCTGG

C57BL/6  GCGAAAACCTTTGATTTCTTTCATGGGATATACAAAGACATCTTGAGCCAGAACTACAGGTGCCAGCCTCTGAGCTTTCAGGACAGGAGTCTCCAGGGCA
BALB/c   GCGAAAACCTTTGATTTCTTTCATGGGATATACAAAGACATCTTGAGCCAGAACTACAGGTGCCAGCCTCTGAGCTTTCAGGACAGGAGTCTCCGGGGCA

C57BL/6  AAAAGAGGTGAAACTCTTGGAAAAGAAAGAGACACTGCTAGGACCCATCGGGTCTGCTGGAGGTCTGGGCACCCATGCTCCTGTCATGGCTCACAGGATT
BALB/c   AAAAGAGGTGAAACTCTTGGAAAAGAAAGAGACACTGCTAGGACCCATCGGGTCTGCTGGAGGTCTGGGCACCCATGCTCCTGTCATGGCTCACAGGATT

C57BL/6  GGGGGCTGGACTGCTCTCCCTGCATTTCCTGCAGAAACTTCTGTTCCCGTATTTCTGGGATGATTTCTGGTACTTGCTGAAGGTGGTGCGCTACGGAATT
BALB/c   TGGGGCTGGACTGCTCTCCCTGCATTTCCTGCAGAAACTTCTGTTCCCGTATTTCTGGGATGATTTCTGGTACTTGCTGAAGGTGGTGCGCTACGGAATT

C57BL/6  CAGATGGAGATGTACAAACTGAGGGGGGAGCTGGTCACGGTGCTGGATAAGTTCCTGAGCCACACCAGGAAGCAACCACGGAAAGCCTTCATCATTTATG
BALB/c   CAGATGGAGATGTACAAACTGAGGGGGGAGCTGGTCACGGTGCTGGATAAGTTCCTGAGCCACGCCAGGAAGCAACCACGGAAAGCCTTCATCATTTATG

C57BL/6  AGGGGGACGTCTACACCTACGAGGACGTGGACAAGAGGAGTAACAGAATAGCCCACGCTCTCCTGAACCACTCCTCGCTGAAAAGGGGGGACGTTGTGGC
BALB/c   AGGGGGACGTCTACACCTACGAGGACGTGGACAAGAGGAGTAACAGAATAGCCCACGCTCTCCTGAACCACTCCTCGCTGAAAAGGGGGGACGTTGTGGC

C57BL/6  TTTGTTGATGAGCAACGAGCCCGACTTCGTTCATGTGTGGTTTGGCCTGGCTAAACTGGGCTGCGTGGTGGCCTTCCTCAACTCCAACCTTCGCTTCGAT
BALB/c   TTTGTTGATGAGCAACGAGCCCGACTTCGTTCATGTGTGGTTTGGCCTGGCTAAACTGGGCTGCGTGGTGGCCTTCCTCAACTCCAACCTTCGCTTCGAT

C57BL/6  TCCCTCCTACACTGCATCAACACCTGTGAACCCACTGCCGTGGTGGTGGGCGGAGGTAGAGTATGGACTGTGGTCTGCTGTGCCCAAAGCAAACCAACCT
BALB/c   TCCCTCCTACACTGCATCAACACCTGTGAACCCACTGCCGTGGTGGTGGGCGGAGGTAGAGTATGGACTGTGGTCTGCTATGCCCAAAGCAAACCAACCT

C57BL/6  ATTTCCTGCCCTCCCCTTTCTATCCATACTTCC
BALB/c   ATTTCCTGCCCTCCCCTTTCTATCCATACTTCC
```

## AE

### *Slc27a6*

chr18:58,685,847-58,687,080

```
C57BL/6  ----------------------------------------CTTCAGTA-GGTTGCATAACCAGG-ATCAAAATCTCGACTTCTGATGTGAGGTCTTGATTTGA
BALB/c   ----------------------------------------CTTCAGTAAGGTTGCATAACCAGGGATCAAAATCTCGACTTCTGATGTGAGGTCTTGATTTGA

C57BL/6  TCACAATCCAGCTATGTGGGGACCGTCAGTTACATTGCACACAGTTATAAAAAAGTATATATGTTCAGGCTCTATGAAAACACTACAAAATGGCAACATC
BALB/c   TCACAATCCAGCTATGTGGGGACCGTCAGTTACATTGCACACAGTTATAAAAAAGTATATATGTTCAGGCTCTATGAAAACACTACAAAATGGCAACATC

C57BL/6  AAGATACCAGGCTTTTAATACACACTAATTAGAAAAATGTGTACGATATTATTCTTCCAACAGAGCTGAAGACTCAAATATACTTGCAAAGGGGTCTCCT
BALB/c   AAGATACCAGGCTTTTAATACACACTAATTAGAAAAATGTGTACGATATTATTCTTCCAACAGAGCTGAAGACTCAAATATACTTGCAAAGGGGTCTCCT

C57BL/6  AGCTATGCTCTGGTCTCGGGGTCTGGGTCAGGTTAGTGTTCTCTGCTAATCAAAGAATGGAAATGAAAAGACAGGAAAAATCTCAGGCAGGACAAGTGAC
BALB/c   AGCTATGCTCTGGTCTCGGGGTCTGGGTCAGGTTAGTGTTCTCTGCTAATCAAAGAATGGAAATGAAAAGACAGGAAAAATCTCAGGCAGGACAAGTGAC

C57BL/6  GACAGCCTGAAGATTCTCATTGAGGCTGACAGAAACTGCTGTTTAAAAAAAAAAAAAAAAAAAAAAAAAACTTTATCAGGGAAGAGGAAATGCTACTCATAA
BALB/c   GACAGCCTGAAGATTCTCATTGAGGCTGACAGAAACTGCTGTTTAAAAAAAAAAAAAAAAAAAA----CTTTATCAGGGAAGAGGAAATGCTACTCATAA

C57BL/6  TAAGATAAGAAATGTTGTGGGTCTACCAGGTGTCCACGGTTCCCTTGCTACACATTGCGCGTGCGCGTTGTTTGCTTTCTCGATACTCTGCCGCTTAAAT
BALB/c   TAAGATA-----TGTTGTGGGTCTACCAGGTGTCCACGGTTCCCTTGCTACACATTGCGCGTGCGCGTTGTTTGCTTTCTCGATACTCTGCCGCTTAAAT

C57BL/6  CTCTACAAACAGATGACTTTACAATTCTCCCGTTGCCCGTCTCCCCGTCTCCCCTACTTTAAATGTGACAGCACTTAGTGTGTGAGTTTATTTTAAAAGA
BALB/c   CTCTACAAACAGATGACTTTACAATTCTCCCGTTGCCCGTCTCCCCGTCTCCCCTACTTTAAATGTGACAGCACTTAGTGTGTGAGTTTATTTTAAAAGA

C57BL/6  CTTGAGATTTCTCTTCATGAGAAGTCATGACTTCCGTTCACAGTTAAGGAGCTAATGGTTACTGCCCTGCCCTGTTTTTATAAGCAGTTGCCCTCAGGCT
BALB/c   CTTGAGATTTCTCTTCATGAGAAGTCATGACTTCCGTTCACAGTTAAGGAGCTAATGGTTACTGCCCTGCCCTGTTTTTATAAGCAGTTGCCCTCAGGCT

C57BL/6  GTGCAGAGGAAATGTTCAGGTTAACAACTGATCTGTTGTAACCCTGCAATTACCCTTTAGCCATAAAGGTAATTATTAAAATCAACAGGAAGAGAGGGAA
BALB/c   GTGCAGAGGAAATGTTCAGGTTAACAATTGATCTGTTGTAACCCTGCAATTACCCTTTAGCCATAAAGGTAATTATTAAAATCAACAGGAAGAGAGGGAA

C57BL/6  CGAGCCCCTTTGCAGTTCATAGAAAATTATATGTACAGAGGTACAGAAAGCTAAGGGACCCTGTGAACTTTATTACCTACCCTGTGTTCCATGATACCCT
BALB/c   CGAGCCCCTTTGCAGTTCATAGAAAATTATATGTACAGAGGTACAGAAAGCTAAGGGACCCTGTGAACTTTATTATCTACCCTGTGTTCCGTGATACCCT

C57BL/6  CTCACAGGACACGAATATTTTTCCAGGACATGGCTATTTTTTTTCTTGTATTTATTAGAATGGCTGTTACCACATTCTTATTTTTTCCCCATCTTGGATT
BALB/c   CTCACAGGACACGAATATTTTTCCAGGACATGGCTATTTTTTTTCTTGTATTTATTAGAATGGCTGTTACCACATTCTTATTTTTTCCCCATCTTGGATT

C57BL/6  TAGCAGGGTCGAGAAAGTAACCTTGCTACAAAAGAAAATCATTTTATTAATTTAAATGTATAGTGTAATTTCAAAAGAAGAAATCTAGACAGAGAAGTGA
BALB/c   TAGCAGGGTCGAGAAAGTAACCTTGCTACAAAAGAAAATCATTTTATTAATTTAAATGTATAGTGTAATTTCAAAAGAAGAAACCTAGACAGAGAAGTGA

C57BL/6  ATGAAACAGCGTGGCTCATACCCATCTCTTCTCAATGGAAAGGGGAAAATGACTTAACACATCCTGTGGAGAC
BALB/c   ATGAAACAGCGTGGCTCATACCCATCTCTTCTCAATGGAAAGGGGAAAATGACTTAACACATCCTGTGGAGAC
```

## AF

### *Spint1*

chr2:118,930,972-118,931,837

```
C57BL/6  --------------------------------------------------CAGAACCAGAGTGTGGAGTGCAGGGAGCAGGACACATAGGGCCAGTGCGACA
BALB/c   --------------------------------------------------CAGAACCAGAGTGTGGAGTGCAGGGAGCAGGACACATAGGGCCAGTGCGACA

C57BL/6  GACCCTTTCTCTCCAGGCTTCTTCCTACTAGGACCTATGGTAGACTAAAGGTCTGCTCTTCTAACTGTCCAATAGCGTCCAATAGCTCTCCTTGATCTCT
BALB/c   GACCCTTTCTCTCCAGGCTTCTTCCTACTAGGACCTATGGTAGACTAAGGGTCTGCTCTTCTAACTGTCCAATAGCGTCCAATAGCTCTCCTTGATCTCT

C57BL/6  GGGGAGACCCCAGGTGGGGATTCTGGAAAGTGGGGGATGCTGTGAAGTAGGTATCTCACTTGGTTAGCATTGGAGCCTGTTCCACAAGAACCTTACAGGA
BALB/c   GGGGAGACCCCAGGTGGGGATTCTGGAAAATGGGGGATGCTGTGAAGTAGGTATCTCACTTGGTTAGCATTGGAGCCTGTTCCACAAGAACCTTACAGGA

C57BL/6  CTCCCTTCCCCGGCCAGAACCTGAGTCTCTGGATCTGGGCCCCTTAGTCCTAGGGCCAGCCTTCCTCTTTCGGAAACCCCCAGAAATGGTCCATTTGGCG
BALB/c   CTCCCTTCCCCGGCCAGAACCTGAGTCTCTGGATCTGGGCCGCTTAGTCCTAGGGCCAGCCTTCCTCTTTCGGAAACCCCCAGAAATGGTCCATTTGGCG

C57BL/6  TGAGATAGGCTAGAATGGGGATTTTCCTGGAACATTTGAGGAGCTGGGTAGGGCTCAGGTGTTGTGTGCACTCCCGGCCATTCTCGGCTGTATCAGGGAT
BALB/c   TGAGATAGGCTAGAATGGGGATTTTCCTGGAACATTCGAGGAGCTGGGTAGGGCTCAGGTGTTGTGTGCACTCCCGGCCATTCTCGGCTGTATCAGGGAT

C57BL/6  ACTGTAAGGAAGTCCAGAAGCTTCATGTCGGTGGTCCCCGTGTCAGCCAAGTAGCCTAGAGGGGATGTGTCAAGAAACCTAAGCTGTTGGTTGATTCCTG
BALB/c   ACTGTAAGGAAGTCCAGAAGCTTCATGTCGGTGGTCCCCGTGTCAGCCAAGTAGCCTAGAGGGGATGTGTCAAGAAACCTAAGCTGTTGGTTGATTCCTG

C57BL/6  TGTTGGATGAACGAATCCTTCCCAGAATGGGTGGCTTGGGTCCCTTTTCCCAGCTTGTGGTTAGGCCAGGCAAGAAGACTCCAGCCCAGTAACTTGTTCT
BALB/c   TGTTGGATGAACGAATCCTTCCCAGAATGGGTGGCTTGGGTCCCTTTTCCCAGCTTGTGGTTAGGCCAGGCAAGAAGACTCCAGCCCAGTAACTTGTTCT

C57BL/6  TTCCCTGTGTCTGTGTCCCCTCACCCTGGCTAGTGGCTTGGGAGGCGGAGTCTGCAGTTCCTCAGATATGTCCTGAGGCCTGGTTTTATTGCCATTCCAG
BALB/c   TTCCCTGTGTCTGTGTCCCCTCACCCTGGCTAGTGGCTTGGGAGGCGGAGTCTGCAGTTCCTCAGATATGTCCTGAGGCCTGGTTTTATTGCCATTCCAG

C57BL/6  CCCTTTGCTCTCTGAGGGTAGGAAGAAGGGAGTGGTAGGTAAGGGTAGGGCATTTGCCCCTTGGTCTTGGGGCTTGGTGAGAG----CAGATAGGGGTTG
BALB/c   CCCTTTGCTCTCTGAGGGTAGGAAGAAGGGAGTGGTAGGTTAGGGTAGGGCATTTGCCCCTTGGTCTTGGGGCTTGGTGAGAGAGAGCAGATAGGG-TTG

C57BL/6  GGGGTGGGGCTGCTCCTCAT
BALB/c   GGGGTGGGGCTGCTCCTCAT
```

## AG

### *Zfp568*

chr7:29,691,813-29,692,926

```
C57BL/6  ---------------------------------------GCTATTCTAAGA-CAGAGGTCTTATCTCAGCCTTTAGGGAGCTCGCCCAGGCCTGGTGGTTTT
BALB/c   ---------------------------------------GCTATTCTAAGAACGGAGGTCTTATCTCAGCCTTTAGGGAGCTCGCCCAGGCCTGGTGGTTTT

C57BL/6  ACTGCAGCTTGAAAGAAAGTAAGGCTTGCGGGCATGCAGTAAGATGGGGGCGGCCGTCCAATACCTGGGATCACTCTGGGCTCGGCACAAAGGAAAGCCC
BALB/c   ACTGCAGCTTGAAAGAAAGTAAGGCTTGCGGGCATGCAGTAAGATGGGGGCGGCTGTCCAATACCTGGGATCACTCTGGGCTCGGCACAAAGGAAAGCCC

C57BL/6  GGCTCTTTGTGGTTTGTAGTGCATTCTTTGCCTCTTTTGGGTAACTGACCTGGCAATGGGTAAAACTGGGAGTAGACATAAAAAAATTGTGACCCTCCAG
BALB/c   GGCTCTTTGTGGTTTGTAGTGCATTCTTTGCCTCTTTTGGGTAACTGACCTGGCAATGGGTAAAACTGGGAGTAGACATAAAAAAATTGTGACCCTCCAG

C57BL/6  GTCTTTCTTGGCCAGGCCAGAGATCATTACCAGCCTTGTCCCCATTATCCATCCTGCGTAGAGAATGGAGATTAGGAGAGAACTGGCACCCTGACCAATT
BALB/c   GTCTTTCTTGGCCAGGCCAGAGATCATTACCAGCCTTGTCCCCATTATCCATCCTGCGTAGAGAATGGAGATTAGGAGAGAACTGGCACCCTGACCAATT

C57BL/6  TGCGAATCTAAACCTCACCTCAGAGATTGTTCCCAGGAGCAGAGTTCTGAAGTCACAACCACTCGGGTGACAGGAGCCCCATGGTTACATGCAAGATGCC
BALB/c   TGCGAATCTAAACCTCACCTCAGAGATTGTTCCCAGGAGCAGAGTTCTGAAGTCACAACCACTCGGGTGACAGGAGCCCCATGGTTACATGCAAGATGCC

C57BL/6  ACATAGAGGTAGGCACAGTCTT----------------------------------------------------------------------------------
BALB/c   ACATAGAGGTAGGCACAGTCTTTTTTTTGTTTGTTTGTTTGTTTTTTGTTTTTTTTTTGTTTGTTTTTGTTTTTTTTGAGACAGGGTTTCTCTGTATAGCCC

C57BL/6  ----------------------------------------------------------------------------------------------------
BALB/c   TGGCTGTCCTGGAACTCACTCTGTAGACCAGGCTGGCCTCGAACTCAGAAATCCGCCTGCCTCTGCCTCCCGAGTGCTGAGATTAAAGGCGTGCGCCACC

C57BL/6  --------------------------ACACTTAAGCCACAACTCCACAACGACCAGACATTTCATGTCCTCATAGAAAAGCTATCAGAATGTATTACAAAC
BALB/c   ACGCCCGGCTGGTAGGCACAGTCTTACACTTAAGCCACAACTCCACAACGACCAGACATTTCATGTCCTCATAGAAAAGCTATCAGAATGTATTACAAAC

C57BL/6  AACCACAAAAACTATTCCAGGGCACTGCGTCCGCCCAGTCGCAAGGCTCCTGCCGAGAGGCTGGACAGGCCACTAGCCTTTCTGGGAAATGTAGTCCACA
BALB/c   AACCACAAAAACTATTCCAGGGCACTGCGTCCGCCCAGTCGCAAGGCTCCTGCCGAGAGGCTGGACAGGCCACTAGCCTTTCTGGGAAATGTAGTCCACA

C57BL/6  AAGGGACAGCCACCTAGCGCCCCCAGCGCCTCCTCTGGTTGTTCCGGAAAAGACCGTCCCCGCAGCTGCTCACGTCCAGCGCTGCTTCGCCGGCGGCGTC
BALB/c   AAGGGACAGCCACCTAGCGCCCCCAGCGCCTCCTCTGGTTGTTCCGGAAAAGACCGTCCCCACAGCTGCTCACGTCCAGCGCTGCTTCGCCGCCGGCGTC

C57BL/6  ACCGTCTCGCCGCGCTAATGGGCTCAGCCACCTCGGCGCCTTCGGCCATGACTCCTGTCCCTAGCACAGGAAACGCCAGCCGAGCTCTGACCGGAAGTTG
BALB/c   ACCGTCTCGCGGCGCTAATGGGCTCAGCCACCTCCGCGCCTTCGGCCATGACTCCTGTCCCTAGTACAGGAAACGCCAGCGGAGCTCTGACCGGAAGTTG

C57BL/6  GGTCGGGCCGCGGGTCCTTCTGGGAGCACCTGTGCTCTGGGACATAACGGGCGCGCGCTGCCGCCCGAGTGGAGGGGCTGGCCGTGCGTGGCCCGCATGC
BALB/c   GGTCGGGCCGCGGGTCCTTCTGGGAGCACCTGTGCTCTGGGACATAACGGGCGCGCGCTGCTGCCCGAGTGGAGGGGCTGGCCGTGCGTGGCCCGCATGC

C57BL/6  CGCGTCTCTGAGGTTGCGTCTCGGGAGGTTTTGGCTGGTGGCGCTGCGGGGACCTCTCCACGGGAGAGGCACGGGTGGTCCGGTTTGCGCCGGCCGTGGG
BALB/c   CGCGTCTCTGAGGTTGAGTCTCGGGAGGTTTTGGCTGGTGACGCTGCGGGGACCTCTCCACGGGAGAGGCACGGGTGGTCCGGTTTGCGCCGGCCGTGGG

C57BL/6  CTTCGGGCTGGGCGCGGGCTCTCGGGGGGCTCGGAGTATCCCGGGTCACAGCTTT
BALB/c   CTTCGGGCTGGGCGCGGGCTCTCGGGGGGCTCGGAGTATCCCGG-TCACAGTTTT
```

## AH

### *Tiam2*

chr17:3,352,121-3,353,556

```
C57BL/6  ----------------------------------------------TATTGTTCTCAATGTGGGTTCATATGTGTGAATGTATACATGTCAACAAGATATGTCT
BALB/c   ----------------------------------------------TATTGTTCTCAATGTGGGTTCATATGTGTGAATGTATACATGTCAACAAGATATGTCT

C57BL/6  TAAAATCCAAATAGGGAATTACAAGTTTATCTAACTATAGTAGAATCTGGTCCTTAATAACTCACCAATGGCATATTTAGCTCAGTTCACTTTTTAAAAA
BALB/c   TAAAATCCAAATAGGGAATTACAAGTTTATCTAACTATAGTAGAATCTGGTCCTTAATAACTCACCAATGGCATATTTAGCTCAGTTCACTTTTTAAAAA

C57BL/6  ATTATTGGATATGTGCACTTATGCCACAGTGTGTGTATGGTCAGTGACAGATATCGGTGTCAGTTTGTCCTTCTACCTTGCTGAGGTAGGTGCTGGGTAC
BALB/c   GTTATTGGATATGTGCACTTATGCCACAGTGTGTGTATGGTCAGTGACAGATATCGGTGTCAGTTTGTCCTTCTACCTTGCTGAGGTAGGTGCTGGGTAC

C57BL/6  TGCAGTCTAGCCAGCTTGAGAGCTTTCCGGAGATTTCTTCCTGTCTCAATGTAGAAGTACCAGGATTACAGATGTACCCCACTGTATCTGGCCATTTTAT
BALB/c   TGCAGTCTAGCCAGCTTGAGAGCTTTCCGGAGATTTCTTCCTGTCTCAATGTAGAAGTACCAGGATTACAGATGTACCCCACTGTATCTGGCCATTTTAT

C57BL/6  GTGGCTTCCAGGGTTCATGATGTTTGTACCAGCAAGAAAGTTAATTCACTGAGCCATCTGAAGGACTTCTGTCACATTTTAAAATTAAGAATGTTACCAT
BALB/c   GTGGCTTCCAGGGTTCATGATGTTTGTACCAGCAAGAAAGTTAATTCACTGAGCCATCTGAAGGACTTCTGTCACATTTTAAAAATAAGAATGTTACCAT

C57BL/6  TATAAAGCTCTTGCCACATGTTCAGGATGGTTCAGGATGGGTTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTATTTATATG
BALB/c   TATAAAGCTCTTGCCACATGTTCAGGATGGTTCAGGATGGGTTGAG--------TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTATTTATATG

C57BL/6  TCTGAAAATGTAGGACTTTTGGTCATTCTGTGGAATATAAACACATAAATTTAAGCATGACAACAGAAATGTAATTTGATTCTATTCATCATAAAACAAC
BALB/c   TCTGAAAATGTAAGACTTTTGGTCATTCTGTGGAATATACACACATAAATTTAAGCATGACAACAGAAATGTAATTTGATTCTATTCATCATAAAACAAC

C57BL/6  AACAACAA---AAATCTAGTATAGAAATAACCTAGGGGCTGGAGAGATGGCTCAGAGGTTAAGAGCACTGGCTGCTTCTCCAGAGGTCCTGAGTTCAATT
BALB/c   AACAACAACAAAAATCTAGTATAGAAATAACCTAGGGGCTGGAGAGATGGCTCAGAGGTTAAGAGCACTGGCTGCTTCTCCAGAGGTCCTGAGTTCAATT

C57BL/6  CCCAGCAACCACATGGTGGCTCACAACCATCGGTAATGGGATCTGGTTTCTTCTTCTGGTATGCAGGTATACATACAGGCAGAACACTGTATCCATAGTA
BALB/c   CCCAGCAACCACATGGTGGCTCACAACCATCGGTAATGGGATCTGGTTTCTTCTTCTGGTATGCAGGTATACATACAGGCAGAACACTGTATCCATAGTA

C57BL/6  AATAAAAAAAT----------------------------------------------------------------------------------------
BALB/c   AATAAAAAAAAATCTTTTTTTGGGCTGGAGAGATGGCTCAGTGGTTAAGTAATGGCTGCTCAGCCAGGCATGGTGGAACACGCCTTTATCCCAGCACTCG

C57BL/6  ----------------------------------------------------------------------------------------------------
BALB/c   GGAGGCAGAACAGGCAGATTTCTGAGTTCGAGGCCAGCCTGGTCTACAAAGTGAGTTCCAGGACAGCCAGGGCTACATAGAGAAACCCTGTCTCAAAAAA

C57BL/6  ----------------------------------------------------CTTTTTAAAAAATGACTTGAAAAAAATCTTGTATATTTTTAACATCTTT
BALB/c   CAAACAAACAAACAAACCCAAAAAACAAAAAAACAAAAACAAAAAAAAAAACTTTTTAAAAAATGACTTGAAAAAAATCTTGTATATTTTTAACATCTTT

C57BL/6  CATCAATGGTTACAGTGAGACTTCAAAATACTCTGCCTCTGTTGTATAGCCTTAATTTACAGTAATGTAATACATACTCTGGATAAAATGGCTCCGCATT
BALB/c   CATCAATGGTTACAGTGAGACTTCAAAATACTCTGCCTCTGTTGTATAGCCTTAATTTACAGTAATGTAATACATACTCTGGATAAAATGGCTCCGCATT

C57BL/6  TAGTTTCCAAGGACTTTGAGTGTAAGACTCTGTTGTGCTAACTTAAGGGGGGTGGGAAGCCCGATTCCAACAAGGAACACAGAATGAGGGTATACTCGAT
BALB/c   TAGTTTCCAAGGACTTTGAGTGTAAGACTCTGTTGTGCTAACTTAAGGGGGGTGGGAAGCCCGATTCCAACAAGGAACACAGAATGAGGGTATACTCGAT

C57BL/6  ATTCTGCAAACCATGGCAGCATTTCCTGGGGCTGTTCTAAAGGCAGTTCTGACCATAACATTCCATCTGTGGTTTTCACTGCTTCCCCAAACACAAGGAT
BALB/c   ATTCTGCAAACCATGGCAGCATTTCCTGGGGCTGTTCTAAAGGCAGTTCTGACCATAACATTCCATCTGTGGTTTTCACTGCTTCCCCAAACACAAGGAT

C57BL/6  GATAGGTGTGTATTCTCAGTGCTACAGGGCTTTATAAATGACGCCAGGGCAGCCCTGGTGAGTAGGGCTTTATTCTGTAACCCTTGGAGCCAGACTCCTA
BALB/c   GATAGGTGTGTATTCTCAGTGCTACAGGGCTTTATAAATGACGCCAGGGCAGCCCTGGTGAGTAGGGCTTTATTCTGTAACCCTTGGAGCCAGACTCCTA

C57BL/6  TGGAACTTCGGTGGTCATGGTGGATTGCTGGCCCAACAAGAGCCTAGTCATGGGCAGCCCCGTGTCTCTGAGGCGAGCATGGTGGCAGGCTGTCCCGGGT
BALB/c   TGGAACTTCGGTGGTCATGGTGGATTGCTGGCCCAACAAGAGCCTAGTCATGGGCAGCCCCGTGTCTCTGAGGCGAGCATGGTGGCAGGCTGTCCCGGGT

C57BL/6  AATCACTCCAGGCAGCCCATG
BALB/c   AACCACTCCAGGCAGCCCATG
```

# 13 Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe des Literaturzitats gekennzeichnet.

Weitere Personen waren an der inhaltlich-materiellen Herstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe eines Promotionsberaters oder anderer Personen in Anspruch genommen. Niemand hat von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

---------- ----------

(Schilling Elmar)