

**Epigenetic mechanisms of fine-tuning *FKBP5* gene
expression**

**Dissertation zur Erlangung des Doktorgrades
an der Fakultät für Biologie
der Ludwig-Maximilians-Universität München**

**Vorgelegt von
Tobias Wiechmann
München, 2019**

**Max-Planck-Institut für Psychiatrie und
Ludwig-Maximilians-Universität München**

Tag der Einreichung: 18. Juni 2019

Tag der mündlichen Prüfung: 22. November 2019

1. Gutachter: PD Dr. Mathias V. Schmidt
2. Gutachter: Prof. Dr. Heinrich Leonhardt
3. Gutachter: Prof. Dr. Niels Dingemanse
4. Gutachter: Prof. Dr. Thomas Nägele
5. Gutachter: Prof. Dr. Christof Osman
6. Gutachter: PD Dr. Carsten T. Wotjak

Table of content

1 Summary	3
1 Zusammenfassung	5
2 Abbreviations	7
3 Introduction	11
3.1 GxE interaction and the relevance of <i>FKBP5</i> in stress-related diseases	11
3.1.1 GxE interaction in stress-related psychiatric disorders	12
3.1.2 The role of <i>FKBP5</i> in the stress system	13
3.1.3 <i>FKBP5</i> genetic variations interact with childhood abuse during early life	15
3.2 DNA methylation	18
3.2.1 The relevance of DNA methylation in disease	20
3.2.2 Techniques to measure DNA methylation levels	21
3.3 The 3D genome.....	23
3.3.1 The „C“-Techniques	25
3.3.2 The relevance of chromatin structure in disease	27
3.3.3 Steroid Hormones and the 3D chromatin structure of <i>FKBP5</i>	29
4 Aim of the thesis	33
5 Results	35
5.1 Paper I - HAM-TBS: high-accuracy methylation measurements via targeted bisulfite sequencing	35
5.2 Paper II - Dynamic Changes in DNA Methylation Occur during the First Year of Life in Preterm Infants.....	46
5.3 Paper III – Identification of dynamic glucocorticoid- induced methylation changes at the <i>FKBP5</i> locus.....	56
5.4 Manuscript I - An intronic structural variant in <i>FKBP5</i> affects GRE enhancer activity through the modulation of topological associated domain interaction frequencies and regulates the responsiveness to glucocorticoid exposure.....	71
6 Discussion	115
6.1 Accurate and robust measurement of DNAm levels within the <i>FKBP5</i> locus in patient cohorts.....	116

Table of content

6.2 Embedding of environmental signals into <i>FKBP5</i> DNAm levels	119
6.3 Molecular mechanisms of GxE interactions in the <i>FKBP5</i> locus	124
6.4 Future perspectives towards translation	129
7 References	133
<hr/>	
8 Declaration of Contributions	153
<hr/>	
9 Acknowledgements	155
<hr/>	
10 Statutory declaration and statement	157
<hr/>	
11 Curriculum Vitae	159
<hr/>	

1 Summary

The development of stress-related psychiatric disorders involves the integration of genetic and environmental factors. How these factors are integrated on a molecular level to determine risk or resilience is far from being understood. The long-term integration of both factors is thought to be mediated by epigenetic adaptations. One molecular mechanism for a statistical gene x environment interaction has been described for the *FKBP5* gene (Klengel et al., 2013). This thesis aims to further elaborate our understanding of this molecular mechanism.

First, a new method - High accurate methylation measurements via targeted bisulfite sequencing (HAM-TBS) - was established to accurately and robustly measure DNA methylation (DNAm) levels of key regulatory sites within the *FKBP5* locus in larger cohorts of patients (chapter 5.1 (Paper I), discussed in chapter 6.1). The new possibilities of this method have been used to better understand dynamic methylation changes in *FKBP5* due to early life stress by monitoring dynamic glucocorticoid (GC)-induced methylation changes in healthy individuals upon a GC challenge (chapter 5.2 (Paper II) & chapter 5.3 (Paper III), discussed in chapter 6.2). The data obtained indicate that the exposure to high levels of GCs can be embedded onto DNAm levels. In addition, the underlying dynamics of these GC-induced methylation changes at distinct CpGs were observed to be genotype-dependent (rs1360780). The processes involved in this embedding of environmental signals are not yet fully understood. However there are several lines of evidence that these are active enzymatically driven leading to dynamic epigenetic changes centered around GR binding sites in the *FKBP5* locus.

Furthermore, lymphoblastoid cell lines carrying different alleles for a 3.3 kb large insertion / deletion (INDEL, esv3608688) in intron 1 of *FKBP5* have been used to identify if and how this INDEL in the disease-associated haplotype contributes to shape GC-induced *FKBP5* gene expression changes (chapter 5.4 (Manuscript II), discussed in chapter 6.3). We identified that the stabilization of architectural and enhancer-promoter loops is a common feature of the factors (T-allele, deletion and GR activation by GCs) leading to increased *FKBP5* mRNA expression. This gene expression response is most likely the result of an increased activity of enhancers due to the stabilization of chromatin interactions. However, further experiments manipulating the DNA sequence and DNA methylation levels of *FKBP5* key

Summary

regulatory sites will be necessary to obtain causality. Moreover, further analysis in patients will help to clarify the role of INDEL alleles in the development of stress-related psychiatry phenotypes. Finally, thoughts on future perspectives and potential translations into the clinical routine of the mechanistic insights of *FKBP5* gene expression regulation are discussed (chapter 6.4).

1 Zusammenfassung

Die Entstehung von stressbedingten psychiatrischen Erkrankungen involviert die Integration von genetischen und Umweltfaktoren. Wie diese Faktoren auf der molekularen Ebene integriert und somit Risiko und Resilienz vermitteln, ist weit davon entfernt verstanden zu werden. Man vermutet, dass die dauerhafte Integration beider Faktoren durch epigenetische Anpassungen vermittelt wird. Ein molekularer Mechanismus für eine statistische Gen-Umweltinteraktion wurde für das *FKBP5* Gen beschrieben (Klengel et al., 2013). Die vorliegende Doktorarbeit zielt darauf ab, unser Verständnis des molekularen Mechanismus, der genetische und Umweltfaktoren integriert, besser zu verstehen.

Als erstes wurde eine neue Methode – Hochgradig akkurate Methylierungsmessungen anhand von zielgerichteter Bisulfit-Sequenzierung (HAM-TBS) – etabliert, um die DNA Methylierungslevel von wichtigen regulatorischen Regionen im *FKBP5* Gen akkurat und robust in größeren Patientenkohorten messen zu können (Kapitel 5.1 (Publikation I), diskutiert in Kapitel 6.1). Die neuen Möglichkeiten dieser Methode wurden dazu genutzt dynamische Methylierungsveränderungen in *FKBP5*, die aufgrund von frühkindlichem Stress auftreten können, weiter zu untersuchen (Kapitel 5.2 (Publikation II), diskutiert in Kapitel 6.2). Hierbei wurde in gesunden Probanden, die mit Dexamethason behandelt wurden, die zugrunde liegenden Dynamik von Glukokortikoid-induzierten DNA Methylierungsveränderungen in *FKBP5* betrachtet (Kapitel 5.3 (Manuskript I), diskutiert in Kapitel 6.2). Die ermittelten Daten deuten darauf hin, dass die Exposition mit hohen Glukokortikoidspiegeln sich auf das DNA Methylierungslevel niederlegen kann. Des Weiteren kann sich die Dynamik der Glukokortikoid-induzierten Methylierungsveränderungen an bestimmten CpGs aufgrund von unterschiedlichen Genotypen ändern (rs1360780). Die beteiligten Einbettungsprozesse der Umweltsignale sind noch nicht in Gänze verstanden, aber einige Evidenzen deuten darauf hin, dass diese aktiv enzymatisch vermittelt sind und zu dynamischen epigenetischen Veränderungen führen, die um Glukokortikoidrezeptorbindestellen zentriert im *FKBP5* Locus auftreten.

Zusätzlich wurden Lymphoblastoide Zelllinien mit unterschiedlichen Allelen für eine 3.3 kb große insertion / deletion (INDEL, esv3608688) in Intron 1 des *FKBP5* Gens verwendet, um herauszufinden inwiefern die INDEL Allele im krankheitsassoziiertem

Haplotyp daran beteiligt ist, die glukokorticoinduzierten *FKBP5* Genexpressionsveränderungen zu gestalten (Kapitel 5.4 (Manuskript II), diskutiert in Kapitel 6.3). Wir konnten feststellen, dass die Stabilisierung von architektonischen und Enhancer-Promotor Loops ein gemeinsames Merkmal von Faktoren (T-Allele, Deletion und Glukokortikoidrezeptoraktivierung durch Glukokortikoide) ist, die zu einer erhöhten *FKBP5* Expression führen. Diese Genexpressionsantwort ist wahrscheinlich ein Ergebnis einer erhöhten Enhanceraktivität aufgrund der Stabilisierung von Chromatininteraktionen. Nichtsdestotrotz sind weitere Experimente notwendig, welche die DNA Sequenz und DNA Methylierungslevel von regulatorisch wichtigen Regionen des *FKBP5* Locus manipulieren, um Kausalität zu erlangen. Außerdem werden weitere Analysen von Patientendaten dazu beitragen, die Rolle der INDEL-Allele innerhalb der Entwicklung von stressbedingten psychiatrischen Phänotypen aufzuklären.

Abschließend, werden Zukunftsperspektiven und potenzielle Möglichkeiten der Translation in den klinischen Alltag diskutiert (Kapitel 6.4).

2 Abbreviations

3C = Chromatin Conformation Capture

4C = Circularized Chromosome Conformation Capture

5C = Chromosome Conformation Capture Carbon Copy

5caC = 5-Carboxylcytosine

5fC = 5-Formylcytosine

5hmC = 5-Hydroxymethylation

5mC = 5-Methylcytosine

ACTH = Adrenocorticotropin Hormone

APOBEC = Apolipoprotein B mRNA Editing Enzyme

AS = Angelman Syndrome

AVP = Vasopressin

BER = Base Excision Repair

BWS = Beckwith-Wiedemann Syndrome

CpGs = Cytosine Guanine Dinucleotides

CRH = Corticotropin Releasing Hormone

CRHR1 = Corticotropin Releasing Hormone Receptor 1

CRISPR = Clustered Regularly Interspaced Short Palindromic Repeats

CT = Chromosome Territory

CTCF = CCCTC-Binding Factor

DEX = Dexamethasone

DNAm = DNA methylation

DNMTs = DNA Methyltransferases

ELS = Early-Life Stress

eRNA = Enhancer RNA

Abbreviations

EWAS = Epigenome-Wide Association Study

FDA = Food and Drug Administration

FISH = Fluorescence *In Situ* Hybridization

FKBP5 = FK506 Binding Protein 5

GC = Glucocorticoids

GR = Glucocorticoid Receptor

GRE = Glucocorticoid Responsive Element

GWAS = Genome-Wide Association Study

GxE = Gene-Environment Interaction

H3K27ac = K27-acetylated Histone H3

H3K36me3 = K36-trimethylated Histone H3

H3K4me = K4-methylated Histone H3

H3K4me3 = K4-trimethylated Histone H3

HAM-TBS = High Accurate Methylation Measurements via Targeted Bisulfite Sequencing

HPA axis = Hypothalamic-Pituitary-Adrenal axis

HPCs = Hippocampal Progenitor Cells

INDEL = Insertion / Deletion

LAD = Lamina-Associated Domain

LCLs = Lymphoblastoid Cell Lines

LD = Linkage Disequilibrium

MDD = Major Depressive Disorder

MR = Mineralocorticoid Receptor

mRNA = Messenger RNA

PCR = Polymerase Chain Reaction

PTSD = Post Traumatic Stress Disorder

Abbreviations

PVN = Paraventricular Nucleus

PWS = Prader-Willi Syndrome

qPCR = Quantitative PCR

RNAPII = RNA Polymerase II

SINE = Short Interspersed Nuclear Elements

SNP = Single-Nucleotide Polymorphism

TAD = Topological Associating Domain

TBS = Targeted Bisulfite Sequencing

TET = Ten-Eleven Translocation Methylcytosine Dioxygenases

TFs = Transcription Factors

tRNA = Transfer RNA

TSS = Transcription Start Site

3 Introduction

3.1 GxE interaction and the relevance of *FKBP5* in stress-related diseases

The vast success of Darwin's evolutionary model and Mendelian genetics pushed the idea that traits can be acquired during life and transmitted to future generations into the background. "Epigenetics, or the study of heritable (mitotic and/or meiotic) changes in gene activity that are not brought about by changes in the DNA sequence..." (Van Soom et al., 2014) now refreshes the idea proposed by Lamarck. One method for identifying hypothesized gene-environment (GxE) interactions is epidemiological research. Epidemiology brought evidences that a GxE interactions are potentially involved in the development of diseases, but lack to explain the molecular mechanisms of GxE interactions (Caspi and Moffitt, 2006). Furthermore, twin studies (Bell and Spector, 2011; Bell et al., 2012; Petronis, 2006) and inbred mouse strains (Weaver et al., 2004) were used to investigate the influence of non-genetic factors onto phenotypes due to the advantage that monozygotic twins and inbred mouse strains have identical genomes. In spite of the fact that some GxE interactions have been reported for mood and anxiety disorders (Binder et al., 2008; Caspi et al., 2010; Heim et al., 2009; Xie et al., 2010), the molecular link how genetic and environmental factors can interact together within specific mechanisms to alter cell function and more remote behavioral phenotypes remains largely unknown (Figure 1). Since epigenetic mechanisms are potentially playing a role in the etiology of various human diseases (Handel et al., 2010) such as depression (McGowan et al., 2009), major psychosis (Mill et al., 2008), autism (Schanen, 2006), cancer (Esteller, 2007; Varambally et al., 2008), asthma (Adcock et al., 2005) and obesity (Gerken et al., 2007), it has been proposed that epigenetic modifications and mechanisms such as DNA methylation, histone modifications, non-coding RNA and chromatin conformation changes (Levenson and Sweatt, 2005; Maddox et al., 2013; Mill and Petronis, 2007; Miller et al., 2008; Mitchell et al., 2014; Tsankova et al., 2007) are involved in mediating GxE interactions by changing the expression of genes implicated in stress-related psychiatric diseases (Jaenisch and Bird, 2003; Yehuda and LeDoux, 2007). Epigenetic changes due to GxE interaction effects have been suggested to be installed during sensitive periods (e.g. development) and to

remain stable over time (Klengel et al., 2013; Klengel et al., 2014). It is discussed that molecular and behavioral adaptations due to epigenetic changes can be inherited across generations via gametes and maternal behavior (Bohacek et al., 2013; Daxinger and Whitelaw, 2012; Franklin et al., 2010; Gapp et al., 2014; Yehuda et al., 2016).

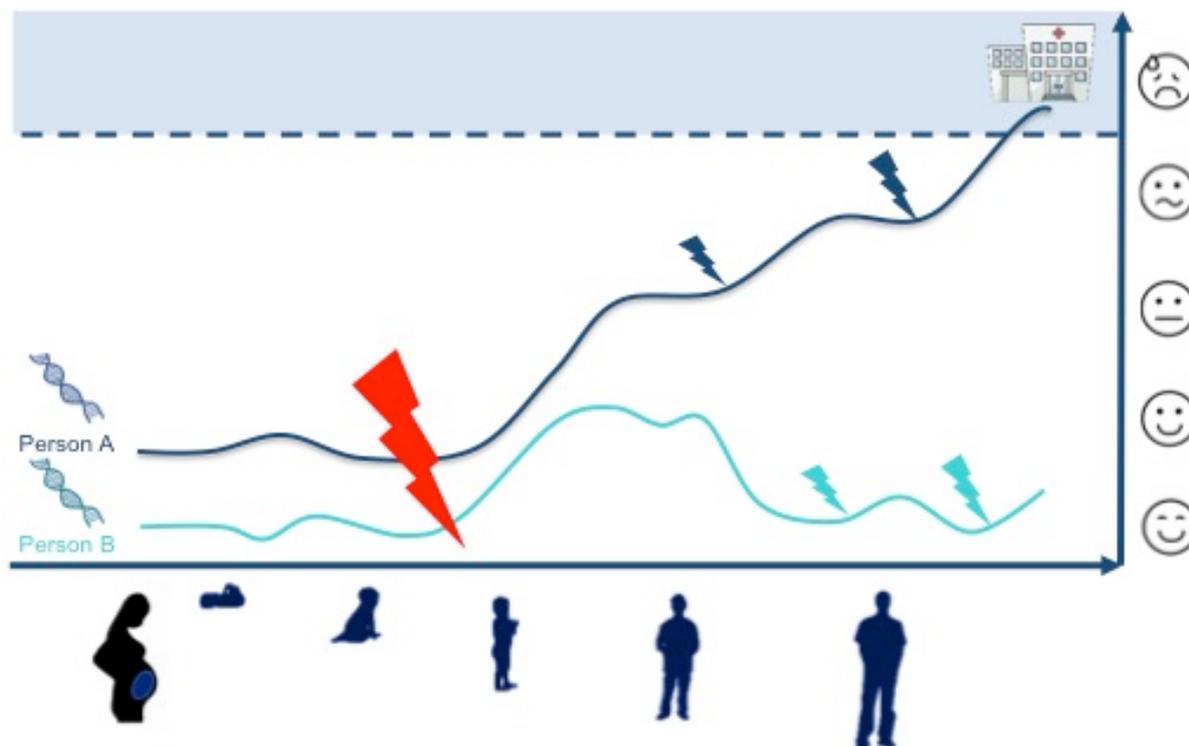


Figure 1: The development of stress-related psychiatric disorders involves the integration of genetic and environmental factors

The importance of integrating genes and the environment when investigating stress-related psychiatric disorders has been shown in epidemiological & genetic studies. Both individuals (Person A and B) start with a different genetic predisposition to develop a stress disorder. After severe trauma during childhood person B does not develop clinical relevant symptoms when struggling under daily life hassles. Whereas person A does develop symptoms. The question is how genetic and environmental factors are integrated to determine risk and resilience and what are the molecular mechanisms, which prime the two persons onto different trajectories after early-life stress (ELS). It is believed that GCs are involved in this priming event and might mediate the impact of environmental factors under the background of genetic variants onto epigenetic layers. Source: Tobias Wiechmann, 2019

3.1.1 GxE interaction in stress-related psychiatric disorders

It has been shown that the risk of developing a stress-related psychiatric disorder like Post Traumatic Stress Disorder (PTSD) or depression not only depends on genetic predisposition of having risk or protective genetic variants, but as well environmental factors like exposure to stressful or traumatic life events (Caspi and Moffitt, 2006; Kendler et al., 1999; Kendler et al., 1995; Klengel et al., 2013; Molnar et al., 2001).

Genes which have shown to be part of GxE are often involved in the stress hormone system such as the *glucocorticoid receptor (GR, NR3C1)*, *corticotrophin releasing*

hormone receptor 1 (CRHR1) and *FK506 binding protein 5 (FKBP5)*, which play a key role in the regulation of the hypothalamic-pituitary-adrenal (HPA) axis (Binder, 2009; Bradley et al., 2008; Klengel and Binder, 2015; Klengel et al., 2013; Polanczyk et al., 2009; Zannas and Binder, 2014). A dysregulated stress hormone system is a strong risk factor for developing stress-related psychiatric disorders and therefore regulating factors of this system display promising therapeutic targets for the treatment of stress-regulated diseases like PTSD or depression (Binder, 2009; Klengel et al., 2013). An increasing number of studies have linked interactions between *FKBP5* genotypes and stressors with diverse disease-related phenotypes (Zannas and Binder, 2014). Phenotypes examined to date include Major Depressive Disorder (MDD) or depressive symptoms (Appel et al., 2011; Dackis et al., 2012; Kohrt et al., 2015; VanZomeren-Dohm et al., 2015; Zimmermann et al., 2011), PTSD and related phenotypes (Binder et al., 2008; Boscarino et al., 2012; Klengel et al., 2013; Koenen et al., 2005; Xie et al., 2010), suicidality (Roy et al., 2010; Roy et al., 2012), aggression and violent behaviors (Bevilacqua et al., 2012), psychosis (Collip et al., 2013), cognitive performance (Hernaus et al., 2014), and general physical illness (Lessard and Holman, 2014). In total 31 independent studies comprising 39 cohorts including a total of 31 000 subjects have been carried out so far, with the majority of them showing that the alleles associated with higher *FKBP5* induction and prolonged cortisol responses are also the alleles associated with higher disease risk (Matosin et al., 2018).

3.1.2 The role of *FKBP5* in the stress system

The HPA axis is a neuroendocrine system that controls reactions to stress in mammals (Sapolsky et al., 2000). During this process, the brain and the body need to be adjusted to successfully cope with stressful situations (e.g. make energy resources available to fight or flight). Exposure to a stressor induces the release of corticotrophin releasing hormone (CRH) and vasopressin (AVP) from the paraventricular nucleus (PVN) in the hypothalamus, which stimulates the release of adrenocorticotropin hormone (ACTH) from the anterior pituitary, in turn inducing the release of glucocorticoids (GC) from the cortex of the adrenal glands into the blood system. GCs act through two mostly intracellular receptors, the mineralocorticoid receptor (MR) and GR. The high affinity MR is mainly implicated in the appraisal

processes and the acute onset of the stress response, whereas the lower affinity GR promotes adaptation to recovery from the stress (De Kloet et al., 1998). GR is activated by cortisol, binds to the DNA at glucocorticoid responsive elements (GRE) and acts as a transcription factor to initiate the transcription of several genes, which can be involved in the regulation of the HPA axis in many tissues (Russell et al., 2010) (Figure 2). In the brain (mainly the hypothalamus), the activation of GR by GCs induces a negative feedback loop repressing the HPA axis and results in the termination of the stress response. At the cellular level, the activity of GR is also regulated in the cytoplasm by a multi-protein complex including heat shock proteins and FK506 proteins such as FKBP51 and FKBP52 (Pratt and Toft, 1997). At baseline, in both humans and mice, *FKBP5* expression levels differ across tissues, being markedly elevated in metabolically active tissues, such as adipocytes, peripheral blood cells as well as the hippocampus (<http://biogps.org/>). In the rodent brain, *FKBP5* has the highest expression levels in the hippocampus, with much lower expression in other brain regions (Scharf et al., 2011). Substantial variability of *FKBP5* expression is similarly observed in the human brain, with high levels noted, for example, in the hippocampus but low levels in the hypothalamus (<http://human.brain-map.org/>). *FKBP5* gene expression is induced by GR via GREs located within the *FKBP5* locus. A function of FKBP51 is to inactivate GR in the cytoplasm, constituting in an ultra-short negative feedback loop, which regulates GR sensitivity. *FKBP5* induction can vary across individuals and has been proposed as a marker of GR sensitivity (Kelly et al., 2012; Menke et al., 2012; Vermeer et al., 2003). In humans, an eightfold increase in *FKBP5* messenger RNA (mRNA) has been observed in peripheral blood cells 3 h after oral administration of 1.5 mg of the GR agonist dexamethasone (DEX). *FKBP5* is the most robustly induced transcript (Menke et al., 2012). Similar robust induction by GR activation has been noted in omental and subcutaneous adipose tissues (Pereira et al., 2014). Following stimulation with dexamethasone or stress exposure, *FKBP5* expression is dramatically increased in a number of brain regions (Scharf et al., 2011), with the largest changes observed in the amygdala and the PVN, whereas the hippocampus shows a much less pronounced *FKBP5* induction. This likely reflects the high baseline levels of *FKBP5* in this brain region that may confer relative GR resistance. By modulating GR signaling, *FKBP5* has the potential to modulate the actions of

Introduction

glucocorticoids, which are hormones with pleiotropic effects that can affect essentially every tissue (Chrousos and Gold, 1992; Nicolaides et al., 2014)(Figure 2).

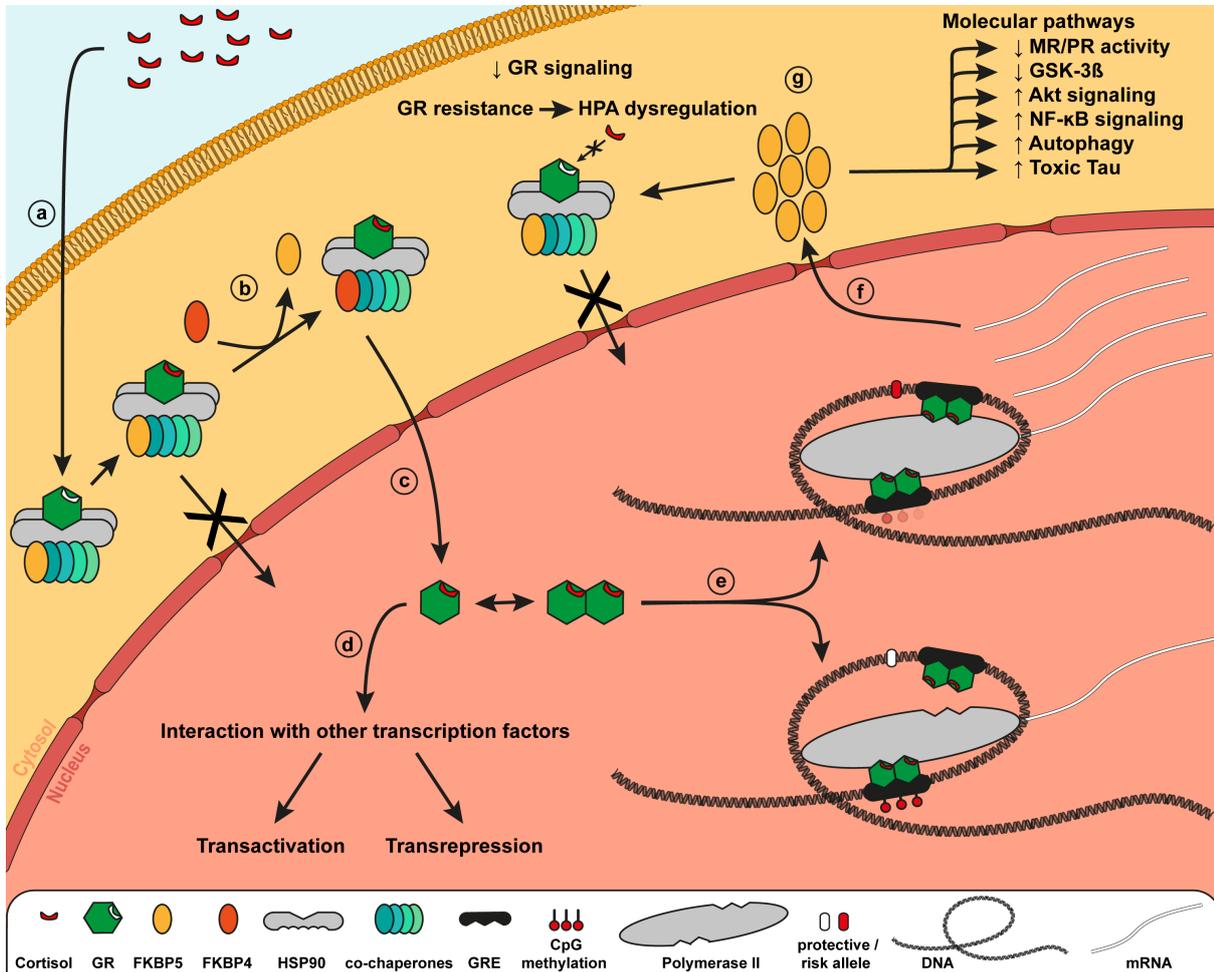


Figure 2: Schematic representation of the molecular events involved in glucocorticoid-mediated FKBP5 induction, the resulting intracellular negative feedback loop, and effects on other biological processes.

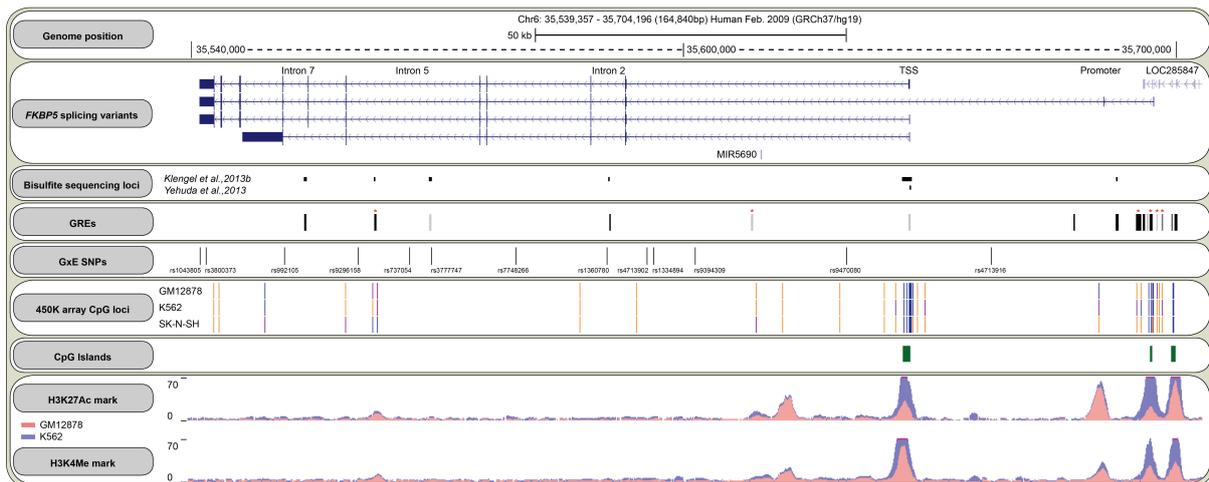
Glucocorticoids enter the cytoplasm (a) and activate the glucocorticoid receptor (GR) complex. FKBP51 binding to the complex reduces affinity of glucocorticoids to the GR and delays translocation of the GR to the nucleus. However, exchange of FKBP51 (FKBP5) for FKBP52 (FKBP4) (b) results in GR translocation to the nucleus (c). The GR can either interact as a monomer with other transcription factors (d) or form a homodimer that binds to DNA at glucocorticoid response elements. Overall, GR functions result in transactivation or transrepression of a large number of genes. The FKBP5 gene is highly responsive to GR, but responsiveness depends on FKBP5 polymorphisms and methylation status (e). The synthesized FKBP5 mRNA translocates to the cytoplasm (f) where it is translated into FKBP51 protein. FKBP51 then inhibits GR activity not only forming an ultra-short, intracellular negative feedback loop of GR signaling but also modulating several other biological pathways (g). Source: adopted from Zannas et al., 2015

3.1.3 FKBP5 genetic variations interact with childhood abuse during early life

The human locus of *FKBP5* (Nair et al., 1997) is located on the short arm of chromosome 6 (6p21.31), consists of 13 exons, 12 intron and spans around 155 kb

Introduction

(Figure 3). The gene transcription is steroid-regulated and mediated by binding of the GR to GREs, which are located in a region spanning over 100 kb and range from upstream of *FKBP5* promoter to introns 2, 5, and 7 of the gene (Paakinaho et al., 2010). With the use of tagging experiments and, more recently, next-generation sequencing, *FKBP5* variants have been described in detail (Ellsworth et al., 2013a; Ellsworth et al., 2013b; Pelleymounter et al., 2011).



Shown tracks are from the UCSC genome browser (<https://genome.ucsc.edu>). *FKBP5* splicing variants are derived from the RefSeq Genes Track. Bisulfite sequencing loci are shown as a custom track and are based on Klengel and Binder (2013) and Yehuda et al. (2013). As shown, these CpG sites are distinct from CpGs covered by the Illumina 450K array (shown for two blood cell lines and one neuroblastoma cell line based on ENCODE/HAIB; warm colors, high methylation; cold colors, low methylation levels). Glucocorticoid response elements (GREs) are displayed for A549 and ECC-1 cells and are derived from transcription factor (glucocorticoid receptor) ChIP-sequencing data of the ENCODE project. Conserved GREs are highlighted with red asterisk and are derived from the HMR Conserved Transcription Factor Binding Sites track (z-score cutoff: 1.64). GxE SNPs represent SNPs identified in gene-environment interaction studies and are mapped based on the Common SNPs(142) track. The H3K27Ac Mark and H3K4Me3 tracks show data from two different blood cells lines derived from the ENCODE project. Source: adopted from Zannas et al., 2015

The best-characterized polymorphisms comprise a haplotype that spans the whole gene, contains up to 18 single-nucleotide polymorphisms (SNPs) in strong linkage disequilibrium (LD) in Caucasians ($r^2 > 40.8$, distance 4500 kb, 1000 genomes next-generation sequencing project), and is commonly tagged by rs3800373, rs9296158, or rs1360780. This haplotype has been associated with heightened induction of *FKBP5* mRNA in response to GR activation (Binder et al., 2004) and interacts with childhood abuse to predict adult PTSD (Binder et al., 2008). This functionality was shown to be likely conferred by rs1360780, a SNP located in an enhancer region, 488 base pairs (bp) away from a functional GRE in intron 2 of the gene, with the risk T-allele facilitating GR-mediated induction of *FKBP5* mRNA (Klengel et al., 2013). However, other polymorphisms within or outside of this *FKBP5* haplotype may also

have functional effects, either at baseline or following transcriptional induction (Ellsworth et al., 2013a; Roy et al., 2010).

A mechanism for the interaction of childhood abuse and genetic variants of the *FKBP5* gene, focusing on rs1360780 variants was proposed by Klengel et al. (2013). In this GxE mechanism, functional polymorphisms within GRE in intron 2 can alter the 3D chromatin structure of the *FKBP5* locus and install childhood trauma dependent epigenetic modification. Here, demethylation of cytosine guanine dinucleotides (CpGs) within a GRE of intron 7, is leading to increased *FKBP5* expression and dysregulation of the stress hormone system resulting in increased risk of developing stress-related disorders like PTSD (Figure 2). By using chromatin conformation capture experiments (3C) it has been shown that the “risk allele” of rs1360780 enables the structural interaction of downstream enhancers located in GRE of intron 2 with the Transcription Start Site (TSS), but not in carriers of the “protective allele”. This structural interaction leads to a stronger *FKBP5* mRNA expression by GR, possibly due to the fact that the “risk allele” generates a TATA-box binding motive, which allows to bind more TATA-box binding proteins and subsequently facilitates the binding of enhancer regions to RNA polymerase II transcription complex. Even minor stressors would therefore result in a prolonged cortisol release in “risk” but not in “protective allele” carriers due to an impaired negative feedback loop. Upon childhood abuse, enhanced cortisol levels result in strong GR activation in risk allele carriers and can trigger demethylation of distal functional GREs in *FKBP5* intron 7, further promoting the *FKBP5* mRNA transcription. Demethylation of CpGs in GREs of *FKBP5* due to childhood abuse and activation of GR was reported by Klengel et al. (2013) in peripheral blood cells and hippocampal progenitor cells. However, whether or not there is a change in methylation levels of CpGs in the *FKBP5* locus around other GREs or functional DNA binding loci has not been investigated so far and has to be further tested. A multistep mechanism model how a targeted demethylation may be implemented in transcriptional activation by GR was proposed for GREs at the *Tat* gene by Grange et al. (2001). In this model, stable GR binding to GREs and transcriptional stimulation by exposure to GCs is achieved by chromatin remodeling, local demethylation of CpGs and at GREs and C/EBP recruitment. However, it remains unclear which proteins mediate CpG demethylation within this mechanism but the involvement of DNA repair proteins has been proposed (Kress et al., 2006). Transient demethylation

due to GR binding (Thomassin et al., 2001) can possibly change to a long term demethylation of the *Fkbp5* gene displaying a potential mark to excessive stress and glucocorticoid exposure (Lee et al., 2011). The higher transcriptional activation of *FKBP5* by GR in risk-allele carrier strengthens the intracellular negative feedback loop, which leads to GR resistance and a deregulated stress hormone system. Moreover, gene expression can be altered in GR responsive systems such as the immune system. Alteration of neuronal circuits as structural changes in the hippocampus, higher the risk for stress-related psychiatric diseases like PTSD. The same gene environment interaction, childhood abuse and *FKBP5* polymorphisms, have been associated with other disorders implying to share the same molecular mechanisms reviewed by Zannas and Binder (2014).

3.2 DNA methylation

DNA methylation (DNAm) is the covalent addition of a methyl group at the 5-carbon ring of cytosine, resulting in 5-methylcytosine (5mC). In the mammalian genome, this occurs predominantly in the context of CpG dinucleotides (Bird, 2002). Non-CpG methylation are rarely detected in somatic mammalian tissues but have been observed in the context of CpNpG or CpA in mouse embryonic stem cells and plants (Tost, 2010). 5mC accounts for approximately 1% of total DNA bases with a high degree (70-80%) of CpGs being methylated (Ehrlich et al., 1982). In mammalian genomes CpGs are generally underrepresented most likely due to the increased mutation rates at CpG sites (10 - 50 times higher than other transitional mutations) which led to their depletion and gain of TpG / CpG dinucleotides during evolution (Bird, 1980; Swartz et al., 1962; Tost, 2010). A possible explanation for the increased mutation rates is the spontaneously occurring deamination of methylated CpGs to TpGs and imperfect repair mechanisms (Coulondre et al., 1978; Shin et al., 2014; Waters and Swann, 1998). CpG-rich clusters of 1 - 4 kb length known as CpG Islands, which are often found in the promoter region of genes, escaped from this evolutionary pressure most likely to their low degree in methylation and more efficient repair mechanisms when unmethylated cytosines are spontaneously deaminated to Uracil (Bird et al., 1985; Shin et al., 2014).

A prerequisite to understand the functionality of DNAm is the distribution of DNAm patterns and that these patterns can vary in time (e.g. development) and space (cell-

type-specific methylation patterns) (Bird, 2002; Luo et al., 2018). In general, DNAm fulfills its function in the context of several other epigenetic marks influencing gene expression and serving further purposes as genomic imprinting, X-chromosome imprinting, definition of cellular identity, aging, maintenance of genomic stability and is thought to be able to serve as a memory for environmental inputs even across generations (Ambrosi et al., 2017; Tost, 2010). Although an inverse correlation of DNAm at promoters and transcriptional activity has been observed, its causal role in suppressing gene expression alone or influence in transcription factor binding to DNA is still unclear and highly context-dependent (Ambrosi et al., 2017; Shin et al., 2014). DNAm at promoters is thought to act repressive either direct by blocking binding of transcription factors (TFs) (Iguchi-Arigo and Schaffner, 1989) or indirect by recruiting proteins with methyl-binding domains, which block TFs from binding via the introduction of repressing domains (Nan et al., 1998). The recent developments enabling a precise manipulation of DNAm at specific target sites using the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) / Cas9 System (Liu et al., 2016; Luo et al., 2018) has the potential to shed further light onto the causal function of DNAm.

Key players in the establishment and maintenance of DNA methylation patterns are DNA methyltransferases (DNMTs). DNMTs can be divided into *de novo* methyltransferases (DNMT3 A, B) which prefer to methylate unmethylated DNA templates as substrates especially during the development and maintenance DNA, and methyltransferases (DNMT1) which prefer hemimethylated DNA templates as substrates during the replication process (Goll and Bestor, 2005). The erasure of DNA methylation patterns can be achieved during a passive demethylation process mainly observed in early development during embryogenesis and an active demethylation pathway, which either can be achieved via a direct DNA repair process (apolipoprotein B mRNA editing enzyme (APOBEC), not shown in mammalian cells) or the oxidation of 5mC to 5-hydroxymethylation (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) via Ten-eleven translocation methylcytosine dioxygenases (TET) 1-3 (Ambrosi et al., 2017; Bergman and Cedar, 2013; Shin et al., 2014; Wu and Zhang, 2014). These modifications are subsequently removed by glycosylases leaving an apyrimidinic acid, which is replaced by a base excision repair (BER). Whether these oxidized forms of 5mC are only intermediate steps of the demethylation pathway or have themselves distinct function as

epigenetic marks is current topic of investigation and debated (Shin et al., 2014; Wu and Zhang, 2014). It is also speculated that local demethylation can be introduced by specific transcription factor binding (Luo et al., 2018). In this model TFs lead to methylation depletion by acting as “pioneer TFs” to induce local epigenetic remodeling (Wapinski et al., 2013) or recruit cytosine demethylation machinery like TET enzymes.

3.2.1 The relevance of DNA methylation in disease

Diseases in which epigenetic modifications play a role often involve genes being epigenetically regulated or are part of the epigenetic machinery (reader / writer / eraser) to maintain and propagate epigenetic modifications during development and cell division (Tost, 2010). Therefore disease phenotypes can display aberrant methylation profiles at single genes but also be very complex and reveal global methylation changes (reviewed in Robertson (2005); Robertson and Wolffe (2000); Tost (2010)).

The Fragile X syndrome is an example of a “repeat-instability disease”, in which aberrant methylation patterns are involved in the gene expression regulation of a single gene. In case of the Fragile X syndrome, the loss of FMR1 gene expression during the development is probably caused by a triplet-repeat expansion (CGG) and gain of methylation at the *FMR1* promoter (Dor and Cedar, 2018; Robertson and Wolffe, 2000). Recently, it has been shown that the demethylation (introduced by the CRISPR/Cas9 system together with Tet1) of CGG repeats is sufficient to reactivate the FMR1 gene and rescue the FXS-related cellular phenotypes in *in vitro*-derived FXS neurons (Liu et al., 2018).

Imprinting disorders like the Beckwith-Wiedemann syndrome (BWS), Prader-Willi syndrome (PWS) or Angelman syndrome (AS) are examples in which altered DNA methylation patterns in important regulatory regions can lead to the loss of normal allele-specific gene expression patterns. Diseases caused by a loss of imprinting are often phenotypically more variable due to the fact that the regulatory region controls several imprinted genes. The molecular rationale for the allele-specific regulation of imprinted genes often include the access of distant enhancers to promoters, which can be regulated by the zinc-finger protein CCCTC-binding factor (CTCF) (Robertson, 2005). CTCF binding to its binding site is methylation sensitive and therefore can block the access of enhancers to promoters due to the methylation

status of the unmethylated allele and established allele-specific gene regulation from the maternal or paternal allele. For example, the molecular phenotype of the Rett syndrome involves the transcriptional deregulation of several hundreds of genes and is an example for defects of a methylation reader. This syndrome is caused by mutations in the methyl-binding protein MeCP2, which binds to methylated DNA and leads to repression of gene expression by recruiting transcriptional co-repressors (Klose and Bird, 2006).

Cancer is one of the best-documented diseases with an epigenetic component and shown to exhibit the most complex aberrant methylation profiles. In general, tumor cells reveal global hypomethylation but also local hypermethylation of CpG islands at promoters (Baylin and Jones, 2011; Jones and Baylin, 2007). Recent studies better defined the monitored demethylation in cancer cells and linked them to regions associated to the nuclear envelop lamina (Berman et al., 2011). Although the underlying mechanisms of this demethylation are unknown, it is thought that the maintenance machinery is too slow (abnormally high recruitment of Dnmt1 to *de novo* targeted CpG islands) and the replication rate is very high in cancer cells and therefore a passive demethylation might be observed (Bergman and Cedar, 2013; O'Hagan et al., 2011). The pathophysiological role of the DNA methylation changes in cancer are yet not clear, but it has been shown that *de novo* methylation may be installed in cancer cells to repress tumor-suppressor genes (Bender et al., 1998; Laird et al., 1995; Tsai et al., 2012).

3.2.2 Techniques to measure DNA methylation levels

In the beginning of the 1980s the earliest techniques to measure DNA methylation were based on the separation of methylated and unmethylated cytosines via chromatography and enabled to assess the total amount of 5mC in genomes (Harrison and Parle-McDermott, 2011). Until the late 1990 various methods combined polymerase chain reaction (PCR) amplification with the use of methylation-sensitive restriction enzymes, immunoprecipitation or bisulfite sequencing to detect DNA methylation. Nowadays, the most commonly used methods to measure DNA methylation at base pair resolution are based on the finding that sodium bisulfite can be used to convert unmethylated cytosine to uracil (Hayatsu et al., 1970) and that the conversion rate for 5mC is much slower (Wang et al., 1980). Frommer et al. (1992)

described that these different conversion rates can be used to analyze DNA methylation profiles of genomes. This realization revolutionized the way one could obtain DNAm levels (Harrison and Parle-McDermott, 2011). Classical whole genome DNA sequencing strategies were based on the cloning of DNA fragments into bacterial vectors, amplification and subsequent Sanger sequencing (Prober et al., 1987). Pyrosequencing bypassed the time-consuming subcloning-step (Tost and Gut, 2007). The coupling of bisulfite conversion and subsequent detection of 5mC sites via microarrays/ beadchips and high-throughput sequencing increased the throughput genome-wide DNAm analysis. Today, there are a great variety of methods investigating DNAm in different dimensions of resolution and genomic regions covered ranging from genome wide assays to targeted approaches (Laird, 2010; Masser et al., 2018; Shin et al., 2014). The strength and weaknesses of most of these approaches have been described by Barros-Silva et al. (2018); Laird (2010). The newest methods to analyze DNA methylation target challenges in the field to uncover the diversity of methylation patterns in single cells (Luo et al., 2018; Masser et al., 2018), detection of other types of DNA modifications (Shin et al., 2014) or to develop sequencing strategies to avoid the bisulfite conversion step e.g Nanopore sequencing (Burgess, 2017; Schatz, 2017). Major challenges in the development of DNA methylation assays for biomarker research and clinical application are robustness, accuracy and costs per sample. A community wide benchmark study by the Blueprint consortium showed that amplicon bisulfite sequencing and bisulfite pyrosequencing had the best all-around performance as DNA methylation assays for biomarker development and clinical applications (consortium, 2016).

Targeted bisulfite sequencing (TBS) offers a candidate approach to perform such studies with high resolution by increasing depth of read coverage per CpG to detect small changes in DNA methylation in a cost-efficient manner. Recently, a few applications of TBS have been developed with differences in accuracy, throughput and library preparation (Bernstein et al., 2015; Chen et al., 2017; Masser et al., 2013; Masser et al., 2015; Tost and Gut, 2007). Assessing DNA modifications with high accuracy and sensitivity in candidate loci would increase the power to detect and replicate the embedding of GxE interactions onto epigenetic layers as well as perform time course experiments in large numbers of samples to understand the stability of the environmentally induced changes during development. In addition, changes related to environmental exposure such as adverse life events and

psychopathology are often only present in specific cell types, although most studies rely on more complex tissues such as post-mortem brain or blood samples. Detecting these effects in mixed tissues requires high accuracy in order to detect small changes emerging from a small number of cells.

3.3 The 3D genome

The genome contains heritable information of eukaryotic organisms and is stored in the nucleus of cells. In order to fit the human genome with an approximate size of 3×10^9 bp into the space of a nucleus efficient compaction of the DNA is necessary. The organization of the DNA has not to only be efficiently condensed for chromosome separation during mitosis but also enable a dynamic access of its information for the specific expression of genes during the development and cell fate decisions (Sati and Cavalli, 2017). A vast amount of data suggests that the interphase genome is not randomly folded but rather is organized into at least 4 hierarchical layers of different scales (reviewed in Bonev and Cavalli, 2016; Buchwalter et al., 2019; Gibcus and Dekker, 2013; Krumm and Duan, 2018; Uhler and Shivashankar, 2017; Yu and Ren, 2017). At the largest scale (ca. 100000 kb), each chromosome occupies a specific space in the nucleus, which is called chromosome territory (CT) (reviewed in Cremer and Cremer, 2010) and overlaps between chromosomes only appear at its borders (Branco and Pombo, 2006). CTs are not randomly distributed in the nucleus but follow specific rules. First, CT positions are partially conserved after mitosis (Parada et al., 2003). Second, CT's have a radial positioning preference, in general small gene-rich chromosomes tend so be placed in the center, whereas larger gene-poor chromosomes are situated at the nuclear periphery (Croft et al., 1999). Third, early-replicating loci and active genes have the tendency to be in the nuclear interior, while late-replicating loci and repressed genes preferably occupy the nuclear periphery (Grasser et al., 2008; Takizawa et al., 2008). Chromatin interacting with the lamina of the nucleus is referred to as lamina-associated domains (LADs) and consists mostly of transcriptionally silent chromatin enriched with marks of heterochromatin (reviewed in Buchwalter et al., 2019). Compartments (scale ca. 10000 kb) can be divided into A and B compartments, which reflect the general chromatin state of the loci and genes located within these separated regions (Lieberman-Aiden et al., 2009; Rao et al., 2014; Wang et al., 2016). The A compartments harbor early replicating loci with a

high density of genes of which exhibit strong gene expression. Conversely, late replicating loci and loci often situated at LADs are found in the B compartment (Ryba et al., 2010). Interestingly, during stem cell differentiation around 60% of the human genome display an A/B compartment switch (Schmitt et al., 2016) and therefore partitioning in a compartment is to a certain degree cell type specific. Topological associating domains (TADs) partition the genome into blocks (scale 100 – 1000 kb), which are described as a region in which loci show a higher contact frequency compared to the rest of the genome which have been found in human, mouse (Dixon et al., 2012; Nora et al., 2012) and non-mammalian genomes like *Drosophila* (Sexton et al., 2012) zebrafish (Gomez-Marin et al., 2015), *Caenorhabditis elegans* (Crane et al., 2015) and yeast (Hsieh et al., 2015). In *Arabidopsis thaliana* (Wang et al., 2015) these domains were not found, indicating that in plants different mechanisms may be present to structure the chromatin. The positioning of the TADs in the human and mouse embryonic stem cells overlay in 50-70% of all cases, indicating that the location of TADs in mammalian genomes is evolutionary conserved (Crane et al., 2015). Although it is not fully understood how the boundaries of TADs are defined, they are enriched with multiple factors such as CTCF binding sites, housekeeping genes, transcription start sites, short interspersed nuclear elements (SINE), transfer RNA (tRNA), K4-trimethylated histone H3 (H3K4me3), K36-trimethylated histone H3 (H3K36m3) (Yu and Ren, 2017). Especially interesting is the occurrence of CTCF, which is the case for 75% of all TADs in mouse embryonic stem cells (Dixon et al. 2012). CTCF is known to be a key player in genome organization with the ability to form homodimers generating long-range looping which can enable or block promoter-enhancer interactions (Phillips and Corces, 2009). The loop extrusion model tries to explain the formation of TADs with these CTCF-mediated long-range loopings and a chromatin motor complex such as the cohesion complex which loads onto the DNA and extrudes a loop until it is stalled by binding of CTCF at convergent oriented binding sites (Sanborn et al., 2015). It has to be noticed that a distinct mechanism independent of CTCF might be responsible for TADs lacking CTCF occupancy (Yu and Ren, 2017). TADs do not only act as structural but also functional units that possibly coordinate gene expression by blocking, guiding or facilitating promoter enhancer loops (Dekker and Heard, 2015; Dixon et al., 2015; Ma et al., 2015; Nora et al., 2012; Symmons et al., 2016; Symmons et al., 2014; Zhan et al., 2017). At the lowest level long-range chromatin interactions can connect genomic loci separated

by several 100 kb of distance. These loops can be classified into two types. The first type consists of structural loops associated with CTCF binding, which for example form TADs. The second type of loops consists of functional loops driving gene expression enabling to build complex networks of enhancer-promoter, promoter-promoter or enhancer-enhancer interactions. These functional loops are often associated with enhancer-associated factors as well as the cohesion or mediator complex (Ji et al., 2016; Phillips-Cremins and Corces, 2013). Interestingly, there are indications that structural loops and functional loops can be organized in foci, which often appear to be spatially linked (Tang et al., 2015). The nuclear architecture can be dynamically regulated, which is key to fulfill its functional roles and some examples of reorganization occur within distinct biological processes such as mitosis, X-chromosome inactivation, circadian oscillating transcription, pluripotency and differentiation (Yu and Ren, 2017). The importance of time as the fourth dimension becomes clear when regarding the spatial and temporal organization of nuclear protein structures (reviewed in Hemmerich et al., 2011). Protein structures like DNA replication sites, the nucleolus, telomeres, centromeres and nuclear lamina can be stable for seconds up to several hours and that the stability can be modulated depending on the structures' function.

3.3.1 The „C“-Techniques

First insights into the organization of genomes were obtained with microscopy-based techniques e.g. fluorescence *in situ* hybridization (FISH), which allowed the direct observation of spatial distance between genomic loci and their dynamic movement in single cells. However, these microscopy-based imaging tools are limited in throughput (single cells), resolution (defined by the microscope, observation of higher order structures) and genome coverage (simultaneous observation of >30 genomic locations (Wang et al., 2016)). The second principal to investigate the 3D genome organization is a molecular-biology-based approach developed by Dekker et al. (2002). This technique is called chromosome conformation capture (3C) and the rationale of 3C-based methods is the conversion of chromatin interactions into ligation products (reviewed by Dekker and Mirny, 2013). Loci in physical proximity in the genome (estimated range 10-100 nm) are stabilized by cross-linking DNA-Protein interactions with the use of cross-linking agents such as formaldehyde. Subsequently, the chromatin is fragmented by restriction digestion or sonication,

followed by ligation of cross-linked fragments. Afterwards the hybrid DNA is purified and the ligation events are quantified by PCR using locus specific primers or sequencing over the ligation site. Due to the fact that an individual cell can provide maximal 2 data points, populations of cells are used to obtain a mean interaction profile of the contacts in the genome. A plethora of 3C-based methods have been developed (reviewed by Davies et al., 2017; Krumm and Duan, 2018; Li et al., 2018; Sajan and Hawkins, 2012; Simonis et al., 2007; van Steensel and Dekker, 2010) to detect chromatin interaction at different resolution (key factor: Fragmentation method of cross-linked chromatin e.g. 6 or 4-base pair restriction enzymes, MNase, DNase and sonication), sensitivity (key factor: number of observations needed to produce interaction profiles) but also with the focus on interactions mediated by specific proteins (e. g ChIA-PET; Enrichment of protein-mediated integrations ameliorated by immunoprecipitation using specific antibodies against factors of interest). High-resolution interaction profiles from few target loci can be achieved using circularized chromosome conformation capture (4C) or Capture-C. If less information about a region of interest is known, chromosome conformation capture carbon copy (5C) can be used detecting all interactions among multiple selected loci. The 3D structure of whole mammalian genomes can be analyzed using Hi-C. The development of *in situ* ligation helped to increase the resolution of Hi-C approaches enabling to visualize chromatin loopings within TAD structures (Rao et al., 2014). Despite the fact, that 3C-based assays have reached a level of maturity and approachability to use them as general tools, current limitations of this approach are the bias of interactions which can be analyzed due to formaldehyde cross-linking, uneven distribution of restriction sites in the genome and interpretation of 3D data visualized in 2D plots (Dekker, 2006; Dekker and Mirny, 2013; Simonis et al., 2007).

The two general approaches - microscopy and 3C - to study genome organization therefore complement each other and should be utilized simultaneously. While there is a good overlay of observations from both techniques onto nuclear organization, some discrepancies at single loci have been described (Dekker, 2016; Williamson et al., 2014). Recent developments in both fields such as STORM (Rust et al., 2006) and single cell 3C approaches (Nagano et al., 2015; Nagano et al., 2017; Ramani et al., 2017; Stevens et al., 2017) will help to disentangle the complexity and dynamics of the 3D genome organization but most promising will be multidisciplinary approaches like the 4D Nucleome project (Yu and Ren, 2017) and the

implementation of the CRISPR-Cas9 tools such as CRISPR-GO (Wang et al., 2018b).

3.3.2 The relevance of chromatin structure in disease

The 3D architecture of the human genome is a hierarchically layered structure that undergoes dynamic restructuring to enable a healthy development. A disruption of the 3D genome at or across different scales can lead to disease and often involves the mutation in architectural important genes or regulatory elements of gene expression (Chakraborty and Ay, 2018; Li et al., 2018). Studies investigating the molecular mechanisms of cancer formation show that alterations at chromosomal territories, compartments, TADs and chromatin loops are potentially involved to develop the disease phenotype (Chakraborty and Ay, 2018). At the level of CTs it has been indicated that the dynamic intermingling at the territorial borders has been correlated with translocation frequencies in human cells, which may play a role in various cancers (Branco and Pombo, 2006). A change of compartments has been shown for important genes in breast cancer (Barutcu et al., 2015) and in a heart failure model (Rosa-Garrido et al., 2017). Moreover, there is evidence for higher mutation rates in the B over the A compartment in prostate cancer cells (Schuster-Bockler and Lehner, 2012). Up to now most studies linking 3D architecture with disease reveal a disruption of architectural loops (CTCF-CTCF interactions forming TADs) or functional loops (e.g. enhancer-promoter loops reviewed in Chakraborty and Ay (2018); Krijger and de Laat (2016); Krumm and Duan (2018)). The disruption of TADs due to structural variants (reviewed in Spielmann et al., 2018) generic disruption (Hnisz et al., 2016) or epigenetic inactivation (Flavahan et al., 2016; Victoria-Acosta et al., 2015) of architectural protein binding sites can lead to rewiring of enhancer-promoter interactions, which would normally be isolated from each other and therefore can alter the gene expression pattern of disease-relevant genes. Intra-TAD duplications can lead to duplications or loss of enhancer's that can lead to a sex reversal phenotype (Cox et al., 2011) (enhancer duplications or loss in the SOX9 TAD), B cell lymphoma (Chapuy et al., 2013) (Duplication of a super-enhancer in the BCL6 TAD) and Polydactyly (Lupianez et al., 2015) (Limb enhancer duplication in the SHH TAD). Digit abnormalities in the Cooks syndrome is an example in which a TAD boundary is duplicated forming a new regulatory unit called "Neo-TAD" (Cox et al., 2011). The deletion of a TAD boundary can lead to the fusion of two TADs and

enables the rewiring of enhancer to gene promoters leading to abnormal gene expression patterns (e.g. congenital limb malformation caused by a CTCF-associated boundary loss at the EPHA4 (Lupianez et al., 2015)). The inversion or translocation of TAD boundaries can lead to a restructuring of TADs within the genome, which is called “TAD shuffling” causing the adaption or loss of enhancers to their target genes. Two diseases in which this TAD abnormality occurs, is the F-Syndrome (Lupianez et al., 2015) exhibiting syndactyly or atypical Rett syndrome (Goubau et al., 2013). Similar to architectural loops forming TADs, genetic variants can disrupt or establish functional loops and alter the gene expression of genes and potentially leading to disease phenotypes. The dysregulation of gene expression can be achieved by enhancer formation (T cell acute lymphoblastic leukemia (Mansour et al., 2014)), deletion (α -thalassemia (Fritsch et al., 1979), weakening (holoprocencephaly (Jeong et al., 2006), strengthening (preaxial polydactyly type II (Lettice et al., 2003), duplicated (lung adenocarcinoma (Zhang et al., 2016)). An example of a disease in which a SNP can lead to the introduction of a new promoter or deletion of a promoter is α -thalassemia (De Gobbi et al., 2006; Lower et al., 2009). The catalogue of known human disease-associated variants has dramatically expanded due to the advent of Genome-wide association studies (GWAS) (Visscher et al., 2017). Variants found linked to disease by GWAS are often located in non-coding regions and 3C methods can be applied to identify candidate genes, which can be affected by the variant. A multi-omics strategy to discover the molecular mechanisms causing disease of associated variants has been proposed by Krijger and de Laat (2016). This approach uses all disease-associated variants and haplotype-based SNP imputation to identify all associated variants. Afterwards, the epigenetic data is used to co-localize candidate causal variants with potential regulatory sequences. Subsequently, genes are identified which are located in the same TAD using HI-C data and physical links of candidate variants to genes are monitored with high-resolution contact maps e.g. obtained from 4C. The next step, is to establish a functional link of the variant to the gene by allele specific gene expression analysis. To finally confirm the molecular mechanism site-directed genome-editing is required and the use of a model organism can confirm that the effect of the variant is relevant for the disease. In order to better investigate molecular mechanisms in psychiatric disorders, the PsychENCODE Consortium has

generated a comprehensive resource integrating multi-omics data including Hi-C maps derived from the adult brain (Wang et al., 2018a).

3.3.3 Steroid Hormones and the 3D chromatin structure of *FKBP5*

The regulation of hormone responsive genes exhibits a shared mode of action. The sequence of hormone response elements like GREs are ubiquitously distributed in the genome, but the presence of the binding motif is not sufficient for GR binding but rather highly selective and cell type specific (John et al., 2011; Uhlenhaut et al., 2013) mainly due to chromatin accessibility but also to a lesser extent by DNA methylation at GREs (Wiench et al., 2011). Upon activation hormone receptors like GR do not only bind directly to the proximal promoter regions of responsive genes but in the majority of cases, to intragenic and far distance sites that correspond to enhancer regions (John et al., 2011; Le Dily and Beato, 2018; Reddy et al., 2009). It has been shown that enhancers regulate GC-mediated activation or repression by the recruitment of transcription factors (reviewed by Grbesa and Hakim (2017)) and not only affect genes in close proximity but also in far distance by the formation of long-range chromatin loops (Biddie, 2011; Hakim et al., 2009; Harmston and Lenhard, 2013). This regulation can involve a complex network of multiple promoters and enhancers (Jin et al., 2013; Le Dily and Beato, 2015; Sanyal et al., 2012). Due to the promiscuous behavior of enhancers (Krijger and de Laat, 2016), TADs facilitate the correct organization of these structures by restraining their actions within a sub-megabase scale (Le Dily and Beato, 2015). It seems that the functional loopings are dynamically regulated upon hormone exposure but stay restricted within the stable TAD boundaries (Jin et al., 2013; Le Dily and Beato, 2015). Interestingly, pre-existing functional loops seem to favor the response to an external stimuli (Grbesa and Hakim, 2017; Jin et al., 2013), indicating that the gene expression response is not only regulated by the establishment and disintegration (Fullwood et al., 2009; Perillo et al., 2008; Quintin et al., 2014) but also by the stability of functional loops (Hakim et al., 2011; Le Dily et al., 2014). In this regard, a dysregulation of architectural proteins/ loops could lead to changes in the stability of functional interactions and lead to an altered response in gene expression (Antony et al., 2015; Quintin et al., 2014; Seitan et al., 2013). Therefore the picture emerges that the 3D genome structure is an important factor for the adaptation to an external stimulus and that the complex interplay of functional and architectural chromatin interactions is a dynamic

playground to integrate external changes in the light of previous experiences but also inherited genetic factors of the individual to enable an adequate reaction.

FKBP5 is a glucocorticoid-regulated gene in which a molecular mechanism for the integration of genetic and environmental factors has been described (see section 3.1.3). To understand how genotypes can influence the complex fine-tuning process of *FKBP5* gene expression regulation in health and disease, one has to understand the molecular regulatory landscape of this gene. An *in silico* analysis comprising the key features of the *FKBP5* locus structure using available literature and datasets (Consortium, 2004; Paakinaho et al., 2010; Rao et al., 2014; Tang et al., 2015) (Figure 4) showed that the topology of the *FKBP5* locus is formed by CTCF, which generates an architectural chromatin loop, the TAD, in which RNA polymerase II (RNAPII) and other proteins like TFs (e.g. GR) further bind and substructures the locus to regulate *FKBP5* gene expression. Active epigenetic marks, such as K4-methylated histone H3 (H3K4me) and K27-acetylated histone H3 (H3K27ac) and H3K36me3 indicate that the locus conformation is already pre-existing before exposure to steroids in blood, prostate and lung epithelial cancer cells (Jaaskelainen et al., 2011) and therefore poised for transcription.

In the same haplotype of *FKBP5*, which comprises the disease-associated functional SNP rs1360780, a larger structural variant esv3608688 can be found (Pelleymounter et al., 2011). This 3.3 kb large insertion/ deletion (INDEL) is located in intron 1 of the *FKBP5* gene and the deletion was found in 65.385 % and the insertion 34.615% of all alleles screened by the 1000 genomes project. What is unknown so far, is how the combination of different genetic factors like rs1360780 and INDEL allele in intron 1 affects the chromatin structure and therefore changes gene expression under basal and stress conditions and which chromatin status is more responsive to induction by glucocorticoids and possibly induce other epigenetic changes like methylation levels of functional elements in the *FKBP5* locus.

Introduction

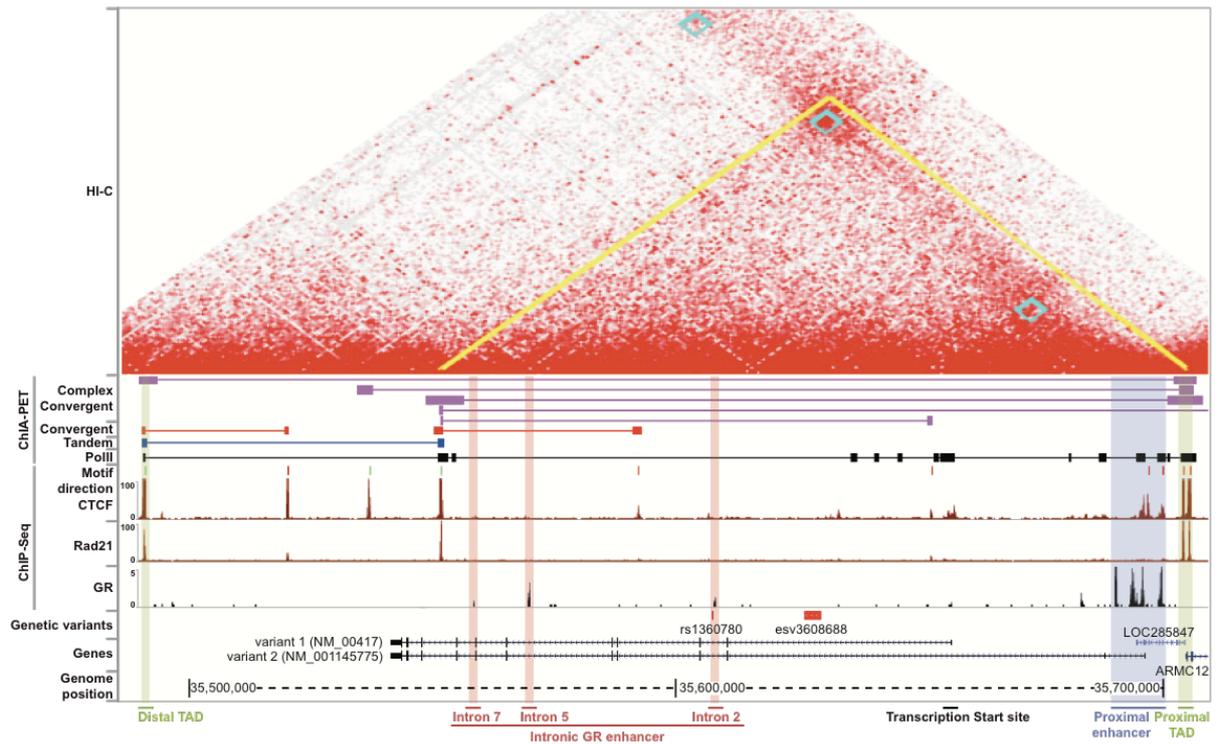


Figure 4: *In silico* analysis of the 3D structure of the FKBP5 locus.

Genome browser shot and Hi-C profile illustrating the FKBP5 locus 3D structure (hg19 / chr6:35487554-35718452) and the location of the genetic variants rs1360780 & esv3608688 in relevance to key functional sites using reference data. ChIP-Seq – The binding of the Glucocorticoid receptor (GR; ENCODE GR HAIB TFBS track from A549 cells), CTCF and Cohesin (indicated by its subunit Rad21)(both ENCODE SYDH TFBS tracks from GM12878) at the FKBP5 locus. ChIA-PET – Chromatin interactions mediated by PolII and CTCF at the FKBP5 locus (Tang et al., 2015)). CTCF interactions are classified in convergent (loop formed by CTCF sites with opposing motif directions), complex convergent (loop formed by CTCF sites with opposing motif directions, between the interaction CTCF sites are other CTCF sites forming chromatin interactions) and tandem loops (loop formed by CTCF sites with the same motif directions) according to Tang et al. (2015). Hi-C – The FKBP5 locus topologically associated domain structure (TAD) in GM12878 using data from Rao et al. (2014) (visualized by Juciebox from the Aiden-Lieberman lab (Durand et al., 2016), in situ Mbol primary + replicate) Source: Tobias Wiechmann, 2019

4 Aim of the thesis

The development of stress-related psychiatric disorders involves the integration of genetic and environmental factors. How these factors are integrated on a molecular level to determine risk or resilience is far from being understood. The long-term integration of both factors is thought to be mediated by epigenetic adaptations. One molecular mechanism for a statistical GxE interaction has been described for the *FKBP5* gene. In order to further investigate how genetic variants and environmental impacts are integrated at a molecular level to regulate *FKBP5* gene expression, this thesis encompassed the following objectives:

Objective 1: Develop a method to accurately measure DNA methylation levels within key regulatory sites of the *FKBP5* gene in large cohorts of patients

This objective was addressed by the development of highly accurate methylation measurements via targeted bisulfite sequencing (HAM-TBS) (Paper 1, chapter 5.1). The establishment, validation and key features of this method are discussed in Section 6.1.

Objective 2: Investigate how environmental signals can relate to *FKBP5* DNA methylation levels

This objective was addressed by studying dynamic methylation changes in preterm infants (Paper II, chapter 5.2). The new possibilities of HAM-TBS have been used to better understand the underlying dynamics of GC-induced DNAm changes in *FKBP5* in healthy individuals upon a DEX challenge (Paper III, chapter 5.3). The implications from Paper II & Paper III are discussed in chapter 6.2.

Objective 3: Investigate, how a 3.3 kb large INDEL (esv3608688) located in intron 1 can modulate GR-dependent gene expression with focus on alterations on the 3D chromatin structure and DNA methylation levels at functional sites of the *FKBP5* gene.

This objective was addressed by studying the effects of the INDEL allele onto *FKBP5* gene expression utilizing various quantitative PCR (qPCR) assays, onto *FKBP5* chromatin structure by application of a 4C approach and *FKBP5* DNAm levels using HAM-TBS in lymphoblastoid cell lines (LCLs) (Manuscript I, chapter 5.4). The findings and the derived model of how the INDEL allele potentially regulates GC-induced gene expression changes are discussed in chapter 6.3.

5 Results

5.1 Paper I - HAM-TBS: high-accuracy methylation measurements via targeted bisulfite sequencing

Roeh, S.*, T. Wiechmann*, S. Sauer, M. Kodel, E. B. Binder and N. Provencal (2018) (* shared first authors)

Originally published in:

Epigenetics Chromatin. 2018 Jul 4;11(1):39. doi: 10.1186/s13072-018-0209-x.

PMID: 29973294

Additional supplementary material:

Additional supplementary figures and tables are freely available for download at the publisher's website (<https://doi.org/10.1186/s13072-018-0209-x>).

METHODOLOGY

Open Access



HAM-TBS: high-accuracy methylation measurements via targeted bisulfite sequencing

Simone Roeh^{1†}, Tobias Wiechmann^{1†}, Susann Sauer¹, Maik Ködel¹, Elisabeth B. Binder^{1,2} and Nadine Provençal^{1,3,4*}

Abstract

Background: The ability to accurately and efficiently measure DNA methylation is critical to advance the understanding of this epigenetic mechanism and its contribution to common diseases. Here, we present a highly accurate method to measure methylation using bisulfite sequencing (termed HAM-TBS). This novel method is able to assess DNA methylation in multiple samples with high accuracy in a cost-effective manner. We developed this assay for the *FKBP5* locus, an important gene in the regulation of the stress system and previously linked to stress-related disorders, but the method is applicable to any locus of interest.

Results: HAM-TBS enables multiplexed analyses of up to 96 samples and regions spanning 10 kb using the Illumina MiSeq. It incorporates a triplicate bisulfite conversion step, pooled target enrichment via PCR, PCR-free library preparation and a minimum coverage of 1000x. TBS was able to resolve DNA methylation levels with a mean accuracy of 0.72%. Using this method, we designed and validated a targeted panel to specifically assess regulatory regions within the *FKBP5* locus that are not covered in commercially available DNA methylation arrays.

Conclusions: HAM-TBS represents a highly accurate, medium-throughput sequencing approach for robust detection of DNA methylation changes in specific target regions.

Keywords: Targeted bisulfite sequencing, DNA methylation, Next-generation sequencing, 5-methylcytosine, *FKBP5*

Background

DNA methylation is the covalent addition of a methyl group at the 5-carbon ring of cytosine, resulting in 5-methylcytosine (5mC). In the mammalian genome, this occurs predominantly in the context of CpG dinucleotides. It is one of several epigenetic marks influencing gene expression and serving multiple other purposes such as genomic imprinting, X chromosome inactivation and maintenance of genomic stability [1, 2]. Aberrant regulation of the establishment, maintenance, erasure or recognition of DNA methylation has been associated with a range of disease phenotypes [3, 4]. In addition, lasting effects of environmental risk factors may be reflected by changes in DNA methylation [5]. The need

to measure DNA methylation in large human cohorts in a cost-effective manner is therefore of increasing interest for research in epidemiology and medicine [6].

Assessing DNA modifications with high accuracy and sensitivity in candidate loci would increase the power to detect and replicate such effects as well as to perform time course experiments in large numbers of samples to understand the stability of the environmentally induced changes during development. In addition, changes related to specific environmental exposure may only be present in specific cell types, although most studies rely on more complex tissues such as postmortem brain or blood samples. Assessing these effects in mixed tissues requires high accuracy in order to detect small changes emerging from a small number of cells. DNA bisulfite treatment followed by next-generation sequencing enabled the quantification of DNA methylation marks at single-base resolution. However, genome-wide bisulfite sequencing, although the best approach to identify DNA

*Correspondence: Nadine_Provençal@sfu.ca

[†]Simone Roeh and Tobias Wiechmann shared first authorship

³ Faculty of Health Sciences, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada

Full list of author information is available at the end of the article

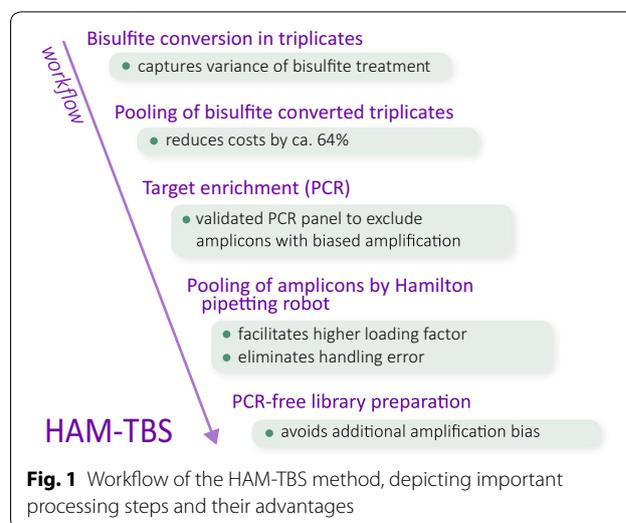


modifications, is still too cost intensive to be applied to large human cohorts at the coverage needed ($>60\times$) to detect differentially methylated sites [6]. Another set of accurate and cost-efficient measurement methods for DNA methylation at single CpG level are Illumina DNA methylation arrays. However, the ones currently available lack coverage in key enhancer regions that are important for environmentally driven changes and have a relatively small number of probes ($\sim 10\text{--}13$) covering each site. Targeted bisulfite sequencing (TBS) offers a candidate approach to perform such studies with high resolution by increasing depth of read coverage per CpG to detect small changes in DNA methylation in a cost-efficient manner. Recently, few applications of TBS have been developed with differences in accuracy, throughput and library preparation [7–10]. Our TBS approach focuses on the *FKBP5* gene, which encodes the FK506-binding protein (FKBP51), a co-chaperone tightly involved in stress regulation. Genetic and epigenetic factors have repeatedly been shown to increase the activity of this gene and associated with increased stress-reactivity and psychiatric disorders [11]. We have previously reported allele-specific demethylation of CpG sites located in intronic enhancer regions of *FKBP5* specific to posttraumatic stress disorder (PTSD) in patients who had experienced child abuse [12]. These gene \times environment interactions (G \times E) may be mediated by differential susceptibility to adversity-induced changes in DNA methylation in specific enhancers. Current methods do not cover the relevant enhancer regions of *FKBP5* affected by stress exposure. A highly accurate, cost- and time-efficient method to investigate *FKBP5* DNA methylation in a large number of samples is thus critical to gain more insight into how DNA methylation changes may mediate these G \times E. In this manuscript, we present a cost-effective, high-accuracy methylation measurement TBS (HAM-TBS) method to assess the regulatory regions of the *FKBP5* locus. Incorporating a triplicate bisulfite conversion step, PCR-free library preparation and rigorous quality control (validation of PCR target sites, $>95\%$ bisulfite conversion efficiency and $1000\times$ coverage minimum) ensures that our method is extremely robust (Fig. 1). Medium throughput and handling accuracy of up to 96 samples spanning approximately 10 kb is facilitated by embedding the Hamilton pipetting robot and TapeStation with the Illumina MiSeq sequencer.

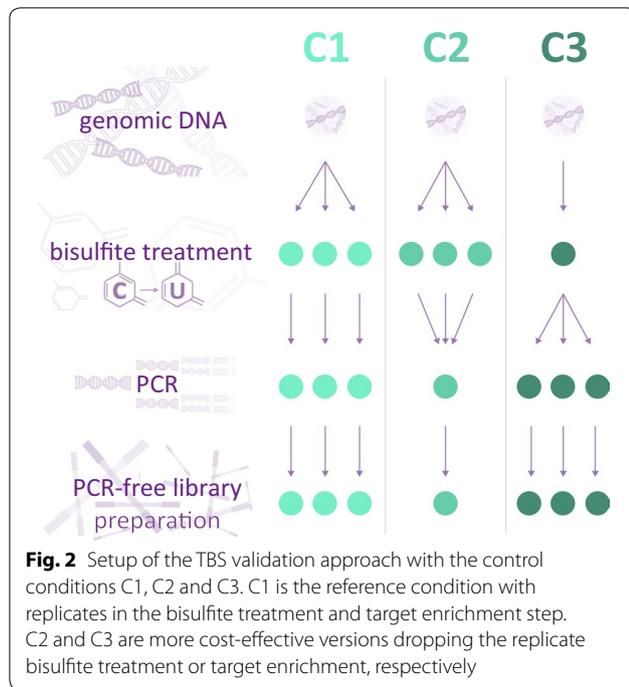
Results

QC, validation and optimization of the HAM-TBS method

TBS is based on bisulfite conversion coupled with targeted enrichment via PCR, library preparation for



sequencing and subsequent quantification of methylation levels. All steps are necessary and may influence the outcome by introducing bias to the assessment of methylation levels or by insufficient quality control of the data. The standard approach to minimize potential biases before sequencing is to produce replicates and assess the mean methylation levels during the analysis. In order to design a highly accurate yet cost-effective approach that is amenable to multiplexing, we assess at which step (bisulfite conversion or amplification) and to what extent technical variability would be introduced, as well as which quality control steps need to be performed on the sequencing data to ensure a robust analysis. To this end, we assessed the methylation level of 0, 25, 50, 75, 100% in vitro methylated bacterial artificial chromosome (BAC) control DNA for 3 different combinations of pooling strategies during the bisulfite treatment and PCR amplification (Fig. 2). Condition 1 (C1) assessed the methylation levels of control DNA using triplicate bisulfite treatments and PCR amplification for each replicate. C1 was considered the standard reference condition since each step was performed in triplicates. In condition 2 (C2), triplicate bisulfite treatments were pooled to perform one PCR amplification reducing the costs by approximately 64%. Finally, in condition 3 (C3) one bisulfite treatment of the control DNAs was performed followed by 3 separate PCR amplifications to assess the extent of the target enrichment bias. A smaller panel of 11 different PCRs (Fig. 3) within the *FKBP5* locus (see table in Additional file 1) served as basis for this analysis. Before comparing the three conditions, the collected sequencing data were subjected to three quality control steps in order to ensure accurate assessment of minimal methylation levels as well as small changes between samples.



1. *Bisulfite conversion rate > 95%*. We assessed the bisulfite conversion rate per sample and per amplicon and excluded rates lower than 95% from the analysis.
2. *Removal of PCR artefacts* During the target amplification, the PCR occasionally introduces artefacts presenting non-existent CpG sequences in the target region. They present at very low coverage and extreme levels of methylation (~0 or ~100%). In order to not exclude potential SNPs giving rise to CpGs, we removed artefacts on this basis rather than limiting the analysis to known CpGs according to the reference genome.
3. *Minimum coverage of 1000×* Higher sequencing depth and coverage of the CpGs yields higher accuracy of the methylation quantification. In order to determine the right balance between sequencing depth and thereby cost and sufficient accuracy, we took random subsamples of varying sequencing depth of an in silico created library representing methylation levels from 0 to 100% and assessed the standard deviation for each level of methylation with respect to coverage (Fig. 4a). To find a meaningful cutoff for coverage, we considered the trade-off between sum of the average standard deviation per amplicon (cost) present in various levels of coverage (Fig. 4b). In accordance with previous findings [7], we identified 1000× coverage as a useful cutoff for our analysis, as the gain in accuracy with increasing coverage above this threshold is low and 1000× is rea-

sonable to achieve for a larger locus, e.g., 9 kb in the *FKBP5* panel.

All PCRs for our validation experiment showed bisulfite conversion levels >99%. After QC, a total of 40 CpG spread across 7 amplicons remained in our analysis (1 PCR failed due to coverage <1000×, 1 showed non-linear amplification and coverage <1000×, 2 showed non-linear amplification). Methylation levels were very similar between all 3 conditions with an average error of <1% when comparing absolute methylation levels of C2 and C3 versus C1 (Fig. 5b). We calculated the R^2 values for each assessed CpG across the titration levels and used the mean per amplicon to compare the 3 conditions. R^2 is a measure for assessing linearity of amplification of the methylation signal, which is crucial when quantifying methylation changes in, e.g., cohort studies. Again, all conditions showed very high mean R^2 values above 0.99 (Fig. 5a). This confirms that all conditions are suitable for high-accuracy methylation detection. The introduced biases in our workflow, based on the control DNA, are minimal and enable very accurate methylation quantification even without including triplicates for the bisulfite conversion or target amplification. However, opposed to the target amplification, we cannot exclude slightly elevated variance of the bisulfite conversion on non-in vitro methylated DNA from, e.g., patients. Therefore, we chose to use C2 for our HAM-TBS method. While it still maintains a triplicate bisulfite conversion step, it is the most cost-effective of the tested conditions, an important factor when processing many samples from cohort studies.

Comparison of the technical accuracy of pyrosequencing to TBS

Next, we aimed to compare TBS to pyrosequencing, the reference method used for targeted DNA methylation analysis. We assessed the methylation levels of 5 CpGs within PCR_5 and PCR_11 measured by pyrosequencing as well as using HAM-TBS with the C1 protocol. The methylation analysis using pyrosequencing showed a high mean standard deviation of 4.68% with a maximum SD of 14.56%. The analysis using next-generation sequencing with C1 showed a much lower mean standard deviation of 0.72% with a maximum SD of 1.83%. This demonstrates a significantly lower technical variation and therefore higher accuracy when assessing methylation levels using a TBS approach.

Development of an extensive HAM-TBS *FKBP5* panel covering relevant regulatory sites

FKBP5 is an important gene in the field of psychiatry. The gene is larger than 100 kb rendering the assessment of the full locus including the adjacent up- and downstream

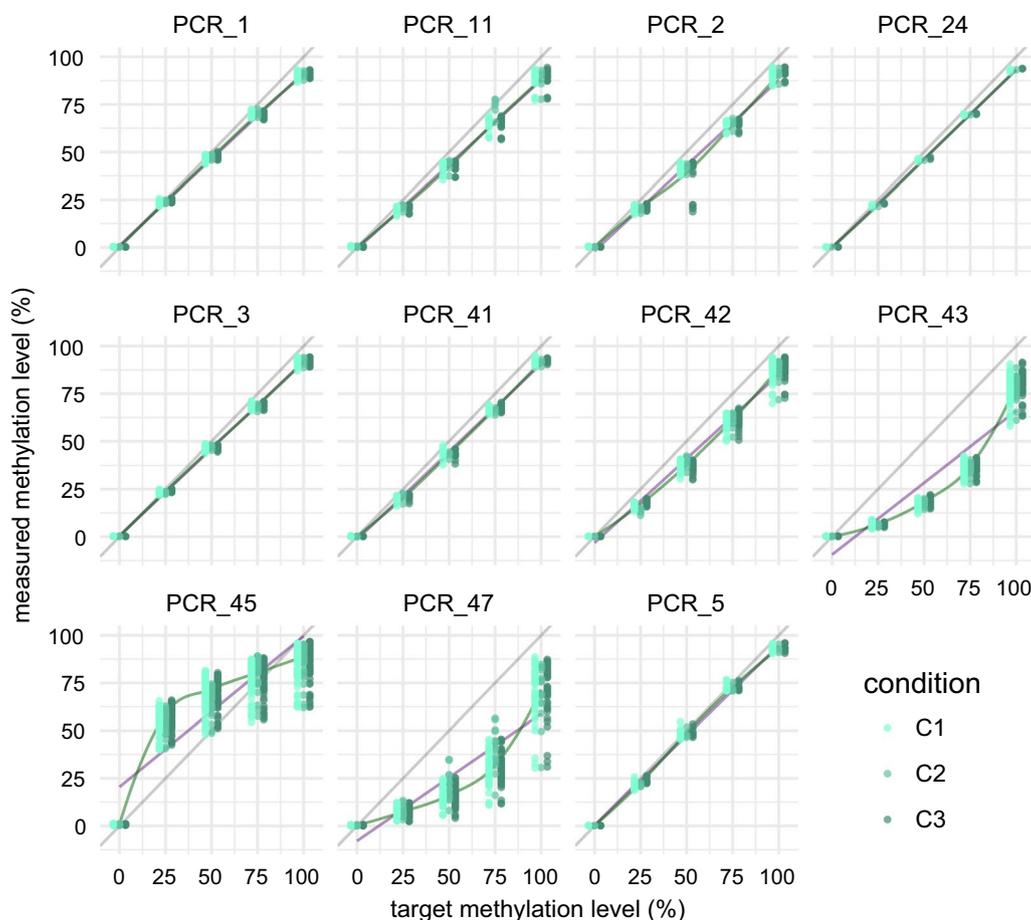


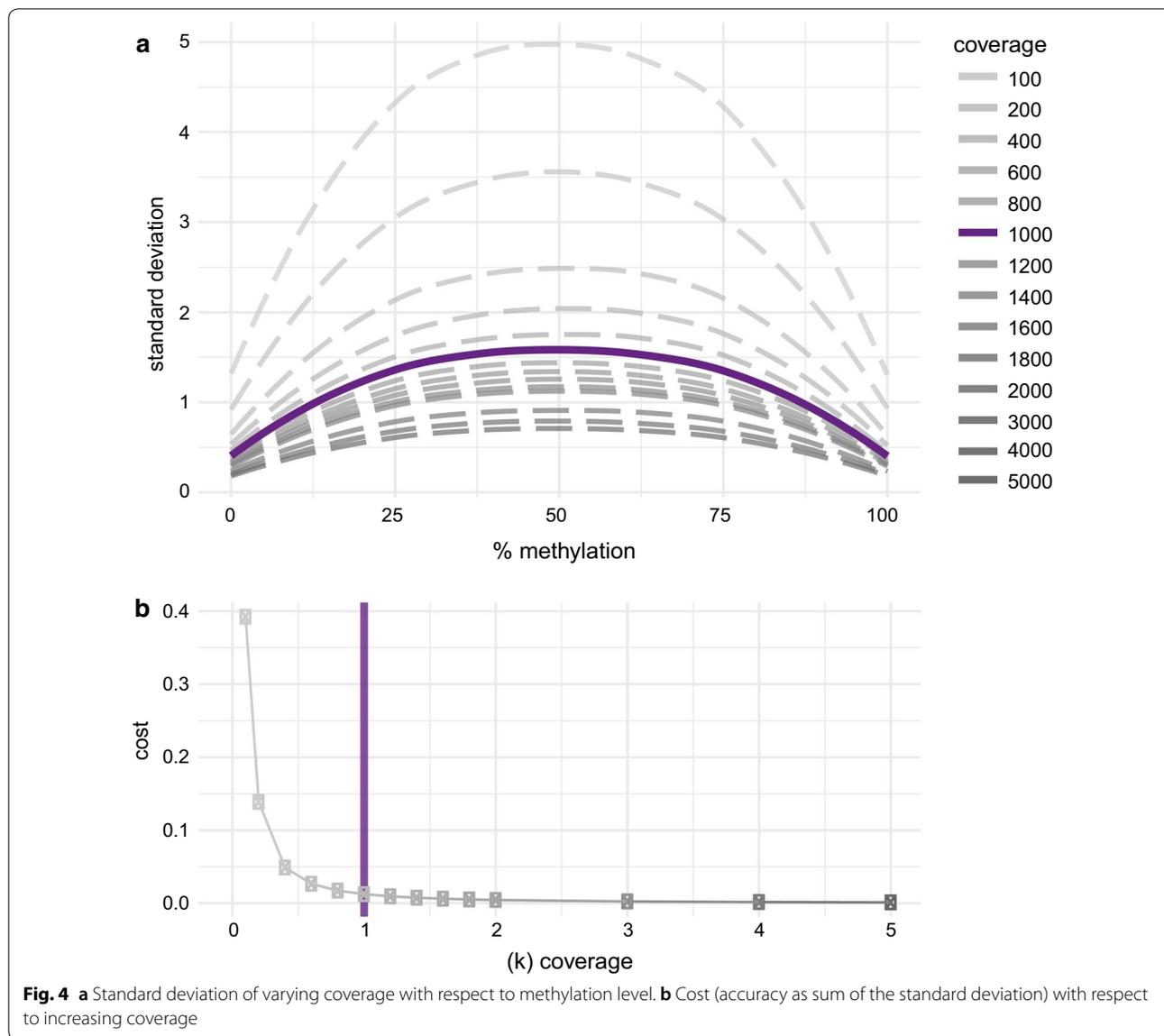
Fig. 3 Methylation quantification of the control DNA used to evaluate the technical variability. Linear regression line (purple), Loess fit line (green). PCR_3 was excluded due to low coverage, PCR_47 was excluded due to low coverage and nonlinear amplification, and PCR_43 and PCR_45 were excluded due to nonlinear amplification

regions unfeasible and too cost intensive for TBS methods. We thus restricted our analysis to functionally relevant sites of interest to ensure compatibility with targeted measurement methods and enable the assessment in large cohorts. To this end, we designed and validated a comprehensive amplicon panel (Fig. 6) including the TSS, TAD boundaries, intergenic and proximal enhancers as well as *GR* and *CTCF* binding sites (see methods for further details). The resulting HAM-TBS *FKBP5* panel is composed of 29 amplicons passing our QC's threshold (described above) and covering 315 CpGs across the locus. The sequencing data showed sufficient bisulfite conversion for all amplicons when performed on control DNA using C2. In total, 27 of the amplicons included in the panel presented good linearity (see figure in Additional file 2) across the assessed methylation levels. Two amplicons located near the TSS showed a mild PCR bias, where methylation levels were lower than expected for the 50% and 75% controls (PCR_7, PCR_9).

These amplicons have a very high CpG content of >25%; hence, CpGs in the primer could not be avoided. It has been previously shown that methylation levels in this region are very low (<5%) across tissues [12], so that any bias at higher methylation levels would not impair accurate quantification of this region. We thus incorporated sites located in this region in the panel, but they should be used with caution if higher methylation levels are observed. PCR_26 of the HAM-TBS *FKBP5* panel is located in the *H19* locus [13] which is an imprinted gene and serves as an internal positive control with an expected methylation level ~50%.

Application and costs

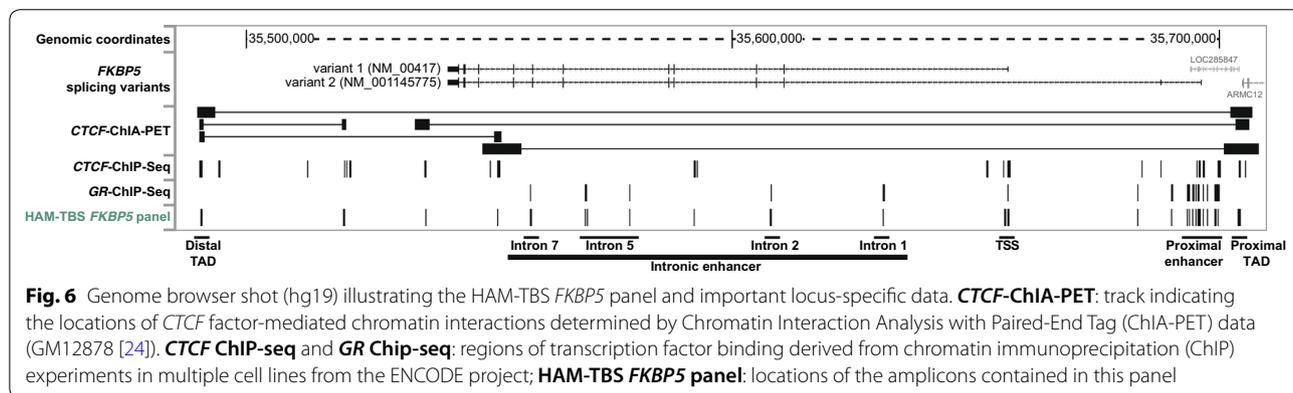
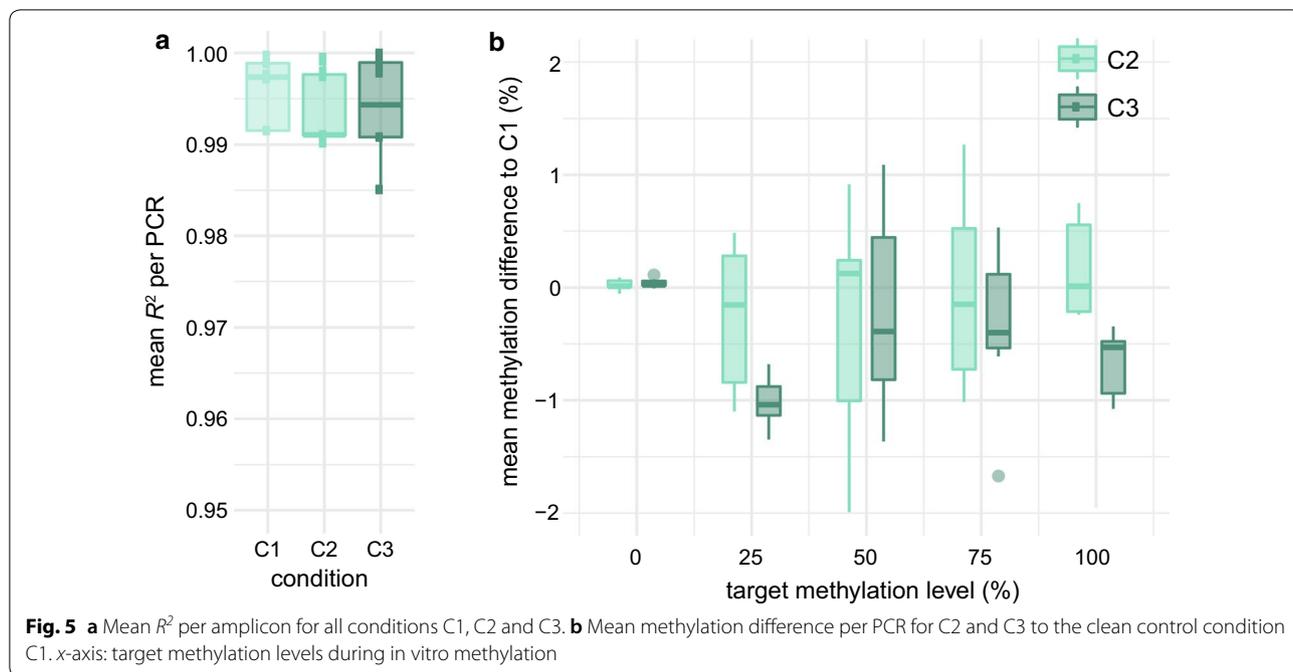
The HAM-TBS method can be multiplexed up to 96 samples in a medium-throughput manner. To demonstrate the applicability of our approach, quality control statistics of data derived from an experiment containing 95 blood samples from patients and the full *FKBP5*



panel of 29 amplicons are described here. After reads mapping and methylation calling, we identified PCR artefacts comprising ~1% of the methylation sites and removed them from the data, and 9 samples in 1 PCR showed insufficient bisulfite conversion rates (<95%) and were also removed. Two loci were identified as SNPs giving rise to a CpG sites in patients. In total, 91% of sample x amplicon data passed our filtering criteria. 27 amplicons passed QC with sufficient coverage and quality in >75% of samples, while two amplicons were dropped due to <1000× coverage (Additional file 3A, B). The control amplicon spanning the H19 imprinted locus for which methylation level is known to be ~50% [14] shows the expected methylation profile

in all samples (Additional file 3C). HAM-TBS approach allowed the quantification of 276 methylation sites for 95 samples in one single MiSeq run.

An assessment of the relative costs for each of the main reagents for this experiment containing 96 samples (95 patients and unmethylated control) with increasing number of amplicons assessed is depicted in Additional file 4. The quantifications using TapeStation and the PCR-free library preparation are the two most cost-intensive steps. The proportion of costs for the amplicon quantification using the TapeStation increases with the higher amount of amplicons investigated, while relative costs for the library preparation and sequencing chemistry decrease with the inclusion of more amplicons.



Discussion

We developed a targeted medium-throughput approach for measuring DNA methylation levels in multiple samples in parallel. This method enables cost-efficient high-resolution methylation measurements of target loci in cohorts of patients and probands at the *FKBP5* gene, a locus with large interest in the psychiatric and psychological community [11]. This cost-efficient, accurate method to determine *FKBP5* methylation levels would thus serve a large number of researchers. Our method is positioned between whole genome bisulfite sequencing and targeted approaches as pyrosequencing. The first is expensive and yields lower coverage and accuracy of single CpGs; the latter only allows to assess very small regions at a time and can generate significant variance between

replicates. HAM-TBS enables the analysis of a targeted but larger region (~10 kb) at high resolution and low costs. DNA methylation studies in large cohorts, investigating the impact of environment or association with disease status in mixed tissues, necessitate high accuracy at single-site resolution. In fact, TBS was able to resolve methylation levels with a mean accuracy of 0.72%. A high level of accuracy was maintained in more cost-efficient approaches using only one PCR amplification round. By pooling triplicate bisulfite treatments prior to PCR amplification, we can account for variance introduced by the bisulfite treatment but also reduce costs and hands-on time during the target amplification.

The accuracy of the method benefits from a PCR-free library preparation and rigorous quality control (prior

evaluation of linear PCR amplification of the target site, bisulfite conversion efficiency >95% and read coverage minimum of 1000×). Nonetheless, a proper assessment of possible amplification biases due to the choice of amplicon location in the design step is critical. Some loci can show nonlinear amplification curves, which renders them inappropriate for methylation quantification. Adjustment of primer design and PCR conditions may help solve this issue, but for some loci optimization may not be possible. For instance, in CpG islands with high CpG density, we observed that amplification curves were not linear, revealing a bias which became more pronounced as the level of methylation increased. Differential methylation results from these sites should be interpreted with caution and perhaps require additional replication. Besides validating each amplicon prior to usage, including controls such as in vitro unmethylated DNA, water and endogenous hemimethylated region, the H19 locus, during each HAM-TBS experiment is important and enables quality checks for each step of the protocol.

Additionally, reaching 1000× coverage is an important step to provide high resolution on methylation changes [8]. However, accurate quantification and pooling of many amplicons across multiple samples while reaching sufficient coverage of all regions has limitations. In theory, even though the MiSeq can handle a much higher loading factor (amplicons × samples) of almost 20,000 (disregarding uneven pooling of libraries, filtering of reads due to low quality or high amounts of PhiX), a maximum of 2500–3000 has proven to be feasible with minimal dropout rates. Assuming multiplexing of 96 samples and 25 amplicons at an average length of 400 bp, a region of approximately 10 kb can be comfortably covered with this approach. Notably, we streamlined the method to handle loading factors >2000 by implementation of Agilent's TapeStation and a pipetting robot for quantification and pooling of amplicons. Besides the throughput, this improves the robustness of the workflow. Our approach is designed to match the specifications of the Illumina MiSeq with its ability to run for 600 cycles resulting in 300 bp-long paired-end reads. This enables full-length coverage of amplicons up to a length of 600 bp. While our approach can be applied to different sequencers, such as the Illumina HiSeq for example, it would be necessary to design shorter amplicons due to the current limits of the sequencing chemistry. Using another sequencer, it is important to mention the index hopping phenomenon on the Illumina platforms [15]. It is less present on the MiSeq compared to other machines with pattern flow cells as our data show consistent levels of methylation close to 0% across all in vitro unmethylated control samples indicating no issue with

this specific bias. Nonetheless, it should be kept in mind that approaches like unique dual indexes when available or Illumina's Free Adapter Blocking Reagent are recommendable and gain importance, especially when using a different Illumina sequencer.

In the past years, only few TBS methods have been developed [8–10] with different methodological foci. Thus far, Bernstein et al. [10] allows a panel of 48 indices, while the approach by Chen et al. [9] could allow for a multiplexing rate of 1536 samples due to custom-made barcodes, but in practice only 478 have been used to date. In the latter method, the high multiplexing capacity comes at the cost of an additional PCR step potentially introducing additional bias. Moreover, increasing the number of samples needs to be weighed against the size of the target region in order to ensure sufficient coverage. We identified 1000× coverage as an optimal cutoff in terms of accuracy and cost in agreement with a publication by Masser et al. [8]. In the above-described study by Chen et al. [9], 100× was used as minimum cutoff. Based on our in silico analysis (Fig. 4a), this would lead to less accurate quantification of methylation levels. Besides the number of samples that can be processed, the size of the region of interest is also an important factor to be considered. The method by Masser et al. [8] has been applied to 2 amplicons (233 and 320 bp), while Chen et al. enable the assessment of larger loci around 10 kb—comparable to our HAM-TBS approach. Lastly, amplification-based library preparation methods have been adapted by most TBS approaches. At this point, HAM-TBS utilizes a PCR-free library preparation to avoid adding amplification biases.

Finally, using the optimized HAM-TBS workflow, we designed a panel comprising 29 amplicons to accurately assess methylation within the *FKBP5* locus using HAM-TBS. This panel covers ~9 kb and targets important regulatory regions of the *FKBP5* gene including the TSS, intergenic and proximal enhancers and TAD boundaries including *CTCF* binding sites. The HAM-TBS method and the *FKBP5* panel present valuable tools for epigenetic studies in which a highly accurate assessment of methylation levels is critical such as GxE studies in psychiatric research. It allows cost-efficient quantification of methylation in larger cohorts with optimized hands-on time due to automatization.

Conclusion

The presented method HAM-TBS offers a robust and low-cost method for researchers interested in DNA methylation measurements of specific target regions. In addition, we supply a validated panel of 29 amplicons to assess methylation levels of important regulatory regions

in the *FKBP5* locus, a gene of great interest in the field of psychiatry.

Methods

Generation of in vitro methylated control DNA

All primers designed for bisulfite PCR were first tested on in vitro methylated DNA to assess amplification efficiency and bias. For PCRs within the *FKBP5* gene, an in vitro methylated BAC (RP11-282I23, BACPAC) was used to generate control DNA. For PCRs outside the *FKBP5* locus (PCR_26, PCR_34, PCR_35), genomic DNA extracted from whole blood was amplified using the REPLI-g Mini Kit (QIAGEN GmbH, Hilden, Germany) to generate unmethylated DNA. 100% methylated DNA was achieved using in vitro methylation with M.SssI methyltransferase. After a first incubation (3 µg DNA, 0.5 µl SAM (32 mM), 1 µl M.SssI (20 U/µl, 40 µl NEB buffer 2 [10×], diluted with ddH₂O up to 400 µl) of 4 h at 37°C, 1 µl of M.SssI (20 U/µl) and 1 µl of SAM (32 mM) were added, and a second 4-h incubation was performed. Subsequently, the reaction was purified using the nucleotide removal kit (QIAGEN GmbH, Hilden, Germany). In vitro methylation was repeated with the eluted DNA for a second time. 25, 50 and 75% methylated control DNA was obtained by mixing 0 and 100% DNAs. In vitro methylation of control DNA was checked via pyrosequencing.

Bisulfite treatment of DNA

We used the EZ DNA Methylation Kit (Zymo Research, Irvine, CA) in column and plate format depending on the amount of DNA and throughput needed. Between 200 and 500 ng was used as input DNA and processed according to the manufacturer's instructions. DNA was eluted twice in 10 µl elution buffer which recovered over 90% of the input DNA after bisulfite conversion when using the column format. In order to quantify bisulfite treated DNA, we use a spectrophotometer with RNA quantification settings.

Target enrichment and amplicon pooling

The amplification of target locations from converted DNA (20 ng per amplicon) was achieved using the TaKaRa EpiTaq HS Polymerase (Clontech, Mountain View, CA; final concentration: 0.025 U/l), bisulfite-specific primers (final concentration of each primer: 0.4 M) and a touchdown cycling protocol with 49 cycles [for more details (see table in Additional file 5 and section HAM-TBS *FKBP5* panel). The amplicons of all PCR reactions were quantified using the Agilent 2200 TapeStation (Agilent Technologies, Waldbronn, Germany) and equimolar pooled with the Hamilton pipetting robot. After speed-vacuum and resuspension in 50 µl, a double-size selection was applied using Agencourt AMPure XP beads

(Beckman Coulter GmbH, Krefeld, Germany) to remove excess of primers and genomic DNA.

Control samples

For every TBS run, we included three different controls. First, up to three water controls in order to monitor cross-contamination with DNA and detect if the plate was accidentally rotated. Second, an unmethylated control DNA as a positive control and to detect failed steps throughout the workflow. And third, the H19 locus which is an imprinted region and presents with methylation levels ~50% as a positive control for bisulfite conversion in genomic DNA and detect outliers in patient samples. An amplicon located at this locus is incorporated in the *FKBP5* panel.

Library preparation and sequencing

For library generation, Illumina TruSeq DNA PCR-Free HT Library Prep Kit (Illumina, San Diego, CA) was used according to the manufacturer's standard protocol and obtained high-quality libraries using 500 ng of starting material (during optimization, input amounts as low as 100 ng were tested and showed no loss of quality on the QC level). Qubit 1.0 (Thermo Fisher Scientific Inc., Schwerte, Germany) was used for quantification, Agilent's 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany) for quality assessment and Kapa HIFI Library quantification kit (Kapa Biosystems Inc., Wilmington, MA) for final quantification before pooling. Libraries were pooled equimolarly. Sequencing of the libraries was performed on an Illumina MiSeq using Reagent Kit v3 (Illumina, San Diego, CA; 600 cycles) in paired-end mode, with 30% PhiX added.

Sequencing data processing

First, read quality was verified using FastQC [16]. Adapter sequences were trimmed using cutadapt v.1.9.1 [17]. For alignment to a restricted reference of hg19 based on the PCR locations, Bismark v.0.15.0 [18] was used. Due to the 600-cycle sequencing chemistry, PCRs shorter than 600 bp produce overlapping paired-end reads. Using an in-house developed Perl script, we trimmed low-quality overlapping ends. Quantification of methylation levels in CpG and CHH context was performed using the R package methylKit [19] with a minimum quality score of 20. The methylation calls were subjected to 3 quality control steps. First, we considered CHH levels for each sample and excluded samples if the conversion was less than 95% efficient. Second, we filtered PCR artefacts introduced by PCR amplification errors giving rise to CpG sites in some reads. As we do not restrict the analysis to known CpG sites, every read indicating the presence of a CpG will be considered and the information extracted. These

artefacts mostly present at very low levels of coverage and 0 or 100% methylation. Lastly, according to our coverage cutoff, we excluded CpG sites supported by less than 1000 reads. Subsequent analysis comparing methylation levels from the conditions C1, C2 and C3 as well as data from pyrosequencing was performed in R.

Coverage considerations

When performing a sequencing experiment, one will usually sequence part of the generated library and quantify the methylation levels on this basis rather than sequence the whole library to see the true level within. Therefore, each sequencing experiment corresponds to drawing a random subset of a certain size (sequencing depth) of the whole library and can be viewed as a subsampling problem. Depending on the sequencing depth, this will yield a different level of accuracy of the methylation levels. We created a dataset simulating CpGs methylated at levels from 0 to 100% supported by 100,000 “fragments” each. Therefore, e.g., for 10% methylation level, a set 10,000 Cs and 90,000 Ts was created. Accordingly, sets for 0–100% methylation were created. Using a bootstrapping approach, we drew 1000 random subsets of varying sequencing coverage (100, 200, 400, ..., 2000, 3000, 4000, 5000) from each set representing a certain level of methylation and the standard deviation (SD) was calculated. As a proxy for the increase in accuracy versus increase in sequencing depth (costs), the combined SD was divided by the sequencing depth. Of note, this is in concordance with results from the same analysis on highly covered amplicon data from our laboratory (data not shown).

Pyrosequencing

Methylation analysis by pyrosequencing of 5 CpGs covered within PCR_5 (CpG 35607969, CpG 35608022) and PCR_11 (CpG 35690280, CpG 35690318, CpG 35690365) was performed in triplicates on BAC control DNA. Bisulfite conversion of in vitro methylated control DNA was applied as described above. Target enrichment by PCR was achieved with a biotinylated reverse primer but otherwise performed as described above. Pre-treatment of PCR amplicons was facilitated with the PyroMark Q96 Vacuum Workstation (QIAGEN GmbH, Hilden, Germany). Sequencing of *FKBP5* CpGs was performed on a PyroMark Q96 ID system using PyroMark Gold Q96 reagents (QIAGEN GmbH, Hilden, Germany) and sequencing primers according to Klengel et al. [12]: P4 S1 (TTT GGAGTAGTAGGTTAAA) GRE3 S1 MPI (GGGAAT TATGAGGTTG). The PyroMark Q96 ID Software 2.5

(QIAGEN GmbH, Hilden, Germany) was used for data analyses.

HAM-TBS *FKBP5* panel

We designed 29 primer pairs (see table in Additional file 5) using BiSearch [20, 21] targeting the *FKBP5* locus. Initially, 32 PCRs were included, but 3 PCRs were not selected for the panel due to QC failure. The excluded amplicons showed nonlinear amplification due to an elevated GC content in the region. Positions of amplicons covering glucocorticoid response elements (GREs) were selected from Klengel et al. [12] and the *GR* ChIP-Seq from the ENCODE project [22]. Amplicons covering *CTCF* binding sites were selected using HI-C peaks [23], *CTCF*-ChIA-Pet interactions from a lymphoblastoid cell line (GM12878, Tang et al. [24]) and *CTCF* ChIP-Seq information from the ENCODE project [22]. Lastly, amplicons located near the TSS were included in the panel. Only primers without CpGs in their sequence were chosen, with the exception of 2 amplicons close to the TSS where this could not be avoided due to the high CpG content of the region. The selected amplicons ranged from 200 to 450 bp in length.

Additional files

Additional file 1. A table containing the genomic coordinates of a smaller panel of 11 amplicons located within the *FKBP5* locus. These amplicons were used to assess the source of potential biases as well as variability between replicates.

Additional file 2. A figure displaying the bias assessment for all amplicons comprising the *FKBP5* HAM-TBS panel.

Additional file 3. A figure displaying the QC statistics of a HAM-TBS experiment with 95 samples using the *FKBP5* panel and methylation levels of the H19 locus.

Additional file 4. A figure displaying the relative and absolute costs for a HAM-TBS experiment with 96 samples.

Additional file 5. A table containing the genomic coordinates of the *FKBP5* HAM-TBS panel including primer sequences and cycling conditions.

Authors' contributions

TW, SR, NP, EBB contributed to experimental design. TW, SS, MK performed wet lab work. SR performed the data analyses. SR, TW, NP, EBB prepared the manuscript. All authors read and approved the final manuscript.

Author details

¹ Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Kraepelinstr. 2-10, 80804 Munich, Germany. ² Department of Psychiatry and Behavioral Sciences, Emory University Medical School, 12 Executive Park Dr NE #200, Atlanta, GA 30329, USA. ³ Faculty of Health Sciences, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada. ⁴ Healthy Starts Theme, BC Children's Hospital Research Institute, 938 West 28th Avenue, Vancouver, BC V5Z 4H4, Canada.

Acknowledgements

The authors would like to thank Jessica Keverne for professional English editing and formatting and Monika Rex-Haffner for her help regarding the sequencing of the libraries and Stoyo Karamihalev for graphics support.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Funding

This study was funded by the BMBF Grant Berlin-LCS (FKZ 01KR1301B) to EB and an ERC starting Grant (GxE molmech, Grant 281338) within the FP7 funding scheme of the EU to EB and fellowship from Canadian Institute of Health Research (CIHR) to NP.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 4 April 2018 Accepted: 28 June 2018

Published online: 04 July 2018

References

- Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev.* 2002;16:6–21.
- Ehrlich M, Wang RY. 5-methylcytosine in eukaryotic DNA. *Science.* 1981;212:1350–7.
- Portela A, Esteller M. Epigenetic modifications and human disease. *Nat Biotechnol.* 2010;28:1057–68.
- Weng YL, An R, Shin J, Song H, Ming GL. DNA modifications and neurological disorders. *Neurotherapeutics.* 2013;10:556–67.
- Lam LL, Emberly E, Fraser HB, Neumann SM, Chen E, Miller GE, et al. Factors underlying variable DNA methylation in a human community cohort. *Proc Natl Acad Sci USA.* 2012;109(2):17253–60.
- Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, et al. DNA methylation in Newborns and Maternal Smoking in Pregnancy: genome-wide Consortium Meta-analysis. *Am J Hum Genet.* 2016;98:680–96.
- Ziller MJ, Hansen KD, Meissner A, Aryee MJ. Coverage recommendations for methylation analysis by whole-genome bisulfite sequencing. *Nat Methods.* 2015;12:230–2.
- Masser DR, Berg AS, Freeman WM. Focused, high accuracy 5-methylcytosine quantitation with base resolution by benchtop next-generation sequencing. *Epigenet Chromatin.* 2013;6:33.
- Masser DR, Stanford DR, Freeman WM. Targeted DNA methylation analysis by next-generation sequencing. *J Vis Exp.* 2015. <https://doi.org/10.3791/52488>.
- Bernstein DL, Kameswaran V, Le Lay JE, Sheaffer KL, Kaestner KH. The BisPCR(2) method for targeted bisulfite sequencing. *Epigenet Chromatin.* 2015;8:27.
- Chen GG, Gross JA, Lutz PE, Vaillancourt K, Maussion G, Bramouille A, et al. Medium throughput bisulfite sequencing for accurate detection of 5-methylcytosine and 5-hydroxymethylcytosine. *BMC Genom.* 2017;18:96.
- Zannas AS, Wiechmann T, Gassen NC, Binder EB. Gene-stress-epigenetic regulation of FKBP5: clinical and translational implications. *Neuropsychopharmacology.* 2016;41:261–74.
- Klengel T, Mehta D, Anacker C, Rex-Haffner M, Pruessner JC, Pariante CM, et al. Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nat Neurosci.* 2013;16:33–41.
- Huang RC, Galati JC, Burrows S, Beilin LJ, Li X, Pennell CE, et al. DNA methylation of the IGF2/H19 imprinting control region and adiposity distribution in young adults. *Clin Epigenet.* 2012;4(1):21.
- Illumina, Inc. Effects of index misassignment on multiplexing and downstream analysis. <https://www.illumina.com/content/dam/illumina-marketing/documents/products/whitepapers/index-hopping-white-paper-770-2017-004.pdf>.
- Andrews SR. FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>. 2010.
- Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 2011;17:10–2.
- Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics.* 2011;27:1571–2.
- Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* 2012;13:R87.
- Aranyi T, Váradi A, Simon I, Tusnády GE. The BiSearch web server. *BMC Bioinf.* 2006;7:431.
- Tusnády GE, Simon I, Váradi A, Aranyi T. BiSearch: primer-design and search tool for PCR on bisulfite-treated genomes. *Nucleic Acids Res.* 2005;33:e9.
- ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012;489:57–74.
- Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell.* 2015;159:1665–80.
- Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, Szalaj P, et al. CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell.* 2015;163:1611–27.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



5.2 Paper II - Dynamic Changes in DNA Methylation Occur during the First Year of Life in Preterm Infants

Piyasena, C., J. Cartier, N. Provencal, **T. Wiechmann**, B. Khulan, R. Sunderasan, G. Menon, J. R. Seckl, R. M. Reynolds, E. B. Binder and A. J. Drake (2016).

Originally published in:

Front Endocrinol (Lausanne). 2016 Dec 15;7:158. doi: 10.3389/fendo.2016.00158.
eCollection 2016. PMID: 28018293

Additional supplementary material:

Additional supplementary figures and tables are freely available for download at the publisher's website (<https://doi.org/10.3389/fendo.2018.000>)



Dynamic Changes in DNA Methylation Occur during the First Year of Life in Preterm Infants

Chinthika Piyasena^{1,2}, Jessy Cartier¹, Nadine Provençal³, Tobias Wiechmann³, Batbayar Khulan¹, Raju Sunderesan², Gopi Menon², Jonathan R. Seckl¹, Rebecca M. Reynolds¹, Elisabeth B. Binder³ and Amanda J. Drake^{1*}

¹British Heart Foundation Centre for Cardiovascular Science, The Queen's Medical Research Institute, University of Edinburgh, Edinburgh, UK, ²Neonatal Unit, Simpson Centre for Reproductive Health, Royal Infirmary of Edinburgh, Edinburgh, UK, ³Department of Translational Research in Psychiatry, Max-Planck Institute of Psychiatry, Munich, Germany

OPEN ACCESS

Edited by:

Susan Ozanne,
University of Cambridge, UK

Reviewed by:

Kazuhiro Takahashi,
Tohoku University, Japan
Daniel Vaiman,
French Institute of Health and
Medical Research, France

*Correspondence:

Amanda J. Drake
mandy.drake@ed.ac.uk

Specialty section:

This article was submitted to
Neuroendocrine Science,
a section of the journal
Frontiers in Endocrinology

Received: 17 October 2016

Accepted: 29 November 2016

Published: 15 December 2016

Citation:

Piyasena C, Cartier J, Provençal N,
Wiechmann T, Khulan B,
Sunderesan R, Menon G, Seckl JR,
Reynolds RM, Binder EB and
Drake AJ (2016) Dynamic Changes in
DNA Methylation Occur during the
First Year of Life in Preterm Infants.
Front. Endocrinol. 7:158.
doi: 10.3389/fendo.2016.00158

Background: Preterm birth associates with a substantially increased risk of later cardiovascular disease and neurodevelopmental disorders. Understanding underlying mechanisms will facilitate the development of screening and intervention strategies to reduce disease risk. Changes in DNA methylation have been proposed as one mechanism linking the early environment with later disease risk. We tested the hypothesis that preterm birth associates with altered DNA methylation in genes encoding insulin-like growth factor 2 (IGF2) and FK506-binding protein 5 (FKBP5), which appear particularly vulnerable to early life adversity.

Methods: Fifty preterm infants were seen and assessed at birth, term equivalent age, 3 months and 1-year corrected ages; 40 term infants were seen at birth, 3 months and 1 year. Saliva was collected for DNA extraction at birth, term, and 1 year. Pyrosequencing of bisulfite-converted DNA was performed to measure DNA methylation at specific CpG sites within the *IGF2* and *FKBP5* loci.

Results: Weight and head circumference was reduced in preterm infants at all time points. Preterm infants had a higher percentage body fat at term-corrected age, but this difference was not persistent. DNA methylation at the differentially methylated region (DMR) of *IGF2* (*IGF2DMR2*) and *FKBP5* was lower in preterm infants at birth- and term-corrected age compared to term infants at birth. *IGF2DMR2* and *FKBP5* methylation was related to birthweight SD score in preterm infants. Among preterm infants, social deprivation was an independent contributor toward reducing DNA methylation at *IGF2DMR2* at birth- and term-corrected age and maternal smoking was associated with reduced DNA methylation at *FKBP5* at birth. There were no persistent differences in DNA methylation at 1 year of age.

Conclusion: Changes in DNA methylation were identified at key regions of *IGF2/H19* and *FKBP5* in preterm infants in early life. Potential contributing factors include maternal smoking and social deprivation. However, these changes did not persist at 1 year of age and further longitudinal studies are required to determine any associations between altered DNA methylation in the perinatal period of individuals born preterm and their long-term health.

Keywords: prematurity, DNA methylation, IGF2, FKBP5, glucocorticoids

INTRODUCTION

Epidemiological evidence linking low birthweight with an increased risk of cardiovascular disease as well as developmental neuropsychiatric disorders (1) has led to the concept of “early life programming.” This proposes that exposure to adverse conditions during critical stages of early development results in a change in the offspring structural and functional phenotype (2). Preterm birth acts as a profound challenge in early life. There is now substantial evidence that prematurity associates with risk factors for cardiovascular disease in adulthood, including hypertension and insulin resistance (3–5). Furthermore, preterm birth is closely associated with neurodevelopmental disorders including cognitive impairment and autism spectrum disorder (6). This has important implications for public health, since worldwide, 15 million infants are born preterm every year, and survival rates have increased markedly over recent years (7).

Factors acting during intrauterine development, which may be important in mediating programming effects in infants born small at term, include undernutrition and glucocorticoid overexposure. These may be of particular importance in infants born preterm. In addition, preterm infants are vulnerable to these factors acting in early *postnatal* life (4). Following birth, many preterm infants develop a cumulative protein and energy deficit and exhibit early postnatal growth failure (8). Preterm infants are additionally exposed to repeated stressful and often painful procedures during a period of rapid brain maturation, and several studies have shown an impact of these procedures on neurodevelopment and hypothalamic–pituitary–adrenal (HPA) axis activity (9, 10).

Understanding the mechanisms by which prematurity associates with long-term effects on health would facilitate the development of effective screening and intervention strategies. Changes in DNA methylation have been proposed as one mechanism linking early life events and later disease risk (11), and genome-wide profiling has revealed DNA methylation differences between preterm and term infants in early life (12–14). Exposure to an adverse environment in early life has repeatedly been associated with altered DNA methylation at a gene of particular importance for fetal growth: the imprinted gene insulin-like growth factor 2 (*IGF2*). *IGF2* is a key growth factor, particularly in early development. *IGF2* expression is controlled by DNA methylation at a number of differentially methylated regions (DMRs) (15) and altered *IGF2* DNA methylation has been reported following exposure to altered maternal nutrition including severe famine (16, 17).

The early life environment can also impact on the normal functioning of the HPA axis, with implications for neurodevelopment. Exposure to an adverse environment pre- or postnatally has been associated with altered DNA methylation at a number of genes important in determining HPA axis function, including the glucocorticoid receptor (*GR*) and 11 β -hydroxysteroid dehydrogenase type 2 (*11 β -HSD2*). Differences in DNA methylation at *GR* and *11 β -HSD2* in placenta have been reported in association with infant behavioral development (18); however, we have previously reported that DNA methylation at *GR* and *11 β -HSD2* is extremely low in individuals exposed to an adverse early life environment and is, therefore, unlikely to impact on

gene expression (17). FK506-binding protein 5 (*FKBP5*) encodes a co-chaperone of *GR* and is induced following stress exposure through *GR* binding to specific genomic response elements; in turn, *FKBP5* protein binds to the *GR* complex, reducing its affinity for cortisol and decreasing nuclear translocation (19). Thus, it is an important component of the stress response. A functional polymorphism in *FKBP5* intron 2 alters mRNA and protein induction following *GR* activation (20), such that the allele associated with stronger *FKBP5* mRNA induction associates with *GR* resistance and an increased risk of a number of psychiatric disorders following childhood trauma (20). Exposure to childhood trauma leads to allele-specific epigenetic changes, with a *GR*-binding-induced decrease in DNA methylation within a functional glucocorticoid response element in intron 7, specifically, in carriers of the risk allele (20). Further, DNA methylation of *FKBP5* in placenta associates with infant arousal scores (21).

In this study, we tested the hypothesis that preterm birth, a profound stressor in early life, associates with altered DNA methylation at the candidate loci *IGF2* and *FKBP5*, which may be particularly vulnerable to early life adversity, and examined whether any changes were persistent over the first year of life.

MATERIALS AND METHODS

Cohort

Fifty preterm (<32 weeks gestation) and 40 term (37–42 weeks gestation) infants were recruited within the first week of life from the Simpson Center for Reproductive Health, Edinburgh, UK, with informed written parental consent. Ethical approval was obtained from the South East Scotland Research Ethics Committee (Reference 11/AL/0329). NHS management approval was obtained (Lothian R&D Project number 2011/R/NE/03). Perinatal samples were collected under the Edinburgh Reproductive Tissue BioBank (ERTBB) (West of Scotland Research Ethics Service Reference 09/S0704/3). All parents gave written informed consent and all studies were performed in accordance with the declaration of Helsinki. Infant samples were collected under the framework of the ERTBB following an amendment to ethical approval (Reference AM07/1). Demographic details were obtained during clinic visits and from hospital records. All of the preterm infants were admitted to the neonatal unit; of these, six infants died. In three of these infants, saliva for buccal cells was not collected after birth due to clinical instability and DNA was of poor quality in a fourth infant. Five of the term infants were also admitted to the neonatal unit for a short period (respiratory distress syndrome, weight loss, and hemolysis from Rhesus isoimmunization) but none required follow-up. Preterm infants were seen at birth, at term-corrected age, and at 3 months and 1-year corrected ages; term infants were seen at birth, 3 months, and 1 year. All visits occurred in the afternoon, supervised by one researcher (Chinthika Piyasena).

Growth and Body Composition

Weight, length, and occipitofrontal head circumference (OFC) was measured by one trained researcher. Percentage body fat mass

was measured by air displacement plethysmography in preterm infants at term-corrected age and 3 months corrected age, and in term infants at birth and 3 months using the PEAPOD Body Composition System (COSMED, Chicago, IL, USA). Skin fold thickness (subscapular and triceps) was measured at 1 year by the same trained researcher. Term infants were measured at a median of 2 days (range 0–8) after birth, at 3 months (13.3 weeks; range 10.4–16.9), and 1 year (52.3 weeks; range 48.1–57.4). Preterm infants were measured at term-corrected age at a median of 40 (range 35 + 0 to 44 + 1)-corrected weeks, at 3 months corrected (13.4 weeks; range 10.3–18.3) and 1 year corrected (53.9 weeks; range 52.1–68.7).

Analysis of DNA Methylation

Saliva was collected using the Oragene DNA (OG-250) kits and saliva sponges CS-1 and extracted using prepIT-2LP (DNA Genotek, Ottawa, ON, Canada). DNA was quantified using the Qubit 2.0 Fluorometer (Life Technologies, Paisley, UK). Five hundred nanograms of DNA were bisulfite converted using the EZ DNA Methylation Gold Kit (Zymo Research Corporation, CA, USA). Pyrosequencing was performed to analyze DNA methylation for DMRs controlling *IGF2* expression: *IGF2* DMR2 (*IGF2DMR2*, $n = 9$ CpGs) and the H19 imprinting control region (*H19ICR*, $n = 8$ CpGs) as previously described (22). Primers were purchased from Invitrogen (Life Technologies, Paisley, UK). DNA was amplified using the AmpliTaq Gold 360 kit (Applied Biosystems, Warrington, UK) and pyrosequencing performed using PyroMark Q24Gold reagents on a PyroMark Q24 Pyrosequencer (Qiagen, Crawley, UK). Data were analyzed using PyroMark Q24 1.0.10. Percentage DNA methylation is expressed as the average across all CpGs in each of the two loci in *IGF2*.

Methylation analysis of three CpGs in *FKBP5* intron 7, two of which are located in consensus GRE motif (CpG 2 and 3), was performed in triplicates using a protocol adapted from Klengel et al. (20). One hundred twenty nanograms genomic DNA was bisulfite converted using the EZ DNA Methylation Kit (Zymo Research Corporation, CA, USA). Bisulfite converted DNA was amplified in a 50 μ l reaction mix (4–10 μ l DNA; each bisulfite specific Primer with a final concentration of 0.2 μ M, *FKBP5*int7_P1_F: GTTGTTTTTTGGAAATTTAAGGTAATTG, and *FKBP5*int7_P1_R_biot: biotin-TCTCTTACCTCCAACACTACTACTAAAA) using the Kapa HIFI Uracil + Hot start Ready Mix (Kapa Biosystems Inc., Wilmington, DE, USA). Cycling conditions of the touchdown PCR were 98°C for 5 min, 2 \times (98°C – 40 s, 62°C – 30 s, 72°C – 60 s), 5 \times (98°C – 40 s, 60°C – 30 s, 72°C – 60 s), 8 \times (98°C – 40 s, 58°C – 30 s, 72°C – 60 s), 34 \times (98°C – 40 s, 56°C – 30 s, 72°C – 60 s), 72°C for 1 min and cooling to 4°C. Pyrosequencing of *FKBP5* CpGs was performed on a PyroMark Q96 ID system (QIAGEN GmbH, Hilden) using PyroMark Q96Gold reagents with the following sequencing primer: *FKBP5*int7_P1_S2: 5'-GTT GATATATAGGAATAAAATAAGA-3' for CpG1 and CpG2 and *FKBP5*int7_P1_S3: 5'-TGGAGTTATAGTGTAGGTTTT-3' for CpG3. PyroMark Q96 ID Software 2.5 (QIAGEN GmbH, Hilden) was used to calculate percentage methylation.

The legends for **Figures 1** and **2** state the number of values/measurements analyzed, excluding samples, which were discarded

because of insufficient DNA and/or poor pyrosequencing quality. Additionally, we were unable to collect saliva from several babies for various reasons, although these babies still underwent anthropometry.

FKBP5 SNP Genotyping

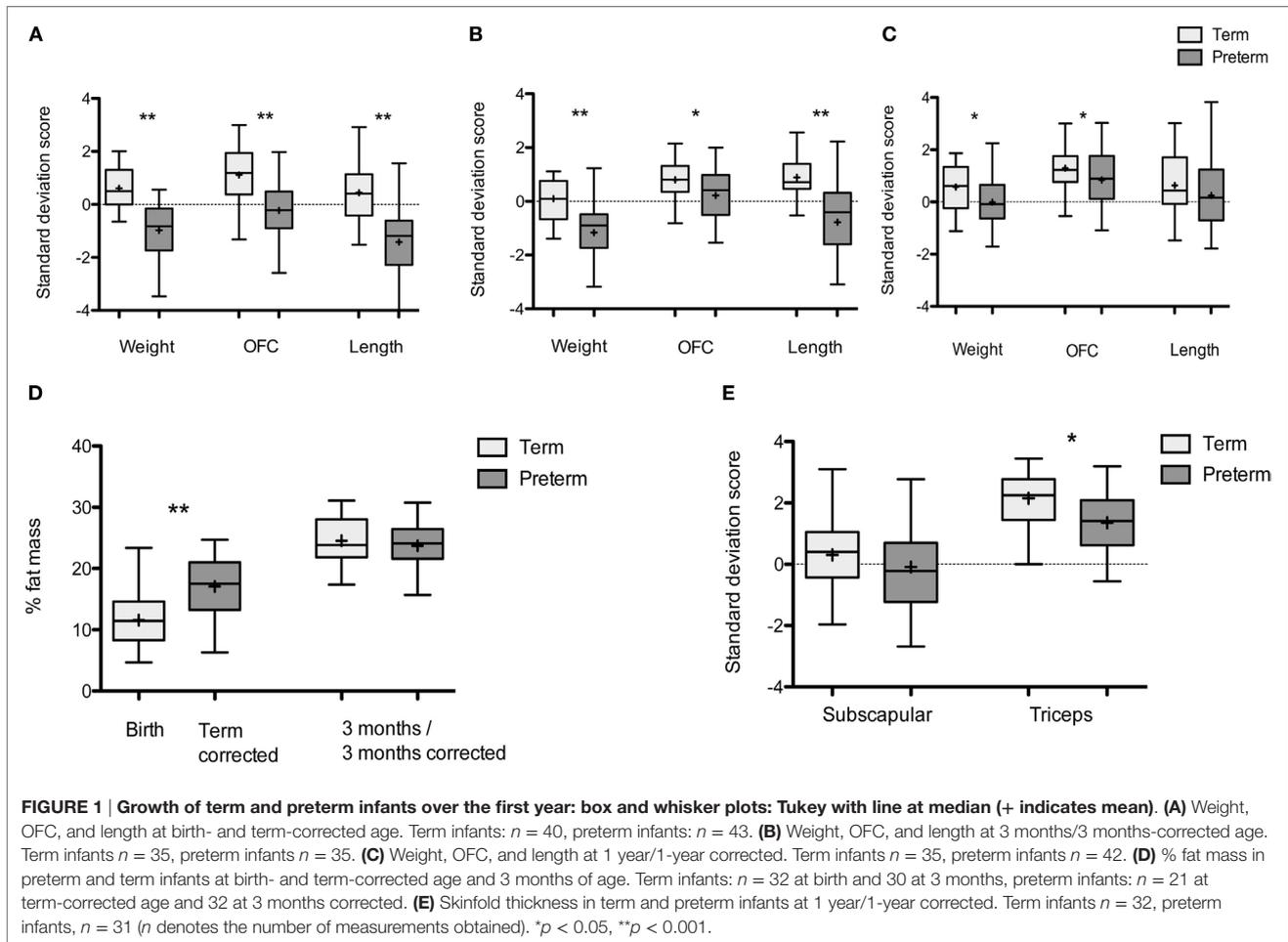
FKBP5 rs1360780 genotyping was performed on a Roche LightCycler 480 System using a TaqMan SNP Genotyping Assay (Applied Biosystems). Thermal cycling conditions were 95°C for 10 min, 45 \times (95°C – 15 s, 60°C – 1 min, 50°C – 30 s). Genotypes were called using endpoint genotyping LightCycler 480 Software version 1.5. Genotypes were in Hardy–Weinberg equilibrium ($p = 0.17$). Participants were divided into protective-genotype (CC) and risk-allele carriers (CT or TT).

Covariates

Covariates that could confound the association or be in the causal pathway were added into the model in a hierarchical manner: maternal smoking, male gender, and social deprivation for *FKBP5* and *IGF2/H19*. Additionally, breast milk at 3 months for *IGF2/H19* at 1 year and risk-allele carriage for *FKBP5* at all time points. Social deprivation was coded as deprivation category (DEPCAT) scores based on the mother's postcode at booking and obtained from the "Carstairs scores for Scottish postcode sectors from the 2001 Census." In this system, scoring is based on the material affluence/deprivation of the area in which a person lives. Postcode sectors are allocated a DEPCAT score, derived from four sets of information: overcrowding, male unemployment, car ownership, and the proportion of people in households in social class 4 or 5 and scores range from 1 to 7 where 7 indicates the worst social deprivation. Maternal smoking was categorized as current smoker, never smoked, former (stopped pre-pregnancy), or former (stopped during pregnancy). Breast milk at 3 months indicates whether or not the infant was receiving any breast milk at 3 months corrected age.

Statistics

Weight, length, OFC, skin fold thickness, and weight gain were adjusted for age and gender by converting to SD scores (SDS) (z -scores) using LMSgrowth, a Microsoft Excel Add-in to access growth references that define the UK-WHO growth charts (23). Demographic, clinical characteristics, and risk-allele carriage between preterm and term infants were compared using independent samples t -testing and chi-square analyses, as appropriate. Multivariate linear regression was used to assess variation in body composition and to test the hypothesis that preterm birth is associated with altered DNA methylation. Outcome variables were percentage DNA methylation at birth, term-corrected age, and 1 year. Unstandardized regression (β) coefficients from these models indicate the change in percentage methylation associated with prematurity and a one-unit change in the other predictors. Paired samples t -testing was used where appropriate. One-way ANOVA with Dunnett's *post hoc t* testing was used to test the effect of maternal smoking on DNA methylation at *FKBP5*. Statistical significance for all analyses was set at $p < 0.05$ (two-tailed).



RESULTS

Preterm and Term Infant Demographics

Characteristics of the cohort are shown in **Table 1**. Birth weight SDS was lower in preterm infants $p < 0.001$, and there were more males in the preterm group, $p = 0.006$. There was no significant difference between the frequency of *FKBP5* risk-allele carriage between the groups ($p = 0.43$). Maternal age, body mass index (BMI), DEPCAT scores, folic acid in first trimester, and maternal smoking were different between the two groups.

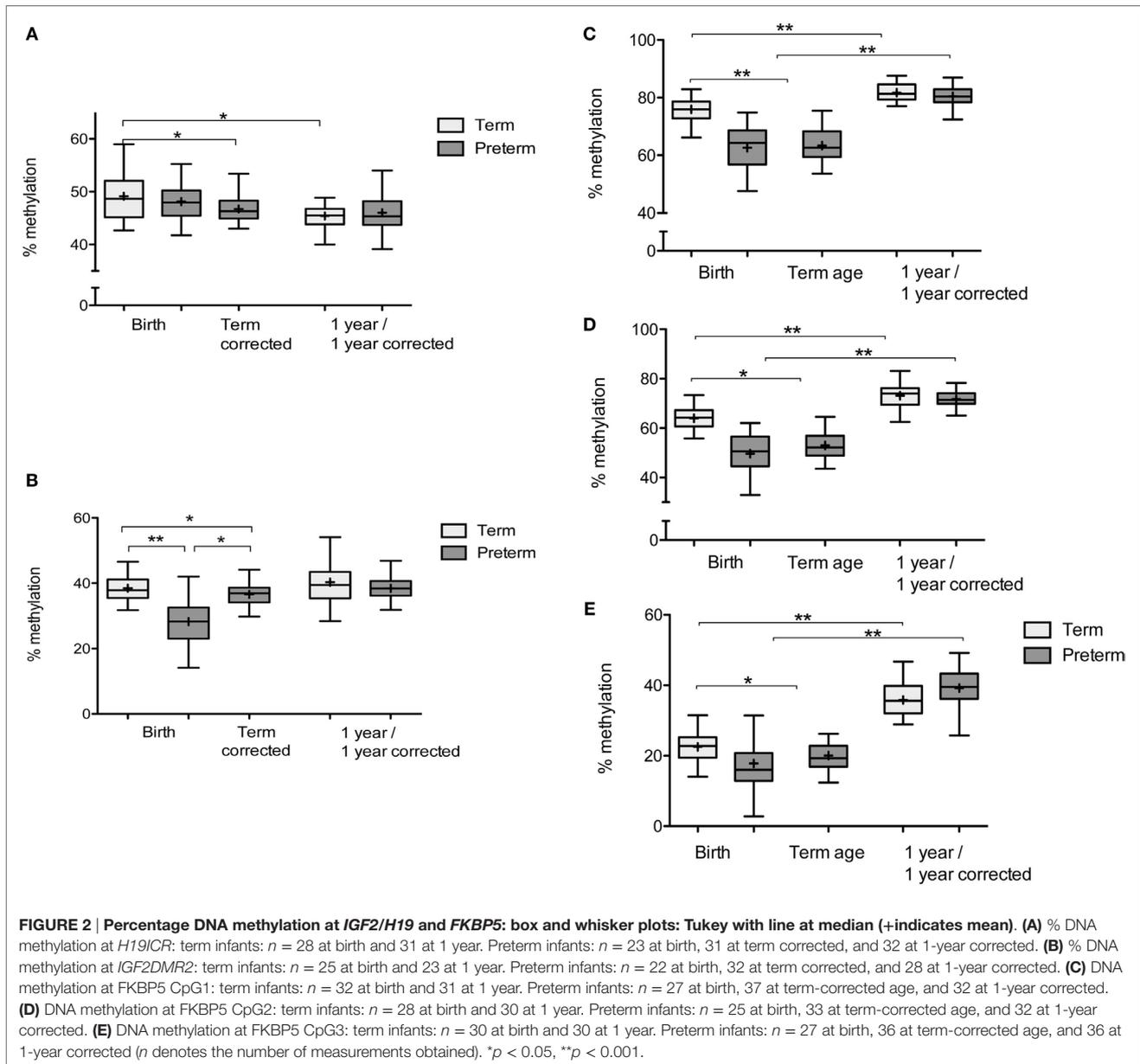
Growth and Body Composition

Weight and OFC was reduced in preterm infants at all time points compared to term infants (**Figures 1A–C**). Preterm infants were shorter than term infants at term-corrected age and 3 months corrected age, but not at 1-year corrected (**Figures 1A–C**). Preterm infants had a higher percentage body fat at term-corrected age when compared with term infants at birth (**Figure 1D**) (mean difference 5.5%, 95% CI [8.0, 3.0], $p < 0.001$) and this remained significant following adjustment for gender ($\beta = 5.7$, 95% CI [3.1, 8.3], $p < 0.001$). The difference in percentage body fat did not persist at 3 months/3 months corrected, including when adjusted for gender and breast milk intake ($\beta = -0.3$, 95% CI [-2.8, 2.2],

$p = 0.82$) (**Figure 1D**). At 1-year corrected, preterm infants had lower triceps skin fold thickness SDS than term infants, adjusted for breast milk exposure at 3 months ($\beta = -0.4$, 95% CI [-1.3, -0.3], $p = 0.008$), but there were no differences in subscapular skin fold thickness (**Figure 1E**).

Prematurity and DNA Methylation

Mean DNA methylation at *IGF2DMR2* was lower in preterm infants at birth compared to term infants at birth and this remained significant in adjusted analyses ($\beta = -11.2$, 95% CI [-15.2, -7.3], $p < 0.001$) (**Figure 2B**). There was no difference in mean DNA methylation at the *H19ICR* between preterm and term infants at birth after adjusted analysis ($\beta = -1.3$, 95% CI [-3.8, 1.2], $p = 0.3$) (**Figure 2A**). In preterm infants at term-corrected age, DNA methylation was reduced at *IGF2DMR2* and *H19ICR* compared to infants born at term, and this remained significant in adjusted analyses ($\beta = -2.8$, 95% CI [-5.0, -0.6], $p = 0.01$ and $\beta = -2.3$, 95% CI [-4.6, -0.1], $p = 0.049$, respectively) (**Figures 2A,B**). However, the significance was attenuated when social deprivation was added into the model for *IGF2DMR2* ($\beta = -2.1$, 95% CI [-4.3, 0.05], $p = 0.055$). At 1 year of age, there was no difference in mean DNA methylation in adjusted analyses at either *IGF2DMR2* ($\beta = -0.3$, 95%



CI $[-4.0, 3.3]$, $p = 0.86$) or *H19ICR* ($\beta = 0.8$, 95% CI $[-1.5, 3.0]$, $p = 0.49$) (Figures 2A,B).

The presence of the *FKBP5* risk allele was associated with higher DNA methylation at *FKBP5* CpG3 at birth ($\beta = 4.0$, 95% CI $[0.3, 7.6]$, $p = 0.03$) and term-corrected age ($\beta = 2.5$, 95% CI $[0.4, 4.5]$, $p = 0.02$) across all infants. DNA methylation at CpG1, CpG2, and CpG3 was lower in preterm infants at birth and at term-corrected age when compared to term infants at birth, including after adjusted analysis (Table 2). There was an increase in DNA methylation at all three CpGs in both groups during the first year (Figures 2C–E) and, by 1 year of age, there were no persistent differences in mean methylation at any CpG between preterm and term infants (Table 2). The presence/absence of

the *FKBP5* risk allele as a covariate did not alter the findings at 1 year and there was no moderation of the relationship between prematurity and DNA methylation by the presence of the risk allele at any time.

Body Composition and DNA Methylation

IGF2DMR2 methylation was highly significantly related to birthweight SDS in preterm ($R = 0.7$, $p < 0.001$), but not term infants. DNA methylation across *FKBP5* was also positively associated with birthweight SDS ($\beta = 0.3$, 95% CI $[0.3, 5.4]$, $p = 0.03$) when adjusted for prematurity, but only in the preterm infants when analyzed separately ($R = 0.4$, $p = 0.04$). There were no other significant relationships between DNA methylation at *IGF2DMR2*,

TABLE 1 | Characteristics of the study participants.

	Term	Preterm	p-Value
Infant characteristics			
Gestation at birth, weeks	40.2 ± 1.1	28.5 ± 2.1	<0.001
Birth weight, g	3,649 ± 517	1,136 ± 350	<0.001
Birth weight SDS	0.44 ± 1.0	-0.44 ± 0.9	<0.001
Male, n (%)	15 (38)	35 (70)	0.006
Death, n (%)	0	6 (12)	
Bronchopulmonary dysplasia, n (%)	0	14 (31)	
Laser for retinopathy of prematurity, n (%)	0	5 (11)	
Necrotizing enterocolitis, n (%)	0	6 (13.3)	
Intraventricular hemorrhage, n (%)	0	3 (6)	
Periventricular leucomalacia, n (%)	0	2 (4.2)	
Late onset sepsis, n (%)	0	17 (37)	
TPN duration, days	0	17.8 ± 21.8	
Any breast milk at 3 months, n (%)	28 (75.7)	9 (22.5)	<0.001
<i>FKBP5</i> risk allele carriage, n (%)	21 (52.5)	28 (60.9)	0.43
Maternal characteristics			
Age, years	35.2 ± 4.6	31.3 ± 6.2	0.001
Body mass index at booking, kg/m ²	24.3 ± 3.1	27.2 ± 6.9	0.016
High DEPCAT score, n (%)	11 (27.5)	27 (61.4)	0.002
Caucasian ethnicity, n (%)	40 (100)	41 (93.8)	0.09
Smoking, n (%)			
Current	0	10 (22.7)	
Former – stopped during pregnancy	2 (5)	4 (9.1)	0.5
Former – stopped pre-pregnancy	12 (30)	11 (25)	0.61
Never	26 (65)	18 (40.9)	0.03
Primiparity, n (%)	21 (52.5)	29 (65.9)	0.69
Folic acid during first trimester, n (%)	40 (100)	38 (86.3)	<0.001
Assisted reproduction, n (%)	0	6 (13.6)	
Multiple pregnancy, n (%)	0	9 (20.4)	
Hypertension or preeclampsia, n (%)	0	10 (22.7)	
Diabetes during pregnancy, n (%)	0	3 (6.8)	
Antenatal steroids, n (%)			
None	40 (100)	3 (6.8)	<0.001
Incomplete course	–	11 (25)	
Complete course	–	30 (68.2)	
Antenatal magnesium sulfate, n (%)	0	10 (22.7)	
Cesarean section, n (%)	28 (70)	23 (52.2)	0.1
Labor, n (%)	18 (37.5)	27 (61.4)	0.13
Age of partner	35.3 ± 5.2	33.1 ± 6.9	0.1

Plus-minus values are means ± SD. n = 40 full term and 50 preterm infants. n = 44 for mothers of preterm infants. Bronchopulmonary dysplasia (BPD) is defined as need for respiratory support and/or supplemental oxygen at 36 weeks' postmenstrual age to maintain oxygen saturations of 90% or more. A diagnosis of necrotizing enterocolitis was applied when cases achieved Bell stage 2 or greater. Intraventricular hemorrhage is defined as only grades III or IV events. The definition of late onset sepsis is taken from the Vermont Oxford Network Manual of Operations, release 16.3. Five preterm infants died before 36 weeks' postmenstrual age (mean 24.9 weeks' gestational age at birth, mean 692 g birthweight). A further preterm infant died at term-corrected age from severe BPD (26-weeks' gestation at birth weighing 1,000 g). In three of these infants, saliva for buccal cells was not collected after birth due to clinical instability and DNA was of poor quality in a fourth infant. Three term infants and six preterm infants were not seen at 3 months. High DEPCAT was regarded as scores 4–7. SDS, SD score; DEPCAT, deprivation category.

H19ICR, or *FKBP5* and weight or measures of body composition in either preterm or term infants at any other time point (Table 3).

Social Deprivation, Maternal Smoking, and DNA Methylation

Social deprivation was an independent contributor toward reducing DNA methylation at *IGF2DMR2* in preterm infants at birth

($\beta = -1.6$, 95% CI [-2.8, -0.3], $p = 0.02$) and at term-corrected age ($\beta = -0.9$, 95% CI [-1.5, -0.2], $p = 0.02$) although this did not persist at 1-year corrected age ($p = 0.07$). Social deprivation was also independently associated with a reduction in DNA methylation at *FKBP5* CpG2 in preterm infants at term-corrected age ($\beta = -1.5$, 95% CI [-2.9, -0.02], $p = 0.047$). Finally, maternal smoking was independently associated with a marked reduction in DNA methylation in preterm infants at *FKBP5* CpG1 ($\beta = -2.6$, 95% CI [-4.4, -0.7], $p = 0.008$) and CpG2 ($\beta = -2.1$, 95% CI [-4.1, -0.03], $p = 0.047$) at birth, compared to term infants. Additionally, maternal smoking was associated with a reduction in DNA methylation at birth in both groups at these loci (Table 4).

DISCUSSION

Preterm infants demonstrated a growth trajectory comparable to that reported in previous studies, such that they were lighter than term infants during the first year of life (24, 25). Although preterm infants had increased percentage fat mass at term-corrected age, there was no persisting difference at 3 months. Previous studies have reported increased and/or altered fat distribution in preterm infants at term-corrected age (26–28), with some reporting that these differences had resolved by 3 months (28). Despite this apparent resolution of excess adiposity in early life, preterm infants do show abnormalities of body composition in young adulthood, with higher adiposity, ectopic lipid deposition, and increased intra-abdominal fat (29).

DNA methylation at *IGF2DMR2* and *FKBP5* intron 7 was markedly lower in preterm infants at birth in comparison to term-born infants, and this was still the case at term-corrected age. Although we did not test the functional consequences of these alterations in DNA methylation, reduced *IGF2DMR2* DNA methylation would be predicted to reduce *IGF2* expression with potential implications for early growth. Reduced methylation of *FKBP5* intron 7 CpGs is associated with higher induction of *FKBP5* by GR activation, leading to increased GR resistance (20); however, whether this might play a role in the HPA axis dysregulation seen in preterm infants (9) is unclear. Social deprivation was an independent predictor of reduced methylation at *IGF2DMR2* at birth and term-corrected age. Inter-related factors such as smoking that accompany socioeconomic deprivation can impact on DNA methylation in adults (30), including at *FKBP5* (31). Recent studies in children and adults exposed to cigarette smoke *in utero* have reported alterations in global and site-specific DNA methylation (32–35), and our results showing that maternal smoking was independently associated with a marked reduction in DNA methylation at *FKBP5* at birth suggest that this extends to effects on *FKBP5*.

There were no persistent differences in DNA methylation at 1 year of age at the DMRs of *IGF2* and at *FKBP5*. Our data support an epigenome-wide association study (EWAS), which demonstrated that although there were many differences in DNA methylation between preterm and term babies at birth, these had largely resolved by 18 years of age. However, DNA methylation differences did persist at a subset of CpGs (12). This

TABLE 2 | DNA methylation at FKBP5 in preterm and term infants.

A	CpG1		CpG2		CpG3	
	β	<i>p</i>	β	<i>p</i>	β	<i>p</i>
Preterm infants at birth vs. term infants at birth	-12.1 [-15.6, -8.8]	<0.001	-12.9 [-16.9, -9.0]	<0.001	-5.2 [-9.3, -1.2]	0.01
Preterm infants at term-corrected age vs. term infants at birth	-11.7 [-14.6, -8.7]	<0.001	-7 [-11.5, -2.5]	0.003	-3 [-5.3, -0.7]	0.01
Preterm infants at 1-year corrected vs. term infants at 1 year	-1.9 [-4.0, 0.2]	0.07	-2 [-4.4, 0.4]	0.1	3.1 [0.8, 6.0]	0.05

β values represent unstandardized regression coefficients adjusting for male gender, maternal smoking, social deprivation score, and risk-allele carriage. Values in square brackets represent 95% confidence intervals.

TABLE 3 | Correlations between percentage DNA methylation at IGF2DMR2, H19ICR, and FKBP5 with weight SDS and percentage body fat.

	IGF2DMR		H19ICR		FKBP5 CpG1		FKBP5 CpG2		FKBP5 CpG3		Intron average	
	<i>R</i>	<i>p</i>	<i>R</i>	<i>p</i>	<i>R</i>	<i>p</i>	<i>R</i>	<i>p</i>	<i>R</i>	<i>p</i>	<i>R</i>	<i>p</i>
% DNA methylation vs. weight SDS												
Term infants at birth vs. birthweight SDS	0	0.98	-0.2	0.44	0	0.83	0.2	0.43	-0.2	0.22	0.2	0.37
Preterm infants at birth vs. birthweight SDS	0.7	<0.001	0.2	0.41	0.6	<0.001	0.5	0.01	0.4	0.04	0.4	0.04
Preterm infants at term age vs. weight SDS	0	0.96	0.2	0.37	0.2	0.34	-0.3	0.16	-0.1	0.53	0.1	0.73
Term infants at 1 year vs. weight SDS	-0.1	0.72	0.2	0.28	0	0.89	0.4	0.05	0.4	0.06	0.1	0.79
Preterm infants at 1 year corrected vs. weight SDS	-0.06	0.75	0.27	0.14	-0.2	0.29	-0.1	0.51	-0.2	0.22	0	0.86
% DNA methylation vs. % body fat												
Term infants at birth vs. % fat	0	0.94	-0.2	0.29	0	0.9	0	0.98	-0.2	0.4	0.2	0.27
Preterm infants at term age vs. % fat	-0.2	0.51	-0.2	0.62	0.3	0.17	0.2	0.47	0.2	0.48	-0.1	0.81

TABLE 4 | DNA methylation at FKBP5 according to maternal smoking status.

Smoking status	CpG1			CpG2			CpG3		
	Mean	SEM	<i>p</i> -Value	Mean	SEM	<i>p</i> -Value	Mean	SEM	<i>p</i> -Value
Birth (term and preterm infants)									
Never	73.6	1.1		59.9	1.2		20.2	1.0	
Former stopped pre-pregnancy	66.1	2.4	0.007	56.4	2.7	0.49	21.4	2.3	0.91
Former stopped during pregnancy	68.8	3.0	0.47	57.0	5.3	0.90	20.8	3.4	1.00
Current	57.1	4.7	0.001	42.5	3.9	0.002	16.9	2.3	0.71
Term age (preterm infants)									
Never	62.9	1.4		53.1	1.6		20.4	1.0	
Former stopped pre-pregnancy	65.9	1.7	0.51	54.1	1.6	0.96	20.6	1.7	1.00
Former stopped during pregnancy	62.2	5.4	1.00	55.1	1.4	0.95	19.2	1.9	0.97
Current	61.6	2.0	0.95	50.6	2.1	0.71	18.1	1.0	0.66
1 year (term and preterm infants)									
Never	80.6	0.5		72.6	0.8		37.8	0.8	
Former stopped pre-pregnancy	81.4	1.1	0.82	72.1	0.9	0.97	37.2	1.5	0.97
Former stopped during pregnancy	83.6	1.4	0.23	72.5	1.6	1.00	36.4	2.6	0.91
Current	80.5	1.9	1.00	72.5	1.6	1.00	40.0	3.1	0.79

Smoking status is recorded as never (regarded as the control group for one-way ANOVA and Dunnett's post hoc *t* testing), former smokers who stopped before pregnancy, former smokers who stopped during pregnancy, and current (smokers who continued throughout pregnancy).

EWAS did not identify changes at IGF2 or FKBP5, which may reflect that it was small ($n = 12$ per group) and studied DNA methylation changes in blood spots, rather than buccal DNA. Our findings are also in agreement with longitudinal EWASs, which show that DNA methylation undergoes developmental changes during childhood (36). Notably, for FKBP5, DNA methylation increased between birth and 1 year in both term and preterm infants, and DNA methylation at CpG3 was still significantly lower at 1 year of age in comparison to levels described in adulthood (20), suggesting that there are ongoing

changes in DNA methylation at this locus through childhood. Since studies clearly show that DNA methylation changes through infancy and childhood, it may be that the lower levels of DNA methylation in preterm infants at birth is a normal finding for infants at this gestation. For obvious reasons, we are unable to compare DNA methylation in saliva in infants of comparable gestation, who remained *in utero*. There were some persistent differences between term infants and preterm infants at term-corrected age, and we suggest that this may reflect differences between the intrauterine maturation in term-born infants and

factors acting during the extra-uterine period to which preterm infants are exposed.

Alternative/additional explanations for the differences in DNA methylation between preterm and term infants in very early life could include the altered nutritional state of infants born preterm (37, 38); and/or glucocorticoid overexposure, since preterm infants experience both *in utero* exposure to synthetic glucocorticoids and significant early postnatal “stress.” *In vivo* and *in vitro* studies have shown that glucocorticoid exposure leads to stable DNA demethylation at these specific sites within the *FKBP5* locus (20, 39), and several studies have now shown that exposure to trauma during childhood associates with allele-specific demethylation at *FKBP5* in adulthood (20, 40). Although childhood abuse impacts on DNA methylation in carriers of the risk allele (20), we found no additional effects of the presence or absence of the risk allele. This is in agreement with a lack of effect of the risk allele on DNA methylation patterns following severe parental trauma (40), supporting the concept that these effects are specific to the timing of exposure (20), and it may be that the stresses around preterm birth occur outside this “vulnerable” period. Finally, it is also possible that differences in cell subtype populations in the saliva from preterm and term infants may impact on DNA methylation (41).

In summary, we found changes in DNA methylation at key regions of *IGF2/H19* and *FKBP5* in this cohort of preterm infants who did not achieve the reference for either growth or body composition by the time they reached 1 year of age. Further, we identified a number of contributing factors including maternal smoking and social deprivation. We found no persisting differences at 1 year of age. Whether further differences will emerge over longer follow-up is unknown; alterations in DNA methylation at key DMRs controlling the expression of *IGF2* have been reported in adults exposed to altered nutrition or prematurity (16, 17, 42); however, these were not longitudinal studies, so that any differences may have arisen postnatally. Additionally, these studies may have been complicated by the phenomenon of reverse causation, where the development of disease leads to changes in DNA methylation rather than *vice versa* (41). Further longitudinal studies are required to understand any potential long-term effects of early differences in DNA methylation in the perinatal period on the health of individuals born preterm.

REFERENCES

- Risnes KR, Vatten LJ, Baker JL, Jameson K, Sovio U, Kajantie E, et al. Birthweight and mortality in adulthood: a systematic review and meta-analysis. *Int J Epidemiol* (2011) 40(3):647–61. doi:10.1093/ije/dyq267
- Gluckman PD, Hanson MA. Living with the past: evolution, development, and patterns of disease. *Science* (2004) 305(5691):1733–6. doi:10.1126/science.1095292
- Parkinson JRC, Hyde MJ, Gale C, Santhakumaran S, Modi N. Preterm birth and the metabolic syndrome in adult life: a systematic review and meta-analysis. *Pediatrics* (2013) 131(4):e1240–63. doi:10.1542/peds.2012-2177
- Bayman E, Drake AJ, Piyasena C. Prematurity and programming of cardiovascular disease risk: a future challenge for public health? *Arch Dis Child Fetal Neonatal Ed* (2014) 99(6):F510–4. doi:10.1136/archdischild-2014-306742
- Tinnion R, Gillone J, Cheatham T, Embleton N. Preterm birth and subsequent insulin sensitivity: a systematic review. *Arch Dis Child* (2014) 99(4):362–8. doi:10.1136/archdischild-2013-304615
- Johnson S, Hollis C, Kochhar P, Hennessy E, Wolke D, Marlow N. Psychiatric disorders in extremely preterm children: longitudinal finding at age 11 years in the EPICure study. *J Am Acad Child Adolesc Psychiatry* (2010) 49(5):453–63.e.1. doi:10.1016/j.jaac.2010.02.002
- Blencowe H, Cousens S, Chou D, Oestergaard M, Say L, Moller A-B, et al. Born too soon: the global epidemiology of 15 million preterm births. *Reprod Health* (2013) 10(Suppl 1):S2. doi:10.1186/1742-4755-10-S1-S2
- Embleton NE, Pang N, Cooke RJ. Postnatal malnutrition and growth retardation: an inevitable consequence of current recommendations in preterm infants? *Pediatrics* (2001) 107(2):270–3. doi:10.1542/peds.107.2.270

ETHICS STATEMENT

Infants were recruited within the first week of life from the Simpson Centre for Reproductive Health, Edinburgh, UK, with informed written parental consent. Ethical approval was obtained from the South East Scotland Research Ethics Committee (Reference 11/AL/0329). NHS management approval was obtained (Lothian R&D Project number 2011/R/NE/03). Perinatal samples were collected under the Edinburgh Reproductive Tissue BioBank (ERTBB) (West of Scotland Research Ethics Service Reference 09/S0704/3). All parents gave written informed consent and all studies were performed in accordance with the declaration of Helsinki. Infant samples were collected under the framework of the ERTBB following an amendment to ethical approval (Reference AM07/1).

AUTHOR CONTRIBUTIONS

CP and AD conceived the study. CP, GM, JS, and AD designed the study. CP, JC, BK, RS, NP, TW, RR, EB, and AD performed the study and analyses. CP and AD wrote the initial manuscript draft and all authors revised it critically for intellectual content. All authors gave final approval of the version to be published.

ACKNOWLEDGMENTS

Our thanks go to Karen French for technical assistance, to the staff of the Wellcome Trust Children’s Clinical Research Facility, Edinburgh for their assistance with clinical studies, and to Susann Sauer for excellent technical assistance with *FKBP5* methylation analysis. Finally, our thanks go to the families who took part in this study.

FUNDING

Funding for the study was provided by a Scottish Senior Clinical Fellowship (SCD/09 to AD) and the MRC (G0501904 to AD). The purchase of the PEAPOD was supported by an equipment grant from Diabetes UK. *FKBP5* methylation analysis was supported by a European Research Council starting grant (grant# 281338, GxE molmech) within the FP7 framework to EB. NP was supported by a fellowship from the Canadian Institute of Health Research. We also acknowledge the support of the British Heart Foundation and the Edinburgh Reproductive Tissue BioBank.

9. Grunau RE, Haley DW, Whitfield MF, Weinberg J, Yu W, Thiessen P. Altered basal cortisol levels at 3, 6, 8 and 18 months in infants born at extremely low gestational age. *J Pediatr* (2007) 150(2):151–6. doi:10.1016/j.jpeds.2006.10.053
10. Brummelte S, Grunau RE, Zaidman-Zait A, Weinberg J, Nordstokke D, Cepeda IL. Cortisol levels in relation to maternal interaction and child internalizing behavior in preterm and full-term children at 18 months corrected age. *Dev Psychobiol* (2011) 53:184–95. doi:10.1002/dev.20511
11. Godfrey KM, Lillycrop KA, Burdge GC, Gluckman PD, Hanson MA. Epigenetic mechanisms and the mismatch concept of the developmental origins of health and disease. *Pediatr Res* (2007) 61(5 Pt 2):5R–10R. doi:10.1203/pdr.0b013e318045bedb
12. Cruickshank M, Oshlack A, Theda C, Davis P, Martino D, Sheehan P, et al. Analysis of epigenetic changes in survivors of preterm birth reveals the effect of gestational age and evidence for a long term legacy. *Genome Med* (2013) 5(10):96. doi:10.1186/gm500
13. Parets SE, Conneely KN, Kilaru V, Fortunato SJ, Syed TA, Saade G, et al. Fetal DNA methylation associates with early spontaneous preterm birth and gestational age. *PLoS One* (2013) 8(6):e67489. doi:10.1371/journal.pone.0067489
14. Sparrow S, Manning JR, Cartier J, Anblagan D, Bastin ME, Piyasena C, et al. Epigenomic profiling of preterm infants reveals DNA methylation differences at sites associated with neural function. *Transl Psychiatry* (2016) 6:e716. doi:10.1038/tp.2015.210
15. Ollikainen M, Smith KR, Joo EJ-H, Ng HK, Andronikos R, Novakovic B, et al. DNA methylation analysis of multiple tissues from newborn twins reveals both genetic and intrauterine components to variation in the human neonatal epigenome. *Hum Mol Genet* (2010) 19(21):4176–88. doi:10.1093/hmg/ddq336
16. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc Natl Acad Sci U S A* (2008) 105(44):17046–9. doi:10.1073/pnas.0806560105
17. Drake AJ, McPherson RC, Godfrey KM, Cooper C, Lillycrop KA, Hanson MA, et al. An unbalanced maternal diet in pregnancy associates with offspring epigenetic changes in genes controlling glucocorticoid action and foetal growth. *Clin Endocrinol* (2012) 77(6):808–15. doi:10.1111/j.1365-2265.2012.04453.x
18. Paquette AG, Lester BM, Lesseur C, Armstrong DA, Guerin DJ, Appleton AA, et al. Placental epigenetic patterning of glucocorticoid response genes is associated with infant neurodevelopment. *Epigenomics* (2015) 7(5):767–79. doi:10.2217/epi.15.28
19. Zannas AS, Binder EB. Gene-environment interactions at the FKBP5 locus: sensitive periods, mechanisms and pleiotropism. *Genes Brain Behav* (2014) 13(1):25–37. doi:10.1111/gbb.12104
20. Klengel T, Mehta D, Anacker C, Rex-Haffner M, Pruessner JC, Pariante CM, et al. Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nat Neurosci* (2013) 16(1):33–41. doi:10.1038/nn.3275
21. Paquette AG, Lester BM, Koestler DC, Lesseur C, Armstrong DA, Marsit CJ. Placental FKBP5 genetic and epigenetic variation is associated with infant neurobehavioral outcomes in the RICHs cohort. *PLoS One* (2014) 9(8):e104913. doi:10.1371/journal.pone.0104913
22. Dupont JM, Tost J, Jammes H, Gut IG. De novo quantitative bisulphite sequencing using the pyrosequencing technology. *Anal Biochem* (2004) 333(1):119–27. doi:10.1016/j.ab.2004.05.007
23. Pan H, Cole TJ. *LMSgrowth, a Microsoft Excel Add-In to Access Growth References Based on the LMS Method. Version 2.77*. (2012). Available from: <http://www.healthforallchildren.co.uk>
24. Hack M, Schluchter M, Cartar L, Rahman M, Cuttler L, Borawski E. Growth of very low birth weight infants to age 20 years. *Pediatrics* (2003) 112(1):e30–8. doi:10.1542/peds.112.1.e30
25. Cole TJ, Statnikov Y, Santhakumaran S, Pan H, Modi N; Neonatal Data Analysis Unit and the Preterm Growth Investigator Group, et al. Birth weight and longitudinal growth in infants born below 32 weeks' gestation: a UK population study. *Arch Dis Child Fetal Neonatal Ed* (2014) 99(1):F34–40. doi:10.1136/archdischild-2012-303536
26. Uthaya S, Thomas EL, Hamilton G, Dore CJ, Bell J, Modi N. Altered adiposity after extremely preterm birth. *Pediatr Res* (2005) 57:211–5. doi:10.1203/01.PDR.0000148284.58934.1C
27. Roggero P, Gianni ML, Amato O, Orsi A, Piemontese P, Morlacchi L, et al. Is term newborn body composition being achieved postnatally in preterm infants? *Early Hum Dev* (2009) 85(6):349–52. doi:10.1016/j.earlhumdev.2008.12.011
28. Ramel SE, Gray HL, Ode KL, Younge N, Georgieff MK, Demerath EW. Body composition changes in preterm infants following hospital discharge: comparison with term infants. *J Pediatr Gastroenterol Nutr* (2011) 53(3):333–8. doi:10.1097/MPG.0b013e3182243aa7
29. Thomas EL, Parkinson JR, Hyde MJ, Yap IKS, Holmes E, Dore CJ, et al. Aberrant adiposity and ectopic lipid deposition characterize the adult phenotype of the preterm infant. *Pediatr Res* (2011) 70(5):507–12. doi:10.1038/pr.2011.732
30. Elliott HR, Tillin T, McArdle WL, Ho K, Duggirala A, Frayling TM, et al. Differences in smoking associated DNA methylation patterns in South Asians and Europeans. *Clin Epigenetics* (2014) 6(1):4. doi:10.1186/1868-7083-6-4
31. Dogan MV, Lei MK, Beach SR, Brody GH, Philibert RA. Alcohol and tobacco consumption alter hypothalamic pituitary adrenal axis DNA methylation. *Psychoneuroendocrinology* (2016) 66:176–84. doi:10.1016/j.psyneuen.2016.01.018
32. Flom JD, Ferris JS, Liao Y, Tehranifar P, Richards CB, Cho YH, et al. Prenatal smoke exposure and genomic DNA methylation in a multiethnic birth cohort. *Cancer Epidemiol Biomarkers Prev* (2011) 20(12):2518–23. doi:10.1158/1055-9965.epi-11-0553
33. Murphy SK, Adigun A, Huang Z, Overcash F, Wang F, Jirtle RL, et al. Gender-specific methylation differences in relation to prenatal exposure to cigarette smoke. *Gene* (2012) 494(1):36–43. doi:10.1016/j.gene.2011.11.062
34. Novakovic B, Ryan J, Pereira N, Boughton B, Craig JM, Saffery R. Postnatal stability, tissue, and time specific effects of *AHRR* methylation change in response to maternal smoking in pregnancy. *Epigenetics* (2014) 9(3):377–86. doi:10.4161/epi.27248
35. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, et al. DNA methylation in newborns and maternal smoking in pregnancy: genome-wide consortium meta-analysis. *Am J Hum Genet* (2016) 98(4):680–96. doi:10.1016/j.ajhg.2016.02.019
36. Simpkin AJ, Hemani G, Suderman M, Gaunt TR, Lyttleton O, McArdle WL, et al. Prenatal and early life influences on epigenetic age in children: a study of mother-offspring pairs from two cohort studies. *Hum Mol Genet* (2016) 25(1):191–201. doi:10.1093/hmg/ddv456
37. Waterland RA, Lin J-R, Smith CA, Jirtle RL. Post-weaning diet affects genomic imprinting at the insulin-like growth factor 2 (*Igf2*) locus. *Hum Mol Genet* (2006) 15(5):705–16. doi:10.1093/hmg/ddi484
38. Rijlaarsdam J, Cecil CA, Walton E, Mesirow MS, Relton CL, Gaunt TR, et al. Prenatal unhealthy diet, insulin-like growth factor 2 gene (*IGF2*) methylation, and attention deficit hyperactivity disorder symptoms in youth with early-onset conduct problems. *J Child Psychol Psychiatry* (2016). doi:10.1111/jcpp.12589
39. Lee RS, Tamashiro KLK, Yang X, Purcell RH, Harvey A, Willour VL, et al. Chronic corticosterone exposure increases expression and decreases deoxyribonucleic acid methylation of *Fkbp5* in mice. *Endocrinology* (2010) 151(9):4332–43. doi:10.1210/en.2010-0225
40. Yehuda R, Daskalakis NP, Bierer LM, Bader HN, Klengel T, Holsboer F, et al. Holocaust exposure induced intergenerational effects on FKBP5 methylation. *Biol Psychiatry* (2015) 80(5):372–80. doi:10.1016/j.biopsych.2015.08.005
41. Birney E, Smith GD, Grealley JM. Epigenome-wide association studies and the interpretation of disease -omics. *PLoS Genet* (2016) 12(6):e1006105. doi:10.1371/journal.pgen.1006105
42. Wehkalampi K, Muurinen M, Wirta SB, Hannula-Jouppi K, Hovi P, Järvenpää A-L, et al. Altered methylation of *IGF2* locus 20 years after preterm birth at very low birth weight. *PLoS One* (2013) 8(6):e67379. doi:10.1371/journal.pone.0067379

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Piyasena, Cartier, Provençal, Wiechmann, Khulan, Sunderesan, Menon, Seckl, Reynolds, Binder and Drake. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

5.3 Paper III – Identification of dynamic glucocorticoid-induced methylation changes at the *FKBP5* locus

Wiechmann T., S. Roeh, S. Sauer, D. Czamara, J. Arloth, M. Ködel, M. Beintner, L. Knop, A. Menke, E. B. Binder and N. Provençal

Originally published in:

Clin Epigenetics. 2019 May 23;11(1):83. doi: 10.1186/s13148-019-0682-5.

PMID: 31122292

Additional supplementary material:

Additional supplementary figures and tables are freely available for download at the publisher's website (<https://doi.org/10.1186/s13148-019-0682-5>).

RESEARCH

Open Access



Identification of dynamic glucocorticoid-induced methylation changes at the *FKBP5* locus

Tobias Wiechmann¹ , Simone Röh¹, Susann Sauer¹, Darina Czamara¹, Janine Arloth^{1,7}, Maik Ködel¹, Madita Beintner¹, Lisanne Knop¹, Andreas Menke^{2,3}, Elisabeth B. Binder^{1,4*} and Nadine Provençal^{1,5,6*}

Abstract

Background: Epigenetic mechanisms may play a major role in the biological embedding of early-life stress (ELS). One proposed mechanism is that glucocorticoid (GC) release following ELS exposure induces long-lasting alterations in DNA methylation (DNAm) of important regulatory genes of the stress response. Here, we investigate the dynamics of GC-dependent methylation changes in key regulatory regions of the *FKBP5* locus in which ELS-associated DNAm changes have been reported.

Results: We repeatedly measured DNAm in human peripheral blood samples from 2 independent cohorts exposed to the GC agonist dexamethasone (DEX) using a targeted bisulfite sequencing approach, complemented by data from Illumina 450K arrays. We detected differentially methylated CpGs in enhancers co-localizing with GC receptor binding sites after acute DEX treatment (1 h, 3 h, 6 h), which returned to baseline levels within 23 h. These changes withstood correction for immune cell count differences. While we observed main effects of sex, age, body mass index, smoking, and depression symptoms on *FKBP5* methylation levels, only the functional *FKBP5* SNP (rs1360780) moderated the dynamic changes following DEX. This genotype effect was observed in both cohorts and included sites previously shown to be associated with ELS.

Conclusion: Our study highlights that DNAm levels within regulatory regions of the *FKBP5* locus show dynamic changes following a GC challenge and suggest that factors influencing the dynamics of this regulation may contribute to the previously reported alterations in DNAm associated with current and past ELS exposure.

Keywords: DNA methylation, *FKBP5*, Glucocorticoid receptor, Early-life stress, Targeted bisulfite sequencing, Dexamethasone

Background

Epidemiological studies indicate a combined contribution of genetic and environmental factors in the risk for psychiatric diseases, which converge to alter gene regulation and consequently cell function [1]. Evidence suggests that epigenetic mechanisms play a major role in embedding environmental risk, including early-life adversity, but our understanding of the underlying mechanisms is limited. Epigenetic mechanisms encompass post-translational modifications of histone proteins and

chemical modifications of single nucleotides (most commonly in the form of methylation at cytosine guanine dinucleotides (CpGs)), which alter chromatin structure, and thus accessibility of the DNA to transcriptional regulators.

Even DNA methylation (DNAm), a stable chemical modification, undergoes highly dynamic regulation in post-mitotic cells. This property makes DNAm a suitable molecular mechanism to encode the impact of environmental cues in post-mitotic tissue [2, 3]. A mechanism that likely contributes to such dynamic, environmentally triggered DNAm changes is transcription factor-mediated DNA demethylation [4]. One example is local demethylation of glucocorticoid response elements (GREs) following activation of the glucocorticoid receptor (GR), a nuclear

* Correspondence: binder@psych.mpg.de; nadine_provençal@sfu.ca

¹Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Kraepelinstr. 2-10, 80804 Munich, Germany

Full list of author information is available at the end of the article



transcription factor [5]. The GR is activated by the glucocorticoid (GC) cortisol, a major mediator of the stress response.

Stress, especially in the form of early adverse life trauma, is a major environmental risk factor for psychiatric disorders [1, 6, 7]. Excessive GC release after stress exposure may induce long-lasting DNAm changes, thereby contributing to the biological embedding of risk trajectories. The mechanism of GR-induced local demethylation is not fully understood, but activation of DNA repair machinery is implicated. Demethylation of GREs facilitates the transcriptional effects of the GR on the target gene [8, 9].

FKBP5 is a stress-responsive gene and co-chaperone protein of GR. Increased activation of this gene by genetic or epigenetic factors has been repeatedly associated with increased stress-sensitivity and risk for psychiatric disorders in both animal and human studies (see [10, 11] for review). We have previously reported on GR-sensitive CpGs in GREs of the *FKBP5* locus. These are located in a functional GRE in intron 7 of the gene. Chromatin conformation capture experiments confirmed an interaction of this intronic enhancer with the transcription start site (TSS) of *FKBP5*. Reporter gene assays also demonstrated that higher DNAm of this enhancer region was associated with lower transcriptional activation of *FKBP5* by GCs [12]. Relative reduction of DNAm in this region has been reported both in peripheral blood and buccal cells of adults as well as children exposed to childhood trauma and in a hippocampal neuronal progenitor cell (HPC) line following exposure to GCs [12–16]. Changes in DNAm following exposure to child abuse seemed to be accentuated in individuals carrying the minor allele of a functional genetic variant in this locus (rs1360780). This variant, located in close proximity to a GRE in intron 2, alters the 3D conformation of the locus. The minor allele generates a TATA-box binding site which allows binding of this intron enhancer to the transcription start site. This is also associated with higher *FKBP5* induction following GR activation. We and others have shown that increased *FKBP5* leads to reduced GR sensitivity and impaired negative feedback regulation of the stress hormone axis [17, 18]. In fact, minor allele carriers have repeatedly been shown to have prolonged cortisol release following stress exposure [11]. Finally, this functional allele has consistently been shown to increase risk for a range of psychiatric disorders with exposure to early adversity [11, 19–21], suggesting that gene x environment interactions at the level of epigenetic regulation may contribute to disease risk.

These studies suggest that the CpGs associated with early trauma exposure may also be responsive to GCs and that increased GR activation with trauma may lead to DNAm changes of the sites. However, direct evidence

for this is so far missing. Furthermore, due to the limitation of the previously used pyrosequencing-based DNAm assessment of these enhancers, only a small number of CpGs had been investigated. Transcriptional regulatory sites of *FKBP5* are distributed throughout the locus and include several upstream, downstream, and intronic enhancer regions with GREs [22] as well as CCCTC-binding factor (CTCF) sites in addition to the TSS. CTCF creates boundaries between topologically associating domains (TADs) in chromosomes, and within these domains, CTCF facilitates interactions between transcription regulatory sequences [23, 24]. The extent of GR-associated DNAm changes in different categories of regulatory elements within the *FKBP5* locus has not yet been explored.

Here, we investigate the changes of DNAm following exposure to the selective GR agonist dexamethasone (DEX) in peripheral blood cells over 24 h and in relation to rs1360780 genotype, in two independent cohorts. DNAm levels were assessed using a high-accuracy methylation measurements via targeted bisulfite sequencing (HAM-TBS) approach [25], which extensively covers CpG sites located in the different categories of regulatory elements in the *FKBP5* locus. The changes are also compared to data generated by the widely used Illumina methylation arrays.

Results

DEX-induced dynamic changes at the *FKBP5* locus in human peripheral blood (study 1)

In order to test if GR activation is associated with changes in DNAm in vivo, we first analyzed serial blood samples from 19 subjects exposed to a single oral dose (1.5 mg) of DEX (see Table 1 for demographic details).

DEX-induced changes in *ACTH*, cortisol, and *FKBP5* mRNA levels

Analysis of serum adrenocorticotropin (ACTH) and cortisol levels showed the expected suppression following

Table 1 Description of study 1 and 2 subjects

	Study 1	Study 2
Samples	19	89
Male	19	67
Female	0	22
rs1360780 genotype	CC = 6; CT = 6; TT = 7	CC = 50; CT = 30; TT = 9
Time points of blood draw after DEX	0 h, 1 h, 3 h, 6 h, 23 h	0 h, 3 h, 18–24 h
Age (mean ± SD)	25.4 ± 2.9	41.6 ± 14.0
BMI (mean ± SD)	N/A	25.1 ± 3.8
Smoking score (mean ± SD)	N/A	−0.6 ± 4.9
Major depressive disorder	0	59 (M = 38; F = 21)

DEX administration with maximal effects observed at 3 and 6 h post-treatment. In addition, DEX induced a 4.2- and 4.0-fold increase in *FKBP5* mRNA levels after 3 and 6 h of treatment, respectively, and returned to baseline level after 23 h (Fig. 1a).

DEX-induced transient DNAm changes—dynamics

DNAm was analyzed using our HAM-TBS technique [25], where a total of 25 amplicons covering 228 CpGs at 5 time points from baseline (0 h) to 1, 3, 6, and 23 h following DEX administration passed QC. Three amplicons located in the proximal enhancer did not pass QC due to increased CHH methylation levels (PCR18) or low coverage (PCR 28 and PCR29; < 1000 reads, Additional file 6: Table S1). DNAm analysis across the loci at baseline revealed low methylation at the TSS and higher methylation within the gene body and 3' and 5' flanking regions (Fig. 1b “% Methylation baseline” track).

Following an acute dose of DEX, 44 CpG sites showed significant changes in DNAm over all time points ($FDR \leq 0.05$ and absolute delta methylation (T_i -baseline) $\geq |1\%$); Fig. 1b “Max. Δ % methylation” track, shaded regions; Additional file 7: Table S2). Significant DEX-induced differential DNAm was seen as early as 1 h after treatment ($n = 17$ sites, mean absolute Δ methylation = $|2.4\%$), and the largest effects were observed after 3 and 6 h ($n = 40$ sites, mean absolute Δ methylation = $|4.4\%$ and $|5.4\%$, respectively) with a range from -17 to $+0\%$. Seventy-four percent of the sites, however, showed decreases in DNAm levels following DEX treatment ranging from -17 to -1% compared to baseline. For the majority of the sites, DNAm levels returned to baseline after 23 h of treatment while only 8 sites remained differentially methylated at $FDR < 0.05$ with a small change compared to baseline (mean absolute Δ methylation = $|1.8\%$).

DEX-induced transient DNAm changes—localization

DEX-induced differentially methylated CpG sites (DMCs) were found in the proximal and intronic enhancers and co-localized with ENCODE GR binding sites and those located at the chromatin interaction blocks overlapped with ENCODE CTCF binding sites (examples are shown in Fig. 1c). Within the 82 CpGs analyzed surrounding the TSS, no DMCs were observed. Out of the 129 CpGs located in GR binding sites, 36 (28%) showed DEX-induced changes whereas only 8 (10%) out of 83 sites located in CTCF binding sites showed changes after DEX. Most of the DMCs located in GR binding sites ($n = 30$) showed reduction of DNAm following DEX (mean Δ methylation $-3.8 \pm 3.3\%$) with the exception of 6 sites located in the proximal enhancer ($n = 4$), intron 5 ($n = 1$) and intron 2 ($n = 1$) showing increased DNAm (mean Δ methylation $+2.9 \pm 1.3\%$).

DMCs within CTCF binding sites located in TAD boundaries and intron 3 showed increase in DNAm ($n = 5$ sites, mean Δ methylation $+3.6 \pm 3.3\%$) whereas those located in the proximal enhancer showed decrease in DNAm ($n = 3$ sites; mean Δ methylation $-3.0 \pm 1.1\%$). Demethylated CpGs in the proximal enhancer overlap with both GR and CTCF binding sites (see Fig. 1c).

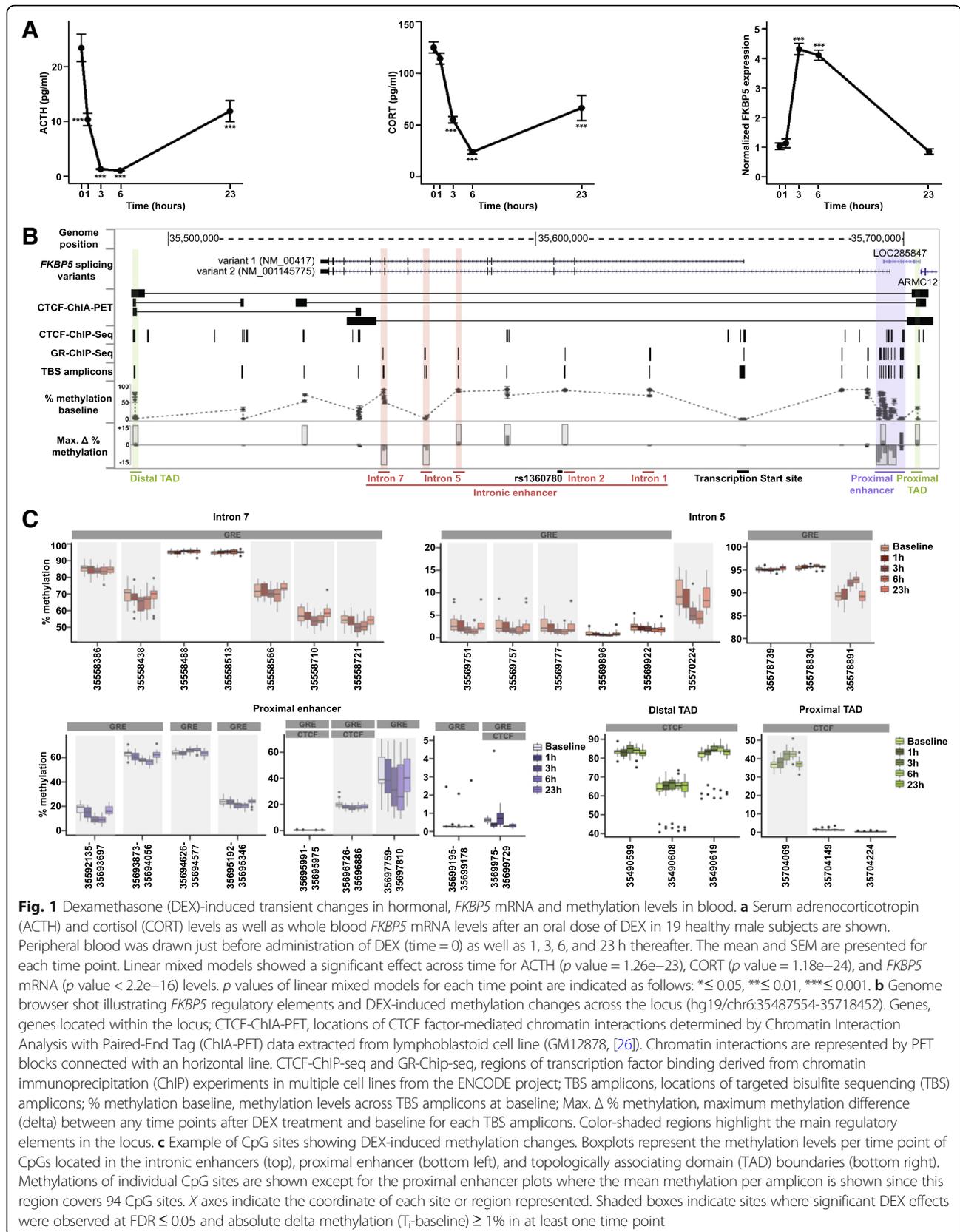
To assess whether changes in DNAm might directly affect binding of GR and CTCF to DNA, we mapped the changes to their relative distance to GR and CTCF consensus binding motifs. DNAm of CpG sites within these motifs has previously shown to impair or decrease transcription factor binding [27, 28]. We used predicted DNA binding motif locations for GR and CTCF from [29] (<http://compbio.mit.edu/encode-motifs/>). Selecting CpGs within ± 50 bp of the consensus motif sequences ($n = 16$ for GR and $n = 9$ for CTCF), we observed that CpG sites directly in CTCF motifs consistently displayed very low DNAm levels ($0.57 \pm 0.10\%$) whereas those in NR3C1 motifs showed intermediate levels at baseline with a high variation ($39.98 \pm 18.43\%$; Fig. 2a). For CTCF motif regions, higher DNAm was observed at more distal sites at the edges of the motif. DEX-induced DMCs were found directly in GR motifs ($n = 4$) whereas none were observed within CTCF motif but at the edges of this motif ($n = 2$, Fig. 2b).

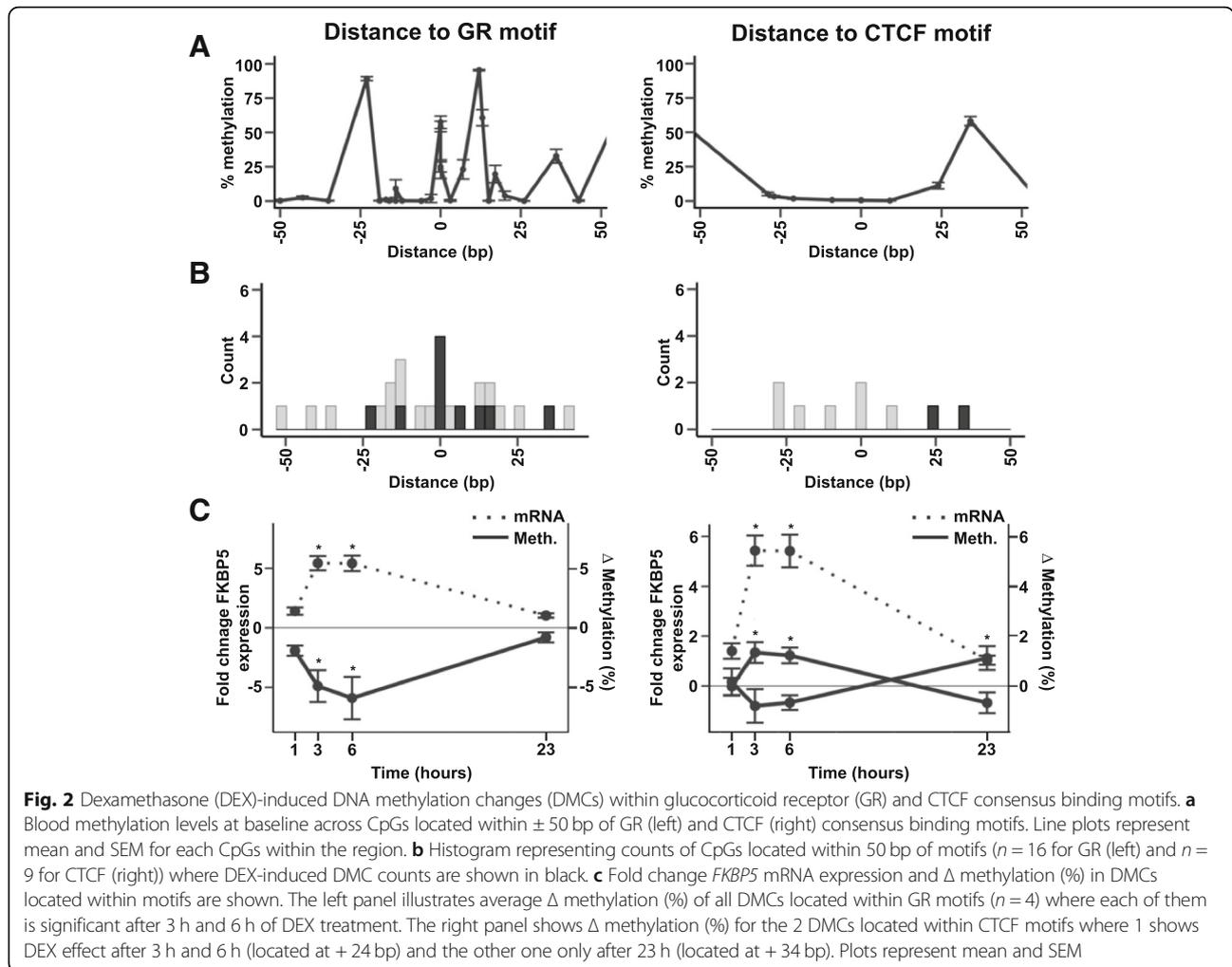
Increases in *FKBP5* mRNA levels occurred in parallel with the decrease of DNAm for 10 DMCs located within 50 bp of GR binding motifs (Fig. 2c and Additional file 6: Table S1).

Validation of DEX-induced DNAm changes in blood in an independent sample ($n = 89$, study 2)

Using study 2, see Table 1 for demographic details, we replicated the findings in an independent sample. Similar to study 1 and has been reported previously [30], DEX treatment induced a significant decrease in CORT and ACTH levels as well as an increase in *FKBP5* mRNA levels (Fig. 3a).

Ten amplicons covering 50 CpG sites (including 25 with significant DEX effect in study 1) were selected. Of these 50 sites, 21 showed significant changes ($FDR \leq 0.05$ and absolute Δ methylation $\geq |1\%$) in DNAm after DEX treatment validating 19 sites from the first study (Fig. 3b). Similar to the effects observed in the first study, DNAm changes were seen after 3 h of treatment (mean absolute Δ methylation = $|4.3\%$), and for most sites ($n = 13$), DNAm returned to baseline after 24 h. Eight DMCs remained significant after 24 h of treatment but showed a much smaller effect (mean absolute Δ methylation = $|1.8\%$). As observed before, the majority of sites (76%) show decrease in DNAm levels following DEX treatment with a range from -10 to -1% compared to baseline.

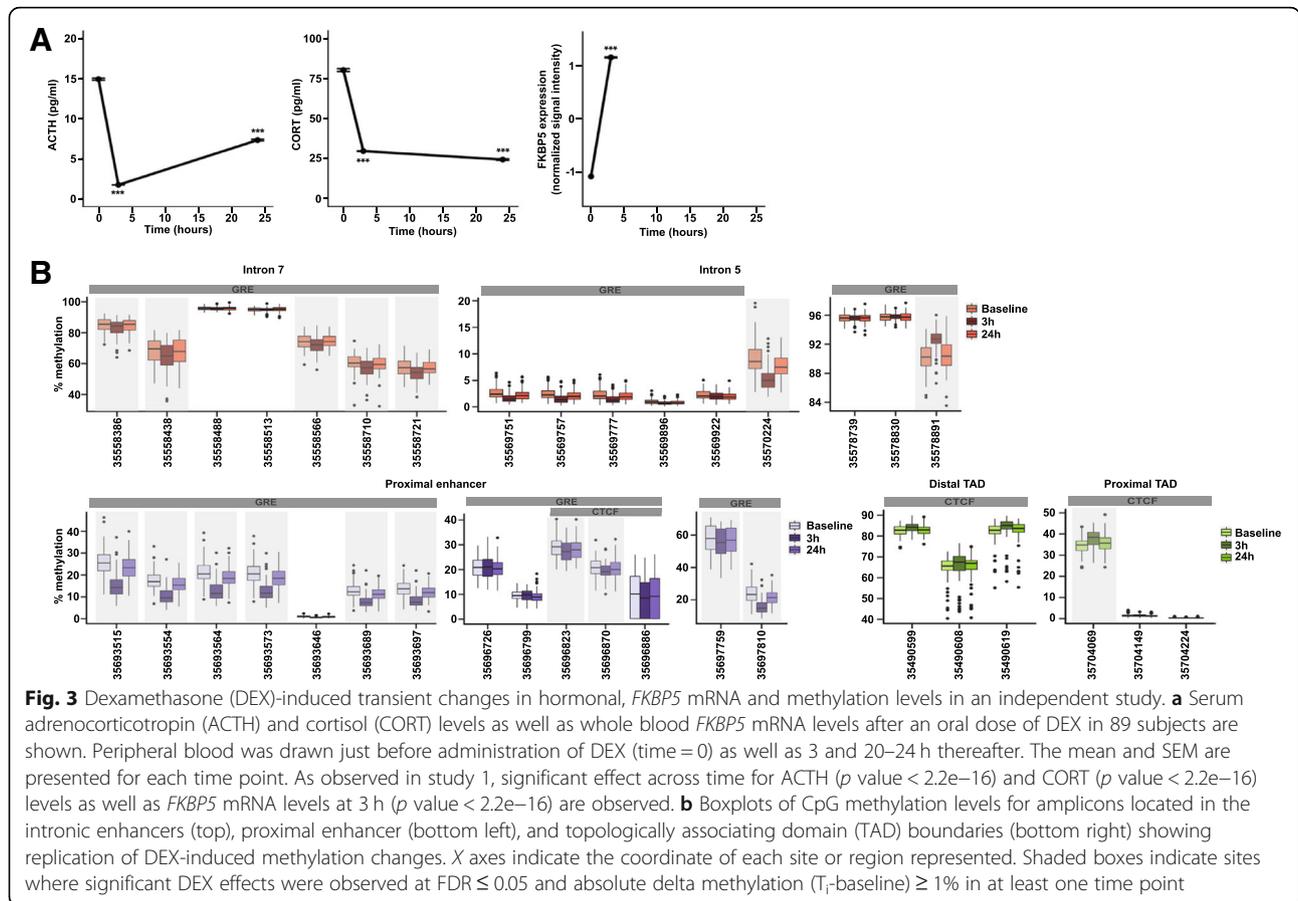




Inter-individual variability influencing DEX-induced methylation changes

Changes in DNA methylation in peripheral blood may reflect changes in immune cell composition. In study 2, we had data on blood cell counts (BCCs) as well as estimated immune cell types from the Illumina 450K array data. BCCs changed over time following DEX (see Additional file 1: Figure S1A) and lymphocyte counts significantly correlated with 9 DMCs (Additional file 8: Table S3). However, when we corrected for lymphocyte counts in the models testing DEX effects on DNAm across time, all 9 sites remained significant ($FDR \leq 0.05$, Additional file 8: Table S3). Since both DNAm and lymphocytes change over time with DEX, we next assessed how much of the variance in DNAm may still be explained by differences in BCCs. Comparing the standardized coefficients of lymphocyte counts change to the time changes in DNAm in a linear mixed model (LMM), we observed a significantly larger absolute coefficient at the 3-h time point (> 2.8 times larger) for all the associated sites for the changes in DNAm vs. the changes in lymphocytes (Additional file 8: Table S3 and Additional file 1: Figure S1B

showing the residuals of the null model correcting for lymphocyte count). Using estimated cell types from the Illumina 450K array data, we did not observe any effects of DEX on differential cell proportions (see Additional file 1: Figure S1C). Together, these analyses suggest that change in immune cell counts with DEX are likely not a major confounder of our results. In addition to immune cell counts, we also assessed the effects of other possible confounders including sex, age, ethnicity, body mass index (BMI), smoking, and depression symptoms. We observed significant main effects of age (7 sites), sex (5 sites), smoking (6 sites), BMI (8 sites), and major depressive disorder (MDD) (4 sites) on methylation but no significant interactions with DEX (Additional file 6: Table S1 and see Additional file 2: Figure S2A–C for examples). For MDD associations with DNAm, only men were included in the analysis ($n = 67$). We have previously reported no change at baseline for *FKBP5* mRNA levels between MDD patients and controls in a sub-sample of men from study 2 (29 cases and 31 controls, [30]). For the larger sample used here to test associations with MDD status in men (38 cases and 29 controls), there is also no



difference in *FKBP5* mRNA at baseline (p value > 0.05 and fold change ≥ 1.15). In contrast to the mRNA levels, out of the four *FKBP5* CpG sites associated with MDD over the three time points, three sites show a difference at baseline (p value ≤ 0.05 and absolute delta methylation (cases-controls) $\geq 1\%$) with less DNAm observed in cases (range from -1.4 to -2.2% , Additional file 6: Table S1).

Furthermore, we performed stepwise regression analyses using the Akaike Information Criterion (AIC) to select the main covariates influencing DNA methylation changes after DEX. This analysis was performed for the 17 CpG sites for which significant associations with DNA methylation were observed. Based on the AICs as well as the estimates of DNA methylation change after 3 h, correcting for smoking score gave the best models for all CpG sites (AIC smaller and largest variance explained). Sixteen out of seventeen sites remained significant ($FDR < 0.05$) after correcting for smoking (Additional file 10: Table S5). Adding the other covariates did not affect the significance of these 16 sites but increased AIC for most of them.

Effects of *FKBP5* genotype

We and others have previously described allele-specific DNAm changes (lower methylation) in intron 7 of the

FKBP5 gene in peripheral blood cells associated with exposure to child abuse only in carriers of the minor/risk T allele of rs1360780 [11]. Therefore, we investigated whether rs1360780 genotype (CC compared to CT/TT = high mRNA induction and disease risk) had an effect on the observed significant DEX-associated methylation changes in both studies. Genotype effect on DEX-induced methylation changes over time was tested using a model testing both additive as well as interactive effects. In the first study, we observed 17 CpGs showing significant interaction ($n = 13$) or additive effects ($n = 6$) on DNAm changes over time (Additional file 6: Table S1). Despite the different timelines, 2 CpGs showed significant genotype-dependent dynamic differences in both studies, cg3558710 located in intron 7 GRE and cg3570224 located in intron 5 next to the GRE (see Additional file 3: Figure S3). We next tested whether the direction of effects in the two different genotype conformations (CC vs CT/TT) was the same in study 1 and 2. When investigating all 50 CpGs common to both studies, 27 showed the same direction of effect for genotype x DEX at time point 3 h in study 2 as compared to the effect observed overall time points in study 1, which is more than expected by chance. This was not the case for effects of genotype on DNAm

differences at the time point 24 h in study 2. We then subdivided this analysis by regulatory regions. For CpGs within intronic and proximal enhancer GREs ($n = 30$), we observed a concordance of the direction of effects for genotype \times time interaction between the two studies significantly more than expected chance ($p = 0.049$), but this was not the case for the 20 CpGs annotated to TADs. Overall, T carriers displayed more methylation changes over time, with differences to CC genotype carriers ranging from 5.01 to 0.01%. This analysis supports that *FKBP5* rs1360780 genotypes associate with a differential DNAm sensitivity to GR activation within GREs but not in TADs.

Usefulness of Illumina methylation array for assessing DEX-induced DNA methylation changes in *FKBP5*

Most studies investigating DNA methylation in peripheral blood in large cohorts use Illumina methylation arrays. Overlapping Illumina 450K methylation data were available at baseline and 3 h post-DEX administration in study 2 ($n = 106$ subjects, [30]) allowing to assess the extent of coverage of DEX-reactive CpGs on these arrays. The 450K array covers 56 CpGs in the *FKBP5* locus (hg19/chr6:35487554-35718452, see Additional file 4: Figure S4 and Additional file 6: Table S1 for further details) which are located mainly around the TSS ($n = 12$) and the proximal enhancer ($n = 15$) with 12 CpGs in TAD boundaries and only sparse coverage within the gene body (11 CpGs) and 3' end (6 CpGs). None of the intronic GREs showing the biggest change in response to DEX are covered by the array. Analysis of 450K CpG sites that passed QC (52 sites) identified 13 DMCs following DEX ($FDR \leq 0.05$ and absolute Δ methylation $\geq |1\%$) identifying 9 additional DMCs not covered by HAM-TBS. Overlapping sites between TBS and 450K displayed high correlation in both datasets (17 CpGs with 0.96 mean correlation) where similar DNAm changes following DEX were observed (Additional file 4: Figure S4).

DEX-reactive sites reside in enhancer regions with cross-tissue relevance

To understand whether the observed changes could have cross-tissue relevance, as initially shown for the GRE in intron 7 [12], we compared the chromatin state segmentations from the Roadmap Consortium (<http://www.roadmapepigenomics.org/>) for blood/immune cells ($n = 29$ tissues), brain cells ($n = 10$), and fibroblasts ($n = 5$ tissues) (Additional file 5: Figure S5A) for the investigated regions within the *FKBP5* locus. Common active TSS marks at the TSS and marks indicating active transcription over the gene body support the well-documented active transcription of *FKBP5* across these tissues. Focusing on the regions in *FKBP5* that showed the most prominent effects of DEX on DNAm (GRE in intron 7, intron 5, and the

proximal enhancer), we observe that most of these regulatory elements show similar chromatin states, suggesting the comparable regulatory impact of these regions across these tissues (see Additional file 5: Figure S5B).

Discussion

Here, we investigated DNAm changes in response to GR activation in the *FKBP5* locus, a gene in which DNAm changes have been shown in association with exposure to childhood trauma in both children and adults [12, 14–16]. We here show that *FKBP5* DNAm in specific enhancers is highly responsive to GR activation by DEX. We observe dynamic methylation changes over time in longitudinal samples from two independent human studies. Significant effects of DEX exposure were detected as early as 1 h following oral ingestion of DEX, with maximal effects 3–6 h later. Most changes returned to baseline within 23 h. These effects remained significant when correcting for immune cell types as well as additional covariates such as age, sex, BMI, and depression status.

The sites dynamically responsive to an acute DEX challenge in blood overlap with sites correlating with 30-day cortisol load in healthy subjects [31] as well as CpGs differentially methylated in Cushing's syndrome patients, a disorder characterized by excess secretion of cortisol. This has been shown for CpGs within intronic GREs, especially the intron 2 and 7 GREs, for which differences were observed not only between patients with active Cushing's syndrome and controls but also in patients with cured Cushing's syndrome [32]. A second study using 450K data reported cg25114611, located in a GRE of the proximal enhancer, to show significantly lower methylation in patients with long-term remission of Cushing's syndrome [33]. In our study, this site showed lower methylation following DEX in blood that remained significant after 23 h. Such overlap suggests that GR activation may—under specific circumstances—result in lasting DNAm marks. In the above studies, patients had been in remission for 7–13 years [32, 33]. Such factors contributing to more lasting changes could be the level and duration of GR activation as well as developmental timing. We have previously reported that lasting changes in DNAm following DEX treatment in hippocampal progenitor cells were observed when cells were treated during proliferation and differentiation but not when they were treated post-differentiation [12].

The DEX-responsive sites also overlap with those reported to be demethylated in both children and adults exposed to early trauma [11]. Although these sites do not show durable demethylation after 23 h, the overlap suggests that these adversity-related epigenetic changes in *FKBP5* GREs could be mediated via GR-dependent effects through prolonged exposure to GCs, as observed in children exposed to trauma [13].

Our results give new insight into the previously described allele-specific methylation changes of intron 7 GRE CpGs that may contribute to the *FKBP5* x early trauma associations with risk for a number of psychiatric disorders ([10], for review). Several studies report that lower methylation of *FKBP5* GRE intron 7 is more pronounced in child abuse-exposed individuals carrying haplotypes tagged by the T allele of rs1360780 [12, 15, 34]. This functional allele has been associated with increased transcriptional activation of *FKBP5* by GR [12, 35]. Here, we show that healthy individuals carrying this allele have a different dynamic of DNAm changes following GR activation in GREs of *FKBP5*. Such a difference in epigenetic dynamics may also contribute to the fact that environmental risk factors linked to stress hormone activation, such as early adversity, could have more lasting effects in individuals carrying this specific allele. Such allele-specific differences in the dynamics of DNAm change may relate to the reported allele-specific differences in gene transcription. As illustrated in Fig. 2c, dynamic DNAm changes in *FKBP5* GREs inversely correlate with changes in RNA expression.

As expected from the literature, methylation levels of CTCF binding sites at the *FKBP5* locus were low ($0.57 \pm 0.10\%$). In general, it has been shown that CTCF occupancy is inversely correlated with DNA methylation and that DNA methylation at CpGs located directly in the core binding motif can inhibit CTCF binding [27, 36–40]. A gain of methylation directly at CTCF binding sites can lead to loss of CTCF binding and therefore disruption of chromatin interactions which can lead to a dysregulated gene expression [37]. DEX-induced DMCs were not found directly at the core motif of the CTCF binding sites but at close distance and show small changes ($< 2\%$), Fig. 2B). The upstream and downstream flanking regions around the CTCF core motif have been implicated in influencing CTCF binding stability [41]. However, how such a small change in DNA methylation at these flanking regions influences CTCF occupancy has not yet been experimentally addressed so that their exact role remains unknown. Overall, our data suggests that the DEX-induced methylation effects concentrate on GR and not CTCF binding sites. These changes in DNA methylation may thus alter enhancer function (as shown in reporter gene assays [12]) but will not likely result in more profound 3D chromatin changes, such as loop disruptions.

In addition to genetically induced altered dynamics of DEX responsivity, other factors may also contribute to long-term effects on *FKBP5* methylation in the context of early adversity. Factors associated with child abuse such as smoking, BMI, and depression all had main effects on *FKBP5* methylation, including in dynamically responsive sites (see Additional file 6: Table S1 and Additional file 2: Figure S2). In addition, age also had a main effect on some

CpGs. The limited age range of our cohorts, however, prohibits to analyze the influence of age, including childhood and adolescence, in more depth. Whether such differences in baseline methylation contribute to the long-term effect of early adversity on *FKBP5* methylation needs to be investigated in more detail in longitudinal cohorts. A limitation of this study is the lack of information on early or more recent adversity. Future studies will need to address the influence of these environmental factors on GR-induced DNA methylation dynamics.

Overall, the effects of DEX on *FKBP5* methylation were mostly in the direction of lower methylation following GR activation. In fact, the observation that DEX can induce DNA methylation changes is not unique to the *FKBP5* locus and has been observed at different genomic loci [42]. Several mechanisms could contribute to differences in DNAm. One is a change in cell type composition, favoring cell types with no methylation at these sites within mixed tissues. While this is a possible explanation for the changes observed, the fact that these effects withstand correction for changes in cell composition over time (see Additional file 1: Figure S1B and Additional file 8: Table S3) and that DEX has no significant effect on specific immune cell types estimated from the genome-wide DNAm data suggests that at least some of these effects likely happen within specific cells. Re-assessing our results in sorted cells would give more information into which cell types or cell characteristics are associated with the highest epigenetic reactivity in humans. A study in mice, using genome-wide bisulfite sequencing after cell sorting, observed GC-induced methylation changes primarily in blood T cells [43]. On the other hand, mapping of enhancers across many tissues, including different immune cells (see Additional file 5: Figure S5), suggests that most of the GR-responsive enhancers exert a shared function and may thus show similar epigenetic responses to GR activation across different cell types and tissues.

A reduction in DNA methylation following GR activation could also be mediated via a transcription factor binding-mediated DNA demethylation which has been reported for GR binding to GREs [5]. The mechanisms for this targeted DNA demethylation are not fully understood, but mechanisms involving DNA repair have been proposed [9]. Similar to the GR-induced transient changes observed here in blood, rapid cyclical pattern of DNAm in response to estrogen stimulation in breast cancer cells has been reported. The ER α -responsive gene *pS2* undergoes rapid demethylation and remethylation cycles following activation of transcription with estrogen [44]. The authors [44] implicated a coordinated binding of DNA methyltransferases, glycosylase, and base excision repair proteins in these processes. The process of demethylation of the *pS2* promoter investigated in the

above paper is thought to involve Dnmt3a/b that is able to deaminate 5mC. The resulting abasic site (AP site) of this deamination is subsequently repaired by recruiting p68, TDG, and BER proteins (AP endonuclease, DNA polymerase β , and DNA ligase I). The rapid GR-induced demethylation followed by remethylation within 23 h observed here in blood cells may occur via similar mechanisms given the reported kinetics of this enzymatic process. In addition, when aggregating data from our study for all GREs and mapping DNAm changes to the distance from the consensus GR binding site, we observe high levels of methylation within the consensus binding site and these central CpGs are also the ones with dynamic reduction following GR activation (see Fig. 2a). These observations would support GRE-centric active DNA demethylation. The mechanisms that would then associate with the more lasting changes in remitted Cushing's patients and in individuals exposed to childhood trauma could relate to different actions of DNA-methyl binding proteins such as MeCP2 and polycomb complexes that would interfere with DNA-driven demethylation/remethylation [45].

Conclusions

Taken together, these data provide novel insight into possible mechanisms of stress and trauma-related changes in DNAm and gene \times stress interactions, suggesting a role of GR-dependent methylation changes at least for a subset of the effects. These effects are best investigated using targeted approaches, such as HAM-TBS [25], as most of the reactive enhancer CpGs are not covered on the current Illumina arrays. The observed dynamics of these changes in peripheral blood have consequences on epigenetic association studies in humans, where controlling for cortisol plasma levels appears to be an important factor. Given that dynamic changes in DNAm that can be induced by a single dose of DEX and given their overlapping sites correlating with 30-day cortisol load as well as with lasting changes observed in patients with Cushing's syndrome, critical questions arise for the long-term epigenetic consequences of the therapeutic use of GCs. Additional research in larger samples, with different exposure lengths and intensity, different tissues, and different developmental stages, will be necessary to better understand this phenomenon on a genome-wide and organism-wide level. Cataloging the moderation of these GR-induced epigenetic effects by common gene variants may further help in identifying genes contributing to risk and resilience to stress-related psychiatric disorders.

Methods

Study samples

Study 1

Healthy male participants ($n = 26$, age 25.4 ± 2.9) were given 1.5 mg of DEX orally at 12:00, see [46] for more

details on the study samples and [30, 47–50] for the choice of dose. Peripheral blood was drawn just before administration of DEX as well as 1, 3, 6, and 23 h thereafter. Nineteen of the 26 samples were selected for TBS for a balanced rs1360780 genotype distribution (7 subjects with TT genotype, 6 with TC, and 6 with the CC genotype).

Study 2

The second sample consisted of 89 Caucasian subjects and were also exposed to 1.5 mg of DEX orally as previously described in [42]. Here, DEX was administered at 6 pm and blood draws occurred immediately before the dose of DEX and then 3 hours as well as ~ 18 hours (from 17.5 to 21 h) later. The subset comprised of 30 healthy probands (female = 1; male = 29) and 59 inpatients with depressive disorders (female = 21; male = 38) with an age of 41.64 ± 13.96 (mean age \pm SD).

The demographics of both studies are reported in Table 1.

DNA and RNA extraction of study samples

For both studies, DNA was extracted from frozen EDTA blood using the Elmer Chemagic 360 Instrument (PerkinElmer chemagen Technologie GmbH, Baesweiler, Germany) in combination with the chemagic DNA Blood Kit special 400 (PerkinElmer chemagen Technologie GmbH, Baesweiler, Germany). Thirty-three blood samples of study 2 were only collected in PAXgene tubes (PreAnalytiX GmbH, Hombrechtikon, Switzerland) for the ~ 18 h time point. For these samples, DNA was extracted from PAXgene tubes using PAXgene Blood DNA Kit (QIAGEN GmbH, Hilden, Germany). Blood for RNA was stored in PAXgene tubes, and RNA was extracted using the PAXgene Blood RNA Kit (QIAGEN GmbH, Hilden, Germany). All samples had an RNA integrity number of at least 7.0.

Genotyping

Study 1

All participants were genotyped for the rs1360780 allele using hybridization probes (forward primer: CCTTATTC-TATAGCTGCAAGTCCC, reverse primer: TCTGAATAT-TACCAGGATGCTGAG, rs1360780_LC: Red640-AAAT TCTTACTTGCTACTGCTGGCACAAGAGA-Phosphate, rs1360780_FL: CAGAAGGCTTTCACATAAGCAAAGT-TACACAAAAC-Fluorescein). Genomic DNA was amplified using the LightCycler 480 Genotyping Master mix (Roche, Mannheim, Germany) with the following cycling conditions: 95 °C, 10 s; 45 \times (95 °C, 1 s; 56 °C, 10 s; 72 °C, 10 s); 95 °C, 1 min; and 40 °C, 1 min, ramped up to 85 °C using a ramp rate of 0.57 °C/s and one acquisition per °C on a LightCycler480 II Instrument (Roche, Mannheim, Germany).

Study 2

Genotyping for this cohort is described in [47] and was based on Illumina 660K genotyping arrays.

Assessment of endocrine and immune measures

Cortisol and ACTH levels were assessed as described in [48, 51]. For the measurement of plasma cortisol concentrations, a radioimmunoassay kit was used (INC Biomedicals, Carlson, CA). Plasma ACTH concentrations were assessed by automated electrochemiluminescence immunoassay using Roche Cobas immunoassay analyzer (Roche, Basel, Switzerland). In the second study, additionally, plasma DEX levels were measured at the 3-h and 20–24-h time point using mass spectrometry as described in [48] and differential blood cell counts evaluated at all three time points as reported in [51].

Gene expression analysis via quantitative real-time PCR

Study 1

FKBP5 mRNA expression levels in blood samples of the first study were assessed as follows. The generation of cDNA was achieved using SuperScript™ II reverse transcriptase (Thermo Scientific Inc., Schwerte, Germany). Subsequently, the cDNA was amplified in duplicates in a LightCycler 480 Instrument II (Roche, Mannheim, Germany) using the LightCycler 480 SYBR Green I Master kit (Roche, Mannheim, Germany) and primer spanning Exon 10-11 (forward primer: AAAAGGCCAAGGAGCA-CAAC, reverse primer: TTGAGGAGGGCCGAGTTC; cycling conditions: 95 °C, 10 s; 45× (95°,10 s; 58 °C, 10 s; 72 °C, 10 s)). C_t values were used to calculate relative expression levels according to [52] normalized on *YWHAZ* expression (Universal ProbeLibrary probe #77; cycling conditions according to the manufacturer's recommendations; Roche, Mannheim, Germany) for normalization and mean assay efficiencies.

Study 2

RNA expression from study 2 was done using Illumina HT12v4 arrays previously described in [48]. In this study, blood RNA samples were only available for two time points (baseline and 3 h after DEX exposure).

Assessment of DNA methylation in study samples

DNA methylation levels in both studies were analyzed using the HAM-TBS approach, comprising an optimized PCR panel of 28 amplicons in the *FKBP5* locus [25]. These data were complemented by Illumina 450K methylation arrays. An overview of the methylation data obtained can be found in Additional file 6: Table S1.

Study 1

HAM-TBS [25] on the *FKBP5* locus was run for all 28 amplicons on DNA from each blood sample (baseline, 1

h, 3 h, 6 h, and 23 h post-DEX administration) in a single sequencing run with 302 CpG sites analyzed.

Study 2

FKBP5 locus DNAm levels were assessed using both HAM-TBS [25] (10 amplicons covering 50 CpGs from each blood sample (baseline, 3 h and 18 h post-DEX administration)) and 450K methylation arrays (56 CpGs at baseline and 3 h post-DEX administration).

Targeted bisulfite sequencing of the *FKBP5* locus

This method has been described in detail in [25] and offers good performance of amplicon bisulfite sequencing assays in a technology comparison by the Blueprint consortium [53].

Amplicon selection and amplification by PCR

We optimized the amplifications of 28 regions covering 302 CpGs within GR and/or CTCF binding sites as well as the transcription start site of the *FKBP5* locus (see Additional file 9: Table S4 for primers and mapping of the amplicons). In order to reduce cost and maximize the number of samples per sequencing run, triplicate bisulfite treatments were performed for each sample and then pooled to run one PCR amplification per amplicon [25]. Overall, 200 ng to 500 ng of DNA was used per sample and bisulfite treated using the EZ DNA Methylation Kit (Zymo Research, Irvine, CA). Twenty nanograms of bisulfite-converted DNA was then used for each PCR amplification employing Takara EpiTaq HS Polymerase (Clontech, Saint-Germain-en-Laye, France) and 49 amplification cycles. PCR amplicons were then quantified with the Agilent 2200 TapeStation (Agilent Technologies, Waldbronn, Germany) and pooled in equimolar quantities for each sample. AMPure XP beads (Beckman Coulter, Krefeld, Germany) were used for a double size selection (200–500 bp) to remove primer dimers and high molecular DNA fragments.

Sequencing

Libraries were generated using the TruSeq DNA PCR-Free HT Library Prep Kit (Illumina, San Diego, CA) according to the manufacturer's instructions. Each library was quantified with the Qubit® 1.0 (Thermo Fisher Scientific Inc., Schwerte, Germany), normalized to 4 nM and pooled. Library concentration and fragment sizes were checked via Agilent's 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany) and quantitative PCR using the Kapa HIFI Library quantification kit (Kapa Biosystems, Wilmington, MA). Paired-end sequencing was performed on an Illumina MiSeq Instrument (Illumina, San Diego, CA) with their MiSeq Reagent Kit v3 (2× 300 cycles) with the addition of 30% of PhiX Library.

Sequencing data processing

The quality of the sequencing reads was checked with FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>), and Illumina adapter sequences were removed using Cutadapt v.1.9.1. Bismark v.0.15.0 was used for the alignment to a restricted reference limited to our PCR targets. In order to stitch paired-end reads, an in-house Perl script has been developed to remove the low-quality ends of the paired-end reads if they overlapped. The methylation levels for all CpGs, CHGs, and CHHs were quantified using the R package methylKit. The resulting DNAm calls were submitted to a 3-step quality control. First, PCR artifacts introducing false CpGs of low coverage at 0 or 100% methylation level were removed. Second, CHH methylation levels were analyzed, and samples with insufficient bisulfite conversion rate (< 95%) were removed. Finally, CpG sites with a coverage lower than 1000 reads were excluded.

Illumina 450K methylation arrays

Study 2

Illumina 450K arrays were processed as described in [42]. Smoking scores were predicted from DNAm data as described in [54] as this information was not available for all subjects. Blood cell ratios were estimated from the DNAm data using the Houseman algorithm [55]. Normalized beta values of 52 CpG sites located within the *FKBP5* locus (hg19/chr6:35,487,554-35,718,452) were extracted from the 425,883 probes that passed quality control (QC).

Statistical analysis

DEX effects in study 1 and study 2

Linear mixed models (LMMs) were used to assess the effects of DEX treatment over time on either ACTH, cortisol, *FKBP5* expression, or DNAm levels for each CpG sites in both studies. All models were run adjusting for intra-individual variability as random effect using the “lmer” function of the Lme4 package in R [56]. *p* values were calculated using the Wald chi-square test from the Car package [57]. False discovery rate (FDR) was applied to correct for multiple testing on methylation *p* values. Post hoc analysis comparing each time point to baseline was ran using LMMs for each site showing significant results from above (FDR ≤ 0.05) to determine at which time point the effect was observed. Differentially methylated CpG sites (DMCs) were called when FDR from the mixed model was ≤ 0.05 and absolute mean methylation differences between significant time points and baseline (*p* value ≤ 0.05) were ≥ 1%.

The power to detect significant DNAm changes after DEX administration was performed using the function “powerSim” from the R package SIMR [56] with 100 simulations for each CpG sites profiled in both studies.

These analyses used a significance (alpha) level of 0.05 and minimum effect of absolute methylation difference ≥ 1% between 3 or 6 h and baseline for study 1 (delta T_3-T_0 and T_6-T_0) and 3 hours and baseline for study 2 (delta T_3-T_0). In our discovery sample (study 1), an average power of 96.6% (bootstrap 95% CI = (94.7, 97.9)) overall the 228 CpG sites was predicted to detect a minimum difference of 1% in methylation after 3 and/or 6 h of DEX administration. Over all the 50 CpG sites profiled in the replication cohort (study 2), an average power of 88.8% (bootstrap 95% CI = (81.2, 93.4)) was predicted to detect a minimum difference of 1% in methylation after 3 h of DEX administration. These results indicate that both cohorts have sufficient power (> 80%) to detect a minimum difference of 1% in methylation after DEX administration with our repeated measures design (5 sampling times in study 1 and 3 in study 2). Although the power in both studies is sufficient to detect a 1% change in methylation, much of the effect observed was larger than 3% (for 66% and 62% of the total significant sites in study 1 and 2, respectively).

In addition, parametric bootstrap using the “bootMer” function of the “lme4 package” in R using 100 simulations, for the mixed models of the 50 CpG sites profiled in study 2, was performed. The bootstrap results including the measures of bias and standard error as well as confidence intervals are given in Additional file 11: Table S6. This analysis revealed that the results are stable as the 95% confidence intervals from the 3-h and 24-h time points indicate a change in DNA methylation for the all sites identified with LMM (FDR ≤ 0.05 and absolute delta methylation ≥ 1%).

Inter-individual factors influencing DEX-induced DNAm changes

To assess inter-individual factors influencing the observed changes in DNAm following DEX administration, each CpG site showing significant DEX effects in study 2 was tested (*n* = 21 CpGs). LMMs were used to assess the association between DNAm change over time and blood cell counts, age, sex, smoking, BMI, and MDD status for each CpG sites. All models were run adjusting for intra-individual variability as random effect using the “lmer” function of the Lme4 package in R [56]. *p* values were calculated using the Wald chi-square test from the Car package [57]. Stepwise regression analysis was also performed on 17 sites showing association with either age, sex, smoking, BMI, or MDD status to select the main confounding variables influencing DNAm change over time. AICs and the DNAm estimates at 3 h of these models were used to select the best model (Additional file 10: Table S5).

FKBP5 genotype effect on DEX-induced methylation changes over time was first assessed on 44 CpG sites in study 1 and 21 sites in study 2 showing DNAm changes at any time point post-administration of DEX. LMMs

were used as described above. Methylation for each CpG was regressed against the main effect of time (DEX) and rs1360780 risk allele (CC or CT/TT) with and without the interaction term of genotype * time point while adjusting for intra-individual variability. *p* values of the additive and interaction effects for each time points were calculated using the Wald chi-square test.

We assessed if the direction of effects was concordant across studies based on the binomial distribution. Assuming that a CpG site shows the same direction of effect in both studies by chance with a probability of 0.5, we determined the probability to observe the present or even a higher number of CpG sites with concordant directions.

Additional files

Additional file 1: Figure S1. Change in blood cell counts after DEX administration. A) Actual blood cell counts at baseline and after DEX administration for granulocytes, monocytes, and lymphocytes in 54 subjects from study 2. B) Boxplot of DNAm residuals from a null model correcting for associated variance in lymphocyte counts across time in 54 subjects from study 2. Post hoc analysis correcting for lymphocyte counts revealed significant change in DNAm after 3 h of DEX for all sites (*p* value < 0.1e–18). C) Predicted blood cell proportions from 450K methylation data in study 2 using the Houseman algorithm [55]. (PDF 260 kb)

Additional file 2: Figure S2. DEX-induced changes in DNAm are also influenced by factors associated with early life adversity. Examples of three CpG sites were significant associations with fixed factors including age, sex, BMI, smoking score, and major depression were observed. (PDF 150 kb)

Additional file 3: Figure S3. CpG sites with significant genotype-dependent dynamic methylation differences in both studies. Effects of rs1360780 genotype on DEX-induced DNA methylation changes in 2 sites located in intron 7 and 5 enhancers. The % methylation levels for rs1360780 risk allele carriers CT/TT and CC carriers following DEX exposure are shown for each study. Methylation of CpG 35558710 shows significant interaction effect at 23 h in study 1 ($\chi^2 = 5.69$, *p* value = 0.02) and additive effect at 3 h in study 2 ($\chi^2 = 4.15$, *p* value = 0.04) with risk allele genotype. Significant interactions between risk allele genotype and DEX on methylation were observed for CpG 35570224 at 6 h and 23 h post-treatment in study 1 ($\chi^2 = 7.59$, *p* value = 0.006 and $\chi^2 = 6.0$, *p* value = 0.01) and at 24 h in study 2 ($\chi^2 = 4.36$, *p* value = 0.04). Points and error bars represent mean and SEM for each genotype. (PDF 101 kb)

Additional file 4: Figure S4. Replication of dexamethasone (DEX)-induced methylation changes (*n* = 106 subjects) analyzed by Illumina 450K arrays. A) Genome browser shot illustrating the location of TBS amplicons assessed as well as the location of the 450K Illumina probes within the *FKBP5* locus (hg19/chr6:35487554-35718452). *CTCF-ChIA-PET*-track indicating the locations of CTCF factor mediated chromatin interactions determined by Chromatin Interaction Analysis with Paired-End Tag (ChIA-PET) data extracted from lymphoblastoid cell line (GM12878, [26]). Chromatin interactions are represented by PET blocks connected with an horizontal line; *CTCF-ChIP-seq* and *GR-ChIP-seq*—regions of transcription factor binding derived from chromatin immunoprecipitation (ChIP) experiments in multiple cell lines from the ENCODE project; *blood TBS amplicons*—locations of targeted bisulfite sequencing (TBS) amplicons assessed in blood of study 1; *450K probe locations*—locations of Illumina probes from the 450K array. B) Boxplot of DNAm levels using TBS or Illumina 450K approach for the overlapping CpG sites showing methylation changes after DEX using TBS. *p* values of linear mixed models for each time point compared to baseline or vehicle are indicated as follows: * ≤ 0.05 , ** ≤ 0.01 , *** ≤ 0.001 . Note that although cg125114611 show significant DEX effect using 450K array, this site has a methylation change after DEX of –0.4% which did not reach our threshold of [1%]. (PDF 480 kb)

Additional file 5: Figure S5. Comparison of chromatin states in *FKBP5* across brain, immune/blood, and fibroblasts. A) Genome browser shot illustrating the chromatin states of the *FKBP5* locus (hg19 / chr6:35487554-35718452) across brain, immune/blood, and fibroblasts. *FKBP5 splicing variants*—visualization of two splicing variants of *FKBP5*; *TBS amplicons*—locations of targeted bisulfite sequencing (TBS) amplicons; *450K probe locations*—locations of Illumina probes from the 450K array; *CTCF-ChIP-seq* and *GR-ChIP-seq*—regions of transcription factor binding derived from chromatin immunoprecipitation (ChIP) experiments in multiple cell lines from the ENCODE project; *CTCF-ChIA-PET* and *PoII-ChIA-PET*—track indicating the locations of CTCF factor or PolII mediated chromatin interactions determined by Chromatin Interaction Analysis with Paired-End Tag (ChIA-PET) data extracted from lymphoblastoid cell line (GM12878, [26]). Chromatin interactions are represented by PET blocks connected with an horizontal line; *Ensembl Reg Build*—overview of the Ensembl regulatory build which represents an annotation of regions likely to be involved in gene regulation; *ChroHMM*—this track displays the chromatin state segmentation of the *FKBP5* locus for selected brain, immune/blood, and fibroblast cells from the Roadmap Consortium. The primary core marks segmentation has been used which visualize predicted functional elements as 15 states, which are displayed at the bottom of the figure. B) Quantification of the 15 chromatin states at key regulatory regions (transcription start site (TSS), topologically associating domains (TAD), proximal Enhancer (proxE), and intronic GREs (iGRE)) of the *FKBP5* locus. Chromatin states were averaged over brain (*n* = 10), immune/blood (*n* = 29), and fibroblasts (*n* = 5) cells. (PDF 390 kb)

Additional file 6: Table S1. Details on the CpG sites assessed in *FKBP5* locus and summary of the results obtained using HAM-TBS and Illumina 450K array in both studies. (XLSX 78 kb)

Additional file 7: Table S2. Summary statistic from linear mixed models testing the change in methylation after DEX for each CpG sites assessed in study 1 (*n* = 228) including post-hoc analysis for each time point. In bold are sites with significant DEX-induced methylation change. (XLSX 234 kb)

Additional file 8: Table S3. Summary statistic from linear mixed models testing the change in methylation after DEX including lymphocyte cells counts as covariate for each CpG sites associated with change in lymphocyte counts (*n* = 9) in study 2. (XLSX 50 kb)

Additional file 9: Table S4. Location of HAM-TBS amplicons and primer sequences used to analyze *FKBP5* CpGs. (XLSX 15 kb)

Additional file 10: Table S5. Results from stepwise regression analyses comparing the LMM models without covariate, with smoking score alone, or with all the associated covariates performed on 17 CpG sites showing association with either age, sex, smoking, BMI, or MDD status in study 2. (XLSX 35 kb)

Additional file 11: Table S6. Summary statistic from the linear mixed models testing the change in methylation after DEX for 50 CpG sites profiled in study 2 as well as the measures of bias, standard error, and confidence intervals using bootstraps with 100 simulations. (XLSX 46 kb)

Abbreviations

4-OHT: 4-Hydroxytamoxifen; BCCs: Blood cell counts; BER: Base excision repair; ChIA-PET: Chromatin Interaction Analysis with Paired-End Tag; ChIP: Chromatin immunoprecipitation; CpGs: Cytosine guanine dinucleotides; CTCF: CCCTC-binding factor; DEX: Dexamethasone; DMC: Differentially methylated CpG site; DNAm: DNA methylation; DNMTs: DNA methyltransferases; ELS: Early-life stress; GC: Glucocorticoid; GR: Glucocorticoid receptor; GRE: Glucocorticoid response element; HAM-TBS: High-accuracy methylation measurements via targeted bisulfite sequencing; HPC: Hippocampal progenitor cell; HW: Hardy-Weinberg equilibrium; LMM: Linear mixed model; MAF: Minor allele frequency; MDD: Major depressive disorder; PBMCs: Peripheral blood mononuclear cells; PcG: Polycomb complexes; SVA: Surrogate variable analysis; SVs: Surrogate variables; TAD: Topologically associating domains; TETs: Ten-eleven translocation methylcytosine dioxygenase proteins; TSS: The transcription start site

Acknowledgements

The authors would like to thank Monika Rex-Haffner for her help with the sequencing the libraries as well as to all individuals who agreed and provided

blood samples for this study. Moreover, the authors would like to thank Manfred Uhr and his team for their support in the assessment of ACTH and cortisol levels in the human blood samples.

Funding

This study was funded by the BMBF grant Berlin-LCS (FKZ 01KR1301B) to EBB and an ERC starting grant (GxE molmech, grant # 281338) within the FP7 funding scheme of the EU to EBB. NP was funded by a research fellowship from the Canadian Institute of Health Research (CIHR).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

TW, SS, MK, and MB performed the wet lab work. SR performed the sequencing data processing. JA performed the processing of 450K data. TW, DC, and NP analyzed and illustrated the data. MB, LK, and AM collected and processed the study samples. TW, NP, and EBB prepared the manuscript. NP and EBB conceptualized and supervised the study. All authors read and approved the final manuscript.

Authors' information

Not applicable

Ethics approval and consent to participate

All studies were approved by the local ethics committee of the Medical School of the Ludwig Maximilians University, and all participants gave informed consent.

Consent for publication

Not applicable

Competing interests

EBB is co-inventor of the European patent application *FKBP5*: a novel target for antidepressant therapy *European Patent# EP 1687443 B1* and receives a research grant from Böhlinger Ingelheim to investigate *FKBP5* as a candidate target in psychiatric disorders. The other authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Kraepelinstr. 2-10, 80804 Munich, Germany. ²Department of Psychiatry, Psychosomatics and Psychotherapy, University Hospital of Würzburg, Würzburg, Germany. ³Comprehensive Heart Failure Center, University Hospital of Würzburg, Würzburg, Germany. ⁴Department of Psychiatry and Behavioral Sciences, Emory University Medical School, Atlanta, GA, USA. ⁵Faculty of Health Sciences, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada. ⁶BC Children's Hospital Research Institute, Vancouver, BC, Canada. ⁷Institute of Computational Biology, Helmholtz Zentrum München, Neuherberg, Germany.

Received: 27 November 2018 Accepted: 9 May 2019

Published online: 23 May 2019

References

- Kessler RC, Davis CG, Kendler KS. Childhood adversity and adult psychiatric disorder in the US National Comorbidity Survey. *Psychol Med*. 1997;27:1101–19.
- Sweatt JD. The emerging field of neuroepigenetics. *Neuron*. 2013;80:624–32.
- Weng YL, An R, Shin J, Song H, Ming GL. DNA modifications and neurological disorders. *Neurotherapeutics*. 2013;10:556–67.
- Feldmann A, Ivanek R, Murr R, Gaidatzis D, Burger L, Schubeler D. Transcription factor occupancy can mediate active turnover of DNA methylation at regulatory regions. *PLoS Genet*. 2013;9:e1003994.
- Thomassin H, Flavin M, Espinas ML, Grange T. Glucocorticoid-induced DNA demethylation and gene memory during development. *EMBO J*. 2001;20:1974–83.
- Kendler KS, Karkowski LM, Prescott CA. Causal relationship between stressful life events and the onset of major depression. *Am J Psychiatry*. 1999;156:837–41.
- Molnar BE, Buka SL, Kessler RC. Child sexual abuse and subsequent psychopathology: results from the National Comorbidity Survey. *Am J Public Health*. 2001;91:753–60.
- Kress C, Thomassin H, Grange T. Local DNA demethylation in vertebrates: how could it be performed and targeted? *FEBS Lett*. 2001;494:135–40.
- Kress C, Thomassin H, Grange T. Active cytosine demethylation triggered by a nuclear receptor involves DNA strand breaks. *Proc Natl Acad Sci U S A*. 2006;103:11112–7.
- Zannas AS, Wiechmann T, Gassen NC, Binder EB. Gene-stress-epigenetic regulation of FKBP5: clinical and translational implications. *Neuropsychopharmacology*. 2016;41:261–74.
- Matosin N, Halldorsdottir T, Binder EB. Understanding the molecular mechanisms underpinning gene by environment interactions in psychiatric disorders: the FKBP5 model. *Biol Psychiatry*. 2018;83:821–30.
- Klengel T, Mehta D, Anacker C, Rex-Haffner M, Pruessner JC, Pariante CM, Pace TW, Mercer KB, Mayberg HS, Bradley B, et al. Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nat Neurosci*. 2013;16:33–41.
- Klengel T, Binder EB. FKBP5 allele-specific epigenetic modification in gene by environment interaction. *Neuropsychopharmacology*. 2015;40:244–6.
- Non AL, Hollister BM, Humphreys KL, Childebayeva A, Esteves K, Zeanah CH, Fox NA, Nelson CA, Drury SS. DNA methylation at stress-related genes is associated with exposure to early life institutionalization. *Am J Phys Anthropol*. 2016;161:84–93.
- Tyrka AR, Ridout KK, Parade SH, Paquette A, Marsit CJ, Seifer R. Childhood maltreatment and methylation of FKBP5 binding protein 5 gene (FKBP5). *Dev Psychopathol*. 2015;27:1637–45.
- Yehuda R, Daskalakis NP, Bierer LM, Bader HN, Klengel T, Holsboer F, Binder EB. Holocaust exposure induced intergenerational effects on FKBP5 methylation. *Biol Psychiatry*. 2016;80:372–80.
- Buchmann AF, Holz N, Boecker R, Blomeyer D, Rietschel M, Witt SH, Schmidt MH, Esser G, Banaschewski T, Brandeis D, et al. Moderating role of FKBP5 genotype in the impact of childhood adversity on cortisol stress response during adulthood. *Eur Neuropsychopharmacol*. 2014;24:837–45.
- Ising M, Depping AM, Siebertz A, Lucae S, Unschuld PG, Kloiber S, Horstmann S, Uhr M, Muller-Myhsok B, Holsboer F. Polymorphisms in the FKBP5 gene region modulate recovery from psychosocial stress in healthy controls. *Eur J Neurosci*. 2008;28:389–98.
- Halldorsdottir T, Binder EB. Gene x environment interactions: from molecular mechanisms to behavior. *Annu Rev Psychol*. 2017;68:215–41.
- Hawn SE, Sheerin CM, Lind MJ, Hicks TA, Marraccini ME, Bountress K, Bacanu SA, Nugent NR, Amstadter AB. GxE effects of FKBP5 and traumatic life events on PTSD: a meta-analysis. *J Affect Disord*. 2019;243:455–62.
- Wang Q, Shelton RC, Dwivedi Y. Interaction between early-life stress and FKBP5 gene variants in major depressive disorder and post-traumatic stress disorder: a systematic review and meta-analysis. *J Affect Disord*. 2018;225:422–8.
- Paakinaho V, Makkonen H, Jaaskelainen T, Palvimo JJ. Glucocorticoid receptor activates poised FKBP5 locus through long-distance interactions. *Mol Endocrinol*. 2010;24:511–25.
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 2012;485:376–80.
- Ong CT, Corces VG. CTCF: an architectural protein bridging genome topology and function. *Nat Rev Genet*. 2014;15:234–46.
- Roeh S, Wiechmann T, Sauer S, Kodel M, Binder EB, Provencal N. HAM-TBS: high-accuracy methylation measurements via targeted bisulfite sequencing. *Epigenetics Chromatin*. 2018;11:39.
- Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, Szalaj P, Trzaskoma P, Magalska A, Wlodarczyk J, Ruszczycki B, et al. CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell*. 2015;163:1611–27.
- Renda M, Baglivo I, Burgess-Beusse B, Esposito S, Fattorusso R, Felsenfeld G, Pedone PV. Critical DNA binding interactions of the insulator protein CTCF: a small number of zinc fingers mediate strong binding, and a single finger-DNA interaction controls binding at imprinted loci. *J Biol Chem*. 2007;282:33336–45.

28. Wiench M, John S, Baek S, Johnson TA, Sung MH, Escobar T, Simmons CA, Pearce KH, Biddie SC, Sabo PJ, et al. DNA methylation status predicts cell type-specific enhancer activity. *EMBO J*. 2011;30:3028–39.
29. Kheradpour P, Kellis M. Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. *Nucleic Acids Res*. 2014;42:2976–87.
30. Menke A, Arloth J, Putz B, Weber P, Klengel T, Mehta D, Gonik M, Rex-Haffner M, Rubel J, Uhr M, et al. Dexamethasone stimulated gene expression in peripheral blood is a sensitive marker for glucocorticoid receptor resistance in depressed patients. *Neuropsychopharmacology*. 2012; 37:1455–64.
31. Lee RS, Mahon PB, Zandi PP, McCaul ME, Yang X, Bali U, Wand GS. DNA methylation and sex-specific expression of FKBP5 as correlates of one-month bedtime cortisol levels in healthy individuals. *Psychoneuroendocrinology*. 2018;97:164–73.
32. Resmini E, Santos A, Aulinas A, Webb SM, Vives-Gilbert Y, Cox O, Wand G, Lee RS. Reduced DNA methylation of FKBP5 in Cushing's syndrome. *Endocrine*. 2016;54:768–77.
33. Glad CA, Andersson-Assarsson JC, Berglund P, Bergthorsdottir R, Ragnarsson O, Johannsson G. Reduced DNA methylation and psychopathology following endogenous hypercortisolism - a genome-wide study. *Sci Rep*. 2017;7:44445.
34. Parade SH, Parent J, Rabemananjara K, Seifer R, Marsit CJ, Yang BZ, Zhang H, Tyrka AR. Change in FK506 binding protein 5 (FKBP5) methylation over time among preschoolers with adversity. *Dev Psychopathol*. 2017;29:1627–34.
35. Yeo S, Enoch MA, Gorodetsky E, Akhtar L, Schuebel K, Roy A, Goldman D. The influence of FKBP5 genotype on expression of FKBP5 and other glucocorticoid-regulated genes, dependent on trauma exposure. *Genes Brain Behav*. 2017;16:223–32.
36. Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature*. 2000;405:482–5.
37. Flavahan WA, Drier Y, Liao BB, Gillespie SM, Venteicher AS, Stemmer-Rachamimov AO, Suva ML, Bernstein BE. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature*. 2016;529:110–4.
38. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature*. 2000;405:486–9.
39. Kanduri C, Pant V, Loukinov D, Pugacheva E, Qi CF, Wolffe A, Ohlsson R, Lobanenkov VV. Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. *Curr Biol*. 2000;10:853–6.
40. Wang H, Maurano MT, Qu H, Varley KE, Gertz J, Pauli F, Lee K, Canfield T, Weaver M, Sandstrom R, et al. Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res*. 2012;22:1680–8.
41. Nakahashi H, Kieffer Kwon KR, Resch W, Vian L, Dose M, Stavreva D, Hakim O, Pruett N, Nelson S, Yamane A, et al. A genome-wide map of CTCF multivalency redefines the CTCF code. *Cell Rep*. 2013;3:1678–89.
42. Zannas AS, Arloth J, Carrillo-Roa T, Iurato S, Roh S, Ressler KJ, Nemeroff CB, Smith AK, Bradley B, Heim C, et al. Lifetime stress accelerates epigenetic aging in an urban, African American cohort: relevance of glucocorticoid signaling. *Genome Biol*. 2015;16:266.
43. Seifuddin F, Wand G, Cox O, Pirooznia M, Moody L, Yang X, Tai J, Boersma G, Tamashiro K, Zandi P, Lee R. Genome-wide Methyl-Seq analysis of blood-brain targets of glucocorticoid exposure. *Epigenetics*. 2017;0.
44. Metivier R, Gallais R, Tiffocche C, Le Peron C, Jurkowska RZ, Carmouche RP, Ibberson D, Barath P, Demay F, Reid G, et al. Cyclical DNA methylation of a transcriptionally active promoter. *Nature*. 2008;452:45–50.
45. Murgatroyd C, Spengler D. Polycomb binding precedes early-life stress responsive DNA methylation at the Avp enhancer. *PLoS One*. 2014;9:e90277.
46. Volk N, Pape JC, Engel M, Zannas AS, Cattaneo N, Cattaneo A, Binder EB, Chen A. Amygdalar microRNA-15a is essential for coping with chronic stress. *Cell Rep*. 2016;17:1882–91.
47. Arloth J, Bogdan R, Weber P, Frishman G, Menke A, Wagner KV, Balsevich G, Schmidt MV, Karbalai N, Czamara D, et al. Genetic differences in the immediate transcriptome response to stress predict risk-related brain function and psychiatric disorders. *Neuron*. 2015;86:1189–202.
48. Menke A, Arloth J, Best J, Namendorf C, Gerlach T, Czamara D, Lucae S, Dunlop BW, Crowe TM, Garlow SJ, et al. Time-dependent effects of dexamethasone plasma concentrations on glucocorticoid receptor challenge tests. *Psychoneuroendocrinology*. 2016;69:161–71.
49. Heuser I, Yassouridis A, Holsboer F. The combined dexamethasone/CRH test: a refined laboratory test for psychiatric disorders. *J Psychiatr Res*. 1994;28:341–56.
50. Hundt W, Zimmermann U, Pottig M, Spring K, Holsboer F. The combined dexamethasone-suppression/CRH-stimulation test in alcoholics during and after acute withdrawal. *Alcohol Clin Exp Res*. 2001;25:687–91.
51. Ramp C, Eichelkraut A, Best J, Czamara D, Rex-Haffner M, Uhr M, Binder EB, Menke A. Sex-related differential response to dexamethasone in endocrine and immune measures in depressed in-patients and healthy controls. *J Psychiatr Res*. 2018;98:107–15.
52. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 2001;29:e45.
53. consortium B. Quantitative comparison of DNA methylation assays for biomarker development and clinical applications. *Nat Biotechnol*. 2016;34: 726–37.
54. Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, Guan W, Xu T, Elks CE, Aslibekyan S, et al. Epigenetic signatures of cigarette smoking. *Circ Cardiovasc Genet*. 2016;9:436–47.
55. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*. 2012;13:86.
56. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Software*. 2015;67.
57. Fox J, Weisberg S. An R companion to applied regression; 2011.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



5.4 Manuscript I - An intronic structural variant in *FKBP5* affects GRE enhancer activity through the modulation of topological associated domain interaction frequencies and regulates the responsiveness to glucocorticoid exposure

Wiechmann T., S. Roeh, D. Czamara, T. Klengel, S. Sauer, T. Rein, N. Provençal, E. B. Binder

Manuscript ahead of submission

An intronic structural variant in *FKBP5* affects GRE enhancer activity through the modulation of topologically associated domain interaction frequencies and regulates the responsiveness to glucocorticoid exposure

Tobias Wiechmann¹, Simone Röh¹, Darina Czamara¹, Torsten Klengel², Susann Sauer¹, Theo Rein¹, Nadine Provençal^{1,3}, Elisabeth B. Binder^{1,4}

¹Dep. of Translational Research in Psychiatry, MPI of Psychiatry, Munich, Germany

²Dep. of Psychiatry, McLean Hospital, Harvard Medical School, Belmont, MA, USA

³Faculty of Health Sciences, Simon Fraser University and BC Children's Hospital Research Institute, Vancouver, BC, Canada

⁴Dep. of Psychiatry and Behavioral Sciences, Emory University Medical School, Atlanta, GA, USA

Abstract

The risk for developing stress-related diseases is shaped by the interaction of environmental and genetic risk factors. The underlying molecular mechanisms integrating environmental and genetic cues have not yet been fully elucidated. We aimed to test, whether the allele of an intronic 3.3 kb large variant esv3608688 (insertion / deletion), located within in a disease-associated haplotype at the *FKBP5* locus, have an effect on glucocorticoid-dependent gene regulation and how this can be explained on a molecular level. We observed that the deletion allele was associated with a higher induction of *FKBP5* expression after activation of the glucocorticoid receptor (GR) in lymphoblastoid cell lines (LCLs). Furthermore, we could show that the esv3608688 allele can moderate the effect on *FKBP5* gene expression regulation of a previously described functional single nucleotide polymorphism rs1360780 (Klengel et al, 2013). We identified that the stabilization of architectural and enhancer-promoter loops is a common feature of the factors (T-allele, deletion and GR activation) leading to increased *FKBP5* mRNA expression. In addition, we observed differentially methylated sites depending on the allele status for esv3608688 within key regulatory sites of *FKBP5*. Our data proposes a model, that this gene expression response is most likely result of an increased activity of GRE enhancers due to the stabilization of chromatin interactions. This study highlights molecular mechanism how genetic and environmental factors can be integrated fine-tune *FKBP5* expression levels, potentially involved in leading to *FKBP5* disinhibition and therefore shape the risk for developing a stress-related psychiatric disease.

Keywords: Chromatin structure / DNA methylation / *FKBP5* / Gene by environment / Glucocorticoids

Introduction

The development of stress-related psychiatric disorders involves the integration of genetic and environmental factors (Caspi & Moffitt, 2006; Kendler et al, 1999; Kendler et al, 1995; Klengel et al, 2013; Molnar et al, 2001). It has been shown by epidemiological, family and molecular genetic studies that genetic predisposition as well as stressful or traumatic life events are important risk factors for psychiatric disorders (Caspi & Moffitt, 2006).

The molecular basis of the gene x environment interactions leading to disease are far away from being fully understood but necessary to translate the results from genetic association studies to the clinic (Krijger & de Laat, 2016). However, epigenetic mechanisms are potentially playing a role in the etiology of various human diseases (Adcock et al, 2005; Esteller, 2007; Gerken et al, 2007; Handel et al, 2010; Varambally et al, 2008; Weksberg et al, 2002) including depression (McGowan et al, 2009), Posttraumatic Stress Disorder (PTSD) (Klengel et al, 2013), major psychosis (Mill et al, 2008) and autism (Schanen, 2006). It is thought that epigenetic modifications and mechanisms such as DNA methylation, histone modifications, non-coding RNA and chromatin conformation changes (Levenson & Sweatt, 2005; Maddox et al, 2013; Mill & Petronis, 2007; Miller et al, 2008; Mitchell et al, 2014; Tsankova et al, 2007) are involved in mediating these interactions between the genotype and the environment (GxE) by changing the expression of genes implicated in stress-related psychiatric diseases (Jaenisch & Bird, 2003; Yehuda & LeDoux, 2007). Epigenetic changes due to GxE interaction effects are supposed to be installed during sensitive periods (e.g. early development) and remain stable over time (Klengel et al, 2013; Klengel et al, 2014).

In terms of stress-related diseases, genes which have shown to be part of GxE are often involved in the stress hormone system such as the glucocorticoid receptor (GR, *NR3C1*), corticotropin releasing hormone receptor 1 (*CRHR1*) and FK506 binding protein 5 (*FKBP5*), which play a key role in the regulation of the HPA axis (Binder, 2009; Bradley et al, 2008; Klengel et al, 2013; Klengel et al, 2014; Polanczyk et al, 2009; Zannas & Binder, 2014). A dysregulated stress hormone system is a strong risk factor for developing stress-related psychiatric disorders and therefore regulating factors of this system display promising therapeutic targets for the treatment of stress-regulated diseases like PTSD or Depression (Binder, 2009; Klengel et al, 2013). A great number of studies (comprising over 12 000 subjects) have linked

interactions between *FKBP5* genotypes and pathological phenotypes (for review see Matosin et al, 2018; Zannas et al, 2016) and often reported that genotypes associated with higher risk for the pathological phenotypes are also associated with higher *FKBP5* induction and prolonged cortisol responses (Zannas et al, 2016). In fact, a function of the co-chaperone FKBP51 is to inactivate GR in the cytoplasm, constituting in an ultra-short negative feedback loop, which regulates GR signaling and through this modulates the actions of glucocorticoids (Zannas et al, 2016).

To understand how genotypes can influence the complex fine-tuning process of *FKBP5* gene expression regulation in health and disease, one has to understand the molecular regulatory landscape of this gene. A key study by Paakinaho et al (2010) highlighted that GR elements in long distance from the promoter but also in the introns of the *FKBP5* gene, are very important to understand gene expression regulation and includes the consideration of the locus chromatin structure. The development of the C-Techniques enabled researchers to get more insights into how the genome folds in 3D and lead to new concepts about structural and functional units of chromosomes (for review see Krijger & de Laat, 2016). This development enabled to bring the genetic variations associated with disease, which are often located in the non-coding parts of the genome, into the context of the 3D regulation of genes (Krijger & de Laat, 2016). One of the most prominent single nucleotide polymorphism associated with psychiatric disorders in *FKBP5* is rs1360780 (Binder et al, 2008; Binder et al, 2004; Matosin et al, 2018; Zannas et al, 2016). This polymorphism lies in close proximity to a GRE in intron 2 and alters the mRNA and protein production after GR activation via an altered 3D chromatin structure (Klengel et al, 2014).

The T-allele of rs1360780 forms a TATA-box binding motif, which enables an efficient interaction of the GRE in intron 2 and the transcription start site leading to *FKBP5* mRNA induction after GR activation. The C-allele on the other hand destabilizes the enhancer-promoter interaction due to the lack of the TATA-box binding motif, which leads to a decreased *FKBP5* induction (Klengel et al, 2013). Moreover, we showed evidence for a GxE interaction of rs1360780 allele and childhood abuse, which leads to allele-specific demethylation of CpGs in a GRE motif in intron 7 in T-allele carrier that were exposed to childhood abuse. The change in methylation further leads to an increased inducibility of the *FKBP5* gene after GC exposure and has long-term effects on the stress hormone system, immune cell

function and brain areas important for stress-regulation (Klengel et al, 2013). These experiments highlighted the potential that allele-specific epigenetic modifications can contribute to GxE interactions and moderate the long-term effects of stress and subsequently the likelihood of developing stress-related psychiatric phenotypes during the lifetime.

In the same haplotype of *FKBP5*, which comprises the rs1360780 polymorphism, a larger structural variant can be found (Pelleymounter et al, 2011). This 3.3 kb large insertion / deletion (INDEL) is located in intron 1 of the *FKBP5* gene. Here, we want to investigate if the structural variant has an influence on *FKBP5* gene expression regulation and if the effects mediated by rs1360780 are influenced by the INDEL allele status. Moreover, we asked how these effects onto gene expression regulation can be explained on a molecular level and if the context of INDEL allele are important for the development of psychiatric disease phenotypes.

To answer these questions, we investigated the effects of the allele combinations of the two genetic variants onto *FKBP5* gene expression induction after Dexamethasone (DEX) treatment in LCLs using an allele-specific gene expression assay for *FKBP5*. To explore, how these transcriptional effects on *FKBP5* induction are reflected on the DNA level we analyzed changes of the chromatin structure (4C-Seq) and DNA methylation (HAM-TBS).

Results

The INDEL alleles modulate GC-dependent *FKBP5* expression induction & modulate rs1360780 effects

In order to better understand the temporal pattern of the *FKBP5* induction due to GR activation, we designed a qPCR assay to detect nascent RNA (see method section and supplemental table 1). We choose two LCLs cell lines (both heterozygote for rs1360780, GM19098 is homozygote for the insertion and GM18516 homozygote for the deletion) and performed a time response experiment under continuous DEX (100nM) stimulation for 1 h, 2 h, 3 h, 4 h, 8 h and 24 h (Figure 1 A). The treatment with DEX results in a fast stimulation of nascent *FKBP5* RNA, which peaks at 2 to 3 h (Mean fold change: 4.698 ± 1.016 & 4.674 ± 0.559 respectively) and then stabilizes after 4 h at a fold change of 3.779 ± 0.697 , which slightly decreases to 3.386 ± 0.705 after 24 h. The mRNA of *FKBP5* strongly increases after 2-4 h after continuous DEX treatment and stabilizes at a fold change at of 2.963 ± 0.176 and 3.336 ± 0.339 after 4 h and 8 h respectively. As observed for the nascent RNA, *FKBP5* mRNA slightly reduces to a fold change of 2.711 ± 0.2218 after 24 h.

Subsequently, we asked, whether the alleles of the INDEL have the potential to regulate the GC-dependent induction of *FKBP5* mRNA and if there is interplay of the INDEL and rs1360780 alleles on *FKBP5* gene expression regulation. We choose to DEX stimulate (4 h, 100 nM DEX) 10 LCLs, which were heterozygote for rs1360780 allele but homozygote for esv3608688 (deletion n = 5, insertion n = 5). Using a hydrolysis probe qPCR assay that detects total *FKBP5* mRNA expression, we observed as expected that GR activation via DEX leads to an induction of *FKBP5* mRNA (see Figure 1 B). Moreover, we observed that the deletion is associated with increased *FKBP5* induction over the insertion due to the DEX treatment (Mean fold change difference = 1.08, P-Value < 0.01 see Figure 1 B). We developed an allele-specific gene expression qPCR setup to discriminate the transcripts produced from the rs1360780 T- or C-allele under a homozygote structural variant background. Due to the fact that the rs1360780 lies within an intron of *FKBP5*, we designed our allele-specific probes on rs3800373, which is in high LD with rs1360780 and located in the 3'UTR of the *FKBP5* mRNA (see method section and supplementary table 1). The proportions of rs1360780 allele under an insertion or deletion background were obtained using the allele-specific qPCR setup in the 10 LCLs and DEX treatment. We

found that more *FKBP5* transcripts are generated from the T-allele over C-allele in each of the conditions (Mean Proportion Difference $_{INS_VEH} = 0.09$, P-Value < 0.001; Mean Proportion Difference $_{INS_DEX} = 0.16$, P-Value < 0.001; Mean Proportion Difference $_{DEL_VEH} = 0.21$, P-Value < 0.001; Mean Proportion Difference $_{DEL_DEX} = 0.22$, P-Value < 0.001; see Figure 1 C). Furthermore, the highest proportion of T-allele transcripts is achieved under a deletion background, leading to a significant genotype effect under vehicle conditions (Mean Proportion Difference $_{VEH} = 0.06$, P-Value < 0.05) but not significant when treated with DEX (Figure 1 C). Subsequently, we analyzed the allele-specific expression levels of *FKBP5* (Figure 1 D). The total *FKBP5* transcript levels depended on the allele combination of the two variants and we observed an additive genotype effect (Figure 1 D). The lowest GC-induced *FKBP5* transcript levels are detected in the combination of the C + insertion alleles, whereas the highest transcript levels are observed under the combination of the T + deletion alleles (Fold change $_{C-INS} = 2.10 \pm 0.30 > \text{Fold change}_{T-INS} = 3.00 \pm 0.50 > \text{Fold change}_{T-DEL} = 3.68 \pm 0.79$; Mean Difference $_{T-INS \text{ vs } C-INS} = 0.85$, P-value < 0.001; Mean Difference $_{T-DEL \text{ vs } C-INS} = 1.58$, P-value < 0.001; Mean Difference $_{T-DEL \text{ vs } T-INS} = 0.73$, P-value < 0.01).

FKBP5 responsiveness to glucocorticoids is modulated by structural variant allele via an altered TAD structure

To understand how the structural variant can influence the complex fine-tuning process of *FKBP5* gene expression regulation, one has to understand the molecular regulatory landscape of this gene. Therefore, we performed an *in silico* analysis comprising the key features of the *FKBP5* locus structure using available literature and datasets (Consortium, 2004; Paakinaho et al, 2010; Rao et al, 2014; Tang et al, 2015) (Figure 2). The topology of the *FKBP5* locus is formed by CCCTC-binding factor (CTCF), which generates a architectural chromatin loop, the topologically associating domain (TAD), in which RNAPII and other proteins like transcription factors (e.g. GR) further bind and substructure the locus to regulate *FKBP5* gene expression during the development, in different tissues and in response to environmental influences. In order to investigate how the esv3608688 allele can affect *FKBP5* gene expression we used a circularized chromosome conformation

capture (4C) sequencing approach, to obtain high-resolution interaction profiles (van de Werken et al, 2012b). We designed one viewpoint on the proximal TAD boundary (hg19 / chr6:35704131-35704628, see Figure 3) and a second on the transcription start site (TSS) (hg19 / chr6:35655866-35656827, see Figure 4 & supplementary Figure 2) using DpnII and NlaIII as restriction enzymes. This enabled us to obtain, the interaction profile from the proximal TAD boundary (to detect effects on TAD interactions) and TSS (to detect effects on promoter-enhancer interactions) from the same 4C DNA samples. We choose to generate high-resolution interaction profiles in triplicates of the two viewpoints from GM18516 (homozygote deletion, heterozygote SNP) and GM19098 (homozygote insertion, heterozygote SNP) under vehicle (EtOH) and DEX treatment (100nM, 4h) conditions. The 4C profiles from the TAD boundary and TSS are reproducible between replicates and specific interactions described in the literature are detected (for further details see supplementary Figure 1). Due to the fact that we can use the same restriction enzymes for the definition of the viewpoints, we detect aligned interaction frequency peaks between both viewpoints. As an example, the interactions at the intragenic TAD boundary and intronic GRE at intron 5 are observable in both viewpoints (Supplemental Figure 1). This indicates that the TAD boundary and the TSS are linked to some degree through the 3D conformation of the locus, highlighting the finding from Tang et al (2015) that RNAPII transcription factories are spatially associated with the CTCF / Cohesin foci.

FKBP5 is known to be a pre-poised locus (Paakinaho et al, 2010), therefore the chromatin structure of the locus is already established before the stimulation with glucocorticoids, which is reflected with a visible interaction profile in the VEH samples (Supplemental Figure 1). To identify INDEL allele and DEX treatment effects onto *FKBP5* interaction profiles we used the 4C-cker pipeline (Raviram et al, 2016). In our quantitative analysis we especially focused on aggregates of differentially interacting regions (DIRs) at an FDR = 0.05 to less likely interpret false positive regulated sites as being regulated. The reason is that interacting regions to the viewpoint are reflected by an increase and decrease of interaction frequencies (shape of Gaussian distribution) and therefore, if the FDR is not too stringent, called DIRs of true regulated interactions should also appear in an aggregate shape of a Gaussian distribution. As expected, we detect differentially interacting regions (DIRs) at the INDEL position in both viewpoints and conditions. The 4C interaction profiles of the insertion allele show an interaction frequency peak at the position of the INDEL position in both

viewpoints (Figure 3 & 4). This indicates that the region of the INDEL allele is in proximity to the TSS and the CTCF foci due to the 3D conformation of the locus. Moreover, aggregates of DIRs are detected at the intronic CTCF to the TAD viewpoint, which are significantly increased under the deletion allele (Figure 3, Quantification – Genotype tracks). Additionally, DIRs in the same direction are detected from the intronic CTCF binding site over the gene body to the TAD viewpoint (Figure 3, Quantification – Genotype tracks). This indicates a more intense TAD forming CTCF interaction from the locus proximal to the intronic site under the deletion allele in comparison to the interaction profile of the insertion allele. A similar distribution of DIRs was found when comparing the interaction profiles of the two genotypes of the TSS viewpoint (Figure 4, Quantification – Genotype tracks). But in this viewpoint, no clear aggregates of DIRs at the intronic CTCF binding site are observable, indicating that the DIRs detected by the TSS viewpoint may originate from the increased TAD forming CTCF interaction under the deletion allele which is detected by the TAD viewpoint. The DEX treatment leads to an increase of interaction frequencies from the TAD viewpoint to the intronic GRE binding site in intron 5, irrespectively from the INDEL allele status (Figure 3, Quantification – Treatment tracks). Aggregates of DIRs are observable due to the DEX treatment at the proximal long-range GRE enhancer when quantifying the interaction profiles of the TSS viewpoint (Figure 4, Quantification – Treatment tracks).

In summary, the quantification of the high resolution interaction profiles of the TAD boundary and TSS viewpoint showed that the INDEL allele have the potential to modulate the TAD forming CTCF interaction (deletion allele increased the interaction frequencies) and that the DEX treatment leads to an increase of interacting frequencies over the whole gene body (TAD viewpoint) and specifically at the long-range GRE enhancer (TSS viewpoint).

The allele of the structural variant affect the methylation levels of key functional sites in the *FKBP5* locus

To further characterize the influence of the structural variant onto the DNA level, we aimed to investigate the DNA methylation levels of CpGs in key functional sites (TAD boundaries, TSS, intronic GREs and proximal GRE enhancer; see Figure 5 & supplementary table 3) in the *FKBP5* locus and regions within and around the

structural variant. Therefore, we analyzed the methylation levels of these regions in 10 LCLs carrying the insertion or deletion ($n = 5$, INDEL = INS / INS, rs1360780 = T / C; $n = 5$, INDEL = DEL / DEL, rs1360780 = T / C) with and without a DEX treatment (4h, 100 nM). We found that the allele of esv3608688 had an effect on the methylation level of 36 CpGs under the VEH (Total: 33 CpGs, 23 CpGs hypermethylated & 10 CpGs hypomethylated under the insertion allele) and DEX (Total: 27 CpGs, 17 CpGs hypomethylated & 10 CpGs hypomethylated under the insertion allele) condition with an overlap of 24 CpGs between both conditions (Figure 5, see Δ Methylation track). We observed a mean genotype effect size of $|19.5 \pm 6.8|$ % (Max: 41.65, Min: -23.32 when $\Delta\text{Meth}_{\text{genotype}} = \text{INS}_{\text{VEH}} - \text{DEL}_{\text{VEH}}$) and $|20.2 \pm 7.2|$ % (Max: 39.6, Min: -25.7 when $\Delta\text{Meth}_{\text{genotype}} = \text{INS}_{\text{DEX}} - \text{DEL}_{\text{DEX}}$) under the VEH and DEX condition, respectively.

The CpGs within the insertion allele (h19 / chr6:35626434-35629744) and the distal flank (hg19/chr6:35625879-35626155) showed a high level of methylation (Insertion_{VEH} = 90.7 ± 8.5 %, $n = 33$; Distal flank Insertion_{VEH} = 94.6 ± 2.0 %, $n = 14$), whereas at the proximal flank (hg19/chr6:35630048-35631736) the methylation levels were more variable (Proximal flank Insertion_{VEH} = 80.7 ± 20.4 %, $n = 14$). At the flanking regions of the INDEL, we found 4 CpGs (1 CpG at the distal flank & 3 CpGs at the proximal flank), which showed increased methylation levels in the presents of the insertion allele in comparison to the deletion allele (Mean effect size_{VEH} = 24.0 ± 9.7 , $n = 4$; Mean effect size_{DEX} = 28.9 ± 9.4 , $n = 3$). One region between the INDEL and TSS (hg19 / chr6:35635518-35636055) included in our methylation analysis showed an array of CpGs located next to each other to be regulated by the INDEL allele (see Figure 5 B). 9 out of 10 CpGs in this regions showed a significant higher methylation levels under the insertion allele and VEH conditions with a mean effect size of 16.6 ± 2.9 %. When treating with DEX the number of CpGs with a genotype effect reduced to 5 out of 10 CpGs with a mean effect size of 18.3 ± 3.1 %. Further genotype effects on methylation were found close or at GREs in intron 5 (a total of 3 CpGs hypomethylated under the insertion allele (Mean effect size_{VEH} = 19.0 ± 1.1 % ($n = 2$) & Mean effect size_{DEX} = 17.3 ± 2.8 % ($n = 3$)) and proximal Enhancer (a total of 3 CpG hypermethylated (Mean effect size_{VEH} = 16.9 ± 0.6 % ($n = 3$) & Mean effect size_{DEX} = 15.7 ± 1.2 % ($n = 2$)) and 8 CpGs hypomethylated (Mean effect size_{VEH} = 19.0 ± 3.0 % ($n = 7$) & Mean effect size_{DEX} = 18.6 ± 3.7 % ($n = 6$)) under the insertion allele (see Figure 5 A & B). Interestingly, 10 out of the total 11 CpGs

showing lower methylation levels due to the insertion allele in both treatment conditions are located in regions in with known GR binding sites. Before TSS we found 1 CpG, which showed a hypermethylation of 14.4% under the insertion when treated with DEX but not in the VEH condition. Moreover, we found genotype effects onto the methylation of 8 CpGs close to CTCF binding sites at the proximal TAD boundary (a total of 2 CpGs hypermethylated under the insertion allele (Mean effect size $_{VEH} = 18.0 \pm 5.7 \%$ (n = 2) & Effect size $_{DEX} = 20.6 \%$ (n = 1)) & distal TAD boundary (a total of 2 CpGs; 1 CpG hypomethylated and 1 CpGs hypermethylated under the insertion allele (Mean effect size $_{VEH} = |15.7 \pm 2.5| \%$ (n = 2) & Mean effect size $_{DEX} = |16.1 \pm 0.1| \%$ (n = 2)), after the TSS (a total of 2 CpGs hypomethylated under the insertion allele (Mean effect size $_{VEH} = 39.1 \pm 3.6 \%$ (n = 2) & Mean effect size $_{DEX} = 37.7 \pm 2.7 \%$ (n = 2)) and after the 3'UTR (a total of 2 CpGs hypomethylated under the insertion allele (Mean effect size $_{VEH} = 15.0 \pm 0.1 \%$ (n = 2) & Mean effect size $_{DEX} = 15.6 \pm 2.3 \%$ (n = 2)) (see Figure 5 A). We did not detect any significant DEX treatment effects onto the DNA methylation within the *FKBP5* locus. Overall, we observed genotype effects onto DNA methylation at key functional sites in the *FKBP5* locus (proximal & distal TAD boundaries, TSS, GRE in intron 5 and GREs at the proximal enhancer), around the position of the INDEL and in between the region of the INDEL and TSS. This indicates that the allele of the INDEL have the potential to lead to adaptations on the DNA methylation level in close proximity to the INDEL position but also in greater distance within the *FKBP5* locus.

Discussion

Here we provide evidence that the alleles of a 3.3 kb large INDEL within a disease-associated haplotype in the *FKBP5* locus can modulate *FKBP5* expression response due to GC excess and therefore potentially to shape the risk for developing stress-related psychiatric disorders. Our data suggest that the INDEL allele not only affect the gene expression regulation of *FKBP5* (Figure 1 B) but also modulates the effect of the polymorphism rs1360780 under glucocorticoid exposure in LCLs (Figure 1 C). By using allele-specific gene expression assays and heterozygote LCLs for rs1360780 allele we could show that a higher proportion of mRNA is generated from the T-allele in comparison to the C-allele. This is in line with the results from Klengel et al (2013), who showed that the T-allele has a higher potential to induce gene expression in gene reporter assays. Interestingly, this differential gene expression response of rs1360780 allele is only observed in cells carrying the insertion allele. Moreover, we monitored the highest proportions of transcripts from the T-allele under the background of the deletion allele. In addition, the highest levels of total *FKBP5* transcripts under the excess of GCs were observed in LCLs combining the T- and deletion allele (Figure 1 D). These data indicate that the INDEL allele do not only moderate the differential effect of rs1360780 allele on *FKBP5* expression, moreover both variants affect total *FKBP5* transcript levels in an additive manner upon GC stimulation. The effects of rs1360780 allele onto *FKBP5* expression can be explained on a molecular level by the establishment of an enhancer-promoter loop due to the formation of a TATA-box binding motif by the T-allele (Klengel et al, 2013). This leads to higher proportions of mRNA at baseline and even higher mRNA proportions from the T-allele under GC stimulation.

In order to better understand how the large INDEL allele can modulate the rs136780 effects on a molecular level, we obtained high-resolution interaction profiles of the *FKBP5* locus TSS and TAD applying a 4C approach. A benefit of our 4C viewpoint design was the simultaneous monitoring of interaction profiles of both viewpoints in the same sample. This was achieved by choosing the same 4 base pair restriction enzymes for the definition of both viewpoints. The interaction profiles obtained from LCLs showed locus interactions that are in line with the reference literature and revealed architectural loops mediated by CTCF but also enhancer-promoter loops between the proximal GR and intronic GR-associated enhancers to the TSS (Klengel and Binder (2013); Paakinaho et al (2010), CTCF-ChiA-PET, PolIII-ChiA-PET from

Tang et al (2015), Hi-C data from Rao et al (2014), see supplemental Figure 1). A comparison of the interaction profiles from cells carrying the insertion or deletion allele showed by interaction frequency peaks at position of the INDEL that the variant can be in close proximity to the TSS and TAD boundary (Figure 3 & 4). This involvement in the *FKBP5* locus architectural and functional loop network indicates a potential ability to modulate gene expression, which is a first explanation for the INDEL expression effects monitored in these cells. Moreover, using the 4C-cker pipeline (Raviram et al, 2016) and DESeq2 (Love et al, 2014) were able to quantify the differences between the interaction profiles between the INDEL allele and DEX stimulation (Figure 3 & 4). The stimulation with DEX did not change the overall *FKBP5* interaction profiles indicating the described pre-established nature as a pre-poised gene (Jaaskelainen et al, 2011; Paakinaho et al, 2010). However, we quantified an increase of interaction frequencies from the TSS to the proximal GR-associated enhancer. This most likely indicates the observed gene expression induction due to GC excess through the long-range GR enhancer. In addition to the specific stabilization of promoter-enhancer looping, we detect an increase of interaction frequencies from the proximal TAD boundary over the *FKBP5* locus upon DEX stimulation (Figure 3). Together with the observation of increased interaction frequencies of the proximal TAD to the intronic TAD boundary in cells harboring the deletion allele, we conclude that the INDEL allele and GC excess may affect the stability of the architectural CTCF mediated-loop linking genotype and environmental signal. This interaction of genotype and environmental on the level of chromatin architecture could explain the *FKBP5* expression results monitored in (Figure 1 B). In fact, the concept that the stability of loops is important for regulating gene expression has been described (Hakim et al, 2011; Le Dily et al, 2014) and that pre-existing loops favor the response to external stimuli (Grbesa & Hakim, 2017; Jin et al, 2013). In regard to the idea that a dysregulation of architectural proteins / loops could lead to changes in the stability of promoter-enhancer interactions and lead to an altered response in gene expression (Antony et al, 2015; Quintin et al, 2014; Seitan et al, 2013), we reason the following model for the integration of INDEL allele and GC excess to modulate *FKBP5* gene expression (Figure 6).

The structural basis of *FKBP5* gene expression regulation is shaped by CTCF generating architectural loopings, which are described as TAD structure observed in Hi-C data (Lieberman-Aiden et al, 2009; Tang et al, 2015), see Figure 2). These TAD

loopings aid the formation of promoter-enhancer interactions, by restricting the promoter-enhancer looping within the TAD boundaries. For *FKBP5* pre-existing proximal & intronic GR enhancer loopings to the promoter (Paakinaho et al, 2010) are important to mediate the gene expression response after GR activation by GCs. The GC-induced expression changes are associated with increased interaction frequencies between long-range promoter-enhancer loops but also from the proximal TAD boundary to the body of the *FKBP5* gene. Genetic variants in *FKBP5* locus can modulate the response to GR activation by affecting either GR-associated promoter-enhancer loops or architectural CTCF mediated interactions. For rs1360780 a loop between a GR enhancer in intron 2 and the promoter is established due to the formation of an additional TATA-box binding motif of the T-allele, which can increase the GC-regulated *FKBP5* induction. The 3.3 kb INDEL on the other hand is associated to moderate the stabilization of the architectural loop between the proximal TAD and intronic TAD boundary. In this scenario, the deletion is associated to lead to increased interaction frequencies between TAD boundaries and production of *FKBP5* transcripts. The common feature in this model is that genetic variants and GR activation are associated to stabilizing loops within the *FKBP5* locus which is reflected by increased interaction frequencies and additive effects onto gene expression. In other words, the activity of the GR enhancers responsible for *FKBP5* gene expression regulation is on the one hand regulated by the activation of the hormone receptor itself but also their probability to find the *FKBP5* promoter. This probability is regulated by genetic variants modulating the stability of the architectural interactions (esv3608688 allele) but also promoter-enhancer interactions (rs1360780 allele). Indications for the concept that the enhancer activity is regulated by the overarching TAD stability have been found by Simonis et al (2007) by manipulating the *Shh* locus. To prove the causality of this concept for the *FKBP5* locus similar locus manipulation experiments using the CRISPR/Cas9 system need to be performed.

The involvement of other chromatin features defining of the activity of enhancers is plausible and DNA methylation levels have been proposed to be instructive for the activity of GR enhancers (Wiench et al, 2011). Using the HAM-TBS workflow we assessed the DNA methylation levels of 46 sites (see Figure 5) covering key functional regions of the *FKBP5* locus (TSS, TAD boundaries, proximal and intronic GR enhancer), the region of the INDEL and other sites between TSS and INDEL of

LCLs. The CpGs within the insertion allele show high methylation levels (Insertion_{VEH} = 90.7 ± 8.5 %, n = 33). We observed significant genotype effects onto DNA methylation levels of 36 CpGs at proximal & distal TAD boundaries, TSS, GRE in intron 5 and GREs at the proximal enhancer. Effects onto methylation levels due to the INDEL allele have been monitored in both directions and depend on the location of the respective CpGs. Moreover, we detected a genotype effect onto methylation levels in an array of CpGs located in intron 1 (hg19 / chr6:35635518-35636055). This region showed increased methylation levels in the background of the insertion allele and has not yet been described to include important regulatory sites. However, a potential role as an enhancer of this region in LCLs is implicated by the chromatin states of the ENCODE project (Consortium, 2012) using multiple epigenetic information and hidden markov models (Figure 5 A, ENCODE ChromHMM Genome segmentation track, Ernst and Kellis (2010); Hoffman et al (2012); Hoffman et al (2013). No DEX-induced methylation changes have been detected in the LCLs. Overall, next to genotype effects of the INDEL allele onto chromatin interactions we detect DNA methylation levels changes at key functional sites involved in the formation of architectural and promoter-enhancer loops indicating a potential integration of both epigenetic layers. The absence of DEX-inducible methylation changes in the LCLs could be explained on the one hand by the chosen treatment conditions or tissue-specific inability of the LCLs to respond to GR activation. To further dissect the importance of the INDEL allele onto chromatin interactions and DNA methylation levels to regulate *FKBP5* expression, experiments combining the manipulation of the DNA sequence but also the methylation levels of key regulatory sites of *FKBP5* should be applied. Moreover, the observation of eRNA being produced from active enhancers (Mikhaylichenko et al, 2018) could be used to detect *in vivo* enhancer activity instructed by the production of eRNA from the respective enhancer in *FKBP5*. This would help to fine map the contribution of each specific GR-associated enhancer and better understand which manipulations of the DNA sequence or DNA methylation levels affect the activity of the *FKBP5* enhancers. Although, there are preliminary data in PTSD patients that the INDEL allele moderate the symptom severity in a GxE manner (not shown here), further analyses need to be performed to enhance the evidence for the INDEL allele to moderate the development of stress related disease phenotypes. Due to the development of the

HAM-TBS (Roeh et al, 2018) workflow, there is now the possibility to investigate the INDEL effect onto the methylation level in patients.

In summary, the data presented in this chapter extends the proposed molecular mechanism (Klengel et al, 2013) integrating genetic and environmental factors at the *FKBP5* locus. We identified that the stabilization of architectural and enhancer-promoter loops is a common feature of the factors (T-allele, deletion and GR activation by GCs) leading to increased *FKBP5* mRNA expression. We and others (Symmons et al (2016) and discussed by Beagrie and Pombo (2016); Hakim et al (2011); Le Dily et al (2014)) find indications that this gene expression response is most likely result of an increased activity of enhancers due to the stabilization of chromatin interactions. However, further experiments manipulating the DNA sequence and DNA methylation levels of *FKBP5* key regulatory sites will be necessary to obtain causality. Moreover, further analysis in patients will help to clarify the role of INDEL allele in the development of stress-related phenotypes.

Materials and Methods

Cell culture of LCLs

LCLs (GM20356, GM20278, HG00634, HG00442, HG00114, HG00140, GM18516, GM19098, GM19371, GM19317) were purchased from the Coriell Institute for Medical Research (Camden, New Jersey, USA) and grown in RPMI 1640 medium (FG-1385, Biochrom GmbH, Berlin, GER) with 10 % FCS (Thermo Fisher scientific Inc., Schwerte, Germany) and 1% Antibiotic/Antimycotic (Thermo Fisher scientific Inc., Schwerte, Germany) in an incubator under 37 °C and 5 % CO₂ conditions. RNA and DNA was extracted from 0.8 million of cells. For RNA extraction, cells were pelleted and lyzed in 750 µl Trizol and RNA was isolated using the RNeasy mini kit (Qiagen GmbH, Hilden). DNA was extracted using the NucleoSpin® Tissue Kit (Macherey-Nagel GmbH & Co.KG, Dueren, GER) following manufacturer instructions.

Detection of *FKBP5* gene expression

Allele-specific *FKBP5* gene expression: The design of the allele-specific gene expression setup (see table 1 for primer sequences) was inspired by Soldner et al (2016). In order to obtain the ratio of mRNA generated from the rs1360780 allele (G / T) of heterozygote cells, we designed hydrolysis probes on the SNP rs3800373 (C / A) which is located in the 3'UTR of the *FKBP5* mRNA and in full LD with rs1360780 allele (rs1360780 G – rs3800373 C, rs1360780 T – rs3800373 A) (Binder et al, 2004). RNA was isolated using Qiazol (QIAGEN GmbH, Hilden, Germany) in combination with the RNeasy Mini Kit (QIAGEN GmbH, Hilden, Germany). Subsequently, DNA contamination in the RNA samples was removed by DNase digestion using the RNase-Free DNase set (QIAGEN GmbH, Hilden, Germany) and purified with the RNeasy MinElute Cleanup Kit (QIAGEN GmbH, Hilden, Germany). After bringing the RNAs to the same concentration with RNase-free water, cDNA was generated using the Superscript II reverse transcriptase (Thermo Scientific Inc., Schwerte, Germany). In order to test for DNA contamination some samples underwent an additional cDNA generation without adding reverse transcriptase (RT), these samples are referred as no RT controls. cDNAs were amplified in triplicates via real-time PCR using double-quenched hybridization probes (IDT PrimeTime® assays) with Taqman® fast advanced master mix (Thermo Scientific Inc., Schwerte,

Germany). All reactions were performed in a final volume of 10 µl on the LightCycler 480 Instrument II (Roche, Mannheim, Germany). Allele-specific multiplexing assays contained of a primer pair (final concentration of each primer = 0.5 µM) and two double-quenched hydrolysis probes (final concentration of each probe = 0.125 µM, Primer to Probe ratio = 2:1) conjugated to different fluorophores (FAM / HEX) dependent on the rs3800373 allele (FAM = C; HEX = A). Cycling condition for the allele-specific multiplexing assay was: 50 °C – 2 min, 95 °C – 20 s and 65x (95 °C – 3 s, 65 °C – 30 s). Cycling conditions for predesigned IDT PrimeTime® assays (Hs.PT.58.20523859 (*FKBP5*), Hs.PT.39a.22214836 (*GAPDH*) and Hs.PT.58.4154200 (*YWHAZ*)) were: 50 °C – 2 min, 95 °C – 20 s and 65x (95 °C – 3 s, 60 °C – 30 s). Each amplification plate included non-template controls for each assay (cDNA substituted by water in amplification mix used as negative control to detect cross-contamination).

For the analysis of relative quantification of allele-specific *FKBP5* expression after DEX treatment we used - similar to Soldner et al (2016) - the Pfaffl method (Pfaffl, 2001) that incorporated the following information: Total *FKBP5* expression (Hs.PT.58.20523859 used as reference assay for total *FKBP5* mRNA), allele-specific *FKBP5* multiplexing assay (target assay to determine proportion of mRNA species from each allele), two stable calibrator genes (Hs.PT.39a.22214836 (*GAPDH*) and Hs.PT.58.4154200 (*YWHAZ*)) and a 1 to 10 dilution of a RNA sample (to generate standard curve and calculate assay efficiencies). Cp values were extracted using the LightCycler 480 Instrument II analysis software. Optionally, we also included two DNA samples with known allele status for rs3800373 / rs1360780 to increase the reproducibility of setting the thresholds for each run within the LightCycler 480 Instrument II analysis software. Mean Cp values were calculated for each sample and assay. Allele-specific *FKBP5* expression proportions were calculated as follows: First, Cp values from Target assay (Cps from FAM (rs1360780 T) and HEX (rs1360780 C)) were subtracted from Cp values of reference assay. Second, the relative change is computed using the respective assay efficiencies. Third, relative changes are calculated for single alleles and normalized to obtain the allele-specific *FKBP5* expression proportions. Finally, to calculate allele-specific gene expression the total *FKBP5* expression (relative to the vehicle samples and normalized on mean calibrator genes) was computed and multiplied with allele-specific *FKBP5* expression proportions.

$$(1) \text{Target (T or C)} - \text{Reference (Cp)} = \Delta \text{Cp (T or C)}$$

$$(2) \text{Assay Efficiency}^{-\Delta \text{Cp (T or C allele)}} = \text{relative change (T or C)}$$

$$(3) \text{relative change (T or C)} * 0.5 = \text{relative change}_{\text{single allele}} \text{ (T or C)}$$

$$(4) \frac{\text{relative change}_{\text{single allele}} \text{ (T or C allele)}}{\text{relative change}_{\text{single allele}} \text{ (T)} + \text{relative change}_{\text{single allele}} \text{ (C)}}$$

$$= \text{Normalized allele proportion (T or C)}$$

$$(5) \text{relative change}_{\text{total FKBP5}} * \text{Normalized allele proportion (T or C)}$$

$$= \text{Allele specific FKBP5 expression (T or C)}$$

Nascent FKBP5 RNA: In order to detect nascent *FKBP5* RNA transcripts we designed primer and a double-quenched-probe targeting the intron carrying the SNP rs1360780 (see supplementary table 1 for primer sequences). RNA extraction and cDNA generation was achieved as described in the section allele-specific *FKBP5* gene expression. cDNAs were amplified in duplicates via real-time PCR using Taqman® fast advanced master mix (Thermo Scientific Inc., Schwerte, Germany). All reactions were performed in a final volume of 10 µl on the LightCycler 480 Instrument II (Roche, Mannheim, Germany). Cycling condition for the nascent *FKBP5* RNA assay was: 50 °C – 2 min, 95 °C – 20 s and 65x (95 °C – 3 s, 60 °C – 30 s). Each amplification plate included non-template controls for each assay (cDNA substituted by water in amplification mix used as negative control to detect cross-contamination) and a 1 to 10 dilution of a RNA sample (to generate standard curve and calculate assay efficiencies). For the analysis of nascent *FKBP5* RNA expression after a DEX treatment in comparison to the *FKBP5* mRNA expression (Hs.PT.58.20523859) we used the Pfaffl method (Pfaffl, 2001).

Circularized chromosome conformation capture (4C-Seq)

In order to investigate the effects of the INDEL allele onto the *FKBP5* chromatin structure we applied the high-resolution 4C-Seq approach developed by the de Laat group (Splinter et al, 2012; van de Werken et al, 2012a; van de Werken et al, 2012b), which uses 4bp cutter restriction enzymes to generate 4C template DNA. The 4C-ker method was used to analyze 4C-Seq data (Raviram et al, 2016).

Design of FKBP5 viewpoints: We designed primer for two viewpoints according to van de Werken et al (2012a) (see supplementary table 2). The first viewpoint was designed on the *FKBP5* transcription start site (hg19 / chr6: 35655865 - 35656827) and the second on proximal TAD boundary (hg19 / chr6: 35704130 - 35704628). Due to the fact that both viewpoints are defined by the same restriction enzymes DpnII and NlaIII, we are able to monitor the interaction profiles from both viewpoints of the same 4C DNA template.

4C template DNA generation and library preparation: The 4C DNA was generated with some slight adjustments according to Splinter et al (2012). In brief, 10 Mio. cells re-suspended in PBS/10% FCS (Thermo Scientific Inc., Schwerte, Germany), were fixated with formaldehyde (2% final concentration, (Thermo Scientific Inc., Schwerte, Germany) for 10 min on RT while tumbling. The cross-linking reaction was quenched on ice using glycine. After re-suspending fixed cells in cold lysis buffer, 50 strokes were applied with a cell douncer (Pestle A). DpnII (New England Biolabs GmbH, Frankfurt am Main, Germany) was used for the first restriction digestion and digestion efficiency was checked via qPCR using the LightCycler 480 Instrument II and LightCycler480 SYBR Green I Master Mix (Roche, Mannheim, Germany) according to Hagege et al (2007) (for primer sequences see table 1; Cycling condition: 95 °C – 5 min, 45x (95 °C – 10 s, 60 °C – 10 s, 72 °C – 10 s), 40°C – 10s). After in-solution ligation using T4 DNA ligase (New England Biolabs GmbH, Frankfurt am Main, Germany), cross-links and RNA were removed with Proteinase K (Merck KGaA, Darmstadt, Germany) and RNase A (Merck KGaA, Darmstadt, Germany) respectively. 3C DNA was cleaned up using Phenol-chloroform and DNA precipitation. NlaIII (New England Biolabs GmbH, Frankfurt am Main, Germany) was used for a second round of digestion to trim DNA circles and ligated in solution again. After DNA precipitation 4C DNA was cleaned up using the nucleotide removal kit (Qiagen GmbH, Hilden). The Expand Long Template Polymerase (Roche, Mannheim, Germany) was used to amplify 3.2 µg 4C DNA with viewpoint-specific primer introducing Illumina adapter sequences (single-indexed) in 16 separate reactions. The cycling condition for the TSS viewpoint was: 94 °C – 2 min, 29x (94 °C – 10 s, 64.9 °C – 60 s, 68°C – 3 min), 68°C – 5 min, 4 °C – ∞. The cycling condition for the proximal TAD viewpoint was: 94 °C – 2 min, 29x (94 °C – 10 s, 62.8 °C – 60 s, 68°C – 3 min), 68°C – 5 min, 4 °C – ∞. After pooling the 16 PCR reactions, AMPure XP beads (1.2x beads, Beckman Coulter, Krefeld, Germany) were used to clean up

PCR products. The purity and average amplicon size was obtained using a bioanalyzer (Agilent Technologies, Waldbronn, Germany) and DNA 7500 Chips. The Kapa Library qPCR quantification kit (Kapa Biosystems, Wilmington, MA) was used to calculate the molarity of 4C libraries. 4C libraries adjusted to 4 nM and pooled for sequencing.

Sequencing: Single-end sequencing (125 bp) was performed on an Illumina HiSeq 2500 Instrument (Illumina, San Diego, CA).

Sequencing data processing: The quality of the sequencing reads was checked with FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>). The sequence reads contained the primer sequence until the primary restriction enzyme recognition sequence (DpnII) and the captured fragments to the bait. The primer sequence until the DpnII recognition sequence was trimmed away from the read. The sequence of the capture fragments was kept until the appearance of the first (DpnII) or second (NlaIII) restriction enzyme recognition sequence. Afterwards the capture sequences were aligned to a reduced DpnII digested 4C reference genome of hg19 using Bowtie2 (Dryden et al, 2014). Each experiment needed to consist of > 1 million reads and > 40% of the capture sequences needed to align in cis (chromosome 6). All reads within 1000 bp around the viewpoints were discarded. Subsequently capture fragments were used to perform the near-bait analysis of the 4C-ker pipeline using default parameters (Raviram et al, 2016). Normalized read counts and quantitative analysis of the 4C signal were obtained using DESeq2 (Love et al, 2014). An FDR cut-off of 0.05 was used for the quantitative analysis.

Targeted bisulfite sequencing

Method development of the HAM-TBS approach has been described in detail in Roeh et al (2018). Below is a brief summary of the most salient methodological features.

Amplicon selection and amplification by PCR: We assessed in the 10 LCLs the methylation levels of 46 regions covering 306 CpGs within GR and/or CTCF binding sites as well as the around the INDEL of the *FKBP5* TAD (see Figure 5 & supplementary table 3) using the HAM-TBS approach (Roeh et al, 2018). Overall, 200 ng to 500 ng of DNA was used per sample and bisulfite treated using the EZ DNA Methylation Kit (Zymo Research, Irvine, CA). 20 ng of bisulfite converted DNA

were then used for each PCR amplification employing Takara EpiTaq HS Polymerase (Clontech, Saint-Germain-en-Laye, France) and 49 amplification cycles. PCR amplicons were then quantified with the Agilent 2200 TapeStation (Agilent Technologies, Waldbronn, Germany) and pooled in equimolar quantities for each sample. AMPure XP beads (Beckman Coulter, Krefeld, Germany) were used for a double size selection (200-500 bp) to remove primer dimers and high molecular DNA fragments.

Sequencing: Libraries were generated using the TruSeq DNA PCR-Free HT Library Prep Kit (Illumina, San Diego, CA) according to the manufacturer's instructions. Each library was quantified with the Qubit® 1.0 (Thermo Fisher Scientific Inc., Schwerte, Germany), normalized to 4 nM and pooled. Library concentration and fragment sizes were checked via Agilent's 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany) and quantitative PCR using the Kapa HIFI Library quantification kit (Kapa Biosystems, Wilmington, MA). Paired-end sequencing was performed on an Illumina MiSeq Instrument (Illumina, San Diego, CA) with their MiSeq Reagent Kit v3 (2 x 300-cycles) with the addition of 15% of PhiX Library.

Sequencing data processing: The quality of the sequencing reads was checked with FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>) and Illumina adapter sequences were removed using Cutadapt v.1.9.1. Bismark v.0.15.0 was used for the alignment to a restricted reference limited to our PCR targets. In order to stitch paired-end reads, an in-house Perl script has been developed to remove the low-quality ends of the paired-end reads if they overlapped. The methylation levels for all CpGs, CHGs and CHHs were quantified using the R package methylKit. The resulting DNAm calls were submitted to a quality control. PCR artifacts introducing false CpGs of low coverage at 0 or 100% methylation level were removed. CHH methylation levels were analyzed, and samples with insufficient bisulfite conversion rate (< 95%) were removed.

Genotyping

Genotyping PCR INDEL: The allele of the 3.3 kb large INDEL in intron 1 of the *FKBP5* gene were detected by separate PCR reactions using two primer pairs (see supplementary table 1 & supplementary Figure 4). The primer of the first pair were designed left (del10left1F) and right (del10span1R) from the INDEL position.

Amplification was carried out using LongAmp Taq polymerase (New England Biolabs GmbH, Frankfurt am Main, Germany) on a GeneAmp PCR System 9700 (Thermo Fisher Scientific Inc., Schwerte, Germany) with 50 ng genomic DNA as template under the following cycling conditions: 94°C – 6 min, 30x (94°C – 30s, 58°C – 45s, 65°C – 6 min); 65°C – 15 min. The second pair of primer consisted of the same right primer (del10span1R) but a left primer (del10right2F), which binds within the insertion allele. Amplification was carried out as described above and the following conditions: 94°C –6 min, 30x (94°C – 30s, 58°C – 45s, 65°C – 4 min), 65°C – 15 min. Both PCR reactions of each sample were loaded next to each other onto an 0.8% agarose gel, separated by electrophoresis and visualized with an E-Box VX2 (Vilber Lourmat Deutschland GmbH, Eberhardzell, Germany)(see supplementary Figure 1). Thus, a sample carrying the deletion allele homozygously, result in a 2275 bp amplicon in the first PCR reation and no band in the second PCR. A sample, heterozygote for the INDEL allele, displays two amplicons for the first PCR, one at 2275 bp and another band at 5588bp. For the second PCR this heterozygote sample shows a band at 1732 bp.

Statistical analysis

Nascent & allele-specific qPCR: GraphPad Prism 5 was used to calculate statistical analyses. To assess genotype and treatment effects onto gene expression, a one-way or two-way ANOVA followed by bonferroni correction for multiple comparisons has been used with replicate normalized gene expression fold changes or allele proportions form each sample as input.

Targeted Bisulfite Sequencing: GraphPad Prism 5 was used to calculate statistical analyses. To assess genotype and treatment effects onto DNA methylation, a two-way ANOVA followed by a bonferroni correction for multiple comparisons has been used with replicate methylation values form each sample as input.

Acknowledgements

The authors would like to thank Monika Rex-Haffner for her help with the sequencing the libraries as well as to Mira Jakovcevski for helpful discussions. This study was funded by the ERC starting grant (GxE molmech, grant # 281338) within the FP7

funding scheme of the EU to EBB, NP was funded by a research fellowship from the Canadian Institute of Health Research (CIHR).

Author contributions

TW designed and performed the wetlab experiments. TK designed the genotyping qPCR assays for esv3608688. SS performed optimization and validation experiments of allele-specific gene expression assays. SR performed the processing of the sequencing data. TW, DC performed the data analysis. TW & EBB wrote the manuscript. NP, TR & EBB supervised the project.

Conflict of interest

EBB is co-inventor of the European patent application *FKBP5*: a novel target for antidepressant therapy. *European Patent# EP 1687443 B1* and receives a research grant from Böhringer Ingelheim to investigate *FKBP5* as a candidate target in psychiatric disorders. The authors declare no further conflict of interest.

References

Adcock IM, Ito K, Barnes PJ (2005) Histone deacetylation: an important mechanism in inflammatory lung diseases. *COPD* 2: 445-455

Antony J, Dasgupta T, Rhodes JM, McEwan MV, Print CG, O'Sullivan JM, Horsfield JA (2015) Cohesin modulates transcription of estrogen-responsive genes. *Biochim Biophys Acta* 1849: 257-269

Beagrie RA, Pombo A (2016) Examining Topological Domain Influence on Enhancer Function. *Dev Cell* 39: 523-524

Binder EB (2009) The role of FKBP5, a co-chaperone of the glucocorticoid receptor in the pathogenesis and therapy of affective and anxiety disorders. *Psychoneuroendocrinology* 34 Suppl 1: S186-195

Binder EB, Bradley RG, Liu W, Epstein MP, Deveau TC, Mercer KB, Tang Y, Gillespie CF, Heim CM, Nemeroff CB et al (2008) Association of FKBP5 polymorphisms and childhood abuse with risk of posttraumatic stress disorder symptoms in adults. *JAMA* 299: 1291-1305

Binder EB, Salyakina D, Lichtner P, Wochnik GM, Ising M, Putz B, Papiol S, Seaman S, Lucae S, Kohli MA et al (2004) Polymorphisms in FKBP5 are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. *Nat Genet* 36: 1319-1325

Bradley RG, Binder EB, Epstein MP, Tang Y, Nair HP, Liu W, Gillespie CF, Berg T, Evces M, Newport DJ et al (2008) Influence of child abuse on adult depression: moderation by the corticotropin-releasing hormone receptor gene. *Arch Gen Psychiatry* 65: 190-200

Caspi A, Moffitt TE (2006) Gene-environment interactions in psychiatry: joining forces with neuroscience. *Nat Rev Neurosci* 7: 583-590

Consortium EP (2004) The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* 306: 636-640

Consortium EP (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489: 57-74

Dryden NH, Broome LR, Dudbridge F, Johnson N, Orr N, Schoenfelder S, Nagano T, Andrews S, Wingett S, Kozarewa I et al (2014) Unbiased analysis of potential targets of breast cancer susceptibility loci by Capture Hi-C. *Genome Res* 24: 1854-1868

Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL (2016) Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst* 3: 99-101

Ernst J, Kellis M (2010) Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat Biotechnol* 28: 817-825

Esteller M (2007) Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet* 8: 286-298

Gerken T, Girard CA, Tung YC, Webby CJ, Saudek V, Hewitson KS, Yeo GS, McDonough MA, Cunliffe S, McNeill LA et al (2007) The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science* 318: 1469-1472

Grbesa I, Hakim O (2017) Genomic effects of glucocorticoids. *Protoplasma* 254: 1175-1185

Hagege H, Klous P, Braem C, Splinter E, Dekker J, Cathala G, de Laat W, Forne T (2007) Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat Protoc* 2: 1722-1733

Hakim O, Sung MH, Voss TC, Splinter E, John S, Sabo PJ, Thurman RE, Stamatoyannopoulos JA, de Laat W, Hager GL (2011) Diverse gene reprogramming events occur in the same spatial clusters of distal regulatory elements. *Genome Res* 21: 697-706

Handel AE, Ebers GC, Ramagopalan SV (2010) Epigenetics: molecular mechanisms and implications for disease. *Trends Mol Med* 16: 7-16

Hoffman MM, Buske OJ, Wang J, Weng Z, Bilmes JA, Noble WS (2012) Unsupervised pattern discovery in human chromatin structure through genomic segmentation. *Nat Methods* 9: 473-476

Hoffman MM, Ernst J, Wilder SP, Kundaje A, Harris RS, Libbrecht M, Giardine B, Ellenbogen PM, Bilmes JA, Birney E et al (2013) Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Res* 41: 827-841

Jaaskelainen T, Makkonen H, Palvimo JJ (2011) Steroid up-regulation of FKBP51 and its role in hormone signaling. *Curr Opin Pharmacol* 11: 326-331

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

Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY, Yen CA, Schmitt AD, Espinoza CA, Ren B (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 503: 290-294

Kendler KS, Karkowski LM, Prescott CA (1999) Causal relationship between stressful life events and the onset of major depression. *Am J Psychiatry* 156: 837-841

Kendler KS, Kessler RC, Walters EE, MacLean C, Neale MC, Heath AC, Eaves LJ (1995) Stressful life events, genetic liability, and onset of an episode of major depression in women. *Am J Psychiatry* 152: 833-842

Klengel T, Binder EB (2013) Allele-specific epigenetic modification: a molecular mechanism for gene-environment interactions in stress-related psychiatric disorders? *Epigenomics* 5: 109-112

Klengel T, Mehta D, Anacker C, Rex-Haffner M, Pruessner JC, Pariante CM, Pace TW, Mercer KB, Mayberg HS, Bradley B et al (2013) Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nat Neurosci* 16: 33-41

Klengel T, Pape J, Binder EB, Mehta D (2014) The role of DNA methylation in stress-related psychiatric disorders. *Neuropharmacology* 80: 115-132

Krijger PH, de Laat W (2016) Regulation of disease-associated gene expression in the 3D genome. *Nat Rev Mol Cell Biol* 17: 771-782

Le Dily F, Bau D, Pohl A, Vicent GP, Serra F, Soronellas D, Castellano G, Wright RH, Ballare C, Filion G et al (2014) Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev* 28: 2151-2162

Levenson JM, Sweatt JD (2005) Epigenetic mechanisms in memory formation. *Nat Rev Neurosci* 6: 108-118

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO et al (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326: 289-293

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15: 550

Maddox SA, Schafe GE, Ressler KJ (2013) Exploring epigenetic regulation of fear memory and biomarkers associated with post-traumatic stress disorder. *Front Psychiatry* 4: 62

Matosin N, Halldorsdottir T, Binder EB (2018) Understanding the Molecular Mechanisms Underpinning Gene by Environment Interactions in Psychiatric Disorders: The FKBP5 Model. *Biol Psychiatry* 83: 821-830

McGowan PO, Sasaki A, D'Alessio AC, Dymov S, Labonte B, Szyf M, Turecki G, Meaney MJ (2009) Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nat Neurosci* 12: 342-348

Mikhaylichenko O, Bondarenko V, Harnett D, Schor IE, Males M, Viales RR, Furlong EEM (2018) The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. *Genes Dev* 32: 42-57

Mill J, Petronis A (2007) Molecular studies of major depressive disorder: the epigenetic perspective. *Mol Psychiatry* 12: 799-814

Mill J, Tang T, Kaminsky Z, Khare T, Yazdanpanah S, Bouchard L, Jia P, Assadzadeh A, Flanagan J, Schumacher A et al (2008) Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. *Am J Hum Genet* 82: 696-711

Miller CA, Campbell SL, Sweatt JD (2008) DNA methylation and histone acetylation work in concert to regulate memory formation and synaptic plasticity. *Neurobiol Learn Mem* 89: 599-603

Mitchell AC, Bharadwaj R, Whittle C, Krueger W, Mirnics K, Hurd Y, Rasmussen T, Akbarian S (2014) The genome in three dimensions: a new frontier in human brain research. *Biol Psychiatry* 75: 961-969

Molnar BE, Buka SL, Kessler RC (2001) Child sexual abuse and subsequent psychopathology: results from the National Comorbidity Survey. *Am J Public Health* 91: 753-760

Paakinaho V, Makkonen H, Jaaskelainen T, Palvimo JJ (2010) Glucocorticoid receptor activates poised FKBP51 locus through long-distance interactions. *Mol Endocrinol* 24: 511-525

Pelleymounter LL, Moon I, Johnson JA, Laederach A, Halvorsen M, Eckloff B, Abo R, Rossetti S (2011) A novel application of pattern recognition for accurate SNP and indel discovery from high-throughput data: targeted resequencing of the glucocorticoid receptor co-chaperone FKBP5 in a Caucasian population. *Mol Genet Metab* 104: 457-469

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29: e45

Polanczyk G, Caspi A, Williams B, Price TS, Danese A, Sugden K, Uher R, Poulton R, Moffitt TE (2009) Protective effect of CRHR1 gene variants on the development of adult depression following childhood maltreatment: replication and extension. *Arch Gen Psychiatry* 66: 978-985

Quintin J, Le Peron C, Palierne G, Bizot M, Cunha S, Serandour AA, Avner S, Henry C, Percevault F, Belaud-Rotureau MA et al (2014) Dynamic estrogen receptor interactomes control estrogen-responsive trefoil Factor (TFF) locus cell-specific activities. *Mol Cell Biol* 34: 2418-2436

Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES et al (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159: 1665-1680

Raviram R, Rocha PP, Muller CL, Miraldi ER, Badri S, Fu Y, Swanzey E, Proudhon C, Snetkova V, Bonneau R et al (2016) 4C-ker: A Method to Reproducibly Identify

Genome-Wide Interactions Captured by 4C-Seq Experiments. PLoS Comput Biol 12: e1004780

Roeh S, Wiechmann T, Sauer S, Kodel M, Binder EB, Provencal N (2018) HAM-TBS: high-accuracy methylation measurements via targeted bisulfite sequencing. Epigenetics Chromatin 11: 39

Schanen NC (2006) Epigenetics of autism spectrum disorders. Hum Mol Genet 15 Spec No 2: R138-150

Seitan VC, Faure AJ, Zhan Y, McCord RP, Lajoie BR, Ing-Simmons E, Lenhard B, Giorgetti L, Heard E, Fisher AG et al (2013) Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. Genome Res 23: 2066-2077

Simonis M, Kooren J, de Laat W (2007) An evaluation of 3C-based methods to capture DNA interactions. Nat Methods 4: 895-901

Soldner F, Stelzer Y, Shivalila CS, Abraham BJ, Latourelle JC, Barrasa MI, Goldmann J, Myers RH, Young RA, Jaenisch R (2016) Parkinson-associated risk variant in distal enhancer of alpha-synuclein modulates target gene expression. Nature 533: 95-99

Splinter E, de Wit E, van de Werken HJ, Klous P, de Laat W (2012) Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. Methods 58: 221-230

Symmons O, Pan L, Remeseiro S, Aktas T, Klein F, Huber W, Spitz F (2016) The Shh Topological Domain Facilitates the Action of Remote Enhancers by Reducing the Effects of Genomic Distances. Dev Cell 39: 529-543

Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, Szalaj P, Trzaskoma P, Magalska A, Wlodarczyk J, Rusczycki B et al (2015) CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. *Cell* 163: 1611-1627

Tsankova N, Renthal W, Kumar A, Nestler EJ (2007) Epigenetic regulation in psychiatric disorders. *Nat Rev Neurosci* 8: 355-367

van de Werken HJ, de Vree PJ, Splinter E, Holwerda SJ, Klous P, de Wit E, de Laat W (2012a) 4C technology: protocols and data analysis. *Methods Enzymol* 513: 89-112

van de Werken HJ, Landan G, Holwerda SJ, Hoichman M, Klous P, Chachik R, Splinter E, Valdes-Quezada C, Oz Y, Bouwman BA et al (2012b) Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat Methods* 9: 969-972

Varambally S, Cao Q, Mani RS, Shankar S, Wang X, Ateeq B, Laxman B, Cao X, Jing X, Ramnarayanan K et al (2008) Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. *Science* 322: 1695-1699

Weksberg R, Shuman C, Caluseriu O, Smith AC, Fei YL, Nishikawa J, Stockley TL, Best L, Chitayat D, Olney A et al (2002) Discordant KCNQ1OT1 imprinting in sets of monozygotic twins discordant for Beckwith-Wiedemann syndrome. *Hum Mol Genet* 11: 1317-1325

Wiench M, John S, Baek S, Johnson TA, Sung MH, Escobar T, Simmons CA, Pearce KH, Biddie SC, Sabo PJ et al (2011) DNA methylation status predicts cell type-specific enhancer activity. *EMBO J* 30: 3028-3039

Yehuda R, LeDoux J (2007) Response variation following trauma: a translational neuroscience approach to understanding PTSD. *Neuron* 56: 19-32

Zannas AS, Binder EB (2014) Gene-environment interactions at the FKBP5 locus: sensitive periods, mechanisms and pleiotropism. *Genes Brain Behav* 13: 25-37

Zannas AS, Wiechmann T, Gassen NC, Binder EB (2016) Gene-Stress-Epigenetic Regulation of FKBP5: Clinical and Translational Implications. *Neuropsychopharmacology* 41: 261-274

Figures

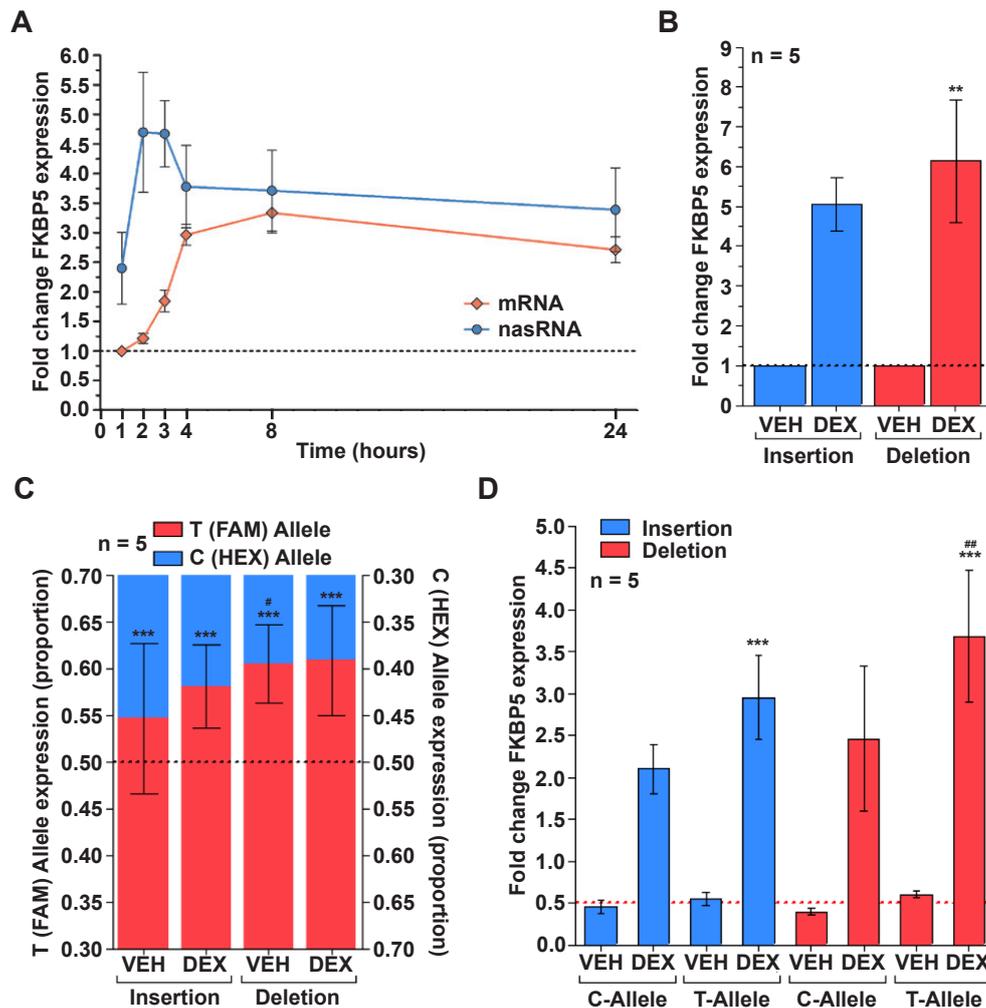


Figure 1: Glucocorticoid-dependent *FKBP5* expression and its modulation by esv3608688 and rs1360780 allele in Lymphoblastoid cell lines.

(A) *FKBP5* expression of nascent (nasRNA) and mature RNA (mRNA) over a time period of 24 h under continuous Dexamethasone treatment (DEX, 100 nM) of two Lymphoblastoid cell lines (LCL; GM19098 & GM18516). **(B)** Comparison of *FKBP5* mRNA expression after DEX treatment (4 h, 100 nM) of 10 LCLs carrying the insertion or deletion. ** indicates a genotype effect onto *FKBP5* expression after DEX treatment (p-value = 0.01). **(C)** Proportions of *FKBP5* mRNA produced from the rs1360780 alleles under the background of the insertion or deletion in 10 LCLs. *** indicate a genotype effect onto *FKBP5* mRNA of the rs1360780 allele (p-value = 0.001); # indicates a genotype effect of the INDEL allele onto the *FKBP5* proportions produced from the rs1360780 allele (p-value = 0.05). **(D)** Allele-specific *FKBP5* mRNA expression for rs1360780 under the background of the insertion or deletion

obtained in 10 LCLs. *** indicate a genotype effect of rs1360780 allele onto *FKBP5* expression levels under the same INDEL background (p-value = 0.001). ## indicates a genotype effect of the INDEL allele onto the *FKBP5* mRNA expression levels generated from the T-allele of rs1360780 (p-value = 0.01).

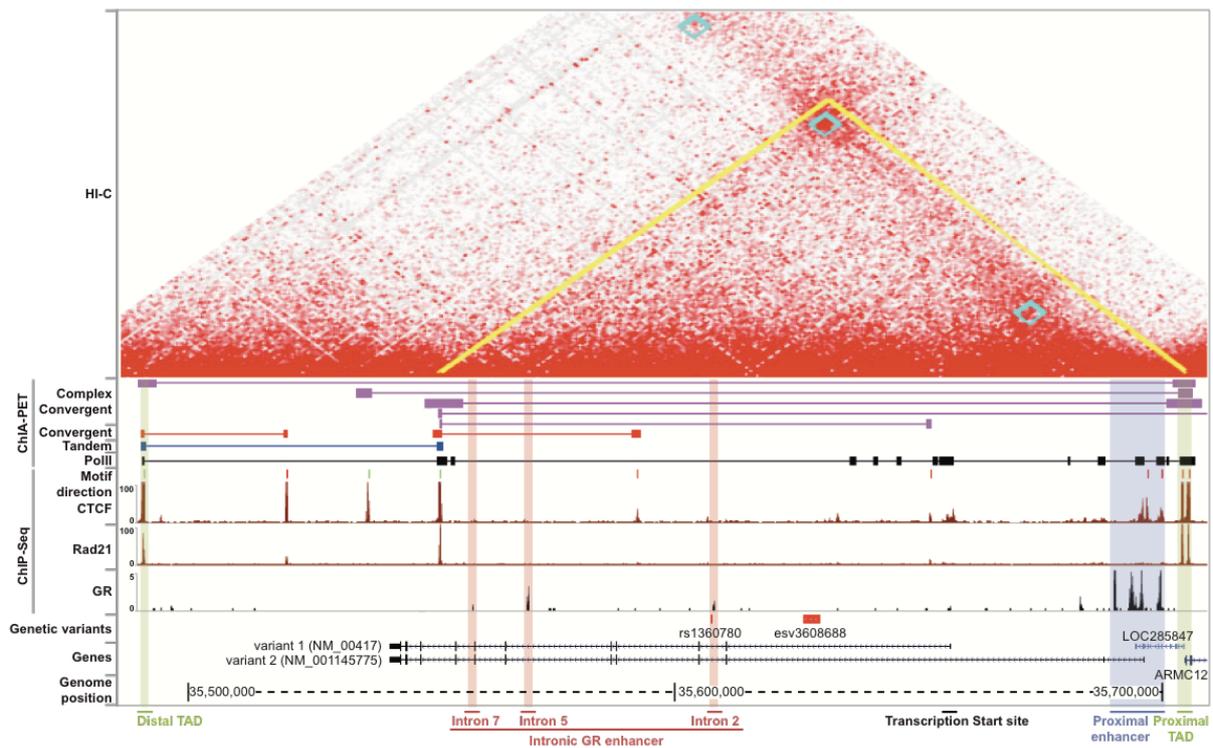


Figure 2: *In silico* analysis of the 3D structure of the *FKBP5* locus.

Genome browser shot and HI-C profile illustrating the *FKBP5* locus 3D structure (hg19 / chr6:35487554-35718452) and the location of the genetic variants rs1360780 & esv3608688 in relevance to key functional sites using reference data. *ChIP-Seq* – The binding of the Glucocorticoid receptor (GR; ENCODE GR HAIB TFBS track from A549 cells), CTCF and Cohesin (indicated by its subunit Rad21)(both ENCODE SYDH TFBS tracks from GM12878) at the *FKBP5* locus. *ChIA-PET* – Chromatin interactions mediated by PolII and CTCF at the *FKBP5* locus (data from Tang et al (2015)). CTCF interactions are classified in convergent (loop formed by CTCF sites with opposing motif directions), complex convergent (loop formed by CTCF sites with opposing motif directions, between the interaction CTCF sites are other CTCF sites forming chromatin interactions) and tandem loops (loop formed by CTCF sites with the same motif directions) according to Tang et al (2015). *HI-C* – The *FKBP5* locus

topologically associated domain structure (TAD) in GM12878 using data from Rao et al (2014) (visualized by Juciebox from the Aiden-Lieberman lab (Durand et al (2016), in situ Mbol primary + replicate).

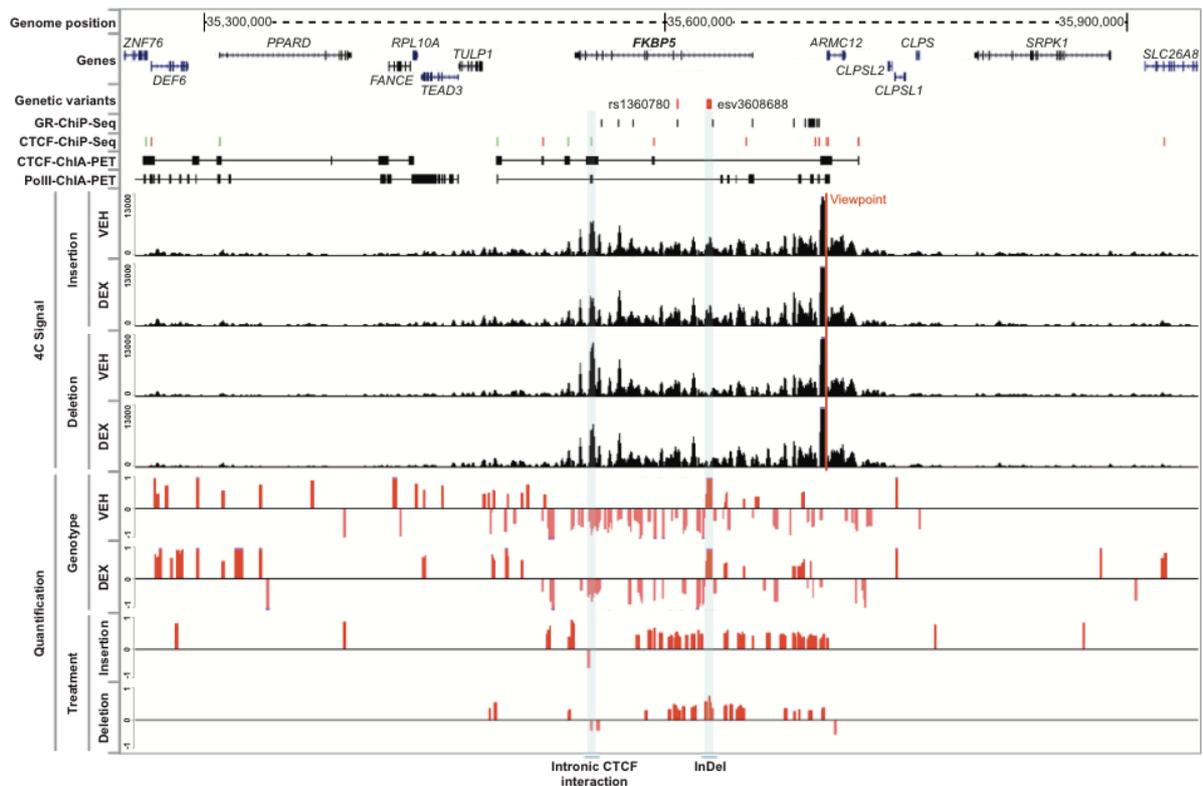


Figure 3: 4C interaction profiles and quantification of LCLs carrying the insertion or deletion allele for esv3608688 using the TAD boundary as viewpoint.

Genome browser shot illustrating the *FKBP5* locus structure (hg19 / chr6:35247364-35946252) and highlighting the location of the INDEL esv3608688 and intronic CTCF interaction involved in forming the *FKBP5* topologically associated domain structure (TAD). The reference data was obtained from the ENCODE project (GR & CTCF peaks from Txn Factor ChiP track, for CTCF peaks green indicates CTCF motif in forward or red in reverse orientation; ChiA-PET from CTCF and PoIII was derived from the GIS ChIA-PET track). The 4C Signal was generated from a viewpoint designed on the TAD boundary, in order to obtain the interaction structure of the *FKBP5* TAD under vehicle and Dexamethasone treatment (4h, 100 nM) of the GM19098 (insertion) and GM18516 (deletion) LCLs. The triplicate 4C data of the two LCLs were processed and quantified using the near-bait cis analysis with standard

parameter of the 4C-ker pipeline (Raviram et al, 2016) and FDR cut-off of 0.05 for the quantification with DESeq2 (Love et al, 2014). The quantification data is displayed in a \log_2 transformation.

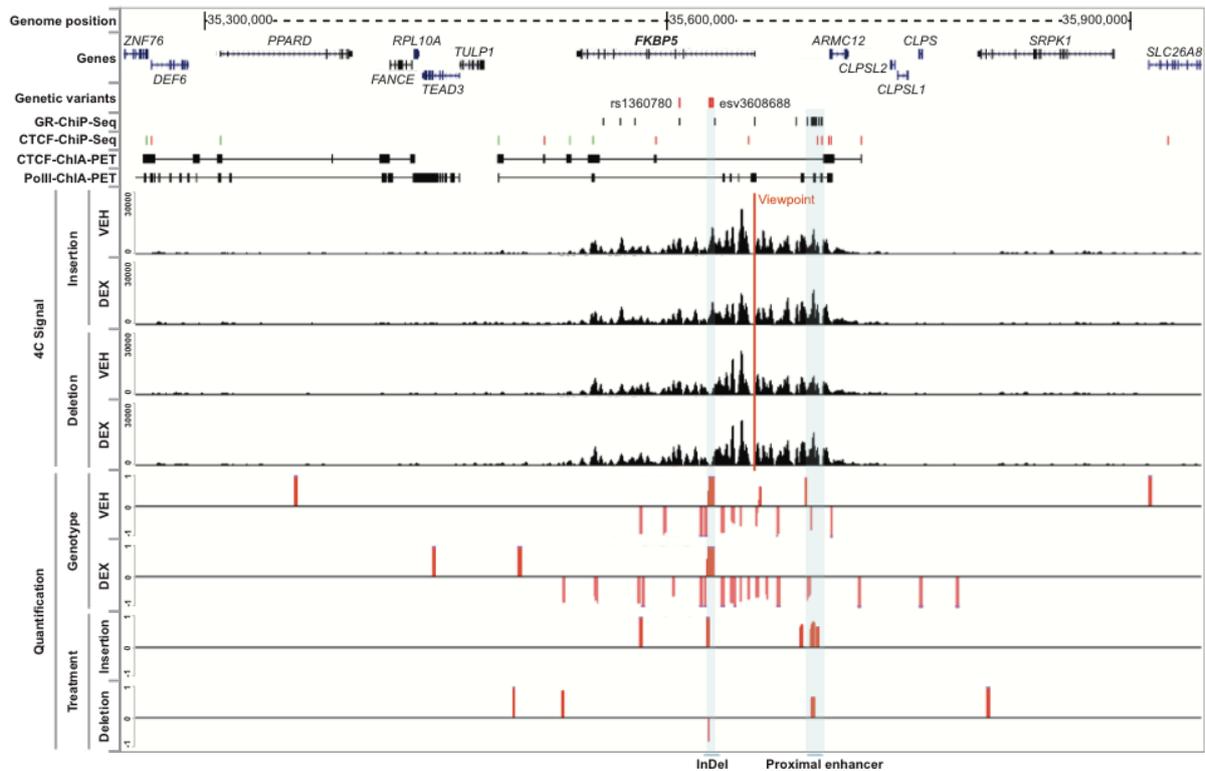


Figure 4: 4C interaction profiles and quantification of LCLs carrying the insertion or deletion allele for esv3608688 using the TSS as viewpoint.

Genome browser shot illustrating the *FKBP5* locus structure (hg19 / chr6:35247364-35946252) and highlighting the location of the INDEL esv3608688 and the proximal long-range enhancer containing multiple GR binding sites. The reference data was obtained from the ENCODE project (GR & CTCF peaks from Txn Factor ChiP track, for CTCF peaks green indicates CTCF motif in forward or red in reverse orientation; ChiA-PET from CTCF and PolIII was derived from the GIS ChIA-PET track). The 4C Signal was generated from a viewpoint designed on the transcription start site (TSS), in order to obtain the interaction structure to the TSS under vehicle and Dexamethasone treatment (4h, 100 nM) of the GM19098 (insertion) and GM18516 (deletion) LCLs. The triplicate 4C data of the two LCLs were processed and quantified using the near-bait cis analysis with standard parameter of the 4C-ker pipeline (Raviram et al, 2016) and FDR cut-off of 0.05 for the quantification with

DESeq2 (Love et al, 2014). The quantification data is displayed in a \log_2 transformation.

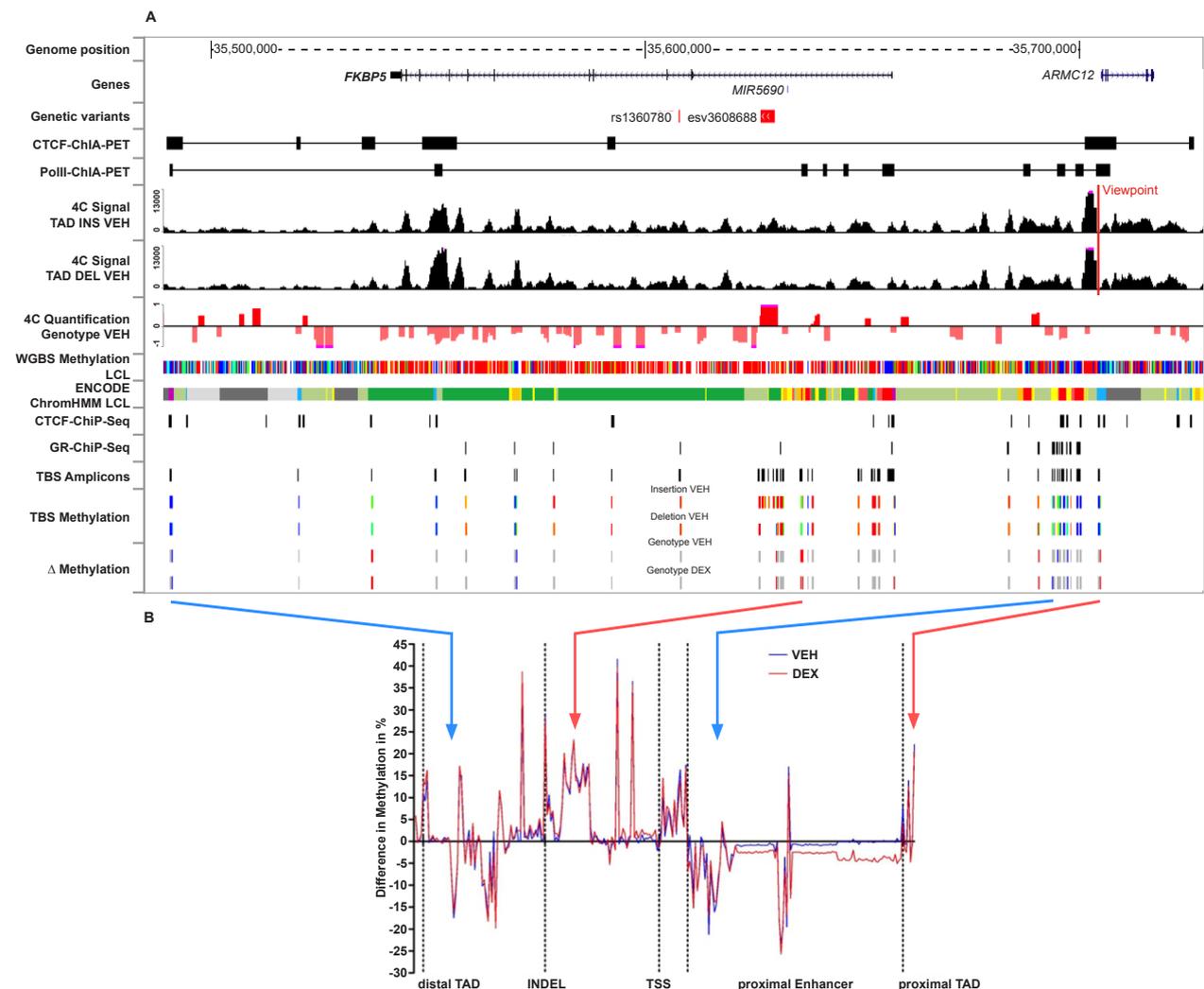
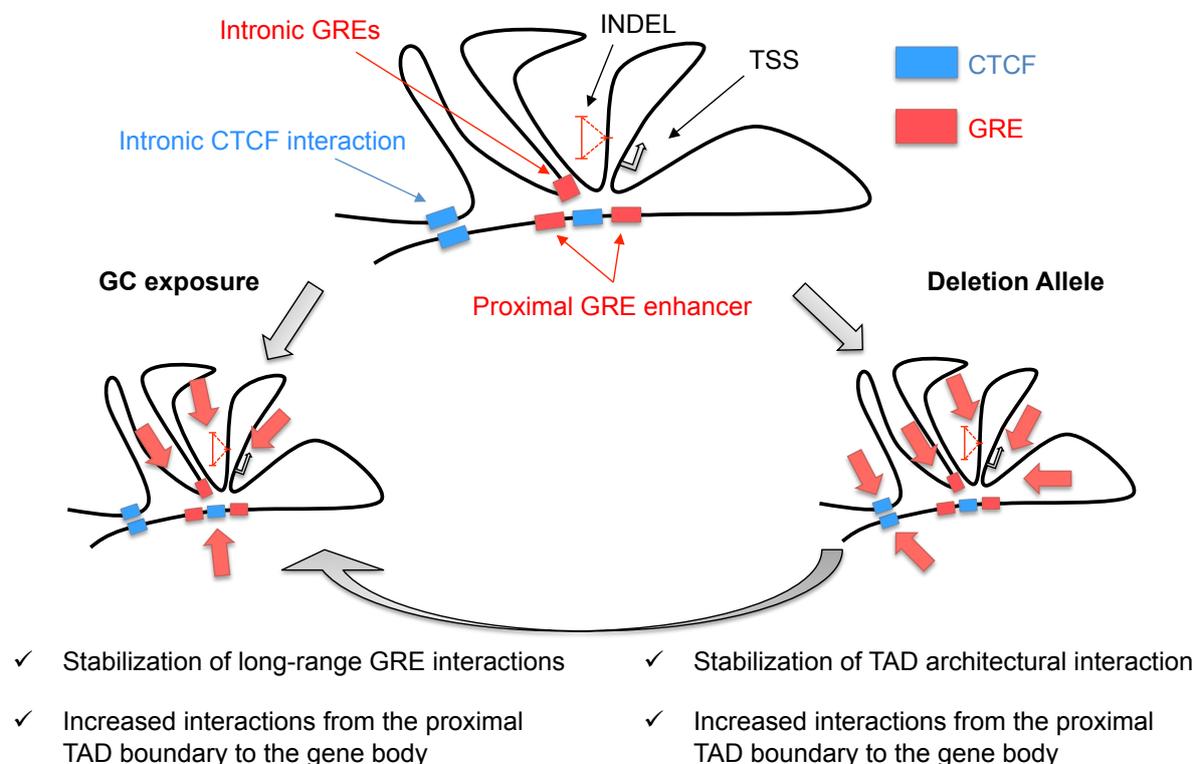


Figure 5: INDEL esv3608688 effect on *FKBP5* methylation in LCLs

(A) Genome browser shot illustrating the *FKBP5* locus structure (hg19 / chr6:35479489-35729448) and the location of the genetic variants rs1360780 & esv3608688. The reference data was obtained from the ENCODE project (GR & CTCF peaks from Txn Factor ChiP track; ChiA-PET from CTCF and PolII was derived from the GIS ChIA-PET track, GM12878 WGBS data was obtained from GSM2308632, low methylation levels are in by cold colors and high methylation levels by warm colors; ENCODE ChromHMM displays the Genome segmentation track of GM12878, summarizing 15 chromatin states which e.g. indicate active promoters (red), enhancer (yellow/orange), transcription (green) but also repression by polycomb (grey). *4C Signal* - The 4C Signal was generated from a viewpoint

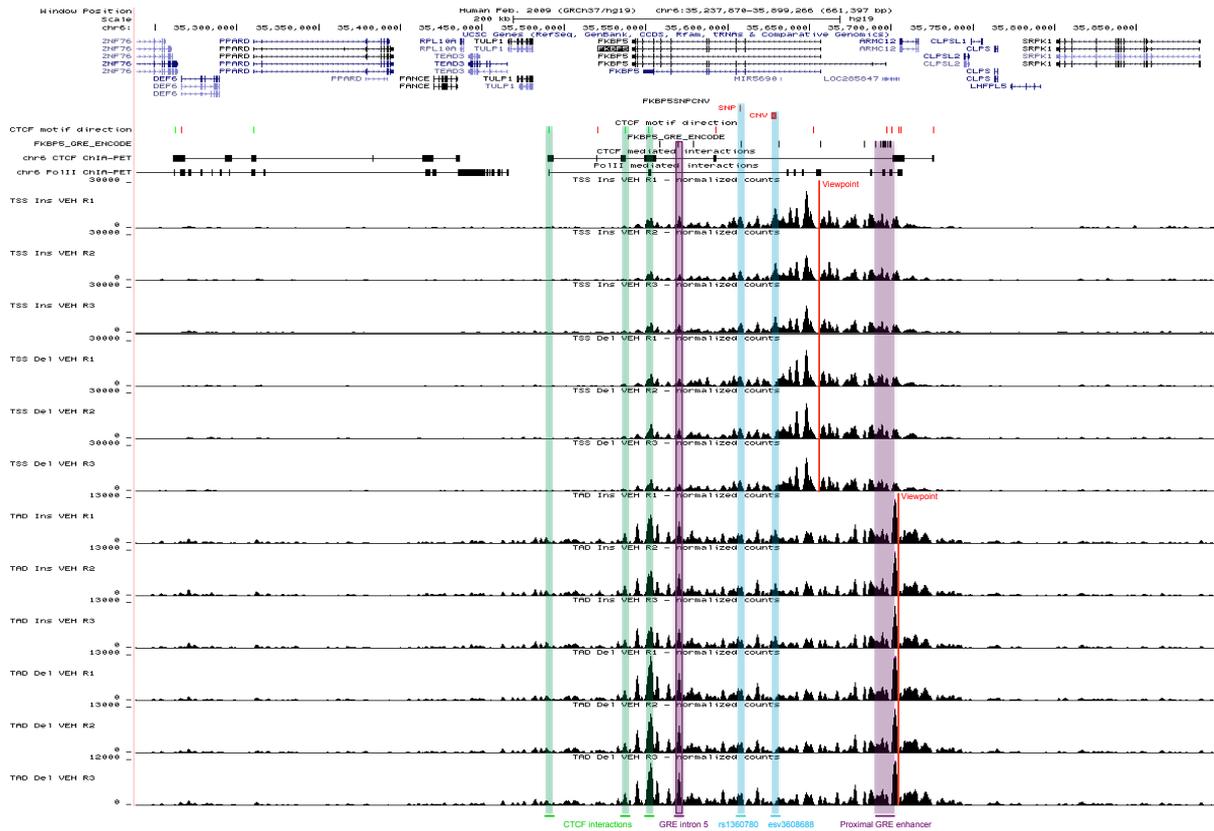
designed on the TAD boundary, in order to obtain the interaction structure of the *FKBP5* TAD under vehicle conditions of the GM19098 (insertion) and GM18516 (deletion) LCLs. *4C quantification* - The 4C quantification data is displayed in a \log_2 transformation and a FDR cut-off of 0.05 was used. *TBS amplicons* - indicate the localization of the amplicons for the assessment of DNA methylation levels at the *FKBP5* locus. *TBS Methylation* - Displays the assessed DNA methylation levels of CpGs at selected sites for *FKBP5* of 10 LCLs which carry the insertion or deletion (mean methylation level of each group, n = 5). Low methylation levels are in by cold colors and high methylation levels by warm colors Δ *Methylation* - Difference of DNA methylation levels between LCLs carrying the insertion / deletion. Significant methylation levels ($p > 0.05$, two-way ANOVA followed by bonferroni correction for multiple testing) are indicated in blue (significant reduced methylation levels in insertion LCLs) or red (significant increased methylation levels in insertion LCLs). **(B)** Detailed Plot of methylation differences due to the INDEL allele of aggregates CpGs in all amplicons distributed over the *FKBP5* locus.



The INDEL allele in *FKBP5* affect GRE enhancer activity through the modulation of TAD interaction frequencies and regulate the responsiveness to glucocorticoid exposure

Figure 6: A model how glucocorticoids and INDEL esv3608688 influence *FKBP5* expression by modulating its underlying chromatin structure

The structural basis of *FKBP5* gene expression regulation is shaped by CTCF generating architectural loopings, which are described as TAD structure. These TAD loopings aid the formation of promoter-enhancer interactions, by restricting the promoter-enhancer looping within the TAD boundaries. For *FKBP5* pre-existing proximal & intronic GR enhancer loopings to the promoter are important to mediate the gene expression response after GR activation by GCs. The GC-induced expression changes are associated with increased interaction frequencies between long-range promoter-enhancer loops but also from the proximal TAD boundary to the body of the *FKBP5* gene. Genetic variants in *FKBP5* locus can modulate the response to GR activation by affecting either GR-associated promoter-enhancer loops or architectural CTCF mediated interactions. The 3.3 kb INDEL is associated to moderate the stabilization of the architectural loop between the 5'TAD and intronic TAD boundary. In this scenario, the deletion is associated to lead to increased interaction frequencies between TAD boundaries and production of *FKBP5* transcripts. The common feature in this model is that genetic variants and GR activation are associated to stabilize loops within the *FKBP5* locus which is reflected by increased interaction frequencies and additive effects onto gene expression. In other words, the activity of the GR enhancers responsible for *FKBP5* gene expression regulation is on the one hand regulated by the activation of the hormone receptor itself but also their probability to find the *FKBP5* promoter. This probability is regulated by genetic variants modulating the stability of the architectural interactions (INDEL allele) but also promoter-enhancer interactions (rs1360780 allele).



Supplementary Figure 1: Triplicate 4C interaction profiles from the TSS and TAD viewpoint under vehicle conditions.

Genome browser shot illustrating the *FKBP5* locus (hg19 / chr6:35237870-35899266) and the location of the genetic variants rs1360780 & esv3608688 in relevance to key functional sites using reference data. The reference data was obtained from the ENCODE project (GR & CTCF peaks from Txn Factor ChiP track, for CTCF peaks green indicates CTCF motif in forward or red in reverse orientation; ChIA-PET from CTCF and PolIII was derived from the GIS ChIA-PET track). The 4C Signal was generated from two viewpoints designed on the transcription start site (TSS) and topologically associated domain (TAD) boundary. The triplicate 4C data of the two LCLs were processed using the 4C-ker pipeline (Raviram et al, 2016). Examples of regions known to interact with the viewpoints or are important in this study are highlighted in different colors.

Tables and their legends

Supplemental Table 1: Primer used for qPCR

Allele-specific *FKBP5* gene expression

rs3800373_F:	GCAACTGCGTGTCAAACC
rs3800373_R:	GAAC TTTATTTACTACTCCTCTATCATGC
rs3800373_FAM:	/56-FAM/AG AAG AGC A/ZEN/A CTA TTT ATT TGT CAA CCC TAC AGA T/3IABkFQ/
rs3800373_HEX:	/5HEX/AG AAG AGC A/ZEN/A CTA TTT ATT TGT CAA CAC TAC AGA T/3IABkFQ/

Nascent *FKBP5* RNA

FK5hnR_F:	GTA AAC AGA GGC AGG GAG ATG
FK5hnR_R:	AGG CTT CTC AAC CTT GTA GC
FK5hnR_Probe:	/56-FAM/ACA GCC TGC /ZEN/AAA GTC TCC ACT GTT /3IABkFQ/

4C RE efficiency

DpnII_P1_F:	ACGTCCCAGAACAGAAGAGC
DpnII_P1_R:	CAGCGCCGTCTAACAGGT
DpnII_P2_F:	ATTCTATAGCTGCAAGTCCC
DpnII_P2_R:	GAAAGCCTTCTGTGCCTG
NlaIII_P1_F:	GTAAGCCTTATTCCACTGGA
NlaIII_P1_R:	G TTCAGGACAGTGTGTTCA

Supplemental Table 2: 4C-Seq primers used in this study

Viewpoint	First RE	Second RE	Reading Primer	Non-reading primer
<i>FKBP5</i> TSS	DpnII	NlaIII	GCGCCGTCTAACAGGT GAT	CCCATCTCCGTGGCCA TG
<i>FKBP5</i> 5'TAD	DpnII	NlaIII	GGCCCTCCCATATGTG ATC	AGTACAGTGGCGTGTC AG

Supplemental Table 3: TBS PCR information

Please see separate File.

Supplemental Table 4: Primer used for qPCR

INDEL allele genotyping PCR

Del10left1F:	CCCAGGATAATCTCCCCATT
Del10span1R:	GGTCCTCTTTGTGCTCTTGC
Del10right2F:	CACTGCTATGAACATTTTTGTGC

6 Discussion

The development of stress-related psychiatric disorders involves the complex integration of genetic and environmental factors (Zannas and Binder, 2014). There are a large number of studies (31 independent studies comprising 39 cohorts with more than 31000 individuals) linking common genetic variants within a haplotype spanning the *FKBP5* locus and environmental impacts with diverse psychiatric phenotypes such as MDD or PTSD (Matosin et al., 2018; Zannas et al., 2016). Interestingly, only a few studies detect a main genotype effect of *FKBP5* variants with increased risk for psychiatric disorder, but most of them show evidence that the risk is shaped by an interaction of *FKBP5* variants and early life adversity. The molecular mechanisms of such statistical gene by environment interaction, are far from being fully understood, and improved understanding of these molecular mechanism and their contribution to pathogenesis would have strong implications for decreasing the burden on individuals and societies suffering from psychiatric diseases. A first indication for a molecular mechanism integrating *FKBP5* variants and early childhood adversity was described by Klengel et al. (2013), highlighting that the SNP rs136780 is the functional variant in the disease associated haplotype, which can modulate enhancer-promoter chromatin contacts and is accompanied by allele-specific local demethylation of CpGs in a specific GRE enhancer. This thesis aims to further elaborate our understanding of this molecular mechanism integrating genetic and environmental factors by enabling new possibilities to investigate how environmental impacts are embedded into *FKBP5* DNA methylation patterns and how other more complex variants in the disease associated haplotype of *FKBP5* can impact the response to stressors. First, a new method - High accurate methylation measurements via targeted bisulfite sequencing (HAM-TBS) - was established to accurately and robustly measure DNAm levels in key regulatory sites of *FKBP5* in larger cohorts of patients (Paper I and further discussed in Chapter 6.1). Second, the new possibilities of this method have been used to better understand the underlying dynamics of GC-induced DNAm changes in *FKBP5* in healthy individuals upon a DEX challenge (Paper III) and preterm infants (Paper II, implications from Paper II & Paper III are further discussed in Chapter 6.2). Third, lymphoblastoid cell lines carrying different alleles for a 3.3 kb large INDEL in intron 1 within the *FKBP5* gene have been used to identify if and how this structural variant in the disease-associated

haplotype contributes to shape GC-induced *FKBP5* gene expression changes (Manuscript I and further discussed in Chapter 6.3). Finally, thoughts on future perspectives and potential translations into the clinical routine of the mechanistic insights of *FKBP5* gene expression regulation are discussed in Chapter 6.4.

6.1 Accurate and robust measurement of DNAm levels within the *FKBP5* locus in patient cohorts

A requirement to investigate epigenetic mechanisms and their contribution to common disease is the ability to accurately and robustly detect epigenetic marks such as DNAm. This is especially important for complex diseases, which involve multiple genetic and environmental factors. Therefore effect sizes of the single factors contributing to the development of complex diseases are rather small and pathological situations favorably emerge due to the interplay of these factors. In addition, mixed tissues like peripheral blood or saliva are often the only material available in larger patient cohorts. In mixed tissues, cell type specific methylation effects can appear and be masked due to the high abundance of DNA from non-responsive cell types and therefore lead to small changes in methylation levels. In the light of this realization and the need of DNAm measurement in large human cohorts for research in epidemiology and medicine (Joubert et al., 2016), we developed the HAM-TBS method that assesses DNAm levels at a high accuracy and cost-efficient manner. Moreover, due to the high interest in the field of psychiatry for *FKBP5*, we established and validated a PCR panel for monitoring methylation levels of important regulatory regions within the *FKBP5* locus.

We chose a TBS (bisulfite conversion of DNA and subsequent target enrichment via PCR) workflow coupled with high-throughput sequencing since this enables the detection of methylation levels at base-pair resolution. Furthermore, TBS workflows have shown to be one of the best techniques to obtain comparable results between laboratories (consortium, 2016). The HAM-TBS is placed between whole genome bisulfite sequencing, which is cost intensive and due to the coverage needed (>60x) to detect differentially methylated sites (Ziller et al., 2015), and Pyrosequencing (Tost and Gut, 2007), which enables the assessment of DNA methylation levels accurately but lacks the throughput capabilities to assess multiple regions in larger cohorts. Illumina DNA methylation arrays are an example to obtain methylation levels at base

pair resolution in a cost-efficient manner but lack the coverage in key enhancer regions which are important for *FKBP5* gene expression regulation and epigenetic mechanisms contributing to the development of psychiatric diseases (Klengel and Binder, 2013; Matosin et al., 2018; Zannas et al., 2016).

In order to clarify which steps within the TBS workflow introduce the most variability onto the DNA methylation measurement we performed a validation experiment using *in vitro* methylated control DNA and different workflow conditions (replicates at the bisulfite treatment, PCR or library preparation step of the workflow). The results of this experiment indicate that each of the conditions were suitable to detect methylation levels at a high accuracy. Due to the experience that sample-to-sample variation of non-*in vitro* generated control samples (e.g. patient samples) can not be excluded, we recommend to use triplicate bisulfite treatment of each patient DNA which can be pooled and further processed as a pooled sample. Next, we compared the technical accuracy of the HAM-TBS workflow to Pyrosequencing, the gold standard method for assessing DNA methylation via targeted bisulfite sequencing. We found that the accuracy of the HAM-TBS workflow was higher than Pyrosequencing (Pyrosequencing mean SD = 4.68%, max. SD = 14.56% and HAM-TBS mean SD = 0.72%, max. SD = 1.83%). To our knowledge, this accuracy for the measurement of DNA methylation is the highest reported so far in the literature. This level of accuracy for DNA methylation measurement could not only be useful for investigating small differences in patient samples but also be beneficial other fields like forensics (Vidaki and Kayser, 2018). The differences in accuracy can potentially be explained by the underlying detection principles of both techniques per se. TBS using Pyrosequencing deducts methylation levels by comparing the intensities of fluorescence signals under consideration of other reference intensities within the sequencing run. Over the length of the sequencing, the enzyme and nucleotides are consumed which leads to a decrease in signal intensities over the length of the run. We observe that this affects the accuracy of the methylation measurement and therefore we only use short sequencing (<30 dispersions) runs when using Pyrosequencing. It has to be noted that there are also further considerations to be aware of when designing Pyrosequencing assays to accurately obtain methylation levels such as long stretches of T-nucleotides that can affect the performance of the sequencing run after the long T-stretch. On the other hand, using the sequencing by synthesis technology from Illumina, methylation levels are calculated by comparing

the read counts of T's and C's at a distinct position. While the detection of the base itself is also based on a fluorescent signal the actual methylation level is not. In order to obtain suitable sequencing quality of the read, we filter out reads with a Phred score lower than 30. Moreover, to ensure an accurate detection of the methylation levels we established a three step quality control. First, we ensure a sufficient bisulfite conversion rate of 95 % per sample and amplicon. Second, PCR artifacts are removed and potential SNPs are monitored. Third, we and others found that a minimum coverage of 1000 reads per CpG is recommended for accurate methylation quantification (Masser et al., 2013). In comparison to other TBS workflows (Bernstein et al., 2015; Chen et al., 2017; Masser et al., 2013; Masser et al., 2015), we believe that reducing the amount of PCR steps within the workflow (we only use one PCR step for target amplification but eliminate the PCR step during the library preparation by ligating the adapters to the amplicons) helps to increase the accuracy of the methylation measurement due to the avoidance of potential amplification biases.

Another benefit of using the sequencing by synthesis technology is the upscaling potential. Our HAM-TBS workflow currently allows to analyze 96 samples of targeted regions (about 10 kb) in one MiSeq run. This is only possible due to the implementation of the Agilent Tape station for amplicon quantification and the use of a liquid handling robot for amplicon pooling into the workflow. The use of the Tape station enables to quantify amplicon band concentrations without the detection of other DNA contaminants such as Primer dimer or unspecific bands in the PCR reaction. Therefore samples can be pooled using a pipetting robot prior to clean up of the PCR products with paramagnetic beads. This saves the amount of cost intensive beads. Furthermore, using a pipetting robot does not only free hands-on time but also avoids the problem of potential mis-pipetting and therefore increases the robustness of the workflow.

Current limitations of the HAM-TBS are the need of PCR for target amplification but also the need to sequence longer reads to cover all CpGs within PCR amplicons. Performing single PCR reactions for each loci of interest in a large amount of samples is time and cost intensive. Furthermore, we observe PCR design limitations and amplification biases especially in CpG rich regions of the genome. Due to the harsh bisulfite treatment amplifying in target loci over 400 bp is challenging and introduces undesirable sample-to-sample variability. In addition, up to now Illumina sequencer allow a maximal read length of 300 bp (MiSeq which is limited to 25

million reads), which ideally would enable to cover 600 bp amplicons. Possible solutions to overcome some of these limitations are the optimization of multiplex PCR reactions or the use of other enrichment strategies such as capture-based methods (Brinkman et al., 2010; Sun et al., 2015).

Due to the importance of *FKBP5* in the field of psychiatry we designed and validated a PCR panel to accurately measure DNA methylation levels of key regulatory sites within the *FKBP5* locus. We designed 29 amplicons that cover the TSS, TAD boundaries, intronic and proximal enhancers as well as GR and CTCF binding sites. Not only does the panel include sites that are potentially important gene expression regulation and have shown to be implicated in disease mechanisms (Klengel and Binder, 2013; Yehuda et al., 2013) but also a control PCR (h19, imprinted region should show a DNA methylation level around 50%, (Boissonnas et al., 2010)) which serve as an control for the DNA measurement of each sample. Each PCR was validated by an initial gradient PCR run to find robust cycling conditions. Moreover, using *in vitro* methylated control DNA we checked the PCRs in the panel for amplification biases.

In summary, the HAM-TBS workflow is a versatile tool to accurately and cost efficiently measure DNA methylation levels of target loci which has the potential not only to enhance research of epigenetic mechanisms but also may influence other fields like forensics (Vidaki and Kayser, 2018). In this regard, the ability to detect small effects onto DNA methylation levels can lead to a more in deep analysis of disease mechanism including DNA methylation marks. On the other hand, the increased throughput of this method enables to measure methylation levels in larger patient cohorts and longitudinal studies. The community interested in the *FKBP5* will benefit from the *FKBP5* PCR panel since it enables the investigation of key regulatory sites which are not covered by the widespread commercial DNA methylation arrays used to cost efficiently assess methylation levels in cohorts.

6.2 Embedding of environmental signals into *FKBP5* DNAm levels

The concept that environmental signals especially during sensitive phases can have an impact onto the phenotype and are mediated by epigenetic marks such as DNAm is well established (Szyf, 2013) and prominent examples are the agouti mouse model

(Jirtle, 2014; Waterland and Jirtle, 2003) or twin studies in humans (Fraga et al., 2005). Although it has been thought that the DNAm patterns in cells are set after dynamic changes during embryogenesis, it now becomes clear that postnatal methylation changes can occur due to environmental signals (Dor and Cedar, 2018). Postnatal changes of DNAm has for example been shown in liver as a result of nutritional changes after birth (Cannon et al., 2016; Ehara et al., 2015), male liver through testosterone secretion (Reizel et al., 2015) and breast tissue after the first pregnancy (Dos Santos et al., 2015). The common conclusion in these examples is that DNAm changes are installed by hormone exposure and maintained to serve as a kind of memory of hormone exposure even after hormone levels are normalized again (Dor and Cedar, 2018). A similar link is potentially observed when individuals are exposed to ELS. In regard of this concept, the release of GCs into the blood circulation after stress exposure could integrate the response including DNAm changes in many tissues, including the brain (Szyf, 2013). In fact, the embedding of ELS into the epigenetic layer in central and peripheral tissues has been described in rodents and humans for several genes of the HPA axis (*NR3C1*, *Crh*, *Avp*, *Crfr2*, *FKBP5*) but also for candidate genes outside the stress axis (*SLC6A4*, *rRNA promoter*, *BDNF*, *Reelin*, *Gad1*) (Klengel and Binder, 2013; Roth et al., 2009) and reviewed in Jawahar et al. (2015). Epidemiological studies find associations that early adversity can impact the susceptibility to develop a psychiatric disorder later in life and that the sensitivity to ELS can be mediated by *FKBP5* genotypes (Matosin et al., 2018; Zannas et al., 2016). Moreover, the combination of ELS and *FKBP5* has been associated with the occurrence of long-term stable alterations in methylation levels in distinct enhancer regions of *FKBP5* (Klengel et al., 2013). The open question is whether the methylation changes in *FKBP5* are already established directly after stress exposure in childhood, similarly as discussed above for other hormones, or if these are installed later in life when pathological phenotypes manifest. The observation that *FKBP5* methylation differences between preterm infants and term babies at birth were resolved later in life (Paper II) are an indication for dynamic methylation changes after ELS. Dynamic changes have also been observed by an epigenome-wide association study (EWAS) (Cruickshank et al., 2013) that highlighted many differences in DNAm (1555 sites) that were resolved at the age of 18 years, but did also show that the methylation level difference of CpGs at some loci persisted (*PCSK9*, *TRIM71*, *SLC44A4*, *GPC6*, *NFYA* and *EGR1*). In this

EWAS, CpGs at *FKBP5* did not reveal any difference between preterm and term babies. This can on the one hand be explained by missing probes for important regulatory sites within *FKBP5* such as intron 7 CpGs on the Illumina methylation array used in the EWAS study but, on the other hand, the use of DNA from different tissue (Paper II used buccal DNA and the EWAS study by Cruickshank et al. (2013) blood spots). Therefore the question arises how cells memorize the early exposure to high levels of stress in light of the dynamic methylation changes. In fact, we know very little about the variation of methylation marks in specific loci within the same individual (Zhang et al., 2013), especially under the influence of an environmental impact. One factor making it difficult to understand how early adversity can impact the susceptibility to develop a psychiatric disorder later in life is that this impact is influenced by a number of factors such as the nature of stressor, exposure in sensitive phases, severity, cumulative exposure effects but also biological factors as gender, age and genotypes (Jawahar et al., 2015). In this regard, using pharmacological challenges to mimic native GC exposure during stressful situations can be a valuable model to better understand the physiology of the embedding of environmental input into epigenetic layers. Therefore, to investigate the underlying dynamics of GC-induced methylation changes in *FKBP5* we exposed healthy individuals to a DEX challenge (Paper III).

The development of the HAM-TBS workflow enabled us to investigate GC-induced methylation changes at the *FKBP5* locus at a higher accuracy and coverage of key regulatory sites (TSS, TAD boundaries, intronic and proximal Enhancer, and GR and CTCF binding sites). Due to the increased throughput we were able to test the methylation patterns (302 CpGs covered by 28 amplicons) in peripheral blood of 19 healthy individuals after an acute DEX treatment (single oral dose of 1.5 mg DEX) for 5 timepoints (baseline, 1 h, 3 h, 6 h, 23 h) and validated the DEX-induced effects in a larger independent sample (n = 89, baseline and 3 h after DEX). We detected GC-induced methylation changes at 44 CpGs within the *FKBP5* locus including CpGs in intron 7 and intron 5 which were shown to be associated with ELS (Harms et al., 2017; Klengel et al., 2013) but also in healthy individuals after a 30-day cortisol load (CpGs in intron 2 & intron 7; Lee et al. (2018)). The effects sizes ranged from -17% to +10% methylation difference, however, 74% of the CpGs analyzed showed decreased methylation levels after the exposure to DEX, which tend to co-localize with GR binding sites at the proximal and intronic GR enhancers. Similarly, studies in

rodents using continuous GC excess and DEX treatment of HPCs also observed reduced methylation levels in *FKBP5* co-localizing with GR binding sites (Klengel et al., 2013; Lee et al., 2010; Seifuddin et al., 2017). Interestingly, we monitored GC-induced DMCs after 1 h of the DEX exposure, which peaked between 3 h to 6 h and returned to baseline within 23 h and are inversely correlated with *FKBP5* expression. Only 8 of the 44 DMCs, which revealed these dynamics remained significant after 23 h and showed a much smaller effect size (mean absolute Δ methylation = |1.8%|). Moreover, we were able to show that allele of rs1360780 could moderate these methylation level dynamics supporting the idea that *FKBP5* genotypes associate with differential sensitivity to GR activation as observed for differential effects of rs1360780 allele onto disease symptom severity (Klengel et al., 2013; Lee et al., 2010; Seifuddin et al., 2017), response to antidepressant treatment and recurrence of depressive episodes (Binder et al., 2004) but also neuroendocrine endophenotypes and other clinically relevant outcomes (reviewed in Matosin et al. (2018); Zannas et al. (2016)).

Several mechanisms could contribute to the dynamics in DNAm observed in blood of (Paper III). On the one hand by controlling for cell type composition before and after DEX treatment we could show that at least some of the methylation changes were likely to appear within specific cells. Sorting cells and re-assessing our results in the sorted cells would give more information, which cells types contribute to the effects monitored. In fact, a study in mice could show by cell sorting and genome-wide bisulfite sequencing that the observed GC-induced methylation changes were observed primarily in blood T-cells (Seifuddin et al., 2017). On the other hand, mapping enhancers within the *FKBP5* locus across many tissues, using chromatin states analyzed by the ENCODE project, suggest that most of the GR-responsive enhancers exert a shared function and thus may show a similar epigenetic response to GR activation across different cell types and tissues. Due to the observation that many of the methylation changes appeared within or very close to GR binding motifs, a reduction in methylation following GC exposure could be mediated by a demethylation pathway involving transcription factor binding of the GR to GREs which has been reported by Thomassin et al. (2001) and Wiench et al. (2011). The mechanisms of this targeted demethylation remain unknown, but the involvement of DNA repair has been proposed (Kress et al., 2006). Interestingly, similar to the transient GC-induced methylation changes at the *FKBP5* locus in blood rapid cyclical

methylation patterns have been reported for the pS2 gene in response to estrogen stimulation in breast cancer cells (Metivier et al., 2008). This study implicated that for the de- and remethylation processes at pS2 DNA methyltransferases, glycosylases and base excision repair proteins are involved. Given the similar reported kinetics of DNAm changes due to hormone exposure, the same enzymatic processes could be involved in GR-induced demethylation. We and others detected hormonal-induced methylation changes at specific CpGs which can appear transiently and be partially remethylated (Wiench et al., 2011) but also long lasting (Cannon et al., 2016; Dos Santos et al., 2015; Ehara et al., 2015; Klengel et al., 2013; Reizel et al., 2015). The question of how the dynamics of de- and remethylation are regulated remains open, but the importance of timing (e. g. development) and the local chromatin structure of the CpGs (e.g tissue specific binding of TF next to hormone receptor) has been discussed by Luo et al. (2018); Wiench et al. (2011). Therefore, hormone-induced methylation changes associated to different TF e.g. GR or STAT5 (Dos Santos et al., 2015) could involve different processes of de- and remethylation, but also similar pathways, which lead to other dynamics. The processes of de- and remethylation could potentially be interfered by DNA-methyl binding proteins such as MeCP2 and polycomb complexes (Murgatroyd and Spengler, 2014) changing the methylation dynamics due to hormones from transient to long-term stable.

In summary, the data presented in this thesis favor the hypothesis that environmental impacts can be embedded into epigenetic layers through the action of hormones, which have the potential to mediate the response to the environment across central and peripheral tissues. The processes involved in this embedding of environmental signals are not yet fully understood but several lines of evidence exist that these are active enzymatically driven that can lead to dynamic epigenetic changes and have to be seen in the light of temporal & spatial factors that can lead to transient but also long term stable epigenetic adjustments. However, to understand how an environmental signal - such as childhood stress - together with other important factors - like genotypes - can prime the underlying epigenetic processes to lead to disease phenotypes later in life, further research on these epigenetic processes has to be done. Essential in this regard will be the ability to model environmental signals on a molecular level, but also an animal or cell model to observe the embedding into epigenetic layers and manipulate the implicated processes.

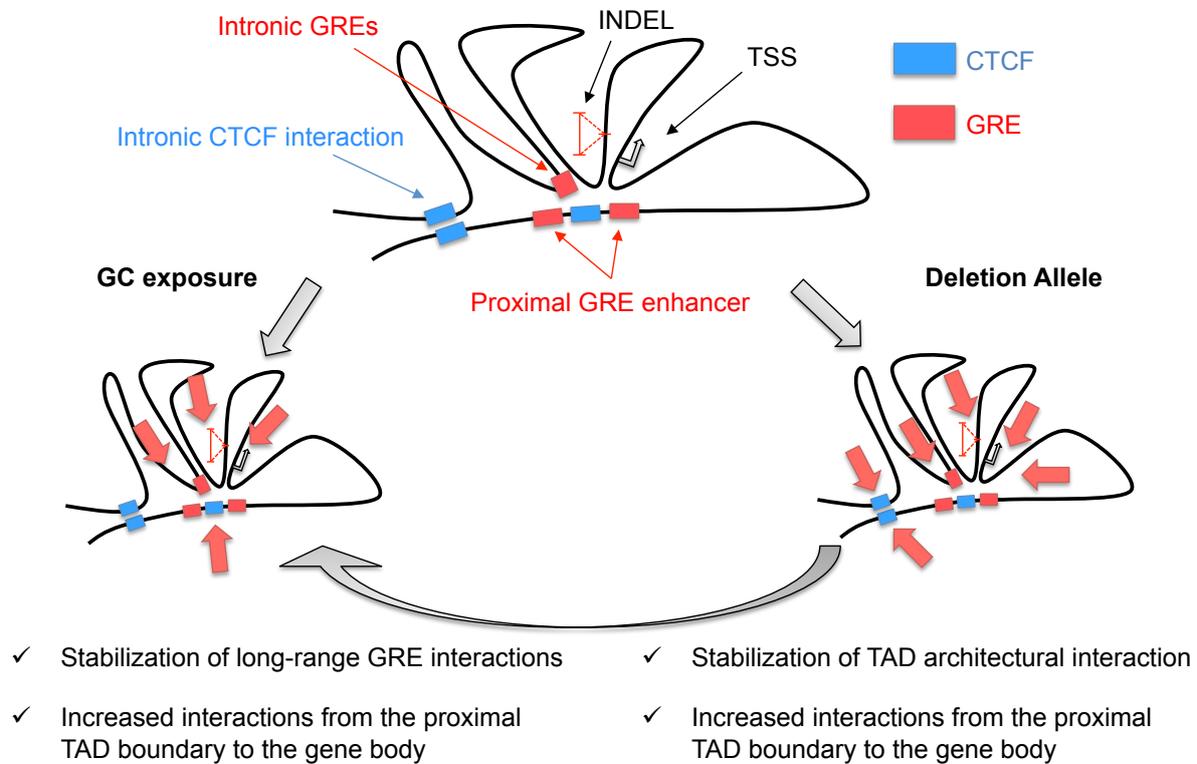
6.3 Molecular mechanisms of GxE interactions in the *FKBP5* locus

In the same disease-associated haplotype of *FKBP5* that comprises the SNP rs1360780, which is involved in a GxE mechanism shaping the risk for stress-related psychiatric disorders (Klengel et al., 2013), a larger structural variant can be found (Pelleymounter et al., 2011). This 3.3 kb INDEL (esv3608688) is located in intron 1 of the *FKBP5* gene. The aim of this thesis was to better understand, how the combination of rs1360780 and INDEL allele in intron 1 influence the chromatin structure and therefore the potentially affect *FKBP5* gene expression under basal and stress conditions. Moreover, the aim is to investigate which allele combination is more responsive to induction by glucocorticoids and possibly induce other epigenetic changes like DNAm levels of functional elements in the *FKBP5* locus.

By using allele-specific gene expression assays and heterozygote LCLs for rs1360780 allele we could show that a higher proportion of mRNA is generated from the T-allele in comparison to the C-allele. This is in line with the results from Klengel et al. (2013), who showed that the T-allele has a higher potential to induce gene expression in gene reporter assays. Furthermore, under the excess of GCs the proportion of mRNA generated from the T-allele is further increased in the LCLs. Interestingly, this differential gene expression response of rs1360780 allele is only observed in cells carrying the insertion allele. Moreover, we monitored the highest proportions of transcripts from the T-allele under the background of the deletion allele. In addition, the highest levels of total *FKBP5* transcripts under the excess of GCs were observed in LCLs combining the T- and deletion allele. These data indicate that the INDEL allele does not only moderate the differential effect of rs1360780 allele on *FKBP5* expression, but moreover both variants affect total *FKBP5* transcript levels in an additive manner upon GC stimulation. The effects of rs1360780 allele onto *FKBP5* expression can be explained on a molecular level by the establishment of an enhancer-promoter loop due to the formation of a TATA-box binding motif by the T-allele (Klengel et al., 2013). This leads to higher proportions of mRNA at baseline and even higher mRNA proportions from the T-allele under GC stimulation. In order to better understand how the large INDEL allele can modulate the rs136780 effects on a molecular level, we obtained high-resolution interaction profiles of the *FKBP5* locus TSS and TAD applying a 4C approach. A benefit of our

4C viewpoint design was the simultaneous monitoring of interaction profiles of both viewpoints in the same sample. This was achieved by choosing the same 4 base pair restriction enzymes for the definition of both viewpoints. The interaction profiles obtained from LCLs showed locus interactions that are in line with the reference literature and revealed architectural loops mediated by CTCF but also enhancer-promoter loops between the proximal GR and intronic GR-associated enhancers to the TSS ((Klengel et al., 2013; Paakinaho et al., 2010), CTCF-ChiA-PET, PolIII-ChiA-PET from Tang et al. (2015), Hi-C (Rao et al., 2014)). A comparison of the interaction profiles from cells carrying the insertion or deletion allele showed by interaction frequency peaks at position of the INDEL that the variant can be in close proximity to the TSS and TAD boundary. This involvement in the *FKBP5* locus architectural and functional loop network indicates a potential ability to modulate gene expression, which is a first explanation for the INDEL expression effects monitored in these cells. Moreover, using the 4C-cker pipeline (Raviram et al., 2016) and DESeq (Love et al., 2014), we were able to quantify the differences between the interaction profiles between the INDEL allele and DEX stimulation. The stimulation with DEX did not change the overall *FKBP5* interaction profiles indicating the described pre-established nature as a pre-poised gene (Jaaskelainen et al., 2011; Paakinaho et al., 2010). However, we quantified an increase of interaction frequencies from the TSS to the proximal GR associated enhancer. This most likely indicates the observed gene expression induction due to GC excess through the long-range GR enhancer. In addition to the specific stabilization of promoter-enhancer looping, we detect an increase of interaction frequencies from the proximal TAD boundary over the *FKBP5* locus upon DEX stimulation. Together with the observation of increased interaction frequencies of the proximal to the intronic TAD boundary in cells harboring the deletion allele, we conclude that the INDEL allele and GC excess may affect the stability of the architectural CTCF mediated loop linking genotype and environmental signal. This interaction of genotype and environment on the level of chromatin architecture could explain the *FKBP5* expression results monitored. In fact, the concept that the stability of loops is important for regulating gene expression (Hakim et al., 2011; Le Dily et al., 2014) and that pre-existing loops favor the response to external stimuli (Grbesa and Hakim, 2017; Jin et al., 2013) has been described. In regard to the idea that a dysregulation of architectural proteins / loops could lead to changes in the stability of promoter-enhancer interactions and altered response in

gene expression (Antony et al., 2015; Quintin et al., 2014; Seitan et al., 2013), we reason the following model for the integration of INDEL allele and GC excess to modulate *FKBP5* gene expression (Figure 5).



The INDEL allele in *FKBP5* affect GRE enhancer activity through the modulation of TAD interaction frequencies and regulate the responsiveness to glucocorticoid exposure

Figure 5: A model how glucocorticoids and INDEL esv3608688 influence *FKBP5* expression by modulating its underlying chromatin structure

Source: Tobias Wiechmann, 2019

The structural basis of *FKBP5* gene expression regulation is shaped by CTCF generating architectural loopings, which are described as TAD structure observed in Hi-C data (Lieberman-Aiden et al., 2009; Tang et al., 2015). These TAD loopings aid the formation of promoter-enhancer interactions, by restricting the promoter-enhancer looping within the TAD boundaries. For *FKBP5* pre-existing proximal & intronic GR enhancer loopings to the promoter (Paakinaho et al., 2010) are important to mediate the gene expression response after GR activation by GCs. The GC-induced expression changes are associated with increased interaction frequencies between long-range promoter-enhancer loops but also from the proximal TAD boundary to the body of the *FKBP5* gene. Genetic variants in *FKBP5* locus can modulate the response to GR activation by affecting either GR-associated promoter-

enhancer loops or architectural CTCF mediated interactions. For rs1360780, a loop between a GR enhancer in intron 2 and the promoter is established due to the formation of an additional TATA-box binding motif of the T-allele, which can increase the GC regulated *FKBP5* induction (Klengel et al., 2013). The 3.3 kb INDEL on the other hand is associated to moderate the stabilization of the architectural loop between the proximal TAD and intronic TAD boundary. In this scenario, the deletion is associated to lead to increased interaction frequencies between TAD boundaries and production of *FKBP5* transcripts. The common feature in this model is that genetic variants and GR activation are associated to stabilizing loops within the *FKBP5* locus which is reflected by increased interaction frequencies and additive effects onto gene expression. In other words, the activity of the GR enhancers responsible for *FKBP5* gene expression regulation is on the one hand regulated by the activation of the hormone receptor itself but also their probability to find the *FKBP5* promoter. This probability is regulated by genetic variants modulating the stability of the architectural interactions (INDEL esv3608688 allele) but also promoter-enhancer interactions (rs1360780 allele). Indications for the concept that the enhancer activity is regulated by the overarching TAD stability have been found by (Symmons et al., 2016) by manipulating the *Shh* locus. To prove the causality of this concept for the *FKBP5* locus similar locus manipulation experiments using the CRISPR/Cas9 system need to be performed.

The involvement of other chromatin features defining of the activity of enhancers is plausible and DNAm levels have been proposed to be instructive for the activity of GR enhancers (Wiench et al., 2011). Using the HAM-TBS workflow we assessed the DNAm levels of 46 sites covering key functional regions of the *FKBP5* locus (TSS, TAD boundaries, proximal and intronic GR enhancer), the region of the INDEL and other sites between TSS and INDEL of LCLs. The CpGs within the insertion allele show high methylation levels ($\text{Insertion}_{\text{VEH}} = 90.7 \pm 8.5 \%$, $n = 33$). We observed significant genotype effects onto DNAm levels of 36 CpGs at proximal & distal TAD boundaries, TSS, GRE in intron 5 and GREs at the proximal enhancer. Effects onto methylation levels due to the INDEL allele have been monitored in both directions and depend on the location of the respective CpGs. Moreover, we detected a genotype effect onto methylation levels in an array of CpGs located in intron 1 (hg19 / chr6:35635518-35636055). This region showed increased methylation levels in the background of the insertion allele and has not yet been described to include

important regulatory sites. However, a potential role as an enhancer of this region in LCLs is implicated by the chromatin states of the ENCODE project using multiple epigenetic information and hidden markov models (Ernst and Kellis, 2012). No DEX-induced methylation changes have been detected in the LCLs. Next to genotype effects of the INDEL allele onto chromatin interactions we detect DNA methylation levels changes at key functional sites involved in the formation of architectural and promoter-enhancer loops indicating a potential integration of both epigenetic layers. The absence of DEX-inducible methylation changes in the LCLs could be explained, on the one hand, by the chosen treatment conditions or tissue specific inability of the LCLs to respond to GR activation. To further dissect the importance of the INDEL allele onto chromatin interactions and DNA methylation levels to regulate *FKBP5* expression, experiments combining the manipulation of the DNA sequence but also the methylation levels of key regulatory sites of *FKBP5* should be applied. Moreover, the observation of eRNA being produced from active enhancers (Mikhaylichenko et al., 2018) could be used to detect *in vivo* enhancer activity instructed by the production of eRNA from the respective enhancer in *FKBP5*. This would help to fine-map the contribution of each specific GR-associated enhancer and better understand which manipulations of the DNA sequence or DNA methylation levels affect the activity of the *FKBP5* enhancers.

Although, there are preliminary findings from data obtained from PTSD patients that the INDEL allele moderate the symptom severity in a GxE manner, further analyses need to be performed to enhance the evidence for the INDEL allele to moderate the development of stress related disease phenotypes. Due to the development of the HAM-TBS workflow, there is now the possibility to investigate the INDEL effect onto the methylation level in patients.

In summary, the data presented in this chapter extends the proposed molecular mechanism (Klengel et al., 2013) integrating genetic and environmental factors at the *FKBP5* locus. We identified that the stabilization of architectural and enhancer-promoter loops is a common feature of the factors (T-allele, deletion and GR activation by GCs) leading to increased *FKBP5* mRNA expression. We and others (Symmons et al. (2016) and discussed by Beagrie and Pombo (2016), (Hakim et al., 2011); Le Dily et al. (2014)) find indications that this gene expression response is most likely result of an increased activity of enhancers due to the stabilization of chromatin interactions. However, further experiments manipulating the DNA

sequence and DNA methylation levels of *FKBP5* key regulatory sites will be necessary to obtain causality. Moreover, further analysis in patients will help to clarify the role of INDEL allele in the development of stress-related phenotypes.

6.4 Future perspectives towards translation

The syndrome-based disease classification has been identified to be one of the major challenges in psychiatry (Stephan et al., 2016). In order to enable the translation of the rapidly increasing knowledge of our genome and the regulation of its output into the clinical routine, we need to overcome the over-dependency of subject questionnaires and favor an integrative classification system that is also based on biological information of the individuals. Promising approaches to establish such integrative systems are the Research Domain Criteria (RDoC) Initiative of the National Institute of Mental Health (NIH) or the BeCOME study (Biological Classification of mental disorders) of the Max-Planck Institute for Psychiatry (<https://www.psych.mpg.de/become>). The work of Consortia is especially important to enable the integration of multi-omics data and promote its availability to better investigate complex diseases. In the context of psychiatry the PsychENCODE Consortium recently developed a comprehensive functional genomic resource for the adult human brain (Wang et al., 2018a). This resource includes various epigenomic datasets highlighting the importance of understanding epigenetics in psychiatry. Biomarkers enabling patient stratification developed by molecular and biological insights of disease mechanisms could have a beneficial impact onto clinical decision-making and improve prevention and personalized treatment of individuals. The only U.S. Food and Drug Administration (FDA)-approved genetic biomarkers in the field of psychiatry is the genotyping of CYP2D6 and CYP2C19 allele (Winner et al., 2013). In this case the biomarker is used to detect inter-individual differences of antidepressant metabolism and therefore the choice and dose of antidepressant can be reflected by the clinician accordingly. Other promising candidate genes to serve as a biomarker in the context of stress-related disorders are *FKBP5*, *DDX6*, *B2M*, *LAIR1*, *RTN4* and *NUB1* (Le-Niculescu et al., 2019). *FKBP5* gene expression regulation has been shown to involve complex interactions of genetic variants but also environmental stressors that can lead to *FKBP5* disinhibition. This disinhibition has been shown to co-occur in a great amount of aberrant phenotypes in rodents and

humans (Matosin et al., 2018; Zannas et al., 2016). Therefore, a promising treatment intervention for at least a subset of patients suffering from stress-related disorders would be the blockage of *FKBP5*. The development of selective FKBP51 antagonist (Gaali et al., 2015) showed first encouraging results in rodents by reducing anxiety and increasing stress-coping (Gaali et al., 2015; Hartmann et al., 2015) as well as increasing GR sensitivity and promoting the negative feedback of HPA axis regulation (Gaali et al., 2015). A caveat for the use of FKBP51 antagonist is the pleiotropic function of this chaperone in several signaling pathways (Zannas et al., 2016). However, FKBP51 blockage may also lead to adverse consequences owing to the association that *FKBP5* genotypes leading to the increased response might also favor antidepressant responses (Binder et al., 2004; Zou et al., 2010) and that FKBP51 can promote antidepressant activity (Gassen et al., 2016; Gassen et al., 2014). In this regard, the same alleles but in different environmental contexts can act negatively but also positively, therefore the term “risk” or “protective” *FKBP5* allele have to be seen in context of the specific environmental setting and may be better termed “environmental sensitive” or “plasticity” alleles (Klengel and Binder, 2013). An example of this is the observation that rs1360780 T-allele carriers, which were exposed to childhood and adult trauma revealed an increased likelihood to exhibit a psychiatric disease phenotype. However, when T-allele carriers were exposed to adult trauma only, this likelihood was decreased (Klengel and Binder, 2013; Klengel et al., 2013). Key questions are, if the environmental sensitivity of these alleles in can be mediated by epigenetic changes and if we can steer those to achieve beneficial outcomes for individuals. The described GC-induced dynamic methylation changes and the moderation by rs1360780 allele (Paper III) can be seen as an indicator for the potential of epigenetic marks being involved in the differential response to the environmental context. The ability to manipulate the environmental sensitivity to favor the response to the environment onto trajectories of beneficial outcomes could be a great opportunity for the treatment of patients. *FKBP5* genotypes and DNA methylation levels could serve as a marker for the current status of the sensitivity to the environment and help to monitor the effectiveness of interventions to steer the sensitivity to facilitate beneficial outcomes. To translate the concept of steering environmental sensitivity of individuals, one has to gain further insights into the epigenetic molecular mechanisms involved. Further experiments should address, if the environmental sensitivity can be modulated in animal models linking the changes

in environmental sensitivity to behavioral outputs. In addition, delivering the tools for biomarker research to apply this concept in a clinical setting.

Potential *FKBP5* biomarker for a biologically-based stratification of patients to enable an enhanced personalized treatment of patients with stress-related diseases will probably include the detection of multiple *FKBP5* variants but also epigenetic markers like DNA methylation levels of distinct CpGs (Figure 6). The future progress of tools for biomarker research such as HAM-TBS and the *FKBP5* panel will be essential to enable the development of epigenetic-based biomarkers.

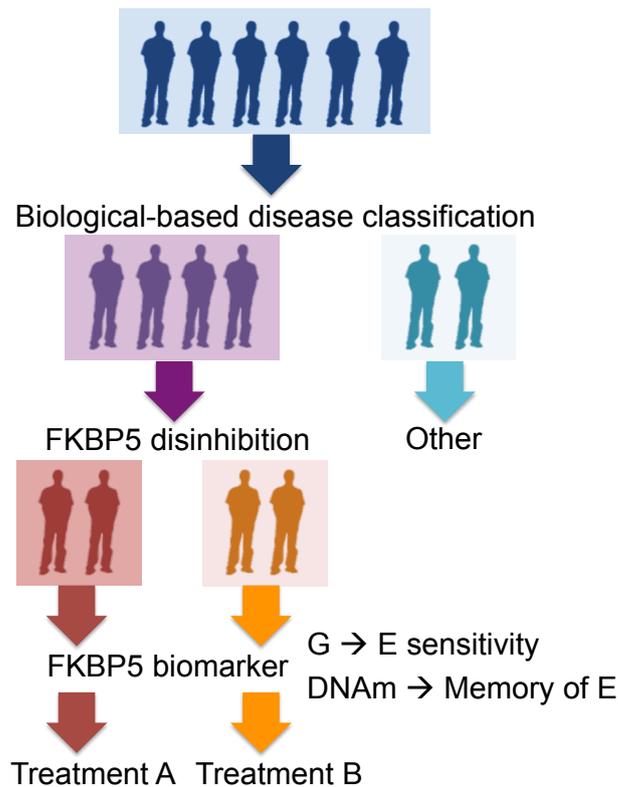


Figure 6: A biological-based disease classification system and potential FKBP5 biomarker have a great potential to enhance translation in psychiatry

In order to enable the translation of the increasing knowledge of our genome and the regulation of its output, a shift from a syndrome-based disease classification system to an integrated biological-based system seems promising. This classification system will potentially include FKBP5 disinhibition to indicate a substructure of patients in which the blockage of FKBP5 or manipulation of environmental sensitivity is promising. FKBP5 biomarker should include genotypes but also DNA methylation levels to instruct the best treatment choice for patients displaying FKBP5 disinhibition. Source: Tobias Wiechmann, 2019

7 References

- Adcock, I.M., Ito, K., and Barnes, P.J. (2005). Histone deacetylation: an important mechanism in inflammatory lung diseases. *COPD* 2, 445-455.
- Ambrosi, C., Manzo, M., and Baubec, T. (2017). Dynamics and Context-Dependent Roles of DNA Methylation. *J Mol Biol* 429, 1459-1475.
- Antony, J., Dasgupta, T., Rhodes, J.M., McEwan, M.V., Print, C.G., O'Sullivan, J.M., and Horsfield, J.A. (2015). Cohesin modulates transcription of estrogen-responsive genes. *Biochim Biophys Acta* 1849, 257-269.
- Appel, K., Schwahn, C., Mahler, J., Schulz, A., Spitzer, C., Fenske, K., Stender, J., Barnow, S., John, U., Teumer, A., *et al.* (2011). Moderation of adult depression by a polymorphism in the FKBP5 gene and childhood physical abuse in the general population. *Neuropsychopharmacology* 36, 1982-1991.
- Barros-Silva, D., Marques, C.J., Henrique, R., and Jeronimo, C. (2018). Profiling DNA Methylation Based on Next-Generation Sequencing Approaches: New Insights and Clinical Applications. *Genes (Basel)* 9, 429-440.
- Barutcu, A.R., Lajoie, B.R., McCord, R.P., Tye, C.E., Hong, D., Messier, T.L., Browne, G., van Wijnen, A.J., Lian, J.B., Stein, J.L., *et al.* (2015). Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells. *Genome Biol* 16, 214.
- Baylin, S.B., and Jones, P.A. (2011). A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer* 11, 726-734.
- Beagrie, R.A., and Pombo, A. (2016). Examining Topological Domain Influence on Enhancer Function. *Dev Cell* 39, 523-524.
- Bell, J.T., and Spector, T.D. (2011). A twin approach to unraveling epigenetics. *Trends Genet* 27, 116-125.
- Bell, J.T., Tsai, P.C., Yang, T.P., Pidsley, R., Nisbet, J., Glass, D., Mangino, M., Zhai, G., Zhang, F., Valdes, A., *et al.* (2012). Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet* 8, e1002629.
- Bender, C.M., Pao, M.M., and Jones, P.A. (1998). Inhibition of DNA methylation by 5-aza-2'-deoxycytidine suppresses the growth of human tumor cell lines. *Cancer Res* 58, 95-101.
- Bergman, Y., and Cedar, H. (2013). DNA methylation dynamics in health and disease. *Nat Struct Mol Biol* 20, 274-281.
- Berman, B.P., Weisenberger, D.J., Aman, J.F., Hinoue, T., Ramjan, Z., Liu, Y., Noushmehr, H., Lange, C.P., van Dijk, C.M., Tollenaar, R.A., *et al.* (2011). Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. *Nat Genet* 44, 40-46.

References

- Bernstein, D.L., Kameswaran, V., Le Lay, J.E., Sheaffer, K.L., and Kaestner, K.H. (2015). The BisPCR(2) method for targeted bisulfite sequencing. *Epigenetics Chromatin* 8, 27.
- Bevilacqua, L., Carli, V., Sarchiapone, M., George, D.K., Goldman, D., Roy, A., and Enoch, M.A. (2012). Interaction between FKBP5 and childhood trauma and risk of aggressive behavior. *Arch Gen Psychiatry* 69, 62-70.
- Biddie, S.C. (2011). Chromatin architecture and the regulation of nuclear receptor inducible transcription. *J Neuroendocrinol* 23, 94-106.
- Binder, E.B. (2009). The role of FKBP5, a co-chaperone of the glucocorticoid receptor in the pathogenesis and therapy of affective and anxiety disorders. *Psychoneuroendocrinology* 34 Suppl 1, S186-195.
- Binder, E.B., Bradley, R.G., Liu, W., Epstein, M.P., Deveau, T.C., Mercer, K.B., Tang, Y., Gillespie, C.F., Heim, C.M., Nemeroff, C.B., *et al.* (2008). Association of FKBP5 polymorphisms and childhood abuse with risk of posttraumatic stress disorder symptoms in adults. *JAMA* 299, 1291-1305.
- Binder, E.B., Salyakina, D., Lichtner, P., Wochnik, G.M., Ising, M., Putz, B., Papiol, S., Seaman, S., Lucae, S., Kohli, M.A., *et al.* (2004). Polymorphisms in FKBP5 are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. *Nat Genet* 36, 1319-1325.
- Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev* 16, 6-21.
- Bird, A., Taggart, M., Frommer, M., Miller, O.J., and Macleod, D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell* 40, 91-99.
- Bird, A.P. (1980). DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res* 8, 1499-1504.
- Bohacek, J., Gapp, K., Saab, B.J., and Mansuy, I.M. (2013). Transgenerational epigenetic effects on brain functions. *Biol Psychiatry* 73, 313-320.
- Boissonnas, C.C., Abdalaoui, H.E., Haelewyn, V., Fauque, P., Dupont, J.M., Gut, I., Vaiman, D., Jouannet, P., Tost, J., and Jammes, H. (2010). Specific epigenetic alterations of IGF2-H19 locus in spermatozoa from infertile men. *Eur J Hum Genet* 18, 73-80.
- Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. *Nat Rev Genet* 17, 661-678.
- Boscarino, J.A., Erlich, P.M., Hoffman, S.N., and Zhang, X. (2012). Higher FKBP5, COMT, CHRNA5, and CRHR1 allele burdens are associated with PTSD and interact with trauma exposure: implications for neuropsychiatric research and treatment. *Neuropsychiatr Dis Treat* 8, 131-139.
- Bradley, R.G., Binder, E.B., Epstein, M.P., Tang, Y., Nair, H.P., Liu, W., Gillespie, C.F., Berg, T., Evces, M., Newport, D.J., *et al.* (2008). Influence of child abuse on

References

adult depression: moderation by the corticotropin-releasing hormone receptor gene. *Arch Gen Psychiatry* 65, 190-200.

Branco, M.R., and Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol* 4, e138.

Brinkman, A.B., Simmer, F., Ma, K., Kaan, A., Zhu, J., and Stunnenberg, H.G. (2010). Whole-genome DNA methylation profiling using MethylCap-seq. *Methods* 52, 232-236.

Buchwalter, A., Kaneshiro, J.M., and Hetzer, M.W. (2019). Coaching from the sidelines: the nuclear periphery in genome regulation. *Nat Rev Genet* 20, 39-50.

Burgess, D.J. (2017). Epigenetics: Rich pore methods for DNA methylation detection. *Nat Rev Genet* 18, 209.

Cannon, M.V., Pilarowski, G., Liu, X., and Serre, D. (2016). Extensive Epigenetic Changes Accompany Terminal Differentiation of Mouse Hepatocytes After Birth. *G3 (Bethesda)* 6, 3701-3709.

Caspi, A., Hariri, A.R., Holmes, A., Uher, R., and Moffitt, T.E. (2010). Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *Am J Psychiatry* 167, 509-527.

Caspi, A., and Moffitt, T.E. (2006). Gene-environment interactions in psychiatry: joining forces with neuroscience. *Nat Rev Neurosci* 7, 583-590.

Chakraborty, A., and Ay, F. (2018). The role of 3D genome organization in disease: From compartments to single nucleotides. *Semin Cell Dev Biol*.

Chapuy, B., McKeown, M.R., Lin, C.Y., Monti, S., Roemer, M.G., Qi, J., Rahl, P.B., Sun, H.H., Yeda, K.T., Doench, J.G., *et al.* (2013). Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer Cell* 24, 777-790.

Chen, G.G., Gross, J.A., Lutz, P.E., Vaillancourt, K., Maussion, G., Bramouille, A., Theroux, J.F., Gardini, E.S., Ehler, U., Bourret, G., *et al.* (2017). Medium throughput bisulfite sequencing for accurate detection of 5-methylcytosine and 5-hydroxymethylcytosine. *BMC Genomics* 18, 96.

Chrousos, G.P., and Gold, P.W. (1992). The concepts of stress and stress system disorders. Overview of physical and behavioral homeostasis. *JAMA* 267, 1244-1252.

Collip, D., Myin-Germeys, I., Wichers, M., Jacobs, N., Derom, C., Thiery, E., Lataster, T., Simons, C., Delespaul, P., Marcelis, M., *et al.* (2013). FKBP5 as a possible moderator of the psychosis-inducing effects of childhood trauma. *Br J Psychiatry* 202, 261-268.

consortium, B. (2016). Quantitative comparison of DNA methylation assays for biomarker development and clinical applications. *Nat Biotechnol* 34, 726-737.

References

- Consortium, E.P. (2004). The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* *306*, 636-640.
- Coulondre, C., Miller, J.H., Farabaugh, P.J., and Gilbert, W. (1978). Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature* *274*, 775-780.
- Cox, J.J., Willatt, L., Homfray, T., and Woods, C.G. (2011). A SOX9 duplication and familial 46,XX developmental testicular disorder. *N Engl J Med* *364*, 91-93.
- Crane, E., Bian, Q., McCord, R.P., Lajoie, B.R., Wheeler, B.S., Ralston, E.J., Uzawa, S., Dekker, J., and Meyer, B.J. (2015). Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* *523*, 240-244.
- Cremer, T., and Cremer, M. (2010). Chromosome territories. *Cold Spring Harb Perspect Biol* *2*, a003889.
- Croft, J.A., Bridger, J.M., Boyle, S., Perry, P., Teague, P., and Bickmore, W.A. (1999). Differences in the localization and morphology of chromosomes in the human nucleus. *J Cell Biol* *145*, 1119-1131.
- Cruickshank, M.N., Oshlack, A., Theda, C., Davis, P.G., Martino, D., Sheehan, P., Dai, Y., Saffery, R., Doyle, L.W., and Craig, J.M. (2013). Analysis of epigenetic changes in survivors of preterm birth reveals the effect of gestational age and evidence for a long term legacy. *Genome Med* *5*, 96.
- Dackis, M.N., Rogosch, F.A., Oshri, A., and Cicchetti, D. (2012). The role of limbic system irritability in linking history of childhood maltreatment and psychiatric outcomes in low-income, high-risk women: moderation by FK506 binding protein 5 haplotype. *Dev Psychopathol* *24*, 1237-1252.
- Davies, J.O., Oudelaar, A.M., Higgs, D.R., and Hughes, J.R. (2017). How best to identify chromosomal interactions: a comparison of approaches. *Nat Methods* *14*, 125-134.
- Daxinger, L., and Whitelaw, E. (2012). Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nat Rev Genet* *13*, 153-162.
- De Gobbi, M., Viprakasit, V., Hughes, J.R., Fisher, C., Buckle, V.J., Ayyub, H., Gibbons, R.J., Vernimmen, D., Yoshinaga, Y., de Jong, P., *et al.* (2006). A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. *Science* *312*, 1215-1217.
- De Kloet, E.R., Vreugdenhil, E., Oitzl, M.S., and Joels, M. (1998). Brain corticosteroid receptor balance in health and disease. *Endocr Rev* *19*, 269-301.
- Dekker, J. (2006). The three 'C' s of chromosome conformation capture: controls, controls, controls. *Nat Methods* *3*, 17-21.
- Dekker, J. (2016). Mapping the 3D genome: Aiming for consilience. *Nat Rev Mol Cell Biol* *17*, 741-742.
- Dekker, J., and Heard, E. (2015). Structural and functional diversity of Topologically Associating Domains. *FEBS Lett* *589*, 2877-2884.

References

- Dekker, J., and Mirny, L. (2013). Biological techniques: Chromosomes captured one by one. *Nature* *502*, 45-46.
- Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. *Science* *295*, 1306-1311.
- Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., *et al.* (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature* *518*, 331-336.
- Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* *485*, 376-380.
- Dor, Y., and Cedar, H. (2018). Principles of DNA methylation and their implications for biology and medicine. *Lancet* *392*, 777-786.
- Dos Santos, C.O., Dolzhenko, E., Hodges, E., Smith, A.D., and Hannon, G.J. (2015). An epigenetic memory of pregnancy in the mouse mammary gland. *Cell Rep* *11*, 1102-1109.
- Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst* *3*, 99-101.
- Ehara, T., Kamei, Y., Yuan, X., Takahashi, M., Kanai, S., Tamura, E., Tsujimoto, K., Tamiya, T., Nakagawa, Y., Shimano, H., *et al.* (2015). Ligand-activated PPAR α -dependent DNA demethylation regulates the fatty acid beta-oxidation genes in the postnatal liver. *Diabetes* *64*, 775-784.
- Ehrlich, M., Gama-Sosa, M.A., Huang, L.H., Midgett, R.M., Kuo, K.C., McCune, R.A., and Gehrke, C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res* *10*, 2709-2721.
- Ellsworth, K.A., Eckloff, B.W., Li, L., Moon, I., Fridley, B.L., Jenkins, G.D., Carlson, E., Brisbin, A., Abo, R., Bamlet, W., *et al.* (2013a). Contribution of FKBP5 genetic variation to gemcitabine treatment and survival in pancreatic adenocarcinoma. *PLoS One* *8*, e70216.
- Ellsworth, K.A., Moon, I., Eckloff, B.W., Fridley, B.L., Jenkins, G.D., Batzler, A., Biernacka, J.M., Abo, R., Brisbin, A., Ji, Y., *et al.* (2013b). FKBP5 genetic variation: association with selective serotonin reuptake inhibitor treatment outcomes in major depressive disorder. *Pharmacogenet Genomics* *23*, 156-166.
- Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* *9*, 215-216.
- Esteller, M. (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet* *8*, 286-298.
- Flavahan, W.A., Drier, Y., Liau, B.B., Gillespie, S.M., Venteicher, A.S., Stemmer-Rachamimov, A.O., Suva, M.L., and Bernstein, B.E. (2016). Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* *529*, 110-114.

References

- Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suner, D., Cigudosa, J.C., Urioste, M., Benitez, J., *et al.* (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* *102*, 10604-10609.
- Franklin, T.B., Russig, H., Weiss, I.C., Graff, J., Linder, N., Michalon, A., Vizi, S., and Mansuy, I.M. (2010). Epigenetic transmission of the impact of early stress across generations. *Biol Psychiatry* *68*, 408-415.
- Fritsch, E.F., Lawn, R.M., and Maniatis, T. (1979). Characterisation of deletions which affect the expression of fetal globin genes in man. *Nature* *279*, 598-603.
- Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L., and Paul, C.L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A* *89*, 1827-1831.
- Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H., *et al.* (2009). An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature* *462*, 58-64.
- Gaali, S., Kirschner, A., Cuboni, S., Hartmann, J., Kozany, C., Balsevich, G., Namendorf, C., Fernandez-Vizarra, P., Sippel, C., Zannas, A.S., *et al.* (2015). Selective inhibitors of the FK506-binding protein 51 by induced fit. *Nat Chem Biol* *11*, 33-37.
- Gapp, K., Soldado-Magraner, S., Alvarez-Sanchez, M., Bohacek, J., Vernaz, G., Shu, H., Franklin, T.B., Wolfer, D., and Mansuy, I.M. (2014). Early life stress in fathers improves behavioural flexibility in their offspring. *Nat Commun* *5*, 5466.
- Gassen, N.C., Hartmann, J., Zannas, A.S., Kretschmar, A., Zschocke, J., Maccarrone, G., Hafner, K., Zellner, A., Kollmannsberger, L.K., Wagner, K.V., *et al.* (2016). FKBP51 inhibits GSK3beta and augments the effects of distinct psychotropic medications. *Mol Psychiatry* *21*, 277-289.
- Gassen, N.C., Hartmann, J., Zschocke, J., Stepan, J., Hafner, K., Zellner, A., Kirmeier, T., Kollmannsberger, L., Wagner, K.V., Dedic, N., *et al.* (2014). Association of FKBP51 with priming of autophagy pathways and mediation of antidepressant treatment response: evidence in cells, mice, and humans. *PLoS Med* *11*, e1001755.
- Gerken, T., Girard, C.A., Tung, Y.C., Webby, C.J., Saudek, V., Hewitson, K.S., Yeo, G.S., McDonough, M.A., Cunliffe, S., McNeill, L.A., *et al.* (2007). The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science* *318*, 1469-1472.
- Gibcus, J.H., and Dekker, J. (2013). The hierarchy of the 3D genome. *Mol Cell* *49*, 773-782.
- Goll, M.G., and Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* *74*, 481-514.
- Gomez-Marin, C., Tena, J.J., Acemel, R.D., Lopez-Mayorga, M., Naranjo, S., de la Calle-Mustienes, E., Maeso, I., Beccari, L., Aneas, I., Viernas, E., *et al.* (2015).

References

- Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. *Proc Natl Acad Sci U S A* *112*, 7542-7547.
- Goubau, C., Devriendt, K., Van der Aa, N., Crepel, A., Wieczorek, D., Kleefstra, T., Willemsen, M.H., Rauch, A., Tzschach, A., de Ravel, T., *et al.* (2013). Platelet defects in congenital variant of Rett syndrome patients with FOXP1 mutations or reduced expression due to a position effect at 14q12. *Eur J Hum Genet* *21*, 1349-1355.
- Grange, T., Cappabianca, L., Flavin, M., Sassi, H., and Thomassin, H. (2001). In vivo analysis of the model tyrosine aminotransferase gene reveals multiple sequential steps in glucocorticoid receptor action. *Oncogene* *20*, 3028-3038.
- Grasser, F., Neusser, M., Fiegler, H., Thormeyer, T., Cremer, M., Carter, N.P., Cremer, T., and Muller, S. (2008). Replication-timing-correlated spatial chromatin arrangements in cancer and in primate interphase nuclei. *J Cell Sci* *121*, 1876-1886.
- Grbesa, I., and Hakim, O. (2017). Genomic effects of glucocorticoids. *Protoplasma* *254*, 1175-1185.
- Hakim, O., John, S., Ling, J.Q., Biddie, S.C., Hoffman, A.R., and Hager, G.L. (2009). Glucocorticoid receptor activation of the Ciz1-Lcn2 locus by long range interactions. *J Biol Chem* *284*, 6048-6052.
- Hakim, O., Sung, M.H., Voss, T.C., Splinter, E., John, S., Sabo, P.J., Thurman, R.E., Stamatoyannopoulos, J.A., de Laat, W., and Hager, G.L. (2011). Diverse gene reprogramming events occur in the same spatial clusters of distal regulatory elements. *Genome Res* *21*, 697-706.
- Handel, A.E., Ebers, G.C., and Ramagopalan, S.V. (2010). Epigenetics: molecular mechanisms and implications for disease. *Trends Mol Med* *16*, 7-16.
- Harms, M.B., Birn, R., Provencal, N., Wiechmann, T., Binder, E.B., Giakas, S.W., Roeber, B.J., and Pollak, S.D. (2017). Early life stress, FK506 binding protein 5 gene (FKBP5) methylation, and inhibition-related prefrontal function: A prospective longitudinal study. *Dev Psychopathol* *29*, 1895-1903.
- Harmston, N., and Lenhard, B. (2013). Chromatin and epigenetic features of long-range gene regulation. *Nucleic Acids Res* *41*, 7185-7199.
- Harrison, A., and Parle-McDermott, A. (2011). DNA methylation: a timeline of methods and applications. *Front Genet* *2*, 74.
- Hartmann, J., Wagner, K.V., Gaali, S., Kirschner, A., Kozany, C., Ruhter, G., Dedic, N., Hausl, A.S., Hoeijmakers, L., Westerholz, S., *et al.* (2015). Pharmacological Inhibition of the Psychiatric Risk Factor FKBP51 Has Anxiolytic Properties. *J Neurosci* *35*, 9007-9016.
- Hayatsu, H., Wataya, Y., Kai, K., and Iida, S. (1970). Reaction of sodium bisulfite with uracil, cytosine, and their derivatives. *Biochemistry* *9*, 2858-2865.

References

- Heim, C., Nater, U.M., Maloney, E., Boneva, R., Jones, J.F., and Reeves, W.C. (2009). Childhood trauma and risk for chronic fatigue syndrome: association with neuroendocrine dysfunction. *Arch Gen Psychiatry* 66, 72-80.
- Hemmerich, P., Schmiedeberg, L., and Diekmann, S. (2011). Dynamic as well as stable protein interactions contribute to genome function and maintenance. *Chromosome Res* 19, 131-151.
- Hernaus, D., van Winkel, R., Gronenschild, E., Habets, P., Kenis, G., Marcelis, M., van Os, J., Myin-Germeys, I., Collip, D., for Genetic, R., *et al.* (2014). Brain-derived neurotrophic factor/FK506-binding protein 5 genotype by childhood trauma interactions do not impact on hippocampal volume and cognitive performance. *PLoS One* 9, e92722.
- Hnisz, D., Weintraub, A.S., Day, D.S., Valton, A.L., Bak, R.O., Li, C.H., Goldmann, J., Lajoie, B.R., Fan, Z.P., Sigova, A.A., *et al.* (2016). Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* 351, 1454-1458.
- Hsieh, T.H., Weiner, A., Lajoie, B., Dekker, J., Friedman, N., and Rando, O.J. (2015). Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. *Cell* 162, 108-119.
- Iguchi-Ariga, S.M., and Schaffner, W. (1989). CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. *Genes Dev* 3, 612-619.
- Jaaskelainen, T., Makkonen, H., and Palvimo, J.J. (2011). Steroid up-regulation of FKBP51 and its role in hormone signaling. *Curr Opin Pharmacol* 11, 326-331.
- Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 33 *Suppl*, 245-254.
- Jawahar, M.C., Murgatroyd, C., Harrison, E.L., and Baune, B.T. (2015). Epigenetic alterations following early postnatal stress: a review on novel aetiological mechanisms of common psychiatric disorders. *Clin Epigenetics* 7, 122.
- Jeong, Y., El-Jaick, K., Roessler, E., Muenke, M., and Epstein, D.J. (2006). A functional screen for sonic hedgehog regulatory elements across a 1 Mb interval identifies long-range ventral forebrain enhancers. *Development* 133, 761-772.
- Ji, X., Dadon, D.B., Powell, B.E., Fan, Z.P., Borges-Rivera, D., Shachar, S., Weintraub, A.S., Hnisz, D., Pegoraro, G., Lee, T.I., *et al.* (2016). 3D Chromosome Regulatory Landscape of Human Pluripotent Cells. *Cell Stem Cell* 18, 262-275.
- Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 503, 290-294.
- Jirtle, R.L. (2014). The Agouti mouse: a biosensor for environmental epigenomics studies investigating the developmental origins of health and disease. *Epigenomics* 6, 447-450.

References

- John, S., Sabo, P.J., Thurman, R.E., Sung, M.H., Biddie, S.C., Johnson, T.A., Hager, G.L., and Stamatoyannopoulos, J.A. (2011). Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nat Genet* 43, 264-268.
- Jones, P.A., and Baylin, S.B. (2007). The epigenomics of cancer. *Cell* 128, 683-692.
- Joubert, B.R., Felix, J.F., Yousefi, P., Bakulski, K.M., Just, A.C., Breton, C., Reese, S.E., Markunas, C.A., Richmond, R.C., Xu, C.J., *et al.* (2016). DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *Am J Hum Genet* 98, 680-696.
- Kelly, M.M., King, E.M., Rider, C.F., Gwozd, C., Holden, N.S., Eddleston, J., Zuraw, B., Leigh, R., O'Byrne, P.M., and Newton, R. (2012). Corticosteroid-induced gene expression in allergen-challenged asthmatic subjects taking inhaled budesonide. *Br J Pharmacol* 165, 1737-1747.
- Kendler, K.S., Karkowski, L.M., and Prescott, C.A. (1999). Causal relationship between stressful life events and the onset of major depression. *Am J Psychiatry* 156, 837-841.
- Kendler, K.S., Kessler, R.C., Walters, E.E., MacLean, C., Neale, M.C., Heath, A.C., and Eaves, L.J. (1995). Stressful life events, genetic liability, and onset of an episode of major depression in women. *Am J Psychiatry* 152, 833-842.
- Klengel, T., and Binder, E.B. (2013). Allele-specific epigenetic modification: a molecular mechanism for gene-environment interactions in stress-related psychiatric disorders? *Epigenomics* 5, 109-112.
- Klengel, T., and Binder, E.B. (2015). FKBP5 allele-specific epigenetic modification in gene by environment interaction. *Neuropsychopharmacology* 40, 244-246.
- Klengel, T., Mehta, D., Anacker, C., Rex-Haffner, M., Pruessner, J.C., Pariante, C.M., Pace, T.W., Mercer, K.B., Mayberg, H.S., Bradley, B., *et al.* (2013). Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nat Neurosci* 16, 33-41.
- Klengel, T., Pape, J., Binder, E.B., and Mehta, D. (2014). The role of DNA methylation in stress-related psychiatric disorders. *Neuropharmacology* 80, 115-132.
- Klose, R.J., and Bird, A.P. (2006). Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* 31, 89-97.
- Koenen, K.C., Saxe, G., Purcell, S., Smoller, J.W., Bartholomew, D., Miller, A., Hall, E., Kaplow, J., Bosquet, M., Moulton, S., *et al.* (2005). Polymorphisms in FKBP5 are associated with peritraumatic dissociation in medically injured children. *Mol Psychiatry* 10, 1058-1059.
- Kohrt, B.A., Worthman, C.M., Ressler, K.J., Mercer, K.B., Upadhaya, N., Koirala, S., Nepal, M.K., Sharma, V.D., and Binder, E.B. (2015). Cross-cultural gene-environment interactions in depression, post-traumatic stress disorder, and the cortisol awakening response: FKBP5 polymorphisms and childhood trauma in South Asia. *Int Rev Psychiatry* 27, 180-196.

References

- Kress, C., Thomassin, H., and Grange, T. (2006). Active cytosine demethylation triggered by a nuclear receptor involves DNA strand breaks. *Proc Natl Acad Sci U S A* *103*, 11112-11117.
- Krijger, P.H., and de Laat, W. (2016). Regulation of disease-associated gene expression in the 3D genome. *Nat Rev Mol Cell Biol* *17*, 771-782.
- Krumm, A., and Duan, Z. (2018). Understanding the 3D genome: Emerging impacts on human disease. *Semin Cell Dev Biol* *90*, 62-77.
- Laird, P.W. (2010). Principles and challenges of genomewide DNA methylation analysis. *Nat Rev Genet* *11*, 191-203.
- Laird, P.W., Jackson-Grusby, L., Fazeli, A., Dickinson, S.L., Jung, W.E., Li, E., Weinberg, R.A., and Jaenisch, R. (1995). Suppression of intestinal neoplasia by DNA hypomethylation. *Cell* *81*, 197-205.
- Le Dily, F., Bau, D., Pohl, A., Vicent, G.P., Serra, F., Soronellas, D., Castellano, G., Wright, R.H., Ballare, C., Filion, G., *et al.* (2014). Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev* *28*, 2151-2162.
- Le Dily, F., and Beato, M. (2015). TADs as modular and dynamic units for gene regulation by hormones. *FEBS Lett* *589*, 2885-2892.
- Le Dily, F., and Beato, M. (2018). Signaling by Steroid Hormones in the 3D Nuclear Space. *Int J Mol Sci* *19*, 306-321.
- Le-Niculescu, H., Roseberry, K., Levey, D.F., Rogers, J., Kosary, K., Prabha, S., Jones, T., Judd, S., McCormick, M.A., Wessel, A.R., *et al.* (2019). Towards precision medicine for stress disorders: diagnostic biomarkers and targeted drugs. *Mol Psychiatry*.
- Lee, R.S., Mahon, P.B., Zandi, P.P., McCaul, M.E., Yang, X., Bali, U., and Wand, G.S. (2018). DNA methylation and sex-specific expression of FKBP5 as correlates of one-month bedtime cortisol levels in healthy individuals. *Psychoneuroendocrinology* *97*, 164-173.
- Lee, R.S., Tamashiro, K.L., Yang, X., Purcell, R.H., Harvey, A., Willour, V.L., Huo, Y., Rongione, M., Wand, G.S., and Potash, J.B. (2010). Chronic corticosterone exposure increases expression and decreases deoxyribonucleic acid methylation of Fkbp5 in mice. *Endocrinology* *151*, 4332-4343.
- Lee, R.S., Tamashiro, K.L., Yang, X., Purcell, R.H., Huo, Y., Rongione, M., Potash, J.B., and Wand, G.S. (2011). A measure of glucocorticoid load provided by DNA methylation of Fkbp5 in mice. *Psychopharmacology (Berl)* *218*, 303-312.
- Lessard, J., and Holman, E.A. (2014). FKBP5 and CRHR1 polymorphisms moderate the stress-physical health association in a national sample. *Health Psychol* *33*, 1046-1056.
- Lettice, L.A., Heaney, S.J., Purdie, L.A., Li, L., de Beer, P., Oostra, B.A., Goode, D., Elgar, G., Hill, R.E., and de Graaff, E. (2003). A long-range Shh enhancer regulates

References

expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum Mol Genet* **12**, 1725-1735.

Levenson, J.M., and Sweatt, J.D. (2005). Epigenetic mechanisms in memory formation. *Nat Rev Neurosci* **6**, 108-118.

Li, R., Liu, Y., Hou, Y., Gan, J., Wu, P., and Li, C. (2018). 3D genome and its disorganization in diseases. *Cell Biol Toxicol* **34**, 351-365.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragozy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., *et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289-293.

Liu, X.S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czauderna, S., Shu, J., Dadon, D., Young, R.A., and Jaenisch, R. (2016). Editing DNA Methylation in the Mammalian Genome. *Cell* **167**, 233-247 e217.

Liu, X.S., Wu, H., Krzisch, M., Wu, X., Graef, J., Muffat, J., Hnisz, D., Li, C.H., Yuan, B., Xu, C., *et al.* (2018). Rescue of Fragile X Syndrome Neurons by DNA Methylation Editing of the FMR1 Gene. *Cell* **172**, 979-992 e976.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550.

Lower, K.M., Hughes, J.R., De Gobbi, M., Henderson, S., Viprakasit, V., Fisher, C., Goriely, A., Ayyub, H., Sloane-Stanley, J., Vernimmen, D., *et al.* (2009). Adventitious changes in long-range gene expression caused by polymorphic structural variation and promoter competition. *Proc Natl Acad Sci U S A* **106**, 21771-21776.

Luo, C., Hajkova, P., and Ecker, J.R. (2018). Dynamic DNA methylation: In the right place at the right time. *Science* **361**, 1336-1340.

Lupianez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., *et al.* (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012-1025.

Ma, W., Ay, F., Lee, C., Gulsoy, G., Deng, X., Cook, S., Hesson, J., Cavanaugh, C., Ware, C.B., Krumm, A., *et al.* (2015). Fine-scale chromatin interaction maps reveal the cis-regulatory landscape of human lincRNA genes. *Nat Methods* **12**, 71-78.

Maddox, S.A., Schafe, G.E., and Ressler, K.J. (2013). Exploring epigenetic regulation of fear memory and biomarkers associated with post-traumatic stress disorder. *Front Psychiatry* **4**, 62.

Mansour, M.R., Abraham, B.J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A.D., Echin, J., Lawton, L., Sallan, S.E., Silverman, L.B., *et al.* (2014). Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science* **346**, 1373-1377.

References

- Masser, D.R., Berg, A.S., and Freeman, W.M. (2013). Focused, high accuracy 5-methylcytosine quantitation with base resolution by benchtop next-generation sequencing. *Epigenetics Chromatin* 6, 33.
- Masser, D.R., Hadad, N., Porter, H., Stout, M.B., Unnikrishnan, A., Stanford, D.R., and Freeman, W.M. (2018). Analysis of DNA modifications in aging research. *Geroscience* 40, 11-29.
- Masser, D.R., Stanford, D.R., and Freeman, W.M. (2015). Targeted DNA methylation analysis by next-generation sequencing. *J Vis Exp*.
- Matosin, N., Halldorsdottir, T., and Binder, E.B. (2018). Understanding the Molecular Mechanisms Underpinning Gene by Environment Interactions in Psychiatric Disorders: The FKBP5 Model. *Biol Psychiatry* 83, 821-830.
- McGowan, P.O., Sasaki, A., D'Alessio, A.C., Dymov, S., Labonte, B., Szyf, M., Turecki, G., and Meaney, M.J. (2009). Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nat Neurosci* 12, 342-348.
- Menke, A., Arloth, J., Putz, B., Weber, P., Klengel, T., Mehta, D., Gonik, M., Rex-Haffner, M., Rubel, J., Uhr, M., *et al.* (2012). Dexamethasone stimulated gene expression in peripheral blood is a sensitive marker for glucocorticoid receptor resistance in depressed patients. *Neuropsychopharmacology* 37, 1455-1464.
- Metivier, R., Gallais, R., Tiffoche, C., Le Peron, C., Jurkowska, R.Z., Carmouche, R.P., Ibberson, D., Barath, P., Demay, F., Reid, G., *et al.* (2008). Cyclical DNA methylation of a transcriptionally active promoter. *Nature* 452, 45-50.
- Mikhaylichenko, O., Bondarenko, V., Harnett, D., Schor, I.E., Males, M., Viales, R.R., and Furlong, E.E.M. (2018). The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. *Genes Dev* 32, 42-57.
- Mill, J., and Petronis, A. (2007). Molecular studies of major depressive disorder: the epigenetic perspective. *Mol Psychiatry* 12, 799-814.
- Mill, J., Tang, T., Kaminsky, Z., Khare, T., Yazdanpanah, S., Bouchard, L., Jia, P., Assadzadeh, A., Flanagan, J., Schumacher, A., *et al.* (2008). Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. *Am J Hum Genet* 82, 696-711.
- Miller, C.A., Campbell, S.L., and Sweatt, J.D. (2008). DNA methylation and histone acetylation work in concert to regulate memory formation and synaptic plasticity. *Neurobiol Learn Mem* 89, 599-603.
- Mitchell, A.C., Bharadwaj, R., Whittle, C., Krueger, W., Mirnics, K., Hurd, Y., Rasmussen, T., and Akbarian, S. (2014). The genome in three dimensions: a new frontier in human brain research. *Biol Psychiatry* 75, 961-969.
- Molnar, B.E., Buka, S.L., and Kessler, R.C. (2001). Child sexual abuse and subsequent psychopathology: results from the National Comorbidity Survey. *Am J Public Health* 91, 753-760.

References

- Murgatroyd, C., and Spengler, D. (2014). Polycomb binding precedes early-life stress responsive DNA methylation at the Avp enhancer. *PLoS One* *9*, e90277.
- Nagano, T., Lubling, Y., Yaffe, E., Wingett, S.W., Dean, W., Tanay, A., and Fraser, P. (2015). Single-cell Hi-C for genome-wide detection of chromatin interactions that occur simultaneously in a single cell. *Nat Protoc* *10*, 1986-2003.
- Nagano, T., Wingett, S.W., and Fraser, P. (2017). Capturing Three-Dimensional Genome Organization in Individual Cells by Single-Cell Hi-C. *Methods Mol Biol* *1654*, 79-97.
- Nair, S.C., Rimerman, R.A., Toran, E.J., Chen, S., Prapapanich, V., Butts, R.N., and Smith, D.F. (1997). Molecular cloning of human FKBP51 and comparisons of immunophilin interactions with Hsp90 and progesterone receptor. *Mol Cell Biol* *17*, 594-603.
- Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* *393*, 386-389.
- Nicolaidis, N.C., Charmandari, E., Chrousos, G.P., and Kino, T. (2014). Recent advances in the molecular mechanisms determining tissue sensitivity to glucocorticoids: novel mutations, circadian rhythm and ligand-induced repression of the human glucocorticoid receptor. *BMC Endocr Disord* *14*, 71.
- Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., *et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* *485*, 381-385.
- O'Hagan, H.M., Wang, W., Sen, S., Destefano Shields, C., Lee, S.S., Zhang, Y.W., Clements, E.G., Cai, Y., Van Neste, L., Easwaran, H., *et al.* (2011). Oxidative damage targets complexes containing DNA methyltransferases, SIRT1, and polycomb members to promoter CpG Islands. *Cancer Cell* *20*, 606-619.
- Paakinaho, V., Makkonen, H., Jaaskelainen, T., and Palvimo, J.J. (2010). Glucocorticoid receptor activates poised FKBP51 locus through long-distance interactions. *Mol Endocrinol* *24*, 511-525.
- Parada, L.A., Roix, J.J., and Misteli, T. (2003). An uncertainty principle in chromosome positioning. *Trends Cell Biol* *13*, 393-396.
- Pelleymounter, L.L., Moon, I., Johnson, J.A., Laederach, A., Halvorsen, M., Eckloff, B., Abo, R., and Rossetti, S. (2011). A novel application of pattern recognition for accurate SNP and indel discovery from high-throughput data: targeted resequencing of the glucocorticoid receptor co-chaperone FKBP5 in a Caucasian population. *Mol Genet Metab* *104*, 457-469.
- Pereira, M.J., Palming, J., Svensson, M.K., Rizell, M., Dalenback, J., Hammar, M., Fall, T., Sidibeh, C.O., Svensson, P.A., and Eriksson, J.W. (2014). FKBP5 expression in human adipose tissue increases following dexamethasone exposure and is associated with insulin resistance. *Metabolism* *63*, 1198-1208.

References

- Perillo, B., Ombra, M.N., Bertoni, A., Cuozzo, C., Sacchetti, S., Sasso, A., Chiariotti, L., Malorni, A., Abbondanza, C., and Avvedimento, E.V. (2008). DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression. *Science* 319, 202-206.
- Petronis, A. (2006). Epigenetics and twins: three variations on the theme. *Trends Genet* 22, 347-350.
- Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. *Cell* 137, 1194-1211.
- Phillips-Cremins, J.E., and Corces, V.G. (2013). Chromatin insulators: linking genome organization to cellular function. *Mol Cell* 50, 461-474.
- Polanczyk, G., Caspi, A., Williams, B., Price, T.S., Danese, A., Sugden, K., Uher, R., Poulton, R., and Moffitt, T.E. (2009). Protective effect of CRHR1 gene variants on the development of adult depression following childhood maltreatment: replication and extension. *Arch Gen Psychiatry* 66, 978-985.
- Pratt, W.B., and Toft, D.O. (1997). Steroid receptor interactions with heat shock protein and immunophilin chaperones. *Endocr Rev* 18, 306-360.
- Prober, J.M., Trainor, G.L., Dam, R.J., Hobbs, F.W., Robertson, C.W., Zagursky, R.J., Cocuzza, A.J., Jensen, M.A., and Baumeister, K. (1987). A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* 238, 336-341.
- Quintin, J., Le Peron, C., Palierne, G., Bizot, M., Cunha, S., Serandour, A.A., Avner, S., Henry, C., Percevault, F., Belaud-Rotureau, M.A., *et al.* (2014). Dynamic estrogen receptor interactomes control estrogen-responsive trefoil Factor (TFF) locus cell-specific activities. *Mol Cell Biol* 34, 2418-2436.
- Ramani, V., Deng, X., Qiu, R., Gunderson, K.L., Steemers, F.J., Disteche, C.M., Noble, W.S., Duan, Z., and Shendure, J. (2017). Massively multiplex single-cell Hi-C. *Nat Methods* 14, 263-266.
- Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., *et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665-1680.
- Raviram, R., Rocha, P.P., Muller, C.L., Miraldi, E.R., Badri, S., Fu, Y., Swanzey, E., Proudhon, C., Snetkova, V., Bonneau, R., *et al.* (2016). 4C-ker: A Method to Reproducibly Identify Genome-Wide Interactions Captured by 4C-Seq Experiments. *PLoS Comput Biol* 12, e1004780.
- Reddy, T.E., Pauli, F., Sprouse, R.O., Neff, N.F., Newberry, K.M., Garabedian, M.J., and Myers, R.M. (2009). Genomic determination of the glucocorticoid response reveals unexpected mechanisms of gene regulation. *Genome Res* 19, 2163-2171.
- Reizel, Y., Spiro, A., Sabag, O., Skversky, Y., Hecht, M., Keshet, I., Berman, B.P., and Cedar, H. (2015). Gender-specific postnatal demethylation and establishment of epigenetic memory. *Genes Dev* 29, 923-933.

References

- Robertson, K.D. (2005). DNA methylation and human disease. *Nat Rev Genet* 6, 597-610.
- Robertson, K.D., and Wolffe, A.P. (2000). DNA methylation in health and disease. *Nat Rev Genet* 1, 11-19.
- Rosa-Garrido, M., Chapski, D.J., Schmitt, A.D., Kimball, T.H., Karbassi, E., Monte, E., Balderas, E., Pellegrini, M., Shih, T.T., Soehalim, E., *et al.* (2017). High-Resolution Mapping of Chromatin Conformation in Cardiac Myocytes Reveals Structural Remodeling of the Epigenome in Heart Failure. *Circulation* 136, 1613-1625.
- Roth, T.L., Lubin, F.D., Funk, A.J., and Sweatt, J.D. (2009). Lasting epigenetic influence of early-life adversity on the BDNF gene. *Biol Psychiatry* 65, 760-769.
- Roy, A., Gorodetsky, E., Yuan, Q., Goldman, D., and Enoch, M.A. (2010). Interaction of FKBP5, a stress-related gene, with childhood trauma increases the risk for attempting suicide. *Neuropsychopharmacology* 35, 1674-1683.
- Roy, A., Hodgkinson, C.A., Deluca, V., Goldman, D., and Enoch, M.A. (2012). Two HPA axis genes, CRHBP and FKBP5, interact with childhood trauma to increase the risk for suicidal behavior. *J Psychiatr Res* 46, 72-79.
- Russell, G.M., Henley, D.E., Leendertz, J., Douthwaite, J.A., Wood, S.A., Stevens, A., Woltersdorf, W.W., Peeters, B.W., Ruigt, G.S., White, A., *et al.* (2010). Rapid glucocorticoid receptor-mediated inhibition of hypothalamic-pituitary-adrenal ultradian activity in healthy males. *J Neurosci* 30, 6106-6115.
- Rust, M.J., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat Methods* 3, 793-795.
- Ryba, T., Hiratani, I., Lu, J., Itoh, M., Kulik, M., Zhang, J., Schulz, T.C., Robins, A.J., Dalton, S., and Gilbert, D.M. (2010). Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res* 20, 761-770.
- Sajan, S.A., and Hawkins, R.D. (2012). Methods for identifying higher-order chromatin structure. *Annu Rev Genomics Hum Genet* 13, 59-82.
- Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., *et al.* (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc Natl Acad Sci U S A* 112, E6456-6465.
- Sanyal, A., Lajoie, B.R., Jain, G., and Dekker, J. (2012). The long-range interaction landscape of gene promoters. *Nature* 489, 109-113.
- Sapolsky, R.M., Romero, L.M., and Munck, A.U. (2000). How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocr Rev* 21, 55-89.
- Sati, S., and Cavalli, G. (2017). Chromosome conformation capture technologies and their impact in understanding genome function. *Chromosoma* 126, 33-44.

References

- Schanen, N.C. (2006). Epigenetics of autism spectrum disorders. *Hum Mol Genet* 15 *Spec No 2*, R138-150.
- Scharf, S.H., Liebl, C., Binder, E.B., Schmidt, M.V., and Muller, M.B. (2011). Expression and regulation of the Fkbp5 gene in the adult mouse brain. *PLoS One* 6, e16883.
- Schatz, M.C. (2017). Nanopore sequencing meets epigenetics. *Nat Methods* 14, 347-348.
- Schmitt, A.D., Hu, M., Jung, I., Xu, Z., Qiu, Y., Tan, C.L., Li, Y., Lin, S., Lin, Y., Barr, C.L., *et al.* (2016). A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell Rep* 17, 2042-2059.
- Schuster-Bockler, B., and Lehner, B. (2012). Chromatin organization is a major influence on regional mutation rates in human cancer cells. *Nature* 488, 504-507.
- Seifuddin, F., Wand, G., Cox, O., Pirooznia, M., Moody, L., Yang, X., Tai, J., Boersma, G., Tamashiro, K., Zandi, P., *et al.* (2017). Genome-wide Methyl-Seq analysis of blood-brain targets of glucocorticoid exposure. *Epigenetics*, 0.
- Seitan, V.C., Faure, A.J., Zhan, Y., McCord, R.P., Lajoie, B.R., Ing-Simmons, E., Lenhard, B., Giorgetti, L., Heard, E., Fisher, A.G., *et al.* (2013). Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome Res* 23, 2066-2077.
- Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* 148, 458-472.
- Shin, J., Ming, G.L., and Song, H. (2014). DNA modifications in the mammalian brain. *Philos Trans R Soc Lond B Biol Sci* 369.
- Simonis, M., Kooren, J., and de Laat, W. (2007). An evaluation of 3C-based methods to capture DNA interactions. *Nat Methods* 4, 895-901.
- Spielmann, M., Lupianez, D.G., and Mundlos, S. (2018). Structural variation in the 3D genome. *Nat Rev Genet* 19, 453-467.
- Stephan, K.E., Bach, D.R., Fletcher, P.C., Flint, J., Frank, M.J., Friston, K.J., Heinz, A., Huys, Q.J.M., Owen, M.J., Binder, E.B., *et al.* (2016). Charting the landscape of priority problems in psychiatry, part 1: classification and diagnosis. *Lancet Psychiatry* 3, 77-83.
- Stevens, T.J., Lando, D., Basu, S., Atkinson, L.P., Cao, Y., Lee, S.F., Leeb, M., Wohlfahrt, K.J., Boucher, W., O'Shaughnessy-Kirwan, A., *et al.* (2017). 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* 544, 59-64.
- Sun, Z., Cunningham, J., Slager, S., and Kocher, J.P. (2015). Base resolution methylome profiling: considerations in platform selection, data preprocessing and analysis. *Epigenomics* 7, 813-828.

References

- Swartz, M.N., Trautner, T.A., and Kornberg, A. (1962). Enzymatic synthesis of deoxyribonucleic acid. XI. Further studies on nearest neighbor base sequences in deoxyribonucleic acids. *J Biol Chem* 237, 1961-1967.
- Symmons, O., Pan, L., Remeseiro, S., Aktas, T., Klein, F., Huber, W., and Spitz, F. (2016). The Shh Topological Domain Facilitates the Action of Remote Enhancers by Reducing the Effects of Genomic Distances. *Dev Cell* 39, 529-543.
- Symmons, O., Uslu, V.V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., Ettwiller, L., and Spitz, F. (2014). Functional and topological characteristics of mammalian regulatory domains. *Genome Res* 24, 390-400.
- Szyf, M. (2013). DNA methylation, behavior and early life adversity. *J Genet Genomics* 40, 331-338.
- Takizawa, T., Meaburn, K.J., and Misteli, T. (2008). The meaning of gene positioning. *Cell* 135, 9-13.
- Tang, Z., Luo, O.J., Li, X., Zheng, M., Zhu, J.J., Szalaj, P., Trzaskoma, P., Magalska, A., Wlodarczyk, J., Ruszczycski, B., *et al.* (2015). CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. *Cell* 163, 1611-1627.
- Thomassin, H., Flavin, M., Espinas, M.L., and Grange, T. (2001). Glucocorticoid-induced DNA demethylation and gene memory during development. *EMBO J* 20, 1974-1983.
- Tost, J. (2010). DNA methylation: an introduction to the biology and the disease-associated changes of a promising biomarker. *Mol Biotechnol* 44, 71-81.
- Tost, J., and Gut, I.G. (2007). DNA methylation analysis by pyrosequencing. *Nat Protoc* 2, 2265-2275.
- Tsai, H.C., Li, H., Van Neste, L., Cai, Y., Robert, C., Rassool, F.V., Shin, J.J., Harbom, K.M., Beaty, R., Pappou, E., *et al.* (2012). Transient low doses of DNA-demethylating agents exert durable antitumor effects on hematological and epithelial tumor cells. *Cancer Cell* 21, 430-446.
- Tsankova, N., Renthal, W., Kumar, A., and Nestler, E.J. (2007). Epigenetic regulation in psychiatric disorders. *Nat Rev Neurosci* 8, 355-367.
- Uhlenhaut, N.H., Barish, G.D., Yu, R.T., Downes, M., Karunasiri, M., Liddle, C., Schwalie, P., Hubner, N., and Evans, R.M. (2013). Insights into negative regulation by the glucocorticoid receptor from genome-wide profiling of inflammatory cistromes. *Mol Cell* 49, 158-171.
- Uhler, C., and Shivashankar, G.V. (2017). Regulation of genome organization and gene expression by nuclear mechanotransduction. *Nat Rev Mol Cell Biol* 18, 717-727.
- Van Soom, A., Peelman, L., Holt, W.V., and Fazeli, A. (2014). An introduction to epigenetics as the link between genotype and environment: a personal view. *Reprod Domest Anim* 49 *Suppl* 3, 2-10.

References

- van Steensel, B., and Dekker, J. (2010). Genomics tools for unraveling chromosome architecture. *Nat Biotechnol* 28, 1089-1095.
- VanZomeren-Dohm, A.A., Pitula, C.E., Koss, K.J., Thomas, K., and Gunnar, M.R. (2015). FKBP5 moderation of depressive symptoms in peer victimized, post-institutionalized children. *Psychoneuroendocrinology* 51, 426-430.
- Varambally, S., Cao, Q., Mani, R.S., Shankar, S., Wang, X., Ateeq, B., Laxman, B., Cao, X., Jing, X., Ramnarayanan, K., *et al.* (2008). Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. *Science* 322, 1695-1699.
- Vermeer, H., Hendriks-Stegeman, B.I., van der Burg, B., van Buul-Offers, S.C., and Jansen, M. (2003). Glucocorticoid-induced increase in lymphocytic FKBP51 messenger ribonucleic acid expression: a potential marker for glucocorticoid sensitivity, potency, and bioavailability. *J Clin Endocrinol Metab* 88, 277-284.
- Victoria-Acosta, G., Vazquez-Santillan, K., Jimenez-Hernandez, L., Munoz-Galindo, L., Maldonado, V., Martinez-Ruiz, G.U., and Melendez-Zajgla, J. (2015). Epigenetic silencing of the XAF1 gene is mediated by the loss of CTCF binding. *Sci Rep* 5, 14838.
- Vidaki, A., and Kayser, M. (2018). Recent progress, methods and perspectives in forensic epigenetics. *Forensic Sci Int Genet* 37, 180-195.
- Visscher, P.M., Wray, N.R., Zhang, Q., Sklar, P., McCarthy, M.I., Brown, M.A., and Yang, J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet* 101, 5-22.
- Wang, C., Liu, C., Roqueiro, D., Grimm, D., Schwab, R., Becker, C., Lanz, C., and Weigel, D. (2015). Genome-wide analysis of local chromatin packing in *Arabidopsis thaliana*. *Genome Res* 25, 246-256.
- Wang, D., Liu, S., Warrell, J., Won, H., Shi, X., Navarro, F.C.P., Clarke, D., Gu, M., Emani, P., Yang, Y.T., *et al.* (2018a). Comprehensive functional genomic resource and integrative model for the human brain. *Science* 362.
- Wang, H., Xu, X., Nguyen, C.M., Liu, Y., Gao, Y., Lin, X., Daley, T., Kipniss, N.H., La Russa, M., and Qi, L.S. (2018b). CRISPR-Mediated Programmable 3D Genome Positioning and Nuclear Organization. *Cell* 175, 1405-1417 e1414.
- Wang, R.Y., Gehrke, C.W., and Ehrlich, M. (1980). Comparison of bisulfite modification of 5-methyldeoxycytidine and deoxycytidine residues. *Nucleic Acids Res* 8, 4777-4790.
- Wang, S., Su, J.H., Beliveau, B.J., Bintu, B., Moffitt, J.R., Wu, C.T., and Zhuang, X. (2016). Spatial organization of chromatin domains and compartments in single chromosomes. *Science* 353, 598-602.
- Wapinski, O.L., Vierbuchen, T., Qu, K., Lee, Q.Y., Chanda, S., Fuentes, D.R., Giresi, P.G., Ng, Y.H., Marro, S., Neff, N.F., *et al.* (2013). Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. *Cell* 155, 621-635.

References

- Waterland, R.A., and Jirtle, R.L. (2003). Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol Cell Biol* 23, 5293-5300.
- Waters, T.R., and Swann, P.F. (1998). Kinetics of the action of thymine DNA glycosylase. *J Biol Chem* 273, 20007-20014.
- Weaver, I.C., Diorio, J., Seckl, J.R., Szyf, M., and Meaney, M.J. (2004). Early environmental regulation of hippocampal glucocorticoid receptor gene expression: characterization of intracellular mediators and potential genomic target sites. *Ann N Y Acad Sci* 1024, 182-212.
- Wiench, M., John, S., Baek, S., Johnson, T.A., Sung, M.H., Escobar, T., Simmons, C.A., Pearce, K.H., Biddie, S.C., Sabo, P.J., *et al.* (2011). DNA methylation status predicts cell type-specific enhancer activity. *EMBO J* 30, 3028-3039.
- Williamson, I., Berlivet, S., Eskeland, R., Boyle, S., Illingworth, R.S., Paquette, D., Dostie, J., and Bickmore, W.A. (2014). Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence in situ hybridization. *Genes Dev* 28, 2778-2791.
- Winner, J.G., Carhart, J.M., Altar, C.A., Allen, J.D., and Dechairo, B.M. (2013). A prospective, randomized, double-blind study assessing the clinical impact of integrated pharmacogenomic testing for major depressive disorder. *Discov Med* 16, 219-227.
- Wu, H., and Zhang, Y. (2014). Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell* 156, 45-68.
- Xie, P., Kranzler, H.R., Poling, J., Stein, M.B., Anton, R.F., Farrer, L.A., and Gelernter, J. (2010). Interaction of FKBP5 with childhood adversity on risk for post-traumatic stress disorder. *Neuropsychopharmacology* 35, 1684-1692.
- Yehuda, R., Daskalakis, N.P., Bierer, L.M., Bader, H.N., Klengel, T., Holsboer, F., and Binder, E.B. (2016). Holocaust Exposure Induced Intergenerational Effects on FKBP5 Methylation. *Biol Psychiatry* 80, 372-380.
- Yehuda, R., Daskalakis, N.P., Desarnaud, F., Makotkine, I., Lehrner, A.L., Koch, E., Flory, J.D., Buxbaum, J.D., Meaney, M.J., and Bierer, L.M. (2013). Epigenetic Biomarkers as Predictors and Correlates of Symptom Improvement Following Psychotherapy in Combat Veterans with PTSD. *Front Psychiatry* 4, 118.
- Yehuda, R., and LeDoux, J. (2007). Response variation following trauma: a translational neuroscience approach to understanding PTSD. *Neuron* 56, 19-32.
- Yu, M., and Ren, B. (2017). The Three-Dimensional Organization of Mammalian Genomes. *Annu Rev Cell Dev Biol* 33, 265-289.
- Zannas, A.S., and Binder, E.B. (2014). Gene-environment interactions at the FKBP5 locus: sensitive periods, mechanisms and pleiotropism. *Genes Brain Behav* 13, 25-37.

References

- Zannas, A.S., Wiechmann, T., Gassen, N.C., and Binder, E.B. (2016). Gene-Stress-Epigenetic Regulation of FKBP5: Clinical and Translational Implications. *Neuropsychopharmacology* 41, 261-274.
- Zhan, Y., Mariani, L., Barozzi, I., Schulz, E.G., Bluthgen, N., Stadler, M., Tiana, G., and Giorgetti, L. (2017). Reciprocal insulation analysis of Hi-C data shows that TADs represent a functionally but not structurally privileged scale in the hierarchical folding of chromosomes. *Genome Res* 27, 479-490.
- Zhang, T.Y., Labonte, B., Wen, X.L., Turecki, G., and Meaney, M.J. (2013). Epigenetic mechanisms for the early environmental regulation of hippocampal glucocorticoid receptor gene expression in rodents and humans. *Neuropsychopharmacology* 38, 111-123.
- Zhang, X., Choi, P.S., Francis, J.M., Imielinski, M., Watanabe, H., Cherniack, A.D., and Meyerson, M. (2016). Identification of focally amplified lineage-specific super-enhancers in human epithelial cancers. *Nat Genet* 48, 176-182.
- Ziller, M.J., Hansen, K.D., Meissner, A., and Aryee, M.J. (2015). Coverage recommendations for methylation analysis by whole-genome bisulfite sequencing. *Nat Methods* 12, 230-232, 231 p following 232.
- Zimmermann, P., Bruckl, T., Nocon, A., Pfister, H., Binder, E.B., Uhr, M., Lieb, R., Moffitt, T.E., Caspi, A., Holsboer, F., *et al.* (2011). Interaction of FKBP5 gene variants and adverse life events in predicting depression onset: results from a 10-year prospective community study. *Am J Psychiatry* 168, 1107-1116.
- Zou, Y.F., Wang, F., Feng, X.L., Li, W.F., Tao, J.H., Pan, F.M., Huang, F., and Su, H. (2010). Meta-analysis of FKBP5 gene polymorphisms association with treatment response in patients with mood disorders. *Neurosci Lett* 484, 56-61.

8 Declaration of Contributions

Paper I: Roeh, S.*, **T. Wiechmann***, S. Sauer, M. Kodel, E. B. Binder and N. Provençal (2018). "HAM-TBS: high-accuracy methylation measurements via targeted bisulfite sequencing." *Epigenetics Chromatin* **11**(1): 39.

(* shared first authors)

TW, SR, NP, EBB contributed to experimental design. TW, SS, MK performed wet lab work. SR performed the data analyses. SR, TW, NP, EBB prepared the manuscript.

Paper II: Piyasena, C., J. Cartier, N. Provençal, **T. Wiechmann**, B. Khulan, R. Sunderasan, G. Menon, J. R. Seckl, R. M. Reynolds, E. B. Binder and A. J. Drake (2016). "Dynamic Changes in DNA Methylation Occur during the First Year of Life in Preterm Infants." *Front Endocrinol (Lausanne)* **7**: 158.

CP and AD conceived the study. CP, GM, JS, and AD designed the study. CP, JC, BK, RS, NP, TW, RR, EB, and AD performed the study and analyses. CP and AD wrote the initial manuscript draft and all authors revised it critically for intellectual content.

Paper III: Wiechmann T., S. Roeh, S. Sauer, D. Czamara, J. Arloth, M. Ködel, M. Beintner, L. Knop, A. Menke, E. B. Binder and N. Provençal (2018) Identification of dynamic glucocorticoid- induced methylation changes at the *FKBP5* locus. *Submitted to Clinical Epigenetics in Nov. 2018*

TW, SS, MK, MB performed wet lab work. SR performed sequencing data processing. JA performed processing of 450K data. TW, DC and NP analyzed and illustrated the data. MB, LK and AM collected and processed study samples. TW, NP, EBB prepared the manuscript. NP, EBB conceptualized and supervised the study.

Manuscript I: Wiechmann T., S. Roeh, D. Czamara, T. Klengel, S. Sauer, T. Rein, N. Provençal, E. B. Binder (2019). An intronic structural variant in *FKBP5* affects GRE enhancer activity through the modulation of topological associated domain interaction frequencies and regulates the responsiveness to glucocorticoid exposure. *Ahead of submission*

TW designed and performed the wetlab experiments. TK designed the genotyping qPCR assays for esv3608688. SS performed optimization and validation experiments of allele specific gene expression assays. SR performed the processing of the sequencing data. TW, DC performed the data analysis. TW & EBB wrote the manuscript. NP, TR & EBB supervised the project.

München, 2019

.....
(Tobias Wiechmann)

Declaration of contributions

Hiermit bestätige ich die von Herrn Wiechmann angegeben Beiträge zu den einzelnen Publikationen.

München, 2019

.....
(PD Dr. Mathias V. Schmidt)

9 Acknowledgements

This thesis is the result of a challenging but pleasant four-year period, in which I had the opportunity to address some very interesting questions in a fascinating field of research. This thesis in its current form could have never been compiled without a group of people who supported me in many ways. Therefore, I would like to acknowledge this support and express my gratitude to my colleagues, friends and family.

First, I would like to thank Elisabeth Binder for being an extraordinary supervisor and allowing me to work on great projects with a lot of freedom to execute them. Thank you for your supervision, trust and great environment to work in.

I would like to thank Mathias V. Schmidt for his support and guidance as my official thesis advisor. I highly enjoyed our motivating discussions. I want to thank Prof. Dr. Heinrich Leonhardt for his willingness and commitment as a second examiner. Moreover, I appreciate the time and effort of Prof. Dr. Niels Dingemans, Prof. Dr. Thomas Nägele, Prof. Dr. Christof Osman and PD Dr. Carsten T. Wotjak for the evaluation of this thesis. In addition, I would like to acknowledge the continuous support of my thesis advisory committee (Elisabeth Binder, Mathias V. Schmidt, Andreas Ladurner, Theo Rein). I highly value your support and enjoyed our yearly meetings.

I am grateful to all colleagues of the Binder lab and the department of translational research for their invaluable support and the good atmosphere. I would like to thank Nadine for her support and helpful discussions and Simone for providing her bioinformatic expertise throughout the projects of this thesis. I cannot stress enough how grateful and lucky I am for the technical support of Monika, Susi, Maik, Laura and Anna – thank you so much for the fun time in the lab. Britta, thank you for our great time and discussions while commuting to the institute and proofreading of this thesis.

Last but not least I would like to thank my family. I sincerely thank my parents for always believing in me and their unconditional support. I am especially grateful to my wife Svenja for her day-to-day support, endurance to bear with me during this time and proofreading.

10 Statutory declaration and statement

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt ist.

München, den**18.Juni 2019**.....

(Tobias Wiechmann)

Erklärung

Hiermit erkläre ich,

dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist.

dass ich mich anderweitig einer Doktorprüfung ohne Erfolg **nicht** unterzogen habe.

München, den.....**18.Juni 2019**.....

(Tobias Wiechmann)