DNA methylation and microRNA-mediated regulation of gene expression in adolescent Chronic Fatigue Syndrome

Kristine Løkås Jacobsen

Thesis submitted for the degree of
Master of Science in Molecular Bioscience
60 credits

Department of Biosciences, Faculty of Mathematics and Natural Sciences, University of Oslo, Norway

Department of Pediatrics and Adolescent Health,
Akershus University Hospital, Lørenskog, Norway

February 2018

2018

DNA methylation and microRNA-mediated regulation
of gene expression in adolescent Chronic Fatigue Syndrome

Author: Kristine Løkås Jacobsen

http://www.duo.uio.no/

Press: Reprosentralen, Universitetet i Oslo

Acknowledgements

First and foremost, I would like to thank my main supervisor Vegard Bruun Bratholm Wyller for giving me the opportunity to work in this interesting field. Thank you for giving me great guidance and for taking my thoughts and opinions about this project into serious consideration.

Secondly, I would like to thank my co-supervisor Hilde Nilsen for her thoughts and guidance on the molecular background of this project, and for giving great feedback on my writing and interpretation of the data.

A special thanks goes to my lab-supervisor Mai Britt Dahl for excellent laboratory training, for becoming a great friend, and for keeping my spirits up when nothing really went as it should. I also want to thank Chinh Bkrong Nguyen for helping me with some of the lab-work and for giving her opinions on my work.

To my co-supervisor Johannes Gjerstad, I would like to say thank you for seeing my potential and for putting me in contact with Vegard and his research group, as well as helping me with structuring my thesis.

I also would like to thank everyone in the PAEDIA research group at the department of Pediatrics and Adolescent Health at Akershus University Hospital, which I have been so lucky to be a part of during this project. Their support and feedback on my work has been invaluable.

Next, I would like to thank everyone at the EpiGen laboratory at Akershus University Hospital who has helped me during my search of the right methods to use on my data set and let me be part of a great work environment.

Finally, I want to thank my friends and family, who have supported me through the good times and challenges during the period I have worked on this project. A special thanks goes to my husband Trond, who always believes in me and brings out the best in me.

Kristine Løkås Jacobsen,
Jessheim, January 2018

Abstract

Chronic fatigue syndrome (CFS) is a disabling disease of which the pathophysiology is poorly understood. CFS patients have previously been found to differ significantly from healthy subjects in the level of expression of several genes, but little is known about the molecular mechanisms giving rise to these differences. DNA methylation and regulation mediated by microRNAs (miRNAs) are two mechanisms that may affect gene expression. In the present study, DNA methylation and miRNA-mediated regulation in adolescent CFS were investigated using a candidate-gene approach.

Expression levels of miR-223-3p, miR-142-5p, miR-143-3p, miR-486-3p and miR-146a-5p were analyzed by reverse transcription quantitative polymerase chain reaction (RT-qPCR). DNA methylation levels were analyzed by magnetic methylated DNA immunoprecipitation (MagMeDIP), followed by qPCR analysis of *IRF7* and *NR3C1-1A* promoter regions.

There was found no significant differences between CFS patients and healthy controls in the expression levels of any of the miRNAs investigated. Significant hypermethylation of the *IRF7* promoter was detected in CFS patients. There was found a non-significant negative correlation between *IRF7* methylation and steps per day, as well as a non-significant positive correlation between *IRF7* methylation and Chalder Fatigue score (CFQ). No significant differences were found between CFS patients and healthy controls in DNA methylation levels of *NR3C1* promoter 1A, but correlation analyses showed a significant negative relationship between *NR3C1* promoter methylation and steps per day.

In conclusion, there was found no evidence supporting the hypothesis that miRNA mediated regulation of gene expression is part of the pathophysiology of adolescent CFS. Hypermethylation of the *IRF7* promoter region was found in CFS patients, and although the correlation between *IRF7* methylation levels and CFQ was not statistically significant, it indicates a link between *IRF7* promoter hypermethylation and fatigue severity. Taken together, this suggests a role for *IRF7* promoter hypermethylation in CFS.

Table of contents

1 Introduction

1.1 Chronic Fatigue Syndrome

1.1.1 The condition

Chronic Fatigue Syndrome (CFS), also known as Myalgic Encephalomyelitis (ME), is a highly disabling condition characterized by unexplained, long-lasting fatigue accompanied by symptoms such as pain, cognitive impairments and exertion intolerance [21; 96]. The underlying pathophysiology of the condition is unknown, and research on disease mechanisms has been conducted along several tracks. As there are no biological markers currently unique for CFS, diagnosis has largely been done by excluding all other diseases known to provoke similar symptoms. Several case definitions exist, but diagnostic criteria are heavily debated [21]. The disease is relatively common in developed countries, with an estimated prevalence varying from 0,1 to 2,5%, depending on the criteria applied [79]. The condition affects more women than men, with a male to female ratio of approximately 1:4 [36].

1.1.2 CFS in adolescents

There have been noted distinct peaks in the age prevalence of CFS. The first of these peaks is in adolescence, between 10 to 19 years' age [7]. However, most research on CFS worldwide is focused on adult patients. The work in our group focuses on adolescents with CFS.

1.1.3 Genetic research in CFS

Genealogical [2] and twin [14] studies suggest a hereditary predisposition in CFS. Gene expression analyses of peripheral blood mononuclear cells (PBMC) [39; 56] reported significant differences in CFS patients compared to healthy subjects in the level of expression of several genes. Our group previously performed ribonucleic acid (RNA) sequencing (RNA-seq) of RNA isolated from whole blood samples from adolescent CFS patients and healthy controls (HCs) [84]. Significant differences in expression levels were found between the two groups for 176 genes (Appendix 1: Table S1). 14 of the genes found to be differentially expressed were further examined by reverse transcription quantitative polymerase chain

reaction (RT-qPCR), and for two of these (*APOBEC3A* and *PLSCR1*) there was found a statistically significant differential expression in CFS patients. Possible upstream transcriptional regulators of some of the differently expressed genes (DEGs) were identified using the bioinformatics software package Ingenuity Pathway Analysis (IPA) (Table 1.1).

Table 1.1 A selection of upstream transcriptional regulators for the observed dataset based on Ingenuity Pathway Analyses (IPA) mechanistic network enrichment [84].

Upstream regulator	p-value of overlap	Target molecules in dataset
IRF7	0,0001	*CASP4, IFI16, IRF9, NAMPT, PLSCR1, S100A8, TLR8*
STAT6	0,0005	*ACSL1, HIPK2, IFI16, IL1RN, IRF4, IRF9, PLSCR1, ST6GAL1*

1.1.4 Neuroendocrine research in CFS

Previous studies have found several candidate genes for CFS related to catecholaminergic and serotoninergic signaling systems [38; 81; 83]. Enhanced sympathetic and attenuated parasympathetic cardiovascular nervous activity has been reported in studies on adolescents with CFS [120; 121; 124]. CFS has been associated with hypothalamic-pituitary-adrenal (HPA) axis hypofunction, including enhanced negative feedback of the HPA axis (hypersuppression) and blunted HPA axis responsiveness [90]. Some studies have found tendencies of mild hypocortisolism in CFS [85], whereas others have found no significant differences in cortisol levels between CFS patients and controls [123]. HPA axis changes are associated with more prominent symptoms and disability and poorer outcomes to standard CFS treatments [90]. The results from the RNA-seq performed previously in our group showed that DEGs are associated with neuroendocrine markers of altered HPA-axis and autonomic nervous activity [84].

1.1.5 CFS and the immune system

Previous studies indicate modest immunological alterations in CFS, such as low-grade systemic inflammation [8; 108]. There have also been several findings of attenuated natural killer (NK) cell-function in CF patients [13; 73; 76]. A beneficial effect of treatment with the anti- CD20 antibody rituximab was reported, suggesting a role for B cells [37]. However, negative results from a follow-up study were recently reported by the scientists in a Norwegian newspaper [29]. Investigations of plasma cytokine levels have been inconclusive,

with some studies showing no significant difference between CFS patients and healthy controls [116; 125], and other studies suggesting that interleukin (IL)-1 and tumor necrosis factor (TNF) may be implicated in the pathophysiology [75]. Results from RNA-seq showed that genes related to B cell differentiation and survival such as *EBF1*, *CD79A*, *CXCR5*, *TNFRSF13C*, and *FLT3* are down-regulated in CFS (Figure 1.1) [84]. In addition, several genes related to innate antiviral responses and inflammation were up-regulated in CFS.

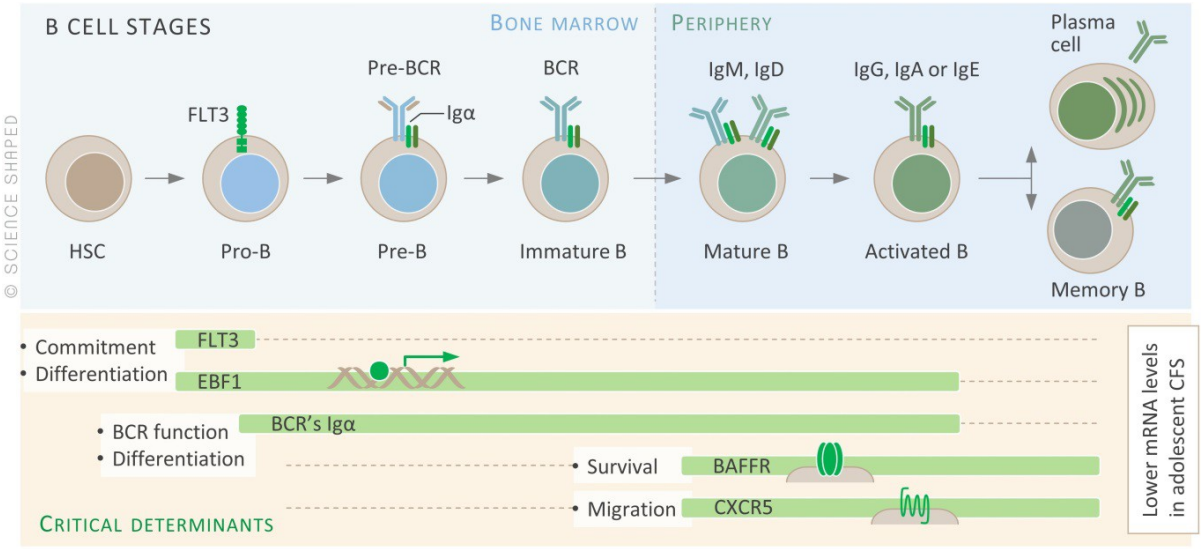


Figure 1.1: Genes related to B-cell commitment, differentiation, survival and migration reported to be down-regulated in adolescent CFS. Figure made by Ellen Tends at Science Shaped.

1.1.6 The 'sustained arousal' model

It has been suggested that all the documented features of CFS may be attributed to an aberrant neurobiological stress response or 'sustained arousal' [122]. The model proposes that certain threats, such as long-lasting infections and psychosocial challenges, may provoke a prolonged stress response, affecting the autonomic nervous system, as well as the endocrine and immune systems. The model suggests that immune alterations are secondary to neuroendocrine alterations, as both sympathetic and parasympathetic nervous activity have complex effects on several immune cells as well as inflammation [89; 127].

1.2 microRNA

1.2.1 microRNA

microRNAs (miRNAs) are small, noncoding RNAs, about 22 nucleotides (nt) in length, which play key roles in posttranscriptional regulation of gene expression [52]. Each miRNA target specific messenger RNA (mRNA) sequences and bind them, causing repression of translation or degradation of the transcript. There are thousands of different miRNAs encoded in the human genome, and it has been estimated that they regulate well over half of the human transcriptome [91].

1.2.2 miRNA biogenesis and storage

The biogenesis of miRNAs is illustrated in Figure 1.2. Pri-miRNAs are miRNA precursor molecules that are transcribed, either from introns of protein-coding genes or from independent genes, and fold into hairpin structures [17]. These molecules are processed in the nucleus by Drosha, an RNase III enzyme which in the presence of a cofactor cleaves the pri-miRNA at the stem of the hairpin structure [57]. The resulting ~70 nt molecule called a pre-miRNA is exported into the cytosol by a nuclear transport receptor called exportin 5 (EXP5). Once in the cytosol, the molecule is further processed by another RNase III enzyme called Dicer. Dicer cleaves the pre-miRNA close to the terminal loop, releasing a ~22 nt miRNA/miRNA* duplex. One, or sometimes both, of the strands in this duplex is further processed. The strand, which is now a mature miRNA, is by the aid of Dicer and other proteins incorporated into a microRNA-induced silencing complex (miRISC) together with AGO and GW182 proteins. The thermodynamic stability of the miRNA duplex ends determines which of the two strands is preferentially loaded into miRISC [88]. The other strand of the duplex is either degraded, or incorporated into a different miRISC. Newly generated miRNAs, translationally repressed mRNAs, AGO- and GW182 proteins may accumulate in discrete cytoplasmic foci known as processing bodies (P-bodies) or glycine-tryptophan bodies (G-W bodies) [60]. These are dynamic structures that are enriched in proteins involved in translational repression, deadenylation, decapping and mRNA degradation, and function in storage as well as decay of repressed mRNAs. Previous data have shown positive correlation between miRNA-mediated mRNA repression and

accumulation of target mRNAs in P bodies, but P-bodies are not essential for miRNA-function, rather a consequence of miRNA-mediated repression [34].

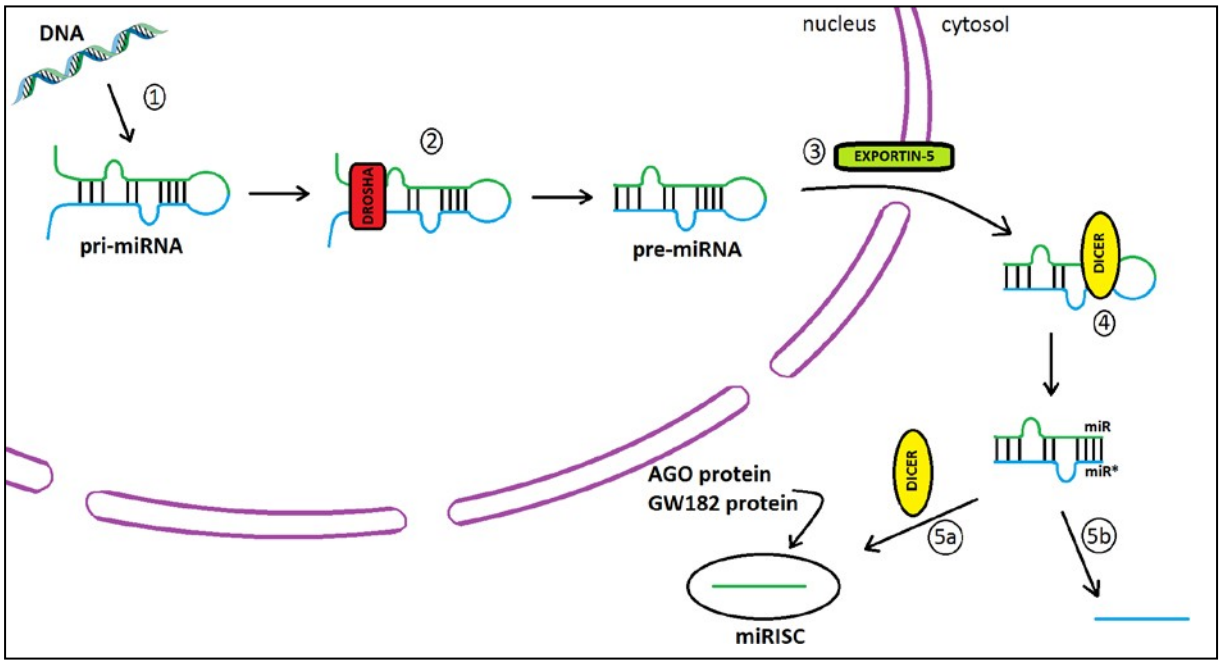


Figure 1.2: Biogenesis of miRNAs. 1: a pri-miRNA is transcribed from DNA. 2: Drosha cleaves the pri-miRNA at the stem of the hairpin structure, resulting in a pre-miRNA.3: the pre-miRNA is exported from the nucleus into the cytosol by the aid of Exportin-5. 4: Dicer cleaves the pre-miRNA close to the terminal loop, releasing a miR/miR*-duplex. 5a: By the aid of Dicer, the miR-strand of the duplex is loaded into a miRISC complex containing AGO and GW182 proteins. 5b: The miR* strand is degraded, or processed individually.

1.2.3 Regulation of miRNA biogenesis

miRNA biogenesis is tightly regulated at several levels, and dysregulation of miRNA is associated with several human diseases [57]. miRNA gene transcription is regulated in a similar manner to that of protein-coding genes, for instance by diverse transcription factors and epigenetic processes. Auto regulatory feedback loops also seem to be important in miRNA-regulation, where a miRNA may directly base-pair with and repress mRNA sequences that encode factors involved in the biogenesis or function of the same miRNA [60]. Some miRNAs are convergently transcribed from both deoxyribonucleic acid (DNA) strands of a single locus, which results in two miRNAs with distinct seed sequences. The seed sequences are the most important mRNA binding sites in the miRNA strand.

Drosha and Dicer operate in complexes with RNA-binding proteins as well as other accessory proteins, and the regulation of these proteins affect the biogenesis of miRNAs [60]. The position of cleavage by Drosha and Dicer determines which nucleotides are in the 3’ and 5’

terminal ends of the miRNAs. The processing is not always uniform, which generates miRNA isoforms with different termini. This may be especially important in the 5' terminal end as this is the part containing the seed sequence.

RNA editing of miRNAs may prevent or enhance Drosha- and Dicer-mediated cleavage as well as preventing export of pre-miRNAs to the cytosol [60]. Editing within miRNA seed sequences may impact target specificity.

1.2.4 miRNA-mRNA interactions

Partial complementarity between a miRNA and its target site is usually sufficient for translational repression [91]. As a consequence, this mediates a complex regulation of gene expression, where a single miRNA can regulate several different mRNAs, and each mRNA can be regulated by multiple miRNAs. The miRNA-mRNA interaction usually contains mismatches and multiple nucleotide bulges [91]. However, residues between nucleotide number 2 and 7, the 'seed region', in the 5'portion of the miRNA are usually perfect or almost perfect complementary to the mRNA. Stability of this pairing is important for the degree of repression [45], although imperfect pairing in this region is sometimes compensated for by extensive 3'end interactions [91]. The mRNA target site is usually located in the 3'-untranslated region (3'-UTR) of a gene, often in multiple copies [52].

1.2.5 miRNA-mediated gene silencing

miRNA-mediated gene silencing occurs through a combination of translational repression, the shortening of mRNA poly(A) tails called deadenylation, hydrolysis of the mRNA 5'-cap structure (decapping) and 5'-to-3'-mRNA degradation [53]. There has also been suggested a role for miRNAs in direct transcriptional regulation within the nucleus [103], however this aspect of miRNA function is not explored in the same extent as post-transcriptional regulation. It has been established that mRNA degradation is the dominant effect of miRNAs at steady state, accounting for 66-90% of the miRNA-mediated silencing observed in cultured mammalian cells [53].

1.2.6 miRNA target degradation

miRNA target degradation is catalyzed by enzymes involved in the cellular 5'-to-3' mRNA decay pathway as shown in figure 1.3. First, mRNAs are deadenylated by the deadenylase complexes PAN2-PAN3 and CCR4-NOT [53]. Following deadenylation, mRNAs are decapped by decapping protein 2 (DCP2) aided by cofactors. Deadenylated and decapped mRNAs are degraded by the cytoplasmic nuclease 5'-to-3' exoribonuclease 1 (XRN1). miRISCs direct the recruitment of this decay machinery via the GW182 adaptor proteins.

The coupling of the steps in this decay pathway seems to be disrupted in oocytes, early embryos, cell-free extracts and probably neuronal cells [53]. This leads to the accumulation of deadenylated and transcriptionally repressed mRNAs and represent potentially reversible silencing of mRNAs.

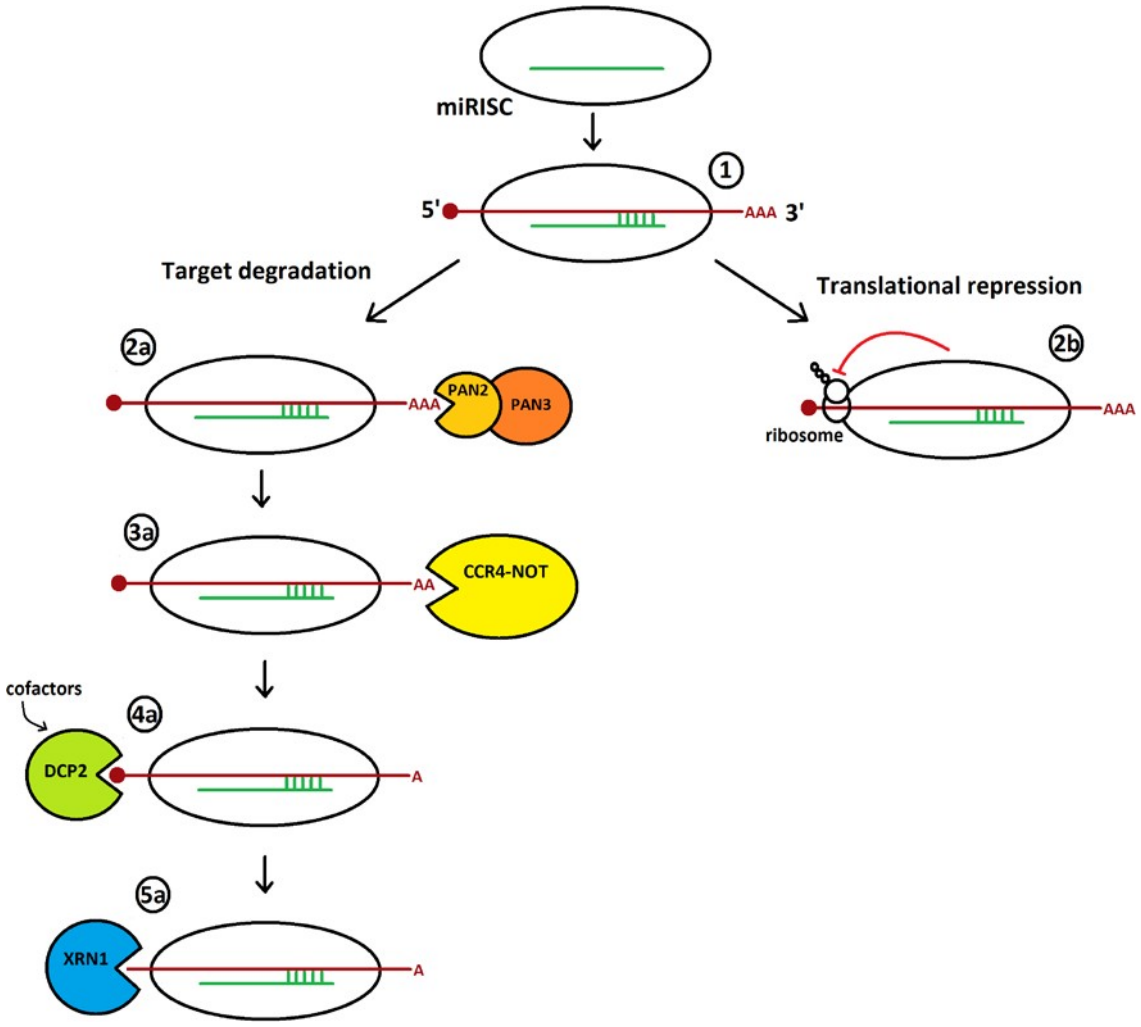


Figure 1.3: miRNA-mediated gene silencing. 1: The miRISC binds an mRNA by partial complementary binding to the miRNA, which leads to either target degradation (a) or translational repression (b) of the mRNA. 2a: The mRNA is deadenylated by the PAN2-PAN3 complex. 3a: The deadenylation process is continued by the CCR4-NOT complex. 4a: mRNA is decapped by DCP2 and its cofactors. 5a: The mRNA is degraded in the 5' to 3' direction by the XRN1. 2b: The miRISC inhibit the cap-dependent initiation of translation.

1.2.7 miRNA-mediated translational repression

Translational repression is observed at early time points after miRNA expression, but has generally weak effects and accounts for only 6-26% of the repression of each endogenous target in mammalian cells [53]. The molecular mechanism of miRNA-mediated translational repression has long been controversial, and the precise mechanism still remains to be solved. However, as shown in figure 1.3, the emerging consensus in the field is that miRNAs inhibit the cap-dependent initiation of translation. Cap-dependent initiation of translation requires the eukaryotic translation initiation factor 4F (eIF4F) complex which interacts with the cap

structure at the 5’ end of the mRNA molecule and recruits the 43S pre-initiation complex. Current models of miRNA-mediated translational repression involve RNA helicases like the initiation factor paralogues eIF4A1 and eIF4A2, and the translational repressor and decapping activator DDX6.

1.2.8 miRNA nomenclature

Mature miRNAs are named with a “miR” prefix followed by a numerical identifier [15]. Animal miRNAs are distinguished from plant miRNAs with the use of a dash symbol between the prefix and the number, and identical miRNAs from different organisms are named with the same identifier except a three letter prefix defining the species name followed by a dash (eg. hsa-miR for human miRNAs). All miRNAs that are mentioned in this text are human unless specified otherwise. Suffixes -3p and -5p indicate two different mature miRNAs generated from the 3’ and 5’ arms of the same pre-miRNA hairpin, respectively.

1.2.9 Circulating miRNA

miRNAs are released from cells into the blood circulation, where they function as signaling molecules between cells [35]. The mechanisms of which the miRNAs are released from the cells are still largely unknown [22], but it has been suggested that they may be secreted in membrane encapsulated particles like exosomes, as well as bound to proteins or as free miRNAs not bound to any biomolecule [35].

1.2.10 miRNA and Chronic Fatigue Syndrome

As previously mentioned, CFS patients have been found to differ significantly from healthy subjects in the level of expression of several genes [39], but little is known about the molecular mechanisms behind these differences. One possible mechanism is dysregulation of miRNAs, which has been associated with a variety of diseases, including CFS [12].

Three miRNAs (miR-127-3p, miR-142-5p and miR-143-3p) were reported to be significantly up-regulated in plasma samples of adult CFS patients compared to healthy controls. This pattern was found using high-throughput sequencing of six CFS patients and six healthy controls followed by confirmative RT-qPCR [11].

Differential expression of miRNAs has also been found in CFS lymphocytes [12; 94]. Brenu et al [12] identified nine differentially expressed miRNAs out of 19 miRNAs analyzed by RT-qPCR in cytotoxic lymphocytes of adult CFS patients. In another study, microarray analyses followed by qPCR validation showed that four miRNAs (miR-99b, miR-330-3p, miR-30c and miR-126) were differentially expressed in PBMCs of adult CFS patients [94].

These studies were performed in adult CFS patients and with relatively small sample sizes. Therefore, studies in larger cohorts are needed in order to elucidate whether differential miRNA expression levels in CFS can provide clues about molecular mechanisms as well as potential new biomarkers of the disease.

1.2.11 Selection of miRNA candidates

IPA was used on RNA-seq data [84] to identify miRNAs that function as upstream regulators of some of the DEGs (Table 1.2). A manual search in the miRNA database miRTarBase confirmed miRNAs that function as upstream regulators of some of the DEGs and regulators (Table 1.2). These miRNAs were selected as candidates for further investigation.

Table 1.2 miRNA regulators of selected DEGs and upstream transcriptional regulators based on IPA mechanistic network enrichment and miRTarBase database search.

miRNA	**Target gene**	**Method of identifying miRNA**
miR-223-3p	*IL1RN*	IPA
miR-16-5p	*FLT3*	IPA
miR-142-5p	*TSPAN3, SLC38A1, DDX5*	miRTarBase
miR-143-3p	*SEICISBP2L*	miRTarBase
miR-486-3p	*BCL7A, OGT*	miRTarBase
miR-146a-5p	*IRF7*	miRTarBase

1.2.12 miR-223-3p

miR-223-3p is known for regulating genes implicated in regulation of cell proliferation, cell cycle regulation, and activation of transcription and replication. miR-223-3p has also been reported to be significantly down-regulated in NK-cells of CFS patients [12].

IL1RN, which was up-regulated in CFS patients compared to HCs in RNA-seq (Appendix 1: Table S1), is one of the target genes regulated by miR-223-3p (Table 1.2). *IL1RN* encodes the IL-1 receptor antagonist (IL-1ra), which binds and inhibits the IL-1 receptor and thereby acts as a potent anti-inflammatory molecule. IL-1 is a highly pro-inflammatory cytokine that

promotes the recruitment of T lymphocytes, neutrophils and macrophages to inflamed tissues [109]. *IL1RN* variants are associated with the severity of several inflammatory/autoimmune diseases such as ulcerative colitis, psoriasis and multiple sclerosis, as well as coronary heart disease.

1.2.13 miR-142-5p

miR-142-5p has been found to be significantly up-regulated in plasma of CFS patients [11]. It is expressed by immune cells and is important for T cell development. Inhibition of miR-142-5p has been shown to cause over-activation of T cells and B cell hyperstimulation by increasing the levels of Cluster of Differentiation 84 (CD84), IL-10, SLAM-associated protein (SAP) and Immunoglobulin G (IgG) [28].

Amongst other genes, miR-142-5p regulates Tetraspanin 3 (*TSPAN3*), Solute Carrier Family 38 Member 1 (*SLC38A1*) and DEAD-box helicase 5 (*DDX5*) (Table 1.2) which are all DEGs found in RNA-seq (Appendix 1: Table S1).

TSPAN3 encodes a membrane protein which has been shown to be required for a normal response to the chemokine SDF1/CXCL12 [61]. SDF1 interacts with the CXCR4 receptor, and this interaction is important for up-regulation of the surface protein CD20 on B-cells in chronic lymphocytic leukemia [92]. It also regulates the proliferation of oligodendrocytes, a process essential for normal myelination and repair [110].

SLC38A1 codes for an important transporter of glutamine, which is an intermediate in the detoxification of ammonia and the production of urea, as well as being a precursor for the synaptic transmitter glutamate [41].

DDX5 encodes DEAD box protein 5, which is an RNA helicase that function in multiple cellular processes, including transcription and ribosome biogenesis [49]. It has also been shown to control Th17 cytokine production and the differentiation of Th17 cells at steady state and in animal models of autoimmunity. *DDX5* plays important roles in several viral infections, like Japanese encephalitis virus [64] and hepatitis B [129]. It is also involved in the control of circadian rhythms [68].

1.2.14 miR-143-3p

miR-143-3p has been found to be significantly up-regulated in plasma of CFS patients [11]. It has been identified as a neutrophil specific miRNA [3] and its expression is up-regulated in cases of heightened erythropoiesis such as in polycythemia [105], a disease often accompanied by fatigue and exertion intolerance.

miR-143-3p regulates *SECISBP2L* (SECIS binding protein 2 like) (Table 1.2), which bind SECIS (Sec insertion sequence) elements present on selenocysteine protein mRNAs [30] and was found to be differentially expressed in CFS patients compared to HCs in RNA-seq (Appendix 1: Table S1).

1.2.15 miR-486-3p

Comparing microarray transcriptome profiles for two datasets of 189 and 208 healthy adults, respectively, Preininger et al [95] idetified nine axes of variation with between 99 and 1028 strongly co-regulated transcripts in common. One of these axes (Axis 3), contain several genes important for B-cell activation. Genes in this axis also overlap largely with several DEGs found in RNA-seq [84]. miR-486 is specified as a miRNA regulating this axis.

miR-486-3p regulates BCL tumor suppressor 7A (*BCL7A)* and O-Linked N-Acetylglucosamine Transferase (*OGT)* (Table 1.2), which are both genes reported to be differentially expressed in CFS patients compared to HCs in RNA-seq (Appendix 1: Table S1).

The specific function of *BCL7A* is not known, but it has been shown to be a target for a gene translocation in a Burkitt lymphoma cell line which leads to disruption of the N-terminal region of the gene product [80]. This is thought to be related to the pathogenesis of a subset of high-grade B cell non-Hodgkin lymphoma.

OGT encodes a glycosyltransferase that catalyzes the addition of a single N-acetylglucosamine in O-glycosidic linkage to serine or threonine residues [112]. It acts as a nutrient sensor, and aberrant *OGT*-activity has been linked to insulin resistance as well as cancer and neurodegenerative diseases such as Alzheimer's disease [63]. It has been suggested as a potential biomarker of maternal stress [112], and modulates circadian clock oscillation [65].

1.2.16 miR-146a-5p

miR-146a-5p has been found to be significantly down-regulated in NK-cells of CFS patients [12]. It has also been suggested to be important in regulation of axis number 3 mentioned above [95].

miR-146a-5p regulates Interferon regulatory factor 7 (*IRF7*) (Table 1.2) which is an upstream regulator of the IRF pathway. Three important genes in this pathway, *IRF9*, *IRF4* and *TLR8* were found to be up-regulated in CFS patients in RNA-seq (Appendix 1: Table S1). *IRF7* encodes a transcription factor which has been shown to play a role in the transcriptional activation of virus-inducible host genes [86]. It plays a critical role in the innate immune response against DNA and RNA viruses, and is associated with Epstein-Barr Virus (EBV) latency.

1.3 Epigenetics

1.3.1 Epigenetics

Epigenetics can be defined as the study of heritable changes in gene function that do not influence or change the underlying DNA sequence [31]. These changes are usually brought about by chemical modifications of chromatin; posttranslational modifications of amino acids on the amino-terminal tail of histones, covalent modifications of DNA bases or histone variants. Most of these modifications are reversible, however relatively stable [4].

1.3.2 DNA methylation

DNA methylation is the addition of a methyl group to a base in DNA [31]. Cytosine methylation is the most common and most excessively researched type of DNA methylation. In this text, 'DNA methylation' will refer to the methylation of cytosine unless otherwise specified. Methylation of cytosine results in the formation of 5-methylcytosine (5mC) (Figure 1.4).

DNA methylation occurs primarily in cytosines that are part of cytosine-phosphate-guanine dinucleotides (CpGs), although cytosines in non-CpG sequences may also be methylated [31]. Methylation is estimated to occur at 70-80% of CpGs throughout the genome [62]. The

remaining unmethylated CpGs are mostly found near gene promoters in dense clusters called CpG islands (CGIs).

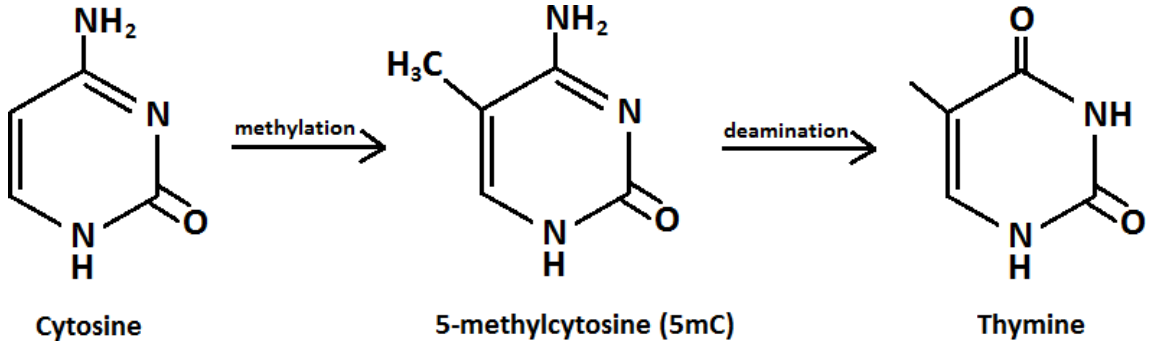


Figure 1.4: Conversion of cytosine to 5-methylcytosine and 5-methylcytosine to thymine. Methylation of cytosine leads to the formation of 5-methylcytosine (5mC). Deamination of 5mC leads to the formation of thymine.

1.3.3 CpG islands

CGIs are short (approximately 1 kb), CpG-rich regions of the genome which are generally refractory to DNA methylation and play central roles in gene regulation [72]. CGIs are often associated with gene promoters, but also exist within gene bodies where their function is largely unknown [54]. CGIs are thought to exist because they are probably never or only transiently methylated in the germline [54]. Methylated sequences in the germline are prone to spontaneous or enzymatic deamination, leading to the conversion of 5mC to thymine (Figure 1.4). The avoidance of germline deamination in CGIs is thought to be the reason for the uneven distribution of CpGs in the genome.

Most promoter-associated CGIs remain largely unmethylated in somatic cells [54]. However, some repressed genes do have methylated promoter CGIs. These are usually restricted to genes at which there is long-term stabilization of repressed states, for instance imprinted genes, genes located on the inactive X chromosome and genes that are exclusively expressed in germ cells.

1.3.4 De novo DNA methylation

In mammals, DNA methylation patterns are established by the de novo methyltransferases DNA methyltransferase (Dnmt) 3A and Dnmt3B [62]. This pattern is largely established

during early embryogenesis, around the time of implantation. During post-implantation development, further epigenetic reprogramming occurs in primordial germ cells. De novo methylation may be directly targeted to specific DNA sequences, or guided by interaction with other DNA binding proteins or RNA interference [31]. However, details on the establishment of methyl cytosine patterns remains largely unknown [114]. Dnmt3A and Dnmt3B enzymes exhibit specificity for de novo methylation of CpG sites, but seem unable to discriminate between primary DNA sequences. This suggests that de novo methylation may be guided by another mechanism, such as chromatin modifications or small RNAs.

1.3.5 Maintenance of DNA methylation patterns

DNA methylation patterns are maintained by the maintenance methyltransferase Dnmt1. This enzyme is associated with replication foci and function by reproducing patterns of methylated and unmethylated CpG sites during DNA replication [114]. These patterns are reproduced with extremely high fidelity; repressed and active transcription states of methylated and unmethylated transgenes have been reported to be stably propagated over long periods of time in cell culture, corresponding to as many as 100 cell divisions [104].

1.3.6 Demethylation

DNA methylation is considered a stable epigenetic mark, but reduced levels of methylation are observed during development [62]. This net loss of methylation can either occur passively, via replication in the absence of functional maintenance methylation pathways, or actively by removal of methylated cytosines.

Active demethylation requires a mechanism that involves cell division or DNA repair and excision of the base rather than the removal of the methyl group directly from the 5mC [54]. In mammals, no effective DNA demethylase has been identified, and the active DNA demethylation pathways are still disputed. Several enzymes have been proposed to be involved in active and passive demethylation. Examples include the ten-eleven translocation (TET) which catalyze the conversion of 5mC to 5-hydroxymethylcytocine, deaminases, and thymine DNA glycosylase (TDG) which is involved in the repair of T:G mismatches caused by 5mC deamination. Another enzyme that may be involved is activation-

induced cytidine deaminase (AID), which mediates cytosine deamination. However, the demethylation role of this enzyme is somewhat controversial [5].

1.3.7 The functions of DNA methylation

DNA methylation is a relatively stable epigenetic mark that is critical for diverse biological processes including gene and transposon silencing, imprinting, and X chromosome inactivation as well as normal development of the organism [62]. It is evolutionary ancient, and DNA methylation defects in mammals are embryonic lethal.

DNA methylation has for a long time been associated with transcriptional repression and gene silencing, as promoters of silenced genes possess significantly more methylated cytosines in comparison to actively transcribed genes [31]. However, lately it has been proposed that the function of DNA methylation varies with context [54]. Recent genome-wide studies of the methylome suggest that the position of the methylation influences its relationship to gene control. DNA methylation in regions associated with transcriptional start sites (TSSs) generally blocks initiation of transcription, while methylation within the gene body may have the opposite effect, stimulating transcription elongation.

1.3.8 DNA methylation on promoter regions

The promoter region of a gene is the DNA sequence that define the transcription start site as well as the direction of transcription [1]. It is typically located directly upstream of the transcription initiation site and is recognized by RNA polymerase and transcription factors. Many eukaryotic genes have a conserved promoter sequence called the TATA box, located 25 to 35 base pairs (bp) upstream of the TSS.

DNA methylation levels in promoter regions directly correlates with transcriptional repression, and several mechanisms have been proposed concerning this relationship [31]. One suggestion is that the modification directly prevents the binding of transcription factors [50]. Another theory is that it attracts mediators of chromatin remodeling like histone-modifying enzymes [9], or other repressors of gene expression.

It has long been a discussion in the field whether silencing or DNA methylation comes first [54]. DNA methylation has been suggested to function as a “lock” to reinforce a silenced state

of X-linked genes. Lock et al [71] showed that methylation of the *Hprt* gene on the inactive X chromosome occurred after the chromosome had been inactivated. Genome-wide studies in cancer cells have also shown that genes with CGI promoters that are already silenced by Polycomb complexes are much more likely to become methylated in cancer compared to other genes [54]. However, other studies suggest that DNA methylation has a more instructive role in initiating rather that reinforcing the silencing [55; 82].

Genes that are CpG-poor at their TSSs have substantial fluctuations in the promoter methylation levels, and there is an apparent inverse relationship between methylation of non-CGIs and expression [54]. However, the details of the role of methylation in controlling non-CGI TSSs are still unknown as research so far has been mostly focused on CGIs.

1.3.9 DNA methylation within the gene body

As mentioned, methylation within the gene body may stimulate transcription elongation [54]. Gene body methylation may also have an impact on splicing.

Methylation in repeat regions is important for chromosomal and genome stability as it suppresses the expression of transposable elements (TEs) while allowing the host gene to undergo transcriptional elongation [54]. DNA methylation is also important in telomeres as well as centromeres to ensure correct attachment of microtubules to centromeres [31].

1.3.10 DNA methylation and histone modifications

There seems to be an intimate communication between DNA methylation and epigenetic modifications of histones [114]. Posttranslational modifications of histones are important for the organization of chromatin structure and regulation of gene transcription [31]. Histones are proteins important for the packing of DNA into nucleosomes, chromatin entities consisting of DNA spooled around an octamer of histones. Each octamer contains two units of each histone H2A, H2B, H3 and H4, which binds in total 146 bp of DNA. Linker DNA of approximately 20 bp connects the nucleosomes and associates with a H1 histone. A variety of histone-modifying enzymes are responsible for a multiplicity of posttranslational modifications on specific serine, lysine and arginine residues on the amino-terminal tail of the histones.

The best characterized modifications are acetylations and methylations of lysine residues on H3 and H4 histones. Acetylation of lysine is associated with transcriptional activation, while methylation of lysine may lead to activation or repression of transcription depending on the amino acid and to which extent it is modified. Histone-modifying enzymes may be targeted to specific DNA sequences or need interaction with intermediates or targeting by RNA interference (RNAi). It is unclear whether histone modifications are correctly replicated during mitosis.

Hypermethylation of CGIs in gene promoters triggers deacetylation of local histones, whereas lower levels of histone acetylation seem to sensitize to targeted DNA methylation.

1.3.11 Environmental effects on epigenetic patterns

Recently, it has been suggested that DNA methylation patterns may change later in life and thus provide a platform through which the environment could sculpt the genome and affect the phenotype [107]. Previous studies imply that environmental exposures may alter the epigenome after birth [78], supported by data suggesting that the DNA methylation pattern in the hippocampus is responsive to environmental exposures even in adults [118]. Studies of hypermethylation events in aging tissues [111], as well as data revealing differences in DNA methylation emerging later in life in monozygotic twins also support this hypothesis. Toxic components of the environment such as nickel, cadmium, ionizing radiation, tobacco smoke and infectious agents are suspected to have an effect on DNA methylation, chromatin organization and histones as essential carriers of specific epigenetic information [44]. Environmental and social stress has also been associated with epigenetic changes, although in rodents [42]. Analysis of DNA methylation within the promoter region of the *NR3C1* gene has shown associations between prenatal stress and increased DNA methylation. Initial studies also suggest that epigenetic modifications, particularly DNA methylation, are associated with the quality of early-life maternal care [101; 108]. In adult animals, chronic stress is associated with reductions in the expression of the *BDNF* gene, and there may be an epigenetic basis to these effects. In rats, immobilization stress combined with exposure to predator odor has been found to induce significant increases in hippocampal DNA methylation within the *BDNF* gene [102] gene. Initial studies indicate that individuals may have differential susceptibility to the effects of stress. In a recent study, stress-susceptible mice were found to have increased levels of *CRF* mRNA in the paraventricular nucleus of the

hypothalamus and decreased DNA methylation within the *CRF* gene [33]. In contrast, stress-resilient mice were found to have no changes in *CRF* mRNA or DNA methylation of this gene.

1.3.12 DNA methylation in disease

Aberrant DNA methylation has been implicated in several human diseases, and is known to be involved in human carcinogenesis [54]. However, known examples of 'true' epigenetic disease are at present limited, as it is often difficult to distinguish causal epigenetic aberrations from those that are merely consequences of disease [77]. The clearest examples of epigenetic disease is a subset of imprinting disorders, in which genes that are normally subject to epigenetic silencing, are improperly expressed [47]. Disease risk may be modulated by environmental factors, although this remains largely a concept until more evidence is found [77].

1.3.13 DNA methylation in CFS

Stable alterations in gene expression have been reported in several studies of CFS, however there have been few studies exploring whether epigenetic modifications may play a role in these alterations. de Vega et al [27] used an epigenome-wide methylation profiling microarray to examine the DNA methylome in PBMCs isolated from CFS patients and compared this to healthy controls. They found an increased abundance of differentially methylated genes related to immune regulation, cellular metabolism and kinase activity. The results were validated by bisulphite pyrosequencing. Using the same methods, the researchers mapped loci that were epigenetically modified in PBMCs of CFS patients and examined their sensitivity to glucocorticoids [26]. They detected 12,608 differentially methylated sites between CFS patients and healthy controls, predominantly localized to cellular metabolism genes. Among CFS patients, glucocorticoid sensitivity was associated with differential methylation at 13 loci.

In the present study, three genes were selected for investigation of methylation levels based on their relevance in CFS. Nuclear Receptor Subfamily 3 Group C Member 1 (*NR3C1)* is an important gene in the HPA axis. *IRF7* was, as mentioned above, identified as the top regulator of DEGs found in RNA-seq (Table 1.1) and is an important gene in innate immunity. *CD79A*

is one of the DEGs found in RNA-seq (Appendix 1: Table S1) and is an important gene in B-cell mediated immunity.

1.3.14 NR3C1

NR3C1 is a gene encoding the glucocorticoid receptor (GR) in humans. It is a complex, epigenetically regulated gene containing multiple, untranslated, alternative first exons clustered into a large non-coding 5'region [6]. Each of these alternative first exons has its own promoter located immediately upstream of the exon. The non-coding 5'region of *NR3C1* is largely thought to be responsible for transcriptional regulation of GR protein levels.

The GR is an important part of the HPA axis, functioning as a receptor for glucocorticoids (e.g. cortisol) produced by the adrenal cortex in response to stress [6]. The GR resides in the cytoplasm of cells in the hippocampus, hypothalamus and anterior pituitary gland until it binds glucocorticoids. Upon binding, the GR translocate to the nucleus where it aids in the regulation of a number of genes encoding immune, inflammatory and metabolic proteins. Binding of glucocorticoids to the GR ultimately participates in a negative feedback loop to stabilize glucocorticoid levels.

One of the proposed mechanisms of HPA axis hypofunction in CFS is enhanced GR sensitivity [115]. One possible mechanism of such enhanced sensitivity is hypomethylation of the promoter region of *NR3C1*. Vangeel et al [115] found a significantly lower methylation of the *NR3C1* promoter 1F region in whole blood from female CFS patients compared to healthy controls. Lower *NR3C1*-1F promoter methylation has also been reported in PBMCs in combat veterans with posttraumatic stress disorder (PTSD), another disorder that has been associated with enhanced negative feedback to the HPA axis [126]. Thus, it would be interesting to see if a similar pattern could be discovered in the present study.

1.3.15 IRF7

IRF7, encoded by the *IRF7* gene, is a multifunctional transcription factor associated with Epstein-Barr virus (EBV) latency and regulation of type I interferon (IFN) responses, which are important in both innate and adaptive immunity [86].

Inactive IRF7 resides in the cytoplasm as a latent form which upon pathogenic infection can be activated by phosphorylation followed by translocation into the nucleus. Inside the

nucleus, it forms a transcriptional complex with other co-activators which bind to promoter regions of target genes to activate transcription. IRF7 is associated with EBV latency and is induced as well as activated by the EBV oncoprotein latent membrane protein-1 (LMP1).

IRF7 is a lymphoid-specific factor, which is constitutively expressed in the cytoplasm in B cells, plasmacytoid dendritic cells (pDCs) and monocytes in the spleen, thymus and peripheral blood lymphocytes. Its expression is potently inducible by type I IFNs and virus infection. The positive regulatory feedback between IRF7 and type I IFNs during antiviral immune responses is the major source for *IRF7* expression in the cell.

IRF7 expression is silenced by promoter hypermethylation [74] and reagents that loosen chromatin structure such as topoisomerase II inhibitors as well as methyltransferase inhibitors can induce expression of *IRF7* [86]. The *IRF7* promoter contains a putative CpG island flanking the TATA box (Appendix 10). This CpG island is endogenously methylated in the human fibroblasts, 2fTGH, which results in *IRF7* gene silencing in these cells. Methylation of the *IRF7* promoter is also found in immortalized fibroblasts from Li-Fraumeni syndrome and in lung cancer cell lines.

IRF7 is known as the “master regulator” of type I IFN production. Type I IFNs constitute a highly pleiotropic cytokine family that participates in immune modulation as well as autoimmune responses, oncogenesis, cellular development and homeostasis [86]. Aberrant production of IFNs is associated with many types of diseases such as cancers, immune disorders and multiple sclerosis. Fine regulation of type 1 IFN production is critical to maintain homeostasis. IRF7 is also required for monocyte differentiation to macrophages, however the mechanism remains unclear. IRF7 remains at low levels in most cell types, and its stability is tissue specific.

Three important genes in the IFN pathway, *IRF9*, *IRF4* and *TLR8* were found to be up-regulated in CFS patients compared to HCs in RNA-seq [84], which may suggest higher activity in this pathway in CFS patients. As IRF7 function as a master regulator of the IFN pathway and its expression may be controlled by promoter methylation, it would be interesting to assess whether the DNA methylation level of this promoter changes with the level of fatigue.

1.3.16 CD79A

Immunoglobulin α (CD79A) is a transmembrane glycoprotein encoded by the *CD79A* (previously called *mb-1)* gene [93]. It belongs to the Ig superfamily and is restricted to B lymphocytes. Together with a similar protein, CD79B, CD79A forms a disulfide-linked heterodimer which functions as a signaling component of the B-cell receptor (BCR). Stimulation of the BCR by an antigen (ag) induces phosphorylation of the CD79A/CD79B cytoplasmic tyrosine residues, initiating a signaling cascade that can result in proliferation, differentiation, or death of the mature B cells. The CD79A/CD79B heterodimer is also part of the pre-BCR complex and is important for progression through early stages of B cell development. Although commitment of hematopoietic stem cells to the B cell lineage does occur in the absence of CD79A, the molecule is required for further development to the pre-B, immature, and mature B cell stages in both mouse and human.

Expression levels of *CD79A* in diffuse large B-cell lymphoma (DLBCL) samples have been found to be significantly correlated with aberrant hypo-methylation at a genome-wide scale [24]. Ha et al [43] found that *CD79A* expression correlated with demethylation in the *CD79A* promoter region and the 5'end of the first intron of *CD79A*-expressing cells, which suggests that DNA methylation in these sites plays a role in expression of this gene.

CD79A was found to be significantly down-regulated in CFS patients compared to HCs in RNA-seq (Appendix 1: Table S1). The mechanism behind this difference in expression remains to be elucidated, but one possible explanation may be DNA methylation differences.

2 Aims

The overall aim of this master project was to investigate possible causes for the reported alterations in gene expression related to CFS. It was hypothesized that alterations in miRNA-mediated regulation and/or epigenetic regulation such as DNA methylation could lead to differential gene expression, and possibly play a part in the pathophysiology of the disease. Relevant miRNA and methylation patterns of selected differentially expressed genes from CFS patients and HCs were analyzed, and differences between groups were investigated. The relationship between miRNA expression/methylation levels and level of fatigue was also examined.

Five sub-goals were defined:

I) Analyze gene expression data to predict miRNAs that may be possible upstream regulators of differentially expressed genes (DEGs) found in RNA-seq [84].

II) Validate differential expression of a selection of DEGs by RT-qPCR.

III) Extract and measure miRNA levels in CFS patients and healthy controls, and analyze miRNA found in sub-goal I as well as relevant miRNA from the literature. Examine whether there are differences in miRNA expression levels between CFS patients and healthy controls.

IV) Extract DNA from CFS patients and healthy controls, and analyze the DNA methylation levels of selected DEGs as well as other genes relevant for CFS. Investigate whether or not there are differences in DNA methylation levels between CFS patients and healthy controls.

V) Examine the relationship between methylation levels and the level of fatigue.

3 Materials and methods

3.1 Study design

3.1.1 Data collection

This study is part of The Norwegian Study of Chronic Fatigue Syndrome in Adolescents: Pathophysiology and Intervention Trial (NorCAPITAL) project conducted at the Department of Pediatrics, Oslo University Hospital. Data was collected from March 2010 to May 2012. Data from a NorCAPITAL sub-cohort of in total 60 participants was used, based on a previous study in our group [84]. Serum samples were harvested and frozen at -80°C. 59 samples in total (37 CFS patients and 22 HCs) were used for miRNA analyses. Whole blood samples were harvested and frozen at -80°C. 54 samples (32 CFS patients and 22 HCs) were used for DNA methylation analyses.

3.1.2 Clinical markers

"Steps per day" was used as a marker of the patients' disability. The *activPAL* accelerometer device (PAL Technologies Ltd, Glasgow, Scotland) was used for monitoring of daily physical activity during 7 consecutive days, as described elsewhere [106], and the main outcome was steps walked per day. Fewer steps walked per day indicated a more impaired functioning.

In addition, total score from the Chalder Fatigue Questionnaire (CFQ) was used as a marker of fatigue severity. In this questionnaire, 11 variables were scored on a Likert scale from 0 to 3, giving a total score of 0-33, where a higher score indicated more severe fatigue [51].

3.2 Validation

3.2.1 Gene expression analysis

RNA extraction and synthesis of complementary DNA (cDNA) were performed by PhD student Chinh Bkrong Nguyen as described elsewhere [84].

Gene expression was analyzed by RT-qPCR. *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase) was used as a reference gene for normalization. Primers for the target mRNAs were designed using the software Primerquest Tool (Integrated DNA Technologies, Coralville, IA, USA). Primers were designed to be complementary to part of the sequences in the genomic DNA, with a minimum melting temperature (Tm) of 57°C, a maximum Tm of 63°C, and a guanine-cytosine (GC) content of 50-60%. A basic local alignment search tool (BLAST) was used to check for identical sequences in other mRNAs to avoid unspecific binding. The primers were synthesized by Thermo Fisher Scientific, Waltham, MA, USA. For primer sequences, see table 3.1.

Table 3.1 Genomic primers used for qPCR. A: Adenosine, C: Cytosine, G: Guanine, T: Thymine, Bp: base pairs

Primer	**Sequence 5'→3'**	**Bp**
FLT3 Forward	AGT GGC TCA TGC CTG TAA TC	20
FLT3 Reverse	AGA AGG GTT TCA CCG TGT TAG	21
IL1RN Forward	CCT CAG CAA CAC TCC TAT TG	20
IL1RN Reverse	GAT GCA CAC TAC CCA GTT AC	20
GAPDH Forward	CCAACTGCTTAGCACCCCT	19
GAPDH Reverse	TGGCATGGACTGTGGTCAT	19

Individual primer assays were investigated by RT-qPCR analysis. All reactions were performed in a 10 µL reaction volume on 96-well plates utilizing the 7900 HT Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). For each gene to be investigated a master mix was prepared, containing the EvaGreen Sso Fast Master mix (Biorad Laboratories, CA, USA) and primer assays specific for the target genes. The cDNA as well as the master mix were loaded onto the 96-well qPCR plate, which then was sealed with a plastic film and centrifuged. The qPCR reaction was set up with the following schedule: 95°C for 10 min followed by 45 cycles of 95°C for 5 sec and 60°C for 15 sec. For protocol, see appendix 2.

The EvaGreen dye is constructed of two monomeric DNA-binding dyes linked by a flexible spacer. When DNA is available, the inactive, looped conformation shifts to a random conformation that binds DNA and emits fluorescence at 520nm. This was used to quantify the amount of PCR product for each cycle. The amount of fluorescence in each well per cycle was used to make an amplification plot (Figure 3.1). The threshold value of fluorescent signal was defined at the approximate middle of the linear phase of the amplification curve (logarithmic scale). The threshold cycle (Ct) value was determined by which cycle each specific sample reached the threshold value of fluorescence.

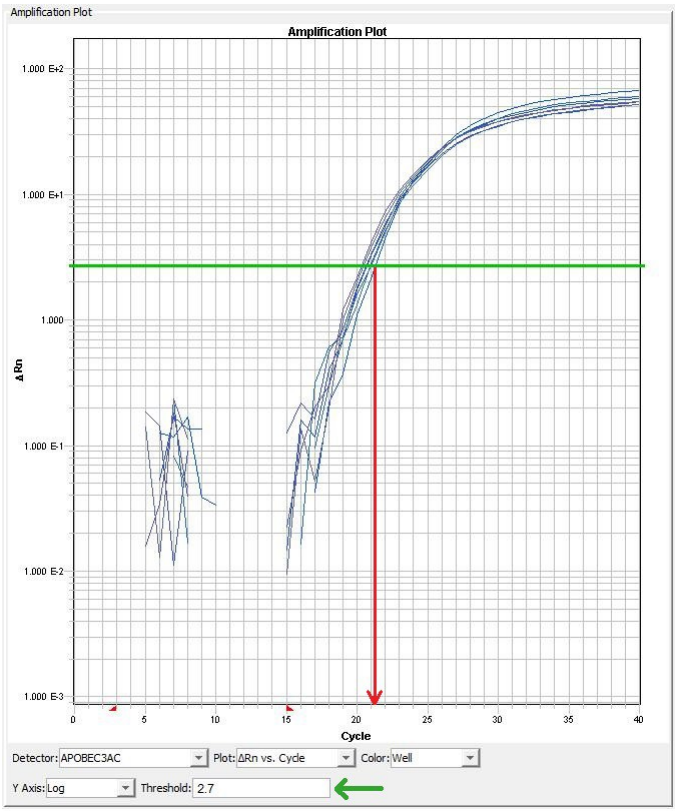


Figure 3.1: RT-qPCR amplification plot of GAPDH. Each curve corresponds to one sample (one well on the PCR plate). The y-axis shows the ΔRn value, a parameter which corresponds to the amount of PCR product produced, and the x-axis shows the number of PCR cycles. The threshold value of fluorescence (pointed by the green arrow) is defined at the approximate middle of the linear part of the amplification curve and is shown on the plot as a green line. The threshold cycle (Ct) value of each sample is determined by at which cycle the specific sample reaches the threshold value of fluorescence (marked by the red arrow).

3.3 miRNA

3.3.1 miRNA extraction

Total miRNA was extracted from serum samples using a miRNeasy Serum/Plasma kit (Qiagen, Hilden, Germany). All samples were centrifuged and the supernatant collected to remove remaining cell debris. QIAzol lysis reagent and carrier RNA (MS2) were added to each sample, before mixing and incubating the samples at room temperature (RT). 5nM Synthetic *C. elegans* microRNA (cel-miR-39) (Qiagen, Hilden, Germany) with no human homolog was added to each sample in identical amounts (the spike-in-step) to serve as reference gene in following analyses. Chloroform was added to the samples, followed by thorough mixing and incubation at RT. Centrifugation of the samples resulted in the division

into three phases; a lower organic phase containing proteins and lipids, an interface containing DNA, and an upper aqueous phase containing RNA. The upper aqueous phase was transferred to new Eppendorf tubes and absolute ethanol (EtOH) was added to the samples. After mixing, the samples were transferred to RNeasy Mini Spin columns in collection tubes and centrifuged, resulting in binding of miRNA to the column. The column was washed with RWT- and RPE-buffer and dried. miRNA was eluted from the column by centrifugation using nuclease-free water. For protocol, see appendix 3.

miRNA concentrations were measured for a small selection of the samples using a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

3.3.2 RT-qPCR

Total miRNA was converted by reverse transcription to cDNA using miRCURY LNA™ microRNA PCR, Polyadenylation and cDNA synthesis kit II (Exiqon, Qiagen, Hilden, Germany). A reaction mix was made containing reverse transcriptase, a primer, deoxynucleotides and RNase inhibitor as well as 2 µl of miRNA sample. The reaction was carried out in an Eppendorf Mastercycler ep Gradient S Thermal Cycler (Eppendorf, Hamburg, Germany) following a program of 42°C for 60 min, 95°C for 5 min and then samples were kept at 4°C. cDNA was diluted 10 times in nuclease-free water. For protocol, see appendix 4.

Expression levels of specific miRNAs were measured by RT-qPCR. Cel-miR-39 was used as a reference. Primer pair sets for the target miRNAs were supplied by Exiqon, Qiagen, Hilden, Germany. Primer target sequences are shown in table 3.2.

Table 3.2 miRNA primer sets used for qPCR. A: Adenosine, C: Cytosine, G: Guanine, T: Thymine, Bp: base pairs

Primer	**Sequence 5'→3'**	**Bp**
hsa-miR-486-3p primer set	CGGGGCAGCUCAGUACAGGAU	21
cel-miR-39-3p primer set	UCACCGGGUGUAAAUCAGCUUG	22
hsa-miR-223-3p primer set	UGUCAGUUUGUCAAAUACCCCA	22
hsa-miR-142-5p primer set	CAUAAAGUAGAAAGCACUACU	21
hsa-miR-143-3p primer set	UGAGAUGAAGCACUGUAGCUC	21
hsa-miR-146a-5p primer set	UGAGAACUGAAUUCCAUGGGUU	22

Individual miRNA assay analyses were performed in a 10 µL reaction volume on a 7900 HT Real-Time PCR System (Applied Biosystems, Foster city, CA, USA). Each reaction contained SYBR green master mix (Exiqon, Qiagen, Hilden, Germany), primers and cDNA. They were loaded onto a 96-well qPCR plate, which then was sealed with a plastic film and centrifuged. The qPCR reaction was set up with the following schedule: 95°C for 10 min followed by 45 cycles of 95°C for 10 sec and 60°C for 1 min. For protocol, see appendix 5.

The SYBR green dye emits fluorescence at 520nm when incorporated into a double-stranded DNA molecule, and was therefore used to quantify the amount of PCR product for each cycle. The amount of fluorescence in each well per cycle was used to make an amplification plot as shown in figure 3.1 and described above.

3.4 DNA methylation

3.4.1 DNA isolation

Total DNA was isolated and extracted from whole blood using a DNeasy kit (Qiagen, Hilden, Germany), and a modified optimized protocol. For each sample, 2 ml of anticoagulated blood was centrifuged and the supernatant containing blood plasma was removed. The remaining pellet containing white blood cells was washed and suspended in phosphate buffered saline (PBS). Proteinase K and lysis buffer (Buffer AL) was added to each sample, followed by thorough mixing and incubation at 56°C. Absolute EtOH was then added to the samples, followed by mixing and centrifugation in spin columns. Flow-through containing cellular debris was discarded and the column was washed several times followed by centrifugation steps. DNA was then eluted from the column by centrifugation using an elution buffer (Buffer AE). For protocol, see appendix 6.

3.4.2 Fragmentation of DNA

Fragmentation of the DNA samples was done at a core facility (Norwegian Radium Hospital, Oslo, Norway). 5 µg of genomic DNA was sheared into fragments of ~100-600 bp using a E220 evolution Focused-ultrasonicator (Covaris, Woburn, MA, USA), with a target peak of 300 bp. Complete shearing conditions are shown in Table 3.3. The fragmentation of each sample was then visualized by D1000 Screen Tape (Agilent Technologies, Santa Clara, CA,

USA), showing peaks between 252-336 bp for all samples (Figure 3.2). For protocol, see appendix 7.

Table 3.3: Shearing conditions for fragmentation of DNA using a E220 evolution Focused-ultrasonicator (Covaris, Woburn, MA, USA)	
microTUBE – 50 µl sample - from 150 to 1,500 bp	
Target BP (Peak)	300
Peak Incident Power (W)	175
Duty Factor	10%
Cycles per Burst	200
Treatment Time (s)	50
Temperature (°C)	7
Water Level – E220	6
Sample volume (µl)	50
E220 - Intensifier (pn500141)	Yes

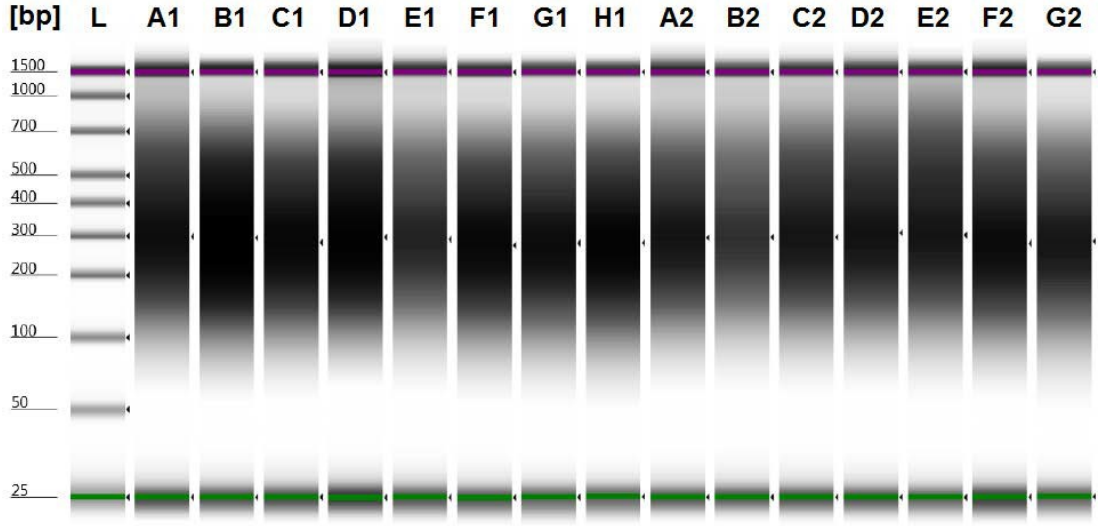


Figure 3.2: Fragmentation visualization of a selection of DNA samples. All samples show fragmentation peaks of approximately 300 bp.

3.4.3 Immunoprecipitation of methylated DNA

Fragments containing methylated DNA were isolated by immunoprecipitation (IP) from fragmented genomic DNA using a MagMeDIP kit (Diagenode, Liege, Belgium) utilizing a 5mC antibody and magnetic beads (Figure 3.3). An IP incubation mix containing DNA sample and buffers was prepared for each sample. Positive (meDNA) and negative (unDNA) controls were added in identical amounts to the IP incubation mix of the two control samples to serve as internal controls in the following analyses. All the IP incubation mixes were incubated at 95°C and chilled on ice. A portion of each IP incubation mix (10% of the IP sample) was kept as an input control sample. The 5mC antibody (ab) was diluted 1:2 and

mixed with buffer to a diluted ab mix which was then added to each IP incubation mix. Magnetic beads were washed and re-suspended and added to each IP sample. The IP samples were incubated overnight (ON) at 4°C on a rotating wheel. Following incubation, all IP samples were thoroughly washed on a magnetic rack before DNA isolation. DNA was isolated from all input samples as well as the IP samples. Each sample was added buffer DIB containing proteinase K and incubated at 55°C followed by incubation at 100°C to release DNA from the magnetic beads. The supernatants containing DNA were then transferred to new tubes and stored in aliquots at -20°C. For protocol, see appendix 8.

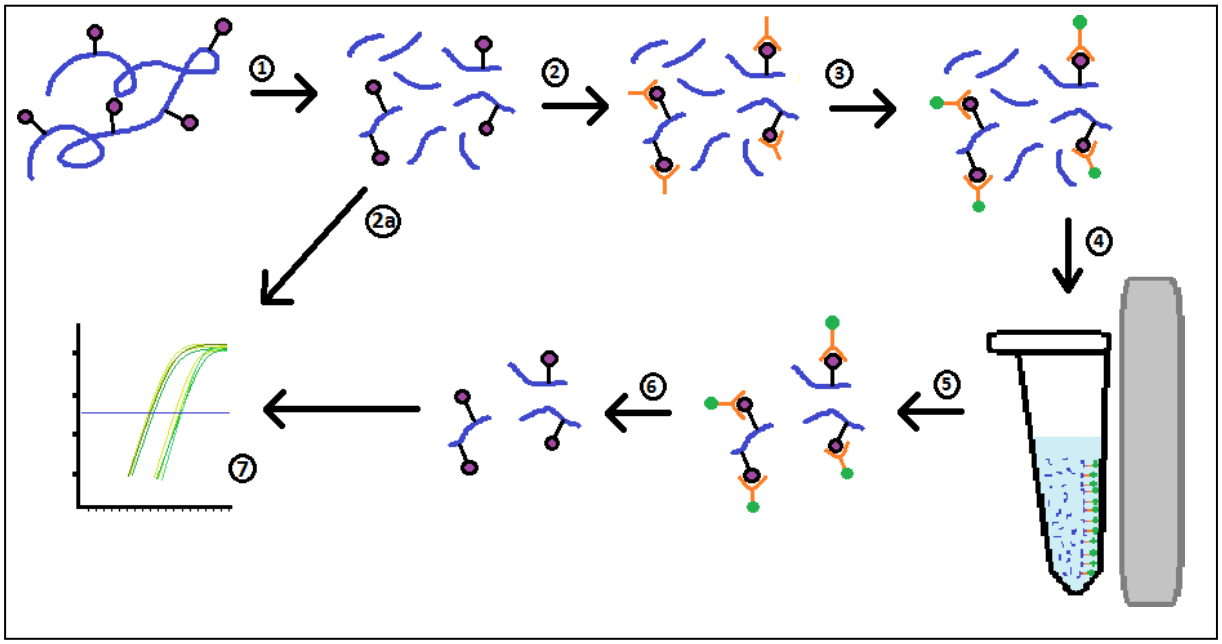


Figure 3.3: Immunoprecipitation of methylated DNA. 1: genomic DNA containing 5mC is sheared into fragments of ~100-600 bp. 2a: genomic DNA fragments are used as input control in qPCR analysis. 2: Ab bind to 5mC. 3: Magnetic beads bind to the ab-DNA complexes. 4: A magnetic rack is used to capture ab-DNA complexes. 5: Sample is washed, resulting in the isolation of ab-DNA complexes. 6: ab-bead complexes are removed. 7: DNA fragments containing 5mC are analyzed by qPCR.

3.4.4 qPCR

Expression levels of specific loci were measured by qPCR. *XIST* was used as a positive (methylated) control, while *GAPDH* served as a negative (non-methylated) control. Two control samples were included in each qPCR plate and detected with positive (meDNA) and negative (unDNA) control primers provided with the MagMeDIP kit. Each IP sample was compared to the corresponding input sample to calculate % methylation. Primers for the target loci were designed using the software Primerquest Tool (Integrated DNA Technologies, Coralville, IA, USA). Primers were designed to be complementary to part of the promoter

region of each gene, with a minimum Tm of 57°C, a maximum Tm of 63°C, and a GC content of 50-60%. BLAST was used to check for identical sequences in other genes to avoid unspecific binding. The primers were synthesized by Thermo Fisher Scientific, Waltham, MA, USA. For primer sequences, see table 3.4.

Individual methylation assay analyses were performed in 25 µL reaction volumes using a 7900 HT Real-Time PCR System (Applied Biosystems, Foster city, CA, USA). Each reaction contained Ssofast EvaGreen supermix (Bio-rad Laboratories Inc., Hercules, CA, USA), primers and DNA. They were loaded onto a 96-well qPCR plate, which then was sealed with a plastic film and centrifuged. The qPCR reaction was set up with the following schedule: 95°C for 7 min followed by 45 cycles of 95°C for 15 sec and 60°C for 1 min and completed with 95°C for 1 min. For protocol, see appendix 9.

Table 3.4: Genomic promoter region primers used for qPCR: A: Adenosine, C: Cytosine, G: Guanine, T: Thymine, Bp: base pairs

Primer	**Sequence 5'→3'**	**Bp**
XIST Forward	GCA GCC ATA TTT CTT ACT CTC T	22
XIST Reverse	ACG CCA TAA AGG GTG TTG	18
GAPDH Forward	ACC GGG AAG GAA ATG AAT G	19
GAPDH Reverse	ATC ACC CGG AGG AGA AAT	18
NR3C1 Forward	CCC TCA TTC TTG TGC CTA TG	20
NR3C1 Reverse	GGT CCA GTG ATT TGG TAT CAG	21
IRF7 Forward	CAC TGT TTA GGT TTC GCT TTC	21
IRF7 Reverse	CGC ACA CAT GAA GTC ACA	18

3.5 Statistics

Data was processed in Microsoft Excel 2010 (Microsoft Corporation, NM, USA), and statistical analyses were conducted in IBM SPSS Statistics 24/25 (International Business Machines, NY, USA). A p-value below 0.05 was considered significant. qPCR data was analyzed by the 2-ΔΔCt method. Shapiro-Wilk's test and visual inspection of histograms, Q-Q plots and box plots were used to assess the distribution of the data. Differences in expression levels and methylation levels were analyzed by comparing groups in a Student t-test where the data was found to be normally distributed. Equality of variances was analyzed by an F-test. Data sets with a non-normal distribution were analyzed using the nonparametric Mann-Whitney U test and Kendall's tau-b correlation analysis. Graphs were made using SigmaPlot 12.5 (Systat Software Inc, San Jose, CA, USA).

4 Results

4.1 Validation

4.1.1 Gene expression analysis

IL1RN and *FLT3* are two of the genes that were found to be differentially expressed in CFS patients by RNA-seq (Appendix 1: Table S1). For each of these genes, a miRNA was identified as a possible upstream regulator (Table 1.2). If *IL1RN* and *FLT3* are differentially expressed in CFS patients, it would be interesting to analyze the corresponding miRNA regulators.

First, we tested whether we could confirm regulation of these genes using a different method than RNA-seq. The RT-qPCR analysis method was chosen for this purpose.

To investigate the expression of the two genes, relative quantification was performed according to the $2^{-\Delta\Delta Ct}$ method. Relative quantity (RQ) values were log-transformed to LOG2-values to be able to compare relative expression levels between the two groups; chronic fatigue syndrome patients (CFS) and healthy controls (HC). Higher LOG2-value was considered a higher relative expression of the gene (Table 4.1). Both *FLT3* and *IL1RN* were normalized against the stably expressed *GAPDH* as a reference gene.

Table 4.1: LOG2- and P-values indicating the difference in expression between groups. Data are given by mean ± standard error of the mean (SEM).

	CFS (LOG2)	**HC (LOG2)**	**P-value**
FLT3	0.841±0.117	0.833±0.092	0.564
IL1RN	3.364±0.263	2.454±0.366	0.093

When examining the distribution of *FLT3* expression data, two outliers with values of 3.159 and 4.018 in the CFS group as well as one outlier with a value of 5.012 in the HC group was detected using the 2.2 interquartile range (IQR) rule (Figure 4.1). These values were excluded from the following analyses. A Shapiro-Wilk's test ($p>0.05$) showed that the *FLT3* expression values were normally distributed for both CFS patients and HCs, with a skewness of -0.381 (SE=0.403) and a kurtosis of -0.361 (SE=0.788) for the CFS patients, and a skewness of 0.561 (SE=0.536) and a kurtosis of -0.540 (SE=1.038) for the HCs. However, as variances were not considered equal between the two groups ($F(1,50)=4.766$, $p=0.034$), the Mann-Whitney U

test was performed to compare the two groups. A boxplot of the *FLT3* expression in the two groups is shown in figure 4.2. The mean value for the CFS group was 0.841, with a SEM of 0.117. The mean value for the HC group was 0.833, with a SEM of 0.092. No significant difference in *FLT3* expression level was found between CFS and HC (p=0.564).

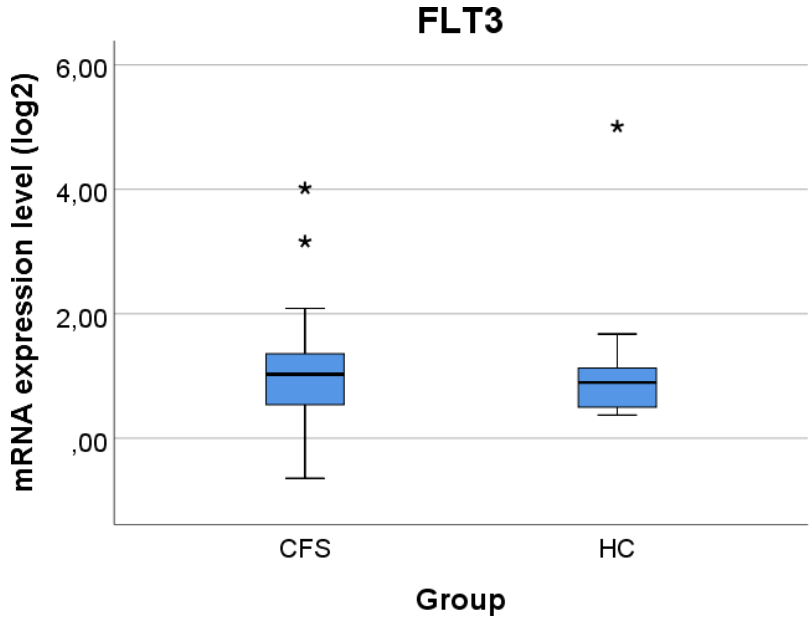


Figure 4.1: Boxplot of the *FLT3* expression level. Two outliers for the CFS group and one outlier for the HC group are marked (*).

A boxplot of the *IL1RN* expression in the two groups is shown in figure 4.3. The mean value for the CFS group was 3.364, with a SEM of 0.263. The mean value for the HC group was 2.454, with a SEM of 0.366. A Shapiro-Wilk's test (p>0.05) showed that the *IL1RN* expression values were normally distributed for HCs, with a skewness of -0.317 (SE=0.616) and a kurtosis of -0.893 (SE=1.191). However, *IL1RN* expression values of the CFS patients were not considered normally distributed (Shapiro-Wilk's test, p<0.05), with a skewness of 0.972 (SE=0.441) and a kurtosis of 1.274 (SE=0,858). Thus, the Mann-Whitney U test was performed to compare the two groups. There was not found a significant difference in *IL1RN* expression level between CFS and HC (p=0.093). However, there is a clear trend indicating increased expression in CFS compared to HC (Table 4.1), which is corresponding to the pattern found in RNA-seq (Appendix 1, Table S1). This, and the relatively low p-value (although not significant), argues for further analyses of the miRNA that regulates this gene.

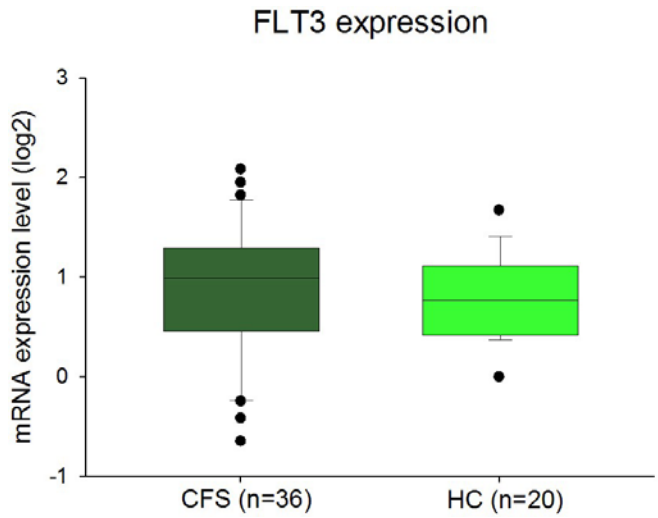


Figure 4.2: Relative gene expression (RT-qPCR) of *FLT3*. mRNA expression values (log2) in CFS patients (CFS) and healthy controls (HC). p=0.564, Mann-Whitney U-test.

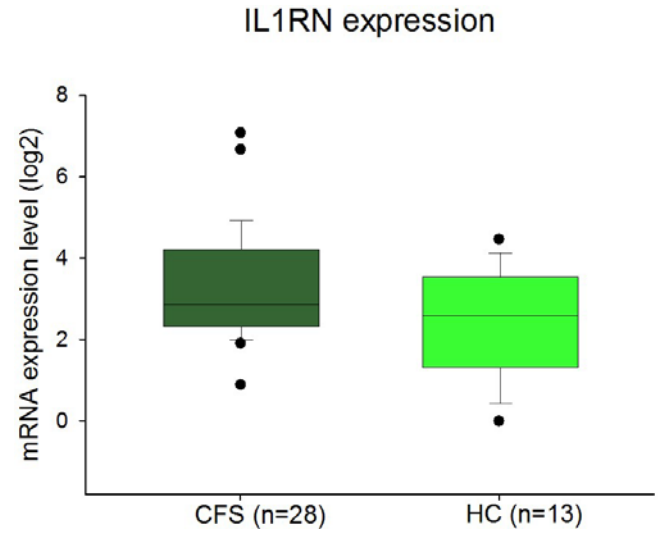


Figure 4.3: Relative gene expression (RT-qPCR) of *IL1RN*. mRNA expression values (log2) in CFS patients (CFS) and healthy controls (HC). p=0.093, Mann-Whitney U-test.

4.2 miRNA

4.2.1 miRNA analysis

Specific miRNAs were chosen for analysis based on their proposed function as regulators of genes that were found to be differentially expressed in CFS patients by RNA-seq (Appendix 1: Table S1), or previously identified as differentially expressed in CFS patients [11; 12].

Due to the spike-in step of the extraction protocol, miRNA concentrations could not be measured in an optimal manner.

The relative expression of miR-223-3p, miR-142-5p, miR-143-3p, miR-486-3p, and miR-146a-5p in CFS patients and HC was determined by RT-qPCR.

The relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method. LOG2-values were used to compare relative expression levels between the two groups. Higher LOG2-value was considered a higher relative expression of the gene (Table 4.2). For miRNA analysis, spiked-in cel-miR-39 was used as a reference for the calculations.

Table 4.2: LOG2 and P-values indicating the difference in miRNA expression between chronic fatigue patients (CFS) and healthy controls (HC). Data are given by mean ± standard error of the mean (SEM).

	CFS (LOG2)	**HC (LOG2)**	**P-value**
miR-223-3p	2.256±0.130	2.255±0.253	0.998
miR-142-5p	1.124±0.187	1.548±0.246	0.190
miR-143-3p	2.055±0.193	1.721±0.277	0.468
miR-486-3p	0.824±0.268	1.265±0.235	0.299
miR-146a-5p	2.456± 0.144	2.397± 0.246	0.825

Figure 4.5 show the expression levels of the 5 miRNAs analyzed.

The mean miR-223-3p expression value for the CFS group was 2.256, with a SEM of 0.130. The mean value for the HC group was 2.255, with a SEM of 0.253. The data was considered normally distributed (Table 4.3), and the variances of the two groups were considered equal ($F(1,57)=1.331$, $p=0.253$). The miRNA expression level of the two groups was compared with the Student t test. There was not found a significant difference in miR-223-3p expression level between CFS and HC ($p=0.998$).

The mean miR-142-5p expression value for the CFS group was 1.124, with a SEM of 0.187. The mean value for the HC group was 1.548, with a SEM of 0.246. Assuming normal distribution (Table 4.3) and equal variance ($F(1,50)=0.021$, $p=0.885$), the miRNA expression level of the two groups was compared with the Student t test. There was found no significant difference in miR-142-5p expression level between CFS and HC ($p=0.190$).

The mean miR-143-3p expression value for the CFS group was 2.05, with a SEM of 0.193. The mean value for the HC group was 1.721, with a SEM of 0.277. One outlier with a value of -4,13 was detected in the HC group using the 2.2 IQR rule (Figure 4.4). This value was excluded from further analyses. The data was considered non-normally distributed for the

CFS group (Table 4.3), and was analyzed by the Mann-Whitney U test. There was not found a significant difference in miR-143-3p expression level between CFS and HC (p=0.468).

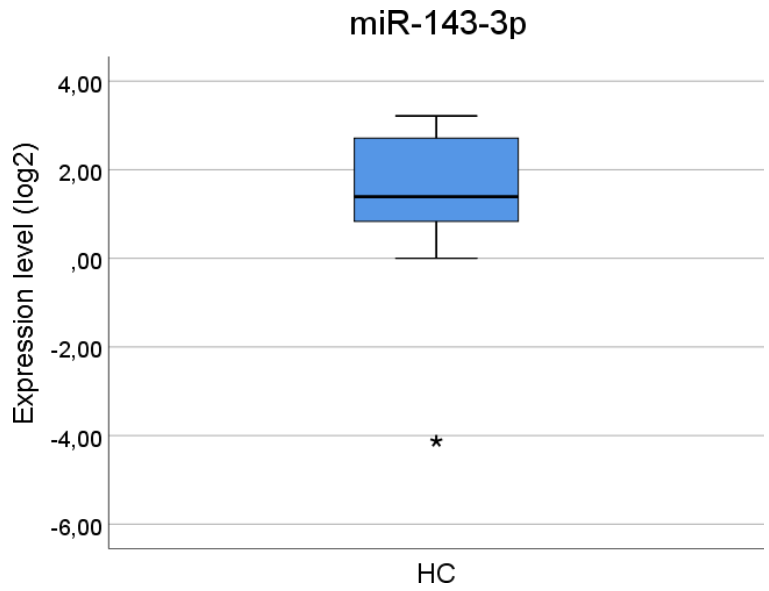
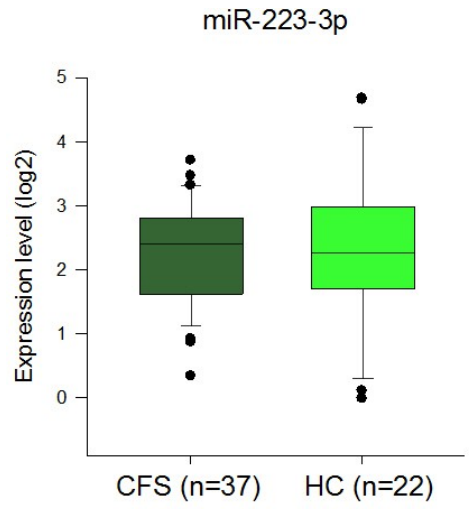


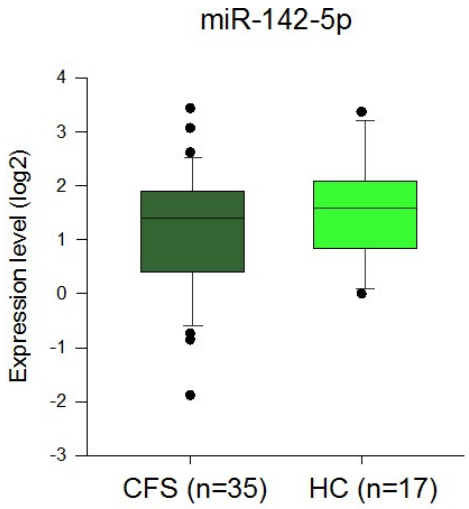
Figure 4.4: Boxplot of the relative gene expression values of miR-143-3p. One outlier for the HC group is marked (*).

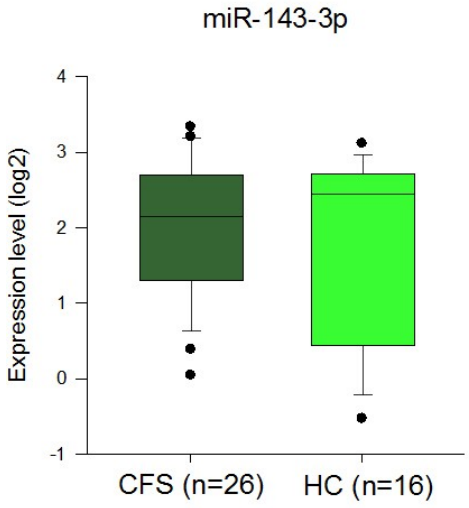
The mean miR-486-3p expression value for the CFS group was 0.824, with a SEM of 0.268. The mean value for the HC group was 1.265, with a SEM of 0.235. The data was considered normally distributed (Table 4.3), and the miRNA expression level of the two groups was compared with the Student t test. The variances of the two groups were considered equal ($F(1,28)=1.140$, $p=0.873$). There was not found a significant difference in miR-486-3p expression level between CFS and HC ($p=0.299$).

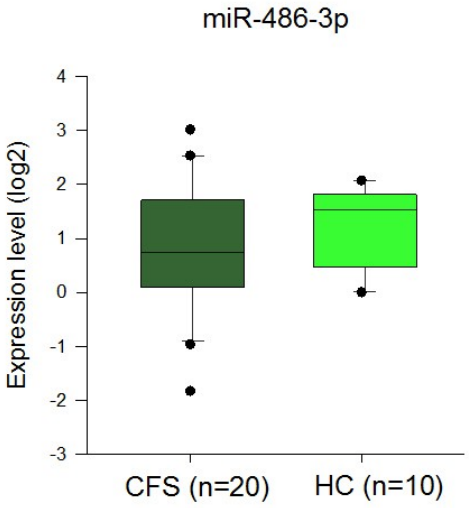
The mean miR-146a-5p expression value for the CFS group was 2.456, with a SEM of 0.144. The mean value for the HC group was 2.397, with a SEM of 0.246. The data was considered normally distributed (Table 4.3), and the miRNA expression level of the two groups was compared with the Student t test. The variances of the two groups were considered equal ($F(1,57)=0.707$, $p=0.404$). There was not found a significant difference in miR-146a-5p expression level between CFS and HC ($p=0.825$).

Table 4.3: Normality tests for miRNA analysis					
miRNA	**Group**	**p (Shapiro-Wilk)**	**Skewness (SE)**	**Kurtosis (SE)**	**Distribution**
miR-223-3p	CFS	0.566	-0.402 (0.388)	-0.379 (0.759)	Normal
	HC	0.259	0.108 (0.491)	0.576 (0.953))	Normal
miR-142-5p	CFS	0.271	-0.798 (0.414)	0.900 (0.809)	Normal
	HC	0.559	0.255 (0.512)	-0.093 (0.992)	Normal
miR-143-3p	CFS	0.008	-1.36 (0.456)	0.606 (0.887)	Non-normal
	HC	0.171	-0.076 (0.564)	-1.357 (1.091)	Normal
miR-486-3p	CFS	0.772	-0.672 (0.524)	0.463 (1.014)	Normal
	HC	0.506	0.349 (0.661)	-1.130 (1.279)	Normal
miR-146a-5p	CFS	0.414	0.172 (0.388)	0.014 (0.759)	Normal
	HC	0.198	0.589 (0.491)	1.156 (0.953)	Normal









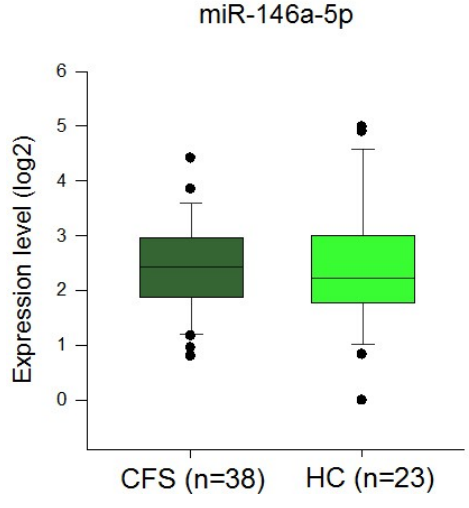


Figure 4.5:Relative gene expression (RT-qPCR) of 5 miRNAs in chronic fatigue syndrome (CFS) patients and healthy controls (HCs).
Expression levels (log2) of miR-223-2p ($p = 0.998$, Student t-test), miR-142-5p ($p = 0.190$, Student t-test), miR-143-3p ($p = 0.468$, Mann-Whitney U test), miR-486-3p ($p = 0.299$, Student t-test), and miR-146a-5p ($p = 0.825$, Student t-test).

4.3 DNA methylation

Three genes were chosen for DNA methylation analysis based on their association to CFS. *IRF7* was identified by IPA analysis to be a top regulator of important DEGs found by RNA-seq (Table 1.1). *NR3C1* is an important part of the HPA axis and has been shown to be hypomethylated in female CFS patiens. *CD79A* was found to be significantly down-regulated in CFS patients compared to HCs in RNA-seq (Appendix 1: S1).

DNA fragments containing 5mC was isolated using methylated DNA immunoprecipitation. To assess the methylation levels of *NR3C1*, *IRF7* and *CD79A* in CFS and HC, qPCR was used. The efficiency of methyl DNA immunoprecipitation of each gene was calculated from qPCR data using the following calculation:

% (meDNA-IP/ Total input)= 2^[(Ct(10%input) - 3.32) - Ct(meDNA-IP)]x 100%

Where 2 is the amplification efficiency, Ct /meDNA-IP) and Ct (10% input) are threshold values obtained from the exponential phase of qPCR for the methyl DNA sample and input sample respectively. The compensatory factor (3.32) is used to take into account the dilution 1:10 of the input. The recovery is the % (meDNA-IP/Total input) and will be referred to as % methylation. This value was used as the main variable in the following analyses.

4.3.1 IRF7

When examining the distribution of *IRF7* methylation data, three outliers with values of 1.0222, 1.5979 and 0.1508 were detected for the CFS group using the 2.2 IQR rule (figure 4.6). These values were excluded from further analyses. A Shapiro-Wilk's test ($p>0.05$) showed that the % methylation values of IRF7 were normally distributed for CFS patients, with a skewness of -0.665 (SE=0.434) and a kurtosis of 0.290 (SE=0.845). However, % methylation values of the HCs were not considered normally distributed (Shapiro-Wilk's test, $p<0.05$), with a skewness of 1.128 (SE=0.491) and a kurtosis of 0.852 (SE=0,953).

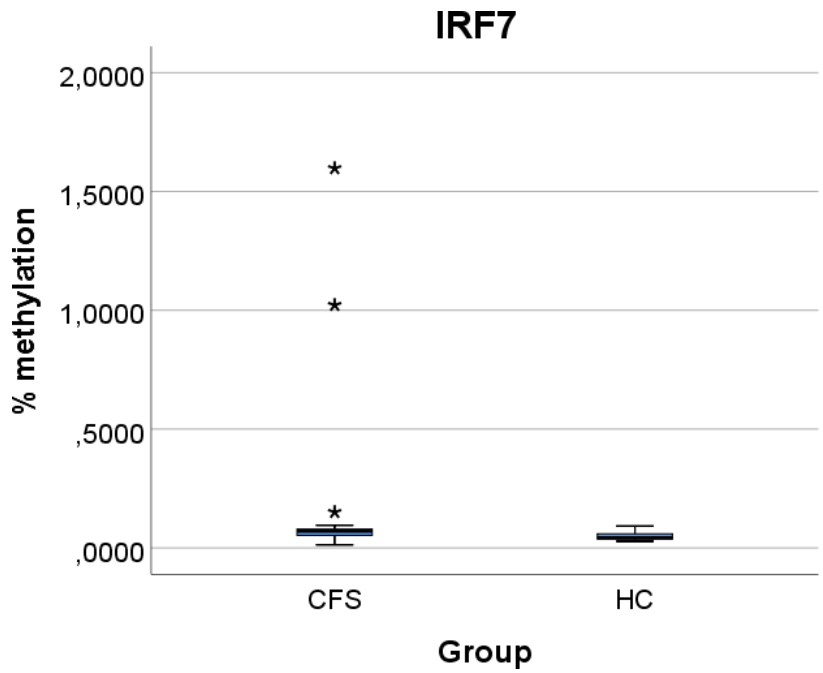


Figure 4.6: Boxplot of the % methylation of IRF7. Three outliers for the CFS group are marked (*).

A boxplot of the % methylation of *IRF7* in the two groups is shown in figure 4.7. The mean value for the CFS group was 0.0626, with a SEM of 0.003. The mean value for the HC group was 0.0486, with a SEM of 0.003. As the HC data was considered non-normally distributed, the Mann-Whitney U test was performed to compare the two groups. There was found a significant difference in *IRF7* methylation level between CFS and HC (p=0.004), where the *IRF7* promoter seem to be significantly hypermethylated in CFS patients compared to controls.

When corrected by gender, the statistical significance disappears for females (p=0.093), although the overall trend still shows higher methylation levels in CFS. For males, the difference is still statistically significant (p=0.035). This is shown in figure 4.8.

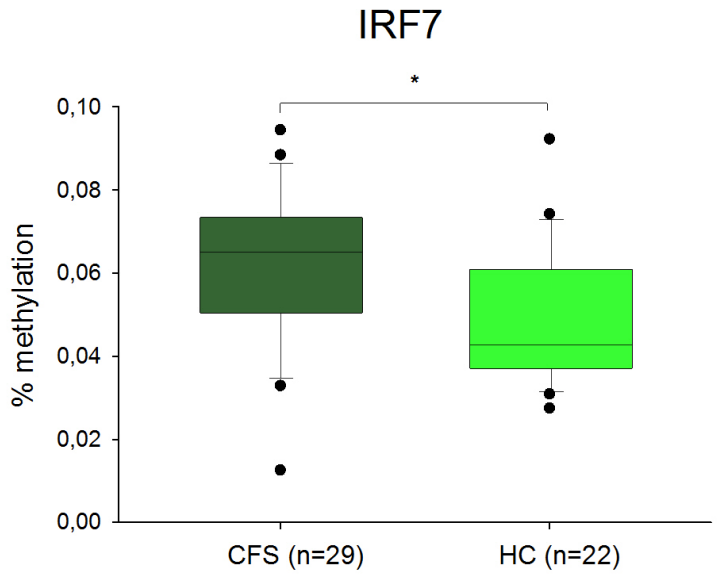


Figure 4.7: Methylation levels of the *IRF7* gene. % Methylation in CFS patients (CFS) and healthy controls (HC). P=0.004, Mann-Whitney U-test.

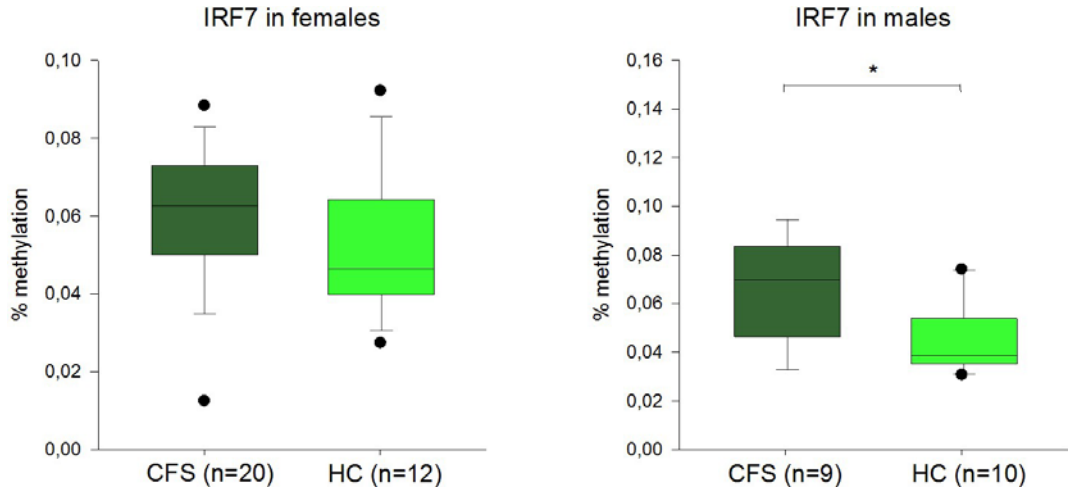


Figure 4.8: Methylation levels of *IRF7* in females and males. % Methylation in CFS patients (CFS) and healthy controls (HC) separated by gender; Females (p=0.093, Mann-Whitney U-test) and males (p=0.035, Mann-Whitney U-test).

Kendall's tau-b correlation analyses were conducted in order to determine the relationship between *IRF7* methylation levels and the level of fatigue in CFS patients. Number of steps per day was used as a marker of disability, and CFQ was used as a marker of fatigue severity, with a higher CFQ indicating a higher level of fatigue. There was found a weak negative correlation between *IRF7* methylation levels and steps per day as shown in figure 4.9, but this relationship was not statistically significant (T_b = -0.130, p= 0.333). There was also found a non-significant positive correlation between *IRF7* methylation levels and CFQ (T_b = 0.226, p= 0.107) as shown in figure 4.10.

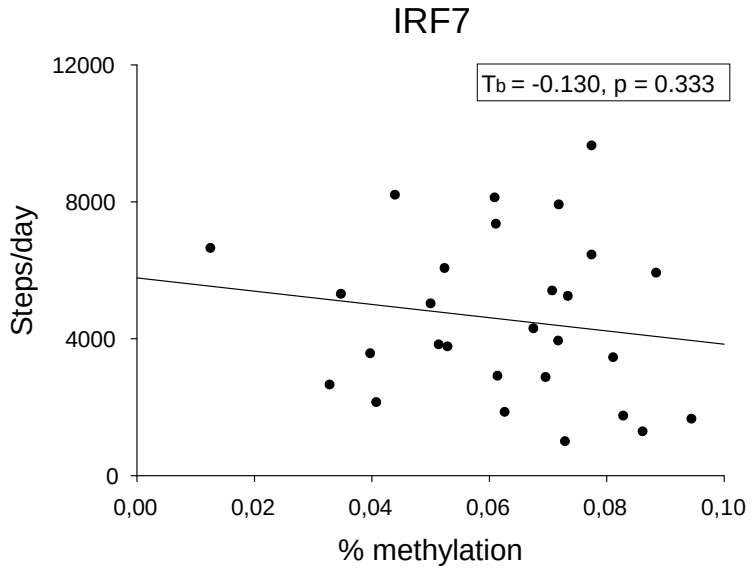


Figure 4.9: Relationship between *IRF7* methylation levels and steps per day for the CFS group.

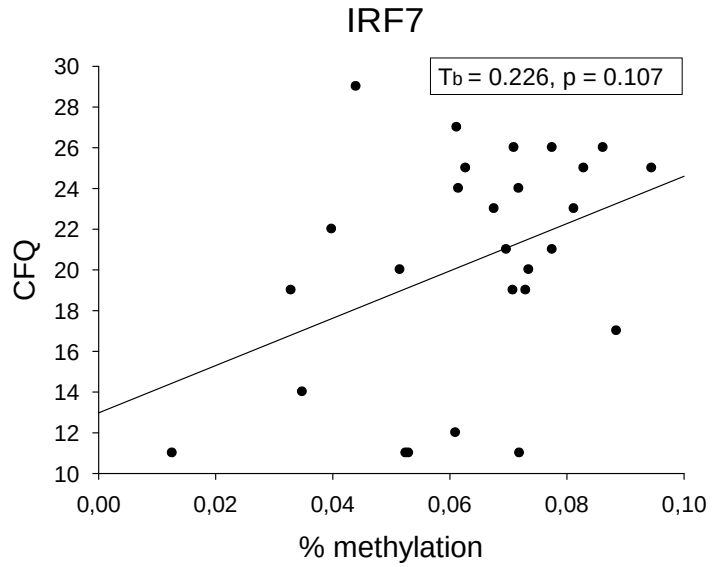


Figure 4.10: Relationship between *IRF7* methylation levels and CFQ for the CFS group.

4.3.2 NR3C1

When examining the distribution of *NR3C1* methylation data, two outliers in the CFS group with values of 1.360 and 3.2595, as well as one outlier in the HC group with a value of 0.5879 were detected using the 2.2 IQR rule (figure 4.11). These values were excluded from further analyses.

A Shapiro-Wilk's test ($p>0.05$) showed that the % methylation values of *NR3C1* were normally distributed for HCs, with a skewness of 0.624 (SE=0.501) and a kurtosis of 0.021

(SE=0.972). However, % methylation values of the CFS patients were not considered normally distributed (Shapiro-Wilk's test, $p<0.05$), with a skewness of 0.885 (SE=0.427) and a kurtosis of 0.330 (SE=0,833).

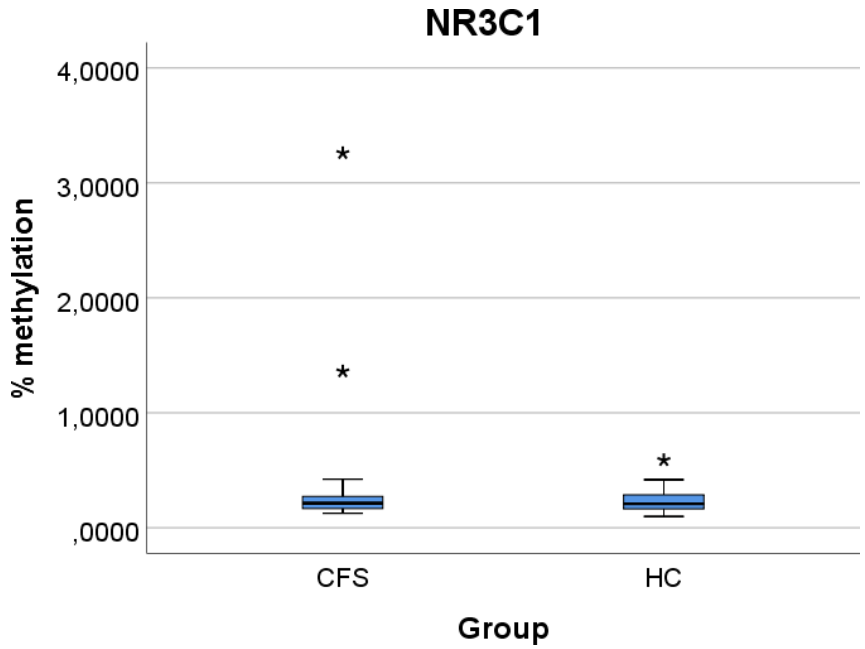


Figure 4.11: Boxplot of the % methylation of NR3C1. Two outliers for the CFS group and one outlier for the HC group are marked (*).

A boxplot of the % methylation of *NR3C1* in the two groups is shown in figure 4.12. The mean value for the CFS group was 0.229, with a SEM of 0.0138. The mean value for the HC group was 0.225, with a SEM of 0.0179. As the CFS data was considered non-normally distributed, the Mann-Whitney U test was performed to compare the two groups. There was found no significant difference in *NR3C1* methylation level between CFS and HC (p=0.833)

There was also found no significant difference when corrected by gender (males p=0.288 and females p=0.706). This is shown in figure 4.13.

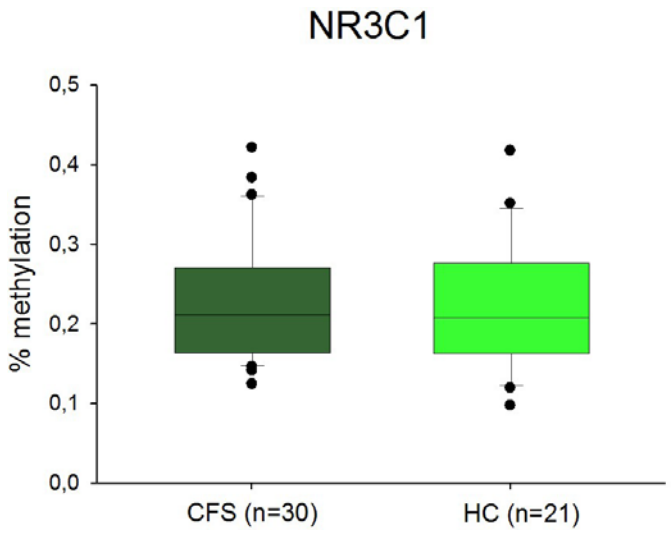


Figure 4.12: Methylation levels of the *NR3C1* gene. % Methylation in CFS patients (CFS) and healthy controls (HC). P=0.833, Mann-Whitney U-test.

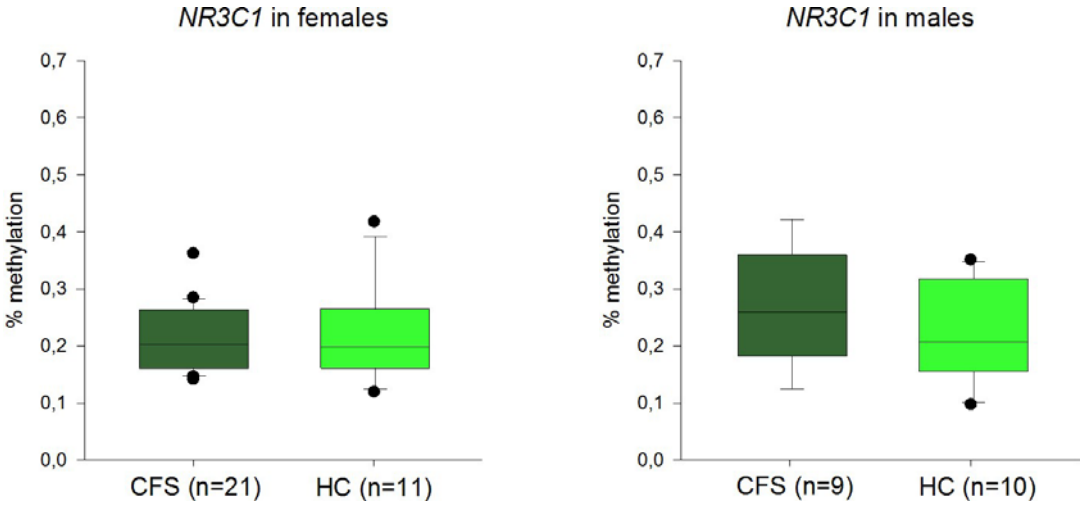


Figure 4.13: Methylation levels of *NR3C1* in females and males. % Methylation in CFS patients (CFS) and healthy controls (HC) separated by gender; Females (p=0.706, Mann-Whitney U-test) and males (p=0.288, Mann-Whitney U-test).

Kendall's tau-b correlation analyses were conducted in order to determine the relationship between *NR3C1* methylation levels and the level of fatigue in CFS patients. As shown in figure 4.14, there was found a negative correlation between *NR3C1* methylation levels and steps per day, which was statistically significant (Tb = -0.281, p= 0.032). There was also found a weak non-significant positive correlation between *NR3C1* methylation levels and CFQ (Tb = 0.144, p= 0.293) as shown in figure 4.15.

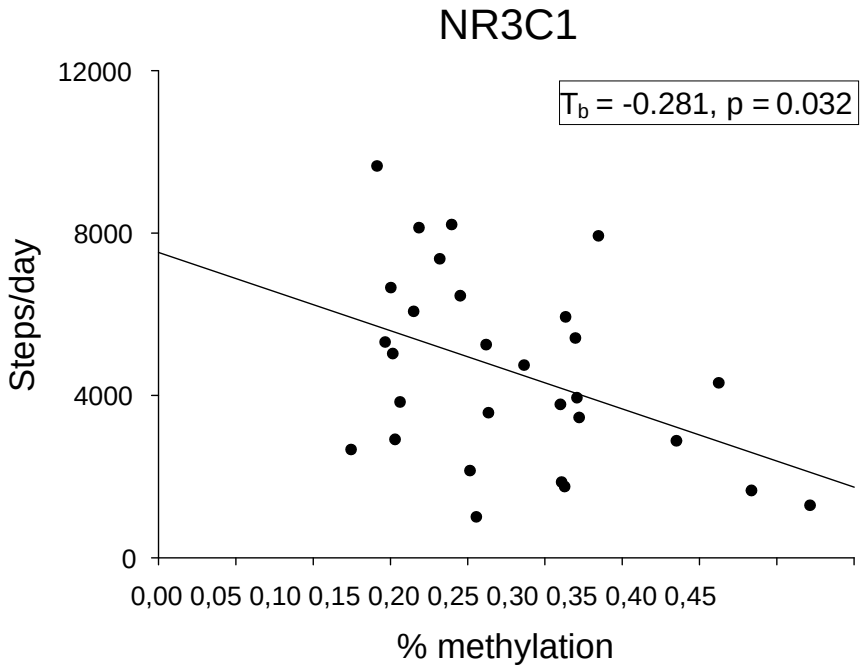


Figure 4.14: Relationship between *NR3C1* methylation level and steps per day in the CFS group.

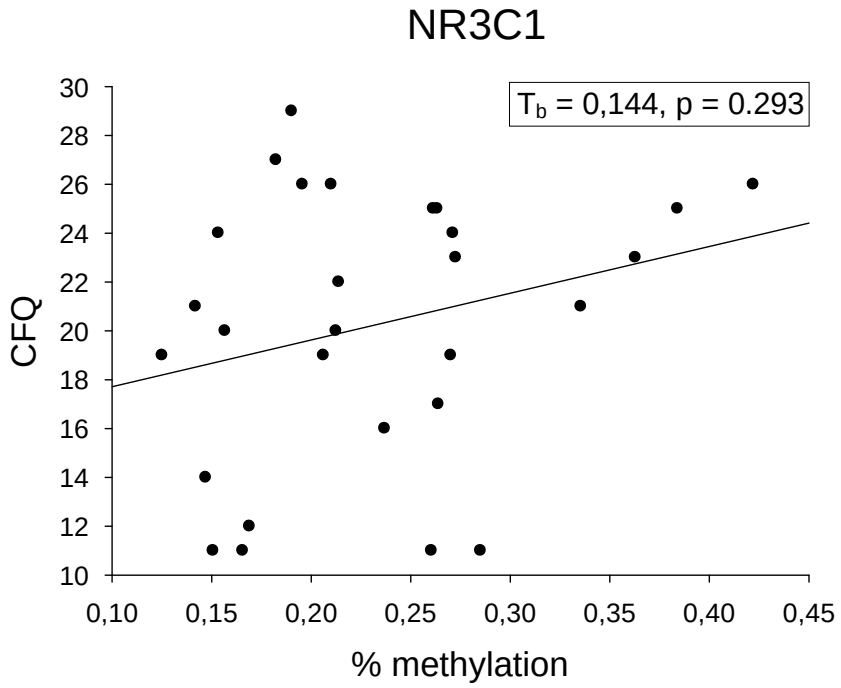


Figure 4.15: Relationship between *NR3C1* methylation level and CFQ in the CFS group.

5 Discussion of methods

5.1 Study design

This study was part of the NorCAPITAL-project. A 'broad' case definition of CFS was applied to this project, requiring 3 months of unexplained, disabling chronic/relapsing fatigue of new onset. No other accompanying symptom criteria were required from patients, but exclusion criteria encompassed somatic and psychiatric co-morbidity, pharmaceutical usage and being bed-ridden. Although wide inclusion criteria may obscure results relating to a subgroup, previous studies of the NorCAPITAL data set do not suggest subgroup differences [106; 125]. The broad inclusion criteria are in agreement with clinical guidelines and previous studies from our group [119-121].

As this is a cross-sectional study where all data are measured at the same time point, it is not possible to draw valid conclusions regarding causality of associations that are found. However, it opens for the study of multiple outcomes and exposures, and it can be useful for generation of further hypotheses.

This study holds a candidate gene approach as opposed to a genome-wide approach. This means that specific genes and miRNAs are chosen and analyzed based on previous studies and knowledge of the phenotype or condition that will be investigated. The advantage of this is that it is relatively quick, easy and cost-effective compared to a genome-wide approach. It is also a good way to test hypotheses and validate previous results. However, this approach is largely limited by its reliance on existing knowledge about the known or assumed biology of the relevant phenotype or disease. It has also been criticized of low replication of results and a limited ability to include all possible causative genes [130]. Genome-wide scanning, on the other hand, usually proceeds without presuppositions regarding the importance of specific functional features of the phenotype or disease, but will usually only give a glancing overview and a complex result with a large number of candidate genes which may be more difficult to understand and analyze. In comparison, the candidate gene approach is very powerful for studying the genetic architecture of complex traits, and is a far more effective and economical method for direct gene discovery.

5.2 Sample material

5.2.1 Serum

In the present study, miRNA levels were analyzed in serum samples. miRNAs have been found to stay remarkably stable in serum for extended periods of time [66]. Previous studies have shown that serum can be stored at -20°C for at least 2-4 years without a significant decrease of miRNA levels [40], and they can probably be stored for even longer time periods at -80°C [35]. They are also resistant to RNase digestion, boiling, pH changes, freeze-thaw cycles etc. [19]. However, it has been shown that the stability of miRNAs in serum may not be as robust as in plasma [35], and that the stability of individual miRNAs in serum may vary [58]. Despite of this, serum is usually the preferred circulating fraction for miRNA detection as plasma contains cellular material that may introduce contaminating cellular miRNAs, as well as anti-coagulants that may inhibit downstream methodologies.

Compared to cell samples, serum samples are easier to harvest and process, clear advantages if the samples in the future will be used for prognosis and/or diagnosis of patients. Because of their robustness, circulating miRNAs have also been proposed as potential biomarkers for disease diagnosis and prognosis [11]. However, there is large inconsistency in miRNA expression when circulating miRNA profiles obtained from different laboratories are directly compared for the same disease [35]. This may be explained by differential sample handling of serum, which may lead to inconsistent and incomparable results.

This study was based on results from RNA-seq analysis which was conducted on whole blood samples. It is therefore reasonable to study circulatory miRNAs as possible regulatory elements of differentially expressed genes, as opposed to miRNAs from other specific tissues.

Previous studies on miRNA expression in CFS have used lymphocytes or plasma as sample material [11; 12; 94]. If possible, we would have analyzed lymphocytes (PMBCs) as well, but this material was not available for analysis. In theory, miRNAs present in plasma and serum should be relatively similar as both contain the same chemical structures, except that serum lacks the clotting factors. However, previous studies have found concentration differences of several miRNAs in serum and plasma, possibly owing to the release of specific miRNAs from red blood cells and platelets during the coagulation process [117]. Because of this, caution

should be taken when comparing the miRNA data generated from this study to previous studies using different sample types.

5.2.2 Whole blood

In the present study, whole blood was used as sample material in the DNA methylation analyses. This is reasonable seeing as also these analyses were based on results from RNA-seq in whole blood samples. Whole blood DNA is also a convenient material to use for analyses, as its collection is noninvasive.

Blood is a commonly used material in epigenomic studies, although its heterogenous nature may lead to interpretation difficulties [48]. Cell composition effects due to the variation in constituent leukocytes in different individuals and at different time points may confound the analyses and lead to inconsistency in the results. Epigenetic state may also be very different in blood compared to different tissues throughout the body. Because of this, it would be preferable to analyze the target tissue of the specific disease to be studied. In CFS, however, there is no clear target tissue although there is increasing evidence that several different immune cells may be affected by the disease [13; 37; 73; 84].

In previous studies on DNA methylation in CFS, PBMCs and whole blood has been used as sample materials [26; 27; 115].

5.3 Gene expression analysis

5.3.1 RT-qPCR for validation of RNA-seq results

RT-qPCR was chosen to validate the *FLT3* and *IL1RN* expression level results found by RNA-seq [84]. RT-qPCR is a widely used technique with high accuracy, sensitivity and rapid execution, which enables the detection of minimal starting amounts of specific nucleic acids [25]. However, this sensitivity leads to the method being subject to a high degree of technical variation that may bias the data. Several steps, including the quality of the mRNA, amplification efficiency and the choice of reliable internal controls (reference genes) may affect the accuracy and interpretation of the results.

Alternatives to the RT-qPCR method include Serial analysis of gene expression (SAGE), Northern and Southern blot [97]. However, these methods require large amounts of input mRNA, and compared to RT-qPCR the SAGE technique is more time-consuming and requires multiple manipulations of the samples, increasing the risk of contaminating the samples. RT-qPCR has for some time been considered the most appropriate method to confirm sequencing- or microarray-generated data. Recently, droplet digital PCR (ddPCR), an approach using water-in-oil droplets combining with high-throughput digital PCR has been developed [46]. In ddPCR, a single sample generates tens of thousands of data points, increasing the power of statistical analysis compared with RT-qPCR. However, ddPCR is a relatively expensive alternative to RT-qPCR.

5.3.2 Normalization method

In gene expression analysis, normalization of the data has proven essential to ensure accurate inference of expression levels [99]. It is a crucial step necessary to minimize the influence of confounding factors. Internal reference genes are often used for normalization of data. To be regarded a reliable reference gene, it must meet several criteria [59]. Most importantly, its expression level must be unaffected by experimental factors. It should also show minimal variability in its expression between tissues and physiological states of the organism. An optimal reference gene should also give a similar ct to the gene of interest. Improper selection of reference genes may result in inaccurate calculations and obscure the biological differences among samples [70].

For normalization of RT-qPCR data in the present study, the housekeeping gene *GAPDH* was used. *GAPDH* is one of the most commonly used reference genes, although its suitability as an internal reference has been questioned because of reported variability in its expression [16]. However, in this study, there was not found a significant difference in *GAPDH* expression levels between the groups (Figure 5.1). *GAPDH* was also the reference gene that was used during validation of other DEGs from RNA-seq [84].

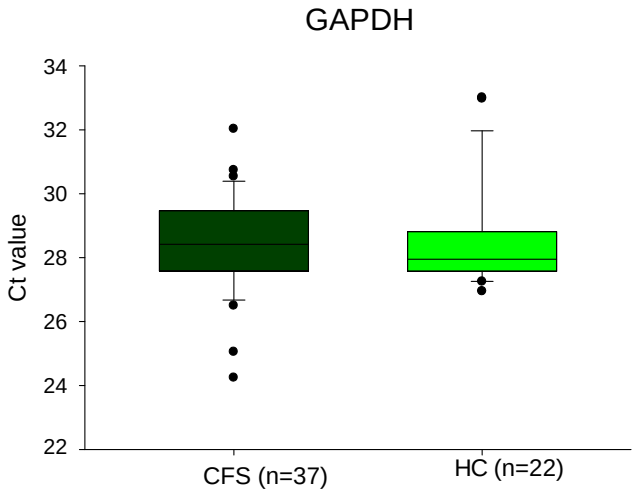


Figure 5.1: *GAPDH* levels from RT-qPCR. Ct values in Chronic fatigue syndrome patients (CFS) and healthy controls (HC). There was found no significant difference between groups (p=0,326, Mann Whitney U-test).

In RNA-seq, normalization of raw read quantification and removal of variation prior to differential expression analyses were processed following the 'remove unwanted variation g' (RUVg) method [84]. This normalization strategy adjusts for nuisance technical effects by performing factor analysis on suitable sets of control genes [99], and may lead to more accurate estimates of expression fold-changes and tests of differential expression compared to normalization methods using single reference points.

There is a possibility that differences in normalization procedures between the two methods partly explain why we were unable to validate the results from RNA-seq using RT-qPCR.

5.4 miRNA analysis

5.4.1 RT-qPCR for miRNA analysis

In this study, the RT-qPCR method was chosen to analyze the expression levels of specific miRNAs, using SYBR-based miRNA assays. As described above, SYBR-based assays rely on the enhanced fluorescence of SYBR Green dye when it intercalates into the double stranded DNA PCR product. An alternative to this would be TaqMan® assays, which uses a specific probe and the fluorescent resonance energy transfer (FRET) system [18]. A disadvantage of SYBR-based assays is the production of false-positive signals from nonspecific reaction products that may be caused by primer dimers or DNA contamination.

TaqMan® based detection is generally considered more specific and reproducible than SYBR-based detection, but it requires the synthesis of different probes for each sequence and is therefore a more costly alternative.

The short sequence of miRNA nucleotides makes it difficult to design conventional RT-qPCR primer assays. This challenge is solved by the addition of a universal primer sequence which is incorporated into the RT primer to extend the length of the cDNA molecule during reverse transcription [18]. The cDNA is then amplified with a miRNA-specific primer on one end and a universal primer on the other end in real-time PCR. Other challenges in miRNA detection and quantitation include the sequence similarities of miRNAs, and the existence of pri-miRNAs and pre-miRNAs in addition to the mature miRNAs. This makes it especially important to use specific primers that will discriminate between the different miRNAs and their immature forms, and is the reason why we chose to use pre-designed primer sets that have already been validated for our target miRNAs instead of designing our own primers.

5.5 DNA methylation analysis

5.5.1 Locus-specific methylation analysis

Several methods can be used to analyze locus-specific methylation, including bisulfite-based PCR, restriction-enzyme based analysis and immunoprecipitation of methylated fragments [20]. Methylated DNA immunoprecipitation (MeDIP) is an efficient and cost-effective method in which methylated DNA fragments are enriched by an antibody which recognizes and binds 5mC. This allows for the analysis of both CpG and non-CpG methylation. MeDIP products can then be used for PCR analysis of individual loci, or genome-wide analysis.

There are several commercial kits available for MeDIP analysis. Prior to this study, several different kits and protocols were tested to find the most optimal kit for our analyses.

Methylated DNA Immunoprecipitation (MeDIP) Kit – DNA (Abcam, Cambridge, Great Britain) was the first kit to be tested. With this kit, methylated DNA is enriched using a 5-mC specific antibody to immunoprecipitate methylated DNA. The main problem with this kit was that it did not give a sufficient product yield to get a good enough detection on qPCR. A qPCR plot from one of the test runs with this kit is shown in figure 5.2. Methylated DNA

(meDNA) samples have very high ct values of 31-44, even for *XIST* and *LAP3* which are positive methylated controls. Usually, a ct value higher that 35 is considered a negative result. Input samples have ct values ranging from 27-29, which are also high compared to expected values. *GAPDH*, which is considered a housekeeping gene, was, according to the manufacturer, expected to give ct values around 20 in input samples. In addition, the curves are not smooth and parallel. There were done several different attempts to identify the reasons for this problem and to solve them, as shown in Table 5.1. Several test-runs of qPCR were done in attempt to optimize the conditions. However, we were still not able to get sufficient yield from this kit.

Next, we attempted to optimize a microChIP protocol described by Dahl and Collas [23]. This protocol was originally meant for studying protein-DNA interactions on DNA isolated from cells using magnetic beads and immunoprecipitation with an antibody. We attempted to alter the conditions and optimize the protocol to be able to immunoprecipitate methylated DNA from whole blood using a 5-mC antibody. However, even after several attempts of optimization and troubleshooting as shown in Table 5.1, we were not able to get any better results than with the MeDIP kit from Abcam (Figure 5.3).

Next, we tried the Novex Immunoprecipitation (IP) kit Dynabeads Protein G (Thermo Fisher Scientific, MA, USA). This kit is designed for immunoprecipitation of the desired antigen using protein G magnetic beads and a suitable antibody. Different input amount and 5-mC ab from two different manufacturers were tried (Table 5.1), but we did still not get a good enough detection on qPCR for the meDNA samples as shown in figure 5.4. To make sure that the problem was not the primer set for *XIST*, two different primer sets for another methylated gene (*MLH1*) were tried out, but these did not give any better results, as shown in figure 5.5.

The microchip protocol and the IP kit were not originally optimized towards methylated DNA, and this may explain why we were unable to get satisfying results using these protocols. The MeDIP kit from Abcam, on the other hand, was specifically made for enrichment of methylated DNA and should have generated a yield of meDNA applicable for qPCR analysis. According to our troubleshooting, it is not likely that the type or amount of input material was the underlying problem; neither were the sonication conditions or the sonication instrument. Changes in washing and centrifugation steps as well as different input and primers for the qPCR reaction were also attempted without any improvement in the

results. Finally, we can only speculate whether there could be something wrong with any of the components in the specific kit that we used.

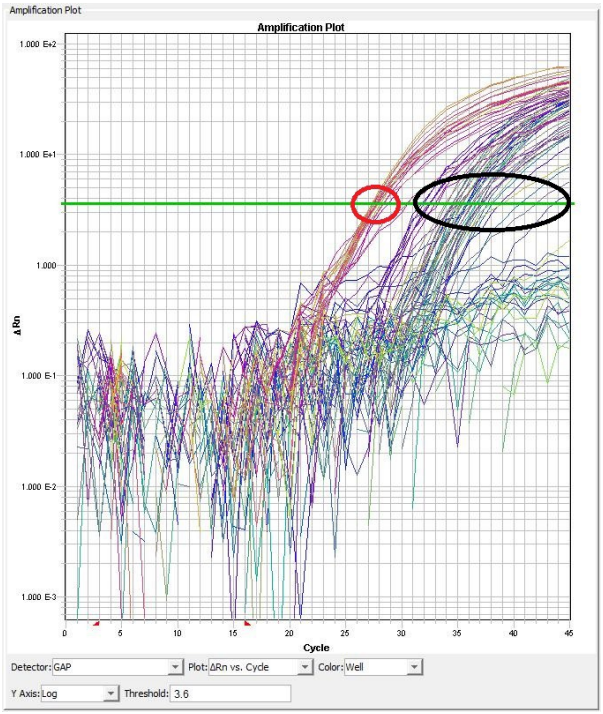


Figure 5.2: A qPCR plot from a test run using the MeDIP kit from Abcam. The black circle shows meDNA samples, with ct values ranging from 31-44. The red circle shows input samples, with ct values ranging from 27-29.

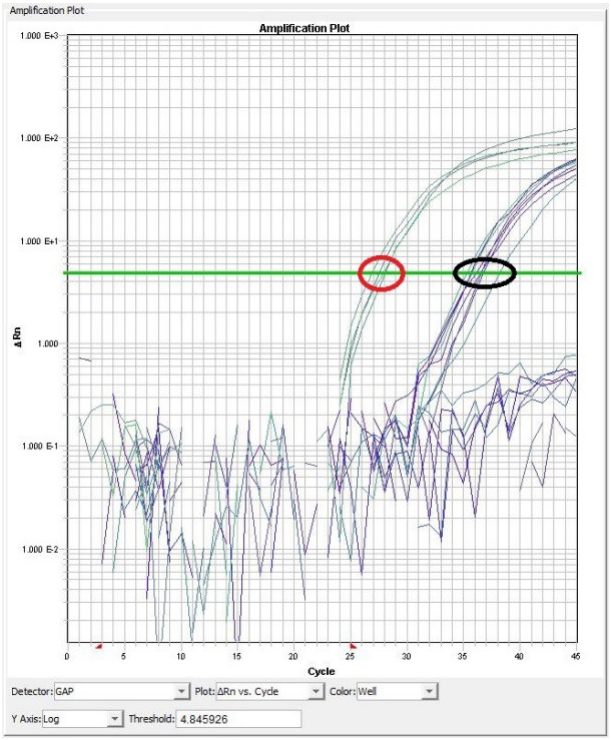


Figure 5.3: A qPCR plot from a test run of the *XIST* gene using an optimized microChIP protocol. The black circle shows meDNA samples, with ct values ranging from 35-38. The red circle shows input samples, with ct values ranging from 27-28.

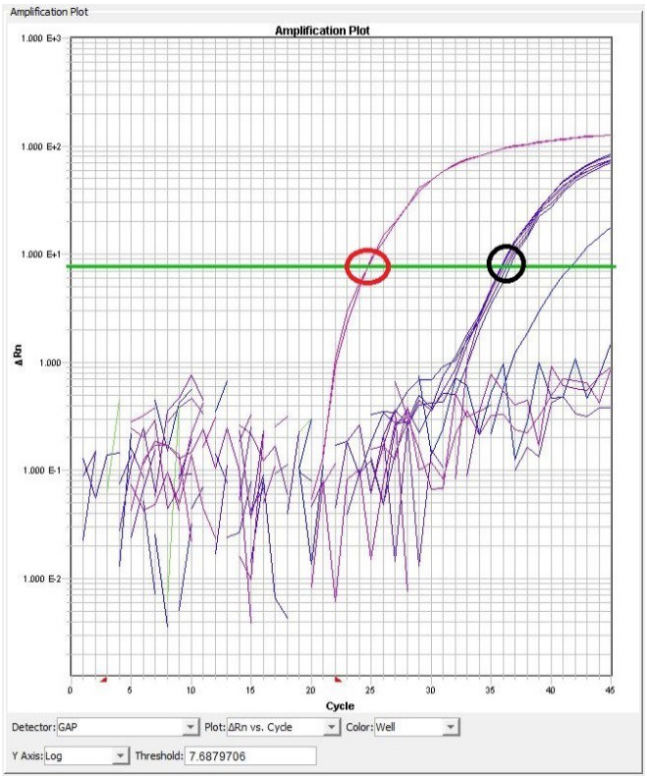


Figure 5.4: A qPCR plot from a test run of the *XIST* gene using the IP kit from Thermo Fisher Scientific. The black circle shows meDNA samples, with ct values between 36-37. The red circle shows input samples, with ct values of approximately 25.

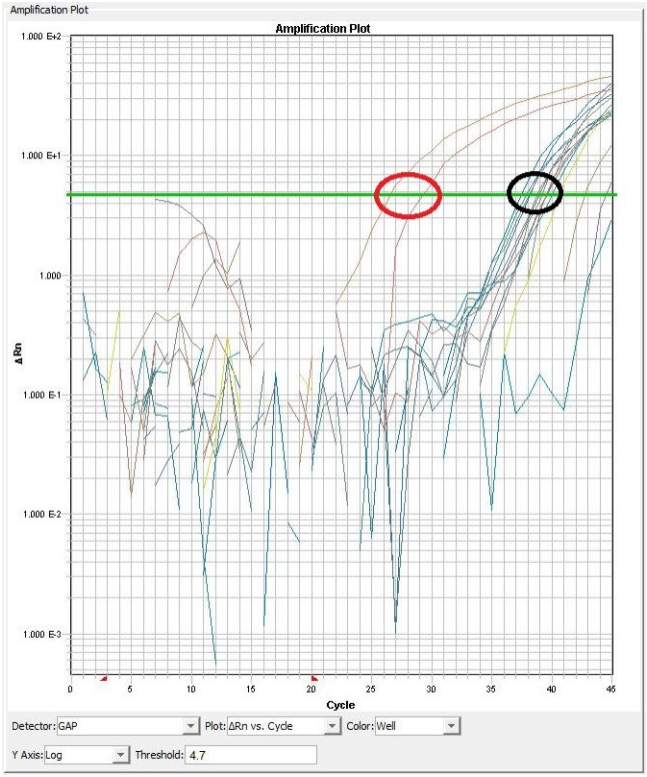


Figure 5.5: A qPCR plot from a test run of the *MLH1* gene using the IP kit from Thermo Fisher Scientific. The black circle shows meDNA samples, with ct values of 37-40. The red circle shows input samples, with ct values between 26 and 30.

Table 5.1: Kits and protocols tested for MeDIP analysis.			
Kit/protocol	**Possible error source**	**Solution attempt**	**Details**
MeDIP kit (Abcam)	DNA input material	*Isolate DNA from different types of tissue.	*DNA from PBMC and buccal cells in addition to whole blood.
	DNA input amount	*Different amounts of DNA input.	*1,5 µg/100 µl and 3 µg/100ul of DNA input as well as 1 µg/100ul as suggested from the manufacturer
	Sonication conditions	*Different sonication instruments *Different sonication conditions	*Bioruptor UCD-300 (Diagenode, Liège, Belgium) and E220 evolution Focused-ultrasonicator (Covaris, Woburn, MA, USA) * Tried cycles with a duration of 30 sec on/off and 15 sec on/off * Tried 10, 15 and 30 cycles * Tried sonication tubes and normal Eppendorf tubes
	Insufficient washing	*Additional washing steps * Increase duration of centrifugation steps	
	qPCR input	* Different amounts of qPCR input * Amplification of meDNA	* Input of 2 µl as well as 1 µl as suggested by the manufacturer * Amplification of meDNA before qPCR run using GenomePlex Single Cell Whole Genome Amplification Kit (Sigma-Aldrich, MO, USA)
	qPCR primers	*Several primers for different genes *Examine melting curves of primers	* Primers for *LAP3, IRF7, GAPDH, NR3C1, IL1RN, XIST, HK3, KCNJ15, CD79A, TNFRSF13C, MLH1* and *STAT6*
microChIP protocol	DNA input amount	*Different amounts of DNA input.	*1 µg and 500 ng of DNA input
	Dilution buffer	*Different dilution buffers.	*Nuclease free water and RIPA chip buffer
	5mC antibody	*Different antibody	*5-mC antibody from Abcam and Diagenode
IP kit (Thermo Fisher Scientific)	DNA input amount	*Different amounts of DNA input.	*500 ng/100 µl, 1µg/100µl and 2µg/100µl
	5mC antibody	*Different antibody *Antibody amount	*5-mC antibody from Abcam and Diagenode. * 2.5 µg, 5 µg and 10 µg
	qPCR primer	* Different primer set for positive methylated control gene	* Two different primer sets for the MLH1 gene

Finally, we ended up using the MagMeDIP kit from Diagenode. This is the same kit that Brebi-Mieville et al [10] observed to give the highest yield compared to two other MeDIP kits. The protocol for this kit was followed relatively precise, except for the fragmentation step, where the protocol recommended shearing the DNA using a Bioruptor sonication device for 10 cycles of 15 sec ON/OFF. Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA) was used to investigate fragment lengths after sonication with Bioruptor UCD-300 (Diagenode, Liège, Belgium). Figure 5.6 and 5.7 show fragment lengths resulting from sonication with different conditions. Figure 5.6 shows the different fragment lengths obtained from sonication for 10 cycles of 15 sec ON/OFF as suggested by manufacturer, and for 15 cycles of 30 sec ON/OFF. Because the optimal range of fragment lengths for MagMeDIP is between 200 and 1000 bp, we concluded that 15 cycles of 30 sec ON/OFF seemed to give a better range of fragments than 10 cycles of 15 sec ON/OFF. However, when we tried to confirm this hypothesis, we got the result shown in figure 5.7. This figure shows that fragments of very different length were obtained for the same DNA sample using the same sonication conditions. Based on this result we concluded that the Bioruptor UCD-300 was not trustworthy enough for our analyses, and our samples were instead sheared at a core facility using a E220 evolution Focused-ultrasonicator (Covaris, Woburn, MA, USA), resulting in fragments of the desired length for the following MagMeDIP procedure.

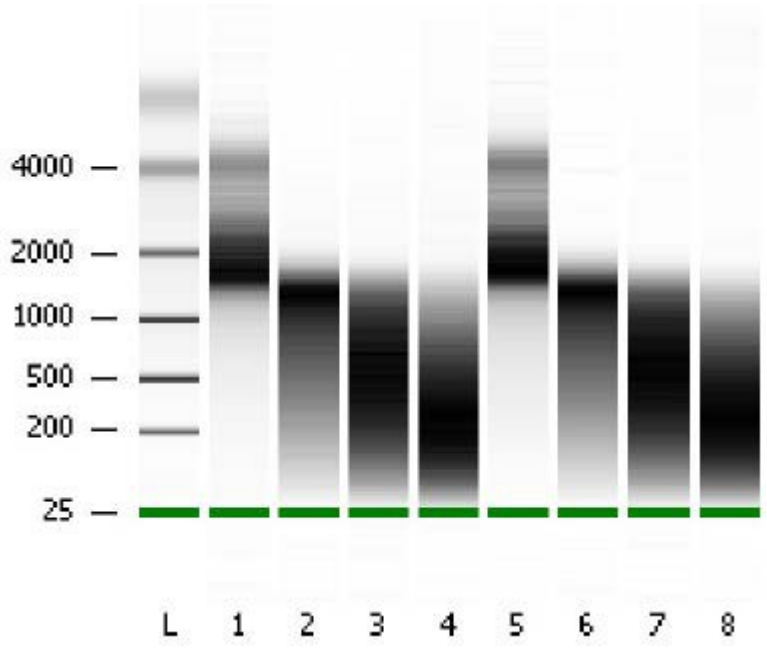


Figure 5.6: A Bioanalyzer run after sonicating DNA with Bioruptor UCD-300. L: Ladder. 1 and 5: Test sample 1; 10 cycles of 15 sec ON/OFF. 2 and 6: Test sample 2; 10 cycles of 15 sec ON/OFF. 3 and 7: Test sample 3; 15 cycles of 30 sec ON/OFF. 4 and 8: Test sample 4; 15 cycles of 30 sec ON/OFF.

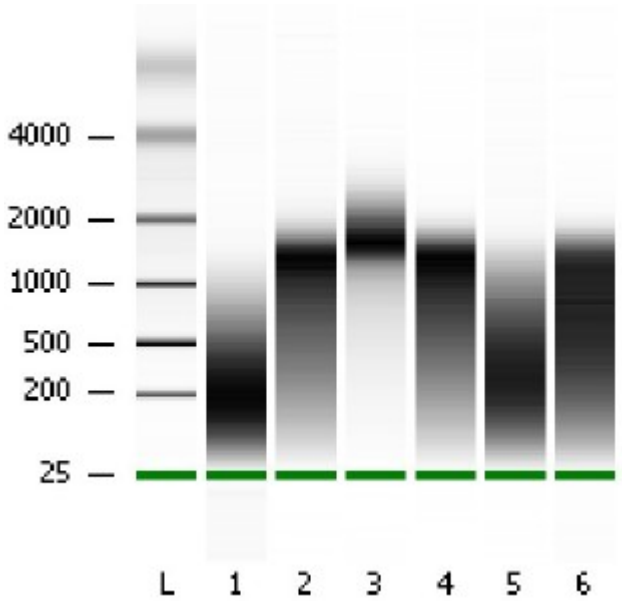


Figure 5.7: A Bioanalyzer run after sonicating DNA with Bioruptor UCD-300. L: Ladder. 1 and 2: Test sample 1; 15 cycles of 30 sec ON/OFF. 3 and 4: Test sample 1; 25 cycles of 15 sec ON/OFF. 5 and 6: Test sample 2; 15 cycles of 30 sec ON/OFF.

5.5.2 Primer design

Primers were initially 'randomly' designed in the promoter region of each gene of interest, choosing primer sequences at a random area in the promoter. For *IRF7* and *CD97A*, this resulted in primers flanking relevant regions for DNA methylation analysis as shown in appendix 10 and appendix 11, respectively. For *NR3C1*, however, expression is controlled by a quite complex mechanism of several first exons with corresponding promoter regions. This means that it is important to design the primers in the promoter region which corresponds to the first exon of interest. Figure 5.8 shows the *NR3C1* structure and the region containing the primer sequences that were used in this study to analyze *NR3C1* methylation. As shown in the figure, the primer sequences are located in the proximal promoter region regulating expression of first intron 1A. However, the promoter region that was previously reported to be relevant in CFS [115] is regulating the first intron 1F as marked in the figure.

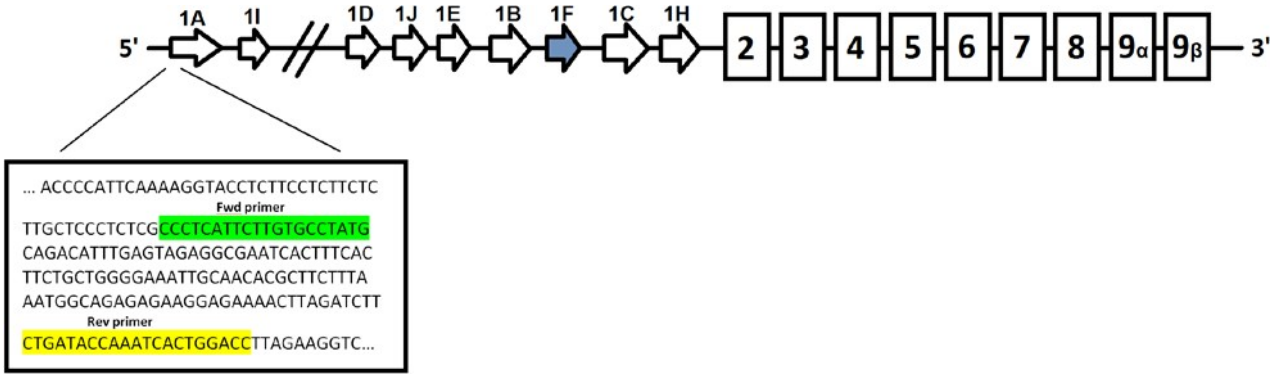


Figure 5.8: *NR3C1* structure and primers used for qPCR. Representation of *NR3C1* exon 2 through 9 marked with boxes, and its alternative first exons marked with arrows. Indicated in the text box is part of the genomic sequence in the distal promoter region, and primers used for qPCR (forward primer in green, reverse primer in yellow). The 1F region is marked in blue. The figure is not to scale and based on Turner et al. [113] and Vangeel et al. [115].

6 Discussion of results

6.1 Gene expression analysis

In the present study we attempted to, by using RT-qPCR, validate results from RNA-seq which indicated gene expression differences between CFS patients and HCs for the genes *FLT3* and *IL1RN*. We were unable to find significant differences in the gene expression levels between the two groups. However, there was found a trend towards higher expression of *IL1RN* in CFS patients compared to HCs. This corresponds to the RNA-seq results for this gene, which showed significantly higher expression of *IL1RN* in CFS patients compared to controls (Appendix 1: Table S1).

The *IL1RN* gene product IL-1ra acts as a potent anti-inflammatory molecule. Higher levels of this molecule in CFS patients would fit with the hypothesis that CFS patients have a "shift" in their immunity, from less specific immune responses and towards more innate immune responses. However, translation of mRNA into protein may be independently regulated, and because of this, changes in mRNA expression only suggest a change in the level of the protein. Wyller et al [125] analyzed IL-1ra levels of the same patient group, using a multiplex cytokine assay on plasma samples. There was not found any differences in IL-1ra levels between CFS patients and HCs.

The poor correspondence between RNA-seq results and RT-qPCR results for the *FLT3* gene may be explained by different primers and different normalization methods between RNA-seq and RT-qPCR. There was also a low concentration of remaining cDNA after RNA-seq, which may have led to less accurate results in RT-qPCR. Because of the small amount of remaining cDNA sample, we were forced to use duplicates instead of triplicates in the qPCR reaction. As not all duplicates were even, we had to remove several data points from the analyses, leading to weaker statistical power of our analyses. Another consequence of the low sample amount was that we were not able to validate our experiments by repeating them.

6.2 miRNA analysis

In this study, we analyzed expression levels of specific miRNA to investigate whether these were different between CFS patients and controls. None of the investigated miRNAs were

found to be significantly differentially expressed in CFS patients compared to controls, although previous studies have shown differential expression of some of them. However, earlier studies on miRNAs in CFS have also been inconclusive. This could be explained by small sample sizes, inappropriate selection of controls, failure to control for confounding factors, different inclusion criteria, or just the fact that CFS is a heterogeneous disease with varying symptoms and severity. Different sample material was also used in the different studies. It is, however, important to note that previous studies on miRNA expression in serum have given inconsistent results in general [35], not only in CFS research. This may, as mentioned above, be because of different handling of the sample material, or different normalization methods.

The trend towards up-regulation of *IL1RN* in CFS patients led to the hypothesis that this rise in expression may be explained by differential regulation of miRNAs. miR-223-3p was a predicted regulator of the *IL1RN* gene and was therefore investigated. miR-223-3p has previously been shown to be down-regulated in CFS patients, although this was in NK-cells, not serum samples. Assuming that this miRNA regulates *IL1RN* in a negative fashion, leading to less expression of *IL1RN*, we would expect to see less expression of miR-223-3p also in our CFS group compared to the HCs. However, we found no differences between the two groups.

miR-142-5p regulates three of the genes that were found to be differentially expressed in RNA-seq; *TSPAN3* and *SLC38A1*, which were up-regulated, and *DDX5*, which was down-regulated. As miR-142-5p has been found to be significantly up-regulated in plasma of CFS patients [11], we expected to see a higher expression of this miRNA in our patient group compared to HCs. No significant difference was found between groups.

Another miRNA that previously have been found to be up-regulated in plasma of CFS patients [11], is miR-143-3p. This miR regulates *SECISBP2L*, which was down-regulated in CFS patients in RNA-seq. Our hypothesis was that this miRNA would be up-regulated in CFS patients, but we found no significant difference between the two groups.

miR-486-3p regulates *BCL7A*, which is down-regulated in CFS patients, and *OGT*, which is up-regulated in CFS patients. No significant difference was found in CFS patients compared to HCs for this miRNA.

miR-146a-5p regulates *IRF7*, which is an important regulator in the IFN pathway. This miRNA was in a previous study found to be down-regulated in NK-cells of CFS patients [12]. We did not find any significant difference in miR-146a-5p expression levels between CFS patients and HCs.

6.3 DNA methylation analysis

In the present study, we analyzed DNA methylation levels of two genes, *IRF7* and *NR3C1*. We wanted to investigate whether these levels were different in CFS patients compared to HCs, and whether methylation level corresponded to the level of fatigue. Methylation levels of *IRF7* were found to be significantly higher in CFS patients compared to HCs. For *NR3C1*, no significant difference was found in methylation level between the two groups. Analysis of a third gene, *CD79A*, was attempted, but we did not have time to find optimal primers for the qPCR analysis of this gene.

We found significantly higher methylation levels of the promoter region of the *IRF7* gene. Hypermethylation of the *IRF7* promoter may result in reduced expression of *IRF7* [74]. As neither expression levels nor protein levels of *IRF7* have been analyzed in this data set, it is not possible to conclude whether *IRF7* expression is affected by this hypermethylation or not. However, the result does indicate that this gene may play a part in CFS and may be interesting to study further.

The IFN pathway, in which *IRF7* has a central function, plays a critical role in the immune response following viral infection. Infections of viral agents such as EBV are a suggested cause or trigger for CFS [8]. *IRF7* is also associated directly with EBV infection and latency. As other important genes in the IFN pathway (*IRF9*, *IRF4* and *TLR8*) have been found to be up-regulated in CFS patients, the hypermethylation of *IRF7*, if resulting in reduced expression of *IRF7*, could represent a compensatory mechanism to retain homeostasis in this pathway.

IRF7 remains at low levels in most cell types, but it is not clear whether promoter methylation is an important mechanism in keeping the expression at a low level. *IRF7* stability is tissue specific, and studies have shown that *IRF7* expression is specific for lymphocytes [86]. Thus, it would be interesting to study *IRF7* promoter methylation in more specific cell populations of lymphocytes.

Although the *IRF7* promoter methylation levels are significantly higher in CFS patients, the levels are relatively low, with a mean value of 0.07% compared to 0.05% for the HCs. This raises the question of whether these differences are functionally relevant or not. However, although methylation levels of *IRF7* are low, they are significantly higher (0.56%±0.05%) than what was detected for *GAPDH*, the negative control gene (0.24%±0.05%).

When corrected by gender, the significant difference disappears for females (p=0.093), but not for males (p=0.035). However, the pattern is similar for both genders, with higher methylation levels in CFS. Although global methylation levels previously have been found to be significantly lower in women compared to men [128], the differences in this analysis are not large enough to indicate that *IRF7* promoter methylation is different between men and women with CFS.

Melting curves can be used to ensure that only the desired amplicon is detected and to indicate potential issues in the qPCR reaction [98]. Melting curves are constructed immediately after PCR; the annealed products are melted at a constant rate, and the decrease in fluorescence is monitored as the strands dissociate. Optimally, the melting curve should consist of one, distinctive peak indicating the melting point of the desired amplicon. The melting curves from the *IRF7* qPCR run is shown in figure 6.1. The curves show a main peak at approximately 90°C, and several smaller peaks at lower temperatures. Because intercalating dyes like EvaGreen will detect any double-stranded DNA, extra peaks in melting curves may represent primer dimers, contaminating DNA, single nucleotide polymorphisms (SNPs) or extra PCR products produced due to mis-annealed primers [67]. However, the presence of extra peaks does not necessarily indicate non-specific amplification; local GC-content differences may cause uneven melting [32]. To confirm that only the desired amplicon is detected, one solution would be to use agarose gel electrophoresis. Software that helps predict melting of a PCR product also exists [32].

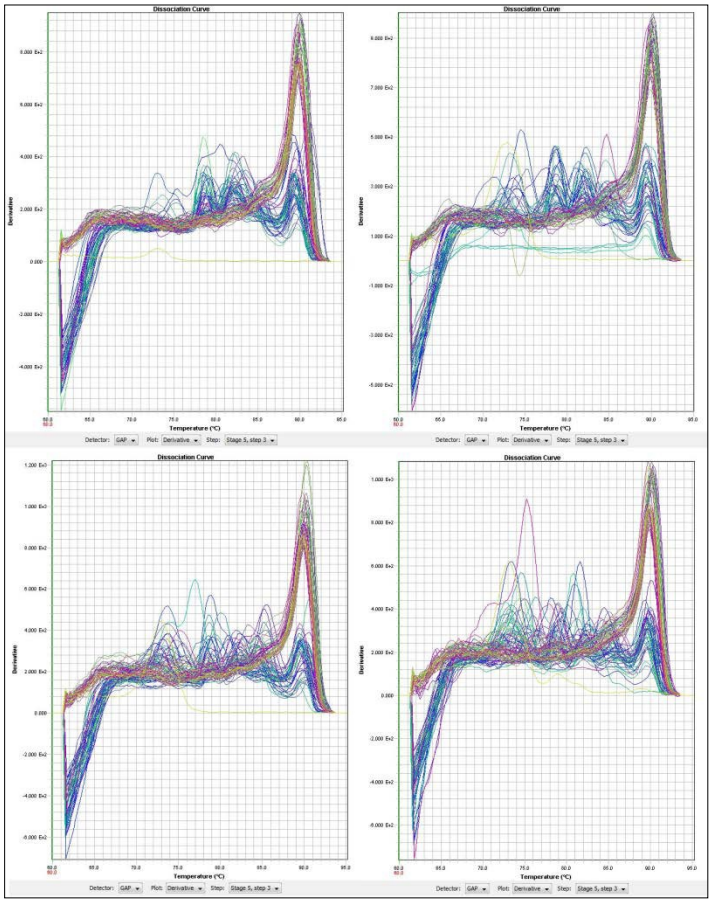


Figure 6.1: Melting curves from the *IRF7* qPCR run. A main peak is present at approximately 90°C. Some of the samples also show several smaller peaks at lower temperatures.

The GR, encoded by *NR3C1*, is an important part of the HPA axis and one of the proposed mechanisms of HPA axis function in CFS is enhanced GR sensitivity. Previous studies have found significantly lower methylation of the *NR3C1* promoter 1F region in whole blood from CFS compared to controls. We hypothesized that we would get similar results in our analysis of methylation levels of the NR3C1 promoter. However, we were only able to analyze methylation levels of the proximal promoter of the 1A region of *NR3C1*, and found no significant differences between the two groups. Little is known about this part of the gene and whether or not it is regulated by DNA methylation.

We attempted to analyze methylation levels of the promoter 1F region of the *NR3C1* gene in addition, but did not have time to find optimal primers for the qPCR analysis of this region.

To investigate possible relationships between promoter methylation and the level of fatigue, correlation analyses were conducted using CFS methylation data and markers of disability (steps per day) and fatigue severity (CFQ). For both *IRF7* and *NR3C1*, there was found a negative correlation with steps per day, but this relationship was significant only for *NR3C1*.

For both genes, there was found a positive but non-significant correlation with CFQ score. Although these correlations were not statistically significant, the positive correlation between *IRF7* and CFQ score as well as the negative correlation between *IRF7* methylation and steps per day strengthens the hypothesis that hypermethylation of the *IRF7* gene may be a feature of CFS.

There was also found a significant negative correlation between steps per day and DNA methylation of *GAPDH* – the negative control (Figure 6.2). Although DNA methylation levels are very low for this gene, this relationship may indicate that physical activity has an impact on DNA methylation at a more general level. A relationship between physical activity and genome-wide changes in DNA methylation has been shown in previous studies, although in muscle [69; 87] and adipose tissue [100]. If there is indeed a relationship between physical activity and genome-wide methylation in blood, this may also explain the significant correlation between NR3C1 and steps per day.

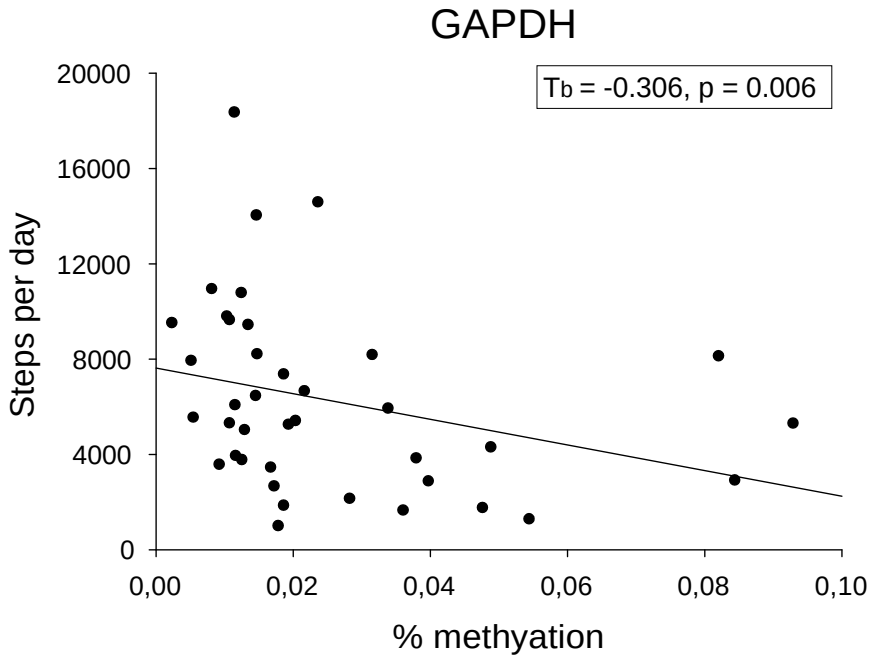


Figure 6.2: Relationship between *GAPDH* methylation level and number of steps per day.

6.4 Future perspectives

In this study, we attempted to investigate some possible mechanisms of regulation of gene expression in CFS. However, there are several other mechanisms of gene expression regulation that may play a part in CFS pathophysiology. It would be interesting to investigate some of these mechanisms as well, for instance other epigenetic mechanisms such as histone modifications.

Using a candidate gene-approach we were able to study specific genes and miRNAs based on existing knowledge, but it would also be interesting to try a whole-genome approach on our sample material to see if we could find other trends or candidate genes. For miRNA analysis, possible methods include smallRNA sequencing or a miRNA microarray. For whole genome methylation profiling, bisulfite sequencing, pyrosequencing or a methylation microarray would be possible methods.

As this is a cross-sectional study, we were not able to say anything about the causality of our results. It would be interesting to look at the same factors in a longitudinal study, preferably over several years and with baseline samples before the patients became ill. This would give a unique possibility to look at causality and generate new knowledge about the pathophysiology of the disease. This would, however, be significantly time- and resource-consuming, which is probably the reason why few studies like this have been executed.

Regarding the methylation analysis, we were not able to investigate the desired regions of the genes *NR3C1* and *CD79A* because of time limitations. If we had more time, we would have designed primers optimized for these regions and maybe get more relevant and interesting results from the methylation analyses. Another thing we would have done if there was time, is controlling for confounding effects. Ideally, the blood samples should have been fractionated based on cell types and studied individually. However, there are also statistical methods available for conducting such analysis [48].

7 Conclusions

I) miR-223-3p and miR-16-5p were predicted as upstream regulators of *IL1RN* and *FLT3*, respectively. As *IL1RN* and *FLT3* were previously found to be differentially expressed in CFS patients, it was hypothesized that these miRNA may be interesting to study further.

II) We were not able to validate differential expression of *IL1RN* and *FLT3* in CFS patients, but our data show a clear trend towards up-regulation of *IL1RN* in CFS patients compared to controls. These results lead to the decision to analyze miR-223-3p, which is a regulator of *IL1RN* expression.

III) No significant differences were found in expression levels of miR-223-3p, miR-142-5p, miR-143-3p, miR-486-3p or miR-146a-5p between CFS patients and healthy controls. This indicates that regulation of gene expression by these miRNAs may not be an important part of the pathophysiology of adolescent CFS.

IV) No significant differences were found between CFS patients and healthy controls in DNA methylation levels of *NR3C1* promoter 1A. Significant hypermethylation of the *IRF7* promoter was detected in CFS patients compared to healthy controls, suggesting a role for *IRF7* promoter hypermethylation in adolescent CFS.

V) There was found a significant negative correlation between *NR3C1-1A* methylation and steps per day, and a non-significant positive correlation between *NR3C1-1A* methylation and CFQ for the CFS group. A negative correlation between *IRF7* methylation and steps per day, and a positive correlation between *IRF7* methylation and CFQ score was also detected. Although not statistically significant, these correlations may indicate a positive relationship between *IRF7* promoter hypermethylation and fatigue severity.

Reference list

[1] Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Molecular Biology of The Cell: Garland Science, 2008.

[2] Albright F, Light K, Light A, Bateman L, Cannon-Albright LA. Evidence for a heritable predisposition to Chronic Fatigue Syndrome. BMC neurology 2011;11:62.

[3] Allantaz F, Cheng DT, Bergauer T, Ravindran P, Rossier MF, Ebeling M, Badi L, Reis B, Bitter H, D'Asaro M, Chiappe A, Sridhar S, Pacheco GD, Burczynski ME, Hochstrasser D, Vonderscher J, Matthes T. Expression profiling of human immune cell subsets identifies miRNA-mRNA regulatory relationships correlated with cell type specific expression. PloS one 2012;7(1):e29979.

[4] Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. Nature reviews Genetics 2016;17(8):487-500.

[5] Ao X, Sa R, Wang J, Dao R, Wang H, Yu H. Activation-induced cytidine deaminase selectively catalyzed active DNA demethylation in pluripotency gene and improved cell reprogramming in bovine SCNT embryo. Cytotechnology 2016;68(6):2637-2648.

[6] Argentieri MA, Nagarajan S, Seddighzadeh B, Baccarelli AA, Shields AE. Epigenetic Pathways in Human Disease: The Impact of DNA Methylation on Stress-Related Pathogenesis and Current Challenges in Biomarker Development. EBioMedicine 2017;18:327-350.

[7] Bakken IJ, Tveito K, Gunnes N, Ghaderi S, Stoltenberg C, Trogstad L, Haberg SE, Magnus P. Two age peaks in the incidence of chronic fatigue syndrome/myalgic encephalomyelitis: a population-based registry study from Norway 2008-2012. BMC medicine 2014;12:167.

[8] Bansal AS, Bradley AS, Bishop KN, Kiani-Alikhan S, Ford B. Chronic fatigue syndrome, the immune system and viral infection. Brain Behav Immun 2012;26(1):24-31.

[9] Bird A. DNA methylation patterns and epigenetic memory. Genes & development 2002;16(1):6-21.

[10] Brebi-Mieville P, Ili-Gangas C, Leal-Rojas P, Noordhuis MG, Soudry E, Perez J, Roa JC, Sidransky D, Guerrero-Preston R. Clinical and public health research using methylated DNA immunoprecipitation (MeDIP): a comparison of commercially available kits to examine differential DNA methylation across the genome. Epigenetics 2012;7(1):106-112.

[11] Brenu EW, Ashton KJ, Batovska J, Staines DR, Marshall-Gradisnik SM. High-throughput sequencing of plasma microRNA in chronic fatigue syndrome/myalgic encephalomyelitis. PloS one 2014;9(9):e102783.

[12] Brenu EW, Ashton KJ, van Driel M, Staines DR, Peterson D, Atkinson GM, Marshall-Gradisnik SM. Cytotoxic lymphocyte microRNAs as prospective biomarkers for Chronic Fatigue Syndrome/Myalgic Encephalomyelitis. J Affect Disord 2012;141(2-3):261-269.

[13] Brenu EW, van Driel ML, Staines DR, Ashton KJ, Ramos SB, Keane J, Klimas NG, Marshall-Gradisnik SM. Immunological abnormalities as potential biomarkers in Chronic Fatigue Syndrome/Myalgic Encephalomyelitis. J Transl Med 2011;9:81.

[14] Buchwald D, Herrell R, Ashton S, Belcourt M, Schmaling K, Sullivan P, Neale M, Goldberg J. A twin study of chronic fatigue. Psychosomatic medicine 2001;63(6):936-943.

[15] Budak H, Bulut R, Kantar M, Alptekin B. MicroRNA nomenclature and the need for a revised naming prescription. Briefings in functional genomics 2016;15(1):65-71.

[16] Bustin SA. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. Journal of molecular endocrinology 2002;29(1):23-39.
[17] Carthew RW, Sontheimer EJ. Origins and Mechanisms of miRNAs and siRNAs. Cell 2009;136(4):642-655.
[18] Chen C, Tan R, Wong L, Fekete R, Halsey J. Quantitation of microRNAs by real-time RT-qPCR. Methods in molecular biology (Clifton, NJ) 2011;687:113-134.
[19] Chen X, Ba Y, Ma L, Cai X, Yin Y, Wang K, Guo J, Zhang Y, Chen J, Guo X, Li Q, Li X, Wang W, Zhang Y, Wang J, Jiang X, Xiang Y, Xu C, Zheng P, Zhang J, Li R, Zhang H, Shang X, Gong T, Ning G, Wang J, Zen K, Zhang J, Zhang CY. Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. Cell research 2008;18(10):997-1006.
[20] Cheung HH, Lee TL, Rennert OM, Chan WY. Methylation profiling using methylated DNA immunoprecipitation and tiling array hybridization. Methods in molecular biology (Clifton, NJ) 2012;825:115-126.
[21] Committee on the Diagnostic Criteria for Myalgic Encephalomyelitis/Chronic Fatigue S, Board on the Health of Select P, Institute of M. The National Academies Collection: Reports funded by National Institutes of Health. Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an Illness. Washington (DC): National Academies Press (US)

[22] Creemers EE, Tijsen AJ, Pinto YM. Circulating MicroRNAs. Novel Biomarkers and Extracellular Communicators in Cardiovascular Disease? 2012;110(3):483-495.
[23] Dahl JA, Collas P. A rapid micro chromatin immunoprecipitation assay (microChIP). Nat Protoc 2008;3(6):1032-1045.
[24] De S, Shaknovich R, Riester M, Elemento O, Geng H, Kormaksson M, Jiang Y, Woolcock B, Johnson N, Polo JM, Cerchietti L, Gascoyne RD, Melnick A, Michor F. Aberration in DNA Methylation in B-Cell Lymphomas Has a Complex Origin and Increases with Disease Severity. PLoS Genetics 2013;9(1):e1003137.
[25] De Spiegelaere W, Dern-Wieloch J, Weigel R, Schumacher V, Schorle H, Nettersheim D, Bergmann M, Brehm R, Kliesch S, Vandekerckhove L, Fink C. Reference gene validation for RT-qPCR, a note on different available software packages. PloS one 2015;10(3):e0122515.
[26] de Vega WC, Herrera S, Vernon SD, McGowan PO. Epigenetic modifications and glucocorticoid sensitivity in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS). BMC medical genomics 2017;10(1):11.
[27] de Vega WC, Vernon SD, McGowan PO. DNA methylation modifications associated with chronic fatigue syndrome. PloS one 2014;9(8):e104757.
[28] Ding S, Liang Y, Zhao M, Liang G, Long H, Zhao S, Wang Y, Yin H, Zhang P, Zhang Q, Lu Q. Decreased microRNA-142-3p/5p expression causes CD4+ T cell activation and B cell hyperstimulation in systemic lupus erythematosus. Arthritis Rheum 2012;64(9):2953-2963.
[29] Dommerud T. Nedslående forskningsresultat for ME-pasienter. Aftenposten. www.aftenposten.no, 2017.
[30] Donovan J, Copeland PR. Selenocysteine insertion sequence binding protein 2L is implicated as a novel post-transcriptional regulator of selenoprotein expression. PloS one 2012;7(4):e35581.
[31] Dupont C, Armant DR, Brenner CA. Epigenetics: Definition, Mechanisms and Clinical Perspective. Seminars in reproductive medicine 2009;27(5):351-357.

[32] Dwight Z, Palais R, Wittwer CT. uMELT: prediction of high-resolution melting curves and dynamic melting profiles of PCR products in a rich web application. Bioinformatics 2011;27(7):1019-1020.
[33] Elliott E, Ezra-Nevo G, Regev L, Neufeld-Cohen A, Chen A. Resilience to social stress coincides with functional DNA methylation of the Crf gene in adult mice. Nature neuroscience 2010;13(11):1351-1353.
[34] Eulalio A, Behm-Ansmant I, Schweizer D, Izaurralde E. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. Molecular and cellular biology 2007;27(11):3970-3981.
[35] Farina NH, Wood ME, Perrapato SD, Francklyn CS, Stein GS, Stein JL, Lian JB. Standardizing analysis of circulating microRNA: clinical and biological relevance. Journal of cellular biochemistry 2014;115(5):805-811.
[36] Faro M, Saez-Francas N, Castro-Marrero J, Aliste L, Fernandez de Sevilla T, Alegre J. Gender differences in chronic fatigue syndrome. Reumatologia clinica 2016;12(2):72-77.
[37] Fluge O, Bruland O, Risa K, Storstein A, Kristoffersen EK, Sapkota D, Naess H, Dahl O, Nyland H, Mella O. Benefit from B-lymphocyte depletion using the anti-CD20 antibody rituximab in chronic fatigue syndrome. A double-blind and placebo-controlled study. PloS one 2011;6(10):e26358.
[38] Goertzel BN, Pennachin C, de Souza Coelho L, Gurbaxani B, Maloney EM, Jones JF. Combinations of single nucleotide polymorphisms in neuroendocrine effector and receptor genes predict chronic fatigue syndrome. Pharmacogenomics 2006;7(3):475-483.
[39] Gow JW, Hagan S, Herzyk P, Cannon C, Behan PO, Chaudhuri A. A gene signature for post-infectious chronic fatigue syndrome. BMC medical genomics 2009;2:38.
[40] Grasedieck S, Scholer N, Bommer M, Niess JH, Tumani H, Rouhi A, Bloehdorn J, Liebisch P, Mertens D, Dohner H, Buske C, Langer C, Kuchenbauer F. Impact of serum storage conditions on microRNA stability. Leukemia 2012;26(11):2414-2416.
[41] Gu S, Roderick HL, Camacho P, Jiang JX. Characterization of an N-system amino acid transporter expressed in retina and its involvement in glutamine transport. The Journal of biological chemistry 2001;276(26):24137-24144.
[42] Gudsnuk K, Champagne FA. Epigenetic Influence of Stress and the Social Environment. ILAR Journal 2012;53(3-4):279-288.
[43] Ha H, Barnoski BL, Sun L, Emanuel BS, Burrows PD. Structure, chromosomal localization, and methylation pattern of the human mb-1 gene. Journal of immunology (Baltimore, Md : 1950) 1994;152(12):5749-5757.
[44] Herceg Z. Epigenetics and cancer: towards an evaluation of the impact of environmental and dietary factors. Mutagenesis 2007;22(2):91-103.
[45] Hibio N, Hino K, Shimizu E, Nagata Y, Ui-Tei K. Stability of miRNA 5'terminal and seed regions is correlated with experimentally observed miRNA-mediated silencing efficacy. Scientific reports 2012;2:996.
[46] Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC, Kitano TK, Hodel MR, Petersen JF, Wyatt PW, Steenblock ER, Shah PH, Bousse LJ, Troup CB, Mellen JC, Wittmann DK, Erndt NG, Cauley TH, Koehler RT, So AP, Dube S, Rose KA, Montesclaros L, Wang S, Stumbo DP, Hodges SP, Romine S, Milanovich FP, White HE, Regan JF, Karlin-Neumann GA, Hindson CM, Saxonov S, Colston BW. High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. Analytical Chemistry 2011;83(22):8604-8610.

[47] Horsthemke B. Mechanisms of imprint dysregulation. American Journal of Medical Genetics Part C: Seminars in Medical Genetics 2010;154C(3):321-328.
[48] Houseman EA, Kim S, Kelsey KT, Wiencke JK. DNA Methylation in Whole Blood: Uses and Challenges. Current environmental health reports 2015;2(2):145-154.
[49] Huang W, Thomas B, Flynn RA, Gavzy SJ, Wu L, Kim SV, Hall JA, Miraldi ER, Ng CP, Rigo F, Meadows S, Montoya NR, Herrera NG, Domingos AI, Rastinejad F, Myers RM, Fuller-Pace FV, Bonneau R, Chang HY, Acuto O, Littman DR. DDX5 and its associated lncRNA Rmrp modulate TH17 cell effector functions. Nature 2015;528(7583):517-522.
[50] Iguchi-Ariga SM, Schaffner W. CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. Genes & development 1989;3(5):612-619.
[51] Jackson C. The Chalder Fatigue Scale (CFQ 11). Occupational medicine (Oxford, England) 2015;65(1):86.
[52] Jackson RJ, Standart N. How do microRNAs regulate gene expression? Sci STKE 2007;2007(367):re1.
[53] Jonas S, Izaurralde E. Towards a molecular understanding of microRNA-mediated gene silencing. Nature reviews Genetics 2015;16(7):421-433.
[54] Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature reviews Genetics 2012;13(7):484-492.
[55] Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, Wolffe AP. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nature genetics 1998;19(2):187-191.
[56] Kaushik N, Fear D, Richards SC, McDermott CR, Nuwaysir EF, Kellam P, Harrison TJ, Wilkinson RJ, Tyrrell DA, Holgate ST, Kerr JR. Gene expression in peripheral blood mononuclear cells from patients with chronic fatigue syndrome. J Clin Pathol 2005;58(8):826-832.
[57] Kim VN, Han J, Siomi MC. Biogenesis of small RNAs in animals. Nature reviews Molecular cell biology 2009;10(2):126-139.
[58] Koberle V, Pleli T, Schmithals C, Augusto Alonso E, Haupenthal J, Bonig H, Peveling-Oberhag J, Biondi RM, Zeuzem S, Kronenberger B, Waidmann O, Piiper A. Differential stability of cell-free circulating microRNAs: implications for their utilization as biomarkers. PloS one 2013;8(9):e75184.
[59] Kozera B, Rapacz M. Reference genes in real-time PCR. Journal of Applied Genetics 2013;54(4):391-406.
[60] Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. Nature reviews Genetics 2010;11(9):597-610.
[61] Kwon HY, Bajaj J, Ito T, Blevins A, Konuma T, Weeks J, Lytle NK, Koechlein CS, Rizzieri D, Chuah C, Oehler VG, Sasik R, Hardiman G, Reya T. Tetraspanin 3 Is Required for the Development and Propagation of Acute Myelogenous Leukemia. Cell stem cell 2015;17(2):152-164.
[62] Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nature reviews Genetics 2010;11(3):204-220.
[63] Lazarus MB, Nam Y, Jiang J, Sliz P, Walker S. Structure of human O-GlcNAc transferase and its complex with a peptide substrate. Nature 2011;469(7331):564-567.
[64] Li C, Ge LL, Li PP, Wang Y, Sun MX, Huang L, Ishag H, Di DD, Shen ZQ, Fan WX, Mao X. The DEAD-box RNA helicase DDX5 acts as a positive regulator of Japanese encephalitis virus replication by binding to viral 3' UTR. Antiviral research 2013;100(2):487-499.

[65] Li MD, Ruan HB, Hughes ME, Lee JS, Singh JP, Jones SP, Nitabach MN, Yang X. O-GlcNAc signaling entrains the circadian clock by inhibiting BMAL1/CLOCK ubiquitination. Cell metabolism 2013;17(2):303-310.
[66] Li Y, Kowdley KV. Method for microRNA isolation from clinical serum samples. Anal Biochem 2012;431(1):69-75.
[67] Liew M, Pryor R, Palais R, Meadows C, Erali M, Lyon E, Wittwer C. Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. Clinical chemistry 2004;50(7):1156-1164.
[68] Lim C, Allada R. Emerging roles for post-transcriptional regulation in circadian clocks. Nature neuroscience 2013;16(11):1544-1550.
[69] Lindholm ME, Marabita F, Gomez-Cabrero D, Rundqvist H, Ekstrom TJ, Tegner J, Sundberg CJ. An integrative analysis reveals coordinated reprogramming of the epigenome and the transcriptome in human skeletal muscle after training. Epigenetics 2014;9(12):1557-1569.
[70] Ling D, Salvaterra PM. Robust RT-qPCR Data Normalization: Validation and Selection of Internal Reference Genes during Post-Experimental Data Analysis. PloS one 2011;6(3):e17762.
[71] Lock LF, Takagi N, Martin GR. Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. Cell;48(1):39-46.
[72] Long HK, King HW, Patient RK, Odom DT, Klose RJ. Protection of CpG islands from DNA methylation is DNA-encoded and evolutionarily conserved. Nucleic acids research 2016;44(14):6693-6706.
[73] Lorusso L, Mikhaylova SV, Capelli E, Ferrari D, Ngonga GK, Ricevuti G. Immunological aspects of chronic fatigue syndrome. Autoimmun Rev 2009;8(4):287-291.
[74] Lu R, Au WC, Yeow WS, Hageman N, Pitha PM. Regulation of the promoter activity of interferon regulatory factor-7 gene. Activation by interferon snd silencing by hypermethylation. The Journal of biological chemistry 2000;275(41):31805-31812.
[75] Maes M, Twisk FN, Ringel K. Inflammatory and cell-mediated immune biomarkers in myalgic encephalomyelitis/chronic fatigue syndrome and depression: inflammatory markers are higher in myalgic encephalomyelitis/chronic fatigue syndrome than in depression. Psychother Psychosom 2012;81(5):286-295.
[76] Maher KJ, Klimas NG, Fletcher MA. Chronic fatigue syndrome is associated with diminished intracellular perforin. Clin Exp Immunol 2005;142(3):505-511.
[77] Martin DI, Cropley JE, Suter CM. Epigenetics in disease: leader or follower? Epigenetics 2011;6(7):843-848.
[78] Meaney MJ, Szyf M. Maternal care as a model for experience-dependent chromatin plasticity? Trends in neurosciences 2005;28(9):456-463.
[79] Mensah FKF, Bansal AS, Ford B, Cambridge G. Chronic fatigue syndrome and the immune system: Where are we now? Neurophysiologie clinique = Clinical neurophysiology 2017;47(2):131-138.
[80] Mertens D, Philippen A, Ruppel M, Allegra D, Bhattacharya N, Tschuch C, Wolf S, Idler I, Zenz T, Stilgenbauer S. Chronic lymphocytic leukemia and 13q14: miRs and more. Leukemia & lymphoma 2009;50(3):502-505.
[81] Meyer B, Nguyen CB, Moen A, Fagermoen E, Sulheim D, Nilsen H, Wyller VB, Gjerstad J. Maintenance of Chronic Fatigue Syndrome (CFS) in Young CFS Patients Is Associated with the 5-HTTLPR and SNP rs25531 A > G Genotype. PloS one 2015;10(10):e0140883.

[82] Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 1998;393(6683):386-389.
[83] Narita M, Nishigami N, Narita N, Yamaguti K, Okado N, Watanabe Y, Kuratsune H. Association between serotonin transporter gene polymorphism and chronic fatigue syndrome. Biochemical and biophysical research communications 2003;311(2):264-266.
[84] Nguyen CB, Alsøe L, Lindvall JM, Sulheim D, Fagermoen E, Winger A, Kaarbø M, Nilsen H, Wyller VB. Whole blood gene expression in adolescent chronic fatigue syndrome: an exploratory cross-sectional study suggesting altered B cell differentiation and survival. Journal of Translational Medicine 2017;15(1):102.
[85] Nijhof SL, Rutten JM, Uiterwaal CS, Bleijenberg G, Kimpen JL, Putte EM. The role of hypocortisolism in chronic fatigue syndrome. Psychoneuroendocrinology 2014;42:199-206.
[86] Ning S, Pagano JS, Barber GN. IRF7: activation, regulation, modification and function. Genes and immunity 2011;12(6):399-414.
[87] Nitert MD, Dayeh T, Volkov P, Elgzyri T, Hall E, Nilsson E, Yang BT, Lang S, Parikh H, Wessman Y, Weishaupt H, Attema J, Abels M, Wierup N, Almgren P, Jansson PA, Ronn T, Hansson O, Eriksson KF, Groop L, Ling C. Impact of an exercise intervention on DNA methylation in skeletal muscle from first-degree relatives of patients with type 2 diabetes. Diabetes 2012;61(12):3322-3332.
[88] Okamura K, Liu N, Lai EC. Distinct mechanisms for microRNA strand selection by Drosophila Argonautes. Mol Cell 2009;36(3):431-444.
[89] Padro CJ, Sanders VM. Neuroendocrine regulation of inflammation. Semin Immunol 2014;26(5):357-368.
[90] Papadopoulos AS, Cleare AJ. Hypothalamic-pituitary-adrenal axis dysfunction in chronic fatigue syndrome. Nat Rev Endocrinol 2011;8(1):22-32.
[91] Pasquinelli AE. MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. Nature reviews Genetics 2012;13(4):271-282.
[92] Pavlasova G, Borsky M, Seda V, Cerna K, Osickova J, Doubek M, Mayer J, Calogero R, Trbusek M, Pospisilova S, Davids MS, Kipps TJ, Brown JR, Mraz M. Ibrutinib inhibits CD20 upregulation on CLL B cells mediated by the CXCR4/SDF-1 axis. Blood 2016;128(12):1609-1613.
[93] Pelanda R, Braun U, Hobeika E, Nussenzweig MC, Reth M. B cell progenitors are arrested in maturation but have intact VDJ recombination in the absence of Ig-alpha and Ig-beta. Journal of immunology (Baltimore, Md : 1950) 2002;169(2):865-872.
[94] Petty RD, McCarthy NE, Le Dieu R, Kerr JR. MicroRNAs hsa-miR-99b, hsa-miR-330, hsa-miR-126 and hsa-miR-30c: Potential Diagnostic Biomarkers in Natural Killer (NK) Cells of Patients with Chronic Fatigue Syndrome (CFS)/ Myalgic Encephalomyelitis (ME). PloS one 2016;11(3):e0150904.
[95] Preininger M, Arafat D, Kim J, Nath AP, Idaghdour Y, Brigham KL, Gibson G. Blood-Informative Transcripts Define Nine Common Axes of Peripheral Blood Gene Expression. PLoS Genetics 2013;9(3):e1003362.
[96] Prins JB, van der Meer JW, Bleijenberg G. Chronic fatigue syndrome. Lancet (London, England) 2006;367(9507):346-355.
[97] Provenzano M, Mocellin S. Complementary techniques: validation of gene expression data by quantitative real time PCR. Advances in experimental medicine and biology 2007;593:66-73.
[98] Pryor RJ, Wittwer CT. Real-time polymerase chain reaction and melting curve analysis. Methods in molecular biology (Clifton, NJ) 2006;336:19-32.

[99] Risso D, Ngai J, Speed TP, Dudoit S. Normalization of RNA-seq data using factor analysis of control genes or samples. Nat Biotechnol 2014;32(9):896-902.
[100] Ronn T, Volkov P, Davegardh C, Dayeh T, Hall E, Olsson AH, Nilsson E, Tornberg A, Dekker Nitert M, Eriksson KF, Jones HA, Groop L, Ling C. A six months exercise intervention influences the genome-wide DNA methylation pattern in human adipose tissue. PLoS Genet 2013;9(6):e1003572.
[101] Roth TL, Lubin FD, Funk AJ, Sweatt JD. Lasting epigenetic influence of early-life adversity on the BDNF gene. Biological psychiatry 2009;65(9):760-769.
[102] Roth TL, Zoladz PR, Sweatt JD, Diamond DM. Epigenetic modification of hippocampal Bdnf DNA in adult rats in an animal model of post-traumatic stress disorder. Journal of psychiatric research 2011;45(7):919-926.
[103] Salmanidis M, Pillman K, Goodall G, Bracken C. Direct transcriptional regulation by nuclear microRNAs. The international journal of biochemistry & cell biology 2014;54:304-311.
[104] Schubeler D, Lorincz MC, Cimbora DM, Telling A, Feng YQ, Bouhassira EE, Groudine M. Genomic targeting of methylated DNA: influence of methylation on transcription, replication, chromatin structure, and histone acetylation. Molecular and cellular biology 2000;20(24):9103-9112.
[105] Slezak S, Jin P, Caruccio L, Ren J, Bennett M, Zia N, Adams S, Wang E, Ascensao J, Schechter G, Stroncek D. Gene and microRNA analysis of neutrophils from patients with polycythemia vera and essential thrombocytosis: down-regulation of micro RNA-1 and -133a. J Transl Med 2009;7:39.
[106] Sulheim D, Fagermoen E, Winger A, Andersen AM, Godang K, Muller F, Rowe PC, Saul JP, Skovlund E, Oie MG, Wyller VB. Disease mechanisms and clonidine treatment in adolescent chronic fatigue syndrome: a combined cross-sectional and randomized clinical trial. JAMA Pediatr 2014;168(4):351-360.
[107] Szyf M, McGowan P, Meaney MJ. The social environment and the epigenome. Environmental and molecular mutagenesis 2008;49(1):46-60.
[108] Szyf M, Weaver IC, Champagne FA, Diorio J, Meaney MJ. Maternal programming of steroid receptor expression and phenotype through DNA methylation in the rat. Frontiers in neuroendocrinology 2005;26(3-4):139-162.
[109] Tazi-Ahnini R, Cox A, McDonagh AJ, Nicklin MJ, di Giovine FS, Timms JM, Messenger AG, Dimitropoulou P, Duff GW, Cork MJ. Genetic analysis of the interleukin-1 receptor antagonist and its homologue IL-1L1 in alopecia areata: strong severity association and possible gene interaction. European journal of immunogenetics : official journal of the British Society for Histocompatibility and Immunogenetics 2002;29(1):25-30.
[110] Tiwari-Woodruff SK, Buznikov AG, Vu TQ, Micevych PE, Chen K, Kornblum HI, Bronstein JM. Osp/Claudin-11 Forms a Complex with a Novel Member of the Tetraspanin Super Family and β1 Integrin and Regulates Proliferation and Migration of Oligodendrocytes. The Journal of Cell Biology 2001;153(2):295-306.
[111] Toyota M, Issa JP. CpG island methylator phenotypes in aging and cancer. Seminars in cancer biology 1999;9(5):349-357.
[112] Trapannone R, Rafie K, van Aalten DM. O-GlcNAc transferase inhibitors: current tools and future challenges. Biochemical Society transactions 2016;44(1):88-93.
[113] Turner JD, Alt SR, Cao L, Vernocchi S, Trifonova S, Battello N, Muller CP. Transcriptional control of the glucocorticoid receptor: CpG islands, epigenetics and more. Biochemical pharmacology 2010;80(12):1860-1868.
[114] Vaissiere T, Sawan C, Herceg Z. Epigenetic interplay between histone modifications and DNA methylation in gene silencing. Mutation research 2008;659(1-2):40-48.

[115] Vangeel E, Van Den Eede F, Hompes T, Izzi B, Del Favero J, Moorkens G, Lambrechts D, Freson K, Claes S. Chronic Fatigue Syndrome and DNA Hypomethylation of the Glucocorticoid Receptor Gene Promoter 1F Region: Associations With HPA Axis Hypofunction and Childhood Trauma. Psychosomatic medicine 2015;77(8):853-862.
[116] Vollmer-Conna U, Cameron B, Hadzi-Pavlovic D, Singletary K, Davenport T, Vernon S, Reeves WC, Hickie I, Wakefield D, Lloyd AR. Postinfective fatigue syndrome is not associated with altered cytokine production. Clin Infect Dis 2007;45(6):732-735.
[117] Wang K, Yuan Y, Cho JH, McClarty S, Baxter D, Galas DJ. Comparing the MicroRNA Spectrum between Serum and Plasma. PloS one 2012;7(7).
[118] Weaver IC, Champagne FA, Brown SE, Dymov S, Sharma S, Meaney MJ, Szyf M. Reversal of maternal programming of stress responses in adult offspring through methyl supplementation: altering epigenetic marking later in life. The Journal of neuroscience : the official journal of the Society for Neuroscience 2005;25(47):11045-11054.
[119] Wyller VB, Barbieri R, Saul JP. Blood pressure variability and closed-loop baroreflex assessment in adolescent chronic fatigue syndrome during supine rest and orthostatic stress. Eur J Appl Physiol 2011;111(3):497-507.
[120] Wyller VB, Barbieri R, Thaulow E, Saul JP. Enhanced vagal withdrawal during mild orthostatic stress in adolescents with chronic fatigue. Ann Noninvasive Electrocardiol 2008;13(1):67-73.
[121] Wyller VB, Due R, Saul JP, Amlie JP, Thaulow E. Usefulness of an abnormal cardiovascular response during low-grade head-up tilt-test for discriminating adolescents with chronic fatigue from healthy controls. Am J Cardiol 2007;99(7):997-1001.
[122] Wyller VB, Eriksen HR, Malterud K. Can sustained arousal explain the Chronic Fatigue Syndrome? Behav Brain Funct 2009;5:10.
[123] Wyller VB, Evang JA, Godang K, Solhjell KK, Bollerslev J. Hormonal alterations in adolescent chronic fatigue syndrome. Acta paediatrica (Oslo, Norway : 1992) 2010;99(5):770-773.
[124] Wyller VB, Godang K, Morkrid L, Saul JP, Thaulow E, Walloe L. Abnormal thermoregulatory responses in adolescents with chronic fatigue syndrome: relation to clinical symptoms. Pediatrics 2007;120(1):e129-137.
[125] Wyller VB, Sorensen O, Sulheim D, Fagermoen E, Ueland T, Mollnes TE. Plasma cytokine expression in adolescent chronic fatigue syndrome. Brain Behav Immun 2015;46:80-86.
[126] Yehuda R, Flory JD, Bierer LM, Henn-Haase C, Lehrner A, Desarnaud F, Makotkine I, Daskalakis NP, Marmar CR, Meaney MJ. Lower methylation of glucocorticoid receptor gene promoter 1F in peripheral blood of veterans with posttraumatic stress disorder. Biological psychiatry 2015;77(4):356-364.
[127] Zen M, Canova M, Campana C, Bettio S, Nalotto L, Rampudda M, Ramonda R, Iaccarino L, Doria A. The kaleidoscope of glucorticoid effects on immune system. Autoimmun Rev 2011;10(6):305-310.
[128] Zhang FF, Cardarelli R, Carroll J, Fulda KG, Kaur M, Gonzalez K, Vishwanatha JK, Santella RM, Morabia A. Significant differences in global genomic DNA methylation by gender and race/ethnicity in peripheral blood. Epigenetics 2011;6(5):623-629.
[129] Zhang H, Xing Z, Mani SK, Bancel B, Durantel D, Zoulim F, Tran EJ, Merle P, Andrisani O. RNA helicase DEAD box protein 5 regulates Polycomb repressive complex 2/Hox transcript antisense intergenic RNA function in hepatitis B virus infection and hepatocarcinogenesis. Hepatology (Baltimore, Md) 2016;64(4):1033-1048.

[130] Zhu M, Zhao S. Candidate Gene Identification Approach: Progress and Challenges. International Journal of Biological Sciences 2007;3(7):420-427.

Abbreviations

5mC	5-methylcytosine
Ab	Antibody
AID	Activation-induced cytidine deaminase
Ag	Antigen
BCL7A	B-cell lymphoma tumor suppressor 7A
BCR	B-cell receptor
BLAST	Basic local alignment search tool
Bp	Base pair
CD84	Cluster of differentiantion 84
cDNA	Complementary DNA
CFS	Chronic Fatigue Syndrome
CFQ	Chalder Fatigue Score
CGI	CpG island
CpG	cytosine-phosphate-guanine dinucleotide
Ct	Threshold cycle
DCP2	Decapping protein 2
ddPCR	Droplet digital PCR
DDX5	DEAD-box helicase 5
DEG	Differentially expressed gene
DLBCL	Diffuse large B-cell lymphoma
DNA	Deoxyribonucleic acid
Dnmnt	DNA methyltransferase
EBV	Epstein-Barr virus
eIF	Eukaryotic translation initiation factor
EtOH	Ethanol
EXP5	Exportin 5
FRET	Fluorescent resonance energy transfer
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GC	Guanine-cytosine
GR	Glucocorticoid receptor
HC	Healthy control
HPA	Hypothalamus-pituitary-adrenal
IFN	Interferon

IgG	Immunoglobulin G
IL	Interleukin
IL-1ra	IL-1 receptor antagonist
IP	Immunoprecipitation
IPA	Ingenuity pathway analysis
IRF	Interferon regulatory factor
IQR	Interquartile range
LMP1	Latent membrane protein-1
MagMeDIP	Magnetic MeDIP
ME	Myalgic Encephalomyelitis
MeDIP	methylated DNA immunoprecipitation
meDNA	Methylated DNA
miRISC	microRNA-induced silencing complex
miRNA	microRNA
mRNA	messenger RNA
NK	Natural killer
NorCAPITAL	The Norwegian Study of Chronic Fatigue Syndrome in Adolescents: Pathophysiology and Intervention Trial
NR3C1	Nuclear receptor subfamily 3 group C member 1
Nt	Nucleotide
OGT	O-linked N-acetylglucosamine transferase
ON	Overnight
PBMC	Peripheral blood mononuclear cells
PCR	Polymerase chain reaction
pDC	Plasmacytoid dendritic cell
PTSD	Posttraumatic stress disorder
qPCR	quantitative PCR
RNA	Ribonucleic acid
RNAi	RNA interference
RNA-seq	RNA sequencing
RQ	Relative quantity
RT-qPCR	Reverse transcription qPCR
RUV	Remove unwanted variation
SAGE	Serial analysis of gene expression
SAP	SLAM-associated protein
SECIS	Sec insertion sequence

SECISBP2L	SECIS binding protein 2 like
SE	Standard error
SEM	Standard error of the mean
SLC38A1	Solute carrier family member 1
SNP	Single nucleotide polymorphism
TDG	Thymine DNA glycosylase
TE	Transposable element
TET	Ten-eleven translocation
Tm	Melting temperature
TNF	Tumor necrosis factor
TSPAN3	Tetraspaning 3
TSS	Transcriptional start site
XRN1	Exoribonuclease 1

Appendices

Appendix 1: Table S1

Table S1. Differentially Expressed Genes and their annotated proteins or gene products in CFS patients as compared to healthy controls, adjusted for age and gender differences across groups and sorted according to foldchange. Retrieved from Nguyen et al. [84]

Ensembl ID	*Gene name*	*Foldchange (log 2)*	*Foldchange*	*p-value (unadjusted)*	*p-value (adjusted)*	*Annotated protein/gene product*
ENSG00000105369	CD79A	-0,2842	0,8212	0,0001	0,0393	CD79a molecule, immunoglobulin-associated alpha
ENSG00000159958	TNFRSF13C	-0,2707	0,8289	0,0001	0,0395	tumor necrosis factor receptor superfamily, member 13C
ENSG00000122025	FLT3	-0,2633	0,8332	0,0006	0,0682	fms-related tyrosine kinase 3
ENSG00000241351	IGKV3_11	-0,2579	0,8363	0,0015	0,0979	Ig kappa chain V region
ENSG00000164330	EBF1	-0,2578	0,8364	0,0004	0,0615	early B-cell factor 1
ENSG00000172264	MACROD2	-0,2548	0,8381	0,0014	0,0928	MACRO domain containing 2
ENSG00000179088	C12orf42	-0,2491	0,8414	0,0001	0,0341	chromosome 12 open reading frame 42
ENSG00000197863	ZNF790	-0,2474	0,8424	0,0002	0,0461	zinc finger protein 790
ENSG00000160683	CXCR5	-0,2383	0,8477	0,0007	0,0735	chemokine (C-X-C motif) receptor 5
ENSG00000110987	BCL7A	-0,2307	0,8522	0,0011	0,0846	B-cell CLL/lymphoma 7A
ENSG00000176533	GNG7	-0,2214	0,8577	0,0007	0,0735	guanine nucleotide binding protein (G protein), gamma 7
ENSG00000171604	CXXC5	-0,2188	0,8593	0,0002	0,0406	CXXC finger 5
ENSG00000126432	PRDX5	-0,2157	0,8611	0,0002	0,0512	peroxiredoxin 5
ENSG00000137265	IRF4	-0,2124	0,8631	0,0012	0,0879	interferon regulatory factor 4
ENSG00000065923	SLC9A7	-0,2086	0,8654	0,0004	0,0606	solute carrier family 9 (sodium/hydrogen exchanger), member 7
ENSG00000184574	LPAR5	-0,2070	0,8663	0,0004	0,0613	lysophosphatidic acid receptor 5
ENSG00000142937	RPS8	-0,1938	0,8743	0,0011	0,0846	ribosomal protein S8 pseudogene 8
ENSG00000131469	RPL27	-0,1886	0,8775	0,0009	0,0771	ribosomal protein L27
ENSG00000124243	BCAS4	-0,1884	0,8776	0,0002	0,0406	breast carcinoma amplified sequence 4
ENSG00000140391	TSPAN3	-0,1815	0,8818	0,0014	0,0915	tetraspanin 3
ENSG00000008282	SYPL1	-0,1738	0,8865	0,0013	0,0906	synaptophysin-like 1
ENSG00000064393	HIPK2	-0,1667	0,8908	0,0005	0,0682	homeodomain interacting protein kinase 2
ENSG00000137054	POLR1E	-0,1665	0,8910	0,0015	0,0969	polymerase (RNA) I polypeptide E, 53kDa
ENSG00000213626	LBH	-0,1620	0,8938	0,0007	0,0735	limb bud and heart development homolog (mouse)
ENSG00000168385	SEPT2	-0,1617	0,8940	0,0008	0,0736	septin 2
ENSG00000130741	EIF2S3	-0,1508	0,9007	0,0000	0,0170	eukaryotic translation initiation factor 2, subunit 3 gamma, 52kDa
ENSG00000169100	SLC25A6	-0,1497	0,9014	0,0011	0,0848	solute carrier family 25 (mitochondrial carrier adenine nucleotide translocator), member 6
ENSG00000167658	EEF2	-0,1489	0,9020	0,0001	0,0341	eukaryotic translation elongation factor 2
ENSG00000139291	TMEM19	-0,1467	0,9033	0,0010	0,0819	transmembrane protein 19
ENSG00000150347	ARID5B	-0,1457	0,9040	0,0012	0,0902	AT rich interactive domain 5B (MRF1-like)
ENSG00000247556	OIP5_AS1	-0,1351	0,9106	0,0009	0,0763	*no annotation*

ENSG00000146842	TMEM209	-0,1333	0,9117	0,0011	0,0872	transmembrane protein 209
ENSG00000168286	THAP11	-0,1291	0,9144	0,0009	0,0794	THAP domain containing 11
ENSG00000161791	FMNL3	-0,1280	0,9151	0,0003	0,0567	formin-like 3
ENSG00000111371	SLC38A1	-0,1236	0,9179	0,0015	0,0987	solute carrier family 38, member 1
ENSG00000197714	ZNF460	-0,1093	0,9270	0,0007	0,0735	zinc finger protein 460
ENSG00000164091	WDR82	-0,1073	0,9283	0,0016	0,0987	WD repeat domain 82
ENSG00000138593	SECISBP2L	-0,1063	0,9289	0,0011	0,0872	SECIS binding protein 2-like
ENSG00000073849	ST6GAL1	-0,1019	0,9318	0,0007	0,0735	ST6 beta-galactosamide alpha-2,6-sialyltranferase 1
ENSG00000165494	PCF11	0,0875	1,0625	0,0007	0,0735	pre cleavate complex 2
ENSG00000100201	DDX17	0,0943	1,0675	0,0008	0,0736	DEAD (Asp-Glu-Ala-Asp) box polypeptide 17
ENSG00000131051	RBM39	0,1010	1,0725	0,0014	0,0915	similar to RNA binding motif protein 39
ENSG00000134186	PRPF38B	0,1112	1,0801	0,0013	0,0906	PRE SPLICING FACTOR 38B
ENSG00000010244	ZNF207	0,1131	1,0816	0,0008	0,0736	zinc finger protein 207
ENSG00000145868	FBXO38	0,1149	1,0829	0,0008	0,0736	F-box protein 38
ENSG00000163516	ANKZF1	0,1163	1,0840	0,0012	0,0902	ankyrin repeat and zinc finger domain containing 1
ENSG00000147854	UHRF2	0,1184	1,0855	0,0012	0,0902	ubiquitin-like with PHD and ring finger domains 2
ENSG00000147162	OGT	0,1199	1,0866	0,0002	0,0461	O-linked N-acetylglucosamine (GlcNAc) transferase
ENSG00000151532	VTI1A	0,1207	1,0873	0,0011	0,0846	vesicle transport through interaction with t-SNAREs homolog 1A
ENSG00000071189	SNX13	0,1224	1,0885	0,0008	0,0736	sorting nexin 13
ENSG00000182944	EWSR1	0,1236	1,0894	0,0013	0,0906	Ewing sarcoma breakpoint region 1
ENSG00000114857	NKTR	0,1264	1,0916	0,0007	0,0735	natural killer-tumor recognition sequence
ENSG00000197601	FAR1	0,1272	1,0922	0,0010	0,0827	fatty acyl CoA reductase 1
ENSG00000116001	TIA1	0,1280	1,0928	0,0012	0,0902	TIA1 cytotoxic granule-associated RNA binding protein
ENSG00000144848	ATG3	0,1289	1,0935	0,0009	0,0794	autophagocytosis associated protein Atg3
ENSG00000133858	ZFC3H1	0,1290	1,0935	0,0000	0,0161	zinc finger, C3H1-type containing
ENSG00000061936	SFSWAP	0,1298	1,0941	0,0010	0,0821	splicing factor, suppressor of white-apricot homolog
ENSG00000197548	ATG7	0,1308	1,0949	0,0003	0,0525	autophagy-related protein 7
ENSG00000164548	TRA2A	0,1309	1,0950	0,0005	0,0682	transformer-2 protein homolog alpha
ENSG00000043462	LCP2	0,1314	1,0953	0,0013	0,0915	lymphocyte cytosolic protein 2
ENSG00000161547	SRSF2	0,1318	1,0957	0,0003	0,0538	splicing factor, arginine/serine-rich 2
ENSG00000108654	DDX5	0,1323	1,0961	0,0016	0,0990	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5
ENSG00000239665	RP11_295P9	0,1324	1,0961	0,0006	0,0682	*no annotation*
ENSG00000132334	PTPRE	0,1328	1,0964	0,0006	0,0699	protein tyrosine phosphatase, receptor type, E
ENSG00000134480	CCNH	0,1337	1,0971	0,0005	0,0682	cyclin H
ENSG00000013441	CLK	0,1350	1,0981	0,0007	0,0735	CDC-like kinase 1
ENSG00000175309	PHYKPL	0,1356	1,0986	0,0006	0,0735	alanine-glyoxylate aminotransferase 2-like 2
ENSG00000115524	SF3B1	0,1359	1,0988	0,0008	0,0736	splicing factor 3b, subunit 1, 155kDa
ENSG00000088448	ANKRD10	0,1368	1,0995	0,0003	0,0555	ankyrin repeat domain 10
ENSG00000100650	SRSF5	0,1379	1,1003	0,0003	0,0555	splicing factor, arginine/serine-rich 5
ENSG00000137504	CREBZF	0,1404	1,1022	0,0008	0,0736	CREB/ATF bZIP transcription factor

ENSG00000163932	PRKCD	0,1404	1,1022	0,0007	0,0735	protein kinase C, delta
ENSG00000132424	PNISR	0,1406	1,1024	0,0001	0,0381	splicing factor, arginine/serine-rich 18
ENSG00000083828	ZNF586	0,1416	1,1031	0,0007	0,0735	zinc finger protein 586
ENSG00000146063	TRIM41	0,1417	1,1032	0,0008	0,0736	tripartite motif-containing 41
ENSG00000107745	MICU1	0,1435	1,1046	0,0014	0,0915	calcium binding atopy-related autoantigen 1
ENSG00000172732	MUS81	0,1460	1,1065	0,0004	0,0643	crossover junction endonuclease MUS81
ENSG00000162231	NXF1	0,1463	1,1067	0,0005	0,0682	nuclear RNA export factor 1
ENSG00000118564	FBXL5	0,1469	1,1072	0,0003	0,0590	F-box and leucine-rich repeat protein 5
ENSG00000215788	TNFRSF25	0,1472	1,1075	0,0011	0,0872	tumor necrosis factor receptor superfamily, member 25
ENSG00000104325	DECR1	0,1485	1,1084	0,0003	0,0590	2,4-dienoyl CoA reductase 1, mitochondrial
ENSG00000101916	TLR8	0,1491	1,1089	0,0013	0,0915	toll-like receptor 8
ENSG00000135164	DMTF1	0,1492	1,1089	0,0001	0,0381	cyclin D binding myb-like transcription factor 1
ENSG00000109046	WSB1	0,1515	1,1107	0,0008	0,0736	WD repeat and SOCS box-containing 1
ENSG00000075884	ARHGAP15	0,1525	1,1115	0,0007	0,0735	Rho GTPase activating protein 15
ENSG00000070269	TMEM260	0,1529	1,1118	0,0008	0,0736	chromosome 14 open reading frame 101
ENSG00000115875	SRSF7	0,1543	1,1129	0,0005	0,0648	splicing factor, arginine/serine-rich 7, 35kDa
ENSG00000230551	CTB_86H12	0,1559	1,1141	0,0004	0,0613	*no annotation*
ENSG00000261490	RP11_448G15	0,1562	1,1143	0,0014	0,0918	*no annotation*
ENSG00000135899	SP110	0,1569	1,1149	0,0016	0,0990	SP110 nuclear body protein
ENSG00000060971	ACAA1	0,1593	1,1167	0,0004	0,0613	acetyl-Coenzyme A acyltransferase 1
ENSG00000141562	NARF	0,1607	1,1178	0,0004	0,0613	nuclear prelamin A recognition factor
ENSG00000140386	SCAPER	0,1610	1,1181	0,0012	0,0879	S-phase cyclin A-associated protein in the ER
ENSG00000104852	SNRNP70	0,1611	1,1181	0,0012	0,0902	small nuclear ribonucleoprotein 70kDa (U1)
ENSG00000100949	RABGGTA	0,1647	1,1209	0,0012	0,0902	Rab geranylgeranyltransferase, alpha subunit
ENSG00000141012	GALNS	0,1648	1,1210	0,0009	0,0759	galactosamine (N-acetyl)-6-sulfate sulfatase
ENSG00000110934	BIN2	0,1664	1,1223	0,0000	0,0276	bridging integrator 2
ENSG00000197070	ARRDC1	0,1664	1,1223	0,0005	0,0647	arrestin domain containing 1
ENSG00000251562	MALAT1	0,1679	1,1234	0,0005	0,0682	metastasis associated lung adenocarcinoma transcript 1
ENSG00000111801	BTN3A3	0,1680	1,1235	0,0001	0,0406	butyrophilin, subfamily 3, member A3
ENSG00000132357	CARD6	0,1694	1,1246	0,0006	0,0735	caspase recruitment domain family, member 6
ENSG00000169045	HNRNPH1	0,1717	1,1264	0,0002	0,0406	heterogeneous nuclear ribonucleoprotein H1 (H)
ENSG00000113240	CLK4	0,1726	1,1271	0,0005	0,0660	CDC-like kinase 4
ENSG00000138964	PARVG	0,1737	1,1279	0,0000	0,0144	parvin, gamma
ENSG00000175471	MCTP1	0,1815	1,1340	0,0002	0,0406	multiple C2 domains, transmembrane 1
ENSG00000267121	CTD_2020K17	0,1820	1,1345	0,0005	0,0648	*no annotation*
ENSG00000163660	CCNL1	0,1826	1,1349	0,0000	0,0276	cyclin L1
ENSG00000071626	DAZAP1	0,1834	1,1356	0,0000	0,0148	DAZ associated protein 1
ENSG00000117616	RSRP1	0,1843	1,1363	0,0002	0,0406	chromosome 1 open reading frame 63
ENSG00000163393	SLC22A15	0,1876	1,1389	0,0016	0,0987	solute carrier family 22, member 15

ENSG00000131042	LILRB2	0,1879	1,1391	0,0014	0,0922	leukocyte immunoglobulin-like receptor, subfamily B, member 2
ENSG00000140379	BCL2A1	0,1894	1,1403	0,0012	0,0898	BCL2-related protein A1
ENSG00000143546	S100A8	0,1909	1,1415	0,0011	0,0859	S100 calcium binding protein A8
ENSG00000237298	TTN_AS1	0,1910	1,1415	0,0001	0,0295	*no annotation*
ENSG00000160796	NBEAL2	0,1932	1,1433	0,0009	0,0752	neurobeachin 2
ENSG00000101104	PABPC1L	0,1939	1,1439	0,0013	0,0915	poly(A) binding protein, cytoplasmic 1-like
ENSG00000144802	NFKBIZ	0,1948	1,1445	0,0000	0,0276	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta
ENSG00000163565	IFI16	0,1963	1,1457	0,0007	0,0735	interferon, gamma-inducible protein 16
ENSG00000248019	FAM13A	0,1967	1,1461	0,0001	0,0381	family with sequence similarity 13, member A
ENSG00000109743	BST1	0,1986	1,1476	0,0007	0,0735	bone marrow stromal cell antigen 1
ENSG00000108465	CDK5RAP3	0,2022	1,1504	0,0000	0,0273	CDK5 regulatory subunit associated protein 3
ENSG00000160593	JAML	0,2023	1,1505	0,0000	0,0063	adhesion molecule, interacts with CXADR antigen 1
ENSG00000168404	MLKL	0,2032	1,1513	0,0006	0,0735	mixed lineage kinase domain-like
ENSG00000196954	CASP4	0,2034	1,1514	0,0009	0,0794	caspase 4, apoptosis-related cysteine peptidase
ENSG00000121060	TRIM25	0,2078	1,1549	0,0003	0,0590	tripartite motif-containing 25
ENSG00000279598	RP11_65L3	0,2080	1,1551	0,0013	0,0915	*no annotation*
ENSG00000185485	SDHAP1	0,2081	1,1552	0,0015	0,0950	similar to succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial precursor
ENSG00000163728	TTC14	0,2099	1,1566	0,0000	0,0161	tetratricopeptide repeat domain 14
ENSG00000166801	FAM111A	0,2107	1,1572	0,0001	0,0341	family with sequence similarity 111, member A
ENSG00000115271	GCA	0,2121	1,1584	0,0000	0,0273	grancalcin, EF-hand calcium binding protein
ENSG00000105835	NAMPT	0,2153	1,1609	0,0004	0,0621	nicotinamide phosphoribosyltransferase
ENSG00000087157	PGS1	0,2181	1,1632	0,0007	0,0735	phosphatidylglycerophosphate synthase 1
ENSG00000100504	PYGL	0,2186	1,1636	0,0001	0,0381	phosphorylase, glycogen, liver
ENSG00000140563	MCTP2	0,2198	1,1646	0,0000	0,0161	multiple C2 domains, transmembrane 2
ENSG00000137752	CASP1	0,2204	1,1651	0,0010	0,0799	caspase 1, apoptosis-related cysteine peptidase (interleukin 1, beta, convertase)
ENSG00000173221	GLRX	0,2205	1,1651	0,0000	0,0005	glutaredoxin (thioltransferase)
ENSG00000150527	CTAGE5	0,2218	1,1662	0,0004	0,0613	CTAGE family, member 5 pseudogene, CTAGE family member
ENSG00000180902	D2HGDH	0,2220	1,1663	0,0003	0,0555	D-2-hydroxyglutarate dehydrogenase
ENSG00000254470	AP5B1	0,2226	1,1668	0,0003	0,0542	AP-5 complex subunit beta 1
ENSG00000129465	RIPK3	0,2228	1,1670	0,0000	0,0136	receptor-interacting serine-threonine kinase 3
ENSG00000255423	EBLN2	0,2258	1,1694	0,0006	0,0682	endogenous bornavirus nucleoprotein
ENSG00000180539	C9orf139	0,2263	1,1698	0,0008	0,0736	chromosome 9 open reading frame 139
ENSG00000278829	RP11_358B23	0,2263	1,1699	0,0007	0,0735	*no annotation*
ENSG00000279884	RP11_48E17	0,2286	1,1717	0,0014	0,0918	*no annotation*
ENSG00000274536	RP6_159A1	0,2313	1,1739	0,0001	0,0381	*no annotation*
ENSG00000204681	GABBR1	0,2315	1,1740	0,0002	0,0460	gamma-aminobutyric acid (GABA) B receptor, 1
ENSG00000134905	CARS2	0,2364	1,1781	0,0000	0,0025	cysteinyl-tRNA synthetase 2, mitochondrial (putative)

ENSG00000110852	CLEC2B	0,2374	1,1789	0,0001	0,0336	C-type lectin domain family 2, member B
ENSG00000270277	RP11_65L3	0,2375	1,1789	0,0006	0,0707	*no annotation*
ENSG00000170956	CEACAM3	0,2492	1,1885	0,0000	0,0148	carcinoembryonic antigen-related cell adhesion molecule 3
ENSG00000245571	AP001258	0,2509	1,1899	0,0003	0,0542	*no annotation*
ENSG00000196358	NTNG2	0,2511	1,1901	0,0016	0,0987	netrin G2
ENSG00000136689	IL1RN	0,2522	1,1910	0,0002	0,0406	interleukin 1 receptor antagonist
ENSG00000215458	FLJ46757	0,2535	1,1921	0,0008	0,0736	hypothetical LOC284837
ENSG00000181381	DDX60L	0,2545	1,1929	0,0002	0,0406	DEAD (Asp-Glu-Ala-Asp) box polypeptide 60-like
ENSG00000172785	CBWD1	0,2552	1,1935	0,0004	0,0643	COBW domain
ENSG00000281106	LINC00282	0,2562	1,1943	0,0007	0,0735	*no annotation*
ENSG00000270426	RP11_326I11	0,2587	1,1964	0,0002	0,0406	*no annotation*
ENSG00000166405	RIC3	0,2616	1,1988	0,0002	0,0406	resistance to inhibitors of cholinesterase 3
ENSG00000213928	IRF9	0,2628	1,1998	0,0004	0,0615	interferon regulatory factor 9
ENSG00000164124	TMEM144	0,2651	1,2017	0,0013	0,0915	transmembrane protein 144
ENSG00000244482	LILRA6	0,2656	1,2021	0,0008	0,0736	leukocyte immunoglobulin-like receptor subfamily A, member 6
ENSG00000151726	ACSL1	0,2662	1,2027	0,0001	0,0397	acyl-CoA synthetase long-chain family member 1
ENSG00000135636	DYSF	0,2687	1,2047	0,0001	0,0393	dysferlin, limb girdle muscular dystrophy 2B (autosomal recessive)
ENSG00000129467	ADCY4	0,2700	1,2058	0,0001	0,0395	adenylate cyclase 4
ENSG00000188313	PLSCR1	0,2737	1,2089	0,0005	0,0677	phospholipid scramblase 1
ENSG00000119121	TRPM6	0,2741	1,2093	0,0001	0,0397	transient receptor potential cation channel, subfamily M, member 6
ENSG00000280832	ST3GAL4	0,2745	1,2096	0,0008	0,0736	CMP-N-acetylneuraminate-beta-galactosamide-alpha-2,3-sialyltransferase
ENSG00000145491	ROPN1L	0,2818	1,2157	0,0002	0,0406	ropporin 1-like
ENSG00000128383	APOBEC3A	0,2826	1,2164	0,0000	0,0276	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3A
ENSG00000239653	PSMD6	0,2850	1,2184	0,0000	0,0025	26 S proteasome non-ATPase regulatory subunit 6
ENSG00000123360	PDE1B	0,2864	1,2196	0,0001	0,0381	phosphodiesterase 1B, calmodulin-dependent
ENSG00000107281	NPDC1	0,2921	1,2244	0,0003	0,0525	neural proliferation, differentiation and control, 1
ENSG00000224950	RP5_1086K	0,2925	1,2247	0,0002	0,0406	*no annotation*
ENSG00000160883	HK3	0,2982	1,2296	0,0000	0,0025	hexokinase 3 (white cell)
ENSG00000157551	KCNJ15	0,3164	1,2452	0,0001	0,0295	potassium inwardly-rectifying channel, subfamily J, member 15

Appendix 2: Procedure for qPCR for validation of gene expression

All reagents were thawed and kept on ice.

A master mix was prepared:

qPCR master mix	
Reagent	**Volume per sample (µL)**
Evagreen Sso Fast Master mix	5,0
ddH_2O	2,0
Primer mix*	1,0
Total volume	8,0

* Primer mix contains forward and reverse primers.

1. 2,0 µl ddH2O were added to the non-template control (NTC) wells on a 96 well plate.
2. 2,0 µl sample cDNA (2,5 ng/µl) were added to the PCR plates in two parallels each.
3. 8,0 µl of the qPCR master mix was loaded to each well.
4. The PCR plate was sealed with a plastic film and sentrifuged at 1500G for 1 min.
5. The qPCR reaction was run at the following schedule:

Duration	**Temperature**	**Cycles**
10 min	95°C	x1
5 sek	95°C	x45
15 sek	60°C	

Appendix 3: Procedure for miRNA extraction using 'miRNeasy Serum/Plasma kit' (Qiagen, Hilden, Germany)

All reagents were thawed and kept on ice.
QIAzol master mix was made:

QIAzol master mix	
Reagent	**Volume per sample**
QIAzol	1000 µL
0,8 µg/µL MS2 RNA	1,25 µL

1. 250 µL of each serum sample was added to a 1,5 ml Eppendorf tube and sentrifuged for 5 min at 3000G and 4°C. 200 µL of the upper phase was then transferred to a new 2,0 ml Eppendorf tube.
2. 1000 µL QIAzol master mix was added to each sample. Samples were mixed by vortexing for 10 sec and spun down before a 5 min incubation at room temperature (RT)
3. 5 µL 5 nM syn-cel-miR-39 was added to each sample.
4. 200 µL chloroform was added to each sample and mixed well by pipetting up and down. Samples were spun down and incubated for 3 min at RT.
5. Samples were sentrifuged for 15 min at 12000 G and 4°C resulting in phase division.
6. 700 µL of the upper aqueous phase was transferred to new 2,0 ml Eppendorf tubes and 1000 µL absolute ethanol (EtOH) was added to each sample and samples were mixed by pipetting up and down.
7. 650 µL of the mix was transferred to an RNeasy Mini Spin column in collection tubes and sentrifuged for 30 sec at 16000G and RT. Flow-through was discarded. This step was repeated three times until the whole sample volume had been through the column.
8. The column was washed once with 700 µL RWT-buffer and sentrifuged for 1 min at 17000 G and RT. Flow-through was discarded.
9. The column was washed three times with 500 µL RPE-buffer and sentrifuged for 1 min at 17000 G and RT. Flow-through was discarded.
10. Colums were transferred to new collection tubes and sentrifuged for 2 min at at 17000 G and RT, followed by drying for 5 min at the bench with open caps.
11. RNA was eluted using 50 µL nuclease-free water applied directly onto the membrane without touching it with the pipette tip followed by incubation at RT for 1 min and 1 min sentrifugation at 17000 G and RT. Flow-through was applied to the membrane once more including incubation and sentrifugation to increase RNA yield.
12. miRNA was stored at -20 °C

Note: In some of the columns, a drop remained at the edge of the membrane after the last sentrifugation. This drop was transferred to the middle of the membrane and the column was sentrifuged one last time with the same conditions as in step 11.

Appendix 4: Procedure for cDNA synthesis using 'miRCURY LNA™ microRNA PCR, Polyadenylation and cDNA synthesis kit II' (Exiqon, Qiagen, Hilden, Germany)

All reagents were thawed and kept on ice.

A master mix was prepared:

cDNA synthesis master mix		
Reagent	**Volume per sample (µL)**	**Volume per –RT sample (µL)**
5x Reaction buffer	2,0	2,0
ddH_2O	5,0	6,0
Reverse transcriptase enzyme mix	1,0	-
Total volume	8,0	8,0

* Primer mix contains forward and reverse primers.

1. The miRNA samples were vortexed and spun down for 10 sec.
2. 2 µl of sample miRNA and 8 µl of master mix were transferred to 0,5 ml Eppendorf tubes. Pooled sample was used for the –RT reaction. Tubes were vortexed and spun down for 10 sec.
3. The reverse transcription reaction was run on the PCR machine at the following program:

Temperature	**Duration**
42°C	60 min
95°C	5 min
4°C	∞

4. cDNA samples were diluted 10x with nuclease free water and stored at -20°C.

Appendix 5: Procedure for qPCR for miRNA analysis

All reagents were thawed and kept on ice.

A master mix was prepared:

qPCR master mix	
Reagent	**Volume per sample (µL)**
SYBR Green master mix	10,0
Primer mix*	2,0
Total volume	12,0

* Primer mix contains forward and reverse primers.

1. 8,0 µl ddH2O were added to the non-template control (NTC) wells on a 96 well plate.
2. 8,0 µl sample cDNA were added to the PCR plates in three parallels each.
3. 12,0 µl of the qPCR master mix was loaded to each well.
4. The PCR plate was sealed with a plastic film and sentrifuged at 1500G for 1 min.
5. The qPCR reaction was run at the following schedule:

Duration	**Temperature**	**Cycles**
10 min	95°C	x1
10 sek	95°C	x45
1 min	60°C	

Appendix 6: Procedure for DNA extraction using 'DNeasy kit' (Qiagen, Hilden, Germany)

All reagents were thawed and kept on ice.

1. 2 ml anticoagulated blood was added to a 2 ml Eppendorf tube and sentrifuged for 15 min at 5000G and room temperature (RT).
2. The supernatant was carefully removed and discarded, and the pellet was washed once with 400 µL PBS. 400 µL PBS was then added again and the pellet was resuspended by pipetting up and down.
3. 40 µL proteinase K and 400 µL Buffer AL was added to the samples, and mixed immediately by vortexing for 10 sec. All samples were incubated at 56°C for 10 min.
4. 400 µL absolute ethanol (EtOH) was added to each sample and mixed by vortexing for 10 sec.
5. 700 µL of the solution was transferred to DNeasy mini spin columns in 2 ml collection tubes and sentrifuged for 1 min at 6000G and RT. Flow-through was discarded.
6. The remaining volume of the solution was transferred to the columns and sentrifuged for 1 min at 6000G and RT. Flow-through was discarded.
7. The column was transferred to a new collection tube and 500 µL of Buffer AW1 was added. Samples were sentrifuged for 1 min at 6000G and RT. Flow-through was discarded.
8. The column was transferred to a new collection tube and 500 µL of Buffer AW2 was added. Samples were sentrifuged for 3 min at 20 000G and RT. The column was removed carefully from the collection tube and flow-through was discarded.
9. The column was transferred to a 1,5 ml Eppendorf tube and 100 µL Buffer AE was applied directly onto the membrane without touching it with the pipette tip followed by incubation at RT for 1 min and 1 min sentrifugation at 6000G and RT. Flow-through was applied to the membrane once more including incubation and sentrifugation to increase RNA yield.
10. DNA was stored at -20 °C

Appendix 7: Procedure for visualization of fragmented samples using D1000 Screen Tape (Agilent Technologies, Santa Clara, CA, USA)

Reagents were equilibrated to room temperature for 30 minutes and vortexed before use.

1. 3 μL D1000 Sample Buffer was mixed with 1 μL D1000 Ladder.
2. DNA samples were prepared by mixing 3 μL D1000 Sample Buffer with 1 μL DNA sample.
3. All samples were spun down and vortexed at 2000 rpm for 1 min.
4. All samples were spun down.
5. Samples were loaded into the 2200 TapeStation instrument and analyzed.

Appendix 8: Procedure for DNA extraction using 'Magnetic methylated DNA Immunoprecipitation kit' (Diagenode, Liege, Belgium)

All reagents were thawed and kept on ice. Magnetic rack was kept on ice.
Bead wash buffer was made and kept on ice:

Bead wash buffer	
Reagent	**Volume(µL)**
Magbuffer A	928
Nuclease free water	4640

1. Immunoprecipitation

1.1 638 µL Magbeads were transferred to a 1,5 ml Eppendorf tube and resuspended by pipetting up and down. The tube was placed in the magnetic rack and supernatant was discarded. 1276 µL bead wash buffer was added and beads were reuspended by pipetting up and down. The tube was placed in the magnetic rack, and supernatant was discarded. Bead wash was repeated once. Beads were resuspended in new 1276 µL bead wash buffer and placed on ice.

1.2 IP incubation mix was prepared in a total of 58 1,5 ml Eppendorf tubes:

IP Incubation mix (IP mix)		
	Samples (54 tubes)	**Pooled samples (4 tubes)**
Reagent	Volume pr tube (µL)	Volume pr tube (µL)
Nuclease free water	48	45
Magbuffer A	24	24
Magbuffer B	6	6
Positive meDNA control	-	1,5
Negative unDNA control	-	1,5
DNA sample	12	12
Total volume	**90**	**90**

1.3 Tubes with IP mix were incubated for 3 min at 95°C and quickly chilled on ice before a quick spin down.

1.4 75 µL (IP sample) and 7,5 µL (input) of each IP mix was transferred to new 1,5 ml Eppendorf tubes. Input tubes were kept at 4°C.

1.5 7 µL of the anti-5mC antibody (ab) was diluted with 14 µL of nuclease free water.

1.6 Antibody mix was prepared (with the reagents added in the following order):

Antibody mix (ab mix)	
Reagent	**Volume (µL)**
Diluted ab	21
Magbuffer A	42
Nuclease free water	147
Magbuffer C	138
Total volume	**348**

1.7 5 µL of ab mix was added to each IP sample and resuspended by pipetting up and down.
1.8 20 µL of washed beads were added to each IP sample and incubated on a rotating wheel at 40 rpm and 4°C over night (ON).

2. Washes

Complete buffer DIB was prepared:

Complete buffer DIB	
Reagent	**Volume (µL)**
Proteinase K	8
Buffer DIB	800

2.1 IP samples were placed in the magnetic rack for 1 min before supernatant was removed and discarded.
2.2 Beads were washed 3x by the following procedure: 150 µL Magwash buffer-1 was added to the beads and resuspended by inverting the tube. Tubes were incubated for 4 min on a rotating wheel at 40 rpm and 4°C, spun down and placed in the magnetic rack. After 1 min in the rack, the supernatant was removed and discarded.
2.3 Beads were washed 1x following the same procedure as above, but using Magwash buffer-2.

3. DNA isolation

3.1 IP samples were removed from the magnetic rack and 100 µL complete DIB was added to each IP sample followed by resuspension by pipetting up and down.
3.2 Input samples were quickly spun down and 92,5 µL of complete DIB was added to each input sample followed by resuspension by pipetting up and down.
3.3 All samples were incubated at 55°C for 15 min.
3.4 All samples were incubated at 100°C for 15 min.
3.5 All samples were spun down and IP samples were placed in the magnetic rack. Supernatants were transferred to new 1,5 ml Eppendorf tubes.
3.6 Methylated DNA was stored in aliquots at -20 °C

Appendix 9: Procedure for qPCR for DNA methylation analysis

All reagents were thawed and kept on ice.

A master mix was prepared:

qPCR master mix	
Reagent	**Volume per sample (µL)**
SsoFast EVAgreen master mix	12,5
ddH_2O	6,5
Primer mix*	1,0
Total volume	20,0

* Primer mix contains forward and reverse primers.

1. 5,0 µl ddH2O were added to the non-template control (NTC) wells on a 96 well plate.
2. 5,0 µl sample meDNA were added to the PCR plates in three parallels each.
3. 20,0 µl of the qPCR master mix was loaded to each well.
4. The PCR plate was sealed with a plastic film and sentrifuged at 1500G for 1 min.
5. The qPCR reaction was run at the following schedule:

Duration	**Temperature**	**Cycles**
7 min	95°C	x1
15 sek	95°C	x45
1 min	60°C	
1 min	95°C	x1

Appendix 10: Promoter sequence and primer sequences for *IRF7*

>NG_029106.1:5001-8445 Homo sapiens interferon regulatory factor 7 (IRF7), RefSeqGene on chromosome 11

TATA box

GAAACTCCCGCCTGGCC**ACCATAAAAGCGCCGG**CCCTCCGCTTCCCCGCGAGACGAAAC

Transcription initiation

TTCCCGTCCCGGCGGCTCTGGCACCCAGGTACTGGGGACCCCAGACCCACGCGGTGCAG

GCCGGGAGCGAGAGCCTCCGTGGGGGCTCCGTGACCCCGGAGGGGTAGAGCCAAGAGCT

GGGGGAGCCTGAGAGATGAGGGTCGGGCGGGGAGGGAGGCGGAGGCGGAGGCGGGGTTC

CGCGGAGCTGAGAACCGGACGGGGTGGGATCGAGGAGGGTG**CG**AAG**CG**C

Fwd primer

CACTGTTTAGGTTT**CG**CTTTCC**CG**GGAGCCTGACC**CG**CCCCTGA**CG**T**CG**CCTTTCC**CG**T

CTCCGCAGGGTCCGGCCTGCGCCTTCCCGCCAGGCCTGGACACTGGTTCAACACC

Rev primer

TGTGACTTCATGTGTGCGCGCCGGCCACACCTGCAGTCACACCTGTAGCCCCCTCTGCC

AAGAGATCCATACCGAGGCAGCGTCGGTGGCT

ACAAGCCCTCAGTCCACACCTGTGGACACCTGTGACACCTGGCCACACGACCTGTGGCC

GCGGCCTGGCGTCTGCTGCGACAGGAGCCCTTACCTCCCCTGTTATAACACCTGACCGC

IRF-7 A

CACCTAACTGCCCCTGCAGAAGGAGCA**ATG**GCCTTGGCTCCTGAGAGGTAAGAGCCCGG

CCCACCCTCTCCAGATGCCAGTCCCCGAGCGCCCTGCAGCCGGCCCTGACTCTCCGCGG

Figure S1: Promoter sequence of IRF7, recovered from National Center for Biotechnology Information (NCBI), U.S. National Library of Medicine, 20.11.2017. The TATA box is highlighted. The transcriptional initiation site, and the transcriptional start site of the IRF-7A isoform are marked with arrows. Some of the CpG sites that are methylated in 2fTGH cells [74] are indicated by bold underlining. The forward primer is marked in green, and the reverse primer is marked in yellow.

Appendix 11: Promoter sequence and primer sequences for *CD79A*

>NG_009619.1:5001-9250 Homo sapiens CD79a molecule (CD79A), RefSeqGene (LRG_42) on chromosome 19

TCCACTCACAGCCTGAAGCATAC**CCGG**CAGGGGCTGTCCCCAGGCCCAACAAGCAAAGGGCCCAGTAGCG

AGGGCCACTGGAGCCCATCT**CCGG**GGGGCTGGGCAGGAAGTAGGGTGGGGTTTGGGGTAGGGATCTGGTA

Fwd primer

CCCTGGGACTGCTGCAACTCAAACTAACCAACCCACTGGGAGAAGATGCCTGGGGGTCCAGGAGTCCTCC

Rev primer

AAGCTCTGCCTGCCACCATCTTCCTCCTCTTCCTGCTGTCTGCTGTCTACCTGGGTATGTGGCCAAAGGG

CAGGAACTGGCGGGAGGTGGGGGAAGCTGTGGAGGCTGCAGAGAGGGCACAGGCAGAGGGAAGGGGGCTC

AGGGAAAGGGGAAGAGGAGGCAGAGGATAGGGGACCCAGGGAAGATGCCTATAGAAATCGTATCTGTGCC

AAGATGGGCCAAGGTGGGGCTGGAGGGAGCCCAGCGAAGGAGAAGGGGCGTCCACAGTCTCACACAGGGA

GGCAGGAGCAAGAGTCACCTCCCCCACCTCCTGTTCCCCACAGGCCAAAATAAGGAACTAAAGTTGCTCT

TGACTGAGCACCAGGGCTGGGGGCAGGAAGGGGACTTAGGGGTAGCAGCATTCAGCGTCTGTCAAGGGGA

GAAAAAGCTTTCTCTGCCTTAAACCTCAGGTGCCTCTCTCTGTTGGGAGTCCCTTCTCAGCACTGGGGGA

Figure S2: Part of promoter sequence of CD79A, recovered from National Center for Biotechnology Information (NCBI), U.S. National Library of Medicine, 24.11.2017. *MspI/HpaII* sites that are specifically demethylated in CD79A-expressing cells [43] are indicated by bold underlining. The forward primer is marked in green, and the reverse primer is marked in yellow.