



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Characterisation of Molecular Causes for
Different ABCD2 Gene Expression
in T-cells and Monocytes“

verfasst von / submitted by

Johannes Binder, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien, 2017 / Vienna 2017

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 066 834

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Masterstudium Molekulare Biologie

Betreut von / Supervisor:

Univ.Prof. Dr. Johannes Berger

Acknowledgement:

I want to thank my supervisor Univ. Prof. Dr. Johannes Berger for providing me with guidance and support during my master's thesis. In addition, I want to thank my colleagues of the department of pathobiology of the nervous system, for the great work atmosphere and helpful discussions.

Table of Contents:

1. Introduction	
1.1. X-linked Adrenoleukodystrophy	6
1.2. Structure and Known Regulators of the <i>ABCD2</i> Promoter	9
2. Aim of the Study	15
3. Materials and Methods	
3.1. Phylogenetic Footprinting	16
3.2. Transcription Factor ChIP-Seq	17
3.3. DNase Hypersensitivity	19
3.4. <i>In Vivo</i> Footprinting	20
4. Results	
4.1. Phylogenetic Footprinting/ChIP-Seq	26
4.2. DNase Hypersensitivity	33
4.3. <i>In Vivo</i> Footprinting	35
5. Discussion	
5.1. <i>In Vivo</i> Footprinting	38
5.2. Phylogenetic Footprinting/ChIP-Seq	40
6. List of Abbreviations	50
7. References	52
8. Appendix	
8.1. List of Figures	63
8.2. List of Tables	63
8.3. Abstract (German)	64
8.4. Abstract (English)	66

1. Introduction:

1.1. X-linked Adrenoleukodystrophy:

X-linked Adrenoleukodystrophy (X-ALD) is a monogenetic disorder caused by mutations in the X-chromosomal gene *ABCD1* (Mosser et al., 1993). The combined birth incidence for hemizygotes and heterozygotes was determined to be around 1 in 16800 new-borns (Bezman et al., 2001; Kemp et al., 2001). This incidence seems to be similar across the world (Wiesinger et al., 2015).

The *ABCD1* gene codes for a peroxisomal ABC-transporter, which is involved in the degradation of very long-chain fatty acids (VLCFAs) (van Roermund et al., 2008). Under healthy conditions VLCFA-CoA esters are transported by ABCD1 across the peroxisomal membrane (Wiesinger et al., 2013). Following this import, VLCFAs are degraded by β -oxidation (see Fig. 1). In X-ALD, due to dysfunctional ABCD1, VLCFAs accumulate in the plasma and tissues of patients (Moser et al., 1981; Wiesinger et al., 2013). It is thought that this excess of VLCFAs destabilizes myelin membranes, which results in progressive axonal demyelination (Moser, 1997).

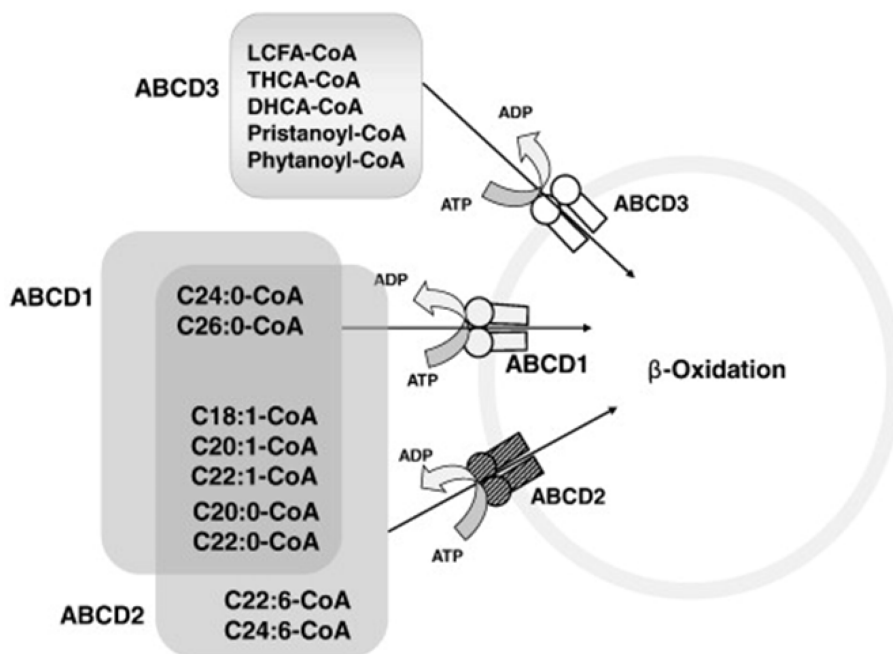


Figure 1: *ABCD1* and the two homologous genes *ABCD2* and *ABCD3*, code for peroxisomal ABC-transporters. *ABCD1* and *ABCD2* have overlapping substrate specificities (Morita and Imanaka, 2012). The transport of VLCFAs into peroxisomes is an essential step for their β -oxidation. This figure is used with the permission of Elsevier.

X-ALD was categorized in two main phenotypes, the milder adrenomyeloneuropathy (AMN) and the more severe inflammatory cerebral adrenoleukodystrophy (cALD) (Berger et al., 2014; Kemp et al., 2012). There is no clear genotype-phenotype correlation, meaning that it is not known why some patients develop AMN, while others develop cALD (Berger et al., 1994; Berger and Gartner, 2006; Smith et al., 1999). In AMN, the progressive loss of isolating myelin in the spinal cord and the periphery causes spastic paraplegia, the main symptom of the disease (Budka et al., 1976). Patients with cALD develop progressive inflammatory lesions in the white matter of the brain and usually die within a few years. Due to unknown reasons, some AMN patients can shift to the inflammatory cALD form (van Geel et al., 2001).

So far, the only viable treatment option for cALD is bone marrow transplantation (Aubourg et al., 1990). The procedure uses hematopoietic stem cells (HSCs) from a healthy donor. These cells are then transplanted into the patient, where they repopulate the bone marrow and differentiate into various hematopoietic cells. HSC transplantation (HSCT) is able to normalize VLCFA levels in patients and has been successfully used for many years now (Shapiro et al., 2000). In some cases, HSCT can lead to graft-versus-host disease, where immune cells of the patient attack donor-derived cells. To prevent such complications it is possible to use autologous HSCs for the therapy of X-ALD (Cartier et al., 2009). This variant uses patient derived HSCs, which are isolated and treated with gene therapy to correct their genotype. After receiving a functional copy of *ABCD1*, the HSCs are transplanted back into the patient (Cartier et al., 2009). Autologous HSCT has the advantage of preventing graft-versus-host disease because all transplanted cells come from the patient.

After a successful HSCT, it can take 6-12 months until the inflammation stops (Peters et al., 2004). Because of that, the procedure is not suitable for patients where cALD was diagnosed at a too late stage of the disease. Due to this narrow time frame, the lack of matching donors and the high costs of the procedure, HSCT is often not a viable treatment option. Furthermore, it appears that HSCT only helps with the inflammatory form of the disease, which means that after a successful transplant, patients can still develop AMN later in life (van Geel et al., 2015).

To help a wider range of patients it would be beneficial to pharmacologically treat the disease (Berger et al., 2010). For this purpose, the homologous gene *ABCD2* has been in the focus of attention. *ABCD2* also codes for a peroxisomal membrane transporter, which has

overlapping substrate specificities with ABCD1 (see Fig. 1) (van Roermund et al., 2011). Because of this overlap, ABCD2 can compensate for ABCD1 when expressed at high enough levels. This was exemplified in X-ALD fibroblasts and *ABCD1*^{-/-} mice, where overexpression of *ABCD2* successfully lowered VLCFA-levels and even reduced neurodegenerative features in the mice (Netik et al., 1999; Pujol et al., 2004). Moreover, it could be demonstrated that endogenous ABCD2 can prevent a more severe metabolic phenotype in selected cell types of the *ABCD1* deficient mouse model (Forss-Petter et al., 1997; Muneer et al., 2014).

Interestingly *ABCD2* has distinct expression levels in several blood cell types, with high amounts in T cells and low amounts in monocytes (see Fig. 2) (Weber et al., 2014b). Furthermore, its levels show an inverse correlation when compared to *ABCD1*. This is very pronounced in T cells and monocytes. While T-cells have rather high levels of *ABCD2* and low levels of *ABCD1* expression, monocytes show an opposite pattern (Weber et al., 2014b). Probably because of that, T cells are less affected in X-ALD because ABCD2 can compensate upon ABCD1 deficiency. Monocytes on the other hand, with their low endogenous ABCD2 levels, cannot compensate for ABCD1 deficiency. It was shown that in X-ALD among several blood cell types, mainly granulocytes and monocytes are affected. This is illustrated by the accumulation of the VLCFA hexacosanoic acid (C26:0) in these cells. Other cell types with higher *ABCD2* expression like B and T cells show almost no accumulation of C26:0 (Weber et al., 2014b). Therefore, a proposed treatment option for X-ALD is the pharmaceutical upregulation of *ABCD2* in monocytes, to normalize VLCFA-levels (Berger et al., 2014, 2010).

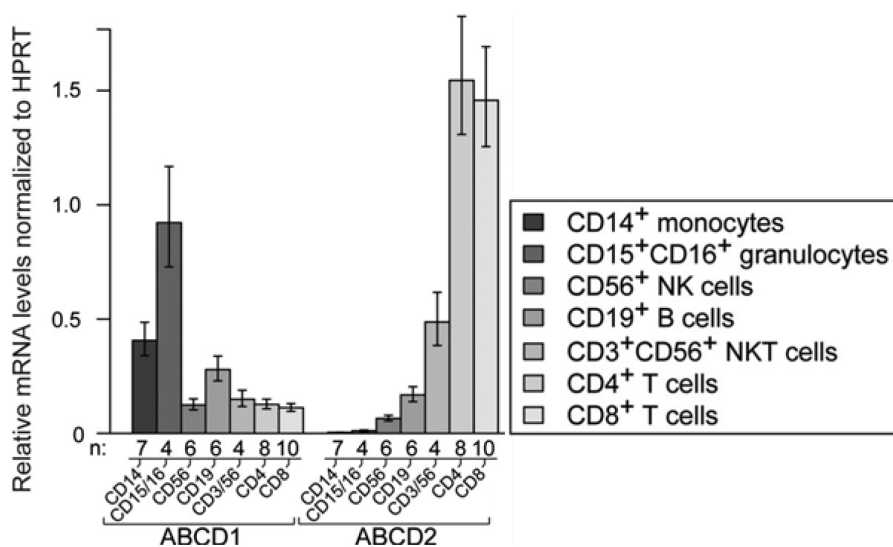


Figure 2: Expression levels of *ABCD1* and *ABCD2* in different immune cells (Weber et al., 2014b).
This figure is used with the permission of Oxford University Press.

1.2. Structure and Known Regulators of the *ABCD2* Promoter:

The promoter of *ABCD2* has been extensively studied for some time now, and considerable progress was made in understanding its regulation (Rampler et al., 2003; Weinhofer et al., 2008, 2002). However so far, attempts to induce *ABCD2* to higher levels have proven to be quite difficult. The reason for this is a complex interplay between various regulatory factors. Some of those factors and their respective mechanisms have been described in the past, including cholesterol levels, nuclear hormone receptors and epigenetic factors (Fourcade et al., 2003; Gondcaille et al., 2005; Rampler et al., 2003; Weber et al., 2014a; Weinhofer et al., 2005, 2002).

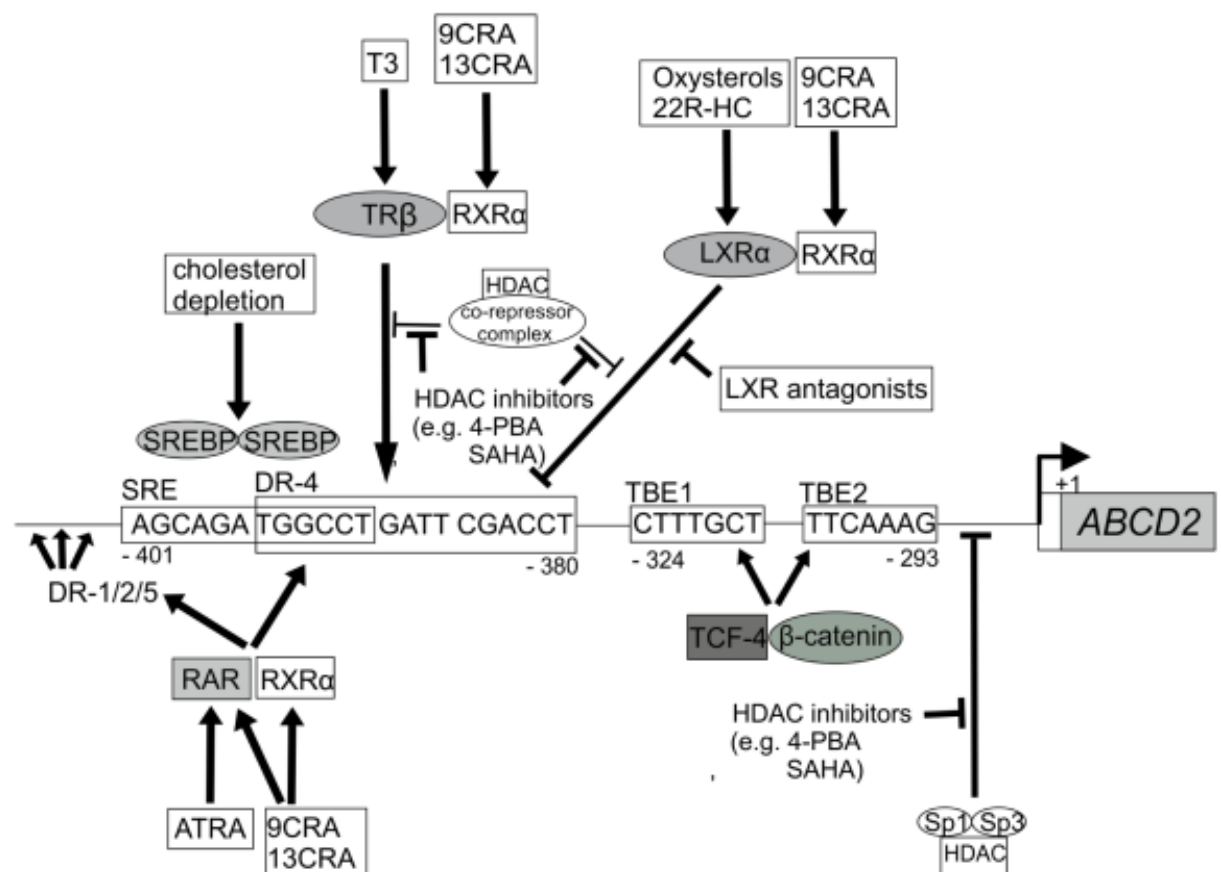


Figure 3: Known regulatory mechanisms at the *ABCD2* promoter.

This figure is used under the Creative Commons Attribution License (Weber et al., 2014a).

The promoter of *ABCD2* contains several functional elements that serve as target sites for regulatory proteins (see Fig. 3) (Weber et al., 2014a). Most of the known regulators bind to the SRE/DR-4 element, which consists of a sterol regulatory element (SRE) overlapping with a direct repeat sequence (DR-4) (Weinhofer et al., 2008).

The SRE serves as a binding site for SRE-binding proteins (SREBPs) that act as cellular cholesterol sensors. At normal sterol concentrations, the SREBP precursor protein is bound to the nuclear envelope and endoplasmic reticulum (Wang et al., 1994). Upon sterol depletion, the precursor protein becomes activated by cleavage and translocates into the nucleus, where it regulates genes involved in the sterol and fatty acid metabolism. In 2002, Weinhofer et al. identified the SRE in the promoter of *ABCD2*. It has been shown that SREBP1a binds directly to the SRE in the human *ABCD2* promoter and is necessary for the efficient induction of *ABCD2* (Weinhofer et al., 2002). In addition, it was shown that cholesterol depletion in X-ALD fibroblasts can increase *ABCD2* expression, which is accompanied by a reduction of VLCFA levels (Weinhofer et al., 2002).

In vivo, the effects of cholesterol lowering drugs are not as clear. There are contradictory results regarding the effect of lovastatin or simvastatin, on VLCFA levels and X-ALD pathology (Cartier et al., 2000; Singh et al., 1998). In a recent clinical study, lovastatin failed to reduce cellular C26:0 levels in erythrocytes and lymphocytes, and lead only to a small transient decrease of plasma VLCFAs (Engelen et al., 2010). Therefore, the authors reasoned that lovastatin is of no clear benefit for the treatment of X-ALD.

Because the SRE overlaps with a DR-4 element, other factors may interfere with SREBP binding and modulate the outcome (Weinhofer et al., 2005). The DR-4 element consists of a direct repeat that is separated by 4 nucleotides. Due to this symmetric structure, various nuclear hormone receptor (NHR) dimers can bind to this site (Aranda and Pascual, 2001). NHRs like the thyroid hormone receptor (TR) or the retinoic acid receptor (RAR), can bind DNA either as homodimers or as heterodimers together with RXR. While both is possible, RXR heterodimers show an increased affinity for their target sequence and a stronger transcriptional activity (Aranda and Pascual, 2001). In the absence of their corresponding ligand, some NHRs including TR and RAR, are bound to their target sequence and repress transcription. These repressive effects are mediated by certain domains of the receptor, or by recruitment of inhibitory proteins (Baniahmad et al., 1992; Casanova et al., 1994).

Upon ligand binding repressive factors are removed. This allows the interaction with transcriptional co-activators, like the histone acetyltransferase p300 to promote gene expression (Aranda and Pascual, 2001).

The NHRs liver X receptor alpha (LXR α), RAR α and TR β , were all reported to bind at the DR-4 element in the *ABCD2* promoter to regulate transcription (Fourcade et al., 2003; Pujol et al., 2000; Weinhofer et al., 2005). To induce *ABCD2*, several treatments that target these receptors were done. For example by treating primary monocytes with the RAR/RXR ligand 13-cis-retinoic acid, mRNA levels were increased significantly (Weber et al., 2014a). Similarly, treatment with the TR ligand tri-iodothyronine (T3) was shown to elevate *ABCD2* expression levels, either by TR α -mediated activation or by preventing TR β -mediated repression (Fourcade et al., 2003; Weinhofer et al., 2008). A repressive effect on *ABCD2* was also shown for LXR α . Accordingly treatment with LXR α antagonists upregulates *ABCD2* expression, while treatment with LXR α agonists achieves the opposite effect (Gondcaille et al., 2014).

Besides the already mentioned transcription factors, peroxisome proliferator-activated receptor-alpha (PPAR α) is also implicated in the regulation of *ABCD2* (Fourcade et al., 2001). PPAR α is an important regulator of the lipid metabolism and is primarily expressed in tissues with high β -oxidation activity (Gervois et al., 2000; Issemann and Green, 1990). To regulate gene expression, PPAR α forms a heterodimer with RXR, which then binds to a peroxisome proliferator hormone response element (PPRE) (Kliwer et al., 1992). Agonists of PPAR α like fenofibrates, can upregulate the expression of *ABCD2* (Berger et al., 1999). Although there are several putative PPREs in the promoter of *ABCD2*, none of them is able to induce expression upon fibrate treatment (Rampler et al., 2003). Therefore, it was proposed that PPAR α exerts its effect indirectly, by lowering cholesterol levels and thereby activating SREBP2. SREBP2 then in turn could mediate the observed induction of *ABCD2* (Gervois et al., 2000; Rampler et al., 2003).

Not only cholesterol, but also some polyunsaturated fatty acids (PUFAs) have an influence on *ABCD2* expression. Mice with a diet deficient in omega-3 fatty acids show a significant increase in hepatic *ABCD2* expression (Leclercq et al., 2008). Since *ABCD2* is also involved in the synthesis of the omega-3 fatty acid DHA (Docosahexaenoic acid), a feedback loop may regulate gene expression to maintain homeostasis (Fourcade et al., 2009; Leclercq et al., 2008). Several transcription factors like PPARs, SREBPs, LXR, NF- κ B and others are regulated

by PUFA levels (Keller et al., 1993; Novak et al., 2003; Sekiya et al., 2003; Yoshikawa et al., 2002). PUFA mediated gene regulation can work, by preventing LXR from binding to the LXR response element. Via this mechanism, PUFAs have a suppressive effect on the expression of the SREBP1c gene (Yoshikawa et al., 2002).

Apart from the SRE/DR-4 element, also other regulatory regions have been described. A few base pairs downstream a CCAAT-element, a GC1-box and two T-cell factor/lymphoid enhancer factor (TCF/LEF)-binding elements (TBEs) are located (Fourcade et al., 2001; Park et al., 2013). The CCAAT-element is used as a binding site for the trimeric nuclear transcription factor-Y (NF-Y), consisting of subunit A, B and C (Gondcaille et al., 2005; Sinha et al., 1995). NF-Y is able to positively influence transcription by interacting with histone acetyltransferases (HATs) like p300 or the p300/CBP associated factor (P/CAF) (Jin and Scotto, 1998; Salsi et al., 2003). HATs can serve as transcriptional co-activators by adding acetyl groups to histone proteins (Brownell et al., 1996). Acetylation masks the positive charge of histones and thereby reduces their interaction with the negatively charged DNA (Kuo and Allis, 1998). As a consequence, DNA is packed less densely which facilitates protein binding and transcription. In the *ABCD2* promoter, it was proposed that such interactions between NF-Y and HATs are used to regulate transcription, but so far no such partner was identified (Gondcaille et al., 2005).

Opposed to HATs, histone deacetylases (HDACs) can remove activating acetylation marks from histones, and thereby mediate a compaction of the chromatin (Kuo and Allis, 1998). Via this mechanism HDACs have a repressive effect on gene expression, which is also used for the regulation of the *ABCD2* gene. In the promoter of *ABCD2* a GC1-box serves as a binding site for the transcription factors SP1 and SP3, which can recruit HDAC1 to represses transcription (Davie, 2003; Doetzlhofer et al., 1999; Gondcaille et al., 2005). Accordingly, treatment with HDAC inhibitors like butyrate or 4-Phenylbutyrate is able to increase *ABCD2* expression rates (Gondcaille et al., 2005). SP1 and SP3 can also bind to the nearby GC2-box. This site however, seems to play no role for HDAC inhibitor dependent activation. Interestingly, the CCAAT-element is necessary for the full effect of HDAC inhibitors, which suggests that the GC1-box and CCAAT-element are both able to interact with HDACs (Gondcaille et al., 2005).

More recently, it was shown that β -catenin and the T cell factor-4 (TCF-4; TCF7L2) regulate *ABCD2* expression by binding at two TCF/LEF-binding elements (TBEs) in a conserved region of the promoter (Park et al., 2013). TCF-4 and β -catenin both belong to the Wnt signalling pathway, which is involved in diverse processes including cellular proliferation, differentiation and tissue homeostasis (Logan and Nusse, 2004; Pinto et al., 2003). In the absence of Wnt signalling, β -catenin is continuously degraded. In this case, TCF proteins are bound to the DNA and silence their target genes via the recruitment of repressive factors like groucho or HDACs (MacDonald et al., 2009). Upon activation of the Wnt pathway, degradation of β -catenin stops and it accumulates in the cytoplasm and nucleus, where it can bind to TCF proteins (Logan and Nusse, 2004). This interaction with β -catenin displaces repressive factors and activates transcription (Daniels and Weis, 2005). For *ABCD2* a strong activation of the promoter was observed after ectopic expression of β -catenin and TCF-4 (Park et al., 2013).

In vivo gene regulation is more complex than simply activating or repressing certain factors. In part, this is because the function of transcription factors can be modulated by posttranslational modifications. The addition of such modifications can alter the activity, stability and other characteristics of proteins (Bannister and Miska, 2000). Since HATs and HDACs do not exclusively target histones, but also other lysine containing substrates, they can produce this effect. P300 mediated acetylation of β -catenin, for example, can favour binding to TCF-4 and reduce interactions with the androgen receptor (Levy et al., 2004). Another example for the effect of such modifications is RXR α . Acetylation by p300 enhances the ability of RXR α to bind DNA, and thereby increases its transcriptional activity (Zhao et al., 2007). Therefore, posttranslational modifications could also play a role in the regulation of *ABCD2*.

Despite all the progress that was made in understanding the promoter of *ABCD2*, its regulation is still not well enough understood. So far, attempts to upregulate the expression of *ABCD2* could not reach the levels found in primary T cells. A reason for this could be yet unknown regulatory elements. Because of their importance for gene expression and for the overall function for the host, regulatory elements are evolutionary conserved (Duret and Bucher, 1997). In general, there is always a certain probability for random mutations to occur. Therefore, the sequence similarity between two species slowly decreases over time.

The rate by which this happens is not identical throughout the genome (Koop, 1995). In non-functional areas of the genome, these mutations can accumulate because they have no severe effects on the fitness of the organism. Conversely, functional elements that are necessary for proper gene expression retain their sequence similarity across greater evolutionary distances. These evolutionary constraints allow it to distinguish functional from non-functional elements (Duret and Bucher, 1997).

Besides the evolutionary conservation, regulatory elements contain also other marks that can be used for their identification. Characteristic features that are often found at these positions range from certain proteins to epigenetic modifications (Heintzman et al., 2007; Visel et al., 2009). A protein that is often found at active enhancers is the histone acetyltransferase p300. P300 acts as a transcriptional co-activator and can be used to identify tissue specific enhancer elements (Hasegawa et al., 1997; Visel et al., 2009).

P300 and other chromatin modifying factors can influence gene expression by depositing epigenetic marks to histone proteins. These posttranslational modifications constitute an additional layer of regulation, positioned on top of the genomic sequence (Strahl and Allis, 2000). Because the genetic code is shared between all cells of the body, such a system is necessary to give rise to the more than 200 cell types found throughout the human body. At enhancers the epigenetic landscape can differ substantially, which reflects cell-type specific differences in gene expression (Heintzman et al., 2009). Active enhancers are characterized by histone 3 lysine 4 mono-methylation (H3K4me1) and lysine 27 acetylation (H3K27ac). Both modifications together were repeatedly used to identify cell type specific enhancers, but it is mainly H3K27ac which distinguishes active from poised enhancers (Creyghton et al., 2010; Heintzman et al., 2009). Histone acetylations like H3K27ac help to make the chromatin more accessible for regulatory factors, and are often found at active promoters and enhancers (Lee et al., 1993). Experimentally, such stretches of open chromatin can be characterized by an increased DNase sensitivity (Gross and Garrard, 1988). While DNase hypersensitivity at promoters is a common feature, sensitivity at distant regulatory regions like enhancers is specific for certain cell types (Xi et al., 2007).

Different usage of such regulatory elements can contribute to cell type specific gene expression (Visel et al., 2009). Therefore, a detailed knowledge of these mechanisms may help to explain the differences of *ABCD2* expression levels between T cells and monocytes.

2. Aim of the Study:

The neurodegenerative disease X-linked Adrenoleukodystrophy (X-ALD) is caused by mutations in the gene *ABCD1*, which codes for a peroxisomal fatty acid transporter. It was shown that the homologous gene *ABCD2* can compensate for *ABCD1* if expressed at high enough levels. Therefore, with this project we aim to better understand the regulation of the *ABCD2* gene promoter. In the future, a detailed understanding of the promoter could help to pharmacologically increase *ABCD2* gene expression levels, as a potential treatment for X-ALD.

3. Materials and Methods:

3.1. Phylogenetic Footprinting:

For the detection of conserved functional regions of the genome, the phylogenetic footprinting procedure can be used (Duret and Bucher, 1997). A phylogenetic footprint is generated, by comparing the sequences of two or more species to determine the degree of conservation between them. Regulatory elements that are important for proper gene expression, show a stronger evolutionary conservation across species (Duret and Bucher, 1997).

Due to the vast amount of data that is produced every day, online databases are filled with useful information. We used pre-existing phylogenetic footprinting data that was accessed via the UCSC (University of California Santa Cruz) genome browser (Karolchik et al., 2003). The genome browser is located on the official website, under the section “Our tools” (<http://genome.ucsc.edu/>). For display the “Human Feb. 2009 (GRCh37/hg19) Assembly” was used at the position “chr12:40,013,187-40,014,311”. With the button “track search” the datasets named “100 vertebrates conservation by PhastCons” and “100 vertebrate conserved elements” can be found and added to the display. The tracks contain information on the sequence conservation across 100 vertebrate species. Among others, the sequences of *Danio rerio* (zebrafish), *Anolis carolinensis* (lizard), *Gallus gallus* (chicken), *Felis catus* (cat), *Pan troglodytes* (chimp) and *Homo sapiens* were used for the generation of this dataset. According to the track description, the data was created with phastCons and phyloP from the PHAST package (<http://compugen.cshl.edu/phast/>).

3.2. Transcription Factor ChIP-Seq:

To understand the function of conserved elements it can be useful to know which transcription factors bind at these sites. Especially chromatin immunoprecipitation sequencing (ChIP-seq) is often used to address this issue. ChIP-seq data is generated by treating cells with a crosslinking agent like formaldehyde. Treatment with formaldehyde connects nearby nucleic acids, proteins and other molecules to preserve the position of proteins on the DNA. In the next step, the crosslinked chromatin is shredded into pieces. With an antibody specific for a protein of interest, protein/DNA clumps are purified from the mixture. After a protease digest, the remaining DNA is sequenced and mapped to the genome. The genomic location and the amount of sequencing reads, produce a characteristic peak pattern. The peaks correspond to sites where the protein of interest was bound. In combination with other data, this technique can provide information about target genes and utilized regulatory mechanisms (Johnson et al., 2007). The results of such ChIP-seq experiments are gathered in online databases.

We used the UCSC (University of California Santa Cruz) genome browser to access the ChIP-seq data (Karolchik et al., 2003). The genome browser is located on the official website, under the section “Our tools” (<http://genome.ucsc.edu/>). For display the “Human Feb. 2009 (GRCh37/hg19) Assembly” was used at the position “chr12:40,010,699-40,018,298”. With the button “track search” the dataset named “Txn Factor ChIP” can be found and added to the display. This track contains the combined data of a large collection of ChIP-seq experiments that were done by the ENCODE project (Karolchik et al., 2003). In total, the dataset contains information about the genomic location of 161 proteins in numerous cell types.

For some proteins of interest, the gene expression levels were determined with the GEO DataSet Browser (<https://www.ncbi.nlm.nih.gov/sites/GDSbrowser>). A search on the official website for the DataSet GDS596, followed by a search for the respective gene name, provides gene expression data for numerous human cell types (GEO accession GSE1133) (Barrett et al., 2013; Su et al., 2004). Some of the mentioned cell types and their corresponding abbreviations are listed in table 1.

Cell Types	
G GM12878; B cell, lymphoblastoid, EBV transf.	K K562; chronic myelogenous leukaemia cell line
g B cell, lymphoblastoid, EBV transf.	H HeLa-S3; cervical carcinoma cell line
u U2OS; osteosarcoma	h HEK293; embryonic kidney cells
1 H1-hESC; embryonic stem cells	p PANC-1; pancreatic carcinoma
L HepG2; hepatocellular carcinoma	s SH-SY5Y; neuroblastoma
M MCF-7; mammary gland, adenocarcinoma	S SK-N-SH_RA; neuroblastoma cell line differentiated with retinoic acid
A A549; epithelial cell line from a lung carcinoma	

Table 1: Mentioned cell types and corresponding abbreviations, taken directly from the UCSC genome browser (Karolchik et al., 2003).

Because T cells and monocytes have very different *ABCD2* expression levels, a comparison between them could reveal distinct regulatory mechanisms (Weber et al., 2014b). Unfortunately, despite the huge amount of ChIP-seq data there is only limited data on primary T cells and monocytes available. Most experiments in the dataset were generated with the lymphoblastoid cell line GM12878. Due to its rather high *ABCD2* expression levels, this cell type may use similar mechanisms as primary T and B cells, for the regulation of the *ABCD2* promoter. Therefore, instead of primary T and B cells the data of GM12878 cells is used. Two cell types with low *ABCD2* expression are K562 and HeLa-S3 cells. K562 is a chronic myeloid leukaemia cell line that shows properties of granulocytic cell types (Lozzio et al., 1981). HeLa is a very commonly used cervical carcinoma cell line. Due to the wide usage of HeLa cells, there is plenty of data available.

3.3. DNase Hypersensitivity:

The degree of chromatin compaction can be investigated by DNase hypersensitivity assays. These assays have been performed worldwide by numerous laboratories. The combined information of these experiments is collected in online databases and is freely available to the public. For our project, the DNase hypersensitivity data was accessed via the UCSC (University of California Santa Cruz) genome browser (Karolchik et al., 2003). The genome browser is located on the official website, under the section “Our tools” (<http://genome.ucsc.edu/>). For display the “Human Feb. 2009 (GRCh37/hg19) Assembly” was used at the position “chr12:40,010,699-40,018,298”. With the button “track search” the dataset named “DNase Clusters” can be found and added to the display. This track contains the combined data of 125 cell lines that were assayed for DNase hypersensitivity by the ENCODE project. The grey bars of the track represent DNase hypersensitive clusters. By clicking on the bars a list of cell types is shown, in which open chromatin was detected at this position. Some of the mentioned cell types are listed in table 2.

Cell Types			
Caco-2	Colorectal adenocarcinoma cells	WI-38	Embryonic lung fibroblast cells
CMK	Acute megakaryocytic leukaemia cells from a donor with Down's syndrome	iPS	Induced pluripotent stem cells derived from skin fibroblasts

Table 2: Selected cell type descriptions, taken directly from the UCSC genome browser (Karolchik et al., 2003).

3.4. In Vivo Footprinting:

To improve our understanding of utilised regulatory regions in T cells and monocytes, an *in vivo* footprinting assay with CD3⁺ and CD14⁺ cells was done. The *in vivo* footprinting procedure creates a snapshot of protein/DNA interactions in the living cell (Hampshire et al., 2007). In general, the procedure was done as outlined by Mueller et al. (2001), with few variations like fluorescent labelling and capillary electrophoresis, to avoid the use of radiolabelled probes (Zianni et al., 2006).

Cell Culture:

To establish the *in vivo* footprinting assay, instead of primary cells, the monocytic cell line THP-1 was used. THP-1 cell aliquots were taken from the liquid nitrogen storage tank. After thawing, the cells were diluted in 20 mL RPMI complete (RPMI-1640 medium supplemented with 10% fetal calf serum (FCS), 1% glutamine, 1% Penicillin/Streptomycin and 0.5% Fungizone). The cells were then incubated in T-75 cell culture flasks (37°C; 5% CO₂). Every 3-4 days the medium was exchanged. For this, the cells were centrifuged in 50 mL Falcon tubes (5 min; 1500 rpm). The resulting pellet was resuspended in 20 mL fresh pre-warmed RPMI complete. The cell suspension was then pipetted back into T-75 cell culture flasks for further incubation (37°C; 5% CO₂). During the establishment of the *in vivo* footprinting procedure, 2*10⁶ cells of this suspension were used in each trial. The cell number was determined with a Casy cell counter.

THP-1 cells that were not used for the procedure, were frozen for long term storage. This was done by centrifugation (5 min; 1500 rpm) and subsequent resuspension of the pellet in filter-sterilized RPMI complete, containing 20% FCS and 10% DMSO. Per 10 million cells, 500 µL of RPMI were used. Aliquots of 500 µL were then pipetted into cryotubes and stored at -80°C in a cryobox filled with isopropanol. After a period of at least 24 h the cryotubes were put in the liquid nitrogen tank.

Isolation of Primary Cells:

For the actual experiment, peripheral blood mononuclear cells (PBMCs) were used. PBMCs were isolated from a leuko-reduction chamber with a Pancoll density gradient (Pan Biotech). CD3⁺ and CD14⁺ cells were purified from the fraction containing the PBMCs, by using MACS-beads (Miltenyi Biotec) according to the manufacturer's protocol. The cell number was determined with a Casy cell counter. For each sample and cell type 2*10⁷ cells were used respectively.

DMS Treatment:

Like in the original protocol, dimethyl sulfate (DMS) is used to methylate unprotected purines in the genome of the cell. For this, the isolated CD3⁺ and CD14⁺ cells were pelleted (500 rpm; 5 min) and resuspended in 1 ml RPMI complete (10% FCS). The cell suspension was pre-warmed to 37°C before adding 10 µl of a freshly prepared 10% DMS solution (in 100% EtOH). DMS can easily pass the plasma membrane, so guanines and adenines become methylated where no protein is bound to the DNA (see Fig. 4) (Mueller et al., 2001). After 1 min incubation at 37°C, the cells were transferred into 49 ml ice-cold phosphate-buffered saline (PBS), mixed and centrifuged (5 min; 500 rpm; 4°C). The pellet was again resuspended in 50 ml ice-cold PBS. Now both *in vivo* and *in vitro* samples were pelleted (5 min; 500 rpm; 4°C) to isolate the DNA with the DNeasy kit (Qiagen) according to protocol.

As a control, DNA of untreated CD3⁺ and CD14⁺ cells was purified to remove interacting proteins and brought to a volume of 200 µl with a vacuum evaporator, and set aside. The *in vitro* samples were concentrated to 175 µl. Then 25 µl of a freshly prepared 1% DMS solution (in dH₂O) were added to the *in vitro* DNA and incubated for 2 min at RT. Treatment of purified DNA with DMS should result in an evenly distributed methylation pattern (see Fig. 4) (Mueller et al., 2001). After that, all samples received 50 µl ice-cold DMS stop buffer. To precipitate the DNA, 3 vol of 100% EtOH (-80°C) were added and put on dry ice for >30 min. After centrifugation (10 min; 4°C) the DNA pellet was washed with 75% EtOH and again centrifuged (10 min; 4°C).

Piperidine Fragmentation:

The DNA was cleaved by adding a freshly prepared piperidine solution (1:10 in dH₂O) to each pellet. After resuspension, the samples were incubated for 30 min at 95°C. Piperidine specifically cleaves at methylated purines, resulting in DNA fragments with varying lengths (see Fig. 4) (Mueller et al., 2001). The piperidine was removed with a vacuum evaporator (1:40 h; 35°C) and the DNA was resuspended in 360 µl TE. For purification 40 µl sodium acetate (pH 5.5) and 2.5 vol of 100% EtOH (-80°C) was added, and put on dry ice for >30 min. After centrifugation (20 min; 4°C) and removal of the supernatant, the precipitation was repeated one more time as described. After that the DNA is washed with 75% EtOH, centrifuged (10 min; 4°C) and the supernatant removed. The pellet was resuspended in 50 µl dH₂O and completely dried in a vacuum evaporator. After resuspension in TE pH 7.5 the concentration was determined.

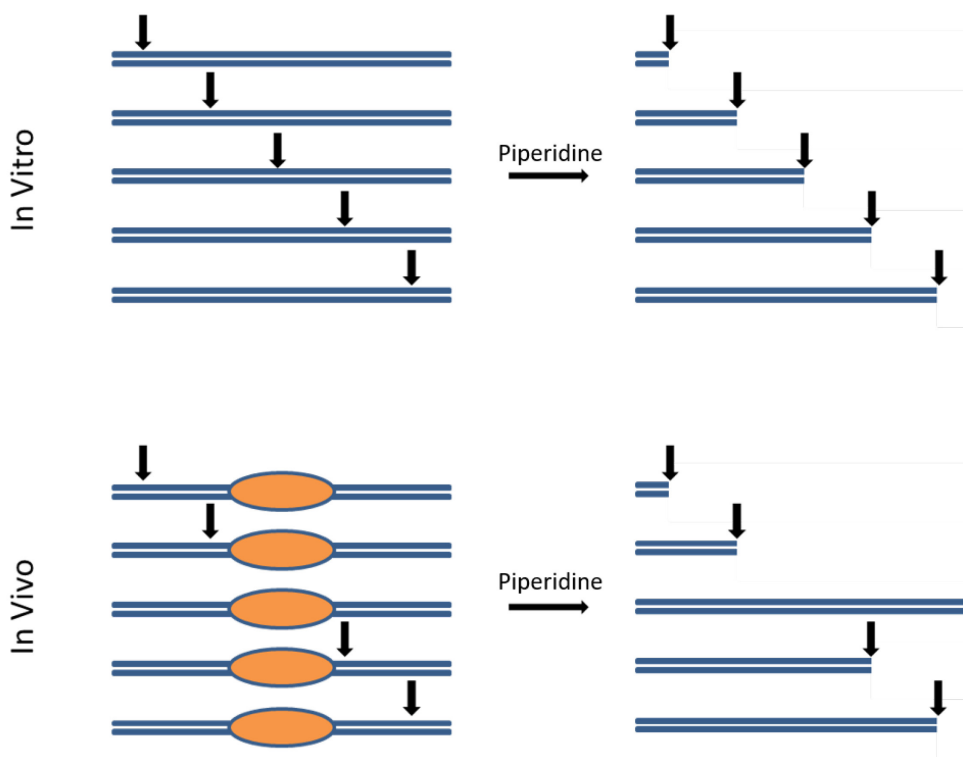


Figure 4: Illustration of the *in vivo* footprinting procedure. DNA is methylated with dimethyl sulfate (DMS) and then cleaved with piperidine (black arrows). Proteins (orange) that are bound to the DNA can protect it from methylation, which results in an underrepresentation of the corresponding fragment size. A comparison between the *in vivo* and *in vitro* sample can reveal sites that were bound by proteins.

Ligation Mediated Polymerase Chain Reaction (LM-PCR):

Before analysis, the fragments were amplified for efficient detection. This is done with a modified version of the polymerase chain reaction, a ligation mediated PCR (LM-PCR). As can be seen in figure 5, an LM-PCR consists of four main parts: First strand synthesis, linker ligation, amplification and end-labelling.

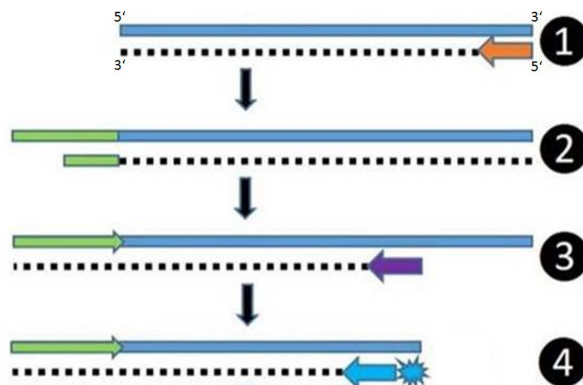


Figure 5: Four main steps of the LM-PCR as outlined by Mueller et al. (2001): (1) First strand synthesis with a gene-specific primer to generate blunt ended fragments. (2) Ligation of a double stranded linker to the blunt end. The linker can only ligate to the cleaved end, because ligation requires a 5' phosphate, which is present on cleaved ends, but not on synthetic oligonucleotides. (3) PCR amplification with a gene and linker specific primer. (4) End-labelling with a 6-FAM labelled primer.

The procedure needs 3 gene specific primers, plus 2 oligonucleotides to generate an asymmetric linker (see Table 3). The primers design is based on the guidelines mentioned in the original protocol (Mueller et al., 2001). In short, Primer 1, 2 and 3 should have increasing melting temperatures (T_m). Additionally, Primer 2 should have a similar T_m as the Linker-Primer (=long linker), as they are used for the amplification step. Primer 3 has to overlap with Primer 2 and extend a few bases, to better compete for the binding site. Primer 3 was labelled with 6-carboxyfluorescein (6-FAM) on the 5' end. The exact sequence of the Linker is not important, as long as it doesn't match any sites in the genome. The short linker is used to stabilize the Linker-Primer and to provide a blunt end on one side. This makes sure that the linker ligates to the fragments in only one direction (Mueller et al., 2001).

Name	Sequence (5'->3')	T _M	BPs from ATG
Primer 1 (OLI2171)	TGTGACAGATGCAGCAGAGC	59,4°C	(-150/-170)
Primer 2 (OLI2172)	CGCTGCATCTACCGGGAATGATT	62,4°C	(-178/-201)
Primer 3 (OLI2176)	CGCTGCATCTACCGGGAATGATTCTCTC	68,0°C	(-178/-206)
Long linker (OLI2174)	GGTGACCCGGGAGATCTGAATTC	64,2°C	
Short linker (OLI2175)	GAATTCAGATC	30,0°C	

Table 3: Sequences and positions of the used oligonucleotides.

For the first strand synthesis 2 µg fragmented DNA, 10 nM Primer 1 (OLI2171), 200 µM dNTPs, buffer (iProof HF) and 0.6 U of iProof polymerase were used in a volume of 30 µL. The PCR had the following settings: 1 cycle with 3 min 98°C (denature), 30 min 62°C (annealing) and 10 min 72°C (extension). Due to the 3'->5' exonuclease activity of iProof, the generated PCR products are blunt ended.

Because the DNA was randomly cleaved with DMS/piperidine, only one side of the fragments is known. The other end depends on the cleavage site. This would be a problem for a normal PCR, because it needs two known sequences on either side of the product for amplification. To solve this problem a double stranded linker (ds-Linker) is ligated to the fragments resulting in a known sequence on this end. To generate the ds-Linker a solution containing 250 mM Tris pH 7.5 and 20 µM of each linker-oligo (OLI2174, OLI2175) was made, and heated it for 5 min to 95°C. Afterwards the solution was gradually cooled down to room temperature over a period of 4 h, and then kept at +4°C over night. The ligation of the ds-Linker to the fragments was done overnight at 17°C. Each reaction contained all of the 1st strand PCR reaction and 1.33 µM ds-Linker, 5% PEG 4000, T4 DNA ligase buffer (Fermentas) and 3 units T4 DNA ligase (Roche) in a final volume of 75 µL. Due to the asymmetric structure of the linker, one side being blunt ended while the other having an overhang, ligation occurs in a known orientation (Mueller et al., 2001).

After ligation, the DNA was purified with sodium acetate and EtOH as described before, however with 0.1 µg yeast tRNA added as a precipitation carrier. After washing with 75% EtOH, the pellets were dried and redissolved in 70 µL dH₂O.

In the next step, the fragments are amplified with a gene- and a linker-specific primer. For this, a reaction with 200 nM linker-primer (OLI2174), 200 nM primer 2 (OLI2172), 200 µM dNTPs, 1.5 mM MgCl₂, iTaq buffer (Biorad), iTaq polymerase (5 U; Biorad) and the entire ligated DNA was combined to a final volume of 100 µL. The PCR had the following settings: 3 min 95°C initial denaturation, then 25 cycles with 30 sec 95°C, 30 sec 60°C, 30 sec 72°C and finally 5 min 72°C. Afterwards samples were kept on ice.

In the final end-labelling step, a PCR with a 6-Carboxyfluorescein (6-FAM) labelled primer is done to label the fragments (Zianni et al., 2006). For this, 5 µL of a premix containing 4 µM primer 3 (OLI2176), 2 mM dNTPs, 1x iTaq-buffer and 1.25 units iTaq-polymerase were added directly to each PCR reaction. The PCR had the following settings: 3 min 95°C initial denaturation, then 5 cycles with 30 sec 95°C, 2 min 63°C, 1 min 72°C and finally 5 min 72°C. Afterwards the PCR products were purified with the GeneJET PCR purification kit (Thermo Scientific) and diluted to a concentration of 100 ng/µL.

Fragment Length Analysis:

Fragment length analysis was done by capillary gel electrophoresis in an external facility. Size standard ILS600 and the filter set Promega-F were used for fragment detection. The received data was analysed with the software Peak Scanner (Applied Biosystems). For analysis, the ratio of *in vivo* and *in vitro* peak area was calculated and normalized to the mean peak area of all peaks in the first 50 bp that correspond to a fragment. To find potential transcription factor binding sites (TFBSs) the sequence was analysed with the prediction software Jaspar (<http://jaspar.genereg.net/>).

Scoring System:

Increased or decreased *in vivo* peak areas received a score of +1 or -1, respectively. Unchanged positions get a score of 0. These scores are added up for each base pair. For examples if the position shows a reduction in 3 out of 3 repetitions, the score is +3. If it is 2 times increased (+2) and once reduced (-1), a score of +1 is given. Because there were only 2 repetitions done of CD14⁺, the score is multiplied by 1,5 for reasons of better illustration.

4. Results:

4.1. Phylogenetic Footprinting/ChIP-Seq:

We used Pre-existing phylogenetic footprinting data for the detection of evolutionary conserved sequences. These potentially functional areas were investigated with a focus on the *ABCD2* gene. Figure 6 shows this phylogenetic footprinting data for the *ABCD2* promoter region. The data contains information from 100 vertebrate species and is displayed in two parts. The upper part of the panel shows the base wise conservation score in green. This score ranges from 0 to 1 and indicates the probability for negative selection of individual nucleotides. Gaps with a low score can also come from insertions or deletions (indels), which create gaps in the alignment. In the first exon, for example, the score drops considerably. This can be explained by several indels at this position. The lower part of the panel shows evolutionary conserved elements. Their degree of conservation is denoted by the log odds ratio (lod), with higher lod scores indicating greater conservation. According to the data, there are several well-conserved areas in the promoter of *ABCD2*. These areas contain all of the known regulatory sites, including the GC1, CCAAT, TBE and SRE/DR-4 elements. For better orientation, the conserved areas which contain such regulatory elements were labelled as section A to D, starting from the promoter. These conserved areas are summarized in table 4, together with their genomic location and potential binding partners.

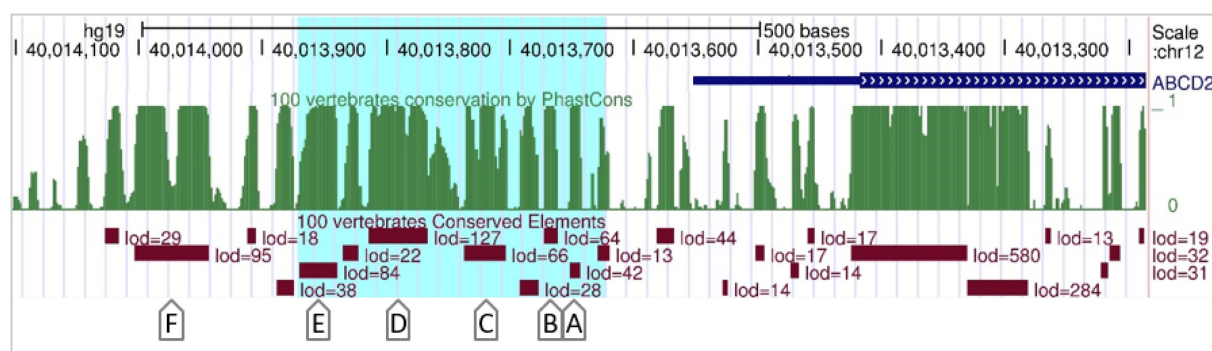


Figure 6: Screenshot of the UCSC Genome Browser, displaying a section of the *ABCD2* promoter (Karolchik et al., 2003). The location of the *ABCD2* gene is illustrated with a blue bar. The Phylogenetic footprinting data of 100 vertebrate species is shown in two parts. The upper track in green, contains the base wise conservation of the sequence. Below conserved areas are shown in red, with the 'lod' score (log odds ratio) indicating their degree of conservation. Higher 'lod' scores signify stronger conservation. The area analysed in the *in vivo* footprinting assay, is highlighted in turquoise.

	Log odd	BPs from ATG	Genomic Location	Elements	TFBSs
A	lod=42	(-227/-235)	chr12:40013644-40013652	GC1-box	SP1/SP3
B	lod=64	(-246/-256)	chr12:40013663-40013673	CCAAT	NF-Y
C	lod=66	(-287/-320)	chr12:40013704-40013737	TBE1/2	TCF-4 (TCF7L2)
D	lod=127	(-351/-398)	chr12:40013768-40013815	SRE/DR-4	SREBF, LXR α , RXR α , MEF2, CTCF
E	lod=84	(-424/-454)	chr12:40013841-40013871		NRF1
F	lod=95	(-528/-587)	chr12:40013945-40014004		RelA (NF- κ B)

Table 4: Conserved regions in the promoter of *ABCD2*. Selected transcription factor binding sites (TFBSs) are listed, together with known binding factors. The data was taken from the UCSC genome browser track “100 vertebrate conserved elements” depicted in figure 6.

The main elements that are known to influence the expression of the *ABCD2* gene are found in the phylogenetic footprinting data. This is a good confirmation of the method, which proves that the data can be used for the detection of regulatory elements. Among section A, B, C and D, section A (GC1-box) has the lowest degree of conservation with a ‘lod’ score of 42. This ‘lod’ score was used as a bottom threshold, for the detection of additional regulatory elements. Elements with a stronger conservation are likely to also play some role for the regulation of the *ABCD2* promoter. Two interesting candidate regions that are well-conserved, are located further upstream of the promoter. These two sites were named section E and F (see Fig. 6). So far no regulatory elements or interacting proteins were described at these two positions.

To better understand the function of conserved regions, the proteins that bind to them were investigated. This was done by analysing available ChIP-seq data. In figure 7, the promoter of the *ABCD2* gene is shown. The bottom part of the graph shows again the phylogenetic footprinting data. Above, selected GM12878 ChIP-seq peaks are shown from proteins with a strong signal in this region. Detected binding sites for these proteins are marked with dotted black lines (see Fig. 7).

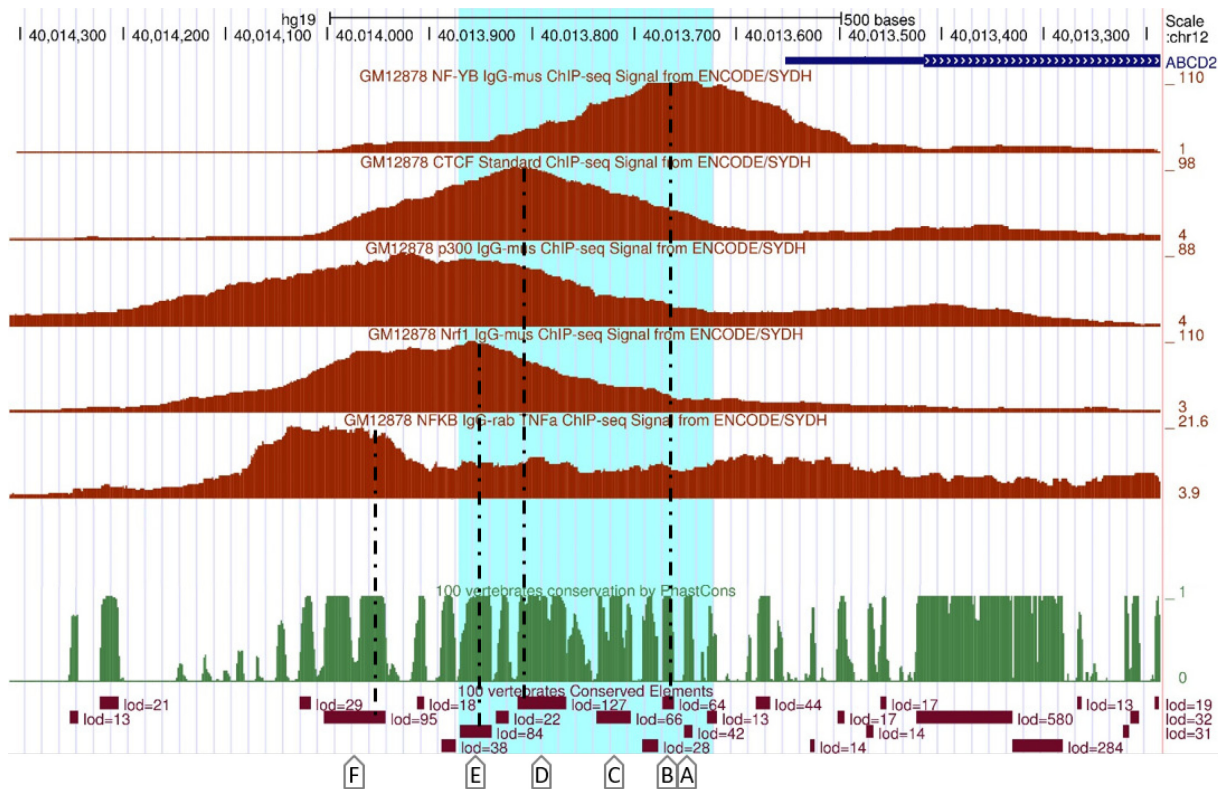


Figure 7: Screenshot of the UCSC Genome Browser, displaying a section of the *ABCD2* promoter (Karolchik et al., 2003). The location of the *ABCD2* gene is illustrated with a blue bar. In the upper part, the ChIP-seq data of selected factors is shown with their name above the peaks. The location of corresponding binding sites is indicated with dotted black lines. The bottom part of the graph shows again the phylogenetic footprinting data depicted in figure 6.

In general, most of the peaks were detected in GM12878. In this cell line, there are clear peaks for NF-YB, CTCF, p300, NRF1, RelA (NF- κ B) and other factors at the promoter. In other cell lines like K562 or HeLa, almost no signals were found at the promoter.

For known regulators of *ABCD2*, the amount of ChIP-seq data is quite limited. The GC1-box, which is located in section A (lod=42), is a binding site for the transcription factors SP1 and SP3 (Gondcaille et al., 2005). For these two factors, no relevant ChIP-seq data was found in the database. Nearby a not so well-conserved area (lod=28) contains the GC2-box. Based on the current knowledge, the GC2-box plays only a minor role for the regulation of the *ABCD2* gene (Gondcaille et al., 2005).

Section B (lod=64) contains the CCAAT element that serves as a binding site for NF-YB. In the ChIP-seq data exactly at this position there is a medium sized peak for NF-YB in GM12878 (see Fig. 7). In K562 this peak is slightly smaller, and in HeLa there is no peak for NF-YB.

Section C (lod=66) contains the two established TCF-binding elements (TBEs). Although there is a peak for the corresponding protein TCF7L2 (TCF4) at this position, the data was generated with HEK293 and PANC-1 cells (see Fig. 8). In HeLa no peak was found at this position. For cell lines like GM12878 or K562, no information is in the database.

Section D (lod=127), the region with the highest amount of conservation, contains the SRE/DR-4 element. For SRE/DR-4 binding factors like TR β , RAR α or LXR α there is no ChIP-seq data available. Only for SREBP and RXR α some information is in the database. Although there is no peak for SREBP in GM12878, the data shows a small accumulation around the promoter (not depicted). This signal is too weak though, to be categorized as a peak. There is no data on SREBP binding for any of the other investigated cell types. For RXR α no signal was detected at the SRE/DR-4. Only a single weak signal from GM12878 is present about 2 kb upstream of the coding sequence (see Fig. 8). Other proteins that have a ChIP-seq peak with a corresponding binding site in section D, are MEF2A, MEF2C and CTCF (see Fig. 8). For MEF2A the only data comes from GM12878 and K562 cells. Among them, MEF2A has a strong signal in GM12878, and a rather weak one in K562 cells. MEF2C shows a clear signal at this position, but was only assayed in GM12878. The binding site for MEF2 is located in the central part of section D, right next to the DR-4 element. The binding site for CTCF is located on the 5' end of section D, where it overlaps with the SRE. CTCF was detected in a wide range of cells, with no clear correlation between signal intensity and gene expression.

Section E (lod=84) contains a binding site for the nuclear respiratory factor-1 (NRF1). In the ChIP-seq data, a clear peak for NRF1 is present in GM12878 cells, and a smaller one in HeLa cells. In K562, there is no peak for NRF1 in the promoter of *ABCD2*.

Section F (lod=95) is the second most conserved area and contains a binding site for RelA (NF- κ B subunit). In TNF α -treated GM12878 cells, a clear peak for RelA was detected at this position. The base wise conservation reveals that section F is composed of two well-conserved sides, with the NF- κ B binding site on the 3' end. On the 5' side multiple potential binding sites were detected for LXR α /RXR α , TCF7L2, PPAR γ and other factors.

On a larger scale, the ChIP-seq peaks are arranged into four clusters. Two of these clusters are located upstream of the promoter (C1, C2), one directly at the promoter (P) and one within the first intron (C3) (see Fig. 8). A schematic illustration of these clusters together with frequently found proteins is depicted in figure 9, with selected ChIP-seq signal intensities summarized in Table 5. Section A-F, which were discussed before, lie within the proximal promoter region where most peaks are located.

In GM12878, where the majority of ChIP-seq peaks were detected, the clusters contain peaks for similar proteins including RUNX3 (Runt Related Transcription Factor-3), MEF2A/C and NF-IC (Nuclear factor-1/C). Additionally, the clusters have peaks for the transcriptional co-activator p300 (EP300), with a clear signal directly at the promoter. The distant clusters C1, C2 and C3, also show small peaks for this factor (see Fig. 8). In relation to neighbouring genes, the p300 peak is rather strong at the promoter of *ABCD2*. For comparison, HeLa and K562 cells have no peak for p300 at the promoter, but show multiple clear signals at other genomic loci (not shown). Overall, these two cell lines have only few ChIP-seq peaks at the promoter of *ABCD2*, which are largely limited to the proximal promoter region.

The C1 cluster is located around 1.9 kb upstream of the coding sequence. In GM12878, ChIP-seq peaks with binding sites for RelA, MEF2A/C, MAZ (MYC Associated Zinc Finger Protein) and YY1 (Yin and Yang-1) are found in this region. Nearby, there is also a weak peak for RXR α . The strongest signal comes from RUNX3, a factor that was only tested in GM12878. In addition, the clusters often contain epigenetic marks and other indicators characteristic for active enhancers. The C1 cluster for example, is located exactly in a valley between two H3K4me1 and H3K27ac peaks (see Fig. 8) (Heintzman et al., 2009, 2007).

The C2 cluster is located around 4.5 kb upstream of the coding sequence. In GM12878, peaks for NF-IC, MEF2A/C, RUNX3 and other factors are found at this position, together with a peak for H3K4me1.

The downstream located C3 cluster lies within the first intron. At this position, the ChIP-seq data from GM12878 contains peaks for YY1, SPI1 (PU.1), NF-IC and other proteins. Although there is only little H3K4 mono-methylation, the area is surrounded by two H3K27ac peaks.

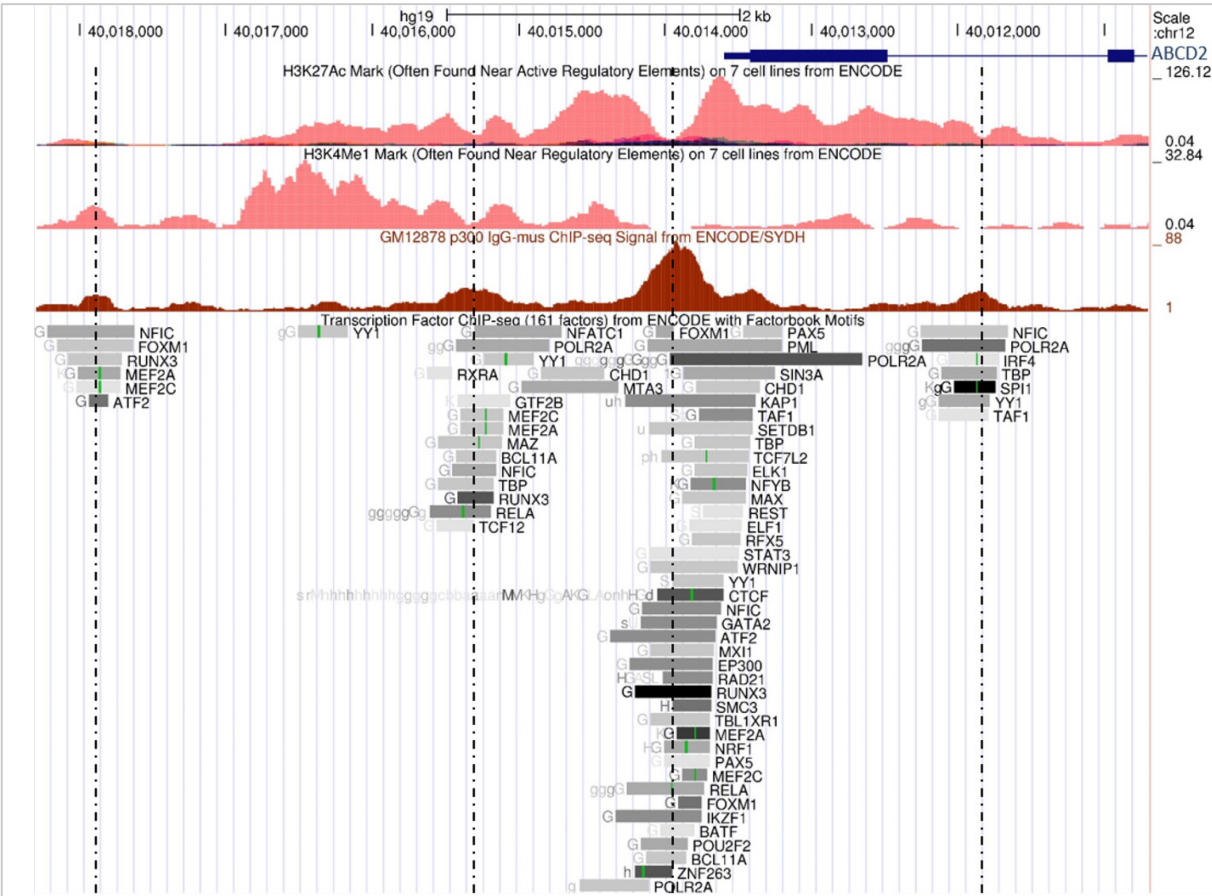


Figure 8: Screenshot of the UCSC Genome Browser displaying an overview of the region around the *ABCD2* promoter. The first 3 lanes contain ChIP-seq data from H3K27ac, H3K4me1 and p300, that was generated with GM12878. The bottom part contains the combined ChIP-seq data of all detected proteins. Protein names are written on the right side of the bars. Letters on the left signify the cell type in which the protein was found (G = GM12878, K = K562, H = HeLa; complete list of abbreviations in table 1) The darkness of a bar is proportional to its peak intensity. The signal intensities of selected factors are summarized in table 5. Short green segments within bars, indicate predicted binding sites for the corresponding factor.

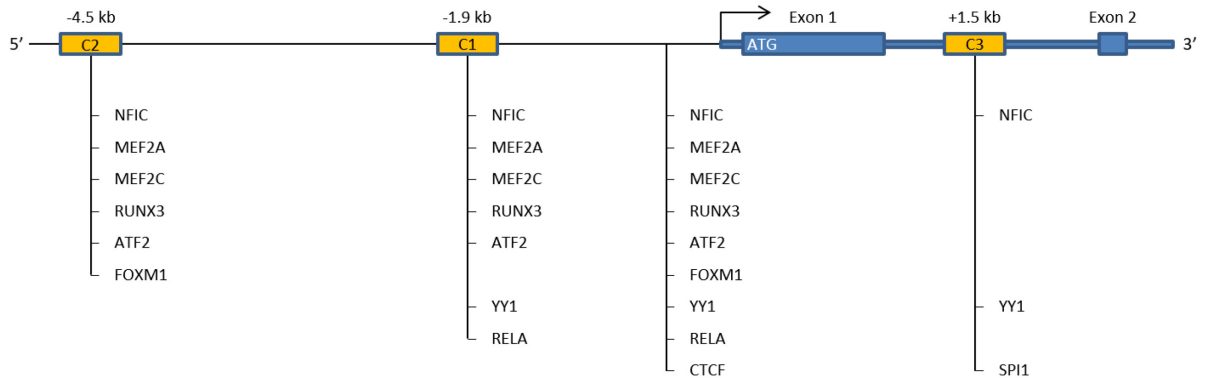


Figure 9: Schematic illustration of the coding strand, containing the promoter of *ABCD2*. The three distant clusters (C1-C3) are indicated by yellow boxes with their distance to the coding sequence noted above. Selected ChIP-seq peaks from figure 8 are listed below.

P	POLR2A	EP300	MEF2A	MEF2C	RUNX3	NF-IC	ATF2	NRF1	NF-YB		
Intensity max. 1000	410 (G)	181 (G)	801 (G) 246 (K)	479 (G)	1000 (G)	377 (G)	238 (G)	371 (G) 194 (H)	408 (G) 346 (K)		
Cell Types w/o signal	G, H, K, L, 1, ...	G, H, K, L, 1, ...	-	-	-	L	1	K, L, 1	H		

C1	POLR2A		MEF2A	MEF2C	RUNX3	NF-IC				YY1	
Intensity max. 1000	300 (G)		227 (G)	170 (G)	616 (G)	381 (G)				167 (G)	
Cell Types w/o signal	G, H, K, L, 1, ...		K	-	-	L				G, K, L, 1, ...	

C2			MEF2A	MEF2C	RUNX3	NF-IC	ATF2				
Intensity max. 1000			335 (G) 161 (K)	159 (G)	261 (G)	301 (G)	581 (G)				
Cell Types w/o signal			-	-	-	L	1				

C3	POLR2A					NF-IC				YY1	SPI1
Intensity max. 1000	408 (G)					253 (G)				271 (G)	1000 (Gg) 388 (K)
Cell Types w/o signal	G, H, K, L, 1, ...					L				G, K, L, 1, ...	-

Table 5: ChIP-seq signal intensities of selected factors at the proximal promoter region (P) and at the three distant ChIP-seq clusters (C1-C3). The corresponding cell type is indicated by letters next to the intensities (G = GM12878, K = K562, H = HeLa, g = GM12891, L = HepG2, 1 = H1-hESC; complete list of cell types in table 1). Only values from the “Txn Factor ChIP”-track, which is shown in figure 8, are included.

4.2. DNase Hypersensitivity:

In addition to epigenetic marks, regulatory elements often overlap with stretches of open chromatin. These regions are characterized by an increased DNase sensitivity, which can be used for the detection of promoters and enhancers (Gross and Garrard, 1988). In figure 10, the DNase hypersensitivity data for the promoter of *ABCD2* is shown. Interestingly, all four ChIP-seq clusters (indicated with dotted black lines) overlap with DNase hypersensitive regions. While the promoter is hypersensitive in almost all of the 125 assayed cell lines, hypersensitivity at distant clusters (C1, C2 and C3) is predominantly found in T and B cells. A complete list of cell types with DNase hypersensitive regions in C1, C2 and C3 can be found in table 6.

At the C1 ChIP-seq cluster, a DNase hypersensitive spot was found in various T and B cell lines and in the neuroblastoma cell line BE2_C. The C2 cluster is DNase hypersensitive in some T and B cell lines, as well as in osteoblasts and several epithelial and fibroblast cell lines. The C3 cluster, which is located in the first intron, is DNase hypersensitive in many T and B cell lines, and also in BE2_C, HL-60 and CD34⁺ cells.

Nearly all investigated cell types are DNase hypersensitive directly at the promoter, with the exception of Caco-2, CMK, iPS and WI-38 (cell type descriptions in table 2). All other analysed cell types, including CD14⁺ monocytes, K562 and HeLa, have open chromatin in this area. This proximal promoter region also contains section A - F from the phylogenetic footprinting data (not shown for reasons of space).

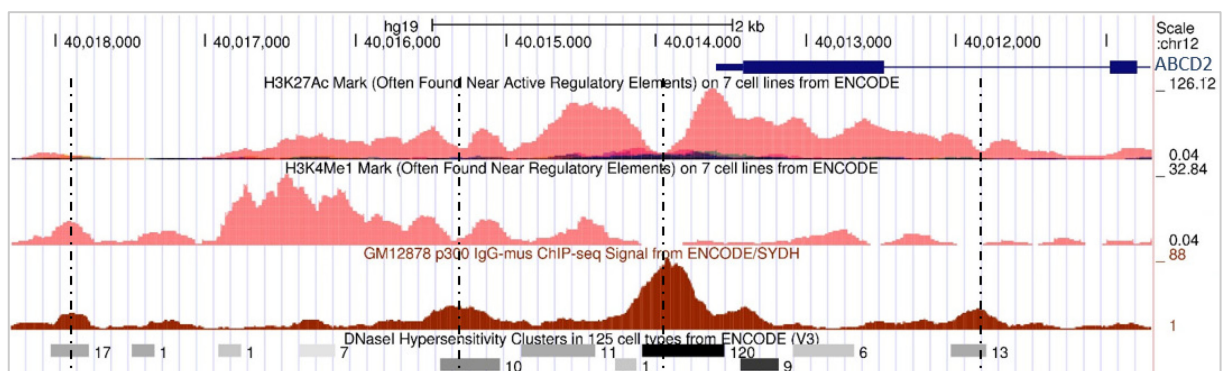


Figure 10: Screenshot of the UCSC Genome Browser displaying an overview of the region around the promoter of *ABCD2*. The first 3 lanes contain ChIP-seq data from H3K27ac, H3K4me and p300, that was generated with GM12878 cells. Below them, a single track shows DNase hypersensitive regions from 125 cell types. The darkness of a bar is proportional to the DNase sensitivity of the area.

Cell Type	Description	C1	C2	C3
Adult CD4 Th0	CD4+ cells enriched for Th0 populations	X		
Th1	primary Th1 T cells	X		X
Th2	primary Th2 T cells	X	X	X
CD20+	B cell	X		X
GM06990	B-lymphocyte, lymphoblastoid (EBV transformed)	X	X	X
GM12864	B-lymphocyte, lymphoblastoid (EBV transformed)	X	X	X
GM12865	B-lymphocyte, lymphoblastoid (EBV transformed)	X		X
GM12878	B-lymphocyte, lymphoblastoid (EBV transformed)	X		X
GM12891	B-lymphocyte, lymphoblastoid (EBV transformed)	X		
GM19238	B-lymphocyte, lymphoblastoid (EBV transformed)			X
GM19239	B-lymphocyte, lymphoblastoid (EBV transformed)			X
GM19240	B-lymphocyte, lymphoblastoid (EBV transformed)			X
CD34+ mobilized	hematopoietic progenitor cells, mobilized			X
HL-60	promyelocytic leukemia cells			X
HAEPiC	amniotic epithelial cells		X	
HCPEpiC	choroid plexus epithelial cells		X	
HEEPiC	esophageal epithelial cells		X	
HIPEpiC	iris pigment epithelial cells		X	
HMEC	mammary epithelial cells		X	
HNPCEpiC	non-pigment ciliary epithelial cells		X	
pHTE	primary tracheal epithelial cells		X	
PrEC	prostate epithelial cell line		X	
RPTEC	renal proximal tubule epithelial cells		X	
SAEC	small airway epithelial cells		X	
FibroP	fibroblasts taken from individuals with Parkinson's disease		X	
HCFAa	cardiac fibroblasts- adult atrial		X	
NHDF-Ad	adult dermal fibroblasts		X	
Osteobl	osteoblasts (NHOst)		X	
BE2_C	neuroblastoma, clone of the SK-N-BE neuroblastoma cell line	X		X

Table 6: Cell types with DNase hypersensitive regions that overlap with the C1, C2 or C3 ChIP-seq cluster. All cell type descriptions were taken directly from the UCSC genome browser.

4.3. *In Vivo* Footprinting:

According to the DNase hypersensitivity data, the proximal promoter region is accessible in most cell types, including T cells, B cells and monocytes. To investigate if T cells and monocytes use different transcription factor binding sites in the promoter of *ABCD2*, the *in vivo* footprinting procedure was used. During the procedure, proteins that are bound to the promoter can protect the DNA from cleavage. This results in distinct fragment compositions that are then analysed. When the data of these *in vivo* processed samples is compared to the data generated with purified DNA (*in vitro*), conclusions can be made about bound regions of the DNA. Between two cell types with different *ABCD2* expression levels, like T cells and monocytes, the usage of regulatory elements in the promoter of the gene might differ. Therefore, a comparison of *in vivo* footprints from these two cell types could reveal responsible elements that are used to either activate or repress the *ABCD2* promoter.

An example for the *in vivo* footprinting data of CD14⁺ monocytes is depicted in figure 11. The blue and white peaks correspond to the amount of fragment that was detected in the *in vivo* and *in vitro* treated samples, respectively. In general, the alignment fits well to the sequence, as exemplified at positions corresponding to thymidines (T). Here for the most part no peaks are present. However, due to unknown reasons there are peaks at positions corresponding to cytosines (C). Where blue peaks (*in vivo*) are lower than white peaks (*in vitro*), a site is located that was potentially protected by a bound protein.

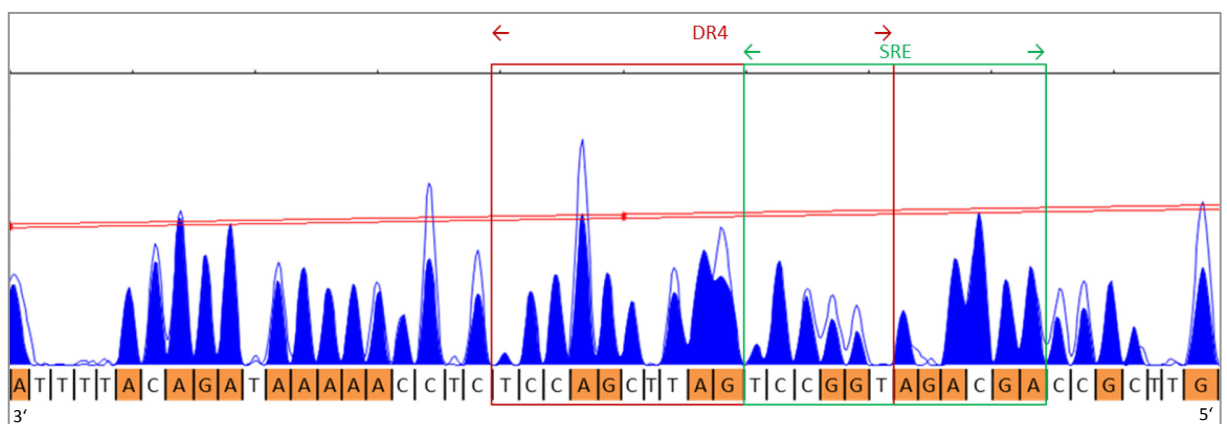


Figure 11: Overlay of *in vivo* and *in vitro* electropherograms of CD14⁺ cells. The depicted region is located 362-420 bp upstream of the start codon. *In vivo* peaks are shown in blue; *in vitro* peaks are shown in white. The SRE and DR-4 elements are labelled in green and red respectively.

The combined data for CD3⁺ and CD14⁺ cells is depicted in figure 12. There are some sites where the *in vivo* samples have a reduced peak area. In the graph, this is indicated by a negative score. In general, a high degree of variation in the data, which is also present between different *in vitro* samples of the same cell type. These samples should have similar patterns, due to the absence of proteins that protect the DNA from cleavage.

To better understand how the sample concentration affects the signal intensity, two dilutions of the same sample were analysed. The results indicate that there is no linear correlation between peak height and fragment concentration (not shown). A 50% reduction of concentration, results in about 10-20% reduced peak area. This reduction is not identical between neighbouring peaks of the same sample and diminishes with increasing fragment length. Only the first 50-100 base pairs have a more or less constant reduction, afterwards the peak area varies too much and conceals the dilution effect. This shows that the peak area is not entirely representative for the amount of each fragment, because it varies considerably even between identical samples.

After several protocol modifications that improved the data quality, these fluctuations still could not be eliminated. In addition, no clear difference is evident at regulatory sites where it is known that certain proteins bind to the sequence (see Fig. 12). The datasets of CD3⁺ and CD14⁺ cells look rather similar, and do not reflect the known differences in gene expression. Because of that no further analysis was done.

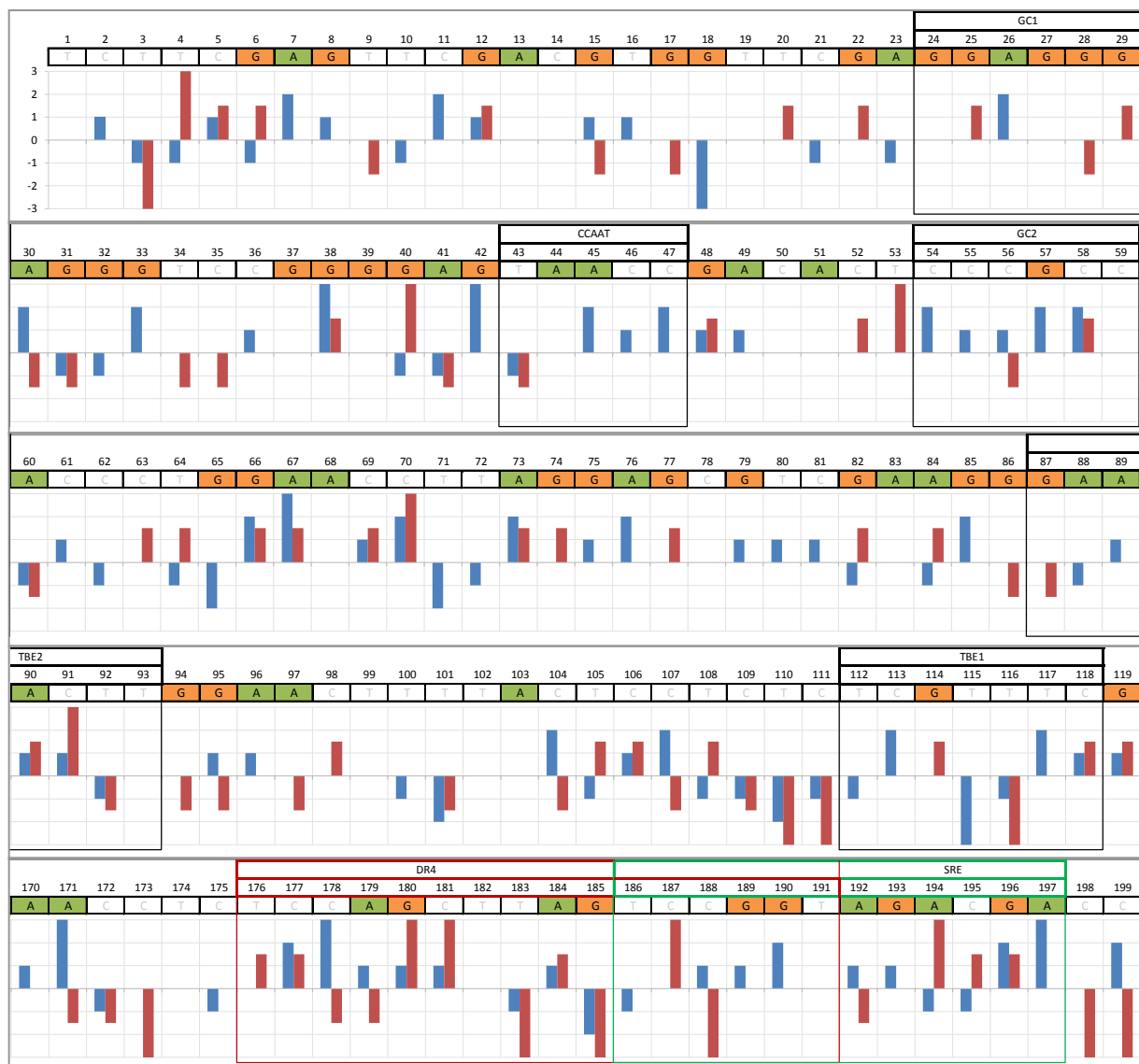


Figure 12: Combined *in vivo* footprinting data of CD3 and CD14 positive cells. The blue (CD3⁺) and red (CD14⁺) bars were calculated with a scoring system detailed in the methods section. Positive values indicate an increased *in vivo* peak area, while negative values indicate a reduction. Known regulatory elements are highlighted with boxes. The regions next to the SRE/DR-4-element are not shown for reasons of space. CD3⁺ n=3; CD14⁺ n=2.

5. Discussion:

5.1. *In Vivo* Footprinting:

During the *in vivo* footprinting procedure several technical problems could be resolved, which greatly improved data quality. Initially, the electropherogram only contained background noise that had too low intensity to stem from real peaks. As it turned out this was due to problems with the PCR. In the original protocol, Vent polymerase was used. After numerous attempts with Vent polymerase it was not possible to reach high enough amplification rates. In addition, unspecific bands appeared when using this polymerase. Since the original protocol is already from 2001, it might be that in the meantime the manufacturing process of Vent or some other thing has been modified. Such a change may not affect a normal PCR, but could become problematic at more sensitive procedures like a footprinting. Because of that, different DNA polymerases were tested. The best results were achieved when using 2 separate polymerases. With iProof for the generation of blunt ends, and iTaq for the amplification and end-labelling steps, much larger and clearer peaks were produced.

After aligning the electropherogram to the sequence, a seemingly random peak distribution appeared. The reason for this is that fragments do not appear in exactly 1 base pair intervals. While some of them travel faster through the capillary, others are slower. After adjusting the width of each individual base pair to the corresponding peak, a correct alignment could be achieved. This is illustrated by the absence of peaks at positions corresponding to thymidines, which confirms that the sequence is positioned properly (see Fig. 11). For some unknown reason though, the electropherogram also contains peaks at positions corresponding to cytosines. According to the protocol, DMS/piperidine should only cut at guanines and adenines (Mueller et al., 2001). To eliminate cytosine peaks a lower temperature during piperidine cleavage was tried, which had no effect on their frequency.

In general, there is a rather high proportion of increased *in vivo* peaks. To improve the results different normalization techniques were tested, as this could skew the ratio towards higher *in vivo* peaks. Initially, only the mean value of peaks corresponding to guanines and adenines was used for normalization. Since DMS/piperidine is reported to cut only at these positions, the signals should originate from cleavage reactions. Other peaks corresponding

to cytosines were not considered due to unknown origin. This normalization technique may cause problems though, because the capillary electrophoresis signal has a better quality in the initial part of the run. With increasing fragment sizes, the peaks become smaller and begin to merge. Because of that, other normalization methods were tested. Using only the first 50 base pairs of the fragments, or the signal of all nucleotides instead of only G/A peaks, wasn't able to produce more plausible results. Also, using the median instead of the mean or any combination of the approaches mentioned above, could not improve the results.

The analysis of diluted samples showed that there are variations even between identical samples. Ideally, dilutions of the same sample should have a constant change in peak area corresponding to the concentration, which is however not the case here. Because the footprinting signals vary in a similar range as the dilution samples, a reliable distinction between real signals and experimental fluctuations is not possible. Additionally, there are some very high peaks in the data that remain unexplained. It is possible that these peaks come from unspecific PCR products. A probable explanation for the fluctuations might be that the used polymerases were not suitable for the procedure. It was reported by Garrity et al. that the previously used combination of Sequenase and Taq polymerase, can lead to variations in fragment amount and wrong bands (Garrity and Wold, 1992). Because there are similar problems in this experiment, it may be that the combination of iProof and iTaq polymerase causes the same issues.

Overall there are some sites where the *in vivo* samples have reduced peak areas, which could indicate transcription factors binding sites that were protected from methylation. There are however no clear reductions at known regulatory sites (e.g. SRE/DR-4). After 3 repetitions it was shown, that the observed differences in peak area are not shared between samples of the same cell type, and are therefore likely to stem from other sources. Bound proteins may have had some impact on the peak area, which is however buried beneath experimental fluctuations. Because of that, it remains uncertain whether the repeatedly observed differences in peak intensity are caused by bound proteins. Thus, in spite of numerous successful improvements, the experiment ultimately had to be stopped. The results of this experiment were not further interpreted and are thus not further discussed within this segment.

5.2. Phylogenetic Footprinting/ChIP-Seq:

The phylogenetic footprinting data provided some interesting information. There are several well-conserved areas in the promoter of *ABCD2*. Among them, all of the known regulatory elements are found that were described in the literature so far. This shows that the method can be used to reliably identify functional regions of the promoter. In the ChIP-seq data there are multiple peaks for proteins, which bind at these evolutionary conserved elements. Among them are established regulators of *ABCD2*, as well as other factors that were not described so far.

At the SRE/DR-4 element in section D, there is a clear peak with a corresponding binding site for the CCCTC-binding factor (CTCF). This peak is found in high and low intensities in various cell types and shows no clear correlation with *ABCD2* levels. This fits to CTCFs general role as a chromosomal organizer, whose effect on gene expression can be either activating or repressive, depending on its interactions (Ong and Corces, 2014). CTCF can form chromatin loops, to bring distant regulatory regions in close proximity of the promoter (Handoko et al., 2011). These loops allow a cross-talk across long distances to regulate transcription. Via these interactions, CTCF can also facilitate the recruitment of regulatory proteins such as p300 to the promoter (Handoko et al., 2011). Among all assayed cell types, GM12878 is the only one where CTCF and p300 are both found at the promoter. Since co-localization of CTCF and p300 is a sign for active transcription, it may be that this interaction contributes to the high *ABCD2* expression rates in GM12878 (Handoko et al., 2011). In cell types with lower expression levels, CTCF may interact with different factors to silence transcription.

Unfortunately, there are no peaks in the ChIP-seq data for other proteins that are known to bind at the SRE/DR-4 element, like TR β , RAR or SREBP. For the most part, this can be attributed to a lack of relevant experiments that were done for these factors. For RXR α , a single weak peak is located close to the C1 cluster. Due to the weak intensity of this signal, it is questionable if this peak is of any relevance for the expression of *ABCD2*.

For the B subunit of the nuclear transcription factor-Y (NF-YB), a ChIP-seq peak was found at the expected position at the CCAAT element, located in section B. NF-Y is a known regulator of *ABCD2* (Gondcaille et al., 2005). However, despite their differences in *ABCD2* expression, NF-YB is found at this position with similar intensities in GM12878 (lymphoblastoid cell line)

and K562 (granulocytic cell line), respectively. The reason behind this may lie in the either repressing or activating capabilities of NF-Y, depending on interacting co-factors. Interactions between NF-Y and HDAC1 were suggested to exert a repressive effect on *ABCD2* expression (Gondcaille et al., 2005). A positive influence on transcription can be mediated by interactions with histone acetyltransferases like p300 (Salsi et al., 2003). Such a mechanism was proposed to promote *ABCD2* expression, however so far the exact protein that mediates this effect was not identified (Gondcaille et al., 2005).

In the ChIP-seq data peaks for p300 and NF-YB are located right next to each other and show a partial overlap. Because of this vicinity, interactions between both factors might occur. Interestingly, the interaction between NF-YB and p300 can be stimulated by cAMP (Faniello et al., 1999). Since *ABCD2* expression and VLCFA β -oxidation are both increased by cAMP pathway activators like forskolin, an enhanced interaction between both proteins might explain the observed effects (Pahan et al., 1998; Pujol et al., 2000).

Besides NF-Y, p300 can also interact with the transcription factor SP1, which binds at the adjacent GC1-box (Billon et al., 1999; Gondcaille et al., 2005). A similar arrangement of NF-Y and SP1 binding sites is found in the promoter of the HDAC1 gene, where ectopic expression of p300 is able to significantly increase promoter activity via these two elements (Schuettengruber et al., 2003). Since p300 is expressed with similar intensities in monocytes and T cells, the recruitment to the promoter may be different between the two cell types (GEO accession GSE1133) (Barrett et al., 2013; Su et al., 2004).

Besides the already known regulatory sites, two additional conserved elements (section E and F) were detected. There are no known regulatory elements at neither position. However, due to their high degree of conservation, it is likely that they have some functional role for the regulation of *ABCD2*. Two proteins that have a peak with corresponding binding sites in section E and F, are the nuclear respiratory factor-1 (NRF1) and RelA (NF- κ B) respectively. Besides NRF1 and RelA, section E and F contain also signals from various other factors. For the majority of these proteins no binding site was detected. This could mean that the binding site simply was not detected, or that these proteins bind to the promoter via other interactions. Recruitment could for example also occur via protein-protein interactions or by recognition of epigenetic modifications (Strahl and Allis, 2000).

NRF1 is a regulator of several nuclear genes involved in the respiratory chain and in the mitochondrial transcription and replication apparatus (Evans and Scarpulla, 1989; Gopalakrishnan and Scarpulla, 1995; Suzuki et al., 1989). NRF1 also regulates the acetyl CoA carboxylase gene (*ACC*) and carnitine palmitoyl transferase 1 gene (*CPT1*) (Adam et al., 2010; Jogl et al., 2004). *ACC* and *CPT1* are both involved in the cellular metabolism of fatty acids. While *ACC* catalyses an important step in their synthesis, *CPT1* is used for the degradation of certain fatty acids. *CPT1* is a mitochondrial membrane transporter that is essential for the import and subsequent β -oxidation of long-chain fatty acids (Jogl et al., 2004; Xia et al., 1997). Likewise, NRF1 might regulate *ABCD2* to control the levels of some fatty acids and to adjust to the metabolic state of the cell. The regulation of NRF1 target genes is often mediated by the peroxisome proliferator-activated receptor gamma co-activator-1 (PGC-1), a key regulator of mitochondrial biogenesis and cellular energy metabolism (Liang and Ward, 2006; Puigserver et al., 1999). Interactions with NRF1 increase the transcriptional activity of PGC-1 (Puigserver et al., 1999; Wu et al., 1999). Besides NRF1, many of the described regulators of *ABCD2* like PPAR, TR β and LXR can interact with PGC-1 (Liang and Ward, 2006). PGC-1 α is also a co-activator of MEF2, a factor which is found several times at the promoter of *ABCD2* (Lin et al., 2002; Michael et al., 2001). Interactions between NRF1 and PGC-1 might provide a link between the metabolic state of a cell and the expression of corresponding genes like *ABCD2*. However, since there is no relevant ChIP-seq data for PGC-1, it is not clear if it is involved in the regulation of *ABCD2*, or if NRF1 interacts with different factors.

In section F, a binding site and a clear ChIP-seq peak for the NF- κ B subunit RelA (p65) is found in TNF α treated GM12878. P65 forms dimers with other REL-homology-domain containing proteins like p50 or p52, to regulate NF- κ B responsive genes (Karin, 1999). Under normal conditions NF- κ B is bound in the cytoplasm by the inhibitory protein I- κ B. Upon external stimuli like TLR activation or proinflammatory cytokines (TNF α , IL-1), the I- κ B kinase (IKK) is activated, and phosphorylates I- κ B (DiDonato et al., 1997). This leads to the degradation of I- κ B, and allows NF- κ B to translocate into the nucleus where it binds its target genes (Chen et al., 1995; Karin, 1999). Besides this mechanism, NF- κ B is also regulated by different MAP kinase pathways, including JNK and p38 (Schulze-Osthoff et al., 1997).

NF- κ B is perhaps best known as a central regulator of inflammation, but is also involved in some other processes such as proliferation and apoptosis (Hoesel and Schmid, 2013).

In addition, NF- κ B can influence cellular energy homeostasis by switching between glycolysis and mitochondrial respiration (Mauro et al., 2011). So far it is not known if *ABCD2* is regulated by NF- κ B. The only relevant ChIP-seq data for p65 comes from TNF α treated GM12878. Treatment with cytokines like TNF α constitutes an artificial setting, which is probably very different from the mechanisms used under physiological conditions. On top of that, there is no information from other cell lines with lower *ABCD2* expression in the database. Therefore, no definitive conclusions can be made about the influence of NF- κ B on the regulation of *ABCD2*.

A potential mechanism involving NF- κ B and other factors that were found at the promoter is an enhanceosome. This multiprotein complex, consists of general transcription factors, gene-specific factors and chromatin modifying proteins (Carey, 1998). There are different forms of enhanceosomes, which vary in their composition. Enhanceosomes can stimulate gene expression by recruiting other co-activators, or directly by interactions with the basal RNA Polymerase II transcriptional machinery (Carey, 1998). Perhaps the best known example is the IFN β -enhanceosome depicted in figure 13. It consists of NF- κ B, ATF2/c-Jun, interferon regulatory factors (IRFs) and other proteins (Vo and Goodman, 2001). These factors recruit a central p300, or the homologous CREB-binding protein (CBP), to the promoter. CBP or p300 can then activate transcription via their HAT domains, and by interactions with a wide range of proteins (Chan and La Thangue, 2001; Merika et al., 1998; Ogryzko et al., 1996).

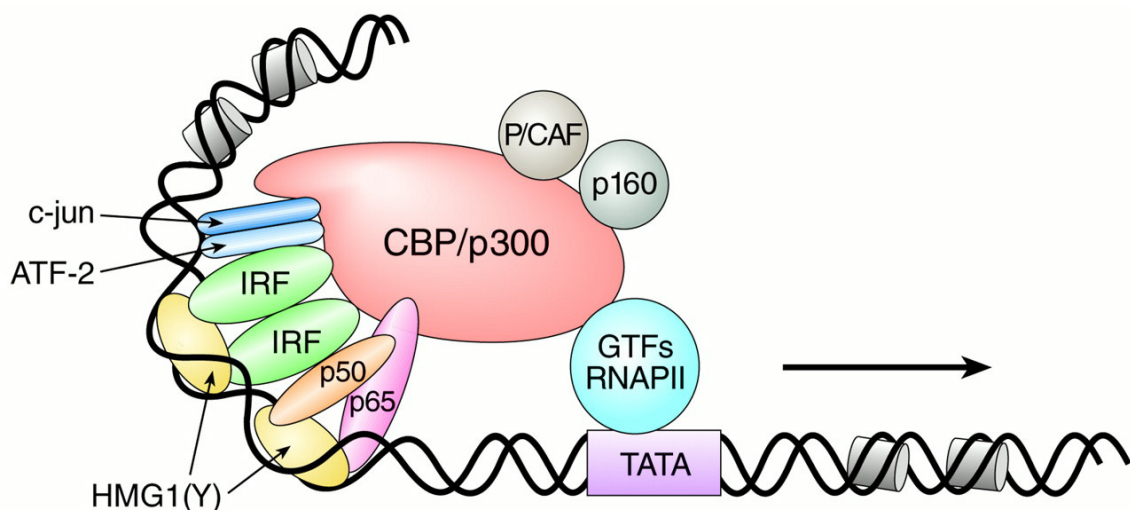


Figure 13: Illustration of the IFN β enhanceosome consisting of general transcription factors and gene-specific factors organized around a central CBP/p300 (Vo and Goodman, 2001). A similar mechanism might be used for the regulation of *ABCD2*.

This figure is used with the permission of the American Society for Biochemistry and Molecular Biology.

Due to the diverse interaction partners of p300, it was proposed that it binds several factors simultaneously and serves as a central adaptor protein (Eckner et al., 1994). In GM12878 the promoter of *ABCD2* contains several ChIP-seq peaks from proteins such as RUNX3, MEF2, ATF2 or RelA, which can interact with p300. This would suggest the usage of a mechanism similar to an enhanceosome, for the regulation of the *ABCD2* gene.

Interestingly many of the ChIP-seq peaks are clustered around two regions located 1.9 kb (C1) and 4.5 kb (C2) upstream of the coding sequence. In addition, about 1.5 kb downstream of the start codon another small cluster of ChIP-seq peaks (C3) is found. The clustering of ChIP-seq peaks around a small region can indicate the presence of underlying regulatory elements like promoters or enhancers (Rye et al., 2011). Based on the co-occurrence of certain histone modifications and DNase hypersensitivity, it is likely for these sites to act as enhancers (Heintzman et al., 2009; Xi et al., 2007). Interestingly, these enhancer specific marks are mostly present in GM12878 cells and closely resemble the ones found in primary B and T cells. This could mean that there are similar regulatory mechanisms used. Because cell-type specific gene expression is in part mediated by enhancers, it may be that these sites are contributing to the different *ABCD2* expression levels in T cells and monocytes.

At two suspected enhancer sites (C1, C2) and at the promoter, binding sites for the myocyte enhancer factor 2 (MEF2) are found. In addition, there are ChIP-seq peaks for MEF2A and MEF2C, which were detected in GM12878 and K562 cells. MEF2 proteins are regulators of various developmental programs and are involved in the differentiation of muscle cells, neurons, lymphocytes and other cell types (Potthoff and Olson, 2007). Vertebrates have four MEF2 protein variants, which are encoded by the paralogous genes *MEF2A*, *B*, *C* and *D* (Potthoff and Olson, 2007; Shore and Sharrocks, 1995). These proteins can form homo- or heterodimers to regulate distinct target genes (Estrella et al., 2015; Ornatsky and McDermott, 1996).

By binding to enhancer regions, MEF2 proteins can influence gene expression (Gossett et al., 1989). The transcriptional activity of MEF2 is regulated by mitogen-activated protein kinase (MAPK) and Ca^{2+} dependent signalling (Han et al., 1997; Lu et al., 2000). Without these signals, MEF2 associates with class IIa HDACs or other repressive proteins like Cabin1 or MITR to silence the target gene (Lu et al., 2000; Sparrow et al., 1999; Youn et al., 1999). Upon signal transduction, this interaction is impaired by phosphorylation of class IIa HDACs,

calcium-dependent inactivation of Cabin1 or other mechanisms (McKinsey et al., 2000; Youn et al., 1999). To more effectively stimulate gene expression, MAPK and Ca^{2+} dependent signalling can cooperate. While the Ca^{2+} /calmodulin-dependent protein kinase (CaMK) alone is able to increase the transcriptional activity of MEF2 by interfering with HDAC recruitment, a much stronger effect is seen with additional MAPK-dependent phosphorylation of MEF2 (Lu et al., 2000).

After the release of repressive proteins, MEF2 is free to interact with activating factors such as p300 or the co-activator GRIP-1 (SRC-2, NCoA-2) (Chen et al., 2000; Youn and Liu, 2000). P300 can then acetylate MEF2, which is necessary for it to effectively bind DNA and induce transcription (Ma et al., 2005). In addition, p300 can bind three MEF2 dimers simultaneously with its highly conserved TAZ2 domain (He et al., 2011). This would allow the formation of a p300/MEF2-enhanceosome. This complex consists of three MEF2 dimers arranged around a central p300 molecule (see Fig. 14). MEF2 regulated genes often have several MEF2 binding sites in their promoter (He et al., 2011; Sandmann et al., 2006). Interestingly, the promoter of *ABCD2* also contains three MEF2 binding sites. This is very similar to the model depicted in figure 14, and would suggest the usage of a similar mechanism for the regulation of *ABCD2*.

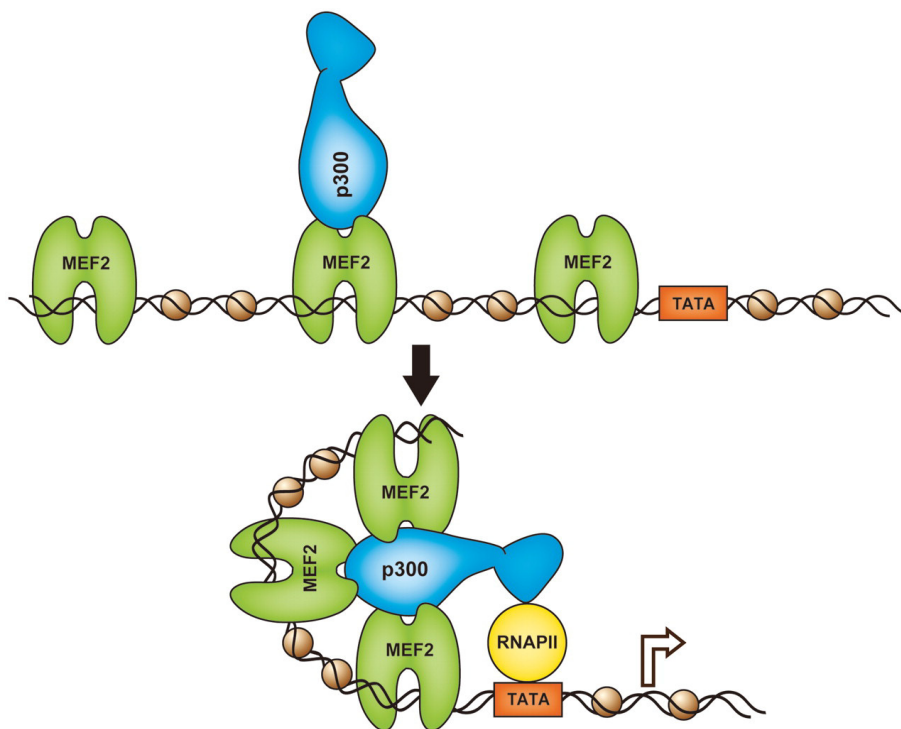


Figure 14: Proposed model of a p300/MEF2 enhanceosome, consisting of three MEF2 dimers bound by a central p300 (He et al., 2011).

This figure is used with the permission of Oxford University Press.

Simultaneous binding of multiple MEF2 dimers located at the promoter and at enhancers could strengthen the recruitment of p300 to the promoter (He et al., 2011). This theory is supported by the presence of p300 ChIP-seq peaks in GM12878, which overlap with all three binding sites of MEF2 (see Fig. 8).

The ChIP-seq data also contains signals from other factors that can interact with either one or both members of the proposed MEF2/p300-complex. In addition, known regulators of *ABCD2* like the thyroid hormone receptor (TR)/retinoid X receptor (RXR)-dimer, can bind MEF2A and p300 (De Luca et al., 2003). Such interactions are used for example, to regulate the expression of the α -MHC gene. In a co-transfection experiment with expression plasmids for TR and RXR, an α -MHC reporter construct was activated by the application of T3 (De Luca et al., 2003). In the same experiment, but with additional MEF2A expression, more than twice as high induction rates were observed. Since T3 treatment is able to effectively increase *ABCD2* expression it would be interesting to see if MEF2 exerts a similar effect on this promoter (Fourcade et al., 2003). Interestingly, one of the MEF2 binding sites of the *ABCD2* promoter is located right next to the DR-4 element within the highly conserved section D. Because of this vicinity, interactions with the DR-4-bound TR/RXR-dimer could occur.

Such interactions between proteins might help to recruit p300 to the *ABCD2* promoter. After the recruitment, p300 could then activate transcription by interactions with additional factors and via its histone acetyl transferase activity (Chan and La Thangue, 2001). In GM12878, the ChIP-seq peaks for p300 are often located directly in a valley between two H3K27ac peaks. Among other lysine residues, p300 has a preference for H3K27 (Tie et al., 2009). Therefore, the observed H3K27ac levels could be a direct consequence of p300 recruitment, which might play an important role for the high *ABCD2* expression in GM12878. Without the recruitment of p300 there may be less H3K27ac and subsequently a stronger chromatin condensation. This could also explain the low *ABCD2* expression in other cell types like K562, where p300 was not found at the promoter.

As a summary, two models for the regulation of the *ABCD2* promoter were created. These models, are based on the analysed ChIP-seq data and are in good agreement with data from DNase hypersensitivity and histone modification assays. In figure 15, a hypothetical enhanceosome complex is shown for cell types with high *ABCD2* expression like T cells, B cells or GM12878. This complex consists of MEF2, p300 and other factors, which gather around the *ABCD2* promoter. The MEF2 proteins directly recognize binding sites at the promoter (P) and at two additional regulatory elements (C1, C2). Via protein-protein interactions, MEF2 and other factors recruit a central p300 to the promoter to form an enhanceosome. The enhanceosome assembly is facilitated by the wide range of p300 interacting factors (Chan and La Thangue, 2001). NF-Y for example, can interact either with HATs or HDACs to activate or silence transcription, respectively (Jin and Scotto, 1998). After the recruitment of p300, the resulting enhanceosome complex could then facilitate the interaction with the RNA polymerase II and stimulate transcription of the *ABCD2* gene.

Based on epigenetic marks, open chromatin and other indicators, it is likely that the C3 element and bound proteins such as YY1 or SPI1, also contribute to the regulation of the *ABCD2* promoter in GM12878. So far, however, is not clear which role they might play for this process.

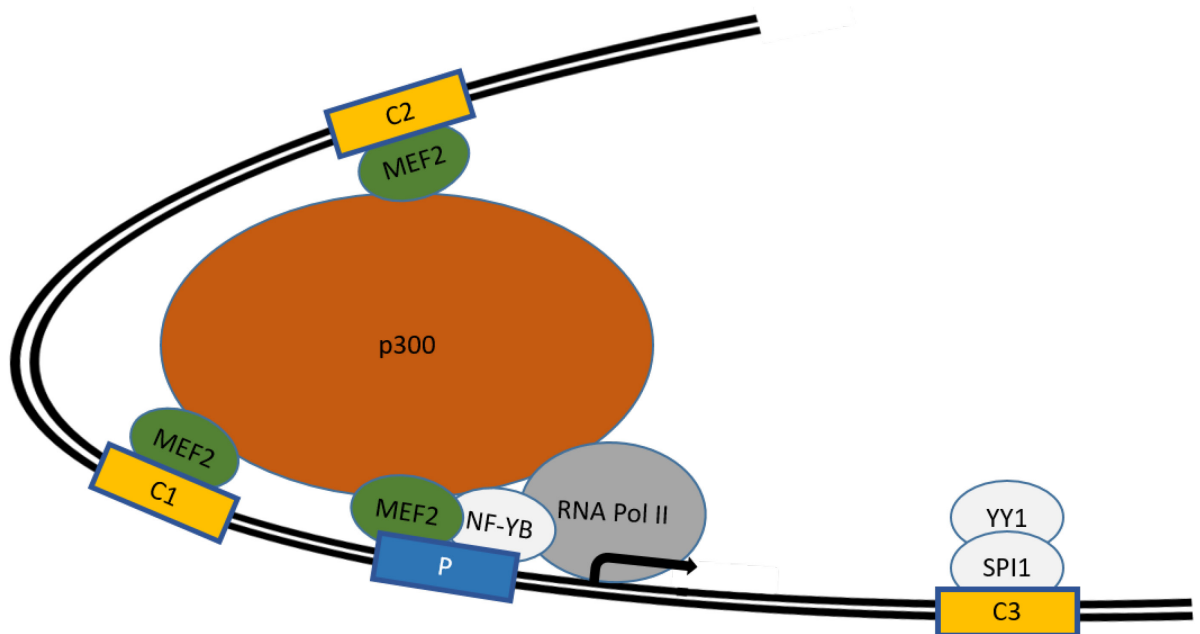


Figure 15: Hypothetical model of an enhanceosome at the *ABCD2* promoter in cell types with high *ABCD2* expression such as T cells, B cells or GM12878. Via protein-protein interactions, a central p300 is recruited to the promoter to facilitate the expression of the *ABCD2* gene.

In figure 16, a model of the *ABCD2* promoter region is shown as envisioned in cell types with low *ABCD2* expression. This model is based on data from the granulocytic cell line K562, which was used representative for granulocytes and monocytes. In the analysed ChIP-seq data of K562, MEF2A was also found at the promoter and at the C2 cluster, but was not detected at the C1 cluster. Compared to GM12878, the peaks have a lower intensity. The reduced MEF2A levels at the *ABCD2* promoter of K562, might be explained by the lack of p300. Via its HAT domain, p300 can acetylate MEF2 and thereby enhance its DNA binding activity (Ma et al., 2005). Thus without p300, MEF2 has a lower affinity for DNA.

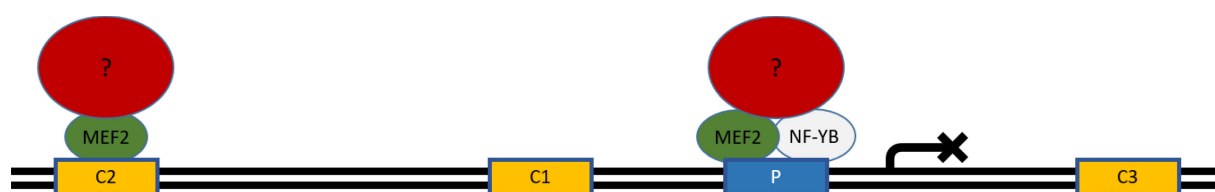


Figure 16: Hypothetical model of the *ABCD2* promoter in cell types with low *ABCD2* expression such as granulocytes, monocytes or K562. Instead of p300, repressive proteins (indicated with “?”) are recruited to the promoter to silence transcription.

Instead of p300, MEF2 might interact with repressive proteins to silence transcription. The repressor Cabin1 for example, binds MEF2 for a sequence-specific recruitment (Han et al., 2003). Cabin1 can then silence the target gene via interactions with the SIN3A/HDAC co-repressor complex (Kadamb et al., 2013; Youn and Liu, 2000). In addition, Cabin1 competes directly with p300 for the binding of MEF2, and thereby also prevents transcription (Youn and Liu, 2000). MEF2 bound repressors like Cabin1 or class IIa HDACs, could contribute to the low activity of the *ABCD2* promoter in granulocytes and monocytes (Weber et al., 2014b). Consequently, it may be that previous attempts of *ABCD2* induction did not reach higher levels, because MEF2 was bound by such repressors. Besides MEF2, also other proteins like NF-Y can interact with HDACs and strengthen their recruitment to the promoter (Jin and Scotto, 1998; Peng and Jahroudi, 2003).

Endogenously, MEF2 is regulated by MAPK and Ca²⁺ dependent signalling (Lu et al., 2000). Therefore, drugs that target these pathways might be useful for the induction of *ABCD2* gene expression. Alternatively, there are also commercially available substances that directly inhibit the interaction between MEF2 and repressors. The small molecule N-(2-aminophenyl)-N'-phenyloctanediamide (BML-210) for example, acts as an HDAC inhibitor (Savickiene et al., 2006). However, unlike traditional HDAC inhibitors, BML-210 does not block the catalytic domain of HDACs directly, but rather prevents their MEF2 dependent recruitment to the promoter (Falkenberg and Johnstone, 2014; Jayathilaka et al., 2012). With such an approach, and simultaneous treatment with activating substances like T3, a stronger activation of the *ABCD2* promoter might be achieved.

The reasons behind different *ABCD2* gene expression levels in T cells and monocytes are still not entirely understood. Besides the mentioned transcription factors, there are many more proteins that bind to the promoter and interact with each other. This complex interplay is likely to contribute to the cell-type specific *ABCD2* expression levels observed in T cells and monocytes. In the future, an improved understanding of these regulatory mechanisms could help with the pharmacological induction of *ABCD2*, as a potential treatment for X-ALD (Berger et al., 2010).

6. List of Abbreviations:

6-FAM	6-carboxyfluorescein
<i>ABCD1</i>	ATP-binding cassette subfamily D member-1
AMN	Adrenomyeloneuropathy
ATF2	Activating transcription factor-2
Cabin1	Calcineurin Binding Protein-1
cALD	cerebral Adrenoleukodystrophy
CBP	CREB binding protein
ChIP-seq	Chromatin immunoprecipitation sequencing
CTCF	CCCTC-binding factor
c-jun	Jun Proto-Oncogene
DHA	Docosahexaenoic acid
DMS	Dimethyl sulfate
DR-4	Direct repeat separated by 4 nucleotides
EP300	E1A binding protein p300
GRIP-1	Glucocorticoid receptor-interacting protein-1
H3K27ac	Histone 3 lysine 27 acetylation
H3K4me1	Histone 3 lysine 4 mono-methylation
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HMG1	High Mobility Group Box 1
HSCT	Hematopoietic stem cell transplantation
IRF	Interferon Regulatory Factor
IKK	I- κ B kinase
I- κ B	Inhibitor of NF- κ B
JNK	C-Jun N-Terminal Kinase
LM-PCR	Ligation mediated polymerase chain reaction
LOD	Log odd ratio

LXR	Liver X receptor
MAZ	MYC Associated Zinc Finger Protein
MEF2	Myocyte enhancer factor-2
MITR	MEF2 interacting transcriptional repressor
NCoA-2	Nuclear receptor coactivator-2
NF-I	Nuclear factor-1
NF-Y	Nuclear transcription factor-Y
NHR	Nuclear hormone receptor
NRF1	Nuclear respiratory factor-1
P/CAF	P300/CBP associated factor
PBMC	Peripheral blood mononuclear cell
PGC-1	Peroxisome proliferator-activated receptor gamma coactivator-1
PPAR	Peroxisome proliferator-activated receptor
PUFA	Polyunsaturated fatty acid
RAR	Retinoic acid receptor
RUNX3	Runt related transcription factor-3
RXR	Retinoid X receptor
SP1	Specificity protein-1
SPI1	Spleen Focus Forming Virus (SFFV) Proviral Integration Oncogene
SRC-2	Steroid receptor coactivator-2
SREBP	Sterol response element binding protein
T3	Tri-iodothyronine
TBE	T-cell factor/lymphoid enhancer factor (TCF/LEF)-binding element
TCF-4	T cell factor-4
TNF α	Tumour necrosis factor alpha
TR	Thyroid hormone receptor
VLCFA	Very long-chain fatty acid
X-ALD	X-linked Adrenoleukodystrophy
YY1	Yin and Yang-1

7. References:

- Adam, T., Opie, L.H., Essop, M.F., 2010. AMPK activation represses the human gene promoter of the cardiac isoform of acetyl-CoA carboxylase: Role of nuclear respiratory factor-1. *Biochem. Biophys. Res. Commun.* 398, 495–499.
- Aranda, A., Pascual, A., 2001. Nuclear hormone receptors and gene expression. *Physiol. Rev.* 81, 1269–1304.
- Aubourg, P., Blanche, S., Jambaque, I., Rocchiccioli, F., Kalifa, G., Naud-Saudreau, C., Rolland, M.O., Debre, M., Chaussain, J.L., Griscelli, C., 1990. Reversal of early neurologic and neuroradiologic manifestations of X-linked adrenoleukodystrophy by bone marrow transplantation. *N. Engl. J. Med.* 322, 1860–1866.
- Baniahmad, A., Köhne, A.C., Renkawitz, R., 1992. A transferable silencing domain is present in the thyroid hormone receptor, in the v-erbA oncogene product and in the retinoic acid receptor. *EMBO J.* 11, 1015.
- Bannister, A.J., Miska, E.A., 2000. Regulation of gene expression by transcription factor acetylation. *Cell. Mol. Life Sci.* 57, 1184–1192.
- Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C.L., Serova, N., Davis, S., Soboleva, A., 2013. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* 41, D991–D995.
- Berger, J., Albet, S., Bentejac, M., Netik, A., Holzinger, A., Roscher, A.A., Bugaut, M., Forss-Petter, S., 1999. The four murine peroxisomal ABC-transporter genes differ in constitutive, inducible and developmental expression. *Eur. J. Biochem.* 265, 719–727.
- Berger, J., Forss-Petter, S., Eichler, F.S., 2014. Pathophysiology of X-linked adrenoleukodystrophy. *Biochimie* 98, 135–142.
- Berger, J., Gartner, J., 2006. X-linked adrenoleukodystrophy: clinical, biochemical and pathogenetic aspects. *Biochim. Biophys. Acta* 1763, 1721–1732.
- Berger, J., Molzer, B., Fae, I., Bernheimer, H., 1994. X-linked adrenoleukodystrophy (ALD): A novel mutation of the ALD gene in 6 members of a family presenting with 5 different phenotypes. *Biochem Biophys Res Commun* 205, 1638–1643.
- Berger, J., Pujol, A., Aubourg, P., Forss-Petter, S., 2010. Current and Future Pharmacological Treatment Strategies in X-Linked Adrenoleukodystrophy. *Brain Pathol.* 20, 845–856.
- Bezman, L., Moser, A.B., Raymond, G. V, Rinaldo, P., Watkins, P.A., Smith, K.D., Kass, N.E., Moser, H.W., 2001. Adrenoleukodystrophy: incidence, new mutation rate, and results of extended family screening. *Ann. Neurol.* 49, 512–517.
- Billon, N., Carlisi, D., Datto, M.B., Grunsven, L.A. van, Watt, A., Wang, X.-F., Rudkin, B.B., 1999. Cooperation of Sp1 and p300 in the induction of the CDK inhibitor p21WAF1/CIP1 during NGF-mediated neuronal differentiation. *Oncogene* 18, 2872–2882.

- Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., Allis, C.D., 1996. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84, 843–851.
- Budka, H., Sluga, E., Heiss, W.D., 1976. Spastic paraplegia associated with Addison's disease: adult variant of adreno-leukodystrophy. *J Neurol* 213, 237–250.
- Carey, M., 1998. The Enhanceosome and Transcriptional Synergy. *Cell* 92, 5–8.
- Cartier, N., Guidoux, S., Rocchiccioli, F., Aubourg, P., 2000. Simvastatin does not normalize very long chain fatty acids in adrenoleukodystrophy mice. *FEBS Lett.* 478, 205–208.
- Cartier, N., Hachein-Bey-Abina, S., Bartholomae, C.C., Veres, G., Schmidt, M., Kutschera, I., 2009. Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science* 326, 818–823.
- Casanova, J., Helmer, E., Selmi-Ruby, S., Qi, J.-S., Au-Fliegner, M., Desai-Yajnik, V., Koudinova, N., Yarm, F., Raaka, B.M., Samuels, H.H., 1994. Functional evidence for ligand-dependent dissociation of thyroid hormone and retinoic acid receptors from an inhibitory cellular factor. *Mol. Cell. Biol.* 14, 5756–5765.
- Chan, H.M., La Thangue, N.B., 2001. p300/CBP proteins: HATs for transcriptional bridges and scaffolds. *J. Cell Sci.* 114, 2363–2373.
- Chen, S.L., Dowhan, D.H., Hosking, B.M., Muscat, G.E.O., 2000. The steroid receptor coactivator, GRIP-1, is necessary for MEF-2C-dependent gene expression and skeletal muscle differentiation. *Genes Dev.* 14, 1209–1228.
- Chen, Z., Hagler, J., Palombella, V.J., Melandri, F., Scherer, D., Ballard, D., Maniatis, T., 1995. Signal-induced site-specific phosphorylation targets I kappa B alpha to the ubiquitin-proteasome pathway. *Genes Dev.* 9, 1586–1597.
- Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., Boyer, L.A., Young, R.A., Jaenisch, R., 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci.* 107, 21931–21936.
- Daniels, D.L., Weis, W.I., 2005. β -catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. *Nat Struct Mol Biol* 12, 364–371.
- Davie, J.R., 2003. Inhibition of histone deacetylase activity by butyrate. *J. Nutr.* 133, 2485S–2493S.
- De Luca, A., Severino, A., De Paolis, P., Cottone, G., De Luca, L., De Falco, M., Porcellini, A., Volpe, M., Condorelli, G., 2003. p300/cAMP-response-element-binding-protein ('CREB')-binding protein (CBP) modulates co-operation between myocyte enhancer factor 2A (MEF2A) and thyroid hormone receptor-retinoid X receptor. *Biochem. J.* 369, 477–484.
- DiDonato, J.A., Hayakawa, M., Rothwarf, D.M., Zandi, E., Karin, M., 1997. A cytokine-responsive I κ B kinase that activates the transcription factor NF- κ B. *Nature* 388, 548–554.

- Doetzlhofer, A., Rotheneder, H., Lagger, G., Koranda, M., Kurtev, V., Brosch, G., Wintersberger, E., Seiser, C., 1999. Histone deacetylase 1 can repress transcription by binding to Sp1. *Mol. Cell. Biol.* 19, 5504–5511.
- Duret, L., Bucher, P., 1997. Searching for regulatory elements in human noncoding sequences. *Curr. Opin. Struct. Biol.* 7, 399–406.
- Eckner, R., Ewen, M.E., Newsome, D., Gerdes, M., DeCaprio, J.A., Lawrence, J.B., Livingston, D.M., 1994. Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. *Genes Dev.* 8, 869–884.
- Engelen, M., Ofman, R., Dijkgraaf, M.G.W., Hijzen, M., van der Wardt, L.A., van Geel, B.M., de Visser, M., Wanders, R.J.A., Poll-The, B.T., Kemp, S., 2010. Lovastatin in X-linked adrenoleukodystrophy. *N. Engl. J. Med.* 362, 276–277.
- Estrella, N.L., Desjardins, C.A., Nocco, S.E., Clark, A.L., Maksimenko, Y., Naya, F.J., 2015. MEF2 Transcription Factors Regulate Distinct Gene Programs in Mammalian Skeletal Muscle Differentiation. *J. Biol. Chem.* 290, 1256–1268.
- Evans, M.J., Scarpulla, R.C., 1989. Interaction of nuclear factors with multiple sites in the somatic cytochrome c promoter. Characterization of upstream NRF-1, ATF, and intron Sp1 recognition sequences. *J. Biol. Chem.* 264, 14361–14368.
- Falkenberg, K.J., Johnstone, R.W., 2014. Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat Rev Drug Discov* 13, 673–691.
- Faniello, M.C., Bevilacqua, M.A., Condorelli, G., de Crombrughe, B., Maity, S.N., Avvedimento, V.E., Cimino, F., Costanzo, F., 1999. The B subunit of the CAAT-binding factor NFY binds the central segment of the Co-activator p300. *J. Biol. Chem.* 274, 7623–7626.
- Forss-Petter, S., Werner, H., Berger, J., Lassmann, H., Molzer, B., Schwab, M.H., Bernheimer, H., Zimmermann, F., Nave, K.A., 1997. Targeted inactivation of the X-linked adrenoleukodystrophy gene in mice. *J. Neurosci. Res.* 50, 829–843.
- Fourcade, S., Ruiz, M., Camps, C., Schluter, A., Houten, S.M., Mooyer, P.A.W., Pampols, T., Dacremont, G., Wanders, R.J.A., Giros, M., Pujol, A., 2009. A key role for the peroxisomal ABCD2 transporter in fatty acid homeostasis. *Am. J. Physiol. Endocrinol. Metab.* 296, E211–221.
- Fourcade, S.S., Savary, S.S., Gondcaille, C., Berger, J., Netik, A., Cadepond, F.F., El Etr, M., Molzer, B., Bugaut, M., 2003. Thyroid hormone induction of the adrenoleukodystrophy-related gene (ABCD2). *Mol. Pharmacol.* 63, 1296–1303.
- Fourcade, Savary, S., Albet, S., Gauthier, D., Gondcaille, C., Pineau, T., Bellenger, J., Bentejac, M., Holzinger, A., Berger, J., Bugaut, M., 2001. Fibrate induction of the adrenoleukodystrophy-related gene (ABCD2): promoter analysis and role of the peroxisome proliferator-activated receptor PPARalpha. *Eur. J. Biochem.* 268, 3490–3500.

- Garrity, P.A., Wold, B.J., 1992. Effects of different DNA polymerases in ligation-mediated PCR: enhanced genomic sequencing and in vivo footprinting. *Proc. Natl. Acad. Sci. U. S. A.* 89, 1021–1025.
- Gervois, P., Torra, I.P., Fruchart, J.C., Staels, B., 2000. Regulation of lipid and lipoprotein metabolism by PPAR activators. *Clin. Chem. Lab. Med.* 38, 3–11.
- Gondcaille, C., Depreter, M., Fourcade, S., Lecca, M.R., Leclercq, S., Martin, P.G.P., Pineau, T., Cadepond, F., ElEtr, M., Bertrand, N., Beley, A., Duclos, S., De Craemer, D., Roels, F., Savary, S., Bugaut, M., 2005. Phenylbutyrate up-regulates the adrenoleukodystrophy-related gene as a nonclassical peroxisome proliferator. *J. Cell Biol.* 169, 93–104.
- Gondcaille, C., Genin, E.C., Lopez, T.E., Dias, A.M.M., Geillon, F., Andreoletti, P., Cherkaoui-Malki, M., Nury, T., Lizard, G., Weinhofer, I., Berger, J., Kase, E.T., Trompier, D., Savary, S., 2014. LXR antagonists induce ABCD2 expression. *Biochim. Biophys. Acta* 1841, 259–266.
- Gopalakrishnan, L., Scarpulla, R.C., 1995. Structure, expression, and chromosomal assignment of the human gene encoding nuclear respiratory factor 1. *J. Biol. Chem.* 270, 18019–18025.
- Gossett, L.A., Kelvin, D.J., Sternberg, E.A., Olson, E.N., 1989. A new myocyte-specific enhancer-binding factor that recognizes a conserved element associated with multiple muscle-specific genes. *Mol. Cell. Biol.* 9, 5022–5033.
- Gross, D.S., Garrard, W.T., 1988. Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.* 57, 159–197.
- Hampshire, A.J., Rusling, D.A., Broughton-Head, V.J., Fox, K.R., 2007. Footprinting: a method for determining the sequence selectivity, affinity and kinetics of DNA-binding ligands. *Methods* 42, 128–140.
- Han, A., Pan, F., Stroud, J.C., Youn, H.-D., Liu, J.O., Chen, L., 2003. Sequence-specific recruitment of transcriptional co-repressor Cabin1 by myocyte enhancer factor-2. *Nature* 422, 730–734.
- Han, J., Jiang, Y., Li, Z., Kravchenko, V. V., Ulevitch, R.J., 1997. Activation of the transcription factor MEF2C by the MAP kinase p38 in inflammation. *Nature* 386, 296–299.
- Handoko, L., Xu, H., Li, G., Ngan, C.Y., Chew, E., Schnapp, M., Lee, C.W.H., Ye, C., Ping, J.L.H., Mulawadi, F., Wong, E., Sheng, J., Zhang, Y., Poh, T., Chan, C.S., Kunarso, G., Shahab, A., Bourque, G., Cacheux-Rataboul, V., Sung, W.-K., Ruan, Y., Wei, C.-L., 2011. CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat Genet* 43, 630–638.
- Hasegawa, K., Meyers, M.B., Kitsis, R.N., 1997. Transcriptional coactivator p300 stimulates cell type-specific gene expression in cardiac myocytes. *J. Biol. Chem.* 272, 20049–20054.
- He, J., Ye, J., Cai, Y., Riquelme, C., Liu, J.O., Liu, X., Han, A., Chen, L., 2011. Structure of p300 bound to MEF2 on DNA reveals a mechanism of enhanceosome assembly. *Nucleic Acids Res.* 39, 4464–4474.

- Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., Ching, K.A., Antosiewicz, J.E., Liu, H., Zhang, X., Green, R.D., Stewart, R., Thomson, J.A., Crawford, G.E., Kellis, M., Ren, B., 2009. Histone Modifications at Human Enhancers Reflect Global Cell Type-Specific Gene Expression. *Nature* 459, 108–112.
- Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., Wang, W., Weng, Z., Green, R.D., Crawford, G.E., Ren, B., 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* 39, 311–318.
- Hoesel, B., Schmid, J.A., 2013. The complexity of NF- κ B signaling in inflammation and cancer. *Mol. Cancer* 12, 86.
- Issemann, I., Green, S., 1990. Activation of a member of the steroid hormone receptor superfamily by peroxisome proliferators. *Nature* 347, 645–650.
- Jayathilaka, N., Han, A., Gaffney, K.J., Dey, R., Jarusiewicz, J.A., Noridomi, K., Philips, M.A., Lei, X., He, J., Ye, J., Gao, T., Petasis, N.A., Chen, L., 2012. Inhibition of the function of class IIa HDACs by blocking their interaction with MEF2. *Nucleic Acids Res.* 40, 5378–5388.
- Jin, S., Scotto, K.W., 1998. Transcriptional Regulation of the MDR1 Gene by Histone Acetyltransferase and Deacetylase Is Mediated by NF- κ B. *Mol. Cell. Biol.* 18, 4377–4384.
- Jogl, G., Hsiao, Y.-S., Tong, L., 2004. Structure and function of carnitine acyltransferases. *Ann. N. Y. Acad. Sci.* 1033, 17–29.
- Johnson, D.S., Mortazavi, A., Myers, R.M., Wold, B., 2007. Genome-wide mapping of in vivo protein-DNA interactions. *Science* 316, 1497–1502.
- Kadamb, R., Mittal, S., Bansal, N., Batra, H., Saluja, D., 2013. Sin3: insight into its transcription regulatory functions. *Eur. J. Cell Biol.* 92, 237–246.
- Karin, M., 1999. How NF- κ B is activated: the role of the I κ B kinase (IKK) complex. *Oncogene* 18, 6867–6874.
- Karolchik, D., Baertsch, R., Diekhans, M., Furey, T.S., Hinrichs, A., Lu, Y.T., Roskin, K.M., Schwartz, M., Sugnet, C.W., Thomas, D.J., Weber, R.J., Haussler, D., Kent, W.J., 2003. The UCSC Genome Browser Database. *Nucleic Acids Res.* 31, 51–54.
- Keller, H., Dreyer, C., Medin, J., Mahfoudi, A., Ozato, K., Wahli, W., 1993. Fatty acids and retinoids control lipid metabolism through activation of peroxisome proliferator-activated receptor-retinoid X receptor heterodimers. *Proc. Natl. Acad. Sci.* 90, 2160–2164.
- Kemp, S., Berger, J., Aubourg, P., 2012. X-linked adrenoleukodystrophy: clinical, metabolic, genetic and pathophysiological aspects. *Biochim. Biophys. Acta* 1822, 1465–1474.
- Kemp, S., Pujol, A., Waterham, H.R., van Geel, B.M., Boehm, C.D., Raymond, G. V., 2001. ABCD1 mutations and the X-linked adrenoleukodystrophy mutation database: role in diagnosis and clinical correlations. *Hum Mutat* 18, 499–515.

- Kliwer, S.A., Umesono, K., Noonan, D.J., Heyman, R.A., Evans, R.M., 1992. Convergence of 9-cis retinoic acid and peroxisome proliferator signalling pathways through heterodimer formation of their receptors. *Nature* 358, 771–774.
- Koop, B.F., 1995. Human and rodent DNA sequence comparisons: a mosaic model of genomic evolution. *Trends Genet.* 11, 367–371.
- Kuo, M.-H., Allis, C.D., 1998. Roles of histone acetyltransferases and deacetylases in gene regulation. *Bioessays* 20, 615–626.
- Leclercq, S., Skrzypski, J., Courvoisier, A., Gondcaille, C., Bonnetain, F., Andre, A., Chardigny, J.-M., Bellenger, S., Bellenger, J., Narce, M., Savary, S., 2008. Effect of dietary polyunsaturated fatty acids on the expression of peroxisomal ABC transporters. *Biochimie* 90, 1602–1607.
- Lee, D.Y., Hayes, J.J., Pruss, D., Wolffe, A.P., 1993. A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* 72, 73–84.
- Levy, L., Wei, Y., Labalette, C., Wu, Y., Renard, C.-A., Buendia, M.A., Neuveut, C., 2004. Acetylation of beta-catenin by p300 regulates beta-catenin-Tcf4 interaction. *Mol. Cell. Biol.* 24, 3404–3414.
- Liang, H., Ward, W.F., 2006. PGC-1 α : a key regulator of energy metabolism. *Adv. Physiol. Educ.* 30, 145–151.
- Lin, J., Wu, H., Tarr, P.T., Zhang, C.-Y., Wu, Z., Boss, O., Michael, L.F., Puigserver, P., Isotani, E., Olson, E.N., Lowell, B.B., Bassel-Duby, R., Spiegelman, B.M., 2002. Transcriptional co-activator PGC-1[α] drives the formation of slow-twitch muscle fibres. *Nature* 418, 797–801.
- Logan, C.Y., Nusse, R., 2004. The Wnt signaling pathway in development and disease. *Annu. Rev. Cell Dev. Biol.* 20, 781–810.
- Lozzio, B.B., Lozzio, C.B., Bamberger, E.G., Feliu, A.S., 1981. A multipotential leukemia cell line (K-562) of human origin. *Proc. Soc. Exp. Biol. Med.* 166, 546–550.
- Lu, J., McKinsey, T.A., Nicol, R.L., Olson, E.N., 2000. Signal-dependent activation of the MEF2 transcription factor by dissociation from histone deacetylases. *Proc. Natl. Acad. Sci.* 97, 4070–4075.
- Ma, K., Chan, J.K.L., Zhu, G., Wu, Z., 2005. Myocyte enhancer factor 2 acetylation by p300 enhances its DNA binding activity, transcriptional activity, and myogenic differentiation. *Mol. Cell. Biol.* 25, 3575–3582.
- MacDonald, B.T., Tamai, K., He, X., 2009. Wnt/ β -catenin signaling: components, mechanisms, and diseases. *Dev. Cell* 17, 9–26.
- Mauro, C., Leow, S.C., Anso, E., Rocha, S., Thotakura, A.K., Tornatore, L., Moretti, M., De Smaele, E., Beg, A.A., Tergaonkar, V., 2011. NF-[κ] B controls energy homeostasis and metabolic adaptation by upregulating mitochondrial respiration. *Nat. Cell Biol.* 13, 1272–1279.

- McKinsey, T.A., Zhang, C.-L., Lu, J., Olson, E.N., 2000. Signal-dependent nuclear export of a histone deacetylase regulates muscle differentiation. *Nature* 408, 106–111.
- Merika, M., Williams, A.J., Chen, G., Collins, T., Thanos, D., 1998. Recruitment of CBP/p300 by the IFN β Enhanceosome Is Required for Synergistic Activation of Transcription. *Mol. Cell* 1, 277–287.
- Michael, L.F., Wu, Z., Cheatham, R.B., Puigserver, P., Adelmant, G., Lehman, J.J., Kelly, D.P., Spiegelman, B.M., 2001. Restoration of insulin-sensitive glucose transporter (GLUT4) gene expression in muscle cells by the transcriptional coactivator PGC-1. *Proc. Natl. Acad. Sci.* 98, 3820–3825.
- Morita, M., Imanaka, T., 2012. Peroxisomal ABC transporters: structure, function and role in disease. *Biochim. Biophys. Acta* 1822, 1387–1396.
- Moser, H.W., 1997. Adrenoleukodystrophy: phenotype, genetics, pathogenesis and therapy. *Brain* 120, 1485–1508.
- Moser, H.W., Moser, A.B., Frayer, K.K., Chen, W., Schulman, J.D., O'Neill, B.P., 1981. Adrenoleukodystrophy: increased plasma content of saturated very long chain fatty acids. *Neurology* 31, 1241.
- Mosser, J., Douar, A.M., Sarde, C.O., Kioschis, P., Feil, R., Moser, H., 1993. Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. *Nature* 361, 25.
- Mueller, P.R., Wold, B., Garritty, P. a, 2001. Ligation-mediated PCR for genomic sequencing and footprinting. *Curr. Protoc. Mol. Biol.* Chapter 15, Unit 15.3.
- Muneer, Z., Wiesinger, C., Voigtländer, T., Werner, H.B., Berger, J., Forss-Petter, S., 2014. Abcd2 Is a Strong Modifier of the Metabolic Impairments in Peritoneal Macrophages of Abcd1-Deficient Mice. *PLoS One* 9, e108655.
- Netik, A., Forss-Petter, S., Holzinger, A., Molzer, B., Unterrainer, G., Berger, J., 1999. Adrenoleukodystrophy-related protein can compensate functionally for adrenoleukodystrophy protein deficiency (X-ALD): implications for therapy. *Hum. Mol. Genet.* 8, 907–913.
- Novak, T.E., Babcock, T.A., Jho, D.H., Helton, W.S., Espat, N.J., 2003. NF- κ B inhibition by ω -3 fatty acids modulates LPS-stimulated macrophage TNF- α transcription. *Am. J. Physiol. Cell. Mol. Physiol.* 284, L84–L89.
- Ogryzko, V. V, Schiltz, R.L., Russanova, V., Howard, B.H., Nakatani, Y., 1996. The Transcriptional Coactivators p300 and CBP Are Histone Acetyltransferases. *Cell* 87, 953–959.
- Ong, C.-T., Corces, V.G., 2014. CTCF: An Architectural Protein Bridging Genome Topology and Function. *Nat. Rev. Genet.* 15, 234–246.
- Ornatsky, O.I., McDermott, J.C., 1996. MEF2 protein expression, DNA binding specificity and complex composition, and transcriptional activity in muscle and non-muscle cells. *J. Biol. Chem.* 271, 24927–24933.

- Pahan, K., Khan, M., Singh, I., 1998. Therapy for X-adrenoleukodystrophy: normalization of very long chain fatty acids and inhibition of induction of cytokines by cAMP. *J. Lipid Res.* 39, 1091–1100.
- Park, C.-Y., Kim, H.-S., Jang, J., Lee, H., Lee, J.S., Yoo, J.-E., Lee, D.R., Kim, D.-W., 2013. ABCD2 is a direct target of beta-catenin and TCF-4: implications for X-linked adrenoleukodystrophy therapy. *PLoS One* 8, e56242.
- Peng, Y., Jahroudi, N., 2003. The NFY Transcription Factor Inhibits von Willebrand Factor Promoter Activation in Non-endothelial Cells through Recruitment of Histone Deacetylases. *J. Biol. Chem.* 278, 8385–8394.
- Peters, C., Charnas, L.R., Tan, Y., Ziegler, R.S., Shapiro, E.G., DeFor, T., Grewal, S.S., Orchard, P.J., Abel, S.L., Goldman, A.I., Ramsay, N.K.C., Dusenbery, K.E., Loes, D.J., Lockman, L.A., Kato, S., Aubourg, P.R., Moser, H.W., Krivit, W., 2004. Cerebral X-linked adrenoleukodystrophy: the international hematopoietic cell transplantation experience from 1982 to 1999. *Blood* 104, 881–888.
- Pinto, D., Gregorieff, A., Begthel, H., Clevers, H., 2003. Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. *Genes Dev.* 17, 1709–1713.
- Potthoff, M.J., Olson, E.N., 2007. MEF2: a central regulator of diverse developmental programs. *Development* 134, 4131–4140.
- Puigserver, P., Adelmant, G., Wu, Z., Fan, M., Xu, J., O'Malley, B., Spiegelman, B.M., 1999. Activation of PPAR γ coactivator-1 through transcription factor docking. *Science* 286, 1368–1371.
- Pujol, A., Ferrer, I., Camps, C., Metzger, E., Hindelang, C., Callizot, N., Ruiz, M., Pampols, T., Giros, M., Mandel, J.L., 2004. Functional overlap between ABCD1 (ALD) and ABCD2 (ALDR) transporters: a therapeutic target for X-adrenoleukodystrophy. *Hum. Mol. Genet.* 13, 2997–3006.
- Pujol, A., Troffer-Charlier, N., Metzger, E., Chimini, G., Mandel, J.L., 2000. Characterization of the adrenoleukodystrophy-related (ALDR, ABCD2) gene promoter: inductibility by retinoic acid and forskolin. *Genomics* 70, 131–139.
- Ramplér, H., Weinhofer, I., Netik, A., Forss-Petter, S., Brown, P.J., Oplinger, J.A., Bugaut, M., Berger, J., 2003. Evaluation of the therapeutic potential of PPAR α agonists for X-linked adrenoleukodystrophy. *Mol. Genet. Metab.* 80, 398–407.
- Rye, M., Sætrom, P., Håndstad, T., Drabløs, F., 2011. Clustered ChIP-Seq-defined transcription factor binding sites and histone modifications map distinct classes of regulatory elements. *BMC Biol.* 9, 80.
- Salsi, V., Caretti, G., Wasner, M., Reinhard, W., Haugwitz, U., Engeland, K., Mantovani, R., 2003. Interactions between p300 and multiple NF-Y trimers govern cyclin B2 promoter function. *J. Biol. Chem.* 278, 6642–6650.
- Sandmann, T., Jensen, L.J., Jakobsen, J.S., Karzynski, M.M., Eichenlaub, M.P., Bork, P., Furlong, E.E.M., 2006. A temporal map of transcription factor activity: mef2 directly regulates target genes at all stages of muscle development. *Dev. Cell* 10, 797–807.

- Savickiene, J., Borutinskaite, V.-V., Treigyte, G., Magnusson, K.-E., Navakauskiene, R., 2006. The novel histone deacetylase inhibitor BML-210 exerts growth inhibitory, proapoptotic and differentiation stimulating effects on the human leukemia cell lines. *Eur. J. Pharmacol.* 549, 9–18.
- Schuettengruber, B., Simboeck, E., Khier, H., Seiser, C., 2003. Autoregulation of mouse histone deacetylase 1 expression. *Mol. Cell. Biol.* 23, 6993–7004.
- Schulze-Osthoff, K., Ferrari, D., Riehemann, K., Wesselborg, S., 1997. Regulation of NF- κ B activation by MAP kinase cascades. *Immunobiology* 198, 35–49.
- Sekiya, M., Yahagi, N., Matsuzaka, T., Najima, Y., Nakakuki, M., Nagai, R., Ishibashi, S., Osuga, J., Yamada, N., Shimano, H., 2003. Polyunsaturated fatty acids ameliorate hepatic steatosis in obese mice by SREBP-1 suppression. *Hepatology* 38, 1529–1539.
- Shapiro, E., Krivit, W., Lockman, L., Jambaque, I., Peters, C., Cowan, M., 2000. Long-term effect of bone-marrow transplantation for childhood-onset cerebral X-linked adrenoleukodystrophy. *Lancet* 356, 713–718.
- Shore, P., Sharrocks, A.D., 1995. The MADS-box family of transcription factors. *Eur. J. Biochem.* 229, 1–13.
- Singh, I., Khan, M., Key, L., Pai, S., 1998. Lovastatin for X-linked adrenoleukodystrophy. *N Engl J Med* 339, 702–703.
- Sinha, S., Maity, S.N., Lu, J., de Crombrughe, B., 1995. Recombinant rat CBF-C, the third subunit of CBF/NFY, allows formation of a protein-DNA complex with CBF-A and CBF-B and with yeast HAP2 and HAP3. *Proc. Natl. Acad. Sci.* 92, 1624–1628.
- Smith, K.D., Kemp, S., Braiterman, L.T., Lu, J.F., Wei, H.M., Geraghty, M., Stetten, G., Bergin, J.S., Pevsner, J., Watkins, P.A., 1999. X-linked adrenoleukodystrophy: genes, mutations, and phenotypes. *Neurochem. Res.* 24, 521–535.
- Sparrow, D.B., Miska, E.A., Langley, E., Reynaud-Deonauth, S., Kotecha, S., Towers, N., Spohr, G., Kouzarides, T., Mohun, T.J., 1999. MEF-2 function is modified by a novel co-repressor, MITR. *EMBO J.* 18, 5085–5098.
- Strahl, B.D., Allis, C.D., 2000. The language of covalent histone modifications. *Nature* 403, 41–45.
- Su, A.I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K.A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., Cooke, M.P., Walker, J.R., Hogenesch, J.B., 2004. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci. U. S. A.* 101, 6062–6067.
- Suzuki, H., Hosokawa, Y., Toda, H., Nishikimi, M., Ozawa, T., 1989. Isolation of a single nuclear gene encoding human ubiquinone-binding protein in complex III of mitochondrial respiratory chain. *Biochem. Biophys. Res. Commun.* 161, 371–378.
- Tie, F., Banerjee, R., Stratton, C.A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M.O., Scacheri, P.C., Harte, P.J., 2009. CBP-mediated acetylation of histone H3 lysine 27 antagonizes *Drosophila* Polycomb silencing. *Development* 136, 3131–3141.

- van Geel, B.M., Bezman, L., Loes, D.J., Moser, H.W., Raymond, G. V, 2001. Evolution of phenotypes in adult male patients with X-linked adrenoleukodystrophy. *Ann Neurol* 49, 186–194.
- van Geel, B.M., Poll-The, B.T., Verrips, A., Boelens, J.-J., Kemp, S., Engelen, M., 2015. Hematopoietic cell transplantation does not prevent myelopathy in X-linked adrenoleukodystrophy: a retrospective study. *J. Inherit. Metab. Dis.* 38, 359–361.
- van Roermund, C.W.T., Visser, W.F., IJlst, L., van Cruchten, A., Boek, M., Kulik, W., 2008. The human peroxisomal ABC half transporter ALDP functions as a homodimer and accepts acyl-CoA esters. *FASEB J* 22, 4201–4208.
- van Roermund, C.W.T., Visser, W.F., IJlst, L., Waterham, H.R., Wanders, R.J.A., 2011. Differential substrate specificities of human ABCD1 and ABCD2 in peroxisomal fatty acid β -oxidation. *Biochim. Biophys. Acta (BBA)-Molecular Cell Biol. Lipids* 1811, 148–152.
- Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., Afzal, V., Ren, B., Rubin, E.M., Pennacchio, L.A., 2009. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457, 854–858.
- Vo, N., Goodman, R.H., 2001. CREB-binding protein and p300 in transcriptional regulation. *J. Biol. Chem.* 276, 13505–13508.
- Wang, X., Sato, R., Brown, M.S., Hua, X., Goldstein, J.L., 1994. SREBP-1, a membrane-bound transcription factor released by sterol-regulated proteolysis. *Cell* 77, 53–62.
- Weber, F.D., Weinhofer, I., Einwich, A., Forss-Petter, S., Muneer, Z., Maier, H., Weber, W.H.A., Berger, J., 2014a. Evaluation of retinoids for induction of the redundant gene ABCD2 as an alternative treatment option in X-linked adrenoleukodystrophy. *PLoS One* 9, e103742.
- Weber, F.D., Wiesinger, C., Forss-Petter, S., Regelsberger, G., Einwich, A., Weber, W.H.A., Koehler, W., Stockinger, H., Berger, J., 2014b. X-linked adrenoleukodystrophy: Very long-chain fatty acid metabolism is severely impaired in monocytes but not in lymphocytes. *Hum. Mol. Genet.* 23, 2542–2550.
- Weinhofer, I., Forss-Petter, S., Žigman, M., Berger, J., 2002. Cholesterol regulates ABCD2 expression: implications for the therapy of X-linked adrenoleukodystrophy. *Hum. Mol. Genet.* 11, 2701–2708.
- Weinhofer, I., Kunze, M., Rampler, H., Bookout, A.L., Forss-Petter, S., Berger, J., 2005. Liver X receptor alpha interferes with SREBP1c-mediated Abcd2 expression. Novel cross-talk in gene regulation. *J. Biol. Chem.* 280, 41243–41251.
- Weinhofer, I., Kunze, M., Rampler, H., Forss-Petter, S., Samarut, J., Plateroti, M., Berger, J., 2008. Distinct modulatory roles for thyroid hormone receptors TRalpha and TRbeta in SREBP1-activated ABCD2 expression. *Eur. J. Cell Biol.* 87, 933–945.
- Wiesinger, C., Eichler, F.S., Berger, J., 2015. The genetic landscape of X-linked adrenoleukodystrophy: inheritance, mutations, modifier genes, and diagnosis. *Appl. Clin. Genet.* 8, 109–121.

- Wiesinger, C., Kunze, M., Regelsberger, G., Forss-Petter, S., Berger, J., 2013. Impaired very long-chain acyl-CoA beta-oxidation in human X-linked adrenoleukodystrophy fibroblasts is a direct consequence of ABCD1 transporter dysfunction. *J. Biol. Chem.* 288, 19269–19279.
- Wu, Z., Puigserver, P., Andersson, U., Zhang, C., Adelmant, G., Mootha, V., Troy, A., Cinti, S., Lowell, B., Scarpulla, R.C., 1999. Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell* 98, 115–124.
- Xi, H., Shulha, H.P., Lin, J.M., Vales, T.R., Fu, Y., Bodine, D.M., McKay, R.D.G., Chenoweth, J.G., Tesar, P.J., Furey, T.S., 2007. Identification and characterization of cell type–specific and ubiquitous chromatin regulatory structures in the human genome. *PLoS Genet* 3, e136.
- Xia, Y., Buja, L.M., Scarpulla, R.C., McMillin, J.B., 1997. Electrical stimulation of neonatal cardiomyocytes results in the sequential activation of nuclear genes governing mitochondrial proliferation and differentiation. *Proc. Natl. Acad. Sci.* 94, 11399–11404.
- Yoshikawa, T., Shimano, H., Yahagi, N., Ide, T., Amemiya-Kudo, M., Matsuzaka, T., Nakakuki, M., Tomita, S., Okazaki, H., Tamura, Y., 2002. Polyunsaturated fatty acids suppress sterol regulatory element-binding protein 1c promoter activity by inhibition of liver X receptor (LXR) binding to LXR response elements. *J. Biol. Chem.* 277, 1705–1711.
- Youn, H.D., Liu, J.O., 2000. Cabin1 represses MEF2-dependent Nur77 expression and T cell apoptosis by controlling association of histone deacetylases and acetylases with MEF2. *Immunity* 13, 85–94.
- Youn, H.D., Sun, L., Prywes, R., Liu, J.O., 1999. Apoptosis of T cells mediated by Ca²⁺-induced release of the transcription factor MEF2. *Science* 286, 790–793.
- Zhao, W.-X., Tian, M., Zhao, B.-X., Li, G.-D., Liu, B., Zhan, Y.-Y., Chen, H.-Z., Wu, Q., 2007. Orphan receptor TR3 attenuates the p300-induced acetylation of retinoid X receptor-alpha. *Mol. Endocrinol.* 21, 2877–2889.
- Zianni, M., Tessanne, K., Merighi, M., Laguna, R., Tabita, F.R., 2006. Identification of the DNA Bases of a DNase I Footprint by the Use of Dye Primer Sequencing on an Automated Capillary DNA Analysis Instrument. *J. Biomol. Tech.* 17, 103–113.

8. Appendix:

8.1. List of Figures:

Figure 1:	Overlapping substrates of ABCD1 and ABCD2	6
Figure 2:	Expression levels of <i>ABCD1</i> and <i>ABCD2</i> in different immune cells	8
Figure 3:	Known regulatory mechanisms at the <i>ABCD2</i> promoter	9
Figure 4:	Illustration of the <i>in vivo</i> footprinting procedure	22
Figure 5:	The four main steps of an LM-PCR	23
Figure 6:	Evolutionary conservation of the <i>ABCD2</i> promoter region	26
Figure 7:	ChIP-seq peaks and binding sites at evolutionary conserved elements	28
Figure 8:	Combined ChIP-seq data for the <i>ABCD2</i> promoter region	31
Figure 9:	Schematic illustration of the <i>ABCD2</i> promoter	31
Figure 10:	DNase hypersensitive regions at the <i>ABCD2</i> promoter	33
Figure 11:	Overlay of an <i>in vivo</i> and <i>in vitro</i> electropherogram of CD14 ⁺ cells	35
Figure 12:	Combined <i>in vivo</i> footprinting data of CD3 ⁺ and CD14 ⁺ cells.	37
Figure 13:	Illustration of the IFN β enhanceosome	43
Figure 14:	Proposed model of an p300/MEF2 enhanceosome	45
Figure 15:	Hypothetical enhanceosome complex at the <i>ABCD2</i> promoter	47
Figure 16:	Model of factors which might prevent the enhanceosome assembly	48

8.2. List of Tables:

Table 1:	Cell types and used abbreviations	18
Table 2:	Description of selected cell types	19
Table 3:	Sequences and positions of the used oligonucleotides	24
Table 4:	Evolutionary conserved regions in the promoter of <i>ABCD2</i>	27
Table 5:	ChIP-seq signal intensities of selected factors	32
Table 6:	Cell types with DNase hypersensitivity at putative regulatory sites	34

8.3. Abstract (German):

Die X-chromosomale Adrenoleukodystrophie (X-ALD) ist eine erbliche Stoffwechselerkrankung, mit einer Häufigkeit von etwa 1 : 16800. Die Krankheit wird durch Mutationen im X-chromosomalen Gen *ABCD1* verursacht, welches für einen peroxisomalen Transporter kodiert. In X-ALD ist mangels funktionalem *ABCD1*, der Transport von CoA-aktivierten überlangkettigen Fettsäuren (ÜLFS) in die Peroxisomen gestört. Aus diesem Grund akkumulieren ÜLFS im Gewebe und Plasma von Patienten. Dies führt zur Demyelinisierung von Nerven und im schlimmsten Fall zu Entzündungsreaktionen im Gehirn, Neurodegeneration und zum Tod. Das homologe Gen *ABCD2* ist fähig den ÜLFS-Gehalt zu reduzieren, wenn es überexprimiert wird. Aus diesem Grund wurde die Induktion von *ABCD2* als eine mögliche Behandlung für X-ALD vorgeschlagen. Leider ist die Regulation der *ABCD2*-Expression bisher noch unzureichend verstanden.

Da *ABCD2* in Immunzellen wie etwa T-Zellen, B-Zellen und Monozyten, in sehr unterschiedlicher Höhe exprimiert wird, kann ein Vergleich dieser Zellen regulatorische Mechanismen aufzeigen. Dazu wurden 3 verschiedene Methoden verwendet. Zuerst wurden phylogenetische Footprinting Daten verwendet um mögliche regulatorische Elemente zu finden die evolutionär konserviert sind. Durch den Vergleich der Sequenzidentität von 100 Wirbeltierarten ist es möglich funktionale von nicht-funktionalen Abschnitten zu unterscheiden. In Kombination mit vorhandenen ChIP-seq Daten können mögliche Bindepartner dieser konservierten Bereiche gefunden werden. Um die Funktion mutmaßlicher regulatorischer Elemente genauer zu bestimmen, wurden ChIP-seq Daten von Histon-Modifikationen untersucht. Posttranslationelle Modifikationen wie etwa Acetylierung oder Methylierung können die Genexpression beeinflussen, indem sie den Verpackungsgrad des Chromatins ändern oder regulatorische Faktoren rekrutieren.

ChIP-seq Signale überspannen oft einen breiten Bereich, wodurch es schwierig ist genau zu sagen wo ein Protein gebunden war. Um ein genaueres Bild von den Bindestellen zu bekommen wurde ein *in vivo* Footprinting Verfahren mit CD3⁺ und CD14⁺ Zellen durchgeführt. Das *in vivo* Footprinting ermöglicht es sehr genau zu bestimmen, welche Bereiche der DNA in der lebenden Zelle von Proteinen gebunden waren. Mit dieser Methode sollen Zelltyp-spezifische Unterschiede gefunden werden, um die Regulation des *ABCD2* Promoters besser zu verstehen.

Unglücklicherweise hatten unsere *in vivo* Footprinting Daten zu starke Schwankungen, wodurch es nicht möglich war eine genaue Aussage über gebundene Stellen zu machen. Trotz mehrerer erfolgreicher Verbesserungen der Datenqualität konnten schlussendlich keine vertrauenswürdigen Ergebnisse erzielt werden. Die Analyse der phylogenetischen Footprinting Daten lieferte jedoch einige interessante Ergebnisse. Neben den bereits bekannten regulatorischen Elementen, enthält der Promoter von *ABCD2* eine Reihe weiterer hoch konservierter Bereiche. Diese Bereiche spielen wahrscheinlich eine Rolle für die Regulation von *ABCD2*.

Zusätzlich befinden sich zwei vermeintliche regulatorische Elemente weiter oberhalb des Promoters. Aufgrund von bestimmten Histon-Modifikationen und zugänglichem Chromatin ist es wahrscheinlich, dass diese beiden Elemente als Enhancer agieren. An beiden Stellen wurde der Myozyten Enhancer Faktor-2 (MEF2) gefunden. Durch Interaktionen mit dem Co-Aktivator p300 könnte MEF2 die Transkription stimulieren. Weitere mögliche Interaktionspartner und deren Effekt auf die Genexpression werden beschrieben, um die Ursachen für zelltypspezifische Unterschiede zu erläutern. Letztendlich soll dies dazu beitragen *ABCD2* zu induzieren, als mögliche Behandlung für X-ALD.

8.4. Abstract (English):

X-linked Adrenoleukodystrophy (X-ALD) is an inherited metabolic disease with a frequency of about 1 in 16800 new-borns. The disease is caused by mutations in the X-chromosomal gene *ABCD1*, which codes for a peroxisomal transporter. In X-ALD, due to *ABCD1* deficiency the peroxisomal import of CoA-activated very long-chain fatty acids (VLCFAs) is impaired. Because of that, VLCFAs are no longer degraded by β -oxidation and accumulate in the plasma and tissues of patients. This causes demyelination of nerves and in the most severe form cerebral inflammation, neurodegeneration and death. The homologous gene *ABCD2* is able to normalize VLCFA-levels upon overexpression, thus induction of *ABCD2* has been suggested as a possible treatment for X-ALD. Unfortunately, the regulation of the *ABCD2* gene expression is still not entirely understood.

Because *ABCD2* is expressed at very different levels in various immune cells like T cells, B cells and monocytes, a comparison between them can reveal regulatory mechanisms. To achieve this, three different approaches were combined. First, phylogenetic footprinting datasets were analysed to find potential regulatory regions that are evolutionary conserved. By comparing the degree of conservation between 100 vertebrate species, it is possible to distinguish between functional and non-functional areas of the promoter. Combined with information from available ChIP-seq datasets, factors that are bound to the promoter are identified. With this approach, it is possible to find candidate factors that are involved in the regulation of a gene. To further clarify the function of putative regulatory regions, ChIP-seq data for histone modifications was used. Posttranslational modifications like acetylations or methylations can affect gene expression by changing the accessibility of the chromatin or by recruiting regulatory factors. Knowledge of these epigenetic marks can help to understand cell type specific differences in the regulation of the *ABCD2* promoter.

Since ChIP-seq peaks often cover broad areas, it can be difficult to exactly pinpoint the corresponding binding site. To get a finer picture of utilized regulatory elements, an *in vivo* footprinting assay with CD3⁺ and CD14⁺ cells was done, from a region with a high density of conserved elements. *In vivo* footprinting helps to identify sites of the DNA that were bound by transcription factors in single base pair resolution. This data can then be used to find cell type specific differences.

Unfortunately, the *in vivo* footprinting procedure did not work as expected. Our data had too high variations to make definite claims about bound regions. Numerous attempts to improve the data quality were successful, but despite these advances no reliable results could be generated. Nonetheless, the phylogenetic footprinting proved helpful for finding putative regulatory regions. Besides the already known elements that were described in the literature, the *ABCD2* promoter contains two additional well-conserved regions. These regions are likely to play a role for the regulation of the *ABCD2* gene.

In addition, two putative regulatory elements were found further upstream of the promoter. Based on the presence of certain histone modifications and open chromatin it is likely for them to act as enhancers. At both putative enhancers the myocyte enhancer factor-2 (MEF2) is found. By interacting with the transcriptional co-activator p300, MEF2 bound enhancers could promote gene expression. In addition, also other potential interaction partners and their effect on gene expression are discussed to shed some light on the reasons for cell type specific gene expression. In the future, this might help to induce *ABCD2*, as a possible treatment for X-ALD.