DISSERTATIONES
MEDICINAE
UNIVERSITATIS
TARTUENSIS
266

MARIO SAARE

The influence of AIRE on gene expression – studies of transcriptional regulatory mechanisms in cell culture systems

DISSERTATIONES MEDICINAE UNIVERSITATIS TARTUENSIS
266

DISSERTATIONES MEDICINAE UNIVERSITATIS TARTUENSIS
266

MARIO SAARE

The influence of AIRE on gene expression – studies of transcriptional regulatory mechanisms in cell culture systems

UNIVERSITY OF TARTU
Press

Molecular Pathology Research Group, Institute of Biomedicine and Translational Medicine, University of Tartu

The dissertation is accepted for commencement of the degree of Doctor of Philosophy in Medicine on the 18th of October, 2017 by the Council of the Faculty of Medicine, University of Tartu, Estonia.

Supervisors:	Pärt Peterson, PhD, Professor of Molecular Immunology, Molecular Pathology, Institute of Biomedicine and Translational Medicine, University of Tartu, Estonia. Ana Rebane, Senior Research Fellow, RNA Biology Research Group, Institute of Biomedicine and Translational Medicine, University of Tartu, Estonia.
Reviewers:	Meeme Utt, PhD, Senior Research Fellow in Immunology, Department of Immunology, Institute of Biomedicine and Translational Medicine, University of Tartu, Estonia Ivar Ilves, PhD, Senior Research Fellow in Biomedicine, Institute of Technology, University of Tartu, Estonia
Opponent:	Mitsuru Matsumoto, PhD, Professor, Division of Molecular Immunology, Institute for Enzyme Research, University of Tokushima, Japan

Commencement: 19th of December, 2017

Publication of this dissertation is granted by the University of Tartu.

This research was supported by the European Union through European Regional Development Fund and the European Social Fund, the Estonian Research Agency, the Estonian Ministry of Education and Research, the Tartu University Development Fund and the Center of Translational Genomics.

European Union
Regional Development Fund

Investing in your future

ISSN 1024-395X
ISBN 978-9949-77-622-1 (print)
ISBN 978-9949-77-621-4 (pdf)

University of Tartu Press
www.tyk.ee

TABLE OF CONTENT

LIST OF ORIGINAL PUBLICATIONS

I Saare, M., Rebane, A., Rajashekar, B., Vilo, J., and Peterson, P. (2012). Autoimmune regulator is acetylated by transcription coactivator CBP/p300. Exp. Cell Res. *318*, 1767–1778.

II Gaetani, M.*, Matafora, V.*, Saare, M., Spiliotopoulos, D., Mollica, L., Quilici, G., Chignola, F., Mannella, V., Zucchelli, C., Peterson, P., et al. (2012). AIRE-PHD fingers are structural hubs to maintain the integrity of chromatin-associated interactome. Nucleic Acids Res. *40*, 11756–11768.

III Guha, M.*, Saare, M.*, Maslovskaja, J.*, Kisand, K., Liiv, I., Haljasorg, U., Tasa, T., Metspalu, A., Milani, L., and Peterson, P. (2017). DNA breaks and chromatin structural changes enhance the transcription of autoimmune regulator target genes. J. Biol. Chem. *292*, 6542–6554.
*These authors contributed equally to this work.

Contribution of Mario Saare to the original publications:

Study I: participation in the study design, performing experiments, analysing the data, creating the figures and writing the paper.

Study II: performing quantitative real-time PCR and *in vitro* GST pull-down experiments, analysing the data, creating the figures and writing the paper.

Study III: participation in the study design, performing 3C experiments, analysing FAIRE-seq and 3C data, creating the figures and writing the paper.

ABBREVIATIONS

14-3-3 proteins	family of proteins defined by their chromatographic migration pattern
3C	chromosome conformation capture
AD	Addison's disease
AIRE	autoimmune regulator
AIRE-Tet	HEK293 cell line with doxycycline-inducible AIRE expression
AP-1	activating protein 1
APS1	autoimmune polyendocrinopathy syndrome 1
ATM	ataxia telangiectasia mutated protein
ATR	ataxia telangiectasia and Rad3 related protein
BAI3	brain-specific angiogenesis inhibitor 3
BAT2	large proline-rich protein BAT2
BCR	B-cell receptor
BIR	baculovirus IAP repeat protein domain
BPTF	bromodomain PHD finger transcription factor
BRCT	tandem breast cancer susceptibility protein domain
BRD4	bromodomain-containing protein 4
CARD	caspase recruitment domain
CBP	CREB-binding protein (CBP)
CD4/8	cluster of differentiation 4 or 8
CDH11	cadherin 11
CDK7/9	cyclin-dependent kinase 7 or 9
CEACAM	carcinoembryonic antigen-related cell adhesion molecule
ChIP	chromatin immunoprecipitation
ChIP-seq	ChIP coupled with high-throughput sequencing
CHST13	carbohydrate sulphotransferase 13
CLDN1	claudin 1
CMC	chronic mucocutaneous candidiasis
CNS1	conserved non-coding sequence 1
CpG	5'-cytosine-phosphate-guanine-3'
CREB	cAMP response element-binding protein
cRNA	complementary RNA
CT	chromosome territory
CTCF	CCCTC-binding factor
Ctcfl	CCCTC-binding factor like protein

CTD	RNAP II C-terminal domain
cTEC	cortical thymic epithelial cell
Ctrl	uninduced/untreated AIRE-Tet cells
CXCR7	C-X-C chemokine receptor type 7
DAPI	4',6-diamidino-2-phenylindole
DC	dendritic cell
DEAF1	deformed epidermal autoregulatory factor 1
DMEM	Dulbeccos's Modified Eagle Medium
DNA-PK	DNA-dependent protein kinase
Dox	doxycycline-treated AIRE-Tet cells
Dox+Etop	doxycycline-induced and etoposide-treated AIRE-Tet cells
DPF3	double PHD fingers 3 protein
ECL	enhanced chemicluminescence
EDTA	ethylenediaminetetraacetic acid
ENCODE	Encyclopedia of DNA Elements
eTAC	extra-thymic AIRE-expressing cell
Etop	etoposide-treated AIRE-Tet cells
Ets	E-twenty-six family transcription factor
FAIRE	formaldehyde-assisted isolation of regulatory elements
FAS	Fas cell surface death receptor
FBS	fetal bovine serum
FOG1	friend of GATA1 protein
FOS	proto-oncogene c-Fos
FOSL1	FOS like antigen 1
FOXO1	forkhead box O1 transcription factor
FOXP3	forkhead box 3 transcription factor
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GATA1	GATA (nucleotide sequence) binding protein 1
γH2AX	histone H2AX phosphorylated at serine 139
GMEB	glucocorticoid modulatory element binding protein
GST	glutathione S-transferase
H3K27me3	histone H3 trimethylated at lysine 27
H3K36me3	histone H3 trimethylated at lysine 36
H3K4me0/3	histone H3 unmethylated or trimethylated at lysine 4
H3K9me3	histone H3 trimethylated at lysine 9
H3S10ph	histone H3 phosphorylated at serine 10
H4K16ac	histone H4 acetylated at lysine 16

H4K20me3	histone H4 trimethylated at lysine 20
HBG2	haemoglobin subunit gamma 2
HEK-AIRE	HEK293 cell line with constitutive AIRE expression
HEK-AIRE. K243/253Q	HEK293 cell line with constitutive AIRE.K243/253Q expression
HEK-YFP	HEK293 cell line with constitutive yellow fluorescent protein expression
HEK293	human embryonic kidney cell line 293
HEK293T	human embryonic kidney cell line 293 expressing the simian vacuolating virus 40 large T antigen
HMGB1	high mobility group protein B1
Hox	homeobox-containing transcription factor
HP	hypoparathyroidism
HP1	heterochromatin protein 1
HPRT1	hypoxanthine phosphoribosyltransferase 1
HRP	horseradish peroxidase
IAP	baculovirus inhibitor of apoptosis
ICD	interchromatin domain
IF	immunofluorescence
IFI16	interferon gamma inducible protein 16
IgG	normal goat serum IgG
IP	immunoprecipitation
Irf4/8	interferon regulatory factor 4 or 8
ISGF3	interferon-stimulated gene factor 3
IVL	involucrin
JAK-STAT	Janus kinase and signal transducer of activator of transcription signalling pathway
Jmjd6	JmjC Domain-Containing Protein 6
K[DN]WK	lysine-[aspartic acid/asparagine]-tryptophan-lysine sequence motif
KAT8	lysine acetyltransferase 8
KMT1A	lysine methyltransferase 1A
KRT73	keratin 73
Ku70/Ku80	70 or 80 kDa subunit of Ku antigen
LPL	lipoprotein lipase
LXXLL	Leu-Xaa-Xaa-Leu-Leu amino acid sequence motif
LY6G6D	lymphocyte antigen 6 family member G6D
mChIP	modified chromatin immunoprecipitation

MHC	major histocompatibility complex
$MHCII^{hi}$	high expression of MHC class 2 molecules
mTEC	medullary thymic epithelial cell
NCBI	National Center for Biotechnology Information
NF-Y	nuclear factor Y
NF-κB	nuclear factor kappa B family of transcription factors
NGS	normal goat serum
NLS	nuclear localisation signal
NucP41/P75	nuclear protein 41 kDa/75 kDa
NUDR	nuclear DEAF1-related protein
P-TEFb	positive transcription elongation factor b
p300	adenovirus early region 1A-associated 300 kDa protein
PAPLN	papilin, proteoglycan like sulphated glycoprotein
PARP1	poly [ADP-ribose] polymerase 1
PBS	phosphate-buffered saline
PCAF	p300/CBP-associated factor
PCR	polymerase chain reaction
PDYN	prodynorphin
PHD1/2	plant homeodomain 1 or 2
PML	promyelocytic leukaemia protein
PRC1/2	polycomb repressive complex 1 or 2
PRR	proline-rich region
PSMD4	proteasome 26S subunit, non-ATPase 4
PTM	post-translational modification
PWWP	proline-tryptophan-tryptophan-proline sequence containing protein domain
qPCR	quantitative polymerase chain reaction
Rad3	protein kinase rad3
RING	really interesting new gene type zinc finger domain
RIPA	radio immunoprecipitation assay
RNAi	RNA interference
RNAP II	RNA polymerase II
Rtt109	regulator of Ty1 transposition protein 109
S100A10	S100 calcium binding protein A10
S100A8	S100 calcium binding protein A8
SAND	SP100, AIRE1, NucP41/P75 and DEAF1 protein domain
SDS-PAG	sodium dodecyl sulphate polyacrylamide gel

SEM	standard error of the mean
Sirt1	sirtuin 1 deacetylase
SIRT2	sirtuin 2 deacetylase
SKP2	S-phase kinase-associated protein 2
Sp1	specificity protein 1
SP100	speckled protein 100 kDa
STAT1/2	transducer and activator of transcription 1 or 2
TAD	topologically associated domain
TAF250	TATA (nucleotide sequence) binding protein-associated factor 250 kDa protein
Tbx21	T-box protein 21
Tcf7	T-cell factor 7
TCR	T-cell receptor
TFIIE/F/H	transcription factor II E, F or H
TOP1	DNA topoisomerase 1
TOP2A/B	DNA topoisomerase 2-alpha or 2-beta
TRA	tissue-restricted antigen
Treg	regulatory T-cell
TSA	Trichostatin A
Ty1	yeast transposon 1
Ube3a	ubiquitin-protein ligase E3A
WB	western blot
YFP	yellow fluorescent protein

1. INTRODUCTION

The immune system has major role in maintaining the homeostasis of the organism by reacting to and eliminating invading pathogens as efficiently as possible, while preventing harmful effects to the body's own cells and tissues. However, the immune effector cells can become dangerous to self and contribute to autoimmunity if the molecular mechanisms of immune tolerance fail for some reason. Autoimmune disorders are mainly complex diseases that correlate with genetic and environmental factors, but the causal order of cellular and molecular events that lead to the pathological condition are often poorly understood.

The autoimmune polyendocrinopathy syndrome 1 (APS1) is a rare exception where the cause can be traced back to the deficiency of a single gene – the autoimmune regulator (*AIRE*) gene. APS1 manifests in adverse immune reactions against multiple organs in conjunction with loss of immunity to fungal infections. The expression of the AIRE protein is almost exclusively limited to the medullary thymic epithelial cells (mTECs) where AIRE acts as transcriptional regulator of thousands of genes that are considered to be tissue-restricted antigens (TRAs). The expressed TRAs are fragmented and the self-peptides are loaded onto the major histocompatibility complex (MHC) proteins, which react with the intact T-cell receptor (TCR) of the developing thymocytes. Thymocytes that have high affinity toward self-peptides undergo apoptosis, which helps to establish and maintain immunological tolerance.

Although the fundamental principles of many of the events outlined above have gained ever deeper insight, the details of the molecular mechanisms governing AIRE-dependent gene expression remain elusive and under intense scientific debate. AIRE is now known to partner with tens of nuclear proteins that affect transcription, chromatin structure, nuclear shuttling and RNA splicing. Furthermore, the structural features of the chromatin, including histone modifications, at AIRE target gene loci and the post-translational modifications of AIRE itself have been associated with AIRE activity. Yet, how all these interactions, structural and biochemical alterations jointly shape the AIRE-mediated gene activation remains to be explored. The modern high-throughput and single cell technologies have confirmed many earlier observations about the transcriptional program present in AIRE-expressing cells and, hopefully, will create new opportunities to solve the puzzle of promiscuous gene expression in the thymus.

The current thesis presents results that establish the role of post-translational acetylation of the AIRE protein in AIRE-dependent transcription, describes the pathological mutation in the second zinc finger domain of AIRE and characterises the AIRE-driven changes in chromatin structure at AIRE target genes and beyond.

2. REVIEW OF LITERATURE

2.1. The mechanisms of immune tolerance

The immune system comprises of two major compartments that are distinguished by their timing and specificity towards infectious agents. First, the cells and molecules of the innate immune system react rapidly to any invading pathogens by responding to molecular patterns and signals that are common to large groups of microorganisms. Second, a persistent immune reaction will eventually activate the components of the adaptive immune system, which detect and eliminate pathogens with great specificity.

The specificity is achieved by a reaction between the B- or T-cell receptor on the cell membrane of a B or T lymphocyte and an antigenic molecule, which the B-cells encounter in the extracellular space and the T-cells recognise as bound to a MHC complex on other cell types, including specialised antigen presenting cells (Kyewski and Klein, 2006; Pelanda and Torres, 2012). The two compartments are tightly connected through signalling molecules and cell-cell interactions, which greatly increase the efficiency of clearing an infection. However, harmless airborne and food compounds or molecules that derive from the organism itself have antigenic properties, but in that case, it is important to exclude them from being targeted by any immune defence response. The organism manages it by applying a multitude of molecular and cellular mechanisms that dampen the effector signals or eliminate the autoreactive lymphocytes altogether resulting in the so-called immune tolerance.

The processes that deprive lymphocytes of activating signals occur mainly in the secondary lymphoid organs and are commonly referred to as peripheral immune tolerance mechanisms. The majority of studies on peripheral tolerance conclude that the lymphocytes survive, but become anergic to their cognate antigen, although increased apoptosis of self-reactive mature T- and B-cells has also been described (Mueller, 2010; Pelanda and Torres, 2012).

The primary lymphoid organs, the thymus and the bone marrow, are responsible for the central immune tolerance, which mainly encompasses the removal or reprogramming of autoreactive lymphocytes during their development before they enter the blood circulation and the lymphatic system. According to estimates, 50–70% of differentiating B lymphocytes in the bone marrow are autoreactive, of which up to 50% undergo B-cell receptor (BCR) editing and the other half are restrained by immunological anergy or ignorance (Grandien et al., 1994; Wardemann et al., 2003).

In the thymus, the developing T lymphocytes have to pass through a more dramatic maturation process including a positive and a negative selection stage where approximately 90% of all thymocytes die by apoptosis (Klein et al., 2014). The positive selection takes place at the thymic cortex where immature cluster of differentiation 4 and 8 (CD4 and CD8, respectively) double-positive thymocytes need to have a fully functional TCR that can interact with a self-

peptide-MHC complex on the surface of the cortical thymic epithelial cells (cTECs), dendritic cells (DCs) or fibroblasts and thereby avoid dying by neglect (Klein et al., 2009). The surviving thymocytes migrate into the thymic medulla where they undergo negative selection where the cells that carry a TCR with high affinity towards self-peptide-MHC complexes on mTECs will receive a death signal, which leads to apoptosis (Starr et al., 2003). Finally, differentiated single-positive CD4 or CD8 thymocytes that have a functional TCR with low affinity for self-peptide-MHC complexes are released to the periphery as mature naive T-cells (Starr et al., 2003). Several studies suggest that in parallel with negative selection the thymic medulla promotes also the emergence of regulatory T-cells (Tregs) from the same pool of autoreactive thymocytes (Apostolou et al., 2002; Fontenot et al., 2005; Itoh et al., 1999; Jordan et al., 2001).

In any case, failure in tolerance may result in an autoimmune disorder where one or multiple tissues are targeted and gradually destroyed by the autoreactive immune cells (Antonia et al., 1995; Asano et al., 1996; Kishimoto and Sprent, 2001; Lesage et al., 2002). Common autoimmune disorders have a polygenic background and the severity of the disease is heavily dependent on a complex interplay between the risk alleles and environmental factors. Yet, few rare autoimmune syndromes follow an autosomal recessive mode of inheritance and are strongly linked to mutations in single key genes, such as in the autoimmune regulator (*AIRE)*, forkhead box P3 (*FOXP3*) transcription factor and the Fas cell surface death receptor (*FAS*) genes (Aaltonen et al., 1997; Chatila et al., 2000; Fisher et al., 1995; Nagamine et al., 1997). Of these, FOXP3 and FAS are necessary for the development and availability of T-cell populations, and AIRE is pivotal in the process of negative selection (Cheng and Anderson, 2012).

2.2. The autoimmune regulator and the autoimmune polyendocrinopathy syndrome 1

The human *AIRE* gene covers a 12.8 kb region on chromosome 21q22.3. Its 14 exons encode a protein of 545 amino acid residues and a molecular weight of 57.7 kDa (Aaltonen et al., 1997; Nagamine et al., 1997) (Figure 1).

Mutations in the *AIRE* gene can cause APS1 which mainly manifests in three pathological conditions: chronic mucocutaneous candidiasis (CMC), hypo-parathyroidism (HP) and Addison's disease (AD) (Neufeld et al., 1981). Usually, the patients suffer from additional clinical complications with large variations in the time of onset and severity (Husebye et al., 2009).

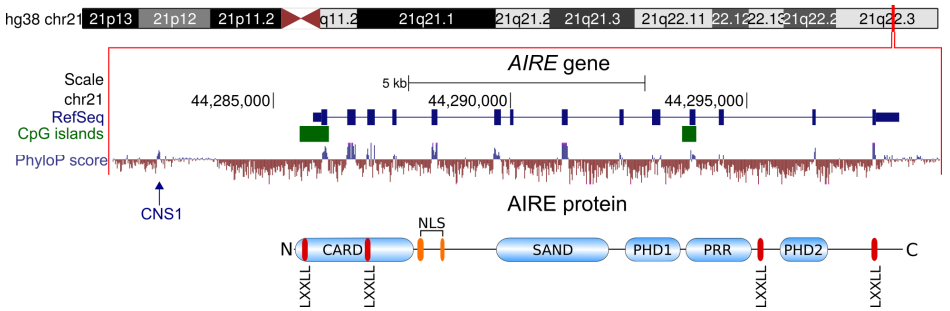

Figure 1. The human *AIRE* gene and protein. The vertical red bar on the chromosome 21 ideogram shows the location of the AIRE gene (human reference genome version 38). The AIRE gene structure is depicted together with the location of 5'-cytosine-phosphate-guanine-3' (CpG) islands and the PhyloP conservation score. The arrow shows the upstream conserved non-coding sequence 1 (CNS1). The AIRE protein contains the following domains and sequence motifs: caspase recruitment domain (CARD); the nuclear localisation signal (NLS); the SP100, AIRE1, NucP41/P75 and DEAF1 (SAND) domain; plant homeodomain 1 and 2 (PHD1 and PHD2); proline-rich region (PRR); the Leu-Xaa-Xaa-Leu-Leu amino acid sequence motifs (LXXLL).

In most cases, APS1 has a classical autosomal recessive mode of inheritance, where only biallelic mutations result in the disease (Ahonen, 1985). To date, more than 100 mutations in the AIRE coding sequence have been uncovered and the number increases due to ever increasing availability of modern sequencing and genotyping technologies (www.hgmd.cf.ac.uk). Some patients display an autosomal dominant pattern of inheritance, for example, the carriers of the SP100, AIRE1, NucP41/P75 and DEAF1 (SAND) domain mutation G228W or six missense mutations found in the plant homeodomain 1 (PHD1) zinc finger domain (Cetani et al., 2001; Oftedal et al., 2015). Based on exome array data, dominant-negative AIRE mutations are rather frequent within the general population suggesting that more organ-specific autoimmune diseases than previously thought of could be due to the expression of a dysfunctional AIRE protein (Oftedal et al., 2015).

2.2.1. The regulation of AIRE expression

The *AIRE* locus is silenced in most tissues and its expression can mainly be detected in mTECs, and to a far lesser extent in thymic DCs, thymic B-cells, extra-thymic AIRE-expressing cells (eTACs), testis, ovary and fetal tissues (Gardner et al., 2008; Halonen et al., 2001; Heino et al., 1999, 2000; Nishikawa et al., 2010; Schaller et al., 2008; Yamano et al., 2015). The *AIRE* promoter region harbours several conserved binding sequences for multiple transcription factors, including specificity protein 1 (Sp1), nuclear factor Y (NF-Y), activating protein 1 (AP-1), and E-twenty-six (Ets) family proteins (Murumägi

et al., 2003, 2006). In addition, the upstream sequence of AIRE contains a CpG island both in humans and mice and its methylation status has been linked to the control of AIRE expression in several cell lines and in thymic epithelial cells (Herzig et al., 2016; Kont et al., 2011; Murumägi et al., 2003) (Figure 1).

Two recent studies published the role of a highly conserved enhancer element ~3 kb upstream of the *AIRE* gene locus that is critically important in mediating the activation of the expression of AIRE by nuclear factor kappa B (NF-κB) signalling (Haljasorg et al., 2015; LaFlam et al., 2015) (Figure 1). Furthermore, mice lacking the enhancer sequence displayed a phenotype similar to AIRE knock-out mice (Haljasorg et al., 2015; LaFlam et al., 2015).

Another layer of complexity in the regulation of AIRE expression was discovered by Yanagihara et al. (2015) who found that the splicing of the AIRE transcript is impaired in mice deficient for the arginine demethylase and lysine hydroxylase called JmjC domain-containing protein 6 (Jmjd6), which is known to modify several splicing regulatory proteins. Although AIRE transcript levels were normal, the retained intron 2 severely hampered the translation of a functional AIRE protein (Yanagihara et al., 2015).

A large-scale study to uncover the complex network of transcriptional regulators that are necessary to facilitate AIRE expression in the thymic cell compartments identified four transcription factors – namely interferon regulatory factor 4 (Irf4), interferon regulatory factor 8 (Irf8), T-box protein 21 (Tbx21) and T-cell factor 7 (Tcf7) – that contribute to the efficient expression of AIRE (Herzig et al., 2016). In addition, this report described the role of the insulator protein CCCTC-binding factor (CTCF) in maintaining a silenced AIRE locus and showed that its eviction by the CCCTC-binding factor like (Ctcfl) protein allowed the recruitment of the aforementioned transcription factors (Herzig et al., 2016).

2.2.2. The structure and cellular localisation of the AIRE protein

The structural domains of AIRE highlight its role in the cell nucleus, particularly in chromatin-related processes (Figure 1). The N-terminal 105 amino acids belong to the caspase recruitment domain (CARD), which mediates the formation of homo-oligomers (Ferguson et al., 2008). The CARD domain is followed by a bipartite nuclear localisation signal (NLS) and putative NLS further upstream (Saltis et al., 2008). The SAND domain is known to bind DNA via the lysine-(aspartic acid/asparagine)-tryptophan-lysine (K[DN]WK) sequence motif and depends on the underlying CpG methylation status (Bottomley et al., 2001; Isaac et al., 2006; Jensik et al., 2014). Two reports have demonstrated that AIRE could bind DNA in sequence-specific manner and the protein-DNA interaction was mapped to the SAND and PHD domains (Kumar et al., 2001; Purohit et al., 2005). However, the canonical motif is missing in AIRE SAND domain and a more recent study showed that the AIRE CARD domain rather than the SAND domain is required to interact with nucleosome-free DNA, which in this case

lacked any sequence-specificity (Maslovskaja et al., 2015). The AIRE protein contains two PHD-type zinc fingers that serve as modules for protein-protein interactions and are essential for the activation of AIRE target genes (Koh et al., 2008; Org et al., 2008; Yang et al., 2013). Notably, the electrostatic surface of the AIRE PHD1 has a negative overall charge that facilitates its binding to histone H3 that is unmethylated at lysine 4 (H3K4me0), which is a characteristic feature of silenced genomic regions (Koh et al., 2008; Org et al., 2008). The PHD2 domain structure resembles that of PHD1, but it has a positively charged surface, which implies that it mediates a different set of protein-protein interactions (Gaetani et al., 2012; Yang et al., 2013). The AIRE protein also has a uncharacterised proline-rich region between the PHD zinc fingers and four LXXLL motifs, which are part of many regulatory proteins in the nucleus and are required for protein-protein interactions, for example, with nuclear receptors (Plevin et al., 2005).

The AIRE protein localises into the cell nucleus and forms dot-like structures called AIRE nuclear bodies, which resemble, but do not overlap with promyelocytic leukaemia protein (PML)-containing nuclear bodies (Akiyoshi et al., 2004; Heino et al., 1999). Mutations in the AIRE CARD domain disrupt the punctate pattern in the nucleus, although other protein domains have also been reported to influence the subcellular localisation of AIRE (Björses et al., 1999; Ferguson et al., 2008; Halonen et al., 2004; Ramsey et al., 2002; Rinderle et al., 1999). AIRE has been shown to interact and colocalise with the acetyltransferase cAMP response element-binding protein (CREB)-binding protein (CBP), which is also known to interact with PML (Doucas et al., 1999; Pitkänen et al., 2000, 2005). CBP and its paralog adenovirus early region 1A-associated 300 kDa protein (p300) are able to post-translationally acetylate AIRE, thereby modulating the transcriptional activity of AIRE (Chuprin et al., 2015; Incani et al., 2014). Intriguingly, AIRE also co-localises with the deacetylase protein sirtuin-1 (Sirt1), which was reported to directly counteract the effort of CBP/p300 to modify AIRE (Chuprin et al., 2015). Similarly to PML nuclear bodies, the core of the AIRE nuclear bodies is devoid of chromatin and RNA polymerase II (RNAP II) suggesting that these structures do not participate in transcription (Pitkänen et al., 2005). Nevertheless, the AIRE nuclear bodies are bound to the nuclear matrix and might affect transcriptional processes indirectly through shaping the chromatin landscape (Abramson et al., 2010; Akiyoshi et al., 2004; Tao et al., 2006).

2.2.3. Transcriptional regulation by AIRE

The AIRE protein is a transcriptional regulator that can potentially activate more than 3000 different genes in a phenomenon called promiscuous gene expression, which allows the AIRE-expressing cells to represent almost the entire protein coding portion of every other tissue in the body (Anderson et al., 2002; Derbinski et al., 2001; Sansom et al., 2014). The molecular mechanisms

behind this extraordinary transcriptional program have been difficult to dissect due to technical challenges that are related to the scarcity of primary cells expressing AIRE. For example, the number of AIRE-positive mTECs in a mouse thymus is less than 50,000 (Anderson and Su, 2016). However, modern single-cell based molecular genomics approaches have started to resolve this issue and give more detailed insights into the gene regulatory mechanisms governing AIRE-dependent gene expression. Single-cell RNA-seq studies have shown that mTECs express the TRA genes with ordered stochasticity meaning that individual mTECs express a somewhat random set of TRA genes, which however tend to cluster on the chromosomes or co-express interchromosomally, but often lack any other obvious common features, such as belonging to the same signalling or metabolic pathway (Brennecke et al., 2015; Meredith et al., 2015). On a population level, this permits mTECs to present a vast array of self-peptides to the developing thymocytes and give the crucial signals to elicit negative selection (Brennecke et al., 2015; Meredith et al., 2015).

Two proteomic screens and many small-scale pull-down experiments with AIRE have identified over 40 interaction partners that could be broadly divided into 4 functional groups – proteins related to transcription, chromatin structure/ remodelling, mRNA splicing and nuclear transport (Abramson et al., 2010; Gaetani et al., 2012). As mentioned above, AIRE partners with the unmethylated histone H3, although this reaction itself has low affinity suggesting that it is a short-term dynamic interaction (Koh et al., 2008; Org et al., 2008). AIRE-responsive genomic regions tend to be also enriched for repressive histone marks, such as histone H3 trimethylated at lysine 27 (H3K27me3), and low levels of acetylated histones or histone H3 trimethylated at lysine 4 (H3K4me3) (Org et al., 2009; Sansom et al., 2014). However, direct binding of AIRE to those modified histones has not been demonstrated. Furthermore, chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-seq) analyses have shown that AIRE can be recruited to almost every gene promoter irrespective of its transcriptional status, but only at loci in a silenced chromatin environment does AIRE display its transcriptional potential (Giraud et al., 2012).

Increasing evidence supports the hypothesis that AIRE facilitates the release of the poised RNAP II at silenced sites by binding to the positive transcription elongation factor b (P-TEFb) and allowing it to phosphorylate the pause factors and RNAP II, which then enters the elongation phase (Giraud et al., 2014; Oven et al., 2007). A more recent study reported that the interaction between AIRE and P-TEFb is, in fact, mediated by the bromodomain-containing protein 4 (BRD4), which is known to recruit P-TEFb through its C-terminal domain (Yoshida et al., 2015). Interestingly, BRD4 bound AIRE via its bromodomains that interacted with lysine residues in the CARD domain that had been acetylated by CBP (Yoshida et al., 2015).

A notably large group of AIRE interaction partners participate in the DNA damage response, such as the DNA topoisomerase 2-alpha (TOP2A), DNA-dependent protein kinase (DNA-PK) and its non-catalytic 70 and 80 kDa

subunits of Ku antigen (Ku70/Ku80), and poly [ADP-ribose] polymerase 1 (PARP1) (Abramson et al., 2010; Gaetani et al., 2012; Liiv et al., 2008; Yang et al., 2013). TOP2A generates transient DNA breaks to resolve torsional tension in supercoiled DNA, which is caused by an elongating DNA or RNA polymerase (Nitiss, 2009). Double-stranded DNA breaks activate the DNA-PK/Ku70/Ku80 complex, which then phosphorylates many nuclear proteins that will initiate the DNA repair process (Callén et al., 2009). DNA-PK can also phosphorylate AIRE at several serine and threonine residues, although the role of phosphorylation in AIRE-dependent gene activation is still under debate (Liiv et al., 2008; Zumer et al., 2012). Nevertheless, mutation analysis of two amino acid positions in the CARD domain – threonine 69 and serine 157 – that are potentially modified by DNA-PK suggested that phosphorylation could be a prerequisite for CBP-dependent acetylation, which is necessary for the interaction with the aforementioned BRD4 (Yoshida et al., 2015).

Additional investigations on AIRE acetylation have revealed a complex interplay between different modified lysine residues and the functionality of AIRE. Acetylation of the lysine residues within the NLS sequence decreases AIRE target gene activation by hindering the localisation of AIRE into the nucleus (Incani et al., 2014). On the other hand, the deacetylase Sirt1 promotes the transcription of AIRE-dependent genes by specifically removing acetyl groups from lysines between the NLS and SAND domain (Chuprin et al., 2015).

Although the mechanistic links and the chain of molecular events regarding AIRE-dependent gene activation have become clearer, the current knowledge cannot explain why this occurs only at certain and not all genes that are targeted by AIRE.

2.3. Linking chromatin structure and transcription

The basic chromatin structure is highly conserved among the eukaryotes from yeast to mammals, although the complexity increases towards the more recent groups of organisms. The low-level unit of the chromatin is the nucleosome core, which consists of a 147 bp DNA sequence wrapped around a histone octamer. The histone octamer is comprised of four pairs of histones H2A, H2B, H3 and H4. Nucleosome cores are connected by a variable-length linker DNA sequence which on average adds up to 200 bp of nucleosomal DNA (Davey et al., 2002; Luger et al., 1997). Nucleosomes are 11 nm in diameter and are thought to be further packaged into 30 nm fibres with the help of linker histone H1, which is structurally distinct from the other four histones (Vignali and Workman, 1998). Gene expression is thought to be tightly linked to the level of chromatin condensation. In general, chromatin is classified into compact and transcriptionally silent heterochromatin and into more accessible and transcriptionally active euchromatin. Modern genomics approaches have started to unravel the finer organisation of the genome allowing to distinguish not only actively transcribed regions from heterochromatin, but also detect the spatio-temporal

regulation of chromatin looping, topologically associated domains, chromosome territories, the nuclear lamina and transcription factories (Bolzer et al., 2005; Dixon et al., 2012; Kadauke and Blobel, 2009; Verschure et al., 2002).

Chromatin loops form between genomic regulatory elements, such as enhancers and promoter, which can be separated by kilo- or megabases of DNA of the same chromosome. The widespread adoption of chromosome conformation capture (3C)-based methods in functional genomics has significantly accelerated the discovery of gene regulatory events. High-resolution maps of chromatin loops reveal the role of promoter-enhancer interactions in transcriptional control (Jin et al., 2013; Lieberman-Aiden et al., 2009; Rao et al., 2014; Sanyal et al., 2012). Evidence from recent reports suggests that enhancer competition determines the rate of transcription of different genes that rely on the availability of the same enhancer (Rao et al., 2014; Sanyal et al., 2012). Often, the position of the enhancer relative to the genes does not matter, and interestingly, as little as 7% of enhancers interact with the closest gene (Rao et al., 2014; Sanyal et al., 2012). Chromatin looping may also depend on the differentiation stage of the cells. For example, transcription factors required for the activation of the α-globin gene are bound to the enhancer located 40 kb away of the locus already at an early proerythroblast stage, although only at the late erythroblast stage does the enhancer interact with the promoter and permit the transcription of α-globin (Vernimmen et al., 2007).

Topologically associated domains (TADs) are large structures consisting of several chromatin loops and usually defined by the relatively frequent intra-TAD interactions as compared to inter-TAD interactions (Dixon et al., 2012; Nora et al., 2012). It has been proposed that TADs could divide the chromatin into regions that have similar activity. Many genes that are co-regulated in development and clustered on chromosomes belong to the same TAD (Nora et al., 2012). Also, genes responding differentially to external stimuli have been shown to group together in different TADs (Le Dily et al., 2014).

Chromatin looping as well as the formation of TADs require architectural proteins, such as the CTCF and cohesin complex, which bind chromatin insulator sequences, thereby blocking promoter-enhancer interactions or heterochromatin spreading (Gaszner and Felsenfeld, 2006). DNA-bound CTCF can dimerise with another distally positioned CTCF, which effectively creates chromatin loops (Guo et al., 2012). Furthermore, ChIP-seq studies have revealed that the CTCF-dependent loops display a cell type-specific pattern and the disruption of the CTCF or cohesin binding sites can significantly alter the expression of nearby genes (Nora et al., 2012; Zuin et al., 2014; Tark-Dame et al., 2014). The depletion of architectural proteins enriched at TAD borders has been found to increase the frequency of inter-TAD interactions (Li et al., 2015).

The non-random nature of the genome organisation extends to whole chromosomes not only at the critical stages of cell division, but also at other cell cycle phases. The phenomenon of chromosomes taking up specific regions in the nucleus with respect to one another and to the nuclear lamina has been termed chromosome territories (CTs) (Cremer and Cremer, 2010). Furthermore, the

position of genes within the CTs tends to be predetermined. Namely, actively transcribed loci are more likely to be found looping out at the periphery of CTs and silenced genomic regions are more frequently located in the interior of CTs (Federico et al., 2008; Kurz et al., 1996; Mahy et al., 2002; Scheuermann et al., 2004). However, additional investigations are necessary to clarify whether gene repositioning is the cause or consequence of altered expression. For example, in human embryonic stem cells, the actively transcribed homeobox (*Hox*) gene loops out of its CT together with genes whose expression remains constant (Morey et al., 2009). Other findings support the hypothesis that loops outside of the CTs represent regions of poised chromatin ready for activation. According to knock-down experiments, the extrusion of the human β-globin locus from its CT prior to activation depends on the activities of the transcription factor GATA nucleotide sequence binding protein 1 (GATA1) and its co-regulator friend of GATA1 (FOG1) protein (Ragoczy et al., 2003). However, RNA interference (RNAi)-mediated depletion of FOG1 after the β-globin gene had looped out of the CT did not affect β-globin expression (Lee et al., 2011).

The radial positioning of CTs and genes therein adds another layer of complexity and gene regulatory features that should be taken into account when studying the relationship between transcription and genome organisation. Notably, it has been observed that chromosomal regions close to the nuclear lamina become transcriptionally silenced, whereas genes located more centrally within the nucleus are actively expressed (Croft et al., 1999; Scheuermann et al., 2004; Tanabe et al., 2002). Furthermore, the latter active transcription sites tend to be highly enriched for proteins belonging to the transcriptional machinery, including phosphorylated RNAP II, transcription and splicing factors, which has been collectively referred to as transcription factories (Branco and Pombo, 2006; Bridger et al., 2005; Cisse et al., 2013; Verschure et al., 2002). Transcription factories tend to occupy the nuclear space between CTs called the interchromatin domain (ICD), which is believed to foster the optimal conditions for efficient transcription by facilitating the colocalisation of genes and regulatory sequences from different CTs, while minimising the likelihood of ectopic contacts (Iborra et al., 1996). Further analyses have found that constitutively active transcription factories are complemented with poised transcription factories that become highly activated in a stimulus-response manner (Ferrai et al., 2010).

Although it is unknown how the compartmentalisation is controlled and what mechanisms rearrange the nuclear organisation in response to external signals, it is clear that the compartmentalisation of the nucleus provides an efficient and quick mode to regulate gene activation and silencing.

2.3.1. The role of topoisomerases in shaping chromatin structure and transcription

A moving polymerase generates positive DNA supercoils ahead of the polymerase complex and negative supercoils behind it (Liu and Wang, 1987). The overwound DNA in front of the polymerase prevents DNA strand separation, which can inhibit the nucleic acid synthesis (Liu and Wang, 1987). The supercoils are detected and bound by DNA topoisomerases that can cleave one or both DNA strands, unwind the DNA thereby relieving torsional stress in the molecule and re-ligate the strands back together (Champoux, 2001). There are two main types of DNA topoisomerases that are classified based on their catalytic activity. Type I topoisomerases (TOP1 in humans) cut a single strand and rotate it around the intact strand releasing both positive and negative supercoils after which the DNA break is sealed (Pommier et al., 1998). Type II topoisomerases (TOP2A and TOP2B in humans) create a double-strand DNA break, pass the intact part of the double helix through the break to remove mainly positive supercoils and close the break (Gale and Osheroff, 1992).

In addition to binding to distorted chromatin structures and altering their topology, the topoisomerase activity, similarly to other DNA damage-causing conditions, changes the histone modification pattern. More specifically, the chromatin at sites of double-stranded DNA breaks, which can be induced by topoisomerases or other DNA-damaging processes, becomes enriched for the histone H2A variant H2AX, which is phosphorylated by either ataxia telangiectasia mutated (ATM), ataxia telangiectasia and Rad3 related (ATR) or DNAP-PK at serine 139 (γH2AX) (Durocher and Jackson, 2001). The region marked by γH2AX can spread around the DNA break for several hundred kilobases (Savic et al., 2009).

In addition to being vitally important in DNA replication and recombination, both types of topoisomerases have also been implicated in RNA transcription (Puc et al., 2017). Transcription has been shown to be attenuated after RNAi-mediated depletion of any of the topoisomerases suggesting that topoisomerases facilitate RNAP II processivity (Brill et al., 1987; Gartenberg and Wang, 1992; Goto and Wang, 1985; Kim and Wang, 1989; Schultz et al., 1992). Indeed, studies with yeast and human cells have shown that TOP1 is enriched at highly expressed genes where it physically binds to the phosphorylated C-terminal domain of the RNAP II catalytic subunit (Baranello et al., 2016; Phatnani et al., 2004; Takahashi et al., 2011). In addition, TOP1 is recruited to transcriptionally active sites through interactions with chromatin remodelling factors (Husain et al., 2016). Importantly, the TOP1-RNAP II interaction stimulates the TOP1 DNA relaxation activity that facilitates promoter escape and elongation past natural pause sites (Baranello et al., 2016).

Intriguingly, TOP1 and TOP2 are critical for the transcription of extremely long genes in human neurons. Investigations have determined that the expression of long genes linked to autism spectrum disorder can be attenuated by TOP1 inhibitors or the knock-down of *TOP1* and *TOP2B* genes (King et al.,

2013). Furthermore, the blocking of TOP1 and TOP2 activity by small-molecule inhibitors reduces the expression of an anti-sense transcript that silences the paternal ubiquitin-protein ligase E3A (Ube3a) allele in Angelman syndrome (Huang et al., 2011).

2.4. Post-translational modifications in transcriptional regulation

Transcription relies on a coordinated series of protein-protein and protein-DNA interactions. All steps in this process are accompanied by the catalysis of chemical moieties onto chromatin proteins, such as histones, or components of the general transcription machinery, including the RNAP II, and transcription factors bound to distal regulatory elements (Chen et al., 2011; Hendriks and Vertegaal, 2016; Phatnani and Greenleaf, 2006; Spange et al., 2009; Suganuma and Workman, 2011).

The best studied post-translational modifications (PTMs) have so far been histone acetylation, methylation and phosphorylation, which have been shown to correlate well with the transcriptional state of a genomic locus (Suganuma and Workman, 2011). For example, the molecular events leading to the activation of the proto-oncogen c-Fos (FOS) like antigen 1 (*FOSL1*) gene in human embryonic kidney cell line 293 (HEK293), involve the phosphorylation of serine 10 on histone H3 (H3S10ph) at the *FOSL1* enhancer allowing the binding of 14-3-3 protein, which is required by the lysine acetyltransferase 8 (KAT8) to acetylate histone H4 lysine 16 (H4K16ac) (Zippo et al., 2009). H3S10ph and H4K16ac enable the binding of BRD4 through its bromodomain, which brings the transcription elongation factor P-TEFb to the promoter (Zippo et al., 2009). In the transcription initiating phase, the RNAP II is already phosphorylated by the cyclin-dependent kinase 7 (CDK7) subunit of the transcription factor II H (TFIIH) at the 5^{th} position of the heptameric repeat on its C-terminal domain (CTD) (Sansó and Fisher, 2013). Only after the cyclin-dependent kinase 9 (CDK9) subunit of the P-TEFb has phosphorylated the 2^{nd} position of the heptad repeats on the CTD does the RNAP II enter the transcription elongation phase (Sansó and Fisher, 2013). Actively transcribed sites acquire additional histone modifications, such as the H3K4me3 in the promoter flanking the transcription start site and histone H3 lysine 36 tri-methylation (H3K36me3) along the gene body (Barski et al., 2007). Although studies in yeast have shown that neither H3K4me3 nor H3K36me3 are required for the transcription, the presence of those modifications does increase the efficiency of the RNA polymerase (Mason and Struhl, 2005; Zhang et al., 2005).

Similarly, gene silencing is accompanied by specific histone modifications that serve as docking sites for regulators that repress transcription. The tri-methylation of histone H3 lysine 9, 20 and 27 (H3K9me3, H4K20me3 and H3K27me3, respectively) are all known to correlate with silenced chromatin

regions (Suganuma and Workman, 2011). Typically, these regions also lack histone acetylation, H3K4me3 and H3K36me3, although all these modifications have been found in various combinations in reciprocal chromatin states (Barski et al., 2007).

In general, the constitutive heterochromatin found at pericentromeric regions, is enriched for H3K9me3, which is bound by the heterochromatin protein 1 (HP1) (Lachner et al., 2001). This interaction is proposed to facilitate chromatin folding and the packaging into higher-order structures (Fan et al., 2004; Maison et al., 2002; Peters et al., 2001; Thiru et al., 2004). Furthermore, the interaction between HP1 and the lysine methyltransferase 1A (KMT1A) is thought to facilitate the spreading of H3K9me3 (Lachner et al., 2001).

Facultative heterochromatin, which is found at developmental and imprinted gene loci, is characterised by the presence of H3K27me3 (Trojer and Reinberg, 2007). This histone mark is established by the polycomb repressive complex 2 (PRC2) and bound by polycomb repressive complex 1 (PRC1) that can actively block ATP-dependent chromatin remodelling and RNAP II activity (Cao et al., 2002; King et al., 2002; Kuzmichev et al., 2002; Levine et al., 2002). The PRC1 also contains E3 ubiquitin ligase activity specifically towards histone H2A lysine 119 whose monoubiquitylation has been shown to correlate with the binding of linker histone H1, which is considered to contribute to the maintenance of a repressive chromatin state (de Napoles et al., 2004; Zhu et al., 2007; Wang et al., 2004).

The proteins that catalyse, bind to and remove the PTMs are often referred to as PTM writers, readers and erasers, respectively, although in many cases these functions are intermingled in the same protein complex (Patel and Wang, 2013). Recent structural analyses have uncovered large families of protein domains that have specialised in interacting with certain PTMs. For example, the chromodomain, Tudor and proline-tryptophan-tryptophan-proline (PWWP) and PHD zinc finger domains bind to methylated lysine or arginine residues by a surface groove pocket recognition mode (Chen et al., 2011; Ruthenburg et al., 2007). The role of these structures in regulating transcription is evidently dependent on the surrounding chromatin context and the interaction partners, because both gene activating and silencing complexes harbour methyl-lysine or -arginine binding functions (Patel and Wang, 2013). Bromodomains, which are part of many acetyltransferases, methyltransferases, chromatin remodelling and co-activator complexes, have high affinity towards acetylated lysines (Dhalluin et al., 1999). Phosphorylated amino acids are bound by the 14-3-3, tandem breast cancer susceptibility (BRCT) and baculovirus inhibitor of apoptosis (IAP) repeat (BIR) domain-containing proteins (Kelly et al., 2010; Mackintosh, 2004; Singh et al., 2012). An additional layer of complexity in the signalling cascades involving PTMs comes from the observations that the PTM-binding domains are found in tandem or in combination, for example the tandem PHD fingers in the double PHD fingers 3 (DPF3) protein or the PHD-bromodomain cassette in the bromodomain PHD finger transcription factor (BPTF) (Lange et al., 2008; Li et al., 2006). Notably, some of these structures are specialised to interact with

unmodified amino acid residues, including the AIRE PHD1 and the PHD finger protein 21A (PHF21A) PHD domain (Koh et al., 2008; Lan et al., 2007; Org et al., 2008).

In summary, numerous studies highlight the involvement of PTMs in the regulation of protein function and suggest a complex crosstalk between PTMs and their readers to maintain the high specificity of gene expression regulation.

2.4.1. Transcription regulation by the acetylation of non-histone proteins

The lysine acetylation has been implicated to modulate the DNA and protein binding, subcellular localisation, protein stability and catalytic activity of a variety of non-histone proteins in both nuclear and cytoplasmic compartments (Glozak et al., 2005).

The earliest reported acetylated non-histone protein was the tumour suppressor p53, which is modified by the acetyltransferase and transcription co-activator protein p300, which leads to an increase in sequence-specific DNA binding by p53 (Gu and Roeder, 1997). Follow-up studies have elaborated on the role of p53 acetylation in transcriptional control of its target genes and revealed an intricate interplay with other post-translational modifications. Namely, acetylation competes with ubiquitylation for the same lysine residues in the p53 protein sequence. Lysine polyubiquitylation marks proteins for proteasomal degradation and, therefore, the mutually exclusive relationship between the two modifications determines the stability of p53 (Ito et al., 2002). Furthermore, the acetylation of lysine 373 by p300 and the subsequent transcription of the cyclin-dependent kinase inhibitor p21 during a DNA damage response are stimulated by the methylation of lysine 372 (Ivanov et al., 2007).

Acetylation can also decrease the activity of a transcription factor. Studies have shown that the acetylation of certain lysine residues of the forkhead box O1 (FOXO1) transcription factor by CBP decreases FOXO1 affinity to DNA and stimulates its subsequent phosphorylation (Matsuzaki et al., 2005). The phosphorylated FOXO1 protein is bound by 14-3-3 proteins and exported out of the nucleus into the cytoplasm where it is polyubiquitylated and degraded by the proteasome (Brunet et al., 1999). Additional studies have shown that the activities of the deacetylase sirtuin 2 (SIRT2) can maintain FOXO1-mediated transcription, which strengthen the notion that acetylation can control FOXO1 activity (Jing et al., 2007).

In addition to stimulating protein phosphorylation of the components of the Janus kinase and signal transducer and activator of transcription (JAK-STAT) pathway (Rawlings et al., 2004), interferon-signalling has also been shown to promote CBP-mediated acetylation of the signal transducer and activator of transcription 2 (STAT2) protein in the cytoplasm, which enables it to dimerise with signal transducer and activator of transcription 1 (STAT1) and form the

interferon-stimulated gene factor 3 (ISGF3) complex (Tang et al., 2007). The acetylated ISGF3 can translocate to the nucleus and activate the transcription of interferon-responsive genes (Tang et al., 2007). Furthermore, the NLS sequences often contain lysine residues that are modified by acetyltransferases. The S-phase kinase-associated protein 2 (SKP2) is acetylated by p300 at its NLS, which promotes the nuclear export of SKP2 (Inuzuka et al., 2012). SKP2 is an E3 ubiquitin ligase that ubiquitylates E-cadherin, which results in the degradation of E-cadherin and thereby contributes to the cell migration and survival in the context of tumorigenesis (Inuzuka et al., 2012). Acetylation of the cytoplasmic enzyme glyceraldehyde-3-phosphate dehydrogenase (GAPDH) by p300/CBP-associated factor (PCAF) allows it to translocate into the nucleus where it participates in transcriptional regulation and DNA repair (Ventura et al., 2010).

Several transcription coactivators, including CBP/p300, KAT8 and the yeast regulator of Ty1 transposition 109 (Rtt109) are able to autoacetylate themselves (Stavropoulos et al., 2008; Thompson et al., 2004; Yuan et al., 2012). The autoacetylation within the activation loop motif of CBP/p300 is triggered by protein dimerisation, which in turn activates the enzyme (Karanam et al., 2006). This sequence of events is further confirmed by findings showing that the deacetylase SIRT2 is able to attenuate p300 activity (Black et al., 2008). The transcription co-activators TATA nucleotide sequence binding protein-associated factor 250 kDa (TAF250), PCAF, and p300 can also acetylate the general transcription factors IIE amd IIF (TFIIE and TFIIF, respecitvely), although the functional consequences of these modifications are unclear (Imhof et al., 1997).

Taken together, post-translational acetylation of transcription regulators plays a multifaceted role in shaping the transcriptional output of the cell. The altered biochemical properties of the acetylated proteins can render them inactive or strengthen their transcriptional potential. A growing amount of evidence highlights the regulatory interplay with other modifications and between acetylation events of different lysine residues on the same protein.

2.5. Summary of the literature

The characterisation of the AIRE protein has greatly helped to gain a deeper insight into the mechanisms of central immune tolerance. Although the key events have been described in detail (Abramson and Goldfarb, 2016; Anderson and Su, 2016), many of the underlying molecular processes that lead to the activation of thousands of TRA genes are still elusive and require more attention.

Proteomic screenings have determined dozens of nuclear proteins that directly bind AIRE (Abramson et al., 2010 and Study III). Many AIRE interaction partners possess enzymatic activities which can activate or inhibit transcriptional regulators or modify chromatin proteins. However, the exact role of most of the partner proteins and their enzymatic activities in shaping the function of AIRE are poorly understood.

Currently, more than 100 mutations have been found along the *AIRE* gene coding sequence that can be causally linked to APS1 (www.hgmd.cf.ac.uk). A large fraction of the detected mutations fall into the functional domains CARD, PHD1 and PHD2. Yet, the molecular basis of the disease-causing mutations and their relationship to the severity of the disease are often unclear and need to be addressed more thoroughly.

Chromatin at actively transcribed regions is considered to be loosely packaged by histones making it easily accessible for transcription-related proteins, including the RNA polymerase II (Campos and Reinberg, 2009). However, AIRE tends to regulate the expression of genes that are located at repressed chromatin regions (Anderson et al., 2002; Derbinski et al., 2001; Sansom et al., 2014). Despite considerable effort from many research groups, the process of how AIRE determines its target gene repertoire has remained enigmatic and is of high scientific interest to date.

The current thesis touches some of the aspects of AIRE-dependent transcription that involve the analysis of post-translational modifications, the effects of missense mutations and changes in the chromatin structure.

3. AIMS OF THE STUDY

The aims of the present study were:

1. To functionally characterise the impact of post-translational acetylation of the AIRE protein by the transcription coactivator and acetyltransferase p300.
2. To investigate the effect of the previously identified APS1-associated mutation C446G in the AIRE PHD2 domain on protein-protein interactions and AIRE-dependent transcriptional regulation.
3. To determine the alterations in chromatin accessibility and CTCF-mediated chromatin conformation that are linked to the expression of AIRE.

4. MATERIALS AND METHODS

4.1. Cloning of reporter and expression plasmids

Plasmids that were used in the transfections are listed in Table 1.

Table 1. The list of expression plasmids that were used in the study.

Plasmid name	Reference
pcAIRE	(Heino et al., 1999)
pcAIRE.K243Q	Described in Study I
pcAIRE.K245Q	Described in Study I
pcAIRE.K253Q	Described in Study I
pcAIRE.K243/245Q	Described in Study I
pcAIRE.K243/253Q	Described in Study I
pcAIRE.K243R	Described in Study I
pcAIRE.K245R	Described in Study I
pcAIRE.K253R	Described in Study I
pcAIRE.K243/245R	Described in Study I
pcAIRE.K243/253R	Described in Study I
pcDNA3-AIRE	Described in Study II
pCDNA3-AIRED297A	Described in Study II
pCDNA3-AIREV301M	Described in Study II
pCDNA3-AIREC446G	Described in Study II
GST-AIRE-C446G	Described in Study II
pBL-INV	(Liiv et al., 2008)
pCMVβ-p300.wt-Myc	Gift from T-P. Yao, Department of Pharmacology and Cancer Biology, Duke University
pCMVβ-p300.DY-Myc	Gift from T-P. Yao, Department of Pharmacology and Cancer Biology, Duke University
pRc/RSV-mCBP-HA-RK	Gift from R. Goodman, Oregon Health and Science Institute
pcDNA3-FLAG/mouse PCAF(WT)	Gift from T. Nakajima, Institute of Medical Science, St. Marianna University School of Medicine
pcDNA3-FLAG/mouse PCAF(ΔHAT)	Gift from T. Nakajima, Institute of Medical Science, St. Marianna University School of Medicine
pd2EYFP-N1	Clontech, Mountain View, CA, USA
pcDNA3.1B(-)Myc/His	Invitrogen, Carlsbad, CA, USA
pGEX-1ZT-SH3	Gift from Dr. K. Saksela, University of Tampere

The acetylation site mutations were created by PCR-based site-directed mutagenesis by using the pcAIRE expression plasmid as template (Heino et al., 1999). The PCR-amplified fragments were cloned into the BamHI/HindIII sites of pcAIRE, which resulted in the following expression plasmids: pcAIRE.K243Q, pcAIRE.K245Q, pcAIRE.K253Q, pcAIRE.K243/245Q, pcAIRE.K243/253Q, pcAIRE.K243R, pcAIRE.K245R, pcAIRE.K253R, pcAIRE.K243/245R and pcAIRE.K243/253R. The primers for the mutagenesis and cloning are listed in Table 2.

Table 2. The list of primers for the site-directed mutagenesis and cloning used in the study.

Mutation	Primer name	Sequence	Template
	F1 (508-529 AIRE cDNA)	GCAGAGCAGCAGCGCCT TCCAC	pcAIRE
	R1 (pcDNA3.1 BGH)	GGCAACTAGAAGGCACA GTCGAGGC	pcAIRE
AIRE.K243Q	K243Q F	CCGGCAGTGGGCAGAAC AAGGC	pcAIRE
	K243Q R	GGCCTTGTTCTGCCCACT GCCG	pcAIRE
AIRE.K245Q	K245Q F	GTGGGAAGAACCAGGCC CGCAG	pcAIRE
	K245Q R	GCTGCGGGCCTGGTTCT TCCCA	pcAIRE
AIRE.K253Q	K253Q F	CAGTGGCCCGCAGCCTC TGGTT	pcAIRE
	K253Q R	ACCAGAGGCTGCGGGCC ACTGC	pcAIRE
AIRE.K243/245Q	K243Q/K245Q F	TGGGCAGAACCAGGCCC GCAGC	pcAIRE.K243Q
	K243Q/K245Q R	CTGCGGGCCTGGTTCTG CCCAC	pcAIRE.K243Q
AIRE.K243/253Q	K253Q F	CAGTGGCCCGCAGCCTC TGGTT	pcAIRE.K243Q
	K253Q R	ACCAGAGGCTGCGGGCC ACTGC	pcAIRE.K243Q
AIRE.K243R	K243R F	CGGCAGTGGGAGGAACA AGGCC	pcAIRE
	K243R R	GGGCCTTGTTCCTCCCAC TGCC	pcAIRE
AIRE.K245R	K245R F	TGGGAAGAACAGGGCCC GCAGC	pcAIRE
	K245R R	TGCTGCGGGCCCTGTTCT TCCC	pcAIRE

Table 2. Continuation

Mutation	Primer name	Sequence	Template
AIRE.K253R	K253R F	AGTGGCCCGAGGCCTCT GGTTC	pcAIRE
	K253R R	AACCAGAGGCCTCGGGC CACTG	pcAIRE
AIRE.K243/245R	K243R/K245R F	TGGGAGGAACAGGGCCC GCAGC	pcAIRE.K243R
	K243R/K245R R	TGCTGCGGGCCCTGTTC CTCCC	pcAIRE.K243R
AIRE.K243/253R	K253R F	AGTGGCCCGAGGCCTCT GGTTC	pcAIRE.K243R
	K253R R	AACCAGAGGCCTCGGGC CACTG	pcAIRE.K243R

4.2. Cell culture and transfection

All cell culture experiments were performed with transiently transfected or established stably expressing human embryonic kidney epithelial cell line HEK293. The latter include HEK293 cells with constitutive AIRE, AIRE.K243/253Q or yellow fluorescent protein expression (HEK-AIRE, HEK-AIRE.K243/253Q and HEK-YFP, respectively) in Study I and HEK293 cells with doxycycline-inducible AIRE expression (AIRE-Tet) in Study III. Except for AIRE-Tet, the cell lines were grown in Dulbeccos's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% antibiotic-antimycotic solution (all components were purchased from PAA Laboratories). The AIRE-Tet cell line was grown in selective DMEM supplemented with 10% tetracycline-negative FBS, 1% antibiotic-antimycotic solution (both were purchased from PAA Laboratories) and 0.15 mg/ml geneticin G418 (Sigma-Aldrich). The cells were kept at 37 °C with 5% CO_2. The cells were transfected at 60–70% confluency with ExGen 500 *in vitro* Transfection Reagent according to manufacturer's protocol (Fermentas). Depending on the following assay, the transfected cells were harvested at 24 or 48 h time point. Each experiment was performed in duplicate and repeated at least three times. Cells transfected with an empty vector were used as negative control samples.

4.3. Antibodies

Antibodies that were used in the western blot (WB), immunoprecipitation (IP), chromatin immunoprecipitation (ChIP) and immunofluorescence (IF) experiments are listed in Table 3.

Table 3. The list of primary and secondary antibodies that were used in the study.

Antibody	Description	Methods	Product reference
Anti-AIRE 6.1	Primary mouse monoclonal to human AIRE protein	WB, IP	Purified in-house
Anti-AIRE	Primary goat polyclonal to human AIRE protein	WB, IF	EB05507, Everest Biotech Ltd
Anti-acetyl lysine	Primary rabbit polyclonal to acetylated lysines	WB, IP	06-933, Upstate/EMD Millipore
Anti-GAPDH	Primary mouse monoclonal [6C5] to human GAPDH protein	WB	ab8245, Abcam
Anti-p300 (C-20)	Primary rabbit polyclonal to human p300 protein	WB	sc-585, Santa Cruz Biotechnology
Anti-tubulin	Primary mouse monoclonal to human tubulin	WB	T9026, Sigma-Aldrich
Anti-histone H3	Primary rabbit polyclonal to histone H3	WB	ab1791, Abcam
Anti-gamma H2A.X	Primary rabbit polyclonal to histone γH2AX	ChIP	ab2893, Abcam
Anti-HMGB1	Primary rabbit polyclonal to HMGB1	ChIP	ab18256, Abcam
Anti-H1	Primary goat polyclonal to histone H1	ChIP	sc-34464, Santa Cruz Biotechnology
Anti-mouse-IgG-HRP	Horseradish peroxidase-conjugated secondary antibody to mouse IgG protein	WB	NA9310, GE Healthcare
Anti-rabbit-IgG-HRP	Horseradish peroxidase-conjugated secondary antibody to rabbit IgG protein	WB	NA9340, GE Healthcare
Alexa Fluor 488 anti-mouse IgG (H+L)	Secondary goat antibody to mouse IgG conjugated with Alexa 488	IF	A11029, Invitrogen
Alexa Fluor 4594 anti-rabbit IgG (H+L)	Secondary goat antibody to rabbit IgG conjugated with Alexa 594	IF	A11037, Invitrogen

4.4. Immunofluorescence and microscopy

Cells were fixed with 3% formaldehyde in phosphate-buffered saline (PBS) for 20 min at 22 °C followed by 2 × 5 min washes with PBS. The fixed cells were permeabilised with 0.5% Triton X-100/1% normal goat serum (NGS) for 10 min at 22 °C followed by 3 × 10 min washes with PBS containing 1% NGS. The cells were incubated with primary and secondary antibodies for 1 h with 2 × 10 min washes with PBS containing 1% NGS in between. Finally, the cells were washed 4 × 10 min including the staining of the nuclei with 4',6-diamidino-2-phenylindole (DAPI, 1:5000 dilution) during the third wash. The cells were covered with Fluorescent Mounting Medium (DakoCytomation) and the staining results were captured with the LSM5 DUO laser confocal microscope (Zeiss) at the Core Facility of Visualisation and Screening at the University of Tartu, Estonia.

4.5. Immunoprecipitation

The cells were scraped off of the 15 cm tissue culture plates and washed twice with ice-cold PBS and the pellet was lysed in radioimmunoprecipitation assay (RIPA) buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% NP-40, 0.5% Na-deoxycholate, 0.1% SDS) supplemented with protease inhibitors (5 µg/ml leupeptin, 5 µg/ml aprotinin, 200 µg/ml 4-(2-aminoethyl)benzenesulfonyl fluoride), phosphatase inhibitors (50 mM NaF, 0.2 mM Na_3VO_4), histone deacetylase inhibitor (10 mM Na-butyrate) and 50 µM $ZnCl_2$ to stabilise AIRE zinc fingers. Lysates were kept on ice for 10 min and clarified by centrifugation at 16000 × g for 10 min at 4 °C. The supernatants were incubated with the AIRE 6.1 antibody (Table 3) or NGS for 2 h at 4 °C. The immune complexes were precipitated with Protein G Sepharose™ 4 Fast Flow (GE Healthcare) for another 2 h at 4 °C. To prepare samples for mass-spectrometry, AIRE was precipitated with cyanogen bromide-activated Sepharose™ 4B (GE Healthcare) conjugated with anti-AIRE 6.1 antibody. The Sepharose beads were washed 3 × 1 ml low salt wash buffer (20 mM Tris-HCl pH 8.1, 150 mM NaCl, 2mM ethylenediaminetetraacetic acid (EDTA), 0.1% SDS, 1% Triton X-100), 2 × 1 ml high salt wash buffer (20 mM Tris-HCl pH 8.1, 500 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100) and 1 × 1 ml TE (10 mM Tris-HCl pH 7.4, 1 mM EDTA). The protein was eluted from the Sepharose beads with 30–40 µl 2× Laemmli buffer (Laemmli, 1970) and heated for 5 min at 95 °C. The immunoprecipitation samples were analysed by western blotting (see section 4.6) and mass-spectrometry, which was performed as a service at the Core Laboratory of Proteomics in the Institute of Technology, University of Tartu, Estonia.

4.6. Western blot

Immunoprecipitation samples or cell lysates were loaded onto an 8% sodium dodecyl sulphate polyacrylamide gel (SDS-PAG) and run for 15 min at 80 V followed by 30 min at 150 V on a Mini-PROTEAN Electrophoresis System (Bio-Rad Laboratories). The separated proteins were transferred to Immobilon^{TM}-P polyvinylidene difluoride membranes (Millipore) for 20 min at 15 V on a Trans-Blot semidry electrophoretic transfer cell (Bio-Rad Laboratories). The membranes were blocked with 5% (w/v) skimmed milk or 5% (w/v) bovine serum albumin in TBST buffer (100 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Tween 20) overnight at 4 °C. The membranes were incubated with primary and secondary antibodies for 1 h with 2 × 10 min washes with 5% skimmed milk/ TBST buffer in between. The final 3 × 10 min washes with 5% skimmed milk/ TBST buffer were followed by 1 min incubation with ECL Advance Western Blotting Reagent (GE Healthcare). The protein signals were captured on X-ray films (Fuji) or by the ImageQuant^{TM}-RT ECL image analysis system (GE Healthcare).

4.7. Protein stability assay

The cells were treated with protein translation inhibitor cycloheximide (40 µg/ml) 24 h after transfection and lysed with 2× Laemmli buffer (Laemmli, 1970) at 0, 2, 4, and 6 h time points. The lysates were processed for western blotting as described in section 4.6. The protein signals were captured and quantified by the ImageQuant^{TM}-RT ECL image analysis system (GE Healthcare). The AIRE protein signals were normalised to the signal of the housekeeping protein GAPDH.

4.8. Luciferase activation assays

The HEK293 cells were grown and transfected as described in section 4.2. Luciferase reporter assays were performed at 48 h time point after transient transfection. The cells were lysed with 1× lysis buffer (Promega) and incubated with the luciferase substrate according to the Luciferase Assay System kit protocol (Promega). Luminescence was counted with Wallac 1420 Victor Multilabel/Plate Reader (Perkin Elmer).

4.9. RNA purification and quantitative real-time PCR

The total RNA from transfected cells was purified with TRIzol (Invitrogen) according to manufacturer's instructions. For complementary DNA synthesis, 3–5 µg of total RNA was processed with the SuperScript III First-Strand Synthesis

System (Invitrogen) according to manufacturer's instructions. Quantitative PCR (qPCR) was performed with the ABI Prism 7900 HT Fast Real-Time PCR System (Applied Biosystems) using the qPCR SYBR Green Core Kit (Eurogentec). All reactions were performed in triplicates and experiments were repeated at least twice. The expression values were normalised to the hypoxanthine-guanine phosphoribosyltransferase 1 (*HPRT1*) housekeeping gene. The fold difference relative to a control sample was assessed by the comparative Ct method (Applied Biosystems). The expression analysis primers are listed in Table 4. At least one primer from each pair was designed to cover an exon-exon boundary to increase the specificity of the PCR reactions.

Table 4. List of expression analysis primers used in the study.

Gene	Primer name	Sequence
S100A8	Hu_S100A8_F	CTCAGTATATCAGGAAAAAGGGTGCAGAC
	Hu_S100A8_R	CACGCCCATCTTTATCACCAGAATGAG
IVL	Hu_IVL_F	GCCTTACTGTGAGTCTGGTTGACA
	Hu_IVL_R	GGAGGAACAGTCTTGAGGAGCT
HBG2	Hu_HBG2_F	CATAAAGCACCTGGATGATCTC
	Hu_HBG2_R	CAGGAGCTTGAAGTTCTCAG
PDYN	Hu_PDYN1_F	TGCCTTTGTTCTATTTTTGCAGGT
	Hu_PDYN1_R	CAGCAATTCCTGCGGCTTG
KRT73	Hu_KRT73_F	GAGTGCAGGATGTCCGGAGAATA
	Hu_KRT73_R	TTGCTGAATCCAAAGCCAGCC
CEACAM5	Hu_CEA5_F	CCTGGATGTCCTCTATGGGC
	Hu_CEA5_R	TACTGCGGGGATGGGTTAGA
PSMD4	Hu_PSMD4_F	GAAGGTGAAAGAGACTCA
	Hu_PSMD4_R	GTCATACTGCTTAGGTCA
HPRT	Hu_HPRT_F	GACTTTGCTTTCCTTGGTCAGG
	Hu_HPRT_R	AGTCTGGCTTATATCCAACACTTCG
BAI3	BAI3 F	GCAGCGATGCCCTGCACCTT
	BAI3 R	GGTGCCTGTGGCATTCAGGGG
CDH11	CDH11 F	CGTGCTTGTGGGCAGGCTTCA
	CDH11 R	TGTCCACCGCCTGAGCCATCA
IFI16	hIFI16_exp_F	CTGTGAGGAAGGAGATAAACTG
	hIFI16_exp_R	TCTTGATGACCTTGATGTGAC
CHST13	CHST13 F	CTCCCTGCGCCCGGCATTTG
	CHST13 R	CGTGAGTGGCGGCTACAGGC
PAPLN	PAPLN F	TGACTGCAGCGCGGAGTGTG
	PAPLN R	CCCTGCCTTCCAGCGCTTGG
CXCR7	CXCR7 F	CTGCAGCCAGCAGAGCTCACA
	CXCR7 R	GAGGCGGGCAATCAAATGACCTCC

Table 2. Continuation

Gene	Primer name	Sequence
CLDN1	h_CLDN1_expr_F	AATTCTATGACCCTATGACCC
	h_CLDN1_expr_R	GACAGGAACAGCAAAGTAGG
LY6G6D	hLY6G6D_exp_F	TGGAACAGATCAAGCTACCT
	hLY6G6D_exp_R	CTCTGTCTCCACTTGATTGC
LPL	hLPL_exp_F	TGCTTGAGTTGTAGAAAGAACC
	hLPL_exp_R	TTGGTAATGGAAGACTTTGTAGG
BAT2	h_BAT2_expr_for	CCAGAGCAAATCTTACCCAG
	h_BAT2_expr-rev	TACTGTAACCTGCTGGAGAG
S100A10	S100A10_for	TCCCTGGATTTTTGGAAAATCAAAAAGACC
	S100A10_rev	CCCGCAATTAGGGAAAAGAAGCTCTGGAA

4.10. Gene expression array

Total RNA from HEK-AIRE, HEK-AIRE.K243/253Q and HEK-YFP cells was isolated with TRIzol (Invitrogen) according to manufacturer's instructions. Samples were treated with Turbo DNase (Ambion) to remove traces of genomic DNA and purified with QIAquick PCR Purification Kit (Qiagen). The labelled complementary RNA (cRNA) was prepared with Illumina TotalPrep RNA Amplification Kit (Ambion) and hybridised to Illumina HumanWG-6_V2_0_R2 BeadChip expression arrays (lllumina). The quality of the RNA and labelled cRNA was assessed with the 2100 Bioanalyzer (Agilent). The samples were collected from three independent experiments. The raw signal intensities were normalised with the rank-invariant normalisation method implemented in the BeadStudio Gene Expression Module v3.3.7 (Illumina). Genes with an absolute differential score above 13, which corresponds to a p-value threshold of 0.05 ($-10 \times \log_{10}0.05$), and an absolute fold change cut-off of 2 were considered to be differentially expressed. Gene and sample groups were analysed with unsupervised hierarchical clustering by using Manhattan distance and complete linkage criteria as implemented in the Multi Experiment Viewer software version 4.0 (http://mev.tm4.org/). The tissue specificity of the differentially expressed genes was analysed using the RefSeq and Unigene gene lists in the TiGER database (Liu et al., 2008). The statistical significance of the enrichment of TRA genes among differentially expressed genes was assessed with chi-squared test by considering the number of all genes (19646) and the fraction of TRA genes (5524) found on both Illumina arrays and in the TiGER database (p-value threshold – 0.05). The genomic gene clusters were analysed with the CROC software with default options except that the minimum number of interesting genes in a cluster was set to two (Pignatelli et al., 2009). The statistical significance of the genomic clustering was calculated as follows: 1) 10,000 random gene sets size-matched to the observed differentially expressed gene list were generated based on the human genome build 36.1/hg18 provided by the National Center for

Biotechnology Information (NCBI); 2) each random gene set was analysed for genomic clustering with the CROC software; 3) the number of occurrences where the number of gene clusters found in random gene sets was equal or greater than the number of observed gene clusters was divided by the number of random gene sets to give a p-value.

4.11. GST pull-down assays

The glutathione S-transferase (GST)-tagged proteins were expressed in the NovaXG *E.coli* strain (Novagen) by addition of 0.8 mM isopropyl β-D-1-thiogalactopyranoside and 0.2 mM $ZnCl_2$. The proteins were purified using Glutathione Sepharose™ 4B (GE Healthcare) as described in (Org et al., 2008). Sepharose-bound proteins (2–10 µg) were incubated for 4 h at 4 °C with 20 µg calf thymus histones (Sigma-Aldrich) in binding buffer containing 50 mM Tris–HCl pH 7.5, 1M NaCl, 1% NP-40, 0.5 mM EDTA and 1× protease inhibitor cocktail (Sigma-Aldrich). The samples were washed with 8 × 1 ml binding buffer and run on 15% SDS-PAG for 20 min at 100 V and followed by Coomassie G-250 staining.

4.12. Chromatin immunoprecipitation

ChIP experiments were performed using 10^7 cells per experiment according to a previously described protocol with slight modifications (Dahl and Collas, 2007). Briefly, the uninduced and doxycycline-induced AIRE-Tet cells, either untreated or treated with etoposide, were cross-linked with 1% formaldehyde for 10 min at room temperature, and the formaldehyde was quenched by the addition of glycine to a final concentration of 0.125 M. Cells were lysed in 200 µl lysis buffer (50 mM Tris-HCl, pH 8, 10 mM EDTA, 1% SDS, 1× Halt Protease Inhibitor Cocktail, Thermo Scientific). The samples were diluted 10 times in RIPA buffer containing 1× Halt Protease Inhibitor Cocktail. Sonication was performed with the Diagenode Bioruptor for 15 min using 30 sec on-off cycles at a high voltage setting. After chromatin preparation, ChIP was performed with the IP-Star Compact Automated System (Diagenode) using the ChIP Direct method, which requires antibody-coated magnetic beads that are added to the sheared chromatin (the protocol and magnetic beads are from Diagenode). ChIP-grade antibodies against high mobility group B1 (HMGB1), γH2AX and histone H1 are listed in Table 3. The ChIP signals were measured by qPCR and the values were log-transformed, standardised and scaled according to the procedure described in (Willems et al., 2008). Experimental groups were compared with the two-sample t-test implemented in R statistical software. Primers for the ChIP DNA analysis are listed in Table 5.

Table 5. List of ChIP primers used in the study.

Region	Primer name	Sequence
S100A8 promoter	ch_SA81_F	TGTGCTGGGTCCCCAATGGC
	ch_SA81_R	GCTGCTTGGGGTCCCTCTGC
S100A8 control	ch_SA81_con_F	TGGCTTTGGTCTCGCCGTCTAAGTAA
	ch_SA81_con_R	TGGTGGGTTCAAGGTGCACTGTAGAT
IVL promoter	ch_IVL_F	CCAATCCTTTAGATATGGTACACAG
	ch_IVL_R	TCCCCAGGTCTCTGGTTCTT
IVL control	ch_IVL_con_F	TGTTTGTGTTGTGCAAGGCCGAGA
	ch_IVL_con_R	AGGAACATTTTGTCAGGCCAAGGCT
PSMD4 promoter	ch_PSMD4_F	GATAGTCCCGGGTTACCAC
	ch_PSMD4_R	TGTAGCTAAAGACAGACCCG
GAPDH promoter	ch_GAPDH_F	CCCGTCCTTGACTCCCTAGT
	ch_GAPDH_R	GGGGGAAGGGACTGAGATT

4.13. Formaldehyde-assisted isolation of regulatory elements

The FAIRE experiments were done with the uninduced and doxycycline-induced AIRE-Tet cells that were either left untreated or treated with etoposide. The cells were grown on 10 cm tissue culture plates to near full confluence, which corresponds to ~10^7 cells. The cells were cross-linked with 1% formaldehyde for 10 min at 22 °C followed by quenching the cross-linking reaction by adding glycine (final concentration 0.125 M) to the medium. The cells were lysed with 200 µl of lysis buffer (50 mM Tris-HCl (pH 8), 10 mM EDTA, 1% SDS, and 1× Halt protease inhibitor mixture (Thermo Scientific)). Cellular debris was removed by spinning the lysates at 20,000 × g for 10 min at 4 °C. The input samples (10% of lysates) were treated with 10 µg of DNase-free RNase A (Thermo) for 30 min at 37 °C, followed by 10 µg of proteinase K incubation overnight at 65 °C. The DNA from the remaining lysates and the proteinase K-treated input samples was prepared with standard phenol-chloroform extraction and additionally purified with a Qiagen PCR purification kit.

4.14. Chromatin conformation capture

Uninduced and doxycycline-induced AIRE-Tet cells were processed according to the protocol in (Hagège et al., 2007), except that the Sau3AI restriction enzyme (purchased from Thermo Fisher) was used in the DNA digestion step. The chromatin interactions were calculated as the ratio of the qPCR signal in

the doxycycline-induced sample relative to the signal in the uninduced control sample. The anchor region was selected based on the results of high throughput sequencing of the FAIRE samples. The 3C signals were measured by qPCR, and the values were log-transformed, standardised, and scaled according to the procedure described in (Willems et al., 2008). Statistical significance of the $\log_2$ fold change was assessed with a one-sample t-test implemented in R statistical software. Primers for the 3C qPCR analysis were designed according to the guidelines in (Hagège et al., 2007) and are listed in Table 6.

Table 6. The list of 3C primers used in the study.

Region	Primer name	Sequence
CTCF site 10 forward	3c_anchor_site10_F	CTCCCCAAGCTCTAACAACCAA
CTCF site 10 reverse	3c_anchor_site10_R	CTCTTTGCACCTCAGTCCTCTC
CTCF site 1	3c_site1_F	GACTTAGAGGCTTCAGTCATCATCC
CTCF site 2	3c_site2_F	TAAGGAGCAAGGAGACCAGGAG
CTCF site 3	3c_site3_F	CTTCCCTTGGCCATTTCCCA
CTCF site 4	3c_site4_F	CAAATTCCCGTCCTCAAGCAATC
CTCF site 5	3c_site5_F	GAAATTAGCCTCACTGAGTCACTGT
CTCF site 6	3c_site5_R	GAGCTGGGAAATAACACTCACACTA
CTCF site 7	3c_site7_F	AGTTGGTAGGAGCGACTTTAGAAAT
CTCF site 8	3c_site8_F	CCTATGACCCTTAGCCTCTCTGAG
CTCF site 9	3c_site9_F	AGATTGTGGTCTTATGTCAGGTCAA
CTCF site 11	3c_site11_R	TGGTGGACAGGAGGGAAGTG
CTCF site 12	3c_site12_F	TTTGCTAAGGAAGTGGAGGTGGA
CTCF site 13	3c_site13_R	CTTGTGGAAACTCTGAGAACTGCAT
CTCF site 14	3c_site14_R	CGGAGAACTGCTTCACAATCTTTAA
CTCF site 15	3c_site15_R	AAAATGAAGCGACTTGTCCAGG
CTCF site 16	3c_site16_F	CACATATTCCCAACAACTCTGCAAG
CTCF site 17	3c_site17_F	TGGAAGTAACTGTCAGAGAGAGCT
CTCF site 18	3c_site18_F	AAAATATAGAAATATGGGGCCGGGC
CTCF site 19	3c_site19_F	CACTTATTCCCAACAACTCTGCAAG
CTCF site 20	3c_site20_F	ACAAACAAACTCAGGCTGTAAAGAC
CTCF site 21/22	3c_site21/22_F	GTGCAGCTAGATGGTCAGTCC
CTCF site 23	3c_site23_R	GCCAATTTAGATTTACCTGCCCC
CTCF site 24	3c_site24_R	AGTGTAAACAACGGTGCTTTTAACA

4.15. High-throughput sequencing and data analysis

The FAIRE-seq libraries were prepared following the Illumina Truseq ChIP sample preparation instructions and sequenced on an Illumina HiSeq 2000 machine to produce 100 bp paired-end reads. The quality control of the sequencing results was performed with the FastQC software (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Sequencing adapter sequences and low quality base calls (Phred score < 30) were removed with Trim Galore! (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). The remaining paired-end reads were mapped to the human genome build 37/hg19 povided by the NCBI with bowtie2 in "local" mode (Langmead and Salzberg, 2012). Properly paired reads with a mapping quality of > 30 were used for detecting differentially accessible chromatin regions with the csaw package for R statistical software (Lun and Smyth, 2016). The differentially accessible regions were annotated with the ChIPseeker R package (Yu et al., 2015). The statistical significance of overlap between different sets of genomic intervals was assessed with the GenometriCorr package for R statistical software (Favorov et al., 2012).

5. RESULTS

5.1. The post-translational acetylation of the AIRE protein by the transcription co-activator protein p300 (Study I)

Previous studies had established that CBP was a major interaction partner for AIRE (Pitkänen et al., 2000). CBP is a well-known transcription co-activator that in addition to mediating various protein-protein interactions possesses acetyltransferase activity. Interestingly, the human genome encodes a paralog of CBP, which is named p300 and has both redundant and unique functionality compared to CBP (Vo and Goodman, 2001). These observations led to the investigation of the post-translational acetylation of AIRE by CBP or p300 by a series of overexpression and immunoprecipitation assays.

The expression plasmids of CBP, p300, p300DY (an acetyltransferase-deficient form of p300) and PCAF were transfected into the stable AIRE-expressing HEK-AIRE cell line (Org et al., 2008). Immunoprecipitation was performed with anti-AIRE antibody and the signal of acetylated lysines was detected with anti-acetyl lysine specific antibody. These steps were also performed with HEK-AIRE cells that were transfected with YFP-encoding expression plasmid, which served as a negative control experiment, and with HEK-AIRE cells that were treated with Trichostatin A (TSA), a compound that is known to inhibit protein deacetylase activity, thereby increase the overall level of protein acetylation and could act as a positive control in this experimental setup. The ectopic expression of CBP, p300, PCAF, and the TSA treatment resulted in a strong total protein acetylation signal, while the ectopic expression of p300DY, similarly to YFP, did not show any significant increase in the fraction of acetylated proteins (Figure 2A, input lanes).

Furthermore, a distinct acetylation signal was visible at the molecular weight corresponding to AIRE protein after the overexpression of p300, which was weakly present or undetectable in other conditions (Figure 2A, IP lanes).

These findings encouraged to perform a mass-spectrometric analysis to compare the protein extracts from HEK-AIRE cells transiently transfected with p300 or CBP with protein extracts from untransfected cells. Altogether 12 acetylated lysine residues (50% of all lysine residues in AIRE protein) were identified from which many were detectable only after the overexpression of p300 (Figure 2B, Table 7).

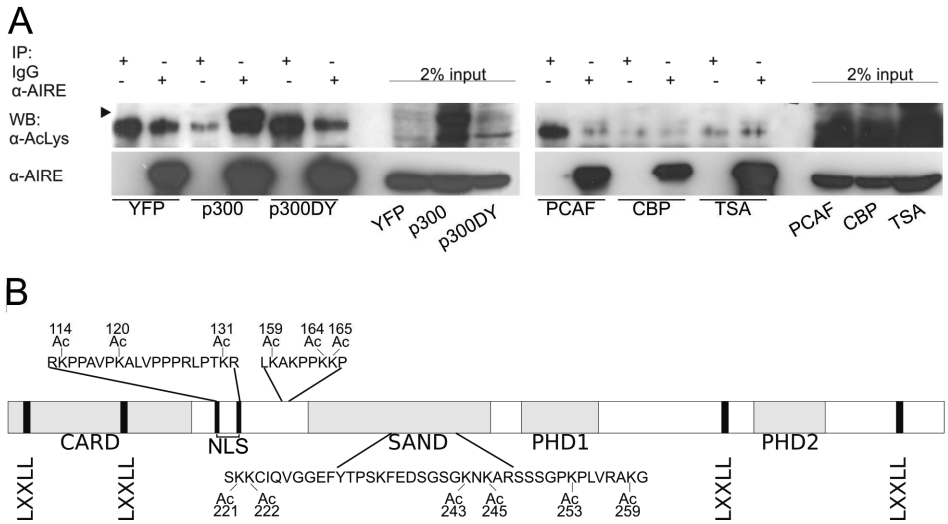

Figure 2. The AIRE protein is the substrate for acetylation by p300. (A) Detection of acetylated AIRE in HEK-AIRE cells after overexpression of CBP, PCAF, p300, p300DY or TSA treatment. Cells transfected with YFP were included as negative control. AIRE was immunoprecipitated with anti-AIRE 6.1 antibody (α-AIRE) and a control IP was performed with normal goat serum IgG (IgG) for every sample. Immunocomplexes were resolved on SDS-PAG and acetylated AIRE was detected with anti-AcLys antibody (α-AcLys) in western blot. The arrow indicates the position of AIRE. Western blot with anti-AIRE 6.1 antibody was used to determine the presence of the AIRE protein. The 2 % input refers to the amount of whole cell lysate loaded onto the gel. (B) Protein domain diagram displaying all acetylated lysines in AIRE according to the mass-spectrometric analysis (CARD, caspase recruitment domain; NLS, nuclear localisation signal; SAND, the Sp100, AIRE-1, NucP41/P75 and DEAF1 domain; PHD, plant homeodomain zinc finger). The figure is adapted from Study I.

Table 7. Acetylated lysine residues in the AIRE protein determined by mass-spectrometry. Roman numerals refer to experimental replicates.

	K114	K120	K131	K159	K164	K165	K221	K222	K243	K245	K253	K259
AIRE I									+		+	+
AIRE II									+		+	
AIRE + p300 I	+	+	+	+	+		+				+	
AIRE + p300 II	+	+			+	+		+	+	+	+	+
AIRE + CBP			+				+	+	+		+	+

The overexpression of CBP resulted in fewer acetylated lysines, although all of them overlapped with the sites found in at least one sample with p300 overexpression. Altogether 3 sites were acetylated in cells without any ectopic expression of p300 or CBP suggesting that AIRE is also a substrate for endogenous acetyltransferases.

Three acetylated lysines, K114, K120, and K131, were found in the known AIRE NLS while the acetylated residues K159, K164, and K165 were mapped to putative bipartite NLS (Ilmarinen et al., 2006; Saltis et al., 2008) (Figure 2B). The remaining six modified residues, K221, K222, K243, K245, K253, and K259 were located in the SAND domain, which has been characterised as a DNA binding domain in many other proteins, but not in AIRE (Figure 2B). In conclusion, these results show that AIRE can be acetylated by ectopically expressed p300 and CBP as well as endogenous acetyltransferases in HEK293 cells.

5.2. The effect of acetylation site mutations in the AIRE SAND domain on protein localisation (Study I)

The SAND domain is a DNA binding domain in several transcriptional regulators (Bottomley et al., 2001; Isaac et al., 2006; Jensik et al., 2014). Although DNA binding by AIRE SAND domain has been disputable, the acetylation of almost all lysines in AIRE SAND domain prompted to ask whether it has any effect on the localisation of the AIRE protein and its function as a transcriptional regulator. Subsequent experiments included a set of AIRE expression plasmids where the most frequently acetylated lysines in the AIRE SAND domain – K243 and K253 together with K245 in-between – were substituted with either glutamine, which has a neutral charge mimicking an acetylated position, or arginine, which preserves the positive charge similar to an unmodified lysine residue (Choudhary et al., 2009; Hecht et al., 1995; Li et al., 2002; Wang and Hayes, 2008). The expressed sets of AIRE proteins were named AIRE Q and AIRE R mutants, respectively.

Immunofluorescence analyses have established that the AIRE protein forms punctate structures in the cell nucleus, which have been named AIRE nuclear bodies, and that AIRE colocalises with the transcriptional coactivator CBP (Pitkänen et al., 2005). In this study, we showed that AIRE also colocalises with endogenous p300, a paralog of CBP, in HEK-AIRE cells (Figure 3A).

The K-to-Q or K-to-R mutations in the AIRE SAND domain did not disturb the colocalisation of AIRE and p300. However, the AIRE Q mutants concentrated into fewer but larger nuclear bodies while the AIRE R mutants formed nuclear bodies characteristic of wild-type AIRE, both in terms of number and size (Figure 3B and C).

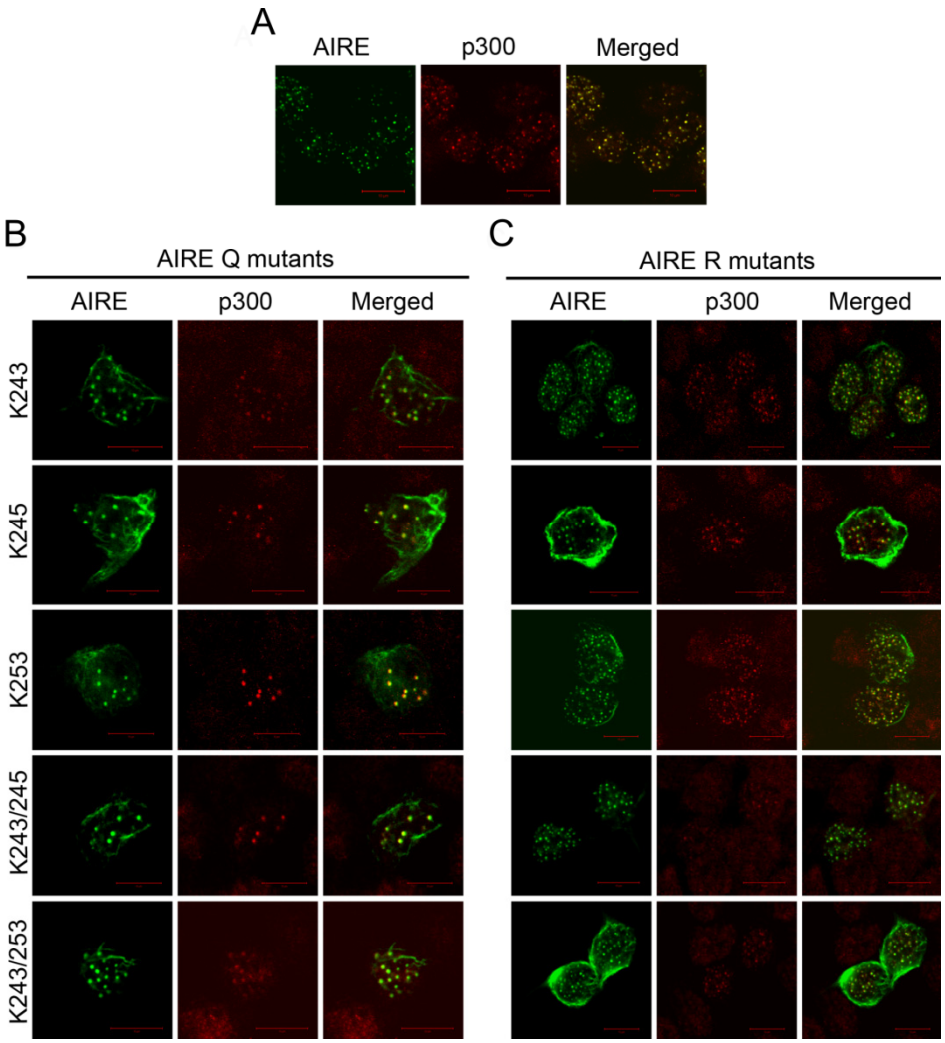

Figure 3. Wild-type and mutated AIRE colocalise with endogenous p300, but display differences in the size and number of nuclear bodies in HEK293 cells. (A) Colocalisation of stably transfected wild-type AIRE and endogenous p300 in HEK-AIRE. (B) Colocalisation of transiently transfected AIRE Q mutants and endogenous p300 in HEK293 cells. (C) Colocalisation of transiently transfected AIRE R mutants and endogenous p300 in HEK293 cells. AIRE and p300 were detected with anti-AIRE and anti-p300 antibodies, respectively. Scale bar corresponds to 10 µm. The figure is adapted from Study I.

5.3. The effect of overexpression of p300 on AIRE protein stability and localisation (Study I)

Lysine acetylation is known to prolong protein half-life by preventing lysine polyubiquitylation, which is recognised by the proteasomal system as a protein degradation signal (Glozak et al., 2005). An earlier study had shown that AIRE is ubiquitylated and that the inhibition of the proteasome increases the AIRE protein level in the nucleus (Akiyoshi et al., 2004). To test whether the acetyltransferase activity of p300 could affect the AIRE protein levels, HEK-AIRE cells were transfected with either wild-type or acetyltransferase-deficient p300 and treated with the protein translation inhibitor cycloheximide.

The analysis of the AIRE protein abundance in the cell lysates at several time points showed that the AIRE protein is stabilised only after the overexpression of the wild-type p300, while the level of AIRE in cells expressing p300DY or YFP (negative control) declined about 50% in 6 h (Figure 4A).

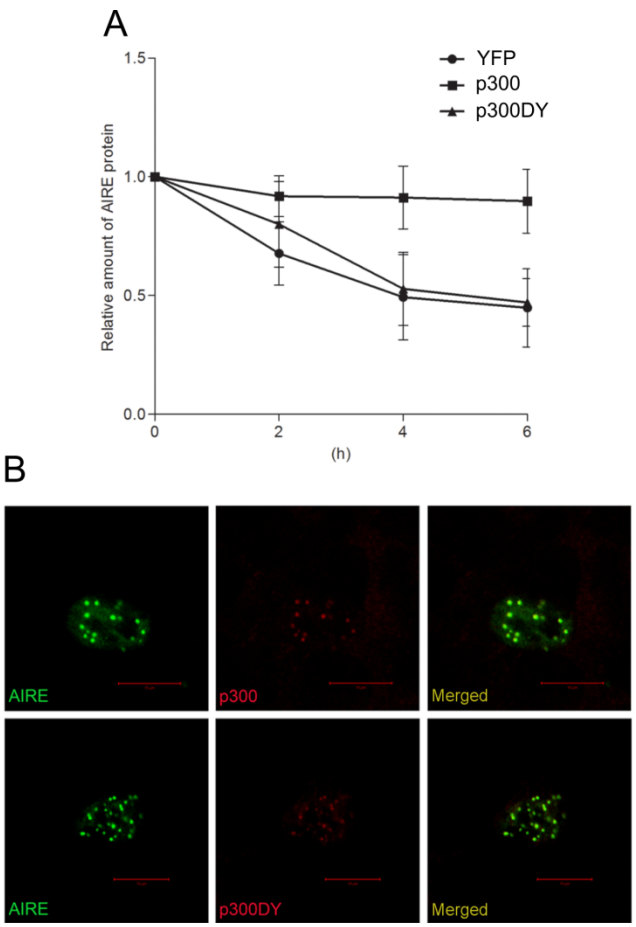

Figure 4. Overexpressed p300 stabilises AIRE and localises to AIRE nuclear bodies in HEK-AIRE cells. (A) The stability of AIRE after the overexpression of p300 in HEK-AIRE cell line. Cells transfected with p300, p300DY or YFP were incubated with translation inhibitor cycloheximide and lysed at different time points. Lysates were subjected to western blotting and AIRE signal intensities were normalised to GAPDH expression. The graphs represent three independent experiments ± standard error of the mean (SEM). (B) Co-localisation of AIRE with wild-type p300 or the acetyltransferase-deficient p300DY in HEK-AIRE cells. AIRE and p300 were detected with anti-AIRE (green) and anti-p300 antibodies (red), respectively. Scale bar corresponds to 10 µm. The figure is adapted from Study I.

Interestingly, the overexpression of p300, and not p300DY, resulted in larger and fewer AIRE nuclear bodies resembling the localisation of the AIRE Q mutants (Figure 4B).

5.4. The impact of acetylation site mutations and p300 overexpression on the transcriptional activity of AIRE (Study I)

A series of luciferase activation assays and expression analyses of endogenous genes by quantitative PCR were performed to test whether the transcriptional activity of AIRE is altered alongside with the observed changes in protein localisation and stability due to p300 overexpression or mutations in the AIRE SAND domain. In these experiments, the expression of two known AIRE-dependent genes – involucrin (*IVL*) and S100 calcium binding protein A8 (*S100A8*) (Derbinski et al., 2005; Liiv et al., 2008; Org et al., 2009) – was used to assess how AIRE activates its endogenous targets in HEK293 cells. For the luciferase activation assays, a reporter plasmid containing the *IVL* gene promoter upstream of the luciferase coding sequence was used.

Both endogenous and reporter gene activation assays showed that the ectopic expression of p300 with or without its acetyltransferase activity did not significantly influence the AIRE-dependent gene activation in HEK293 cells (Figure 5A and B).

However, the luciferase activity after the overexpression of AIRE Q mutants decreased 1.5–3.4 times compared to the wild-type protein (Figure 5C), while the AIRE R mutants retained their transcriptional activity at levels comparable to that of wild-type AIRE (Figure 5D).

The K-to-Q mutations had an even more drastic effect on the expression of endogenous AIRE target genes by reducing the transactivation of AIRE at least 5 times (Figure 5E). Again, the gene activation mediated by the AIRE R mutants was similar to that of wild-type AIRE (Figure 5F). These data suggest that the acetylation of the lysine residues within the AIRE SAND domain contributes to the transcriptional activity of AIRE. Furthermore, the similar results of wild-type and AIRE R mutants indicated that AIRE might be deacetylated prior to the activation of AIRE-dependent genes.

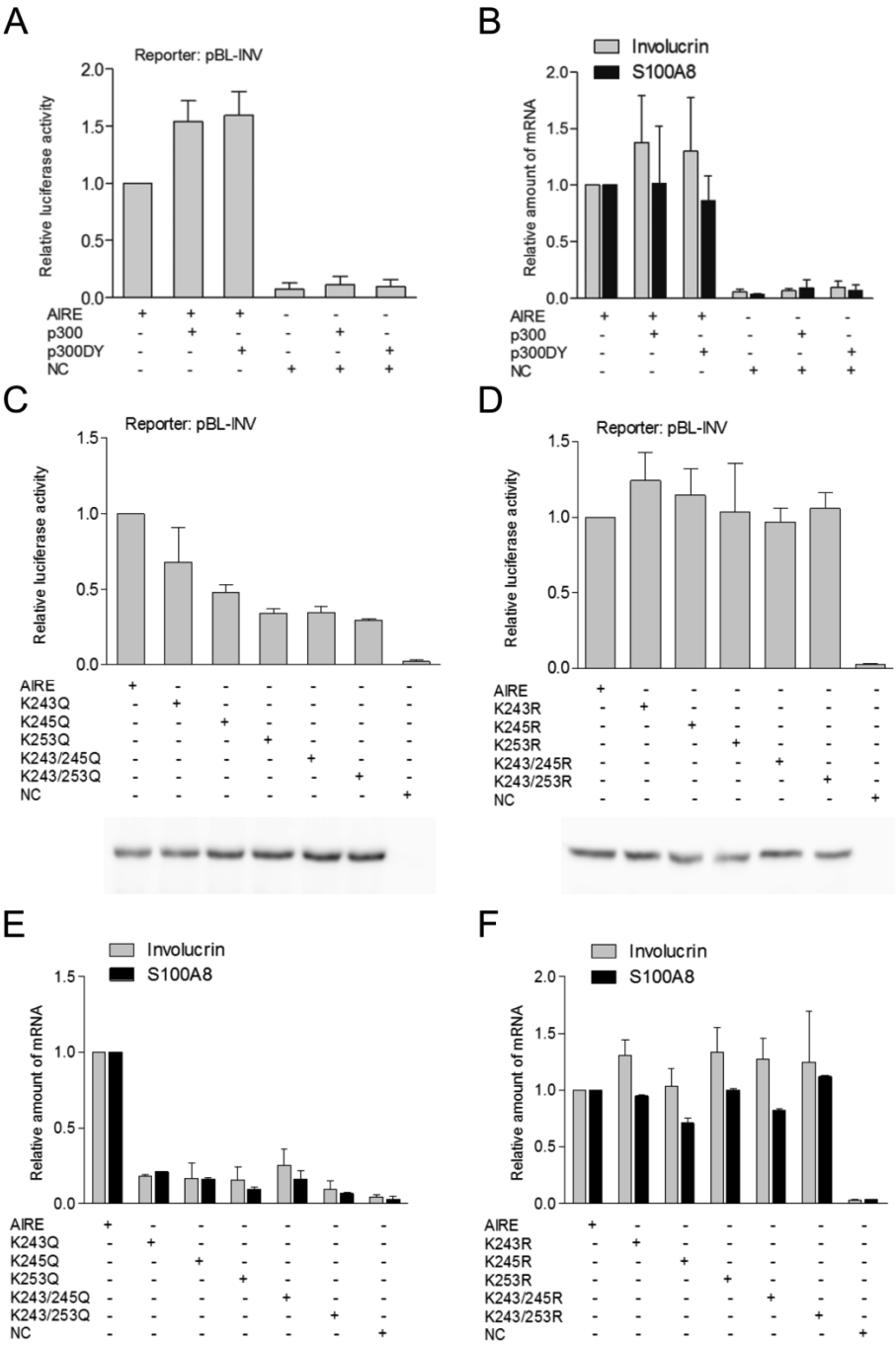
A
Reporter: pBL-INV
Relative luciferase activity
2.0
1.5
1.0
0.5
0.0
AIRE + + + - - -
p300 - + - - + -
p300DY - - + - - +
NC - - - + + +
B
Involucrin
S100A8
Relative amount of mRNA
2.0
1.5
1.0
0.5
0.0
AIRE + + + - - -
p300 - + - - + -
p300DY - - + - - +
NC - - - + + +
C
Reporter: pBL-INV
Relative luciferase activity
1.5
1.0
0.5
0.0
AIRE + - - - - - -
K243Q - + - - - - -
K245Q - - + - - - -
K253Q - - - + - - -
K243/245Q - - - - + - -
K243/253Q - - - - - + -
NC - - - - - - +
D
Reporter: pBL-INV
Relative luciferase activity
1.5
1.0
0.5
0.0
AIRE + - - - - - -
K243R - + - - - - -
K245R - - + - - - -
K253R - - - + - - -
K243/245R - - - - + - -
K243/253R - - - - - + -
NC - - - - - - +
E
Involucrin
S100A8
Relative amount of mRNA
1.5
1.0
0.5
0.0
AIRE + - - - - - -
K243Q - + - - - - -
K245Q - - + - - - -
K253Q - - - + - - -
K243/245Q - - - - + - -
K243/253Q - - - - - + -
NC - - - - - - +
F
Involucrin
S100A8
Relative amount of mRNA
2.0
1.5
1.0
0.5
0.0
AIRE + - - - - - -
K243R - + - - - - -
K245R - - + - - - -
K253R - - - + - - -
K243/245R - - - - + - -
K243/253R - - - - - + -
NC - - - - - - +

Figure 5. The acetylation site mutations, but not the overexpression of p300, affect the kiirustasime rongiletranscriptional activity of AIRE in HEK293 cells. (A) Luciferase reporter gene activity after the overexpression of AIRE with wild-type or acetyltransferase-deficient p300 (p300DY). HEK293 cell lysates were assayed for luciferase expression driven by the AIRE target gene *IVL* promoter in the pBL-INV reporter plasmid 48 h after transfection with AIRE. The results represent means of 3 independent experiments ± SEM. (B) The expression of endogenous AIRE target genes *IVL* and *S100A8* after the overexpression of AIRE with wild-type or acetyltransferase-deficient p300 (p300DY). The mRNA was extracted from HEK293 cells 48 h after transfection. The results represent means of 5 independent qPCR experiments ± SEM. (C) and (D) Luciferase reporter gene activity after the overexpression of AIRE Q and R mutants, respectively. HEK293 cell lysates were assayed for luciferase expression driven by the AIRE target gene *IVL* promoter in the pBL-INV reporter plasmid 48 h after transfection with AIRE. Expression of wild-type and mutant AIRE proteins is shown below the graphs. (E) and (F) The expression of endogenous AIRE target genes *IVL* and *S100A8* after the overexpression of AIRE Q and R mutants, respectively. The mRNA was extracted from HEK293 cells 48 h after transfection. (C–F) Luciferase assay and qPCR results are presented as means of two independent experiments ± SEM. NC, negative control sample (cells transfected with the empty expression vector pcDNA3.1B(–)Myc/His). The figure is adapted from Study I.

5.5. The effect of the K243/K253Q SAND domain mutation on the expression profile of AIRE target genes in HEK293 cells (Study I)

The results with the select target genes prompted to investigate whether the mutations at the acetylation sites in the AIRE SAND domain can interfere with the genome-wide expression of AIRE-dependent genes. To address this question, an additional HEK293 cell line was generated that expressed the AIRE.K243/253Q mutant protein. The mRNA expression profile of the HEK-AIRE.K243/253Q cell line was compared to the previously characterised HEK-AIRE and HEK-YFP (control) cell lines in Illumina array analysis (Org et al., 2008).

The differential expression analysis revealed that 243 genes were significantly upregulated in the HEK-AIRE cells compared to the control cell line. In contrast to the expectation that the HEK-AIRE.K243/253Q cells have a similar transcriptional program to HEK-YFP cells, 190 upregulated genes was detected of which 81 genes were shared with the HEK-AIRE cells (Figure 6A).

However, the fold changes of the differentially expressed genes in the presence of AIRE.K243/253Q seemed to be weaker, which was also confirmed by qPCR on a subset of differentially expressed genes (Figure 6B).

The genes up-regulated in HEK-AIRE or HEK-AIRE.K243/253Q cells were mostly expressed at a very low or undetectable level in HEK-YFP cells (average signal intensity was 176 for AIRE and 133 for K243/253Q). These findings are consistent with a previous report, which concluded that AIRE activates

genes with low initial expression and represses highly expressed genes (Org et al., 2009). In concordance with earlier studies, both AIRE and AIRE.K243/253Q-regulated genes were enriched for TRA genes and based on their location in the genome belonged to chromosomal gene clusters (Derbinski et al., 2005; Johnnidis et al., 2005; Org et al., 2009) (Table 8).

Taken together, the data show that the AIRE Q double mutant still retains transactivation activity, but targets a different set of genes than the wild-type protein.

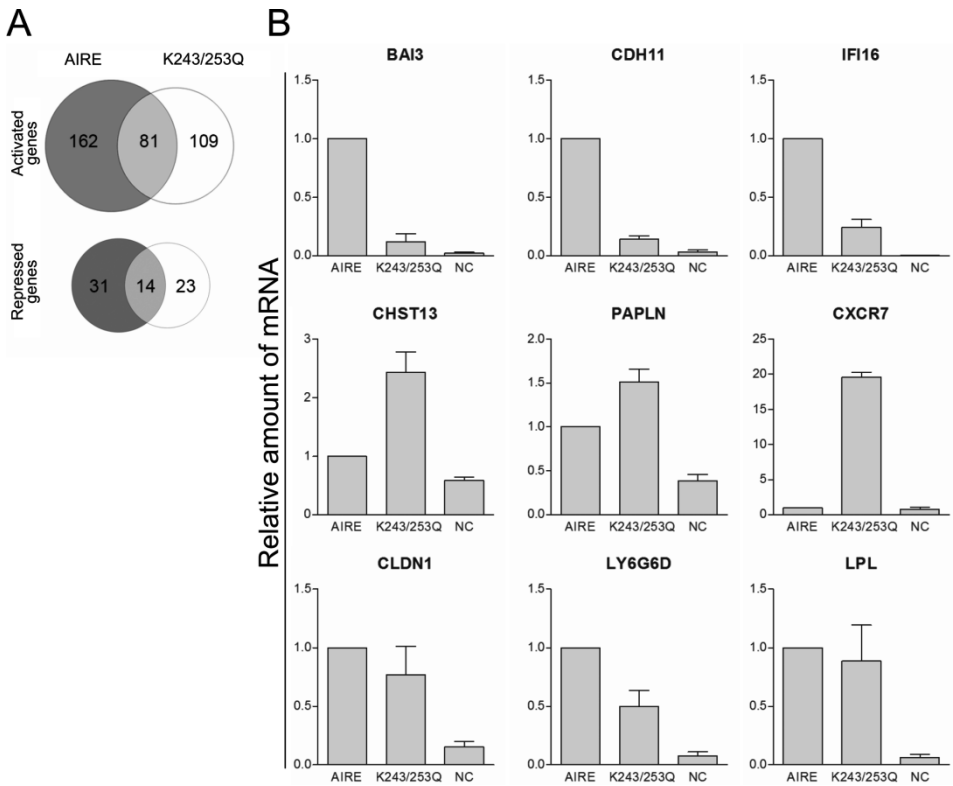

Figure 6. Acetylation mimicking mutations change the set of AIRE-regulated genes. (A) Venn diagram showing gene numbers activated or repressed by either wild type AIRE (dark grey) or AIRE.K243/253Q (white), and their overlap (light grey). (B) Relative mRNA levels of genes activated by wild-type AIRE (top row, *BAI3*, *CDH11*, and *IFI16*), AIRE.K243/253Q (middle row, *CHST13*, *PAPLN*, and *CXCR7*) and both (bottom row, *CLDN1*, *LY6G6D*, and *LPL*). The results are presented as mean of three independent qPCR experiments ± SEM. NC, negative control (HEK-YFP cell line). The figure is adapted from Study I.

Table 8. Enrichment of tissue-specific genes and chromosomal clustering of AIRE-regulated genes.

	Upregulated			Downregulated		
	Tissue-specific	**Total**	**p-value**	**Tissue-specific**	**Total**	**p-value**
AIRE	62	136	5.82E-06	9	28	0.63
AIRE. K243/253Q	40	97	0.004	9	19	0.06
Overlap	31	66	0.0007	6	13	0.15
	Upregulated			**Downregulated**		
	Genes	**Clusters**	**p-value**	**Genes**	**Clusters**	**p-value**
AIRE	214	21	0.017	43	3	0.019
AIRE.K243/253Q	168	16	0.007	31	2	0.034
Overlap	73	6	0.005	13	1	0.053

5.6. The role of the AIRE PHD2 domain mutation C446G in AIRE-dependent transcriptional regulation and protein-protein interactions (Study II)

The AIRE protein has two PHD-type zinc finger domains, named accordingly PHD1 and PHD2, and both have been linked to the role of AIRE as a transcriptional regulator. The PHD1 domain is known to mediate protein-protein interactions with the core nucleosome component histone H3 and mutations that abolish this interaction have been found in APS1 patients (Koh et al., 2008; Org et al., 2008). In addition, the AIRE PHD2 mutation C446G has been discovered in APS1 patients (Wolff et al., 2007). The AIRE PHD2 domain has been shown to participate in gene activation when tethered to the DNA-binding Gal4 domain (Meloni et al., 2008; Uchida et al., 2004). More recently, transgenic mice expressing AIRE that lacks the PHD2 domain displayed weaker expression of the AIRE-dependent TRAs (Yang et al., 2013). However, the molecular events that are associated to the mutations in the AIRE PHD2 and lead to disease have not been described before.

Study II focused on the identification of interaction partners that require either of the intact AIRE PHD domains to form protein complexes with AIRE. In the first part of the study, the structure of the AIRE PHD2 was modelled, which revealed the network of Zn^{2+} ion-coordinating (including the amino acid C446) and hydrophobic interactions that stabilise the protein domain (Study II, Figure 1). The APS1-associated mutation C446G disrupts metal binding and destroys the structure of PHD2 domain (Study II, Figure 1D). Furthermore, immunofluorescence analyses showed that the loss of the PHD2 domain causes the majority of AIRE to retain in the cytosol (Figure 7A) where it forms insoluble aggregates, which are difficult to detect by western blot (Figure 7B). In contrast, the mutations V301M and D297A in PHD1 domain do not affect the subcellular localisation of AIRE (Figure 7A and B).

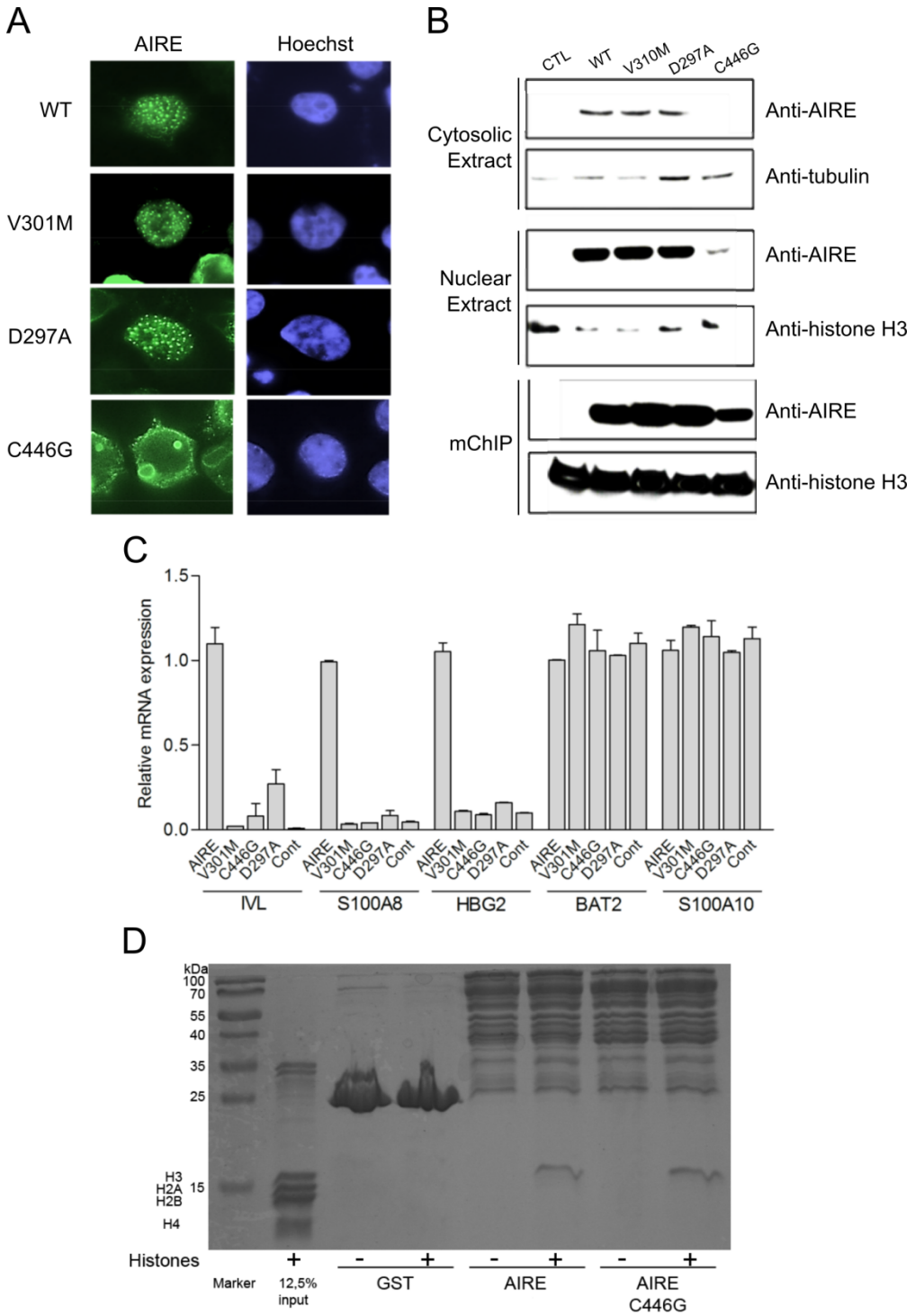
A
AIRE
Hoechst
WT
V301M
D297A
C446G
B
CTL
WT
V310M
D297A
C446G
Cytosolic Extract
Anti-AIRE
Anti-tubulin
Nuclear Extract
Anti-AIRE
Anti-histone H3
mChIP
Anti-AIRE
Anti-histone H3
C
Relative mRNA expression
1.5
1.0
0.5
0.0
AIRE
V301M
C446G
D297A
Cont
IVL
S100A8
HBG2
BAT2
S100A10
D
kDa
100
70
55
40
35
25
15
H3
H2A
H2B
H4
Histones
Marker
12,5% input
GST
AIRE
AIRE C446G

Figure 7. The effects of PHD finger mutations on AIRE subcellular localisation, AIRE-dependent target gene expression and histone binding. (A) AIRE staining (green) in HEK293T cell line transfected with expression plasmids for wild-type AIRE (WT), AIRE V301M, AIRE D297A and AIRE C446G. The nucleus was stained with Hoechst 33258 (blue). (B) Western blot of HEK293T cell lysates transfected with expression plasmids for wild-type AIRE (WT) and PHD finger mutants. The amount of recovered AIRE protein (anti-AIRE antibody, Everest Biotech) in the cytosolic and nuclear extracts as well as after chromatin solubilisation (mChIP) was compared to tubulin (anti-tubulin antibody, Sigma-Aldrich) and histone H3 (anti-histone H3 antibody, Abcam). CTL, negative control sample (cells transfected with the empty expression vector pcDNA3.1B(–)Myc/His). (C) A comparison of the transcriptional activity of wild-type AIRE and AIRE harbouring PHD1 mutations V301M and D297AV or the PHD2 mutation C446G. Relative expression level of AIRE-regulated genes *IVL*, *S100A8* and *HBG2* genes and AIRE-independent genes *BAT2* and *S100A10* was measured 48 h after transfection. The results are presented as means of two independent qPCR experiments ± SEM. Cont, negative control (cells transfected with the empty expression vector pcDNA3.1B(–)Myc/His). (D) The AIRE PHD2 mutant binds histone H3 as efficiently as wild type AIRE. The interaction between calf thymus histones and GST alone (GST), GST-AIRE (AIRE) or GST-AIRE.C446G (AIRE C446G) was visualised by Coomassie G-250 staining. The figure is adapted from Study II.

Decidedly, the structural aberrations related to the mutation C446G strongly reduced the expression of AIRE target genes, such as *IVL*, *S100A8* and the haemoglobin subunit gamma 2 (*HBG2*) (Figure 7C). The drop in gene activation was comparable to the one seen after the ectopic expression of AIRE carrying the PHD1 mutations D297A or V301M (Figure 7C), the former affecting binding to histone H3 and the latter potentially disrupting additional protein-protein interactions as previously reported (Koh et al., 2008; Org et al., 2008). Interestingly, in GST pull-down assays with full-length AIRE protein, the C446G mutation did not affect the binding to histone H3 (Figure 7D) suggesting that the structural integrity of PHD2 is not required for the interaction of AIRE with histone H3, which is in agreement with published data on peptide pull-down assays with AIRE deletion mutants (Koh et al., 2008).

In summary, the results suggest that the C446G mutation in AIRE PHD2 alters the protein structure and subcellular localisation preventing the activation of the expression of AIRE target genes. However, unlike mutations in the PHD1, the C446G mutation in AIRE PHD2 does not influence binding to histone H3, indicating that the two PHD zinc fingers mediate distinct protein-protein interactions.

5.7. The AIRE-dependent changes in chromatin accessibility and conformation near AIRE target genes and beyond (Study III)

The protein structure of AIRE, which consists of several domains that are required to interact with chromatin- and transcription-related proteins, and the ability to activate thousands of silenced genes in any cell type where AIRE is introduced raises the question of what happens to the underlying chromatin landscape during this dramatic shift in the cells' transcriptional program.

Accordingly in Study III, the genome-wide changes in gene expression were integrated with the analyses of chromatin-related molecular events by taking advantage of the AIRE-Tet cell line where AIRE expression can be induced by adding doxycycline to the cells (Liiv et al., 2012). Previously, it was shown that the inhibition of the religation activity of topoisomerase 2 (TOP2) by etoposide could trigger the upregulation of many AIRE-dependent genes (Abramson et al., 2010). Therefore, etoposide treatment was included as one of the experimental conditions in the study.

In Study III, the chromatin immunoprecipitation analyses revealed that the AIRE-expressing AIRE-Tet cells, especially after etoposide treatment become enriched for the histone H2A variant γH2AX at the AIRE target gene promoters (Figure 8A).

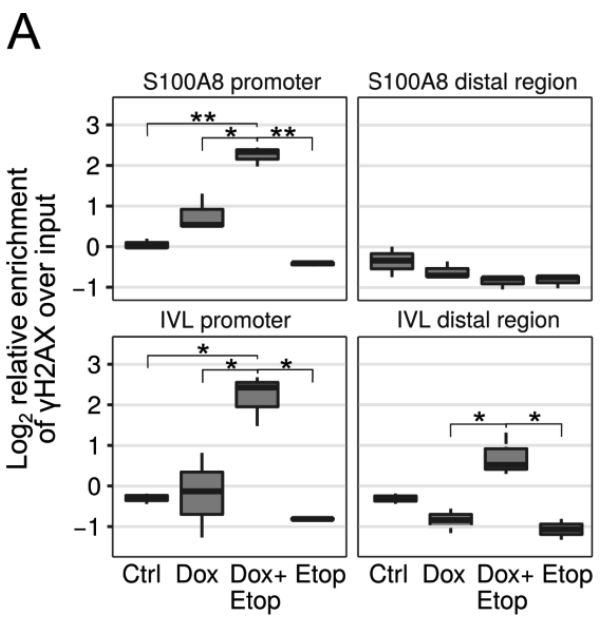

Figure 8.

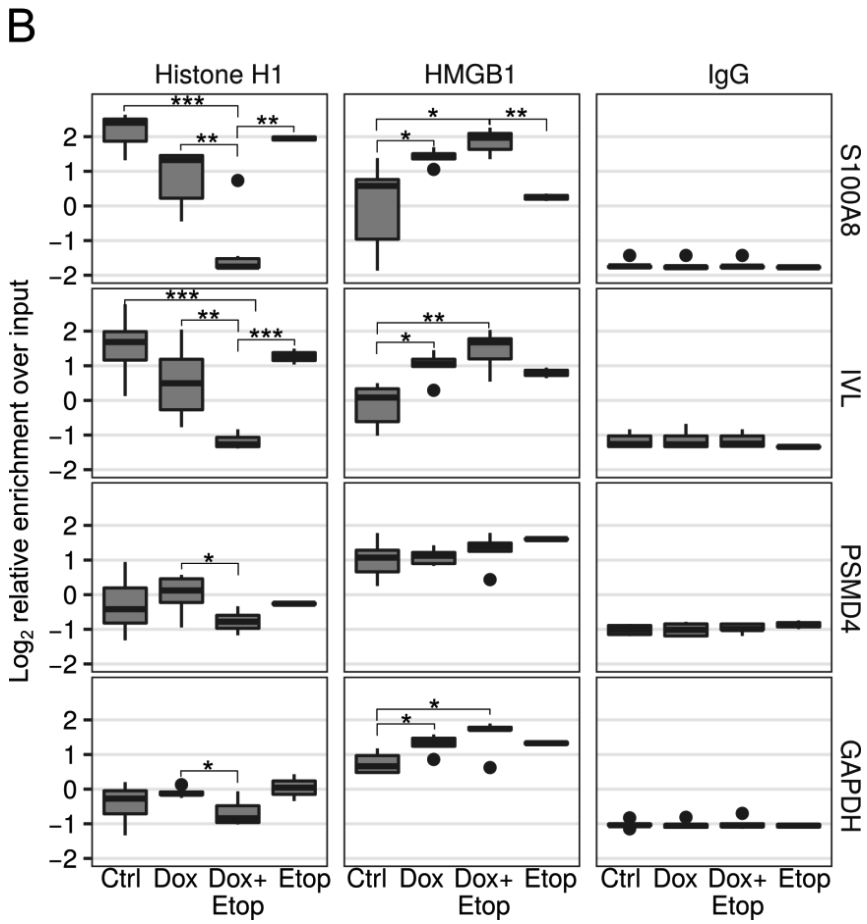

Figure 8. Inhibition of TOP2 with etoposide increases the occurrence of double-stranded DNA breaks and facilitates H1/HMGB1 exchange in AIRE target gene promoters. (A) Assessment of DNA breaks as measured by the enrichment of the histone γH2AX in AIRE-dependent *S100A8* and *IVL* gene promoters and distal control regions in the uninduced/untreated (Ctrl), doxycycline-induced (Dox), doxycycline-induced/etoposide-treated (Dox+Etop), and etoposide-treated (Etop) AIRE-Tet cells using qPCR. The box and whiskers plots show the median and interquartile range of $\log_2$-transformed data from three independent experiments. Whiskers cover data points within a 1.5 × interquartile range. Statistical significance was assessed by two-sample t test (*, $p < 0.05$; **, $p < 0.01$). (B) Analysis of H1 and HMGB1 enrichment at AIRE target gene (*S100A8* and *IVL*) and AIRE-independent gene (*PSMD4* and *GAPDH*) promoters in Ctrl, Dox, Dox+Etop, and Etop AIRE-Tet cells by qPCR. IgG was used as a negative control for ChIP experiments. The box and whiskers plots show the median and interquartile range of $\log_2$-transformed data from five independent experiments. Whiskers cover data points within a 1.5 × interquartile range. Statistical significance was assessed by two-sample t test (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). The figure is adapted from Study III.

Notably, the presence of histone γH2AX signals DNA damage, including DNA breaks that are generated by TOP1 and TOP2 to relieve DNA supercoiling that results from the DNA or RNA polymerase activity (Lowndes and Toh, 2005). The same set of promoter regions underwent an exchange of the linker histone H1, which is considered to be a constituent of silenced chromatin, for the HMGB1 (Figure 8B), an indicator of a more relaxed and accessible chromatin (Cato et al., 2008). These results with the select AIRE target gene loci suggested that AIRE can influence the chromatin structure and prompted to test whether the presence of AIRE coincides with genome-wide changes in chromatin accessibility.

The genome-wide alterations in chromatin structure were analysed by formaldehyde-assisted isolation of regulatory elements (FAIRE) coupled with high-throughput sequencing, which allows the detection of open chromatin regions (Giresi et al., 2007). Approximately 1000 and 1900 genomic regions in either untreated or etoposide-treated AIRE-expressing AIRE-Tet cells (Dox and Dox+Etop, respectively) were found to be differentially accessible compared to the uninduced/untreated control AIRE-Tet cells (Ctrl). In contrast, cells only treated with etoposide (Etop) did not have any FAIRE peaks that were differentially accessible compared with the control AIRE-Tet cells. Interestingly, more than 90% of the FAIRE peaks in Dox and Dox+Etop cells apparently lost their open conformation, which mostly occurred at gene promoters, as assessed by the overlap between the FAIRE peaks and genic/intergenic annotations (Figure 9A).

However, only a small fraction out of all differentially accessible regions in Dox and Dox+Etop cells were close to AIRE-dependent genes (Figure 9B). Instead, a large proportion of differentially accessible regions in Dox and Dox+Etop cells were detected at HEK293-specific binding sites of the chromatin architectural protein CTCF (Figure 9C). Earlier studies have shown that the chromatin loops formed by CTCF can determine the expression genes within and outside of the boundaries of those structures (Nora et al., 2012; Zuin et al., 2014; Tark-Dame et al., 2014). Furthermore, disruption of the CTCF loops can result in misregulation of transcription leading to disease (Katainen et al., 2015). The finding that chromatin accessibility changes at CTCF sites prompted to test whether CTCF-mediated chromatin contacts could be remodelled in the presence of AIRE.

The study concentrated next on a 310-kb genomic region on chr19 containing the carcinoembryonic antigen-related cell adhesion molecule (CEACAM) cluster genes *CEACAM3*, *CEACAM4*, *CEACAM5*, *CEACAM6*, *CEACAM7*, and *CEACAM21,* which are coexpressed with AIRE in human medullary thymic epithelial cells (Cloosen et al., 2007; Pinto et al., 2013). In AIRE-Tet cells, several of the CEACAM genes were differentially expressed in Dox and Dox+Etop cells, and the differentially accessible FAIRE peaks overlapped with CTCF sites within and flanking the gene cluster (Figure 10A).

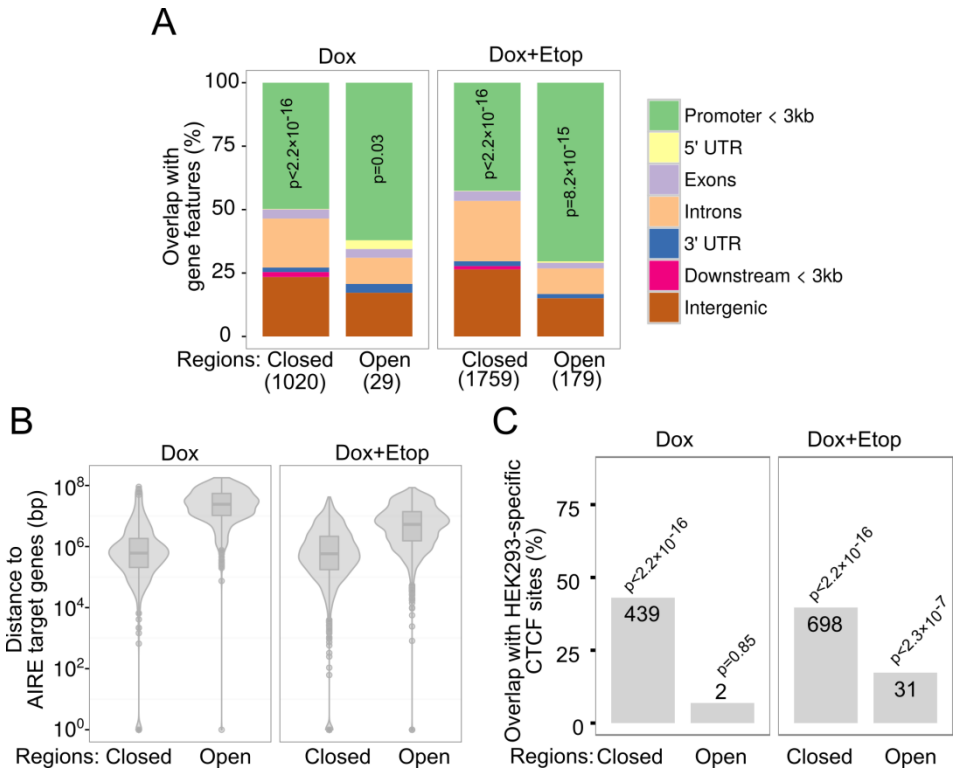

Figure 9. FAIRE-seq reveals AIRE-dependent changes in chromatin accessibility in AIRE-Tet cells. (A) Percentages of FAIRE peaks that overlap with genic and intergenic regions in Dox and Dox+Etop AIRE-Tet cells relative to Ctrl. The x-axis labels refer to the FAIRE peaks that show either loss or gain of chromatin accessibility (Closed and Open, respectively). The numbers of the peaks are shown in brackets. The overlaps between the peaks and promoters up to 3 kb from transcription start site were assessed with the relative distance test implemented in the GenometriCorr package for R statistical software. (B) Distance between differentially expressed genes and closest FAIRE peaks in Dox and Dox+Etop AIRE-Tet cells. The statistical analysis did not determine significant overlap between the genomic regions. (C) Differentially enriched FAIRE peaks show a strong enrichment of CTCF sites in AIRE-Tet cells both with Dox and Dox+Etop treatment. Numbers of FAIRE peaks in the overlap are shown on the bars. The statistical significance of the overlaps between the FAIRE-seq peaks and HEK293-specific CTCF sites was assessed with the relative distance test implemented in the GenometriCorr package for R statistical software. The figure is adapted from Study III.

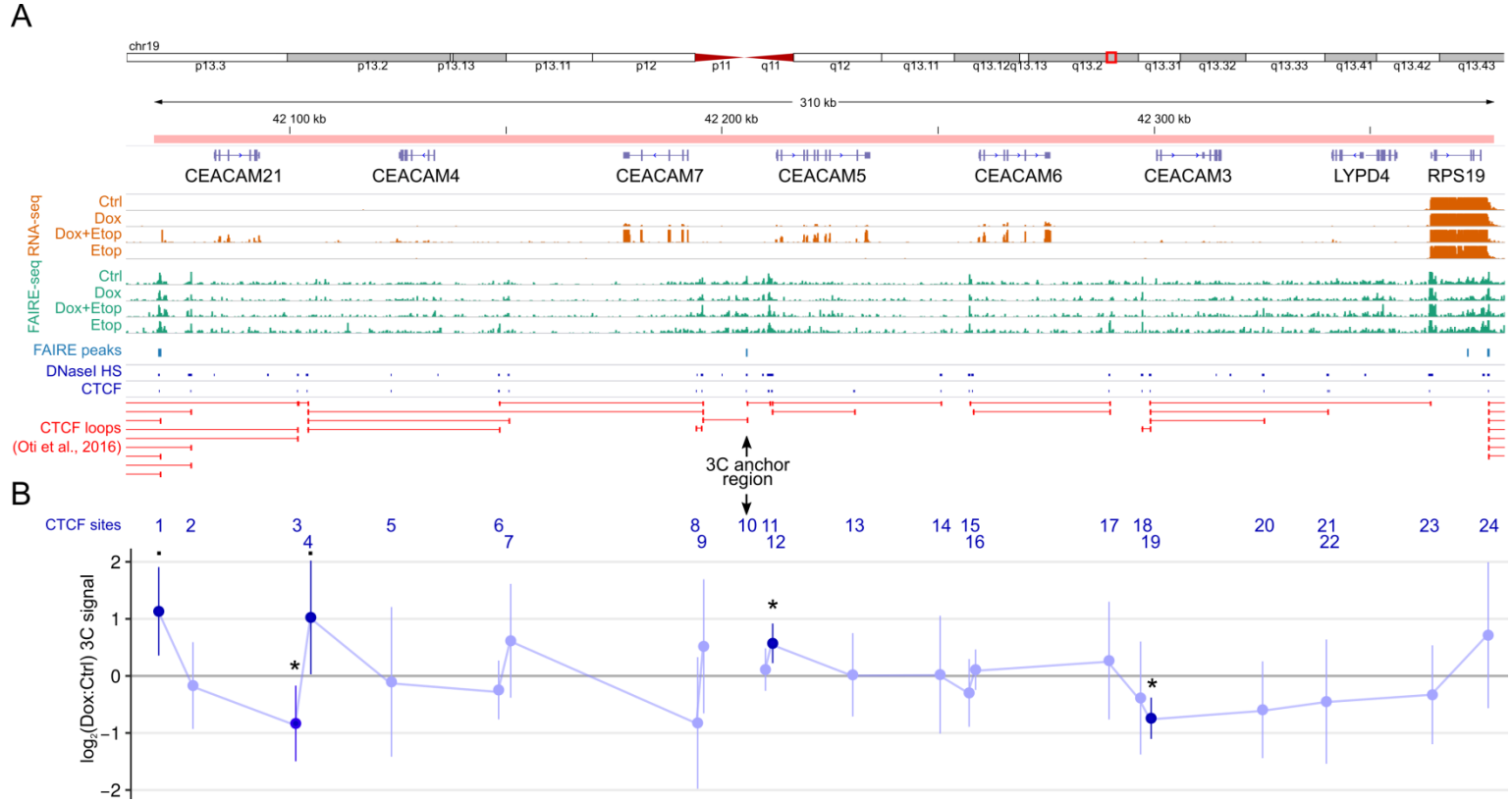

Figure 10. AIRE affects interactions between CTCF binding sites at the CEACAM gene cluster in AIRE-Tet cells. (A) An annotated overview of the genomic region that comprises the CEACAM cluster at chr19:42068524-42378450. The tracks below the gene annotations show the AIRE-Tet-specific gene expression (RNA-seq), chromatin accessibility (FAIRE-seq), differentially enriched FAIRE peaks, HEK293-specific DNase-seq and CTCF ChIP-seq peaks (ENCODE) and computationally predicted CTCF loops (Oti et al., 2016). (B) Chromatin interactions of the 24 CTCF sites in the CEACAM gene cluster depicted in panel A were analysed by 3C. The interaction frequencies were quantified by qPCR and are shown as $\log_2$ fold changes between the signals from Dox and Ctrl AIRE-Tet cells. CTCF site 10 was used as an anchor region in the qPCR assays. The data show the mean ± standard deviation from 5 independent experiments. Statistical significance was assessed by one-sample t-test comparing the $\log_2$ fold changes to the reference value of 0 (▪–$p<0.1$, *–$p<0.05$). The figure is adapted from Study III.

The potential impact of AIRE on CTCF binding and CTCF- mediated chromatin interactions was investigated by chromosome conformation capture (3C) assay, which combines protein cross-linking and proximity ligation of DNA to detect long-range chromatin interactions between genomic loci (Hagège et al., 2007). The FAIRE peak that overlapped with the CTCF binding site 10 between *CEACAM5* and *CEACAM7* genes (Figure 10B), and became less accessible after AIRE induction, was set as the anchor region for the 3C experiments. The chromatin interactions between the anchor region and the remaining 23 CTCF sites within the CEACAM locus were analysed by qPCR using primers that corresponded to Sau3AI-specific restriction fragments that either overlapped or flanked the CTCF sites. The 3C signal from Dox cells was compared to the uninduced AIRE-Tet cells. The results showed that in AIRE-expressing cells, the anchor CTCF site formed stronger chromatin contacts with sites 1 and 4 flanking *CEACAM21* and with site 12 near *CEACAM5*. In addition, contacts with sites 3 and 19 close to the *CEACAM21* and *CEACAM3* genes, respectively, became significantly weaker in AIRE-expressing cells (Figure 10B).

In conclusion, the results suggest that AIRE can affect long-range restructuring of chromatin and influence the partitioning of the CEACAM cluster into chromatin loops.

DISCUSSION

Since its first identification and initial characterisation, the AIRE protein has been associated with many putative functions that contribute to the maintenance of immunological tolerance towards self. AIRE has been linked with the determination of the thymic stromal organisation, thymocyte toleration, regulation of the immune response and apoptosis of parenchymal cells to facilitate cross-presentation of self-antigens (Mathis and Benoist, 2007). The strong expression of AIRE in mTECs, the protein structure and its localisation to the nucleus hinted that AIRE likely binds to chromatin and/or DNA and considering the distinct transcriptomic profile of the mTECs, several studies followed up on the possible role of AIRE in controlling the gene expression program of these cells (Anderson et al., 2002; Derbinski et al., 2001). Indeed, the results causally linked the presence of AIRE to the availability of many, albeit not all, self-antigens in the thymus, which were necessary to avoid autoreactivity by preventing self-recognizing thymocytes from leaving the thymus (Kyewski and Klein, 2006). Despite of the huge effort by many research groups to reveal the molecular mechanisms that make AIRE the ultimate transcription regulator, the promiscuous gene expression and the role of AIRE in establishing it remain to be elucidated.

The molecular characteristics that define the functions of a protein often include post-translational modifications. The findings of Pitkänen et al. (2000 & 2005) showing that the transcription coactivator and acetyltransferase CBP physically interacts with AIRE and that these two proteins cooperate in the regulation of AIRE target genes raised the question about the role of protein acetylation by CBP in shaping AIRE activity. In Study I, the immunoprecipitation and mass-spectrometric analyses found that 12 out 24 lysines were acetylated by CBP and p300. Acetylation occurred at protein sequences that are annotated as the nuclear localisation signals and at the putative DNA-binding SAND domain. These findings suggested that the acetylation of the AIRE NLS could influence the shuttling of AIRE between the cytoplasm and the nucleus, while the acetylation of the SAND domain hinted to a possible effect on protein-DNA or protein-protein interactions.

Although the consequences of the acetylation at the NLS were not analysed in this study, a later report by Incani et al. (2014) concluded, based on the experiments with AIRE acetylation site mutants, that disruption of the acetylation at the NLS prevents the correct localisation of the AIRE protein. Indeed, acetylation of the NLS is a likely mechanism to regulate protein abundance in subcellular compartments given that this phenomenon has been described for many other nuclear proteins (Inuzuka et al., 2012; Matsuzaki et al., 2005; Ventura et al., 2010). Three acetylation sites were detected at lysines 159, 164 and 165, which have been suggested to be part of a second NLS, based on sequence comparison with other nuclear proteins (Saltis et al., 2008) .

However, the function of this sequence has not been tested by mutation analysis, and its role as a second NLS remains speculative.

Study I focused on the acetylation sites in the AIRE SAND domain, which has been proposed to mediate interactions with DNA in other proteins, such as speckled protein 100 kDa (SP100), the nuclear DEAF1-related protein (NUDR), the deformed epidermal autoregulatory factor 1 (DEAF1) and the glucocorticoid modulatory element binding protein (GMEB) (Bottomley et al., 2001; Christensen et al., 1999; Gross and McGinnis, 1996). The DNA-binding property of AIRE has been mapped to the SAND and PHD domains by electromobility shift assays (Kumar et al., 2001; Purohit et al., 2005), but this has not been replicated under more physiological conditions in any other study. Moreover, the AIRE SAND domain lacks the DNA-binding motif K[DN]WK, which is present in other SAND domain-containing proteins (Bottomley et al., 2001). Additionally, Ramsey et al. (2002) reported that mutations in the SAND domain affect the subcellular localisation of AIRE, which together with the aforementioned findings suggest that the AIRE SAND might serve other functions besides interacting with DNA. Therefore, in Study I, acetylation was analysed in the context of protein localisation in AIRE nuclear bodies, protein stability and transcriptional activity.

The protein localisation analysis by immunofluorescence showed that acetylation of AIRE can influence the size and number of AIRE nuclear bodies. Namely, mutations in the AIRE SAND at amino acid positions K243, K245 and K253 that mimic the effect of acetylation rendered AIRE to localise into larger, but fewer nuclear bodies than wild-type AIRE. Similarly, AIRE formed fewer large nuclear bodies after the overexpression of the wild-type p300. In contrast, mutations that mimicked unacetylated lysines in the SAND domain or the overexpression of the acetyltransferase-deficient p300 did not change the size and number of AIRE nuclear bodies compared to the wild-type AIRE. It is important to note that the AIRE protein with mutations mimicking acetylation did not form aggregates, which were observed by Ramsey et al. (2002) in cells expressing the AIRE SAND mutants K221A, K222A, K222E and K253E. This suggests that the mutations used in Study I do not disrupt the correct folding of AIRE. However, the presence of fewer large AIRE nuclear bodies due to acetylation-mimicking mutations or p300 overexpression could be the result of stronger hydrophobic interactions between proteins in the nuclear bodies, which reduce the mobility and availability of AIRE, because the positive charge of the lysine residue is blocked or removed. The AIRE nuclear bodies are bound to the nuclear matrix and regarded as transcriptionally inactive structures (Akiyoshi et al., 2004; Pitkänen et al., 2005; Tao et al., 2006). Notably, the nuclear matrix has been shown to be strongly associated with silenced genomic loci (Croft et al., 1999; Scheuermann et al., 2004; Tanabe et al., 2002). Therefore, the AIRE nuclear bodies are positioned close to the regions where AIRE has the greatest gene activating effect. Hypothetically, the acetylation of AIRE could be required to tighten the contacts with the nuclear matrix, which thereby controls the availability of AIRE and its activity.

In addition to inflicting changes in protein localisation, the overexpression of p300 caused the stabilisation of AIRE protein, which was not observed after the overexpression of the acetyltransferase-deficient p300. Protein stability depends on another post-translational modification called polyubiquitylation, and it has been well established that protein acetylation competes with polyubiquitylation for the same lysine residues (Glozak et al., 2005). Indeed, a report by Akiyoshi et al. (2004) demonstrated that the AIRE protein is polyubiquitylated and that the AIRE nuclear bodies enlarge after the inhibition of the proteasome complex due to the retention of the polyubiquitylated AIRE. Furthermore, a recent study demonstrates that the ubiquitin ligase FBXO3 is responsible for the ubiquitylation of AIRE (Shao et al., 2016). These findings are in line with the protein stability assay results in Study I, which suggest that acetylation by p300 can block possible polyubiquitylation sites in AIRE and thereby prolong its half-life. However, further investigations are required to determine, what is the exact interplay between acetylation and polyubiquitylation and how it affects the availability of AIRE.

Interestingly, the acetylation-mimicking mutations in the SAND domain significantly reduced the transcriptional activity of AIRE, which was analysed by measuring the expression of endogenous AIRE target genes or luciferase reporter activity. In contrast, AIRE containing mutations that mimicked unacetylated lysines retained the full activity of AIRE. By combining these results with the protein localisation data, it could be inferred that the active form of AIRE must be unmodified in the SAND domain and dispersed throughout the nucleus. These findings are in line with a recent study reporting that an active deacetylation by the deacetylase Sirt1 recovers the pool of transcriptionally active AIRE in wild-type mTECs (Chuprin et al., 2015). However, Yoshida et al. (2015) have determined that the transactivating potential of AIRE does depend on the acetylated lysines in the CARD domain, which are a prerequisite for the binding to the bromodomain-containing BRD4 protein that recruits P-TEFb, which eventually activates the poised RNAP II at silenced loci. Thus, gene activation by AIRE seems to require both acetyltransferase and deacetylase activities, albeit targeting different parts of the AIRE protein. How the acetylation and deacetylation events are temporarily and spatially controlled, remains to be explored.

Additionally, microarray analyses showed that the acetylation-mimicking mutations did not repress all AIRE-dependent gene expression, but rather altered the selection of target genes. Furthermore, a sizeable fraction of differentially expressed genes were shared between cells expressing AIRE.K243/253Q and wild-type AIRE compared to the control cell line, which lacked AIRE expression. These results suggest that protein acetylation could be a mechanism that helps to shape the AIRE target gene repertoire. Importantly, single cell PCR and RNA-seq experiments have determined that individual AIRE-positive mTECs have apparently distinct and stochastic transcriptional programs (Brennecke et al., 2015; Meredith et al., 2015; Sansom et al., 2014). One possible explanation for the observed variability of the gene expression could be the level of

acetylated AIRE in the individual cells, but the post-translational modifications of AIRE in primary thymic stromal cells have not been analysed and further testing would be required to confirm or dispute the role of AIRE acetylation in TRA gene activation *in vivo*. In summary, the interaction of AIRE with the ubiquitously expressed transcription coactivators and acetyltransferases CBP/p300 has a broad effect on the function of AIRE.

In Study II, one of the main aims was to structurally and functionally characterise the AIRE PHD2 domain by comparing the intact domain to the APS1 mutation C446G containing PHD2 domain. The results showed that the structure of PHD2 resembles that of a typical PHD domain and the C446G mutation completely destroys its fold leading to protein aggregation in the cytosol and loss of TRA expression. Interestingly, *in vitro* pull-down experiments showed that AIRE C446G mutant could still bind histone H3, which is in agreement with reports showing that the two AIRE PHD domains function independently of each other (Koh et al., 2008; Org et al., 2008). The structural analysis in Study II confutes the computational predictions, which suggested that the PHD2 domain belongs to the really interesting new gene (RING)-type zinc finger domains (Saltis et al., 2008). As a RING-type zinc finger, AIRE PHD2 should have possessed ubiquitin ligase activity, but this has never been observed. However, AIRE has been shown to function as a E3 ligase, but this activity has been mapped to the PHD1 domain instead (Uchida et al., 2004). Yet, several independent reports have concluded that the PHD2 domain is required for transcriptional control of AIRE target genes (Meloni et al., 2008; Uchida et al., 2004; Yang et al., 2013). In the reporter activation experiments conducted by Uchida et al. (2004) and Meloni et al. (2008), the analysed PHD2 mutations disrupted the zinc ion-coordinating cysteines similarly to the C446G mutation in Study II. Although the two studies did not analyse protein localisation, the observed low or absent transactivation in these studies could be attributed to the protein aggregation in the cytoplasm, as seen in Study II. Nevertheless, the PHD2 has a more direct influence on AIRE-dependent expression, which was demonstrated by Yang et al. (2013) using a PHD2 deletion knock-out mouse model. Specifically, AIRE without its PHD2 was still able to localise into the nucleus, but it could not bind many chromatin- and transcription-related proteins that would facilitate AIRE-dependent transcription. In conclusion, growing evidence supports the role of the AIRE PHD2 domain in gene activation. Future studies addressing the AIRE-regulated transcription, would need take advantage of the structural data on AIRE protein domains to design mutation analysis experiments that can unravel the molecular events leading to TRA expression.

A substantial part of Study III addressed the question of what happens to the chromatin landscape near AIRE target gene loci once AIRE starts to express in the cells. First, ChIP experiments demonstrated that the promoter regions of AIRE-dependent genes are enriched for the histone variant γH2AX, which acts as a marker of double-stranded DNA breaks, and HMGB1, which replaces the linker histone H1 at loci that become transcriptionally active (Cato et al., 2008;

Lowndes and Toh, 2005). Collectively, these results show that the AIRE-regulated transcription correlates with a more permissive chromatin structure at target gene loci. These data complement earlier findings that show an increase in histone acetylation and histone H3 lysine 4 methylation levels, both of which are associated with transcriptionally active chromatin (Org et al., 2009).

In addition, the ChIP enrichment of γH2AX and HMGB1 and the loss of histone H1 were amplified by the treatment with the TOP2-specific inhibitor etoposide suggesting that the altered chromatin structure at AIRE target gene loci heavily depends on the activity of TOP2. TOP2A is a major interaction partner of AIRE, and the inhibition of its religation activity by etoposide has been shown to activate several AIRE-dependent genes in HEK293 cells, which has highlighted the role of DNA breaks in AIRE-regulated transcription (Abramson et al., 2010). Notably, Abramson et al. also reported that the AIRE-positive $MHCII^{hi}$-expressing mTECs in wild-type mice have more DNA breaks compared to AIRE knock-out littermates. In summary, the analysis of etoposide treated AIRE-expressing cells reveals a strong link between DNA breaks, AIRE-dependent transcription and chromatin structure.

In Study III, the FAIRE-seq approach was used to test, whether besides the altered enrichment of chromatin proteins, it would be possible to detect changes in open chromatin regions, which would identify transcription factor binding sites that are utilised in AIRE-expressing AIRE-Tet cells. The analysis revealed that many promoter regions become less open in AIRE-expressing cells, but interestingly, relatively few changes were detected within or near AIRE-regulated genes. Although FAIRE-seq does not identify the specific transcription factors, it could hint to some molecular events that are put into motion by AIRE. First, based on published ChIP-seq data, AIRE has been shown to bind to almost all promoters irrespective of the transcriptional status of the genes (Giraud et al., 2012), and therefore, it would be likely that the open chromatin regions are occupied by AIRE itself. This possibility, however, was not addressed in Study III. Second, open chromatin could be taken over by histones or other chromatin-related proteins.

Indeed, a large fraction of AIRE-related structural changes in chromatin occurred at CTCF binding sites that are considered to define genomic boundaries between active and silenced chromosomal regions (Nora et al., 2012; Zuin et al., 2014; Tark-Dame et al., 2014). This observation raised the notion that AIRE-dependent transcriptional control requires the rearrangement of chromatin interactions between distant regions. It would be feasible to speculate that the genomic boundaries have to be redefined in AIRE-expressing cells to allow the transcriptional machinery to enter and interact with the silenced AIRE target gene regions. This hypothesis was analysed with the chromosome conformation capture assay, which showed that the CTCF-mediated chromatin interaction frequencies change at least in the CEACAM gene cluster in AIRE-expressing cells compared to AIRE-negative AIRE-Tet cells. These structural changes at CTCF sites could affect promoter-enhancer interactions and modulate the transcriptional permissiveness of AIRE target genes. However,

further exploration of the epigenetic profiles and genome-wide chromatin interaction maps are required to comprehensively determine the interplay between the underlying chromatin structure and the transcriptional program of AIRE-expressing cells.

In conclusion, the current thesis highlights the role of post-translational acetylation in the stability and transcriptional activity of AIRE. Additionally, structural analysis in combination with protein localisation studies showed that the C446G mutation in the AIRE PHD2 domain causes protein aggregation, which prevents AIRE from entering the cell nucleus, thus effectively blocking AIRE-dependent transcription. Interestingly, the intact AIRE protein not only controls the activation of thousands of silenced genes, but also affects chromatin structure throughout the genome. However, more in-depth studies are needed to elucidate the cause-and-effect relationships that exist between the chromatin conformation and AIRE-regulated gene activation. Presumably, it would lay a solid basis to fully understand the still enigmatic molecular mechanisms of promiscuous gene expression and immune tolerance.

CONCLUSIONS

1. The AIRE protein is extensively acetylated by p300 in HEK293 cells. The acetylated lysines are part of critical sequences in the nuclear localisation signal and the SAND domain. The acetylation affects the stability of the AIRE protein and the size and number of AIRE nuclear bodies. Mutation analyses highlight the role of the acetylated lysines in the transcriptional control of AIRE-dependent genes.
2. The previously identified APS1-associated mutation C446G in AIRE PHD2 domain destroys the structure of the PHD finger and hinders the correct subcellular localisation of AIRE into the nucleus. Although the C446G mutation permits the binding of AIRE to histone H3, which is a critical interaction for the transactivation of AIRE-dependent genes, this mutation abolishes the capacity of AIRE to activate its target genes. This suggests that the correct folding of the PHD2 zinc finger plays a critical role in establishing AIRE-regulated gene expression.
3. The activation of AIRE target genes coincides with changes in chromatin accessibility at promoter regions of AIRE-independent genes: apparently, chromatin accessibility decreases whereas a large fraction of those sites overlaps with CTCF binding sites. The interaction frequency between CTCF binding sites at the CEACAM gene cluster is altered in AIRE-expressing cells. Jointly, this means that AIRE affects large-scale rearrangements of the chromatin structure at AIRE target gene loci.

REFERENCES

Aaltonen, J., Björses, P., Perheentupa, J., Horelli–Kuitunen, N., Palotie, A., Peltonen, L., Lee, Y.S., Francis, F., Henning, S., Thiel, C., et al. (1997). An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. Nat. Genet. *17*, 399–403.

Abramson, J., and Goldfarb, Y. (2016). AIRE: From promiscuous molecular partnerships to promiscuous gene expression. Eur. J. Immunol. *46*, 22–33.

Abramson, J., Giraud, M., Benoist, C., and Mathis, D. (2010). Aire's Partners in the Molecular Control of Immunological Tolerance. Cell *140*, 123–135.

Ahonen, P. (1985). Autoimmune polyendocrinopathy – candidosis – ectodermal dystrophy (APECED): autosomal recessive inheritance. Clin. Genet. *27*, 535–542.

Akiyoshi, H., Hatakeyama, S., Pitkänen, J., Mouri, Y., Doucas, V., Kudoh, J., Tsurugaya, K., Uchida, D., Matsushima, A., Oshikawa, K., et al. (2004). Subcellular Expression of Autoimmune Regulator Is Organized in a Spatiotemporal Manner. J. Biol. Chem. *279*, 33984–33991.

Anderson, M.S., and Su, M.A. (2016). AIRE expands: new roles in immune tolerance and beyond. Nat. Rev. Immunol. *16*, 247–258.

Anderson, M.S., Venanzi, E.S., Klein, L., Chen, Z., Berzins, S.P., Turley, S.J., von Boehmer, H., Bronson, R., Dierich, A., Benoist, C., et al. (2002). Projection of an immunological self shadow within the thymus by the aire protein. Science (80–.). *298*, 1395–1401.

Antonia, S.J., Geiger, T., Miller, J., and Flavell, R.A. (1995). Mechanisms of immune tolerance induction through the thymic expression of a peripheral tissue-specific protein. Int. Immunol. *7*, 715–725.

Apostolou, I., Sarukhan, A., Klein, L., and von Boehmer, H. (2002). Origin of regulatory T cells with known specificity for antigen. Nat. Immunol. *3*, 756–763.

Asano, M., Toda, M., Sakaguchi, N., and Sakaguchi, S. (1996). Autoimmune disease as a consequence of developmental abnormality of a T cell subpopulation. J. Exp. Med. *184*, 387–396.

Baranello, L., Wojtowicz, D., Cui, K., Devaiah, B.N., Chung, H.-J.J., Chan-Salis, K.Y., Guha, R., Wilson, K., Zhang, X., Zhang, H., et al. (2016). RNA Polymerase II Regulates Topoisomerase 1 Activity to Favor Efficient Transcription. Cell *165*, 357–371.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-Resolution Profiling of Histone Methylations in the Human Genome. Cell *129*, 823–837.

Björses, P., Pelto-Huikko, M., Kaukonen, J., Aaltonen, J., Peltonen, L., and Ulmanen, I. (1999). Localization of the APECED protein in distinct nuclear structures. Hum. Mol. Genet. *8*, 259–266.

Black, J.C., Mosley, A., Kitada, T., Washburn, M., and Carey, M. (2008). The SIRT2 Deacetylase Regulates Autoacetylation of p300. Mol. Cell *32*, 449–455.

Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Müller, S., Eils, R., Cremer, C., Speicher, M.R., et al. (2005). Three-Dimensional Maps of All Chromosomes in Human Male Fibroblast Nuclei and Prometaphase Rosettes. PLoS Biol. *3*, 826–842.

Bottomley, M.J., Collard, M.W., Huggenvik, J.I., Liu, Z., Gibson, T.J., and Sattler, M. (2001). The SAND domain structure defines a novel DNA-binding fold in transcriptional regulation. Nat. Struct. Biol. *8*, 626–633.

Branco, M.R., and Pombo, A. (2006). Intermingling of Chromosome Territories in Interphase Suggests Role in Translocations and Transcription-Dependent Associations. PLoS Biol. *4*, 780–788.

Brennecke, P., Reyes, A., Pinto, S., Rattay, K., Nguyen, M., Küchler, R., Huber, W., Kyewski, B., and Steinmetz, L.M. (2015). Single-cell transcriptome analysis reveals coordinated ectopic gene-expression patterns in medullary thymic epithelial cells. Nat. Immunol. *16*, 933–941.

Bridger, J.M., Kalla, C., Wodrich, H., Weitz, S., King, J.A., Khazaie, K., Kräusslich, H.G., and Lichter, P. (2005). Nuclear RNAs confined to a reticular compartment between chromosome territories. Exp. Cell Res. *302*, 180–193.

Brill, S.J., DiNardo, S., Voelkel-Meiman, K., and Sternglanz, R. (1987). Need for DNA topoisomerase activity as a swivel for DNA replication for transcription of ribosomal RNA. Nature *326*, 414–416.

Brunet, A., Bonni, A., Zigmond, M.J., Lin, M.Z., Juo, P., Hu, L.S., Anderson, M.J., Arden, K.C., Blenis, J., and Greenberg, M.E. (1999). Akt Promotes Cell Survival by Phosphorylating and Inhibiting a Forkhead Transcription Factor. Cell *96*, 857–868.

Callén, E., Jankovic, M., Wong, N., Zha, S., Chen, H.-T., Difilippantonio, S., Di Virgilio, M., Heidkamp, G., Alt, F.W., Nussenzweig, A., et al. (2009). Essential Role for DNA-PKcs in DNA Double-Strand Break Repair and Apoptosis in ATM-Deficient Lymphocytes. Mol. Cell *34*, 285–297.

Campos, E.I., and Reinberg, D. (2009). Histones: Annotating Chromatin. Annu. Rev. Genet. *43*, 559–599.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. Science (80–.). *298*, 1039–1043.

Cato, L., Stott, K., Watson, M., and Thomas, J.O. (2008). The Interaction of HMGB1 and Linker Histones Occurs Through their Acidic and Basic Tails. J. Mol. Biol. *384*, 1262–1272.

Cetani, F., Barbesino, G., Borsari, S., Pardi, E., Cianferotti, L., Pinchera, A., and Marcocci, C. (2001). A novel mutation of the autoimmune regulator gene in an Italian kindred with autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy, acting in a dominant fashion and strongly cosegregating with hypothyroid autoimmune thyroiditis. J. Clin. Endocrinol. Metab. *86*, 4747–4752.

Champoux, J.J. (2001). DNA Topoisomerases: Structure, Function, and Mechanism. Annu. Rev. Biochem. *70*, 369–413.

Chatila, T.A., Blaeser, F., Ho, N., Lederman, H.M., Voulgaropoulos, C., Helms, C., and Bowcock, A.M. (2000). JM2, encoding a fork head-related protein, is mutated in X-linked autoimmunity-allergic disregulation syndrome. J. Clin. Invest. *106*, 75–81.

Chen, C., Nott, T.J., Jin, J., and Pawson, T. (2011). Deciphering arginine methylation: Tudor tells the tale. Nat. Rev. Mol. Cell Biol. *12*, 629–642.

Cheng, M.H., and Anderson, M.S. (2012). Monogenic Autoimmunity. Annu. Rev. Immunol. *30*, 393–427.

Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J. V, and Mann, M. (2009). Lysine Acetylation Targets Protein Complexes and Co-Regulates Major Cellular Functions. Science (80–.). *325*, 834–840.

Christensen, J., Cotmore, S.F., and Tattersall, P. (1999). Two new members of the emerging KDWK family of combinatorial transcription modulators bind as a heterodimer to flexibly spaced PuCGPy half-sites. Mol. Cell. Biol. *19*, 7741–7750.

Chuprin, A., Avin, A., Goldfarb, Y., Herzig, Y., Levi, B., Jacob, A., Sela, A., Katz, S., Grossman, M., Guyon, C., et al. (2015). The deacetylase Sirt1 is an essential regulator of Aire-mediated induction of central immunological tolerance. Nat. Immunol. *16*, 737–745.

Cisse, I.I., Izeddin, I., Causse, S.Z., Boudarene, L., Senecal, A., Muresan, L., Dugast-Darzacq, C., Hajj, B., Dahan, M., and Darzacq, X. (2013). Real-Time Dynamics of RNA Polymerase II Clustering in Live Human Cells. Science (80–.). *341*, 664–667.

Cloosen, S., Arnold, J., Thio, M., Bos, G.M.J., Kyewski, B., and Germeraad, W.T. V (2007). Expression of tumor-associated differentiation antigens, MUC1 glycoforms and CEA, in human thymic epithelial cells: Implications for self-tolerance and tumor therapy. Cancer Res. *67*, 3919–3926.

Cremer, T., and Cremer, M. (2010). Chromosome Territories. Cold Spring Harb. Perspect. Biol. *2*, 1–22.

Croft, J.A., Bridger, J.M., Boyle, S., Perry, P., Teague, P., and Bickmore, W.A. (1999). Differences in the localization and morphology of chromosomes in the human nucleus. J. Cell Biol. *145*, 1119–1131.

Dahl, J.A., and Collas, P. (2007). A quick and quantitative chromatin immunoprecipitation assay for small cell samples. Front. Biosci. *12*, 4925–4931.

Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., and Richmond, T.J. (2002). Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9Å Resolution. J. Mol. Biol. *319*, 1097–1113.

de Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M., et al. (2004). Polycomb Group Proteins Ring1A/B Link Ubiquitylation of Histone H2A to Heritable Gene Silencing and X Inactivation. Dev. Cell *7*, 663–676.

Derbinski, J., Gäbler, J., Brors, B., Tierling, S., Jonnakuty, S., Hergenhahn, M., Peltonen, L., Walter, J., and Kyewski, B. (2005). Promiscuous gene expression in thymic epithelial cells is regulated at multiple levels. J. Exp. Med. *202*, 33–45.

Derbinski, J., Schulte, A., Kyewski, B., and Klein, L. (2001). Promiscuous gene expression in medullary thymic epithelial cells mirrors the peripheral self. Nat. Immunol. *2*, 1032–1039.

Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K., and Zhou, M.M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. Nature *399*, 491–496.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Doucas, V., Tini, M., Egan, D.A., and Evans, R.M. (1999). Modulation of CREB binding protein function by the promyelocytic (PML) oncoprotein suggests a role for nuclear bodies in hormone signaling. Proc. Natl. Acad. Sci. U. S. A. *96*, 2627–2632.

Durocher, D., and Jackson, S.P. (2001). DNA-PK, ATM and ATR as sensors of DNA damage: variations on a theme? Curr. Opin. Cell Biol. *13*, 225–231.

Fan, J.Y., Rangasamy, D., Luger, K., and Tremethick, D.J. (2004). H2A.Z alters the nucleosome surface to promote HP1α-mediated chromatin fiber folding. Mol. Cell *16*, 655–661.

Favorov, A., Mularoni, L., Cope, L.M., Medvedeva, Y., Mironov, A.A., Makeev, V.J., and Wheelan, S.J. (2012). Exploring Massive, Genome Scale Datasets with the GenometriCorr Package. PLoS Comput. Biol. *8*, e1002529.

Federico, C., Cantarella, C.D., Di Mare, P., Tosi, S., and Saccone, S. (2008). The radial arrangement of the human chromosome 7 in the lymphocyte cell nucleus is associated with chromosomal band gene density. Chromosoma *117*, 399–410.

Ferguson, B.J., Alexander, C., Rossi, S.W., Liiv, I., Rebane, A., Worth, C.L., Wong, J., Laan, M., Peterson, P., Jenkinson, E.J., et al. (2008). AIRE's CARD Revealed, a New Structure for Central Tolerance Provokes Transcriptional Plasticity. J. Biol. Chem. *283*, 1723–1731.

Ferrai, C., Xie, S.Q., Luraghi, P., Munari, D., Ramirez, F., Branco, M.R., Pombo, A., and Crippa, M.P. (2010). Poised transcription factories prime silent uPA gene prior to activation. PLoS Biol. *8*, e1000270.

Fisher, G.H., Rosenberg, F.J., Straus, S.E., Dale, J.K., Middelton, L.A., Lin, A.Y., Strober, W., Lenardo, M.J., and Puck, J.M. (1995). Dominant interfering fas gene mutations impair apoptosis in a human autoimmune lymphoproliferative syndrome. Cell *81*, 935–946.

Fontenot, J.D., Rasmussen, J.P., Williams, L.M., Dooley, J.L., Farr, A.G., and Rudensky, A.Y. (2005). Regulatory T cell lineage specification by the forkhead transcription factor foxp3. Immunity *22*, 329–341.

Gaetani, M., Matafora, V., Saare, M., Spiliotopoulos, D., Mollica, L., Quilici, G., Chignola, F., Mannella, V., Zucchelli, C., Peterson, P., et al. (2012). AIRE-PHD fingers are structural hubs to maintain the integrity of chromatin-associated interactome. Nucleic Acids Res. *40*, 11756–11768.

Gale, K.C., and Osheroff, N. (1992). Intrinsic Intermolecular DNA Ligation Activity of Eukaryotic Topoisomerase II. J. Biol. Chem. *267*, 12090–12097.

Gardner, J.M., DeVoss, J.J., Friedman, R.S., Wong, D.J., Tan, Y.X., Zhou, X., Johannes, K.P., Su, M.A., Chang, H.Y., Krummel, M.F., et al. (2008). Deletional Tolerance Mediated by Extrathymic Aire-Expressing Cells. Science (80–.). *321*, 843–847.

Gartenberg, M.R., and Wang, J.C. (1992). Positive supercoiling of DNA greatly diminishes mRNA synthesis in yeast. Proc. Natl. Acad. Sci. *89*, 11461–11465.

Gaszner, M., and Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. Nat. Rev. Genet. *7*, 703–713.

Giraud, M., Jmari, N., Du, L., Carallis, F., Nieland, T.J.F., Perez-Campo, F.M., Bensaude, O., Root, D.E., Hacohen, N., Mathis, D., et al. (2014). An RNAi screen for Aire cofactors reveals a role for Hnrnpl in polymerase release and Aire-activated ectopic transcription. Proc. Natl. Acad. Sci. *111*, 1491–1496.

Giraud, M., Yoshida, H., Abramson, J., Rahl, P.B., Young, R. a., Mathis, D., and Benoist, C. (2012). Aire unleashes stalled RNA polymerase to induce ectopic gene expression in thymic epithelial cells. Proc. Natl. Acad. Sci. *109*, 535–540.

Giresi, P.G., Kim, J., McDaniell, R.M., Iyer, V.R., and Lieb, J.D. (2007). FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. Genome Res. *17*, 877–885.

Glozak, M. a, Sengupta, N., Zhang, X., and Seto, E. (2005). Acetylation and deacetylation of non-histone proteins. Gene *363*, 15–23.

Goto, T., and Wang, J.C. (1985). Cloning of yeast TOP1, the gene encoding DNA topoisomerase I, and construction of mutants defective in both DNA topoisomerase I and DNA topoisomerase II. Proc. Natl. Acad. Sci. *82*, 7178–7182.

Grandien, A., Fucs, R., Nobrega, A., Andersson, J., and Coutinho, A. (1994). Negative selection of multireactive B cell clones in normal adult mice. Eur. J. Immunol. *24*, 1345–1352.

Gross, C.T., and McGinnis, W. (1996). DEAF-1, a novel protein that binds an essential region in a Deformed response element. EMBO J. *15*, 1961–1970.

Gu, W., and Roeder, R.G. (1997). Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. Cell *90*, 595–606.

Guo, Y., Monahan, K., Wu, H., Gertz, J., Varley, K.E., Li, W., Myers, R.M., Maniatis, T., and Wu, Q. (2012). CTCF/cohesin-mediated DNA looping is required for protocadherin promoter choice. Proc. Natl. Acad. Sci. *109*, 21081–21086.

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

Haljasorg, U., Bichele, R., Saare, M., Guha, M., Maslovskaja, J., Kõnd, K., Remm, A., Pihlap, M., Tomson, L., Kisand, K., et al. (2015). A highly conserved NF-κB-responsive enhancer is critical for thymic expression of Aire in mice. Eur. J. Immunol. *45*, 3246–3256.

Halonen, M., Kangas, H., Rüppell, T., Ilmarinen, T., Ollila, J., Kolmer, M., Vihinen, M., Palvimo, J., Saarela, J., Ulmanen, I., et al. (2004). APECED-causing mutations in AIRE reveal the functional domains of the protein. Hum. Mutat. *23*, 245–257.

Halonen, M., Pelto-Huikko, M., Eskelin, P., Peltonen, L., Ulmanen, I., and Kolmer, M. (2001). Subcellular location and expression pattern of autoimmune regulator (Aire), the mouse orthologue for human gene defective in autoimmune polyendocrinopathy candidiasis ectodermal dystrophy (APECED). J. Histochem. Cytochem. *49*, 197–208.

Hecht, A.T., Laroche, T., Strahl-Bolsinger, S., Gasser, S.M., and Grunstein, M. (1995). Histone H3 and H4 N termini interact with SIR3 and SIR4 proteins: A molecular model for the formation of heterochromatin in yeast. Cell *80*, 583–592.

Heino, M., Peterson, P., Kudoh, J., Nagamine, K., Lagerstedt, A., Ovod, V., Ranki, A., Rantala, I., Nieminen, M., Tuukkanen, J., et al. (1999). Autoimmune Regulator Is Expressed in the Cells Regulating Immune Tolerance in Thymus Medulla. Biochem. Biophys. Res. Commun. *257*, 821–825.

Heino, M., Peterson, P., Sillanpää, N., Guérin, S., Wu, L., Anderson, G., Scott, H.S., Antonarakis, S.E., Kudoh, J., Shimizu, N., et al. (2000). RNA and protein expression of the murine autoimmune regulator gene (Aire) in normal, RelB-deficient and in NOD mouse. Eur. J. Immunol. *30*, 1884–1893.

Hendriks, I.A., and Vertegaal, A.C.O. (2016). A comprehensive compilation of SUMO proteomics. Nat. Rev. Mol. Cell Biol. *17*, 581–595.

Herzig, Y., Nevo, S., Bornstein, C., Brezis, M.R., Ben-Hur, S., Shkedy, A., Eisenberg-Bord, M., Levi, B., Delacher, M., Goldfarb, Y., et al. (2016). Transcriptional programs that control expression of the autoimmune regulator gene Aire. Nat. Immunol. *18*, 161–172.

Huang, H.-S., Allen, J.A., Mabb, A.M., King, I.F., Miriyala, J., Taylor-Blake, B., Sciaky, N., Dutton, J.W., Lee, H.-M., Chen, X., et al. (2011). Topoisomerase inhibitors unsilence the dormant allele of Ube3a in neurons. Nature *481*, 185–189.

Husain, A., Begum, N.A., Taniguchi, T., Taniguchi, H., Kobayashi, M., and Honjo, T. (2016). Chromatin remodeller SMARCA4 recruits topoisomerase 1 and suppresses transcription-associated genomic instability. Nat. Commun. *7*, 10549.

Husebye, E.S., Perheentupa, J., Rautemaa, R., Ka, O., and Kämpe, O. (2009). Clinical manifestations and management of patients with autoimmune polyendocrine syndrome type I. J. Intern. Med. *265*, 514–529.

Iborra, F.J., Pombo, A., Jackson, D.A., and Cook, P.R. (1996). Active RNA polymerases are localized within discrete transcription "factories" in human nuclei. J. Cell Sci. *109*, 1427–1436.

Ilmarinen, T., Melén, K., Kangas, H., Julkunen, I., Ulmanen, I., and Eskelin, P. (2006). The monopartite nuclear localization signal of autoimmune regulator mediates its nuclear import and interaction with multiple importin α molecules. FEBS J. *273*, 315–324.

Imhof, A., Yang, X.-J., Ogryzko, V. V, Nakatani, Y., Wolffe, A.P., and Ge, H. (1997). Acetylation of general transcription factors by histone acetyltransferases. Curr. Biol. *7*, 689–692.

Incani, F., Serra, M.L., Meloni, A., Cossu, C., Saba, L., Cabras, T., Messana, I., and Rosatelli, M.C. (2014). AIRE acetylation and deacetylation: effect on protein stability and transactivation activity. J. Biomed. Sci. *21*, 85.

Inuzuka, H., Gao, D., Finley, L.W.S., Yang, W., Wan, L., Fukushima, H., Chin, Y.R., Zhai, B., Shaik, S., Lau, A.W., et al. (2012). Acetylation-dependent regulation of Skp2 function. Cell *150*, 179–193.

Isaac, A., Wilcox, K.W., and Taylor, J.L. (2006). SP100B, a repressor of gene expression preferentially binds to DNA with unmethylated CpGs. J. Cell. Biochem. *98*, 1106–1122.

Ito, A., Kawaguchi, Y., Lai, C.-H., Kovacs, J.J., Higashimoto, Y., Appella, E., and Yao, T.-P. (2002). MDM2–HDAC1-mediated deacetylation of p53 is required for its degradation. EMBO J. *21*, 6236–6245.

Itoh, M., Takahashi, T., Sakaguchi, N., Kuniyasu, Y., Shimizu, J., Otsuka, F., and Sakaguchi, S. (1999). Thymus and autoimmunity: production of CD25+CD4+ naturally anergic and suppressive T cells as a key function of the thymus in maintaining immunologic self-tolerance. J. Immunol. *162*, 5317–5326.

Ivanov, G.S., Ivanova, T., Kurash, J., Ivanov, A., Chuikov, S., Gizatullin, F., Herrera-Medina, E.M., Rauscher, F., Reinberg, D., and Barlev, N.A. (2007). Methylation-Acetylation Interplay Activates p53 in Response to DNA Damage. Mol. Cell. Biol. *27*, 6756–6769.

Jensik, P.J., Vargas, J.D., Reardon, S.N., Rajamanickam, S., Huggenvik, J.I., and Collard, M.W. (2014). DEAF1 Binds Unmethylated and Variably Spaced CpG Dinucleotide Motifs. PLoS One *9*, e115908.

Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.-A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature *503*, 290–294.

Jing, E., Gesta, S., and Kahn, C.R. (2007). SIRT2 Regulates Adipocyte Differentiation through FoxO1 Acetylation/Deacetylation. Cell Metab. *6*, 105–114.

Johnnidis, J.B., Venanzi, E.S., Taxman, D.J., Ting, J.P.-Y., Benoist, C.O., and Mathis, D.J. (2005). Chromosomal clustering of genes controlled by the aire transcription factor. Proc. Natl. Acad. Sci. *102*, 7233–7238.

Jordan, M.S., Boesteanu, A., Reed, A.J., Petrone, A.L., Holenbeck, A.E., Lerman, M.A., Naji, A., and Caton, A.J. (2001). Thymic selection of CD4+CD25+ regulatory T cells induced by an agonist self-peptide. Nat. Immunol. *2*, 301–306.

Kadauke, S., and Blobel, G.A. (2009). Chromatin loops in gene regulation. Biochim. Biophys. Acta – Gene Regul. Mech. *1789*, 17–25.

Karanam, B., Jiang, L., Wang, L., Kelleher, N.L., and Cole, P.A. (2006). Kinetic and mass spectrometric analysis of p300 histone acetyltransferase domain autoacetylation. J. Biol. Chem. *281*, 40292–40301.

Katainen, R., Dave, K., Pitkänen, E., Palin, K., Kivioja, T., Välimäki, N., Gylfe, A.E., Ristolainen, H., Hänninen, U.A., Cajuso, T., et al. (2015). CTCF/cohesin-binding sites are frequently mutated in cancer. Nat. Genet. *47*, 818–821.

Kelly, A.E., Ghenoiu, C., Xue, J.Z., Zierhut, C., Kimura, H., and Funabiki, H. (2010). Survivin Reads Phosphorylated Histone H3 Threonine 3 to Activate the Mitotic Kinase Aurora B. Science (80–.). *330*, 235–239.

Kim, R.A., and Wang, J.C. (1989). Function of DNA topoisomerases as replication swivels in Saccharomyces cerevisiae. J. Mol. Biol. *208*, 257–267.

King, I.F., Yandava, C.N., Mabb, A.M., Hsiao, J.S., Huang, H.-S., Pearson, B.L., Calabrese, J.M., Starmer, J., Parker, J.S., Magnuson, T., et al. (2013). Topoisomerases facilitate transcription of long genes linked to autism. Nature *501*, 58–62.

King, I.F.G., Francis, N.J., and Kingston, R.E. (2002). Native and Recombinant Polycomb Group Complexes Establish a Selective Block to Template Accessibility To Repress Transcription In Vitro. Mol. Cell. Biol. *22*, 7919–7928.

Kishimoto, H., and Sprent, J. (2001). A defect in central tolerance in NOD mice. Nat. Immunol. *2*, 1025–1031.

Klein, L., Hinterberger, M., Wirnsberger, G., and Kyewski, B. (2009). Antigen presentation in the thymus for positive selection and central tolerance induction. Nat. Rev. Immunol. *9*, 833–844.

Klein, L., Kyewski, B., Allen, P.M., and Hogquist, K.A. (2014). Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). Nat. Rev. Immunol. *14*, 377–391.

Koh, A.S., Kuo, A.J., Park, S.Y., Cheung, P., Abramson, J., Bua, D., Carney, D., Shoelson, S.E., Gozani, O., Kingston, R.E., et al. (2008). Aire employs a histone-binding module to mediate immunological tolerance, linking chromatin regulation with organ-specific autoimmunity. Proc. Natl. Acad. Sci. U. S. A. *105*, 15878–15883.

Kont, V., Murumägi, A., Tykocinski, L.O., Kinkel, S.A., Webster, K.E., Kisand, K., Tserel, L., Pihlap, M., Ströbel, P., Scott, H.S., et al. (2011). DNA methylation signatures of the AIRE promoter in thymic epithelial cells, thymomas and normal tissues. Mol. Immunol. *49*, 518–526.

Kumar, P.G., Laloraya, M., Wang, C.-Y., Ruan, Q.-G., Davoodi-Semiromi, A., Kao, K.-J., and She, J.-X. (2001). The Autoimmune Regulator (AIRE) Is a DNA-binding Protein. J. Biol. Chem. *276*, 41357–41364.

Kurz, A., Lampel, S., Nickolenko, J.E., Bradl, J., Benner, A., Zirbel, R.M., Cremer, T., and Lichter, P. (1996). Active and inactive genes localize preferentially in the periphery of chromosome territories. J. Cell Biol. *135*, 1195–1205.

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the enhancer of zeste protein. Genes Dev. *16*, 2893–2905.

Kyewski, B., and Klein, L. (2006). A central role for central tolerance. Annu. Rev. Immunol. *24*, 571–606.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116–120.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680–685.

LaFlam, T.N., Seumois, G., Miller, C.N., Lwin, W., Fasano, K.J., Waterfield, M., Proekt, I., Vijayanand, P., and Anderson, M.S. (2015). Identification of a novel cis-regulatory element essential for immune tolerance. J. Exp. Med. *212*, 1993–2002.

Lan, F., Collins, R.E., Cegli, R. De, Alpatov, R., Horton, J.R., Shi, X., Gozani, O., Cheng, X., Shi, Y., De Cegli, R., et al. (2007). Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. Nature *448*, 718–723.

Lange, M., Kaynak, B., Forster, U.B., Tonjes, M., Fischer, J.J., Grimm, C., Schlesinger, J., Just, S., Dunkel, I., Krueger, T., et al. (2008). Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. Genes Dev. *22*, 2370–2384.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

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

Lee, H.-Y., Johnson, K.D., Boyer, M.E., and Bresnick, E.H. (2011). Relocalizing Genetic Loci into Specific Subnuclear Neighborhoods. J. Biol. Chem. *286*, 18834–18844.

Lesage, S., Hartley, S.B., Akkaraju, S., Wilson, J., Townsend, M., and Goodnow, C.C. (2002). Failure to Censor Forbidden Clones of CD4 T Cells in Autoimmune Diabetes. J. Exp. Med. *196*, 1175–1188.

Levine, S.S., Weiss, A., Erdjument-Bromage, H., Shao, Z., Tempst, P., and Kingston, R.E. (2002). The Core of the Polycomb Repressive Complex Is Compositionally and Functionally Conserved in Flies and Humans. Mol. Cell. Biol. *22*, 6070–6078.

Li, H., Ilin, S., Wang, W., Duncan, E.M., Wysocka, J., Allis, C.D., and Patel, D.J. (2006). Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature *442*, 91–95.

Li, L., Lyu, X., Hou, C., Takenaka, N., Nguyen, H.Q., Ong, C.-T., Cubeñas-Potts, C., Hu, M., Lei, E.P., Bosco, G., et al. (2015). Widespread Rearrangement of 3D Chromatin Organization Underlies Polycomb-Mediated Stress-Induced Silencing. Mol. Cell *58*, 216–231.

Li, M., Luo, J., Brooks, C.L., and Gu, W. (2002). Acetylation of p53 Inhibits Its Ubiquitination by Mdm2. J. Biol. Chem. *277*, 50607–50611.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science (80–.). *326*, 289–293.

Liiv, I., Haljasorg, U., Kisand, K., Maslovskaja, J., Laan, M., and Peterson, P. (2012). AIRE-induced apoptosis is associated with nuclear translocation of stress sensor protein GAPDH. Biochem. Biophys. Res. Commun. *423*, 32–37.

Liiv, I., Rebane, A., Org, T., Saare, M., Maslovskaja, J., Kisand, K., Juronen, E., Valmu, L., Bottomley, M.J., Kalkkinen, N., et al. (2008). DNA-PK contributes to the phosphorylation of AIRE: Importance in transcriptional activity. Biochim. Biophys. Acta – Mol. Cell Res. *1783*, 74–83.

Liu, L.F., and Wang, J.C. (1987). Supercoiling of the DNA template during transcription. Proc. Natl. Acad. Sci. *84*, 7024–7027.

Liu, X., Yu, X., Zack, D.J., Zhu, H., and Qian, J. (2008). TiGER: a database for tissue-specific gene expression and regulation. BMC Bioinformatics *9*, 271.

Lowndes, N.F., and Toh, G.W.-L. (2005). DNA Repair: The Importance of Phosphorylating Histone H2AX. Curr. Biol. *15*, R99–R102.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251–260.

Lun, A.T.L., and Smyth, G.K. (2016). csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. Nucleic Acids Res. *44*, e45.

Mackintosh, C. (2004). Dynamic interactions between 14-3-3 proteins and phosphoproteins regulate diverse cellular processes. Biochem. J. *342*, 329–342.

Mahy, N.L., Perry, P.E., Gilchrist, S., Baldock, R.A., and Bickmore, W.A. (2002). Spatial organization of active and inactive genes and noncoding DNA within chromosome territories. J. Cell Biol. *157*, 579–589.

Maison, C., Bailly, D., Peters, A.H.F.M., Quivy, J.-P., Roche, D., Taddei, A., Lachner, M., Jenuwein, T., and Almouzni, G. (2002). Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. Nat. Genet. *30*, 329–334.

Maslovskaja, J., Saare, M., Liiv, I., Rebane, A., and Peterson, P. (2015). Extended HSR/CARD domain mediates AIRE binding to DNA. Biochem. Biophys. Res. Commun. *468*, 913–920.

Mason, P.B., and Struhl, K. (2005). Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo. Mol. Cell *17*, 831–840.

Mathis, D., and Benoist, C. (2007). A decade of AIRE. Nat. Rev. Immunol. *7*, 1–6.

Matsuzaki, H., Daitoku, H., Hatta, M., Aoyama, H., Yoshimochi, K., and Fukamizu, A. (2005). Acetylation of Foxo1 alters its DNA-binding ability and sensitivity to phosphorylation. Proc. Natl. Acad. Sci. *102*, 11278–11283.

Meloni, A., Incani, F., Corda, D., Cao, A., and Rosatelli, M.C. (2008). Role of PHD fingers and COOH-terminal 30 amino acids in AIRE transactivation activity. Mol. Immunol. *45*, 805–809.

Meredith, M., Zemmour, D., Mathis, D., and Benoist, C. (2015). Aire controls gene expression in the thymic epithelium with ordered stochasticity. Nat. Immunol. *16*, 942–949.

Morey, C., Kress, C., and Bickmore, W.A. (2009). Lack of bystander activation shows that localization exterior to chromosome territories is not sufficient to up-regulate gene expression. Genome Res. *19*, 1184–1194.

Mueller, D.L. (2010). Mechanisms maintaining peripheral tolerance. Nat. Immunol. *11*, 21–27.

Murumägi, A., Silvennoinen, O., and Peterson, P. (2006). Ets transcription factors regulate AIRE gene promoter. Biochem. Biophys. Res. Commun. *348*, 768–774.

Murumägi, A., Vähämurto, P., and Peterson, P. (2003). Characterization of Regulatory Elements and Methylation Pattern of the Autoimmune Regulator (AIRE) Promoter. J. Biol. Chem. *278*, 19784–19790.

Nagamine, K., Peterson, P., Scott, H.S., Kudoh, J., Minoshima, S., Heino, M., Krohn, K.J.E., Lalioti, M.D., Mullis, P.E., Antonarakis, S.E., et al. (1997). Positional cloning of the APECED gene. Nat. Genet. *17*, 393–398.

Neufeld, M., Maclaren, N.K., and Blizzard, R.M. (1981). Two types of autoimmune Addison's disease associated with different polyglandular autoimmune (PGA) syndromes. Medicine (Baltimore). *60*, 355–362.

Nishikawa, Y., Hirota, F., Yano, M., Kitajima, H., Miyazaki, J., Kawamoto, H., Mouri, Y., and Matsumoto, M. (2010). Biphasic Aire expression in early embryos and in medullary thymic epithelial cells before end-stage terminal differentiation. J. Exp. Med. *207*, 963–971.

Nitiss, J.L. (2009). DNA topoisomerase II and its growing repertoire of biological functions. Nat. Rev. Cancer *9*, 327–337.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381–385.

Oftedal, B.E., Hellesen, A., Erichsen, M.M., Bratland, E., Vardi, A., Perheentupa, J., Kemp, E.H., Fiskerstrand, T., Viken, M.K., Weetman, A.P., et al. (2015). Dominant Mutations in the Autoimmune Regulator AIRE Are Associated with Common Organ-Specific Autoimmune Diseases. Immunity *42*, 1185–1196.

Org, T., Chignola, F., Hetényi, C., Gaetani, M., Rebane, A., Liiv, I., Maran, U., Mollica, L., Bottomley, M.J., Musco, G., et al. (2008). The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression. EMBO Rep. *9*, 370–376.

Org, T., Rebane, A., Kisand, K., Laan, M., Haljasorg, U., Andreson, R., and Peterson, P. (2009). AIRE activated tissue specific genes have histone modifications associated with inactive chromatin. Hum. Mol. Genet. *18*, 4699–4710.

Oti, M., Falck, J., Huynen, M.A., and Zhou, H. (2016). CTCF-mediated chromatin loops enclose inducible gene regulatory domains. BMC Genomics *17*, 252.

Oven, I., Brdickova, N., Kohoutek, J., Vaupotic, T., Narat, M., Peterlin, B.M., Brdicková, N., Kohoutek, J., Vaupotic, T., Narat, M., et al. (2007). AIRE recruits P-TEFb for transcriptional elongation of target genes in medullary thymic epithelial cells. Mol. Cell. Biol. *27*, 8815–8823.

Patel, D.J., and Wang, Z. (2013). Readout of Epigenetic Modifications. Annu. Rev. Biochem. *82*, 81–118.

Pelanda, R., and Torres, R.M. (2012). Central B-Cell Tolerance: Where Selection Begins. Cold Spring Harb. Perspect. Biol. *4*, a007146.

Peters, A.H.F.M., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schöfer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., et al. (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell *107*, 323–337.

Phatnani, H.P., and Greenleaf, A.L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. Genes Dev. *20*, 2922–2936.

Phatnani, H.P., Jones, J.C., and Greenleaf, A.L. (2004). Expanding the functional repertoire of CTD kinase I and RNA polymerase II: Novel phosphoCTD-associating proteins in the yeast proteome. Biochemistry *43*, 15702–15719.

Pignatelli, M., Serras, F., Moya, A., Guigo, R., and Corominas, M. (2009). CROC: finding chromosomal clusters in eukaryotic genomes. Bioinformatics *25*, 1552–1553.

Pinto, S., Michel, C., Schmidt-Glenewinkel, H., Harder, N., Rohr, K., Wild, S., Brors, B., and Kyewski, B. (2013). Overlapping gene coexpression patterns in human medullary thymic epithelial cells generate self-antigen diversity. Proc. Natl. Acad. Sci. *110*, 3497–3505.

Pitkänen, J., Doucas, V., Sternsdorf, T., Nakajima, T., Aratani, S., Jensen, K., Will, H., Vähämurto, P., Ollila, J., Vihinen, M., et al. (2000). The Autoimmune Regulator Protein Has Transcriptional Transactivating Properties and Interacts with the Common Coactivator CREB-binding Protein. J. Biol. Chem. *275*, 16802–16809.

Pitkänen, J., Rebane, A., Rowell, J., Murumägi, A., Ströbel, P., Möll, K., Saare, M., Heikkilä, J., Doucas, V., Marx, A., et al. (2005). Cooperative activation of transcription by autoimmune regulator AIRE and CBP. Biochem. Biophys. Res. Commun. *333*, 944–953.

Plevin, M.J., Mills, M.M., and Ikura, M. (2005). The LxxLL motif: a multifunctional binding sequence in transcriptional regulation. Trends Biochem. Sci. *30*, 66–69.

Pommier, Y., Pourquier, P., Fan, Y., and Strumberg, D. (1998). Mechanism of action of eukaryotic DNA topoisomerase I and drugs targeted to the enzyme. Biochim. Biophys. Acta – Gene Struct. Expr. *1400*, 83–106.

Puc, J., Aggarwal, A.K., and Rosenfeld, M.G. (2017). Physiological functions of programmed DNA breaks in signal-induced transcription. Nat. Rev. Mol. Cell Biol. *18*, 471–476.

Purohit, S., Kumar, P.G., Laloraya, M., and She, J.-X.X. (2005). Mapping DNA-binding domains of the autoimmune regulator protein. Biochem. Biophys. Res. Commun. *327*, 939–944.

Ragoczy, T., Telling, A., Sawado, T., Groudine, M., and Kosak, S.T. (2003). A genetic analysis of chromosome territory looping: Diverse roles for distal regulatory elements. Chromosom. Res. *11*, 513–525.

Ramsey, C., Bukrinsky, A., and Peltonen, L. (2002). Systematic mutagenesis of the functional domains of AIRE reveals their role in intracellular targeting. Hum. Mol. Genet. *11*, 3299–3308.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell *159*, 1665–1680.

Rawlings, J.S., Rosler, K.M., and Harrison, D.A. (2004). The JAK/STAT signaling pathway. J. Cell Sci. *117*, 1281–1283.

Rinderle, C., Christensen, H.M., Schweiger, S., Lehrach, H., and Yaspo, M.L. (1999). AIRE encodes a nuclear protein co-localizing with cytoskeletal filaments: Altered sub-cellular distribution of mutants lacking the PHD zinc fingers. Hum. Mol. Genet. *8*, 277–290.

Ruthenburg, A.J., Allis, C.D., and Wysocka, J. (2007). Methylation of Lysine 4 on Histone H3: Intricacy of Writing and Reading a Single Epigenetic Mark. Mol. Cell *25*, 15–30.

Saltis, M., Criscitiello, M.F., Ohta, Y., Keefe, M., Trede, N.S., Goitsuka, R., and Flajnik, M.F. (2008). Evolutionarily conserved and divergent regions of the Autoimmune Regulator (Aire) gene: A comparative analysis. Immunogenetics *60*, 105–114.

Sansó, M., and Fisher, R.P. (2013). Pause, play, repeat: CDKs push RNAP II's buttons. Transcription *4*, 146–152.

Sansom, S.N., Shikama-Dorn, N., Zhanybekova, S., Nusspaumer, G., Macaulay, I.C., Deadman, M.E., Heger, A., Ponting, C.P., and Holländer, G.A. (2014). Population and single-cell genomics reveal the Aire dependency, relief from Polycomb silencing, and distribution of self-antigen expression in thymic epithelia. Genome Res. *24*, 1918–1931.

Sanyal, A., Lajoie, B.R., Jain, G., and Dekker, J. (2012). The long-range interaction landscape of gene promoters. Nature *489*, 109–113.

Savic, V., Yin, B., Maas, N.L., Bredemeyer, A.L., Carpenter, A.C., Helmink, B.A., Yang-Iott, K.S., Sleckman, B.P., and Bassing, C.H. (2009). Formation of Dynamic γ-H2AX Domains along Broken DNA Strands Is Distinctly Regulated by ATM and MDC1 and Dependent upon H2AX Densities in Chromatin. Mol. Cell *34*, 298–310.

Schaller, C.E., Wang, C.L., Beck-engeser, G., Goss, L., Scott, H.S., Anderson, M.S., and Wabl, M. (2008). Expression of Aire and the Early Wave of Apoptosis in Spermatogenesis. J. Immunol. *180*, 1338–1343.

Scheuermann, M.O., Tajbakhsh, J., Kurz, A., Saracoglu, K., Eils, R., and Lichter, P. (2004). Topology of genes and nontranscribed sequences in human interphase nuclei. Exp. Cell Res. *301*, 266–279.

Schultz, M.C., Brill, S.J., Ju, Q., Sternglanz, R., and Reeder, R.H. (1992). Topoisomerases and yeast rRNA transcription: negative supercoiling stimulates initiation and topoisomerase activity is required for elongation. Genes Dev. *6*, 1332–1341.

Shao, W., Zumer, K., Fujinaga, K., and Peterlin, B.M. (2016). FBXO3 protein promotes ubiquitylation and transcriptional activity of AIRE (Autoimmune Regulator). J. Biol. Chem. *291*, 17953–17963.

Singh, N., Basnet, H., Wiltshire, T.D., Mohammad, D.H., Thompson, J.R., Heroux, A., Botuyan, M.V., Yaffe, M.B., Couch, F.J., Rosenfeld, M.G., et al. (2012). Dual recognition of phosphoserine and phosphotyrosine in histone variant H2A.X by DNA damage response protein MCPH1. Proc. Natl. Acad. Sci. *109*, 14381–14386.

Spange, S., Wagner, T., Heinzel, T., and Krämer, O.H. (2009). Acetylation of non-histone proteins modulates cellular signalling at multiple levels. Int. J. Biochem. Cell Biol. *41*, 185–198.

Starr, T.K., Jameson, S.C., and Hogquist, K.A. (2003). Positive and negative selection of T cells. Annu. Rev. Immunol. *21*, 139–176.

Stavropoulos, P., Nagy, V., Blobel, G., and Hoelz, A. (2008). Molecular basis for the autoregulation of the protein acetyl transferase Rtt109. Proc. Natl. Acad. Sci. *105*, 12236–12241.

Suganuma, T., and Workman, J.L. (2011). Signals and Combinatorial Functions of Histone Modifications. Annu. Rev. Biochem. *80*, 473–499.

Takahashi, T., Burguiere-Slezak, G., Van der Kemp, P.A., and Boiteux, S. (2011). Topoisomerase 1 provokes the formation of short deletions in repeated sequences upon high transcription in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. *108*, 692–697.

Tanabe, H., Habermann, F.A., Solovei, I., Cremer, M., and Cremer, T. (2002). Non-random radial arrangements of interphase chromosome territories: evolutionary considerations and functional implications. Mutat. Res. Mol. Mech. Mutagen. *504*, 37–45.

Tang, X., Gao, J.-S., Guan, Y., McLane, K.E., Yuan, Z.-L., Ramratnam, B., and Chin, Y.E. (2007). Acetylation-Dependent Signal Transduction for Type I Interferon Receptor. Cell *131*, 93–105.

Tao, Y., Kupfer, R., Stewart, B.J., Williams-Skipp, C., Crowell, C.K., Patel, D.D., Sain, S., and Scheinman, R.I. (2006). AIRE recruits multiple transcriptional components to specific genomic regions through tethering to nuclear matrix. Mol. Immunol. *43*, 335–345.

Tark-Dame, M., Jerabek, H., Manders, E.M.M., Heermann, D.W., and van Driel, R. (2014). Depletion of the Chromatin Looping Proteins CTCF and Cohesin Causes Chromatin Compaction: Insight into Chromatin Folding by Polymer Modelling. PLoS Comput. Biol. *10*, e1003877.

Thiru, A., Nietlispach, D., Mott, H.R., Okuwaki, M., Lyon, D., Nielsen, P.R., Hirshberg, M., Verreault, A., Murzina, N. V, and Laue, E.D. (2004). Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin. EMBO J. *23*, 489–499.

Thompson, P.R., Wang, D., Wang, L., Fulco, M., Pediconi, N., Zhang, D., An, W., Ge, Q., Roeder, R.G., Wong, J., et al. (2004). Regulation of the p300 HAT domain via a novel activation loop. Nat. Struct. Mol. Biol. *11*, 308–315.

Trojer, P., and Reinberg, D. (2007). Facultative Heterochromatin: Is There a Distinctive Molecular Signature? Mol. Cell *28*, 1–13.

Uchida, D., Hatakeyama, S., Matsushima, A., Han, H., Ishido, S., Hotta, H., Kudoh, J., Shimizu, N., Doucas, V., Nakayama, K.I., et al. (2004). AIRE functions as an E3 ubiquitin ligase. J. Exp. Med. *199*, 167–172.

Ventura, M., Mateo, F., Serratosa, J., Salaet, I., Carujo, S., Bachs, O., and Pujol, M.J. (2010). Nuclear translocation of glyceraldehyde-3-phosphate dehydrogenase is regulated by acetylation. Int. J. Biochem. Cell Biol. *42*, 1672–1680.

Vernimmen, D., Gobbi, M. De, Sloane-Stanley, J.A., Wood, W.G., and Higgs, D.R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. EMBO J. *26*, 2041–2051.

Verschure, P.J., van der Kraan, I., Enserink, J.M., Moné, M.J., Manders, E.M.M., and van Driel, R. (2002). Large-scale Chromatin Organization and the Localization of Proteins Involved in Gene Expression in Human Cells. J. Histochem. Cytochem. *50*, 1303–1312.

Vignali, M., and Workman, J.L. (1998). Location and function of linker histones. Nat. Struct. Biol. *5*, 1025–1028.

Vo, N., and Goodman, R.H. (2001). CREB-binding Protein and p300 in Transcriptional Regulation. J. Biol. Chem. *276*, 13505–13508.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. Nature *431*, 873–878.

Wang, X., and Hayes, J.J. (2008). Acetylation Mimics within Individual Core Histone Tail Domains Indicate Distinct Roles in Regulating the Stability of Higher-Order Chromatin Structure. Mol. Cell. Biol. *28*, 227–236.

Wardemann, H., Yurasov, S., Schaefer, A., Young, J.W., Meffre, E., and Nussenzweig, M.C. (2003). Predominant autoantibody production by early human B cell precursors. Science (80–.). *301*, 1374–1377.

Willems, E., Leyns, L., and Vandesompele, J. (2008). Standardization of real-time PCR gene expression data from independent biological replicates. Anal. Biochem. *379*, 127–129.

Wolff, A.S.B., Erichsen, M.M., Meager, A., Magitta, N.F., Myhre, A.G., Bollerslev, J., Fougner, K.J., Lima, K., Knappskog, P.M., and Husebye, E.S. (2007). Autoimmune polyendocrine syndrome type 1 in Norway: Phenotypic variation, autoantibodies, and novel mutations in the autoimmune regulator gene. J. Clin. Endocrinol. Metab. *92*, 595–603.

Yamano, T., Nedjic, J., Hinterberger, M., Steinert, M., Koser, S., Pinto, S., Gerdes, N., Lutgens, E., Ishimaru, N., Busslinger, M., et al. (2015). Thymic B Cells Are Licensed to Present Self Antigens for Central T Cell Tolerance Induction. Immunity *42*, 1048–1061.

Yanagihara, T., Sanematsu, F., Sato, T., Uruno, T., Duan, X., Tomino, T., Harada, Y., Watanabe, M., Wang, Y., Tanaka, Y., et al. (2015). Intronic regulation of Aire expression by Jmjd6 for self-tolerance induction in the thymus. Nat. Commun. *6*, 8820.

Yang, S., Bansal, K., Lopes, J., Benoist, C., and Mathis, D. (2013). Aire's plant homeo-domain(PHD)-2 is critical for induction of immunological tolerance. Proc. Natl. Acad. Sci. *110*, 1833–1838.

Yoshida, H., Bansal, K., Schaefer, U., Chapman, T., Rioja, I., Proekt, I., Anderson, M.S., Prinjha, R.K., Tarakhovsky, A., Benoist, C., et al. (2015). Brd4 bridges the transcrip-

tional regulators, Aire and P-TEFb, to promote elongation of peripheral-tissue antigen transcripts in thymic stromal cells. Proc. Natl. Acad. Sci. *112*, 4448–4457.

Yu, G., Wang, L.-G., and He, Q.-Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics *31*, 2382–2383.

Yuan, H., Rossetto, D., Mellert, H., Dang, W., Srinivasan, M., Johnson, J., Hodawadekar, S., Ding, E.C., Speicher, K., Abshiru, N., et al. (2012). MYST protein acetyltransferase activity requires active site lysine autoacetylation. EMBO J. *31*, 58–70.

Zhang, K., Lin, W., Latham, J.A., Riefler, G.M., Schumacher, J.M., Chan, C., Tatchell, K., Hawke, D.H., Kobayashi, R., and Dent, S.Y.R. (2005). The Set1 methyltransferase opposes Ipl1 Aurora kinase functions in chromosome segregation. Cell *122*, 723–734.

Zhu, P., Zhou, W., Wang, J., Puc, J., Ohgi, K.A., Erdjument-Bromage, H., Tempst, P., Glass, C.K., and Rosenfeld, M.G. (2007). A Histone H2A Deubiquitinase Complex Coordinating Histone Acetylation and H1 Dissociation in Transcriptional Regulation. Mol. Cell *27*, 609–621.

Zippo, A., Serafini, R., Rocchigiani, M., Pennacchini, S., Krepelova, A., and Oliviero, S. (2009). Histone Crosstalk between H3S10ph and H4K16ac Generates a Histone Code that Mediates Transcription Elongation. Cell *4*, 1122–1136.

Zuin, J., Dixon, J.R., van der Reijden, M.I.J.A., Ye, Z., Kolovos, P., Brouwer, R.W.W., van de Corput, M.P.C., van de Werken, H.J.G., Knoch, T.A., van IJcken, W.F.J., et al. (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proc. Natl. Acad. Sci. *111*, 996–1001.

Zumer, K., Low, A.K., Jiang, H., Saksela, K., and Peterlin, B.M. (2012). Unmodified Histone H3K4 and DNA-Dependent Protein Kinase Recruit Autoimmune Regulator to Target Genes. Mol. Cell. Biol. *32*, 1354–1362.

SUMMARY IN ESTONIAN

AIRE mõju geenidekspressioonile – transkriptsiooni regulatsiooni mehhanismi uuringud koekultuuri rakkudes

Immuunsüsteemi põhiline ülesanne on ära tunda ja võidelda patogeenidega, mis üritavad tungida organismi ning seda kahjustada. Samaväärselt oluline on organismi seisukohast immuunsüsteemi võime ära tunda omaenda kehale omaseid või väljastpoolt tulevaid kahjutuid molekule ning hoiduda nende vastu suunatud immuunreaktsiooni tekitamisest. Seda nähtust nimetatakse immuuntolerantsuseks. Üks peamisi immuuntolerantsuse kujunemise organeid on harknääre ehk tüümus, mis vastutab kehaomaseid valke tundvate T-lümfotsüütide elimineerimise eest. Nimelt reageerivad arenevad T-lümfotsüüdid T-raku retseptori kaudu tüümuse medullaarsete epiteelirakkude pinnal olevate koesobivuskompleksidega. Antud kompleksid kannavad enda küljes kehavalkudest pärit peptiide, mille suhtes spetsiifilise T-raku retseptoriga T-lümfotsüüt suunatakse apoptoosi. T-lümfotsüüdid, mis reageerivad nõrgalt või ei reageeri üldse kehaomaste peptiididega, jäävad ellu ning pääsevad küpsete naiivsete T-rakkudena tüümusest välja vere- ja lümfiringesse. Tüümuse medullaarsed epiteelirakud on võimelised avaldama hämmastavalt suurt osa kõigist organismi valke kodeerivatest geenidest, sealjuures ülimalt koespetsiifilisi geene. Paljuski sõltub see n-ö. avatud geenidekspressioon ühe transkriptsiooni reguleeriva valgu olemasolust, milleks on AIRE (*autoimmune regulator*). Mutatsioonid *AIRE* geenis võivad põhjustada 1. tüüpi autoimmuunset polüendokrinopaatia sündroomi (APS1). Kuna AIRE puudumise tõttu ei avaldu koespetsiifilised geenid, siis jäävad autoreaktiivset T-raku retseptorit kandvad T-lümfotsüüdid ellu, pääsevad vere- ja lümfiringesse ning ründavad keha enda kudesid. Lisaks tekivad APS1 korral kudede vastu suunatud autoantikehad. APS1 raskusaste ja sümptomite spekter on väga varieeruvad, kuid kõige sagedamini esinevad patsientidel kõrvalkilpnäärme alatalitlus ja neerupealiste puudulikkus. Seennakkuste vastu võitlemiseks vajalike tsütokiinide vastaste autoantikehade tõttu tekib peaaegu kõigil patsientidel kandidoos.

Alates *AIRE* geeni ja valgu kirjeldamisest 1997. aastal on palju tähelepanu pööratud protsessidele, mis võimaldavad AIRE valgul käivitada tuhandete geenide avaldumise, mis AIRE puudumisel on üldjuhul vaigistatud. Intensiivse uurimistöö tulemusel on selgunud, et AIRE seondub paljude geenidekspressiooni mõjutavate valkudega ja seeläbi kasutab ära teiste valkude omadusi, et reguleerida geenide avaldumist eelkõige transkriptsiooniliselt inaktiivsetes genoomi piirkondades. Kahe iseseisva proteoomiuuringu tulemuste põhjal on kindlaks tehtud, et AIRE interakteerub väga laia valkude repertuaariga, mida võib nende funktsiooni alusel jagada 4 suuremasse rühma. Nendeks on transkriptsioonis, kromatiini struktuuris ja remodelleerimises, RNA splaissingus ja rakutuuma transpordis osalevad valgud. Järk-järgult on selgunud, millistes etappides on need kirjeldatud valgud AIRE poolt reguleeritava transkriptsiooni jaoks vajalikud. AIRE interaktsioonipartneritest üks olulisemaid ning rakutuuma valkudest sagedasemaid on histooni H3 selline variant, mida leidub üldjuhul just vaigistatud genoomsetes piirkondes. AIRE seondub transkriptsiooni elongatsiooni-

faktoritega, mis aitavad RNA polümeraasi aktiveerida, ning rakutuumas paiknevate ensüümidega, mis parandavad polümeraasi liikumise käigus tekkivaid DNA katkeid. Lisaks on AIRE omadust aktiveerida just vaigistatud geene püütud seletada sellega, et AIRE interakteerub transkriptsiooni represseeriva valgukompleksiga, mis osaleb heterokromatiinile omase histoonimodifikatsioonimustri loomisel ning seondub metüülitud CpG dinukleotiididega vaigistatud genoomsetes piirkondades. Siia tuleb veel märkida, et on kirjeldatud ka AIRE seondumist kromatiniseerumata DNA molekuliga, leides, et sel juhul toimub geeni aktivatsioon sõltumatult regulatoorsete järjestuste, nagu promootori või mittetransleeritavate 5' ja 3' järjestuste olemasolust. Sellise interaktsiooni roll koespetsiifiliste geenide transkriptsiooni käivitamisel tüümuse medullaarsetes epiteelirakkudes on veel ebaselge.

Käesoleva uurimistöö esimene osa lähtus varasematest tulemustest, mis näitasid, et AIRE seondub CBP-ga ja seeläbi võimendub AIRE sihtmärkgeenide ekspressioon. Kuna CBP atsetüültransferaasse aktiivsuse rolli selles protsessis ei oldud kirjeldatud, keskenduti esmalt küsimusele, kas AIRE ja CBP interaktsiooni üheks tagajärjeks on AIRE valgu atsetüülimine. Immuunosadestamise ning mass-spektromeetrilise analüüsi põhjal leiti, et vähemalt pooled AIRE valgu järjestuses olevad lüsiinijäägid on CBP ja tema paraloogi, p300, poolt atsetüülitud. Antud katsetest selgus, et atsetüülimine toimub CARD ja SAND domeeni vahelises osas, kuhu jääb tuumalokalisatsioonisignaali järjestus, ning SAND domeeni sees. Edasine tähelepanu koondus atsetüülimisele, mis toimub SAND domeenis, mida varasemates töödes on kirjeldatud kui DNA-ga seonduvat järjestust. Atsetüülimine võib mõjutada valgu lokalisatsiooni, tema stabiilsust ja võimet interakteeruda DNA või teiste valkudega. Leiti, et p300 üleekspressiooni järel muutuvad AIRE tuumakehad suuremaks, kuid nende arv väheneb. Samuti täheldati, et AIRE valk laguneb palju aeglasemalt, vihjates sellele, et atsetüülitud lüsiinijäägid võivad olla substraadiks polüubikvitineerimisreaktsioonidele, mis suunavad valgu lagundamisele proteasoomide poolt. Valitud AIRE sihtmärkgeenide avaldumist p300 üleekspressioon ei mõjutanud, kuid teatud AIRE SAND domeenis olevate atsetüülitavate lüsiinide asendamine glutamiinidega, mis positiivse laengu puudumise ning külgahela struktuuri tõttu mimikeerivad atsetüülitud lüsiine, vähendas tunduvalt AIRE poolt reguleeritavate geenide avaldumist. Samas AIRE säilitas täielikult oma transaktivatsioonivõime, kui lüsiinid olid asendatud arginiinidega, mis sarnaselt atsetüülimata lüsiinidele kannavad positiivset laengut. Ülegenoomi ekspressioonianalüüsiga tuvastati, et atsetüülimist mimikeerivate mutatsioonidega AIRE valk ei kaota täielikult oma transaktivatsioonivõimet, vaid muutub sihtmärkgeenide valik, mis siiski kattub osaliselt ka metsiktüüpi AIRE valgu poolt reguleeritavate geenide nimekirjaga.

Töö teises osas analüüsiti APS1 patsientidel leitud AIRE PHD2 domeenis esineva *missense* mutatsiooni C446G mõju domeeni struktuurile ja AIRE funktsioonile. Selgus, et antud mutatsioon lõhub täielikult PHD2 domeeni kooshoidmiseks vajaminevad sidemed Zn^{2+} ioonide ning aminohapete vahel. Selle tulemuseks on AIRE valk, mis agregeerub tsütosoolis ning ei pääse seetõttu raktuuma, et käivitada seal oma sihtmärkgeenide avaldumine. Huvitaval kombel selgus, et GST liitvalguna ekspresseeritud C446G AIRE mutatsiooniga valk

seondus *in vitro* seondumiskatsetes endiselt histoon H3-ga, mis on kriitiliselt oluline AIRE sihtmärkgeenide aktiveerimiseks. Sellest võib järeldada, et õige lokalisatsiooni korral suudaks antud mutatsiooniga AIRE valk tagada sihtmärkgeenide ekspressiooni. Seda oletust toetavad katsed, mis on tehtud ilma PHD2 domeenita AIRE valguga, mis ei agregeeru ning on endiselt transkriptsiooniliselt aktiivne.

Töö kolmanda osana analüüsiti AIRE ja TOP2A valkude koostoimet AIRE sihtmärkgeenide transkriptsioonile. Varasemalt avaldatud töös leiti, et TOP2A DNA ahelaid ligeeriva ensümaatilise aktiivsuse pärssimine etoposiidiga käivitab sarnase transkriptioonilise programmi, nagu see on AIRE üleekspressiooni korral ja AIRE ning etoposiidi koostoimel ei ole teineteist täiendavat efekti. Käesolevas uuringus leiti, et AIRE valku ekspresseerivates rakkudes võib madala etoposiidi kontsentratsiooni juures toimuda väga võimas AIRE sihtmärkgeenide avaldumise tõus. Etoposiidi töötlus stabiliseerib DNA katked genoomis ja DNA katketega piirkondi iseloomustab histoon γH2AX rikastuse tõus. Antud juhul suurenes γH2AX rikastus AIRE sihtmärkgeenide lähedal. Sihtmärkgeenide promootoralades esines ka rohkem HMGB1 valku, mis iseloomustab transkriptsiooniliselt aktiivset ning avatud kromatiini struktuuri, ning vähenes linkerhistooni H1 hulk, mida seostatakse vaigistatud kromatiini olekuga. Nende tulemuste põhjal tehti ülegenoomi sekveneerimise katse, et tuvastada genoomsed piirkonnad, mille avatus muutub AIRE olemasolust tingitult. Leiti, et AIRE juureolekul ning eriti juhul, kui AIRE-t ekspresseerivaid rakke oli töödeldud etoposiidiga, kaotasid paljud piirkonnad oma avatuse. Eelkõige toimusid muudatused geenide promootoralades, kuid huvitaval kombel oli nende hulgas väga vähe AIRE poolt reguleeritud geenide promootoralasid. Üldiselt toimusid muutused AIRE sihtmärkgeenidest kaugel, kuid need piirkonnad kattusid olulisel määral CTCF valgu seondumispiirkondadega. CTCF valku peetakse oluliseks kromatiini struktuuri loovaks valguks, luues valk-DNA ja valk-valk interaktsioonide kaudu kromatiini lingusid. Järgnevalt testiti, kas mitmeid AIRE sihtmärkgeene sisaldava CEACAM geeniklastri sees ja servades olevate CTCF seondumiskohtade vahel on võimalik detekteerida interaktsioonide sageduse muutuseid, kasutades selleks 3C meetodit. Tulemustest selgus, et AIRE juuresolekul tõepoolest tõuseb teatud interaktsioonide sagedus, samal ajal kui teiste interaktsioonide sagedus langeb.

Kokkuvõtvalt võib öelda, et post-translatsiooniline atsetüülimine mõjutab olulisel määral AIRE valgu stabiilsust ning transkriptsioonilist aktiivust. Lisaks tehti kindlaks, et APS1 patsientidel leiduv C446G mutatsioon AIRE PHD2 domeenis põhjustab AIRE valgu agregeerumist, mis takistab AIRE-l liikuda tsütoplasmast rakutuuma, et seal käivitada oma sihtmärkgeenide ekspressioon. Huvitaval kombel on intaktne AIRE võimeline mitte ainult aktiveerima tuhandete vaigistatud geenide avaldumist, vaid ka mõjutama kogu genoomi ulatuses kromatiini struktuuri. Siiski on vajalikud täpsemad uuringud, et selgitada, millised põhjus-tagajärg seosed valitsevad kromatiini struktuuri ja AIRE poolt reguleeritava transkriptsiooni vahel, mis eeldatavasti looks hea aluse senini enigmaatilisena püsinud avatud geeniekspressiooni ja seeläbi immuuntolerantsuse molekulaarsete mehhanismide täielikuks mõistmiseks.

ACKNOWLEDGMENTS

The research presented in this dissertation was conducted at the Molecular Pathology Research Group in the Institute of Biomedicine and Translational Medicine at the University of Tartu. I would like to thank all the people who supported me throughout my doctorate studies.

I wish to express my sincere gratitude to my supervisor Pärt Peterson for his expert advice and encouragement at every step of my studies and for the opportunity to work on many interesting projects in his group.

I am very grateful to my supervisors senior researcher Ana Rebane and researcher Kaidi Möll for taking me under their wing in the early days of my life in the lab.

I would like to thank my colleagues for their wonderful collaboration. You supported me greatly and were always willing to help me. My special thanks goes to Maire Pihlap for her excellent technical assistants.

I would like to thank my referees, senior researchers Meeme Utt and Ivar Ilves, for their valuable comments and suggestions on the thesis manuscript.

I would like to thank all our collaborators and co-authors of the publications for their contribution in the different projects.

Finally, I would like to thank my friends and family for their loving support and for all the joyful moments spent together.

***

The work was supported by the grants from the Estonian Science Foundation (grants 8169 and 8350), Estonian Research Council funding (grants IUT2-2 and MJD164), Estonian Ministry of Education and Research targeted funding SF0180021s07, the Estonian Academy of Sciences, European Regional Development Fund (Project No.2014–2020.4.01.15-0012), European Regional Development Fund, ERA-Net.Rus project EGIDA and the Archimedes Foundation.

PUBLICATIONS

CURRICULUM VITAE

Name:	Mario Saare
Date of birth:	15th of March 1984
Address:	Molecular Pathology Research Group, Institute of Biomedicine and Translational Medicine University of Tartu Ravila 19, 50411 Tartu Estonia
E-mail:	mario.saare@ut.ee
Telephone:	+3727374186

Education

2003–2006	BSc in Biology, University of Tartu
2006–2008	MSc in Biomedicine, University of Tartu
2008–present	PhD student in Immunology, University of Tartu

Professional employment

2005–2016	Laboratory specialist at the Molecular Pathology Research Group, Institute of Biomedicine and Translational Medicine, University of Tartu
2016–present	Junior research fellow at the Molecular Pathology Research Group, Institute of Biomedicine and Translational Medicine, University of Tartu

Supervised dissertations

2013	Karin Kõnd, Master's degree; supervisors: Mario Saare, Arnold Kristjuhan; The role of the 5' conserved regulatory region on AIRE gene expression; Faculty of Science and Technology, University of Tartu.
2014	Uku Hämarik, Master's degree, supervisors: Mario Saare, Tambet Tõnissoo; Studies on mouse Sp140 gene; Faculty of Science and Technology, University of Tartu

Professional memberships

2007–present	Estonian Society for Immunology and Allergology
2007–present	Estonian Society of Human Genetics

Publications

1. Gaetani, M., Matafora, V., Saare, M., Spiliotopoulos, D., Mollica, L., Quilici, G., Chignola, F., Mannella, V., Zucchelli, C., Peterson, P., et al. (2012). AIRE-PHD fingers are structural hubs to maintain the integrity of chromatin-associated interactome. Nucleic Acids Res. *40*, 11756–11768.

2. Guha, M., Saare, M., Maslovskaja, J., Kisand, K., Liiv, I., Haljasorg, U., Tasa, T., Metspalu, A., Milani, L., and Peterson, P. (2017). DNA breaks and chromatin structural changes enhance the transcription of autoimmune regulator target genes. J. Biol. Chem. *292*, 6542–6554.
3. Haljasorg, U., Bichele, R., Saare, M., Guha, M., Maslovskaja, J., Kõnd, K., Remm, A., Pihlap, M., Tomson, L., Kisand, K., et al. (2015). A highly conserved NF-κB-responsive enhancer is critical for thymic expression of Aire in mice. Eur. J. Immunol. *45*, 3246–3256.
4. Liiv, I., Rebane, A., Org, T., Saare, M., Maslovskaja, J., Kisand, K., Juronen, E., Valmu, L., Bottomley, M.J., Kalkkinen, N., et al. (2008). DNA-PK contributes to the phosphorylation of AIRE: Importance in transcriptional activity. Biochim. Biophys. Acta – Mol. Cell Res. *1783*, 74–83.
5. Limbach, M., Saare, M., Tserel, L., Kisand, K., Eglit, T., Sauer, S., Axelsson, T., Syvänen, A.-C., Metspalu, A., Milani, L., et al. (2016). Epigenetic profiling in CD4+ and CD8+ T cells from Graves' disease patients reveals changes in genes associated with T cell receptor signaling. J. Autoimmun. *67*, 46–56.
6. Maslovskaja, J., Saare, M., Liiv, I., Rebane, A., and Peterson, P. (2015). Extended HSR/CARD domain mediates AIRE binding to DNA. Biochem. Biophys. Res. Commun. *468*, 913–920.
7. Pitkänen, J., Rebane, A., Rowell, J., Murumägi, A., Ströbel, P., Möll, K., Saare, M., Heikkilä, J., Doucas, V., Marx, A., et al. (2005). Cooperative activation of transcription by autoimmune regulator AIRE and CBP. Biochem. Biophys. Res. Commun. *333*, 944–953.
8. Saare, M., Rebane, A., Rajashekar, B., Vilo, J., and Peterson, P. (2012). Autoimmune regulator is acetylated by transcription coactivator CBP/p300. Exp. Cell Res. *318*, 1767–1778.
9. Saare, M., Hämarik, U., Venta, R., Panarina, M., Zucchelli, C., Pihlap, M., Remm, A., Kisand, K., Toots, U., Möll, K., et al. (2015). SP140L, an Evolutionarily Recent Member of the SP100 Family, Is an Autoantigen in Primary Biliary Cirrhosis. J. Immunol. Res. *2015*, 1–17.
10. Tserel, L., Limbach, M., Saare, M., Kisand, K., Metspalu, A., Milani, L., and Peterson, P. (2014). CpG sites associated with NRP1, NRXN2 and miR-29b-2 are hypomethylated in monocytes during ageing. Immun. Ageing *11*, 1.
11. Tserel, L., Kolde, R., Limbach, M., Tretyakov, K., Kasela, S., Kisand, K., Saare, M., Vilo, J., Metspalu, A., Milani, L., et al. (2015). Age-related profiling of DNA methylation in CD8+ T cells reveals changes in immune response and transcriptional regulator genes. Sci. Rep. *5*, 13107.
12. Zucchelli, C., Tamburri, S., Quilici, G., Palagano, E., Berardi, A., Saare, M., Peterson, P., Bachi, A., and Musco, G. (2014). Structure of human Sp140 PHD finger: an atypical fold interacting with Pin1. FEBS J. *281*, 216–231.

ELULOOKIRJELDUS

Nimi:	Mario Saare
Sünniaeg:	15 märts 1984. a
Aadress:	Molekulaarpatoloogia uurimisrühm Bio- ja siirdemeditsiini instituut Tartu Ülikool Ravila 19, 50411 Tartu Eesti Vabariik
E-post:	mario.saare@ut.ee
Telefon:	+3727374186

Haridus

2003–2006	Loodusteaduste bakalaureusekraad (BSc) bioloogias, Tartu Ülikool
2006–2008	Loodusteaduste magistrikraad (MSc) biomeditsiinis, Tartu Ülikool
2008–praegu	Doktorant immunoloogia erialal, arstiteaduse õppekava, Tartu Ülikool

Erialane teenistuskäik

2005–2016	Laborispetsialist, Molekulaarpatoloogia uurimisrühm, Bio- ja siirdemeditsiini instituut, Tartu Ülikool
2016–praegu	Nooremteadur, Molekulaarpatoloogia uurimisrühm, Bio- ja siirdemeditsiini instituut, Tartu Ülikool

Juhendatud väitekirjad

2013	Karin Kõnd, magistrikraad; juhendajad: Mario Saare Arnold Kristjuhan; Konserveerunud 5' regulatoorse ala mõju AIRE geeni ekspressioonile; Loodus- ja tehnoloogiateaduskond, Tartu Ülikool
2014	Uku Hämarik, magistrikraad; juhendajad: Mario Saare, Tambet Tõnissoo; SP140 geeni funktsiooni iseloomustamine; Loodus- ja tehnoloogiateaduskond, Tartu Ülikool

Kuulumine erialaorganisatsioonidesse

2007–praegu	Eesti Immunoloogide ja Allergoloogide Selts
2007–praegu	Eesti Inimesegeneetika Ühing

Publikatsioonid

1. Gaetani, M., Matafora, V., Saare, M., Spiliotopoulos, D., Mollica, L., Quilici, G., Chignola, F., Mannella, V., Zucchelli, C., Peterson, P., et al. (2012). AIRE-PHD fingers are structural hubs to maintain the integrity of chromatin-associated interactome. Nucleic Acids Res. *40*, 11756–11768.

2. Guha, M., Saare, M., Maslovskaja, J., Kisand, K., Liiv, I., Haljasorg, U., Tasa, T., Metspalu, A., Milani, L., and Peterson, P. (2017). DNA breaks and chromatin structural changes enhance the transcription of autoimmune regulator target genes. J. Biol. Chem. *292*, 6542–6554.
3. Haljasorg, U., Bichele, R., Saare, M., Guha, M., Maslovskaja, J., Kõnd, K., Remm, A., Pihlap, M., Tomson, L., Kisand, K., et al. (2015). A highly conserved NF-κB-responsive enhancer is critical for thymic expression of Aire in mice. Eur. J. Immunol. *45*, 3246–3256.
4. Liiv, I., Rebane, A., Org, T., Saare, M., Maslovskaja, J., Kisand, K., Juronen, E., Valmu, L., Bottomley, M.J., Kalkkinen, N., et al. (2008). DNA-PK contributes to the phosphorylation of AIRE: Importance in transcriptional activity. Biochim. Biophys. Acta – Mol. Cell Res. *1783*, 74–83.
5. Limbach, M., Saare, M., Tserel, L., Kisand, K., Eglit, T., Sauer, S., Axelsson, T., Syvänen, A.-C., Metspalu, A., Milani, L., et al. (2016). Epigenetic profiling in CD4+ and CD8+ T cells from Graves' disease patients reveals changes in genes associated with T cell receptor signaling. J. Autoimmun. *67*, 46–56.
6. Maslovskaja, J., Saare, M., Liiv, I., Rebane, A., and Peterson, P. (2015). Extended HSR/CARD domain mediates AIRE binding to DNA. Biochem. Biophys. Res. Commun. *468*, 913–920.
7. Pitkänen, J., Rebane, A., Rowell, J., Murumägi, A., Ströbel, P., Möll, K., Saare, M., Heikkilä, J., Doucas, V., Marx, A., et al. (2005). Cooperative activation of transcription by autoimmune regulator AIRE and CBP. Biochem. Biophys. Res. Commun. *333*, 944–953.
8. Saare, M., Rebane, A., Rajashekar, B., Vilo, J., and Peterson, P. (2012). Autoimmune regulator is acetylated by transcription coactivator CBP/p300. Exp. Cell Res. *318*, 1767–1778.
9. Saare, M., Hämarik, U., Venta, R., Panarina, M., Zucchelli, C., Pihlap, M., Remm, A., Kisand, K., Toots, U., Möll, K., et al. (2015). SP140L, an Evolutionarily Recent Member of the SP100 Family, Is an Autoantigen in Primary Biliary Cirrhosis. J. Immunol. Res. *2015*, 1–17.
10. Tserel, L., Limbach, M., Saare, M., Kisand, K., Metspalu, A., Milani, L., and Peterson, P. (2014). CpG sites associated with NRP1, NRXN2 and miR-29b-2 are hypomethylated in monocytes during ageing. Immun. Ageing *11*, 1.
11. Tserel, L., Kolde, R., Limbach, M., Tretyakov, K., Kasela, S., Kisand, K., Saare, M., Vilo, J., Metspalu, A., Milani, L., et al. (2015). Age-related profiling of DNA methylation in CD8+ T cells reveals changes in immune response and transcriptional regulator genes. Sci. Rep. *5*, 13107.
12. Zucchelli, C., Tamburri, S., Quilici, G., Palagano, E., Berardi, A., Saare, M., Peterson, P., Bachi, A., and Musco, G. (2014). Structure of human Sp140 PHD finger: an atypical fold interacting with Pin1. FEBS J. *281*, 216–231.

DISSERTATIONES MEDICINAE UNIVERSITATIS TARTUENSIS

1. **Heidi-Ingrid Maaroos**. The natural course of gastric ulcer in connection with chronic gastritis and *Helicobacter pylori*. Tartu, 1991.
2. **Mihkel Zilmer**. Na-pump in normal and tumorous brain tissues: Structural, functional and tumorigenesis aspects. Tartu, 1991.
3. **Eero Vasar**. Role of cholecystokinin receptors in the regulation of behaviour and in the action of haloperidol and diazepam. Tartu, 1992.
4. **Tiina Talvik**. Hypoxic-ischaemic brain damage in neonates (clinical, biochemical and brain computed tomographical investigation). Tartu, 1992.
5. **Ants Peetsalu**. Vagotomy in duodenal ulcer disease: A study of gastric acidity, serum pepsinogen I, gastric mucosal histology and *Helicobacter pylori*. Tartu, 1992.
6. **Marika Mikelsaar**. Evaluation of the gastrointestinal microbial ecosystem in health and disease. Tartu, 1992.
7. **Hele Everaus**. Immuno-hormonal interactions in chronic lymphocytic leukaemia and multiple myeloma. Tartu, 1993.
8. **Ruth Mikelsaar**. Etiological factors of diseases in genetically consulted children and newborn screening: dissertation for the commencement of the degree of doctor of medical sciences. Tartu, 1993.
9. **Agu Tamm**. On metabolic action of intestinal microflora: clinical aspects. Tartu, 1993.
10. **Katrin Gross**. Multiple sclerosis in South-Estonia (epidemiological and computed tomographical investigations). Tartu, 1993.
11. **Oivi Uibo**. Childhood coeliac disease in Estonia: occurrence, screening, diagnosis and clinical characterization. Tartu, 1994.
12. **Viiu Tuulik**. The functional disorders of central nervous system of chemistry workers. Tartu, 1994.
13. **Margus Viigimaa**. Primary haemostasis, antiaggregative and anticoagulant treatment of acute myocardial infarction. Tartu, 1994.
14. **Rein Kolk**. Atrial versus ventricular pacing in patients with sick sinus syndrome. Tartu, 1994.
15. **Toomas Podar**. Incidence of childhood onset type 1 diabetes mellitus in Estonia. Tartu, 1994.
16. **Kiira Subi**. The laboratory surveillance of the acute respiratory viral infections in Estonia. Tartu, 1995.
17. **Irja Lutsar**. Infections of the central nervous system in children (epidemiologic, diagnostic and therapeutic aspects, long term outcome). Tartu, 1995.
18. **Aavo Lang**. The role of dopamine, 5-hydroxytryptamine, sigma and NMDA receptors in the action of antipsychotic drugs. Tartu, 1995.
19. **Andrus Arak**. Factors influencing the survival of patients after radical surgery for gastric cancer. Tartu, 1996.

20. **Tõnis Karki**. Quantitative composition of the human lactoflora and method for its examination. Tartu, 1996.
21. **Reet Mändar**. Vaginal microflora during pregnancy and its transmission to newborn. Tartu, 1996.
22. **Triin Remmel**. Primary biliary cirrhosis in Estonia: epidemiology, clinical characterization and prognostication of the course of the disease. Tartu, 1996.
23. **Toomas Kivastik**. Mechanisms of drug addiction: focus on positive reinforcing properties of morphine. Tartu, 1996.
24. **Paavo Pokk**. Stress due to sleep deprivation: focus on $GABA_A$ receptor-chloride ionophore complex. Tartu, 1996.
25. **Kristina Allikmets**. Renin system activity in essential hypertension. Associations with atherothrombogenic cardiovascular risk factors and with the efficacy of calcium antagonist treatment. Tartu, 1996.
26. **Triin Parik**. Oxidative stress in essential hypertension: Associations with metabolic disturbances and the effects of calcium antagonist treatment. Tartu, 1996.
27. **Svetlana Päi**. Factors promoting heterogeneity of the course of rheumatoid arthritis. Tartu, 1997.
28. **Maarike Sallo**. Studies on habitual physical activity and aerobic fitness in 4 to 10 years old children. Tartu, 1997.
29. **Paul Naaber**. *Clostridium difficile* infection and intestinal microbial ecology. Tartu, 1997.
30. **Rein Pähkla**. Studies in pinoline pharmacology. Tartu, 1997.
31. **Andrus Juhan Voitk**. Outpatient laparoscopic cholecystectomy. Tartu, 1997.
32. **Joel Starkopf**. Oxidative stress and ischaemia-reperfusion of the heart. Tartu, 1997.
33. **Janika Kõrv**. Incidence, case-fatality and outcome of stroke. Tartu, 1998.
34. **Ülla Linnamägi**. Changes in local cerebral blood flow and lipid peroxidation following lead exposure in experiment. Tartu, 1998.
35. **Ave Minajeva**. Sarcoplasmic reticulum function: comparison of atrial and ventricular myocardium. Tartu, 1998.
36. **Oleg Milenin**. Reconstruction of cervical part of esophagus by revascularised ileal autografts in dogs. A new complex multistage method. Tartu, 1998.
37. **Sergei Pakriev**. Prevalence of depression, harmful use of alcohol and alcohol dependence among rural population in Udmurtia. Tartu, 1998.
38. **Allen Kaasik**. Thyroid hormone control over β-adrenergic signalling system in rat atria. Tartu, 1998.
39. **Vallo Matto**. Pharmacological studies on anxiogenic and antiaggressive properties of antidepressants. Tartu, 1998.
40. **Maire Vasar**. Allergic diseases and bronchial hyperreactivity in Estonian children in relation to environmental influences. Tartu, 1998.
41. **Kaja Julge**. Humoral immune responses to allergens in early childhood. Tartu, 1998.

42. **Heli Grünberg**. The cardiovascular risk of Estonian schoolchildren. A cross-sectional study of 9-, 12- and 15-year-old children. Tartu, 1998.
43. **Epp Sepp**. Formation of intestinal microbial ecosystem in children. Tartu, 1998.
44. **Mai Ots**. Characteristics of the progression of human and experimental glomerulopathies. Tartu, 1998.
45. **Tiina Ristimäe**. Heart rate variability in patients with coronary artery disease. Tartu, 1998.
46. **Leho Kõiv**. Reaction of the sympatho-adrenal and hypothalamo-pituitary-adrenocortical system in the acute stage of head injury. Tartu, 1998.
47. **Bela Adojaan**. Immune and genetic factors of childhood onset IDDM in Estonia. An epidemiological study. Tartu, 1999.
48. **Jakov Shlik**. Psychophysiological effects of cholecystokinin in humans. Tartu, 1999.
49. **Kai Kisand**. Autoantibodies against dehydrogenases of α-ketoacids. Tartu, 1999.
50. **Toomas Marandi**. Drug treatment of depression in Estonia. Tartu, 1999.
51. **Ants Kask**. Behavioural studies on neuropeptide Y. Tartu, 1999.
52. **Ello-Rahel Karelson**. Modulation of adenylate cyclase activity in the rat hippocampus by neuropeptide galanin and its chimeric analogs. Tartu, 1999.
53. **Tanel Laisaar**. Treatment of pleural empyema — special reference to intrapleural therapy with streptokinase and surgical treatment modalities. Tartu, 1999.
54. **Eve Pihl**. Cardiovascular risk factors in middle-aged former athletes. Tartu, 1999.
55. **Katrin Õunap**. Phenylketonuria in Estonia: incidence, newborn screening, diagnosis, clinical characterization and genotype/phenotype correlation. Tartu, 1999.
56. **Siiri Kõljalg**. *Acinetobacter* – an important nosocomial pathogen. Tartu, 1999.
57. **Helle Karro**. Reproductive health and pregnancy outcome in Estonia: association with different factors. Tartu, 1999.
58. **Heili Varendi**. Behavioral effects observed in human newborns during exposure to naturally occurring odors. Tartu, 1999.
59. **Anneli Beilmann**. Epidemiology of epilepsy in children and adolescents in Estonia. Prevalence, incidence, and clinical characteristics. Tartu, 1999.
60. **Vallo Volke**. Pharmacological and biochemical studies on nitric oxide in the regulation of behaviour. Tartu, 1999.
61. **Pilvi Ilves**. Hypoxic-ischaemic encephalopathy in asphyxiated term infants. A prospective clinical, biochemical, ultrasonographical study. Tartu, 1999.
62. **Anti Kalda**. Oxygen-glucose deprivation-induced neuronal death and its pharmacological prevention in cerebellar granule cells. Tartu, 1999.
63. **Eve-Irene Lepist**. Oral peptide prodrugs – studies on stability and absorption. Tartu, 2000.

64. **Jana Kivastik**. Lung function in Estonian schoolchildren: relationship with anthropometric indices and respiratory symptomas, reference values for dynamic spirometry. Tartu, 2000.
65. **Karin Kull**. Inflammatory bowel disease: an immunogenetic study. Tartu, 2000.
66. **Kaire Innos**. Epidemiological resources in Estonia: data sources, their quality and feasibility of cohort studies. Tartu, 2000.
67. **Tamara Vorobjova**. Immune response to *Helicobacter pylori* and its association with dynamics of chronic gastritis and epithelial cell turnover in antrum and corpus. Tartu, 2001.
68. **Ruth Kalda**. Structure and outcome of family practice quality in the changing health care system of Estonia. Tartu, 2001.
69. **Annika Krüüner**. *Mycobacterium tuberculosis* – spread and drug resistance in Estonia. Tartu, 2001.
70. **Marlit Veldi**. Obstructive Sleep Apnoea: Computerized Endopharyngeal Myotonometry of the Soft Palate and Lingual Musculature. Tartu, 2001.
71. **Anneli Uusküla**. Epidemiology of sexually transmitted diseases in Estonia in 1990–2000. Tartu, 2001.
72. **Ade Kallas**. Characterization of antibodies to coagulation factor VIII. Tartu, 2002.
73. **Heidi Annuk**. Selection of medicinal plants and intestinal lactobacilli as antimicrobil components for functional foods. Tartu, 2002.
74. **Aet Lukmann**. Early rehabilitation of patients with ischaemic heart disease after surgical revascularization of the myocardium: assessment of health-related quality of life, cardiopulmonary reserve and oxidative stress. A clinical study. Tartu, 2002.
75. **Maigi Eisen**. Pathogenesis of Contact Dermatitis: participation of Oxidative Stress. A clinical – biochemical study. Tartu, 2002.
76. **Piret Hussar**. Histology of the post-traumatic bone repair in rats. Elaboration and use of a new standardized experimental model – bicortical perforation of tibia compared to internal fracture and resection osteotomy. Tartu, 2002.
77. **Tõnu Rätsep**. Aneurysmal subarachnoid haemorrhage: Noninvasive monitoring of cerebral haemodynamics. Tartu, 2002.
78. **Marju Herodes**. Quality of life of people with epilepsy in Estonia. Tartu, 2003.
79. **Katre Maasalu**. Changes in bone quality due to age and genetic disorders and their clinical expressions in Estonia. Tartu, 2003.
80. **Toomas Sillakivi**. Perforated peptic ulcer in Estonia: epidemiology, risk factors and relations with *Helicobacter pylori*. Tartu, 2003.
81. **Leena Puksa**. Late responses in motor nerve conduction studies. F and A waves in normal subjects and patients with neuropathies. Tartu, 2003.
82. **Krista Lõivukene**. *Helicobacter pylori* in gastric microbial ecology and its antimicrobial susceptibility pattern. Tartu, 2003.

83. **Helgi Kolk**. Dyspepsia and *Helicobacter pylori* infection: the diagnostic value of symptoms, treatment and follow-up of patients referred for upper gastrointestinal endoscopy by family physicians. Tartu, 2003.
84. **Helena Soomer**. Validation of identification and age estimation methods in forensic odontology. Tartu, 2003.
85. **Kersti Oselin**. Studies on the human MDR1, MRP1, and MRP2 ABC transporters: functional relevance of the genetic polymorphisms in the *MDR1* and *MRP1* gene. Tartu, 2003.
86. **Jaan Soplepmann**. Peptic ulcer haemorrhage in Estonia: epidemiology, prognostic factors, treatment and outcome. Tartu, 2003.
87. **Margot Peetsalu**. Long-term follow-up after vagotomy in duodenal ulcer disease: recurrent ulcer, changes in the function, morphology and *Helicobacter pylori* colonisation of the gastric mucosa. Tartu, 2003.
88. **Kersti Klaamas**. Humoral immune response to *Helicobacter pylori* a study of host-dependent and microbial factors. Tartu, 2003.
89. **Pille Taba**. Epidemiology of Parkinson's disease in Tartu, Estonia. Prevalence, incidence, clinical characteristics, and pharmacoepidemiology. Tartu, 2003.
90. **Alar Veraksitš**. Characterization of behavioural and biochemical phenotype of cholecystokinin-2 receptor deficient mice: changes in the function of the dopamine and endopioidergic system. Tartu, 2003.
91. **Ingrid Kalev**. CC-chemokine receptor 5 (CCR5) gene polymorphism in Estonians and in patients with Type I and Type II diabetes mellitus. Tartu, 2003.
92. **Lumme Kadaja**. Molecular approach to the regulation of mitochondrial function in oxidative muscle cells. Tartu, 2003.
93. **Aive Liigant**. Epidemiology of primary central nervous system tumours in Estonia from 1986 to 1996. Clinical characteristics, incidence, survival and prognostic factors. Tartu, 2004.
94. **Andres, Kulla**. Molecular characteristics of mesenchymal stroma in human astrocytic gliomas. Tartu, 2004.
95. **Mari Järvelaid**. Health damaging risk behaviours in adolescence. Tartu, 2004.
96. **Ülle Pechter**. Progression prevention strategies in chronic renal failure and hypertension. An experimental and clinical study. Tartu, 2004.
97. **Gunnar Tasa**. Polymorphic glutathione S-transferases – biology and role in modifying genetic susceptibility to senile cataract and primary open angle glaucoma. Tartu, 2004.
98. **Tuuli Käämbre**. Intracellular energetic unit: structural and functional aspects. Tartu, 2004.
99. **Vitali Vassiljev**. Influence of nitric oxide syntase inhibitors on the effects of ethanol after acute and chronic ethanol administration and withdrawal. Tartu, 2004.

100. **Aune Rehema**. Assessment of nonhaem ferrous iron and glutathione redox ratio as markers of pathogeneticity of oxidative stress in different clinical groups. Tartu, 2004.
101. **Evelin Seppet**. Interaction of mitochondria and ATPases in oxidative muscle cells in normal and pathological conditions. Tartu, 2004.
102. **Eduard Maron**. Serotonin function in panic disorder: from clinical experiments to brain imaging and genetics. Tartu, 2004.
103. **Marje Oona**. *Helicobacter pylori* infection in children: epidemiological and therapeutic aspects. Tartu, 2004.
104. **Kersti Kokk**. Regulation of active and passive molecular transport in the testis. Tartu, 2005.
105. **Vladimir Järv**. Cross-sectional imaging for pretreatment evaluation and follow-up of pelvic malignant tumours. Tartu, 2005.
106. **Andre Õun**. Epidemiology of adult epilepsy in Tartu, Estonia. Incidence, prevalence and medical treatment. Tartu, 2005.
107. **Piibe Muda**. Homocysteine and hypertension: associations between homocysteine and essential hypertension in treated and untreated hypertensive patients with and without coronary artery disease. Tartu, 2005.
108. **Külli Kingo**. The interleukin-10 family cytokines gene polymorphisms in plaque psoriasis. Tartu, 2005.
109. **Mati Merila**. Anatomy and clinical relevance of the glenohumeral joint capsule and ligaments. Tartu, 2005.
110. **Epp Songisepp**. Evaluation of technological and functional properties of the new probiotic *Lactobacillus fermentum* ME-3. Tartu, 2005.
111. **Tiia Ainla**. Acute myocardial infarction in Estonia: clinical characteristics, management and outcome. Tartu, 2005.
112. **Andres Sell**. Determining the minimum local anaesthetic requirements for hip replacement surgery under spinal anaesthesia – a study employing a spinal catheter. Tartu, 2005.
113. **Tiia Tamme**. Epidemiology of odontogenic tumours in Estonia. Pathogenesis and clinical behaviour of ameloblastoma. Tartu, 2005.
114. **Triine Annus**. Allergy in Estonian schoolchildren: time trends and characteristics. Tartu, 2005.
115. **Tiia Voor**. Microorganisms in infancy and development of allergy: comparison of Estonian and Swedish children. Tartu, 2005.
116. **Priit Kasenõmm**. Indicators for tonsillectomy in adults with recurrent tonsillitis – clinical, microbiological and pathomorphological investigations. Tartu, 2005.
117. **Eva Zusinaite**. Hepatitis C virus: genotype identification and interactions between viral proteases. Tartu, 2005.
118. **Piret Kõll**. Oral lactoflora in chronic periodontitis and periodontal health. Tartu, 2006.
119. **Tiina Stelmach**. Epidemiology of cerebral palsy and unfavourable neurodevelopmental outcome in child population of Tartu city and county, Estonia Prevalence, clinical features and risk factors. Tartu, 2006.

120. **Katrin Pudersell**. Tropane alkaloid production and riboflavine excretion in the field and tissue cultures of henbane (*Hyoscyamus niger* L.). Tartu, 2006.
121. **Külli Jaako**. Studies on the role of neurogenesis in brain plasticity. Tartu, 2006.
122. **Aare Märtson**. Lower limb lengthening: experimental studies of bone regeneration and long-term clinical results. Tartu, 2006.
123. **Heli Tähepõld**. Patient consultation in family medicine. Tartu, 2006.
124. **Stanislav Liskmann**. Peri-implant disease: pathogenesis, diagnosis and treatment in view of both inflammation and oxidative stress profiling. Tartu, 2006.
125. **Ruth Rudissaar**. Neuropharmacology of atypical antipsychotics and an animal model of psychosis. Tartu, 2006.
126. **Helena Andreson**. Diversity of *Helicobacter pylori* genotypes in Estonian patients with chronic inflammatory gastric diseases. Tartu, 2006.
127. **Katrin Pruus**. Mechanism of action of antidepressants: aspects of serotoninergic system and its interaction with glutamate. Tartu, 2006.
128. **Priit Põder**. Clinical and experimental investigation: relationship of ischaemia/reperfusion injury with oxidative stress in abdominal aortic aneurysm repair and in extracranial brain artery endarterectomy and possibilities of protection against ischaemia using a glutathione analogue in a rat model of global brain ischaemia. Tartu, 2006.
129. **Marika Tammaru**. Patient-reported outcome measurement in rheumatoid arthritis. Tartu, 2006.
130. **Tiia Reimand**. Down syndrome in Estonia. Tartu, 2006.
131. **Diva Eensoo**. Risk-taking in traffic and Markers of Risk-Taking Behaviour in Schoolchildren and Car Drivers. Tartu, 2007.
132. **Riina Vibo**. The third stroke registry in Tartu, Estonia from 2001 to 2003: incidence, case-fatality, risk factors and long-term outcome. Tartu, 2007.
133. **Chris Pruunsild**. Juvenile idiopathic arthritis in children in Estonia. Tartu, 2007.
134. **Eve Õiglane-Šlik**. Angelman and Prader-Willi syndromes in Estonia. Tartu, 2007.
135. **Kadri Haller**. Antibodies to follicle stimulating hormone. Significance in female infertility. Tartu, 2007.
136. **Pille Ööpik**. Management of depression in family medicine. Tartu, 2007.
137. **Jaak Kals**. Endothelial function and arterial stiffness in patients with atherosclerosis and in healthy subjects. Tartu, 2007.
138. **Priit Kampus**. Impact of inflammation, oxidative stress and age on arterial stiffness and carotid artery intima-media thickness. Tartu, 2007.
139. **Margus Punab**. Male fertility and its risk factors in Estonia. Tartu, 2007.
140. **Alar Toom**. Heterotopic ossification after total hip arthroplasty: clinical and pathogenetic investigation. Tartu, 2007.

141. **Lea Pehme**. Epidemiology of tuberculosis in Estonia 1991–2003 with special regard to extrapulmonary tuberculosis and delay in diagnosis of pulmonary tuberculosis. Tartu, 2007.
142. **Juri Karjagin**. The pharmacokinetics of metronidazole and meropenem in septic shock. Tartu, 2007.
143. **Inga Talvik**. Inflicted traumatic brain injury shaken baby syndrome in Estonia – epidemiology and outcome. Tartu, 2007.
144. **Tarvo Rajasalu**. Autoimmune diabetes: an immunological study of type 1 diabetes in humans and in a model of experimental diabetes (in RIP-B7.1 mice). Tartu, 2007.
145. **Inga Karu**. Ischaemia-reperfusion injury of the heart during coronary surgery: a clinical study investigating the effect of hyperoxia. Tartu, 2007.
146. **Peeter Padrik**. Renal cell carcinoma: Changes in natural history and treatment of metastatic disease. Tartu, 2007.
147. **Neve Vendt**. Iron deficiency and iron deficiency anaemia in infants aged 9 to 12 months in Estonia. Tartu, 2008.
148. **Lenne-Triin Heidmets**. The effects of neurotoxins on brain plasticity: focus on neural Cell Adhesion Molecule. Tartu, 2008.
149. **Paul Korrovits**. Asymptomatic inflammatory prostatitis: prevalence, etiological factors, diagnostic tools. Tartu, 2008.
150. **Annika Reintam**. Gastrointestinal failure in intensive care patients. Tartu, 2008.
151. **Kristiina Roots**. Cationic regulation of Na-pump in the normal, Alzheimer's and CCK_2 receptor-deficient brain. Tartu, 2008.
152. **Helen Puusepp**. The genetic causes of mental retardation in Estonia: fragile X syndrome and creatine transporter defect. Tartu, 2009.
153. **Kristiina Rull**. Human chorionic gonadotropin beta genes and recurrent miscarriage: expression and variation study. Tartu, 2009.
154. **Margus Eimre**. Organization of energy transfer and feedback regulation in oxidative muscle cells. Tartu, 2009.
155. **Maire Link**. Transcription factors FoxP3 and AIRE: autoantibody associations. Tartu, 2009.
156. **Kai Haldre**. Sexual health and behaviour of young women in Estonia. Tartu, 2009.
157. **Kaur Liivak**. Classical form of congenital adrenal hyperplasia due to 21-hydroxylase deficiency in Estonia: incidence, genotype and phenotype with special attention to short-term growth and 24-hour blood pressure. Tartu, 2009.
158. **Kersti Ehrlich**. Antioxidative glutathione analogues (UPF peptides) – molecular design, structure-activity relationships and testing the protective properties. Tartu, 2009.
159. **Anneli Rätsep**. Type 2 diabetes care in family medicine. Tartu, 2009.
160. **Silver Türk**. Etiopathogenetic aspects of chronic prostatitis: role of mycoplasmas, coryneform bacteria and oxidative stress. Tartu, 2009.

161. **Kaire Heilman**. Risk markers for cardiovascular disease and low bone mineral density in children with type 1 diabetes. Tartu, 2009.
162. **Kristi Rüütel**. HIV-epidemic in Estonia: injecting drug use and quality of life of people living with HIV. Tartu, 2009.
163. **Triin Eller**. Immune markers in major depression and in antidepressive treatment. Tartu, 2009.
164. **Siim Suutre**. The role of TGF-β isoforms and osteoprogenitor cells in the pathogenesis of heterotopic ossification. An experimental and clinical study of hip arthroplasty. Tartu, 2010.
165. **Kai Kliiman**. Highly drug-resistant tuberculosis in Estonia: Risk factors and predictors of poor treatment outcome. Tartu, 2010.
166. **Inga Villa**. Cardiovascular health-related nutrition, physical activity and fitness in Estonia. Tartu, 2010.
167. **Tõnis Org**. Molecular function of the first PHD finger domain of Autoimmune Regulator protein. Tartu, 2010.
168. **Tuuli Metsvaht**. Optimal antibacterial therapy of neonates at risk of early onset sepsis. Tartu, 2010.
169. **Jaanus Kahu**. Kidney transplantation: Studies on donor risk factors and mycophenolate mofetil. Tartu, 2010.
170. **Koit Reimand**. Autoimmunity in reproductive failure: A study on associated autoantibodies and autoantigens. Tartu, 2010.
171. **Mart Kull**. Impact of vitamin D and hypolactasia on bone mineral density: a population based study in Estonia. Tartu, 2010.
172. **Rael Laugesaar**. Stroke in children – epidemiology and risk factors. Tartu, 2010.
173. **Mark Braschinsky**. Epidemiology and quality of life issues of hereditary spastic paraplegia in Estonia and implemention of genetic analysis in everyday neurologic practice. Tartu, 2010.
174. **Kadri Suija**. Major depression in family medicine: associated factors, recurrence and possible intervention. Tartu, 2010.
175. **Jarno Habicht**. Health care utilisation in Estonia: socioeconomic determinants and financial burden of out-of-pocket payments. Tartu, 2010.
176. **Kristi Abram**. The prevalence and risk factors of rosacea. Subjective disease perception of rosacea patients. Tartu, 2010.
177. **Malle Kuum**. Mitochondrial and endoplasmic reticulum cation fluxes: Novel roles in cellular physiology. Tartu, 2010.
178. **Rita Teek**. The genetic causes of early onset hearing loss in Estonian children. Tartu, 2010.
179. **Daisy Volmer**. The development of community pharmacy services in Estonia – public and professional perceptions 1993–2006. Tartu, 2010.
180. **Jelena Lissitsina**. Cytogenetic causes in male infertility. Tartu, 2011.
181. **Delia Lepik**. Comparison of gunshot injuries caused from Tokarev, Makarov and Glock 19 pistols at different firing distances. Tartu, 2011.
182. **Ene-Renate Pähkla**. Factors related to the efficiency of treatment of advanced periodontitis. Tartu, 2011.

183. **Maarja Krass**. L-Arginine pathways and antidepressant action. Tartu, 2011.
184. **Taavi Lai**. Population health measures to support evidence-based health policy in Estonia. Tartu, 2011.
185. **Tiit Salum**. Similarity and difference of temperature-dependence of the brain sodium pump in normal, different neuropathological, and aberrant conditions and its possible reasons. Tartu, 2011.
186. **Tõnu Vooder**. Molecular differences and similarities between histological subtypes of non-small cell lung cancer. Tartu, 2011.
187. **Jelena Štšepetova**. The characterisation of intestinal lactic acid bacteria using bacteriological, biochemical and molecular approaches. Tartu, 2011.
188. **Radko Avi**. Natural polymorphisms and transmitted drug resistance in Estonian HIV-1 CRF06_cpx and its recombinant viruses. Tartu, 2011, 116 p.
189. **Edward Laane**. Multiparameter flow cytometry in haematological malignancies. Tartu, 2011, 152 p.
190. **Triin Jagomägi**. A study of the genetic etiology of nonsyndromic cleft lip and palate. Tartu, 2011, 158 p.
191. **Ivo Laidmäe**. Fibrin glue of fish (*Salmo salar*) origin: immunological study and development of new pharmaceutical preparation. Tartu, 2012, 150 p.
192. **Ülle Parm**. Early mucosal colonisation and its role in prediction of invasive infection in neonates at risk of early onset sepsis. Tartu, 2012, 168 p.
193. **Kaupo Teesalu**. Autoantibodies against desmin and transglutaminase 2 in celiac disease: diagnostic and functional significance. Tartu, 2012, 142 p.
194. **Maksim Zagura**. Biochemical, functional and structural profiling of arterial damage in atherosclerosis. Tartu, 2012, 162 p.
195. **Vivian Kont**. Autoimmune regulator: characterization of thymic gene regulation and promoter methylation. Tartu, 2012, 134 p.
196. **Pirje Hütt**. Functional properties, persistence, safety and efficacy of potential probiotic lactobacilli. Tartu, 2012, 246 p.
197. **Innar Tõru**. Serotonergic modulation of CCK-4- induced panic. Tartu, 2012, 132 p.
198. **Sigrid Vorobjov**. Drug use, related risk behaviour and harm reduction interventions utilization among injecting drug users in Estonia: implications for drug policy. Tartu, 2012, 120 p.
199. **Martin Serg**. Therapeutic aspects of central haemodynamics, arterial stiffness and oxidative stress in hypertension. Tartu, 2012, 156 p.
200. **Jaanika Kumm**. Molecular markers of articular tissues in early knee osteoarthritis: a population-based longitudinal study in middle-aged subjects. Tartu, 2012, 159 p.
201. **Kertu Rünkorg**. Functional changes of dopamine, endopioid and endocannabinoid systems in CCK2 receptor deficient mice. Tartu, 2012, 125 p.
202. **Mai Blöndal**. Changes in the baseline characteristics, management and outcomes of acute myocardial infarction in Estonia. Tartu, 2012, 127 p.

203. **Jana Lass**. Epidemiological and clinical aspects of medicines use in children in Estonia. Tartu, 2012, 170 p.
204. **Kai Truusalu**. Probiotic lactobacilli in experimental persistent *Salmonella* infection. Tartu, 2013, 139 p.
205. **Oksana Jagur**. Temporomandibular joint diagnostic imaging in relation to pain and bone characteristics. Long-term results of arthroscopic treatment. Tartu, 2013, 126 p.
206. **Katrin Sikk**. Manganese-ephedrone intoxication – pathogenesis of neurological damage and clinical symptomatology. Tartu, 2013, 125 p.
207. **Kai Blöndal**. Tuberculosis in Estonia with special emphasis on drug-resistant tuberculosis: Notification rate, disease recurrence and mortality. Tartu, 2013, 151 p.
208. **Marju Puurand**. Oxidative phosphorylation in different diseases of gastric mucosa. Tartu, 2013, 123 p.
209. **Aili Tagoma**. Immune activation in female infertility: Significance of autoantibodies and inflammatory mediators. Tartu, 2013, 135 p.
210. **Liis Sabre**. Epidemiology of traumatic spinal cord injury in Estonia. Brain activation in the acute phase of traumatic spinal cord injury. Tartu, 2013, 135 p.
211. **Merit Lamp**. Genetic susceptibility factors in endometriosis. Tartu, 2013, 125 p.
212. **Erik Salum**. Beneficial effects of vitamin D and angiotensin II receptor blocker on arterial damage. Tartu, 2013, 167 p.
213. **Maire Karelson**. Vitiligo: clinical aspects, quality of life and the role of melanocortin system in pathogenesis. Tartu, 2013, 153 p.
214. **Kuldar Kaljurand**. Prevalence of exfoliation syndrome in Estonia and its clinical significance. Tartu, 2013, 113 p.
215. **Raido Paasma**. Clinical study of methanol poisoning: handling large outbreaks, treatment with antidotes, and long-term outcomes. Tartu, 2013, 96 p.
216. **Anne Kleinberg**. Major depression in Estonia: prevalence, associated factors, and use of health services. Tartu, 2013, 129 p.
217. **Triin Eglit**. Obesity, impaired glucose regulation, metabolic syndrome and their associations with high-molecular-weight adiponectin levels. Tartu, 2014, 115 p.
218. **Kristo Ausmees**. Reproductive function in middle-aged males: Associations with prostate, lifestyle and couple infertility status. Tartu, 2014, 125 p.
219. **Kristi Huik**. The influence of host genetic factors on the susceptibility to HIV and HCV infections among intravenous drug users. Tartu, 2014, 144 p.
220. **Liina Tserel**. Epigenetic profiles of monocytes, monocyte-derived macrophages and dendritic cells. Tartu, 2014, 143 p.
221. **Irina Kerna**. The contribution of *ADAM12* and *CILP* genes to the development of knee osteoarthritis. Tartu, 2014, 152 p.

222. **Ingrid Liiv**. Autoimmune regulator protein interaction with DNA-dependent protein kinase and its role in apoptosis. Tartu, 2014, 143 p.
223. **Liivi Maddison**. Tissue perfusion and metabolism during intra-abdominal hypertension. Tartu, 2014, 103 p.
224. **Krista Ress**. Childhood coeliac disease in Estonia, prevalence in atopic dermatitis and immunological characterisation of coexistence. Tartu, 2014, 124 p.
225. **Kai Muru**. Prenatal screening strategies, long-term outcome of children with marked changes in maternal screening tests and the most common syndromic heart anomalies in Estonia. Tartu, 2014, 189 p.
226. **Kaja Rahu**. Morbidity and mortality among Baltic Chernobyl cleanup workers: a register-based cohort study. Tartu, 2014, 155 p.
227. **Klari Noormets**. The development of diabetes mellitus, fertility and energy metabolism disturbances in a Wfs1-deficient mouse model of Wolfram syndrome. Tartu, 2014, 132 p.
228. **Liis Toome**. Very low gestational age infants in Estonia. Tartu, 2014, 183 p.
229. **Ceith Nikkolo**. Impact of different mesh parameters on chronic pain and foreign body feeling after open inguinal hernia repair. Tartu, 2014, 132 p.
230. **Vadim Brjalin**. Chronic hepatitis C: predictors of treatment response in Estonian patients. Tartu, 2014, 122 p.
231. **Vahur Metsna**. Anterior knee pain in patients following total knee arthroplasty: the prevalence, correlation with patellar cartilage impairment and aspects of patellofemoral congruence. Tartu, 2014, 130 p.
232. **Marju Kase**. Glioblastoma multiforme: possibilities to improve treatment efficacy. Tartu, 2015, 137 p.
233. **Riina Runnel**. Oral health among elementary school children and the effects of polyol candies on the prevention of dental caries. Tartu, 2015, 112 p.
234. **Made Laanpere**. Factors influencing women's sexual health and reproductive choices in Estonia. Tartu, 2015, 176 p.
235. **Andres Lust**. Water mediated solid state transformations of a polymorphic drug – effect on pharmaceutical product performance. Tartu, 2015, 134 p.
236. **Anna Klugman**. Functionality related characterization of pretreated wood lignin, cellulose and polyvinylpyrrolidone for pharmaceutical applications. Tartu, 2015, 156 p.
237. **Triin Laisk-Podar**. Genetic variation as a modulator of susceptibility to female infertility and a source for potential biomarkers. Tartu, 2015, 155 p.
238. **Mailis Tõnisson**. Clinical picture and biochemical changes in blood in children with acute alcohol intoxication. Tartu, 2015, 100 p.
239. **Kadri Tamme**. High volume haemodiafiltration in treatment of severe sepsis – impact on pharmacokinetics of antibiotics and inflammatory response. Tartu, 2015, 133 p.

240. **Kai Part**. Sexual health of young people in Estonia in a social context: the role of school-based sexuality education and youth-friendly counseling services. Tartu, 2015, 203 p.
241. **Urve Paaver**. New perspectives for the amorphization and physical stabilization of poorly water-soluble drugs and understanding their dissolution behavior. Tartu, 2015, 139 p.
242. **Aleksandr Peet**. Intrauterine and postnatal growth in children with HLA-conferred susceptibility to type 1 diabetes. Tartu. 2015, 146 p.
243. **Piret Mitt**. Healthcare-associated infections in Estonia – epidemiology and surveillance of bloodstream and surgical site infections. Tartu, 2015, 145 p.
244. **Merli Saare**. Molecular Profiling of Endometriotic Lesions and Endometria of Endometriosis Patients. Tartu, 2016, 129 p.
245. **Kaja-Triin Laisaar**. People living with HIV in Estonia: Engagement in medical care and methods of increasing adherence to antiretroviral therapy and safe sexual behavior. Tartu, 2016, 132 p.
246. **Eero Merilind**. Primary health care performance: impact of payment and practice-based characteristics. Tartu, 2016, 120 p.
247. **Jaanika Kärner**. Cytokine-specific autoantibodies in AIRE deficiency. Tartu, 2016, 182 p.
248. **Kaido Paapstel**. Metabolomic profile of arterial stiffness and early biomarkers of renal damage in atherosclerosis. Tartu, 2016, 173 p.
249. **Liidia Kiisk**. Long-term nutritional study: anthropometrical and clinico-laboratory assessments in renal replacement therapy patients after intensive nutritional counselling. Tartu, 2016, 207 p.
250. **Georgi Nellis**. The use of excipients in medicines administered to neonates in Europe. Tartu, 2017, 159 p.
251. **Aleksei Rakitin**. Metabolic effects of acute and chronic treatment with valproic acid in people with epilepsy. Tartu, 2017, 125 p.
252. **Eveli Kallas**. The influence of immunological markers to susceptibility to HIV, HBV, and HCV infections among persons who inject drugs. Tartu, 2017, 138 p.
253. **Tiina Freimann**. Musculoskeletal pain among nurses: prevalence, risk factors, and intervention. Tartu, 2017, 125 p.
254. **Evelyn Aaviksoo**. Sickness absence in Estonia: determinants and influence of the sick-pay cut reform. Tartu, 2017, 121 p.
255. **Kalev Nõupuu**. Autosomal-recessive Stargardt disease: phenotypic heterogeneity and genotype-phenotype associations. Tartu, 2017, 131 p.
256. **Ho Duy Binh**. Osteogenesis imperfecta in Vietnam. Tartu, 2017, 125 p.
257. **Uku Haljasorg**. Transcriptional mechanisms in thymic central tolerance. Tartu, 2017, 147 p.
258. **Živile Riispere**. IgA Nephropathy study according to the Oxford Classification: IgA Nephropathy clinical-morphological correlations, disease progression and the effect of renoprotective therapy. Tartu, 2017, 129 p.

259. **Hiie Soeorg**. Coagulase-negative staphylococci in gut of preterm neonates and in breast milk of their mothers. Tartu, 2017, 216 p.
260. **Anne-Mari Anton Willmore**. Silver nanoparticles for cancer research. Tartu, 2017, 132 p.
261. **Ott Laius**. Utilization of osteoporosis medicines, medication adherence and the trend in osteoporosis related hip fractures in Estonia. Tartu, 2017, 134 p.
262. **Alar Aab**. Insights into molecular mechanisms of asthma and atopic dermatitis. Tartu, 2017, 164 p.
263. **Sander Pajusalu**. Genome-wide diagnostics of Mendelian disorders: from chromosomal microarrays to next-generation sequencing. Tartu, 2017, 146 p.
264. **Mikk Jürisson**. Health and economic impact of hip fracture in Estonia. Tartu, 2017, 164 p.
265. **Kaspar Tootsi**. Cardiovascular and metabolomic profiling of osteo-arthritis. Tartu, 2017, 150 p.