DISSERTATIONES
MEDICINAE
UNIVERSITATIS
TARTUENSIS
296

JULIA MASLOVSKAJA

The importance of DNA binding and DNA breaks for AIRE-mediated transcriptional activation

DISSERTATIONES MEDICINAE UNIVERSITATIS TARTUENSIS
296

DISSERTATIONES MEDICINAE UNIVERSITATIS TARTUENSIS
296

JULIA MASLOVSKAJA

The importance of DNA binding and DNA breaks for AIRE-mediated transcriptional activation

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 November 20th, 2019 by the Council of the Faculty of Medicine, University of Tartu, Estonia

Supervisors: Pärt Peterson, PhD, Professor of Molecular Immunology, Molecular Pathology Research Group, Institute of Biomedicine and Translational Medicine, Faculty of Medicine, University of Tartu, Estonia

Ana Rebane, PhD, Professor of Molecular Medicine, RNA Biology Research Group, Institute of Biomedicine and Translational Medicine, Faculty of Medicine, University of Tartu, Estonia

Reviewers: Aleksandr Žarkovski, PhD, Professor of Pharmacology and Toxicology, Institute of Biomedicine and Translational Medicine, Faculty of Medicine, University of Tartu, Estonia

Radko Avi, PhD, Senior Research Fellow in Medical Virology, Institute of Biomedicine and Translational Medicine, Faculty of Medicine, University of Tartu, Estonia

Opponent: Jorma J. Palvimo, PhD, Professor of Medical Biochemistry, Institute of Biomedicine, University of Eastern Finland, Finland

Commencement: January 30th, 2020

Publication of this dissertation is granted by the University of Tartu

ISSN 1024-395X
ISBN 978-9949-03-250-1 (print)
ISBN 978-9949-03-251-8 (pdf)

University of Tartu Press
www.tyk.ee

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

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

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

III Guha, M.[1], Saare, M.[1], Maslovskaja, J.[1], Kisand, K., Liiv, I., Haljasorg, U., Tasa, T., Metspalu, A., Milani, L., Peterson, P., 2017. DNA breaks and chromatin structural changes enhance the transcription of autoimmune regulator target genes. J. Biol. Chem. 292, 6542–6554.

[1]These authors contributed equally to this work

Contribution of Julia Maslovskaja to the original publications:

Study I	participating in the study design, performing all the experiments, analyzing the data, writing the manuscript
Study II	performing cell sorting and qPCR experiments, participating in writing the manuscript
Study III	participating in the study design, performing RNA-seq experiments, participating in the analysis of experimental data and figure preparation, co-writing the manuscript

ABBREVIATIONS

aa	amino acid
AdML	adenovirus major late
AIRE	autoimmune regulator
AIRE-Tet	HEK293 cell line with doxycycline-inducible AIRE expression
APECED	autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy
bp	base pair
BRD4	bromodomain-containing protein 4
CBP	CREB-binding protein
CDK	cyclin-dependent kinase
CTCF	CCCTC-binding factor
CTD	C-terminal domain
cTEC	cortical thymic epithelial cell
Ctrl	uninduced AIRE-Tet cells
DE	differentially expressed
DEU	differential exon usage
DNA-PK(cs)	DNA-dependent protein kinase (catalytic subunit)
Dox	doxycycline-induced AIRE-Tet cells
Dox+Etop	doxycycline-induced and etoposide-treated AIRE-Tet cells
DSB	double-strand break
DSIF	DRB sensitivity-inducing factor
EMSA	electrophoretic mobility shift assay
Etop	etoposide-treated AIRE-Tet cells
FACS	fluorescence-activated cell sorting
GST	glutathione S-transferase
GTF	general transcription factor
γH2A.X	histone H2A.X phosphorylated at serine 139
H3K4me1	histone 3 monomethylated at lysine 4
H3K4me3	histone 3 trimethylated at lysine 4
H3K9me3	histone 3 trimethylated at lysine 9
H3K27ac	histone 3 acetylated at lysine 27
HEK293	human embryonic kidney cell line 293
HMG	high-mobility group
HSR/CARD	homogeneously staining region/caspase recruitment domain
IF	immunofluorescence
IFN	interferon
Inr	initiator element
INS	insulin
IP	immunoprecipitation
IVL	involucrin

KAP-1	KRAB-associated protein-1
Luc	luciferase
MHC	major histocompatibility complex
MNase	micrococcal nuclease
mRNA	messenger RNA
mTEC	medullary thymic epithelial cell
NELF	negative elongation factor
NF-κB	nuclear factor kappa-light-chain-enhancer of activated B cells
NHEJ	non-homologous end joining
NLS	nuclear localization signal
PARP-1	poly(ADP-ribose) polymerase 1
PHD	plant homeodomain
PIC	preinitiation complex
Pol II	RNA polymerase II
P-TEFb	positive transcription elongation factor b
qPCR	quantitative polymerase chain reaction
RNA-seq	RNA sequencing
SAND	Sp100, AIRE-1, NucP41/75, DEAF-1 domain
SD	standard deviation
SEC	super elongation complex
SEM	standard error of the mean
shRNA	short hairpin RNA
SL	stem-loop
snRNP	small nuclear ribonucleoprotein
SRSF	serine/arginine-rich splicing factors
SSB	single-strand break
TCR	T-cell receptor
TF	transcription factor
TOP	DNA topoisomerase
Treg	regulatory T cell
TSA	tissue-specific antigen
TSS	transcription start site
wt	wild-type

1. INTRODUCTION

The most important task of the immune system is to recognize harmful substances and elicit response against them. The ability to make difference between dangerous and harmless molecules comprises the capacity to distinguish between organism's own and foreign constituents. The failure to achieve this task may manifest in autoimmune diseases. Although it seems very intuitive, the discrimination between self and non-self is not a trivial task. The immune system uses several mechanisms to differentiate between foreign substances and its own components. The more straightforward and earlier evolved system is the pattern recognition by the innate immunity – the ability to sense conserved elements of microbes that are not commonly present in the human organism. Such molecular combinations include, for example, components of the bacterial membrane or viral double-stranded RNAs (Janeway and Medzhitov, 2002).

The adaptive immune system exploits another approach – it generates a huge variety of receptors by random shuffling gene segments in immunoglobulin and T-cell receptor (TCR) genomic regions. The generated receptors are able to recognize a large number of different peptides, irrespective of their origin. These randomly generated receptors ensure that even highly mutated pathogens are detected by the immune system. However, the randomly generated repertoire of immunoglobulins and TCRs potentially creates the risk of targeting the body's own tissues. The adaptive immune system is complex: the synthesis of enormous diversity of receptors also requires a set of complicated mechanisms to ensure the elimination or inactivation of cells carrying unsuitable TCR combinations (Abbas and Janeway, 2000; Bonilla and Oettgen, 2010; Litman et al., 2010).

The characterization of the autoimmune regulator (AIRE) protein in 1997 opened a new perspective on how efficient elimination of the cells with the potentially damaging autoreactive receptors is achieved. For the selection mechanism, the presentation of peptides on the surface of thymic stromal cells is required. However, for peptides to reach the cell surface the process should start much earlier with the induction of gene expression. AIRE is one of the most puzzling transcriptional regulators responsible for the removal of potentially self-reactive T cells. AIRE-mediated transcription does not obey conventional rules and is characterized by several unusual features influencing its peculiar choice of target genes. In attempt to characterize the rationale for AIRE transcriptional targets' choice, we investigated the target gene elements determining the activation by AIRE and explored AIRE interaction with DNA and proteins supporting its function. The findings that help to elucidate AIRE-mediated gene activation mechanism not only provide valuable knowledge of the immune system functioning but also highlight the complexity and the diversity of approaches in regulation of eukaryotic transcription.

2. REVIEW OF LITERATURE

2.1. Regulation of transcription

Although all cells of organism contain the same set of genes, cell properties and functions are dictated by the combination of genes being expressed in a cell (Tupler et al., 2001). Therefore, the gene expression is tightly regulated to suit the functional needs of a specific cell. To follow the central dogma of molecular biology (Crick, 1970) the information is transferred from DNA to RNA during the process of transcription and further to protein during translation. Multiple additional steps and factors in both processes influence the functional product amount, including RNA processing, splicing and stability, translation efficiency, post-translational protein modifications, protein activation by proteolysis and protein secretion (Schwanhäusser et al., 2011). To describe cell state, often messenger RNA (mRNA) molecule diversity and quantity is used, suggesting that the abundance of mRNAs correlates with protein molecules (Liu et al., 2016). Transcriptome analysis is often preferable nowadays because it is easier, quicker and cheaper to perform than proteome analysis; however, it should be considered that transcriptome data provides only a slice of information. Recognizing all the additional steps between processed RNA and protein, the discrepancies between mRNA and protein quantity are not surprising (Pradet-Balade et al., 2001; Vogel and Marcotte, 2012). Despite these restrictions, no protein would be made without transcript available. While all layers of control contribute and sophisticate the production of protein, transcriptional control remains one major regulation step (Levine and Tjian, 2003).

2.1.1. Transcription process

Transcription rate is regulated during initiation and elongation steps and is affected by RNA processing and splicing (Venkatesh and Workman, 2015). Eukaryotic RNA polymerase II (Pol II), the enzyme responsible for RNA synthesis of protein-coding genes and most of noncoding regulatory RNAs, is not able to recognize the promoters of target genes by itself. C-terminal domain (CTD) of Pol II largest subunit consist of heptad (Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7) repeated 52 times in vertebrates and can be subjected to modifications important for transcriptional regulation (Harlen and Churchman, 2017). General transcription factors (GTF), such as TFIIA, -B, -D, -E, -F and -H are needed for the recognition of promoter regions and for the initiation of transcription. Together with Pol II they form a basic transcription complex to initiate gene expression (Liu et al., 2013). The formation of preinitiation complex (PIC) is promoted by transcription factors (TFs) and cofactors bound at enhancers, which interact with GTFs and Pol II (Haberle and Stark, 2018) (Figure 1A). Among GTFs, TFIIH possesses helicase module, responsible for double-stranded DNA unwinding and promoter melting (Guzmán and Lis, 1999), and kinase module CDK7 (cyclin-dependent kinase 7), which phosphorylates Ser5 and Ser7 of Pol II CTD (Akoulitchev et al., 1995; Egloff, 2012; Phatnani and Greenleaf, 2006) (Figure 1B). The

Ser5 phosphorylation is necessary for Pol II dissociation from GTFs and advancing downstream – the process, called promoter escape (Buratowski, 2009). Mediator complex and p300 assist in phosphorylation of Pol II to mobilize the promoter escape (Poss et al., 2013). After synthesizing short stretches of nascent RNA, Pol II pauses 30–50 bp downstream of transcription start site (TSS) (Adelman and Lis, 2012). The Ser5 phosphorylation is also needed for recruiting the capping enzymes to a newly synthesized RNA strand (Bentley, 2014). Capping, the attachment of 7-methylguanosine at 5'-end of nascent RNA, is important for the stability of RNA and the prevention of promoter-proximal transcription termination, as uncapped RNAs are degraded by nucleases, and the pausing is aborted consequently (Henriques et al., 2013).

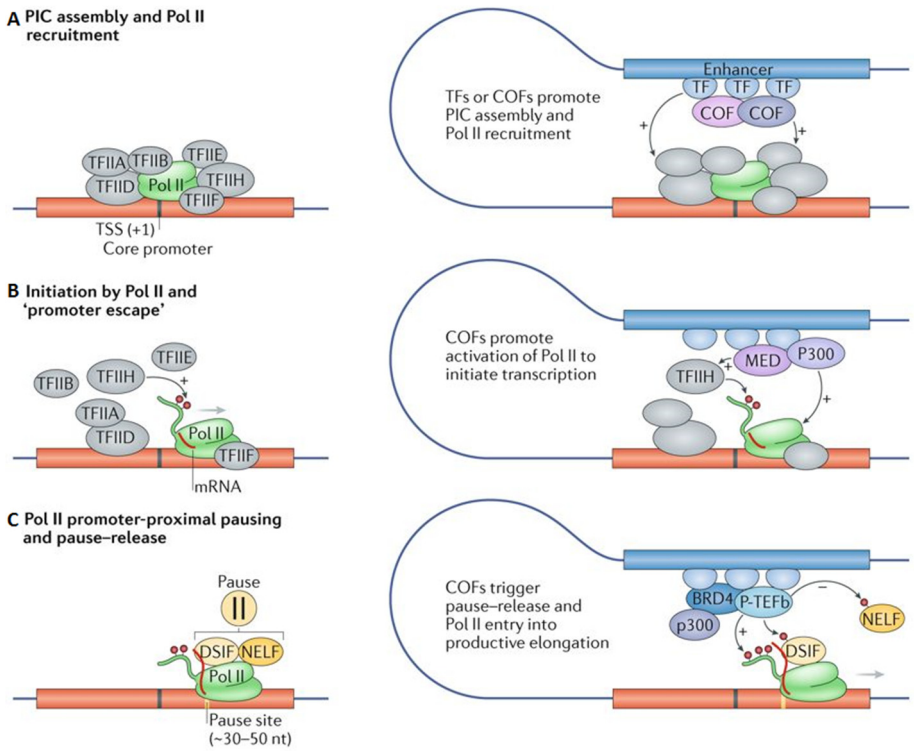

Figure 1. The stages and regulation of transcription. (A) Preinitiation complex (PIC) assembly at core promoter consisting of GTFs and Pol II. Enhancer-bound transcription factors (TF) and cofactors (COF) facilitate PIC formation. **(B)** TFIIH helicase activity is required for DNA duplex unwinding at promoter that enables RNA synthesis from transcription start site (TSS). TFIIH kinase activity is needed for Pol II CTD Ser5 and Ser7 phosphorylation that promotes dissociation from GTFs and promoter escape. Mediator complex (MED) or histone acetyltransferase p300 assist in this process through enhancer-promoter contact. **(C)** Pol II promoter-proximal pausing and pause release. After synthesizing 30-50 nt, Pol II pauses with DSIF and NELF factors stabilizing the complex. P-TEFb evokes pause release by phosphorylating DSIF, NELF and Ser2 of Pol II CTD. This allows the entry of Pol II into productive elongation. BRD4 and p300 reinforce transition into the productive elongation (modified from Haberle and Stark, 2018).

The pausing factors NELF (negative elongation factor) and DSIF (DRB sensitivity-inducing factor) are required for maintaining Pol II pausing and preventing the pause release (Wu et al., 2003) (Figure 1C). The P-TEFb (positive transcription elongation factor) composed of the catalytic subunit cyclin-dependent kinase 9 (CDK9) and its regulatory subunit, cyclin-T1, is required for the release of Pol II into productive elongation (Peterlin and Price, 2006). It phosphorylates DSIF, NELF and Pol II CTD Ser2 (Brès et al., 2008) (Figure 1C). The P-TEFb can be a part of at least three larger complexes: SEC- (super elongation complex), BRD4- (bromodomain-containing protein 4) and 7SK-associated (Chen et al., 2018). 7SK snRNP (small nuclear ribonucleoprotein) together with HEXIM (hexamethylene bisacetamide inducible) proteins sequesters and inhibits a large portion of P-TEFb (90%) (Zhou et al., 2012). In contrast, SEC- and BRD4-associated complexes are active. BRD4 that recognizes acetylated lysine residues, including those at N-terminal ends of histones, through bromodomain is a positive regulator of P-TEFb and augments the phosphorylation of Pol II CTD through the action of CDK9 (Jang et al., 2005) (Figure 1C). The CDK9-independent function of BRD4 was also demonstrated, as BRD4 protein degradation hinders transcription elongation while CDK9 recruitment to target genes is not affected (Winter et al., 2017). Multiple other factors influence the transition of P-TEFb from repressive complexes into active ones, including c-Myc (Rahl et al., 2010), KAP-1 (KRAB-associated protein-1) (also known as TRIM28) (McNamara et al., 2016) and SRSF (serine/arginine-rich splicing factors) (Ji et al., 2013), providing links between chromatin landscape, elongation and splicing. It should be noted that, although the stages of transcription are defined and studied as distinct processes, they are thoroughly interrelated (Moore and Proudfoot, 2009).

Initially, paused Ser5-phosphorylated Pol II was found near TSSs in quickly-inducible genes in the absence of activating signals. The first gene described exploiting such mechanism was heat shock protein 70 (*Hsp70*) in *Drosophila melanogaster* (Gilmour and Lis, 1986). Later, paused Pol II presence has been demonstrated at many promoters (Core et al., 2008), including house-keeping genes (Rougvie and Lis, 1990). In most metazoans, the majority of genes exhibit transient pausing with varying time of Pol II pause, ranging from a few minutes for the majority of genes to an hour for stably paused genes (Chen et al., 2015). Responses to environmental and developmental cues are mediated through Pol II pause-release (Chen et al., 2018).

2.1.2. Chromatin

In addition to complex assembly at TSS and enhancer-promoter contact establishment, the surrounding chromatin structure can influence transcription rate (Fuda et al., 2009). For efficient packaging into chromatin, genomic DNA is wrapped around octamer consisting of four pairs of histones H2A, H2B, H3 and H4, forming a nucleosome unit and providing additional layer for transcriptional regulation via repositioning nucleosomes or modifying histones (Campos and

Reinberg, 2009). In the nucleosome, N-terminal regions of histone molecules are protruding and can be subjected to numerous covalent modifications. All four histone molecules can be methylated, acetylated, phosphorylated, ubiquitylated or modified by some other rare mechanisms (Kouzarides, 2007). The array of enzymes is responsible for establishing modifications and at least as many enzymes for removing them. Most prevailing modifications are histone acetylation and methylation. Enzymes involved in directing these modifications are acetyltransferases (such as CREB-binding protein (CBP)/p300 and histone acetyltransferase 1 (HAT1)) adding acetyl residues to lysines of histones and deacetylases (such as sirtuins (SIRT)) erasing this modification, and lysine methyltransferases, an example being EZH2 (enhancer of zeste homolog 2) belonging to PRC2 (polycomb repressive complex 2) methylating histone H3 at lysine 27 (H3K27me3) and lysine demethylases (such as jumonji D (JMJD) proteins) with opposing function (Kouzarides, 2007).

Certain histone modifications are characteristic for particular regulatory elements and specific processes. For example, enhancers are enriched with histone H3 monomethylated at lysine 4 (H3K4me1) and acetylated at lysine 27 (H3K27ac); active promoters are trimethylated at histone H4 lysine 4 (H3K4me3) and enriched with H3/H4 acetylation; active gene bodies contain ubiquitinylated H2 (H2Bub), H3 trimethylated at lysine 36 (H3K36me3) and lysine 79 (H3K79me3); active promoters with abovementioned modifications contrast repressive promoters characterized by H3 trimethylated at lysine 9 (H3K9me3) (Millar and Grunstein, 2006). Despite strong correlation of histone modifications with abovementioned activities, their functional association is not always clear. For instance, enhancer mark H3K4me1 and active transcription mark H3K4me3 are shown to be dispensable without affecting enhancer and transcriptional activity (Hödl and Basler, 2012; Pollex and Furlong, 2017).

In addition to covalent modifications of histone tails, histone variants exist that differ from canonical histones in a few amino acids or larger sequence stretches (Buschbeck and Hake, 2017). H3.3 and H2A.Z replacement is involved in regulation of transcription (Cairns, 2009), and H2A.X has major role in DNA damage response. The phosphorylation of H2A.X at Ser139 (γH2A.X) is the first step occurring after DNA damage and is accomplished by kinases involved in DNA repair pathways, such as ataxia telangiectasia mutated (ATM) and ATM-Rad3-related (ATR) (Kinner et al., 2008).

In addition to enzymes that modify histones, proteins containing domains recognizing histone modifications are important in mediating various activities. Bromodomain containing proteins (for instance BRD4) recognize acetylated lysines (Dey et al., 2003), chromo- and PHD (plant homeodomain) domains–containing proteins usually recognize some level of lysine methylation. For example, chromodomain-helicase-DNA-binding protein 4 (CHD4) have preference for unmodified histone H3K4 and methylated histone H3K9 (Mansfield et al., 2011), while H3K4me3 is recognized by inhibitor of growth (ING2) (Shi et al., 2006) and bromodomain and PHD domain transcription factor (BPTF) (Li et al., 2006).

Histones are also one example of proteins binding to DNA without sequence specificity. The binding is mediated by positive charges of basic amino acids within histone and negatively charged nucleic acid. Additional examples include protamines used for DNA condensation in sperm cells (Braun, 2001) and architectural proteins participating in DNA compaction in prokaryotes (Dame, 2005). One interesting example of DNA recognition in a non-specific manner is HMG (high mobility group)-box containing proteins that recognize DNA distortions (Štros et al., 2007). Ku proteins, involved in DNA repair, also bind DNA in sequence-independent manner having affinity for free DNA ends (Blier et al., 1993).

2.1.3. Regulatory sequences and transcription factors

A core promoter is a short sequence that embeds TSS, spanning ~50 bp upstream and downstream from TSS (Haberle and Stark, 2018). It serves as a site for the assembly of the transcription machinery consisting of Pol II and GTFs. Core promoters support low basal transcriptional activity that can be modified by chromatin context and distal regulatory elements, such as enhancers. Enhancers bind regulatory proteins – TFs possessing DNA-binding domain, and cofactors that do not directly bind DNA but participate in regulating the transcription through protein-protein interactions (Zabidi and Stark, 2016). Cohesin and Mediator complexes have important roles in enabling long-range interactions bringing into the proximity promoters and enhancers (Chen et al., 2018; Kagey et al., 2010). Despite the importance and the necessity of core promoters, they are not well defined in eukaryotes. TATA-box, the core-promoter motif conserved among organisms, is quite rare and present in only 5% of core promoters in flies (Ohler et al., 2002). Other, more abundant, core-promoter motif Inr (initiator) was initially defined in humans as a dinucleotide combination of a pyrimidine followed by a purine, the most preferred CG, CA, TG (Carninci et al., 2006), followed by later reports that some additional nucleotides are also important for higher informational context (Vo Ngoc et al., 2017). Nucleosome positioning in the proximity of TSS could be decisive for gene expression initiation with higher importance than primary nucleotide sequence, leading to focused or dispersed distribution of TSS (Carninci et al., 2006).

The activation of transcription could be achieved by establishing contact between promoter and enhancer through protein interactions or by binding of transcription factors to the preformed enhancer-promoter interaction (Ghavi-Helm et al., 2014). Enhancers regulate genes in *cis*; however, their location is variable as they can localize upstream, downstream or within introns of genes as well as they do not necessarily act on the closest promoter and can regulate multiple genes (Pennacchio et al., 2013). These features make the identification of enhancers difficult; however, enhancers share common characteristics, such as chromatin modifications H3K4me1 and H3K27ac, and high p300 and Mediator proteins' amount (Chen et al., 2018).

Historically, the term transcription factor could be applied to any protein involved in transcription and capable of influencing gene expression. Currently, this term implies sequence-specific DNA binding and regulation of transcription (Lambert et al., 2018). The sequence specificity of TFs can vary broadly (Mitchell and Tjian, 1989). Around 1500 sequence-specific DNA-binding TFs have been described with the research mostly concentrated on the disease-causing TFs. The most cited transcription factors in the literature are p53, ER (estrogen receptor), FOS, MYC, JUN, SP1 (specificity protein 1), NF-κB (nuclear factor kappa-light-chain-enhancer of activated B cells), CREB1 (cAMP responsive element binding protein 1) (Vaquerizas et al., 2009). The TFs can be classified based on the structure of their DNA-binding domains. Among the DNA-binding protein structures, three types dominate in the human genome; the C2H2 zinc-finger, homeodomain and helix-loop-helix, which together cover over 80% of the transcription factors (Vaquerizas et al., 2009). Besides DNA-binding domain, TFs possess an effector domain that could assist ligand binding (as in hormone-dependent factors), acquire enzymatic activity (for instance, histone-modifying enzymes) or mediate protein-protein interaction. The majority of DNA sequence-specific TFs have short DNA recognition motifs (6-12 bp) enabling a large number of recognition sites in genome, whereas experimentally determined binding sites of TF and recognition sequences only partially overlap (Lambert et al., 2018). Additional mechanism, such as synergy and cooperativity, compensate for low TF specificity (Reiter et al., 2017). Furthermore, experimentally determined binding sites of TFs do not entirely coincide with the genes they regulate (Cusanovich et al., 2014).

More recently, gene expression models were disputed, following the genome-wide analysis reports that many genomic positions outside annotated gene start sites can initiate transcription (Kapranov et al., 2007; The ENCODE Project Consortium, 2007). These include unstable enhancer RNAs produced bidirectionally from active enhancers (Kim et al., 2010) and upstream antisense RNAs produced at promoters in the opposite direction from mRNA (Preker et al., 2008). The divergent transcription at unusual sites is initiated in the similar way as at defined TSS (Core et al., 2014), even including Pol II pause (Henriques et al., 2018). On the other hand, long stable mRNAs could be produced only from a core promoter in the sense direction as those transcripts are stabilized after initiation by the presence of 5'-splice sites and by the absence of premature polyadenylation signals (Almada et al., 2013). As a result of this flexibility, strong activating factors could potentially induce transcription at any sequence, as achieving perfect activation specificity is unjustifiable (Haberle and Stark, 2018).

2.1.4. RNA processing

The sequence of mammalian genes is interrupted by long stretches of introns that are removed from pre-mRNA before it is transported from the nucleus into the cytoplasm for translation. Splicing occurs in concert with transcription, and these processes are mutually influenced by each other (Ji et al., 2013; Lin et al., 2008). In addition to splicing, RNA processing comprises editing, base modification, and most important for the stability, transport and translation of RNA molecule, modifications of 5'- and 3'-ends – capping and polyadenylation (Garneau et al., 2007). Although splicing is the main component for increased isoform diversity (Graveley, 2001), it is now appreciated that over half of the mRNAs in the human genome are also alternatively polyadenylated. Alternative 3'-end untranslated region adds variety to microRNA targeting, export, subcellular localization and tissue-specific expression (Tian et al., 2005).

Formation of the poly(A) tail occurs in a two-step reaction. Firstly, the pre-mRNA is cleaved at a site defined by two signals, a highly conserved upstream AAUAAA polyadenylation sequence and a downstream G/U-rich sequence. Cleavage of the nascent transcripts by CPSF (cleavage and polyadenylation specificity factor) occurs 18–30 bp downstream of a polyadenylation signal (Porrua and Libri, 2015). Thereafter, the poly(A) tail of ~200 adenosines in mammals is added to the 3′-end of the RNA, catalyzed by polyadenylate polymerase (Lutz and Moreira, 2011). Mammalian expression vector plasmids commonly contain the sequence motif AAUAA that promotes both polyadenylation and termination, the most efficient termination signals include additional helper sequences to prevent run-through transcription (Schek et al., 1992).

The metazoan replication-dependent histone mRNAs are unusual as they are the only eukaryotic mRNAs that lack poly(A) tails. These RNAs are produced mainly in the S-phase of somatic cells to supply histones for packaging of newly replicated DNA into chromatin. Typically, the replication-dependent histone mRNAs lack introns and their genes are arranged in clusters (Dávila López and Samuelsson, 2008). Instead of poly(A) tail, they contain an RNA stem–loop (SL) structure close to the 3′-end of the mature RNA, and this 3′-end is generated by cleavage involving the U7 snRNP and protein factors, such as the stem–loop binding protein (SLBP). The processing of SL 3′-end and polyadenylation are related as some protein components are shared between the two systems (Dávila López and Samuelsson, 2008).

2.1.5. Topoisomerases in transcription

Though substantial knowledge is accumulated about TFs and chromatin architecture, little is known about how mechanical barriers, such as tension created by unwinding of double-stranded DNA during the process of transcription, are tackled. The movement of RNA Pol II molecule along the template strand of DNA double-stranded helix creates negative (or underwound) supercoiling

behind, and positive (or overwound) supercoiling ahead of the enzyme (Figure 2). Negative supercoiling facilitates the separation of strands, promoter opening and transcription bubble forming (Ma and Wang, 2016). At the same time, negative supercoiling can lead to the generation of R-loops, three-stranded structures, where the template strand of DNA hybridizes with nascent RNA and the non-template strand is unpaired (Figure 2) (Pommier et al., 2016). R-loops can impede transcription elongation (Tous and Aguilera, 2007), additionally unpaired DNA strand that loops out is more susceptible to DNA damage compromising genomic integrity (Skourti-Stathaki and Proudfoot, 2014). Accumulation of positive supercoiling in front of advancing polymerase mechanically hinders its movement along the DNA strand, possibly assisting Pol II pausing (Roca, 2011).

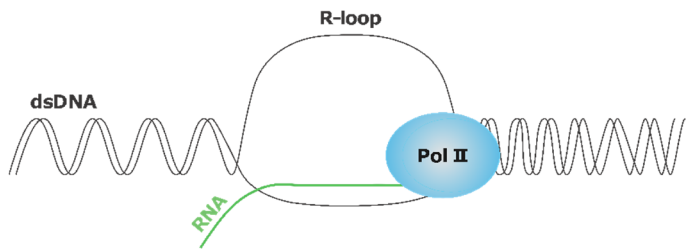

Figure 2. **Negative and positive DNA supercoiling in the process of transcription.** Negative supercoiling facilitates DNA unwinding at promoter area, positive supercoiling accumulating in the front of proceeding Pol II contributes to promoter-proximal pausing.

To resolve torsional tension topoisomerases (TOPs) catalyze the reaction of DNA cleavage and subsequent free DNA ends resealing. Six DNA topoisomerases (TOP1, TOP1mt, TOP2α, TOP2β, TOP3α and TOP3β) exist in human cells (Pommier et al., 2016). Functionally, TOPs are divided into two classes, type I and type II, depending on whether they cleave one or two strands of DNA, respectively. Type I TOPs cleave one strand of double-stranded helix, and are divided based on mechanism how they relieve torsional stress, which is either by passing one strand through the break (IA) or by rotating around the other strand (IB) (Chen et al., 2013). Type II TOPs cut both strands of a DNA duplex, allowing a second intact duplex to pass through (Baranello et al., 2013) (Figure 3). Both types relax DNA by producing transient covalent bonds between conserved tyrosine of protein and phosphate of DNA; the break is quickly re-ligated after relieving topological tension (Deweese and Osheroff, 2009). Human enzymes TOP1 and TOP1mt belong to type IB; TOP3α and TOP3β to type IA; and TOP2α and TOP2β to type IIA. Reflecting the essential role of TOPs in various cellular processes, the mouse knock-outs for *Top1* and *Top2a* are lethal at embryo stage, and *Top2b* knock-out dies at birth (Pommier et al., 2016). TOP2α is absolutely required for chromosome segregation, TOP2β is indispensable for transcription in differentiated, non-dividing cells (Pommier et al., 2016).

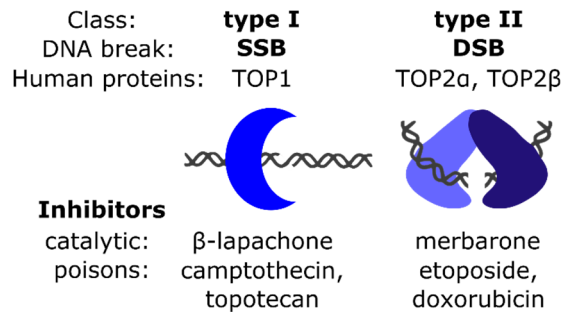

Figure 3. TOP type I and II mechanism of action. Human TOP1 promotes single-strand break (SSB) formation and resolves topological tension by rotating (type IB), TOP2 generates double-strand breaks (DSB) creating the possibility of passing one DNA duplex through the other (type IIA). The examples of catalytic and poison inhibitors of TOP1 and TOP2 are given.

The activity of TOPs can be negatively affected by two different mechanisms. If the catalytic activity is blocked, the creation of breaks is inhibited (Pommier, 2013). Alternatively, if the re-ligation after DNA cleavage reaction is prevented, TOP is fixed to DNA and the break existence is prolonged, which may lead to apoptosis (Collins et al., 2001). Therapeutic TOP inhibitors use the mechanism of trapping protein-DNA complexes (enzyme poisons) rather than inhibiting the catalytic activity. One of the well-known TOP2 poisons, etoposide, derived from *Podophyllum* plant (Hande, 1998), stabilizes protein-DNA complex, hampers re-ligation, and creates DSBs (Montecucco and Biamonti, 2007). Excessive number of DNA breaks leads to cytotoxicity, and therefore etoposide and other analogous TOP inhibitors are widely used in cancer treatments (Hande, 1998). While etoposide is long known and thoroughly studied as the agent causing cytotoxicity and cell apoptosis (Hande, 1998), its impact on transcription is poorly explored (Montecucco et al., 2015).

The functional role of each TOP enzyme depends on the type of topological task, surrounding chromatin landscape, and the presence of accessory factors. TOP1 and TOP2 differ in their substrates influencing distinct processes. For example, TOP2 is more efficient in relaxation of chromatin (Salceda et al., 2006), whereas TOP1 is not proficient on nucleosomal templates and rather operates on accessible DNA molecules (Zobeck et al., 2010). In addition, the transcriptional activity of low-output promoters is enabled by relieving the topological stress by TOP1, whereas both TOP1 and TOP2 are needed to handle the torsional forces induced by the transcription of highly active promoters (Kouzine et al., 2013).

The chromatin looping that brings distal regulatory elements into the proximity of promoter may also create topological tension that requires TOP activity. TOP1 has been shown to associate with the transcription of β-globin locus control

region (LCR), one of the best-studied examples of enhancer-promoter interaction in gene regulation (Rosenberg et al., 2013). Similarly, TOP2 was detected at tissue-specific enhancer of the immunoglobulin kappa gene (Cockerill and Garrard, 1986). TOP2 together with cohesin and CTCF (CCCTC-binding factor) proteins also delineates the boundaries of topological domains (Uusküla-Reimand et al., 2016). Considering the wide distribution and the tight involvement of TOPs with many cellular processes, it is surprising that their transcriptional effects are understudied. Only several reports claim TOPs involvement in transcription (Kawano et al., 2010), demonstrating TOP effect on long (King et al., 2013), inflammatory (Rialdi et al., 2016), neuronal (Appiah et al., 2007; Sano et al., 2008), highly expressed (Kouzine et al., 2013), and hormone-dependent genes (Ju et al., 2006).

2.1.6. DNA-PK complex in transcription

Though it seems counterintuitive, several reports have suggested that the generation of breaks and involvement of DNA damage repair mechanisms may be necessary for promotion of efficient transcription (Bunch, 2016). On one hand, DNA breaks are potential source of cellular damage, on the other hand, they are necessary to resolve DNA tension generated by replication, recombination and transcription (Puc et al., 2017).

DNA-PK (DNA-dependent protein kinase) function is mostly associated with non-homologous end joining (NHEJ) DNA repair pathway, which is the most prevalent type of DSB repair in mammals (Lieber et al., 2003). DNA-PK, a serine/threonine protein kinase complex, is composed of a heterodimer of Ku proteins (Ku70/Ku80) and the catalytic subunit DNA-PKcs (Figure 4). Ku proteins bind to double-stranded DNA ends with high affinity in a sequence-independent manner via a central ring formed by Ku70 and Ku80 subunits (Fell and Schild-Poulter, 2015). When bound, Ku proteins recruit DNA-PKcs forming a heterotrimeric complex (Burma and Chen, 2004). As the name implies, DNA-PK is able to phosphorylate DNA-associated proteins at serine/threonine residues, including itself at several sites, and this kinase activity is induced upon DNA-PK binding with DNA (Pawelczak et al., 2011). Besides NHEJ process, DNA-PK is also implicated in V(D)J recombination, and mutation in the *PRKDC* gene encoding the DNA-PKcs protein results in SCID (severe combined immunodeficiency) phenotype lacking T and B cells in mice and humans (Blunt et al., 1995; van der Burg et al., 2009).

Surprisingly, Ku80 protein involvement in reinitiation of transcription (Woodard et al., 2001) as well as its association with Pol II elongation sites (Mo and Dynan, 2002) was identified independently of its role in DNA repair.

An intriguing work showed the necessity of TOP2β-mediated DSBs at promoter of pS2 gene in human breast adenocarcinoma MCF7 cells for estrogen-induced transcription activation (Ju et al., 2006). Estrogen stimulation induced recruitment of TOP2β, DNA-PK, Ku70/80, PARP-1 (poly(ADP-ribose) poly-

merase 1), CBP and Pol II to the promoter and evoked nucleosomal changes. Subsequent works corroborated the concept that DNA cleavage is required for transcriptional activation and demonstrated this mechanism in neuronal early-response genes (Appiah et al., 2007), neuronal genes in AT-rich genomic environment (Sano et al., 2008), developmentally regulated genes in murine brains (Lyu et al., 2006) and in nuclear receptor-mediated transcription (Trotter et al., 2015).

TOP-mediated transient breaks are short-lived as re-ligation process is very fast (Deweese and Osheroff, 2009), but prolonged misalignment of DNA ends bound to the enzyme can lead to DNA lesion with genotoxic potential (McClendon and Osheroff, 2007). This has led to the hypothesis of scheduled physiological controlled DNA breaks, however, it is not clear how discrimination between harmful and beneficial DNA breaks is accomplished (Puc et al., 2017). One possibility is that the presence of certain proteins near actively transcribed genes directs the choice of DNA repair pathway, as it is known that different types of DNA lesions, for example induced by ionizing radiation or alkylating agents, require specific repair pathways (Helleday et al., 2008). One example how dangerous consequences of DNA breaks near transcribed genes are avoided is tyrosyl-DNA phosphodiesterase 2 (TDP2) requirement in NHEJ process of TOP2-mediated transcription in breast cancer cells, where TDP2-dependent NHEJ suppresses chromosomal translocations (Gómez-Herreros et al., 2017).

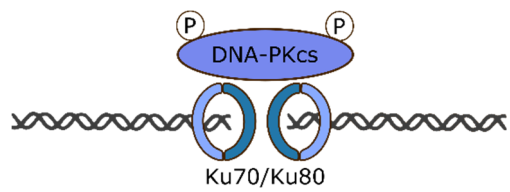

Figure 4. DNA-PK heterotrimeric complex formation at the site of double-strand DNA break. Ku70/Ku80 dimers cover free ends of DNA and interact with DNA-PK catalytic subunit (cs), that is able to phosphorylate (P) itself and many other proteins at Ser/Thr residues upon binding to DNA.

2.2. T-cell tolerance

The immune tolerance can be divided into the central and peripheral parts to highlight their different roles and anatomical locations. The central part embeds the primary lymphoid organs, the thymus and bone marrow, in which the maturation processes of the immune cells occur. The peripheral part consists of the secondary lymphoid organs, lymph nodes and the spleen, as well as aggregates of lymphoid tissue in mucosa, skin, bronchi and gut, and is necessary for the efficient antigen presentation and cell activation (Waldmann, 2016; Xing and Hogquist, 2012). Though the distinction between central and peripheral immune

tolerance appears strict, the involvement and the interplay of the whole immune system is necessary to achieve normal immune response (Parkin and Cohen, 2001).

2.2.1. Thymus

T cells complete their maturation program in the thymus and for that are called thymocytes during the differentiation process. The histological areas of the thymic cortex (outer area) and medulla (inner area) can be discriminated and are functionally devoted to different stages of T-cell selection. The cortical and medullary thymic epithelial cells (cTECs and mTECs) (Nitta and Suzuki, 2016) and different subtypes of dendritic cells (Lopes et al., 2015; Oh and Shin, 2015) are the major players in the thymic selection processes. Nevertheless, the cellular composition of the thymus is more heterogeneous, including B cells (Yamano et al., 2015), macrophages that digest apoptotic cells, fibroblasts (Gray et al., 2007; Sun et al., 2015), endothelial, myoid (Mesnard-Rouiller et al., 2004) and recently identified epithelial tuft-like cells (Bornstein et al., 2018; Miller et al., 2018).

Thymus originates from the third pharyngeal pouch of an embryo (Rodewald, 2008), its organogenesis and proper histological structure formation is critically dependent on Forkhead box N1 (*FoxN1*) gene (Vaidya et al., 2016). Both cTECs and mTECs evolve from the endodermal germ layer (Rodewald, 2008) and share common TEC progenitors (Rossi et al., 2006). mTEC-committed progenitors can be defined by the expression of the tight junction proteins claudin-3 and -4 (Cldn4 and 4) (Hamazaki et al., 2007), and their further differentiation heavily depends on the signals provided by maturing thymocytes (Lopes et al., 2015).

Thymus has an unusual property to decrease in size and lose its functional mass with the age (Singh and Singh, 1979). Because of involution process, thymus was thought to have rudimentary role, however, it is demonstrated that even in elderly thymus produces competent T cells, although in smaller amounts (Jamieson et al., 1999). Healthy organism compensates this diminished capacity by division of naïve cells to maintain the population (Murray et al., 2003), while circulating memory cells defend from recurrent infections (Aspinall and Andrew, 2000). However, this remaining capability appears critical if organism is subjected to immunological insults as radiation, HIV infection or chemotherapy (Lynch et al., 2009). On contrary, neonatal period is when manipulations with thymus have radical consequences. Neonatally thymectomized mice are unable to mount immune response against infections or reject foreign tissues (Miller, 1962). At the same time, allogenic antigens introduced at this time window are perceived as self-molecules (Billingham et al., 1953).

2.2.2. Central tolerance: positive and negative selection

To complete the differentiation into T cells, progenitors migrate from the bone marrow via blood into the thymus (Petrie and Zúñiga-Pflücker, 2007). An important step in T-cell maturation is the generation of TCR by randomly rearranging genomic segments via V(D)J recombination (Spits, 2002). The successful accomplishment of the maturation presumes surmounting two barriers. First, T cells have to pass a positive selection that takes place in the cortex of the thymus and assures survival of the cells recognizing organism's major histocompatibility complex (MHC). After the positive selection, the survived cells travel into the medulla for the second challenge – the negative selection, where the cells recognizing self-antigens presented by MHC on the surface of thymic stroma are deleted via apoptosis (Figure 5). To enhance the presentation of produced peptides to maturing thymocytes, several approaches are possible – to expand the cell population presenting antigens by transfer of antigens from mTECs to B cells and dendritic cells (Hubert et al., 2011; Klein et al., 2011), by cell membrane exchange (Millet et al., 2008), by releasing exosomes (Skogberg et al., 2015), as well as to intensify autophagy in TECs (Nedjic et al., 2008). As a result of the thymic selection processes, the functional immune system generates T cells, which express major histocompatibility complex (MHC)-restricted TCRs and at the same time are tolerant to self-antigens.

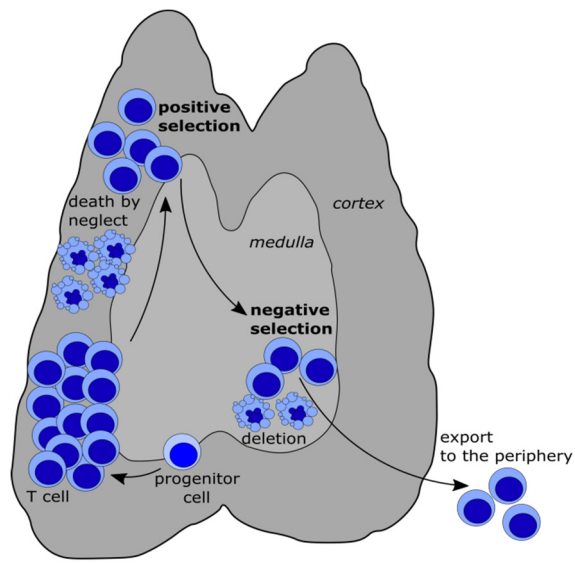

Figure 5. Positive and negative selection in the thymus. Immature T cells in the thymus, also called thymocytes, develop from hematopoietic progenitor cells and initially pass positive selection in the cortex, which assures survival of thymocytes recognizing organism's major histocompatibility complex (MHC). Survived T cells move into the medulla where cells recognizing self-antigens presented by MHC on the surface of thymic stroma are deleted via apoptosis. T cells that successfully completed negative selection are exported to the periphery.

The majority of differentiating thymocytes with randomly rearranged receptors are unsuccessful in recognizing MHC complex as they cannot form functional TCR on the cell membrane (Hogquist et al., 2005). Without receiving a positive survival signal, T cells are determined to the death by neglect (Kyewski and Klein, 2006; Palmer, 2003). Approximately 90% of cells die at this stage (Palmer, 2003). The fate of minority survived T cells, having membrane receptors capable of interacting properly with MHC-peptide complex, depends on the presence of antigens in the thymus (Malhotra et al., 2016) and the strength of interaction with MHC complex (Klein et al., 2014). Three major outcomes of negative selection are possible (Figure 6). T cells survive and exit to the periphery if none or minimal amount of antigen they can potentially recognize is present in the thymus (Figure 6A). Those T cells can later elicit immune response if they encounter a foreign antigen (bacterial, viral or modified self-molecules) in activating conditions in the periphery. When self-antigen is present, two contrasting scenarios may occur. Firstly, a T cell can commit to a regulatory T-cell (Treg) lineage with the ability to suppress autoimmune response in the periphery should it be inappropriately induced (Millar and Ohashi, 2016) (Figure 6B). This mechanism is preferentially used to avoid autoimmune attack against antigens specifically expressed in certain organs such as tissue-specific antigens (TSAs). Alternatively, thymocytes with TCRs engaging very strongly with MHC-peptide complex undergo apoptosis within the thymus (Palmer, 2003; Venanzi et al., 2004) (Figure 6C). Antigens evoking such response can be housekeeping proteins expressed by the majority of cells, circulating antigens in the blood, or antigens expressed by mTECs at high level (Malhotra et al., 2016).

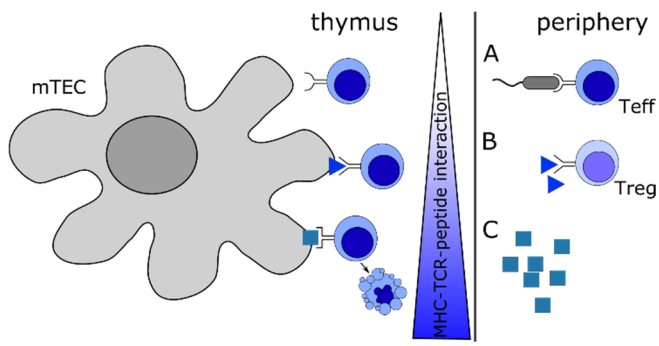

Figure 6. Three outcomes of mTEC-thymocyte interaction. When no antigen is present in the thymus, a naïve T cell exits to the periphery with potential to become an effector T cell (Teff) and to elicit immune response to foreign antigens **(A)**. If the TCR interacts with MHC-peptide with intermediate strength, a cell is directed to regulatory T-cell (Treg) lineage with the ability to dampen inappropriate immune reaction in the periphery **(B)**. If the TCR binds to the peptide with strong affinity, a cell commits apoptotic death in the thymus **(C)**. Antigens in (B) and (C) are derived from self-proteins and present in both the thymus and the periphery, an antigen in (A) is foreign and absent in the thymus.

The unusual capacity of mTECs to express tissue-specific genes was initially demonstrated with an insulin promoter-driven transgene in murine models, as the expression was detectable in the thymus in addition to the pancreas (Heath et al., 1992; von Herrath, Dockter, & Oldstone, 1994), and endogenous insulin expression was demonstrated later in human (Vafiadis et al., 1997), mouse and rat thymi. Other pancreatic enzymes (Antonia et al., 1995) and hormones from other organs were also detected in the thymus (Fuller and Verity, 1989; Geenen et al., 1986). The importance of this mechanism was initially underestimated as it was believed that TSAs from distant locations would not reach the thymus and are not presented to the thymocytes (Kyewski and Klein, 2006). The discovery of the *AIRE* gene and characterization of its protein functions in the late 1990s renewed emphasis on central tolerance. The findings demonstrated that AIRE triggers the transcription of thousands TSAs and explained how peptides derived from peripherally restricted sites are expressed in the thymus, and thus can be recognized by thymocytes (Anderson, 2002).

2.3. AIRE

The human *AIRE* gene is located on chromosome 21 and consists of 14 exons encoding 545 aa long protein with molecular mass of 57.5 kDa. The mouse and human Aire/AIRE proteins share essential domains and sequence homology (73%) (Blechschmidt et al., 1999; Mittaz et al., 1999).

2.3.1. Mutations in *AIRE* gene cause APECED autoimmune disease

In 1997, two research groups identified the *AIRE* gene by studying families with rare disease named **a**utoimmune **p**oly**e**ndocrinopathy-**c**andidiasis-**e**ctodermal **d**ystrophy (APECED), also known as autoimmune polyendocrinopathy syndrome 1 (APS1) (OMIM 240300, Orphanet 3453) (Nagamine et al., 1997; The Finnish-German APECED Consortium, 1997). The disease is characterized by several autoimmune manifestations against many (predominantly endocrine) organs leading to multiple variable symptoms including candidiasis, hypoparathyroidism, Addison's disease, hypothyroidism, diabetes mellitus, autoimmune hepatitis, gastritis with pernicious anemia, ovarian and testicular failure, together with ectodermal manifestations such as alopecia, vitiligo, dental enamel dysplasia and nail dystrophy (Kisand and Peterson, 2011). The mucocutaneous *Candida* infection together with autoimmune reactions against parathyroid and adrenal cortex are the three most frequent symptoms, and the presence of two of them satisfies the criteria for diagnosis establishment (Ahonen et al., 1990). Consequently, the most frequent symptoms patients display are recurrent candidiasis, muscle cramps and paresthesia due to hypocalcemia seen in hypoparathyroidism, hypotension and fatigue due to the lack of adrenal hormones in Addison's disease

(Husebye et al., 2009). The first reports of disease as co-occurrence of chronic candidiasis, hypoparathyroidism and Addison's disease in children (Leonard, 1946; Sutphin et al., 1943; Thorpe and Handley, 1929) were described long before the genetic cause was discovered. While the first symptoms emerge in the early childhood, the variable combination of symptoms and the differences in the disease course can delay the correct diagnosis for many years (Mazza et al., 2011).

The patients develop multiple autoantibodies, of which some are associated with specific symptoms, for instance, autoantibodies against steroidogenic cytochromes (steroid 17-α- and 21-hydroxylases, side-chain cleavage enzyme) are present together with adrenocortical and hypogonadal dysfunction (Kisand and Peterson, 2011). However, the correlation between autoantibodies and disease manifestations is not always strict; despite high prevalence of autoantibodies specific for autoimmune hepatitis (anti-AADC – aromatic L-amino acid decarboxylase), the disease itself is quite rare in APECED patients (Kisand and Peterson, 2011). Almost all patients develop autoantibodies against type 1 interferon (IFN)-ω and –α (Wolff et al., 2013), although the reason for the early appearance of these neutralizing, high-titer autoantibodies is not known (Kisand et al., 2011). The redefining and expanding diagnostic criteria with inclusion of anti-IFN-ω testing could help to prevent the delay of diagnosis and provide better management for patients (Ferre et al., 2016; Antonella Meloni et al., 2008). Moreover, assays for anti-IFN antibodies can complement genetic testing and provide precise answers if unusual mutations are present. The autoimmune component of candidiasis was not obvious until the demonstration of neutralizing autoantibodies against Th17-related interleukins IL-17A, IL17F and IL-22, which presence is associated with fungal infection in most of APECED patients (Kisand et al., 2010).

APECED is more common among populations with founder effects such as Finns (1:25000), Persian Jews (1:9000) and Sardinians (1:14000). The disease has lower prevalence (1:100000) in other geographical regions (Kisand and Peterson, 2011). Over a hundred disease-causing mutations have been described at the moment (Abramson and Husebye, 2016); however, the prevalence of some hotspot mutations dominates among patients from certain populations. For example, a mutation R139X is prevalent among Sardinian and R257X among Finnish patients; both mutations result in premature stop codon and cause the deletion of important protein domains. The deletion of 13 bp in exon 8 (964del13) is common in APECED patients originating from different countries (Bruserud et al., 2016). The correlations between phenotypic features and AIRE mutations are not clear and are difficult to decipher due to the variability of symptoms (Capalbo et al., 2012). Nevertheless, certain features seem to be more prevalent in distinct populations, such as patients with the major Finnish mutation R257X are more prone to type 1 diabetes and candidiasis (Fierabracci, 2016), whereas *Candida* infection in Persian Jews with Y85C mutation is relatively rare (Kisand et al., 2011). The pattern of recessive inheritance is typical for the majority of APECED mutations including those described above (Vogel et al., 2002).

However, a dominant mutation of G228W was described in Italian family with autoimmune thyroiditis (Cetani et al., 2001) and was shown to cause autoimmunity in mice with additional unusual neurological features (Su et al., 2008). The dominant mutations in the *AIRE* gene, reported in 2015 (Oftedal et al., 2015), are linked to variable autoimmune phenotypes, for instance pernicious anemia or vitiligo, not satisfying the criteria for APECED diagnosis. These dominant mutations are missense mutations affecting PHD1 finger (for instance, C311Y, V301M, D312N) and showed dominant-negative effect on the expression of AIRE-dependent genes in transfected cells (Oftedal et al., 2015).

2.3.2. AIRE role in thymic tolerance

Several important roles for AIRE in central tolerance were suggested including induction of promiscuous gene expression (Anderson, 2002; Liston et al., 2003), influence on alternative splicing (Keane et al., 2014), presentation and transfer of self-antigens (Hubert et al., 2011), and shaping the repertoire of regulatory T cells (Malchow et al., 2013). The transcriptional activation of genes is considered to be the key function of AIRE. Multiple studies have shown that AIRE is able to activate luciferase and other reporters with various promoters as well as endogenous genes in cell culture transfections (Björses et al., 2000; The Finnish-German APECED Consortium, 1997). In mTECs, AIRE activates a plethora of genes, which are otherwise expressed in specific cell types and are tightly regulated by tissue-specific TFs (Anderson, 2002; Derbinski et al., 2001). The unusual capacity of AIRE to promiscuously activate a large number of genes allows maturing thymocytes to be selected for self-antigens (Anderson, 2002). As for now, the most plausible explanation for APECED disease manifestations is that due to decreased transcription of peripheral antigens, the maturing thymocytes do not get access to many self-antigens and autoreactive T cells escape to the periphery causing autoimmune destruction. This has been mechanistically demonstrated in Aire-deficient mouse models with certain antigens including eye antigen IRBP (interphotoreceptor retinoid-binding protein) (DeVoss et al., 2006) and stomach antigen mucin 6 (Gavanescu et al., 2007).

AIRE location on chromosome 21 makes it an interesting candidate in association with autoimmunity seen in Down syndrome patients with trisomy of the corresponding chromosome (Rabinowe et al., 1989). In one report the diminished AIRE expression has been demonstrated despite the excess of chromosomal material (Giménez-Barcons et al., 2014; Lima et al., 2011), while others observed increased mRNA levels of AIRE and its target genes (Skogberg et al., 2014). In both cases normal histological structure of the thymus was disrupted. Interestingly, the development of autoimmunity in mice with augmented Aire expression has been described (Nishijima et al., 2018), though the targeted tissue in this mouse model was muscle and not endocrine organs.

Aire-deficient mice share similar phenotypic autoimmune features with APECED patients, although target organs of autoimmune attack vary and diffe-

rent antibody spectrum is present (Hubert et al., 2009). This may reflect the inter-species differences of the immune systems, gestational courses, or other features (Kisand et al., 2014; Pöntynen et al., 2006). As a convenient model organism, Aire-deficient mice have been extensively used to explore the function of AIRE, however, the reported phenotypic differences should be taken into consideration and conclusions should be cautiously extrapolated (Peterson et al., 2008). The generation of AIRE-deficient rat model could help to decipher APECED disease features, as in those animals APECED similar manifestations are more pronounced (Ossart et al., 2018).

mTEC cell population can be divided further on the basis of MHC class II expression level into $mTEC^{hi}$ and $mTEC^{lo}$ subsets. $mTEC^{hi}$ is more mature stage derived from $mTEC^{lo}$ and besides high level of MHC class II acquires surface expression of co-stimulatory molecules CD80 and CD86, supporting effective antigen presentation (Derbinski et al., 2005). All mTECs are in principle committed to express Aire (Kawano et al., 2015), however, as it occurs at certain differentiation stage, only around 50-60% of $mTEC^{hi}$ cell population express Aire in mice at the given moment (Gray et al., 2007). The expression of AIRE in $mTEC^{hi}$ cells in the thymus is dependent on receptor activator of nuclear factor κB (RANK) expressed on the surface of mTECs and ligands for this receptor provided by hematopoietic cells (Bichele et al., 2016; Rossi et al., 2007). An enhancer situated ~3kb upstream of the *Aire* coding sequence in mice is crucial for mediating RANK signaling and thymic expression of Aire (Haljasorg et al., 2015). Interestingly, Aire expression in the testes is preserved after deleting enhancer sequence, assuming that other regulatory mechanisms exist in reproductive system. Other TFs important for Aire expression were characterized, including interferon regulatory factor (Irf) 4, Irf8, T-box 21 (Tbx21), T-cell factor 7 (Tcf7) and transcriptional repressor Ctcfl (CTCF-like) acting through binding to Aire promoter (Herzig et al., 2016). Using NOD (non-obese diabetic) mice models with time and dose controllable switch of Aire expression, the Aire protein was shown to be of critical importance during perinatal stage as turning off Aire in mice after day 21 did not impact later immune status (Guerau-de-Arellano et al., 2009).

Almost all Aire-positive mTECs are post-mitotic and show high turnover rate (Gray et al., 2007). The stage with high Aire expression, prompting TSA transcription and high level of molecules involved in presentation on the cell surface, might be the culmination point for mTECs but not the last one. Post-Aire cells with decreased expression of Aire, TSAs, MHC class II and CD80/86 molecules, but upregulated keratinocyte-specific genes, are described (Metzger et al., 2013; Wang et al., 2012). After loss of nuclei these post-Aire cells form the cornified structures of Hassal's corpuscles in the thymus.

Although the expression of AIRE is most prominent in thymic medullary cells, there are other cell types of lymphoid and reproductive systems where the presence of lower amounts of the AIRE protein and mRNA has been demonstrated, including stromal and other cells in secondary lymphoid organs (Gardner et al., 2013, 2008; Yamano et al., 2019), B cells (Yamano et al., 2015), dendritic

cells (Fergusson et al., 2019; Kogawa et al., 2002), testes (Schaller et al., 2008) and early embryos (Nishikawa et al., 2010). The role and importance of AIRE in other cell types require further investigation.

Some hints that AIRE is not the only one responsible for the expression of whole scope of antigens in mTECs appeared, as not all autoimmunity-provoking TSAs are downregulated in Aire-deficient thymus (Kuroda et al., 2005; Niki et al., 2006). Another transcription factor, also implicated in nervous system development, Fezf2 (FEZ [forebrain embryonic zinc-finger] family zinc finger 2) was demonstrated to regulate the expression of TSAs in the thymus (Takaba et al., 2015).

Deviation of a thymocyte to Treg lineage is not entirely clear but is thought to occur if intermediate affinity signal is received during peptide-MHC-TCR interaction (Klein et al., 2019). The initial studies reported no differences between adult Aire wild-type (wt) and knock-out mice in frequency and function of Treg compartment (Lei et al., 2011; Yang et al., 2015) and autoimmunity developing after the simultaneous co-transfer of $Aire^{+/+}$ and $Aire^{-/-}$ thymi into nude mouse hosts (Anderson et al., 2005; Kuroda et al., 2005), contradicting the mechanism of dominant tolerance. However, certain repertoire of Treg selection is dependent on Aire (Malchow et al., 2013), whether this relates to specific antigens (Leonard et al., 2017) or the antigen presentation way (Perry et al., 2014). Tumor-associated Tregs also depend on Aire (Malchow et al., 2016, 2013), providing potential therapeutic implication of transient blockade of central tolerance to enhance anti-tumor response (Bakhru et al., 2017; Khan et al., 2014).

2.3.3. AIRE protein domains

AIRE has an unusual combination of functional domains, including domains common for transcriptional regulator. It contains HSR/CARD (**h**omogenously **s**tained **r**egion/**ca**spase **r**ecruitment **d**omain) that is present in apoptosis-related protease enzymes and is responsible for oligomerization and interaction with other proteins; **n**uclear **l**ocalization **s**ignal (NLS) responsible for driving the protein into the nucleus; SAND (**S**p100, **A**IRE-1, **N**ucP41/75, **D**EAF-1) domain with potential ability to bind DNA; two PHD (**p**lant **h**omeo**d**omain) fingers characteristic for transcriptional regulators separated by **p**roline-**r**ich **r**egion (PRR); and four LxxLL motifs that mediate protein-protein interactions associated with transcriptional regulation (Plevin et al., 2005) (Figure 7). Such set of functional units strongly suggests its role in transcription in the nucleus.

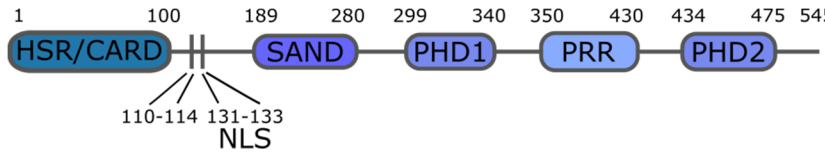

Figure 7. The human AIRE protein domains. The AIRE protein contains the following domains: homogenously stained region/caspase recruitment domain (HSR/ CARD), the nuclear localization signal (NLS), SAND (Sp100, AIRE-1, NucP41/75, DEAF-1) domain, plant homeodomains 1 and 2 (PHD1 and PHD2), proline-rich region (PRR), and four interspersed LxxLL (L – Leu, x – any amino acid) sequence motifs (not depicted). The numbers of amino acids are indicated.

AIRE was described to locate in large >670 kDa complexes (Halonen et al., 2004) and form characteristic dots in the nucleus (Björses et al., 1999). Interestingly, N-terminal domain of AIRE is responsible for oligomerization (Pitkänen et al., 2000), copious interactions with other proteins, and harbors many patient mutations that disrupt dotted localization pattern in the nucleus producing diffuse staining instead (Halonen et al., 2004). N-terminal part of AIRE (aa 1-100) was initially named the homogenously staining region (HSR) domain and later found to possess high structural similarity with the caspase recruitment domain (CARD) (Ferguson et al., 2008). AIRE ability to evoke apoptosis was demonstrated in mTECs (Gray et al., 2007) and spermatogonia (Schaller et al., 2008).

AIRE has nuclear signal and predominantly localizes inside the nucleus with some cytoplasmic staining (Rinderle et al., 1999). Initially, the NLS was found between amino acids 113-133 (The Finnish-German APECED Consortium, 1997), but later the mutagenesis studies demonstrated that only short 131-133 amino acid stretch is necessary to transport the AIRE protein into the nucleus (Ilmarinen et al., 2006). Some authors argue that AIRE NLS potentially still functions as bipartite with upstream conserved aa 110-111 replacing mutated aa 113-114 (Saltis et al., 2008). Consensus sequence of monopartite NLS at aa 159-167 was determined by comparative analysis of evolutionary conserved motifs, but this potential NLS was never confirmed functionally (Saltis et al., 2008). In addition, undefined NLS in the C-terminus was proposed as constructs lacking N-terminal domain were still efficiently transported into the nucleus (Pitkänen et al., 2001), however, never mapped to certain amino acids and not confirmed by others (Ramsey et al., 2002).

SAND domain has been claimed to be responsible for DNA binding; however, the core recognition amino acid sequence in the AIRE protein is modified and the results of AIRE SAND domain binding to DNA are inconclusive (Kumar et al., 2001; Purohit et al., 2005; Zumer et al., 2012). The lysines in NLS, SAND, and the area between them are reported to be acetylated by CBP/p300; acetylation of

K243 and K253 leads to the formation of larger nuclear bodies and the decrease of AIRE transactivation activity (Saare et al., 2012)

The interaction of PHD1 with histone H3 nonmethylated at K4 was first described by Org *et al* (Org et al., 2008) and later confirmed by Koh *et al* (Koh et al., 2010). This finding explained how AIRE is able to trigger the induction of low expressed genes characterized by the absence of H3K4me3. However, as PHD1 mutation did not abolish the upregulation of all genes (Org et al., 2008) and histone demethylase treatment only partially changed the set of genes activated by AIRE (Koh et al., 2010), other determinants should have a role in selection of the target genes. Earlier, PHD1 was also proposed to have E3-ubiquitin ligase function (Uchida et al., 2004).

A role of PHD2 is not well known and has been associated with different functions. Yang *et al* (Yang et al., 2013) demonstrated that PHD2 is required for binding chromatin structure proteins, while Gaetani *et al* (Gaetani et al., 2012) suggested that it is important for proper AIRE localization to nuclear bodies.

Some authors emphasize the importance of C-terminus attributing AIRE transactivation domain (TAD) properties to 30 (Alessandra Meloni et al., 2008) or 57 C-terminal aa (Žumer et al., 2011).

2.3.4. AIRE interacting proteins

Proteins interacting with AIRE provide great interest, as they potentially unveil the mechanism of AIRE-mediated regulation. The cooperation with CBP, ubiquitous transcriptional coactivator possessing histone acetyltransferase activity, was one the earliest described (Pitkänen et al., 2000). While AIRE and CBP synergistically activated promoters, their co-localization *in vivo* was rare. Interestingly, histone deacetylase Sirt1 with the opposite function was also claimed to be important for AIRE functioning (Chuprin et al., 2015). Among other proteins interacting with AIRE the following were detected: E3 SUMO ligase PIAS1 (protein inhibitor of activated STAT1) (Ilmarinen et al., 2008), the apoptosis- and transcription-implicated protein DAXX (death domain-associated protein) (Meloni et al., 2010), p63 necessary for effective MHC class II expression in mTECs (Tonooka et al., 2009), ATF7ip-MBD1 (activating transcription factor 7 interacting protein–methyl CpG binding protein 1), both linked to repressive epigenetic markers – ATF7ip is a cofactor implicated in the generation of the repressive histone mark H3K9me3, and MBD1 recognizes methylated CpGs (Waterfield et al., 2014).

Several proteins crucial for transcriptional elongation have been shown to interact with AIRE. P-TEFb, protein decisive for switching from promoter-proximal pausing to elongation stage, was demonstrated to be recruited by AIRE to its target genes (Oven et al., 2007) with subsequent evidence that patient mutation in AIRE disrupting its interaction with P-TEFb results in the disease (Žumer et al., 2011). BRD4 that binds free P-TEFb and recruits it to the promoters was also demonstrated to be critical for AIRE functioning (Yoshida et al., 2015), although

others demonstrated that AIRE interaction with P-TEFb is BRD4-independent and bromodomain inhibitor JQ1 does not decrease transcriptional activity of AIRE (Huang et al., 2018). In addition to elongation stage proteins, AIRE interacts extensively with splicing proteins (Abramson et al., 2010; Giraud et al., 2014) with reports claiming that AIRE influences target genes' alternative splicing (Keane et al., 2014; St-Pierre et al., 2015). One example of elongation and RNA processing connections in the context of AIRE is HNRNPL protein, an RNA processing factor that belongs to heterogeneous nuclear ribonucleoprotein (hnRNP) family, that interacts with 7SK-containing P-TEFb complexes ensuring AIRE initial contact with inactive P-TEFb (Giraud et al., 2014).

The identification of AIRE interacting proteins was taken to the next level when thorough screening using stable-isotope labelling by amino acids in cell culture (SILAC) method was undertaken by two groups (Abramson et al., 2010; Gaetani et al., 2012), with variety of proteins being identified as partners of AIRE. In general, four functional groups were outlined including nuclear transport, chromatin binding, transcription and pre-mRNA processing, and complex centered on DNA-PK, including DNA-PK, PARP-1, TOP2A, FACT (facilitates chromatin transcription), and Pol II was characterized to be important in AIRE-mediated activation (Abramson et al., 2010). Various DNA repair proteins constituting complex together with DNA-PK suggest that not only phosphorylating activity of DNA-PK plays role in AIRE-mediated transcriptional activation. Interestingly, mutual tight links between DNA repair proteins and chromatin structure modifiers are established with PARP-1 implicated in chromatin modulation, acting at enhancers or insulators and functioning as a transcriptional co-regulator (Kraus, 2008; Ray Chaudhuri and Nussenzweig, 2017) and histone chaperone FACT coordinating H2A.X signaling and DNA repair (Piquet et al., 2018). Other proteins interacting with AIRE such as KAP-1 (Bunch et al., 2014; Iyengar and Farnham, 2011; Ziv et al., 2006) and CHD4 (Larsen et al., 2010; Mansfield et al., 2011) are extensively characterized in chromatin remodeling processes.

Some proteins demonstrated before as AIRE partners, such as CBP and P-TEFb, were not confirmed in Abramson *et al*. study and the set of interacting proteins differed between Abramson *et al*. and Gaetani *et al*. studies, probably reflecting transient nature of protein-protein contacts, methodological differences in cell lysate treatments (micrococcal nuclease (MNase) treatment versus mild sonication), overexpression of proteins, or cell culture conditions. Of note, ubiquitin was detected as persistent interactor in Gaetani *et al*. study, reinforcing the role of AIRE in apoptotic processes (Liiv et al., 2012; Schaller et al., 2008). Additionally, ubiquitinylation is applied in various cellular processes including transcription-related regulation, histone ubiquitinylation and DNA damage response (Muratani and Tansey, 2003; Schwertman et al., 2016; Strzyz, 2016).

Intriguing association of AIRE with mitotic spindle via its last LxxLL motif in embryonic stem cells was described and may provide new evidence about unrecognized AIRE function in dividing cells (Gu et al., 2017).

2.3.5. AIRE is an unusual transcriptional activator

Although it would be very tempting to study AIRE function in native environment, the thymic epithelial cells lose quickly AIRE expression when cultured *in vitro*, reflecting the requirement for three-dimensional structure and specific signals (Kont et al., 2008; Peterson et al., 2008). Therefore, several model systems have been employed to study AIRE functions, which reveal that the set of genes upregulated by AIRE depends heavily on the context and varies in different cell lines and mouse strains (Guerau-de-Arellano et al., 2008; Jiang et al., 2005; Venanzi et al., 2008). Even if the selection of target genes varies in different cell types, the transcriptional activating effect is undeniable in every context. To date, the established common features of AIRE target genes are their initial low expression and chromatin context (Org et al., 2009), in particular the enrichment of repressive chromatin marks H3K27me3 and H3K9me3, and depletion of an active chromatin mark H3K4me3 (Handel et al., 2018). The interaction with nonmethylated H3K4 was detected, although the choice of AIRE target genes could not be explained completely through this interaction (Koh et al., 2010; Org et al., 2008). In contrast, AIRE DNA binding sites were depleted in repressive chromatin marks and enriched for promoter and enhancer associated chromatin modifications (Handel et al., 2018).

Thus, the most straightforward way to explain gene activation assuming that target genes share DNA sequence-specific motifs, enabling protein binding and subsequent transcriptional activation, is not applicable to AIRE regulatory mechanism. Although few reports demonstrated DNA binding via the SAND or PHD domains and even sequence preferences (Kumar et al., 2001; Purohit et al., 2005), the results of other research groups have shown no direct binding of the AIRE SAND domain to DNA (Zumer et al., 2012), suggesting other proteins' involvement. The classical KDWK DNA-binding motif of the SAND domain is changed to KNKA in AIRE, and mutating every amino acid in this sequence did not diminish AIRE transactivation, provoking a new hypothesis that DNA-PK is mediating binding of AIRE to DNA (Zumer et al., 2012). The notion that DNA-PK catalytic inhibitor NU7441 did not decrease AIRE-mediated activation, highlighted the importance of DNA-PK structural role and not catalytic phosphorylation for AIRE-mediated transactivation. Interestingly, early-acting elements of NHEJ were associated with the AIRE protein, but not late-acting ones as X-ray repair cross-complementing protein 4 (XRCC4) (Abramson et al., 2010), assuming an alternative role for DNA-PK in AIRE-mediated transcription not acting through the same NHEJ pathways and involving different elements.

Current knowledge favors AIRE involvement in the elongation stage of transcription rather than initiation. Supporting evidence for this statement is interaction with P-TEFb, ability to induce transcription from the reporters supporting elongation but not initiation (Oven et al., 2007), and first exon transcripts' abundance in Aire knock-out mice (Giraud et al., 2012), indicating the inability of Pol II to transit into elongation phase without Aire partaking near TSS of TSA.

Initial studies of AIRE-mediated activation mechanism were accomplished analyzing cell populations in bulk, producing the summarized result of individual cells. Recent single-cell RNA-seq technology has enabled to explore how individual mTECs contribute to the cell population pattern. The most outstanding features of transcriptional regulation by AIRE is the high number of genes it induces and the stochastic choice of upregulated genes. The mature Aire-positive murine mTEC cell population expresses 87% (19293) of protein-coding genes, probably the highest proportion of genes yet found to be expressed in any cell type (Sansom et al., 2014). Yet, only a small proportion of mTECs (1–3%) expresses a particular TSA (Pinto et al., 2013), which highlights the stochastic mechanism and creates mosaic expression pattern, when individual cell with few antigens expressed contributes to a very diverse pool of antigens in a cell population. The majority of protein-coding genes, including TSAs, are covered in a couple of hundreds of mTECs, limiting the number of cells thymocytes have to sample (Miragaia et al., 2018). Mosaic expression also avoids excessive burden to a cell, as TSAs account for 10-20% of all the genes expressed in a single cell (Brennecke et al., 2015; Miragaia et al., 2018). The genes regulated by Aire do not share common genomic position, biological function or transcriptional regulation (Meredith et al., 2015). The repertoires of TSAs expressed in single mTECs did not exhibit significant enrichment for any particular peripheral tissues refuting the developmental model where each mature mTEC cell commits to certain peripheral lineage program (Gillard and Farr, 2005). Supporting this observation, insulin expression in mTECs persists even if TFs required for pancreatic expression are knocked out (Danso-Abeam et al., 2013; Villaseñor et al., 2008), arguing for completely different regulation mechanisms in the thymus and differentiated tissues.

Clustering of AIRE target genes was remarked early on (Derbinski et al., 2005; Gotter, 2004; Johnnidis et al., 2005), evoking hypotheses of epigenetic marks or spatial organization determining AIRE target gene choice. Single-cell PCR analysis revealed that cluster gene regulation in the thymus and differentiated tissue is cardinally different (Derbinski et al., 2008) with casein gene locus demonstrating simultaneous expression of milk proteins in mammary gland, but with rarely observed co-expression in mTECs. Single-cell RNA-seq data enabled precise deciphering of cluster regulation in individual mTECs. Co-expression patterns were detected with 51% of Aire-induced transcripts grouped into clusters in cell population. Interestingly, most correlations were detected between chromosomes and those clusters were not reproduced between mice (Meredith et al., 2015). In other work, Aire-dependent, but not TSAs dependent on other TFs, clustered, and clustering was more pronounced with Aire-dependent but not with Aire-enhanced genes (Miragaia et al., 2018), pointing that Aire-dependent genes require additional level of regulation provided by cluster formation in order to be induced.

2.4. Summary of the literature

The decoding of human genome provided several surprises, unveiling the large fraction of non-coding sequences and only ~21000 protein-coding genes, suggesting that complexity of regulation cannot be explained by protein-coding genes only and intergenic sequences require heightened attention (Frazer, 2012). Estimatingly, 51,8% of the human genome is transcribed, but only 1,2% encodes proteins (Ransohoff et al., 2017). A large group of non-coding RNAs has been described with steadily growing list of biological roles, not only reflecting transcriptional noise (Geisler and Coller, 2013; Jacquier, 2009). Diverse group of regulatory elements (promoters, enhancers, insulators) appeared explaining the mechanism of disease-associated variants not linking to protein-coding sequences (Kellis et al., 2014; Maston et al., 2006).

The patterns of transcription derived from prokaryotes did not seem to be entirely applicable to eukaryotic systems (Kornberg, 1999). Eukaryotic transcription is pervasive and stochastic in nature, transcription factor binding sequence motif alone does not determine the actual binding and the binding does not guarantee the regulation of gene. DNA packaging to nucleosomes and spatial organization of the nucleus provide additional layers of regulation in eukaryotes. Chromatin remodeling, transcription and RNA processing, including splicing and polyadenylation, are tightly coupled processes that influence each other.

The AIRE protein expressed in medullary thymic epithelial cells is an unusual transcriptional regulator. AIRE is able to upregulate large plethora of genes in different cell types, doing so in random fashion with the set of genes activated depending on the cell type. The genes regulated by AIRE are characterized by low initial expression level, repressive chromatin markers and tissue-specificity. Such abundant stochastic pattern of regulation is termed promiscuous gene expression.

This unconventional model of regulation becomes especially useful in the thymus, which is primary lymphoid organ, where T cells complete their maturation. T cells during differentiation are exposed to as many self-antigens as possible to purge autoreactive cells recognizing them, and in this way to avoid autoimmune reactions in the future. The *AIRE* gene identification transformed the understanding of central tolerance. Though in the thymus the unexpected expression of proteins, commonly found only in particular organs, was noticed earlier, the demonstration of AIRE ability to induce the expression of vast scope of genes, creating the reflection of its own self, shifted the paradigm of T-cell tolerance. The mutations in the *AIRE* gene evoke autoimmune APECED disease, unique in its monogenic inheritance. As it is different from the majority of autoimmune diseases where the combination of environmental factors and slight gene changes contribute to the phenotype, APECED provides valuable model to decipher immune system functioning.

Despite continuously accumulating information and improved understanding of AIRE functional mechanism, the complete comprehension is still not achieved. The questions of current interest still are – what determinants dictate the choice of AIRE target genes and how these chosen genes are activated.

3. AIMS OF THE STUDY

The main objectives of the present study were:

1. To determine the gene elements required for AIRE-mediated transcriptional activation.
2. To study the DNA binding properties of AIRE.
3. To characterize the role of topoisomerases and DNA repair proteins in AIRE-mediated transcription.

4. MATERIALS AND METHODS

4.1. Cell culture, treatments and transfections

The human embryonic kidney (HEK)293 cell line was cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum and 1% antibiotic-antimycotic (10 000 U penicillin, 10 mg streptomycin, 25 μg amphotericin B) (all components from PAA Laboratories Inc) at 37°C in humidified atmosphere of 5% CO_2 in air. All transfections were performed with TurboFect™ Transfection Reagent (Thermo Scientific) according to the manufacturer's protocol. The generation of stable HEK293 cells expressing wt AIRE and YFP was described previously (Org et al., 2009).

Doxycycline-inducible AIRE expression in the HEK293 cell line (AIRE-Tet) has been described previously (Liiv et al., 2012). Cells were cultured in DMEM supplemented with 10% tetracycline-negative fetal calf serum, 100 units/ml penicillin/streptomycin, and 0.15 mg/ml G418 for at least 24 h before any additional treatments. Untreated cells served as a negative control material, and AIRE expression was induced with 1.5 μM doxycycline (631311, Clontech). After 24 h, the control and induced cells were either mock-treated with DMSO or 2 μM etoposide (E1383, Sigma, reconstituted in DMSO) for another 24 h. In some experiments, the cells were additionally incubated with 1 μM merbarone (M2070, Sigma), 1 μM camptothecin (C9911, Sigma), or 1 μM β-lapachone (L2307, Sigma) for 24 h.

4.2. Plasmids

The following full-length AIRE expression constructs were used: pcAIRE (Heino et al., 1999), pcAIRE-D312A (Org et al., 2008), pcAIRE-K243/245Q (Saare et al., 2012), and pcAIRE-L28P (Ferguson et al., 2008). pcAIRE-R15C, A21V, L29P, V80L, and L93R were generated by PCR-based mutagenesis, K83E was amplified from pSI-AIRE-K83E (Pitkänen et al., 2001), the generated fragments were cloned into the EcoRI/HindIII sites of pcDNA3.1B(-)Myc/ His. To generate pcHSR/CARD (aa 1–109) and pcHSR/CARD+ (aa 1–143), the corresponding regions were PCR amplified from pcAIRE and cloned into the NotI/EcoRI sites of pcDNA3.1B(-)Myc/His. pcHSR/CARD+R113A/K114A was generated by PCR-based site-directed mutagenesis. pcDNA3.1B(-)Myc/His (Invitrogen), pGL3-Basic (Promega), pBL-KS (a gift from prof. K. Saksela, University of Tampere, Finland) and pd2EYFP-N1 (Clontech) were used as the controls. The lack of promoter sequences in pGL3-Basic and pBL-KS plasmids was confirmed by GenBank and Eukaryotic Promoter Database searches. The luciferase (Luc) reporter pBL-INV (pBL-IVL in this study) containing promoter area of involucrin (3737 bp) was generated by cutting out the promoter area fragment from pTZhINV-nlbgal plasmid (a gift from A. Männik, FitBiotech, Estonia) and cloned into the HindIII site of pBL-KS. IVL-pr (bp -259/+1), IVL-pr-in

(-259/+1235 bp), IVL-in (-27/+1235 bp), INS-pr (-326/+30 bp), INS-pr-in (-326/+238 bp) and INS-in (-5/+238 bp) were PCR amplified from pBL-IVL or human genomic DNA and cloned into HindIII/BamHI sites of pBL-KS vector. The AdML (adenovirus major late) intron was amplified from the sequence of the AdML splicing substrate (Mishler et al., 2008) and cloned into the BamHI site of IVL-pr or INS-pr. H1-SL, H1-mutSL and MS2-SL containing histone modified stem-loops in a T7-based vector with the Luc reporter gene were gifts from B. Marzluff and were described previously (Gallie et al., 1996; Sànchez et al., 2002) as *luc*-SL_{wt}, *luc*-tetraloop and Luc-MS2, respectively. To generate GST-HSR/CARD (aa 1–109) and GST-HSR/CARD-L28P, the corresponding regions were PCR amplified from pcAIRE and pcAIRE-L28P, respectively, and cloned into the EcoRI/XhoI sites of pGEX-1λT vector. GST-HSR/CARD+ (aa 1–138) (previously named as GST-AIRE (1–138)) was previously described (Pitkänen et al., 2000). Plasmid pGST-PHD1 was generated by insertion of PCR-amplified sequence encoding human AIRE aa 293–354 into the NcoI and KpnI restriction sites of the pETM-30 vector (EMBL-Heidelberg protein expression and purification laboratory). pcDNA-p50 and pcDNA-FLAG-p65 were a gift from prof. E. Kalkhoven from Utrecht University. The 3xNF-kBLuc reporter plasmid was a gift from prof. K. Saksela. All generated constructs were verified by sequencing. The primers used for plasmid generation are listed in Table 1.

Table 1. List of oligonucleotides used in the experiments presented in this work.

Oligo name	**Sequence**
Plasmid cloning and mutagenesis	
hAIRE EcoRI F	TTTTGAATTCACCATGGCGACGGACGCGGCGCT
hAIRE HindIII STOP R	CATTAAGCTTAGGAGGGGAAGGGGGC
R15C F	GAGGCTGCACTGCACGGAGAT
R15C R	ATCTCCGTGCAGTGCAGCCTC
A21V F	GATCGCGGTGGTCGTGGACAGC
A21V R	GCTGTCCACGACCACCGCGATC
L29P F	TTCCCACTGCCGCACGCGCTG
L29P R	CAGGGCGTGCGGGCAGTGGGAA
V80L F	CTTCTGGAGGTTGCTGTTCAA
V80L R	TTGAACAGCAACCTCCAGAAG
L93R F	GGCGGCAGCCCATCCTGGA
L93R R	TCCAGGATGGGCTGCCGCC
pcHSR/CARD F	TTTTGCGGCCGCATGGCGACGGACGCGGCG
pcHSR/CARD R	TTTGAATTCGGGCTGGCTGAGGTCCACA
pcHSR/CARD+ F	TTTTGCGGCCGCATGGCGACGGACGCGGCG
pcHSR/CARD+ R	TTTGAATTCTGGCGCGGCAGCTCGAGC
R113A/K114A F	AAGGGGGCGGCGCCCCCGGCC
R113A/K114A R	GGCCGGGGGCGCCGCCCCCTT
GST-HSR/CARD F	TTTTGAATTCATGGCGACGGACGCGGCG
GST-HSR/CARD R	TTTTGTCGACTCAGGGCTGGCTGAGGTCCAC

Oligo name	**Sequence**
IVL-pr F	ATAAAGCTTTGCCTAGGTCAGAAAAGCAT
IVL-pr R	ATAGGATCCTTTGGTGGAGTGCTGAGC
IVL-pr-in F	ATAAAGCTTTGCCTAGGTCAGAAAAGCAT
IVL-pr-in R	AATTGGATCCTTAGAAGCTACTGTCAACCTGAA
IVL-in F	AATTAAGCTTCTGCTCAGCTCAGCACTCCACCAA
IVL-in R	AATTGGATCCTTAGAAGCTACTGTCAACCTGAA
IVL-pr-ADin F	AATTGGATCCACTCTCTTCCGCATCGCTGTCTGC
IVL-pr-ADin R	AATTGGATCCGGCTCTTACCGTTCGGAGGCCGAC
INS-pr F	GCGAAGCTTTCTCCTGGTCTAATGTGGAA
INS-pr R	CGCGGATCCCTCTTCTGATGCAGCCTGTC
INS-pr-in F	GCGAAGCTTTCTCCTGGTCTAATGTGGAA
INS-pr-in R	AATTGGATCCGGCAGAAGGACAGTGATCTGGGA
INS-in F	GCGAAGCTTCCCTCAGCCCTCCAGGACAGGCTGCA
INS-in R	AATTGGATCC GGCAGAAGGACAGTGATCTGGGA
INS-pr-ADin F	AATTGGATCCACTCTCTTCCGCATCGCTGTCTGC
INS-pr-ADin R	AATTGGATCCGGCTCTTACCGTTCGGAGGCCGAC
ChIP analysis	
IVL pr F	TGCTTAAGATGCCTGTGGTG
IVL pr R	TGAAGGTGATGGACAGGTTTC
IVL 3' F	TAGGTGGCCCGTCTCATCTGTGAA
IVL 3' R	ATTGGGGTCATTGGGGTTGGCACT
S100A8 pr F	GGCCTGACCACCAATGCAGGG
S100A8 pr R	CTGGGCTGCTGGCATCCACT
S100A8 3' F	AATCTTGACAGAAAAAGGGTGCAGACG
S100A8 3' R	GGGCTGCCACGCCCATCTTT
B2M pr F	CCTTGTCCTGATTGGCTGGG
B2M pr R	CCTTGTCCTGATTGGCTGGG
B2M 3' F	TCTCTTGCACTCAAAGCTTGTT
B2M 3' R	TCCCCCAAATTCTAAGCAGAGT
HPRT1 pr F	CCTGCTTCTCCTCAGCTTCA
HPRT1 pr R	GGGAAAGCCGAGAGGTTCG
HPRT1 3' F	GGGAAAGCCGAGAGGTTCG
HPRT1 3' R	GGGAACTGCTGACAAAGATTCAC
Luc F	AGGTCTTCCCGACGATGACG
Luc R	GGCGACGTAATCCACGATCTC
Expression analysis	
S100A8 F	CTCAGTATATCAGGAAAAAGGGTGCAGAC
S100A8 R	CACGCCCATCTTTATCACCAGAATGAG
IVL F	GCCTTACTGTGAGTCTGGTTGACA
IVL R	GGAGGAACAGTCTTGAGGAGCT
HBG2 F	CATAAAGCACCTGGATGATCTC
HBG2 R	CAGGAGCTTGAAGTTCTCAG
KRT73 F	GAGTGCAGGATGTCCGGAGAATA
KRT73 R	TTGCTGAATCCAAAGCCAGCC
DMBT1 F	AGAACCCAGCAAAATGGGGA
DMBT1 R	TTGGGATCCACCCACCTGTA

Oligo name	Sequence
PSMD4 F	GAAGGTGAAAGAGACTCA
PSMD4 R	GTCATACTGCTTAGGTCA
CEACAM5 F	CCTGGATGTCCTCTATGGGC
CEACAM5 R	TACTGCGGGGATGGGTTAGA
HPRT1 F	GACTTTGCTTTCCTTGGTCAGG
HPRT1 R	AGTCTGGCTTATATCCAACACTTCG

4.3. Luciferase reporter assay and quantitative PCR

Cells were seeded at 20–30% confluency in 24-well plates 24 h before the transfection. A total of 1 μg of plasmid (with a 1/10 expression/reporter plasmid ratio) per well was used. Cells were harvested 48 h later, and Luc activity was measured using a Luciferase Assay System (Promega) according to the manufacturer's instructions. The transfections were performed in duplicate, and experiments were repeated at least two times. Luminescence was counted with a Wallac 1420 Victor Multilabel/Plate Reader (Perkin Elmer).

RNA was purified with TRIzol-chloroform extraction, treated with TURBO DNaseI (Thermo Fisher Scientific) and reverse transcribed with SuperScript III (Life Technologies). The quantitative PCR was performed in triplicate using Maxima SYBR Green/ROX qPCR Master Mix (Thermo Scientific) and a ViiA 7 Real-Time PCR System (Life Technologies). Relative gene expression levels were calculated using the comparative Ct method (according to Applied Biosystems), where the relative expression is calculated as $2^{-\Delta\Delta Ct}$ and Ct represents threshold cycle value. *HPRT1* (hypoxanthine phosphoribosyltransferase 1) was used as a house-keeping gene for normalization. The primers used for expression analysis are listed in Table 1.

4.4. Chromatin immunoprecipitation

A total of $5x10^6$ transfected HEK293 cells were used in a single chromatin immunoprecipitation (ChIP) reaction with the following antibodies: mouse monoclonal anti-c-Myc (ab32, Abcam), rabbit polyclonal anti-Histone H3 (ab1791, Abcam) and rabbit IgG (Chemicon International), as a negative control. ChIP was performed as described in (Org et al., 2008). Briefly, transfected HEK293 cells were cross-linked with 1% formaldehyde for 10 min at room temperature, and formaldehyde was quenched by addition of glycine to a final concentration of 0.125 M. The cells were lyzed in 20mM Tris-HCl, 0,5% NP-40, 1% SDS, 10mM EDTA and 1x Halt Protease Inhibitor Cocktail containing buffer and sonicated with the Diagenode Bioruptor for 15 min using 30 sec on-off cycles at a high voltage setting. After 10x dilution with 150mM NaCl, 20mM Tris-HCl, 1% Triton X-100, 2mM EDTA and 1x Halt Protease Inhibitor Cocktail buffer, the sonicated chromatin solution was pre-cleared with

protein G agarose beads, that were pre-absorbed with 100 µg/ml bovine serum albumin and 500 µg/ml sheared salmon sperm DNA. The pre-cleared solution was incubated with the corresponding immunoprecipitating antibody at 4°C overnight, after which blocked protein G agarose beads were added. After thorough washing with low salt wash buffer (150 mM NaCl, 20 mM Tris-HCl, 0,1% SDS, 1% Triton X-100, 2 mM EDTA), high salt wash buffer (500 mM NaCl, 20 mM Tris-HCl, 0,1% SDS, 1% Triton X-100, 2 mM EDTA) and LiCl wash buffer (0,25 M LiCl, 10 mM Tris-HCl, 1% deoxycolate, 1% NP-40, 1 mM EDTA), the complexes were eluted from beads by shaking for 15 min in 1% SDS, 0,1 M $NaHCO_3$ buffer at room temperature. After cross-link reversal and proteinase K treatment the DNA was phenol:chloroform extracted and analyzed by quantitative PCR as described in section 4.3. Datasets of each primer pairs were normalized to ChIP input values and then the Ct values obtained from AIRE-control cells with nonspecific antibody were subtracted from Ct values obtained with specific antibody. The ChIP primers' sequences are listed in Table 1.

4.5. The expression and purification of GST fusion proteins, GST-pulldown and mass-spectrometry

The GST (glutathione S-transferase)-tagged proteins were expressed in *Escherichia coli* BL21-DE3 strain by induction with 0.4 mM IPTG for 4 h. The PHD1-containing fusion protein was expressed in the presence of 0.1 mM $ZnCl_2$. The proteins were purified using Glutathione Sepharose 4B (Amersham Biosciences) according to the manufacturer's instructions except that 1% N-laurylsarcosine and 3% Triton X-100 were added to increase the solubility of the proteins and 50 µM $ZnCl_2$ was included when PHD1-containing fusion protein was purified. The purified proteins were verified by resolving on SDS-PAGE followed by staining with Coomassie Blue.

50–100 µg of nuclear extract, prepared from human monocyte cell line THP-1, was incubated with 20–25 µg of GST-fusion proteins bound to 25 µl of packed sepharose beads in buffer B1 (10 mM HEPES pH 8.0, 150 mM NaCl, 0.7 mM $MgCl_2$, 12.5% glycerol, 0.1 mM EDTA, 25 µM $ZnCl_2$, 0.5 mM DTT and proteinase inhibitor mix) overnight at 4 °C. The beads were washed intensively with buffer B1, the bound proteins were eluted and separated on SDS-PAGE. The specific protein bands were cut out and analyzed with mass spectrometry. Mass mapping of the peptides generated was performed with an Ultraflex™MALDI-TOF/TOF mass spectrometer (Bruker-Daltonics, Bremen, Germany) equipped with a nitrogen laser in a positive ion reflector mode using α-cyano-4-hydroxycinnamic acid as the matrix. The MALDI spectra were externally calibrated with the standard peptide mixture from Bruker-Daltonics (Bremen, Germany). In case of peptide fragmentation analysis, a peptide from above mass mapping analysis was selected as a precursor ion and subjected for further MS/MS fragmentation in the MALDI-TOF/TOF lift-mode. Database searches

were carried out by either Mascot peptide map fingerprint or Mascot MS/MS ions search (http://www.matrixscience.com/).

4.6. Thymic stromal cell isolation from mice

C57BL/6J background wild type mice were maintained at the mouse facility of the Institute of Molecular and Cell Biology, Tartu University. Small cuts were made into the capsules of thymi dissected from eight to twelve 4-week-old mice and thymocytes were released by repetitive pipeting. The remaining thymic fragments were incubated in 0.5 mg/ml dispase/collagenase (Roche) and 1.5 μg/ml DNase I (AppliChem) in PBS at 37°C for 20 min, with gentle agitation using a Pasteur pipette every 5 min. Released cells were collected to separate fractions and fresh enzyme solution was added four times. The cells were resuspended in 5mM EDTA in 10 ml of RPMI-1460. For CD45 depletion CD45 MicroBeads (Miltenyi Biotec) were used according to manufacturers instructions. For cortical epithelial cell (cTEC) isolation, the CD45− cells were stained with H213-HB Ab (anti-CDR1) followed by goat anti-rat IgG Microbeads (Miltenyi Biotech) and AutoMACS separation (isolation mode: Possel-S). The positive fraction (CDR1+) contained magnetically bound cTECs. For mTEC isolation CDR1− cells were stained with G8.8 (anti-EpCAM, generated from a G8.8 hybridoma cell line) followed by goat anti-rat IgG Microbeads (Miltenyi Biotec) and separation as previously described. The purity of mTEC and cTEC was >80% as assessed by staining with anti-CD45 FITC (30F11, Miltenyi Biotech) and anti-I-A^b PE (AF6-120.1, BD Biosciences) using FACSCalibur flow cytometer (BD Biosciences).

4.7. Electrophoretic mobility shift assays (EMSA)

Electrophoretic mobility shift assay (EMSA) was performed according to a previously described protocol (Murumägi et al., 2003). The proteins were probed with a γ-^{32}P-dATP-labeled duplex oligonucleotide containing the FLAG tag coding sequence: 5'-AATTGAATTCGATTACAAGGACGACGATGACAAG TAG CT TAAGTTAA-3' without any known recognition elements for transcription factor binding. EMSA reactions were performed in 10 μL of 1× binding buffer (5 mM $MgCl_2$, 2.5 mM EDTA, 2.5 mM DTT, 125 mM NaCl, 50 mM Tris–HCl [pH 7.5], 10% glycerol) containing 5 μg of recombinant protein, 1 μg poly-dIdC and 20 pmol of γ-^{32}P-dATP-labeled duplex oligonucleotide. EMSA in agarose gel was performed as previously described (Stros and Reich, 1998) using pcDNA3, circular or linearized with HindIII, as a probe. The reaction was performed with 100 ng of plasmid and 1 μg of recombinant protein in 0.14 mM NaCl, 20 mM Tris/HCl, pH 7.5, 0.2 mM EDTA, 5 mM dithiothritol solution. The samples were subjected to electrophoresis for 15–18 h at 3 V/cm in 1% agarose gel at 4°C and stained with 0.5 μg/ml ethidium bromide.

4.8. Immunofluorescence

The cells were fixed with 3% formaldehyde and permeabilized with 0,5% Triton X-100 in the presence of 1% normal goat serum (NGS) as a blocking reagent. After that, the cells were incubated with primary, mouse monoclonal anti-Myc (ab32, Abcam), and afterwards secondary, goat anti-mouse IgG conjugated with Alexa-488 (A11029, Invitrogen), antibodies for 1 h at room temperature. The washes were performed with 1% NGS in PBS between and after antibody incubation. Nuclei were stained with DAPI (4′,6-diamidino-2-phenylindole) (1 μg/ml) and fluorescent mounting medium (Dako) was used for mounting. The images were captured with the LSM5 DUO confocal microscope (Zeiss) and processed with LSM Image Browser software (Zeiss).

4.9. RNA-seq sample preparation and data analysis

RNA was isolated with the miRNAeasy Mini Kit (Qiagen) using on-column DNase digestion from 10^6 AIRE-Tet cells induced with 1.5 μM doxycycline for 48 h, 2 μM etoposide for 24 h, or a combination of both. Uninduced DMSO-treated cells were used as a negative control (Ctrl). Two independent biological replicates from each treatment were prepared from 500 ng of RNA with RNA integrity number (RIN)>9. Sequencing libraries were generated with the TruSeq Stranded Total RNA LT (with Ribo Zero Gold) Sample Prep Kit (Illumina) according to the instructions of the manufacturer, with the exception that 12 PCR cycles were used for the DNA enrichment step. Paired-end sequencing (2 × 100 bp) was performed with Illumina HiSeq 2000, which generated 80–100 × 10^6 sequenced fragments per sample. Prior to alignment, the reads were trimmed to remove adapter sequences and bases with a Phred score lower than 30 using Trim Galore!. Reads were mapped to the human reference genome GRCh37 (Ensembl release 75) using STAR aligner version 2.4.2a with the two-pass mode (Dobin et al., 2013). Read count tables were generated with the STAR aligner option -quantMode geneCounts. Differential gene expression between the experimental samples was analyzed with DESeq2 (Love et al., 2015). Read count tables and group comparisons for differential exon usage were performed with DEXSeq (Anders et al., 2012). Determination of tissue specificity was based on data from the Human Protein Atlas (Fagerberg et al., 2014). Chromosomal clusters of genes were detected with the command line tool CROC (Pignatelli et al., 2009), where a sliding window approach combined with hypergeometric distribution is used to evaluate the statistical significance of the predicted clusters. In our analysis we applied the gene-based type of sliding window with size of seven genes and minimum number of genes that define a cluster set to three genes.

4.10. Immunoprecipitations and Western blot analysis

For co-immunoprecipitations either whole-cell extract (Study II) or nuclear extract with addition of MNase (Study III) was prepared from transfected (Study II) or inducible HEK293 cells (Study III). In study II, immunoprecipitations were performed with mouse monoclonal anti-DNA-PK 4F10C5 (BD Pharmingen), anti-AIRE 6.1 (Heino et al., 1999), anti-HuR 3A2 (a gift from J. Steitz, Yale University, USA), anti-GAPDH 6C5 (Ambion), rabbit polyclonal anti-GST sc-459 (SantaCruzBiotech). Rabbit anti-TOP2A (Abcam, ab2987) or IgG control (Diagenode, AIP-103-110) antibodies were used in study III. Protein G-sepharose beads were added to lysate-antibody mixtures for capturing antibody complexes. After washing the beads, the immunoprecipitates were eluted from the beads with Laemmli buffer at 99 °C for 10 min, separated on SDS-PAGE, and Western blot was performed with anti-AIRE 6.1 (Study II) or mouse polyclonal anti-AIRE (1:10,000, sc-17985, Santa Cruz Biotechnology), rabbit polyclonal anti-TOP2a (1:1000, ab2987, Abcam) or mouse monoclonal anti-β-actin (1:10,000, A5441, Sigma) (Study III). The more detailed protocols could be found in the methods' section of the corresponding studies.

For treatments, RNaseA/RNaseT mix (4 μg and 10 U per ml, respectively) (Fermentas), ethidium bromide (0.4 mg/ml) or 1 U of MNase (Fermentas) in 50 μl of digestion buffer (10 mM HEPES pH 7.0, 4 mM $CaCl_2$, 50 mM NaCl, 0.1 mM $ZnCl_2$) was used.

4.11. TOP2A knockdown

Uninduced and doxycycline-induced AIRE-Tet cells, either untreated or treated with etoposide, were grown on 6-well plates and transfected with four SureSilencing shRNA vectors along with one scrambled negative control shRNA vector (KH01520N, SABiosciences) using Turbofect reagent (Thermo Scientific). The efficiency of knockdown was checked by Western blotting with anti-TOPA antibody 48 h after transfection. Anti-β-actin antibody (A5441, Sigma) was used to determine protein loading.

4.12. TUNEL assay

Uninduced and doxycycline-induced AIRE-Tet cells, either untreated or treated with etoposide, camptothecin, merbarone, and β-lapachone were fixed with 4% formaldehyde. The occurrence of DNA breaks was quantitatively assessed with the APO-BrdU TUNEL (terminal deoxynucleotidyl transferase dUTP nick end labeling) assay (Pharmingen). Briefly, the cells were labeled with Br-dUTP using TdT enzyme for 60 min at 37 °C and then incubated with FITC-labeled anti-BrdU antibody for 30 min at room temperature. After washing with PBS, the cells were analyzed by LSR Fortessa using FACSDiva software (both from BD Bio-

sciences). The efficacy of the assay was determined with the positive and negative control samples included in the kit.

4.1.3. Statisctics

In Study I and II luciferase, ChIP and qPCR experiments were performed as two to four biological replicates with two to three technical replicates. Data were anlayzed with GraphPad software and presented as mean ± SEM (standard error of the mean).

In Study III qPCR values were log10 transformed and mean ± SD (standard deviation) from three to six experiments was presented. Statistical significance was determined by applying an unpaired t-test. The results were considered significant at p-value less than 0.05. RNA-seq data were analyzed DESeq2 and DEXSeq software packages.

5. RESULTS

5.1. AIRE activates reporter plasmids without promoter elements

AIRE has been shown to have roles in various steps of gene regulation, including chromatin recognition (Koh et al., 2008; Org et al., 2008), transcriptional elongation (Giraud et al., 2012; Oven et al., 2007), pre-mRNA processing and splicing (Abramson et al., 2010; Keane et al., 2014), and microRNA-mediated regulation (Macedo et al., 2013).

To identify the gene elements that are important for AIRE-mediated gene activation, we assessed AIRE influence on transcriptional activation while eliminating defined regulatory features in reporter plasmid. We chose insulin (*INS*) and involucrin (*IVL*) as marker genes because they have been repeatedly demonstrated as AIRE targets, by comparing *Aire* wt and knock-out mice (Anderson, 2002) and by transfecting HEK293 cells with reporter plasmids (Org et al., 2008). We cloned gene promoter regions of human *INS* and *IVL* in front of Luc sequence in pBL-KS plasmid (IVL-pr and INS-pr, Figure 8A). As negative controls, we used pBL-KS and pGL3, both without any promoter sequence inserted. While the effect of AIRE on activation of reporter constructs containing promoter sequences was predictable, the AIRE ability to activate plasmids without promoter elements was an unexpected result (Figure 8B). Although the level of Luc activity in case of promoterless plasmids was weaker in both AIRE positive and negative conditions, the fold change was comparable between plasmids with and without promoters (Figure 8C). To exclude the possibility of this finding being artificial due to overexpression of transcriptional activators or reporter plasmid leakiness, we tested whether NF-κB activating subunits can induce the expression of pBL-KS. In contrast to AIRE, p65/p50 dimer did not have activating effect on pBL-KS plasmid; however, it successfully increased the expression of reporter gene having NF-κB sites upstream in the plasmid (3xNF-κB) (Figure 8D, 8E).

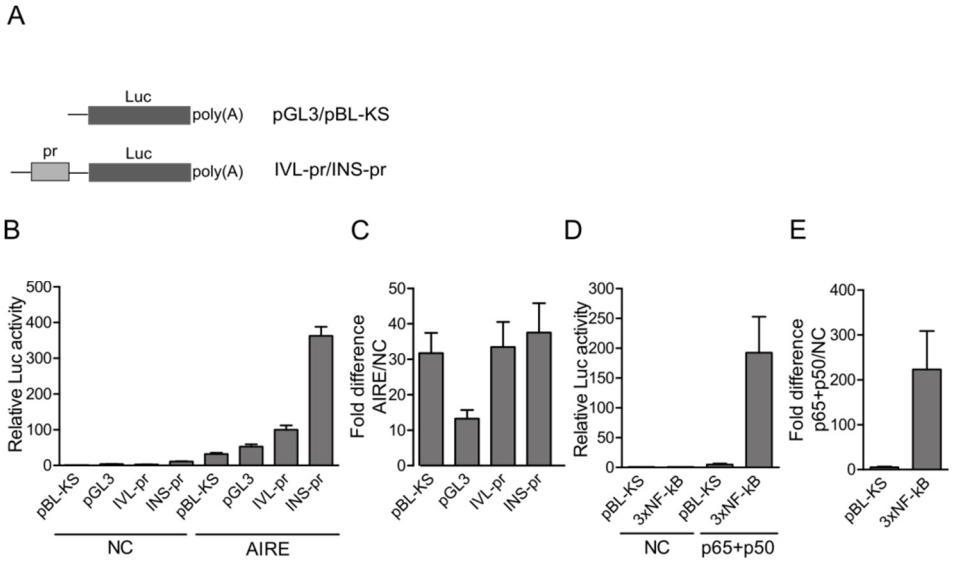

Figure 8. The presence of a promoter does not impact AIRE-mediated gene activation capacity. (A) Schematic representation of the Luc reporter plasmids. pr – promoter, Luc – luciferase coding sequence. **(B)** pcAIRE (AIRE) or pcDNA3 (NC – negative control) were transfected into HEK293 cells together with the indicated Luc reporter plasmids. Luc activity is shown relative to the value of pBL-KS in the NC transfection (= 1). **(C)** Same data as in (B), given as the fold difference in AIRE-expressing and NC samples. **(D)** p65/p50 or YFP (NC) expression plasmids were transfected together with pBL-KS or a Luc reporter containing three upstream NF-kB binding sites. Luc activity is shown relative to the value of pBL-KS in the NC transfection (= 1). **(E)** Same data as in (D) given as the fold difference in p65/p50 expressing and NC samples. **(B-E)** The data are shown as the mean ± SEM (standard error of the mean) of at least two experiments, each performed in duplicate.

5.2. The presence of introns or polyadenylation signals does not influence AIRE-mediated activation

Transcription and RNA splicing/processing are tightly interconnected processes influencing each other, with initial processing steps taking place as soon as a newly synthesized RNA molecule emerges and RNA polymerase II CTD tail serves as a docking site for processing enzymes (Kornblihtt et al., 2004; Moore and Proudfoot, 2009). As AIRE interaction with splicing factors suggested its involvement in pre-mRNA processing events (Abramson et al., 2010), we aimed to test whether the intron regions have impact on AIRE-mediated transcription.

To elucidate whether intronic sequences would enhance AIRE-mediated positive effect on gene expression, we designed the plasmids containing the first intron together with flanking exonic regions of *IVL* and *INS* genes (IVL-pr-in and INS-pr-in). To exclude the putative effect of gene-specific regulatory sequences in intronic regions, we additionally generated plasmid, containing AdML as a

model intron in combination with *IVL* and *INS* promoters (IVL-pr-ADin and INS-pr-ADin) (Figure 9A). While endogenous intron of the *IVL* gene increased Luc expression higher compared to the IVL-pr, the *INS* intron had the opposite effect compared to INS-pr (Figure 9B), suggesting that *INS* first intron contains inhibitory sequence. In addition, INS-pr initially showed higher expression than IVL-pr (Figure 8B, 9B). The adenoviral model intron demonstrated the highest activation as compared to *INS* and *IVL* promoter containing constructs (Figure 9B). Despite differences in level of initial Luc activity, fold changes in response to AIRE were comparable between plasmids with and without intronic sequences (Figure 9C). Supporting the previous observation that promoter is not obligatory for the transcriptional activation, the constructs containing only intronic sequences without promoter showed similar activation fold change with promoter-containing reporters (Figure 9C). The RNA produced from intron-containing constructs was effectively spliced in all cases (Study I Figure 1I).

As the usage of intronic sequence did not alter the ability of AIRE to activate the expression from plasmid, we next aimed to study whether the activation by AIRE depends on polyadenylation. We tested if AIRE is able to activate plasmids containing histone 1 (H1) stem-loop (SL) downstream of the Luc stop codon instead of conventional AAUAAA polyadenylation signal. We used three reporters with firefly Luc coding region under the control of the bacteriophage T7 promoter that initially had been used for *in vitro* transcription; the first one, H1 SL was 32 bp consensus SL generated from phylogenetic comparison and positioned 27 bp downstream of the Luc stop codon, which is similar to the spacing present in histone mRNAs. The effect of H1 SL on the expression is comparable with a poly(A) tail (Gallie et al., 1996). The second, H1-mutSL was a tetraloop that makes use of an unusual base pair for increased thermodynamic stability. It introduces the stable structure at the 3'-end of mRNA; however, does not fulfill the role of functional SL having no positive impact on the expression (Gallie et al., 1996). The third, MS2-SL construct contained the binding site for bacteriophage MS2 coat protein. Despite the lack of eukaryotic promoter and tetraloop structures nonfunctional in 3'-end processing, we observed activation with all stem-loop variants tested, as none of these stem-loop structures hindered transcription or translation in AIRE-positive HEK293 cells (Figure 9D, 9E).

Thus, the gene elements that we tested did not overall affect AIRE-mediated activation of transfected reporters in HEK293 cells, indicating that AIRE-mediated gene expression activation is not stringently restricted to any sequence-specific or regulatory elements.

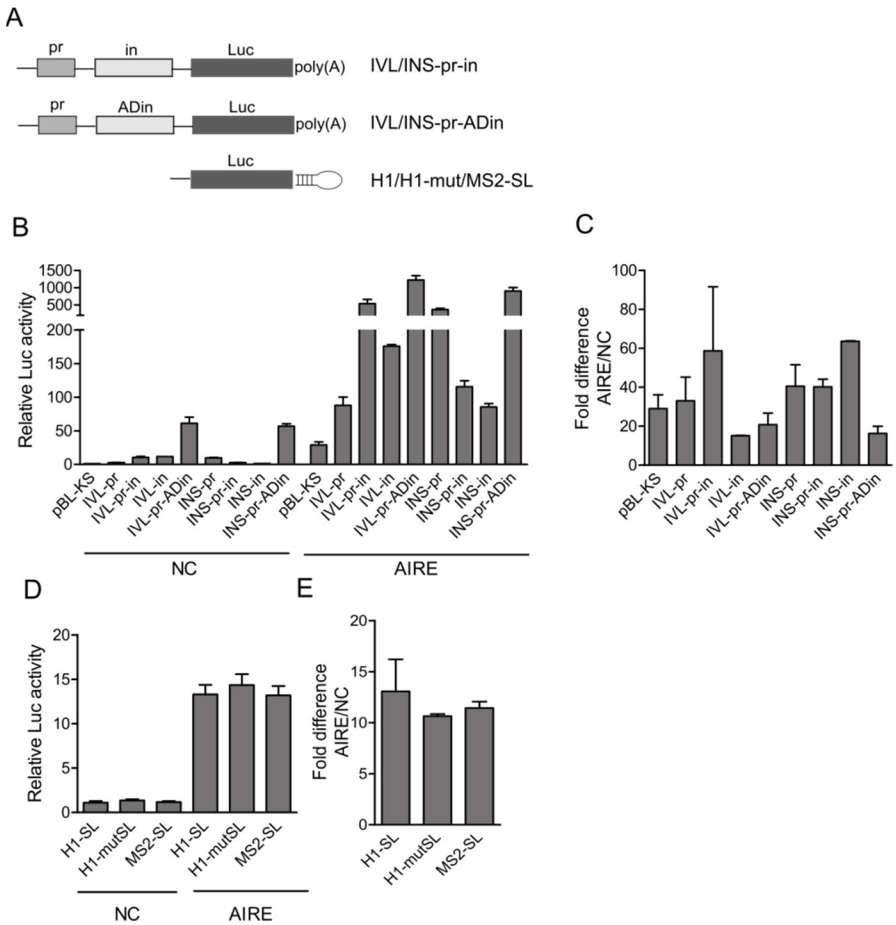

Figure 9. The presence of RNA processing elements does not influence AIRE-mediated gene activation. (A) Schematic representation of the Luc reporter plasmids. pr – promoter, Luc – luciferase coding sequence, in – intron, ADin – adenoviral major late intron, SL – stem-loop. **(B)** pcAIRE (AIRE) or pcDNA3 (NC – negative control) were transfected into HEK293 cells together with the indicated Luc reporter plasmids. Luc activity is shown relative to the value of pBL-KS in the NC transfection (=1). **(C)** Same data as in (B), given as the fold difference between reporter activity in AIRE expressing and NC samples. **(D)** pcAIRE (AIRE) or pcDNA3 (NC – negative control) were transfected into HEK293 cells together with the indicated Luc reporter plasmids. Luc activity is shown relative to the value of H1-SL in the NC transfection (=1). **(E)** Same data as in (D), shown as the fold difference between reporter activity in AIRE expressing and NC samples. **(B-D)** The data are shown as the mean ± SEM of at least two experiments each performed in duplicate.

5.3. Transfected AIRE strongly binds plasmid DNA and mutations in HSR/CARD domain disrupt DNA binding

Intrigued by the ability of AIRE to activate plasmid reporters regardless of the lack of promoter and RNA processing elements, we applied chromatin immunoprecipitation (ChIP) technique to assess the AIRE protein binding to plasmid DNA. For the expression of AIRE, we used a plasmid encoding Myc-tagged AIRE enabling to introduce mutations to AIRE to assess which domains are responsible for binding to plasmid DNA.

AIRE PHD1 interaction with histone H3 nonmethylated at K4 has been found to be at least one of determinants directing AIRE to genomic DNA in chromatin context (Koh et al., 2010; Org et al., 2008). In case of plasmid DNA, which is devoid of nucleosomes and has very low histone content, we surprisingly could observe a very strong AIRE binding (Figure 10A). Accordingly, despite high level of H3 on genomic DNA, the binding of AIRE was much lower to endogenous target gene promoter *S100A8* as compared to plasmid DNA (Figure 10B).

To specify which AIRE protein domains play the prevailing role in a strong binding to plasmid DNA, we tested three mutations positioned in different domains in context of full-length Myc-tagged AIRE: L28P in HSR/CARD domain, K243/245Q in SAND and D312A in PHD1. Each of them had been previously shown to decrease the activation capacity of AIRE on plasmid reporters (Org et al., 2008; Pitkänen et al., 2001; Saare et al., 2012). In our experiments with plasmid reporter, only the L28P mutation showed drastic negative effect on AIRE binding, while others had similar protein enrichment to wt AIRE (Figure 10C). Next, we focused on the HSR/CARD domain and tested other known point mutations, observing a strong correlation between plasmid DNA binding and transcriptional activity. AIRE, containing the mutations shown to maintain transcriptional activity (A21V, V80L, K83E), though causing APECED symptoms (Halonen et al., 2004), demonstrated the similar binding strength as wt AIRE. At the same time, AIRE containing HSR/CARD point mutations that disrupted transcriptional activation (R15C, L28P, L29P, L93R) (Halonen et al., 2004) showed very low enrichment in ChIP experiments (Figure 10D).

These results demonstrate that the integrity of the HSR/CARD domain is essential for AIRE interaction with transfected plasmid DNA and that mutations reducing DNA binding cause transcriptional inactivation.

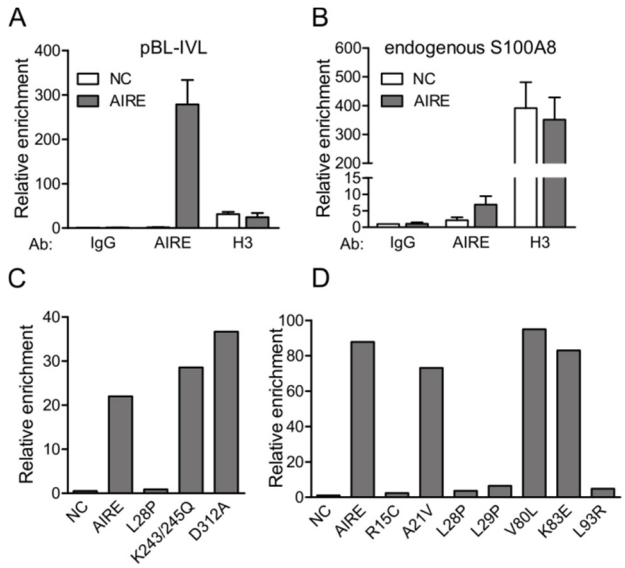

Figure 10. The HSR/CARD domain is essential for AIRE binding to transfected plasmid. (A, B) HEK293 cells were transfected with pcDNA3 (NC – negative control) or pcAIRE (AIRE) followed by ChIP with anti-Myc (AIRE), anti-H3 or anti-IgG antibodies. The enrichment of wt AIRE and H3 on the transfected pBL-IVL reporter plasmid (A) and the endogenous target gene S100A8 (B) was calculated relative to the background signal of anti-IgG (=1). The data is shown as mean ± SEM of three experiments. **(C, D)** AIRE ChIP with anti-Myc antibody was performed with HEK293 cells transfected with AIRE expression plasmids containing indicated mutations and pBL-IVL plasmid as target. IVL promoter primers were used for qPCR. The data are shown as relative enrichment of AIRE-to-IgG from a single representative experiment.

5.4. AIRE N-terminal HSR/CARD region and amino acids R113 and K114 are responsible for DNA binding

To investigate whether the HSR/CARD domain is sufficient for DNA binding, we generated aa 1-143 HSR/CARD+ fragment comprising the HSR/CARD region aa 1-100 together with 43 aa downstream covering the NLS signals to warrant nuclear transport, and performed ChIP with generated construct. The HSR/CARD+ fragment appeared to bind plasmid DNA even more strongly than full-length AIRE (Figure 11A), that might be due to smaller size, the absence of other domains possibly impeding this interaction in the full-length protein, better transport into the nucleus and more uniform nuclear localization compared to the speckled pattern of full-length AIRE (Figure 11F). Both full-length AIRE and HSR/CARD+ fragment showed weaker binding compared to plasmid when tested in genomic settings; however, the pattern of HSR/CARD+ stronger binding over full-length AIRE was preserved (Figure 11B). While full-length AIRE bound with similar strength to target and house-keeping genes as well as promoter and 3’ gene regions, HSR/CARD+ fragment showed the highest enrichment at

house-keeping gene promoters β_2-microglobulin (*B2M)* and *HPRT1* with lower levels at their 3' gene regions. The enrichment at target genes (*IVL* and *S100A8*) was moderate and similar in both promoter and 3' regions (Figure 11B). The differences in enrichment on various genomic locations observed with HSR/ CARD+ fragment were not recurred with full-length AIRE possibly due to generally lower levels of binding, preventing distinction of interacting protein amounts. The preferred binding of HSR/ CARD+ fragment to promoters of active genes is in line with a previous study (Giraud et al., 2012), where full-length AIRE binding was assessed in genome-wide ChIP experiment and was found to accumulate at TSSs of Pol II-rich genes.

To explore DNA binding capacity of different HSR/CARD+ regions we generated and purified GST-tagged AIRE fragments from bacterial lysates, and tested them in EMSA experiments with oligonucleotides or plasmid DNA. With both oligonucleotide and circular or linearized plasmid DNA, we detected the shift with 1-138 aa long fragment but not with 1-109 aa (Figure 11C, 11D), suggesting that binding activity could be appended to the stretch containing the amino acids 113-133 where bipartite nuclear signal was initially localized (The Finnish-German APECED Consortium, 1997). In later study, it was established that only the latter (131-133 aa) of two parts (Ilmarinen et al., 2006) was necessary for transport into the nucleus rendering the nuclear signal monopartite. As amino acids 113 and 114 are well conserved between organisms (Study I Figure 4A), we next tested how DNA binding is affected if positively charged arginine 113 and lysine 114 are mutated to neutral alanines. HSR/CARD+ R113A/K114A construct was not able to bind to pBL-IVL plasmid demonstrating low enrichment comparable with negative control, while full-length wt AIRE and wt HSR/CARD+ bound efficiently (Figure 11E). Next, we checked if mutated HSR/CARD+ was localized inside the nucleus, performing immunofluorescence staining. Both HSR/CARD+ and HSR/CARD+ R113A/K114A were effectively transported into the nucleus, although their localization was more uniform skipping the nucleoli, compared to the speckled pattern of full-length AIRE (Figure 11F). While we usually observed some cytoplasmic fiber staining with transfected AIRE, it was not visible with shorter fragments; HSR/CARD+ R113A/K114A also tended to form some larger aggregates inside the nucleus. HSR/CARD without nuclear signal was retained in the cytoplasm (Figure 11F).

In conclusion, these data indicate that the integrity of the HSR/CARD domain and amino acids R113 and K114 are important for AIRE DNA binding to plasmid DNA.

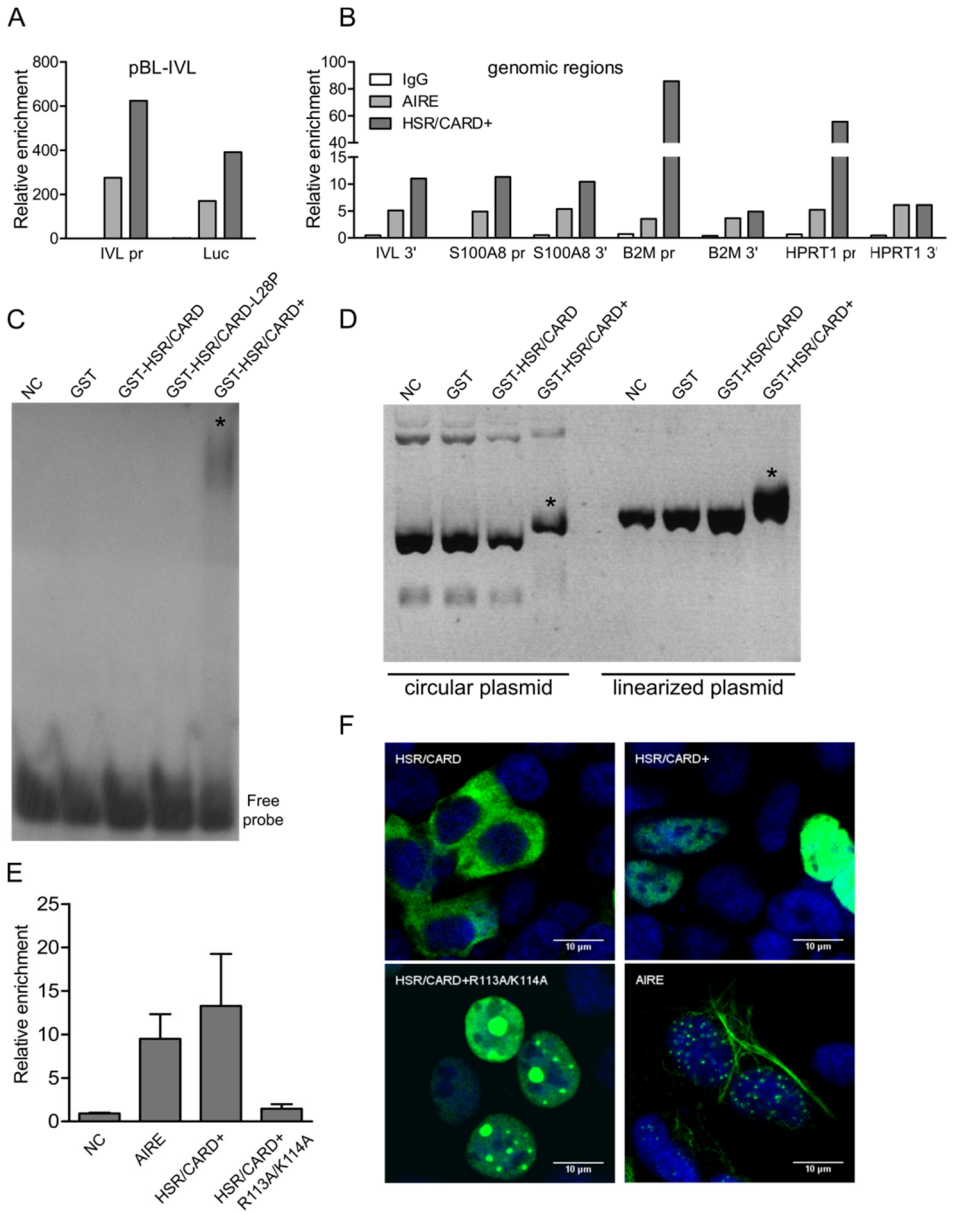

Figure 11. AIRE N-terminal region aa 1-143 binds directly to DNA and aa R113 and K114 are important for the DNA binding activity. (A, B) HEK293 cells were transfected with pBL-IVL and pcAIRE (AIRE) or pcHSR/CARD+ (HSR/CARD+) followed by ChIP with anti-Myc for AIRE and HSR/CARD+ or anti-IgG (IgG) antibodies. qPCR was performed with the primers recognizing the reporter plasmid (IVL pr, Luc) or genomic sequences of AIRE target (IVL and S100A8) and housekeeping genes (B2M, HPRT1). Promoter (pr) and 3' regions were tested in genomic regions. **(C)** An autoradiogram of the polyacrylamide gel from EMSA with indicated recombinant protein fragments and ^{32}P-labeled probe alone. The shifted complex is marked with an asterisk. **(D)**

An ethidium bromide-stained agarose gel image showing EMSA with the indicated recombinant protein fragments incubated with circular or linearized pcDNA3. The shifted complexes are marked with asterisks. NC indicates reaction mixture without protein. **(E)** HEK293 cells were transfected with pcDNA3 (NC – negative control), pcAIRE (AIRE), HSR/CARD+ or HSR/CARD+R113A/K114A together with pBL-IVL reporter plasmid, and ChIP was performed with anti-Myc antibody. The data is shown as mean ± SEM of relative enrichment of AIRE-to-IgG from two experiments (NC = 1). **(F)** Immunofluorescence analysis of HEK293 cells transfected with Myc-tagged AIRE and indicated fragments. The cells were stained with DAPI (blue) and anti-Myc Alexa488 (green).

5.5. AIRE interacts with DNA-PK via PHD1 domain

Before extensive screening for AIRE interacting proteins (Abramson et al., 2010; Gaetani et al., 2012), only few proteins had been shown to interact with AIRE. One intriguing interacting partner emerged when we performed GST pull-down in nuclear extract from monocytic THP-1 cells using GST-tagged AIRE PHD1 (Figure 12A). Three protein bands with approximate weight of 70, 80 and 250 kDa appeared when separated on SDS-PAGE and stained with Coomassie (Figure 12B). The mass spectrometry analysis of corresponding bands excised from the gel identified them as Ku70, Ku80 and DNA-PKcs (DNA-dependent protein kinase, catalytic subunit), respectively (Study II Table 1). The full-length AIRE was confirmed to interact with DNA-PKcs by co-immunoprecipitation with anti-DNA-PKcs antibody in pcAIRE-transfected HEK293 cells (Study II Figure 2A). This interaction proved to be specific as other antibodies against GST, GAPDH (glyceraldehyde 3-phosphate dehydrogenase) and RNA-binding HuR (human antigen R) proteins were not able to precipitate AIRE. Anti-AIRE antibody was used as a positive control, as AIRE forms homomultimers. To test this interaction for DNA and RNA dependency, we repeated the co-immunoprecipitation experiment in the presence of ethidium bromide, RNaseA/T or MNase. While ethidium bromide and RNase did not influence the binding strength, micrococcal nuclease weakened the interaction (Study II Figure 2C), indicating that DNA conformation and RNA presence is not necessary for this interaction; however, the chromatin can mediate this interaction partly.

DNA-PK is expressed in most tissues (Moll et al., 1999). To assess whether DNA-PK and AIRE interaction can occur *in vivo,* we measured DNA-PK expression level in murine thymic stromal cells. We isolated mTECs from murine thymi based on the cell surface marker EpCAM and tested the expression of DNA-PK, AIRE and two AIRE target genes (Figure 12C, D). DNA-PK was highly expressed in mTECs together with AIRE (Figure 12D). The expression of IVL and LOR was lower, but considerably enriched in mTECs compared to cTECs (Figure 12C).

While DNA-PK is well described in the DNA repair processes, it was initially characterized in transcriptional complexes (Dvir et al., 1992; Jackson et al., 1990), where its phosphorylation activity was associated with positive regulation of transcription. We next demonstrated that the AIRE protein is phosphorylated

by DNA-PK at threonine 68 and serine 156 and that replacing these amino acids with alanines diminished AIRE-mediated activation (Study II Figure 7). While pull-down via the PHD1 domain discovered DNA-PK as potential partner, its possible role by phosphorylating residues in AIRE N-terminal part suggested more extensive contact.

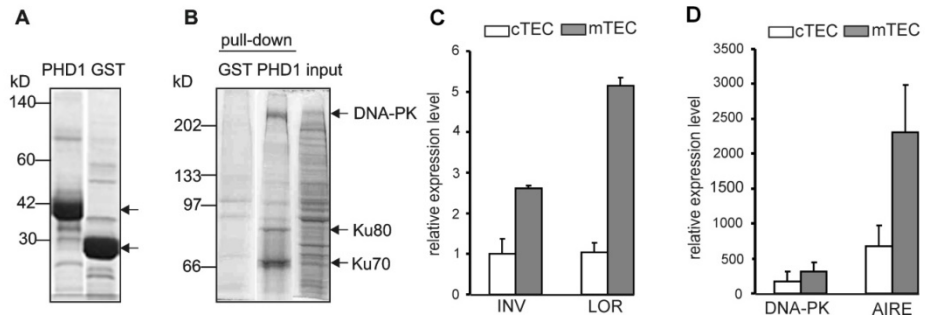

Figure 12. AIRE PHD1 domain as well as full-length AIRE interacts with Ku70, Ku 80 and DNA-PKcs. (A) Coomassie staining of purified GST-AIRE-PHD1 and GST, indicated by arrows. **(B)** Affinity purification of AIRE-PHD1 binding proteins from THP-1 nuclear extract. Three protein bands that specifically bound to AIRE PHD1 and were analyzed by MALDI-TOF are indicated by arrows. Input represents 5% of the amount of nuclear extract used in the pull-down assay. **(C, D)** DNA-PK is expressed in mTECs. The expression of DNA-PK, AIRE and AIRE target genes involucrin and loricrin was assessed by qPCR in murine cTECs and mTECs. The relative expression levels are compared to involucrin level in cTEC cell population (=1), the values (± SEM) are mean of two independent quantitative RT-PCR reactions both performed as triplicate measurements.

5.6. TOP2 inhibitor etoposide enhances the expression of AIRE target genes

A more recent screening of AIRE partners determined many interacting proteins, subdivided into four large groups of nuclear transport, chromatin structure, pre-mRNA processing and transcription (Abramson et al., 2010). Among others, a large complex centered on DNA-PK was described containing besides AIRE and DNA-PK, Ku80, Ku70, PARP-1, H2A.X and TOP2A. These early-acting elements of the NHEJ DNA repair machinery were suggested to be recruited after TOP2A generated DSB was stabilized by AIRE. As a subset of genes activated by TOP2A inhibitor etoposide and AIRE overlapped tightly, the hypothesis that AIRE acts similarly to TOP cleavage-stabilizing inhibitors was proposed (Abramson et al., 2010).

To assess whether the effect of AIRE on gene expression may be due to its action through DNA repair machinery, we next applied a total RNA-seq technique to describe in parallel the genes activated by AIRE and TOP2 inhibitor

etoposide that prolongs DNA break existence. For the expression of AIRE, we used doxycycline-inducible HEK293 cell line (AIRE-Tet) (Liiv et al., 2012). A total RNA from four conditions, DMSO-treated AIRE-Tet cells as negative control (Ctrl), AIRE-Tet cells induced with 1.5 μM doxycycline for 48 h (Dox), 2 μM etoposide for 24 h (Etop), or a combination of both (Dox+Etop) was sequenced. After AIRE induction, we detected the upregulation of 691 genes and only one gene, encoding a long noncoding RNA *SNORD3A* (small nucleolar RNA C/D box 3A), down-regulated (Figure 13B, upper panel). Thus, we observed a strong activation by AIRE in HEK293 cells, where 99,9% of influenced genes were upregulated.

The addition of etoposide to AIRE-positive cells increased further the number of activated genes (6118); however, considerable fraction of down-regulated genes appeared (2215) (Figure 13B, middle panel). When cells without AIRE were treated with etoposide, almost equal parts of genes were up- and down regulated (2452 and 1961, respectively) (Figure 13A, 13B, lower panel). Interestingly, almost all genes activated by AIRE were also upregulated in AIRE plus etoposide conditions (687 of 691, 99,4%) (Figure 13A, 13D). Simultaneously, the higher upregulation of 691 AIRE target genes was achieved in AIRE plus etoposide (Dox+Etop) cells compared to AIRE positive cells (Dox) (Figure 13E) and the same set of genes did not show any change in expression when treated with etoposide only (Etop) (Figure 13E). Supporting this observation, we noticed that genes activated separately by AIRE (Dox) and etoposide (Etop) overlapped very marginally (Figure 13D), indicating that AIRE and etoposide mode of action is mechanistically different. Among other dissimilarities, we observed that the initial expression level of AIRE activated genes was low (Figure 13B, upper panel), while etoposide-influenced genes were characterized by diverse initial expression (Figure 13B, lower panel). In summary, when AIRE positive cells were treated with etoposide, the influence of AIRE prevailed with 73% of differentially expressed (DE) genes being upregulated, of which 81% were initially expressed at low level. In alignment with this, the tissue-specific genes were enriched in AIRE positive and etoposide treated AIRE positive cells, while etoposide-treated cells did not show any preference (Figure 13C). The distinct profiles of gene expression changes in AIRE-positive cells (Figure 13B, upper panel) and in etoposide treated cells (Figure 13B, lower panel) suggests that AIRE function does not mimic that of TOP2 inhibitor.

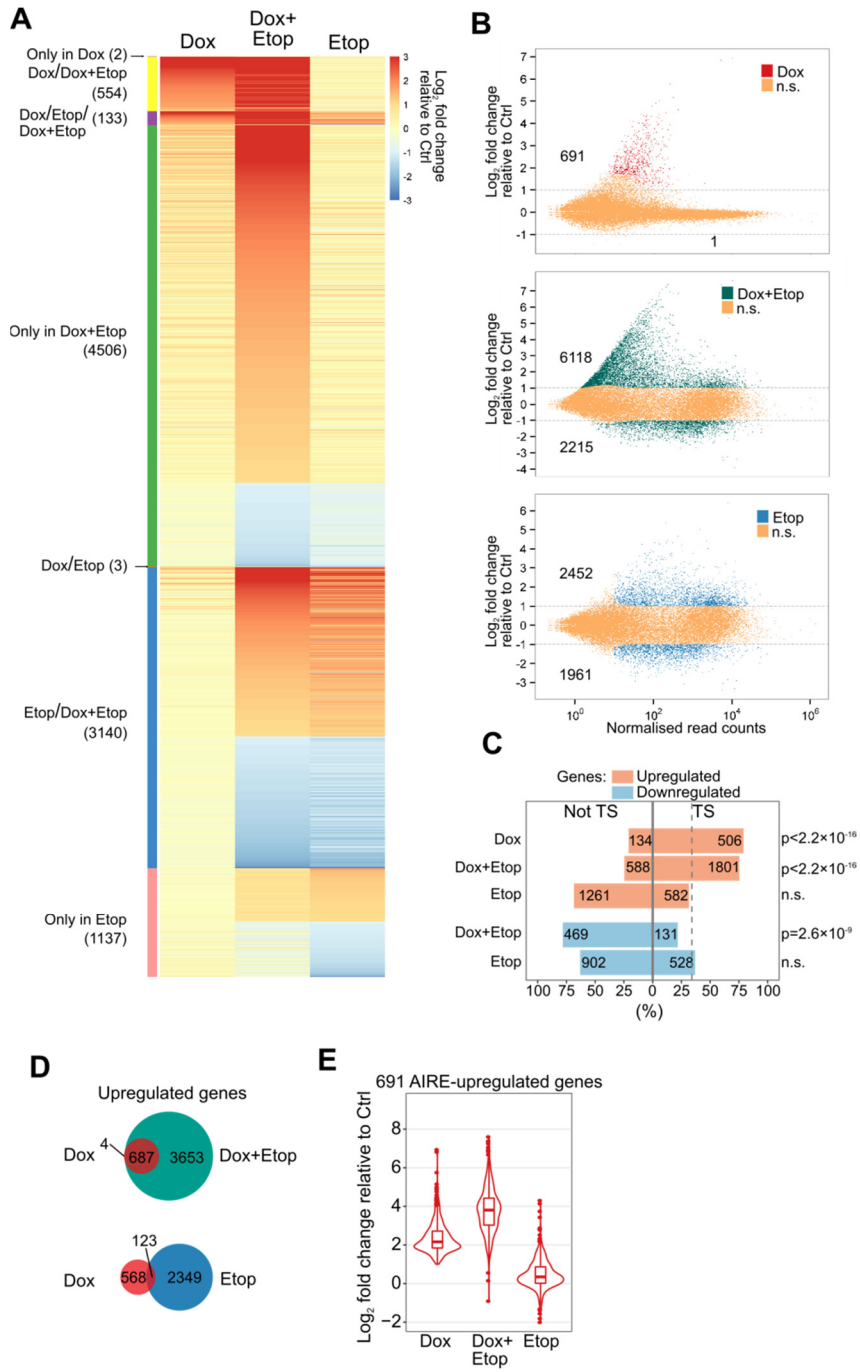

Figure 13. Etoposide treatment expands the repertoire of AIRE target genes and enhances their expression in AIRE-Tet cells. (A, B) Assessment of genome-wide gene expression changes with RNA-seq after doxycycline (Dox), etoposide (Etop), or Dox+Etop treatment in the AIRE-Tet cell line relative to control (Ctrl). **(A)** Heatmap represents log2-

fold changes of differentially expressed (DE) genes across all conditions relative to Ctrl. The text and color bar on the left show the division of genes between indicated conditions. **(B)** Differential gene expression profiles. The colors and numbers depict DE genes with an adjusted $p < 0.1$ and absolute log2-fold change > 1 (gray lines); n.s., not significant. **(C)** The percentage and the number of tissue-specific (TS) and broadly expressed (not TS) genes among the DE genes. The dashed line marks the percentage of TS genes in the genome (34%). **(D)** The number of unique and overlapping upregulated genes in Dox-, Etop-, and Dox+Etop-treated AIRE-Tet cells. **(E)** The distribution of the effect sizes of the 691 AIRE-upregulated genes after Dox, Dox+Etop, or Etop treatment.

5.7. AIRE and etoposide induce alternative splicing in subsets of genes different from differentially expressed genes

The idea that AIRE can influence splicing to expand the scope of different protein isoforms to be presented to maturing thymocytes seems plausible. Indeed, several studies had demonstrated that AIRE enhances the production of spliced transcripts by interacting with pre-mRNA processing proteins and co-localizing with nuclear speckles-containing splicing machinery (Abramson et al., 2010; Giraud et al., 2014; Oven et al., 2007). From our RNA-seq dataset, it revealed that the number of genes featuring differential exon usage (DEU) following AIRE induction was in approximately the same range (568 genes) with DE genes. Adding etoposide influenced positively alternative splicing, as the number of DEU genes increased in both Dox+Etop and Etop cells, but the increase was very similar (1989 and 1850 DEU genes, respectively) (Figure 14A). Furthermore, a large fraction of the DEU genes found in Dox cells overlapped with DEU genes both in Dox+Etop and Etop cells (428 of 568 and 406 of 568, respectively) (Figure 14A), demonstrating different behavior from what we have seen with DE genes (Figure 13D). This suggests that the DEU genes are not synergistically affected by AIRE and etoposide combination, and the effect on splicing is functionally more similar between AIRE and etoposide the one on transcription. Additionally, there was a very little overlap between the DE and DEU genes in Dox samples (50 of 568 genes) (Figure 14B). The overlap was larger in Dox+Etop (654 of 1989) and Etop (582 of 1580) cells (Figure 14B), but this was mostly due to a greater number of DEU genes in these populations. In general, the alternatively spliced genes tended to be highly expressed (Figure 14C) and were less likely to represent tissue-specific genes as compared with AIRE target genes (Study III supplemental Fig. 2B). However, it should be noted that the initial low expression of AIRE target genes might hinder the analysis of alternative splicing, as the higher expression of the gene, the more potential material for alternative splicing available. In summary, the influence of etoposide on AIRE-mediated transcription and pre-mRNA splicing had different patterns.

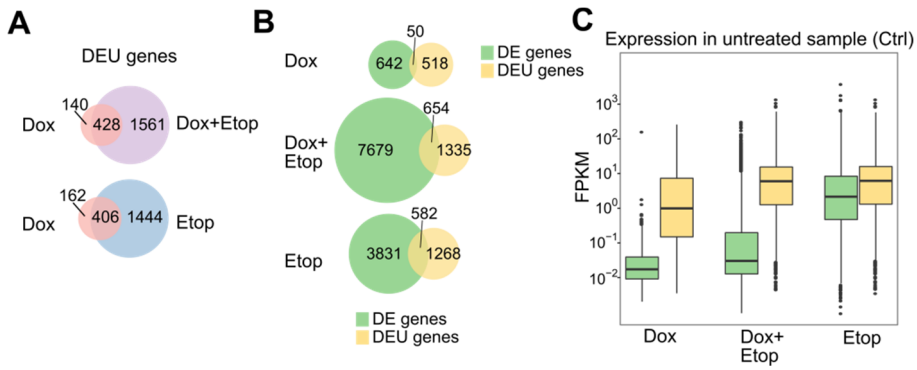

Figure 14. AIRE and etoposide induced differential exon usage (DEU) mainly in non-DE genes. (A) The number of unique and overlapping DEU genes in Dox-, Etop-, and Dox+Etop-treated AIRE-Tet cells. **(B)** The overlap between DE and DEU genes under Dox, Dox+Etop, and Etop conditions. **(C)** Comparison of the initial expression in negative control (Ctrl) of DE and DEU genes in indicated conditions. The y axis shows the normalized expression level in fragments per kilobase of exon per million mapped reads (FPKM).

5.8. AIRE target genes are organized in clusters

The early notification that AIRE-activated genes are organized in clusters (Derbinski et al., 2005; Johnnidis et al., 2005) lead to the speculation that epigenetic or spatial organization mechanisms could be involved in AIRE-mediated regulation. Later, single cell sequencing revealed that in individual cell, the genes forming the cluster are not activated simultaneously, but rather individual genes from different clusters are induced in each cell (Meredith et al., 2015; Sansom et al., 2014).

We tested whether DE genes in our conditions localize randomly in genome or tend to accumulate in clusters. We characterized cluster formation by the number of clusters, size (less or more than 5 genes), the total number of genes combined into clusters and the probability that cluster is not generated by random juxtaposition of genes in genome ($-\log_{10}$ of FDR-adjusted p-value). The highest number of genes organized in clusters and the highest number of clusters with tendency to longer clusters was observed within upregulated genes in Dox+Etop condition (Figure 15). In total, 44% of DE genes that were induced in Dox+Etop condition were in clusters (an average 4.7 genes per cluster). Among the downregulated genes in Dox+Etop cells, their tendency to cluster was much weaker (Figure 15). The clusters in Dox cells, although shorter in length, were characterized by higher p-values (Figure 15). Altogether, we observed highly significant chromosomal clustering of 22% of AIRE-induced genes with an average of 4.5 genes per cluster.

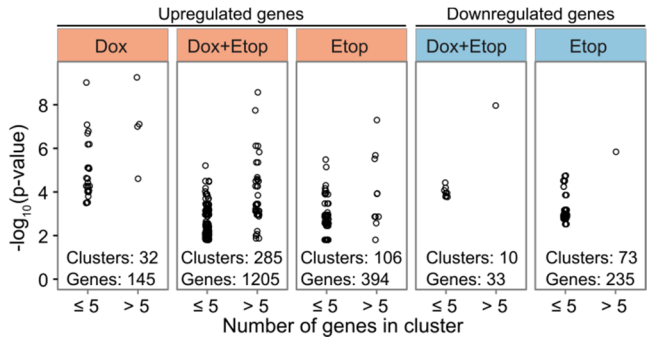

Figure 15. Clusters of closely located genes among the up- and downregulated genes in each sample relative to the Ctrl. The y-axis represents the FDR-adjusted p-values of clustering.

5.9. TOP2A silencing and AIRE HSR/CARD domain disruption decrease target gene activation

We next used short-hairpin (sh)RNA to knock down TOP2A expression in AIRE-Tet cells. The effect of gene transcription was significantly reduced in cells transfected with a TOP2A-specific shRNA (shTOP2A4) compared with the scrambled control shRNA (Scrambled) that retained the ability to enhance AIRE target gene expression (Figure 16A). While almost all of the tested AIRE target genes demonstrated reduced activation in the cells treated with shTOP2A in Dox and Dox+Etop conditions, AIRE-independent gene *PSMD4* showed lower expression in Dox+Etop conditions only. These results indicate that AIRE target genes, but also other genes are dependent on TOP2A function.

The HSR/CARD domain integrity was shown to be crucial for the AIRE protein homomultimerization as well as interaction with DNA (Study I) and other proteins (Abramson and Goldfarb, 2016). Concordantly, wt AIRE, but not AIRE with L28P mutation introduced into the HSR/CARD domain, interacted with TOP2 (Figure 16B, 16C). The same mutation also prevented target gene activation with and without etoposide treatment (Figure 16D). Thus, these data indicate that the integrity of HSR/CARD is important for the interaction with TOP2A, DNA or other proteins in this complex and the disruption of these contacts prevents the activation of AIRE target genes. The MNase treatment of lysates prior to IP does not exclude the possibility of indirect contact between TOP2A and AIRE through DNA and other proteins present in the mixture, as L28P mutation in HSR/CARD domain hinders interaction with DNA (Study I) and possibly other proteins.

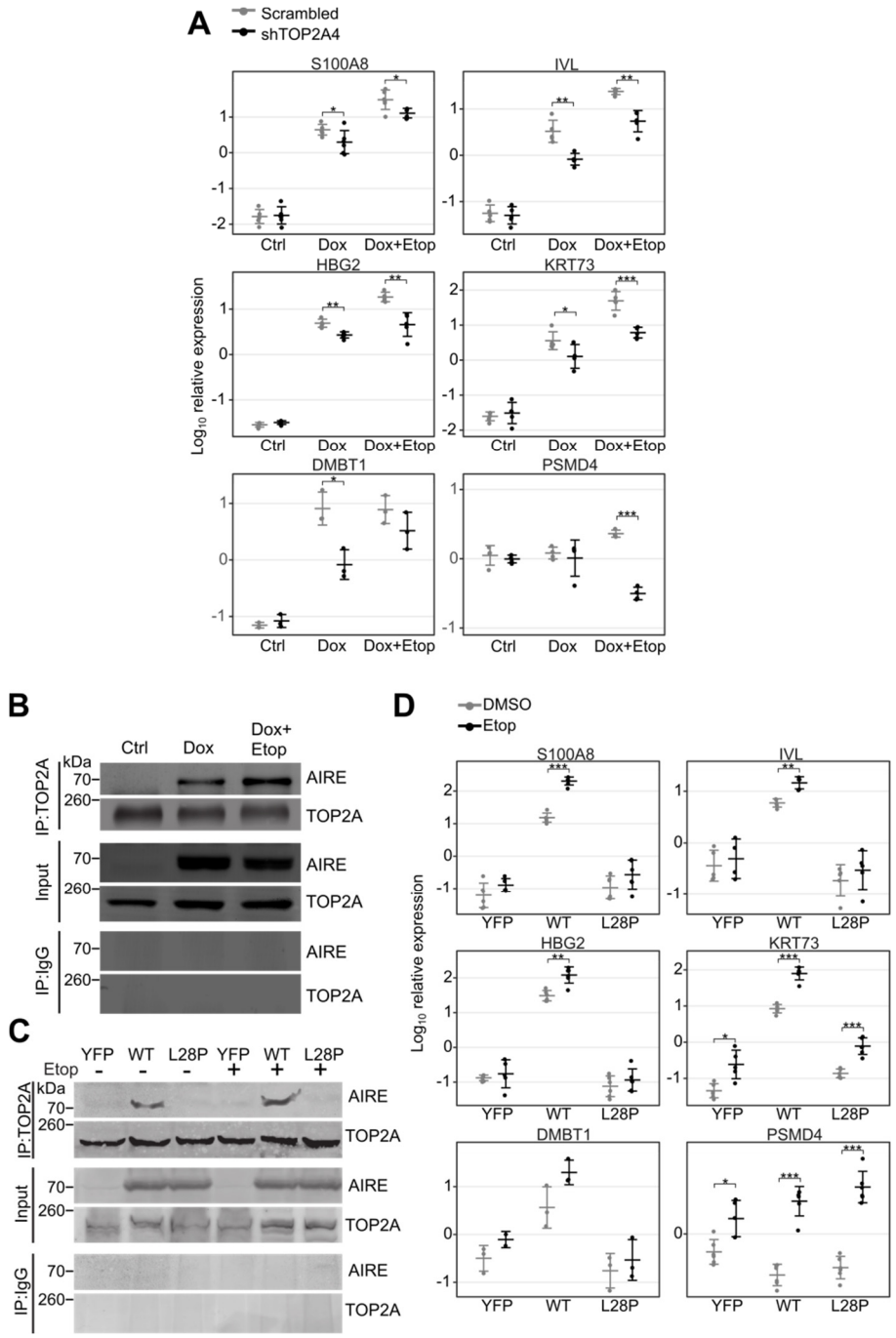

Figure 16. Knockdown of TOP2A reduces the activation of AIRE target genes, and AIRE HSR/CARD domain mutation abolishes the AIRE-TOP2A interaction. (A) The expression analysis of indicated AIRE-dependent genes and an AIRE-independent *PSMD4* gene in AIRE-Tet cells using qPCR after shRNA-mediated knockdown of

TOP2A. **(B)** The analysis of the AIRE and TOP2A interaction in Ctrl, Dox, and Dox+Etop AIRE-Tet cells by co-IP. **(C)** The co-IP of AIRE and TOP2A interaction in HEK293 cells stably expressing YFP, wt AIRE, or mutant AIRE (L28P) with or without etoposide treatment. Anti-TOP2A and anti-IgG antibody as a negative control were used for immunoprecipitation followed by Western blot with anti-TOP2A and anti-AIRE antibodies. Input was used as a protein loading control. **(D)** The expression analysis of indicated AIRE-dependent genes and an AIRE-independent gene *PSMD4* in HEK293 cells stably expressing YFP, wt AIRE, or mutant AIRE (L28P). **(A, D)** Log10-transformed data points are shown as mean ± SD (standard deviation) and are from three to six independent experiments. Statistical significance was assessed by unpaired *t*-test (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

5.10. TOP1 inhibition with DNA break formation amplifies AIRE-mediated gene activation

TOPs resolve topological stress created by molecular processes requiring DNA double-stranded helix unwinding. TOP1 creates transient single strand break (SSB) that facilitates the tension relief by spinning one DNA strand around the other, TOP2 generates double strand break and mediates the passage of one DNA duplex through the other (Figure 3). Using different chemicals, it is possible to inhibit catalytic activity of both types of TOP before DNA break formation, leaving DNA structure intact, or after DNA cleavage, prolonging DNA break existence. Thus, we next applied etoposide to inhibit TOP2, and camptothecin to inhibit TOP1 after the brake formation, as well as merbarone for TOP2 and β-lapachone for TOP1, to inhibit TOP catalytic activity without DNA lesion (Figure 17A). In accordance with the type of inhibition, we detected higher percentage of cells with Br-dUTP-labeled breaks by flow cytometry in etoposide and camptothecin conditions; while merbarone and β-lapachone–treated cells were similar to control (Figure 17B). In parallel with DSB or SSB formation, the activation of AIRE target genes *S100A8* and *CEACAM5* increased with addition of TOP poisons etoposide or camptothecin while merbarone and β-lapachone did not influence these genes (Figure 17C).

This experiment demonstrates that rather the generation of single- or double-strand breaks is important for AIRE-mediated activation than TOP specificity and activity. Whether single-strand breaks can be transformed into double-strand breaks in post-mitotic AIRE-positive cells *in vivo* and whether TOP1- and TOP2-mediated breaks involve the same pathways associated with activation by AIRE requires further investigation.

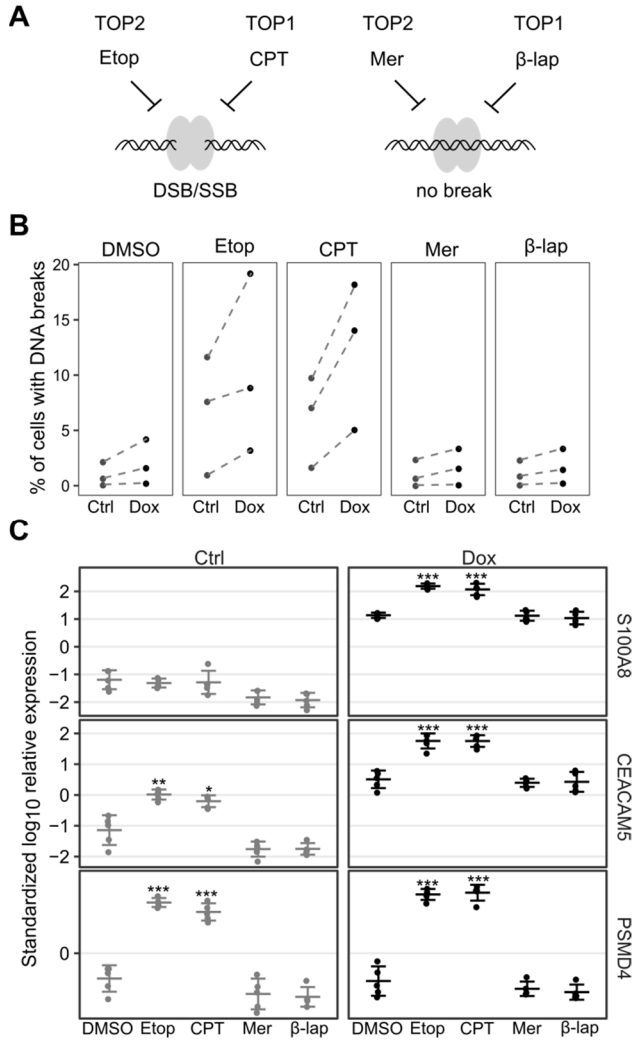

Figure 17. AIRE-dependent gene expression relies on the DNA cleavage activity of TOP2 or TOP1. (A) Schematic of the mechanisms of TOP1 and TOP2 inhibition by etoposide (Etop), camptothecin (CPT), merbarone (Mer), and β-lapachone (β-lap). The gray ellipses and the helix represent the TOP homodimer and the DNA molecule, respectively. SSB, single-strand break; DSB, double-strand break. **(B)** Percentage of uninduced (Ctrl) and induced (Dox) AIRE-Tet cells with DNA breaks after mock treatment with DMSO, TOP2-specific (Etop and Mer), and TOP1-specific inhibitors (CPT and β-lap) using Br-dUTP-labeled nicked DNA measured by FACS. The data are derived from three independent experiments. Dashed lines connect experimental groups from the same experiment. **(C)** The expression of the AIRE target genes *S100A8* and *CEACAM5* and the AIRE-independent gene *PSMD4* after DMSO, Etop, CPT, Mer, and β-lap treatment in Ctrl and Dox AIRE-Tet cells. Log10-transformed data (mean ± *SD*) are from four to five independent experiments. Unpaired *t*-test is performed relative to the mock (DMSO) treatment separately in Ctrl and Dox conditions (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

6. DISCUSSION

6.1. AIRE ability to activate reporters without regulatory sequences is in line with its universal capacity to augment gene expression

Our first finding in this study was that AIRE is able to activate Luc reporters without promoter sequences (Study I). This unusual ability of AIRE could be interpreted as the capacity to efficiently use short cryptic sequences for transcription activation or to push pre-formed initiation complexes outside gene promoters into productive elongation (Venters and Pugh, 2013). That could also explain described AIRE usage of alternative TSSs (Villaseñor et al., 2008). Though, in our total RNA-seq experiments we could not detect transcripts outside annotated gene starts, which could be due to their unstable nature (Davis and Ares, 2006; Thompson and Parker, 2006). The promoter strength is decisive for AIRE-mediated activation. While reporters containing none or weak promoters, such as endogenous target gene, are activated by AIRE (Study I), strong promoters (SV40 and CMV) are not subjected to AIRE influence (Giraud et al., 2014).

Our current (Study I) and many previous works (Mathis and Benoist, 2009; Peterson et al., 2008) suggest that AIRE does not have any sequence specificity for DNA binding, despite some reports claiming inclination for certain motifs (Kumar et al., 2001; Purohit et al., 2005; Ruan et al., 2007). Both genomic sequences and reporter vectors could contain cryptic transcription initiation sites (Lemp et al., 2012; Vopálenský et al., 2008) due to the indiscriminate nature of eukaryotic transcription. Scanning for specific elements might not be useful as Inr similar dinucleotide combinations could occur very frequently. Transcription in the mammalian genome can be initiated at many loci, with productive transcription being limited to certain functional product containing genes (Kapranov et al., 2007). The biological significance of such pervasive transcription is not completely understood (Jacquier, 2009), but could explain why we observe reporter induction by AIRE with plasmids without any regulatory elements.

AIRE-induced promiscuous expression has been shown to have strong features of stochasticity in thymic epithelial cells (Guerau-de-Arellano et al., 2008; Jiang et al., 2005; Venanzi et al., 2008; Villaseñor et al., 2008). However, eukaryotic gene expression is not deterministic in general (Blake et al., 2003; Elowitz et al., 2002). Stochastic mRNA synthesis in genetically identical mammalian cell population was demonstrated by counting individual molecules of RNAs produced from the reporter integrated in the genome of Chinese hamster ovary (CHO) cells (Raj et al., 2006). Accordingly, it has been proposed that mRNAs are synthesized in short but intensive transcription bursts with burst size regulated by promoter type and burst frequency by enhancer (Haberle and Stark, 2018). Modulating transcription burst frequency through enhancer is more prevalent way to regulate the intensity of transcription (Fukaya et al., 2016). In light of this

knowledge, AIRE seems to be a unique factor that has capacity to induce stochastic pervasive background transcription potentially present in every cell.

The presence of intron can drastically increase the effectiveness of transcription, for example, beta globin intron intensifies RNA production hundredfold in plasmid reporter (Buchman and Berg, 1988). Alternative splicing and polyadenylation are the ways to expand the repertoire of isoforms, and presentation of those different protein forms in the thymus can be crucial for avoiding autoimmunity (Klein et al., 2000). Considering the interaction of AIRE with splicing and RNA processing machinery (Abramson et al., 2010), we tested how the presence of introns and poly(A) sequences influences AIRE-mediated transcriptional activation. We found that AIRE can activate to similar extent intronless plasmids and reporters containing histone stem-loops instead of poly(A) signal. In our RNA-seq experiments, we found that AIRE enhanced alternative splicing (Study III), but alternative exon usage was noted in a subset of genes different from differentially expressed genes. While transcription and splicing are tightly interconnected processes (Herzel et al., 2017; Kornblihtt et al., 2004) with splicing taking place as soon as nascent transcript emerges and exon inclusion is influenced by the speed of progressing Pol II (de la Mata et al., 2003), it is intriguing to observe such division into two separate gene sets.

6.2. AIRE binding to DNA

We characterized the DNA-binding ability of AIRE protein mediated by extended HSR/CARD+ domain (1–143 aa). Within this domain, mutations that disrupt the integrity of HSR/CARD (1–100 aa) and mutations of amino acids 113-114 outside the HSR/CARD domain drastically reduced the binding (Study I). Although amino acids 113-114 in AIRE protein were earlier described to be responsible for nuclear transport (The Finnish-German APECED Consortium, 1997), their mutations do not interfere with AIRE nuclear localization (Ilmarinen et al., 2006). In other proteins, DNA-binding domain and NLS signal overlap has been described, which could be evolutionarily beneficial, as DNA-binding proteins should be transported into and kept in the nucleus to exert their functions (Boulikas, 1994; Cokol et al., 2000).

Several studies have shown or suggested AIRE binding to DNA. The first report about AIRE binding to DNA was published in 2001, when PHD1 with potential leucine zipper was proposed to be responsible for AIRE dimerization and DNA interaction (Kumar et al., 2001). In this study, AIRE monomers were not able to bind oligonucleotides, but dimers and tetramers obtained by oxidation and refolding method, usually applied to isolate protein aggregates from bacterial inclusion bodies, showed preferential binding to GG in A/T rich environment (Kumar et al., 2001). Albeit the binding to DNA was firstly attributed to PHD zinc fingers, the function was later associated with SAND domain (Purohit et al., 2005). Though, as in a more recent work, the mutations introduced into the potential DNA-binding motif of SAND domain did not interrupt DNA binding, a

strong interaction of AIRE with DNA was alternatively proposed to be mediated via DNA-PK (Zumer et al., 2012). However, EMSA experiments shown here (Study I) and by others (Koh et al., 2008) found AIRE binding to DNA to be independent of DNA-PK or other proteins. Furthermore, EMSA experiments demonstrated AIRE interaction with both reconstituted mononucleosomes and the free DNA used to reconstitute them, but with the greater affinity towards naked DNA (Koh et al., 2008). Importantly, the removal of PHD1, PHD2 or SAND domains did not disrupt this binding (Koh et al., 2008), which is in line with our results demonstrating that AIRE DNA binding is mediated by extended HSR/CARD+ domain. Although initial transactivation experiments in 2000 were conducted tethering of Gal4-AIRE fusion protein to reporter plasmids via Gal4-binding domain, the phenomenon that AIRE was able to activate the same plasmids even without Gal4 fragment was noticed (Björses et al., 2000; Pitkänen et al., 2000). While Gal4 tethering was still used in later experiments (Žumer et al., 2011) to achieve stronger activation (up to hundredfold) (Björses et al., 2000), our results confirm that AIRE is able to activate reporter genes efficiently (10-30-fold), even without fused additional DNA binding domains (Study I).

Several studies have demonstrated that AIRE target genes are characterized by low level of initial expression and repressive chromatin marks H3K27me3 and H3K9me3 (Handel et al., 2018; Sansom et al., 2014) as well as lack of active marker H3K4me3 (Koh et al., 2010; Org et al., 2008). In contrast, ChIP-on-ChIP experiments showed that the strongest AIRE binding at TSS of genes with high expression and high Pol II levels, however, the expression of these genes was not influenced by AIRE (Giraud et al., 2012). These ChIP-on-ChIP results could reflect the ubiquitous preferential binding of AIRE to accessible chromatin, which is characteristic to many TFs (Lambert et al., 2018). The interaction with accessible chromatin suggests that for the specific target gene activation, AIRE requires additional factors, which are present at genes with low expression level, or just indicates the inability of AIRE to enhance the expression of highly expressed genes. Considering the reported capacity of AIRE to release paused Pol II, the recirculation of transcription machinery on highly expressed genes could be already so effective, that AIRE might not be able to increase it. In our ChIP experiments, the full-length AIRE enrichment did not show differences between AIRE target and tested housekeeping genes (Study I). This could be because of differences in chromatin structure of HEK293 cells as compared with mTECs, altered modification pattern of overexpressed transfected AIRE versus endogenous AIRE, or relative insensitivity of qPCR-based analysis of ChIP material. On the other hand, ChIP with extended HSR/CARD+ domain demonstrated the highest binding to promoters of highly expressed genes and not near target genes (Study I), which indicates that previously observed AIRE preferential binding to TSSs of highly expressed genes (Giraud et al., 2012) may also reflect the capacity of AIRE to directly bind to DNA via HSR/CARD+ domain.

AIRE ability to favor paused Pol II into productive elongation was demonstrated by applying specifically designed plasmids that differ in ability to recruit and load Pol II at promoters and thus to induce initiation and/or elongation (Oven

et al., 2007). This finding was supported by detecting on wide scale first exon skewing of Aire target genes with more transcripts produced from the first exon relative to the whole gene in Aire knock-out mice, thus demonstrating Pol II promoter-proximal pausing in the absence of Aire (Giraud et al., 2012). However, we could not detect increased Pol II CTD Ser2 modification accompanying the transition from initiation to elongation stage on AIRE target genes in genomic context (data not shown). Elevated levels of Pol II phosphorylated at Ser2, especially near 3'-end, were reported on plasmid target when AIRE protein was tethered by Gal4 (Žumer et al., 2011). We could detect high amount of AIRE protein bound to naked plasmid DNA devoid of nucleosomes and histones, and the ability of AIRE to activate reporters in these conditions (Study I). Interestingly, the addition of etoposide did not enhance the expression of plasmid reporters further (data not shown), indicating that additional DNA breaks created by inhibiting TOPs and thereby increased amount of chromatin-free DNA may own more importance in genomic context.

At first glance, AIRE deviates from traditional TFs, though, accumulating evidence of eukaryotic transcription suggests that simple model of DNA recognition motif dictating specific binding sites near the regulated genes may not be necessarily the case for the majority of eukaryotic TFs (Lambert et al., 2018).

6.3. AIRE-mediated transcriptional activation involves DNA damage repair machinery

In the current study, we identified AIRE interaction with DNA-PK, a protein mainly associated with DNA damage repair (Study II). The functional significance of AIRE interaction with DNA-PK has been intriguing question. Despite the extensive studies on DNA-PK role in DNA repair, the first reports about the DNA-PK protein described its involvement in transcription by phosphorylating SP1 (Jackson et al., 1990) and paused RNA Pol II (Dvir et al., 1992). We demonstrate here the ability of DNA-PK to phosphorylate AIRE at N-terminal amino acids T68 and S156 and show that mutations of these amino acids to alanines decrease the target gene activation (Study II). In subsequent work, inhibition of DNA-PK catalytic activity with NU7441 did not suppress AIRE-mediated transcriptional activation, invoking alternative hypothesis of DNA-PK requirement for AIRE binding to DNA (Zumer et al., 2012). Later, thorough screening of AIRE interacting proteins revealed several other proteins to be involved in DNA damage repair processes, including helicases, TOPs and ribosyltransferases to associate with AIRE (Abramson et al., 2010; Gaetani et al., 2012). Abramson *et al*. proposed that AIRE mimics the action of etoposide by stabilizing TOP2A-created breaks that activate DNA-PK and other DNA repair associated proteins. Of note, this hypothesis was made because etoposide but no other DNA break-inducing compounds, for instance, H_2O_2, demonstrated the activating effect on AIRE target genes (Abramson et al., 2010).

To test the hypothesis about similarity of AIRE and etoposide effects, we used HEK293 cell line with inducible AIRE and studied the effects of etoposide, AIRE and their co-influence in these cells. In whole transcriptome analysis with doxycycline-inducible HEK293 cell line, we observed a strong activating effect of AIRE with 691 genes upregulated and only one downregulated (Study III). Predominant transcription activation was described earlier by comparison mTECs of Aire-deficient and Aire wt mice, with a small fraction of genes down-regulated (Meredith et al., 2015; Sansom et al., 2014). More downregulated genes can be explained by prolonged presence of AIRE in the cells, AIRE effect on mTEC differentiation or other pathways involved in mediating AIRE impact. In contrast, recent report demonstrated AIRE repressive effect at chromatin accessibility and transcriptional levels (Koh et al., 2018). Although we could observe chromatin compaction at sites distant from AIRE target genes (Study III Figure 6A, 6B), we detected undeniable upregulating effect of AIRE on its target genes in HEK293 cells.

We observed the enhancing effect of etoposide on AIRE target genes, increasing the expression level and expanding the repertoire of activated genes (Study III). We could not observe activation-augmenting effect of etoposide when assaying for plasmid reporter transcription in AIRE-positive cells (data not shown), implying that chromatin context is important for etoposide-mediated gene expression changes. Interestingly, we ourselves could not detect the synergistic effect of combining AIRE and etoposide when etoposide was applied before and AIRE expression induced after that (data not shown). This highlights the etoposide importance for AIRE-mediated activation in genomic landscape and suggests that DSB are not important for AIRE recruitment. As AIRE-dependent genes are embedded within relatively closed chromatin, it is possible that etoposide capacity to stabilize DSBs in this context is more influential on transcription rate than in case of plasmid where transcription machinery can process relatively easily as soon as recruited.

Besides TOP2 inhibitor etoposide, we observed similar activating effect of AIRE target genes when TOP1 inhibitor camptothecin was used while topoisomerase inhibitors that block TOP1 or TOP2 catalytic activities entirely without DSB formation did not affect AIRE-induced gene expression activation (Study III). This indicates that stabilization of either SSBs or DSBs or the prolonged presence of TOP1 or TOP2 have positive effect on AIRE-mediated gene activation. Interestingly, TOP1 was shown to be involved in initial complex assembly at superenhancers and enriched at H3K27ac rich distant sites, at the same time TOP2A was accumulated at TSS and involved in subsequent transcriptional events (Bansal et al., 2017). The same study found TOP inhibitors to have repressive effect on genes upregulated by AIRE (Bansal et al., 2017), which was seemingly in conflict with earlier findings of Abramson *et al* and ours (Study III; Abramson et al., 2010). The differences between the studies could stem from the way etoposide was introduced – adding to the cell culture medium or injecting intraperitoneally into the mice for three days (Bansal et al., 2017).

To our knowledge, no whole transcriptome screening was performed with etoposide earlier, although individual targets are reported (Collins et al., 2001; Tammaro et al., 2013). We observed a surprisingly large proportion of up-regulated genes when uninduced HEK293 cells were treated with only etoposide at low dose (2μM) (2452 genes activated (55,6 %) and 1961 downregulated (44,4 %)), unlikely to be attributed to secondary effects, such as stress response (Study III). Considering that etoposide induces DSB formation, such large number of activated genes was not expected. In RNA-seq experiments, we also detected a considerable set of genes, where alternative splicing was influenced by etoposide. Etoposide influence on alternative splicing was demonstrated in proteomic analyses, where etoposide treatment of osteosarcoma cells evoked modifications of proteins involved in RNA metabolism (Beli et al., 2012), including dephosphorylation of SRSF1 (Montecucco and Biamonti, 2013), leading to alternative splicing of genes involved in the choice between pro- and anti-apoptotic pathways (Montecucco et al., 2015). Assessing thoroughly transcriptional and co-transcriptional effects of topoisomerase inhibitors could bring new knowledge and attenuate their usage in cancer therapies.

L28P mutation in HSR/CARD domain disrupts binding to plasmid DNA (Study I) as well as interaction with TOP2A (Study III). MNase treatment of lysates before IP does not exclude the participation of DNA or other proteins mediating AIRE-TOP2A interaction, as HSR/CARD domain is important for oligomerization and copious contacts with other proteins and DNA (Study I; Pitkänen et al., 2000). The potential order of complex formation between AIRE, TOP2A and DNA is not defined, and initial event is not determined. However, in a previous report, the treatment with DNA intercalator ethidium bromide disrupted complex formation, suggesting that complexes are not pre-formed but depend on DNA presence (Bansal et al., 2017). In our experiments, we observed that MNase treatment, but not ethidium bromide, weakened AIRE─DNA-PK complex formation (Study II).

6.4. Stochastic nature of AIRE-mediated activation

One of the first analysis of target gene expression at the single cell level revealed a contrasting pattern of casein locus regulation in mTECs and mammary gland epithelial cells (Derbinski et al., 2008). While tight co-expression of milk proteins was observed in mammary gland cells, mTECs did not comply with the same rules and rarely co-expressed the same set of genes in each individual cell. This observation emphasized again the difference between stochastic nature of promiscuous gene expression in mTECs and cell lineage-specific gene control. The tendency to accumulate in clusters could reflect the maturation program of mTECs, as gene clustering is also distinguishable in Aire-negative mTEChi cells, although the number of clusters and genes per cluster are reduced (Derbinski et al., 2005). We could observe clustering of AIRE target genes in HEK293 cells as well (Study III), indicating that the feature is not exclusively inherent for mTECs.

The propensity to form clusters could refer to AIRE ability to delineate transcriptional boundaries, occupying superenhancers (Bansal et al., 2017) or interacting with CTCF and other insulators. The formed cluster increases the probability of expression and predisposes genes for transcription induction; however, the final productive expression may be achieved by additional mechanisms with random selectivity.

Single-cell RNA-seq identified microclusters co-occuring in very few cells, suggesting these these cells developed from the common progenitor sharing same epigenetic signatures (Meredith et al., 2015). The identified clusters contained activated genes localized on different chromosomes, suggesting that inter-chromosomal contacts in mTECs are more prevalent, than *cis*-interactions (Meredith et al., 2015). Spatial organization could play critical role in gene expression (Dixon et al., 2016; Vermunt et al., 2019). AIRE forms distinct speckles inside the nuclei (Ramsey et al., 2002) and is shown to associate tightly with nuclear matrix resulting in non-random localization of AIRE-responsive genes (Tao et al., 2006). There are several possibilities to create various territories inside the nucleus, starting from every chromosome occupying certain territory (Cremer and Cremer, 2010) ending with diverse *cis*-regulatory elements including enhancers, silencers, insulators, tethering elements that provide boundaries and scaffold for spatial architecture (Spitz and Furlong, 2012).

Despite differences in AIRE actions on plasmid target and in genomic context, the *in vitro* studies with transfected AIRE in different cell types have provided invaluable insights into AIRE functioning mechanisms. Although some results are conflicting and not all of them are readily translatable to *in vivo* situations, the pieces of AIRE transcriptional puzzle are coming together. Based on current knowledge and our three studies, we propose a following model for AIRE action as one option. First, when AIRE expression is activated, it localizes into the nucleus and binds both at active and inactive genes. Direct binding of extended HSR/CARD+ domain to DNA enables interaction with genes with less packed chromatin, and repressive chromatin marks consolidate binding via PHD1 to closed chromatin. AIRE binding is accompanied by recruitment of TOP proteins that cleave DNA and by activation of DNA-PK and other DNA repair proteins. As a result, AIRE gets phosphorylated, chromatin more relaxed, and interaction with P-TEFb and BRD4 established via superenhancer contact, culminating in Pol II pause-release and productive transcription. Many other interacting proteins support these chromatin and spatial changes. While functionally the importance of AIRE binding to inactive chromatin is more obvious, the rationale of interaction with active genes has yet to be explored. One option may be to increase the number of splice variants, as we detected differential exon usage in the set of genes with higher initial expression and different from DE genes. The group of proteins interacting with AIRE and AIRE domains' role could differ depending whether AIRE is bound near highly expressed gene or in more silent genomic context. These surroundings and partners may dictate if AIRE behaves as transcriptional activator or rather modifies alternative splicing.

Advanced technological methods, such as single cell chromatin immunoprecipitation (Clark et al., 2016), proximity RNA-seq (Morf et al., 2017), high-throughput chromosome conformation capture enabling detecting interchromosomal contacts (4C and beyond) (Simonis et al., 2006; Zhao et al., 2006) could offer new insights on AIRE function inside the nucleus. Various reports have indicated that AIRE takes advantage of numerous partner proteins in chromatin recognition, protein interaction, DNA binding and topological organization to achieve stochastic expression of genes in individual cells enabling promiscuous expression in the population of mTECs. Delineating AIRE function promises deeper apprehension not only on the fundamental processes of the immune system but also on transcriptional mechanisms in the eukaryotic cell.

7. CONCLUSIONS

1) AIRE can induce the expression from plasmid reporters in the presence or absence of promoter, intron or polyadenylation sequences, indicating that AIRE influence on gene expression occurs at transcriptional level and is not restricted to particular regulatory elements.
2) N-terminal part of AIRE containing extended HSR/CARD+ domain (aa 1–143) has capacity to bind plasmid and genomic DNA in ChIP experiments. Mutations in aa R113 and K114, originally attributed to NLS, disrupt the binding to DNA, while transport to the nucleus is preserved, suggesting that these aa together with HSR/CARD domain are responsible for AIRE binding to DNA. Extended HSR/CARD+ domain binds DNA in EMSA without the participation of other proteins, proposing that the binding to DNA is direct. ChIP experiments demonstrate strong AIRE binding to plasmid DNA devoid of nucleosomes. The capacity of extended HSR/ CARD+ to bind strongly to plasmid and genomic DNA with higher enrichment near promoter areas of highly expressed genes suggest that extended HSR/CARD+ is another important determinant for locating AIRE to the chromatin besides the previously reported interaction of AIRE PHD1 with nonmethylated histone H3.
3) AIRE interacts with proteins involved in DNA repair, including DNA-PK, which demonstrates capacity to mediate phosphorylation of aa T68 and S156, important for transcriptional activity of AIRE. The topoisomerase inhibitors, including etoposide, that trap topoisomerase after DNA cleavage and prolong DNA breaks, enhance AIRE-mediated transcription and expand the repertoire of AIRE target genes. At the same time, the inhibitors blocking topoisomerase enzymatic activity before the break formation did not exhibit this behavior, indicating that DNA breaks contribute to AIRE-mediated transcriptional activation. The genes upregulated in the presence of AIRE together with etoposide have features characteristic for AIRE target genes, including low initial expression, tissue-specificity and accumulation into clusters. Accordingly, topoisomerase 2A silencing by shRNA decreases activation of AIRE target genes both with and without etoposide indicating that the presence of DNA breaks or chromatin-free DNA has positive effect on AIRE ability to activate the transcription. The L28P mutation in the HSR/CARD domain abolishes the effect of AIRE-mediated activation in the presence and absence of etoposide, indicating the importance of HSR/CARD domain integrity for oligomerization, AIRE DNA binding activity, and recruitment of other proteins, including topoisomerases.

8. REFERENCES

Abbas, A.K., Janeway, C.A., 2000. Immunology: Improving on Nature in the Twenty-First Century. Cell 100, 129–138.

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

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

Abramson, J., Husebye, E.S., 2016. Autoimmune regulator and self-tolerance – molecular and clinical aspects. Immunol. Rev. 271, 127–140.

Adelman, K., Lis, J.T., 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat. Rev. Genet. 13, 720–31.

Ahonen, P., Myllärniemi, S., Sipilä, I., Perheentupa, J., 1990. Clinical Variation of Autoimmune Polyendocrinopathy–Candidiasis–Ectodermal Dystrophy (APECED) in a Series of 68 Patients. N. Engl. J. Med. 322, 1829–1836.

Akoulitchev, S., Mäkelä, T.P., Weinberg, R.A., Reinberg, D., 1995. Requirement for TFIIH kinase activity in transcription by RNA polymerase II. Nature 377, 557–560.

Almada, A.E., Wu, X., Kriz, A.J., Burge, C.B., Sharp, P.A., 2013. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. Nature 499, 360–363.

Anders, S., Reyes, A., Huber, W., 2012. Detecting differential usage of exons from RNA-seq data. Genome Res. 22, 2008–2017.

Anderson, M.S., 2002. Projection of an Immunological Self Shadow Within the Thymus by the Aire Protein. Science 298, 1395–1401.

Anderson, M.S., Venanzi, E.S., Chen, Z., Berzins, S.P., Benoist, C., Mathis, D., 2005. The cellular mechanism of Aire control of T cell tolerance. Immunity 23, 227–239.

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

Appiah, K.A., Li, X., Ptacek, T.S., Brown, E.E., Edberg, J.C., Kandiyoti, D., Lovejoy, C.O., Montoya, M.J., Madabhushi, R., Gao, F., Pfenning, A.R., Pan, L., Yamakawa, S., Seo, J., Rueda, R., Phan, T.X., Yamakawa, H., Pao, P.-C., Stott, R.T., Gjoneska, E., Nott, A., Cho, S., Kellis, M., Tsai, L.-H., Goodnow, C.C., Fujimura, J.H., Duster, T., Rajagopalan, R., Nelson, N.C., Bevan, V., Learmonth, M., Blanco, P., Shlumukova, M., Sargent, C. a, Jobling, M. a, Affara, N., Hurles, M.E., 2007. Activity-Induced DNA Breaks Govern the Expression of Neuronal Early-Response Genes. Cell 43, 118–125.

Aspinall, R., Andrew, D., 2000. Thymic Involution in Aging. J. Clin. Immunol. 20, 250–256.

Bakhru, P., Zhu, M.-L., Wang, H.-H., Hong, L.K., Khan, I., Mouchess, M., Gulati, A.S., Starmer, J., Hou, Y., Sailer, D., Lee, S., Zhao, F., Kirkwood, J.M., Moschos, S., Fong, L., Anderson, M.S., Su, M.A., 2017. Combination central tolerance and peripheral checkpoint blockade unleashes antimelanoma immunity. JCI Insight 2, e93265.

Bansal, K., Yoshida, H., Benoist, C., Mathis, D., 2017. The transcriptional regulator Aire binds to and activates super-enhancers. Nat. Immunol. 18, 263–273.

Baranello, L., Kouzine, F., Levens, D., 2013. DNA Topoisomerases. Transcription 4, 232–237.

Beli, P., Lukashchuk, N., Wagner, S.A., Weinert, B.T., Olsen, J.V., Baskcomb, L., Mann, M., Jackson, S.P., Choudhary, C., 2012. Proteomic Investigations Reveal a Role for

RNA Processing Factor THRAP3 in the DNA Damage Response. Mol. Cell 46, 212–225.
Bentley, D.L., 2014. Coupling mRNA processing with transcription in time and space. Nat. Rev. Genet. 15, 163–175.
Bichele, R., Kisand, K., Peterson, P., Laan, M., 2016. TNF superfamily members play distinct roles in shaping the thymic stromal microenvironment. Mol. Immunol. 72, 92–102.
Billingham, R.E., Brent, L., Medawar, P.B., 1953. Actively acquired tolerance of foreign cells. Nature 172, 603–6.
Björses, P., Halonen, M., Palvimo, J.J., Kolmer, M., Aaltonen, J., Ellonen, P., Perheentupa, J., Ulmanen, I., Peltonen, L., 2000. Mutations in the AIRE gene: effects on subcellular location and transactivation function of the autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy protein. Am. J. Hum. Genet. 66, 378–392.
Björses, P., Pelto-Huikko, M., Kaukonen, J., Aaltonen, J., Peltonen, L., Ulmanen, I., 1999. Localization of the APECED protein in distinct nuclear structures. Hum Mol Genet 8, 259–266.
Blake, W.J., KÆrn, M., Cantor, C.R., Collins, J.J., 2003. Noise in eukaryotic gene expression. Nature 422, 633–637.
Blechschmidt, K., Schweiger, M., Wertz, K., Poulson, R., Christensen, H.M., Rosenthal, A., Lehrach, H., Yaspo, M.L., 1999. The mouse Aire gene: comparative genomic sequencing, gene organization, and expression. Genome Res. 9, 158–66.
Blier, P.R., Griffith, A.J., Craft, J., Hardin, J.A., 1993. Binding of Ku protein to DNA. Measurement of affinity for ends and demonstration of binding to nicks. J. Biol. Chem. 268, 7594–7601.
Blunt, T., Finnie, N.J., Taccioli, G.E., Smith, G.C.M., Demengeot, J., Gottlieb, T.M., Mizuta, R., Varghese, A.J., Alt, F.W., Jeggo, P.A., Jackson, S.P., 1995. Defective DNA-Dependent Protein Kinase Activity Is Linked to V(D)J Recombination and DNA Repair Defects Associated with the Murine scid Mutation. Cell 80, 813–823.
Bonilla, F.A., Oettgen, H.C., 2010. Adaptive immunity. J. Allergy Clin. Immunol. 125, S33–S40.
Bornstein, C., Nevo, S., Giladi, A., Kadouri, N., Pouzolles, M., Gerbe, F., David, E., Machado, A., Chuprin, A., Tóth, B., Goldberg, O., Itzkovitz, S., Taylor, N., Jay, P., Zimmermann, V.S., Abramson, J., Amit, I., 2018. Single-cell mapping of the thymic stroma identifies IL-25-producing tuft epithelial cells. Nature 559, 622–626.
Boulikas, T., 1994. Putative nuclear localization signals (NLS) in protein transcription factors. J. Cell. Biochem. 55, 32–58.
Braun, R.E., 2001. Packaging paternal chromosomes with protamine. Nat. Genet. 28, 10–12.
Brennecke, P., Reyes, A., Pinto, S., Rattay, K., Nguyen, M., Küchler, R., Huber, W., Kyewski, B., Steinmetz, L.M., 2015. Single-cell transcriptome analysis reveals coordinated ectopic gene-expression patterns in medullary thymic epithelial cells. Nat. Immunol. 16, 933–941.
Brès, V., Yoh, S.M., Jones, K.A., 2008. The multi-tasking P-TEFb complex. Curr. Opin. Cell Biol. 20, 334–340.
Bruserud, Ø., Oftedal, B.E., Wolff, A.B., Husebye, E.S., 2016. AIRE-mutations and autoimmune disease. Curr. Opin. Immunol. 43, 8–15.
Buchman, A.R., Berg, P., 1988. Comparison of intron-dependent and intron-independent gene expression. Mol. Cell. Biol. 8, 4395–405.

Bunch, H., 2016. Role of genome guardian proteins in transcriptional elongation. FEBS Lett. 590, 1064–1075.
Bunch, H., Zheng, X., Burkholder, A., Dillon, S.T., Motola, S., Birrane, G., Ebmeier, C.C., Levine, S., Fargo, D., Hu, G., Taatjes, D.J., Calderwood, S.K., 2014. TRIM28 regulates RNA polymerase II promoter-proximal pausing and pause release. Nat. Struct. Mol. Biol. 21, 876–83.
Buratowski, S., 2009. Progression through the RNA Polymerase II CTD Cycle. Mol. Cell 36, 541–546.
Burma, S., Chen, D.J., 2004. Role of DNA–PK in the cellular response to DNA double-strand breaks. DNA Repair 3, 909–918.
Buschbeck, M., Hake, S.B., 2017. Variants of core histones and their roles in cell fate decisions, development and cancer. Nat. Rev. Mol. Cell Biol. 18, 299–314.
Cairns, B.R., 2009. The logic of chromatin architecture and remodelling at promoters. Nature 461, 193–198.
Campos, E.I., Reinberg, D., 2009. Histones: Annotating Chromatin. Annu. Rev. Genet. 43, 559–599.
Capalbo, D., De Martino, L., Giardino, G., Di Mase, R., Di Donato, I., Parenti, G., Vajro, P., Pignata, C., Salerno, M., 2012. Autoimmune polyendocrinopathy candidiasis ectodermal dystrophy: Insights into genotype-phenotype correlation. Int. J. Endocrinol. 2012, 1–9.
Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C.A.M., Taylor, M.S., Engström, P.G., Frith, M.C., Forrest, A.R.R., Alkema, W.B., Tan, S.L., Plessy, C., Kodzius, R., Ravasi, T., Kasukawa, T., Fukuda, S., Kanamori-Katayama, M., Kitazume, Y., Kawaji, H., Kai, C., Nakamura, M., Konno, H., Nakano, K., Mottagui-Tabar, S., Arner, P., Chesi, A., Gustincich, S., Persichetti, F., Suzuki, H., Grimmond, S.M., Wells, C.A., Orlando, V., Wahlestedt, C., Liu, E.T., Harbers, M., Kawai, J., Bajic, V.B., Hume, D.A., Hayashizaki, Y., 2006. Genome-wide analysis of mammalian promoter architecture and evolution. Nat. Genet. 38, 626–635.
Cetani, F., Barbesino, G., Borsari, S., Pardi, E., Cianferotti, L., Pinchera, A., 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.
Chen, F., Gao, X., Shilatifard, A., 2015. Stably paused genes revealed through inhibition of transcription initiation by the TFIIH inhibitor triptolide. Genes Dev. 29, 39–47.
Chen, F.X., Smith, E.R., Shilatifard, A., 2018. Born to run: Control of transcription elongation by RNA polymerase II. Nat. Rev. Mol. Cell Biol. 19, 464–478.
Chen, S.H., Chan, N.-L., Hsieh, T., 2013. New Mechanistic and Functional Insights into DNA Topoisomerases. Annu. Rev. Biochem. 82, 139–170.
Chuprin, A., Avin, A., Goldfarb, Y., Herzig, Y., Levi, B., Jacob, A., Sela, A., Katz, S., Grossman, M., Guyon, C., Rathaus, M., Cohen, H.Y., Sagi, I., Giraud, M., McBurney, M.W., Husebye, E.S., Abramson, J., 2015. The deacetylase Sirt1 is an essential regulator of Aire-mediated induction of central immunological tolerance. Nat. Immunol. 16, 737–745.
Clark, S.J., Lee, H.J., Smallwood, S.A., Kelsey, G., Reik, W., 2016. Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity. Genome Biol. 17, 72.

Cockerill, P.N., Garrard, W.T., 1986. Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. Cell 44, 273–282.
Cokol, M., Nair, R., Rost, B., 2000. Finding nuclear localization signals. EMBO Rep. 1, 411–5.
Collins, I., Weber, A., Levens, D., 2001. Transcriptional consequences of topoisomerase inhibition. Mol. Cell. Biol. 21, 8437–51.
Core, L.J., Martins, A.L., Danko, C.G., Waters, C.T., Siepel, A., Lis, J.T., 2014. Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. Nat. Genet. 46, 1311–20.
Core, L.J., Waterfall, J.J., Lis, J.T., 2008. Nascent RNA Sequencing Reveals Widespread Pausing and Divergent Initiation at Human Promoters. Science 322, 1845–1848.
Cremer, T., Cremer, M., 2010. Chromosome territories. Cold Spring Harb. Perspect. Biol. 2, a003889.
Crick, F., 1970. Central Dogma of Molecular Biology. Nature 227, 561–563.
Cusanovich, D.A., Pavlovic, B., Pritchard, J.K., Gilad, Y., 2014. The Functional Consequences of Variation in Transcription Factor Binding. PLoS Genet. 10, e1004226.
Dame, R.T., 2005. The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. Mol. Microbiol. 56, 858–870.
Danso-Abeam, D., Staats, K.A., Franckaert, D., Van Den Bosch, L., Liston, A., Gray, D.H.D., Dooley, J., 2013. Aire mediates thymic expression and tolerance of pancreatic antigens via an unconventional transcriptional mechanism. Eur. J. Immunol. 43, 75–84.
Dávila López, M., Samuelsson, T., 2008. Early evolution of histone mRNA 3' end processing. RNA 14, 1–10.
Davis, C.A., Ares, M., 2006. Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. 103, 3262–3267.
de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D., Kornblihtt, A.R., 2003. A Slow RNA Polymerase II Affects Alternative Splicing In Vivo. Mol. Cell 12, 525–532.
Derbinski, J., Gäbler, J., Brors, B., Tierling, S., Jonnakuty, S., Hergenhahn, M., Peltonen, L., Walter, J., Kyewski, B., 2005. Promiscuous gene expression in thymic epithelial cells is regulated at multiple levels. J Exp Med 202, 33–45.
Derbinski, J., Pinto, S., Rösch, S., Hexel, K., Kyewski, B., 2008. Promiscuous gene expression patterns in single medullary thymic epithelial cells argue for a stochastic mechanism. Proc Natl Acad Sci U S A 105, 657–662.
Derbinski, J., Schulte, A., Kyewski, B., Klein, L., 2001. Promiscuous gene expression in medullary thymic epithelial cells mirrors the peripheral self. Nat. Immunol. 2, 1032–9.
DeVoss, J., Hou, Y., Johannes, K., Lu, W., Liou, G.I., Rinn, J., Chang, H., Caspi, R.R., Caspi, R., Fong, L., Anderson, M.S., 2006. Spontaneous autoimmunity prevented by thymic expression of a single self-antigen. J. Exp. Med. 203, 2727–35.
Deweese, J.E., Osheroff, N., 2009. The DNA cleavage reaction of topoisomerase II: Wolf in sheep's clothing. Nucleic Acids Res. 37, 738–748.
Dey, A., Chitsaz, F., Abbasi, A., Misteli, T., Ozato, K., 2003. The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. Proc. Natl. Acad. Sci. 100, 8758–8763.

Dixon, J.R., Gorkin, D.U., Ren, B., 2016. Chromatin Domains: The Unit of Chromosome Organization. Mol. Cell 62, 668–680.
Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.
Dvir, A., Peterson, S.R., Knuth, M.W., Lu, H., Dynan, W.S., 1992. Ku autoantigen is the regulatory component of a template-associated protein kinase that phosphorylates RNA polymerase II. Proc. Natl. Acad. Sci. 89, 11920–11924.
Egloff, S., 2012. Role of Ser7 phosphorylation of the CTD during transcription of snRNA genes. RNA Biol. 9, 1033–1038.
Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S., 2002. Stochastic gene expression in a single cell. Science 297, 1183–6.
Fagerberg, L., Hallström, B.M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., Habuka, M., Tahmasebpoor, S., Danielsson, A., Edlund, K., Asplund, A., Sjöstedt, E., Lundberg, E., Szigyarto, C.A.-K., Skogs, M., Takanen, J.O., Berling, H., Tegel, H., Mulder, J., Nilsson, P., Schwenk, J.M., Lindskog, C., Danielsson, F., Mardinoglu, A., Sivertsson, A., von Feilitzen, K., Forsberg, M., Zwahlen, M., Olsson, I., Navani, S., Huss, M., Nielsen, J., Ponten, F., Uhlén, M., 2014. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol. Cell. Proteomics 13, 397–406.
Fell, V.L., Schild-Poulter, C., 2015. The Ku heterodimer: Function in DNA repair and beyond. Mutat. Res. Mutat. Res. 763, 15–29.
Ferguson, B.J., Alexander, C., Rossi, S.W., Liiv, I., Rebane, A., Worth, C.L., Wong, J., Laan, M., rt Peterson, P., Jenkinson, E.J., Anderson, G., Scott, H.S., Cooke, A., Rich, T., Hall, E., Peterson, P., Jenkinson, E.J., Anderson, G., Scott, H.S., Cooke, A., Rich, T., 2008. AIRE's CARD revealed, a new structure for central tolerance provokes transcriptional plasticity. J. Biol. Chem. 283, 1723–31.
Fergusson, J.R., Morgan, M.D., Bruchard, M., Huitema, L., Heesters, B.A., van Unen, V., van Hamburg, J.P., van der Wel, N.N., Picavet, D., Koning, F., Tas, S.W., Anderson, M.S., Marioni, J.C., Holländer, G.A., Spits, H., 2019. Maturing Human CD127+ CCR7+ PDL1+ Dendritic Cells Express AIRE in the Absence of Tissue Restricted Antigens. Front. Immunol. 9.
Ferre, E.M.N., Rose, S.R., Rosenzweig, S.D., Burbelo, P.D., Romito, K.R., Niemela, J.E., Rosen, L.B., Break, T.J., Gu, W., Hunsberger, S., Browne, S.K., Hsu, A.P., Rampertaap, S., Swamydas, M., Collar, A.L., Kong, H.H., Lee, C.-C.R., Chascsa, D., Simcox, T., Pham, A., Bondici, A., Natarajan, M., Monsale, J., Kleiner, D.E., Quezado, M., Alevizos, I., Moutsopoulos, N.M., Yockey, L., Frein, C., Soldatos, A., Calvo, K.R., Adjemian, J., Similuk, M.N., Lang, D.M., Stone, K.D., Uzel, G., Kopp, J.B., Bishop, R.J., Holland, S.M., Olivier, K.N., Fleisher, T.A., Heller, T., Winer, K.K., Lionakis, M.S., 2016. Redefined clinical features and diagnostic criteria in autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy. JCI insight 1, e88782.
Fierabracci, A., 2016. Type 1 Diabetes in Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy Syndrome (APECED): A "Rare" Manifestation in a "Rare" Disease. Int. J. Mol. Sci. 17, 1106.
Frazer, K.A., 2012. Decoding the human genome. Genome Res. 22, 1599–1601.
Fuda, N.J., Ardehali, M.B., Lis, J.T., 2009. Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature 461, 186–192.

Fukaya, T., Lim, B., Levine, M., 2016. Enhancer Control of Transcriptional Bursting. Cell 166, 358–368.
Fuller, P.J., Verity, K., 1989. Somatostatin gene expression in the thymus gland. J. Immunol. 143, 1015–7.
Gaetani, M., Matafora, V., Saare, M., Spiliotopoulos, D., Mollica, L., Quilici, G., Chignola, F., Mannella, V., Zucchelli, C., Peterson, P., Bachi, A., Musco, G., 2012. AIRE-PHD fingers are structural hubs to maintain the integrity of chromatin-associated interactome. Nucleic Acids Res 40, 11756–11768.
Gallie, D.R., Lewis, N.J., Marzluff, W.F., 1996. The histone 3′-terminal stem-loop is necessary for translation in Chinese hamster ovary cells. Nucleic Acids Res. 24, 1954–1962.
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., Anderson, M.S., 2008. Deletional tolerance mediated by extrathymic Aire-expressing cells. Science 321, 843–7.
Gardner, J.M., Metzger, T.C., McMahon, E.J., Au-Yeung, B.B., Krawisz, A.K., Lu, W., Price, J.D., Johannes, K.P., Satpathy, A.T., Murphy, K.M., Tarbell, K.V., Weiss, A., Anderson, M.S., 2013. Extrathymic Aire-Expressing Cells Are a Distinct Bone Marrow-Derived Population that Induce Functional Inactivation of CD4+ T Cells. Immunity 39, 560–572.
Garneau, N.L., Wilusz, J., Wilusz, C.J., 2007. The highways and byways of mRNA decay. Nat. Rev. Mol. Cell Biol. 8, 113–126.
Gavanescu, I., Kessler, B., Ploegh, H., Benoist, C., Mathis, D., 2007. Loss of Aire-dependent thymic expression of a peripheral tissue antigen renders it a target of autoimmunity. Proc. Natl. Acad. Sci. 104, 4583–4587.
Geenen, V., Legros, J.J., Franchimont, P., Baudrihaye, M., Defresne, M.P., Boniver, J., 1986. The neuroendocrine thymus: Coexistence of oxytocin and neurophysin in the human thymus. Science 232, 508–511.
Geisler, S., Coller, J., 2013. RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. Nat. Rev. Mol. Cell Biol. 14, 699–712.
Ghavi-Helm, Y., Klein, F.A., Pakozdi, T., Ciglar, L., Noordermeer, D., Huber, W., Furlong, E.E.M., 2014. Enhancer loops appear stable during development and are associated with paused polymerase. Nature 512, 96–100.
Gillard, G.O., Farr, A.G., 2005. Contrasting models of promiscuous gene expression by thymic epithelium. J. Exp. Med. 202, 15–19.
Gilmour, D.S., Lis, J.T., 1986. RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in Drosophila melanogaster cells. Mol. Cell. Biol. 6, 3984–9.
Giménez-Barcons, M., Casteràs, A., Armengol, M. del P., Porta, E., Correa, P.A., Marín, A., Pujol-Borrell, R., Colobran, R., 2014. Autoimmune Predisposition in Down Syndrome May Result from a Partial Central Tolerance Failure due to Insufficient Intrathymic Expression of AIRE and Peripheral Antigens. J. Immunol. 193, 3872–3879.
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., Benoist, C., 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., Benoist, C., 2012. Aire unleashes stalled RNA polymerase to induce ectopic gene expression in thymic epithelial cells. Proc. Natl. Acad. Sci. 109, 535–540.

Gómez-Herreros, F., Zagnoli-Vieira, G., Ntai, I., Martínez-Macías, M.I., Anderson, R.M., Herrero-Ruíz, A., Caldecott, K.W., 2017. TDP2 suppresses chromosomal translocations induced by DNA topoisomerase II during gene transcription. Nat. Commun. 8, 233.

Gotter, J., 2004. Medullary Epithelial Cells of the Human Thymus Express a Highly Diverse Selection of Tissue-specific Genes Colocalized in Chromosomal Clusters. J. Exp. Med. 199, 155–166.

Graveley, B.R., 2001. Alternative splicing: increasing diversity in the proteomic world. Trends Genet. 17, 100–107.

Gray, D., Abramson, J., Benoist, C., Mathis, D., 2007. Proliferative arrest and rapid turnover of thymic epithelial cells expressing Aire. J Exp Med 204, 2521–2528.

Gray, D.H.D., Tull, D., Ueno, T., Seach, N., Classon, B.J., Chidgey, A., McConville, M.J., Boyd, R.L., 2007. A unique thymic fibroblast population revealed by the monoclonal antibody MTS-15. J. Immunol. 178, 4956–65.

Gu, B., Lambert, J.-P., Cockburn, K., Gingras, A.-C., Rossant, J., 2017. AIRE is a critical spindle-associated protein in embryonic stem cells. Elife 6, e28131.

Guerau-de-Arellano, M., Martinic, M., Benoist, C., Mathis, D., 2009. Neonatal tolerance revisited: a perinatal window for Aire control of autoimmunity. J. Exp. Med. 206, 1245–1252.

Guerau-de-Arellano, M., Mathis, D., Benoist, C., 2008. Transcriptional impact of Aire varies with cell type. Proc. Natl. Acad. Sci. 105, 14011–14016.

Guzmán, E., Lis, J.T., 1999. Transcription factor TFIIH is required for promoter melting in vivo. Mol. Cell. Biol. 19, 5652–8.

Haberle, V., Stark, A., 2018. Eukaryotic core promoters and the functional basis of transcription initiation. Nat. Rev. Mol. Cell Biol. 1–17.

Haljasorg, U., Bichele, R., Saare, M., Guha, M., Maslovskaja, J., Kõnd, K., Remm, A., Pihlap, M., Tomson, L., Kisand, K., Laan, M., Peterson, P., 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., Eskelin, P., 2004. APECED-causing mutations in AIRE reveal the functional domains of the protein. Hum Mutat 23, 245–257.

Hamazaki, Y., Fujita, H., Kobayashi, T., Choi, Y., Scott, H.S., Matsumoto, M., Minato, N., 2007. Medullary thymic epithelial cells expressing Aire represent a unique lineage derived from cells expressing claudin. Nat. Immunol. 8, 304–311.

Hande, K.R., 1998. Etoposide: Four decades of development of a topoisomerase II inhibitor. Eur. J. Cancer 34, 1514–1521.

Handel, A.E., Shikama-Dorn, N., Zhanybekova, S., Maio, S., Graedel, A.N., Zuklys, S., Ponting, C.P., Holländer, G.A., 2018. Comprehensively Profiling the Chromatin Architecture of Tissue Restricted Antigen Expression in Thymic Epithelial Cells Over Development. Front. Immunol. 9, 2120.

Harlen, K.M., Churchman, L.S., 2017. The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. Nat. Rev. Mol. Cell Biol. 18, 263–273.

Heath, W.R., Allison, J., Hoffmann, M.W., Schönrich, G., Hämmerling, G., Arnold, B., Miller, J.F., 1992. Autoimmune diabetes as a consequence of locally produced interleukin-2. Nature 359, 547–9.

Heino, M., Peterson, P., Kudoh, J., Nagamine, K., Lagerstedt, A., Ovod, V., Ranki, A., Rantala, I., Nieminen, M., Tuukkanen, J., Scott, H.S., Antonarakis, S.E., Shimizu, N.,

Krohn, K., 1999. Autoimmune Regulator Is Expressed in the Cells Regulating Immune Tolerance in Thymus Medulla. Biochem. Biophys. Res. Commun. 257, 821–825.
Helleday, T., Petermann, E., Lundin, C., Hodgson, B., Sharma, R.A., 2008. DNA repair pathways as targets for cancer therapy. Nat. Rev. Cancer 8, 193–204.
Henriques, T., Gilchrist, D.A., Nechaev, S., Bern, M., Muse, G.W., Burkholder, A., Fargo, D.C., Adelman, K., 2013. Stable Pausing by RNA Polymerase II Provides an Opportunity to Target and Integrate Regulatory Signals. Mol. Cell 52, 517–528.
Henriques, T., Scruggs, B.S., Inouye, M.O., Muse, G.W., Williams, L.H., Burkholder, A.B., Lavender, C.A., Fargo, D.C., Adelman, K., 2018. Widespread transcriptional pausing and elongation control at enhancers. Genes Dev. 32, 26–41.
Herzel, L., Ottoz, D.S.M., Alpert, T., Neugebauer, K.M., 2017. Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. Nat. Rev. Mol. Cell Biol. 18, 637–650.
Herzig, Y., Nevo, S., Bornstein, C., Brezis, M.R., Ben-Hur, S., Shkedy, A., Eisenberg-Bord, M., Levi, B., Delacher, M., Goldfarb, Y., David, E., Weinberger, L., Viukov, S., Ben-Dor, S., Giraud, M., Hanna, J.H., Breiling, A., Lyko, F., Amit, I., Feuerer, M., Abramson, J., 2016. Transcriptional programs that control expression of the autoimmune regulator gene Aire. Nat. Immunol. 18, 1–14.
Hödl, M., Basler, K., 2012. Transcription in the Absence of Histone H3.2 and H3K4 Methylation. Curr. Biol. 22, 2253–2257.
Hogquist, K.A., Baldwin, T.A., Jameson, S.C., 2005. Central tolerance: learning self-control in the thymus. Nat. Rev. Immunol. 5, 772–782.
Huang, F., Shao, W., Fujinaga, K., Peterlin, B.M., 2018. Bromodomain-containing protein 4-independent transcriptional activation by autoimmune regulator (AIRE) and NF-κB. J. Biol. Chem. 293, 4993–5004.
Hubert, F.-X., Kinkel, S.A., Davey, G.M., Phipson, B., Mueller, S.N., Liston, A., Proietto, A.I., Cannon, P.Z.F., Forehan, S., Smyth, G.K., Wu, L., Goodnow, C.C., Carbone, F.R., Scott, H.S., Heath, W.R., 2011. Aire regulates the transfer of antigen from mTECs to dendritic cells for induction of thymic tolerance. Blood 118, 2462–72.
Hubert, F.-X.X., Kinkel, S.A., Crewther, P.E., Cannon, P.Z.F., Webster, K.E., Link, M., Uibo, R., O'Bryan, M.K., Meager, A., Forehan, S.P., Smyth, G.K., Mittaz, L., Anto-narakis, S.E., Peterson, P., Heath, W.R., Scott, H.S., 2009. Aire-deficient C57BL/6 mice mimicking the common human 13-base pair deletion mutation present with only a mild autoimmune phenotype. J Immunol 182, 3902–3918.
Husebye, E.S., Perheentupa, J., Rautemaa, R., Kämpe, O., 2009. Clinical manifestations and management of patients with autoimmune polyendocrine syndrome type I. J. Intern. Med. 265, 514–29.
Ilmarinen, T., Kangas, H., Kytömaa, T., Eskelin, P., Saharinen, J., Seeler, J.-S., Tanhuan-pää, K., Chan, F.Y.-L., Slattery, R.M., Alakurtti, K., Palvimo, J.J., Ulmanen, I., 2008. Functional interaction of AIRE with PIAS1 in transcriptional regulation. Mol. Immunol. 45, 1847–1862.
Ilmarinen, T., Melén, K., Kangas, H., Julkunen, I., Ulmanen, I., Eskelin, P., 2006. The monopartite nuclear localization signal of autoimmune regulator mediates its nuclear import and interaction with multiple importin alpha molecules. FEBS J 273, 315–324.
Iyengar, S., Farnham, P.J., 2011. KAP1 protein: an enigmatic master regulator of the genome. J. Biol. Chem. 286, 26267–76.
Jackson, S.P., MacDonald, J.J., Lees-Miller, S., Tjian, R., 1990. GC box binding induces phosphorylation of Sp1 by a DNA-dependent protein kinase. Cell 63, 155–165.

Jacquier, A., 2009. The complex eukaryotic transcriptome: Unexpected pervasive transcription and novel small RNAs. Nat. Rev. Genet. 10, 833–844.
Jamieson, B.D., Douek, D.C., Killian, S., Hultin, L.E., Scripture-Adams, D.D., Giorgi, J. V, Marelli, D., Koup, R.A., Zack, J.A., 1999. Generation of Functional Thymocytes in the Human Adult. Immunity 10, 569–575.
Janeway, C.A., Medzhitov, R., 2002. Innate Immune Recognition. Annu. Rev. Immunol. 20, 197–216.
Jang, M.K., Mochizuki, K., Zhou, M., Jeong, H.-S., Brady, J.N., Ozato, K., 2005. The Bromodomain Protein Brd4 Is a Positive Regulatory Component of P-TEFb and Stimulates RNA Polymerase II-Dependent Transcription. Mol. Cell 19, 523–534.
Ji, X., Zhou, Y., Pandit, S., Huang, J., Li, H., Lin, C.Y., Xiao, R., Burge, C.B., Fu, X.-D., 2013. SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. Cell 153, 855–68.
Jiang, W., Anderson, M.S., Bronson, R., Mathis, D., Benoist, C., 2005. Modifier loci condition autoimmunity provoked by Aire deficiency. J Exp Med 202, 805–815.
Johnnidis, J.B., Venanzi, E.S., Taxman, D.J., Ting, J.P.-Y., Benoist, C.O., Mathis, D.J., 2005. Chromosomal clustering of genes controlled by the aire transcription factor. Proc. Natl. Acad. Sci. 102, 7233–7238.
Ju, B.-G., Lunyak, V. V, Perissi, V., Garcia-Bassets, I., Rose, D.W., Glass, C.K., Rosenfeld, M.G., 2006. A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription. Science 312, 1798–1802.
Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., Taatjes, D.J., Dekker, J., Young, R.A., 2010. Mediator and cohesin connect gene expression and chromatin architecture. Nature 467, 430–435.
Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Duttagupta, R., Willingham, A.T., Stadler, P.F., Hertel, J., Hackermuller, J., Hofacker, I.L., Bell, I., Cheung, E., Drenkow, J., Dumais, E., Patel, S., Helt, G., Ganesh, M., Ghosh, S., Piccolboni, A., Sementchenko, V., Tammana, H., Gingeras, T.R., 2007. RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. Science 316, 1484–1488.
Kawano, H., Nishijima, H., Morimoto, J., Hirota, F., Morita, R., Mouri, Y., Nishioka, Y., Matsumoto, M., 2015. Aire Expression Is Inherent to Most Medullary Thymic Epithelial Cells during Their Differentiation Program. J. Immunol. 195, 5149–5158.
Kawano, S., Miyaji, M., Ichiyasu, S., Tsutsui, K.M.K., Tsutsui, K.M.K., 2010. Regulation of DNA Topoisomerase IIbeta through RNA-dependent association with heterogeneous nuclear ribonucleoprotein U (hnRNP U). J. Biol. Chem. 285, 26451–60.
Keane, P., Ceredig, R., Seoighe, C., 2014. Promiscuous mRNA splicing under the control of AIRE in medullary thymic epithelial cells. Bioinformatics 31, 986–990.
Kellis, M., Wold, B., Snyder, M.P., Bernstein, B.E., Kundaje, A., Marinov, G.K., Ward, L.D., Birney, E., Crawford, G.E., Dekker, J., Dunham, I., Elnitski, L.L., Farnham, P.J., Feingold, E.A., Gerstein, M., Giddings, M.C., Gilbert, D.M., Gingeras, T.R., Green, E.D., Guigo, R., Hubbard, T., Kent, J., Lieb, J.D., Myers, R.M., Pazin, M.J., Ren, B., Stamatoyannopoulos, J.A., Weng, Z., White, K.P., Hardison, R.C., 2014. Defining functional DNA elements in the human genome. Proc. Natl. Acad. Sci. 111, 6131–6138.
Khan, I.S., Mouchess, M.L., Zhu, M.-L., Conley, B., Fasano, K.J., Hou, Y., Fong, L., Su, M.A., Anderson, M.S., 2014. Enhancement of an anti-tumor immune response by transient blockade of central T cell tolerance. J. Exp. Med. 211, 761–8.

Kim, T.-K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., Markenscoff-Papadimitriou, E., Kuhl, D., Bito, H., Worley, P.F., Kreiman, G., Greenberg, M.E., 2010. Widespread transcription at neuronal activity-regulated enhancers. Nature 465, 182–187.

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., Chamberlain, S.J., Philpot, B.D., Zylka, M.J., 2013. Topoisomerases facilitate transcription of long genes linked to autism. Nature 501, 58–62.

Kinner, A., Wu, W., Staudt, C., Iliakis, G., 2008. Gamma-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin. Nucleic Acids Res. 36, 5678–5694.

Kisand, K., Lilic, D., Casanova, J.L., Peterson, P., Meager, A., Willcox, N., 2011. Mucocutaneous candidiasis and autoimmunity against cytokines in APECED and thymoma patients: clinical and pathogenetic implications. Eur J Immunol 41, 1517–1527.

Kisand, K., Peterson, P., 2011. Autoimmune polyendocrinopathy candidiasis ectodermal dystrophy: known and novel aspects of the syndrome. Ann. N. Y. Acad. Sci. 1246, 77–91.

Kisand, K., Peterson, P., Laan, M., 2014. Lymphopenia-Induced Proliferation in Aire-Deficient Mice Helps to Explain Their Autoimmunity and Differences from Human Patients. Front Immunol 5, 51.

Kisand, K., Wolff, A.S.B., Podkrajsek, K.T., Tserel, L., Link, M., Kisand, K. V, Ersvaer, E., Perheentupa, J., Erichsen, M.M., Bratanic, N., Meloni, A., Cetani, F., Perniola, R., Ergun-Longmire, B., Maclaren, N., Krohn, K.J.E., Pura, M., Schalke, B., Ströbel, P., Leite, M.I., Battelino, T., Husebye, E.S., Peterson, P., Willcox, N., Meager, A., 2010. Chronic mucocutaneous candidiasis in APECED or thymoma patients correlates with autoimmunity to Th17-associated cytokines. J. Exp. Med. 207, 299–308.

Klein, L., Hinterberger, M., von Rohrscheidt, J., Aichinger, M., 2011. Autonomous versus dendritic cell-dependent contributions of medullary thymic epithelial cells to central tolerance. Trends Immunol. 32, 188–193.

Klein, L., Klugmann, M., Nave, K.-A., Tuohy, V.K., Kyewski, B., 2000. Shaping of the autoreactive T-cell repertoire by a splice variant of self protein expressed in thymic epithelial cells. Nat. Med. 6, 56–61.

Klein, L., Kyewski, B., Allen, P.M., 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.

Klein, L., Robey, E.A., Hsieh, C.-S., 2019. Central CD4+ T cell tolerance: deletion versus regulatory T cell differentiation. Nat. Rev. Immunol. 19, 7–18.

Kogawa, K., Nagafuchi, S., Katsuta, H., Kudoh, J., Tamiya, S., Sakai, Y., Shimizu, N., Harada, M., 2002. Expression of AIRE gene in peripheral monocyte/dendritic cell lineage. Immunol. Lett. 80, 195–198.

Koh, A.S., Kingston, R.E., Benoist, C., Mathis, D., 2010. Global relevance of Aire binding to hypomethylated lysine-4 of histone-3. Proc Natl Acad Sci U S A 107, 13016–13021.

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., Benoist, C., Mathis, D., 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.

Koh, A.S., Miller, E.L., Buenrostro, J.D., Moskowitz, D.M., Wang, J., Greenleaf, W.J., Chang, H.Y., Crabtree, G.R., 2018. Rapid chromatin repression by Aire provides precise control of immune tolerance. Nat. Immunol. 19, 162–172.

Kont, V., Laan, M., Kisand, K., Merits, A., Scott, H.S., Peterson, P., 2008. Modulation of Aire regulates the expression of tissue-restricted antigens. Mol. Immunol. 45, 25–33.

Kornberg, R.D., 1999. Eukaryotic transcriptional control. Trends Biochem. Sci. 24, M46–M49.

Kornblihtt, A.R., de la Mata, M., Fededa, J.P., Muñoz, M.J., Nogues, G., 2004. Multiple links between transcription and splicing. RNA 10, 1489–1498.

Kouzarides, T., 2007. Chromatin Modifications and Their Function. Cell 128, 693–705.

Kouzine, F., Gupta, A., Baranello, L., Wojtowicz, D., Ben-Aissa, K., Liu, J., Przytycka, T.M., Levens, D., 2013. Transcription-dependent dynamic supercoiling is a short-range genomic force. Nat. Struct. Mol. Biol. 20, 396–403.

Kraus, W.L., 2008. Transcriptional control by PARP-1: chromatin modulation, enhancer-binding, coregulation, and insulation. Curr. Opin. Cell Biol. 20, 294–302.

Kumar, P.G., Laloraya, M., Wang, C.Y., Ruan, Q.G., Davoodi-Semiromi, A., Kao, K.J., She, J.X., 2001. The autoimmune regulator (AIRE) is a DNA-binding protein. J Biol Chem 276, 41357–41364.

Kuroda, N., Mitani, T., Takeda, N., Ishimaru, N., Arakaki, R., Hayashi, Y., Bando, Y., Izumi, K., Takahashi, T., Nomura, T., Sakaguchi, S., Ueno, T., Takahama, Y., Uchida, D., Sun, S., Kajiura, F., Mouri, Y., Han, H., Matsushima, A., Yamada, G., Matsumoto, M., 2005. Development of autoimmunity against transcriptionally unrepressed target antigen in the thymus of Aire-deficient mice. J Immunol 174, 1862–1870.

Kyewski, B., Klein, L., 2006. A Central Role for Central Tolerance. Annu. Rev. Immunol. 24, 571–606.

Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., Weirauch, M.T., 2018. The Human Transcription Factors. Cell 172, 650–665.

Larsen, D.H., Poinsignon, C., Gudjonsson, T., Dinant, C., Payne, M.R., Hari, F.J., Rendtlew Danielsen, J.M., Menard, P., Sand, J.C., Stucki, M., Lukas, C., Bartek, J., Andersen, J.S., Lukas, J., 2010. The chromatin-remodeling factor CHD4 coordinates signaling and repair after DNA damage. J. Cell Biol. 190, 731–40.

Lei, Y., Ripen, A.M., Ishimaru, N., Ohigashi, I., Nagasawa, T., Jeker, L.T., Bösl, M.R., Holländer, G.A., Hayashi, Y., Malefyt, R. de W., Nitta, T., Takahama, Y., 2011. Aire-dependent production of XCL1 mediates medullary accumulation of thymic dendritic cells and contributes to regulatory T cell development. J. Exp. Med. 208, 383–94.

Lemp, N.A., Hiraoka, K., Kasahara, N., Logg, C.R., 2012. Cryptic transcripts from a ubiquitous plasmid origin of replication confound tests for cis-regulatory function. Nucleic Acids Res. 40, 7280–7290.

Leonard, J.D., Gilmore, D.C., Dileepan, T., Nawrocka, W.I., Chao, J.L., Schoenbach, M.H., Jenkins, M.K., Adams, E.J., Savage, P.A., 2017. Identification of Natural Regulatory T Cell Epitopes Reveals Convergence on a Dominant Autoantigen. Immunity 47, 107–117.e8.

Leonard, M.F., 1946. Chronic idiopathic hypoparathyroidism with superimposed Addison's disease in a child. J. Clin. Endocrinol. Metab. 6, 493–506.

Levine, M., Tjian, R., 2003. Transcription regulation and animal diversity. Nature 424, 147–151.

Li, H., Ilin, S., Wang, W., Duncan, E.M., Wysocka, J., Allis, C.D., 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.
Lieber, M.R., Ma, Y., Pannicke, U., Schwarz, K., 2003. Mechanism and regulation of human non-homologous DNA end-joining. Nat. Rev. Mol. Cell Biol. 4, 712–720.
Liiv, I., Haljasorg, U., Kisand, K., Maslovskaja, J., Laan, M., Peterson, P., 2012. AIRE-induced apoptosis is associated with nuclear translocation of stress sensor protein GAPDH. Biochem. Biophys. Res. Commun. 423, 32–37.
Lima, F.A., Moreira-Filho, C.A., Ramos, P.L., Brentani, H., Lima, L. de A., Arrais, M., Bento-de-Souza, L.C., Bento-de-Souza, L., Duarte, M.I., Coutinho, A., Carneiro-Sampaio, M., 2011. Decreased AIRE expression and global thymic hypofunction in Down syndrome. J. Immunol. 187, 3422–30.
Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., Fu, X.-D., 2008. The splicing factor SC35 has an active role in transcriptional elongation. Nat. Struct. Mol. Biol. 15, 819–826.
Liston, A., Lesage, S., Wilson, J., Peltonen, L., Goodnow, C.C., 2003. Aire regulates negative selection of organ-specific T cells. Nat. Immunol. 4, 350–354.
Litman, G.W., Rast, J.P., Fugmann, S.D., 2010. The origins of vertebrate adaptive immunity. Nat. Rev. Immunol. 10, 543–53.
Liu, X., Bushnell, D.A., Kornberg, R.D., 2013. RNA polymerase II transcription: structure and mechanism. Biochim. Biophys. Acta – Gene Regul. Mech. 1829, 2–8.
Liu, Y., Beyer, A., Aebersold, R., 2016. On the Dependency of Cellular Protein Levels on mRNA Abundance. Cell 165, 535–550.
Lopes, N., Sergé, A., Ferrier, P., Irla, M., 2015. Thymic Crosstalk Coordinates Medulla Organization and T-Cell Tolerance Induction. Front. Immunol. 6, 365.
Love, M.I., Anders, S., Kim, V., Huber, W., 2015. RNA-Seq workflow: gene-level exploratory analysis and differential expression. F1000Research 4, 1070.
Lutz, C.S., Moreira, A., 2011. Alternative mRNA polyadenylation in eukaryotes: an effective regulator of gene expression. Wiley Interdiscip. Rev. RNA 2, 22–31.
Lynch, H.E., Goldberg, G.L., Chidgey, A., Van den Brink, M.R.M., Boyd, R., Sempowski, G.D., 2009. Thymic involution and immune reconstitution. Trends Immunol. 30, 366–373.
Lyu, Y.L., Lin, C.-P., Azarova, A.M., Cai, L., Wang, J.C., Liu, L.F., 2006. Role of topoisomerase IIbeta in the expression of developmentally regulated genes. Mol. Cell. Biol. 26, 7929–41.
Ma, J., Wang, M.D., 2016. DNA supercoiling during transcription. Biophys. Rev. 8, 75–87.
Macedo, C., Evangelista, A.F., Marques, M.M., Octacílio-Silva, S., Donadi, E.A., Sakamoto-Hojo, E.T., Passos, G.A., 2013. Autoimmune regulator (Aire) controls the expression of microRNAs in medullary thymic epithelial cells. Immunobiology 218, 554–560.
Malchow, S., Leventhal, D.S., Lee, V., Nishi, S., Socci, N.D., Savage, P.A., 2016. Aire Enforces Immune Tolerance by Directing Autoreactive T Cells into the Regulatory T Cell Lineage. Immunity 44, 1–12.
Malchow, S., Leventhal, D.S., Nishi, S., Fischer, B.I., Shen, L., Paner, G.P., Amit, A.S., Kang, C., Geddes, J.E., Allison, J.P., Socci, N.D., Savage, P.A., 2013. Aire-dependent thymic development of tumor-associated regulatory T cells. Science 339, 1219–24.
Malhotra, D., Linehan, J.L., Dileepan, T., Lee, Y.J., Purtha, W.E., Lu, J. V, Nelson, R.W., Fife, B.T., Orr, H.T., Anderson, M.S., Hogquist, K. a, Jenkins, M.K., 2016. Tolerance

is established in polyclonal CD4+ T cells by distinct mechanisms, according to self-peptide expression patterns. Nat. Immunol. 17, 187–195.
Mansfield, R.E., Musselman, C.A., Kwan, A.H., Oliver, S.S., Garske, A.L., Davrazou, F., Denu, J.M., Kutateladze, T.G., Mackay, J.P., 2011. Plant Homeodomain (PHD) Fingers of CHD4 Are Histone H3-binding Modules with Preference for Unmodified H3K4 and Methylated H3K9. J. Biol. Chem. 286, 11779–11791.
Maston, G.A., Evans, S.K., Green, M.R., 2006. Transcriptional Regulatory Elements in the Human Genome. Annu. Rev. Genomics Hum. Genet. 7, 29–59.
Mathis, D., Benoist, C., 2009. Aire. Annu. Rev. Immunol. 27, 287–312.
Mazza, C., Buzi, F., Ortolani, F., Vitali, A., Notarangelo, L.D., Weber, G., Bacchetta, R., Soresina, A., Lougaris, V., Greggio, N.A., Taddio, A., Pasic, S., de Vroede, M., Pac, M., Kilic, S.S., Ozden, S., Rusconi, R., Martino, S., Capalbo, D., Salerno, M., Pignata, C., Radetti, G., Maggiore, G., Plebani, A., Notarangelo, L.D., Badolato, R., 2011. Clinical heterogeneity and diagnostic delay of autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy syndrome. Clin. Immunol. 139, 6–11.
McClendon, A.K., Osheroff, N., 2007. DNA topoisomerase II, genotoxicity, and cancer. Mutat. Res. Mol. Mech. Mutagen. 623, 83–97.
McNamara, R.P., Reeder, J.E., McMillan, E.A., Bacon, C.W., McCann, J.L., D'Orso, I., 2016. KAP1 Recruitment of the 7SK snRNP Complex to Promoters Enables Transcription Elongation by RNA Polymerase II. Mol. Cell 61, 39–53.
Meloni, A., Fiorillo, E., Corda, D., Incani, F., Serra, M.L., Contini, A., Cao, A., Rosatelli, M.C., 2010. DAXX is a new AIRE-interacting protein. J. Biol. Chem. 285, 13012–21.
Meloni, A., Furcas, M., Cetani, F., Marcocci, C., Falorni, A., Perniola, R., Pura, M., Bøe Wolff, A.S., Husebye, E.S., Lilic, D., Ryan, K.R., Gennery, A.R., Cant, A.J., Abinun, M., Spickett, G.P., Arkwright, P.D., Denning, D., Costigan, C., Dominguez, M., McConnell, V., Willcox, N., Meager, A., 2008. Autoantibodies against Type I Interferons as an Additional Diagnostic Criterion for Autoimmune Polyendocrine Syndrome Type I. J. Clin. Endocrinol. Metab. 93, 4389–4397.
Meloni, A., Incani, F., Corda, D., Cao, A., 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., Benoist, C., 2015. Aire controls gene expression in the thymic epithelium with ordered stochasticity. Nat Immunol 16, 942–949.
Mesnard-Rouiller, L., Bismuth, J., Wakkach, A., Poëa-Guyon, S., Berrih-Aknin, S., 2004. Thymic myoid cells express high levels of muscle genes. J. Neuroimmunol. 148, 97–105.
Metzger, T.C., Khan, I.S., Gardner, J.M., Mouchess, M.L., Johannes, K.P., Krawisz, A.K., Skrzypczynska, K.M., Anderson, M.S., 2013. Lineage Tracing and Cell Ablation Identify a Post-Aire-Expressing Thymic Epithelial Cell Population. Cell Rep. 5, 166–179.
Millar, C.B., Grunstein, M., 2006. Genome-wide patterns of histone modifications in yeast. Nat. Rev. Mol. Cell Biol. 7, 657–666.
Millar, D.G., Ohashi, P.S., 2016. Central tolerance: what you see is what you don't get! Nat. Immunol. 17, 115–6.
Miller, C.N., Proekt, I., von Moltke, J., Wells, K.L., Rajpurkar, A.R., Wang, H., Rattay, K., Khan, I.S., Metzger, T.C., Pollack, J.L., Fries, A.C., Lwin, W.W., Wigton, E.J., Parent, A. V., Kyewski, B., Erle, D.J., Hogquist, K.A., Steinmetz, L.M., Locksley,

R.M., Anderson, M.S., 2018. Thymic tuft cells promote an IL-4-enriched medulla and shape thymocyte development. Nature 559, 627–631.
Miller, J.F.A.P., 1962. Effect of neonatal thymectomy on the immunological responsiveness of the mouse. Proc. R. Soc. London. Ser. B. Biol. Sci. 156, 415–428.
Millet, V., Naquet, P., Guinamard, R.R., 2008. Intercellular MHC transfer between thymic epithelial and dendritic cells. Eur. J. Immunol. 38, 1257–1263.
Miragaia, R.J., Zhang, X., Gomes, T., Svensson, V., Ilicic, T., Henriksson, J., Kar, G., Lönnberg, T., 2018. Single-cell RNA-sequencing resolves self-antigen expression during mTEC development. Sci. Rep. 8, 685.
Mishler, D.M., Christ, A.B., Steitz, J.A., 2008. Flexibility in the site of exon junction complex deposition revealed by functional group and RNA secondary structure alterations in the splicing substrate. RNA 14, 2657–70.
Mitchell, P.J., Tjian, R., 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science 245, 371–8.
Mittaz, L., Rossier, C., Heino, M., Peterson, P., Krohn, K.J.E., Gos, A., Morris, M.A., Kudoh, J., Shimizu, N., Antonarakis, S.E., Scott, H.S., 1999. Isolation and Characterization of the Mouse Aire Gene. Biochem. Biophys. Res. Commun. 255, 483–490.
Mo, X., Dynan, W.S., 2002. Subnuclear localization of Ku protein: functional association with RNA polymerase II elongation sites. Mol. Cell. Biol. 22, 8088–99.
Moll, U., Lau, R., Sypes, M.A., Gupta, M.M., Anderson, C.W., 1999. DNA-PK, the DNA-activated protein kinase, is differentially expressed in normal and malignant human tissues. Oncogene 18, 3114–3126.
Montecucco, A., Biamonti, G., 2007. Cellular response to etoposide treatment. Cancer Lett. 252, 9–18.
Montecucco, A., Biamonti, G., 2013. Pre-mRNA processing factors meet the DNA damage response. Front. Genet. 4, 102.
Montecucco, A., Zanetta, F., Biamonti, G., 2015. Molecular mechanisms of etoposide. EXCLI J. 14, 95–108.
Moore, M.J., Proudfoot, N.J., 2009. Pre-mRNA processing reaches back to transcription and ahead to translation. Cell 136, 688–700.
Morf, J.J., Wingett, S.W., Farabella, I., Cairns, J., Furlan-Magaril, M., Liu, X., Craig, F.F., Andrews, S., Marti-Renom, M.A., Fraser, P., 2017. Spatial RNA proximities reveal a bipartite nuclear transcriptome and territories of differential density and transcription elongation rates. bioRxiv 196147.
Muratani, M., Tansey, W.P., 2003. How the ubiquitin–proteasome system controls transcription. Nat. Rev. Mol. Cell Biol. 4, 192–201.
Murray, J.M., Kaufmann, G.R., Hodgkin, P.D., Lewin, S.R., Kelleher, A.D., Davenport, M.P., Zaunders, J.J., 2003. Naive T cells are maintained by thymic output in early ages but by proliferation without phenotypic change after age twenty. Immunol. Cell Biol. 81, 487–495.
Murumägi, A., Vähämurto, P., Peterson, P., 2003. Characterization of regulatory elements and methylation pattern of the autoimmune regulator (AIRE) promoter. J. Biol. Chem. 278, 19784–90.
Nagamine, K., Peterson, P., Scott, H.S., Kudoh, J., Minoshima, S., Heino, M., Krohn, K.J., Lalioti, M.D., Mullis, P.E., Antonarakis, S.E., Kawasaki, K., Asakawa, S., Ito, F., Shimizu, N., 1997. Positional cloning of the APECED gene. Nat Genet 17, 393–398.

Nedjic, J., Aichinger, M., Emmerich, J., Mizushima, N., Klein, L., 2008. Autophagy in thymic epithelium shapes the T-cell repertoire and is essential for tolerance. Nature 455, 396–400.

Niki, S., Oshikawa, K., Mouri, Y., Hirota, F., Matsushima, A., Yano, M., Han, H., Bando, Y., Izumi, K., Matsumoto, M., Nakayama, K.I., Kuroda, N., Matsumoto, M., 2006. Alteration of intra-pancreatic target-organ specificity by abrogation of Aire in NOD mice. J. Clin. Invest. 116, 1292–301.

Nishijima, H., Kajimoto, T., Matsuoka, Y., Mouri, Y., Morimoto, J., Matsumoto, M., Kawano, H., Nishioka, Y., Uehara, H., Izumi, K., Tsuneyama, K., Okazaki, I., Okazaki, T., Hosomichi, K., Shiraki, A., Shibutani, M., Mitsumori, K., Matsumoto, M., 2018. Paradoxical development of polymyositis-like autoimmunity through augmented expression of autoimmune regulator (AIRE). J. Autoimmun. 86, 75–92.

Nishikawa, Y., Hirota, F., Yano, M., Kitajima, H., Miyazaki, J., Kawamoto, H., Mouri, Y., 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.

Nitta, T., Suzuki, H., 2016. Thymic stromal cell subsets for T cell development. Cell. Mol. Life Sci. 73, 1021–1037.

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., Fleishman, S.J., Banka, S., Newman, W.G., Sewell, W. a C., Sozaeva, L.S., Zayats, T., Haugarvoll, K., Orlova, E.M., Haavik, J., Johansson, S., Knappskog, P.M., Løvås, K., Wolff, A.S.B., Abramson, J., Husebye, E.S., 2015. Dominant Mutations in the Autoimmune Regulator AIRE Are Associated with Common Organ-Specific Autoimmune Diseases. Immunity 42, 1185–96.

Oh, J., Shin, J.-S., 2015. The Role of Dendritic Cells in Central Tolerance. Immune Netw. 15, 111.

Ohler, U., Liao, G., Niemann, H., Rubin, G.M., 2002. Computational analysis of core promoters in the Drosophila genome. Genome Biol. 3, RESEARCH0087.

Org, T., Chignola, F., Hetényi, C., Gaetani, M., Rebane, A., Liiv, I., Maran, U., Mollica, L., Bottomley, M.J., Musco, G., Peterson, P., 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., Peterson, P., 2009. AIRE activated tissue specific genes have histone modifications associated with inactive chromatin. Hum. Mol. Genet. 18, 4699–710.

Ossart, J., Moreau, A., Autrusseau, E., Ménoret, S., Martin, J.C., Besnard, M., Ouisse, L.-H., Tesson, L., Flippe, L., Kisand, K., Peterson, P., Hubert, F.-X., Anegon, I., Josien, R., Guillonneau, C., 2018. Breakdown of Immune Tolerance in AIRE-Deficient Rats Induces a Severe Autoimmune Polyendocrinopathy–Candidiasis–Ectodermal Dystrophy–like Autoimmune Disease. J. Immunol. 201, 874–887.

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

Palmer, E., 2003. Negative selection – clearing out the bad apples from the T-cell repertoire. Nat. Rev. Immunol. 3, 383–391.

Parkin, J., Cohen, B., 2001. An overview of the immune system. Lancet 357, 1777–1789.

Pawelczak, K.S., Bennett, S.M., Turchi, J.J., 2011. Coordination of DNA-PK activation and nuclease processing of DNA termini in NHEJ. Antioxid. Redox Signal. 14, 2531–43.
Pennacchio, L.A., Bickmore, W., Dean, A., Nobrega, M.A., Bejerano, G., 2013. Enhancers: five essential questions. Nat. Rev. Genet. 14, 288–95.
Perry, J.S.A., Lio, C.W.J., Kau, A.L., Nutsch, K., Yang, Z., Gordon, J.I., Murphy, K.M., Hsieh, C.S., 2014. Distinct contributions of Aire and antigen-presenting-cell subsets to the generation of self-tolerance in the thymus. Immunity 41, 414–426.
Peterlin, B.M., Price, D.H., 2006. Controlling the Elongation Phase of Transcription with P-TEFb. Mol. Cell 23, 297–305.
Peterson, P., Org, T., Rebane, A., 2008. Transcriptional regulation by AIRE: molecular mechanisms of central tolerance. Nat Rev Immunol 8, 948–957.
Petrie, H.T., Zúñiga-Pflücker, J.C., 2007. Zoned out: functional mapping of stromal signaling microenvironments in the thymus. Annu. Rev. Immunol. 25, 649–679.
Phatnani, H.P., Greenleaf, A.L., 2006. Phosphorylation and functions of the RNA polymerase II CTD. Genes Dev. 20, 2922–2936.
Pignatelli, M., Serras, F., Moya, A., Guigo, R., 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., Kyewski, B., 2013. Overlapping gene coexpression patterns in human medullary thymic epithelial cells generate self-antigen diversity. Proc. Natl. Acad. Sci. 110, E3497–E3505.
Piquet, S., Parc, F. Le, Bai, S.-K., Chevallier, O., Adam, S., Polo, S.E., 2018. The histone chaperone FACT coordinates H2A.X-dependent signaling and repair of DNA damage. bioRxiv 264952.
Pitkänen, J., Doucas, V., Sternsdorf, T., Nakajima, T., Aratani, S., Jensen, K., Will, H., Vähämurto, P., Ollila, J., Vihinen, M., Scott, H.S., Antonarakis, S.E., Kudoh, J., Shimizu, N., Krohn, K., Peterson, P., 2000. The autoimmune regulator protein has transcriptional transactivating properties and interacts with the common coactivator CREB-binding protein. J. Biol. Chem. 275, 16802–9.
Pitkänen, J., Vähämurto, P., Krohn, K., Peterson, P., 2001. Subcellular localization of the autoimmune regulator protein. characterization of nuclear targeting and transcriptional activation domain. J. Biol. Chem. 276, 19597–19602.
Plevin, M.J., Mills, M.M., Ikura, M., 2005. The LxxLL motif: a multifunctional binding sequence in transcriptional regulation. Trends Biochem. Sci. 30, 66–9.
Pollex, T., Furlong, E.E.M., 2017. Correlation Does Not Imply Causation: Histone Methyltransferases, but Not Histone Methylation, SET the Stage for Enhancer Activation. Mol. Cell 66, 439–441.
Pommier, Y., 2013. Drugging Topoisomerases: Lessons and Challenges. ACS Chem. Biol 8, 41.
Pommier, Y., Sun, Y., Huang, S.N., Nitiss, J.L., 2016. Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. Nat. Rev. Mol. Cell Biol. 17, 703–721.
Pöntynen, N., Miettinen, A., Petteri Arstila, T., Kämpe, O., Alimohammadi, M., Vaarala, O., Peltonen, L., Ulmanen, I., 2006. Aire deficient mice do not develop the same profile of tissue-specific autoantibodies as APECED patients. J. Autoimmun. 27, 96–104.
Porrua, O., Libri, D., 2015. Transcription termination and the control of the transcriptome: why, where and how to stop. Nat. Rev. Mol. Cell Biol. 16, 190–202.

Poss, Z.C., Ebmeier, C.C., Taatjes, D.J., 2013. The Mediator complex and transcription regulation. Crit. Rev. Biochem. Mol. Biol. 48, 575–608.
Pradet-Balade, B., Boulmé, F., Beug, H., Müllner, E.W., Garcia-Sanz, J.A., 2001. Translation control: Bridging the gap between genomics and proteomics? Trends Biochem. Sci. 26, 225–229.
Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M.S., Mapendano, C.K., Schierup, M.H., Jensen, T.H., 2008. RNA exosome depletion reveals transcription upstream of active human promoters. Science 322, 1851–4.
Puc, J., Aggarwal, A.K., 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., She, J.X., 2005. Mapping DNA-binding domains of the autoimmune regulator protein. Biochem Biophys Res Commun 327, 939–944.
Rabinowe, S.L., Rubin, I.L., George, K.L., Adri, M.N.S., Eisenbarth, G.S., 1989. Trisomy 21 (Down's syndrome): Autoimmunity, aging and monoclonal antibody-defined T-cell abnormalities. J. Autoimmun. 2, 25–30.
Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., Young, R.A., 2010. c-Myc Regulates Transcriptional Pause Release. Cell 141, 432–445.
Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., Tyagi, S., 2006. Stochastic mRNA Synthesis in Mammalian Cells. PLoS Biol. 4, e309.
Ramsey, C., Bukrinsky, A., Peltonen, L., 2002. Systematic mutagenesis of the functional domains of AIRE reveals their role in intracellular targeting. Hum Mol Genet 11, 3299–3308.
Ransohoff, J.D., Wei, Y., Khavari, P.A., 2017. The functions and unique features of long intergenic non-coding RNA. Nat. Rev. Mol. Cell Biol. 19, 143–157.
Ray Chaudhuri, A., Nussenzweig, A., 2017. The multifaceted roles of PARP1 in DNA repair and chromatin remodelling. Nat. Rev. Mol. Cell Biol. 18, 610–621.
Reiter, F., Wienerroither, S., Stark, A., 2017. Combinatorial function of transcription factors and cofactors. Curr. Opin. Genet. Dev. 43, 73–81.
Rialdi, A., Campisi, L., Zhao, N., Lagda, A.C., Pietzsch, C., Ho, J.S.Y., Martinez-Gil, L., Fenouil, R., Chen, X., Edwards, M., Metreveli, G., Jordan, S., Peralta, Z., Munoz-Fontela, C., Bouvier, N., Merad, M., Jin, J., Weirauch, M., Heinz, S., Benner, C., van Bakel, H., Basler, C., Garcia-Sastre, A., Bukreyev, A., Marazzi, I., 2016. Topoisomerase 1 inhibition suppresses inflammatory genes and protects from death by inflammation. Science 352, aad7993-aad7993.
Rinderle, C., Christensen, H.-M., Schweiger, S., Lehrach, H., 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.
Roca, J., 2011. Transcriptional inhibition by DNA torsional stress. Transcription 2, 82–85.
Rodewald, H.-R., 2008. Thymus organogenesis. Annu. Rev. Immunol. 26, 355–388.
Rosenberg, M., Fan, A.X., Lin, I.-J., Liang, S.Y., Bungert, J., 2013. Cell-cycle specific association of transcription factors and RNA polymerase II with the human β-globin gene locus. J. Cell. Biochem. 114, 1997–2006.
Rossi, S.W., Jenkinson, W.E., Anderson, G., Jenkinson, E.J., 2006. Clonal analysis reveals a common progenitor for thymic cortical and medullary epithelium. Nature 441, 988–991.

Rossi, S.W., Kim, M.-Y., Leibbrandt, A., Parnell, S.M., Jenkinson, W.E., Glanville, S.H., McConnell, F.M., Scott, H.S., Penninger, J.M., Jenkinson, E.J., Lane, P.J.L., Anderson, G., 2007. RANK signals from CD4(+)3(−)inducer cells regulate development of Aire-expressing epithelial cells in the thymic medulla. J. Exp. Med. 204, 1267–1272.

Rougvie, A.E., Lis, J.T., 1990. Postinitiation transcriptional control in Drosophila melanogaster. Mol. Cell. Biol. 10, 6041–6045.

Ruan, Q.-G., Tung, K., Eisenman, D., Setiady, Y., Eckenrode, S., Yi, B., Purohit, S., Zheng, W.-P., Zhang, Y., Peltonen, L., She, J.-X., 2007. The autoimmune regulator directly controls the expression of genes critical for thymic epithelial function. J. Immunol. 178, 7173–80.

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

Salceda, J., Ferná Ndez, X., Roca, J., 2006. Topoisomerase II, not topoisomerase I, is the proficient relaxase of nucleosomal DNA. EMBO J. 25, 2575–2583.

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

Sànchez, R., Marzluff, W.F., Sa, R., 2002. The Stem-Loop Binding Protein Is Required for Efficient Translation of Histone mRNA In Vivo and In Vitro The Stem-Loop Binding Protein Is Required for Efficient Translation of Histone mRNA In Vivo and In Vitro. Society 22, 7093–7104.

Sano, K., Miyaji-Yamaguchi, M., Tsutsui, K.M., Tsutsui, K., 2008. Topoisomerase IIbeta activates a subset of neuronal genes that are repressed in AT-rich genomic environment. PLoS One 3, e4103.

Sansom, S.N., Shikama-Dorn, N., Zhanybekova, S., Nusspaumer, G., Macaulay, I.C., Deadman, M.E., Heger, A., Ponting, C.P., 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.

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

Schek, N., Cooke, C., Alwine, J.C., 1992. Definition of the upstream efficiency element of the simian virus 40 late polyadenylation signal by using in vitro analyses. Mol. Cell. Biol. 12, 5386–93.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., Selbach, M., 2011. Global quantification of mammalian gene expression control. Nature 473, 337–342.

Schwertman, P., Bekker-Jensen, S., Mailand, N., 2016. Regulation of DNA double-strand break repair by ubiquitin and ubiquitin-like modifiers. Nat. Rev. Mol. Cell Biol. 17, 379–394.

Shi, X., Hong, T., Walter, K.L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Peña, P., Lan, F., Kaadige, M.R., Lacoste, N., Cayrou, C., Davrazou, F., Saha, A., Cairns, B.R., Ayer, D.E., Kutateladze, T.G., Shi, Y., Côté, J., Chua, K.F., Gozani, O., 2006. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature 442, 96–99.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., de Laat, W., 2006. Nuclear organization of active and inactive chromatin domains

uncovered by chromosome conformation capture–on-chip (4C). Nat. Genet. 38, 1348–1354.
Singh, J., Singh, A.K., 1979. Age-related changes in human thymus. Clin Exp Immunol 37, 507–11.
Skogberg, G., Lundberg, V., Berglund, M., Gudmundsdottir, J., Telemo, E., Lindgren, S., Ekwall, O., 2015. Human thymic epithelial primary cells produce exosomes carrying tissue-restricted antigens. Immunol. Cell Biol. 93, 727–734.
Skogberg, G., Lundberg, V., Lindgren, S., Gudmundsdottir, J., Sandstrom, K., Kampe, O., Anneren, G., Gustafsson, J., Sunnegardh, J., van der Post, S., Telemo, E., Berglund, M., Ekwall, O., 2014. Altered Expression of Autoimmune Regulator in Infant Down Syndrome Thymus, a Possible Contributor to an Autoimmune Phenotype. J. Immunol. 193, 2187–2195.
Skourti-Stathaki, K., Proudfoot, N.J., 2014. A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. Genes Dev. 28, 1384–96.
Spits, H., 2002. Development of αβ T cells in the human thymus. Nat. Rev. Immunol. 2, 760–772.
Spitz, F., Furlong, E.E.M., 2012. Transcription factors: from enhancer binding to developmental control. Nat. Rev. Genet. 13, 613–626.
St-Pierre, C., Trofimov, A., Brochu, S., Lemieux, S., Perreault, C., 2015. Differential Features of AIRE-Induced and AIRE-Independent Promiscuous Gene Expression in Thymic Epithelial Cells. J Immunol 195, 498–506.
Štros, M., Launholt, D., Grasser, K.D., 2007. The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins. Cell. Mol. Life Sci. 64, 2590–2606.
Stros, M., Reich, J., 1998. Formation of large nucleoprotein complexes upon binding of the high-mobility-group (HMG) box B-domain of HMG1 protein to supercoiled DNA. Eur. J. Biochem. 251, 427–34.
Strzyz, P., 2016. Controlling ubiquitylation at DNA lesions. Nat. Rev. Mol. Cell Biol. 17, 329–329.
Su, M.A., Giang, K., Žumer, K., Jiang, H., Oven, I., Rinn, J.L., Devoss, J.J., Johannes, K.P.A., Lu, W., Gardner, J., Chang, A., Bubulya, P., Chang, H.Y., Matija Peterlin, B., Anderson, M.S., Zumer, K., Jiang, H., Oven, I., Rinn, J.L., Devoss, J.J., Johannes, K.P.A., Lu, W., Gardner, J., Chang, A., Bubulya, P., Chang, H.Y., Peterlin, B.M., Anderson, M.S., 2008. Mechanisms of an autoimmunity syndrome in mice caused by a dominant mutation in Aire. J Clin Invest 118, 1712–1726.
Sun, L., Sun, C., Liang, Z., Li, H., Chen, L., Luo, H., Zhang, H., Ding, P., Sun, X., Qin, Z., Zhao, Y., 2015. FSP1+fibroblast subpopulation is essential for the maintenance and regeneration of medullary thymic epithelial cells. Sci. Rep. 5, 1–16.
Sutphin, A., Fuller Albright, M.D., McCune, D.J., 1943. Five Cases (Three in Siblings) of Idiopathic Hypoparathyroidism Associated with Moniliasis 1. J. Clin. Endocrinol. 3, 625–634.
Takaba, H., Morishita, Y., Tomofuji, Y., Danks, L., Nitta, T., Komatsu, N., Kodama, T., Takayanagi, H., 2015. Fezf2 Orchestrates a Thymic Program of Self-Antigen Expression for Immune Tolerance. Cell 163, 975–987.
Tammaro, M., Barr, P., Ricci, B., Yan, H., 2013. Replication-dependent and transcription-dependent mechanisms of DNA double-strand break induction by the topoisomerase 2-targeting drug etoposide. PLoS One 8, e79202.

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

The ENCODE Project Consortium, 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447, 799–816.

The Finnish-German APECED Consortium, 1997. An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. Nat. Genet. 17, 399–403.

Thompson, D.M., Parker, R., 2006. Cytoplasmic Decay of Intergenic Transcripts in Saccharomyces cerevisiae. Mol. Cell. Biol. 27, 92–101.

Thorpe, E.S., Handley, H.E., 1929. Chronic tetany and chronic mycelial stomatitis in a child aged four and one-half years. Arch. Pediatr. Adolesc. Med. 38, 328–338.

Tian, B., Hu, J., Zhang, H., Lutz, C.S., 2005. A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res. 33, 201–212.

Tonooka, A., Kubo, T., Ichimiya, S., Tamura, Y., Ilmarinen, T., Ulmanen, I., Kimura, S., Yokoyama, S., Takano, Y., Kikuchi, T., Sato, N., 2009. Wild-type AIRE cooperates with p63 in HLA class II expression of medullary thymic stromal cells. Biochem. Biophys. Res. Commun. 379, 765–770.

Tous, C., Aguilera, A., 2007. Impairment of transcription elongation by R-loops in vitro. Biochem. Biophys. Res. Commun. 360, 428–432.

Trotter, K.W., King, H.A., Archer, T.K., 2015. Glucocorticoid Receptor Transcriptional Activation via the BRG1-Dependent Recruitment of TOP2β and Ku70/86. Mol. Cell. Biol. 35, 2799–2817.

Tupler, R., Perini, G., Green, M.R., 2001. Expressing the human genome. Nature 409, 832–833.

Uchida, D., Hatakeyama, S., Matsushima, A., Han, H., Ishido, S., Hotta, H., Kudoh, J., Shimizu, N., Doucas, V., Nakayama, K.I., Kuroda, N., Matsumoto, M., 2004. AIRE Functions As an E3 Ubiquitin Ligase. J. Exp. Med. 199, 167–172.

Uusküla-Reimand, L., Hou, H., Samavarchi-Tehrani, P., Rudan, M.V., Liang, M., Medina-Rivera, A., Mohammed, H., Schmidt, D., Schwalie, P., Young, E.J., Reimand, J., Hadjur, S., Gingras, A.-C., Wilson, M.D., 2016. Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders. Genome Biol. 17, 182.

Vafiadis, P., Bennett, S.T., Todd, J.A., Nadeau, J., Grabs, R., Goodyer, C.G., Wickramasinghe, S., Colle, E., Polychronakos, C., 1997. Insulin expression in human thymus is modulated by INS VNTR alleles at the IDDM2 locus. Nat. Genet. 15, 289–292.

Vaidya, H.J., Briones Leon, A., Blackburn, C.C., 2016. FOXN1 in thymus organogenesis and development. Eur. J. Immunol. 46, 1826–1837.

van der Burg, M., van Dongen, J.J.M., van Gent, D.C., 2009. DNA-PKcs deficiency in human: Long predicted, finally found. Curr. Opin. Allergy Clin. Immunol. 9, 503–509.

Vaquerizas, J.M., Kummerfeld, S.K., Teichmann, S.A., Luscombe, N.M., 2009. A census of human transcription factors: function, expression and evolution. Nat. Rev. Genet. 10, 252–263.

Venanzi, E.S., Benoist, C., Mathis, D., 2004. Good riddance: thymocyte clonal deletion prevents autoimmunity. Curr. Opin. Immunol. 16, 197–202.

Venanzi, E.S., Melamed, R., Mathis, D., Benoist, C., 2008. The variable immunological self: genetic variation and nongenetic noise in Aire-regulated transcription. Proc Natl Acad Sci U S A 105, 15860–15865.
Venkatesh, S., Workman, J.L., 2015. Histone exchange, chromatin structure and the regulation of transcription. Nat. Rev. Mol. Cell Biol. 16, 178–189.
Venters, B.J., Pugh, B.F., 2013. Genomic organization of human transcription initiation complexes. Nature 502, 53–8.
Vermunt, M.W., Zhang, D., Blobel, G.A., 2019. The interdependence of gene-regulatory elements and the 3D genome. J. Cell Biol. 218, 12–26.
Villaseñor, J., Besse, W., Benoist, C., Mathis, D., 2008. Ectopic expression of peripheral-tissue antigens in the thymic epithelium: probabilistic, monoallelic, misinitiated. Proc Natl Acad Sci U S A 105, 15854–15859.
Vo Ngoc, L., Cassidy, C.J., Huang, C.Y., Duttke, S.H.C., Kadonaga, J.T., 2017. The human initiator is a distinct and abundant element that is precisely positioned in focused core promoters. Genes Dev. 31, 6–11.
Vogel, A., Strassburg, C.P., Obermayer-Straub, P., Brabant, G., Manns, M.P., 2002. The genetic background of autoimmune polyendocrinopathy–candidiasis–ectodermal dystrophy and its autoimmune disease components. J. Mol. Med. 80, 201–211.
Vogel, C., Marcotte, E.M., 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat. Rev. Genet. 13, 227–232.
von Herrath, M.G., Dockter, J., Oldstone, M.B.A., 1994. How virus induces a rapid or slow onset insulin-dependent diabetes mellitus in a transgenic model. Immunity 1, 231–242.
Vopálenský, V., Masek, T., Horváth, O., Vicenová, B., Mokrejs, M., Pospísek, M., 2008. Firefly luciferase gene contains a cryptic promoter. RNA 14, 1720–1729.
Waldmann, H., 2016. Mechanisms of immunological tolerance. Clin. Biochem. 49, 324–328.
Wang, X., Laan, M., Bichele, R., Kisand, K., Scott, H.S., Peterson, P., 2012. Post-Aire Maturation of Thymic Medullary Epithelial Cells Involves Selective Expression of Keratinocyte-Specific Autoantigens. Front. Immunol. 3, 19.
Waterfield, M., Khan, I.S., Cortez, J.T., Fan, U., Metzger, T., Greer, A., Fasano, K., Martinez-Llordella, M., Pollack, J.L., Erle, D.J., Su, M., Anderson, M.S., 2014. The transcriptional regulator Aire coopts the repressive ATF7ip-MBD1 complex for the induction of immunotolerance. Nat Immunol 15, 258–265.
Winter, G.E., Mayer, A., Buckley, D.L., Erb, M.A., Roderick, J.E., Vittori, S., Reyes, J.M., di Iulio, J., Souza, A., Ott, C.J., Roberts, J.M., Zeid, R., Scott, T.G., Paulk, J., Lachance, K., Olson, C.M., Dastjerdi, S., Bauer, S., Lin, C.Y., Gray, N.S., Kelliher, M.A., Churchman, L.S., Bradner, J.E., 2017. BET Bromodomain Proteins Function as Master Transcription Elongation Factors Independent of CDK9 Recruitment. Mol. Cell 67, 5–18.
Wolff, A.S., Sarkadi, A.K., Maródi, L., Kärner, J., Orlova, E., Oftedal, B.E., Kisand, K., Oláh, E., Meloni, A., Myhre, A.G., Husebye, E.S., Motaghedi, R., Perheentupa, J., Peterson, P., Willcox, N., Meager, A., 2013. Anti-cytokine autoantibodies preceding onset of autoimmune polyendocrine syndrome type I features in early childhood. J Clin Immunol 33, 1341–1348.
Woodard, R.L., Lee, K.J., Huang, J., Dynan, W.S., 2001. Distinct roles for Ku protein in transcriptional reinitiation and DNA repair. J. Biol. Chem. 276, 15423–33.
Wu, C.-H., Yamaguchi, Y., Benjamin, L.R., Horvat-Gordon, M., Washinsky, J., Enerly, E., Larsson, J., Lambertsson, A., Handa, H., Gilmour, D., 2003. NELF and DSIF

cause promoter proximal pausing on the hsp70 promoter in Drosophila. Genes Dev. 17, 1402–14.
Xing, Y., Hogquist, K.A., 2012. T-Cell Tolerance: Central and Peripheral. Cold Spring Harb. Perspect. Biol. 4, a006957.
Yamano, T., Dobeš, J., Vobořil, M., Steinert, M., Brabec, T., Ziętara, N., Dobešová, M., Ohnmacht, C., Laan, M., Peterson, P., Benes, V., Sedláček, R., Hanayama, R., Kolář, M., Klein, L., Filipp, D., 2019. Aire-expressing ILC3-like cells in the lymph node display potent APC features. J. Exp. Med. 216, 1027–1037.
Yamano, T., Nedjic, J., Hinterberger, M., Steinert, M., Koser, S., Pinto, S., Gerdes, N., Lutgens, E., Ishimaru, N., Busslinger, M., Brors, B., Kyewski, B., Klein, L., 2015. Thymic B Cells Are Licensed to Present Self Antigens for Central T Cell Tolerance Induction. Immunity 42, 1048–1061.
Yang, S., Bansal, K., Lopes, J., Benoist, C., Mathis, D., 2013. Aire's plant homeodomain (PHD)-2 is critical for induction of immunological tolerance. Proc Natl Acad Sci U S A 110, 1833–1838.
Yang, S., Fujikado, N., Kolodin, D., Benoist, C., Mathis, D., 2015. Regulatory T cells generated early in life play a distinct role in maintaining self-tolerance. Science 348, 589–594.
Yoshida, H., Bansal, K., Schaefer, U., Chapman, T., Rioja, I., Proekt, I., Anderson, M.S., Prinjha, R.K., Tarakhovsky, A., Benoist, C., Mathis, D., 2015. Brd4 bridges the transcriptional regulators, Aire and P-TEFb, to promote elongation of peripheral-tissue antigen transcripts in thymic stromal cells. Proc. Natl. Acad. Sci. 112, E4448–E4457.
Zabidi, M.A., Stark, A., 2016. Regulatory Enhancer–Core-Promoter Communication via Transcription Factors and Cofactors. Trends Genet. 32, 801–814.
Zhao, Z., Tavoosidana, G., Sjölinder, M., Göndör, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Singh Sandhu, K., Singh, U., Pant, V., Tiwari, V., Kurukuti, S., Ohlsson, R., 2006. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat. Genet. 38, 1341–1347.
Zhou, Q., Li, T., Price, D.H., 2012. RNA Polymerase II Elongation Control. Annu. Rev. Biochem. 81, 119–143.
Ziv, Y., Bielopolski, D., Galanty, Y., Lukas, C., Taya, Y., Schultz, D.C., Lukas, J., Bekker-Jensen, S., Bartek, J., Shiloh, Y., 2006. Chromatin relaxation in response to DNA double-strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. Nat. Cell Biol. 8, 870–876.
Zobeck, K.L., Buckley, M.S., Zipfel, W.R., Lis, J.T., 2010. Recruitment Timing and Dynamics of Transcription Factors at the Hsp70 Loci in Living Cells. Mol. Cell 40, 965–975.
Zumer, K., Low, A.K., Jiang, H., Saksela, K., Peterlin, B.M., 2012. Unmodified Histone H3K4 and DNA-Dependent Protein Kinase Recruit Autoimmune Regulator to Target Genes. Mol. Cell. Biol. 32, 1354–1362.
Žumer, K., Plemenitaš, A., Saksela, K., Peterlin, B.M., 2011. Patient mutation in AIRE disrupts P-TEFb binding and target gene transcription. Nucleic Acids Res. 39, 7908–7919.

SUMMARY IN ESTONIAN

DNA seondumise ja DNA katkete olulisus AIRE-vahendatud transkriptsiooni aktivatsioonis

Immuunsüsteemi peamiseks ülesandeks on pakkuda adekvaatset kaitsefunktsiooni samal ajal vältides kahjulike reaktsioone iseenda vastu. Selleks peab immuunsüsteem suutma eristada oma ja võõrast. Vältimatuks vahendiks on siin omandatud immuunsüsteemi rakud – olulisi effektorfunktsioone täitvad B- ja T-lümfotsüüdid, mille pinnal on avaldunud juhuslikult genereeritud struktuuriga retseptorid, mis tagavad väga paljude erinevate molekulide ära tundmise. Ühelt poolt võimaldab juhuslik retseptorite genereerimine tundmatute antigeenide ära tundmist, kuid teiselt poolt on sel moel raske ära hoida enda keha osiste vastast autoimmuunreaktsiooni. Et viimast vältida, läbivad lümfotsüüdid enne vereringesse sattumist keerulised küpsemisprotsessid primaarsetes lümfoidorganites, mis T-rakkude puhul toimub tüümuses ehk harknäärmes. Kaua aega ei saadud aru, kuidas sobilike T-rakkude välja valimine toimub. Siiski märgati, et tüümusel on ebatavaline võime ekspresseerida väga paljusid valke, mis täidavad spetsiifilisi funktsioone teistes organites, nt. insuliin kõhunäärmes. Hiljem tehti kindlaks, et see ongi vajalik organismile sobilike T-rakkude välja valimiseks, sest need T-rakud, mis tunnevad ära oma keha molekulid, suunatakse apoptoosi ja sel viisil välditakse autoreaktiivsete rakkude väljumist tüümusest. Käesolevaks ajaks on näidatud, et tüümusele iseloomuliku laia spektri geenide avaldumise eest on vastutav AIRE (autoimmuunne regulaator) valk. AIRE olulisus avastati analüüsides peresid, mille liikmetel väljendus harvaesinev autoimmuunhaigus APECED (*autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy*), mille põhilisteks tunnusteks on krooniline *Candida* infektsioon limaskestadel, kõrvalkilpnäärme alatalitlus ja neerupealiste puudulikkus.

Organismis on AIRE avaldunud kõige tugevamalt tüümuse säsi epiteelirakkudes. Kuigi AIRE ekspressioon on kirjeldatud ka teistes immuunsüsteemi rakkudes ja sugurakkudes, ei ole tema funktsionaalne roll seal nii domineeriv. Rakusiseselt lokaliseerub AIRE valk peamiselt tuuma, kus ta moodustab iseloomuliku tähnilise mustri. AIRE mõju transkriptsioonile on erinev kanoonilistest transkriptsiooni faktoritest. AIRE poolt vahendatud geenieksprессioon on nn avatud iseloomuga, st et AIRE mõjul aktiveeritakse väga paljud geenid, kusjuures aktiveeritud geenide kogum sõltub raku tüübist või hiire tüvest. Samuti esineb palju alternatiivsete transkriptsiooni start saitide kasutamist ja vermimismehhanismide ignoreerimist. Lisaks on AIRE märklaudgeenidele iseloomulik algne madal ekspressiooni tase, repressiivsete kromatiinimarkerite olemasolu ja koespetsiifilisus, mis tähendab, et AIRE on võime aktiveerida geene, mis teistes kudedes alluvad rangele kontrollile. Näiteks on kirjeldatud AIRE võimet interakteeruda mittemetüleeritud lüsiiniga histoon H3 neljandas positsioonis, mis on oluline madalalt ekspresseerunud geenide tunnus. Seetõttu, võttes kokku need ebatavalised jooned, ongi nimetatud AIRE poolset regulatsiooni tüümuses avatud geenieksprессiooniks.

Antud doktoritöö eesmärk oli välja selgitada mehhanisme, kuidas AIRE poolt vahendatud geenieksрessiooni aktiveeriv efekt on saavutatav ning kas ja millised geenide funktsionaalsed osad määravad AIRE märklaudgeenide valikut. Selleks, et kindlaks teha AIRE poolseks aktivatsiooniks vajalikud järjestuselemendid, konstrueerisime lutsiferaasi reporteri konstruktid, kus geeni aktiveerimiseks olulised osad olid ühe kaupa või eri kombinatsioonides välja jäetud või vahetatud. Sellisel moel testisime AIRE teadaolevate märklaudgeenide, insuliini (*INS*) ja involukriini (*IVL*), promootori, introni ja polüadenüleerimissignaali olulisust AIRE poolsele aktivatsioonile. Üllatavaks tulemuseks osutus, et HEK293 rakkudesse transfekteeritud AIRE on võimeline aktiveerima reportereid, millel puudusid määratletud promootorjärjestused. Kuigi introni olemasolu reporteris üldiselt tõstis lutsiferaasi aktiivsuse taset, oli muutus reporteri avaldumises AIRE-positiivsete ja negatiivsete rakkude vahel sarnane intronita plasmiidiga. Polüadenüleerimissignaali ära vahetamine histooni tüvi-ling järjestusega ei tühistanud samuti AIRE aktiveerimisefekti. Need tulemused viitavad, et AIRE-poolne mõju geenide avaldumisele toimub transkriptsiooni tasemel, kuid ei ole piiratud märklaudgeenide promootorjärjestuste olemasoluga.

Lisaks sellele, et mitmed geenielemendid ei osutunud oluliseks AIRE-vahendatud aktivatsiooni jaoks plasmiidse matriitsi puhul, täheldasime AIRE tugevat seondumist plasmiidse DNA-ga kromatiini immunosadestamise (ChIP) katsetes. AIRE seondumine plasmiidile oli kordades tugevam kui genoomsele DNA-le ning ei sõltunud teiste valkude olemasolust, kuna detekteerisime AIRE N-terminaalse domeeni ja DNA interaktsiooni ka EMSA (*electrophoretic mobility shift assay*) katsetes. DNA seondumise eest vastutavad aminohapped arginiin 113 ja lüsiin 114 asusid N-terminaalses otsas ja nende muteerimine vähendas AIRE interaktsiooni DNA-ga ChIP analüüsis. Varasemalt oli aminohapetele R113 ja K114 omistatud tuumalokalisatsiooni funktsioon, kuid meie ja teiste hilisemad tööd on näidanud, et nende aminohapete muteerimine rakutuuma transporti ei häiri. Lisaks täheldasime, et AIRE laiendatud HSR/CARD (*homogeneously staining region/caspase recruitment domain*) domeen seondub tugevamini kui täispikk AIRE nii plasmiidse kui genoomse DNA-ga. Genoomse DNA puhul oli täheldatav laiendatud HSR/CARD tugevam seondumine kõrgelt ekspresseerunud geenide promootoraladele. On huvitav, et kõrgelt ekspresseerunud geenid ei ole AIRE märklaudgeenid, kuid nende promootoraladel DNA ei ole pakitud nukleosoomidesse ja sarnaneb seetõttu rohkem plasmiidsega, mis tõenäoliselt soodustabki AIRE laiendatud HSR/CARD domeeni seondumist. Transkriptsiooni aktivatsioonile negatiivset mõju avaldavad punktmutatsioonid HSR/ CARD domeenis takistasid ka AIRE seondumist plasmiidile, rõhutades HSR/ CARD terviklikkuse tähtsust DNA-ga seondumisel.

Lisaks võimele aktiveerida palju geene, iseloomustab AIRE-t võime interakteeruda paljude valkudega. Mõned nendest valkudest on olnud ennustatavad lähtudes AIRE funktsioonist, näiteks tuuma transpordis, RNA protsessingus ja transkriptsioonis osalevad valgud. Teised seonduvad valgud ei sobi AIRE funktsiooniga nii loogiliselt. Üks huvitavamaid ja tugevamini interakteeruvaid AIRE partnervalke, mis käesolevas töös leiti, on DNA-PK (*DNA-dependent protein*

kinase), mis mängib olulist rolli DNA katkete parandamisel läbi fosforüleerimisaktiivsuse. Seos transkriptsiooni efektiivsuse ja DNA reparatsiooni vahel, kuigi kirjeldatud varasemalt, ei ole funktsionaalselt lõpuni selge. DNA-PK moodustab osa suuremast AIRE-ga interakteeruvast kompleksist, mille hulka kuuluvad ka topoisomeraasid – ensüümid, mis lahendavad DNA topoloogilist stressi läbi lühiajaliste ühe- ja kaheahelaliste DNA katkete tekke ja parandamise. Topoisomeraasid on olulised paljudes rakulistes protsessides, kaasa arvatud replikatsioonis ja rekombinatsioonis kuid nende osalus transkriptsioonis on vähem uuritud. Hiired, kellel topoisomeraase kodeeritavad geenid on välja lülitatud, ei ole eluvõimelised.

Käesolevas töös näitasime, et topoisomeraaside inhibiitorid, sealhulgas etoposiid, mis toimivad pärast DNA lõikamisreaktsiooni fikseerides topoisomeraasi DNA ahela külge ja takistades religatsiooni, tõstavad AIRE transkriptsioonilist efekti suurendades nii aktiveeritud geenide arvu kui ka nende aktivatsiooni taset. Selleks rakendasime kogu transkriptoomi sekveneerimist (RNA-seq), et hinnata geenide ekspressiooni taset neljal erineval moel töödeldud HEK293 rakuliinis, kus AIRE ekspressioon on doksütsikliiniga indutseeritav – indutseerimata rakud (Ctrl), doksütsükliiniga töödeldud ehk AIRE positiivsed rakud (Dox), etoposiidiga töödeldud rakud (Etop) ja nii etoposiidi kui ka doksütsükliiniga töödeldud rakud (Dox+Etop). Vaatamata etoposiidi varem kirjeldatud toksilisele efektile, märkasime tema tugevat transkriptsiooni aktiveerivat efekti AIRE positiivsetele rakkudele: kui Dox rakkudes AIRE poolt oli aktiveeritud 691 geeni, siis etoposiidi ja AIRE kombinatsioon Dox+Etop rakkudes tõstis aktiveeritud geenide arvu 6118-ni. Dox+Etop tingimustes üles reguleeritud geenid omasid sarnaseid jooni AIRE märklaudgeenidega, kaasa arvatud madal algne ekspressiooni tase, koespetsiifilisus ja klasterdumine. Kui Dox+Etop üles reguleeritud geenikogumi sees sisaldusid Dox tingimustes aktiveeritud geenid peaaegu täielikult, siis AIRE (Dox) ja etoposiidi (Etop) poolt eraldi aktiveeritud geenid kattusid väga vähe. Niisugune käitumismuster viitab, et kuigi etoposiid toetab efektiivselt AIRE-vahendatud transkriptsiooni, siis funktsionaalselt on nende tegutsemisviisid erinevad.

Nii etoposiidi kui ka AIRE juuresolekul märkasime alternatiivselt splaissitud geenide suuremat arvu võrreldes töötlemata rakkudega (Ctrl). Siiski, diferentseaalselt ekspresseerunud ja alternatiivselt splaissitud geenide kogumid erinesid, vihjates et nii AIRE kui ka etoposiidi reguleerimismehhanismid on erinevad transkriptsiooni ja splaissingu puhul ja ei toimi samadele transkriptidele.

Võrreldes erinevate töötlustega rakupopulatsioone, leidsime, et AIRE poolt reguleeritud geenidele on omane klasterduda. Ühe raku sekveneerimistööd näitavad et igas klastris saavad ainult üksikud geenid olla võimendunud AIRE-positiivsetes mTEC-des, mis on erinev samade klastrite regulatsioonist teistes differentseerunud kudedes, kus samasse klastrisse kuuluvad geenid saavad avalduda samaaegselt. Siiski on teatud geenilõigud, mille tõenäosus ekspresseeruda on suurem AIRE juuresolekul.

Kinnitamaks, et etoposiidi mõjutab AIRE-vahendatud transkriptsiooni topoisomeraas 2 kaudu, inhibeerisime TOP2A produktsiooni shRNA-ga, mis viis

AIRE märklaudgeenide aktivatsiooni nõrgenemisele. Samuti häiris etoposiidi aktiveerivat efekti L28P mutatsiooni sisseviimine HSR/CARD domeeni, katkestades TOP2A ja AIRE omavahelist või DNA seondumise kaudu toimuvat interaktsiooni.

Kokkuvõttes näitasime käesolevas doktoritöös AIRE võimet aktiveerida ilma geeni ekspressiooni kontrollitavate elementideta plamiidseid reportereid, s.o. promootori, introni või polüadenüleerimisjärjestuste ära jätmine ei mõjutanud negatiivselt AIRE-vahendatud transkriptsiooni aktivatsiooni. Samal ajal detekteerisime AIRE otsese seondumise DNA-ga, mis oli vahendatud aminohapete arginiini 113 ja lüsiini 114 kaudu ja vajas HSR/CARD domeeni terviklikkust. Teiseks näitasime DNA parandamismehhanismide eest vastutavate valkude interakteerumist AIRE-ga ja topoisomeraasi inhibiitorite positiivset mõju nii AIRE poolt vahendavale kui AIRE-st sõltumatule transkriptsioonile. Oleme veendunud, et käesolevas doktoritöös esitatud tulemused aitavad aru saada kuidas AIRE suudab mõjutada nii paljude geenide suhteliselt juhuliku iseloomuga avaldumist tüümuses. Käesolev töö on seega oluline nii immuunsüsteemi funktsioneerimise seisukohalt kuid ka eukarüootse transkriptsiooni kontekstis.

ACKNOWLEDGMENTS

The studies presented in this dissertation were conducted at the Molecular Pathology Research Group in the Institute of Biomedicine and Translational Medicine at the University of Tartu. I want to thank all current and previous members of MolPat lab for creating one of the best work atmospheres I have ever experienced, the aspiration and support from all of you were immense.

I would like to thank my supervisor Pärt Peterson for accepting me in the lab and providing freedom and possibilities to work with different exciting projects. The advice and non-restrictive guidance provided by Pärt were essential.

I want to thank my supervisor Ana Rebane for teaching me many technical skills and coping mechanisms. The importance of negative controls, the effort to apply critical thinking in all demanding situations and the courage to approach problems – were very valuable lessons that I hopefully learned well.

I want to thank senior researchers Kai Kisand and Martti Laan, who were always ready to share their vast knowledge and to provide emotional support, even when not directly involved in experiments. It is hard to overestimate their contribution to lab functioning.

I am very grateful to all my colleagues from other laboratories and workplaces where I have had a chance to work. I believe I learned something new in every place, enriched my experience and created some good memories.

I want to express gratitude to all co-authors of the articles for sharing their wisdom and additional knowledge.

I am thankful to my reviewers Aleksandr Žarkovski and Radko Avi for taking the time to immerse into the thesis manuscript and to approach it critically.

I want to thank the opponent Jorma Palvimo for finding the time in busy schedule to review my thesis and participate in the public examination.

I am very grateful to all my friends who supported me in challenging situations and never skipped the opportunity to remind me about the thesis.

I want to thank my family for tolerating my never-ending studies, accepting my unexpected career choices and supporting me in them.

PUBLICATIONS

CURRICULUM VITAE

Name:	Julia Maslovskaja
Date of birth:	November 6th, 1982
Address:	Molecular Pathology Research Group Institute of Biomedicine and Translational Medicine University of Tartu Ravila 19, 50411 Tartu Estonia
Telephone:	+372 737 4419
E-mail:	julia.maslovskaja@ut.ee

Education

2014–...	University of Tartu, Medicine
2007–…	University of Tartu, Medicine (Immunology), Doctoral Studies
2005–2007	University of Tartu, Biomedicine, M.Sc.
2000–2004	University of Tartu, Gene Technology (Molecular Diagnostics), B.Sc.
1989–2000	Jõhvi Russian Gymnasium

Professional employment

2012–2016	University of Tartu, Molecular Pathology Research Group, Laboratory Specialist
01.2012–04.2012	Swiss Institute of Allergy and Asthma Research, Davos, Switzerland, Trainee
2011–2012	Icosagen Cell Factory OÜ, Researcher
2005–2011	University of Tartu, Molecular Pathology Research Group, Laboratory Specialist
2004–2005	Estonian Biocentre, Senior Laboratory Technician

Publications

1. Vaher, H., Kivihall, A., Runnel, T., Raam, L., Prans, E., Maslovskaja, J., Abram, K., Kaldvee, B., Mrowietz, U., Weidinger, S., Kingo, K., Rebane, A., 2019. SERPINB2 and miR-146a/b are coordinately regulated and act in the suppression of psoriasis-associated inflammatory responses in keratinocytes. Exp. Dermatol. exd.14049.
2. Vaher, H., Runnel, T., Urgard, E., Aab, A., Carreras Badosa, G., Maslovskaja, J., Abram, K., Raam, L., Kaldvee, B., Annilo, T., Tkaczyk, E.R., Maimets, T., Akdis, C.A., Kingo, K., Rebane, A., 2019. miR-10a-5p is increased in atopic dermatitis and has capacity to inhibit keratinocyte proliferation. Allergy Eur. J. Allergy Clin. Immunol. 2146–2156.

3. Hermann, H., Runnel, T., Aab, A., Baurecht, H., Rodriguez, E., Magilnick, N., Urgard, E., Šahmatova, L., Prans, E., Maslovskaja, J., Abram, K., Karelson, M., Kaldvee, B., Reemann, P., Haljasorg, U., Rückert, B., Wawrzyniak, P., Weichenthal, M., Mrowietz, U., Franke, A., Gieger, C., Barker, J., Trembath, R., Tsoi, L.C., Elder, J.T., Tkaczyk, E.R., Kisand, K., Peterson, P., Kingo, K., Boldin, M., Weidinger, S., Akdis, C.A., Rebane, A., 2017. miR-146b Probably Assists miRNA-146a in the Suppression of Keratinocyte Proliferation and Inflammatory Responses in Psoriasis. J. Invest. Dermatol. 137, 1945–1954.
4. Guha, M., Saare, M., Maslovskaja, J., Kisand, K., Liiv, I., Haljasorg, U., Tasa, T., Metspalu, A., Milani, L., Peterson, P., 2017. DNA breaks and chromatin structural changes enhance the transcription of autoimmune regulator target genes. J. Biol. Chem. 292, 6542–6554.
5. Maslovskaja, J., Saare, M., Liiv, I., Rebane, A., Peterson, P.P., 2015. Extended HSR/CARD domain mediates AIRE binding to DNA. Biochem. Biophys. Res. Commun. 468, 913–920.
6. Haljasorg, U., Bichele, R., Saare, M., Guha, M., Maslovskaja, J., Kõnd, K., Remm, A., Pihlap, M., Tomson, L., Kisand, K., Laan, M., Peterson, P., 2015. A highly conserved NF-κB-responsive enhancer is critical for thymic expression of Aire in mice. Eur. J. Immunol. 45, 3246–3256.
7. Rebane, A., Runnel, T., Aab, A., Maslovskaja, J., Rückert, B., Zimmermann, M., Plaas, M., Kärner, J., Treis, A., Pihlap, M., Haljasorg, U., Hermann, H., Nagy, N., Kemeny, L., Erm, T., Kingo, K., Li, M., Boldin, M.P., Akdis, C.A., 2014. MicroRNA-146a alleviates chronic skin inflammation in atopic dermatitis through suppression of innate immune responses in keratinocytes. J. Allergy Clin. Immunol. 134, 836-847.
8. Kärner, J., Meager, A., Laan, M., Maslovskaja, J., Pihlap, M., Remm, A., Juronen, E., Wolff, A.S.B., Husebye, E.S., Podkrajšek, K.T., Bratanic, N., Battelino, T., Willcox, N., Peterson, P., Kisand, K., 2013. Anti-cytokine autoantibodies suggest pathogenetic links with autoimmune regulator deficiency in humans and mice. Clin. Exp. Immunol. 171, 263–272.
9. Arukuusk, P., Pärnaste, L., Oskolkov, N., Copolovici, D.M., Margus, H., Padari, K., Möll, K., Maslovskaja, J., Tegova, R., Kivi, G., Tover, A., Pooga, M., Ustav, M., Langel, Ü., 2013. New generation of efficient peptide-based vectors, NickFects, for the delivery of nucleic acids. Biochim. Biophys. Acta – Biomembr. 1828, 1365–1373.
10. Liiv, I., Haljasorg, U., Kisand, K., Maslovskaja, J., Laan, M., Peterson, P., 2012. AIRE-induced apoptosis is associated with nuclear translocation of stress sensor protein GAPDH. Biochem. Biophys. Res. Commun. 423, 32–37.
11. Liiv, I., Rebane, A., Org, T., Saare, M., Maslovskaja, J., Kisand, K., Juronen, E., Valmu, L., Bottomley, M.J., Kalkkinen, N., Peterson, P., 2008. DNA-PK contributes to the phosphorylation of AIRE: Importance in transcriptional activity. Biochim. Biophys. Acta – Mol. Cell Res. 1783, 74–83.

ELULOOKIRJELDUS

Nimi: Julia Maslovskaja
Sünniaeg: 6. november, 1982. a.
Aadress: Molekulaarpatoloogia uurimisrühm
Bio- ja siirdemeditsiini instituut
Tartu Ülikool
Ravila 19, 50411 Tartu
Eesti
Telefon: +372 737 4419
E-post: julia.maslovskaja@ut.ee

Hariduskäik

2014–...	Tartu Ülikool, arstiteadus
2007–…	Tartu Ülikool, arstiteadus (immunoloogia), doktorantuur
2005–2007	Tartu Ülikool, biomeditsiin, M.Sc.
2000–2004	Tartu Ülikool, geenitehnoloogia (molekulaardiagnostika), B.Sc.
1989–2000	Jõhvi Vene Gümnaasium

Teenistuskäik

2012–2016	Tartu Ülikool, molekulaarpatoloogia uurimisgrupp, laborispetsialist
01.2012–04.2012	Šveitsi allergia ja astma instituut, Davos, Šveits, külalistudeng
2011–2012	Icosagen Cell Factory OÜ, teadur
2005–2011	Tartu Ülikool, molekulaarpatoloogia uurimisgrupp, laborispetsialist
2004–2005	Eesti Biokeskus, vanemlaborant

Publikatsioonid

1. Vaher, H., Kivihall, A., Runnel, T., Raam, L., Prans, E., Maslovskaja, J., Abram, K., Kaldvee, B., Mrowietz, U., Weidinger, S., Kingo, K., Rebane, A., 2019. SERPINB2 and miR-146a/b are coordinately regulated and act in the suppression of psoriasis-associated inflammatory responses in keratinocytes. Exp. Dermatol. exd.14049.
2. Vaher, H., Runnel, T., Urgard, E., Aab, A., Carreras Badosa, G., Maslovskaja, J., Abram, K., Raam, L., Kaldvee, B., Annilo, T., Tkaczyk, E.R., Maimets, T., Akdis, C.A., Kingo, K., Rebane, A., 2019. miR-10a-5p is increased in atopic dermatitis and has capacity to inhibit keratinocyte proliferation. Allergy Eur. J. Allergy Clin. Immunol. 2146–2156.
3. Hermann, H., Runnel, T., Aab, A., Baurecht, H., Rodriguez, E., Magilnick, N., Urgard, E., Šahmatova, L., Prans, E., Maslovskaja, J., Abram, K.,

Karelson, M., Kaldvee, B., Reemann, P., Haljasorg, U., Rückert, B., Wawrzyniak, P., Weichenthal, M., Mrowietz, U., Franke, A., Gieger, C., Barker, J., Trembath, R., Tsoi, L.C., Elder, J.T., Tkaczyk, E.R., Kisand, K., Peterson, P., Kingo, K., Boldin, M., Weidinger, S., Akdis, C.A., Rebane, A., 2017. miR-146b Probably Assists miRNA-146a in the Suppression of Keratinocyte Proliferation and Inflammatory Responses in Psoriasis. J. Invest. Dermatol. 137, 1945–1954.

4. Guha, M., Saare, M., Maslovskaja, J., Kisand, K., Liiv, I., Haljasorg, U., Tasa, T., Metspalu, A., Milani, L., Peterson, P., 2017. DNA breaks and chromatin structural changes enhance the transcription of autoimmune regulator target genes. J. Biol. Chem. 292, 6542–6554.
5. Maslovskaja, J., Saare, M., Liiv, I., Rebane, A., Peterson, P.P., 2015. Extended HSR/CARD domain mediates AIRE binding to DNA. Biochem. Biophys. Res. Commun. 468, 913–920.
6. Haljasorg, U., Bichele, R., Saare, M., Guha, M., Maslovskaja, J., Kõnd, K., Remm, A., Pihlap, M., Tomson, L., Kisand, K., Laan, M., Peterson, P., 2015. A highly conserved NF-κB-responsive enhancer is critical for thymic expression of Aire in mice. Eur. J. Immunol. 45, 3246–3256.
7. Rebane, A., Runnel, T., Aab, A., Maslovskaja, J., Rückert, B., Zimmermann, M., Plaas, M., Kärner, J., Treis, A., Pihlap, M., Haljasorg, U., Hermann, H., Nagy, N., Kemeny, L., Erm, T., Kingo, K., Li, M., Boldin, M.P., Akdis, C.A., 2014. MicroRNA-146a alleviates chronic skin inflammation in atopic dermatitis through suppression of innate immune responses in keratinocytes. J. Allergy Clin. Immunol. 134, 836–847.
8. Kärner, J., Meager, A., Laan, M., Maslovskaja, J., Pihlap, M., Remm, A., Juronen, E., Wolff, A.S.B., Husebye, E.S., Podkrajšek, K.T., Bratanic, N., Battelino, T., Willcox, N., Peterson, P., Kisand, K., 2013. Anti-cytokine autoantibodies suggest pathogenetic links with autoimmune regulator deficiency in humans and mice. Clin. Exp. Immunol. 171, 263–272.
9. Arukuusk, P., Pärnaste, L., Oskolkov, N., Copolovici, D.M., Margus, H., Padari, K., Möll, K., Maslovskaja, J., Tegova, R., Kivi, G., Tover, A., Pooga, M., Ustav, M., Langel, Ü., 2013. New generation of efficient peptide-based vectors, NickFects, for the delivery of nucleic acids. Biochim. Biophys. Acta – Biomembr. 1828, 1365–1373.
10. Liiv, I., Haljasorg, U., Kisand, K., Maslovskaja, J., Laan, M., Peterson, P., 2012. AIRE-induced apoptosis is associated with nuclear translocation of stress sensor protein GAPDH. Biochem. Biophys. Res. Commun. 423, 32–37.
11. Liiv, I., Rebane, A., Org, T., Saare, M., Maslovskaja, J., Kisand, K., Juronen, E., Valmu, L., Bottomley, M.J., Kalkkinen, N., Peterson, P., 2008. DNA-PK contributes to the phosphorylation of AIRE: Importance in transcriptional activity. Biochim. Biophys. Acta - Mol. Cell Res. 1783, 74–83.

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 osteoarthritis. Tartu, 2017, 150 p.
266. **Mario Saare**. The influence of AIRE on gene expression – studies of transcriptional regulatory mechanisms in cell culture systems. Tartu, 2017, 172 p.
267. **Piia Jõgi**. Epidemiological and clinical characteristics of pertussis in Estonia. Tartu, 2018, 168 p.
268. **Elle Põldoja**. Structure and blood supply of the superior part of the shoulder joint capsule. Tartu, 2018, 116 p.
269. **Minh Son Nguyen**. Oral health status and prevalence of temporomandibular disorders in 65–74-year-olds in Vietnam. Tartu, 2018, 182 p.
270. **Kristian Semjonov**. Development of pharmaceutical quench-cooled molten and melt-electrospun solid dispersions for poorly water-soluble indomethacin. Tartu, 2018, 125 p**.**
271. **Janne Tiigimäe-Saar**. Botulinum neurotoxin type A treatment for sialorrhea in central nervous system diseases. Tartu, 2018, 109 p.
272. **Veiko Vengerfeldt.** Apical periodontitis: prevalence and etiopathogenetic aspects. Tartu, 2018, 150 p.
273. **Rudolf Bichele.** TNF superfamily and AIRE at the crossroads of thymic differentiation and host protection against *Candida albicans* infection. Tartu, 2018, 153 p.
274. **Olga Tšuiko**. Unravelling Chromosomal Instability in Mammalian Preimplantation Embryos Using Single-Cell Genomics. Tartu, 2018, 169 p.
275. **Kärt Kriisa.** Profile of acylcarnitines, inflammation and oxidative stress in first-episode psychosis before and after antipsychotic treatment. Tartu, 2018, 145 p.
276. **Xuan Dung Ho.** Characterization of the genomic profile of osteosarcoma. Tartu, 2018, 144 p.
277. **Karit Reinson.** New Diagnostic Methods for Early Detection of Inborn Errors of Metabolism in Estonia. Tartu, 2018, 201 p.

278. **Mari-Anne Vals.** Congenital N-glycosylation Disorders in Estonia. Tartu, 2019, 148 p.
279. **Liis Kadastik-Eerme**. Parkinson's disease in Estonia: epidemiology, quality of life, clinical characteristics and pharmacotherapy. Tartu, 2019, 202 p.
280. **Hedi Hunt**. Precision targeting of intraperitoneal tumors with peptide-guided nanocarriers. Tartu, 2019, 179 p.
281. **Rando Porosk**. The role of oxidative stress in Wolfram syndrome 1 and hypothermia. Tartu, 2019, 123 p.
282. **Ene-Ly Jõgeda**. The influence of coinfections and host genetic factor on the susceptibility to HIV infection among people who inject drugs. Tartu, 2019, 126 p.
283. **Kristel Ehala-Aleksejev**. The associations between body composition, obesity and obesity-related health and lifestyle conditions with male reproductive function. Tartu, 2019, 138 p.
284. **Aigar Ottas**. The metabolomic profiling of psoriasis, atopic dermatitis and atherosclerosis. Tartu, 2019, 136 p.
285. **Elmira Gurbanova.** Specific characteristics of tuberculosis in low default, but high multidrug–resistance prison setting. Tartu, 2019, 129 p.
286. **Van Thai Nguyeni.** The first study of the treatment outcomes of patients with cleft lip and palate in Central Vietnam. Tartu, 2019, 144 p.
287. **Maria Yakoreva**. Imprinting Disorders in Estonia. Tartu, 2019, 187 p.
288. **Kadri Rekker**. The putative role of microRNAs in endometriosis pathogenesis and potential in diagnostics. Tartu, 2019, 140 p.
289. **Ülle Võhma**. Association between personality traits, clinical characteristics and pharmacological treatment response in panic disorder. Tartu, 2019, 121 p.
290. **Aet Saar**. Acute myocardial infarction in Estonia 2001–2014: towards risk-based prevention and management. Tartu, 2019, 124 p.
291. **Toomas Toomsoo**. Transcranial brain sonography in the Estonian cohort of Parkinson's disease. Tartu, 2019, 114 p.
292. **Lidiia Zhytnik**. Inter- and intrafamilial diversity based on genotype and phenotype correlations of Osteogenesis Imperfecta. Tartu, 2019, 224 p.
293. **Pilleriin Soodla.** Newly HIV-infected people in Estonia: estimation of incidence and transmitted drug resistance. Tartu, 2019, 194 p.
294. **Kristiina Ojamaa.** Epidemiology of gynecological cancer in Estonia. Tartu, 2020, 133 p.
295. **Marianne Saard.** Modern Cognitive and Social Intervention Techniques in Paediatric Neurorehabilitation for Children with Acquired Brain Injury. Tartu, 2020, 168 p.