Epigenetic signals that direct cell type specific interferon beta response in mouse cells

Dissertation

submitted by Markus Muckenhuber

2020

Dissertation

submitted to the
Combined Faculty of Natural Sciences and Mathematics
of the Ruperto Carola University Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by
Markus Muckenhuber, M.Sc.
Born in Steyr, Austria
Oral examination: 14.12.2020

Epigenetic signals that direct cell type specific interferon beta response in mouse cells

Referees: Prof. Dr. Karsten Rippe
Dr. Steeve Boulant

This work was performed from September 2016 to December 2020 under the supervision of Prof. Dr. Karsten Rippe in the Division Chromatin Networks at the DKFZ and the Bioquant Center in Heidelberg, Germany.

Declaration

I hereby declare that I have written the submitted dissertation “Epigenetic signals that direct cell type specific interferon beta response in mouse cells” myself and in this process, have used no other sources or materials than those explicitly indicated. I hereby declare that I have not applied to be examined at any other institution, nor have I used the dissertation in this or any other form at any other institution as an examination paper, nor submitted it to any other faculty as a dissertation.

_____ _____

Place, Date Markus Muckenhuber

Table of contents

Summary

Interferon beta (IFNβ) triggers the JAK-STAT signaling cascade to induce IFN-stimulated genes (ISGs), which is a hallmark of innate immune response against viral infections. The transcription factors STAT1 and STAT2 become activated and, together with IRF9, assemble into the ISGF3 complex. This complex translocates to the nucleus and activates ISGs by binding to its DNA recognition motif. Most cell types have the potential to activate ISGs upon IFNs signaling but embryonic stem cells (ESC) have an attenuated response compared to differentiated cells. However, the exact molecular mechanisms that drive this cell type specific interferon signaling are poorly characterized.

In this thesis, the cell type specific IFNβ response was compared between mouse ESCs and differentiated cells like mouse embryonic fibroblasts (MEFs) that carry the same genome. I tested the hypothesis that the cell type specific differences in IFNβ response originate from distinct epigenetic states by applying a genome-wide multiomics approach: (i) A differential gene expression analysis by RNA sequencing (RNA-seq) of IFNβ stimulation defined a total of 513 ISGs and allowed it to identify cell type specific ISG signatures. The bulk sequencing analysis was complemented with single cell RNA-seq to resolve heterogeneity of gene expression response. (ii) By TF chromatin immunoprecipitation followed by sequencing (ChIP-seq) the STAT1 and STAT2 binding sites were mapped across cell types. (iii) Active chromatin regions were detected with the assay for transposase accessible chromatin with high-throughput sequencing (ATAC-seq). Single cell ATAC-seq was used to identify coregulated enhancers and promoters. (iv) ChIP-seq of histone acetylation (ac) and mono- and tri-methylation (me1, me3) marks at histone lysine residues for H3K4me1, H3K4me3, H3K9ac, H3K27ac, H3K9me3 and H3K27me3 was conducted.

The analysis of this comprehensive data set yielded cell type specific patterns of ISGs, ISGF3 binding and chromatin features. The overall stronger IFNβ response in MEFs could be rationalized by factors from the JAK/STAT signaling cascade being constitutively more strongly expressed. In addition, 33 ISGs in ESCs and 305 ISGs in MEFs were found to be cell type specific and thus candidates for epigenetic regulation. To characterize the underlying mechanism, the genomic location, chromatin context and target genes of ISGF3 were characterized. While 92 ISGF3 sites were shared between ESCs and MEFs, 116 and 184 sites were specific for one cell type and found at promoters and putative enhancers. Based on a co-regulation analysis of single cell ATAC-seq data, many of these enhancers could be linked to specific ISGs.

Furthermore, the analysis revealed that a pre-existing enrichment of H3K4me1 and open chromatin loci at ISGF3 sites was positively correlated with ISGF3 binding while H3K27me3 showed the opposite effect.

In summary, this thesis characterizes the contribution of epigenetic gene regulation mechanisms to the cell type specific IFNβ response and rationalizes how chromatin features direct cell type specific ISGF3 binding. The insight gained opens up new possibilities for targeted interference with interferon response in anti-viral drug development by accounting for the contribution of chromatin to this process.

Zusammenfassung

Interferon beta (IFNβ) führt zur Aktivierung der JAK-STAT-Signalkaskade, die IFN-stimulierte Gene (ISGs) aktiviert und einen Teil der angeborenen Immunantwort auf Infektionen mit Viren darstellt. Dabei werden die Transkriptionsfaktoren STAT1 und STAT2 aktiviert und bilden zusammen mit IRF9 den ISGF3-Komplex. Dieser Komplex wird in den Zellkern transportiert und stimuliert dort die Aktivierung der ISGs, indem er an sein DNA-Erkennungsmotiv bindet. Die meisten Zelltypen haben das Potenzial, ISGs auf IFNs-Signale hin zu aktivieren. Embryonale Stammzellen (ESC) hingegen zeigen eine abgeschwächte Antwort verglichen mit differenzierten Zellen. Die genauen molekularen Mechanismen, welche diese zelltypspezifischen Unterschiede im Interferon-Signalweg steuern, sind nur unzureichend charakterisiert.

In dieser Arbeit wurde die zelltypspezifische IFNβ Antwort zwischen ESCs aus Maus und differenzierten Zellen wie embryonalen Fibroblasten (MEFs) der Maus verglichen. Essentiell ist dabei, dass alle untersuchten Zellen das gleiche Genom haben. Daher überprüfte ich die Hypothese, dass die zelltypspezifischen Unterschiede in der IFNβ Antwort auf Unterschiede in den epigenetischen Signaturen zurückzuführen sind. Dabei wurde ein genomweiter Multiomik-Ansatz verwendet, um folgende Punkt zu untersuchen: (i) Eine differentielle Genexpressionsanalyse in Folge von IFNβ Stimulation wurde mittels RNA-Sequenzierung (RNA-seq) durchgeführt und identifizierte insgesamt 513 ISGs. Des Weiteren konnten zelltypspezifische ISG-Signaturen identifiziert werden. Die Bulk-Sequenzierungsanalyse wurde durch RNA-seq von einzelnen Zellen bestätigt. Zusätzlich wurde die Frage der Heterogenität der Genexpressionsantwort in Folge einer IFNβ Stimulation aufzulösen. (ii) Durch Chromatin-Immunpräzipitation mit anschließender Sequenzierung (ChIP-seq) wurden die STAT1- und STAT2-Bindungsstellen in ESCs und MEFs kartiert. (iii) Aktive Chromatinregionen wurden mit dem Assay für Transposase-zugängliches Chromatin mit Hochdurchsatz-Sequenzierung (ATAC-seq) nachgewiesen. Zusätzlich, wurde Einzelzell-ATAC-seq verwendet, um regulatorische Enhancer mit ihren potenziellen Zielpromotoren über mehrere hundert Kilobasen hinweg zu assoziieren. (iv) Außerdem, wurden durch ChIP-seq Histon-acetylierung (ac) und Mono- und Tri-Methylierung (me1, me3) an Histon-Lysinresten für H3K4me1, H3K4me3, H3K9ac, H3K27ac, H3K9me3 und H3K27me3 bestimmt und genutzt, um zelltypspezifische ISGF3 Bindungsstellen zu charakterisieren.

Die Analyse dieses umfassenden Datensatzes ergab zelltypspezifische Muster von ISGs, ISGF3-Bindung und Chromatinmerkmalen. Die insgesamt stärkste IFNβ

induzierte Aktivierung von ISGs wurde in MEFs beobachtet und durch erhöhte Aktivität von Faktoren der JAK/STAT-Signalkaskade erklärt. Darüber hinaus erwiesen sich 33 ISGs in ESCs und 305 ISGs in MEFs als zelltypspezifisch und damit als Kandidaten für eine mögliche epigenetische Regulation. Um den zugrundeliegenden Mechanismus zu charakterisieren, wurden die genomischen Regionen, deren Chromatinkontext und die Zielgene von ISGF3 charakterisiert. Während 92 ISGF3-Stellen in ESCs und MEFs gebunden wurden, waren 116 nur in ESCs und 184 nur in MEFs gebunden. Diese zelltypespezifischen Bindungsstellen fanden sich an Promotoren und überwiegend an mutmaßlichen Enhancern. Auf der Grundlage einer Koregulationsanalyse von Einzelzell-ATAC-seq-Daten konnten viele dieser Enhancer mit aktivierten ISGs verknüpft werden. Darüber hinaus ergab die Analyse, dass eine bereits vorhandene Anreicherung von H3K4me1 und zugänglichem Chromatin an ISGF3-Bindingstellen positiv mit der ISGF3-Bindung korreliert war, während H3K27me3 den gegenteiligen Effekt zeigte.

Zusammenfassend charakterisiert diese Arbeit den Beitrag der epigenetischen Genregulationsmechanismen zur zelltypspezifischen IFNβ Antwort und erklärt, welchen Beitrag Chromatin zu einer direkten zelltypspezifischen ISGF3-Bindung hat. Die gewonnenen Erkenntnisse eröffnen neue Möglichkeiten zur gezielten Beeinflussung der Interferonantwort bei der Entwicklung antiviraler Medikamente, indem der Beitrag des Chromatins zu diesem Prozess berücksichtigt wird.

List of figures

List of tables

Abbreviations

Abbreviation	*Full name*
∞	infinity
3C	chromatin conformation capture
3D	three dimensional
AB	antibody
ATAC	assay for transposase-accessible chromatin
ATAC-seq	ATAC-sequencing
BHK	baby hamster kidney
BSA	bovine serum albumin
CAs	cellular aggregates
CCL2	c-c motif chemokine ligand 2
CCND2	cyclin D2
CDK8	cyclin-dependent kinase 8
ChIP	chromatin immunoprecipitation
ChIP-seq	ChIP-sequencing
CO_2	carbon dioxide
CRISPR-Cas9	clustered regularly interspaced short palindromic repeats - caspase 9
dCD	double catalytic dead
dKO	double knock out
DNA	deoxyribonucleic acid
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E6/13/19	mouse embryonic day 6/13/19
EB	elution buffer
EDTA	ethylenediaminetetraacetic acid
ESCs	murine embryonic stem cells
EtOH	ethanol
FCS/FBS	fetal calf serum / fetal bovine serum
FPKM	fragments per kilobase million
FRiP	fraction of reads in peaks
GAS	gamma-interferon-activation sites
GBP6	guanylate binding protein family member 6
gDNA	genomic DNA
GO-term	gene ontology term
GTF	general transfer format
H_2O	water
H3	histone 3
H3K27ac	histone 3 with acetylation at the 27th lysine residue
H3K27me3	histone 3 with tri-methylation at the 27th lysine residue
H3K36me3	histone 3 with tri-methylation at the 36th lysine residue
H3K4me1	histone 3 with mono-methylation at the 4th lysine residue
H3K4me3	histone 3 with tri-methylation at the 4th lysine residue
H3K9ac	histone 3 with acetylation at the 9th lysine residue

H3K9me3	histone 3 with tri-methylation at the 9th lysine residue
HAT	histone acetyl transferases
HCl	hydrochloric acid
HDAC	histone deacetylases
IFI205	interferon-activable protein 205-A
IFI27	interferon alpha inducible protein 27
IFN	interferon
IFNAR1/2	interferon alpha and beta receptor subunit 1/2
IFNGR1/2	interferon gamma receptor 1/2
IFNβ	interferon beta
IgG	immunoglobulin G
IL10RA/B	interleukin 10 receptor subunit alpha/ beta
IP	immunoprecipitation
IRF1/2/3/7/8/9	interferon regulatory factor 1/2/3/7/8/9
ISG	IFN-stimulated genes
ISGF3	IFN-stimulated gene factor 3
ISRE	interferon stimulated response element
JAK1/2	Janus kinase 1/2
K	lysine
KDM	histone lysine dimethyl transferase
KMT	histone lysine methyl transferase
LACV	La Crosse Virus
LIF	leukemia inhibitory factor
LPS	lipopolysaccharide
MCMV	mouse cytomegalovirus
MEFs	mouse embryonic fibroblasts
MgCl2	magnesium chloride
MLL3/4	mixed-lineage leukemia protein 3/4
mm10	Mus musculus (house mouse) genome assembly GRCm38
mRNA	messenger RNA
MT-PBS	mouse tonicity phosphate buffered saline
MX2	MX dynamin like GTPase 2
N2	nitrogen
NaCl2	sodium chloride
NGS	next-generation sequencing
NPCs	neuronal progenitor cells
NSC	normalized strand coefficient
NSG2	neuronal vesicle trafficking associated 2
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PE	paired end
PIAS	protein inhibitor of activated STAT
PIC	protease inhibitor cocktail
PMSF	phenylmethylsulphonyl fluoride
PTMs	post translational modification
QC	quality control
RA	retinoic acid

Rb	rabbit
RNA	ribonucleic acid
RNA PolII	RNA polymerase II
RNA-seq	RNA-sequencing
rRNA	ribosomal RNA
RSC	relative strand correlation
RT-qPCR	reverse transcriptase quantitative PCR
RTP4	receptor transporter protein 4
scATAC-seq	single cell ATAC-seq
scRNA-seq	single cell RNA-seq
SE	single end
SOCS	suppressors of cytokine signaling
STAT	signal transducer and activator of transcription
STAT1-p727	STAT1 with phosphorylation at tyrosine 701
STAT1-p701	STAT1 with phosphorylation at serine 727
STAT1/2/3/4	signal transducer and activator of transcription 1/2/3/4
TAD	topological associated domain
TFs	transcription factors
Tn5	transposase
TPM	transcripts per kilobase million
TSS	transcription start site
TYK2	tyrosine kinase 2
USP18	ubiquitin specific peptidase 18
WB	western blot

Abbreviation	*Name of unit*
%	percentage
°C	grad Celsius
A	ampere
bp	base pair
g	g(raviation) force
h	hour(s)
kb	kilobase
kDA	kilo dalton
lfc	fold change of log transformed values
M	million or molar
mg/ml	milligram per milliliter
min	minute(s)
ml	milliliter
mm	millimeter
mM	millimolar
ng	nanogram
nM	nanomolar
O/N	over night
pg	picogram
pval	probability value
rpm	rounds per minute
RT	room temperature (20-21°C)
s	second
U	units
U/ml	units per milliliter
U/µl	units per microliter
V	volt
V	volt
µF	microfarad
µg	microgram
µl	microliters
µm	micrometer
µM	micromolar

1. Introduction

1.1 Chromatin and gene regulation

1.1.1. The essential role of gene regulation

The human and mouse genome encodes for around 20,000 proteins that are expressed in specific patterns in hundreds of different cell types (Daniel *et al.* 2014; Roadmap Epigenomics *et al.* 2015). Furthermore, cells have to be able to respond to environmental cues like nutrients or cytokines (**Figure 1**). Such signals can cause the activation of repressed genes or the silencing of active genes. Once the stimulus is no longer present, the cell's gene expression program can shift back to its initial state or self-maintain the changes via epigenetic mechanisms.

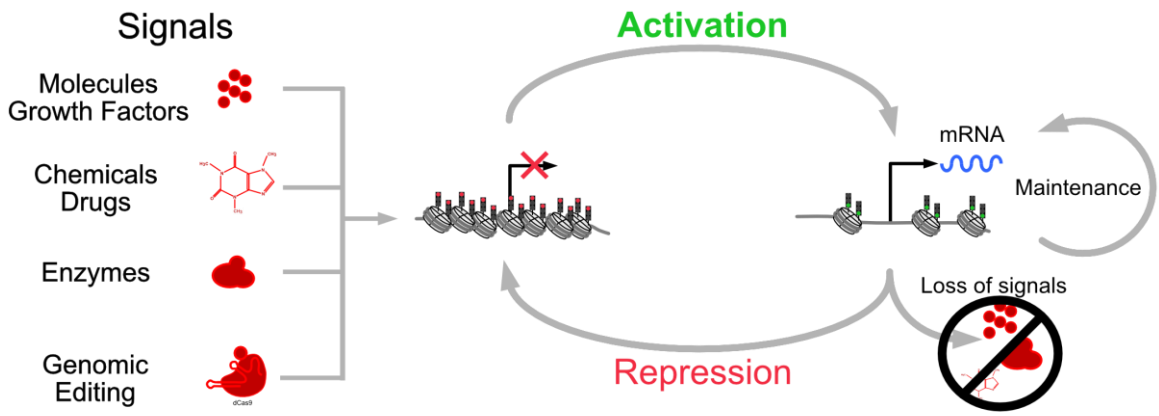

Figure 1: Signals that activate or repress gene expression

Signal molecules like growth factors or interferons, chemicals and drugs are external signals induce changes in gene expression patterns. The activation of specific enzymes or ectopic perturbances like genomic editing can change gene expression in the cell. These signals can cause their target genes to become activated or repressed. After the stimulus is lost either the activated state is maintained via epigenetic mechanisms without further stimulation or the system reverts back to its initial state.

Accordingly, the regulation of gene expression is essential for both the development of the organism and also for its adaption to a given environment. The binding of gene specific transcription factors (TFs) creates an activating or inhibiting environment for transcription initiation by RNA polymerase II (RNAP II), which transcribes messenger RNA (mRNA) of protein coding genes. This level of gene regulation also involves distal regulatory elements called enhancers, which can regulate transcription from a distance of up to megabases away from their target genes (Daniel *et al.* 2014). In addition, gene expression levels can be altered also at the level of mRNA stability

(Liu *et al.* 2014). The precise regulation of transcription is essential for development and survival of an organism and involves a variety of molecular mechanisms.

1.1.2. Chromatin as a genome organizer

An important aspect of gene expression regulation is the organization of the genome into chromatin to control both its compaction and accessibility (**Figure 2A**). The nucleus of a single mammalian cell contains genomic DNA with a total length of around two meters in a volume of around 400 μm^3 (Maul and Deaven 1977; Misteli 2008). In mice, the total DNA is encoded in 20 pairs of chromosomes. Each chromosome is built up a complex mixture of various protein and RNA factor that assemble around the DNA. The basic unit is built up from two copies each of the core histones H2A, H2B, H3 and H4 called histone octamer. DNA is wrapped around this structure in 1.67 left-handed turns with a length of 147 nucleotides followed by about 50 bp of linker DNA between nucleosomes (Luger *et al.* 1997). The combination of the histone octamer and the surrounding DNA is called nucleosome (**Figure 2B**). Non-canonical histone variants like H2A.Z or H3.3 can replace the core histones and provide additional levels of regulation and are enriched in specific genomic regions (Au-Yeung and Horvath 2018) or cell types (Maze *et al.* 2015). Classically, the density of nucleosomes was used to distinguish between packed and loose chromatin. A modern way to classify genomic regions is based on pattern of post-translational modifications of the N-terminal arm of histones (Kouzarides 2007) so called chromatin states (Roadmap Epigenomics *et al.* 2015) (**Figure 2C**). Repressed chromatin states are enriched for marks like H3K27me3 and H3K9me3 and are often associated with no gene expression (**Figure 2D**). Active chromatin states contain actively transcribed promoters and are associated with marks as H3K4me3, H3K9ac and H3K27ac (Bernstein *et al.* 2005) (**Figure 2E/F**). In addition to methylation and acetylation, there are various other modifications presents on histone tails like phosphorylation, ubiquitinylation or sumoylation (Kebede *et al.* 2015; Kouzarides 2007). Promoters make up around 1-3 % of a mammalian genome and the remaining sites are often harboring regulatory elements. The most prominent of these elements are enhancers, which are enriched for H3K4me1 and H3K27ac in the active state (**Figure 2G**). Additionally, these sites have high levels of chromatin accessibility and can be be bound by various TFs. The binding patterns of these TFs contribute to gene expression. On top of that, chromatin is organized in higher level structures, which defined structural domains called topological associated domains (TAD) (**Figure 2H**) (Yu and Ren 2017). Gene regulation often happen within these domains.

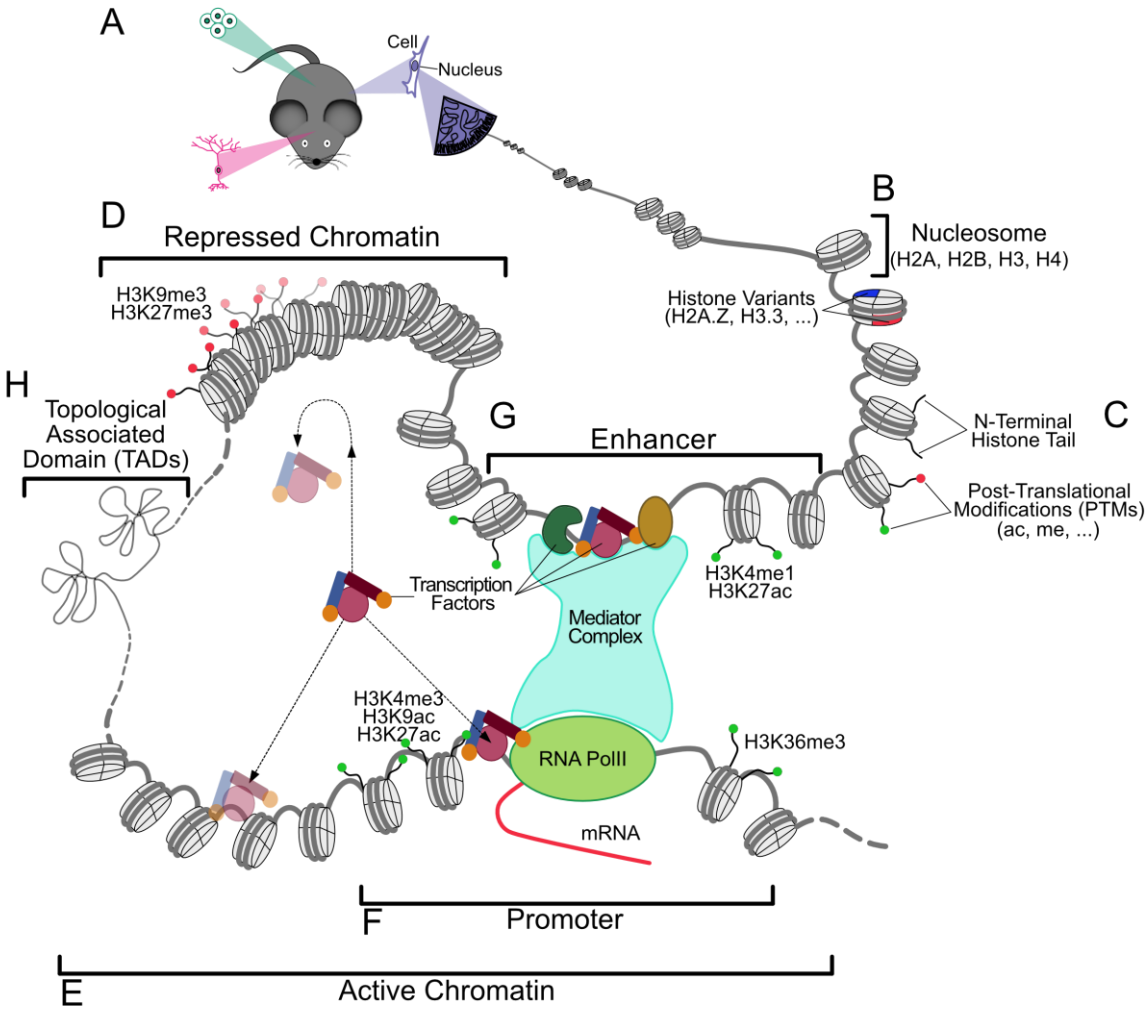

Figure 2: Hierarchical layers of chromatin organization in mouse cells

(A) Visualization of multiple diverse cell types in mouse and a zoom in into the nucleus of one cell. (B) The smallest unit of organization is a complex of DNA with four pairs of histone proteins, H2A, H2B, H3 and H4, called the nucleosome. In addition to the classical canonic histones, multiple histone variants exist like H2A.Z or H3.3. (C) The N-terminal tail of histones can be post-transcriptional modified (PTMs) with methylation, acetylation or other modifications. (D) A high density of nucleosomes forms not accessible, dense chromatin regions defined as repressed chromatin state. The post-translational modifications H3K9me3 and H3K27me3 are located at these genomic sites and the binding of normal transcriptions factors (TF) is inhibited. (E) In contrast, open, accessible genomic loci are called active chromatin states and provide binding sites for TFs. Active promoters and enhancers are examples for these sites. (F) Active promoters are bound by TFs and RNA Polymerase II (RNA PolII), which transcribes genes into mRNA. The histone marks H3K4m3 is found at promoters in general, while the H3K27ac mark allows to distinguish active from poised promoters. (G) Enhancer are labeled with H3K4me1 and H3K27ac when activated. Active enhancers are binding platforms for TFs and via the mediator complex, links to target promoters are established to promote gene induction. (H) Chromatin is organized into higher chromatin structures called topological associated domains (TAD). Figure was inspired by (Aranda *et al.* 2015).

1.1.3. Transcription factors as initiators of gene expression changes

Combinations of histone modifications create chromatin environments to promote the binding of specific TFs (Fulton *et al.* 2009). The basic idea of TFs is that domains of these proteins are able to stably bind to their recognition DNA sequence called motif. Upon their binding to the 6-12bp motif, cofactors, like histone acetyl transferases (HATs), mediator complex or additional TFs, are recruited to induce gene expression. Between 1,500 and 2,765 potential TFs are predicate in the mammalian genome (Lambert *et al.* 2018; Vaquerizas *et al.* 2009). TFs are grouped by their DNA binding domain and many of these are very specific, like the C2H2-zinc finger, homeodomain, basic helix-loop-helix and basic leucine zipper (Johnson and McKnight 1989). In addition, there are also simple and widespread domains used to interact with DNA, like the AT-hook (Aravind and Landsman 1998). The last domain can be found in highly conserved TFs like the STAT and IRF families.

The potential of TFs is shown by single TFs driving the differentiation of cell types (Fong and Tapscott 2013) or reprogramming differentiated cells back into pluripotent stem cells (Takahashi and Yamanaka 2006). The ability to bind can be influenced by the nucleosome occupancy of a potential binding sites (Teif *et al.* 2013). Thereby the binding site can be protected by nucleosomes, the binding motif is not recognized by a TF and the binding is inhibited. Consequently, target genes are not activated and so the accessibility of binding motifs is essential to control gene expression profiles. Recent evidence identified specific TFs as pioneering TFs (Zaret and Carroll 2011). The specific ability of these factors is to bind their motifs even when the DNA sequence is protected by a nucleosome. Such sites are often in sites with repressed chromatin states. As a consequent, these pioneering TFs are creating open chromatin regions and make them accessible for the second wave of TFs (Zaret and Mango 2016). Pioneer TFs are key players to initiate the differentiation of cells (Iwafuchi-Doi and Zaret 2014) and to enhance the response in differentiated cells in response to external signals like IFN (Ghisletti *et al.* 2010). Binding sites for TFs are not randomly distributed over the genome, specific genomic sites are highly enriched for such motifs. These sites are often promoters and enhancers, which are classified based on the combination of histone modifications they harbor.

1.1.4. Distinct chromatin states are defined by histone modifications

At the beginning of the 2000s in addition to the classical genetic code, a "Histone code" was proposed (Jenuwein and Allis 2001). The development of techniques like Chromatin Immunoprecipitation followed by Next Generation Sequencing (ChIP-Seq) allowed to characterize histone marks genome-wide and establish a code of histone modification patterns to identify genomic regions of interest on a new level (Roadmap Epigenomics *et al.* 2015).

The most prominent and best studied genomic regions are regions upstream of the transcription start site of genes called promoter (**Figure 2F**) (Juven-Gershon and Kadonaga 2010; Smale and Kadonaga 2003). A mammalian promoter is up to 1,000 base pairs (bp) in size and the DNA sequence in this region has binding sites for general TFs which are then recruiting RNAP II. The most specific histone mark at promoter regions is H3K4me3 found independent of the activation state of a promoter (Barski *et al.* 2007; Bernstein *et al.* 2005; Santos-Rosa *et al.* 2002). In combination with the histone mark H3K27ac and H3K9ac, it is strongly correlating with active transcription. In a developmental context, the combination of H3K4me3 and the repressive mark H3K27me3 are defining bivalent promoters, which were identified as critical for mammalian differentiation (Barski *et al.* 2007). Depending on the developmental path of the cell, some of these promoters become active, while others are permanently silenced by removing the H3K4me3 mark (Karmodiya *et al.* 2012).

An essential part of gene expression are regulatory elements, so called enhancers (**Figure 2G**). In the mouse genome around 300,000 cis-regulatory elements are annotated (Shen *et al.* 2012). These elements are up to few hundred nucleotides in length and contain specific DNA sequences called motifs which are binding platforms for TF binding (Spitz and Furlong 2012). The combination of TF binding orchestra the expression patterns of their target promoters by enhancing their transcriptional output (Higgs 2020; Juven-Gershon and Kadonaga 2010). These interactions can span megabases of nucleotides (Lettice *et al.* 2003). During the differentiation of cell types, regulatory elements are essential to direct gene expression patterns. Consequently, many enhancers are activated in a cell type specific manner (Furlong and Levine 2018). The link between active enhancer and a target promoter is established by the multi-subunit protein complex mediator of RNA polymerase II transcription (Mediator) complex (Soutourina 2018). Most gene have the possibility to be contacted by multiple enhancers and the transcriptional output can be finetuned by different patterns of enhancer activations.

The annotation of gene coding promoters is well established due to technological developments like ChIP-seq and RNA-seq. The identification of enhancers is profiting from the same technologies. However, the cell type specific nature of enhancers and specificity to certain signals makes it particularly hard to define a comprehensive set of all enhancers in an organism. One approach is to check for the presence of the activator histone acetyltransferase (HAT) p300 or open chromatin (ATAC-seq) to identify enhancer (Bernstein *et al.* 2005; Giresi *et al.* 2007; Mallm *et al.* 2019). Additionally, the detection of enhancements for the histone modifications H3K4me1 is the most prominent mark to identify enhancers (Heintzman *et al.* 2007; Visel *et al.* 2009). The combination with H3K27ac is found on active enhancers only (Creyghton *et al.* 2010), while the cooccurrence with H3K27me3 marks poised enhancers often in a developmental context (Rada-Iglesias *et al.* 2011). The type of modifications and the combination of various marks allows to better understand, how chromatin states impact the gene regulation within a cell. The deposition of histone marks is a dynamic process and many enzymes are involved in placing and removing those marks.

1.1.5. Writers/Readers/Erasers of histone modifications

The histone modifications are mainly placed on the N-terminal tail as post-translational modifications (PTMs). In the context of histones, the enzymes placing such modifications are called “Writers” and the removers are “Erasers” (**Figure 3A/B**). The third essential group of proteins are called “Readers” and these contain a domain to recognizes specific modification (**Figure 3C**). The amino acid lysin (single letter code: K) can be modified by adding up to three methyl groups by enzymes called histone methyl transferases (KMT) (**Figure 3A**). The dual role of histone methylation as activator (H3K4me3, H3K36me3) or repressor (H3K9me2/3, H3K27me3) is indirect and strongly linked with the recruitment of co-factors, which define the role of the methylation mark (Vermeulen *et al.* 2010). Additionally, certain modifications like H3K4me1 are used to mark genomic regions and the combination with other marks define its functionality. Examples of KMTs are the MLL-family like MLL3 or MLL4 and SET-family for creating H3K4me1 (Gu and Lee 2013). SETDB2 and SUV39H1/2 for methylation of H3K9 or EZH1/2 for H3K27me3 (Margueron *et al.* 2008; Volkel and Angrand 2007). The enzymes removing methyl groups are named histone lysine demethylases (KDM) (**Figure 3C**).

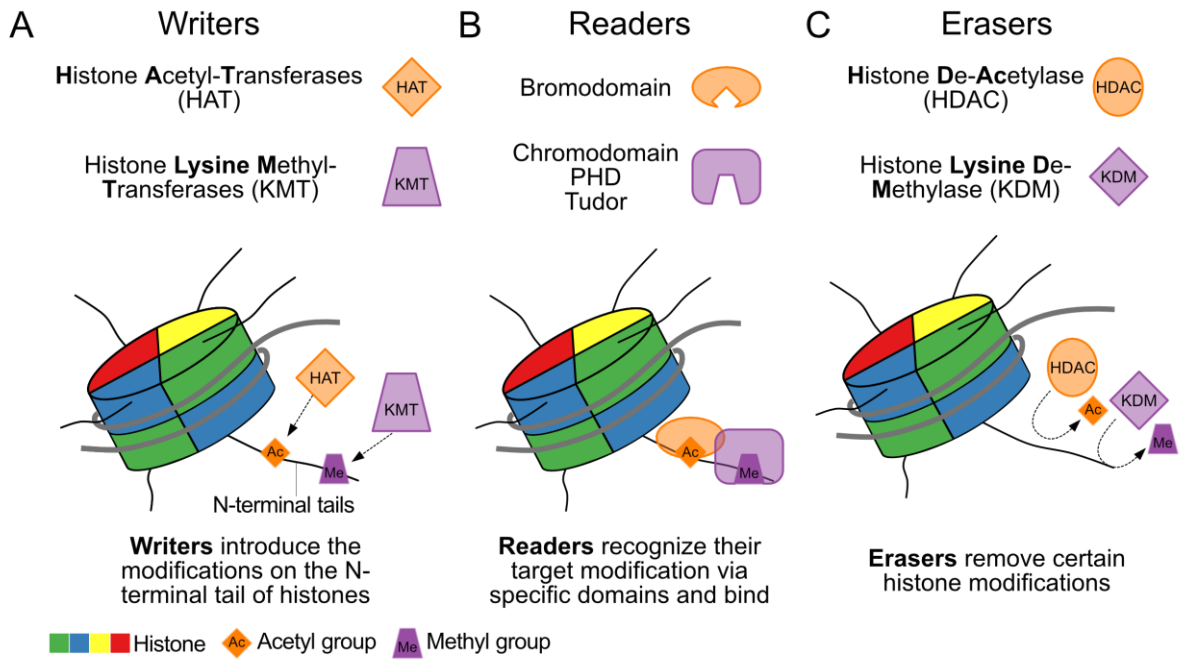

Figure 3: Proteins functioning on histone modifications

(A) "Writers" are enzymes adding a covalent modification on the N-terminal tail of histones. Enzymes placing acetylation (ac) are called histone acetyl transferases (HAT) and for methylations (me) histone lysine methyltransferases (KMT) are responsible. (B) Various enzymes contain specific protein domains to recognize and interact these modifications. The name of these factors is "Readers". The classical domain to recognize acetylations is the bromodomain. Methylation groups are detected by chromodomains, PHDs and tudor domains. (C) Enzymes to remove modifications are called "Erasers". HDACs, histone deacetylases, and KDMs, histone lysine demethylases, are the responsible enzymes.

The presents of acetylation marks are always an activating mark, as it removes the positive charge on histones and consequently reduce the interactions to the negatively charged DNA. The distances between histone tail and DNA expanse and the chromatin opens up. The writer enzymes are called histone acetyl transferases (HAT) (**Figure 3A**) and the most prominent examples are P300, GBP or KAT2A (GCN5) (Dancy and Cole 2015; Thomas and Chiang 2006). The modified lysins are found on position 9 and 27 of histone 3, named H3K9ac and H3K27ac. HDACs, histone deacetylases, are the counterparts of HATs and have mainly repressive functions (**Figure 3C**).

Readers are all proteins with a domain to recognize modified histone tails. Methylation marks are recognized by tandem tudor, chromo, plant homeodomain (PHD) and others (**Figure 3B**) (Yun *et al.* 2011). Acetylations are read by proteins containing bromodomain, double PHD finger and others (**Figure 3B**) (Filippakopoulos *et al.* 2012; Sabari *et al.* 2017).

1.1.6. The function of histone marks in gene regulation

H3K4me1 – Mono-methylation of histone 3 at lysine 4

The modification with a single methyl group (me1) at lysine (K) number 4 on the N-terminal tail of histones (H3K4me1) is found in intergenic regions identified as enhancers (Bernstein *et al.* 2005) (**Figure 4A/B**). Active enhancers gain additional H3K27ac and higher levels of accessibility measured by ATAC-seq (**Figure 4A**) (Creyghton *et al.* 2010). In development, cell type specific enhancers are poised and marked by H3K4me1 and the repressive mark H3K27me3 (**Figure 4B**). In addition, this marks is present in the flanking regions of promoters, while depleted in the exact promoter site (Barski *et al.* 2007). This pattern is found on active promoters, while the H3K4me1 mark at the center of a promoter in combination with H3K27me3 is correlated with repressed gene expression (Cheng *et al.* 2014). The mode of action for H3K4me1 is not completely understood. It is supposed to work by being recognized by other proteins, which then can stably bind and recruit activators like HATs to activate the target region (Jeong *et al.* 2011). However, recent studies also imply that the physical presents of the KMTs MLL3/4 at enhancers is more essential then the positioning of this methylation mark to the histone tail (Dorighi *et al.* 2017).

H3K4me3 – Tri-methylation of histone 3 at lysine 4

The modification with three methyl groups of the same residue is called H3K4me3 modification and is found at promoters in general (Barski *et al.* 2007; Bernstein *et al.* 2005; Santos-Rosa *et al.* 2002). Active promoters harbor a strong H3K4me3, H3K9ac, H3K27ac and ATAC signal directly at the center of a promoter region and are strongly correlated with transcriptional activation (**Figure 4A**). Low levels of H3K4m3 are detected together with repressive H3K27me3 modification at bivalent domains (**Figure 4B**) (Bernstein *et al.* 2006). In a developmental context, these domains keep genes repressed in ESCs and during differentiation specific gene get activated, while others become permanently silenced by altering the chromosomal state (Voigt *et al.* 2013). The function of H3K4me3 at promoters is to allow a faster transcription activation of specific gens upon stimulations (Lauberth *et al.* 2013).

H3K9ac – Acetylation of histone 3 at lysine 9

The acetylation of histone 3 at lysine 9 (H3K9ac) is also strongly correlated with H3K4me3, H3K27ac and ATAC at active promoters (**Figure 4A**) (Karmodiya *et al.* 2012). Further, H3K9ac is often found at CpG rich promoters. In ESCs that H3K9ac correlates with the pluripotency potential and the reprogramming capacity (Hezroni

et al. 2011). Upon differentiation the H3K9ac levels were reported to drop (Krejci *et al.* 2009). Additionally, a role of H3K9ac in the release of paused RNAP II was identified (Gates *et al.* 2017).

H3K9me3 – Tri-methylation of histone 3 at lysine 9

H3K9me3 is the classical mark of repressed regions, transcriptionally silenced genomic sites (**Figure 4C**) (Lehnertz *et al.* 2003). H3K9me3 mainly found in pericentromeric regions and the end of chromosomes called telomers. These sites contain high numbers of repetitive elements like satellite repeats or transposons, which have to be silenced permanently (Magaraki *et al.* 2017; Monaghan *et al.* 2019). The histone mark H3K9me3 is enriched there and these sites are more static to keep genomic stability (Saksouk *et al.* 2015). It is deployed by Suv39h and correlated with repressive DNA methylation at these repeat-rich genomic regions.

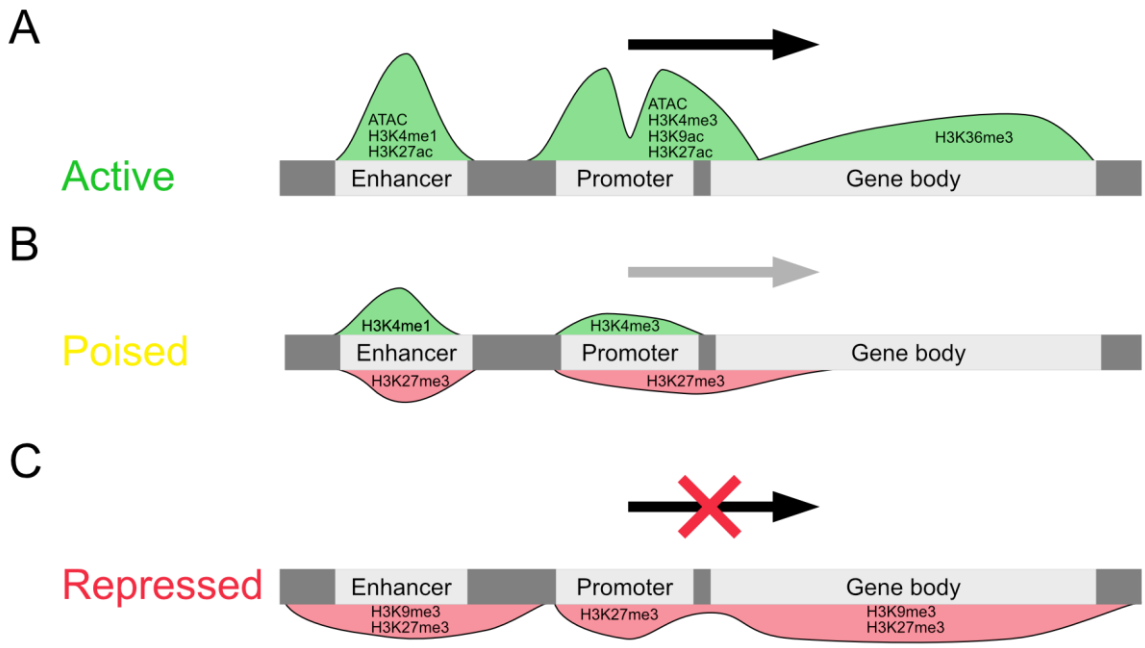

Figure 4: Role of histone modifications on specific genomic features

(A) Active chromatin states at enhancers, promoters and in gene bodies. Enhancers are marked by H3K4me1, H3K27ac and accessible chromatin measured by ATAC. The modifications H3K4me3, H3K9ac, H3K27ac and ATAC are marks of active promoters, while the gene body is enriched for H3K36me3. The green color of the peaks indicating, that all these modifications are correlated with gene activation. (B) In poised state, enhancers are marked by HeK4me1 in combination with the repressive marks H3K27me3. The same repressive mark together with H3K4me3 is found at poised promoter also called bivalent. The green peaks correspond to active marks, while the red peaks are for repressive marks. (C) Enrichment patterns for repressed regions. H3K27me3 is found on all sites in the repressed state. H3K9me3 is often found on repressed sites distal from actual promoters. The red color of the peaks indicates the repressive nature of these marks.

H3K27ac – Acetylation of histone 3 at lysine 27

Like other acetylation marks H3K27ac is an activating modification. It correlates strongly with previously described marks, like H3K9ac or H3K4me1/4, and transcription when occurring at promoters (Ernst *et al.* 2011). A promoter is classified as active when H3K27ac is present (**Figure 4A**). In addition, it is found with H3K4me1 at active enhancers (Creyghton *et al.* 2010). Both marks occur at non-promoter sites also classifying this regulatory element as active (Creyghton *et al.* 2010).

H3K27me3 – Tri-methylation of histone 3 at lysine 27

Tri-methylation of lysine 27 of histone 3 (H3K27me3) is a repressive chromatin mark and associated with the repression of genes via modifying histones at their promoters and gene bodies (**Figure 4C**) (Morey and Helin 2010). Especially its role in the repression of developmental genes is well characterized (Boyer *et al.* 2006). Developmental enhancers are often marked by both H3K4me1 and the repressive H3K27me3 and therefore called poised (**Figure 4B**) (Rada-Iglesias *et al.* 2011). Depending on the developmental signals these enhancers either become activated or completely silenced. Also, developmental promoters carry the H3K27me3 marks together with the activating H3K4me3 (**Figure 4B**) (Bernstein *et al.* 2006). The only known KMT for H3K27me3 is EZH2 (Kuzmichev *et al.* 2002).

1.1.7. Genome-wide approaches to characterize chromatin states

The implementation of next-generation sequencing (NGS) approaches allowed to investigate chromatin states genome-wide. This massive parallel sequencing technologies revolutionized the field of genomics. The basic idea is to determine the exact sequence of selected DNA fragments applied to these platforms. The DNA of interest is fragmented and specific adaptors are added. At this point the samples are called NGS libraries and can be applied to an NGS sequencer. By using nucleotides with different fluorescent labels, the exact sequence of the DNA fragment can be detected. And this is happening for millions of fragments in parallel.

This basic principle can be applied to numerous different applications. The entire mRNA of a sample can be isolated, reverse-transcribed and transformed into an NGS library. This allows to characterize the entire transcriptome (RNA-seq) (**Figure 5A**) (Wilhelm and Landry 2009). DNA and bound proteins can be crosslinked and specific

antibodies can pull down proteins of interest together with bound DNA fragments. These proteins can be TFs like STAT1 (**Figure 5B**) or histone modifications like H3K4me3 (**Figure 5C**). The isolated DNA is the starting point of an NGS library. This approach is called chromatin immunoprecipitation followed by NGS sequencing (ChIP-seq) (Park 2009). Assaying chromatin accessibility genome-wide (ATAC-seq) is a technique using a transposase, which binds and cuts nucleosome free DNA regions (**Figure 5D**) (Buenrostro *et al.* 2015a). This DNA is then isolated and sequenced, revealing a map of accessible chromatin regions. The combination of these and other readouts helped during the last decade to create a detailed picture of the function of genomic regions. In addition to the classical sequencing of bulks or entire populations of cells, novel approaches allow to analyze genome-wide patterns with single cell resolution. Single cell RNA-seq (scRNA-seq) characterizes the transcriptome of thousands of cells (Cusanovich *et al.* 2015) and single cell ATAC-seq (scATAC-seq) to collect information about chromatin accessibility (Buenrostro *et al.* 2015b; Cusanovich *et al.* 2015).

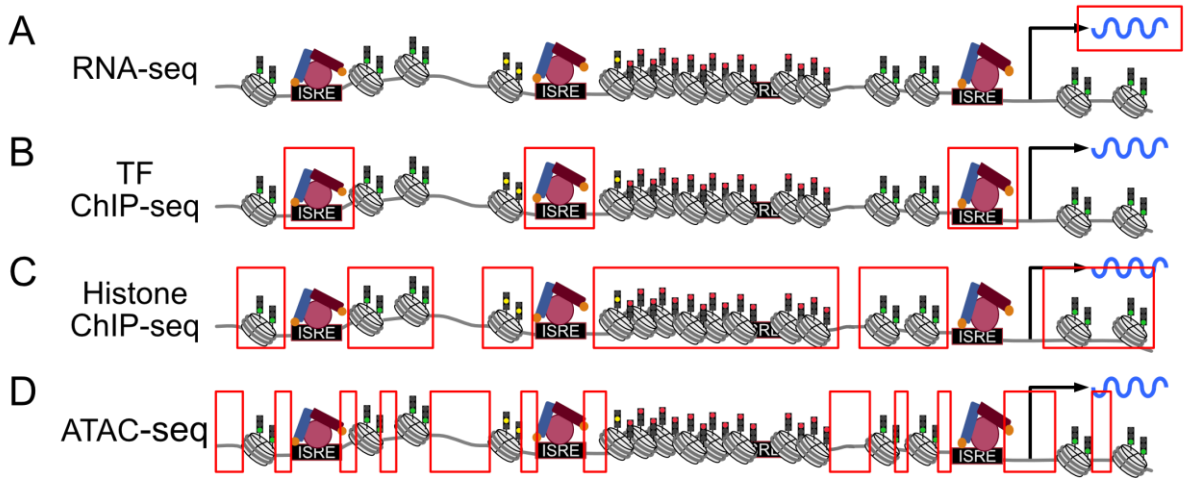

Figure 5: Detected genomic regions by different NGS approaches

Summary of some NGS methods and their analyzed sites on the model of a representative genomic loci. (A) Theoretical genomic regions, which highlights the detected features, the mRNA, by RNA-seq (B) ChIP-seq for TFs identifies aa bound genomic sites bound by this TF. (C) Histone modifications can be detected by ChIP-seq as well. By using different antibodies, different genomic loci are identified. (D) ATAC-seq characterizes open, accessible genomic regions. Read boxes indicate the genomic location, information is gathered for the particular method.

1.2. The unique potential of embryonic stem cells

NGS techniques are often used to identify dynamics of histone marks or TF binding sites in the context of development. This process required the differentiation from potent stem cells into terminally differentiated cell types (Spitz and Furlong 2012; Weissman 2000). The first level of stem cells, totipotent stem cells, are capable to self-renewal and to differentiate into all cell types of the embryo as well as into extra embryonal tissues like placenta (Reik and Surani 2015). The next level are pluripotent stem cells, which are able to give rise to all embryonic cell including more specific stem cells (Martello and Smith 2014). This last group of stem cells are multipotent and can give rise to cell types of specific linages. For example, neuronal stem cells can differentiate into cells of the neuronal branch like neurons or glia cells (Urban and Guillemot 2014).

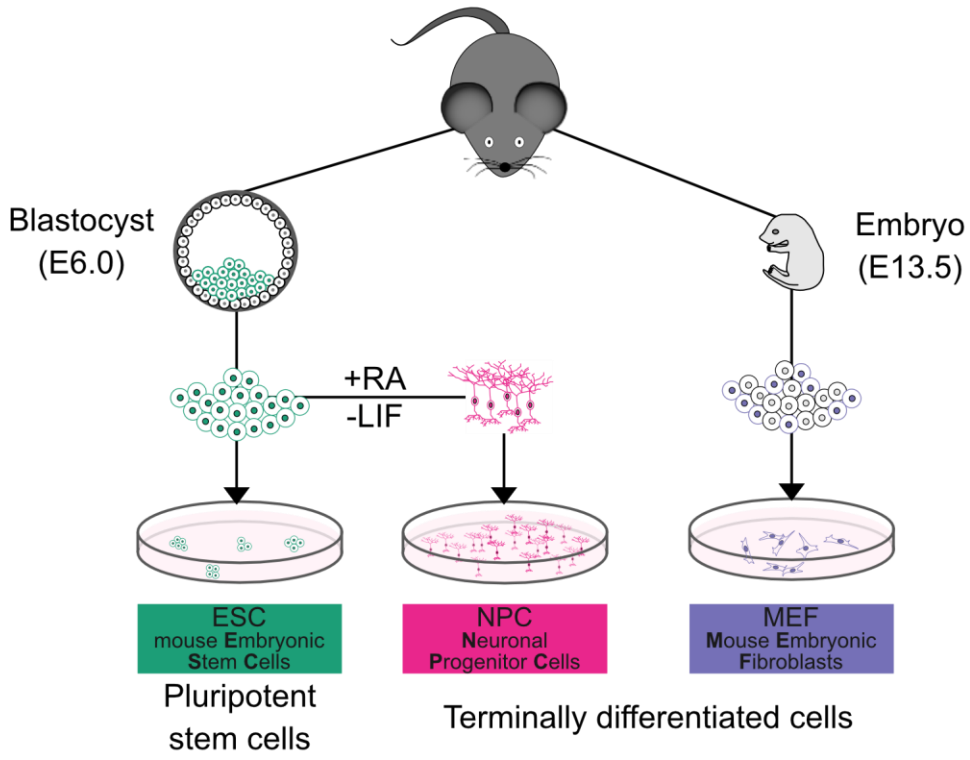

Figure 6: Scheme of the origin of various mouse cell types

The cell lines are isolated from mouse embryos. The blastocyst stage at embryonic day 6 (E6.0) is the source of mouse embryonic stem cells (ESCs) (left). These pluripotent stem cells can be differentiated into neuronal progenitor cells by adding retinoic acid (RA) and removing leukemia inhibiting factor (LIF) from culturing media (middle) (Bibel *et al.* 2007). At the later embryonic developmental stage (E13.5), embryos are isolated, organs and limbs are removed and the trunk is used as source to isolate mouse embryonic fibroblasts (MEFs).

The most commonly used murine stem cells are isolated pluripotent stem cells from the inner cell mass at the blastocyst stage of the mouse embryo, called embryonic stem cells (ESCs) (**Figure 6**) (Weissman 2000). The chromatin of ESCs has specific features, as it is more open with less repressed chromatin regions and more plasticity (Karmodiya *et al.* 2012). These primary cells are fast dividing cells, which preferentially grow in attached 3D clusters. As starting point for many developmental processes ESCs have a lot of bivalent promoter signatures and poised enhancer states. These specific epigenetic marks are required to allow ESCs differentiated into various other cell types, like neurons (Bibel *et al.* 2007) or fibroblasts (Bai *et al.* 2015). Stem cell populations are very valuable for an organism, as they harbor the potential to regenerate damaged tissues during the entire life of an organism. Consequently, stem cells are very potent and unique in many points. Above other features, recent studies identified an underdeveloped innate immune response in ESCs compared to differentiated cells like MEFs (Guo 2017).

A great example for the potential of stem cells is the possibility to differentiate ESCs into neuronal progenitor cells (NPCs) (**Figure 6**). The removal of specific molecules like leukemia inhibiting factor (LIF) that keep ESCs in their stem cell state is combined with adding retinoic acid (RA) to promote the differentiation from ESCs to neuronal progenitor cells (NPCs) (Bibel *et al.* 2007). After the differentiation, these cells are post-mitotic, terminally differentiated and without any additional potential for further differentiation. They morphologically reassemble neurons by the formation of neuron-specific structures like axon and dendrites. Further, NPCs are able to form neuronal networks and are able to create neuronal action potential (Bibel *et al.* 2007).

In contrast, another established murine embryonic cells source is mouse embryonic fibroblasts (MEFs) (**Figure 6**). These primary cells are isolated from mouse embryos at embryonic day 13.5. The isolate trunk of embryos is stimulated to release cells with fibroblast-like features. They do not have the potential to differentiate. Phenotypically they resemble a fibroblast-like state.

1.3. Activation of gene expression upon interferon signaling

1.3.1. Innate immunity is the first level of antiviral response

Mammals are exposed to potential pathogenic organisms on a daily base. To avoid these pathogens to cause an infection, mammals have developed an immune system. The immune system has two parts, the innate and adaptive immune system (Marshall *et al.* 2018). A fast, more unspecific immune system is required to challenge new pathogens from the beginning of the encounter. That's the job of the innate immune system (Turvey and Broide 2010). This system recognizes conserved patterns in pathogens to trigger a faster immune response. This response can cause an inflammatory reaction as well as phagocytosis, the activation of immune cells like neutrophils and macrophages. The innate immune system is much older and highly conserved in animals as well as in plants. The adaptive response is based on the innate immunity and critical, if the first wave of responses was insufficient to clear the pathogen. It has a high level of specificity to a pathogen and is able to remember previous infections (Bonilla and Oettgen 2010). This happens on the costs of time, as the adaptive immunity requires more time when dealing with a new infectious pathogen for the first time.

While pathogens like bacteria or parasites have very unique patterns in their surface to be recognized by the innate immune system, the surface of viruses normally doesn't have such molecules (Marshall *et al.* 2018). Viruses are normally assembled by their host and composed of host cell lipids, which makes it very hard to identify virus particle as hostile. One unique feature of some viruses is the presents of double-stranded RNA (dsRNA). Within a host cell this dsRNA is recognized and a drastic cellular response is triggered. Firstly, the dsRNA will be cut into small pieces and used in the RNA interference machinery to degrade the virus RNA (Maillard *et al.* 2013). Secondly, the host cell starts to produce signaling molecules named cytokines like interferon alpha (IFNα) or interferon beta (IFNβ). Upon release from its host cell, these molecules bind to surface receptors of neighboring cells and trigger the JAK-STAT signaling cascade to activate antiviral genes expression.

1.3.2. Activation of canonical JAK-STAT signaling cascade

The Janus kinase-signal transducer and activator of transcription (JAK-STAT) signaling cascade and its role in the anti-viral innate immune response was discovered

during the 1990s (Darnell *et al.* 1994; Fu *et al.* 1992; Ihle *et al.* 1994; Levy and Darnell 2002; Silvennoinen *et al.* 1993; Stark and Darnell 2012). We can distinguish between three types of interferon signaling. Type I and III IFNs are found in most cells, type II is restricted to immune cells (Stanifer *et al.* 2019). Type I is triggered by interferons, like alpha (α) or beta (β), and recognized by the heterodimer of transmembrane receptors IFNAR1/2 (**Figure 7A**). The associated kinases JAK1 and TYK2 are activated by phosphorylating themselves. Additional targets of this membrane associated kinase family are cytoplasmic STATs (**Figure 7B**). The phosphorylation of tyrosine 701 in STAT1 and tyrosine 689 in STAT2 causes together with IRF9, to form the ISGF3 complex (**Figure 7C**) (Levy and Darnell 2002). In addition, Structural studies imply that the binding between STAT2 and IRF9 is highly specific within these TF families (Rengachari *et al.* 2018).

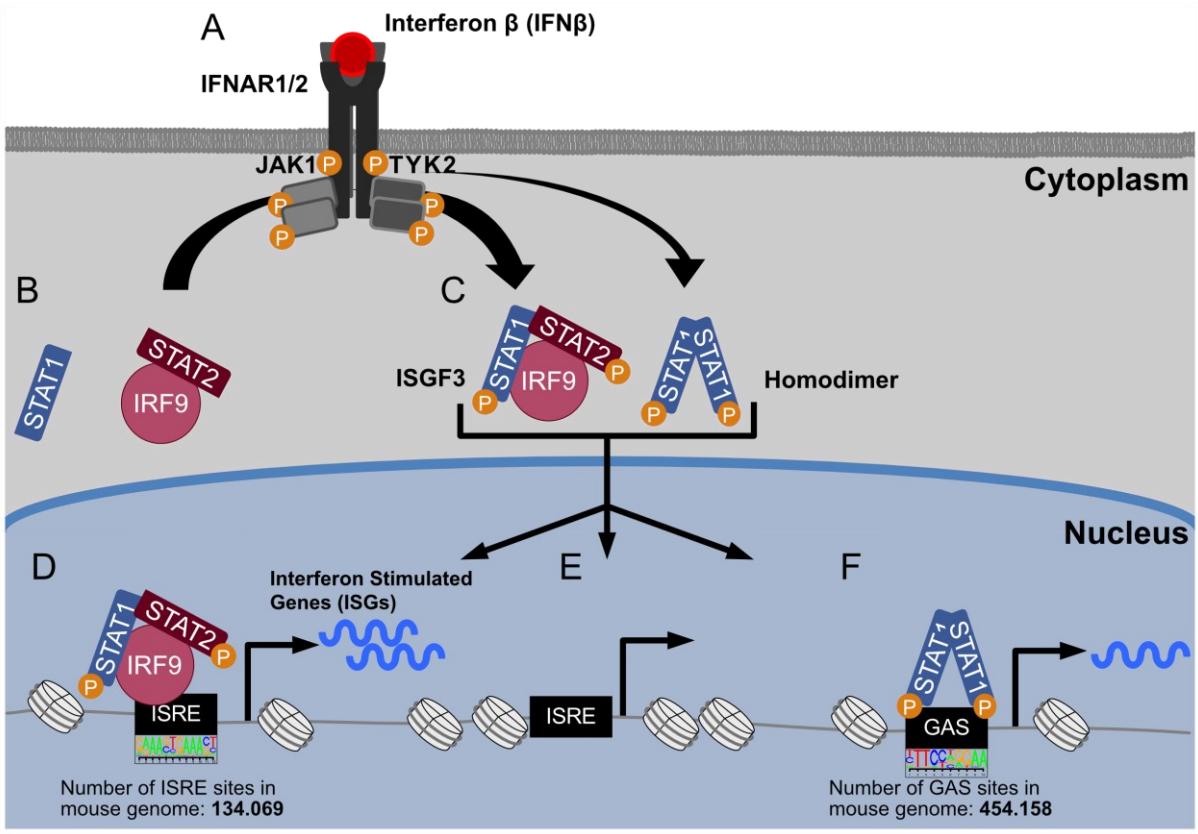

Figure 7: The canonical JAK-STAT signaling cascade triggered by type I interferon

(A) Interferon beta (IFNβ) outside of the cells is recognized by the transmembrane receptors IFNAR1/2. Upon binding, the bound kinases JAK1 and TYK2 start to phosphorylate themselves and become active. (B) The cytoplasmic, potential TFs STAT1 and STAT2/IRF9 are targets of the active kinases. (C) The TFs become phosphorylated and form ISGF3 complex or STAT1 homodimers. These complexes then can translocate to the nucleus. (D) ISGF3 can bind its target motif called ISRE and cause gene to be activated. The mouse genome contains this sequence 134,096 times. (E) Some of these motives are not bound due to cell type specific effects. (F) STAT1 homodimers do recognize instead the GAS motif, which has 454,158 copies in the mouse genome. The binding of STAT1 complexes cause gene activation.

Further the high affinity of STAT2 and IRF9 makes it likely that these factors are pre-bound in the cytoplasm without any IFN stimulation (Rengachari *et al.* 2018). STAT1, STAT2 and IRF9 together as ISGF3, are then translocated to the nucleus and binds to its motifs, which are called ISRE (**Figure 7D**) (Loutfy *et al.* 2003; Schindler *et al.* 1992). Upon the binding of its motif, the TFs stimulate gene expression to promote the antiviral properties of the cell. Activated genes are called interferon stimulated genes (ISGs). Conceptually, STAT2 is the only member of the STAT family which does not recognizing a GAS motifs (Seegert *et al.* 1994). The reason for that is, that it has a unique sequence, a nucellar export signal, which becomes activated upon homodimerization of STAT2 (Melen *et al.* 2001).

STAT2 homodimers are rapidly exported from the nucleus, reducing their probability to find and bind their potential binding motifs. The interactions with STAT1 and IRF9 are required to become an active TF complex. Based on the structure of this complex, IRF9 is the essential part for motif recognition. Consequently, the ISRE motif, part of the IRF-family, is bound. Based on the reference sequence of ISRE, the mouse genome contains 134,069 ISRE motifs and not all potential bindings sites will be bound by activated TFs (**Figure 7E**). In addition, phosphorylated STAT1 can also form a homodimer, which recognizes the STAT1-motif, also called GAS (**Figure 7F**) (Brierley and Fish 2005). The activation of ISGs is the consequence of this binding. Additionally, other enzymes are also able to modify STATs. One of the most prominent modifications, is the phosphorylation of serine 727 of STAT1, which causes enhanced transcriptional induction, and is placed by CDK8 in the nucleus (Bancerek *et al.* 2013).

Type II interferon responses are triggered by the gamma version of interferon and recognized by IFNGR1/2 (interferon gamma receptor 1/2). The kinases are JAK1 and JAK2. They are bound to these receptors and they predominantly lead to the formation of homodimers of STATs. The binding in the nucleus happened at the GAS motif. The last type of interferon signaling is type III via interferon lambda. The receptor composition is different to the other types, as type III uses one unique IFNLR1 and on interleukin receptor, IL10, IL22 or IL26. The kinases JAK1 and TYK2 are associated with these receptors and cause the formation of STAT homodimers. The binding to the GAS motif is conserved with other types of interferon signaling. The differences between the types of interferon response are also characterized by the dynamics of ISG responses and different abundances of receptors (Pervolaraki *et al.* 2018). The aim of these pathways is to enhance the antiviral property of the target cells.

1.3.3. Anti-viral function of interferon stimulated genes (ISGs)

The IFN triggered signaling cascades activate genes to deal with viral infections. ISGs are a hallmark of the innate anti-viral immune response and able to inhibit the virus life cycle at almost every step of its maturation (Schoggins and Rice 2011). Many early studies found between 200 to 500 ISGs to become activated upon IFN stimulation depending on type of IFN and cell type (de Veer *et al.* 2001; Der *et al.* 1998). Among the target genes are directly TFs like STAT1/2, IRF1/7/9, which then further enhance the production of ISGs (Schoggins *et al.* 2011). RTP4 was shown to interfere with viral replication (Dang *et al.* 2018), while in neuronal cells RTP4 seems to be repressive for the TF IRF3 (He *et al.* 2020). The gene family OAS1/2/3, for instance, causes the activation of RNaseL to degrade viral RNA genomes (Kristiansen *et al.* 2011). IFIT genes are able to inhibit the protein translation of the host cell (Mears and Sweeney 2018). The interaction with eukaryotic initiation factor allows to block the initiation of translation. Members of the Guanylate-Binding Protein (GBP) family are responsible for inhibiting the viral protein furin, responsible for viral envelope protein maturation (Braun *et al.* 2019). Instead of inhibiting the virus, some genes cause the infected cell to induce controlled cell death called apoptosis (Kotredes and Gamero 2013). IFI27 is one of these genes inducing apoptosis via the activation of specific caspases (Gytz *et al.* 2017). But at some point, the anti-viral response has to be stopped and the cell has to be able to return to its starting condition. Firstly, this is important as the anti-viral response is toxic to the cell. Blocking essential pathways like translation or inducing apoptotic pathways can be handled to a certain degree but an overstimulation causes tissue damages and ultimately might cause more damage to the organism then the actual infection. Secondly, the cell has to reset its antiviral response potential to be able to respond to future viral infections. Negative feedback loops are the answer to stop the anti-viral response.

1.3.4. Negative feedback loops are part of JAK-STAT signaling

As the activation of the signaling cascade and the downstream effects are fast and severe, it has to be tightly regulated to avoid unnecessary damages and to put the system back in a responsive state. The idea of a negative feedback loop is to activate target genes transiently as part of the anti-viral response but to inhibit the host cell to continue with its anti-viral response. This allows the cell to come back to its starting condition. There are three main classes of STAT induced negative regulators. Suppressors of cytokine signaling (SOCS) target proteins, like JAKs and STATs, for

proteolytic degradation (Kiu and Nicholson 2012). In addition, several SOCSs contain kinase inhibitor regions to block the initial phosphorylation of STATs (Kiu and Nicholson 2012). The protein inhibitor of activated STAT (PIAS) family interacts with STAT complexes upon IFN stimulation. This interaction physically inhibits nuclear STAT complexes to successfully bind DNA (Niu *et al.* 2018). Two members of this family additionally recruit co-repressors like HDACs to ISGs, to remove activating histone acetylation marks (Shuai 2006). The third big group are phosphatases, like SHP1, SHP2 or DUSP2 (Villarino *et al.* 2017), which cause dephosphorylation of tyrosine residues on STATs, JAKs or IFNAR (Kiu and Nicholson 2012). The removal of phosphorylation of STATs causes the complexes to fall part. Additionally, other genes are also involved in these negative loops like USP18. This protein binds to the IFNAR2 receptor by competing with JAK1 (Bliven *et al.* 2018). In addition, the binding with USP18 reduces the affinity of the receptor for IFNα and inhibits the receptor dimerization. Finally, USP18 is responsible for modifying proteins with ISG15, an ISG with similar structure to ubiquitin (Bliven *et al.* 2018). Many levels to regulate the innate immune response are known and one with particular interest for us is the link to chromatin modifying genes.

1.3.5. The role of chromatin in the IFNβ stimulated antiviral response

Evidence was found that chromatin associated factors are required for a functional JAK-STAT signaling cascade and to adapt the ISG response for different cell types (Smale *et al.* 2014). There are diverse interactions of the ISG pathway with chromatin modifiers. STAT2 promotes the recruitment the HAT p300, CBP and the chromatin remodeler BRG1 to promote the induction of ISGs (Liu *et al.* 2002; Loutfy *et al.* 2003). Alternatively, upon IFN stimulation, STAT1 and STAT2 are interreacting with the repressor HDAC1 (Nusinzon and Horvath 2003) ChiP-seq of the p300 defined many enhancers specific activated upon LPS triggered innate immune response (Ghisletti *et al.* 2010). The authors further identified PU.1 as pioneer TF to prime cell type specific enhancers during fibroblast development by binding the enhancers and marking them permanently with H3K4me1. The H3K4me1 mark allows a faster and stronger acetylation of these enhancers upon the next stimulation. The maintenance of this histone mark was called a chromatin-mediated memory mechanism (Ostuni *et al.* 2013). In addition, there was also a correlation found, that LPS-induced gene promoters and enhancers can be suppressed by H3K27me3 marks (De Santa *et al.* 2009). Recent studies linked the binding of STAT complexes with the presence of the non-canonical histone variant H2A.Z, which blocks the binding of the TFs (Au-Yeung

and Horvath 2018). Another histone variant H3.3 was found to play a role in IFN-mediated transcription of ISGs (Tamura *et al.* 2009). Additionally, H3.3 and the histone modification H3K36me3 are associated with epigenetic short time memory. Thereby, previous IFN stimulated cell show a faster and stronger transcriptional response during a second stimulation within 24 h (Kamada *et al.* 2018). In summary, evidence shows that chromatin associated factors impact the ISG response in various cell types.

1.3.6. Specific properties of the innate immune response in ESCs

The chromatin landscape in ESCs is different from many differentiated cell types. Interestingly, in ESCs the type I and III IFNs responses is very unique, although it is thought to be highly conserved between cell type (Sen 2001). As the innate immune response might also be harmful for the affected cells (Kotredes and Gamero 2013), the loss of stem cells might be even more damaging for the organism (Guo 2019). Therefore, the response to viral infections is unique and uncharacteristic in stem cells. Firstly, ESCs do fail to respond to La Crosse virus infections (Wang *et al.* 2013) and a wide range of markers of infectious agents like LPS, while fibroblasts are responsive (Guo *et al.* 2015). In line with that, stem cells are not able to produce interferon themselves (Hong and Carmichael 2013). While human ESCs are not responding to IFNβ due to a high expression of the repressor SOCS1 (Hong and Carmichael 2013). Studies demonstrate that murine pluripotent cells like mouse ESCs do respond to IFN type I because SOCS1 is not expressed (Whyatt *et al.* 1993). Upon differentiation of human ESCs to trophoblast, the ability to respond to IFNβ is re-acquired. However, their gene induction levels are much lower compared to differentiated murine cells such as mouse embryonic fibroblasts (MEF) (Guo 2017; Wang *et al.* 2014). Similar to human stem cells, the differentiation of murine ESCs into fibroblasts enhances the antiviral response (D'Angelo *et al.* 2016). One possibility to explain cell type specific gene expression differences is via differences on the levels of associated factors. Alternatively, the regulation of enhancers impacts gene expression profiles and these elements are known to be highly cell type specific. The role of enhancers can be directed by cell type specific chromatin states, which promotes or represses the binding of specific TFs.

1.4. Scope of the thesis

In the present thesis, I test the hypothesis that the chromatin environment plays an essential role in the cell type specificity of the IFNβ triggered ISG response. Selected ISGs show strong differences in their induction levels between embryonic stem cells and differentiated cells (Wang *et al.* 2014). Thus, there must be an epigenetic contribution to the IFNβ response. Furthermore, the key components like the receptors IFNRA1/2 and kinases JAK1/TYK2 are expressed in mouse ESCs and the cells are responsive to IFNβ (D'Angelo *et al.* 2016; Wang *et al.* 2014). However, the molecular mechanism that lead to an attenuated and/or cell type specific ISG response are only partly understood (Guo 2019).

The key TFs for inducing gene expression in the nucleus upon IFNβ stimulation are the TFs STAT1 and STAT2 (Stark and Darnell 2012). Therefore, genome-wide binding profiles of these TFs are needed to understand their role in the cell type specific ISG responses. While previous studies performed ChIP-seq for STAT1 in HeLa S3 cells (Robertson *et al.* 2007), for STAT2 in B cells (Mostafavi *et al.* 2016) or for STAT1p727 in CDK8 negative neoplasm (Nitulescu *et al.* 2017) a comparison of different cell types with the same genome is currently missing. Furthermore, chromatin states need to be considered as these are key regulators of TF binding (Li *et al.* 2007). Accordingly, I here investigated how ISG induction, STAT TF binding sites and chromatin states are linked. Firstly, a comprehensive comparison of differences of IFNβ stimulated ISG induction patterns in different cell types on a genome-wide scale was conducted by performing RNA-seq experiments in ESCs, NPCs and MEFs at for one (1h) and six hours (6h) of IFNβ treatment. These data sets allowed it to identify ISGs genome-wide for each cell type and to further characterize their cell type specificity. Secondly, the identification of cell type specific ISGs was complemented with the analysis of cell type specific STAT1 and STAT2 binding by ChIP-seq. The combination of the occupancy patterns of these two TFs mapped ISGF3 sites as STAT1 and STAT2 double positive sites in each cell type to dissect ISFG3 binding at promoters, introns and intergenic sites. By conducting a co-regulation analysis of single cell ATAC-seq data, we linked non-promoter bound ISGF3 binding to ISGs. These ISGF3 enhancer elements activated ISGs with similar strength than ISGF3 directly bound at the promoter. Third, the chromatin environment of ISGF3 binding sites was characterized. ESCs have higher chromatin accessibility (Tee and Reinberg 2014), less constitutive heterochromatin (Efroni *et al.* 2008) and differences in nucleosome positioning (Teif *et al.* 2012) compared to differentiated cells like MEFs. These epigenetic differences impact the possibilities of TFs to find and bind their recognition

motives. Consequently, we characterized how chromatin states impact on STAT binding profiles upon IFNβ stimulation.

By addressing the above three main aims, this thesis provides a genome-wide multiomics data set of IFNβ stimulated ESCs in comparison to differentiated cells with the same genome. The cell type specific ISG induction was linked to STAT1 and STAT2 binding. A chromatin signature was identified that is permissive for the binding of the activating ISGF3 complex. It is anticipated that these findings will help to investigate the potential of epigenetic drugs that change the chromatin environment at ISGF3 binding sites. This could help to fine tune the antiviral response in a cell type specific manner.

2. Materials and Methods

2.1. Kits and Reagents

Product name	Company	Reference
96 Well White/Clear Bottom Plate, TC Surface	Thermo Fisher Scientific	165306
Accutase® solution	Sigma-Aldrich	A6964-100ml
AMPure XP Beads	Beckman Coulter	A63881
Bright-Glo™ Luciferase Assay System	Promega	#E2610
Chromium i7 Multiplex Kit	10xGenomics	PN-120262
Chromium i7 Multiplex Kit N, Set A	10xGenomics	PN-1000084
Chromium Single Cell 3' Reagent Kits v2	10xGenomics	PN-120237
Chromium Single Cell ATAC Reagent Kits	10xGenomics	PN-1000110
Clarity™ Western ECL Substrate	Biorad	1705060
Corning® Matrigel® Growth Factor Reduced (GFR) Basement Membrane Matrix	Corning	356230
D5000 Reagents	Agilent	5067-5589
D5000 ScreenTape	Agilent	5067-5588
DMEM, low glucose, pyruvate, no glutamine, no phenol red	Gibco	11880-28
DMEM/F-12, ohne Phenolrot	Gibco	21041025
E-Gel™ EX Agarose Gels, 2%	Invitrogen	G401002
FBS, South America origin, fetal bovine serum, tetracycline free, 0.2 µm sterile filtered	PAN-Biotech	P30-3602
Formaldehyd solution (37 wt. % in H2O, contains 10-15% Methanol)	Sigma-Aldrich	252549
G-5 Supplement (100X)	Gibco	17503012
Gelatin from porcine skin	Sigma-Aldrich	G2500-100g
GlycoBlue™ Coprecipitant (15 mg/mL)	Thermo Fisher Scientific	AM9515
Halt Protease Inhibitor Cocktail (100x)	Thermo Fisher Scientific	TE266456A
High Sensitivity RNA Sample Buffer	Agilent	5067-5580

High Sensitivity RNA ScreenTape	Agilent	5067-5579
High Sensitivity RNA ScreenTape Ladder	Agilent	5067-5581
Interferon β from mouse	Sigma-Aldrich	I9032
Laemmli SDS sample buffer, reducing (6X)	Alfa Aesar	J61337
LUNA™ Cell Counting Slides	Logos Biosystems	L12003
MinElute PCR Purification Kit	Qiagen	28004
Mini-PROTEAN TGX Gels 4-20%	Biorad	4568094
N-2 Supplement (100X)	Gibco	17502048
NEBNext Multiplex Oligos for Illumina	New England Biolabs	E7335S
NEBNext Ultra II Directional RNA Library Prep Kit for Illumina	New England Biolabs	E7760
NEBNext® rRNA Depletion Kit (Human/Mouse/Rat)	New England Biolabs	E6350
NEBNext® Ultra™ II DNA Library Prep Kit for Illumina®	New England Biolabs	E7645
Neurobasal™ Medium	Gibco	21103049
NucleoSpin RNA Plus Kit	Macherey-Nagel	740984.250
Page Ruler Plus Prestained	Thermo Fisher Scientific	26619
Penicillin-Streptomycin, 10,000 U/ml Penicillin, 10 mg/ml Streptomycin	PAN-Biotech	P06-07100
Phosphatase Inhibitor Cocktail 3	Sigma-Aldrich	P0044
Pierce™ BCA™ Protein-Assay	Thermo Fisher Scientific	23225
Pierce™ Methanol-freie Formaldehyd-Ampullen (16%, Methanol-free)	Thermo Fisher Scientific	28906
PowerStem ESPro 1 - Serum-free media for mouse ES cells and knockout mice, (w/o: LIF)	PAN-Biotech	P04-77510K
Proteinase K (20mg/ml)	Genaxxon Bioscience	M3037
Qubit assay tubes	Thermo Fisher Scientific	Q32856
Qubit dsDNA HS assay kit	Thermo Fisher Scientific	Q32854
Retinoic acid	Sigma-Aldrich	R2625
Ribo-Zero rRNA Removal Kit	Illumina	15066012
RiboLock RNase-Inhibitor (40 U/µl)	Thermo Fisher Scientific	EO0381
RNA Clean XP beads	Beckman Coulter	A63987
RQ1 RNase-Free DNase	Promega	M610A

Simple ChIP Enzymatic Chromatin IP Kit	Cell Signaling Technology	9003
Sodium Acetate (3 M), pH 5.5, RNase-free	Thermo Fisher Scientific	AM9740
Sodium clorid (NaCl)	Fisher Scientific	S/3160/65
Sodium dihydrogen phosphate anhydrous (NaH2PO4)	AppliChem	122018.121
Sodium phosphate dibasic (Na2HPO4)	AppliChem	A3599.0500
Stable Glutamine 200mM (100x)	PAN-Biotech	P04-82100
Stempan DMEM, w: L-Glutamine, w: 3.7 g/L NaHCO3, w/o: LIF	PAN-Biotech	P08-50500
TECAN plate reader Infinite® 200 PRO	Tecan	
Trans-Blot® Turbo™ Mini-size LF PVDF Membrane	Biorad	1620263
Trans-Blot® Turbo™ Mini-size LF Transfer Stacks	Biorad	1620263
Trypsin (10x) 0.5 %/EDTA 0.2 % in PBS, w/o: Ca and Mg	PAN-Biotech	P10-024100
Ultra-Low Attachment 75cm² U-Flask	Corning	3814
Venor® Gem Advance	Minerva Biolabs	11 7024

Table 1: Kits and reagents

2.2. Antibodies

Name	Company	Reference	Species	ChIP	WB
H3K4me1	Abcam	ab8895	Rabbit	2µg for 25µg of chromatin	1:500
H3K4me3	Abcam	ab8580	Rabbit	2µg for 25µg of chromatin	1:1000
H3K9ac	Active Motif	39137	Rabbit	10 µl per ChIP	1:1000
H3K9me2	Abcam	ab1220	Mouse	2-4 µg for 25 µg of chromatin	1:1000
H3K9me3	Abcam	ab8898	Rabbit	2-4 µg for 25 µg of chromatin	-
H3K27ac	Abcam	ab4729	Rabbit	2µg for 25µg of chromatin	1:1000
H3K27me3	Abcam	ab6002	Mouse	5-10 µg for 25 µg of chromatin	1:1000
H3K27me3	Active Motif	39155	Rabbit	5 µg per ChIP	1:1000
H3K36me3	Abcam	ab9050	Rabbit	4µg for 10^6 cells.	1:1000
H3	Abcam	ab1791	Rabbit	2µg for 10^6 cells.	1:1000
IgG rabbit	Acris	AB-105-C	Rabbit	2µl	
STAT1	Cell Signaling	#9172	Rabbit	1:50	1:1000
STAT1 p701	Cell Signaling	#7649	Rabbit	1:100	1:1000
STAT1 p727	Cell Signaling	#8826	Rabbit	1:50	1:1000
STAT2	Cell Signaling	#72604	Rabbit	1:50	1:1000
STAT2 p689	Merck	07-224	Rabbit	10µg (1µg/µl)	1:1000
CTCF	Active Motif	61311	Rabbit	5µl (1µg/µl)	1:500
IgG Rb	Cell Signaling	#2729	Rabbit	2µl (µg/µl)	-

Table 2: Antibodies

2.3. Instruments

Name	Company	Reference
E-Gel Safe Imager	Invitrogen	G6500
ChemiDoc XRS+ System	Biorad	1708265
Mini-PROTEAN® Tetra Vertical Electrophoresis Cell	Biorad	1658004
Trans-Blot® Turbo™ Transfer System	Biorad	17001917
2200 TapeStation system	Agilent	G2964AA
Gene Pulser Xcell Total System	Biorad	1652660
EpiShear™ Probe Sonicator	Active Motif	53052
Qubit 2.0 Fluorometer	Thermo Fisher Scientific	Q32866
LUNATM Automated Cell Counter	Logos Biosystems	L10001

Table 3: Instruments

2.4. Cell Culture Media

Cell Line	Growth Media	Company	Reference
mESC Media	PowerStem ESPro1	PAN-Biotech	P04-77510K
	PowerStem ESPro1 Supplements	PAN-Biotech	P04-77510K
	LIF	Self-made	
MEF Media	DMEM	Gibco	11880-28
	10% FCS	PAN-Biotech	P30-3602
	1x L-Glutamin (200mM)	PAN-Biotech	P04-80050
	1x Penicillin-Streptomycin, 10,000 U/ml Penicillin, 10 mg/ml Streptomycin	PAN-Biotech	P06-07050
CA Media	StemPAN	PAN-Biotech	P08-50500
NPC Media	Neuronal base medium	Gibco	21103049
	G5 supplements	Gibco	17503012
	NSC supplements	Gibco	17502048
	DMEM/F12	Gibco	21041-025

Table 4: Cell culture media

2.5. Laboratory methods

2.5.1. Cell culture ESCs and MEFs

ESCs cultured on 0.1 % gelatinized tissue flasks in ESC media. Cells were incubated at 37 °C and 5 % CO_2. Media was changed every second day and cells were split at least twice per week with Accutase. After medium removal, Accutase was added to the plates and incubated for a maximum of 5 min at 37 °C. Three times the volume of media was added, transferred to tubes and centrifuged with 300 g at RT for 5 min. Supernatant was removed and resuspended in fresh media. Cells were counted and platted accordingly in newly gelatinized plates. ESCs cell lines with double knockout (dKO), double catalytic dead (dCD) versions of MLL3/4 and the parental wild type (WT) were received from the Wysocka research group (Stanford University School of Medicine, Stanford, California 94305, USA) (Dorighi *et al.* 2017). Mycoplasma tests were performed and confirmed to be negative. Cells were adapted for five generations to our standard ESC media and cultured as described above. Experiments were preformed from these newly established stocks.

MEFs were culture in MEF medium and incubated at 37 °C and 5 % CO_2. Media was changed every second day and cell were split at least once per week with Trypsin. The procedure was identical to the splitting of ESCs with the following exceptions. Before applying Trypsin, plates were washed two times with PBS. MEFs were cultured on normal, not-gelatinized plates. All media compositions are found in **Table 4**.

2.5.2. Mycoplasma test

Cell lines were regularly tested for mycoplasma contaminations with Venor GeM Advance kit according to their manufacturer's protocol. In short, 500 µl of cell culture media of confluent cells was collected and heated up to 95 °C for 10 min. The sample was centrifuged at 13,000 rpm at RT for 10 min and the supernatant were transferred into fresh tube. 23 µl of Rehydration Buffer was added to PCR tube provided by the kit and then 2 µl of sample was added. The provided positive control was prepared by adding 25 µl Rehydration Buffer. The PCR amplification was done according to protocol and 5 µl of each sample was loaded on a 2 % E-Gel and ran two times 8 min with E-Gel Safe Imager. Gel was analyzed with ChemiDoc XRS+ System. Information for the kit and instruments are summarized in Error! Reference source not found. and **REF _Ref54075802 \h * MERGEFORMAT Table 3**.

2.5.3. Preparation of LIF

COS-7 cells were cultured in LIF-free ESC media till 70 % confluence. Then washed two times with MT-PBS and split cells into two 15 ml Falcon tubes. The cells were centrifuged down at 300 g for 5 min at RT. The pellet was resuspended in 800 µl MT-PBS and transfer into a 4 mm electroporation cuvette. After 5 min at RT, 20 µg pC10 plasmid (encodes for LIF) was added and incubated for 5 min at RT. Cells were pulsed in electroporator (Gene Pulser Xcell Electroporation Systems, 240V, 480µF, exponential decay) and incubated for 5 min at RT. Then the sample was transferred from each cuvette to a T75 flask containing LIF-free ESC media. The media was changed the next day. After three additional days media was collected (15 ml per flask) and fresh media was added. After two more days the second round of media was collected and pool with the first round. The combined mix was filtered with 0.45 µm and 0.2 µm filters. 1 ml aliquots of the filtered media containing LIF were prepared and frozen at -20 °C. New batches of LIF were tested as followed. ESCs were plated in 0.1 % coated 24-well plates and media with various dilutions of the new LIF (no LIF; 1:10,000 to 1:100) was added. As control the previous batch of LIF was added to a well in previously defined dilution for that batch. Morphology of ESCs over seven days were characterized and the lowest concentration of LIF to have optimal ESC growth was defined. Media composition and used instruments are listed in **Table 3** and **Table 4**.

2.5.4. Preparation of interferon beta (IFNβ)

BHK (Baby hamster kidney) cells were cultured in MEF media (**Table 4**) at 37 °C and 5 % CO_2. When reached nearly 100 % confluency, media was removed and attached cells were washed with MT-PBS once. Then cells were starved for 24 h by using modified MEF media, standard DMEM media with only 2 % FCS. The cells start producing IFNβ and release the signaling molecule to the media. The IFNβ enriched media was then collected and sterile filtered (0.45 µm). For long term storage the media was aliquoted and kept at -80 C. The activity was measured by IFNβ-activity assay and benchmarked against bought IFNβ (Sigma) using immortalized murine E19 embryonic small intestinal epithelial cells (IEC) from a transgenic mouse with a Mx2-Luciferase reporter insertion (Schwerk *et al.* 2013). $1.5*10^4$ IECs with a reporter-fusion of Mx2 were seeded into 96-well plates. These cells were treated with different concentrations of the self-made IFNβ stock (0-300µl) and benchmarked against the bought IFNβ (Sigma, 900U/µl). After 24 h a Luciferase assay was performed according to manufactural protocol and analyzed with the Tecan plate reader. The

self-made IFNβ concentration was measured to be 16.6U/µl. This work was performed with technical assistance from Caroline Bauer.

2.5.5. Differentiation of NPC

ESCs were differentiated into NPCs based on the Bible-protocol (Bibel *et al.* 2007). In short, 4-5*10^6 cultured ESCs were split onto a T75 UltraLow-BindingPlates in 15 ml StemPAN media to form cellular aggregates (CAs). Media was changed after two days. At day 4 and day 6 the media was changed to StemPAN + 5µM retinoic acid (RA). On day 7 plates for NPC plating were coated with 1 % Matrigel (100 µl Matrigel in 10 ml DMEM/F12) and placed in the incubator at 37 °C and 5% CO_2 (at least 15 min or O/N). On day 8 CA plates were collected, washed two times with PBS and dissociated for 3 min at 37 °C with 1 ml Accutase per T75 plate. After incubation CAs were carefully pipetted two times with a p1000 and 10 ml DMEM/F12 were added. CAs were spun down for 5 min at 200 g and resuspended in 3 ml per T75 of NPC media. NPCs were counted and plated (9.5*10^6 cells per 150mm plate). Media was changed between day two to four once. Experiments were performed five days after neuronal plating. The compositions of the used cell culture medias are summarized in **Table 4**.

2.5.6. Western blot

Western blot samples were prepared by collecting cells directly out of cell culture and cell number was calculated with Luna Cell Counter. All cells were transferred into 1.5 ml tubes, washed once with PBS and 50 µl of pre-prepared RIPA buffer were added per 0.5*10^6 cells. The mixes were incubated for 60 min on ice, spin down at max speed at 4 °C for 30 min. Supernatant was transferred to a fresh tube and stored at -20 °C.

Western blot buffer	
RIPA buffer	150 mM NaCl 1 % NP40 50 mM Tris-HCl, pH8.0 0.5 % Sodium deoxycholate 0.1 % Sodium dodecyl sulfate (SDS)

The exact protein concentration was determined by using BCA assay. ESC and MEF samples were adjusted to a concentration of 1.0 µg/µl and NPCs to 0.5µg/µl with RIPA buffer and 6x Laemmli Buffer.

Mini-PROTEAN TGX Gels 4-20% were used to run the Western blots. Protein samples were incubated at 95 °C for 10min before usage. 40µg of protein samples and 10µl of ladder Page Ruler Plus Prestained were used per gel. The running chambers were assembled based on manufacturer's guidelines and ran for 130 V for 70 min. The first quality measurements of electrophoresed gels were done with ChemiDoc XRS+ System to check the quality of the samples and the run. Afterwards gels were transferred onto a Trans-Blot® Turbo™ Mini-size LF PVDF Membrane using Trans-Blot® Turbo™ Transfer System with pre-programmed protocols for Mixed MW (Turbo) (7 min, 2.5 A, up to 25 V for protein 5-150 kDa). Membrane was blocked with 5 % BSA in TBST on a shaker at RT for 1 h. The primary antibodies were diluted according to manufactural recommendations in 5 % BSA and incubated at 4 °C O/N on a roller. On the following day, the membrane was washed three times with 1x TBST shaking at RT for 5 min under agitation. Membrane was then incubated with secondary anti-HRP antibody (normally 1:500 diluted in 5 % BSA) at RT for 1 h on a roller. Afterwards the membrane was washed three times with 1x TBST, incubation with Clarity™ Western ECL Substrate for 5min and exposed with ChemiDoc XRS+ System. All used reagents, kits and instruments can be found in Error! Reference s ource not found. and **Table** 3.

Hybridize the Probes to the RNA	Temperature	Time
	95 °C	2 min
	95-22 °C	0.1 °C/s
	2 °C	5 min

2.5.7. AMPure XP Beads purification

Based on the used manufacture's protocol the amount of beads was adjusted. In general, the defined amount of well mixed AMPure XP Beads was added to each sample, vortexed and incubated for 5 min at RT. Beads were placed on the magnetic stand and supernatant was removed. With the tube on the magnetic rack, 200 µl of 80 % EtOH was added without disturbing the beads. The mix was incubated at RT for 30 s. EtOH was removed and the washing step was repeated for a total of two washes. After last wash, the tubes were briefly spun down, placed back at magnetic rack and remaining liquid was removed. For 3 min at RT the beads were dried and resuspended in 1.1x µl water by vortexing. After that the mix was incubated at RT for 5 min and placed back on magnetic rack. After incubation at RT for 2 min, the supernatant was

transferred in a fresh tube. Reference for AMPure XP beads can be found in in Error! R eference source not found..

2.5.8. RNA Isolation of ESCs, ESCs MLL3/4 dKO/dCD, MEFs and NPCs

RNA isolation was performed with NucleoSpin RNA Plus Kit. Per sample one well on a 6-well plate was seeded accordingly. For MEFs after two days, and for NPCs at day 5, the cells were washed two times with PBS and 500 µl LBP was added and the mixes were stored at -80°C. The samples were thawed on ice. Up to 500 µl of the lysate was added onto a NucleoSpin gDNA Removal Column and centrifuged at 11,000 g at RT for 1 min. All flowthrough was transferred into a fresh tube. If needed, this step was repeated until all the sample went through the column. 0.3x volume of the flowthrough of Binding Solution BS was added to each sample, which was around 210 µl. Samples were mixed and 450 µl transferred to NucleoSpin RNA Plus Column. The columns were centrifuged at 11,000 g at RT for 1 min and flowthrough was discarded. This step was repeated until all the sample was loaded onto the column. 200 µl of Buffer WB1 was added for washing. The columns were centrifuged at 11,000 g at RT for 1 min and flowthrough was discarded. 600 µl of Buffer WB2 were added, the columns were centrifuged at 11,000 g at RT for 1 min and flowthrough was discarded. Additionally, 250 µl of Buffer WB2 were added, the columns were centrifuged at 11,000 g at RT for 1 min and flowthrough was discarded. The empty columns were centrifuged at 11,000 g at RT for 1 min to dry. The columns were transferred into a fresh 1.5 ml tube and 30 µl of RNase-free water was added and incubated at RT for 2 min. The columns were centrifuged with 11,000 g at RT for 1 min. The elution step was repeated for a total of two times within the same tube. Concentrations were measured by Qubit RNA HS Assay kit and the quality of RNA was analyzed by TapeStation HS RNA tape. For further details see Error! Reference source n ot found. and **Table** 3.

2.5.9. Depletion of rRNA for ESCs, ESCs MLL3/4 dKO/ dCD and MEFs

Removal of rRNAs from isolated samples of IFNβ stimulated ESCs and MEFs were done following the protocol of Ribo-Zero rRNA Removal Kit. In short, washing was done by adding 225 µl magnetic beads per sample and placed on a magnetic stand. Supernatant was removed and discarded. Beads were removed from stand and washed with 225 µl RNase-free water and vortex. Tubes were put back to magnetic stand and supernatant was removed. This washing step was repeated two times in total. After the last washing step, 60 µl Bead Resuspension Solution were added, vortexed and

65 µl magnetic beads per sample were transferred to a fresh tube. 1 µl of RiboGuard RNase Inhibitor was added per tube. Beads were kept at RT till usage. In the next step, the probes in the removal solution to hybridize to rRNA in the sample was prepared. I used 5 µg total RNA input per sample in a total volume of 26 µl. The samples were prepared in PCR tubes and 10 µl Removal Solution was added to a total volume of 40 µl and incubated for 10 min at 68 °C. The mix was spun down briefly and incubated for 5 min at RT. Then, 40 µl of RNA sample were transferred to the 1.5 ml tube, which contained the 65 µl washed magnetic beads. The solution was mixed and incubated for 5 min at RT. Tubes were placed for 5 min at 50 °C in a thermomixer after placing them back on magnetic stand. The supernatant, which contains the rRNA-free RNA, was transferred to a fresh 1.5 ml tube. The RNA was then cleaned-up by ethanol precipitation. RNase-free water was added to rRNA-free RNA sample to a total volume of 180 µl. In addition, 18 µl 3M sodium acetate and 2 µl of glycogen (10mg/ml) were added and the mix was vortexed. 600 µl of pure ethanol was added, vortex and frozen away O/N. Samples were centrifuged at 10,000 g at 4 °C for 30 min. Discard supernatant and washed by adding 200 µl freshly prepared 70 % ethanol. Samples were centrifuged at 10,000 g at 4 °C for 5 min and supernatant was removed. In total, two washing steps were performed. After the last one, the samples were briefly centrifuged and remaining liquids was removed. Lids were kept open and beads were dried for around 3 min. Beads were resuspended in 30 µl RNase-free water. Concentrations were measured by Qubit RNA HS Assay kit. Information for kits and instruments can be found in Error! Reference source not found. and **Table** 3.

2.5.10. Depletion of rRNA for NPCs

Isolated RNA was treated with DNase and incubated at 37 °C for 30 min. The RNA was purified by precipitation (3x volume of pure EtOH, 1/10 volume of sodium acetate and 2 µl of Glycoblue Coprecipitant) and placed at -20 °C O/N. The samples were centrifuged by 13,000 rpm at 4 °C for 30 min, washed two times with 70% EtOH and eluted in 30 µl water with 1 µl RiboLock. The concentration was measured by Qubit RNA HS Assay and quality of isolated RNA assayed by TapeStation High Sensitivity RNA ScreenTape. 750 ng of DNase-treated RNA was used for rRNA depletion based on NEBNext® rRNA Depletion Kit (Human/Mouse/Rat). The RNA samples were filled up to 12 µl and probes were hybridized to the RNA. Therefore 2 µl of Probe Hybridization Buffer and 1 µl NEBNext rRNA Depletion Solution were added. The mix was pipetted up and down ten times and incubated. Afterwards, the samples were briefly spun down and placed on ice. A RNase H Digestion was performed by adding 2 µl NEBNext RNase H, 2 µl RNase H Reaction Buffer, 1 µl of Nuclease-free Water.

After mixing by pipetting up and down for 10 times, the samples were incubated at 37 °C for 30 min. The sample were then treated with DNase I by adding 5 µl DNase I Reaction Buffer, 2.5 µl DNase I (RNase-free) and 22.5 µl Nuclease-free Water. The samples were mixed well and incubated at 37 °C for 30 min. Finally, the mix was purified by adding RNA Clean XP beads. 110 µl (2.2X) beads were added to the sample and mixed by vortexing. The mix was incubated on ice for 15 min, followed by a 5 min incubation on a magnetic rack to collect the beads. The supernatant was then carefully removed, and the beads were washed two times for 30 s with 80% freshly prepared EtOH. After the second wash, the beads were briefly spun down and placed back on the magnetic rack. The collected liquid was removed and the beads air-dried for a maximum of 3 min. 8 µl of Nuclease-free Water was added to the beads, vortexed and spun down. The mix was incubated at RT for 2min before it was placed back on the magnetic rack. 6 µl of the elution was transferred in a fresh tube and concentrations were measured by Qubit RNA HS Assay kit. For details see Error! Reference source n ot found. and **Table** 3.

2.5.11. RNA-seq library preparation for all samples

Purified rRNA-depleted RNA of ESC, MEF and NPC samples were used to prepare NGS libraries based on the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina. In short, 5 µl of RNA were mixed with 4 µl NEBNext First Strand Synthesis Reaction Buffer and 1 µl Random Primers by pipetting up and down for ten times. The RIN value of all samples were above 7 and therefore the mix were incubated at 94°C for 15 min. Following that, 8 µl of NEBNext Strand Specificity Reagent and 2 µl NEBNext First Strand Synthesis Enzyme Mix were added to 10 µl sample. After mixing by pipetting ten times, the samples were incubated at 25 °C for 10 min, 42 °C for 15 min and 70 °C for 15 min. The sample were placed on ice and 8 µl of NEBNext Second Strand Synthesis Reaction Buffer with dUTP Mix, 4 µl NEBNext Second Strand Synthesis Enzyme Mix and 48 µl Nuclease-free Water was added. The mixture was incubated at 16 °C for 1 h and afterwards purified with 1.8x AMPure Beads. The samples were eluted in 50 µl TE Buffer and directly processed with the End Prep of cDNA library. The sample in 50 µl of sample was mixed with 7 µl NEBNext Ultra II End Prep Reaction Buffer and 3 µl NEBNext Ultra II End Prep Enzyme Mix. After mixing by pipetting ten times, it was incubated at 20 °C for 30 min and then at 65 °C for 30 min. Afterwards the samples were placed on ice and the adaptor ligation was performed. 5-fold dilution of NEBNext Adaptor in Adaptor Dilution Buffer was used as default, for some samples with 10 ng input material a 25-fold dilution was used. 2.5 µl of diluted Adaptor, 1 µl of NEBNext Ligation Enhancer and 30 µl NEBNext Ultra

II Ligation Master Mix were added to the sample. The sample mix was pipetted up and down ten times and further incubated at 20 °C for 15 min. 3 µl USER Enzyme was added and further incubated at 37 °C for 15 min. Afterwards the samples were purified with AMPure XP Beads at a 0.9x ration and eluted in 15 µl 0.1x TE. For the PCR Enrichment of Adaptor Ligation DNA, unique i5 and i7 primers (**Table 7**) were used. 25 µl NEBNext Ultra II Q5 Master Mix, 5 µl of i5 primer and 5 µl of i7 primer were added to the 15 µl sample. The samples were mixed ten times and incubated in a thermocycler. For 50 ng samples a total of 9 x cycles were performed and the 10 ng samples for 11 x cycles. Afterwards the samples were purified with AMPure XP Beads with a 0.9x ratio and eluted 23 µl 0.1x TE. The concentration was measured with Qubit dsDNA HS assay kit and the fragment size was determined with Tapestation D5000 ScreenTape. For details to used chemicals and instruments see Error! Reference s ource not found. and **Table** 3.

PCR Adaptor Ligated DNA	Temperature	Time	Cycles
Initial Denaturation	98 °C	30 s	1 x
Denaturation	98 °C	10 s	9-11 x
Annealing/Extension	65 °C	75 s	9-11 x
Final Extension	65 °C	5 min	1 x
Hold	4 °C	∞	

2.5.12. ChIP of histone modifications in ESCs, MEFs and NPCs

These experiments were performed with technical assistance from Caroline Bauer. ESCs were cultured in 150 mm dished and treated with IFNβ (500U/ml) for 0 h, 1 h or 6 h. Media was removed and cells were detached with Accutase, washed once with 1xPBS/100mM PMSF and crosslinked with 1 % Formaldehyde (1 ml 16 % Formaldehyde with 15 ml 1xPBS) for 10 min at RT. 125 mM Glycine was added and samples were rotated at RT for 5 min. Afterwards three times washed with 1xPBS/100 mM PMSF and cell pellet was resuspended in 10 ml Swelling Buffer and 100 mM PMSF. The mixes were incubated on ice for 10 min and centrifuged with 2000 rpm for 5 min at 4 °C. The cells from two 150mm plates were combined to get a total cell number of around $40*10^6$. Cell pellet was resuspended in 100 µl MNase Buffer and 40 U MNase were added per sample. The Mix was incubated at 37 °C for 15 min. 100 µl 10x sonication buffer and 800 µl water were added on top and again an

incubation step was performed on ice for 5 min. Note, for replicate 1 of the histone ChIP in ESCs, the MNase treatment was not performed. The samples were sonicated with Covaris Sonicator. Therefore, samples were transferred into Covaris tubes (blue cap, 12x24mm) and shared for 15 min (Burst 200; Cycle 20%; Intensity 8). Afterwards, the samples were centrifuged with 13,000 rpm at 4 °C for 15 min. The supernatant was transferred into fresh tube. Chromatin was frozen with liquid N2 and stored then at -80 °C. After quality check of reverse cross-linked sample via 2 % agarose cell, the shared chromatin was around 150 bp and samples were used for the IP. Pre-equilibrate 25 µl Protein G Beads were added to each sample in Covaris sonication buffer and incubated at RT for 10 min. For pre-cleaning, to each sample 25 µl Protein G beads (not pre-equilibrated) and 4 µg IgG AB (rabbit or mouse) were added and incubated rotating at 4 °C for 2 h. Beads were pelleted, and supernatant transferred to fresh tubes. 4 µg of antibodies (or manufacturer's recommended amounts) were added to chromatin samples (**Table 2**) and incubate at 4 °C for 2 h. Then, 25 µl of pre-equilibrated beads were added to the samples and incubate rotating at 4 °C O/N. Beads were washed by rotating at 4 °C for 5 min with Covaris sonication, high-salt, Li-buffer and 2x with 1x TE-buffer. The final elute was done twice with 250 µl elution buffer at 37 °C for 15 min on shaker. Elutions were combined and reverse crosslinked by adding 20 µl 5 M NaCl and incubated at 65 °C O/N. 10 µl EDTA (0.5M), 0.5 µl RNaseA (10 mg/ml) and 50 µl Tris (1M, pH 6.8) were added and incubated at 37 °C for 30 min. Then 2 µl Proteinase K (20 mg/ml) were added and additionally incubated at 55 °C for 2 h. Precipitate purified DNA by adding 1 x volume isopropanol and 1/10 x volume ammoniumacetat. The mixes were placed at -20 °C for 60 min and then centrifuge by 14,800 rpm at 4 °C for 20 min. A washing step with 75 % EtOH was performed once and centrifuged by 14,800 rpm at 4 °C for 1 min. The pellets were air-dried and resuspended in water.

ChIP Buffers	
Sonication Buffer (Covaris)	10 mM Tris-HCl pH 8.0 200 mM NaCl 1 mM EDTA 0.1% Na-Deoxycholate 0.5% n-Lauroylsarcosine 1x Protease Inhibitor
Swelling Buffer	25 mM Hepes pH 7.8 1 mM MgCl2 10 mM KCl 0.1 % NP-40 1 mM DTT 0.5 mM PMSF

	1x Protease Inhibitor
Li Buffer	20 mM Tris-HCl pH 8.0 1 mM EDTA 250 mM LiCl 0.5 % NP-40 0.5 % Na-Deoxycholate 0.5 mM PMSF
High Salt Buffer	50 mM Hepes pH 7.9 500 mM NaCl 1 mM EDTA 1 % Triton-X-100 0.1 % Na-Deoxycholate 0.1 % SDS 0.5 mM PMSF
TE Buffer	10 mM Tris-HCl pH 8.0 1 mM EDTA
Elution Buffer	50 mM Tris-HCl pH 8.0 1 mM EDTA 1 % SDS 50 mM NaHCO3

2.5.13. Library preparation for histone modification ChIPs

The sequencing libraries were made based the Solexa ChIP Sample Prep Protocol (v1.3) by Caroline Bauer (https://www.crg.eu/en/content/genome-analyzer-and-hiseq2000-sample-preparation#a3). The barcoding details can be found in **Table 7**.

2.5.14. ChIP for TFs in ESCs, MEFs and NPCs

ChIP for TFs STAT1 and STAT2 were performed based on the Simple ChIP Enzymatic Chromatin IP Kit from Cell Signaling Technology. The kit requested around $4*10^6$ cells per IP. ESCs were cultured according to standard protocol and 2x150mm dishes were used per condition. The nature and specific growth behavior of ESCs allowed to have around $20*10^6$ cells per 150mm plate. ESCs were treated with self-made IFNβ (16.6 U/µl) to have a final concentration of 500U/ml on the plates. Samples were collected unstimulated (0h), 1h or 6h of IFNβ stimulation. Media was removed and plates were washed 1x MT-PBS. For ESC replicate 1, 277,7 µl 37 % Formaldehyde was added to 10 ml media and the media was put back onto the 150mm dished to cross-link the cells. For all further replicates, as well as all MEF and NPC samples, 16 % Formaldehyde was diluted with 15 ml PBS, and the cells were crosslinked by adding 9 ml 1 % Formaldehyde in PBS to the plates. The crosslinking was performed at RT for 10 min and stopped by adding 1 ml Glycine Solution (10x) and incubated at RT for 5min. All liquid was removed and each dish was washed twice with ice-cold 20 ml MT-

PBS. All plates were put into the fridge (4 °C). Single plates were processed by removing the last washing liquid and adding 2 ml Collection Buffer. Using a cell scratcher, the crosslinked cells were removed from the plates and collected in 15ml tubes on ice. All tubes were centrifuged with 2000 g at 4 °C for 5 min. Supernatant was removed, and cells were resuspended in 10 ml ice-cold Buffer A (used 1 ml per IP). The mix was incubated on ice for 10 min and inverted every 3 min. Another centrifugation with 2,000 g at 4 °C for 5 min was done. Supernatant was removed, and cells were resuspended in 10 ml ice-cold Buffer B (used 1 ml per IP). Samples were centrifuged with 2,000 g at 4 °C for 5 min and resuspended in 1 ml ice-cold Buffer B (100 µl per IP). The samples were transferred to 1.5 ml tube. For ESCs and MEFs no MNase treatment was preformed and tubes were directly centrifuged by 13,000 rpm at 4 °C for 1 min. to NPCs samples 5 µl MNase (0.5 µl per IP) were added and incubated at 37 °C for 20 min. Reaction was stopped by adding 100 µl of 0.5M EDTA (10 µl per IP) and then centrifuged by 13,000 rpm at 4 °C for 1 min. The nuclear pellet was resuspended in 1 ml ChIP Buffer (100µl per IP) and incubated at ice for 10 min. Each sample was spited up to two 1.5 ml tubes containing 500 ml samples, which would be similar to $20*10^6$ input cells. Sonication was done using EpiShear™ Probe Sonicator with certain number of cycles (30 s ON; 30 s OFF), amplitude 50 % at 4 °C. For ESCs and MEFs 10x cycles were performed. For NPCs 5x cycles resulted in the optimal result. Samples were centrifuged by 10,000 rpm at 4 °C for 10 min. Supernatant was saved and conditions were pooled. An aliquot of 50 µl from each condition was taken, 100 µl H_2O and Reverse Crosslinking Buffer was added. Those samples were then incubated at 65 °C O/N and the size of shared chromatin was checked by 2 % E-Gel. The main sample was frozen with liquid N_2 and stored at -80 °C.

After QC, chromatin was thawed on ice and concentration were measured by Qubit. 10 µg of chromatin was used per IP and diluted with ChIP Buffer to a total volume of 500 µl. Tubes were prefilled with ChIP Buffer and the required amount of chromatin was added to get 10 µg per IP. Antibodies were added based on the recommended dilution for ChIP (**Table 2**). Sample were incubated rotating at 4 °C O/N. 50 µl of chromatin were saved as input control. On the next day ChIP-grade Protein G Magnetic Beads was vortexed and 30 µl was added to each IP. Further the mixes were incubated rotating at 4 °C for 2 h. Tubes were placed on a magnetic rack. When the beads were collected on the wall, the supernatant was removed and 1 ml Low Salt Wash Buffer was added. The mixes were incubated at 4 °C for 5 min on a roller. In total, three rounds of washing with Low Salt Wash Buffer were performed. After the last removal of buffer, 1 ml High Salt Wash Buffer was added. Again, the incubation

was done at 4 °C for 5 min on a roller. Afterwards, the tubes were placed back onto the magnetic rack and supernatant was removed. Tubes were briefly centrifuged down, placed back and the remaining liquid was removed. 150 µl Elution Buffer was added to each ChIP sample and incubated at 65 °C for 30 min. In parallel, 150 µl Elution Buffer was added to each input sample and placed at RT till processing. ChIP samples were placed on a magnetic rack and waited until beads were collected at wall. Supernatant was transferred to fresh tubes and saved. Reverse Crosslinking Buffer was added to ChIP samples and input controls. All tubes were incubated at 65 °C O/N.

After cross-linking was reversed, samples were purified as followed. 750 µl DNA Binding Buffer was added and samples were briefly vortexed. A maximum of 450 µl of the mix was transferred to DNA spin column, centrifuged with 18,500 g at RT for 1 min and flowthrough was discarded. These steps were repeated till the entire samples were loaded on the same DNA spin column. 750 ml DNA Wash Buffer was added to columns and centrifuged with 18,500 g at RT for 1 min. Flowthrough was discarded and columns were centrifuged with 18,500 g at RT for 1 min. Columns were placed in fresh 1.5 ml tubes, 50 µl Elution Buffer was added and incubated at RT for 2 min. The samples were centrifuged with 18,500 g at RT for 5 min, columns were discarded afterwards. Samples concentration were measured by Qubit dsDNA HS Assay Kit and samples were stored at -20 °C. Kit and instrumental details were summarized in Error! Reference source not found. and **Table** 3.

ChIP Buffer	
Collection Buffer	2 ml MT-PBS 10 µl PIC (200x) 20 µl Phosphatase Inhibitor (100x)
1% Formaldehyde in media	270 µl 37 % Formaldehyde per 10 ml media
1% Formaldehyde in MT-PBS	1 ml 16 % Formaldehyde in 15 ml MT-PBS
Buffer A (1x)	750 µl H2O 250 µl Buffer A stock (4x) 0.5 µl DTT (1M) 5 µl PIC (200x) 10 µl Phosphatase Inhibitor (100x)
Buffer B (1x)	825 µl H2O 275 µl Buffer B stock (4x) 0.55 µl DTT (1M) 11 µl Phosphatase Inhibitor (100x)
ChIP Buffer (1x)	90 µl H2O 10 µl ChIP Buffer Stock (10x) 0.5 µl PIC (200x) 1 µl Phosphatase Inhibitor (100x)
Reverse Crosslink Buffer	6 µl NaCl (5M) 2 µl Proteinase K (20 mg/ml)
Low Salt Wash Buffer	2700 µl H2O 300 µl ChIP Buffer Stock (10x)

High Salt Wash Buffer	900 µl H2O 100 µl ChIP Buffer Stock (10x) 70 µl NaCl (5M)
Elution Buffer	75 µl H2O 75 µl ChIP Elution Buffer Stock (2x)

2.5.15. Library preparation for ChIP of TFs in ESCs, MEFs and NPCs

The sequencing libraries were prepared using NEBNext Ultra™ II DNA Library Prep Kit for Illumina. The barcodes used for each sample are summarized in **Table 7**. The recommended amount of input DNA for this kit was around 4 ng (range from 500 pg to 1 µg). As most ChIP samples were concentrated too low, we used 40 µl ChIP sample and added 10 µl 1x 10 mM Tris-HCl pH 8.0. For input controls, a 1:10 dilution was made and from this dilution, 4 µg chromatin were added and filled up with 1x 10 mM Tris-HCl pH 8.0 to a total volume of 50 µl. All samples were handled in PCR tubes. For the end repair, 3 µl of NEBNext Ultra II End Prep Enzyme Mix (green) and 7 µl NEBNext Ultra II End Prep Reaction Buffer (green) were added to the fragmented DNA. The samples were mixed by pipetting the entire volume ten times and briefly spun down. The tubes were placed into a thermocycler with lid temperature set to 75 °C at 20 °C for 30 min and then 65 °C for 30 min. For the adaptor ligation a sample amount of less than 5 ng was considered and then the adaptor solution was diluted 1:25 with 10 mM Tris-HCl pH 8.0. To the 60 µl End Prep Reaction Mixture 30 µl NEBNext Ultra II Ligation Master Mix (red), 1µl NEBNext Ligation Enhancer (red) and 2,5 µl NEBNext Adaptor for Illumina (red) was added a total volume of 93.5 µl. The entire volume was mixed by pipetting ten times and the mix was collected by a brief spin. The mixes were incubated at 20 °C for 15 min without heated lid. 3 µl of (red) USER™ Enzyme were added, mixed and further incubated at 37 °C for 15 min with heated lid set to 47 °C. Afterwards the mixes were cleaned up with AMPure XP Beads. The beads were placed at RT for at least 30 min, vortexed for 30 s and then a 0.9x cleanup was performed (see above). In short, 87 µl AMPure XP Beads were added to each mix and the elution was done in 17 µl 10mM Tris-HCl pH 8.0, from which 15 µl per sample was used for enrichment step. The PCR amplification mix was prepared and mixed by pipetting the entire volume ten times. The amplification PCR was run according to protocol. Afterwards samples were purified by performing a 0.9x Cleanup with AMPure beads and eluted in 30 µl of 10 mM Tris-HCl pH 8.0. Concentrations were measured by Qubit dsDNA HS Assay Kit and fragment distributions were analyzed by TapeStation D5000 ScreenTape run. Note, that for replicate 1 of ESCs, initially 11 cycles were run. The amplifications of many samples were insufficient so additional PCR amplification with four or five more cycles was performed. NGS

sequencing was performed with the DKFZ Genomics Core Facility. Error! Reference s ource not found. and **Table** 3 provide details for used kits and tools.

Volume	Reagent
15 µl	Adaptor Ligated DNA Fragments
25 µl	NEBNext Ultra II Q5 Master Mix
5 µl	Index Primer/i7 Primer
5 µl	Index Primer/i5 Primer
50 µl	Total Volume

Cycle Step	Temperature	Time	Repetitions
Initial Denaturation	98 °C	30 s	1 x
Denaturation	98 °C	10 s	3-15 x
Annealing/Extension	65 °C	75 s	3-15 x
Final Extension	65 °C	5 min	1 x
Hold	4 °C	∞	

2.5.16. ATAC-seq for ESCs, MEFs and NPCs

ESCs were plated on 6-well plates and treated for 0 h, 1 h or 6 h with self-made IFNβ (500 U/ml; stock: 16.6 U/µl). Cells were detached using Accutase, collected and washed with 1xMT-PBS. 50,000 cells were transferred into fresh tubes and centrifuged by 800 g at 4 °C for 5 min. For ESCs, the cell pellet (normally hardly visible) was resuspended in 200 µl ATAC lysis buffer and incubate for at RT for 2 min and centrifuged by 800 g at 4 °C for 5 min. The supernatant was discarded, and pellet was resuspended in 20 µl ATAC reaction buffer. Samples were incubated at 37 °C for 30 min. Reaction was stopped by adding 5 µl EDTA (100mM) in Tris-HCl pH 8.0 (final. conc. 20 nM). Library amplification PCR was started according to provided scheme. The resulting mix was purified with AMPure Beads (1.4x) and eluted in 32 µl water. For MEFs and NPCs, the cells were directly resuspended in 25 µl ATAC reaction buffer with Digitonin and incubated at 37 °C for 30 min. The samples were purified with a MinElute PCR Purification Kit and eluted in 12 µl EB. For MEFs and NPCs 11x cycles for the library amplification PCR was performed. Note, the used barcodes for each sample can be found in **Table 7**. The resulting mix was purified with 1.4x AMPure Beads followed by a size selection 0.5x/1.4x AMPure Beads. In short, 25 µl beads were added to the 50 µl sample and incubated at RT for 10 min. Afterwards, samples were placed on a magnetic rack until the beads were collected on the wall. Supernatant was transferred to a fresh tube and 45 µl beads were added. Note, that in combination with

the previous 25 µl beads this summed up to a 1.4x AMPure bead reaction. Samples were incubated at RT for 10 min and then placed on a magnetic rack. The supernatant was removed, and the beads were washed 2x with fresh 80 % EtOH. After the last washing step, the beads were briefly spun down and placed back to the magnetic rack. Remaining liquid was removed, and beads were dried for around 3 min. 21 µl EB was added to beads and incubated at RT for 5min. The tubes were placed back to the magnetic rack and 20µl of library samples were transferred to fresh tubes. Concentration was measured by Qubit dsDNA HS Assay Kit and run TapeStation D5000 ScreenTape to verify size distribution of final library. DNA concentration and mean fragment size (150 bp to 1,000 bp) for calculating required amount for multiplexing.

For ESCs

Volume	Reagent
5 µl	DNA sample
15 µl	H2O
25 µl	NEBNext Polymerase Mix
0.5 µl	Primer 1 (100 µM; Index 701-724)
0.5 µl	Primer 2 (100 µM; Index 501-524)
4 µl	MgCl2 (25mM)
50 µl	Total Volume

For MEFs and NPCs

Volume	Reagent
10 µl	DNA sample
10 µl	H2O
25 µl	NEBNext Polymerase Mix
2.5 µl	Primer 1 (25 µM; Index 701-724)
2.5 µl	Primer 2 (25 µM; Index 501-524)
50 µl	Total Volume

Library amplification PCR	Temperature	Time	Repetitions
Initial Extension	72 °C	5 min	1 x
Initial Denaturation	98 °C	30 s	1 x
Denaturation	98 °C	10 s	11-13 x
Annealing	63 °C	30 s	11-13 x
Extension	72 °C	30 s	11-13 x
Final Extension	72 °C	60 s	1 x
Hold	4 °C	∞	1 x

ATAC Buffers	
MT-PBS	4 mM NaH_2PO_4 16 mM Na_2HPO_4 150 mM NaCl
ATAC lysis buffer	10 mM Tris-HCl pH 7.4 10 mM $NaCl_2$ 3 mM $MgCl_2$ 0.1 % NP-40
ATAC reaction buffer	10 µl 2x Transposase buffer (Illumina) 2.5 µl Tn5 enzyme H_2O
ATAC reaction buffer with Digitonin	9.75 µl H_2O 12.5 µl 2x Transposase buffer (Illumina) 0.5 µl 50x Proteinase Inhibitor 2 µl Tn5 Enzyme (Illumina) 0.25 µl 1 % Digitonin

2.5.17. Single-cell RNA-seq (scRNA-seq) of ESCs

The experiment was performed based on the standard manufacturer's protocol of Chromium Single Cell 3' Reagent Kits v2. In short, $0.4*10^6$ ESCs were seeded on 6-well plate and on the following day treated with treated for 0 h, 1 h or 6 h with self-made IFNβ (500 U/ml; Stock: 16.6 U/µl). Cells were dissociated with Accutase, washed once with MT-PBS and resuspended in 100 µl MT-PBS. After counting, the cell number was adjusted based on the manufactural protocol to recover around 4,000 cells per condition. 50 µl of RT Reagent Mix, 3.8 µl of RT Primer, 2.4 µl of Additive A and 10.0 µl of RT Enzyme Mix were combined and the samples were added to a total volume of 90 µl. The Single Cell 3'Chip in the 10x Chip Holder and 90 µl of Master Mix plus cells were added to row 1. The Single Cell 3' Gel beads were vortexed for 30s and 40 µl were transferred into row 2. In row 3 270 µl of Partitioning Oil was added and the 10x Gasket was attached. The Chromium Controller was run for 6.5 min with the default setup. The Chip holder was placed in a stable 45° angle position and 100 µl of GEMs from the bottom of the well were transferred into a PCR tube and the GEM RT Incubation was performed. Afterwards, 125 µl were slowly added. After 1 min the 125 µl red agent-oil mix was removed from the bottom of the tube. DynaBeads were mixed by vortexing for 30 s and 200 µl were added to the remaining samples into the PCR tubes. The samples were mixed by pipetting five times and incubated at RT for 10 min. Elution Solution I was prepared by combining 98 µl Buffer EB, 1 µl 10 % Tween 20 and 1 µl Additive A. Samples were placed on a magnetic rack und all supernatant was removed. The attached beads were washed twice with 125 µl and once with 200 µl freshly prepared 80 % EtOH for 30 s each time. After the last washing

step, the PCR tubes were briefly spun down and placed again on the magnetic rack to remove remaining EtOH. The beads were air-dried for 1 min, removed from the magnetic rack and eluted in 35.5 µl Elution Solution I. After 1 min incubation time, the samples were placed back on the magnetic rack and after 1 min 35 µl of each purified sample was transferred into fresh PCR tubes.

GEM RT Incubation	Temperature	Time
Step1	53 °C	45:00 min
Step2	85 °C	5:00 min
Step3	4 °C	Hold

The cDNA Amplification Reaction was performed by adding 8 µl Nuclease-Free Water, 50 µl Amplification Master Mix, 5 µl cDNA Additive and 2 µl cDNA Primer Mix to the 35 µl purified GEM-RT product and mixing it by pipetting 15 times. The scDNA amplification was done by running 13 cycles.

cDNA Amplification	Temperature	Time	Cycles
Step 1	98 °C	3:00 min	1 x
Step 2	98 °C	0:15 min	13 x
Step 3	67 °C	0:20 min	13 x
Step 4	72 °C	1:00 min	13 x
Step 5	72 °C	1:00 min	1 x
Step 6	4 °C	Hold	1 x

The mix was purified with an AMPure Beads 0.6x cleanup and elution in 40 µl 10mM Tris-HCl pH 8.0. Concentration of purified GEM-RT product was measured by Qubit dsDNA HS Assay Kit and mean peak size of the sample was determined by TapeStation D5000 ScreenTape.

Samples which passed the quality control were used for the next step of the manufacturer's protocol, the Library Construction. The Fragmentation Mix was prepared by combining 10 µl Fragmentation Enzyme Blend and 5 µl Fragmentation

Buffer and added to 35 µl purified GEM-RT product. The End Repair and A-tailing was performed. Afterwards an AMPure Beads Size Selection 0.6x and a cleanup 0.8x was performed. The samples were eluted in 50 µl 10mM Tris-HCl pH 8.0. Adaptor Ligation Mix was prepared by mixing 20 µl Ligation Buffer, 10 µl DNA Ligase and 2.5 µl Adaptor Mix and added to the 50 µl sample. After mixing the samples were incubated 20 °C for 15 min and a 0.8x AMPure Bead purification was performed. The samples were eluted in 30 µl 10mM Tris-HCl pH 8.0.

End Repair and A-tailing	Temperature	Time
Pre-cool block	4 °C	Hold
End Repair	32 °C	5:00 min
A-tailing	65 °C	30:00 min
Hold	4 °C	Hold

To index the samples, 8 µl Nuclease-Free Water, 50 µl Amplification Master Mix and 2 µl SI-PCR Primer were added to the 30 µl purified samples. Additionally, 10 µl of specific Chromium i7 Sample Index was added (**Table 7**). The Sample Index PCR was performed with 10x cycles and a 0.9x AMPure bead cleanup was performed. The samples were eluted in 35 µl 10mM Tris-HCl pH 8.0. Concentration of scRNA libraries were measured by Qubit dsDNA HS Assay Kit and mean peak size of the sample was determined by TapeStation D5000 ScreenTape. The fragment size of all three libraries were found to be too big (higher than 650bp) and therefore additional 0.9x AMPure bead cleanup was performed. The samples were eluted again in 35 µl 10mM Tris-HCl pH 8.0. Concentration of scRNA libraries were measured by Qubit dsDNA HS Assay Kit and mean peak size of the sample were determined by TapeStation D5000 ScreenTape. The libraries were sequenced by the DKFZ Genomics Core Facility.

Sample Index PCR	Temperature	Time	Cycles
Step 1	98 °C	0:45 min	1 x
Step 2	98 °C	0:20 min	10 x
Step 3	54 °C	0:30 min	10 x
Step 4	72 °C	0:20 min	10 x
Step 5	72 °C	1:00 min	1 x
Step 6	4 °C	Hold	1 x

2.5.18. Single-cell ATAC-seq (scATAC) of ESCs

The experiment was performed based on the standard manufacturer's protocol Chromium Single Cell ATAC Reagent Kits. $1.0*10^6$ ESCs were seeded on 6-well plate and on the following day treated with for 0 h or 6 h with self-made IFNβ (500 U/ml; Stock: 16.6 U/µl). Cells were dissociated with Accutase, washed once with MT-PBS and resuspended in 100 µl MT-PBS. After counting the cells with the Luna cell counter, $1.0*10^6$ cells per condition were transferred in a fresh 2 ml tube, fill up to 1 ml with MT-PBS, centrifuged at 300 g at 4 °C for 5 min and resuspended in 100 µl chilled Lysis buffer. Samples were mixed by pipetting ten times up and down. Then the samples were placed on ice for 5 min. 1 ml of cold 1x Wash Buffer was added, mix-pipetted at least five times and centrifuged 500 g at 4 °C for 5 min. The supernatant was removed from the nuclei pellet. Based on the manufacturer's recommendations, we assumed a loss of around 50 % of all input cells. We aimed for 10,000 cells per condition, therefore, based on the provided Nuclei Stock Concentration Table. With an estimated remaining amount of $0.5*10^6$ cells left, so we added 100 µl Dilution Nuclei Buffer to obtain a concentration of 5.000 nuclei/µl. A 5 µl of this solution was used with the Chromium Single Cell ATAC Reagent Kits protocol.

Nuclei Isolation for scATAC-seq Buffers	Component	Stock	Final	
Diluted Nuclei Buffer	Nuclei Buffer Nuclease-free H2O	20x	1x	50 µl 950 µl
Wash Buffer	Tris-HCl pH 7.4 NaCl2 MgCl2 BSA Tween-20 Nuclease-free H2O	1 M 5 M 1 M 10 % 10 % -	10 mM 10 mM 3 mM 1 % 0.1 % -	100 µl 20 µl 30 µl 1 ml 100 µl 8.75ml
Lysis Buffer	Tris-HCl pH 7.4 NaCl2 MgCl2 Tween-20 NP-40 Digitonin BSA Nuclease-free H2O	1 M 5 M 1 M 10 % 10 % 5 % 10 % -	10 mM 10 mM 3 mM 0.1 % 0.1 % 0.01% 1 % -	50 µl 10 µl 15 µl 50 µl 50 µl 10 µl 500 µl 4.315 ml

The Transposition Mix was prepared by mixing 7 µl of ATAC Buffer with 3 µl ATAC Enzyme and placed on ice. Based on the sample stock concentration (5,000 nuclei/µl) we used 4 µl of our cell stock (based on the Nuclei Concentration Guidelines provided by the kit) together with 1 µl Diluted Nuclei Buffer and 10 µl of the Transposition Mix. The solutions were mixed and incubated on a thermocycler at 37 °C for 60 min with a lid temperature of 50 °C. The Chromium Chip E was placed into a 10x Chip Holder. In addition to my two ESC samples, five more were processes in parallel. The one remaining empty well in the Chromium Chip E was filled with 50 % Glycerol Solution, row labeled 1 with 75 µl, row labeled 2 with 40 µl and row labeled 3 with 240 µl. The Master Mix was prepared by combining 61.5 µl Barcoding Reagent, 1.5 µl Reducing Agent B and 2.0 µl Barcoding Enzyme. A total volume of 65 µl of the Master Mix was added to the 15 µl transposed nuclei and 75 µl of the mix was pipetted into row 1. After the Gel Beads strips were vortexed for 30 s and spun down, 40 µl of those beads were pipetted to row 2. Finally, 240 µl Partitioning Oli was added to row 3 and the 10x Gasket was added on top of the Chip holder. The Chip was carefully placed into the 10x controller and the scATAC program was started and ran successfully for 7 min. The Chip holder was opened and brought into a 45° position. The observed volumes were similar in all wells, therefore 100 µl of the GEMs were transferred into a pre-cooled PCR strip. The mixes were then incubated in a thermocycler running the following program.

ATAC Amplification	Temperature	Time	Cycles
Step 1	72 °C	5:00 min	1 x
Step 2	98 °C	0:30 min	1 x
Step 3	98 °C	0:10 min	12 x
Step 4	59 °C	0:30 min	12 x
Step 5	72 °C	1:00 min	12 x
Step 6	15 °C	Hold	1 x

A total of 125 µl of the Recovery Agent was added to each sample and the tubes were mixed by inverting ten times. A clear separation of two phases was observed and 125 µl of the red Recovery Agent at the bottom of the tubes were carefully removed. 200 µl of a cleanup mix containing 182 µl Cleanup Buffer, 8 µl Dynabeads MyOne SILANE and 5 µl Reducing Agent B were added to each sample. The samples were incubated for 10 min at RT. The tubes were placed on a magnetic rack, supernatant was removed and 300 µl freshly prepared 80 % EtOH was added for washing. The washing step was repeated once with 200 µl 80 % EtOH. Afterwards the tubes were briefly spun down, placed back to the magnetic rack and the remaining supernatant was removed. After this, 40.5 µl Elution Buffer I (98 µl Buffer EB, 1.0 µl 10 % Tween 20, 1.0 µl Reducing Agent B) was added to the beads and the mix was incubated for 2 min at RT. The tubes were placed back to the magnetic rack, and 40 µl supernatant was transferred to fresh tubes for the next steps. Beads were discarded and 48 µl of SPRIselect reagent was added to each sample and incubated at RT for 5 min. The tubes were placed on the magnetic rack and the supernatant was removed. The beads were washed two times with 200 µl 80 % EtOH for 30 s. Afterwards, the tubes were briefly centrifuged and the remaining supernatant was removed, after placing the tubes back to the magnetic rack. 40.5 µl Buffer EB was added and incubated at RT for 2 min. The tubes were placed back on the magnetic rack and 40 µl of supernatant was transferred to fresh tubes. The samples were stored at this save point O/N at -20 °C.

On the next day, the protocol was continued with the library construction part. The Sample Index PCR Mix was prepared by combining 50 µl Amp Mix with 7.5 µl SI-PCR Primer B. 57.5 µl of this mix was added to the 40 µl sample. 2.5 µl individual Chromium i7 Sample Index N, Set A was added to each sample (**Table 7**). The library was then prepared running the following program on a thermocycler.

Library Construction	Temperature	Time	Cycles
Step 1	98 °C	0:45 min	1 x
Step 2	98 °C	0:20 min	11 x
Step 3	67 °C	0:30 min	11 x
Step 4	72 °C	0:20 min	11 x
Step 5	72 °C	1:00 min	1 x
Step 6	4 °C	Hold	1 x

Afterwards, 40 µl SPRIselect reagents was added to the library mix and incubated at RT for 5 min. The tubes were placed onto a magnetic rack and the supernatant was transferred into fresh tubes and 74 µl SPRIselect reagents was added. The first round of beads was discarded. The sample mix was again incubated at RT for 5 min. The tubes were placed on the magnetic racks, supernatant was removed and the beads were washed with 80 % EtOH for 30 s each. After the last wash, the tubes were briefly centrifuged and placed back to the magnetic rack. Remaining liquid was removed and immediately 20.5 µl Buffer EB added. After a 2 min incubation at RT, beads were separated from the liquid using the magnetic rack and 20 µl of the supernatant were transferred in the final collection tube. The final library concentration was measured by Qubit dsDNA HS Assay Kit and mean peak sizes of the samples were determined by TapeStation D5000 ScreenTape. The libraries were sequenced by the DKFZ Genomics Core Facility. References for kits and instruments can be found in Error! R eference source not found. and **Table** 3.

2.6. Software

Tool name	Version	Reference	Parameters
Affinity Designer	1.8.3	Serif Europe	
Awk	4.0.2		-v chr="mt" '$1==chr' infile > outfile
BEDTools	2.27.1	(Quinlan and Hall 2010)	bedtools bamToBed -h; bedtools intersectBed -abam infile -b blacklist -sorted
BEDTools	2.27.1	(Quinlan and Hall 2010)	intersect -v -abam infile -b blacklist -sorted; bamToBed -i infile > outfile
Bowtie	1.2.2	(Langmead *et al.* 2009)	bowtie -t --chunkmbs 256 --best --strata -v 0 -m 1 index infile --sam outfile
Bowtie2	2.3.3.1	(Langmead and Salzberg 2012)	bowtie -p 8 -t --very-sensitive -X 2000 --fr --seed -x index -1 outfile1 -2 outfile2
DAVID	6.8	(Huang da *et al.* 2009)	
FastQC	0.10.0		fastqc -t 8 -o “output folder” “input folder“
HTSeq	0.12.4	(Anders *et al.* 2015)	Python: Htseq-count -m union -f bam -t intron -s no -i gene_id “infile” GRCm38.93_mm10_nesembl_20180727_with_Chr_intronic.gtf
IGV tools	2.3.23	(Thorvaldsdottir *et al.* 2013)	JAVA: igvtools count --windowSize 1 outfile1 outfile2 genome
IGVTools	2.3.23	(Thorvaldsdottir *et al.* 2013)	JAVA: igvtools jav -Xmx2g -jar IGVTools toTDF infile outfile mm
JAVA	1.8.0_252		
MACS2 (ATAC)	2.1.2	(Zhang *et al.* 2008)	Python: macs2 callpeak -t infile -B --nomodel --shift `expr 4 - 28/ 2` --extsize 200 --format BED --broad --broad-cutoff 0.1 --name outfile
MACS2 (ChIP)	2.1.2	(Zhang *et al.* 2008)	Python: macs2 callpeak --treatment infile --control ctl_infile -B --tsize 51 --gsize mm --bw 200 --format BED --pvalue 1e-5 --name outfile
MultiQC	1.7	(Ewels *et al.* 2016)	Python: multiqc -o outfolder infolder1 infolder2
phantompeakqualtools	1.14	(Marinov *et al.* 2014)	Rscript: phantompeak -c=infile.bam -s=-500:5:1500 -odir=outfolder -savp -out=outfile -rf

Python	2.7.12		
R	3.6.0		
REVIGO		(Supek *et al.* 2011)	
RSEM	1.3.0	(Li and Dewey 2011)	rsem-calculate-expression –no-bam-output –bam –forward-prob 0 “infile” ref_mm10_ensembl(index) “outfile”
RSeQC	2.6.6	(Wang *et al.* 2012)	Python: rseqc geneBody_coverage.py -r GRCm38.93_mm10_ensembl_20180727_ with_Chr.gtf -i “infile” -o “outfile”
RStudios	1.2.1335		
Samtools	1.3	(Li *et al.* 2009)	samtools rmdup -s infile outfile; samtools index outfile outfile.bai; samtools sort infile -o outfile
Samtools	1.3	(Li *et al.* 2009)	samtools sort infile -o outfile; samtools index outfile outfile.bai
SICER	0.1.1	(Xu *et al.* 2014)	python: sicer -t infile -c ctl_infile -rt 1 -w 200 -fs 150 -gs 0.7 -g 4 outfile
SortMeRNA	2.1	(Kopylova *et al.* 2012)	sortmerna --ref silva-euk-18s-id95.fasta silva-euk-28s-id98.fasta rfam-5.8s-database-id98.fasta -a 4 –reads “infile” –aligned “rRNA-reads” --other “mRNA-reads” –fastx
STAR	2.5.3a	(Dobin *et al.* 2013)	star –runThreadN 8 –runMode alignReads –runDirPerm All_RWX –genomeDir “star_index/mm10” –readFilesIn “infile” –outSAMtype BAM SortedByCoordinate –outWigType bedGraph –outSAMstrandField intronMotif –outReadsUnmapped Fastx –outFileNamePrefix “outfile_Aligned.out.bam” –quantMode TranscriptomeSAM GeneCounts –outFilterMultimapNmax 20 –sjdbGTFfile GRCm38.93_mm10_ensembl_20180727_ with_Chr.gtf –sjdbOverhang 49
Trimmomatic	0.36	(Bolger *et al.* 2014)	JAVA: trimmomatic PE -threads 8 -phred33 -trimlog infile1 infile2 ILLUMINACLIP:NexteraPE-PE.fa:2:30:10:8:TRUE SLIDINGWINDOW:5:20 MINLEN:25
DiffBind	2.12.0	(Ross-Innes *et al.* 2012)	Rscript
HOMER	4.9	(Heinz *et al.* 2010)	homer findMotifsGenome infile mm10 outfile -size given -mis 4 -len 8,9,10,11,12 -S 10 -bg mus_musculus.GRCm38.Regulatory_Build. regulatory_features

BEDTools	2.27.1	(Quinlan and Hall 2010)	bedtools sort -i infile | merge -i stdin -d 0 -c 2 -o count
BEDTools	2.27.1	(Quinlan and Hall 2010)	bedtools slop -i infile -g mouse -b 1000
BEDTools	2.27.1	(Quinlan and Hall 2010)	bedtools multiBamCov -bams readfile -bed targetregion -q 20

Table 5: Software versions and parameters

2.7. R packages

Package	Package Version	Package	Package Version
DeSeq2	1.24.0	***Seurat***	3.2.0
org.Mm.eg.db	3.8.2	***dplyr***	1.0.0
apeglm	1.6.0	***cowplot***	1.0.0
seqLogo	1.50.0	***scales***	1.1.1
pheatmap	1.0.12	***Cairo***	1.5-12.2
dendsort	0.3.3	***cluster***	2.1.0
GenomicRanges	1.36.4	***data.table***	1.12.2
ggplot2	3.3.2	***DiffBind***	2.12.0
ggpubr	0.4.0	***alluvial***	0.1-2
gridExtra	2.3.	***ggalluvial***	0.12.0
plyr	1.8.6	***RColorBrewer***	1.1-2
ggrepel	0.8.2		

Table 6: R packages with versions

2.8. Computational methods

2.8.1. RNA-pipeline – Annotation, mapping and quality controls

The RNA-seq-libraries were sequenced at the DKFZ Genomics Core facility. As our internal lab default, we performed sequencing of RNA-seq libraries single end (SE) 50 bp on HiSeq4000 and based on this I defined parameters for a computational pipeline for the mapping and annotations. The basic workflow of the pipeline was made by Dr. Nick Kepper. Demultiplexed, zipped fastq-files were the starting point of this pipeline. First, raw read files (fastq) were unzipped, renamed and the overall quality of those files was analyzed by FastQC. Remaining rRNA reads were removed by SortMeRNA followed by second analysis with FastQC. The remaining reads were aligned to ENSEMBLE mouse mm10 (https://www.ensembl.org/Mus_musculus/Info/Index) as reference using STAR resulting in three important output files, Aligned.sorted.ByCoord.out.bam, Aligned.to.Transcriptome.out.bam and ReadsPerGene.ou.tab. The first file was used to create TDF-files to visualize reads in the IGV browser tool and RSeQC to see read distribution over gene bodies as part of the QC. The second file was used by RSEM to calculate normalized read-counts (FPKM and TPM). The third file is a raw read count table (created by HTSeq, which is integrated in STAR), which was used as input for differential gene expression analysis by DeSeq2 in R studios.

Counting of intronic reads was done as follows. The reference GTF file was used to create a new GTF file containing only intronic regions. This new GTF file was used with HTSeq to create non-normalized count tables. Note, this tool was already integrated in STAR and usage of HTSeq with the default parameter on the originally GTF file resulted in the same results. Consequently, the same tool was used to create the regular and intronic count tables. TPM values were calculated by dividing the read counts by the length of each detected gene in kilobases (kb). Next, all counts per condition were summed up and divided by 1.000.000 to get a scaling factor. In the end, each read per kilobase from step one was divided by the scaling factor to become TPM values. Data fitting and visualizations were done with R using RStudios and the packages ggplot2. Parameters of tools and version were summarized in **Table 5** and **Table 6**.

2.8.2. Differential gene expression analysis with DeSeq2

STAR map reads were used as input for DeSeq2 to identify differentially expressed gene between two conditions. The default setup for DeSeq2 was used. For ESCs, MEFs and NPCs unstimulated samples (0 h) were compared with 1 h and 6 h IFNβ stimulated samples. For ESCs MLL3/4 dKO and dCD the comparison was between conditions as well as between cell types. MA plot visualization was done with DeSeq2 features and ISGs lists were extracted. For details see **Table 5** and **Table 6**.

2.8.3. Annotation of scRNA-seq and scATAC-seq with CellRanger

Cellranger, a tool provided by 10xGenomics, was used for demultiplexing and annotating of scRNAN and scATAC data. It was used with default setup. For scATAC, CellRanger was run for sequencing lanes separated and combined.

2.8.4. Analysis of scRNA by SEURAT

The analysis of annotated scRNA data by CellRanger was further analyzed by the R package SEURAT following a provided basal tutorial optimized for this type of data (https://satijalab.org/seurat/v3.2/pbmc3k_tutorial.html).

2.8.5. Analysis of scATAC co-accessibility

Co-accessibility analysis of ISGF3 binding sites with ISG promoters were done by Isabelle Lander, a PhD Student in the AG Rippe. For this analysis ArchR in R Studios was used on the default mouse mm10 reference genome. Cells were filtered using a lower and upper threshold for the number of detected fragments per cell ($10^{3.5}$ and 10^{5}), a transcription start site ratio above four and a ratio of fragments in blacklisted genomic regions to all detected fragments below 0.0225. The default ArchR workflow was used to calculate the co-accessibility scores between sites of interest. These sites were in a one megabase window (+/- 500 kb) around each ISGF3 binding site in ESCs. The binding site was embedded in a 2 kb tile and the rest of the window was further segregated into 2 kb tiles. By random shuffling of these tiles, a background score was calculated for each site and combined to define a threshold for all sites. On average a background score of 0.11 was detected, and by setting a threshold to 0.13, 99% of all detected average background signal was excluded. Therefore, only co-accessibility scores above 0.13 were considered to be real signals and used for analysis.

2.8.6. ChIP pipeline – Annotation, mapping and quality controls

NGS sequencing was performed for ChIP-seq-libraries at the DKFZ Genomics Core facility. As our internal lab default, we performed sequencing of ChIP-seq libraries SE 50bp on HiSeq4000 and based on this the computational pipeline was optimized for the mapping and annotations. The same scaffold as for the RNA-seq pipeline from Dr. Nick Kepper was used. The optimization and parameter settings for the required tools for ChIP-seq analysis was done by me. Demultiplexed raw reads were mapped using Bowtie. Following, by Samtools was used to remove duplicates, sort and index the mapped reads. Peaks were called either with MACS2 (narrow peaks) or SICER (broad peaks) based on Encode recommendations (https://www.encodeproject.org/data-standards/chip-seq/). IGVTools was used to create browser tracks for visualization. For quality measurements the fraction of reads in peaks (FRiP) scores, normalized strand coefficient (NSC) and relative strand correlation (RSC) values (phantompeak tools) for each sample was analyzed. Further FastQC and MultiQC was performed for each sample. Samples with low QC scores and few called peaks were removed from downstream analysis. The exact parameters and tools can be found in **Table 5** and **Table 6**.

2.8.7. ATAC pipeline – Annotation, mapping and quality controls

The ATAC-seq-libraries were sequenced at the DKFZ Genomics Core facility. All samples were sequenced at 50 bp paired end (PE). ESC samples were sequenced on a HiSeq 2000 while MEFs and NPCs were sequenced on a NovaSeq SE. The backbone of the analysis pipeline was established by Dr. Nick Kepper. The development and maintenance of the ATAC pipeline was done by Lara Klett, PhD student at the AG Rippe. Demultiplexed reads were processed with Trimmomatic to remove adaptor sequences and remaining reads were mapped using Bowtie2 with indices for mouse genome mm10. The mapped reads were sorted, indexed and PCR duplicates removed by Samtools. Reads mapped into regions marked in a blacklist were removed with BEDTools and a quality threshold was applied with Samtools. Each read was shifted by four basepairs at the positive strand and five at the negative. This was needed as the integration site of the transposase was a couple of nucleotides away from the finally detected fragment start site. After that, read files were transformed from bam to bed and all reads mapped to the mitochondrial genome were removed using the basic bash command awk. The remaining reads were used as input for peak calling by MACS2 and visualized with IGVtools. Quality was measured at multiple points by

FastQC, MultiQC, NSC, RSC and FRiP score calculations. The parameters and versions of the used tools was listed in **Table 5** and **Table 6**.

2.8.8. Venn diagrams

For all overlap diagrams (venn diagrams) samples were compared using Venny 2.1 (https://bioinfogp.cnb.csic.es/tools/venny/). The proportional overlaps were visualized by using a web tool (https://www.stefanjol.nl/venny) and used as template to prepare the blots with Affinity Designer.

2.8.9. GO-Term analysis

Functional annotations of differentially expressed gene lists were done by DAVID (https://david.ncifcrf.gov/) to identify gene ontology (GO) terms. The resulting terms were summarized by using REVIGO (http://revigo.irb.hr/).

2.8.10. Establishing differently bound STAT sites by DiffBind

DiffBind was used to combine MACS2 called peaks for different STAT ChIP replicates and the resulting peak files contain significantly different peaks between unstimulated and stimulated samples. Those lists were used for downstream analysis. For details see **Table 5** and **Table 6**.

2.8.11. Motif analysis by HOMER

The TF ChIP-seq peaks predefined by DiffBind were used as input for HOMER findMotifsGenome. As background file a list of regulatory regions annotated for mm10 genome (http://www.ensembl.org/info/genome/funcgen/regulatory_build.html, (Zerbino *et al.* 2015)) was used. Enrichment for known HOMER motifs as well as *de novo* motifs were identified in the provide peak lists. The resulting list of known HOMER motifs were further visualized by RStudios using ggplots2 (**Table 6**). Reference motifs for members of STAT and IRF family were downloaded from HOMER (http://homer.ucsd.edu/homer/motif/HomerMotifDB/homerResults.html).

2.8.12. Normalizing read counts and clustering

List of ISGF3 bound sites in ESCs and MEFs or STAT2 bound sites in ESCs and NPCs were used combined by bash basic commend cat, sorted and merged by BEDtools. Then those sites were centered and expanded by 1000bp up and down streams to create regions covering the surrounding of the TF binding sites. Those defined regions were used as input to count reads of ChIP-seq results of various histone marks and ATAC signals. Counts were normalized by library depth and fragment length. For ChIP data, enrichments over corresponding controls (H3 for histone ChIPs, IgG Rb for TF ChIPs) were calculated. Finally, replicates of the same samples and time points of IFNβ stimulation were merged. Those count tables were used as input for the kmeans clustering to characterize the chromatin environment of TF binding sites. Calculation of Silhouette coefficient was done to identify the optimal number of clusters was performed in RStudios (**Table 6**).

2.8.13. GREAT

The peak files for STAT complex from DiffBind were used as input file for GREAT (http://great.stanford.edu/public/html/; 4.0.4) and the provided Mouse GRCm38 (mm10) annotation was used for the analysis. The parameter “Single nearest gene” was used. The resulting list of nearest genes was visually optimized in RStudios with ggplots2 (**Table 6**).

3. Results

3.1. IFNβ dependent gene expression patterns

3.1.1. IFNβ induced hundreds of genes in ESCs, MEFs and NPCs

Mouse embryonic stem cells (ESCs), ESC-derived neuronal progenitor cells (NPCs) and mouse embryonic fibroblasts (MEFs) were treated with interferon beta (IFNβ) for 1 h and 6 h to characterize the gene response by RNA-seq on a genome-wide scale. Transcriptional profiles of 1 h or 6 h IFNβ treated cells were compared to unstimulated controls (0 h).

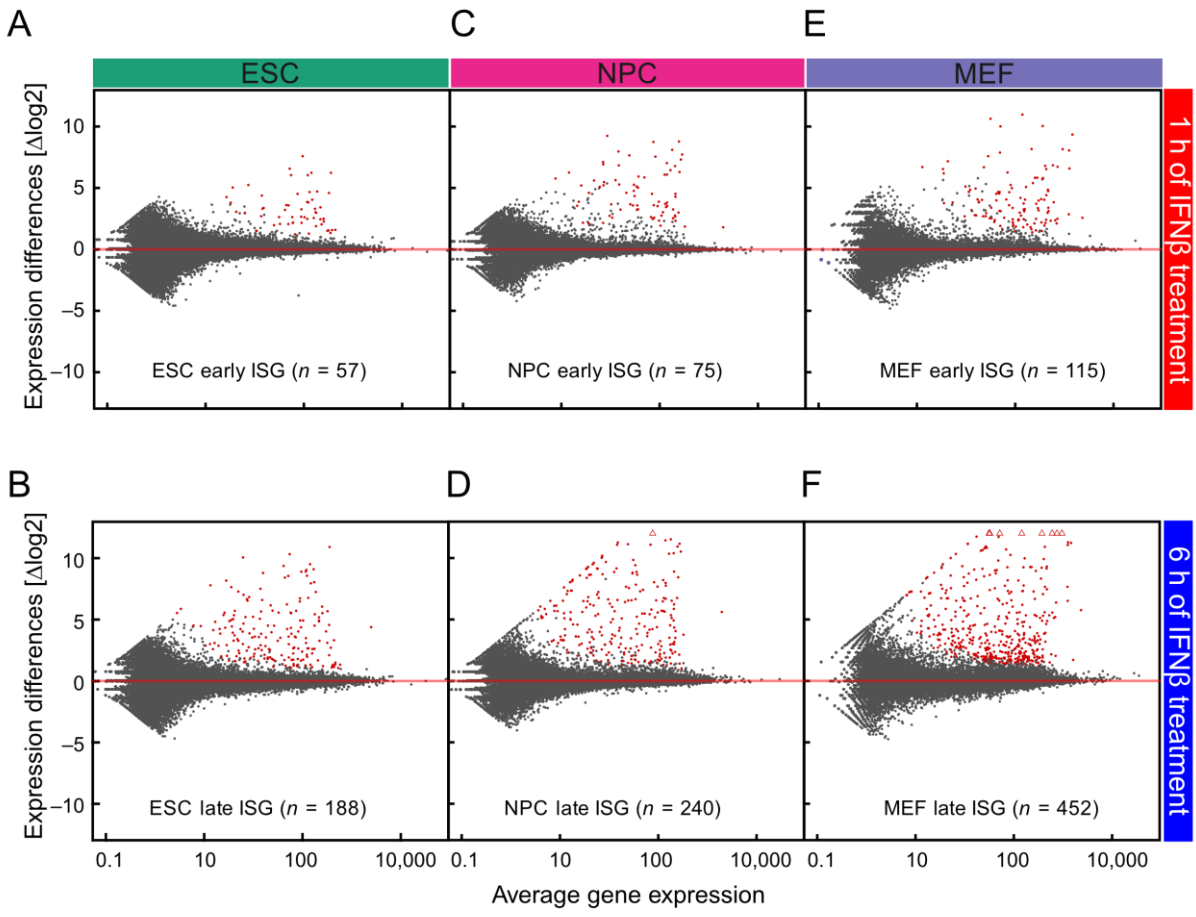

Figure 8: Differential gene expression analysis upon IFNβ treatment.

MA plot visualization of differentially expressed genes between IFNβ stimulation and untreated controls by DESeq2. Gene with significantly different expression levels (padj<0.05 and 1.5 - fold upregulated) were marked red. The x-axis showed the averaged gene expression of a gene over all samples. The y-axis represented the log2 fold change between both conditions. (A) In ESCs a total of 57 genes were found to be upregulated after 1 h IFNβ stimulation and 188 after 6 h stimulation in ESCs (B). (C) 75 genes were significantly upregulated in NPCs after 1 h and (D) 240 ISGs at the later stimulation time point. (E) Simulations in MEFs at 1 h and (F) 6 h, resulted in 115 early and 452 late responders.

Genes with differential expression levels (padj<0.5) and upregulated by at least 1.5-fold were identified as interferon stimulated genes (ISGs) by DESeq2. In ESCs, I identified 57 genes upregulated after 1 h and 188 after 6 h (**Figure 8 A/B**). In NPCs more ISGs were found after 1 h (n=75) and 6 h (n=240) of treatment compared to ESCs (**Figure 8 C/D**). The highest number of induced genes were identified in MEFs, where 115 genes were significantly upregulated after 1 h and 452 after 6 h (**Figure 8 E/F**). No significantly downregulated gene was identified in any cell line over the time course.

The overlap of early and late responding genes in ESCs resulted in 54 of 57 early ISGs to be found still upregulated after 6 h of stimulation (**Figure 9A**). Only 3 genes were only found differentially expressed at the 1 h time point. In total, I identified 191 genes to be responsive in ESCs. The same pattern was observed in NPCs and MEFs (**Figure 9B/C**). 71 out of 74 early ISGs in NPCs and 104 of 115 in MEFs were also responding at the later time point. Very few genes, 4 and 11, were found to be specifically upregulated at 1 h but not further at 6 h. In all three cell lines, the majority of genes, between 134 in ESCs and 348 in MEFs, were found to be only responding only at 6 h. In summary, the majority of early responding ISGs were also upregulated at 6 h of IFNβ stimulation.

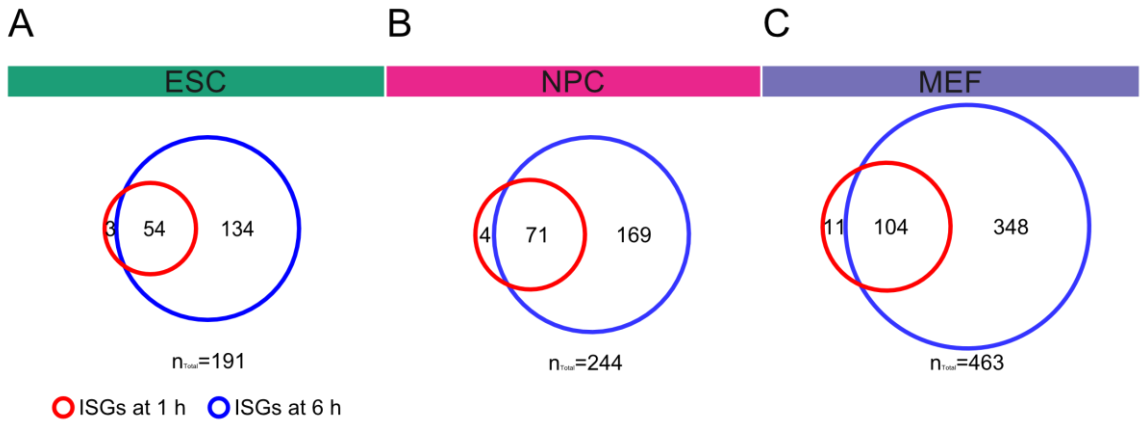

Figure 9: Overlap of early and late responding ISGs in three cell types

Overlaps of ISGs upregulated at 1 h (red) or 6 h (blue) in ESCs (A), NPCs (B) and MEFs (C). In total 191 ISGs were identified in ESCs, 244 in NPCs and 463 in MEFs.

3.1.2. Gene ontology (GO) terms reveal enrichment of innate immunity terms

As next step, we validated the ISG responses of each cell type by gene ontology (GO) term analysis. Thereby, enriched GO-terms were identified by DAVID (Huang da *et al.* 2009) and summarized by REVIGO (Supek *et al.* 2011) (**Figure 10**). The terms "Defense Response to Virus", "ISG15-Protein Conjugation" and "Negative Regulation Of Viral Genome Replication" were enriched in ESCs and MEFs (**Figure 10A**, yellow term). The first two were also found in NPCs. The terms "Purine Nucleotide Biosynthesis" and "Lipoprotein Metabolism" (green term) were shared between ESCs and NPCs, reflecting the close connection of those two cell types. In MEFs "Positive Regulation Of Type I Interferon Production" and "Cellular Response To Interferon-Beta" were the strongest enriched GO-terms (**Figure 10B**).

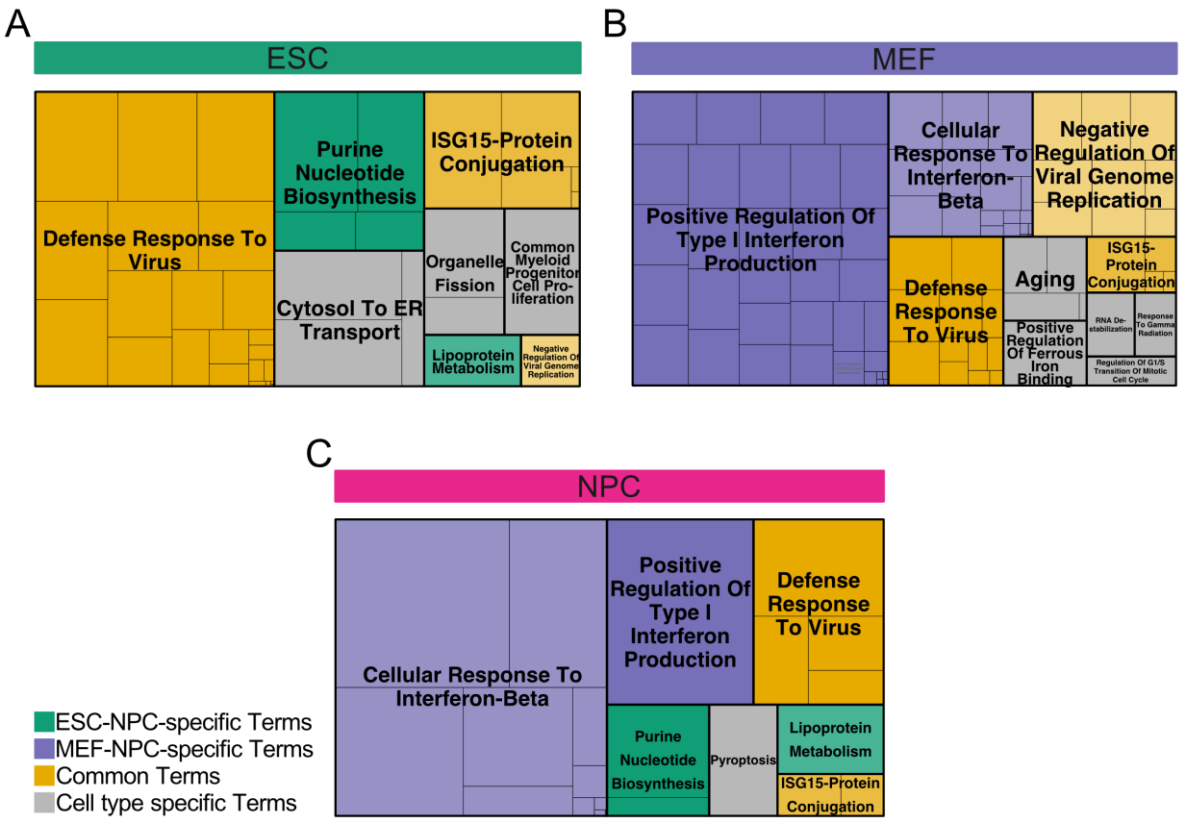

Figure 10: GO-term analysis of ISGs in ESCs, NPCs and MEFs

GO-Term analyses were performed by providing ISG list from DESeq2 to the webtool DAVID followed by REVIGO (Supek *et al.* 2011). (A) ESCs were enriched for shared (yellow), shared with NPCs (green) and unique terms (grey). (B) In MEFs other innate immune relevant terms were identified (purple) and shared with NPCs. Common and other unique terms were found as well. (C) The ISGs in NPCs represented a mix of shared terms with ESCs or MEFs only but also common with both other cell types.

Those two were shared with NPCs but not with ESCs. NPCs exhibited terms from both the cell type of origin (ESCs) as well from second differentiated cell type (MEFs), underlining the intermediate position of NPCs in respect to interferon responses (**Figure 10C**). In summary, the overall responses in all three cell types were strongly associated with innate immunity supported by shared GO-terms. However, we also identified cell type specific differences between ESCs and MEFs, while NPCs showed an intermediate of terms from both other cell types. As the majority of upregulated genes were confirmed to be innate immune related, I further investigated the cell type specificity of ISGs.

3.1.3. ISGs common in all three cell types were identified upon IFNβ treatment

The cell type specific ISG response was characterized by overlapping the ISG lists from the ESCs, NPCs and MEFs (**Figure 11**). I identified between 143 to 227 shared ISGs between different combinations of cell types (**Figure 11A/B/C**, yellow part). As the highest number of ISGs was found in MEFs, the comparisons also resulted in most MEF-specific ISGs. The comparison between NPCs and MEFs showed 17 NPC-specific genes (**Figure 11B**). ESCs showed in both comparisons with differentiated cells comparable numbers of 33 and 48 ESC-specific ISGs (**Figure 11A/C**). The comparison of ESCs to NPCs resulted in a set of 101 NPC-specific ISGs and 48 ESC-specific ISGs (**Figure 11C**). The overlap of all three cell types, showed that ESC-specific and NPC-specific ISGs were exclusively defined by the overlap with MEFs (**Figure 11D**).

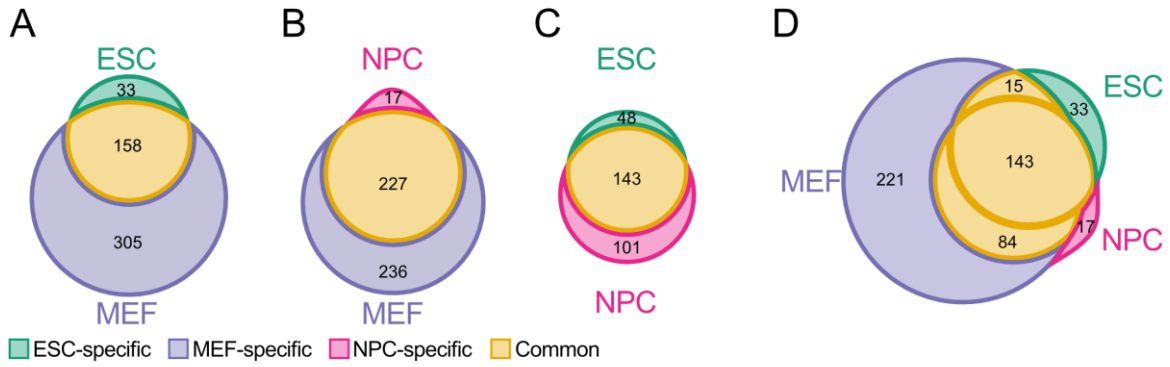

Figure 11: Cell type specific ISG patterns

(A) The overlap of ISGs detected in ESCs and MEFs. (B) Overlap between the two differentiated cell types NPCs and MEFs. (C) ISGs from ESCs compared with ISGs from NPCs. (D) Comparisons between all three cell types resulted in 33 ESC-specific, 17 NPC-specific and 221 MEF-specific ISGs. The majority of 242 ISGs were shared between cell types.

Most NPC-specific ISGs from the ESC comparison (**Figure 11C**) were also found in MEFs. In summary, this data showed a high overlap of ISGs between all three cell types and a high number of MEF-specific ISGs. Nevertheless, certain ISGs failed to respond in differentiated cells and were consider to be ESC-specific (n=33).

I defined a threshold to distinguish between expressed and silenced genes. This analysis was done using normalized RNA read counts called TPMs (Transcripts Per Kilobase Million). TPMs allow a comparison between conditions as these values are additionally normalized with a scaling factor. First, the TPM values of the entire transcriptome were plotted as histogram for each cell type in unstimulated conditions. Two populations of genes were identified underlying the bimodal distributions. In ESCs (**Figure 12A**), a strong peak was found for low expressed values, indicating a high number of genes with no or low TPM values. A distinct second peak represented a number of expressed genes. The TPM value of the point of intersection was defined as threshold to distinguish not expressed genes from expressed gene. For example, in ESCs the point of intersect was at 1.212 (natural-log scale), which represents a TPM value of 3.36. The same was performed for MEFs (**Figure 12B**) and NPCs (**Figure 12C**). In those cell types, also the two expected populations were found and a TPM threshold was identified at 2.89 in MEFs and 3.00 in NPCs.

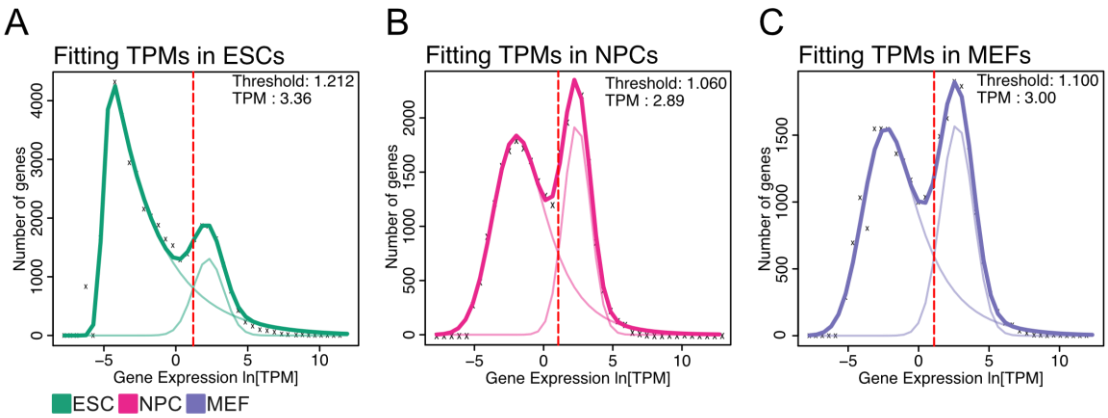

Figure 12: Determination of thresholds to distinguish expression states of genes

(A) Histogram fit of normalized read counts (TPMs) in unstimulated ESCs. The range of gene expression values were binned in 50 parts on the x-axis and plotted as natural logarithm. The y-axis showed the number of genes per bin. The tick green curve was plotted onto the resulting histogram and two curves, thin green lines, were fitted with two Gaussian distributions to explain the data distribution. The intercept was marked by a dotted red line and represented the threshold between actively expressed and repressed genes. The same approach was performed for NPCs (B) and MEFs (C).

In both differentiated cell types the expressed gene population was greater than the non-expressed gene population. In summary, thresholds to classify gene expression levels were identified on a cell type specific level. These results enabled to characterize the ISG dynamics in more depth.

The set of common ISGs allowed us to investigate the dynamics and gene induction levels in ESCs, NPCs and MEFs. Despite the fact, that the overall number of ISG were lower in ESCs, I was wondering, if the expression levels of common ISGs were different as well. Therefore, I characterized a set of well-established ISGs, which were upregulated in all three cell lines. These ISGs were grouped based on their basal gene expression levels. The basal gene expression levels (black) of *Irf9*, *Stat1* and *Stat2* were above the calculated TPM threshold for expressed genes (red line) in all three cell types (**Figure 13A**).

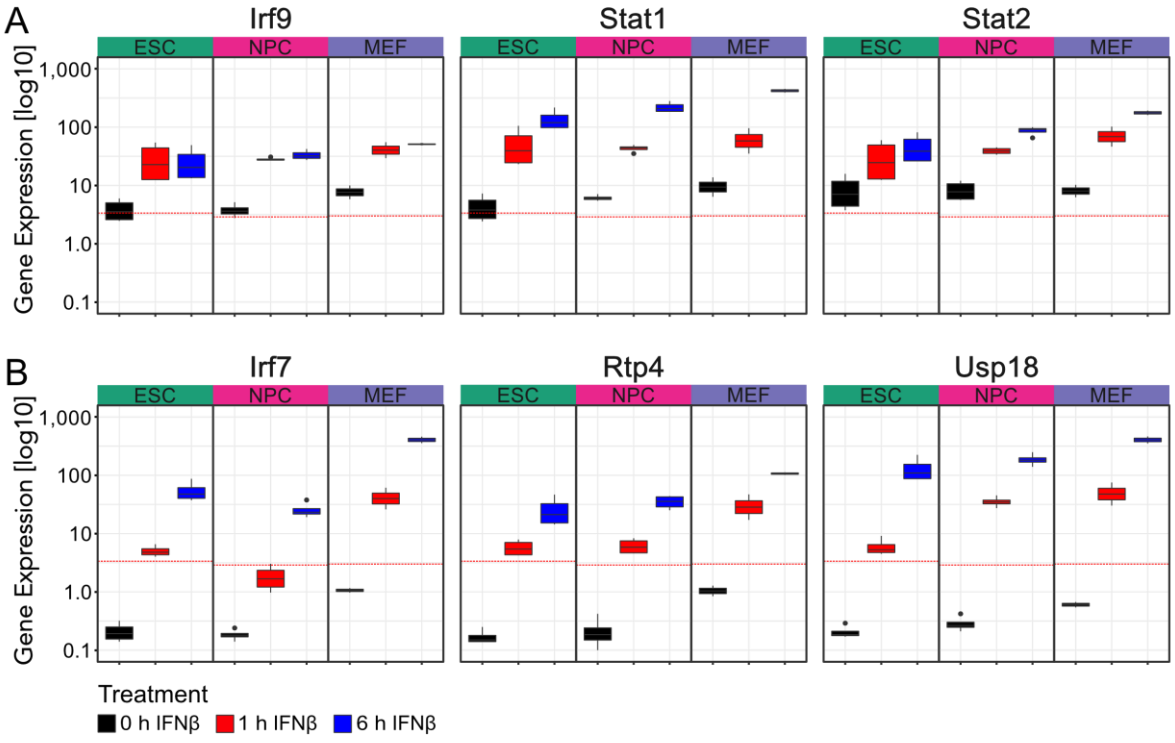

Figure 13: Examples for ISGs common to all three cell types

Boxplots of TPM values selected genes over IFNβ time course in ESCs, NPCs and MEFs. Unstimulated conditions (0 h) were colored in black, 1 h of IFNβ in red and 6 h in blue. The dotted red, horizontal lines marked the previously defined expression threshold for the corresponding cell type. (A) The genes *Irf9*, *Stat1* and *Stat2* were found to be expressed before stimulation and got upregulated upon IFNβ treatment. (B) *Irf7*, *Rtp4* and *Usp18* were silenced in untreated cells and found to be activated in all three cell types.

The expression patterns upon stimulation were highly comparable between cell types for all three genes. *Stat1* and *Start2* RNA levels reached a 5 to 10-fold higher level at 6 h of IFNβ treatment in MEFs compared to ESCs. The attenuated ISG response in ESCs was reflected in the induction levels of this three ISGs. For *Irf9* this difference was not as pronounced. The induction levels of NPCs were found in-between ESCs and MEFs. In contrast, *Irf7*, *Rtp4* and *Usp18* had basal gene expression levels below the expression threshold and were considered to be repressed before IFNβ stimulation (**Figure 13B**). Upon stimulation, these ISGs got strongly induced after 1 h and 6 h of IFNβ treatment. Here, the induction levels after 1 h of IFNβ treatment resulted in a much stronger induction in MEFs compared to ESCs. Additionally, the 6 h RNA levels were higher as well. Although, the dynamics of gene induction were comparable, the expression levels in ESCs were found to be attenuated by a magnitude higher than in MEFs and NPCs. We tested multiple hypothesis to explain this observation.

3.1.4. IFNβ response occurred at the gene induction stage

The RNA-seq analysis conducted above assesses steady state levels of spliced and processed mRNAs. In order to test if IFNβ response involved changes in mRNA stability, I conducted a differential gene expression analysis of nascent mRNAs. I hypothesized that on nascent RNA levels, the gene inductions of ESCs was more similar to MEFs and that regulations after gene inductions contributed to the attenuated ISG response observed previously. Therefore, intronic reads were counted, which are unique features of nascent mRNAs. The same bulk RNA-seq data were used for new counting on a self-assembled file containing only intronic sites only. Thereby, only levels of nascent mRNAs were detected and used for differential gene expression analysis. After 6 h of IFNβ stimulation, 77 ISGs in ESCs, 121 in MEFs and 440 in MEFs (**Figure 14A**). In combination with the ISGs upregulated at 1 h of stimulation, a total of 82 ISGs in ESCs, 128 in NPCs and 453 in MEFs were found to be. This showed that the total number of ISGs in ESCs was lower than in the regular RNA-seq analysis. Most of these genes were shared between cell types (**Figure 14B**), similar to the classical differentially gene expression analysis (**Figure 11**). In summary, the lower numbers of induced ISGs in ESCs were found on nascent RNA levels as well. Thus, we conclude that the lower mRNA levels in ESCs occurred by decreasing the transcription initiation of these genes but that variations in ISG RNA stability were negligible.

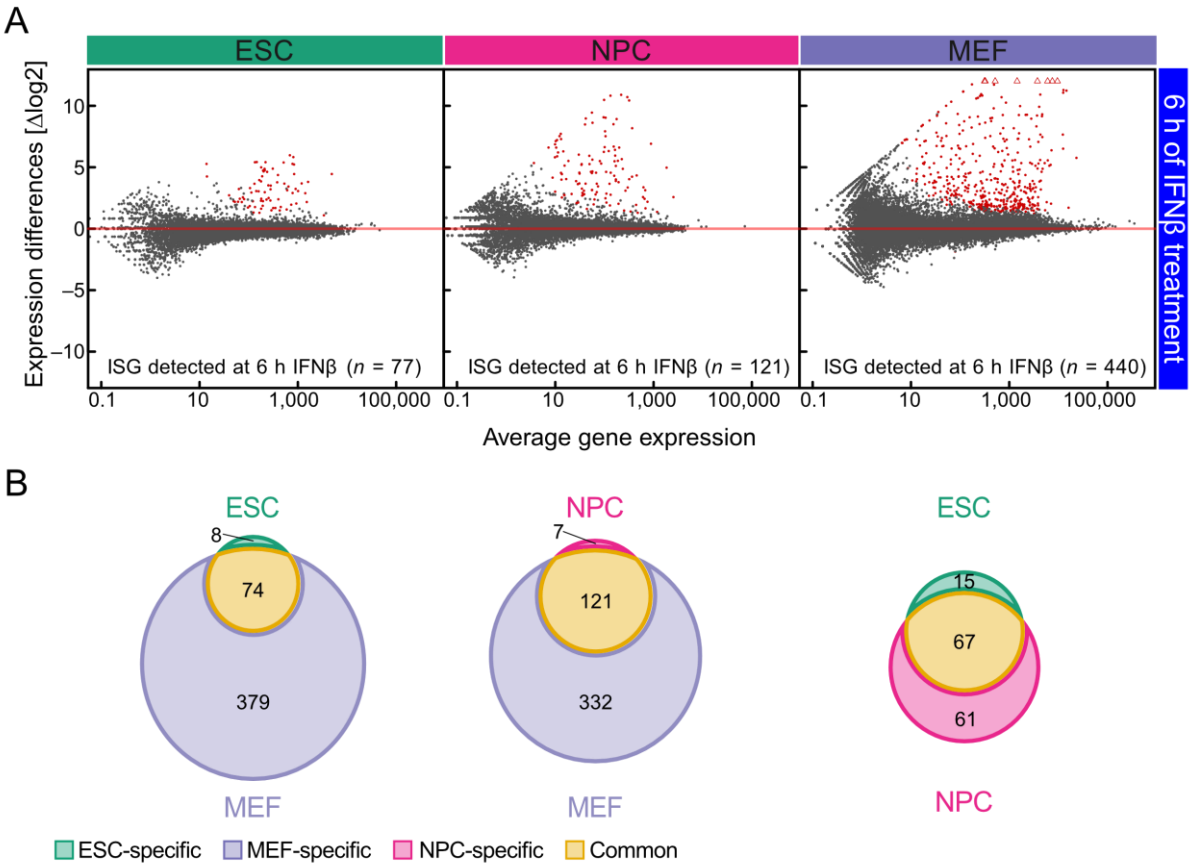

Figure 14: Nascent mRNA expression analysis

(A) MA plot visualization of differentially expressed nascent genes between 6 h IFNβ stimulation and untreated controls (0 h) by DESeq2. Significantly different expressed genes (padj<0.05 & 1.5 - fold upregulated) were marked red. (B) Overlaps of nascent ISG lists from ESCs with MEFs (left), NPCs with MEFs (middle) and ESCs with NPCs (right).

3.1.5. Single cell analysis revealed a mostly homogenous response

Next, it was tested if the transcriptional response in ESCs was homogeneous or if the observed upregulation of ISGs arises from a subset of very strong responding cells while others do not respond. The heterogeneity of innate immune response was assessed by single cell RNA sequencing (scRNA-seq) using the drop-seq methods on the Chromium platform from 10x Genomics. ESCs were treated with IFNβ for 1 h and 6 h and scRNA-seq libraries were generated and sequenced. For quality filtering, cells within a certain percentage of mitochondrial reads (2.5 % < accepted cells < 7.5 %) (**Figure 15A**) and number of detected genes (2000 < accepted cells < 6500) were selected (**Figure 15B**), yielding 1,332 cells for time point 0 h, 2,085 cells for 1 h and 4,825 for 6 h IFNβ stimulation. Therefore, the optimal number of principal components (PC) was determined and the first 20 PCs from the entire transcriptome were used for the reduction of dimensions (**Figure 15C**). For the bulk ISG focused analysis the first 8 PCs were used (**Figure 15D**).

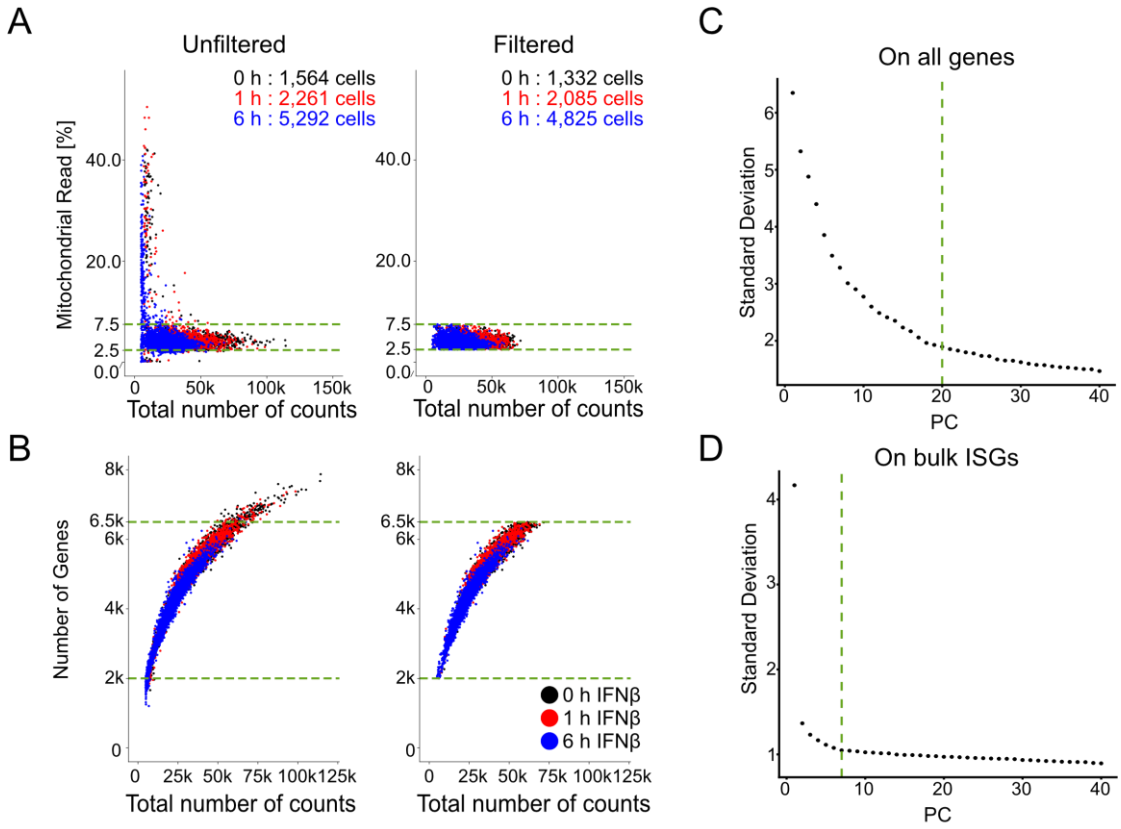

Figure 15: Quality assessment for scRNA-seq data of IFNβ stimulated ESCs

(A) The percentage of reads mapped to the mitochondrial genome were plotted against the total number of reads per cell. All cells with less than 2.5 % or more than 7.5 % mitochondrial reads (green lines) were removed (right). (B) The number of detected genes per cell was plotted against the total number of reads for all cells. The unfiltered cells (left) showed all cells and and the remaining cells after the application of all QC cutoffs (right). More than 2,000 and less then 6,500 genes were the cutoffs to keep a cell in the analysis (green lines). (C) Elbow plots showing the explained standard deviation of each principal component (PC) using the top 2,000 most variable genes from the entire transcriptome for the analysis. Number of finally used PCs for downstream analysis were set to 20 (green line). (D) Same as (C) but for ISGs instead of the entire transcriptome. ISGs were based on the bulk RNA-seq ISG list identified in ESCs.

Low-dimensional embedding based on the 2,000 most variable genes of the entire transcriptome resulted in the clear separation of cells treated for 6 h IFNβ (blue) from other conditions (**Figure 16A**). No separation between unstimulated (0 h, black) and 1 h IFNβ (red) cells was observed. Indicating, that the stimulation with IFNβ does not result in strong enough upregulation of ISGs on single cell levels to distinguished these two groups in reduced dimensions. The conclusion was that the transcriptional changes were sufficient to separate the 6 h from the 1 h and 0 h cells but failed to separate the early time point.

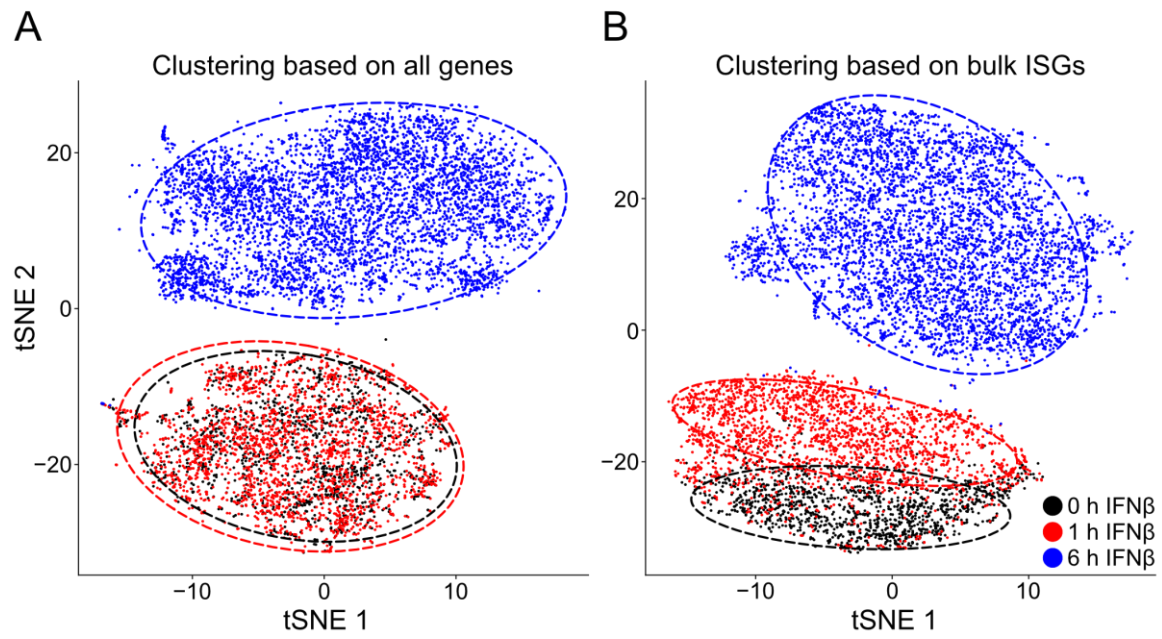

Figure 16: Low-dimensional embedding of ESCs treated with IFNβ from scRNA-seq

(A) A t-SNE representation of the low-dimensional embedding of ESCs stimulated with IFNβ based on the entire transcriptome. The coloring represented the treatment condition. Circles were added manually to visualize the location of the majority of cells for each group. (B) Same as in (A). The low-dimensional embedding was based on 191 ISGs, previously identified in bulk RNA-seq data of ESCs.

As next step, we aimed to identify a pattern of separation between 0 h and 1 h cells. Therefore, we performed the embedding based on previously identified list of bulk ISGs instead of the entire transcriptome. With this approach, a separation within the mix embedding of 0 h and 1 h cells were observed (**Figure 16B**). The unstimulated cells (black) separated partially from the 1 h time point (red), which ended up in-between untreated and the 6 h IFNβ treated cells (blue). The initial response after 1 h was heterogeneous, but after the longer stimulation (6 h) all cells responded homogeneously. Separated grouping of the 6 h time point was also seen with this approach, showing that the main factor behind this separation was the induction of ISGs. This was confirmed by analyzing the gene induction of specific ISGs. Genes, identified as ISGs in all three cell types by bulk RNA data (common), strongly responded in single cell RNA data as well (**Figure 17A**). A strong upregulation for *Stat1*, *Stat2 Irf9* and *Rtp4* was already observed after 1 h and further after 6 h. *Irf7* and *Usp18* were detected strongly activated after 6 h only. This showed that strong effects occurred robustly in many cells at 6 h, while the earlier time point showed many cells not or very weakly responding. Cell type specific ISGs like *Ifi27*, an ESC-specific gene, also displayed a strong response in the scRNA-seq analysis (**Figure 17B**).

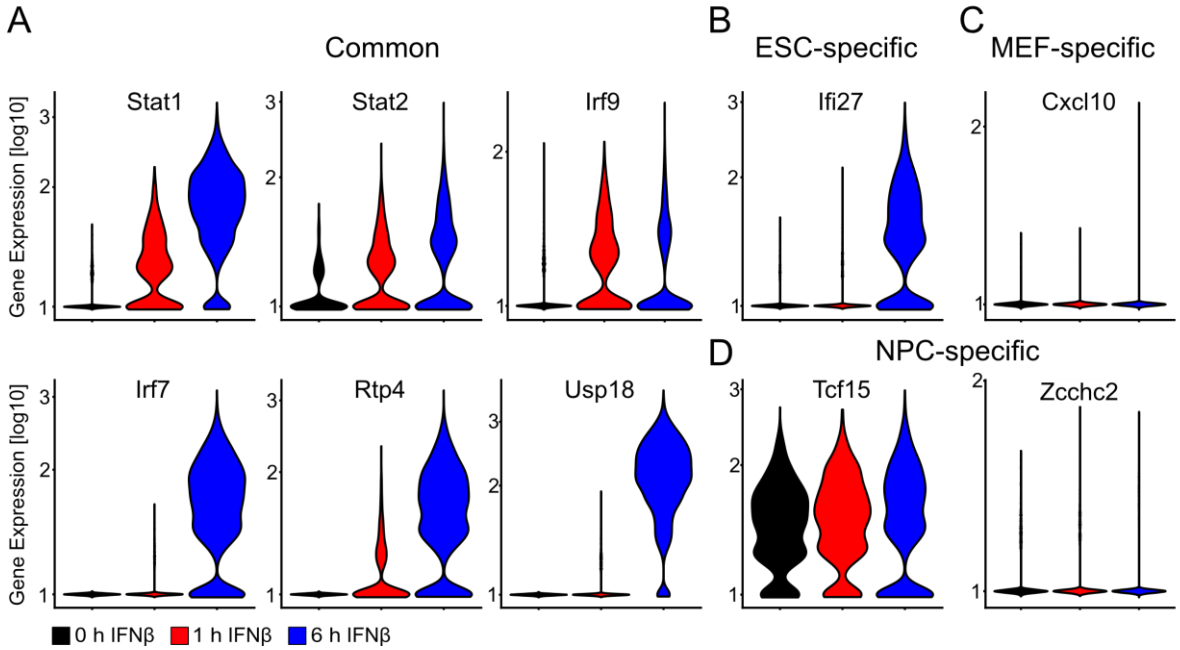

Figure 17: Expression dynamics of ISGs in ESCs by scRNA-seq

(A) Violin plot of log-normalized gene counts for ISGs over an IFNβ treatment time course. Unstimulated cells (black), 1 h (red) and 6 h (bleu) of treatments were plotted for various genes. *Stat1*, *Stat2*, *Irf9*, *Irf7*, *Rtp4* and *Usp18* were identified as ISGs in multiple cell types and responded strongly after 6 h. The ESC-specific genes *Ifi27* (B) got activated, while MEF-specific *Cxcl10* (C) and the NPC-specific genes *Tcf15* and *Zcchc2* (D) did not respond in ESCs.

In contrast, neither the MEF-specific gene *Cxcl10* (**Figure 17C**) nor the NPC-specific *Tcf15* and *Zcchc2* genes showed any upregulation upon IFNβ stimulation (**Figure 17D**). In summary, the IFNβ stimulation caused upregulation of genes at 1 h and 6 h, similar to bulk data. The initial responses were found to be heterogeneous but turned to be more homogeneous response at the later time point. The cell type specific ISGs pattern was confirmed in the single cell data, although many of those genes identified in bulk were not detected in single cell data.

3.1.6. Gene expression of IFN receptors and kinases was reduced in ESCs

Next, we hypothesized, that the abundancy of receptors and kinases involved in the canonical JAK-STAT signaling cascade impact the ISG induction levels. Therefore, the basal gene expression levels (0 h IFNβ) of genes involved in the JAK-STAT signaling cascade were analyzed in bulk RNA data (**Figure 18**). The receptors *Ifnar1/2*, *Ifngr1/2* and the two kinases *Jak1/2* were significantly differential expressed between cell types. While ESCs did show the lowest levels, MEFs had the highest mRNA levels for those genes. The third kinase *Tyk2*, was found to be differentially expressed between only MEFs and NPCs. Additionally, the kinase *Cdk8* was reported to be

essential for the phosphorylation of STAT1 at serine 727. Here, it was found to be significantly higher expressed in MEFs than in ESCs. Despite the differences on the levels of receptors and kinases, no differences in the gene expression levels of all three associated TFs, *Irf9*, *Stat1* and *Stat2*, were found. In conclusion, we found higher gene expression levels of IFN associated receptors and kinases in differentiated cells compared to stem cells. No difference in the basal expression levels of TFs were found, which resulted in the hypothesis that the overall level of activated TFs was the bottle neck in the cascade.

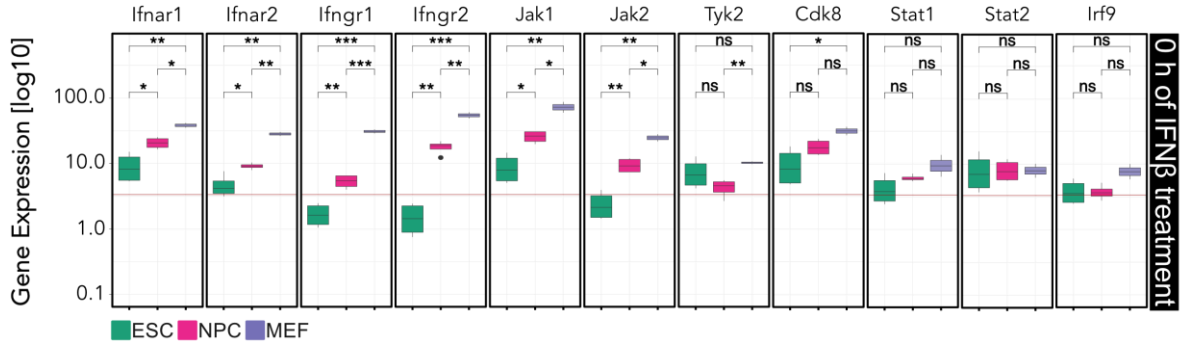

Figure 18: Basal gene expression levels of genes from the JAK-STAT signaling cascade

Boxplots of normalized RNA counts (TPMs) of unstimulated (0 h) samples for genes associated with the JAK-STAT signaling cascade. *Ifnar1/2* do form the classical receptor to recognized IFNα and IFNβ and are involved in type I interferon response. *Ifngr1/2* recognize interferon gamma and is essential for type II interferon signaling. The next level in the signaling cascade are the kinases *Jak1*, *Jak2* and *Tyk2*. Another kinase *Cdk8* is required for adding phosphorylation marks in the nucleus to *Stat1* followed by the set of TFs *Stat1*, *Stat2* and *Irf9*. Significance was analyzed using paired samples t-test.

3.1.7. Lower STAT protein levels and weaker $STAT1_{p727}$ mark were found in ESCs

The measured gene expression levels reflected the mRNA levels but not necessarily the protein levels. In addition, it is known that phosphorylation, posttranslational modifications, played a key role on multiple levels of the IFN signaling cascade. Especially, the phosphorylation at STAT1 Tyrosine 701 ($STAT1_{p701}$) and Serine 727 ($STAT1_{p727}$) are important for the activation of STAT complexes. The higher basal expression levels of JAK kinases might also impact the phosphorylation status of STATs upon stimulation. Therefore, ESCs and MEFs were stimulated with IFNβ for 1 h and 6 h to analyze protein levels of total STAT1 and STAT2. In addition, the two key phosphorylation marks, $STAT1_{p701}$ and $STAT1_{p727}$, were characterized. The insignificant differences between ESCs and MEFs on RNA levels before stimulation (0 h) (**Figure 18**) were not confirmed on protein levels for STAT1 (**Figure 19A**) and

STAT2 (**Figure 19B**). For both STATs the basal protein levels were higher in MEFs. Upon stimulation with IFNβ, a clear induction of both TFs was observed. The dimerization phosphorylation mark $STAT1_{p701}$ showed no signal before stimulation, a very strong and specific signal at 1 h (**Figure 19C**). Interestingly, in both cell types, the mark was nearly removed after 6 h of stimulation. The phosphorylation mark $STAT1_{p727}$ is known to be associated with transcription. This mark showed a weaker signal in ESCs especially at 1 h of treatment (**Figure 19D**).

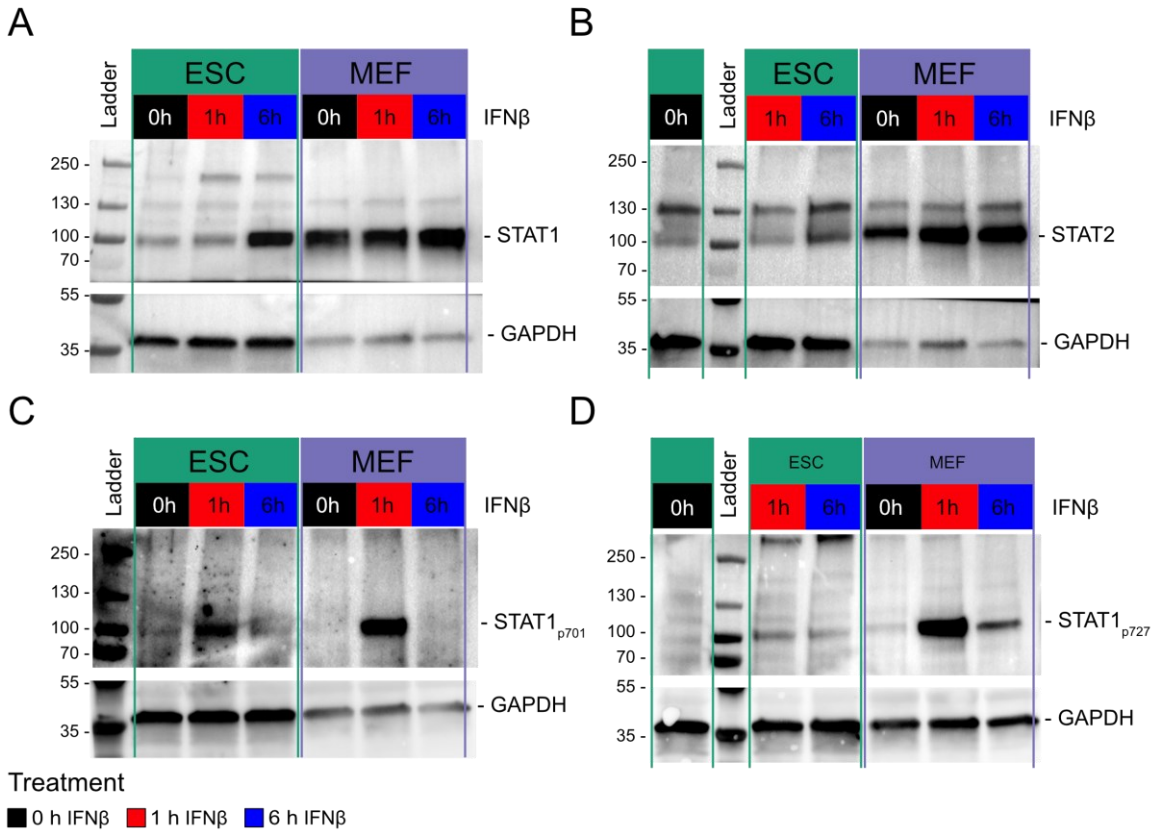

Figure 19: Protein levels of STATs were lower in ESCs before and after IFNβ treatment

Western blot analysis of ESCs and MEFs treated with IFNβ for 1 h and 6 h. GAPDH was used as housekeeping control gene. (A) Staining of STAT1 showed a higher basal protein level and stronger induction at 1 h and 6 h in MEFs. (B) Same results were found for STAT2. (C) $STAT1_{p701}$ was only detected at 1 h of treatment in both cell lines. (D) $STAT1_{p727}$ was strongly enriched at 1 h in MEFs and also depleted at 6 h. In ESCs the induction level was lower and more similar between 1 h and 6 h.

In MEFs this mark, similar to $STAT1_{p701}$, was strongly depleted but not completely removed after 6 h in MEFs, while it was lower in ESCs at both treatment time points. In summary, I observed the induction of both TFs STAT1 and STAT2 over the time course on protein level. The overall levels of the TFs were higher in MEFs compared to ESCs. The dimerization mark $STAT1_{p701}$ did peak in both cell types at 1 h and was

almost completely removed at the last time point. The enhancing mark at $STAT1_{p727}$ was much stronger in MEFs compared with ESCs. In conclusion, the presence of the enhancing phosphorylation mark at Serine 727 could cause the stronger induction of ISGs in MEFs compared to other cell types.

3.1.8. Specific sets of ISGs are induced in ESCs and MEFs

The previous results showed that ESCs had less components of the JAK-STAT signaling cascade. This might be an explanation for the attenuated immune response in ESCs for common ISGs. However, there was a set of cell type specific ISGs that comprised 33 (ESC), 17 (NPC) and 221 (MEF) genes (**Figure 11**). The different response is illustrated in **Figure 20** for selected genes. After 6 h of stimulation *Ccnd2*, *Ifi27* and *Nsg2* were significantly stronger expressed in ESCs. In NPCs, *Ccnd2* and *Nsg2* were constitutively expressed and while the lowly expressed *Ifi27* showed a small expression increase after 6 h. In MEFs, all three genes were not induced. Three MEF-specific genes *Ccl2*, *Gbp6* and *Ifi205* were strongly upregulated upon IFNβ stimulation (**Figure 20B**). They responded strongly at 1 h and 6 h independently from their basal gene expression levels in MEFs. *Gpb6* and *Ifi205* also responded in NPCs albeit at a lower level. *Gpb6* showed a similar pattern of induction in ESCs. *Lama3*, *Tcf15* and *Zcchc2* represented the group of 17 NPC-specific ISGs. In all three cell types those genes were lowly expressed (**Figure 20C**). The induction pattern in NPCs showed minor levels of upregulation. The induction of *Lama3* was significant but still below the classification threshold for an active gene, while *Tcf15* and *Zcchc2* became expressed. The expression levels of these three genes in ESCs and MEFs was found to be low and not significantly altered upon IFNβ treatment. An exception was the induction of *Zcchc2* in MEFs. NPC-specific ISGs displayed induction profiles that were similar to some ISGs in ESCs like *Lama3* while other like *Zcchc2* resembled the behavior in MEFs. Only very few genes like *Tcf15* showed induction patterns that were specific for NPCs. In summary, cell type specific genes were often lowly expressed in the other cell lines. Thereby, they would have had the potential to be activated and upregulated, however they failed to do so. Based on this data, we decided to further investigate the cell type specific responses of ESCs and MEFs. I especially focused on genes, not responding in MEFs, the cell type with the highest activation potential.

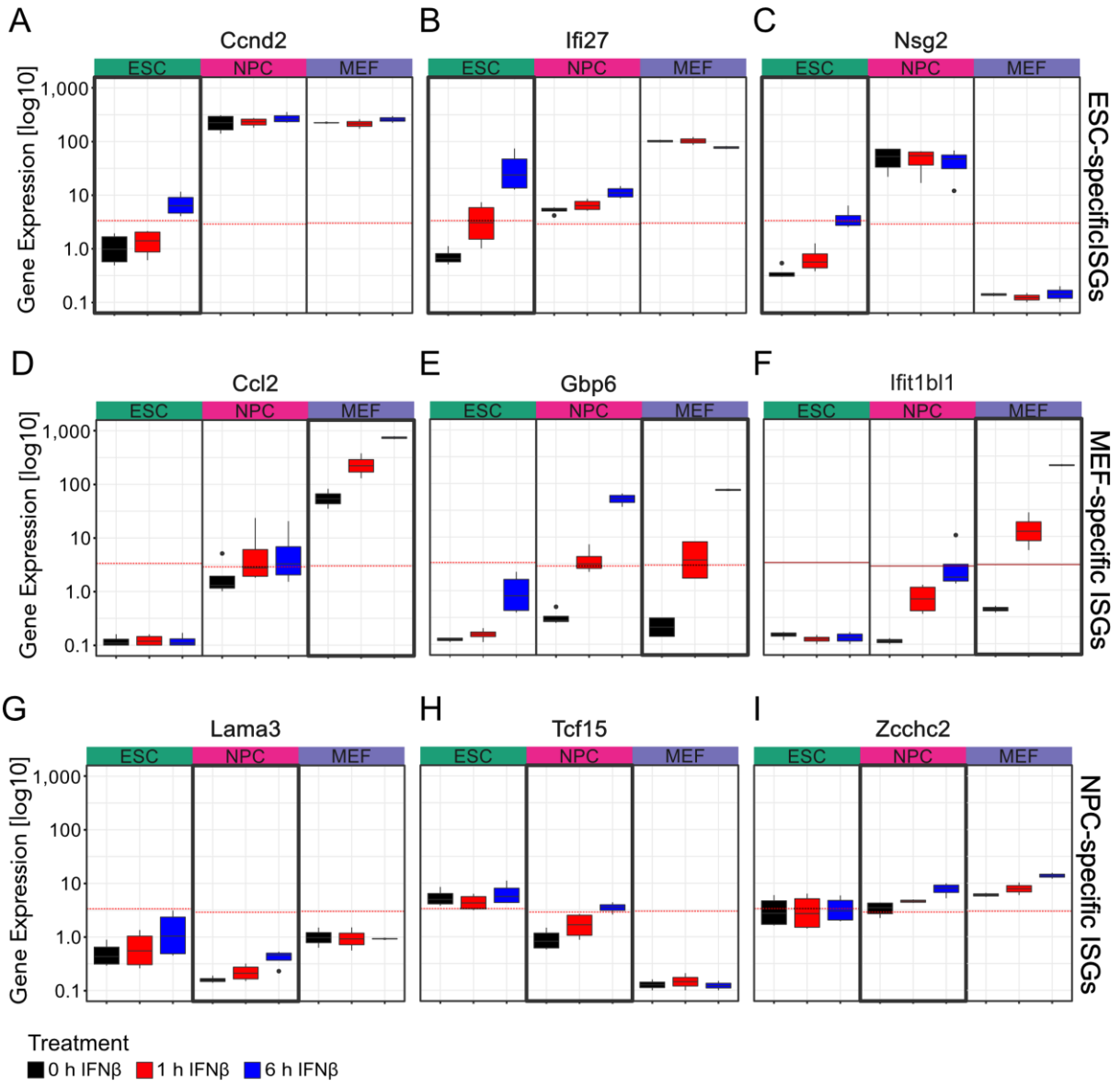

Figure 20: Induction patterns of cell type specific ISGs

Boxplots of normalized RNA read counts (TPM) values for various genes over IFNβ time course in ESCs, NPCs and MEFs showing the expression values in multiple replicates. Unstimulated condition (0 h) was colored in black, 1 h of IFNβ treatment in red and 6 h in blue. The dotted red, horizontal lines marked the previously defined expression threshold for active genes in the respective cell type. (A) The genes *Ccnd2*, *Ifi27* and Nsg2 were found to be induced only in ESCs upon IFNβ treatment. (B) The genes *Ccl2*, *Gbp6* and *Ifi205* were identified as ISGs in MEFs. (C) *Lama3*, *Tcf15* and *Zcchc2* were identified in NPCs only.

3.2. Function of STAT complexes on ISG induction patterns

3.2.1. Binding sites of STAT1 and STAT2 were identified by ChIP-seq

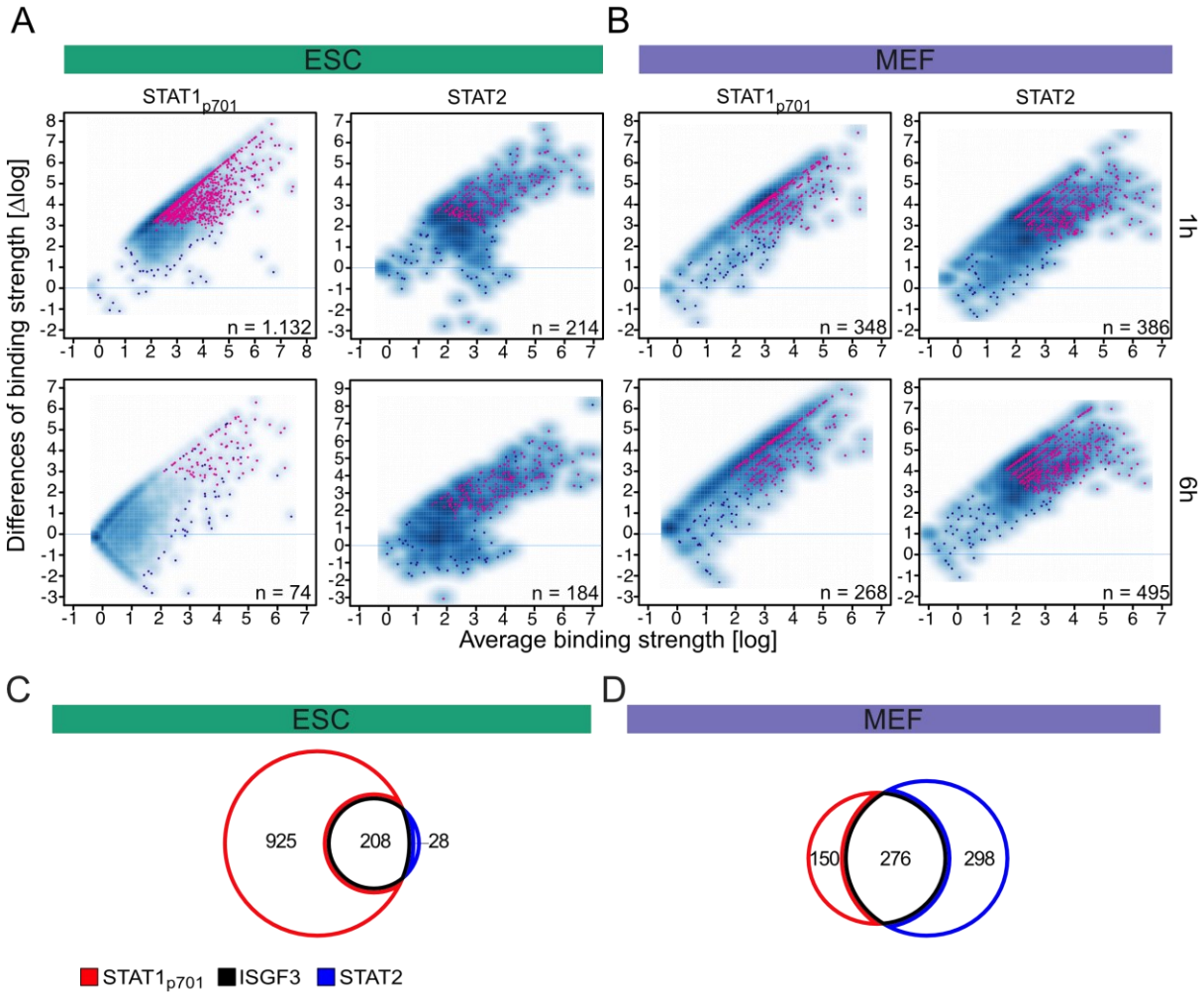

Figure 21: Identification of different STAT1/2 bound sites by ChIP-seq

(A) MA plot of enrichment signals for $STAT1_{p701}$ (left) or STAT2 (right) in ESCs on previously called peaks from MACS2. Each blue dot represented a peak and the y-axis reflected the enriched signal in comparison to unstimulated control sample. The purple marked dots represented the significantly differential enriched sites. Top row showed the signal detected at 1 h of IFNβ and the bottom one the 6 h comparison. The numbers on the right, bottom corner show the number of differentially enriched peaks. (B) Same as in (A) in MEFs. (C) Overlap of defined peak lists of $STAT1_{p701}$ and STAT2 in ESCs. The used input lists were combined results from the 1 h and 6 h comparison from (A). (D) Same as (C) for MEFs.

To investigate the cell type specific differences of ISG induction, genome-wide binding sites of STAT1 and STAT2 were mapped in ESCs and MEFs by ChIP-seq. At least three biological replicates for $STAT1_{p701}$ and were acquired and used to identify STAT

binding sites at 1 h and 6 h in comparison to the unstimulated (0 h) control (**Figure 21**). Purple marked sites in the MA-plots highlighted significantly enriched sites. These selected sites were used in all further down-stream analyses (**Figure 21**). The analysis yielded 1,132 (1 h) and 74 (6 h) $STAT1_{p701}$ peaks and 214 (1h) and 184 (6 h) STAT2 peaks in ESCs (**Figure 21A**). In MEFs more peaks for STAT2 were found, namely 386 and 495 peaks at the 1 h and 6 h time points, respectively. The numbers for $STAT1_{p701}$ were in a similar range as in ESCs at 348 (1 h) and 495 (6 h) peaks.

Next, the binding sites identified at the 1 h and 6 h time points were combined and intersected between $STAT1_{p701}$ and STAT2. It revealed 208 (ESC) and 276 (MEF) co-bound sites that are referred to as ISGF3 bound sites in the following (**Figure 21C, 14D**). The sites that were only bound by $STAT1_{p701}$ amounted to 925 (ESC) and 150 (MEFs) while the corresponding fraction of STAT2 only sites comprised 28 (ESCs) and 298 (MEF) sites. In conclusion, hundreds of IFNβ induced STAT1 and STAT2 binding sites were identified in ESCs and MEFs. The combination of these two TF signals allowed to characterize ISGF3 binding sites with high confidence.

3.2.2. Specificities of STAT peaks were validated by motif enrichments

Next, I wanted to validate the quality of identified peak sets by investigating the enriched motifs found in the center of each peak set. Previous studies of the JAK-STAT signaling cascade identified the DNA sequence of binding motif for STAT and IRF TFs. This information was used to identify enriched motifs in our predefined peaks by HOMER (Heinz *et al.* 2010). In both ESCs and MEFs, the STAT-family motifs (STAT1, Stat3, Stat3+il21, STAT4, STAT5) were enriched at STAT1 peaks (**Figure 22A/B**), but not in the ISGF3 peak sets. These sites were associated with IRF-family motifs (IRF1, IRF2, IRF3, IRF8, ISRE) (**Figure 22C/D**). The same motifs were found in the center of STAT2 peaks (**Figure 22E/F**). In MEFs, the ISGF3 and STAT2 peaks were more strongly enriched for PU.1 motifs (PU.1:IRF8, PU.1-Irf) than their ESC counterparts.

For the quantification of absolute numbers of motif abundancy, peaks with either a STAT family motif (red box), IRF family motif (blue box) or both (red/blue box) were counted. Peaks without any of those motifs were labeled black. For the STAT1 peaks, at total of 65.8 % (54.3 % + 11.5 %) harbored an expected STAT family motif in ESCs (**Figure 22A**). In MEFs, 85.4 % (72.0 % + 13.4 %) of all STAT1 peaks contained the

expected motif (**Figure 22B**). In ESCs, IRF motifs were found 82.7 % of ISGF3 peaks (42.8 % + 39.9 %), in MEFs the rate was even higher with 90.2 % (42.4 % + 47.8 %) (**Figure 22C/D**). Similar numbers were found for STAT2 peaks, where IRF motifs were present in 85.8 % (57.1 % + 28.6 %) in ESCs and 87.6 % (62.4 % + 25.2 %) in MEFs (**Figure 22E/F**).

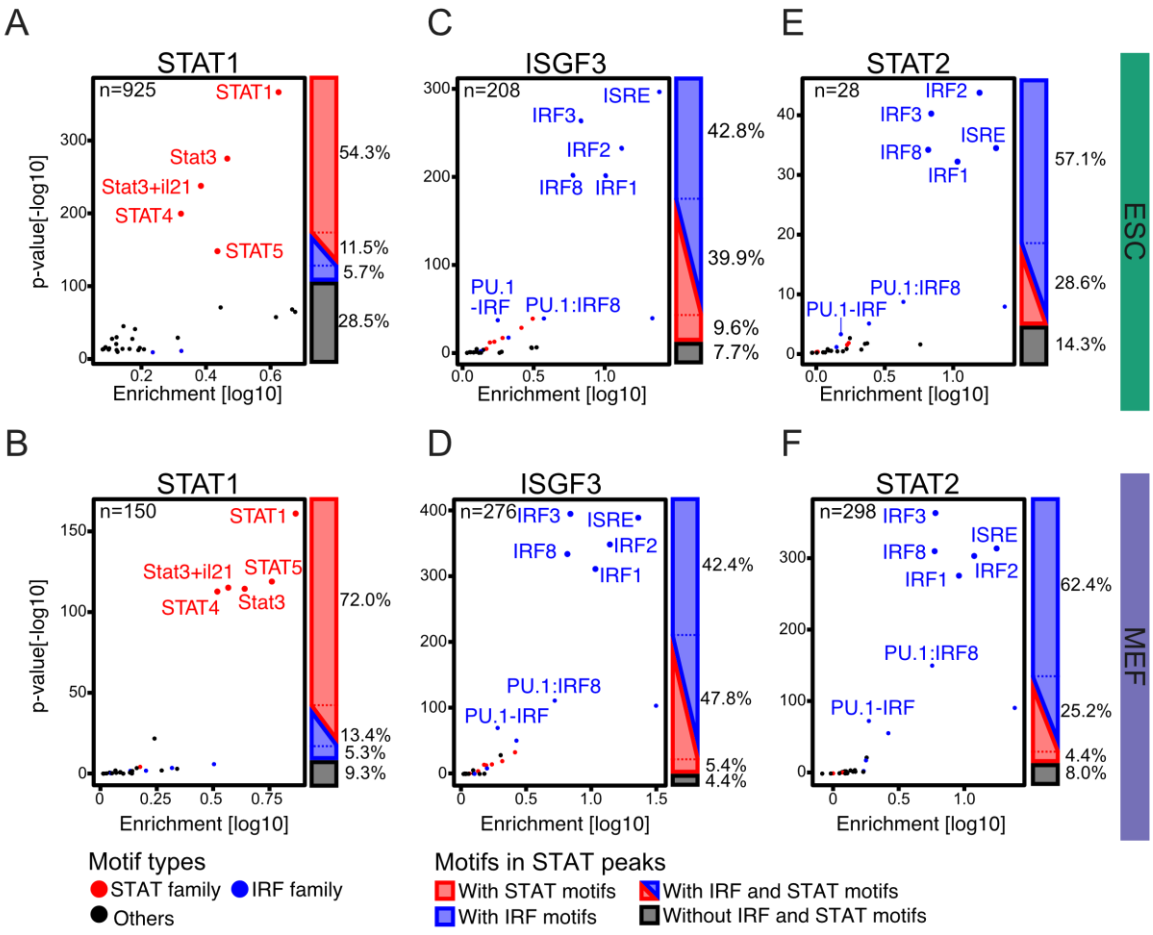

Figure 22: Motif enrichments and quantifications for STAT complexes

Scatterplot visualizing the detected binding motif enrichment over background in STAT peaks and the magnitude of the associated p-value by HOMER. Names of top five (A/B) or top five plus PU.1 (C-F) enriched motifs were shown in the plot. STAT family motifs were colored red, IRF in blue and other black. The bar plot represented the proportion of peaks with a STAT motif in red, an IRF motif in blue, both of them in red\blue or none of those motifs in black. (A) In ESCs, 925 STAT1 peaks were enriched for STAT1, Stat3, Stat3+il20, STAT4 and STAT5. (B) In MEFs, the STAT1 peak set contained 150 sites, which were enriched for the same motifs. (C) The 208 ISGF3 sites in ESCs were enriched for ISRE, IRF3, IRF2, IRF8 and IRF1. (D) In MEFs, 276 ISGF3 sites were identified in which same motifs were enriched. For STAT2 peaks in ESCs (28 sites; E) and MEFs (298 sites; F) the exact same motifs were found.

In addition, two IRF motifs associated with the pioneer TF motif of PU.1 were labeled. These motifs were found in ISGF3 and STAT2 peak sets only and significantly stronger

enriched in MEFs compared to ESCs. In summary, the predefined peak sets harbored the expected motif in at least 65.0 % of all peaks, demonstrating the quality and specificity of the identified binding sites. The identified motif families justified the separation of STAT1 only peaks from STAT1/STAT2 double positive and were in line with literature. The binding of STAT1 sites would represent the binding of a homodimer recognizing the STAT1 motif. In addition, sites bound by STAT1 and STAT2 represented ISGF3 binding sites and were bound to IRF motifs. However, the enrichment for all ISGF3 and STAT2 complexes identified five IRF motifs to a similar level in our data. For STATs, the STAT1 motif was by far the strongest enriched hit but additionally four STAT motifs were found to be enriched as well. This raised the question, how these motifs differ in their DNA sequence composition.

Therefore, the annotated sequences of all highlighted motifs were analyzed. The expected motif for STAT1 homodimers is named STAT1 in the reference annotation, but is also known as GAS motif. (**Figure 23A, top**). The palindromic central motif (TTCCNGGAA) is flanked by three nucleotides at the beginning and two at the end, which resulted in a total length of 14 nucleotides. The similar height of these five nucleotides at these positions indicates a high level of variance. The central motif was found to be conserved for STAT3, STAT3 + il20 and STAT4 motifs (**Figure 23A**). STAT5 was the only motif with an alteration in the core sequence. In summary, based on the described motifs similarities, the presence of one of those motifs in the detected ChIP binding site was considered to be a validation for a specific TF binding event.

Similar observations were made for the ISRE motifs, the expected binding motif of ISGF3 (**Figure 23B, top**). This motif represents the binding motif for the third member of the ISGF3 complex, IRF9. The central sequence (TTTCNNTTTC) is supported with two more flexible nucleotides at the beginning of the ISRE motif, resulting a total length of twelve nucleotides. The very same sequences were found for the motifs of other members of the IRF family (Irf1, Irf2, Irf3, Irf8) (**Figure 23B**). In contrast to STAT motifs, the IRF family showed a conserved motif length but slight variations within the motifs were found. In summary, also IRF motifs showed a high similarity within and therefore the downstream analyses referred always to the motif families instead of single motif.

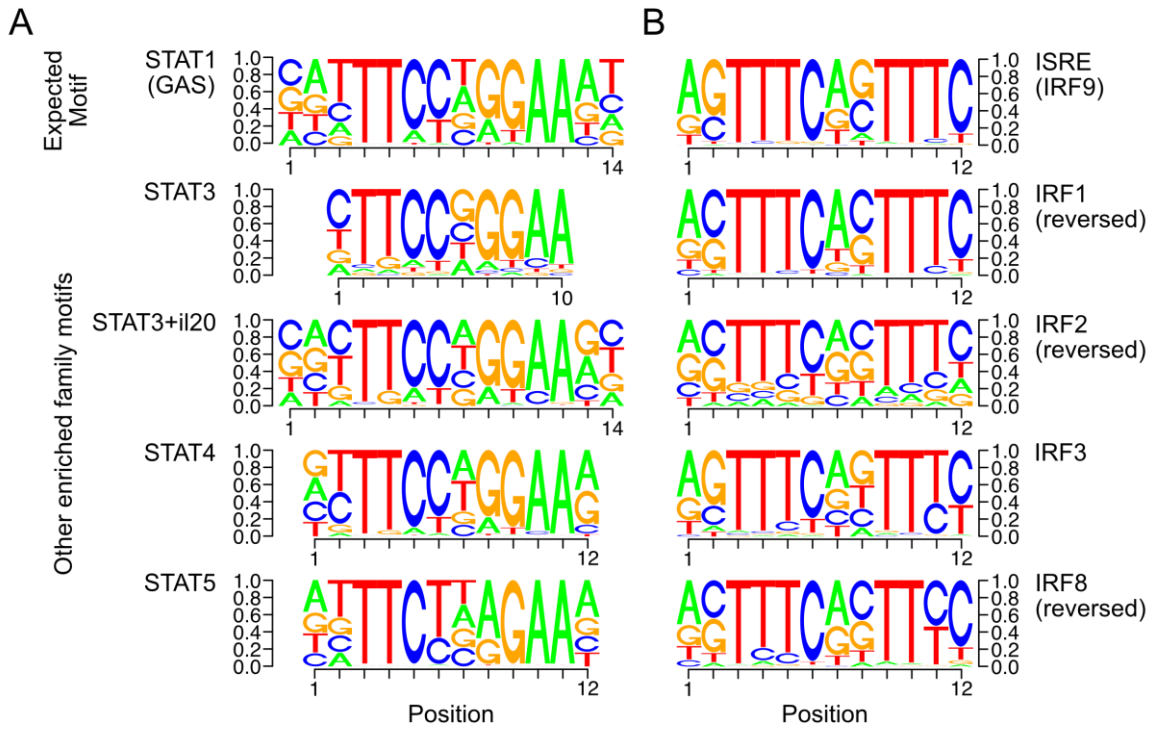

Figure 23: Sequences of top enriched STAT and IRF motifs identified in STAT peaks

Position weight matrices (PWMs) of various motifs found to be enriched in STAT ChIP-seq peaks. Based on literature, the STAT1 (GAS) motif was bound by STAT1 and the ISRE by ISGF3. For each position of the annotated sequence the probability of each nucleotide was depictured by the size of the single letter code of the nucleotide. (A) Sequences of the top five STAT motifs, STAT1, STAT3, STAT3+il20, STAT4 and STAT5 were shown. (B) The IRF motifs, ISRE, IRF1, IRF2, IRF3 and IRF8 were depictured.

3.2.3. Majority of STAT peaks were at non-promoter sites

Motivated by the validation of different STAT peak sets by motif enrichments, I tested how the distribution of STAT peak sets was over genomic loci (**Figure 24**). Firstly, a list of ENSEMBL annotated transcription start sites (TSS) was used to identify TF binding events directly at the transcription start site of genes. The remaining sites were overlapped with a list of exons followed by a list of introns from the same reference. Independent of the cell type, a minority (15.2 % to 38.0 %) of peaks with STAT1 and STAT2 binding sites were found to be directly at TSS. In contrast, between 41.3 % to 48.9 % of ISGF3 complexes were bound at promoters (**Figure 24, middle**). Binding in exons occurred rarely over all analyzed binding site peaks. Intronic and intergenic binding events, often functioning as enhancers, was highly abundant with more than 50 % in all groups. In summary, the binding patterns over genomic features was highly similar between cell types for ISGF3 and STAT2 peaks. In MEFs, the observed pattern for STAT1 followed the distributions of other complexes. In ESCs, the relative binding to intergenic elements was strongly enriched compared to all other complexes. The majority of binding events happened at regulatory elements,

pointed out a potential role of these sites in the STAT-depended innate immunity. The binding of an activating TF to a promoter normally should result in the induction of gene expression of the genes. Interestingly, in ESCs we identified 233 STAT bound promoters but only 191 ISGs. In MEFs, similar numbers of 242 bound promoters and 463 ISGs were found. I concluded from these numbers, that the binding of a STAT complex to a promoter not necessarily caused a significant induction of gene expression. We wanted to further characterize the effects of STAT complex binding to promoters and consequently took a closer look at these bound promoters.

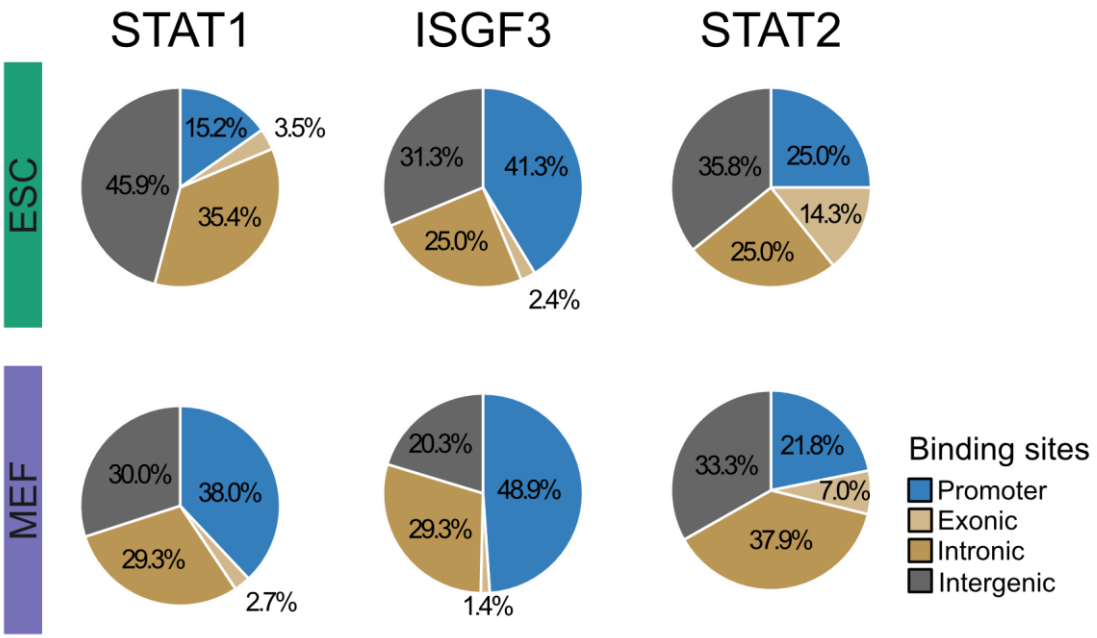

Figure 24: Binding patterns of STAT complexes over specific genomic loci

Pie charts of the overlaps between STAT complex binding and essential genomic features. These features contained TSS annotation, exons, introns and intergenic positions. The top row showed the distributions in ESCs and the lower one in MEFs.

3.2.4. ISG promoters mainly bound by ISGF3 complexes

The association of TF binding with gene activation was done by linking STAT bound TSS to gene expression levels. In total numbers, most promoters (n=165) were bound by STAT1 complexes in ESCs but only a small proportion, 10 out of 165, were responding to the IFNβ stimulation (**Figure 25A**). In ESCs, 69 out of 96 ISGF3 bound promoters showed a transcriptionally respond (**Figure 25B**). Also, only 5 out of 10 STAT2-bound TSS were transcriptionally induced (**Figure 25C**). The overall picture was confirmed with the binding data of STATs in MEFs (**Figure 25D**). STAT1 and

STAT2 peaks were found at TSS with similar numbers, 77 and 75. STAT2 binding still caused upregulation of 34 genes (**Figure 25F**), while STAT1 binding mainly failed to do so. In contrast, a total of 105 of 148 ISGF3-bound TSS showed significant induction of gene expression levels in respective genes (**Figure 25E**). In MEFs, ISGF3 also bound the most TSS compared with the other complexes. In conclusion, binding of STATs to promoters not necessarily caused gene activation. The majority of ISGs was bound by ISGF3, while many other gene promoters were bound by STAT1, but failed to response. On the contrary, STAT2 bound promoters often showed transcriptional responding. I wondered, if a high basal gene expression level of a promoter correlated with STAT binding.

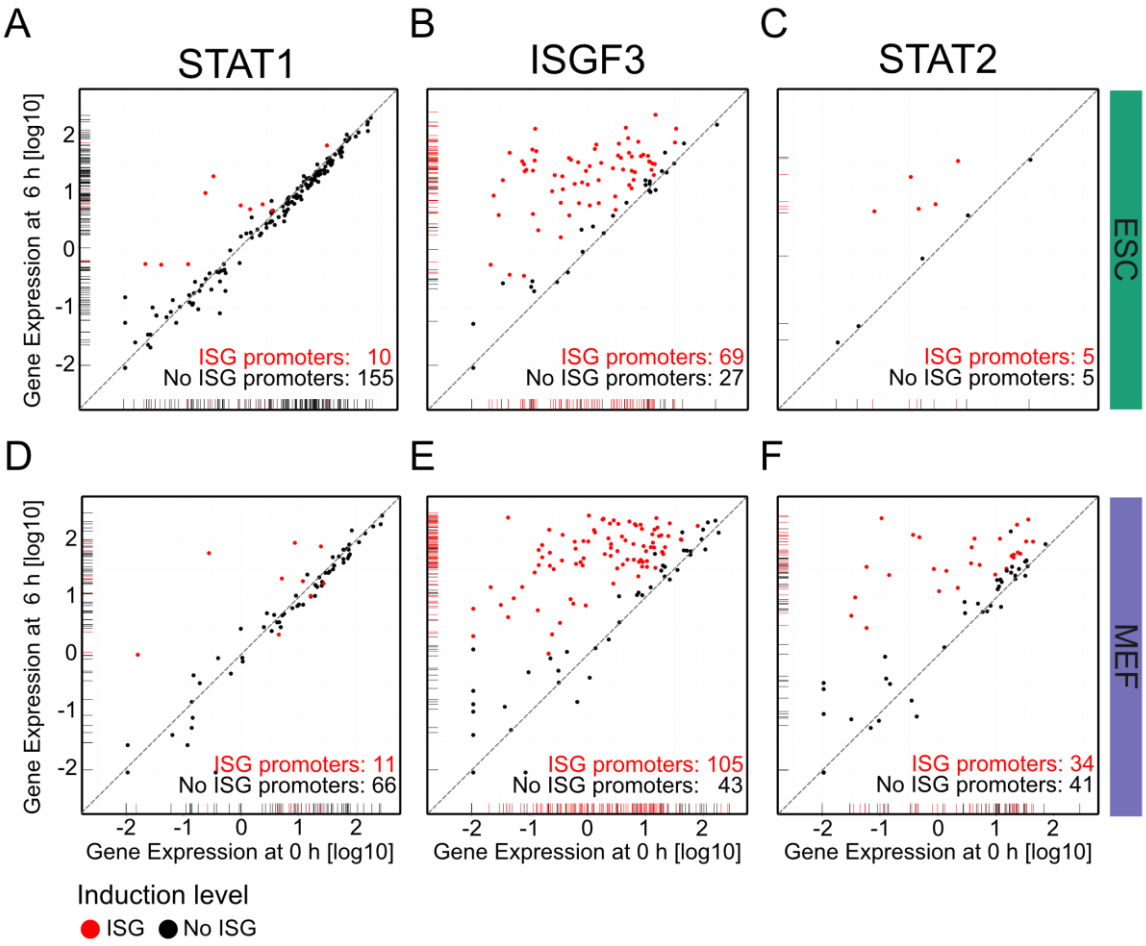

Figure 25: Gene induction changes upon STAT bindings to promoters

Scatter plots showed normalized RNA counts (TPMs) of unstimulated (0 h) verses 6 h of IFNβ treatment cells. Each dot represented a gene bound on at least one TSS by the corresponding STAT complex. Red dots were genes previously identified as ISGs, while black dots showed genes not significantly upregulated by the IFNβ treatment. Right bottom corner showed the number of STAT-bound promoters in two groups. Genes identified as ISGs were colored red, the rest in black. Induction patterns in ESCs for promoter bound by STAT1 (A), ISGF3 (B) and STAT2 (C). The lower row, showed gene promoters bound by STAT1 (D), ISGF3 (E) and STAT2 (F) in MEFs.

3.2.5. Highly expressed genes were not induced by STAT binding at their promoters

The fact, that many promoters were bound by STATs but did not induce their transcription, raised the question, if the basal gene expression levels of these genes had an influence of this behavior. To test this, the respective gene expression levels of all promoters bound by a STAT complex were plotted as density plot separately for the three time points.

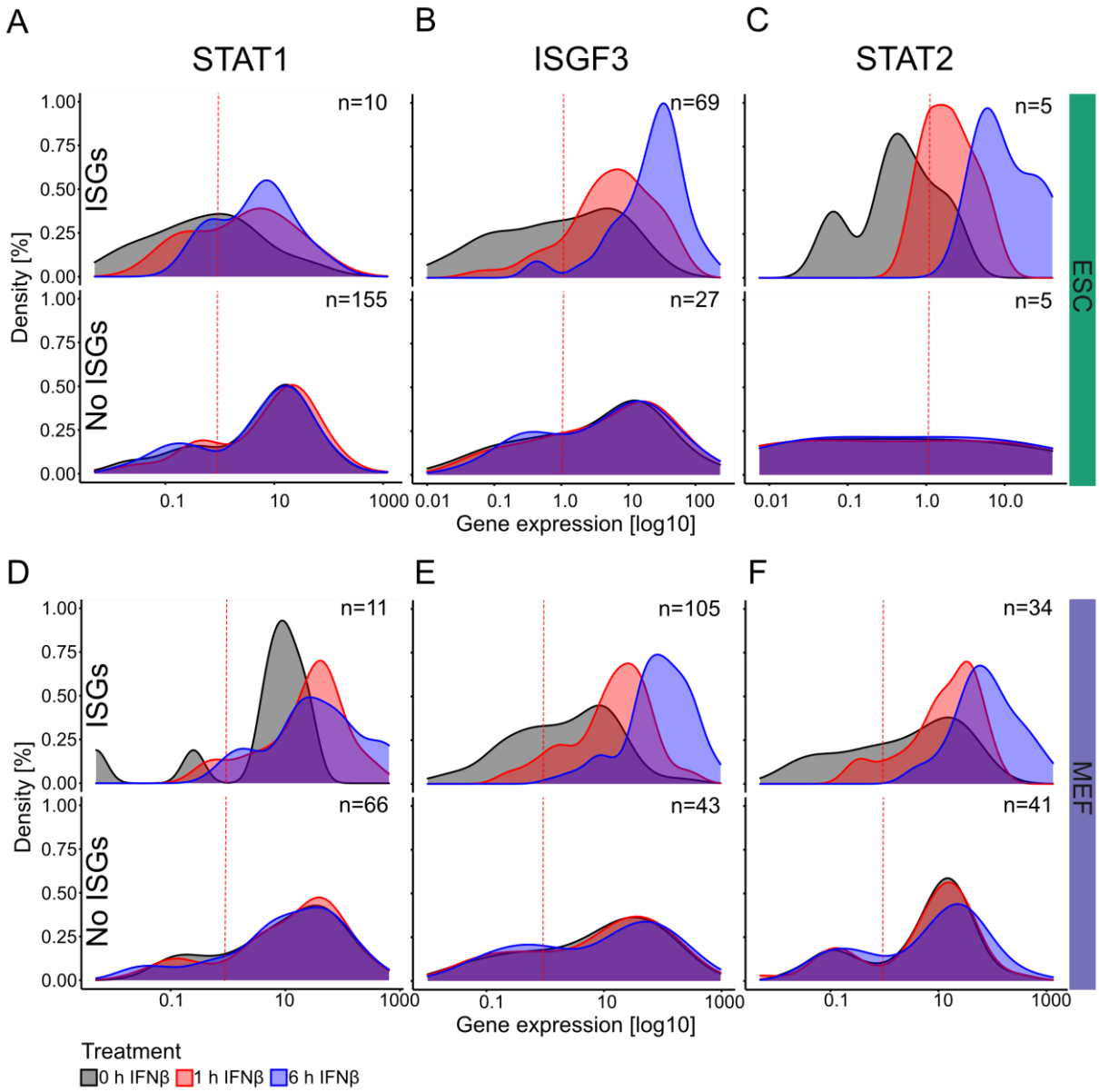

Figure 26: Dynamics of gene expression patterns upon STAT binding at promoters

Density plots of normalized RNA expression values (TPM) at 0 h, 1 h and 6 h of IFNβ treatment for genes bound by a STAT complex. Significant transcriptionally responding genes were called "ISGs" (upper) and not induced ones were called "No ISGs" (lower). The dynamics of gene expression in ESCs for STAT1 (A), ISGF3 (B) and STAT2 (C) bound genes were plotted. The same plots in MEFs for STAT1 (D), ISGF3 (E) and STAT2 (F).

Strong induction of gene expression for STAT bound ISGs was observed for all three complexes, while ISGF3 and STAT2 bound promoters showed the most robust induction of gene expression (**Figure 26A/B/C, upper**). In ESCs, the majority of non-ISG promoters bound by STAT1 had a high basal expression level (black) and no change in the overall gene expression patterns was observed during the treatment (red and blue) (**Figure 26A, lower**). The same was found for ISGF3 bound non-ISGs. For STAT2, the low numbers made it difficult to characterize them further and draw a conclusion (**Figure 26B/C, lower**). In MEFs, a highly similar result was observed. Most non-ISGs bound by any STAT complex, had a higher basal gene expression level and upon stimulation with IFNβ, no alterations of the basal patterns were observed (**Figure 26D/E/F, lower**). The shift to higher gene expression levels for ISGs, bound by any STAT complex, were consistent with findings in ESCs (**Figure 26D/E/F, upper**). In conclusion, high gene expression levels might be preventive to further induce gene expression via STAT complexes. ISGs had the tendency to have lower to no basal gene expression levels. However, some non-responding genes did also have lower basal expression levels. Therefore, another level of regulation was needed to be investigated to explain their failure to respond.

3.2.6. ISGF3 binding to ISG promoters caused faster and stronger activation

As only a fraction of ISGs were directly bound by STATs at their promoter, the question was raised if there were any differences in the intensity of gene expression levels between unbound and bound ISGs. This was addressed by comparing the average RNA levels of bound versus unbound ISGs over the time course of IFNβ stimulation. In ESCs, no differences were observed in the expression levels between bound versus not bound ISGs by either STAT1 or STAT2 (**Figure 27A/C**). In contrast, ISGs bound by ISGF3 showed a higher basal gene expression then their unbound controls (**Figure 27B**). Additionally, the binding of ISGF3 to an ISG caused a significantly stronger upregulation after 1 h. At the 6 h time point the induction was still higher than for unbound ISGs. The observed faster and stronger gene induction for ISGF3 bound sites was identically found in MEFs (**Figure 27E**). No differences before stimulation were observed in MEFs in contrast to ESCs. STAT2 bound sites followed the observed ISGF3 pattern, resulting in a faster and stronger response of bound ISGs, in MEFs (**Figure 27F**). The few STAT1 ISGs showed no difference to their unbound counterparts, except for a slightly stronger response after 1 h of IFNβ stimulation.

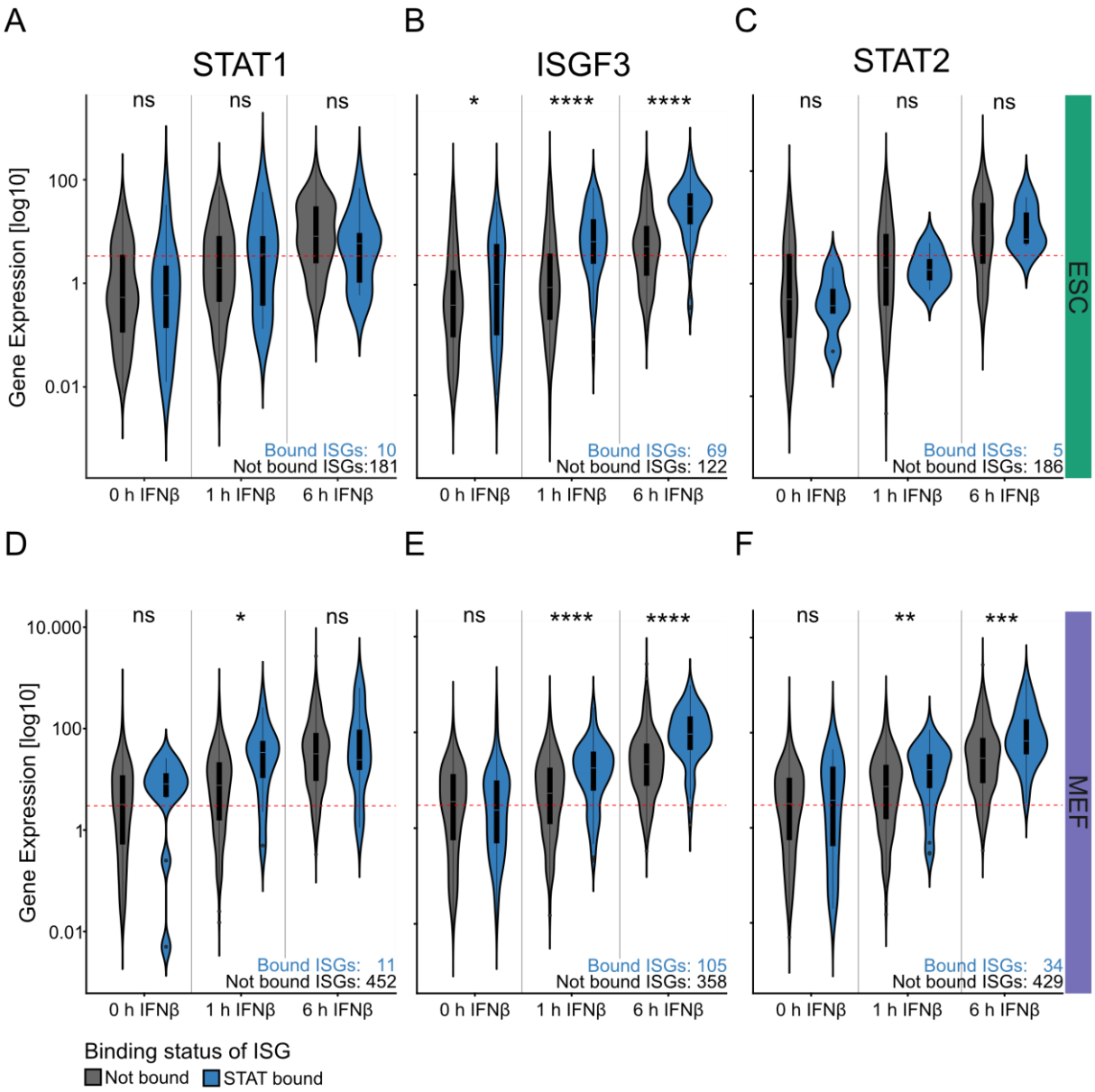

Figure 27: Effects of STAT binding to ISG promoters in comparision to not-bound ISGs

Violin plots of normalized RNA read counts (TPMs) for unbound ISGs (black) and STAT bound (blue). Each plot contained the comparison at the 0 h, 1 h and 6 h time points. The numbers of ISGs not-bound and bound by STAT in each plot were shown in the bottom right corner. Statistical testing between conditions of one time point were done with a Wilcoxon test. The horizontal, dotted red line showed the calculated expression threshold from **Figure 12**. (A) STAT1 complex bound ISGs compared to unbound ISGs in ESCs. (B) ISGF3 bound genes versus not bound ISGs and the same plot for STAT2 (C) in ESCs. STAT1 (D), ISGF3 (E) and STAT2 (F) comparisons of unbound and bound ISGs in MEFs.

In summary, in both cell types the binding of ISGF3 to ISGs caused a stronger and faster transcriptional response after 1 h of IFNβ treatment. After 6 h, the induction levels of those ISGs was significantly higher than for unbound ISGs. In MEFs, this pattern was also observed for STAT2 bound sites. The direct targeting of ISGF3 complexes to promoters caused gene induction, but these sites made up less than 50%

of all detected sites (**Figure 24**). Consequently, we hypothesized, that many STAT complexes bound non-promoter sites would also influenced the observed ISG induction patterns.

3.2.7. Nearest-gene approach was insufficient to characterize non-promoter ISGF3s

A well characterized level of gene regulation is via regulatory elements like enhancers, potentially up to megabases away from target promoters. Therefore, I characterized the contribution of these sites to the observed ISG patterns, since I found the majority of STAT bound sites were located at non-promoter regions. As ISGF3 was the most potent activator in all used cell types, the following analysis was focused on the ISGF3 complex. The established nearest-gene-approached links non-promoter binding of TFs with the nearest gene promoter. In ESCs, all, apart from one, ISGF3 binding sites were found not further then 500 kb away from at least one gene promoter (**Figure 28A**). The detected list of nearest genes was overlapped with the ISG list from ESCs. An overlap of 70 genes was found (**Figure 28B**), which was one additional ISGF3 link to a target ISG, than previously found (**Figure 25**). Therefore, we concluded that this approach was insufficient to further enlighten the activation potential of non-promoter binding of ISGF3 in ISG response.

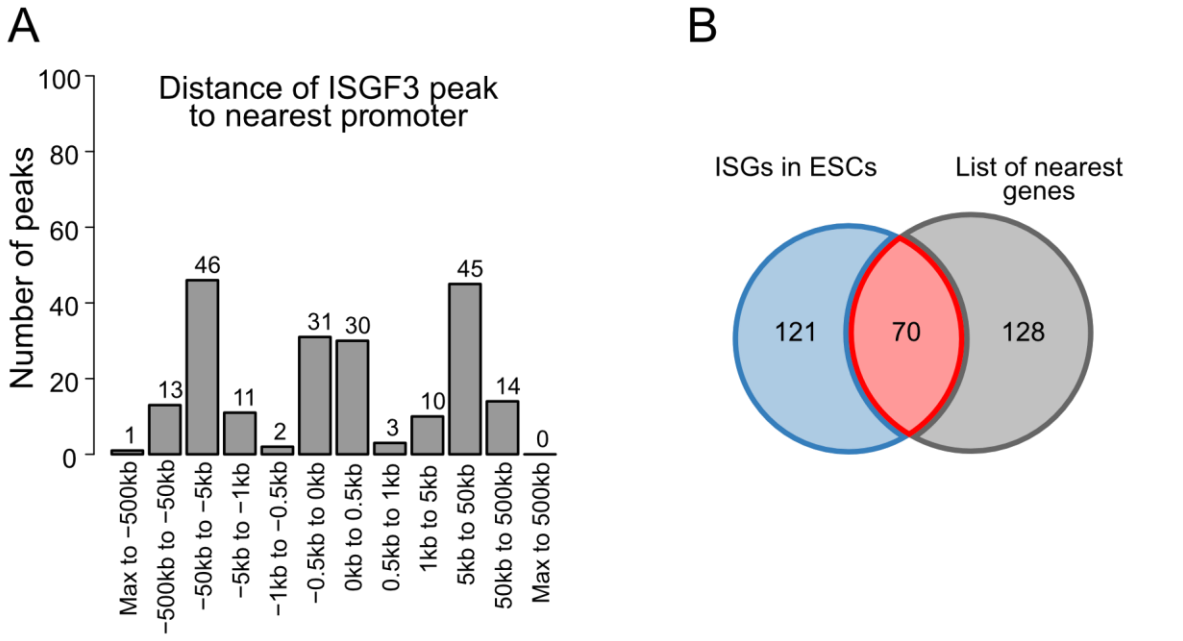

Figure 28: Nearest gene approach to link ISGF3 binding to regulated ISGs

(A) Histogram of distances for ISGF3 binding sites to the nearest gene promoter. Distances were calculated by GREAT (McLean *et al.* 2010). (B) Overlap of ISG list in ESCs from RNA-seq with list of nearest genes to an ISGF3 peak. The list of nearest genes was prepared by GREAT.

3.2.8. ISGF3 sites can be linked to ISGs via a co-regulation analysis by scATAC

For ISGF3 sites not located at promoters it is difficult to assign a target ISG. The measurements of chromatin accessibility on single cell level (scATAC-seq) were conducted. By mapping the genomic loci that were simultaneously open in a given single cell, co-accessibility maps in ESCs were generated. These maps were then used to compute pairwise correlation coefficients between two regions of interest to link ISGF3 binding sites with their targeted promoters. The overall bulk ATAC-seq signal of all ISGF3 sites in ESCs was very similar to the aggregated scATAC-seq data as a "pseudo bulk" signal (**Figure 29**). The majority of sites had a solid ATAC enrichment before stimulation (0 h), which increased upon IFNβ treatment (1 h, 6 h) (**Figure 29A**). Similar dynamics were observed in unstimulated (0 h) and 6 h treated ESCs by scATAC. Thus, the ISGF3 sites were already primed and accessible. Upon the stimulation of IFNβ the accessibility was further increased (**Figure 29B**).

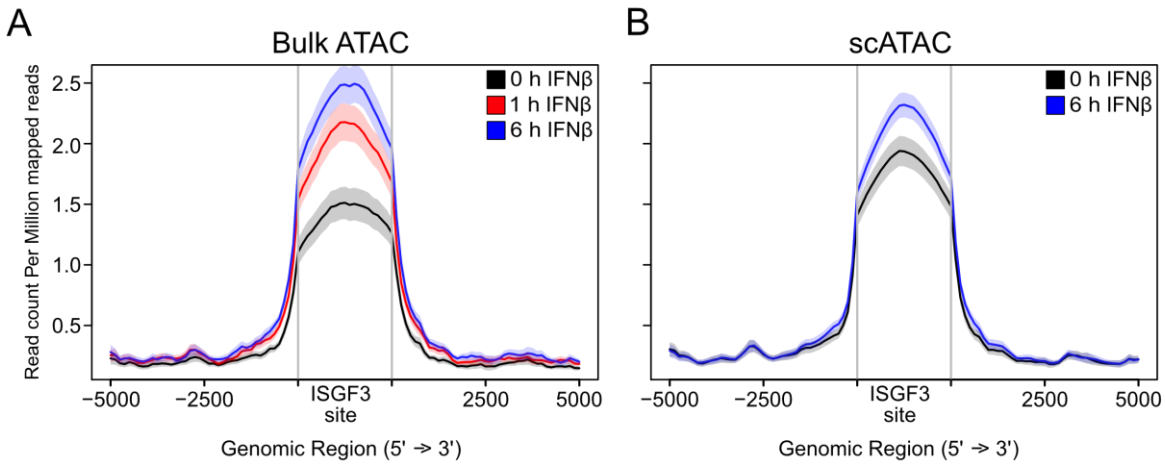

Figure 29: Chromatin accessibility at ISGF3 binding sites upon IFNβ stimulation

Genome coverage tracks of ATAC signals at ISGF3 binding sites in ESCs. Normalized enrichment counts were plotted for regions +/- 5 kb around ISGF3 binding sites. The unstimulated signal (0 h, black), 1 h (red) and 6 h of IFNβ treatment (blue). (A) bulk ATAC data. (B) of scATAC data.

Next, we computed co-accessibility maps from the scATAC-seq data. An exemplary region was shown for an intronic ISGF3 binding site in the proximity of the *Uba7* ISG (**Figure 30A**). The binding site had few contact interactions before stimulation (0 h) within the depicted genomic region. Upon IFNβ stimulation, a strong increased number of links to other genomic regions were detected. A direct link was established to the promoter of *Uba7*. The promoter of this ISG gained an ATAC peak upon IFNβ stimulation.

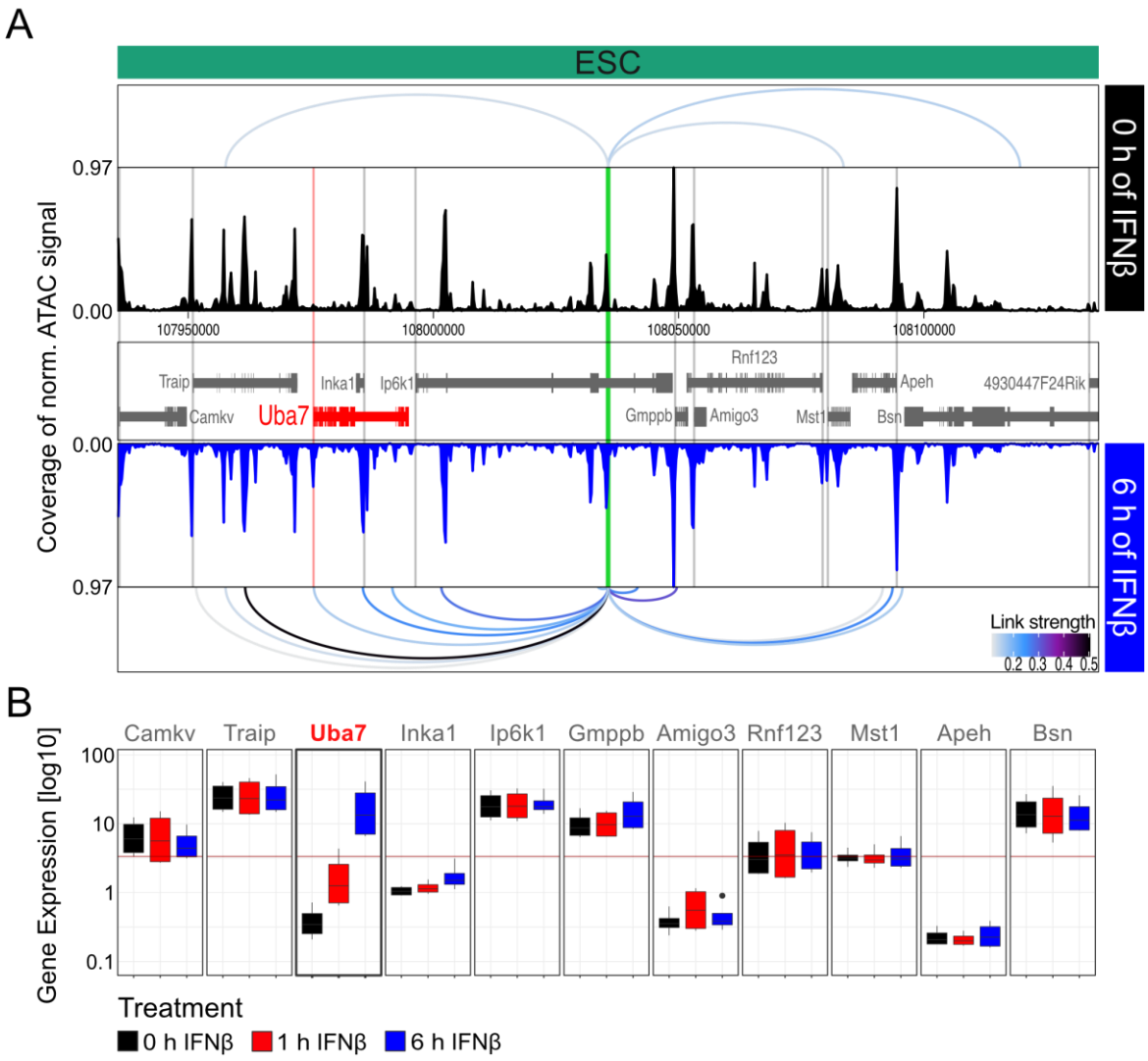

Figure 30: Linkage of an intronic ISGF3 binding site to the promoter Uba7

(A) Browser tracks of pseudo bluk from scATAC-seq from ESCs for an ISGF3 binding site (green line) in the proximity of the ISG *Uba7*. Region of +/- 120kb around the ISGF3 site were split into 2 kb tiles and the co-accessibility scores above the threshold of 0.13 were labeled by loops. The upper part showed the detected links between the ISGF3 site to other genomic regions (by ArchR). Followed by the actually pseudo bulk of scATAC-seq track and the gene annotations of this regions. The ISG *Uba7* was marked in red and all other genes in grey. The promoters of genes were highlighted, *Uba7* by a red line and remaining genes by grey lines. The lower part contained the browser tracks of 6 h IFNβ treatment and the detected co-accessibility links. The coloring of the co-accessibility links is based on the co-accessibility score, which is calculated from a correlation coefficient. (B) Box plots of normalized bulk RNA counts (TPM) for all genes in the genomic regions for at 0 h, 1 h and 6 h of IFNβ treatment. The red line showed the defined expression threshold in ESCs.

The bulk RNA-seq of all gene within the observed genomic window showed clearly that only *Uba7* responded to the ISGF3 binding event, while all other genes in the genomic regions were unaffected (**Figure 30B**). Links of the ISGF3 binding site to the promoters of other expressed genes like *Traip* and *Gmppb* were established as well. However, they did not respond by increasing their gene expression. We conclude

that newly established co-regulated links that occurred between non-promoters ISFG3 binding sites and ISG promoters can be identified from the scATAC-seq analysis. However, there were a general increase of co-occurring chromatin opening observed that also involved changes that appear to be not correlated with induction in the gene expression level.

3.2.9. ISGF3 bound sites gained interactions upon IFNβ stimulation

As next step, the observations for *Uba7* were validated on a genome-wide scale. The total number of detected co-accessibility links from ISGF3 sites increased from 5,436 at the 0 h time point to 7,852 after 6 h IFNβ treatment (**Figure 31A**).

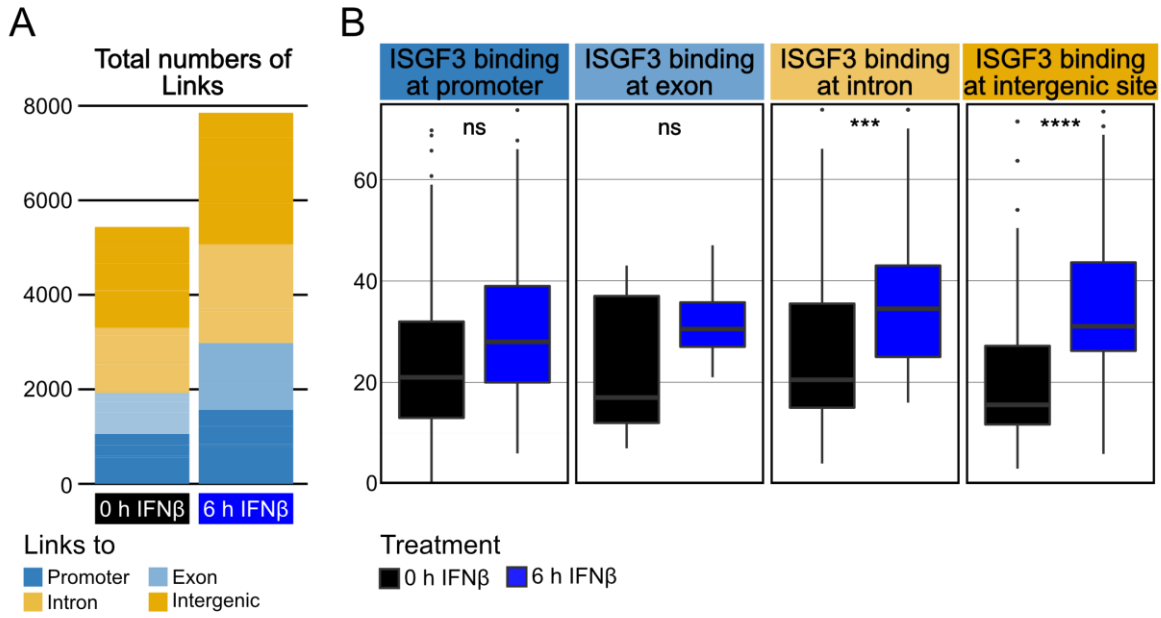

Figure 31: IFNβ induced changes of co-accessibility from ISGF3 binding sites

(A) Barplots represented the total number of links from ISGF3 sites at time point 0 h and 6 h of IFNβ treatment. The boxes were categorized based on the area they linked to. Links to promoter were colored in blue, to exons in light blue and to introns in yellow. The remaining links were not associated with one of those genomic features and therefore be called intergenic links (orange). (B) Separated analysis of the number of links from ISGF3 sites at promoters, exons, introns or intergenic sites. The unstimulated conditions (black) was compared to 6 h of IFNβ stimulation. Significance was measured by Wilcoxon test.

The gain in co-accessibility links for ISGF3 binding sites at promoters and exons was not significantly between the two time points. In contrast, the number of co-accessible links from ISGF3 binding sites at introns and intergenic regions was significantly higher at 6 h IFNβ treatment (**Figure 31B**). In conclusion, upon IFNβ stimulation ISGF3 binding sites at intronic and intergenic regions significantly gained more interactions. This result pointed out that regulatory elements like enhancers did play

an important role in the regulation of ISGs in ESCs. Therefore, the question remained, how many more ISGs were linked to an intergenic ISGF3.

3.2.10. Non-promoter ISGF3 bound sites linked to ISGs caused faster ISG response

As first step, we linked each ISGF3 binding site to ISGs. Thereby, three different groups were identified (**Figure 32A**). The first group, were ISGs with an ISGF3 binding events at their promoters ("Direct promoter bound", blue). Second group, were ISGs with a link from a distal ISGF3 peak, located at the promoter of another ISG, a promoter-promoter link ("Link from distal ISG promoter", green). The last group, were ISGs connected to a non-promoter ISGF3 binding sites ("Link from distal non-promoter", yellow). For the first group, 64 ISGF3 binding events at ISG promoters activated 69 ISGs (**Figure 32B**). Further, these 64 ISGF3 binding events linked to 37 distal ISG promoters ("Link from distal ISG promoter", green). Finally, 40 non-promoters bound ISGF3 sites had an ISGs in their proximity. Out of these binding sites, a subset established co-accessibility links to 28 distal ISGs ("Link from distal non-promoter", yellow). Interestingly, half of all identified ISGF3 binding sites (n=104) were more then 1Mb away from the next ISGs and therefore could not be directly linked to a target gene. The overlap of these three groups of links, resulted in 38 ISG promoters with one binding of ISGF3 at their promoters (**Figure 32C**). 21 of those ISGs had an additional link from another distal ISG promoter in the proximity. 6 ISGs directly bound with ISGF3 at their promoter had an additional link to an intergenic ISGF3 binding site. A group of 4 ISGs were found in all three groups. A remaining list of 12 ISGs were found to be only linked via another distal ISGF3 bound ISG promoter. While 18 were found to be linked only to a non-promoter binding event of ISGF3.

As next step, we investigated if the differential regulatory process of ISGF3 binding effected the induction patterns of the ISGs (**Figure 32D**), which was an expansion of **Figure 27B**. Therefore, we used only genes linked by one type of ISGF3. These groups were "Direct promoter bound" (n=38), "Link from distal ISG promoter" (n=12) and "Link from distal non-promoter" (n=28). In addition, all ISGs with more than one link to an ISGF3 binding sites were grouped as "Multiple links" (n=31; 21+6+4). Finally, all ISGs without any connection to an ISGF3 binding event were plotted as "No link to ISGF3" (n=92).

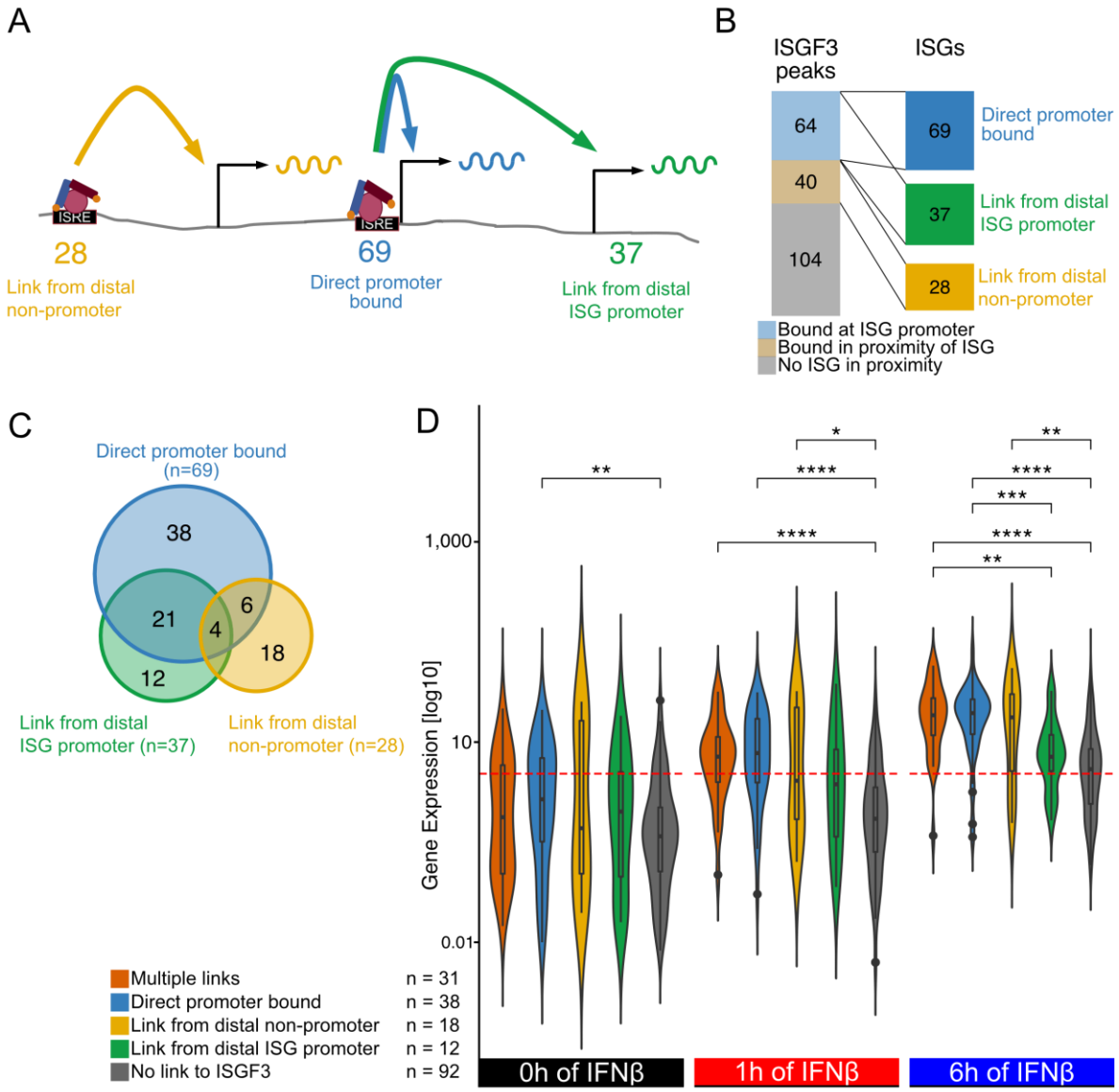

Figure 32: Effects of non-promoter ISGF3 binding links onto ISG induction patterns

(A) Model of three basic types of links from ISGF3 binding sites to ISG promoters. (B) Characterization of regulatory connections between ISGF3 binding sites and ISGs. The central bar represents all ISGF3 peaks in ESCs categorized in three groups. Peaks at promoters (n=64, light blue), peaks with an ISG promoter within +/- 500kb (n=40, light brown) and no ISG within the observed window of +/- 500kb (n=104, dark grey). From the viewpoint of ISGs, 69 ISGs were regulated by an ISGF3 binding event directly at the ISG promoter ("Directly promoter bound", blue). The promoter of 37 ISGs exhibited a co-accessible link to an ISGF3 peak in a different ISG promoter or distant enhancer region ("Link from distal ISG promoter", green). For 28 ISGs a co-accessible link between a distant ISGF3 peak and the gene body of the ISG was detected ("Link from distal non-promoter", yellow). (C) Overlap of the three previously described ISG groups. (D) Violin plots of normalized read counts (TPM) for five different ISG groups, depending on their ISGF3 binding state, over three IFNβ treatment time points. The three described groups were used, but all sites with more than one link from ISGF3 were grouped as "Multiple links" (orange). In addition, all remaining ISGs identified in bulk RNA-seq in ESCs and without any link to ISGF3 were plotted as "No link to ISGF3" group (grey). Significance were measured by Wilcoxon test.

Before IFNβ stimulation, “Direct promoter bound” ISGs had a higher basal gene expression levels compared to “No link to ISGF3” ISGs. After 1 h of IFNβ stimulation the “Direct promoter bound” and “Multiple links” ISGs were higher expressed than the “No link to ISGF3” ISGs. The third group of “Link from distal non-promoter” ISGs also resulted in a significantly higher gene expression then “No link to ISGF3” ISGs. After 6 h of IFNβ stimulation, the “Direct promoter bound” and “Multiple links” categories were higher expressed then the ISGs “Link from distal ISG promoter” and the “No link to ISGF3” groups. The “Link from distal ISG promoter” ISGs showed a significantly higher level of expression then “No link to ISGF3” ISGs. In summary, scATAC revealed new regulatory links between ISGF3 binding at regulatory elements and allowed to link more ISGs to ISGF3 binding. “Link from distal non-promoter” ISGs responded faster and stronger then unbound ISGs, highlighting the important role of non-promoter binding of ISGF3 in the innate immune response. The link from an “Link from distal non-promoter” ISGF3 binding had the same effects on ISGs inductions then the “Direct promoter bound” ISGF3 group. This showed how essential it is to characterize the binding sites of STAT1 and STAT2 on a genome-wide scale. As next step, I investigated the function of chromatin states on the binding behaviors of the STAT TFs.

3.3. Chromatin state dependent ISGF3 binding

The previous findings highlighted the impact of STAT complexes on the induction of gene expression. The aim was to understand the effects of STAT binding to a transcriptional read-out. However, all these processes required a stable binding of STATs at the specific genomic positions. I successfully identified these sites by TF ChIP-seq and were able to validate these sites by motif enrichments and chromatin accessibility levels. On average 170,074 (+/- 27,042) and 194,316 (+/- 41,719) accessibility peaks were called from ATAC-seq data in ESCs and MEFs, respectively. Furthermore, 815,638 IRF-motifs and 1,741,024 STAT-motifs were annotated in total in the mouse genome (IRF motifs number referred to ISRE, IRF1, IRF2, IRF3 and IRF8; STAT motifs referred to STAT1, STAT3, STAT3+il20, STAT4 and STAT5). In contrast, a total of 1,161 STAT binding sites in ESCs and 724 in MEFs were identified. We hypothesized that additional regulatory mechanisms were contributing to finetune the distinct ISGF3 binding events and their consequences in ISG response upon IFNβ stimulation. I postulated that different chromatin features had an important impact on STAT binding in a cell type specific manner and were essential to understand cell type specific transcriptional regulations.

3.3.1. A large fraction of ISGF3 binding sites was cell type specific

As first step, I focused on the cell type specific binding pattern of STAT complexes. The overlap of STAT peaks between cell types resulted in a moderate overlay (**Figure 33A**). Only 38 STAT1 and three for STAT2 peaks were shared between ESCs and MEFs. The vast majority of detected peaks were found to be cell type specific. ISGF3 exhibited with 92 peaks the highest fraction of shared peaks (**Figure 33**), which resulted in the 116 ESC-specific and 184 MEF-specific ISGF3 peaks. In addition, for ISGF3 the majority of common peaks were located at promoters (76.1 %), while 64.1 % to 86.2 % of cell type specific binding events happened in non-promoter regions. In summary, the cell type specific response to IFNβ was even more pronounced on the level of TF binding events than on gene expression levels, which could not be explained by motif sequences or chromatin accessibility data. Despite the similarities in the activation potential on gene expression levels, the strongly cell type specific binding profile of ISGF3 raised the question, how the binding of ISGF3 was regulated.

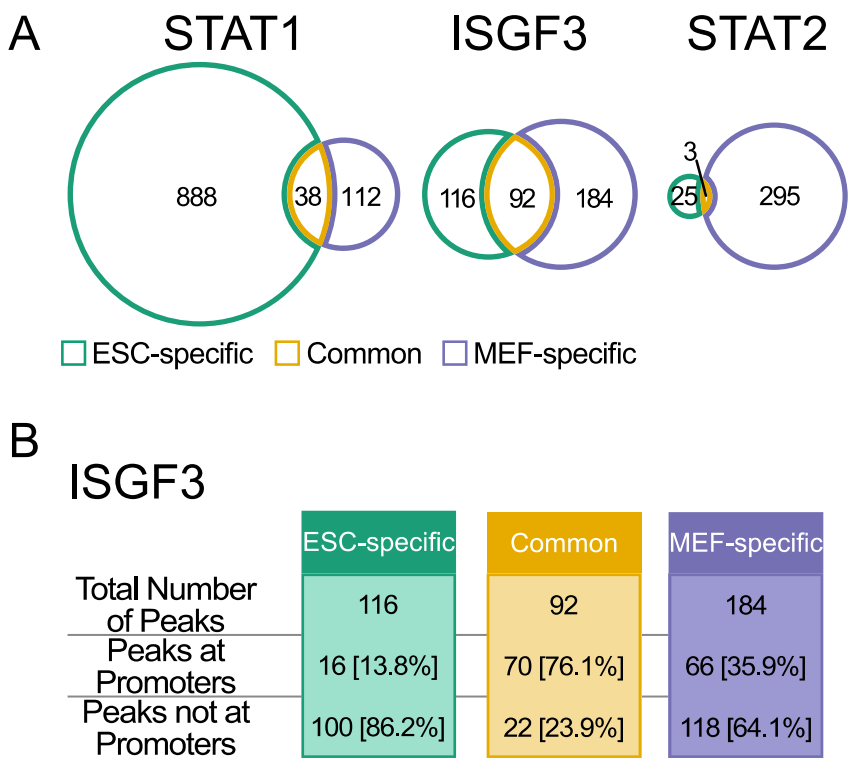

Figure 33: Overlap of STAT complexes in ESCs and MEFs after IFNβ induction

(A) List of STAT1 binding sites in ESCs were overlapped with STAT1 binding sites in MEFs. The same was done for ISGF3 and STAT2. (B) Overlap of ISGF3 bindings sites with transcription start site list.

3.3.2. Specific ISG promoters were enriched for different chromatin marks

The previously described ChIP-seq experiments for STAT2 and the phosphorylation mark at Tyrosine 701 in STAT1 ($STAT1_{p701}$) were expanded for multiple histone modifications. Six different chromatin marks (H3K4me1, H3K4me3, H3K9ac, H3K9me3, H3K27ac and H3K27me3) were analyzed by ChIP-seq experiments to understand the chromatin environment. Additionally, the bulk ATAC-seq was added to the analysis. A visualization of those data sets can be found in **Figure 34** with all data sets for three exemplary regions. The promoter (+/-500bp) of the ESC-specific ISG *Ifi27*, the gene *Usp18*, found in both cell types, and the MEF-specific ISG *Gbp6* were shown. Note that all three promoters did harbor the same binding motif, ISRE, but the binding of ISGF3 was not observed in all conditions. In ESCs, *Ifi27* was not expressed before stimulation, but became activated upon treatment (**Figure 34A, left**). The binding of $STAT1_{p701}$ and STAT2 was found in combination with the enrichment of activating marks like H3K4me3, H3K27ac and ATAC. Consequently, the mRNA levels were induced. In MEFs, this gene was expressed and the binding of the TFs did not impact the expression levels (**Figure 34B, left**). Usp18 was induced in

both cell types. The binding of STATs and the enrichments of all activating marks was observed in ESCs and MEFs (**Figure 34A/B, middle**). The MEF-specific ISG *Gbp6* was not bound by any TF in ESCs and no activating marks were observed (**Figure 34A, right**). The promoter region was enriched for the H3K4me1 mark. In MEFs this gene showed a strong increase of enrichment for $STAT1_{p701}$ and STAT2 in combination with a gain of ATAC (**Figure 34B, right**).

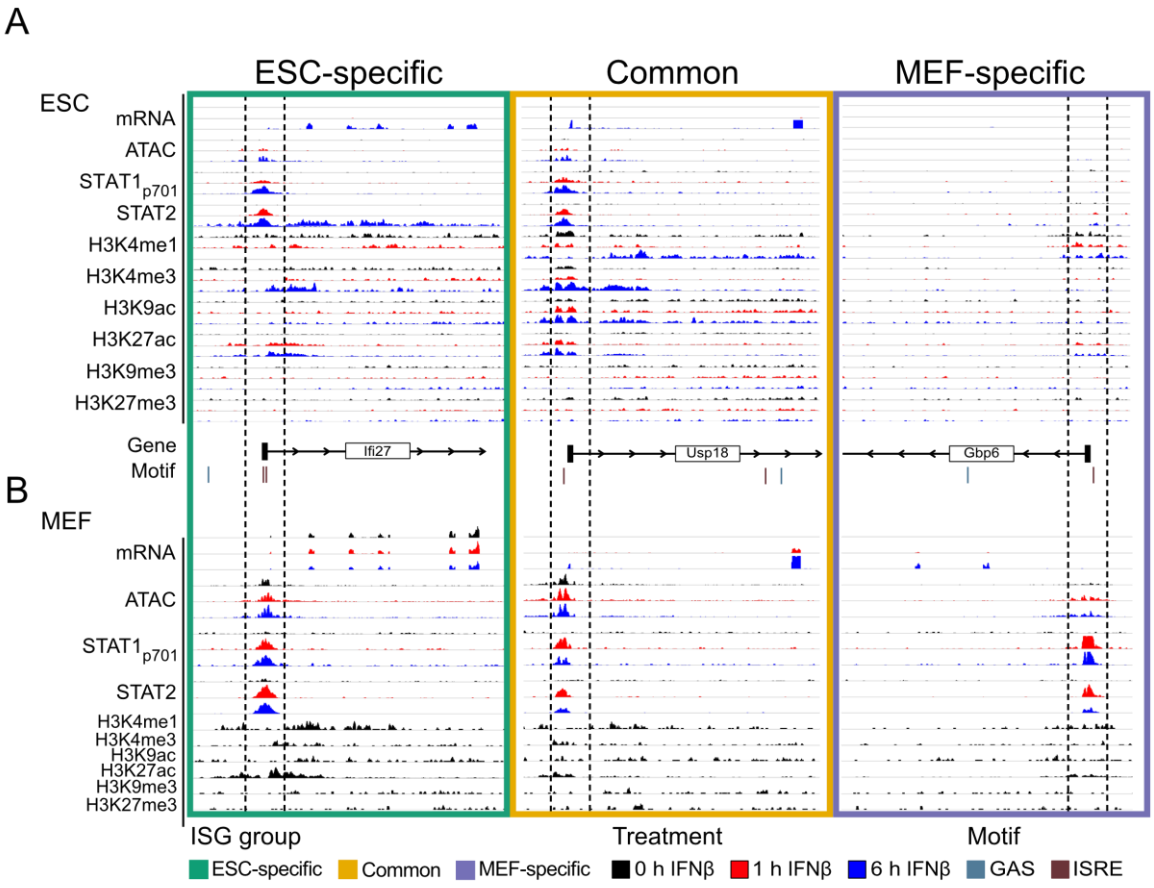

Figure 34: Browser tracks of the multiomics read-out for representative ISGs

Browser tracks by IGV tools for multiple genome-wide readouts over IFNβ stimulation. Unstimulated control was colored in black, 1 h IFNβ in red and 6 h in blue. The read outs were RNA-seq (mRNA), ATAC-seq (ATAC), TF ChIP-seq ($STAT1_{p701}$, STAT2), histone mark ChIP-seq (H3K4me1, H3K4me3, H3K9ac, H3k27ac, H3K9me3, H3K27me3), ENSEMBLE gene annotation and binding motifs GAS (blue) or ISRE (red). For each mark the same scale was used over the time course or set to a minimal scale of 10 normalized read counts. Binding sites for both STAT1 and STAT2 at ISG promoter were in the center of the black-dotted box, which covered a region of +/-500bp around the binding site. In the center of these regions an ISRE motif was found. The ISG group was defined by using the overlaps from Figure 11. *Ifi27* represented a gene upregulated only in ESCs (ESC-specific), while *Usp18* was found to be an ISG in all cell types (common) and the MEF-specific gene *Gbp6* was shown on the very right (MEF-specific). The data from treated ESCs (A) and from MEFs (B) was shown here. Note that the histone modifications in MEFs were done in unstimulated cells only.

In summary, ChIP-seq experiments were successfully performed and the promoter of previously identified ISGs were bound by TFs in a cell type specific manner. Binding events of STATs were observed at target motifs and correlated with enrichment of activating marks. We hypothesized that the chromatin state of STAT binding sites differed on a genome wide scale and contributed to the induction of ISGs by promoting stable binding of STATs.

3.3.3. ISGF3 sites have cell type specific chromatin states before induction

Alterations of chromatin state via histone modifications are known to be essential regulators of gene expression activation. Upregulation of ISGs was most strongly correlated with the binding of ISGF3. This was why we set the focus on understanding how chromatin alterations impact the binding of this STAT complex. First, all binding sites for ISGF3 from ESCs and MEFs were combined to a list of 392 potential ISGF3 sites (**Figure 35A**). In an +/- 1 kb window around each ISGF3 biding site, the chromatin environment of unstimulated cells (0 h) was analyzed via a read counting approach. The idea was to characterize a promotive chromatin state for ISGF3 binding. ChIP-seq of four active chromatin marks (H3K4me1, H3K4me3, H3K9ac, H3K27ac), two repressive marks (H3K9me3, H3K27me3) and chromatin accessibility data via ATAC-seq was used for an unsupervised k-means clustering. The clustering was performed on counted reads from 0 h, 1 h, 6 h IFNβ treated ESCs and unstimulated (0 h) MEFs of all 392 ISGF3 sites together. Thereby it was guaranteed that the cluster assignment was comparable between cell types and conditions. To identify the optimal number of clusters, the silhouette coefficient was calculated and a cluster number of five was identified as optimal (**Figure 35B**). Note, the highest coefficient was found with two clusters, which would result in a separation of active from repressive chromatin marks. This would from a statistical point of view right, but from a biological point of view, the aim was to identify various groups with different combinations of marks and that was why I decided to use the second highest coefficient.

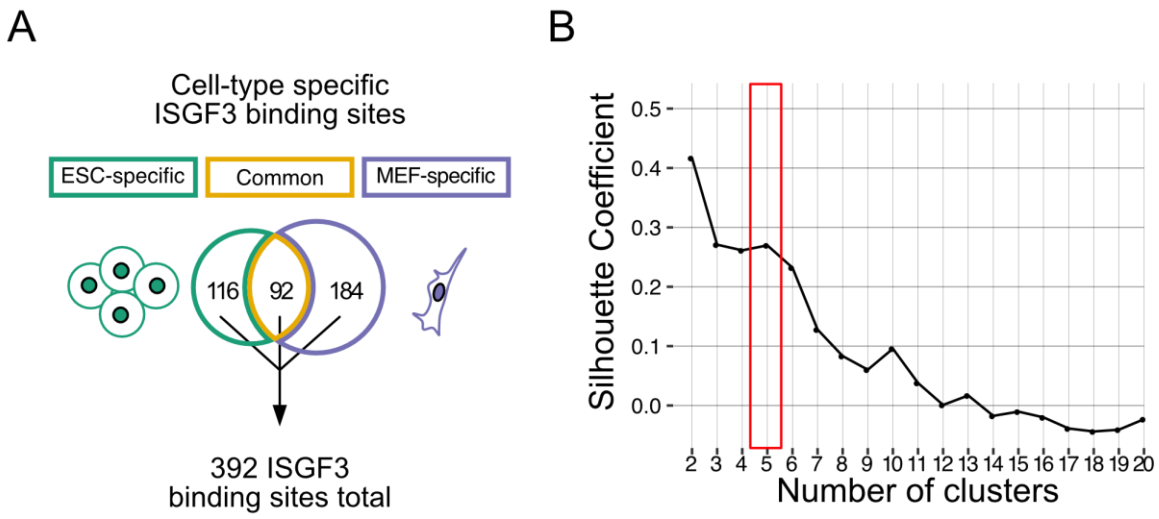

Figure 35: Preparation of ISGF3 peak sets and selection of cluster numbers

(A) Schematic to explain the combination of ISGF3 binding sites from ESCs and MEFs into a list of potential ISGF3 binding sites, containing 392 sites in total. (B) Calculation of Silhouette Coefficient to identify the optimal number of clusters for an unsupervised k-means clustering of the chromatin states for potential ISGF3 binding sites.

The five clusters of ISGF3 sites with distinct chromatin states were visualized as heatmap. The five clusters showed chromatin patterns, which reflected defined biological chromatin states. The classical promoter mark H3K4me3 was strongest enriched together with H3K27ac, H3K9ac and ATAC in the first cluster, which resembled an "Active Promoter" state (**Figure 36A**). The combination of H3K4me1, H3K27ac and ATAC were classical marks for "Active Enhancers" in the second cluster. Note that promoters of weakly expressed genes might harbor the same combination of marks with additional low levels of H3K4me3. The presence of repressive chromatin marks like H3K27me3 and H3K9me3 in combination with H3K4me1 and H3K27ac was classified/categorized as "Bivalent" cluster. The "Poised" cluster was enriched for H3K4me1 only. The classical "Repressed" cluster showed strongest signal for both repressive chromatin marks, H3K27me3 and H3K9me3.

Looking at the individual ISGF3 sites in all cluster, we found that the smallest cluster was the "Bivalent" followed by the "Active Promoter" state, while the remaining three were uniformly distributed (**Figure 36B**). The distribution of ISGF3 sites over the five chromatin states was depicted separately for untreated ESCs and MEFs (**Figure 36C**). In ESCs, more than 50 % of all sites were either in the "Repressed" (n=102) or "Poised" (n=118) chromatin state clusters. The "Bivalent" cluster was the smallest with only 22 ISGF3 sites. The numbers of "Active Enhancers" (n=87) and "Active

Promoter" in ESCs were comparable with the sites from MEFs in these clusters (n=54, n=90). In MEFs, more ISGF3 sites were found to possess the "Bivalent" chromatin state (n=76), while the two repressed chromatin state clusters were smaller compared to ESCs (n=91, n=81). In summary, the clustering of ISGF3 site chromatin states revealed patterns associated with five established chromatin states. The overall distributions of chromatin states over ISGF3 sites were comparable between ESCs and MEFs with minor alterations.

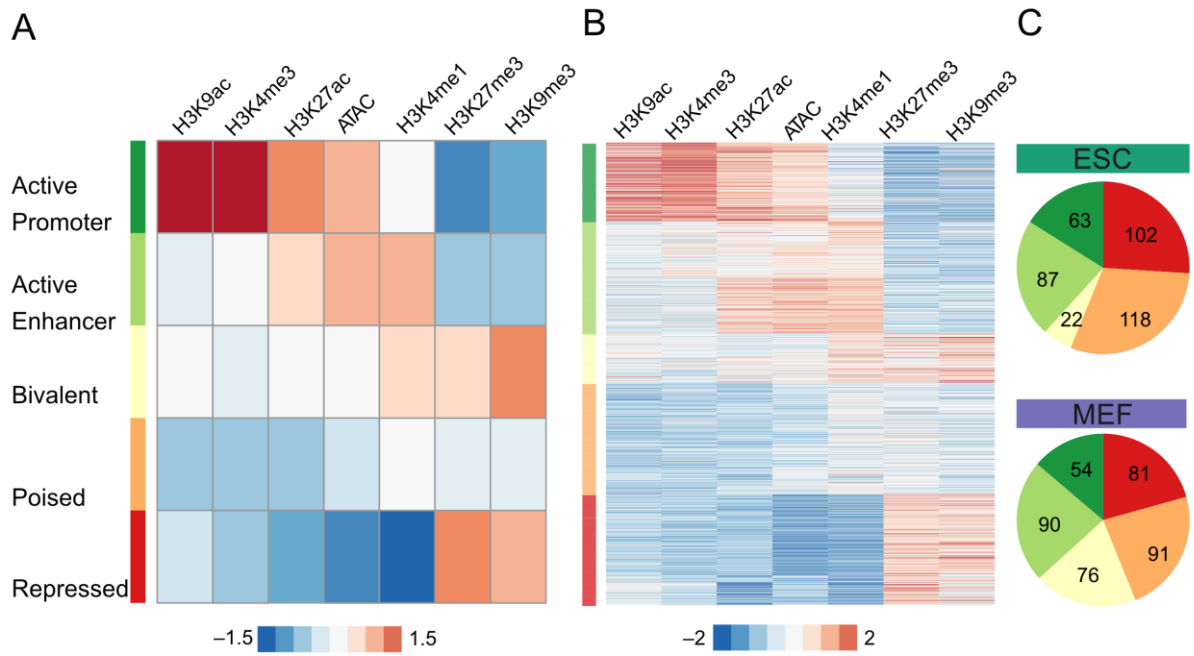

Figure 36: Assignment of ISGF3 binding sites to chromatin states

(A) Unsupervised k-means clustering of normalized read counts of histone modifications (H3K4me1, H3K4me3, H3K27ac, H3K9ac, H3K27me3, H3K9me3) and chromatin accessibility (ATAC) data in untreated ESCs and MEFs at 392 ISGF3 sites. The averaged signal was used for coloring each field. The rows were ordered based on the averaged H3K27ac signal. The columns were ordered manually. The normalized counts were z-transformed and log transformed. (B) Same data as in (A) with all sites in each cluster. (C) Piecharts of chromatin states of the potential 392 ISGF3 binding sites in 0 h ESCs (upper) and 0 h MEFs (lower).

As next step, the chromatin states of ISGF3 sites in ESCs and MEFs were linked for each site. The resulting figure showed a highly dynamic picture, pointing out a high number ISGF3 binding sites with highly different chromatin states between cell types (**Figure 37A**). In each chromatin state cluster, a solid number of ISGF3 sites did not change chromatin state between the cell types. Many ISGF3 binding sites moved from the "Active Enhancer" state in ESCs to "Poised" or "Repressed" chromatin state in MEFs. The reversed picture was found for MEFs. Many repressed sites in ESCs were found in more active clusters like "Active Enhancer" or "Bivalent" in MEFs. Setting a

focus on cell type specific binding events, we found a 3-fold reduction of ESC-specific binding events in "Active Enhancers" (51 to 16) from ESCs to MEFs and a 5-fold induction of "Repressed" sites (11 to 53) in MEFs (**Figure 37B**). A similar picture was found in MEFs, MEF-specific sites were strongly enriched in the "Repressed" cluster in ESCs, (84 to 21) while more of these sites showed "Active Enhancer" (17 to 38) and "Bivalent" (3 to 56) characteristics in MEFs (**Figure 37C**). In summary, we identified a highly diverse chromatin environment of ISGF3 binding sites in ESCs and MEFs. Linking the different chromatin states of those sites, revealed that cell type specific sites were enriched in enhancer cluster in the corresponding cell type, while present in the other cell type, an enrichment in the "Repressed" clusters were found. Thus, the main chromatin changes that determine the cell type specific binding of ISGF3 occur for the "Active Enhancer" and the "Repressed" chromatin state.

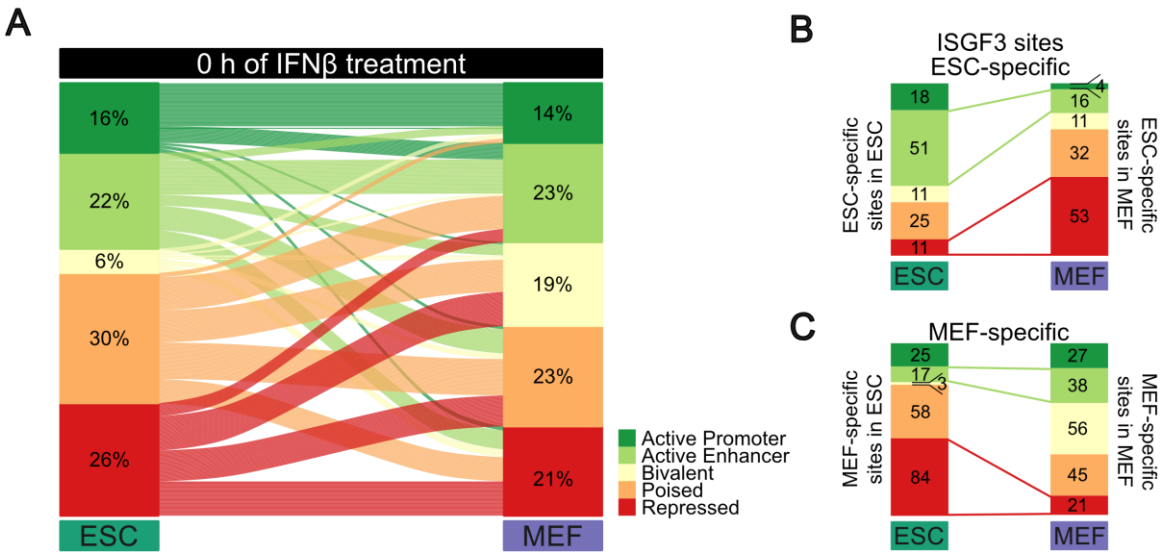

Figure 37: Chromatin state changes of ISGF3 sites in ESCs and MEFs

(A) Alluvial diagram linking chromate states assignments of ISGF3 sites in untreated ESCs to their respective states in untreated MEFs. The left pillar represented the distribution of 392 ISGF3 binding sites over the five states in ESCs and the right one in MEFs. Each line between the pillars was linking the chromatin state in ESCs to the same ISGF3 site in MEFs. Coloring of the connecting lines was based on ESC chromatin state assignments. (B) Chromatin state distribution of 118 ESC-specific ISGF3 sites identified in ESCs (left) and MEFs (right). The green lines link the strongest decreased cluster and the red lines the strongest gained cluster. (C) Same as (B) but for 184 MEF-specific ISGF3 sites.

3.3.4. Accessible chromatin and H3K4me1 mark were permissive for ISGF3 binding

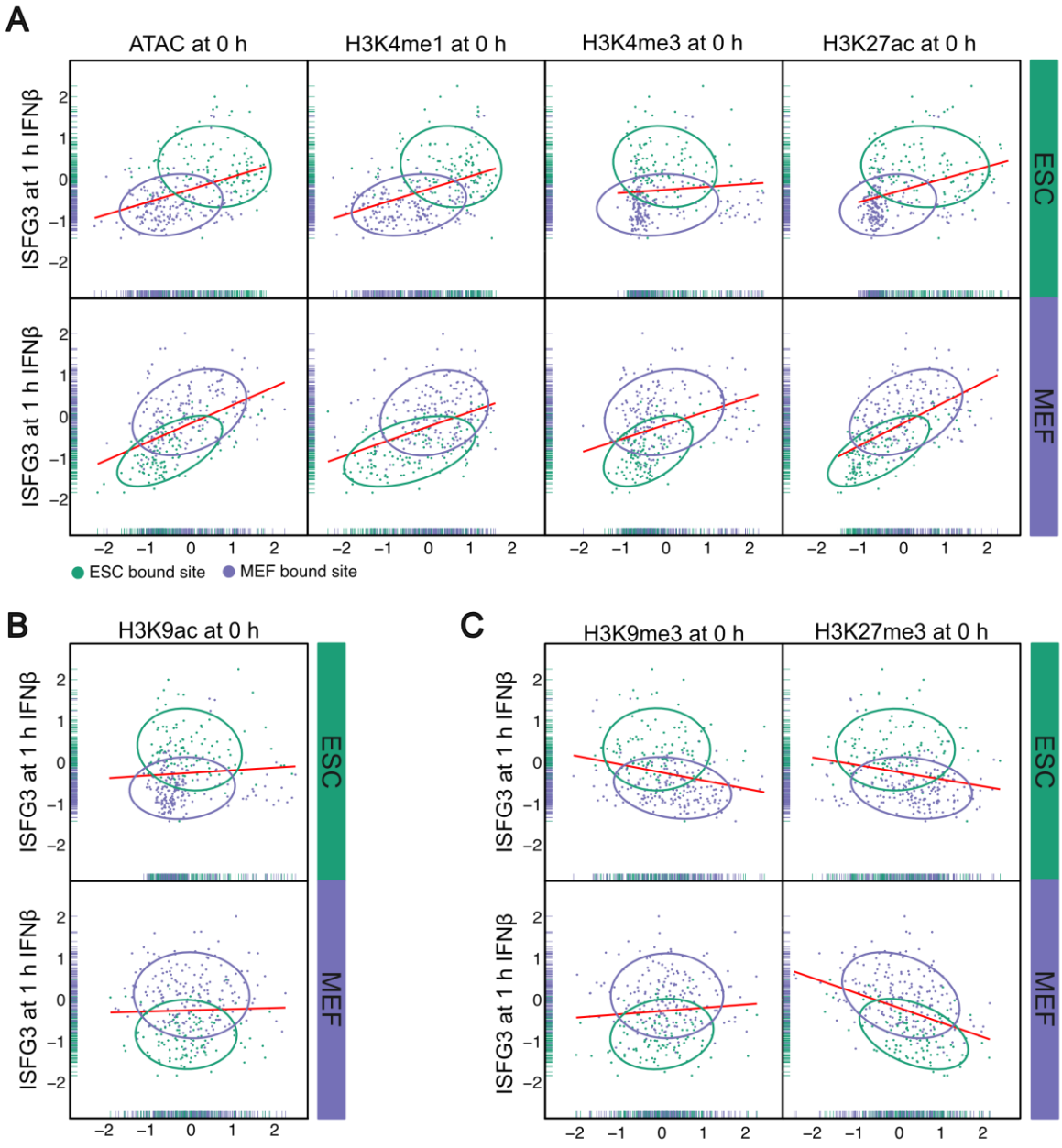

Figure 38: Correlations between ISGF3 binding and chromatin features.

Scatterplots of normalized read counts of ISGF3 binding from ChIP-seq at 1 h of IFNβ treatment against chromatin marks before stimulation (0 h). ESC-specific ISGF3 sites were labeled in green and MEF-specific in violet. Upper plot showed rest in ESCs and the lower one in MEFs. Red lines marked correlation coefficient of all data points by Spearman. (A) Correlation with chromatin accessibility (ATAC) and active chromatin marks H3K4me1, H3K4me3 and H3K27ac. (B) Promoter specific mark H3K9ac was correlated with ISGF3 and (C) repressive marks H3K9me3 and H3K9me3.

Based on the previous analysis, I hypothesized that a chromatin fingerprint in unstimulated cells directs the cell type specific binding of ISGF3 complexes upon

IFNβ stimulation. Scatterplots were generated to correlate the binding of ISGF3 after 1 h of IFNβ stimulation with various chromatin marks before treatment (**Figure 38**). The ESC-specific sites (green) showed higher basal levels of ATAC, H3K4me1, H3K4me3, H3K27ac and H3K9ac in ESCs (**Figure 38A/B**, upper). MEF-specific sites (violet) exhibited low levels of those marks and of ISGF3 binding in ESCs. A positive correlation between ISGF3 binding was only found with ATAC, H3K4me1 and H3K27ac. A very similar picture was found for MEF-specific sites in MEFs (**Figure 38A/B**, lower), with the exception that also H3K4me3 showed a positive correlation with ISGF3 binding. The repressive marks were even anti-correlated with the TF binding, except for H3K9m3 in MEFs (**Figure 38C**).

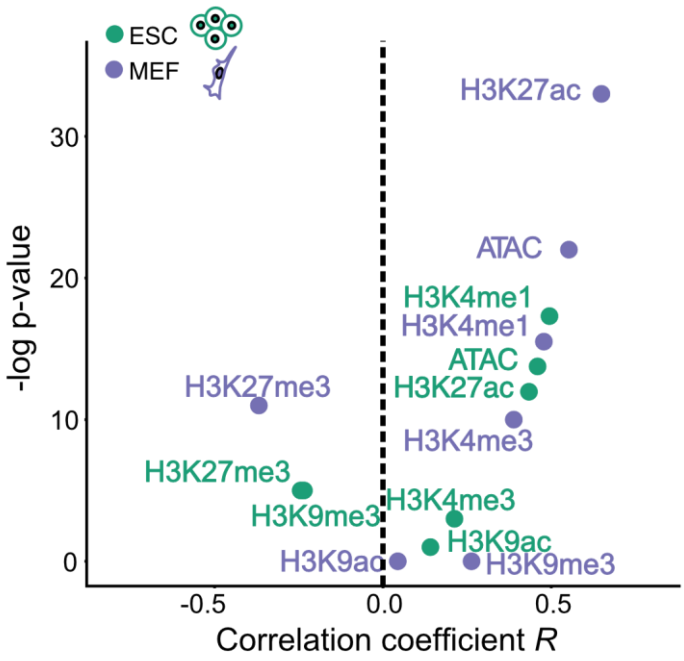

Figure 39: Identification of permissive chromatin signature for ISGF3 binding

Scatter plot of the p-value versus correlation coefficient between ESC-specific and MEF-specific ISGF3 sites for various marks. Signal from ESCs were marked green and from MEFs were labeled in purple. The strongest positive correlation of ISGF3 binding were found for H3K4me1, H3K27ac and ATAC while pre-existing H3K27me3 had a negative effect.

For both cell types, the enrichment scores for each histone mark at all ISGF3 sites were compared. The differences between histone mark enrichments in each cell type were tested and p-values calculated. These p-values were used together with the correlation coefficient from the previous analysis to define a finger print, which is permissive for ISGF3 binding. The analysis revealed that ATAC and H3K4me1 were found a most robust markers for ISGF3 binding in both cell types (**Figure 39**). H3K27ac as well was strongly enriched for ESCs but weaker for MEFs. In addition, the repressive mark H3K27me3 showed the best anti-correlation to ISGF3 binding. In conclusion, we suppose a combination of the presence of H3K4me1 with a certain level of chromatin accessibility and the absence of repressive marks as H3K27me3 to be promotive for ISGF3 binding.

3.3.5. Depletion of MLL3/4 did not impact ISG response in ESCs

To test the previous finding, we hypothesized that the depletion of a H3K4me1 histone methyltransferase could impact the ISGs induction in ESCs. Therefore, I studied mutant ESC cell lines which carried either a double catalytic dead version (dCD) of MLL3 and MLL4 or a double knock out (dKO) of those two enzymes (Dorighi *et al.* 2017). These cell lines and the corresponding WT control were treated for 1 h and 6 h with IFNβ. As initial experiment, RNA-seq was performed. The induction of ISGs was similar to the previously descirbed effects in ESCs (**Figure 40**). In the WT, I identified 106 upregulated genes, 132 in the dCD mutant and 130 in the dKO mutant. The observed induction levels at 6 h were stronger in the WT cells, while the mutants resulted in more upregulated genes. In summary, the number of detected ISGs and their induction levels were comparable to ESC WT cells.

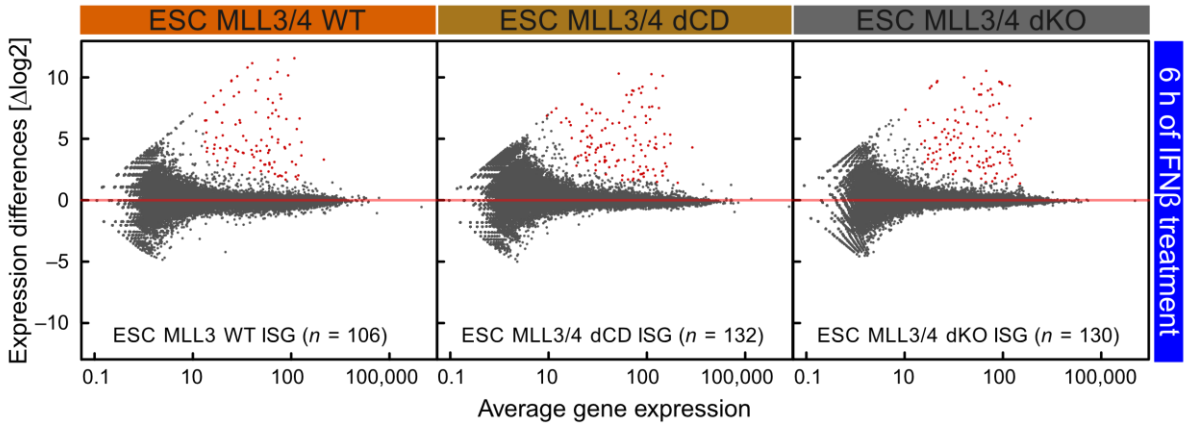

Figure 40: ISG induction in MLL3/4 mutant and WT ESCs

MA plot of differentially expressed genes between IFNβ stimulation and untreated controls by DESeq2. Significantly different expressed genes (padj<0.05 & 1.5 - fold upregulated) were marked in red. The ESC MLL3/4 WT cells showed 106 upregulated genes. The mutated line ESC MLL3/4 dCD resulted in 132 ISGs and the ESC MLL3/4 dKO in 130 induced genes upon IFNβ treatment.

The overlaps of the three ISG lists showed that the majority of 99 genes were upregulated independently of the MLL3/4 state (**Figure 41A**). 39 genes were found in the dCD and dKO mutants only, of while 11 were specific for the dCD and 10 for dKO. A single gene was found to be specifically enriched in WT cells (**Figure 41B/C**). The basally expressed TFs *Irf9*, *Stat1* and *Stat2* as well as the initially inactive genes *Irf7*, *Rtp4* and *Usp18* did not show any differences in induction between cell types. Minor differences like the stronger 1 h gene induction of *Usp18* in the dKO or the higher 6 h levels of *Rtp4* (**Figure 41C**) did not justify further experiments with this particular cell lines. In conclusion, the mutants of MLL3/4 were not sufficient to

impact the induction of ISGs in ESCs. Consequently, no further experiments were conducted to characterize the impact of MLL3/4 perturbation on IFNβ induced ISG expression.

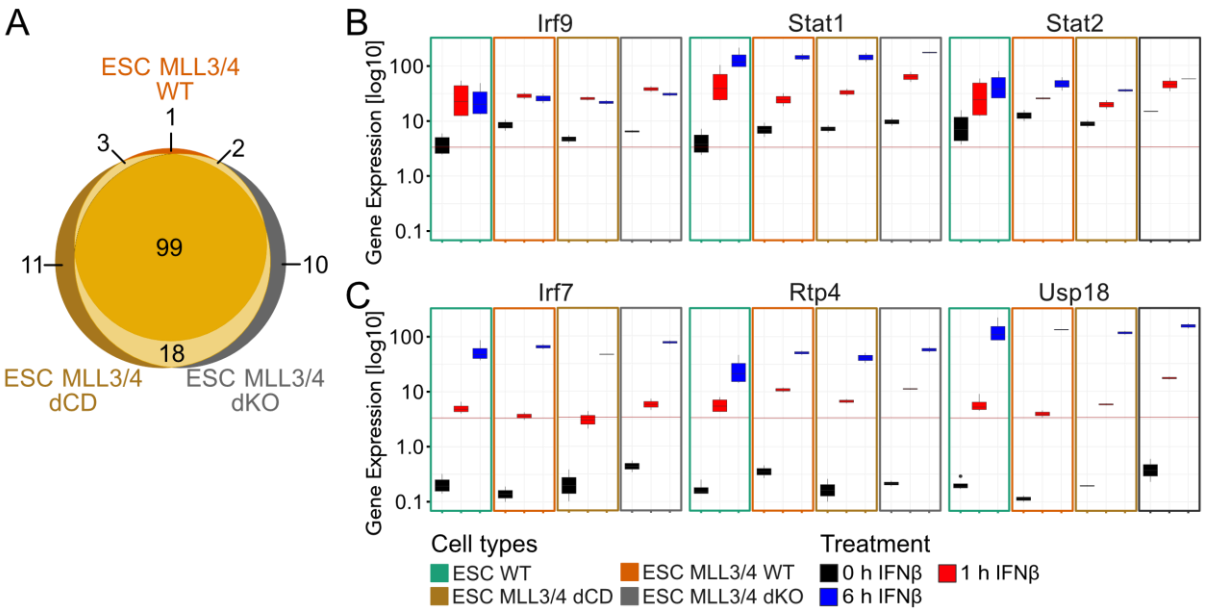

Figure 41: ISG induction levels in WT and MLL3/4 mutant ESCs

(A) Venn diagram of ISG lists from ESC MLL3/4 WT, dCD and dKO. (B) Boxplots of normalized RNA read counts (TPM) for expressed common ISGs *Irf9*, *Stat1* and *Stat2* in ESC MLL3/4 WT, dCD and dKO and the previously used ESCs. Red lines indicated the expression threshold for ESCs. (C) Same as in (B) for not basally expressed ISGs *Irf7*, *Rtp4* and *Usp18*.

4. Discussion

The induction of ISGs via the type I IFN system is an essential part of the innate immunity. As first line of defense against virus infections every cell type has the potential to respond to such threats. However, stem cells have a unique, uncharacteristic response compared to differentiated cells. In this thesis, I investigated the cell type specific responses of ESCs compared to differentiated NPCs and MEFs cells with an multiomics sequencing approach. The cell type specific ISG patterns were identified on a genome-wide level. The strongest overall IFNβ response was found to in MEFs while all cell types displayed cell type specific ISG activation and STAT1 and STAT2 binding patterns. ISGF3 complexes at promoters were highly correlated with ISG induction while sites that had only STAT1 bound show little activation capacity. However, many binding sites were intergenic or intronic at putative enhancers and could be partially linked to ISGs inductions via a scATAC-seq co-accessibility analysis. By correlating chromatin features with STAT binding, it was revealed how chromatin context, ISFG3 binding and ISGs are connected, which provides insight into the cell type specific innate response.

4.1. IFNβ dependent gene expression patterns

4.1.1. IFNβ induces ISGs response in a cell type specific manner

The first studies to characterize the ISG response upon IFN stimulation identified between 200 to 500 upregulated genes in different cell types (de Veer *et al.* 2001; Der *et al.* 1998). Other studies treated eleven hemopoietic mouse cell types with recombinant IFNα and detected a total of 975 genes in at least one cell type (Mostafavi *et al.* 2016). A set of 166 ISGs were identified in all cell types. In my experiments, ESCs treatment with IFNβ induced the upregulation of 191 genes. In differentiated cells, between 244 ISGs in NPCs and 463 in MEFs were found to be induced. 143 ISGs were identified as common responding genes in all our cell types, which was in the same range as previously reported numbers. Recent studies reported that ESCs had an underdeveloped response to interferons in contrast to differentiated cells like mouse embryonic fibroblasts (MEF) (Guo *et al.* 2015; Wang *et al.* 2014). The Wang *et al.* demonstrated that ESCs in response to La Crosse virus (LACV) failed to induce the two ISGs *Isg15* and *Oas1a*. The combination of LACV and IFNβ caused the gene to respond up to a 7-fold increase. The same setup in a MEFs resulted in induction of 40 to 100-fold for LACV only and 45 to 125-fold for LACV and IFNβ (Wang *et al.* 2014).

The inductions patterns of common ISGs were highly similar between cell types but the maximum level of expression was 10- to 100-fold higher in MEFs than in ESCs. Gene regulation can also occur both at the level gene induction or mRNA stability (Liu *et al.* 2014). Accordingly, I characterized both the steady state and the nascent RNA levels of ISGs in ESCs and MEFs. Similar pattern on nascent mRNA levels were identified as for mature mRNAs. Consequently, in the context of IFNβ induced gene expression, no evidence was found that the attenuated ISG levels in ESCs were linked to mRNA stability. I conclude that the regulation of ISGs takes place predominantly at the level of gene induction rather than by changing mRNA stability.

4.1.2. ESCs showed a homogeneous response upon IFNβ induction

Previous studies showed by single cell RNA-seq that upon West Nile Virus infections, only a subset of the cell population started expression IFNβ mRNAs (O'Neal *et al.* 2019). The authors concluded, that in their system, the response to a virus infection was heterogeneous but no detailed analysis was done to characterize the response upon the production of IFNβ. I tested the hypothesis that only a small fraction of ESCs responded to IFNβ stimulation with similar induction levels as in MEFs while the majority of ESCs failed to do respond. The transcriptional heterogeneity upon IFNβ stimulation was investigated in ESCs and she scRNA-seq approach revealed that the entire ESC population was responsive after 6 h of IFNβ treatment. The original heterogeneous response after 1 h was transient and turned into a very homogeneous response after 6 h. Based on inspection of selected ISG targets from bulk RNA-seq data, I concluded that many induced ISGs at 1 h were lowly expressed and not detected in the majority of cells. At the later time point, the response became more homogeneous and the selected ISGs become clearly induced in a high fraction of cells. A clear separation of 6 h IFNβ treated cells from the other two conditions was revealed. Conclusively, the induction of ISGs was sufficient to clearly separate the 6 h cells from the early (1 h) and unstimulated (0 h) time point. In general, the single cell experiments validated the detected ISG pattern in ESCs and showed that the entire ESC population responded homogeneously after 6 h of IFNβ treatment. Consequently, the attenuated ISG induction in ESCs was not promoted by a heterogeneous response on single cell level.

4.1.3. Attenuated ISG response in ESCs was associated with lower STAT levels

It was shown by RT-qPCR that ESCs express main components of the IFN signaling pathway with only the interferon-α/β receptor *Ifnar1* being significantly downregulated while *Stat2*, *Tyk2* and *Irf9* were upregulated (Wang *et al.* 2014). Based on comprehensive RNA-seq in this thesis, I confirmed the downregulation of *Ifnar1* while the differences for *Stat2*, *Tyk2* and *Irf9* were not significant on the $p < 0.01$ significance level. However, we additional detected a strong downregulation of *Ifnar1/2*, the receptors for IFNγ, and the *Jak1/2* kinases in ESCs relative to differentiated cells. Furthermore, the basal protein levels of STAT1 and STAT2 were lower in ESCs. Thus, a globally reduced IFN response can be assigned to lower levels of key components of the signaling pathways. To some extent this might be related to the lack of basal levels of IFNα/β/γ as ESCs are unable to express interferons themselves (Wang *et al.* 2013). Additionally, a positive feedback loop that could amplify the IFN response is also lacking in ESCs (Wang *et al.* 2013). In addition, the phosphorylation at serine 727 of STAT1 was enriched in MEFs in contrast to ESCs. This modification is associated with transcription and found mainly on promoter bound STAT1 (Bancerek *et al.* 2013). The nuclear kinase CDK8 is required for this PTM and its RNA level was lower in ESCs as well. The lower level of $STAT_{p727}$ could be an additional possibility of contributions to the attenuated response in ESCs.

Literature provides additional hypotheses to explain the attenuated antiviral response in ESCs (Guo 2019). The hematopoietic pioneer TF PU.1, normally found in macrophages, creates new enhancers during differentiation (Ghisletti *et al.* 2010). These sites become marked by accessible chromatin and the enrichment for H3K4me1 mark and the maintenance of these features allow a faster antiviral response upon future exposer to virus or interferon. In addition, the overexpression of PU.1 in fibroblasts is sufficient to create new enhancers sites, known before to be only functional in macrophages (Ostuni *et al.* 2013). Further, these enhancers were specifically bound by the TFs STAT1/6 upon LPS stimulation. The idea is that ESCs could be in a naïve state as they never have been exposed to a pioneering factor like PU.1. In contrast, MEFs were isolated at a much later time point during murine development, would have a higher chance of being exposed to such a pre-stimulation, which could be linked to differentiation signals. To test this hypothesis, the overexpression of a potent pioneering TF like PU.1 in ESCs would be required. At first, the ISG induction levels would have to be investigated, followed by the characterization of STAT binding profiles. Alternatively, multiple rounds of IFN

stimulation could help to train ESCs to become a stronger responder. Such a training effect, presumably based on short time epigenetic memory, was shown for IFNβ stimulations in fibroblasts (Gupta *et al.* 2015). In this publication, the authors demonstrate that multiple rounds of stimulations cause faster and stronger ISG induction and this was linked to H3K36me3 and H3.3 levels.

In summary, MEFs do have a higher potential to active ISGs, which explains the attenuated ISG response in ESCs. However, I identified a set of 33 ISGs only responding in ESCs and in none of the differentiated cell types. I hypothesized that other gene regulatory mechanisms were responsible for the ESC-specific induction of these genes. Cell type specific chromatin states might direct the binding of STAT complexes. In order to rationalize the differences in the ISG patterns, I wanted to investigate the binding sites for STAT1 and STAT2 upon IFNβ stimulation in ESCs and MEFs.

4.2. Function of STAT complexes on ISG induction patterns

I performed ChIP-seq experiments for STAT1 and STAT2 to characterize the binding sites of STAT1 homodimers and ISGF3 on a genome-wide scale. In previous studies, the binding of STAT1 was mapped in HeLa S3 cells (Robertson *et al.* 2007) and STAT2 in B cells (Mostafavi *et al.* 2016). In this thesis, the binding of STAT1 and STAT2 in ESCs and MEFs were characterized over a IFNβ treatment time course (0 h-1 h-6 h).

4.2.1. Subgrouping of STAT complex was confirmed by motif enrichments

An essential point for all ChIP-seq experiments is to validate the used antibody (AB). In addition, validate on the bench site, a key computational way to validate the observed enrichment patterns, is by characterizing the motifs found in the ChIP-seq peaks. For the STAT family, these motifs are known and a great stat point to validate the quality of a ChIP-seq experiment. The STAT1 homodimer is binding to the GAS motif (STAT1 motif), while the ISGF3 complex is recognizing the ISRE sites (Stark and Darnell 2012). For the analysis in this thesis I defined sites enriched by STAT1 as homodimer binding sites and co-enrichments for STAT1 and STAT2 as ISGF3. Recently, it was shown that STAT2 and IRF9 are pre-bound in the cytoplasm independent of phosphorylation and IFN stimulation (Rengachari *et al.* 2018).

Therefore, I concluded that the detection of STAT2 indicates the presence of STAT2-IRF9. This was confirmed by motif analysis, which showed that 85.7 % of STAT2 binding sites in ESCs and 87.6 % in MEFs were enriched for IRF motifs in contrast to STAT1 only peaks, were STAT motifs were found in 65.8 % and 85.4 % of peaks. Based on motif analysis, we confirmed IRF9 as the driving force behind the DNA binding of ISGF3 complexes, as 82.7 % in ESCs and 90.2 % in MEFs harbored an IRF motif. These results were in line with the literature and confirmed the quality of the ChIP-seq data (Loutfy *et al.* 2003). Interestingly, ISGF3 peak sites did also contain STAT motifs, between 49.5 % and 53.2 % of all cases. This makes sense, as these genes can be targeted by other STAT complexes like STAT1 homodimers upon IFN stimulations. Consequently, the same ISGs are able to be targeted by multiple TFs. Recently, it was reported that Type I and III interferon response in the same cell type caused the same ISGs to respond with different kinetics (Pervolaraki *et al.* 2018). In line with this finding, the presence of multiple recognition motifs for STAT or IRF TFs in ISG promoter would be a robust solution to ensure the activation of these antiviral genes as part of the innate immunity. I also identified STAT2 peaks without STAT1 enrichment, especially in MEFs these peaks were enriched. This can be explained as evidence was found that other complexes containing STAT2 are existing but have limited activating functions like STAT2-STAT3, STAT2-STAT6 or STAT2-STAT6-IRF9 in various cell types (Blaszczyk *et al.* 2016). Further, STAT2-IRF9 complex without a second STAT family member are translocated into the nucleus independent of IFN stimulation. In human fibrosarcoma cells this complex is responsible for the basal gene expression of ISGs (Blaszczyk *et al.* 2015). In my data, I did not find evidence of STAT2 binding before the IFNβ stimulus. In summary, by motif analysis I confirmed the specificity of identified STAT complexes, which were enriched for motifs described in literature. Next, I wanted to characterize where in the mouse genome the binding happened.

4.2.2. Promoter binding of ISGF3 induced gene expression

A TF binding event at a promoter is normally linked with the activation of the gene. (Lee and Young 2013; Singh *et al.* 2014). Our data showed, that STAT1, ISGF3 and STAT2 complexes were bound to promoters. For ISGF3 the fraction of promoter binding was highest with 41.3 % to 48.9 % for ESCs and MEFs, respectively. This was followed by STAT1 homodimer binding in MEFs with 38.0 %. For STAT2 in both cell types and STAT1 in ESCs the fraction was below 25.0 %. When characterizing the effects of these binding events, I found ISGF3 was mainly responsible for ISG induction, where out of 86 bound promoters, 69 were responding. In MEFs, 105 out

of 148 genes were induced. In addition, in MEFs STAT2 also caused many genes to respond (34 of 75). On the one hand, these results were in line with the original literature, where ISGF3 was identified as the main activator of type I IFN response (Au-Yeung *et al.* 2013). On the other hand, many promoters were enriched for STAT binding but failed to respond to IFNβ stimulation. Consequently, I conclude that STAT binding to a promoter is not sufficient for ISG upregulation but that ISGF3 was the main activator upon IFNβ stimulation in ESCs and MEFs.

Many non-responding genes bound by a STAT complex had high basal expression in the analyzed cell types. One possibility would be, that high basal expression would prevent these genes to get further induced upon STAT binding as their transcriptional potential was saturated. This might be partially true but for many of these genes no tendency of induction was found. It seemed that the binding has no effect at all. This could point the requirement of (unknown) co-activators that were absent in the particular cases studied. The missing co-activators for some genes, would also explain, why repressed or lowly expressed but STAT bound genes were not responding, although they had the potential to do so. Alternatively, the binding could be spurious and the mere consequence of highly accessible chromatin, which is a hallmark of active genes. ChIP-seq data provides a snapshot of a cell population and enriches then signals. TFs are scanning the genome for their recognition motif by 3D diffusion and local motions (Suter 2020). Active promoters with highly accessible chromatin environments are easier to access for searching TFs. As many promoters are enriched for TF binding motifs, it would be likely, that STATs find a potential binding motif in these regions by chance. In summary, we could link 84 (10 by STAT1, 69 by ISGF3, 5 by STAT2) out of 191 ISGs (44.0 %) with direct binding of STAT complexes to a promoter in ESCs. In MEFs, 150 (11 by STAT1, 105 by ISGF3, 34 by STAT2) of 463 IGSs (32.4 %) were associated with TF binding and again ISGF3 binding was found at the majority of promoters. STAT-bound promoters, which did not respond, might be a consequence of highly open chromatin, missing co-factors or saturated gene expression levels. As the activation potential at STAT complex at promoters was diverse, I focused on the main activator, ISGF3, and investigate the activation potential of non-promoter binding sites.

4.2.3. The contribution of non-promoter bound ISGF3 to ISG induction

More than 50% of genomic regions bound by STATs were found in intergenic or intronic sites. Binding sites in such regions are normally associated with binding to regulatory elements like enhancer (Agrawal *et al.* 2010). The observed STAT binding patterns in both, ESCs and MEFs, were pointing to a functional role of regulatory elements as part of the antiviral response. A previous study linked binding of STATs to enhancer regulation in T cell populations (Vahedi *et al.* 2012). There are challenges associated with these binding events, as it is not straightforward to links those STAT binding sites to a target gene. Enhancer-promoter interaction can be millions of nucleotides away from each other (Furlong and Levine 2018). A classical approach is linking the intergenic binding site with the nearest genes and assumes that this gene is regulated. A comprehensive study concluded that only around 7 % of looping interaction are with the nearest gene and that consequently genomic proximity is not a reliable predictor (Sanyal *et al.* 2012). In our data set, we identified 69 ISGs with ISGF3 complex bound directly at the promoter. The nearest gene approach resulted in 70 ISGs including the promoter bound sites. In other words, one additional ISGs was identified with the nearest gene approach. Consequently, for this data set it was insufficient to use a proximity approach to characterize the function of non-promoter bound ISGF3 to the innate immunity.

A more promising way to link genomic loci was introduced with the chromatin conformation capture (3C) technology (Dekker *et al.* 2002) and the genome-wide advanced NGS-based method HiC (Belton *et al.* 2012). HiC analyses allow to investigate interaction between any two genomic loci and nearly independent of their distance. It was used to characterize topological associated domains (TADs), higher order chromatin structures and also to identify enhancer and promoter interactions (Kieffer-Kwon *et al.* 2013). One limitation of these technologies is the resolution, which is limited to 500 bp (Wang *et al.* 2018). This means, that HiC can identify interactions between two regions each with a length of 500 bp. Often data sets have a resolution of up to 10 kb or even higher, which is sufficient to define TADs but makes it difficult to pinpoint distinct enhancer-promoter interactions. Mammalian promoters are between 100 bp and 1,000 bp in length, while enhancers can be between 10 bp and 1,000bp (Li and Wunderlich 2017). The exact binding motifs for members of the STAT or IRF family are 8-12 bp in length. Consequently, the identification of regulatory elements in data sets with 1 kb+ resolutions might be insufficient. The novel technology of single cell ATAC-seq can be used to enhance the

resolution and identify promoter-enhancer links (Buenrostro *et al.* 2015b; Cusanovich *et al.* 2015). This is achieved by using the information from thousands of cells and computing correlations of accessibility between sites of interest. Co-regulated sites are more likely to appear in the same cell and are consequently highly coregulated. By using the ISGF3 binding sites in ESCs as anchor point, we identified a strong gain of interactions for these sites upon IFNβ treatment with ArchR (https://www.archrproject.com/). The strongest gain of interaction was found for intergenic and intronic sites. As these sites were often enhancers, we identified links of these sites to ISG promoters and characterized their induction patterns. With this approach we were able to identify 30 ISGs with links from distal ISGF3 binding sites up to 500 kb away. We further demonstrated that the link from a non-promoter ISGF3 binding can upregulate the ISGs to the same level as an ISGF3 binding directly at the promoter. In summary, this result showed that enhancer bindings of STAT complexes are important for ISG regulation upon IFNβ stimulation. However, a validation experiment would be needed to strengthen this finding. A CRISPR-Cas9 based tools allow to complete deletion of such an enhancer and investigate the ISG induction patterns (Moorthy and Mitchell 2016). It would be interesting to see, if the enhancer is completely responsible for activating its target ISG or part of a regulatory network. Deletion of the binding platform of ISGF3 would be one way to go. Alternatively, the chromatin state of these sites could be targeted. In this thesis, I focused to investigate, how chromatin states impact the cell type specific ISGF3 binding events.

4.3. Chromatin state dependent ISGF3 binding

The comparison of STAT binding events between ESCs and MEFs revealed that the majority of bound sites were indeed cell type specific. Genomic loci bound by ISGF3 in ESCs failed to be recognized or stable bound in MEFs upon the same type of IFN stimulation. Consequently, the regulation of the binding of STAT complexes is essential to understand transcriptional responses and I tested the role of chromatin states for the binding patterns of STATs.

4.3.1. A fraction of ISRE motifs was bound by ISGF3

The focus was set on the characterization of binding sites for the main activator ISGF3. The motivation of investigating the chromatin states for ISGf3 binding sites was supported by the total number of possible ISGF3 binding motifs compared to the numbers of detected ISGF3 peaks. Originally, the ISGF3 complex was found to

recognize a DNA motif named ISRE element (Loutfy *et al.* 2003; Schindler *et al.* 1992). Based on HOMER motif data base, the ISRE motif is present in the mouse genome 134,069 times. Previous studies, of B splenocytes treated in vivo for 2 h with IFNα, revealed for STAT2 6,703 peaks (Mostafavi *et al.* 2016). The specificity of these peaks was not validated via motif analysis, thereby it is not possible to precisely know how many ISRE sites were bound. In best case, all 6,703 STAT2 peaks contained an ISRE sites, which would result in a maximum of around 5.0 % of all existing ISRE elements to be bound at the same time. Considering this numbers, mechanisms to guide ISGF3 complexes to find the "right" motives are required. Originally, we aimed to establish a chromatin fingerprint for all ISRE elements to understand, why and how a specific ISRE subset was bound. As our chromatin data set contained the most abundant and best characterized marks, there are many more histone modifications or regulatory mechanism to be considered. On example are non-canonical histone variants. It was shown that the variants H3.3 is promotive for ISG expression (Tamura *et al.* 2009) and H2A.Z blocks the binding of STATs (Au-Yeung and Horvath 2018). Consequently, we decided to take opportunity of our cell type specific ISGF3 binding date Instead of taking all possible binding sites into consideration for chromatin characterization, we used the combined ISGF3 binding sites from ESCs and MEFs, for which we know that ISGF3 binding can occur. I focused on dissecting the differences in chromatin environment of cell type specific binding sites to identify, why ISGF3 bound in one cell type but not in the other one.

4.3.2. Genomic regions defined by histone mark abundancy

An TF binding site is typical less than 50 bp in length, depleted of nucleosomes and consequently marked as open chromatin (Nie *et al.* 2014). The nucleosomes in the surrounding are marked with specific histone marks. I aimed to characterize the chromatin environment of ISGF3 binding sites by analyzing the histone marks of the nucleosomes in close proximity. Genomic regions were categorized using the following the guidelines characterized by the Epigenome Roadmap Consortium (Roadmap Epigenomics *et al.* 2015). In short, (I) "Active promoters" are enriched for H3K4me3, H3K9ac and H3K27ac ("Active Promoter") (Ernst *et al.* 2011), while (II) Active enhancers are marked by H3K4me1 and H3K27ac ("Active Enhancer") (Creyghton *et al.* 2010). (III) Enhancers with H3K4me1 only are in a poised state ("Poised") (Creyghton *et al.* 2010). (IV) The presence of active marks like H3K4me3 and repressive marks like H3K27me3 is called bivalent ("Bivalent") (Bernstein *et al.* 2006), while (V) the presence of only repressive marks like H3K9me3 or H3K27me3 denote regions as repressed chromatin sites ("Repressed") (Lehnertz *et al.* 2003;

Morey and Helin 2010). These five groups represent five different chromatin states before stimulation located in the surrounding of ISGf3 binding sites.

The distribution of ISGF3 binding sites in ESCs and MEFs over this five clusters looked very similar for the “Active Promoter” and “Active Enhancer” group in terms of total numbers. In MEFs more sites were found in the “Bivalent” group and less in the “Repressed”. Both are enriched for repressive chromatin marks (H3K9me3 and H3K27me3), but the “Bivalent” has a higher level of H3K4me1 and H3K4me3. It is known that pioneering TFs like PU.1 are able to activate repressed enhancers during differentiation (Ghisletti *et al.* 2010). These sites are permanently marked with H3K4me1 and a higher level of accessibility. Consequently, upon the repeated activation of specific TFs like STATs these sites can readily be bound, which would accelerate the deposition of active marks like H3K27ac. For IFNβ this is not known but for LPS this effect has been demonstrated (Ghisletti *et al.* 2010). As LPS also triggers an antiviral response, I we hypothesize that a similar mechanism would apply to the IFNβ stimulation pathway. In MEFs, the higher abundance of ISGF3 sites in the “Bivalent” cluster could be a consequence of such a developmental priming.

4.3.3. Function of ISGF3 bound enhancers in ISG induction

In general, for more than 50 % of ISGF3 sites the associated chromatin state was changed between ESCs and MEFs. Especially for the cell type specific ISGF3 sites, many active enhancers were bound upon IFNβ stimulation. In the context of innate immunity, Natural Killer cells show higher gain of ATAC signals at intergenic and intronic regions upon mouse cytomegalovirus (MCMV) infection (Lau *et al.* 2018). In this cellular context, a highly dynamic change of chromatin states is described during the maturation of this immune cells. Interestingly, a positive feedback loop activated the signaling molecules interleukin 12/18 and cause the activation of STAT4, which then binds these newly accessible intergenic sites (Rapp *et al.* 2017). These regulatory sites stay then marked by H3K4me1 to allow a faster and stronger future response (Ostuni *et al.* 2013). The authors are calling that an epigenetic memory effect. PU.1 is the master regulator in macrophages and pioneering TF, which is responsible for these new enhancers. The potential of PU.1 allows even to activate macrophage specific enhancers and genes in fibroblasts by overexpressing this factor (Ghisletti *et al.* 2010).

In MEFs, the MEF-specific ISGF3 sites were enriched for "Active Enhancer" and "Poised" clusters, i.e. sites enriched for H3K4me1 (**Figure 42**). I hypothesize that a potent pioneering TF like PU.1 could be responsible for marking these sites during in vivo differentiation. Upon IFNβ treatment, theses marked enhancer sites were targeted by activated STAT1 and STAT2 and contributed to the enhanced ISG induction levels observed in MEFs. In ESCs, such a priming event was not apparent, yet, and most of the MEF-specific sites were marked only with the repressive chromatin marks H3K27me3 and H3K9me3 but not H3K4me1 (**Figure 42**). The absence of H3K4me1 was then sufficient to prevent the binding of ISGF3.

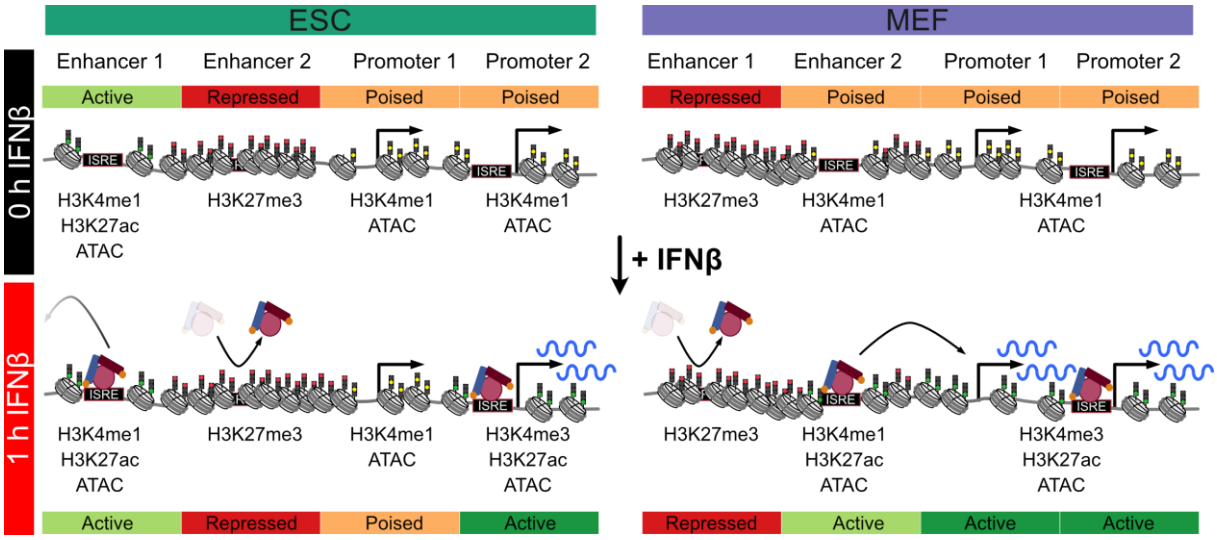

Figure 42: Model for cell type specific ISG induction via altered ISGF3 binding

Summary of gene induction via ISGF3 binding. In ESCS, common ESC-specific enhancers are in an active state and marked by H3K4me1, H3K27ac and ATATC. MEF-specific enhancers are repressed by H3K27me3. ISG promoter are in a poised state with low levels of H3K4me1, H3K4me3 and low levels of accessibility. In MEFs, ESC-specific enhancers are repressed, and MEF-specific enhancers are not active but marked by H3K4me1. ISG promoters are in similar states as in ESCs. Upon IFN stimulation, accessible sites are bound by ISGF3 complex. Some ISGs have ISRE motifs in their promoter and ISGF3 can bind directly. Other ISGs are activated by enhancer bound ISGF3. In ESCs some enhancer bound sites might not affect ISG inductions, while in MEFs specific enhancers interact with target ISGs.

The picture of the distribution over the five defined groups for ESC-specific ISGF3 binding sites looked similar. In ESCs, many active enhancers were bound and most of these sites were lost in MEFs (**Figure 42**). In addition, these ESC-specific binding sites were repressed in MEFs. ESC-specific regulatory elements were silenced during differentiation into fibroblasts, this would explain that these sites had a "Repressed" chromatin signature in MEFs. In ESCs, the highest number of ISGF3 binding events happened in the "Active Enhancers" cluster before the IFNβ stimulation. "Active Enhancers" are highly accessible and harbor binding sites for many TFs including

STAT and IRF motifs (Spitz and Furlong 2012). This could facilitate recognition and binding of these sites by active complexes. As an “Active Enhancers” were already marked with H3K27ac, the binding of additional activators would have minor influence on its activation state. One could think of this as a way to dilute the TFs concentration by offering them additional binding sites. Consequently, the “right” binding sites were less frequently bound and, in the end, this could contribute to the attenuated ISG induction levels in ESCs. For both cell types, ESCs and MEFs, regulatory elements like enhancers were identified as an essential part of the innate immune response. My analyses support a model in which enhancers both act as storage points of cell type specific epigenetic memory in ESCs and function as amplifiers of the response to viral infections in the ISG gene network in MEFs. Enhancer sites with high STAT binding potential could also act as ‘sponge sites’ to, for example, attenuate the antiviral response under conditions in ESCs where levels of active STAT complexes become limiting.

4.3.4. H3K4me1 and accessible chromatin facilitate ISGF3 binding

In our analysis, we identified H3K4me1, chromatin accessibility (ATAC) and H3K27ac as promotive for ISGF3 binding, while H3K27me3 was inhibiting. In MEFs, H3K27ac showed the highest correlation with ISGF3 binding. However, it was shown that acetylation marks are much more dynamic than methylation marks. Active genes are often enriched for both, HATs and HDACs, which results in a high turnover of acetylation marks (Wang *et al.* 2009). In the plant alfalfa, the half-life of acetylation marks on different histone tails was found to be around 30 min (Waterborg and Kapros 2002). In contrast, the stability and half-life of methylation marks like H3K4me1 and H3K36me3 is much longer. Accordingly, these marks are being discussed as modulators of epigenetic memory (Kamada *et al.* 2018; Ostuni *et al.* 2013). Chromatin accessibility can change quickly but requires the recruitment of chromatin remodeler complexes and involves ATP-dependent processes (Clapier *et al.* 2017). The longer life time of histone methylation and chromatin accessibility makes these marks better candidates for a factor that regulates locus specific TF binding. Consequently, we focused on H3K4me1 and chromatin accessibility signal as marks permissive for ISGF3 binding.

To test the role of H3K4me1 in ISGF3 binding patterns, I hypothesized that the perturbation of this methylation mark would impact the ISG response. A recent study used ESCs with a catalytic dead (dCD) or knock out (dKO) of the KMTs MLL3 and

MLL4 to characterize their role in enhancer maintenance (Dorighi *et al.* 2017). MLL3/4 are identified to place mono-methylations onto histone 3 and are strongest enriched at non-promoter sites. This publication pointed to a coactivator role of MLL3/4 at enhancers independent of their methyltransferase function. Here, the exact same cells were treated with IFNβ to investigate a role H3K4me1 placed by MLL3/4 in the antiviral response. We tested, if a depletion of the H3K4me1 marks was associated with weakened ISG response and found no substantial changes in the number and amplitude of induced ISGs. This was confirmed in absolute numbers, as the fewest numbers of ISGs were identified in the WT control. Also, the dynamics of previously identified ISGs like Irf7, Rtp4 or Usp18 were nearly identical between WT and mutants. We concluded that there is no impact of mutants of MLL3/4 onto the induction of ISGs via IFNβ stimulation in ESCs. On explanation for this result is the redundancy of KMTs. In addition to MLL3/4, other known KMTs for H3K4 are MLL1/2, SETD1A/B, SETD7 and PRDM9 (Husmann and Gozani 2019). Another possibility is that the role of enhancers is more pronounced in MEFs than in ESCs, for which MLL3/4 knockout models were not available. Overall, more ISGF3 sites were marked in MEFs with this histone mark and a potential removal at primed enhancers could impact the ISG induction levels. To further investigate the role of KMTs, a systematic screen should be done to identify which enzyme is responsible for placing and maintaining the H3K4me1 mark at ISGF3 binding sites. As a next step, CRISPR-Cas9 technologies could be used to delete or create enzymatic dead versions of this KMT and then investigate the antiviral response dynamics in this cellular setup. The fusion of CRISPR-Cas9 with the KDM LSD1 is a way to remove specifically H3K4me1 from enhancers and cause repression of these sites (Kearns *et al.* 2015).

Instead of perturbing the histone methylation, the chromatin accessibility could be altered to test the role of this feature on ISGF3 binding. Upon IFN stimulation the chromatin remodeler BAF is required to allow a complete activation of ISGs (Liu *et al.* 2002; Qiao *et al.* 2013). As these multi-enzyme complexes are essential for gene regulation and cell survival, perturbations like knock out models of key factors of these complexes are very difficult to establish and have normally a lot of side effects (Langst and Manelyte 2015). One possibility to overcome this, would be to force a chromatin region of interest into a less accessibly chromatin state. This can be done by applying fusions of dead Cas9 (dCas9) with domains that locally induce heterochromatic marks, such as the KRAB domain (Yeo *et al.* 2018). The fusion construct can be specifically recruited to nearly every genomic locus, which is a specific feature of the CRISPR-dCas9 system. A similar approach was recently done

for the gene Plod2, which became resilient to TGFβ1 stimulation (Gjaltema *et al.* 2020). In our system, a specific set of non-promoters bound site, validated by scATAC co-accessibility analysis could be targeted with CRISPR-Cas9 and heterochromatic regions be established. The resulting reduction of chromatin accessibility could be sufficient to reduce the induction pattern of linked ISGs.

4.4. Conclusion

In this thesis, the cell type specific differences in the IFNβ response of ESCs and differentiated MEF and NPCs were dissected by integrating different genome-wide sequencing methods. Further specific features of the previously described attenuated antiviral response in ESCs were identified in comparison to MEFs (Wang *et al.* 2014). These provide insight into which regulatory steps can dampen the innate immune response. For example, phosphorylation of STAT1 upon stimulation was reduced in ESCs. Accordingly, it will be informative to test, if a stimulation of this activity, e.g. by treatment of ESCs with phosphatase inhibitors, would induce a stronger IFNβ response.

The data provided by this thesis provide a number of novel links between chromatin features and ISGF3 binding that need to be dissected in further studies. One important aspect is it to assess whether the correlations identified here represent causal links by perturbing chromatin features around ISREs to promote or inhibit ISFG3 binding. One approach could be the artificial recruitment of histone modifiers to place repressive chromatin marks like H3K9m3 or H3K27me3 that are predicted to prevent ISGF3 binding and consequently silence the target ISGs. Another approach would be to deplete the H3K4me1 mark, which, in combination high chromatin accessibility, is associated with a chromatin state permissive for ISGF3 binding. Such an experiment would represent also an excellent validation for the detected co-accessibility links between ISRF3 enhancers and target genes and further highlight the role of regulatory elements in the innate immunity. The validation approach tested in the thesis to partially deplete the H3K4me1 mark by knockout of the MLL3/4 KMTs was inconclusive. A caveat of these experiments is the high abundance of different KMTs in addition to MLL3/4 that could substitute for the H3K4 methylation activity of the knocked-out enzymes. Thus, it will be important to use more efficient approaches to decrease H3K4me1 levels, for example by inhibiting all KMTs that set H3K4me1. In addition, other chromatin features identified here could be perturbed to modulate the innate immunity. An increasing number of so called “epigenetic drugs” is available that inhibit enzymes that set or remove histone acetylation and methylation marks (Brien *et al.* 2016; Hauser *et al.* 2018). Of particular interest is the effect HDAC inhibitors that result in the hyper-acetylation of the genome and render chromatin more accessible (Gorisch *et al.* 2005; Toth *et al.* 2004). These drugs elicit a complex cellular response that could involve making additional ISRE sequences accessible to ISGF3 binding. Under such conditions, the IFNβ-triggered gene induction could be

enhanced. On the other hand, the activation of STAT TFs has been linked to cancer cell proliferation, survival and invasiveness (Yu *et al.* 2009). Thus, it would be important to investigate how a treatment with HDAC inhibitors affects IFN signaling, STAT binding cancer related phenotypes. Finally, it was shown that a sustained type I interferon signaling can cause resistance to immune therapy in melanoma and non-small cell lung cancer (Jacquelot *et al.* 2019). A combinatorial treatment with epigenetic drugs to target the chromatin states and to inhibit STAT binding could thus provide novel treatment options to overcome the observed resistances. Thus, the insights on the interplay of epigenetic signaling, STAT binding and ISG induction obtained in this thesis could help to identify key epigenetic regulators for modulate the innate immune response in cancer cells. In addition, the comprehensive multiomics data set acquired here provides a valuable resource for future projects to validate or supplement ISGs or STAT binding profiles.

5. References

Agrawal P, Heimbruch KE, Rao S. 2010. *Genome-Wide Maps of Transcription Regulatory Elements and Transcription Enhancers in Development and Disease*. Wiley.

Anders S, Pyl PT, Huber W; 2015; *HTSeq--a Python framework to work with high-throughput sequencing data*; Bioinformatics 31: 166-169.

Aranda S, Mas G, Di Croce L; 2015; *Regulation of gene transcription by Polycomb proteins*; Sci Adv 1: e1500737.

Aravind L, Landsman D; 1998; *AT-hook motifs identified in a wide variety of DNA-binding proteins*; Nucleic Acids Res 26: 4413-4421.

Au-Yeung N, Horvath CM; 2018; *Histone H2A.Z Suppression of Interferon-Stimulated Transcription and Antiviral Immunity Is Modulated by GCN5 and BRD2*; ISCIENCE 6: 68-82.

Au-Yeung N, Mandhana R, Horvath CM; 2013; *Transcriptional regulation by STAT1 and STAT2 in the interferon JAK-STAT pathway*; JAKSTAT 2: e23931.

Bai F, Ho Lim C, Jia J, Santostefano K, Simmons C, Kasahara H, Wu W, Terada N, Jin S; 2015; *Directed Differentiation of Embryonic Stem Cells Into Cardiomyocytes by Bacterial Injection of Defined Transcription Factors*; Sci Rep 5: 15014.

Bancerek J, Poss ZC, Steinparzer I, Sedlyarov V, Pfaffenwimmer T, Mikulic I, Dolken L, Strobl B, Muller M, Taatjes DJ *et al.*; 2013; *CDK8 kinase phosphorylates transcription factor STAT1 to selectively regulate the interferon response*; Immunity 38: 250-262.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K; 2007; *High-resolution profiling of histone methylations in the human genome*; Cell 129: 823-837.

Belton JM, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J; 2012; *Hi-C: a comprehensive technique to capture the conformation of genomes*; Methods 58: 268-276.

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, 3rd, Gingeras TR *et al.*; 2005; *Genomic maps and comparative analysis of histone modifications in human and mouse*; Cell 120: 169-181.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K *et al.*; 2006; *A bivalent chromatin structure marks key developmental genes in embryonic stem cells*; Cell 125: 315-326.

Bibel M, Richter J, Lacroix E, Barde YA; 2007; *Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells*; Nat Protoc 2: 1034-1043.

Blaszczyk K, Nowicka H, Kostyrko K, Antonczyk A, Wesoly J, Bluyssen HA; 2016; *The unique role of STAT2 in constitutive and IFN-induced transcription and antiviral responses*; Cytokine Growth Factor Rev 29: 71-81.

Blaszczyk K, Olejnik A, Nowicka H, Ozgyin L, Chen YL, Chmielewski S, Kostyrko K, Wesoly J, Balint BL, Lee CK *et al.*; 2015; *STAT2/IRF9 directs a prolonged ISGF3-like transcriptional response and antiviral activity in the absence of STAT1*; Biochem J 466: 511-524.

Bliven S, Lafita A, Parker A, Capitani G, Duarte JM; 2018; *Automated evaluation of quaternary structures from protein crystals*; PLoS Comput Biol 14: e1006104.

Bolger AM, Lohse M, Usadel B; 2014; *Trimmomatic: a flexible trimmer for Illumina sequence data*; Bioinformatics 30: 2114-2120.

Bonilla FA, Oettgen HC; 2010; *Adaptive immunity*; J Allergy Clin Immunol 125: S33-40.

Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK *et al.*; 2006; *Polycomb complexes repress developmental regulators in murine embryonic stem cells*; Nature 441: 349-353.

Braun E, Hotter D, Koepke L, Zech F, Gross R, Sparrer KMJ, Muller JA, Pfaller CK, Heusinger E, Wombacher R *et al.*; 2019; *Guanylate-Binding Proteins 2 and 5 Exert Broad Antiviral Activity by Inhibiting Furin-Mediated Processing of Viral Envelope Proteins*; Cell Rep 27: 2092-2104 e2010.

Brien GL, Valerio DG, Armstrong SA; 2016; *Exploiting the Epigenome to Control Cancer-Promoting Gene-Expression Programs*; Cancer Cell 29: 464-476.

Brierley MM, Fish EN; 2005; *Stats: multifaceted regulators of transcription*; J Interferon Cytokine Res 25: 733-744.

Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. 2015a. *ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide*. John Wiley & Sons, Inc., Hoboken, NJ, USA.

Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, Chang HY, Greenleaf WJ; 2015b; *Single-cell chromatin accessibility reveals principles of regulatory variation*; Nature 523: 486-490.

Cheng J, Blum R, Bowman C, Hu D, Shilatifard A, Shen S, Dynlacht BD; 2014; *A role for H3K4 monomethylation in gene repression and partitioning of chromatin readers*; Mol Cell 53: 979-992.

Clapier CR, Iwasa J, Cairns BR, Peterson CL; 2017; *Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes*; Nat Rev Mol Cell Biol 18: 407-422.

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA *et al.*; 2010; *Histone H3K27ac separates active from poised enhancers and predicts developmental state*; Proc Natl Acad Sci U S A 107: 21931-21936.

Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, Steemers FJ, Trapnell C, Shendure J; 2015; *Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing*; Science 348: 910-914.

D'Angelo W, Acharya D, Wang R, Wang J, Gurung C, Chen B, Bai F, Guo YL; 2016; *Development of Antiviral Innate Immunity During In Vitro Differentiation of Mouse Embryonic Stem Cells*; Stem Cells Dev 25: 648-659.

Dancy BM, Cole PA; 2015; *Protein lysine acetylation by p300/CBP*; Chem Rev 115: 2419-2452.

Dang W, Xu L, Yin Y, Chen S, Wang W, Hakim MS, Chang KO, Peppelenbosch MP, Pan Q; 2018; *IRF-1, RIG-I and MDA5 display potent antiviral activities against norovirus coordinately induced by different types of interferons*; Antiviral Res 155: 48-59.

Daniel B, Nagy G, Nagy L; 2014; *The intriguing complexities of mammalian gene regulation: how to link enhancers to regulated genes. Are we there yet?*; FEBS Lett 588: 2379-2391.

Darnell JE, Jr., Kerr IM, Stark GR; 1994; *Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins*; Science 264: 1415-1421.

De Santa F, Narang V, Yap ZH, Tusi BK, Burgold T, Austenaa L, Bucci G, Caganova M, Notarbartolo S, Casola S *et al.*; 2009; *Jmjd3 contributes to the control of gene expression in LPS-activated macrophages*; EMBO J 28: 3341-3352.

de Veer MJ, Holko M, Frevel M, Walker E, Der S, Paranjape JM, Silverman RH, Williams BR; 2001; *Functional classification of interferon-stimulated genes identified using microarrays*; J Leukoc Biol 69: 912-920.

Dekker J, Rippe K, Dekker M, Kleckner N; 2002; *Capturing chromosome conformation*; Science 295: 1306-1311.

Der SD, Zhou AM, Williams BRG, Silverman RH; 1998; *Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays*; Proc Natl Acad Sci USA 95: 15623-15628.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR; 2013; *STAR: ultrafast universal RNA-seq aligner*; Bioinformatics 29: 15-21.

Dorighi KM, Swigut T, Henriques T, Bhanu NV, Scruggs BS, Nady N, Still CD, 2nd, Garcia BA, Adelman K, Wysocka J; 2017; *Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation*; Mol Cell 66: 568-576 e564.

Efroni S, Duttagupta R, Cheng J, Dehghani H, Hoeppner DJ, Dash C, Bazett-Jones DP, Le Grice S, McKay RD, Buetow KH *et al.*; 2008; *Global transcription in pluripotent embryonic stem cells*; Cell Stem Cell 2: 437-447.

Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M *et al.*; 2011; *Mapping and analysis of chromatin state dynamics in nine human cell types*; Nature 473: 43-49.

Ewels P, Magnusson M, Lundin S, Kaller M; 2016; *MultiQC: summarize analysis results for multiple tools and samples in a single report*; Bioinformatics 32: 3047-3048.

Filippakopoulos P, Picaud S, Mangos M, Keates T, Lambert JP, Barsyte-Lovejoy D, Felletar I, Volkmer R, Muller S, Pawson T *et al.*; 2012; *Histone recognition and large-scale structural analysis of the human bromodomain family*; Cell 149: 214-231.

Fong AP, Tapscott SJ; 2013; *Skeletal muscle programming and re-programming*; Curr Opin Genet Dev 23: 568-573.

Fu XY, Schindler C, Improta T, Aebersold R, Darnell JE, Jr.; 1992; *The proteins of ISGF-3, the interferon alpha-induced transcriptional activator, define a gene family involved in signal transduction*; Proc Natl Acad Sci U S A 89: 7840-7843.

Fulton DL, Sundararajan S, Badis G, Hughes TR, Wasserman WW, Roach JC, Sladek R; 2009; *TFCat: the curated catalog of mouse and human transcription factors*; Genome Biol 10: R29.

Furlong EEM, Levine M; 2018; *Developmental enhancers and chromosome topology*; Science 361: 1341-1345.

Gates LA, Shi J, Rohira AD, Feng Q, Zhu B, Bedford MT, Sagum CA, Jung SY, Qin J, Tsai MJ *et al.*; 2017; *Acetylation on histone H3 lysine 9 mediates a switch from transcription initiation to elongation*; J Biol Chem 292: 14456-14472.

Ghisletti S, Barozzi I, Mietton F, Polletti S, De Santa F, Venturini E, Gregory L, Lonie L, Chew A, Wei CL *et al.*; 2010; *Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages*; Immunity 32: 317-328.

Giresi PG, Kim J, McDaniell RM, Iyer VR, Lieb JD; 2007; *FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin*; Genome Res 17: 877-885.

Gjaltema RAF, Goubert D, Huisman C, Pilar Garcia Tobilla CD, Koncz M, Jellema PG, Wu D, Brouwer U, Kiss A, Verschure PJ *et al.*; 2020; *KRAB-Induced Heterochromatin Effectively Silences PLOD2 Gene Expression in Somatic Cells and is Resilient to TGFbeta1 Activation*; Int J Mol Sci 21.

Gorisch SM, Wachsmuth M, Toth KF, Lichter P, Rippe K; 2005; *Histone acetylation increases chromatin accessibility*; J Cell Sci 118: 5825-5834.

Gu B, Lee MG; 2013; *Histone H3 lysine 4 methyltransferases and demethylases in self-renewal and differentiation of stem cells*; Cell Biosci 3: 39.

Guo YL; 2017; *Utilization of different anti-viral mechanisms by mammalian embryonic stem cells and differentiated cells*; Immunol Cell Biol 95: 17-23.

-; 2019; *The underdeveloped innate immunity in embryonic stem cells: The molecular basis and biological perspectives from early embryogenesis*; Am J Reprod Immunol 81: e13089.

Guo YL, Carmichael GG, Wang R, Hong X, Acharya D, Huang F, Bai F; 2015; *Attenuated Innate Immunity in Embryonic Stem Cells and Its Implications in Developmental Biology and Regenerative Medicine*; Stem Cells 33: 3165-3173.

Gupta M, Shin DM, Ramakrishna L, Goussetis DJ, Platanias LC, Xiong H, Morse HC, 3rd, Ozato K; 2015; *IRF8 directs stress-induced autophagy in macrophages and promotes clearance of Listeria monocytogenes*; Nat Commun 6: 6379.

Gytz H, Hansen MF, Skovbjerg S, Kristensen AC, Horlyck S, Jensen MB, Fredborg M, Markert LD, McMillan NA, Christensen EI *et al.*; 2017; *Apoptotic properties of the type 1 interferon induced family of human mitochondrial membrane ISG12 proteins*; Biol Cell 109: 94-112.
Hauser AT, Robaa D, Jung M; 2018; *Epigenetic small molecule modulators of histone and DNA methylation*; Curr Opin Chem Biol 45: 73-85.
He X, Ashbrook AW, Du Y, Wu J, Hoffmann HH, Zhang C, Xia L, Peng YC, Tumas KC, Singh BK *et al.*; 2020; *RTP4 inhibits IFN-I response and enhances experimental cerebral malaria and neuropathology*; Proc Natl Acad Sci U S A 117: 19465-19474.
Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA *et al.*; 2007; *Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome*; Nat Genet 39: 311-318.
Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK; 2010; *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities*; Mol Cell 38: 576-589.
Hezroni H, Tzchori I, Davidi A, Mattout A, Biran A, Nissim-Rafinia M, Westphal H, Meshorer E; 2011; *H3K9 histone acetylation predicts pluripotency and reprogramming capacity of ES cells*; Nucleus 2: 300-309.
Higgs DR; 2020; *Enhancer-promoter interactions and transcription*; Nat Genet 52: 470-471.
Hong XX, Carmichael GG; 2013; *Innate immunity in pluripotent human cells: attenuated response to interferon-beta*; J Biol Chem 288: 16196-16205.
Huang da W, Sherman BT, Lempicki RA; 2009; *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources*; Nat Protoc 4: 44-57.
Husmann D, Gozani O; 2019; *Histone lysine methyltransferases in biology and disease*; Nat Struct Mol Biol 26: 880-889.
Ihle JN, Witthuhn BA, Quelle FW, Yamamoto K, Thierfelder WE, Kreider B, Silvennoinen O; 1994; *Signaling by the cytokine receptor superfamily: JAKs and STATs*; Trends Biochem Sci 19: 222-227.
Iwafuchi-Doi M, Zaret KS; 2014; *Pioneer transcription factors in cell reprogramming*; Genes Dev 28: 2679-2692.
Jacquelot N, Yamazaki T, Roberti MP, Duong CPM, Andrews MC, Verlingue L, Ferrere G, Becharef S, Vetizou M, Daillere R *et al.*; 2019; *Sustained Type I interferon signaling as a mechanism of resistance to PD-1 blockade*; Cell Res 29: 846-861.
Jenuwein T, Allis CD; 2001; *Translating the histone code*; Science 293: 1074-1080.
Jeong KW, Kim K, Situ AJ, Ulmer TS, An W, Stallcup MR; 2011; *Recognition of enhancer element-specific histone methylation by TIP60 in transcriptional activation*; Nat Struct Mol Biol 18: 1358-1365.
Johnson PF, McKnight SL; 1989; *Eukaryotic transcriptional regulatory proteins*; Annu Rev Biochem 58: 799-839.
Juven-Gershon T, Kadonaga JT; 2010; *Regulation of gene expression via the core promoter and the basal transcriptional machinery*; Dev Biol 339: 225-229.
Kamada R, Yang W, Zhang Y, Patel MC, Yang Y, Ouda R, Dey A, Wakabayashi Y, Sakaguchi K, Fujita T *et al.*; 2018; *Interferon stimulation creates chromatin marks and establishes transcriptional memory*; Proc Natl Acad Sci U S A 115: E9162-E9171.
Karmodiya K, Krebs AR, Oulad-Abdelghani M, Kimura H, Tora L; 2012; *H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells*; BMC Genomics 13: 424.
Kearns NA, Pham H, Tabak B, Genga RM, Silverstein NJ, Garber M, Maehr R; 2015; *Functional annotation of native enhancers with a Cas9-histone demethylase fusion*; Nat Methods 12: 401-403.
Kebede AF, Schneider R, Daujat S; 2015; *Novel types and sites of histone modifications emerge as players in the transcriptional regulation contest*; FEBS J 282: 1658-1674.
Kieffer-Kwon KR, Tang Z, Mathe E, Qian J, Sung MH, Li G, Resch W, Baek S, Pruett N, Grontved L *et al.*; 2013; *Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation*; Cell 155: 1507-1520.
Kiu H, Nicholson SE; 2012; *Biology and significance of the JAK/STAT signalling pathways*; Growth Factors 30: 88-106.
Kopylova E, Noe L, Touzet H; 2012; *SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data*; Bioinformatics 28: 3211-3217.
Kotredes KP, Gamero AM; 2013; *Interferons as inducers of apoptosis in malignant cells*; J Interferon Cytokine Res 33: 162-170.
Kouzarides T; 2007; *Chromatin modifications and their function*; Cell 128: 693-705.
Krejci J, Uhlirova R, Galiova G, Kozubek S, Smigova J, Bartova E; 2009; *Genome-wide reduction in H3K9 acetylation during human embryonic stem cell differentiation*; J Cell Physiol 219: 677-687.
Kristiansen H, Gad HH, Eskildsen-Larsen S, Despres P, Hartmann R; 2011; *The oligoadenylate synthetase family: an ancient protein family with multiple antiviral activities*; J Interferon Cytokine Res 31: 41-47.
Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D; 2002; *Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein*; Genes Dev 16: 2893-2905.
Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, Chen X, Taipale J, Hughes TR, Weirauch MT; 2018; *The Human Transcription Factors*; Cell 172: 650-665.
Langmead B, Salzberg SL; 2012; *Fast gapped-read alignment with Bowtie 2*; Nat Methods 9: 357-359.
Langmead B, Trapnell C, Pop M, Salzberg SL; 2009; *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome*; Genome Biol 10: R25.
Langst G, Manelyte L; 2015; *Chromatin Remodelers: From Function to Dysfunction*; Genes (Basel) 6: 299-324.

Lau CM, Adams NM, Geary CD, Weizman OE, Rapp M, Pritykin Y, Leslie CS, Sun JC; 2018; *Epigenetic control of innate and adaptive immune memory*; Nat Immunol 19: 963-972.
Lauberth SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH, Roeder RG; 2013; *H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation*; Cell 152: 1021-1036.
Lee TI, Young RA; 2013; *Transcriptional regulation and its misregulation in disease*; Cell 152: 1237-1251.
Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH; 2003; *Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin*; Curr Biol 13: 1192-1200.
Lettice LA, Heaney SJ, Purdie LA, Li L, de Beer P, Oostra BA, Goode D, Elgar G, Hill RE, de Graaff E; 2003; *A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly*; Hum Mol Genet 12: 1725-1735.
Levy DE, Darnell JE, Jr.; 2002; *Stats: transcriptional control and biological impact*; Nat Rev Mol Cell Biol 3: 651-662.
Li B, Carey M, Workman JL; 2007; *The role of chromatin during transcription*; Cell 128: 707-719.
Li B, Dewey CN; 2011; *RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome*; BMC Bioinformatics 12: 323.
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S; 2009; *The Sequence Alignment/Map format and SAMtools*; Bioinformatics 25: 2078-2079.
Li L, Wunderlich Z; 2017; *An Enhancer's Length and Composition Are Shaped by Its Regulatory Task*; Front Genet 8: 63.
Liu H, Kang H, Liu R, Chen X, Zhao K; 2002; *Maximal induction of a subset of interferon target genes requires the chromatin-remodeling activity of the BAF complex*; Mol Cell Biol 22: 6471-6479.
Liu H, Luo M, Wen JK; 2014; *mRNA stability in the nucleus*; J Zhejiang Univ Sci B 15: 444-454.
Loutfy MR, Blatt LM, Siminovitch KA, Ward S, Wolff B, Lho H, Pham DH, Deif H, LaMere EA, Chang M *et al.*; 2003; *Interferon alfacon-1 plus corticosteroids in severe acute respiratory syndrome: a preliminary study*; JAMA 290: 3222-3228.
Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ; 1997; *Crystal structure of the nucleosome core particle at 2.8 A resolution*; Nature 389: 251-260.
Magaraki A, van der Heijden G, Sleddens-Linkels E, Magarakis L, van Cappellen WA, Peters A, Gribnau J, Baarends WM, Eijpe M; 2017; *Silencing markers are retained on pericentric heterochromatin during murine primordial germ cell development*; Epigenetics Chromatin 10: 11.
Maillard PV, Ciaudo C, Marchais A, Li Y, Jay F, Ding SW, Voinnet O; 2013; *Antiviral RNA interference in mammalian cells*; Science 342: 235-238.
Mallm JP, Iskar M, Ishaque N, Klett LC, Kugler SJ, Muino JM, Teif VB, Poos AM, Grossmann S, Erdel F *et al.*; 2019; *Linking aberrant chromatin features in chronic lymphocytic leukemia to transcription factor networks*; Mol Syst Biol 15: e8339.
Margueron R, Li G, Sarma K, Blais A, Zavadil J, Woodcock CL, Dynlacht BD, Reinberg D; 2008; *Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms*; Mol Cell 32: 503-518.
Marinov GK, Kundaje A, Park PJ, Wold BJ; 2014; *Large-scale quality analysis of published ChIP-seq data*; G3 (Bethesda) 4: 209-223.
Marshall JS, Warrington R, Watson W, Kim HL; 2018; *An introduction to immunology and immunopathology*; Allergy Asthma Clin Immunol 14: 49.
Martello G, Smith A; 2014; *The nature of embryonic stem cells*; Annu Rev Cell Dev Biol 30: 647-675.
Maul GG, Deaven L; 1977; *Quantitative determination of nuclear pore complexes in cycling cells with differing DNA content*; J Cell Biol 73: 748-760.
Maze I, Wenderski W, Noh KM, Bagot RC, Tzavaras N, Purushothaman I, Elsasser SJ, Guo Y, Ionete C, Hurd YL *et al.*; 2015; *Critical Role of Histone Turnover in Neuronal Transcription and Plasticity*; Neuron 87: 77-94.
McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G; 2010; *GREAT improves functional interpretation of cis-regulatory regions*; Nat Biotechnol 28: 495-501.
Mears HV, Sweeney TR; 2018; *Better together: the role of IFIT protein-protein interactions in the antiviral response*; J Gen Virol 99: 1463-1477.
Melen K, Kinnunen L, Julkunen I; 2001; *Arginine/lysine-rich structural element is involved in interferon-induced nuclear import of STATs*; J Biol Chem 276: 16447-16455.
Misteli T; 2008; *Physiological importance of RNA and protein mobility in the cell nucleus*; Histochem Cell Biol 129: 5-11.
Monaghan L, Massett ME, Bunschoten RP, Hoose A, Pirvan PA, Liskamp RMJ, Jorgensen HG, Huang X; 2019; *The Emerging Role of H3K9me3 as a Potential Therapeutic Target in Acute Myeloid Leukemia*; Front Oncol 9: 705.
Moorthy SD, Mitchell JA; 2016; *Generating CRISPR/Cas9 Mediated Monoallelic Deletions to Study Enhancer Function in Mouse Embryonic Stem Cells*; J Vis Exp: e53552.
Morey L, Helin K; 2010; *Polycomb group protein-mediated repression of transcription*; Trends Biochem Sci 35: 323-332.
Mostafavi S, Yoshida H, Moodley D, LeBoite H, Rothamel K, Raj T, Ye CJ, Chevrier N, Zhang SY, Feng T *et al.*; 2016; *Parsing the Interferon Transcriptional Network and Its Disease Associations*; Cell 164: 564-578.
Nie Y, Cheng X, Chen J, Sun X; 2014; *Nucleosome organization in the vicinity of transcription factor binding sites in the human genome*; BMC Genomics 15: 493.

Nitulescu, II, Meyer SC, Wen QJ, Crispino JD, Lemieux ME, Levine RL, Pelish HE, Shair MD; 2017; *Mediator Kinase Phosphorylation of STAT1 S727 Promotes Growth of Neoplasms With JAK-STAT Activation*; EBioMedicine 26: 112-125.
Niu GJ, Xu JD, Yuan WJ, Sun JJ, Yang MC, He ZH, Zhao XF, Wang JX; 2018; *Protein Inhibitor of Activated STAT (PIAS) Negatively Regulates the JAK/STAT Pathway by Inhibiting STAT Phosphorylation and Translocation*; Front Immunol 9: 2392.
Nusinzon I, Horvath CM; 2003; *Interferon-stimulated transcription and innate antiviral immunity require deacetylase activity and histone deacetylase 1*; Proc Natl Acad Sci U S A 100: 14742-14747.
O'Neal JT, Upadhyay AA, Wolabaugh A, Patel NB, Bosinger SE, Suthar MS; 2019; *West Nile Virus-Inclusive Single-Cell RNA Sequencing Reveals Heterogeneity in the Type I Interferon Response within Single Cells*; J Virol 93.
Ostuni R, Piccolo V, Barozzi I, Polletti S, Termanini A, Bonifacio S, Curina A, Prosperini E, Ghisletti S, Natoli G; 2013; *Latent enhancers activated by stimulation in differentiated cells*; Cell 152: 157-171.
Park PJ; 2009; *ChIP-seq: advantages and challenges of a maturing technology*; Nat Rev Genet 10: 669-680.
Pervolaraki K, Rastgou Talemi S, Albrecht D, Bormann F, Bamford C, Mendoza JL, Garcia KC, McLauchlan J, Hofer T, Stanifer ML *et al.*; 2018; *Differential induction of interferon stimulated genes between type I and type III interferons is independent of interferon receptor abundance*; PLoS Pathog 14: e1007420.
Qiao Y, Giannopoulou EG, Chan CH, Park SH, Gong S, Chen J, Hu X, Elemento O, Ivashkiv LB; 2013; *Synergistic activation of inflammatory cytokine genes by interferon-gamma-induced chromatin remodeling and toll-like receptor signaling*; Immunity 39: 454-469.
Quinlan AR, Hall IM; 2010; *BEDTools: a flexible suite of utilities for comparing genomic features*; Bioinformatics 26: 841-842.
Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J; 2011; *A unique chromatin signature uncovers early developmental enhancers in humans*; Nature 470: 279-283.
Rapp M, Lau CM, Adams NM, Weizman OE, O'Sullivan TE, Geary CD, Sun JC; 2017; *Core-binding factor beta and Runx transcription factors promote adaptive natural killer cell responses*; Sci Immunol 2.
Reik W, Surani MA; 2015; *Germline and Pluripotent Stem Cells*; Cold Spring Harb Perspect Biol 7.
Rengachari S, Groiss S, Devos JM, Caron E, Grandvaux N, Panne D; 2018; *Structural basis of STAT2 recognition by IRF9 reveals molecular insights into ISGF3 function*; Proc Natl Acad Sci U S A 115: E601-E609.
Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J *et al.*; 2015; *Integrative analysis of 111 reference human epigenomes*; Nature 518: 317-330.
Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, Euskirchen G, Bernier B, Varhol R, Delaney A *et al.*; 2007; *Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing*; Nat Methods 4: 651-657.
Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR *et al.*; 2012; *Differential oestrogen receptor binding is associated with clinical outcome in breast cancer*; Nature 481: 389-393.
Sabari BR, Zhang D, Allis CD, Zhao Y; 2017; *Metabolic regulation of gene expression through histone acylations*; Nat Rev Mol Cell Biol 18: 90-101.
Saksouk N, Simboeck E, Dejardin J; 2015; *Constitutive heterochromatin formation and transcription in mammals*; Epigenetics Chromatin 8: 3.
Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T; 2002; *Active genes are tri-methylated at K4 of histone H3*; Nature 419: 407-411.
Sanyal A, Lajoie BR, Jain G, Dekker J; 2012; *The long-range interaction landscape of gene promoters*; Nature 489: 109-113.
Schindler C, Shuai K, Prezioso VR, Darnell JE, Jr.; 1992; *Interferon-dependent tyrosine phosphorylation of a latent cytoplasmic transcription factor*; Science 257: 809-813.
Schoggins JW, Rice CM; 2011; *Interferon-stimulated genes and their antiviral effector functions*; Curr Opin Virol 1: 519-525.
Schoggins JW, Wilson SJ, Panis M, Murphy MY, Jones CT, Bieniasz P, Rice CM; 2011; *A diverse range of gene products are effectors of the type I interferon antiviral response*; Nature 472: 481-485.
Schwerk J, Koster M, Hauser H, Rohde M, Fulde M, Hornef MW, May T; 2013; *Generation of mouse small intestinal epithelial cell lines that allow the analysis of specific innate immune functions*; PLoS ONE 8: e72700.
Seegert D, Strehlow I, Klose B, Levy DE, Schindler C, Decker T; 1994; *A novel interferon-alpha-regulated, DNA-binding protein participates in the regulation of the IFP53/tryptophanyl-tRNA synthetase gene*; J Biol Chem 269: 8590-8595.
Sen GC; 2001; *Viruses and interferons*; Annu Rev Microbiol 55: 255-281.
Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, Wagner U, Dixon J, Lee L, Lobanenkov VV *et al.*; 2012; *A map of the cis-regulatory sequences in the mouse genome*; Nature 488: 116-120.
Shuai K; 2006; *Regulation of cytokine signaling pathways by PIAS proteins*; Cell Res 16: 196-202.
Silvennoinen O, Ihle JN, Schlessinger J, Levy DE; 1993; *Interferon-induced nuclear signalling by Jak protein tyrosine kinases*; Nature 366: 583-585.
Singh H, Khan AA, Dinner AR; 2014; *Gene regulatory networks in the immune system*; Trends Immunol 35: 211-218.
Smale ST, Kadonaga JT; 2003; *The RNA polymerase II core promoter*; Annu Rev Biochem 72: 449-479.

Smale ST, Tarakhovsky A, Natoli G; 2014; *Chromatin contributions to the regulation of innate immunity*; Annu Rev Immunol 32: 489-511.
Soutourina J; 2018; *Transcription regulation by the Mediator complex*; Nat Rev Mol Cell Biol 19: 262-274.
Spitz F, Furlong EE; 2012; *Transcription factors: from enhancer binding to developmental control*; Nat Rev Genet 13: 613-626.
Stanifer ML, Pervolaraki K, Boulant S; 2019; *Differential Regulation of Type I and Type III Interferon Signaling*; Int J Mol Sci 20.
Stark GR, Darnell JE, Jr.; 2012; *The JAK-STAT pathway at twenty*; Immunity 36: 503-514.
Supek F, Bosnjak M, Skunca N, Smuc T; 2011; *REVIGO summarizes and visualizes long lists of gene ontology terms*; PLoS ONE 6: e21800.
Suter DM; 2020; *Transcription Factors and DNA Play Hide and Seek*; Trends Cell Biol 30: 491-500.
Takahashi K, Yamanaka S; 2006; *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors*; Cell 126: 663-676.
Tamura T, Smith M, Kanno T, Dasenbrock H, Nishiyama A, Ozato K; 2009; *Inducible deposition of the histone variant H3.3 in interferon-stimulated genes*; J Biol Chem 284: 12217-12225.
Tee WW, Reinberg D; 2014; *Chromatin features and the epigenetic regulation of pluripotency states in ESCs*; Development 141: 2376-2390.
Teif VB, Erdel F, Beshnova DA, Vainshtein Y, Mallm JP, Rippe K; 2013; *Taking into account nucleosomes for predicting gene expression*; Methods 62: 26-38.
Teif VB, Vainshtein Y, Caudron-Herger M, Mallm JP, Marth C, Hofer T, Rippe K; 2012; *Genome-wide nucleosome positioning during embryonic stem cell development*; Nat Struct Mol Biol 19: 1185-1192.
Thomas MC, Chiang CM; 2006; *The general transcription machinery and general cofactors*; Crit Rev Biochem Mol Biol 41: 105-178.
Thorvaldsdottir H, Robinson JT, Mesirov JP; 2013; *Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration*; Brief Bioinform 14: 178-192.
Toth KF, Knoch TA, Wachsmuth M, Frank-Stohr M, Stohr M, Bacher CP, Muller G, Rippe K; 2004; *Trichostatin A-induced histone acetylation causes decondensation of interphase chromatin*; J Cell Sci 117: 4277-4287.
Turvey SE, Broide DH; 2010; *Innate immunity*; J Allergy Clin Immunol 125: S24-32.
Urban N, Guillemot F; 2014; *Neurogenesis in the embryonic and adult brain: same regulators, different roles*; Front Cell Neurosci 8: 396.
Vahedi G, Takahashi H, Nakayamada S, Sun HW, Sartorelli V, Kanno Y, O'Shea JJ; 2012; *STATs shape the active enhancer landscape of T cell populations*; Cell 151: 981-993.
Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM; 2009; *A census of human transcription factors: function, expression and evolution*; Nat Rev Genet 10: 252-263.
Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG *et al.*; 2010; *Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers*; Cell 142: 967-980.
Villarino AV, Kanno Y, O'Shea JJ; 2017; *Mechanisms and consequences of Jak-STAT signaling in the immune system*; Nat Immunol 18: 374-384.
Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F *et al.*; 2009; *ChIP-seq accurately predicts tissue-specific activity of enhancers*; Nature 457: 854-858.
Voigt P, Tee WW, Reinberg D; 2013; *A double take on bivalent promoters*; Genes Dev 27: 1318-1338.
Volkel P, Angrand PO; 2007; *The control of histone lysine methylation in epigenetic regulation*; Biochimie 89: 1-20.
Wang L, Wang S, Li W; 2012; *RSeQC: quality control of RNA-seq experiments*; Bioinformatics 28: 2184-2185.
Wang Q, Sun Q, Czajkowsky DM, Shao Z; 2018; *Sub-kb Hi-C in D. melanogaster reveals conserved characteristics of TADs between insect and mammalian cells*; Nat Commun 9: 188.
Wang R, Wang J, Acharya D, Paul AM, Bai F, Huang F, Guo YL; 2014; *Antiviral responses in mouse embryonic stem cells: differential development of cellular mechanisms in type I interferon production and response*; J Biol Chem 289: 25186-25198.
Wang R, Wang J, Paul AM, Acharya D, Bai F, Huang F, Guo YL; 2013; *Mouse embryonic stem cells are deficient in type I interferon expression in response to viral infections and double-stranded RNA*; J Biol Chem 288: 15926-15936.
Wang Z, Zang C, Cui K, Schones DE, Barski A, Peng W, Zhao K; 2009; *Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes*; Cell 138: 1019-1031.
Waterborg JH, Kapros T; 2002; *Kinetic analysis of histone acetylation turnover and Trichostatin A induced hyper- and hypoacetylation in alfalfa*; Biochem Cell Biol 80: 279-293.
Weissman IL; 2000; *Stem cells: units of development, units of regeneration, and units in evolution*; Cell 100: 157-168.
Whyatt LM, Duwel A, Smith AG, Rathjen PD; 1993; *The responsiveness of embryonic stem cells to alpha and beta interferons provides the basis of an inducible expression system for analysis of developmental control genes*; Mol Cell Biol 13: 7971-7976.
Wilhelm BT, Landry JR; 2009; *RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing*; Methods 48: 249-257.
Xu S, Grullon S, Ge K, Peng W; 2014; *Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells*; Methods Mol Biol 1150: 97-111.
Yeo NC, Chavez A, Lance-Byrne A, Chan Y, Menn D, Milanova D, Kuo CC, Guo X, Sharma S, Tung A *et al.*; 2018; *An enhanced CRISPR repressor for targeted mammalian gene regulation*; Nat Methods 15: 611-616.

Yu H, Pardoll D, Jove R; 2009; *STATs in cancer inflammation and immunity: a leading role for STAT3*; Nat Rev Cancer 9: 798-809.
Yu M, Ren B; 2017; *The Three-Dimensional Organization of Mammalian Genomes*; Annu Rev Cell Dev Biol 33: 265-289.
Yun M, Wu J, Workman JL, Li B; 2011; *Readers of histone modifications*; Cell Res 21: 564-578.
Zaret KS, Carroll JS; 2011; *Pioneer transcription factors: establishing competence for gene expression*; Genes Dev 25: 2227-2241.
Zaret KS, Mango SE; 2016; *Pioneer transcription factors, chromatin dynamics, and cell fate control*; Curr Opin Genet Dev 37: 76-81.
Zerbino DR, Wilder SP, Johnson N, Juettemann T, Flicek PR; 2015; *The ensembl regulatory build*; Genome Biol 16: 56.
Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W *et al.*; 2008; *Model-based analysis of ChIP-Seq (MACS)*; Genome Biol 9: R137.

6. Appendix

Approach	HiSeq -ID	Sample Name	Barcode Type	Index - i7 Name	Index - i7 Sequence	Index - i5 Name	Index - i5 Sequence
ATAC	2622	ESC_Ola_IFNb_0h_Rep1	3	i701	TAAGGCGA	Universal	-
ATAC	2623	ESC_Ola_IFNb_0h_Rep2	3	i702	CGTACTAG	Universal	-
ATAC	2624	ESC_Ola_IFNb_1h_Rep1	3	i703	AGGCAGAA	Universal	-
ATAC	2625	ESC_Ola_IFNb_1h_Rep2	3	i704	TCCTGAGC	Universal	-
ATAC	2626	ESC_Ola_IFNb_6h_Rep1	3	i705	GGACTCCT	Universal	-
ATAC	2627	ESC_Ola_IFNb_6h_Rep2	3	i706	TAGGCATG	Universal	-
ATAC	3513	MEF_Ola_IFNb_0h_Rep1	3	i701	TAAGGCGA	i501	TAGATCGC
ATAC	3514	MEF_Ola_IFNb_1h_Rep1	3	i702	CGTACTAG	i502	CTCTCTAT
ATAC	3515	MEF_Ola_IFNb_6h_Rep1	3	i703	AGGCAGAA	i503	TATCCTCT
ATAC	3516	MEF_Ola_IFNb_0h_Rep2	3	i704	TCCTGAGC	i504	AGAGTAGA
ATAC	3517	MEF_Ola_IFNb_1h_Rep2	3	i705	GGACTCCT	i505	GTAAGGAG
ATAC	3518	MEF_Ola_IFNb_6h_Rep2	3	i706	TAGGCATG	i506	ACTGCATA
ATAC	3532	NPC_Ola_IFNb_0h_Rep1	3	i701	TAAGGCGA	i501	TAGATCGC
ATAC	3533	NPC_Ola_IFNb_1h_Rep1	3	i702	CGTACTAG	i502	CTCTCTAT
ATAC	3534	NPC_Ola_IFNb_6h_Rep1	3	i703	AGGCAGAA	i503	TATCCTCT
ATAC	3535	NPC_Ola_IFNb_0h_Rep2	3	i704	TCCTGAGC	i504	AGAGTAGA
ATAC	3536	NPC_Ola_IFNb_1h_Rep2	3	i705	GGACTCCT	i505	GTAAGGAG
ATAC	3537	NPC_Ola_IFNb_6h_Rep2	3	i706	TAGGCATG	i506	ACTGCATA
ATAC	3538	NPC_Ola_IFNb_0h_Rep3	3	i707	CTCTCTAC	i507	AAGGAGTA
ATAC	3539	NPC_Ola_IFNb_1h_Rep3	3	i708	CAGAGAGG	i508	CTAAGCCT
ATAC	3540	NPC_Ola_IFNb_6h_Rep3	3	i709	GCTACGCT	i509	TGGAAATC
ATAC	3541	NPC_Ola_IFNb_0h_Rep4	3	i710	CGAGGCTG	i510	AACATGAT
ATAC	3542	NPC_Ola_IFNb_1h_Rep4	3	i711	AAGAGGCA	i511	TGATGAAA
ATAC	3543	NPC_Ola_IFNb_6h_Rep4	3	i712	GTAGAGGA	i512	GTCGGACT
ChIP Histone	1396	ESC_Ola_H3K4me1_IFNb_0h_Rep1	1	R701	ATCACG	Universal	-
ChIP Histone	1397	ESC_Ola_H3K4me3_IFNb_0h_Rep1	1	R702	CGATGT	Universal	-
ChIP Histone	1399	ESC_Ola_H3K9ac_IFNb_0h_Rep1	1	R703	TTAGGC	Universal	-
ChIP Histone	1400	ESC_Ola_H3K9me3_IFNb_0h_Rep1	1	R706	GCCAAT	Universal	-
ChIP Histone	1401	ESC_Ola_H3K36me3_IFNb_0h_Rep1	1	R707	CAGATC	Universal	-
ChIP Histone	1402	ESC_Ola_H3K27me3_IFNb_0h_Rep1	1	R708	ACTTGA	Universal	-
ChIP Histone	1403	ESC_Ola_H3K27ac_IFNb_0h_Rep1	1	R709	GATCAG	Universal	-
ChIP Histone	1404	ESC_Ola_H3_IFNb_0h_Rep1	1	R710	TAGCTT	Universal	-
ChIP Histone	1405	ESC_Ola_IgG_Rb_IFNb_0h_Rep1	1	R711	GGCTAC	Universal	-
ChIP Histone	1406	ESC_Ola_Input_IFNb_0h_Rep1	1	R712	CTTGTA	Universal	-
ChIP Histone	1407	ESC_Ola_H3K4me1_IFNb_1h_Rep1	1	R701	ATCACG	Universal	-
ChIP Histone	1408	ESC_Ola_H3K4me3_IFNb_1h_Rep1	1	R702	CGATGT	Universal	-
ChIP Histone	1409	ESC_Ola_H3K9ac_IFNb_1h_Rep1	1	R703	TTAGGC	Universal	-
ChIP Histone	1410	ESC_Ola_H3K9me3_IFNb_1h_Rep1	1	R706	GCCAAT	Universal	-
ChIP Histone	1411	ESC_Ola_H3K36me3_IFNb_1h_Rep1	1	R707	CAGATC	Universal	-
ChIP Histone	1412	ESC_Ola_H3K27me3_IFNb_1h_Rep1	1	R708	ACTTGA	Universal	-
ChIP Histone	1413	ESC_Ola_H3K27ac_IFNb_1h_Rep1	1	R709	GATCAG	Universal	-
ChIP Histone	1414	ESC_Ola_H3_IFNb_1h_Rep1	1	R710	TAGCTT	Universal	-
ChIP Histone	1415	ESC_Ola_IgG_Rb_IFNb_1h_Rep1	1	R711	GGCTAC	Universal	-
ChIP Histone	1416	ESC_Ola_Input_IFNb_1h_Rep1	1	R712	CTTGTA	Universal	-
ChIP Histone	1417	ESC_Ola_H3K4me1_IFNb_6h_Rep1	1	R701	ATCACG	Universal	-
ChIP Histone	1418	ESC_Ola_H3K4me3_IFNb_6h_Rep1	1	R702	CGATGT	Universal	-
ChIP Histone	1419	ESC_Ola_H3K9ac_IFNb_6h_Rep1	1	R703	TTAGGC	Universal	-
ChIP Histone	1420	ESC_Ola_H3K9me3_IFNb_6h_Rep1	1	R706	GCCAAT	Universal	-
ChIP Histone	1421	ESC_Ola_H3K36me3_IFNb_6h_Rep1	1	R707	CAGATC	Universal	-
ChIP Histone	1422	ESC_Ola_H3K27me3_IFNb_6h_Rep1	1	R708	ACTTGA	Universal	-
ChIP Histone	1423	ESC_Ola_H3K27ac_IFNb_6h_Rep1	1	R709	GATCAG	Universal	-
ChIP Histone	1424	ESC_Ola_H3_IFNb_6h_Rep1	1	R710	TAGCTT	Universal	-
ChIP Histone	1425	ESC_Ola_IgG_Rb_IFNb_6h_Rep1	1	R711	GGCTAC	Universal	-
ChIP Histone	1426	ESC_Ola_Input_IFNb_6h_Rep1	1	R712	CTTGTA	Universal	-
ChIP Histone	2060	ESC_Ola_H3K4me1_IFNb_0h_Rep2	1	R702	CGATGT	Universal	-
ChIP Histone	2061	ESC_Ola_H3K4me3_IFNb_0h_Rep2	1	R703	TTAGGC	Universal	-
ChIP Histone	2062	ESC_Ola_H3K9ac_IFNb_0h_Rep2	1	R704	TGACCA	Universal	-
ChIP Histone	2063	ESC_Ola_H3K9me2_IFNb_0h_Rep2	1	R701	ATCACG	Universal	-
ChIP Histone	2064	ESC_Ola_H3K9me3_IFNb_0h_Rep2	1	R705	ACAGTG	Universal	-
ChIP Histone	2065	ESC_Ola_H3K27ac_IFNb_0h_Rep2	1	R707	CAGATC	Universal	-
ChIP Histone	2066	ESC_Ola_H3K27me3_IFNb_0h_Rep2	1	R706	GCCAAT	Universal	-
ChIP Histone	2067	ESC_Ola_H3K36me3_IFNb_0h_Rep2	1	R704	TGACCA	Universal	-
ChIP Histone	2068	ESC_Ola_H3_IFNb_0h_Rep2	1	R709	GATCAG	Universal	-
ChIP Histone	2069	ESC_Ola_IgG_Rb_IFNb_0h_Rep2	1	R710	TAGCTT	Universal	-
ChIP Histone	2070	ESC_Ola_IgG_M_IFNb_0h_Rep2	1	R711	GGCTAC	Universal	-
ChIP Histone	2071	ESC_Ola_Input_IFNb_0h_Rep2	1	R712	CTTGTA	Universal	-
ChIP Histone	2072	ESC_Ola_H3K4me1_IFNb_1h_Rep2	1	R701	ATCACG	Universal	-
ChIP Histone	2073	ESC_Ola_H3K4me3_IFNb_1h_Rep2	1	R702	CGATGT	Universal	-
ChIP Histone	2074	ESC_Ola_H3K9ac_IFNb_1h_Rep2	1	R703	TTAGGC	Universal	-
ChIP Histone	2075	ESC_Ola_H3K9me2_IFNb_1h_Rep2	1	R704	TGACCA	Universal	-
ChIP Histone	2076	ESC_Ola_H3K9me3_IFNb_1h_Rep2	1	R705	ACAGTG	Universal	-
ChIP Histone	2077	ESC_Ola_H3K27ac_IFNb_1h_Rep2	1	R706	GCCAAT	Universal	-
ChIP Histone	2078	ESC_Ola_H3K27me3_IFNb_1h_Rep2	1	R707	CAGATC	Universal	-
ChIP Histone	2079	ESC_Ola_H3K36me3_IFNb_1h_Rep2	1	R708	ACTTGA	Universal	-
ChIP Histone	2080	ESC_Ola_H3_IFNb_1h_Rep2	1	R709	GATCAG	Universal	-
ChIP Histone	2081	ESC_Ola_IgG_Rb_IFNb_1h_Rep2	1	R710	TAGCTT	Universal	-
ChIP Histone	2082	ESC_Ola_IgG_M_IFNb_1h_Rep2	1	R711	GGCTAC	Universal	-
ChIP Histone	2083	ESC_Ola_Input_IFNb_1h_Rep2	1	R712	CTTGTA	Universal	-
ChIP Histone	2084	ESC_Ola_H3K4me1_IFNb_6h_Rep2	1	R701	ATCACG	Universal	-
ChIP Histone	2085	ESC_Ola_H3K4me3_IFNb_6h_Rep2	1	R702	CGATGT	Universal	-
ChIP Histone	2086	ESC_Ola_H3K9ac_IFNb_6h_Rep2	1	R703	TTAGGC	Universal	-

ChIP Histone	2087	ESC_Ola_H3K9me2_IFNb_6h_Rep2	1	R704	TGACCA	Universal	-
ChIP Histone	2088	ESC_Ola_H3K9me3_IFNb_6h_Rep2	1	R705	ACAGTG	Universal	-
ChIP Histone	2089	ESC_Ola_H3K27ac_IFNb_6h_Rep2	1	R706	GCCAAT	Universal	-
ChIP Histone	2090	ESC_Ola_H3K27me3_IFNb_6h_Rep2	1	R707	CAGATC	Universal	-
ChIP Histone	2091	ESC_Ola_H3K36me3_IFNb_6h_Rep2	1	R708	ACTTGA	Universal	-
ChIP Histone	2092	ESC_Ola_H3_IFNb_6h_Rep2	1	R709	GATCAG	Universal	-
ChIP Histone	2093	ESC_Ola_IgG_Rb_IFNb_6h_Rep2	1	R710	TAGCTT	Universal	-
ChIP Histone	2094	ESC_Ola_IgG_M_IFNb_6h_Rep2	1	R711	GGCTAC	Universal	-
ChIP Histone	2095	ESC_Ola_Input_IFNb_6h_Rep2	1	R712	CTTGTA	Universal	-
ChIP Histone	2123	ESC_Ola_H3K9me2_IFNb_0h_Rep3	1	R701	ATCACG	Universal	-
ChIP Histone	2124	ESC_Ola_H3K9ac_IFNb_0h_Rep3	1	R702	CGATGT	Universal	-
ChIP Histone	2125	ESC_Ola_H3K9me3_IFNb_0h_Rep3	1	R703	TTAGGC	Universal	-
ChIP Histone	2126	ESC_Ola_H3K27me3_IFNb_0h_Rep3	1	R706	GCCAAT	Universal	-
ChIP Histone	2127	ESC_Ola_H3_IFNb_0h_Rep3	1	R707	CAGATC	Universal	-
ChIP Histone	2128	ESC_Ola_IgG_Rb_IFNb_0h_Rep3	1	R709	GATCAG	Universal	-
ChIP Histone	2129	ESC_Ola_IgG_M_IFNb_0h_Rep3	1	R710	TAGCTT	Universal	-
ChIP Histone	2130	ESC_Ola_Input_IFNb_0h_Rep3	1	R712	CTTGTA	Universal	-
ChIP Histone	2131	ESC_Ola_H3K9me2_IFNb_1h_Rep3	1	R702	CGATGT	Universal	-
ChIP Histone	2132	ESC_Ola_H3K9ac_IFNb_1h_Rep3	1	R701	ATCACG	Universal	-
ChIP Histone	2133	ESC_Ola_H3K9me3_IFNb_1h_Rep3	1	R703	TTAGGC	Universal	-
ChIP Histone	2134	ESC_Ola_H3K27me3_IFNb_1h_Rep3	1	R704	TGACCA	Universal	-
ChIP Histone	2135	ESC_Ola_H3_IFNb_1h_Rep3	1	R705	ACAGTG	Universal	-
ChIP Histone	2136	ESC_Ola_IgG_Rb_IFNb_1h_Rep3	1	R706	GCCAAT	Universal	-
ChIP Histone	2137	ESC_Ola_IgG_M_IFNb_1h_Rep3	1	R707	CAGATC	Universal	-
ChIP Histone	2138	ESC_Ola_Input_IFNb_1h_Rep3	1	R708	ACTTGA	Universal	-
ChIP Histone	2139	ESC_Ola_H3K9me2_IFNb_6h_Rep3	1	R702	CGATGT	Universal	-
ChIP Histone	2140	ESC_Ola_H3K9ac_IFNb_6h_Rep3	1	R701	ATCACG	Universal	-
ChIP Histone	2141	ESC_Ola_H3K9me3_IFNb_6h_Rep3	1	R703	TTAGGC	Universal	-
ChIP Histone	2142	ESC_Ola_H3K27me3_IFNb_6h_Rep3	1	R704	TGACCA	Universal	-
ChIP Histone	2143	ESC_Ola_H3_IFNb_6h_Rep3	1	R705	ACAGTG	Universal	-
ChIP Histone	2144	ESC_Ola_IgG_Rb_IFNb_6h_Rep3	1	R706	GCCAAT	Universal	-
ChIP Histone	2145	ESC_Ola_IgG_M_IFNb_6h_Rep3	1	R707	CAGATC	Universal	-
ChIP Histone	2146	ESC_Ola_Input_IFNb_6h_Rep3	1	R708	ACTTGA	Universal	-
ChIP Histone	243	MEF_Ola_H3K4me3_0h_Rep1	1	R702	CGATGT	Universal	-
ChIP Histone	245	MEF_Ola_H3K9me3_0h_Rep1	1	R704	TGACCA	Universal	-
ChIP Histone	251	MEF_Ola_H3K9me3_0h_Rep2	1	R710	TAGCTT	Universal	-
ChIP Histone	253	MEF_Ola_H3K4me3_0h_Rep2	1	R712	CTTGTA	Universal	-
ChIP Histone	280	MEF_Ola_H3K4me1_0h_Rep1	1	R710	TAGCTT	Universal	-
ChIP Histone	289	MEF_Ola_H3K4me1_0h_Rep2	1	R709	GATCAG	Universal	-
ChIP Histone	290	MEF_Ola_H3K36me3_0h_Rep1	1	R711	GGCTAC	Universal	-
ChIP Histone	291	MEF_Ola_H3K36me3_0h_Rep2	1	R712	CTTGTA	Universal	-
ChIP Histone	322	MEF_Ola_H3K9ac_0h_Rep1	1	R706	GCCAAT	Universal	-
ChIP Histone	323	MEF_Ola_H3K9ac_0h_Rep2	1	R708	ACTTGA	Universal	-
ChIP Histone	350	MEF_Ola_H3K27me3_0h_Rep1	1	R705	ACAGTG	Universal	-
ChIP Histone	351	MEF_Ola_H3K27me3_0h_Rep2	1	R706	GCCAAT	Universal	-
ChIP Histone	352	MEF_Ola_H3K27ac_0h_Rep1	1	R707	CAGATC	Universal	-
ChIP Histone	353	MEF_Ola_H3K27ac_0h_Rep2	1	R708	ACTTGA	Universal	-
ChIP Histone	417	MEF_Ola_H3_0h_Rep1	1	R705	ACAGTG	Universal	-
ChIP Histone	418	MEF_Ola_H3_0h_Rep2	1	R706	GCCAAT	Universal	-
ChIP Histone	419	MEF_Ola_Input_0h_Rep1	1	R707	CAGATC	Universal	-
ChIP Histone	420	MEF_Ola_Input_0h_Rep2	1	R708	ACTTGA	Universal	-
ChIP Histone	244	NPC_Ola_H3K4me3_0h_Rep1	1	R703	TTAGGC	Universal	-
ChIP Histone	247	NPC_Ola_H3K9me3_0h_Rep5	1	R706	GCCAAT	Universal	-
ChIP Histone	248	NPC_Ola_H3K9me3_0h_Rep6	1	R707	CAGATC	Universal	-
ChIP Histone	249	NPC_Ola_H3K4me3_0h_Rep2	1	R708	ACTTGA	Universal	-
ChIP Histone	285	NPC_Ola_H3K4me1_0h_Rep1	1	R705	ACAGTG	Universal	-
ChIP Histone	286	NPC_Ola_H3K4me1_0h_Rep2	1	R706	GCCAAT	Universal	-
ChIP Histone	287	NPC_Ola_H3K36me3_0h_Rep1	1	R707	CAGATC	Universal	-
ChIP Histone	288	NPC_Ola_H3K36me3_0h_Rep2	1	R708	ACTTGA	Universal	-
ChIP Histone	354	NPC_Ola_H3K27me3_0h_Rep3	1	R709	GATCAG	Universal	-
ChIP Histone	355	NPC_Ola_H3K27me3_0h_Rep4	1	R710	TAGCTT	Universal	-
ChIP Histone	356	NPC_Ola_H3K27ac_0h_Rep1	1	R711	GGCTAC	Universal	-
ChIP Histone	357	NPC_Ola_H3K27ac_0h_Rep2	1	R712	CTTGTA	Universal	-
ChIP Histone	1724	NPC_Ola_H3K9ac_0h_Rep1	1	R702	CGATGT	Universal	-
ChIP Histone	1725	NPC_Ola_H3K9me3_0h_Rep1	1	R706	GCCAAT	Universal	-
ChIP Histone	1726	NPC_Ola_H3K27me3_0h_Rep1	1	R707	CAGATC	Universal	-
ChIP Histone	1727	NPC_Ola_H3_0h_Rep1	1	R709	GATCAG	Universal	-
ChIP Histone	1728	NPC_Ola_IgG_Rb_0h_Rep1	1	R711	GGCTAC	Universal	-
ChIP Histone	1729	NPC_Ola_Input_0h_Rep1	1	R712	CTTGTA	Universal	-
ChIP Histone	1730	NPC_Ola_H3K9ac_0h_Rep2	1	R702	CGATGT	Universal	-
ChIP Histone	1731	NPC_Ola_H3K9me3_0h_Rep2	1	R706	GCCAAT	Universal	-
ChIP Histone	1732	NPC_Ola_H3K27me3_0h_Rep2	1	R707	CAGATC	Universal	-
ChIP Histone	1733	NPC_Ola_H3_0h_Rep2	1	R709	GATCAG	Universal	-
ChIP Histone	1734	NPC_Ola_IgG_Rb_0h_Rep2	1	R711	GGCTAC	Universal	-
ChIP Histone	1735	NPC_Ola_Input_0h_Rep2	1	R712	CTTGTA	Universal	-
ChIP Histone	2004	NPC_Ola_Input_0h_Rep3	1	R701	ATCACG	Universal	-
ChIP Histone	2005	NPC_Ola_H3_0h_Rep3	1	R702	CGATGT	Universal	-
ChIP Histone	2006	NPC_Ola_IgG_Rb_0h_Rep3	1	R704	TGACCA	Universal	-
ChIP Histone	2007	NPC_Ola_H3K9me3_0h_Rep3	1	R708	ACTTGA	Universal	-
ChIP Histone	2008	NPC_Ola_Input_0h_Rep4	1	R709	GATCAG	Universal	-
ChIP Histone	2009	NPC_Ola_H3_0h_Rep4	1	R710	TAGCTT	Universal	-
ChIP Histone	2010	NPC_Ola_IgG_Rb_0h_Rep4	1	R711	GGCTAC	Universal	-
ChIP Histone	2011	NPC_Ola_H3K9me3_0h_Rep4	1	R712	CTTGTA	Universal	-
CHIP TF	2716	ESC_Ola_Stat1_CST_IFNb_0h_Rep1	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	2717	ESC_Ola_Stat1_CST_IFNb_1h_Rep1	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	2718	ESC_Ola_Stat1_CST_IFNb_6h_Rep1	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	2719	ESC_Ola_Stat2_CST_IFNb_1h_Rep1	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	2720	ESC_Ola_Stat2_CST_IFNb_6h_Rep1	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	2721	ESC_Ola_Stat2_CST_IFNb_h_Rep1	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	2722	ESC_Ola_Stat1_p701_CST_IFNb_0h_Rep1	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	2723	ESC_Ola_Stat1_p701_CST_IFNb_1h_Rep1	2	i702	TCCGGAGA	i502	ATAGAGGC

CHIP TF	2724	ESC_Ola_Stat1_p701_CST_IFNb_6h_Rep1	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	2725	ESC_Ola_Stat1_p727_CST_IFNb_0h_Rep1	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	2726	ESC_Ola_Stat1_p727_CST_IFNb_1h_Rep1	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	2727	ESC_Ola_Stat1_p727_CST_IFNb_6h_Rep1	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	2728	ESC_Ola_Stat2_p690_IFNb_0h_Rep1	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	2729	ESC_Ola_Stat2_p690_IFNb_1h_Rep1	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	2730	ESC_Ola_Stat2_p690_IFNb_6h_Rep1	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	2731	ESC_Ola_CTCF_IFNb_0h_Rep1	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	2732	ESC_Ola_CTCF_IFNb_1h_Rep1	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	2733	ESC_Ola_CTCF_IFNb_6h_Rep1	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	2734	ESC_Ola_IgG_Rb_CST_IFNb_0h_Rep1	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	2735	ESC_Ola_IgG_Rb_CST_IFNb_1h_Rep1	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	2736	ESC_Ola_IgG_Rb_CST_IFNb_6h_Rep1	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	2737	ESC_Ola_Input_CST_IFNb_0h_Rep1	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	2738	ESC_Ola_Input_CST_IFNb_1h_Rep1	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	2739	ESC_Ola_Input_CST_IFNb_6h_Rep1	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	2798	ESC_Ola_Stat1_CST_IFNb_0h_Rep2	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	2799	ESC_Ola_Stat1_CST_IFNb_1h_Rep2	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	2800	ESC_Ola_Stat1_CST_IFNb_6h_Rep2	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	2801	ESC_Ola_Stat2_CST_IFNb_0h_Rep2	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	2802	ESC_Ola_Stat2_CST_IFNb_1h_Rep2	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	2803	ESC_Ola_Stat2_CST_IFNb_6h_Rep2	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	2804	ESC_Ola_Input_CST_IFNb_0h_Rep2	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	2805	ESC_Ola_Input_CST_IFNb_1h_Rep1_2	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	2806	ESC_Ola_Stat1_p701_CST_IFNb_0h_Rep2	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	2807	ESC_Ola_Stat1_p701_CST_IFNb_1h_Rep2	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	2808	ESC_Ola_Stat1_p701_CST_IFNb_6h_Rep2	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	2809	ESC_Ola_Stat1_p727_CST_IFNb_0h_Rep2	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	2810	ESC_Ola_Stat1_p727_CST_IFNb_1h_Rep2	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	2811	ESC_Ola_Stat1_p727_CST_IFNb_6h_Rep2	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	2812	ESC_Ola_Input_CST_IFNb_1h_Rep2	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	2813	ESC_Ola_Input_CST_IFNb_0h_Rep1_2	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	2814	ESC_Ola_CTCF_CST_IFNb_0h_Rep2	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	2815	ESC_Ola_CTCF_CST_IFNb_1h_Rep2	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	2816	ESC_Ola_CTCF_CST_IFNb_6h_Rep2	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	2817	ESC_Ola_IgG_Rb_CST_IFNb_0h_Rep2	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	2818	ESC_Ola_IgG_Rb_CST_IFNb_1h_Rep2	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	2819	ESC_Ola_IgG_Rb_CST_IFNb_6h_Rep2	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	2820	ESC_Ola_Input_CST_IFNb_6h_Rep2	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	2821	ESC_Ola_Stat1_p701_CST_IFNb_6h_Rep1_2	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3260	MEF_Ola_Stat1_CST_0h_IFNb_Rep1	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3261	MEF_Ola_Stat1_CST_1h_IFNb_Rep1	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3262	MEF_Ola_Stat1_CST_6h_IFNb_Rep1	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3263	MEF_Ola_Stat1_p701_CST_0h_IFNb_Rep1	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3264	MEF_Ola_Stat1_p701_CST_1h_IFNb_Rep1	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3265	MEF_Ola_Stat1_p701_CST_6h_IFNb_Rep1	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	3266	MEF_Ola_Stat2_CST_0h_IFNb_Rep1	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	3267	MEF_Ola_Stat2_CST_1h_IFNb_Rep1	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	3268	MEF_Ola_Stat2_CST_6h_IFNb_Rep1	2	i709	CGGCTATG	i509	CTCTGGAT
CHIP TF	3269	MEF_Ola_IgG_Rb_CST_0h_IFNb_Rep1	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3270	MEF_Ola_IgG_Rb_CST_1h_IFNb_Rep1	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3271	MEF_Ola_IgG_Rb_CST_6h_IFNb_Rep1	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3272	MEF_Ola_Input_CST_0h_IFNb_Rep1	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3273	MEF_Ola_Input_CST_1h_IFNb_Rep1	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3274	MEF_Ola_Input_CST_6h_IFNb_Rep1	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	3275	MEF_Ola_Stat1_CST_0h_IFNb_Rep2	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	3276	MEF_Ola_Stat1_CST_1h_IFNb_Rep2	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	3277	MEF_Ola_Stat1_CST_6h_IFNb_Rep2	2	i709	CGGCTATG	i509	CTCTGGAT
CHIP TF	3278	MEF_Ola_Stat1_p701_CST_0h_IFNb_Rep2	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3279	MEF_Ola_Stat1_p701_CST_1h_IFNb_Rep2	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3280	MEF_Ola_Stat1_p701_CST_6h_IFNb_Rep2	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3281	MEF_Ola_Stat2_CST_0h_IFNb_Rep2	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3282	MEF_Ola_Stat2_CST_1h_IFNb_Rep2	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3283	MEF_Ola_Stat2_CST_6h_IFNb_Rep2	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	3284	MEF_Ola_IgG_Rb_CST_0h_IFNb_Rep2	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	3285	MEF_Ola_IgG_Rb_CST_1h_IFNb_Rep2	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	3286	MEF_Ola_IgG_Rb_CST_6h_IFNb_Rep2	2	i709	CGGCTATG	i509	CTCTGGAT
CHIP TF	3287	MEF_Ola_Input_CST_0h_IFNb_Rep2	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3288	MEF_Ola_Input_CST_1h_IFNb_Rep2	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3289	MEF_Ola_Input_CST_6h_IFNb_Rep2	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3290	ESC_Ola_Stat1_CST_IFNb_0h_Rep3	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3291	ESC_Ola_Stat1_CST_IFNb_1h_Rep3	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3292	ESC_Ola_Stat1_CST_IFNb_6h_Rep3	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	3293	ESC_Ola_Stat1_p701_CST_IFNb_0h_Rep3	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	3294	ESC_Ola_Stat1_p701_CST_IFNb_1h_Rep3	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	3295	ESC_Ola_Stat1_p701_CST_IFNb_6h_Rep3	2	i709	CGGCTATG	i509	CTCTGGAT
CHIP TF	3296	ESC_Ola_Stat2_CST_IFNb_0h_Rep3	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3297	ESC_Ola_Stat2_CST_IFNb_1h_Rep3	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3298	ESC_Ola_Stat2_CST_IFNb_6h_Rep3	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3299	ESC_Ola_IgG_Rb_CST_IFNb_0h_Rep3	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3300	ESC_Ola_IgG_Rb_CST_IFNb_1h_Rep3	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3301	ESC_Ola_IgG_Rb_CST_IFNb_6h_Rep3	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	3302	ESC_Ola_Input_CST_IFNb_0h_Rep3	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	3303	ESC_Ola_Input_CST_IFNb_1h_Rep3	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	3304	ESC_Ola_Input_CST_IFNb_6h_Rep3	2	i709	CGGCTATG	i509	CTCTGGAT
CHIP TF	3363	ESC_Ola_Stat1_CST_IFNb_0h_Rep4	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3364	ESC_Ola_Stat1_p701_CST_IFNb_0h_Rep4	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3365	ESC_Ola_Stat2_CST_IFNb_0h_Rep4	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3366	ESC_Ola_IgG_Rb_CST_IFNb_0h_Rep4	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3367	ESC_Ola_Input_CST_IFNb_0h_Rep4	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3368	ESC_Ola_Stat1_CST_IFNb_1h_Rep4	2	i706	GAATTCGT	i506	TAATCTTA

CHIP TF	3369	ESC_Ola_Stat1_p701_CST_IFNb_1h_Rep4	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	3370	ESC_Ola_Stat2_CST_IFNb_1h_Rep4	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	3371	ESC_Ola_IgG_Rb_CST_IFNb_1h_Rep4	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3372	ESC_Ola_Input_CST_IFNb_1h_Rep4	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3373	ESC_Ola_Stat1_CST_IFNb_6h_Rep4	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3374	ESC_Ola_Stat1_p701_CST_IFNb_6h_Rep4	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3375	ESC_Ola_Stat2_CST_IFNb_6h_Rep4	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3376	ESC_Ola_IgG_Rb_CST_IFNb_6h_Rep4	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	3377	ESC_Ola_Input_CST_IFNb_6h_Rep4	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	3378	ESC_Ola_Stat1_CST_IFNb_0h_Rep5	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	3379	ESC_Ola_Stat1_p701_CST_IFNb_0h_Rep5	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3380	ESC_Ola_Stat2_CST_IFNb_0h_Rep5	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3381	ESC_Ola_IgG_Rb_CST_IFNb_0h_Rep5	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3382	ESC_Ola_Input_CST_IFNb_0h_Rep5	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3383	ESC_Ola_Stat1_CST_IFNb_1h_Rep5	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3384	ESC_Ola_Stat1_p701_CST_IFNb_1h_Rep5	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	3385	ESC_Ola_Stat2_CST_IFNb_1h_Rep5	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	3386	ESC_Ola_IgG_Rb_CST_IFNb_1h_Rep5	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	3387	ESC_Ola_Input_CST_IFNb_1h_Rep5	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3388	ESC_Ola_Stat1_CST_IFNb_1h_Rep6	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3389	ESC_Ola_Stat1_p701_CST_IFNb_1h_Rep6	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3390	ESC_Ola_Stat2_CST_IFNb_1h_Rep6	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3391	ESC_Ola_IgG_Rb_CST_IFNb_1h_Rep6	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3392	ESC_Ola_Input_CST_IFNb_1h_Rep6	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	3393	ESC_Ola_Stat1_CST_IFNb_6h_Rep5	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	3394	ESC_Ola_Stat1_p701_CST_IFNb_6h_Rep5	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	3395	ESC_Ola_Stat2_CST_IFNb_6h_Rep5	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3396	ESC_Ola_IgG_Rb_CST_IFNb_6h_Rep5	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3397	ESC_Ola_Input_CST_IFNb_6h_Rep5	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3398	ESC_Ola_Stat1_CST_IFNb_6h_Rep6	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3399	ESC_Ola_Stat1_p701_CST_IFNb_6h_Rep6	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3400	ESC_Ola_Stat2_CST_IFNb_6h_Rep6	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	3401	ESC_Ola_IgG_Rb_CST_IFNb_6h_Rep6	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	3402	ESC_Ola_Input_CST_IFNb_6h_Rep6	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	3556	NPC_Ola_Stat1_CST_0h_IFNb_Rep1	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3557	NPC_Ola_Stat1_CST_1h_IFNb_Rep1	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3558	NPC_Ola_Stat1_CST_6h_IFNb_Rep1	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3559	NPC_Ola_Stat1_p701_CST_0h_IFNb_Rep1	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3560	NPC_Ola_Stat1_p701_CST_1h_IFNb_Rep1	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3561	NPC_Ola_Stat1_p701_CST_6h_IFNb_Rep1	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	3562	NPC_Ola_Stat2_CST_0h_IFNb_Rep1	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	3563	NPC_Ola_Stat2_CST_1h_IFNb_Rep1	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	3564	NPC_Ola_Stat2_CST_6h_IFNb_Rep1	2	i709	CGGCTATG	i509	CTCTGGAT
CHIP TF	3565	NPC_Ola_CTCF_CST_0h_IFNb_Rep1	2	i710	TCCGCGAA	i510	TCGCCTTA
CHIP TF	3566	NPC_Ola_CTCF_CST_1h_IFNb_Rep1	2	i711	TCTCGCGC	i511	ACTGATCG
CHIP TF	3567	NPC_Ola_CTCF_CST_6h_IFNb_Rep1	2	i712	AGCGATAG	i512	GAGCCTTA
CHIP TF	3568	NPC_Ola_IgG_Rb_CST_0h_IFNb_Rep1	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3569	NPC_Ola_IgG_Rb_CST_1h_IFNb_Rep1	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3570	NPC_Ola_IgG_Rb_CST_6h_IFNb_Rep1	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3571	NPC_Ola_Input_CST_0h_IFNb_Rep1	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3572	NPC_Ola_Input_CST_1h_IFNb_Rep1	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3573	NPC_Ola_Input_CST_6h_IFNb_Rep1	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	3574	NPC_Ola_Stat1_CST_0h_IFNb_Rep2	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	3575	NPC_Ola_Stat1_CST_1h_IFNb_Rep2	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	3576	NPC_Ola_Stat1_CST_6h_IFNb_Rep2	2	i709	CGGCTATG	i509	CTCTGGAT
CHIP TF	3577	NPC_Ola_Stat1_p701_CST_0h_IFNb_Rep2	2	i710	TCCGCGAA	i510	TCGCCTTA
CHIP TF	3578	NPC_Ola_Stat1_p701_CST_1h_IFNb_Rep2	2	i711	TCTCGCGC	i511	ACTGATCG
CHIP TF	3579	NPC_Ola_Stat1_p701_CST_6h_IFNb_Rep2	2	i712	AGCGATAG	i512	GAGCCTTA
CHIP TF	3580	NPC_Ola_Stat2_CST_0h_IFNb_Rep2	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3581	NPC_Ola_Stat2_CST_1h_IFNb_Rep2	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3582	NPC_Ola_Stat2_CST_6h_IFNb_Rep2	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3583	NPC_Ola_CTCF_CST_0h_IFNb_Rep2	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3584	NPC_Ola_CTCF_CST_1h_IFNb_Rep2	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3585	NPC_Ola_CTCF_CST_6h_IFNb_Rep2	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	3586	NPC_Ola_IgG_Rb_CST_0h_IFNb_Rep2	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	3587	NPC_Ola_IgG_Rb_CST_1h_IFNb_Rep2	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	3588	NPC_Ola_IgG_Rb_CST_6h_IFNb_Rep2	2	i709	CGGCTATG	i509	CTCTGGAT
CHIP TF	3589	NPC_Ola_Input_CST_0h_IFNb_Rep2	2	i710	TCCGCGAA	i510	TCGCCTTA
CHIP TF	3590	NPC_Ola_Input_CST_1h_IFNb_Rep2	2	i711	TCTCGCGC	i511	ACTGATCG
CHIP TF	3591	NPC_Ola_Input_CST_6h_IFNb_Rep2	2	i712	AGCGATAG	i512	GAGCCTTA
CHIP TF	3592	NPC_Ola_Stat1_CST_0h_IFNb_Rep3	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3593	NPC_Ola_Stat1_CST_1h_IFNb_Rep3	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3594	NPC_Ola_Stat1_CST_6h_IFNb_Rep3	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3595	NPC_Ola_Stat1_p701_CST_0h_IFNb_Rep3	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3596	NPC_Ola_Stat1_p701_CST_1h_IFNb_Rep3	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3597	NPC_Ola_Stat1_p701_CST_6h_IFNb_Rep3	2	i706	GAATTCGT	i506	TAATCTTA
CHIP TF	3598	NPC_Ola_Stat2_CST_0h_IFNb_Rep3	2	i707	CTGAAGCT	i507	CAGGACGT
CHIP TF	3599	NPC_Ola_Stat2_CST_1h_IFNb_Rep3	2	i708	TAATGCGC	i508	GTACTGAC
CHIP TF	3600	NPC_Ola_Stat2_CST_6h_IFNb_Rep3	2	i709	CGGCTATG	i509	CTCTGGAT
CHIP TF	3601	NPC_Ola_CTCF_CST_0h_IFNb_Rep3	2	i710	TCCGCGAA	i510	TCGCCTTA
CHIP TF	3602	NPC_Ola_CTCF_CST_1h_IFNb_Rep3	2	i711	TCTCGCGC	i511	ACTGATCG
CHIP TF	3603	NPC_Ola_CTCF_CST_6h_IFNb_Rep3	2	i712	AGCGATAG	i512	GAGCCTTA
CHIP TF	3604	NPC_Ola_IgG_Rb_CST_0h_IFNb_Rep3	2	i701	ATTACTCG	i501	TATAGCCT
CHIP TF	3605	NPC_Ola_IgG_Rb_CST_1h_IFNb_Rep3	2	i702	TCCGGAGA	i502	ATAGAGGC
CHIP TF	3606	NPC_Ola_IgG_Rb_CST_6h_IFNb_Rep3	2	i703	CGCTCATT	i503	CCTATCCT
CHIP TF	3607	NPC_Ola_Input_CST_0h_IFNb_Rep3	2	i704	GAGATTCC	i504	GGCTCTGA
CHIP TF	3608	NPC_Ola_Input_CST_1h_IFNb_Rep3	2	i705	ATTCAGAA	i505	AGGCGAAG
CHIP TF	3609	NPC_Ola_Input_CST_6h_IFNb_Rep3	2	i706	GAATTCGT	i506	TAATCTTA
RNA	737	ESC_Ola_IFNb_0h_Rep1	1	R702	CGATGT	Universal	-
RNA	738	ESC_Ola_IFNb_0h_Rep2	1	R704	TGACCA	Universal	-
RNA	739	ESC_Ola_IFNb_1h_Rep1	1	R705	ACAGTG	Universal	-

RNA	740	ESC_Ola_IFNb_1h_Rep2	1	R706	GCCAAT	Universal	-
RNA	741	ESC_Ola_IFNb_6h_Rep1	1	R707	CAGATC	Universal	-
RNA	742	ESC_Ola_IFNb_6h_Rep2	1	R712	CTTGTA	Universal	-
RNA	2160	ESC_Ola_IFNb_0h_Rep3	1	R702	CGATGT	Universal	-
RNA	2161	ESC_Ola_IFNb_0h_Rep4	1	R704	TGACCA	Universal	-
RNA	2162	ESC_Ola_IFNb_1h_Rep3	1	R706	GCCAAT	Universal	-
RNA	2163	ESC_Ola_IFNb_1h_Rep4	1	R708	ACTTGA	Universal	-
RNA	2164	ESC_Ola_IFNb_6h_Rep3	1	R709	GATCAG	Universal	-
RNA	2165	ESC_Ola_IFNb_6h_Rep4	1	R712	CTTGTA	Universal	-
RNA	3073	ESC_R1_WT_IFNb_0h_Rep1	2	i701	ATTACTCG	i501	TATAGCCT
RNA	3074	ESC_R1_WT_IFNb_1h_Rep1	2	i702	TCCGGAGA	i502	ATAGAGGC
RNA	3075	ESC_R1_WT_IFNb_6h_Rep1	2	i703	CGCTCATT	i503	CCTATCCT
RNA	3076	ESC_R1_WT_IFNb_0h_Rep3	2	i704	GAGATTCC	i504	GGCTCTGA
RNA	3077	ESC_R1_WT_IFNb_1h_Rep3	2	i705	ATTCAGAA	i505	AGGCGAAG
RNA	3078	ESC_R1_WT_IFNb_6h_Rep3	2	i706	GAATTCGT	i506	TAATCTTA
RNA	3079	ESC_R1_dKO_IFNb_0h_Rep2	2	i707	CTGAAGCT	i507	CAGGACGT
RNA	3080	ESC_R1_dKO_IFNb_1h_Rep2	2	i708	TAATGCGC	i508	GTACTGAC
RNA	3081	ESC_R1_dKO_IFNb_6h_Rep2	2	i709	CGGCTATG	i509	CTCTGGAT
RNA	3094	ESC_R1_dKO_IFNb_0h_Rep3	2	i702	TCCGGAGA	i502	ATAGAGGC
RNA	3095	ESC_R1_dKO_IFNb_1h_Rep3	2	i703	CGCTCATT	i503	CCTATCCT
RNA	3096	ESC_R1_dKO_IFNb_6h_Rep3	2	i704	GAGATTCC	i504	GGCTCTGA
RNA	3085	ESC_R1_dCD_IFNb_0h_Rep2	2	i703	CGCTCATT	i503	CCTATCCT
RNA	3086	ESC_R1_dCD_IFNb_1h_Rep2	2	i704	GAGATTCC	i504	GGCTCTGA
RNA	3087	ESC_R1_dCD_IFNb_6h_Rep2	2	i705	ATTCAGAA	i505	AGGCGAAG
RNA	3088	ESC_R1_dCD_IFNb_0h_Rep3	2	i706	GAATTCGT	i506	TAATCTTA
RNA	3089	ESC_R1_dCD_IFNb_1h_Rep3	2	i707	CTGAAGCT	i507	CAGGACGT
RNA	3090	ESC_R1_dCD_IFNb_6h_Rep3	2	i708	TAATGCGC	i508	GTACTGAC
RNA	3082	MEF_Ola_IFNb_0h_Rep1	2	i710	TCCGCGAA	i510	TCGCCTTA
RNA	3083	MEF_Ola_IFNb_1h_Rep1	2	i701	ATTACTCG	i501	TATAGCCT
RNA	3084	MEF_Ola_IFNb_6h_Rep1	2	i702	TCCGGAGA	i502	ATAGAGGC
RNA	3091	MEF_Ola_IFNb_0h_Rep2	2	i709	CGGCTATG	i509	CTCTGGAT
RNA	3092	MEF_Ola_IFNb_1h_Rep2	2	i710	TCCGCGAA	i510	TCGCCTTA
RNA	3093	MEF_Ola_IFNb_6h_Rep2	2	i701	ATTACTCG	i501	TATAGCCT
RNA	3097	NPC_Ola_IFNb_0h_Rep1	2	i705	ATTCAGAA	i505	AGGCGAAG
RNA	3098	NPC_Ola_IFNb_1h_Rep1	2	i706	GAATTCGT	i506	TAATCTTA
RNA	3099	NPC_Ola_IFNb_6h_Rep1	2	i707	CTGAAGCT	i507	CAGGACGT
RNA	3100	NPC_Ola_IFNb_0h_Rep2	2	i708	TAATGCGC	i508	GTACTGAC
RNA	3101	NPC_Ola_IFNb_1h_Rep2	2	i709	CGGCTATG	i509	CTCTGGAT
RNA	3102	NPC_Ola_IFNb_6h_Rep2	2	i710	TCCGCGAA	i510	TCGCCTTA
scATAC	3310	ESC_Ola_IFNb_0h_Rep1	Chromium i7 Multiplex Kit N, Set A	SI-NA-A1	AAACGGCG | CCTACCAT | GGCGTTTC | TTGTAAGA	Not specific	NNNNNNNN
scATAC	3311	ESC_Ola_IFNb_6h_Rep1	Chromium i7 Multiplex Kit N, Set A	SI-NA-A2	AGCCCTTT | CAAGTCCA | GTGAGAAG | TCTTAGGC	Not specific	NNNNNNNN
scRNA	2655	ESC_Ola_IFNb_0h_Rep1	Chromium i7 Multiplex Kit	SI-GA-A6	CGCTATGT | GCTGTCCA | TTGAGATC | AAACCGAG	Not specific	NNNNNNNN
scRNA	2657	ESC_Ola_IFNb_1h_Rep1	Chromium i7 Multiplex Kit	SI-GA-A7	ACAGAGGT | TATAGTTG | CGGTCCCA | GTCCTAAC	Not specific	NNNNNNNN
scRNA	2658	ESC_Ola_IFNb_6h_Rep1	Chromium i7 Multiplex Kit	SI-GA-A8	GCATCTCC | TGTAAGGT | CTGCGATG | AACGTCAA	Not specific	NNNNNNNN

Table 7: NGS sequencing primers for all approaches

Short cuts for origin of barcodes.
1: TruSight Tumor 15; Identical to: NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 1); E7335S
2: TruSeq HT; Based on NEBNext® Multiplex Oligos for Illumina® (Dual Index Primers Set 1); E7600S
3: Nextera Index Kit

7. Danksagung

Allem vorweg ein riesengroßes **Dankeschön** an alle Menschen, die mich auf dem Weg zum Doktor begleitet und unterschützt haben.

Zuallererst möchte ich Karsten danken.
Dafür, dass du mir die Möglichkeit gegeben hast, eine PhD in deiner Arbeitsgruppe zu machen. Dafür, dass du mich über die Jahre durchgehend gefordert und gefördert hast. Dafür, dass du mir die Möglichkeit gegeben hast, mich selbst weiterzuentwickeln. Dafür, dass du mich bis zur Abgabe konstruktive unterstützt hast. Danke!

Natürlich sind großartige Arbeitskollegen ein essentieller Bestanteil um erfolgreich einen PhD abzuschließen. An alle, die da waren und weitergezogen sind. An alle, die gekommen sind und noch hier sind. Es ist und war großartig in dieser konstruktiven, freundschaftlichen Atmosphäre zu arbeiten. Danke.
Und auch für die Kaffeepausen, die Kaffeemaschine, die Kommentare nach den Group Meetings, das Erdulden der vielen Fragen, die Testtalks, Cocktails, das neue Gruppenfoto auf der Webpage, die Überwindung einer Singstar-Phobie, Gamenights, Cocktails (nochmals) und alles andere...
Speziellen Dank für die Unterstützung auf allen Ebenen an Lara, Fabian, Delia, Lukas, Jorge, Isabelle, Alexandra, Philipp, Kathi, Robin und Armin. Auch an Caro, Sabrina, Nick und Simon, ohne deren Hilfe vieles nicht so glatt laufen würde, wie es das tut.

A big thanks to Dr. Steeve Boulant, for supporting me over the years and being part of my TAC and thesis committee.

I would also like to thank Prof. Dr. Ralf Bartenschlager and Prof. Dr. Ursula Klingmüller, and for agreeing to be part of my PhD thesis committee. In addition, I would also like to thank Dr. Julien Bethune for his support as TAC member.

For Denise. Inspiring scientist and brave woman. For providing unlimited support. Thank you!

Meine Eltern möchte ich für die Unterstützung über all die Jahre danken. Ich bin sehr glücklich, dass ihr mich immer unterschützt habt, egal was ich mir eingebildet habe. Vom Anfang des Studiums bis zum PhD. Einfach Danke. Bleibt wie ihr seid.

Und um eine alte Weisheit von mir aufzugreifen: Liebe Sarah, ich bin auch froh, dass keine Doktorarbeit angefangen hast, das hat den Druck schon etwas rausgenommen. Und auch Dieter, Sylvia, Sylvia, Fritz und Markus für euch Unterschätzung über die Jahre. Liebste Anna Oma, ich finde es besonders schön und großartig, dass wir diesen Abschluss miteinander feiern können. Ich hoffe sehr, es wird noch viele solcher Gelegenheiten geben.

Seit dem ersten Tag der PhD Selection am DKFZ kennen wir uns und es freut mich sehr, dass wir auch gemeinsam unsre PhDs abschließen. Es war eine gute, stressige und einmalig schöne Zeit. Ich bin gespannt, wo wir alle landen werden. Danke Ellie, David, Tanvi, Micha, Tobi und Nina. Pooja, Alex, Janick und Ines, ich bin super froh, dass wir alle während unserer Zeit am DKFZ Freunde geworden sind.

Angela, JJ, Annabelle, Pat und Tina. Das Beste, was mir aus der Zeit in Mainz geblieben ist. Bin sehr froh, dass wir uns getroffen haben und noch immer Freunde sind. Außerdem für eure Unterstützung über all die Jahre hinweg. In guten wie in schlechten Zeiten.

Sylvie, vom Bachelor in Wien bis zum PhD am DKFZ, könnte ja fast meine, das war geplant. Es ist toll, dass es dich gibt. Bleib einfach so wie du bist (also cremig). Danke!

Betty, wie wir mit großer Sorge festgestellt haben, haben wir die 7 Jahre schon lange überschritten. Heißt, wohl wir kriegen uns nicht mehr los. Kann mir schlimmeres vorstellen. Und irgendwie ist das schön. Danke!

Daniel, Manu und Phips. Immer eine Freude euch zu sehen. Ich fühle mich immer herzlichst willkommen. Und Daniel THX für den wöchentlichen Support, hoffe wir halten das am Laufen.

Sachen die gesagt werden müssen:

HAUBE! Get used to it! – Trester-Bitch! – "*Getting crushed by a wending machine*" – Manuel analyzing – "Bad Austrian!" – "*Egoista di prima categoria*" – Einfach mal rumpöbeln!

DANKE!
Muckenhuber – Ende.