

Transcriptional and epigenetic regulation of differentially activated macrophages

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Wolfgang Krebs

aus Vorra, Deutschland

Bonn, März 2016

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. med. Joachim L. Schultze
2. Gutachter: Prof. Dr. Waldemar Kolanus

Tag der Promotion: 27.10.2016

Erscheinungsjahr: 2016

Eidesstattliche Erklärung

Hiermit versichere ich, dass die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne die Benutzung anderer als der angegebenen Quellen angefertigt wurde. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind gemäß §6 der Promotionsordnung vom 17.06.2011 als solche kenntlich gemacht.

Teile dieser Arbeit wurden bereits in Form von wissenschaftlichen Artikeln in den Journals *Nucleic Acids Research* (results part 5.1) und *Cell Research* (results part 5.2) mit der Hilfe der angegebenen Autoren veröffentlicht:

“Optimization of transcription factor binding map accuracy utilizing knockout-mouse models”. Krebs, W.*, Schmidt, S.V.*, Goren, A., De Nardo, D., Labzin, L., Bovier, A., Ulas, T., Theis, H., Kraut, M., Latz, E., Beyer M., Schultze JL.

Nucleic Acids Research, 2014, Vol. 42, No. 21 13051–13060

* shared first authors

“The transcriptional regulator network of human inflammatory macrophages is defined by open chromatin”. Schmidt, S.V.*, Krebs, W.*, Ulas, T.*, Xue, J., Bassler, K., Günther, P., Hardt, AL., Schultze, H., Sander, J., Klee, K., Theis, H., Kraut, M., Beyer, M., Schultze, JL.

Cell Res. 2016 Feb;26(2):151-70. doi: 10.1038/cr.2016.1. Epub 2016 Jan 5.

* shared first authors

RNA-seq von Makrophagen wurden von Michael Kraut und Laura Bohmann durchgeführt. Die weitere Rohdatenanalyse wurde von Jil Sander vorgenommen. Außerdem wurden die Makrophagen und ChIP-seq Experimente für H3K4me3 mit der Hilfe von Stephanie Riesenberg etabliert. Die bioinformatische Datenanalyse wurde mit Hilfe von Dr. Susanne V. Schmidt, Dr. Thomas Ulas, Jia Xue und Jil Sander ausgeführt.

Vorra, den 28.03.2016

.....

Wolfgang Krebs

Acknowledgements

I would like to express my deepest gratitude to Professor Dr. Joachim Schultze for the chance to work in his excellent group and for the constant support and fresh ideas making the presented work possible. I also want to thank Professor Eicke Latz and Professor Anton Bovier for their support allowing the successful publication of my scientific results.

Furthermore, I want to thank all Postdocs in the Schultze group for their constant attendance to give competent support no matter if it was day or night, especially Dr. Susanne V. Schmidt and Dr. Marc Beyer. I want to give special thanks to Dr. Astrid Draffehn and Dr. Thomas Ulas giving me advice in uncountable situations.

I also want to thank all Phd students of the Genomics and Immunoregulation group at the LIMES Institute, too many to name them all, which were available for constructive discussions and help in the lab and beyond. Especially, Jil Sander and Jia Xue were always there to help me solve “binary” problems. Furthermore, the daily work in the wet lab was enriched by the company of Maren Mai, Daniel Sommer, Andrea Nino, Yasser Thabet and Lisa Schmidleithner.

I also want to mention Larisa Labzin, Dominic De Nardo, Stefanie Riesenberger, Laura Bohmann and Michael Kraut for their support in generating sequencing data sets and other members of neighboring laboratories in the LIMES Institute for their help over the last years.

Last but not least, without my family, Michael and Stefan the presented work would not have been possible. Thank you for your constant support.

This work was financially supported by the German Research Foundation [SFB 704, SFB 832, INST 217/575-1, INST 217/576-1, INST 217/577-1].

Contents

ACKNOWLEDGEMENTS	3
ABBREVIATIONS	9
SUMMARY	12
1. INTRODUCTION	14
1.1 THE INNATE IMMUNE SYSTEM	14
1.2 MACROPHAGE ACTIVATION AND PLASTICITY	14
1.2.1 <i>PU.1 - A pioneer transcription factor in macrophage differentiation</i>	17
1.3 EPIGENETICS	19
1.3.1 <i>DNA methylation</i>	20
1.3.2 <i>DNA structure</i>	20
1.3.3 <i>RNA mediated processes</i>	21
1.3.4 <i>Histone modifications</i>	22
1.3.5 <i>Super enhancers – A distinct class of cis-regulatory sites</i>	24
1.3.6 <i>Epigenetic regulation in macrophages</i>	25
1.4 NEXT GENERATION SEQUENCING TECHNOLOGY	27
1.4.1 <i>Next generation sequencing method variations</i>	28
1.4.2 <i>Technical considerations for ChIP-seq experiments</i>	31
2. OBJECTIVES	35
3. MATERIALS	36
3.1. CHEMICALS AND REAGENTS	36
3.2. CYTOKINES.....	37
3.3. ANTIBODIES	38
3.4. RT-PCR PRIMERS	38
3.5. PLASTIC WARE	39
3.6. EQUIPMENT.....	40
3.7. BUFFERS AND MEDIA	40
3.8. KITS	41
3.9. SOFTWARE.....	42
3.10. NEXT GENERATION SEQUENCING DATA SETS	43
4. METHODS	44
4.1. ISOLATION OF MONOCYTES.....	44
4.2. GENERATION AND ACTIVATION OF HUMAN MONOCYTE DERIVED MACROPHAGES	45
4.3. GENERATION OF MURINE BONE-MARROW DERIVED MACROPHAGES.....	46
4.4. FLOW CYTOMETRY.....	46

4.5. SEMI-QUANTITATIVE REAL TIME PCR	47
4.6. ISOLATION OF RNA FOLLOWING RNA-SEQUENCING	48
4.7. CHIP FOLLOWING DEEP SEQUENCING	49
4.7.1. <i>Native ChIP (N-ChIP)</i>	50
4.7.2. <i>Cross-linked ChIP (X-ChIP)</i>	51
4.8. BIOINFORMATICS DATA ANALYSIS	52
4.8.1. <i>Conversion of published data sets</i>	52
4.8.2. <i>Alignment to reference genome assembly</i>	52
4.8.3. <i>Peak Calling</i>	53
4.8.4. <i>Peak annotation and tag distributions</i>	55
4.8.5. <i>De novo motif enrichment analysis</i>	56
4.8.6. <i>Gene ontology enrichments</i>	56
4.8.7. <i>Identification of “Hyper-ChIPable regions”</i>	57
4.8.8. <i>Epigenetic promoter and enhancer class identification</i>	58
4.8.9. <i>Correlation of gene expression to enhancer and promoter classification</i>	59
4.8.10. <i>Generation of co-regulation networks for transcriptional regulators and macrophage core signature genes</i>	59
5. RESULTS	61
5.1. KNOCKOUT MOUSE MODELS - A TOOL FOR THE OPTIMIZATION OF TF CHIP-SEQ EXPERIMENTS	61
5.1.1. <i>Bioinformatic processing during “Knockout implemented normalization method” (KOIN)</i>	62
5.1.2. <i>KOIN identifies and excludes false-positive ChIP-seq signals</i>	65
5.1.3. <i>Altered signal-to-noise ratios can increase peak numbers</i>	69
5.1.4. <i>Effect of KOIN on “hyper-ChIPable regions”</i>	71
5.1.5. <i>Increased identification of ChIP protein specific binding motifs after KOIN correction</i>	72
5.1.6. <i>KOIN significantly enhances biological interpretation of TF ChIP-seq data</i>	76
5.2. EPIGENETIC CHANGES DURING ACTIVATION OF PRIMARY HUMAN MACROPHAGES	78
5.2.1. <i>ChIP-seq validation procedures</i>	79
5.2.2. <i>Optimization of next generation sequencing library construction</i>	81
5.2.3. <i>Epigenetic classification of specific promoter and enhancer states during human macrophage activation</i>	83
5.2.4. <i>Epigenetic core program in differentially activated macrophages</i>	93
5.2.5. <i>Influence of super enhancers onto macrophage activation</i>	100
5.2.6. <i>Epigenetic regulation of central macrophage transcriptional regulators in human</i>	105
5.2.7. <i>Epigenetic and transcriptional regulation in human tissue-defined TR networks</i>	110
5.2.8. <i>Transcriptional regulator networks in tissue macrophages</i>	113
5.2.9. <i>Model of epigenetic regulation in human macrophages in an inflammatory model</i>	118
6. DISCUSSION	120

6.1. FALSE-POSITIVE AND NEGATIVE SIGNALS IN CHIP-SEQ EXPERIMENTS.....	120
6.2. CORRECT BIOLOGICAL INTERPRETATION REQUIRES KOIN CORRECTION.....	122
6.3. KNOCKOUT DATA SETS - OPTIMAL CONTROL FOR TF CHIP-SEQ EXPERIMENTS	123
6.4. A COMMON EPIGENETIC CORE PROGRAM DEFINES THE BASIC TRANSCRIPTIONAL LANDSCAPE IN MACROPHAGES.....	124
6.5. PROFOUND CHANGES AT PROMOTER OR CIS-REGULATORY SITES ALTER THE TRANSCRIPTIONAL PROGRAM IN ACTIVATED HUMAN MACROPHAGES.....	125
6.6. SPECIALIZED TRANSCRIPTIONAL CONTROL OF TRS IN ACTIVATED HUMAN MACROPHAGES	128
6.7. EPIGENETICS - THE “NEXT GENERATION” TOOL FOR THE IDENTIFICATION OF REGULATORY GENOMIC SITES.....	131
7. REFERENCES	135
ZUSAMMENFASSUNG.....	155

Figures

Figure 1. Schema of DNA/chromatin structure and site of histone modification	23
Figure 2. Schematic overview of variances influencing ChIP-seq data files.	33
Figure 3. Workflow for primary human macrophage generation and differentiation.	46
Figure 4. Flowchart depicting the generation of false-positive corrected peak files (KOIN method).	63
Figure 5. Schema of ChIP-seq data analysis using either the KOIN method or the standard method.	64
Figure 6. Independent ChIP-seq experiments reveal false-positive signals after KOIN calculations.	66
Figure 7. Genomic annotation of KOIN corrected and uncorrected peaks reveal a high distribution of false-positive peaks to various genomic positions.	67
Figure 8. False-positive signals are depleted by KOIN correction in TF ChIP-seq data sets.	68
Figure 9. Representative genomic positions for true-positive and false-positive ChIP-seq peaks differ in KO signal intensities.	69
Figure 10. KOIN correction improves signal-to-noise ratios and corresponding statistical significance for KOIN peaks.	70
Figure 11. Examples of Hyper-ChIPable regions with highly enriched ChIP-seq signals.	72
Figure 12. TF motif analysis is significantly improved after KOIN correction.	73
Figure 13. Absent enrichment for specific TF motifs in uncorrected KO data sets.	74
Figure 14. Enriched WT motifs are exclusively found in WT data sets compared to KO results.	75
Figure 15. GO-term enrichment analysis for promoter regions is significantly improved after KOIN correction.	77
Figure 16. KOIN also improves GOEA analysis at <i>cis</i> -regulatory genomic regions.	78
Figure 17. FACS analysis of surface marker expression in activated macrophages.	79
Figure 18. Quality controls for successful ChIP experiments.	80
Figure 19. Sequencing results for ChIP-seq libraries constructed with variable amounts of DNA input material.	82
Figure 20. Representative size distribution of macrophage ChIP-seq library.	83
Figure 21. Definition of epigenetic states at promoter and enhancer sites.	84
Figure 22. Activation state specific epigenetic promoter states in human macrophages.	85
Figure 23. PU.1 frequencies at stimulus specific accessible M ^{TPP} promoters and GO-terms related to typical macrophage functions.	86
Figure 24. Boxplots of expression values for genes marked with stimulus specific accessible and poised promoters.	87
Figure 25. Identification of specific epigenetic enhancer states and corresponding PU.1 frequencies in four macrophage activation states.	89
Figure 26. RNA-seq expression values of genes marked with activation specific enhancers.	90
Figure 27. Visualization for representative sites in M ^{TPP} with activation specific accessible promoter and strong enhancers.	91
Figure 28. Correlation analysis of promoters and enhancers with different epigenetic activity markings.	92
Figure 29. Common accessible and poised promoters in activated macrophages.	94
Figure 30. Expression values and PU.1 frequencies of common macrophage promoter and enhancer sites.	96

Figure 31. Differential Gene Ontology Enrichment Analysis of CHCS genes marked with common Pa or common Es.	97
Figure 32. Expression values of common macrophage genes with similar epigenetic status at regulatory sites compared to other cell types of the immune system.	98
Figure 33. Motif enrichment analysis of common Pa and Es sites found in four activated macrophages.	99
Figure 34. Network visualization for the epigenetic status of regulatory sites in common macrophage core signature genes.	100
Figure 35. Identification of super enhancers in activated macrophages.	101
Figure 36. Super enhancer characterization.	102
Figure 37. Correlation of gene ontology terms to genes with SE marks.	103
Figure 38. Visualization of genomic loci for macrophage activation state specific SE loci. ...	104
Figure 39. Epigenetic landscape of transcriptional regulators in baseline macrophages.	106
Figure 40. Promoter and enhancer mark distributions for transcriptional regulators in activated macrophages.	107
Figure 41. Epigenetic landscape of transcriptional regulators in activated macrophages.	109
Figure 42. Main features of ChIP-seq datasets for five different human tissues.	111
Figure 43. Epigenetic control of transcriptional regulators in human tissues.	112
Figure 44. Workflow for the generation of TR networks in murine tissue macrophages.	113
Figure 45. Epigenetic regulation of TRs in murine tissue macrophages.	115
Figure 46. Epigenetic landscape of TR in murine tissue macrophages.	117
Figure 47. Model for epigenetic regulation in differentially activated macrophages.	118

Tables

Table 1. Utilized next generation sequencing data sets.	43
Table 2. Cytokines and chemicals used for macrophage activation.	45
Table 3. qRT-PCR reaction mix.	47
Table 4. Standard RT-qPCR program.	48

Abbreviations

AP-1	Activator protein-1
APC	Antigen presenting cells
ATF3	Cyclic AMP-dependent transcription factor
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
cAMP	Cyclic adenosine monophosphate
CD	Cluster of differentiation
CFSE	Carboxyfluorescein diacetate succinimidyl ester
ChIP	Chromatin immunoprecipitation
ChIP-seq	Chromatin immunoprecipitation following deep sequencing
CTLA-4	Cytotoxic T-lymphocyte antigen 4
DC	Dendritic cells
DE	Differential expressed
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphates
EDTA	Ethylenediaminetetraacetic acid
EGTA	ethylene glycol tetraacetic acid
EtBr	Ethidium bromide
FACS	Fluorescence activated cell sorting
FC	Fold of change
FCS	Fetal calf serum
FITC	Fluorescein isothiocyanat
GATA3	GATA-binding protein 3
GFP	Green fluorescent protein
GOEA	Gene ontology enrichment analysis
GREAT	Genomic Regions Enrichment of Annotations Tool
H3K27Ac	Acetylation of lysine 27 on histone protein 3
H3K27me3	Trimethylation of lysine 27 on histone protein 3
H3K4me1	Monomethylation of lysine 27 on histone protein 3
H3K4me3	Trimethylation of lysine 4 on histone protein 3
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HLA	Human leukocyte antigen
HM	Histone modification

IFN	Interferon
IFN γ	Interferon gamma
Ig	Immunoglobulin
IL	Interleukin
IL-4	Interleukin-4
IRF	Interferon response factor
KO	Knockout
KOIN	Knockout implemented normalization
LiCl	Litiumchloride
LPS	Lipopolysaccharide
M Φ	Macrophage
MACS	Magnetic assorted cell sorting
M ^b	Unstimulated macrophages
MHC	Major Histocompatibility Complex
M ^{IFNγ}	IFN γ stimulated inflammatory macrophages
M ^{IL4}	IL-4 stimulated macrophages
MNase	Micrococcal nuclease
M ^{TPP}	TNF- α , PGE2, Pam3 stimulated macrophages
Mf	Macrophages
NaCl	Sodiumchloride
NEB	New england biolabs
NFAT	Nuclear factor of activated T-cells
NF- κ B	Nuclear factor kappa-light-chain enhancer of activated T-cells
NGS	Next generation sequencing
NP-40	Tergitol-type NP-40
P3C	Pam3CSK4
PAMPs	Pathogen- associated molecular patterns
PBMC	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PE	Phycoerythrin
Pfu	Pyrococcus furiosus
PGE2	Prostaglandin E2
PI	Propidium iodide
PMSF	phenylmethylsulfonyl fluoride
PNK	Polynucleotide kinase
PPR	Pathogen recognition receptors
qRT-PCR	Semi quantitative real time PCR

RE	Relative expression
rh	Recombinant human
RIPA buffer	Radioimmunoprecipitation assay buffer
rm	Recombinant murine
RNA	Ribonucleic acid
RNA-Pol II	RNA Polymerase 2
ROS	Reactive oxygen species
SDS	Sodium dodecylsulfate
Seq	Sequencing
siRNA	Small interfering RNA
SRF	Serum response factor
Tag	Computationally aligned ChIP-seq DNA sequence
TBE	Tris/Borate/EDTA
Tconv	Conventional T-cells
TCR	T cell receptor
TE buffer	Tris-EDTA buffer
TF	Transcription factor
TH1	T helper 1
TH17	T helper 17
TLR	Toll like receptor
TNF- α	Tumor necrosis factor- α
TR	Transcriptional regulators
Treg	Regulatory T cells
TSS	Transcriptional start site
TTS	Transcriptional terminating site
WT	Wild type

Summary

Macrophages are an indispensable part of the innate immune system which mediate various functions including host defense against pathogens, metabolism, tissue homeostasis and even developmental processes. These extremely heterogeneous cells can adapt their transcriptional program upon a plethora of stimulatory cues and thus exist in different activation states to facilitate their diverse roles in the body. Corresponding transcriptional changes are established amongst others by transcriptional regulators (TRs) with diverse functions or by complex epigenetic alterations. Next generation sequencing technologies provide excellent experimental methods like ChIP- or RNA-sequencing, with which one can analyze genome wide enrichment properties of DNA binding proteins or the transcriptional activity of genes to elucidate in detail the activation of macrophages on the transcriptional level.

Integrating the KO implemented normalization method (KOIN) into the standard peak calling procedure revealed multiple enhancements for ChIP-seq data analysis. False-positive signals can be eliminated in a tremendous amount, while signal-to-noise ratios are increased in low and even high quality ChIP-seq data sets. Besides the identification and removal of a recently identified special type of false-positive signal called “hyper-ChIPable regions”, the biological interpretation can profoundly benefit from KOIN. Overall, the KOIN method demonstrated its value as new possible gold standard control with various advantages compared to the currently established Input chromatin and IgG ChIP-seq controls.

Furthermore, the ChIP-seq technology allows the definition of 1) different activity states for promoters or *cis*-regulatory regions and 2) important regulators in the establishment and maintenance of the transcriptional landscape by the detection of different covalent posttranslational histone modifications (HM), like acetylation or methylation. Four differentially activated primary human macrophages demonstrated a common epigenetic core program, maintained by various promoter sites. Simultaneously, activation state specific epigenetic differences at promoters, super-enhancer regions and especially at enhancer sites could mediate their specialization upon employed stimulatory signals. Finally, despite the detected epigenetic differences an astonishing fraction of genomic loci was defined by accessible promoter and enhancer markings in macrophage activation states. This was especially demonstrated in co-regulation networks for TRs and revealed an uncoupling of epigenetic and transcriptional control in monocyte-derived activated macrophages associated with cellular plasticity in response to microenvironmental signals. Other additional levels of transcriptional fine-tuning like enhancer RNAs, repressor proteins or the cross-talk between HMs could play an important role in fine-tuning macrophage transcription.

Especially, the cooperative binding of pioneer transcription factors (TF) like PU.1 with other secondary TFs like STAT proteins to these open genomic macrophage loci could represent an additional important switch in macrophage transcription in concert with HMs.

1. Introduction

1.1 The innate immune system

Throughout life an individual faces constant threats by infections with microbial pathogens and viruses or other dangerous biological compounds. Survival depends on the recognition of infectious microbes and an appropriate defense response. Two fundamental parts of the immune system are responsible for this recognition to ensure host survival (1,2). The adaptive immunity, an evolutionary achievement of vertebrates creates specialized and highly specific antigen receptors *de novo* in each organism, which are not-germ line encoded (3). In contrast, the innate immune system is the evolutionary more ancient form of host defense, found in most multicellular organisms and recognizes threats with a defined set of germ line encoded pathogen recognition receptors (PPR) (4). Mainly pathogen associated molecular patterns (PAMPs) but also abnormal self-antigens are recognized by PPRs. They can induce in their secreted form the complement system and lead to the phagocytosis of bound microorganisms by macrophages. Transmembrane receptors like Toll-like receptors (TLRs) recognize specific microbial compounds (5). In humans, ten different members of the TLR family can recognize intra- or extracellular PAMPs. Extracellular TLRs can identify lipoteichoic acid (TLR1/2), lipoproteins (TLR2/6), lipopolysaccharide (TLR4) or flagelin (TLR5). In the membrane of phagocytic compartments localized intracellular TLRs (TLR3/7/9) can recognize nucleic acids (6). All TLR signaling pathways ultimately lead to the expression and activation of the TFs nuclear factor kB (NFkB) and activation protein-1 (AP-1) inducing the transcription of genes necessary for microbial immune defense. If the innate immune system is not able to eliminate microbial infections, the adaptive immunity activates T and B cell mediated immunity, providing specific recognition by clonal amplification of antigen specific receptors and additionally provides memory for faster resolution of repeated challenges with the same microbial antigen. But in most cases, innate immune cells like neutrophils, granulocytes, mast cells, dendritic cells or macrophages respond very rapidly upon a challenge compared to adaptive immune cells and resolve the threat.

1.2 Macrophage activation and plasticity

Macrophages, fundamental effectors of the innate immune system, fulfill many specialized roles in the defense against pathogens, in developmental processes of the body, in tissue homeostasis and even in metabolic functions (7-9). They originate from circulating monocytes or tissue-resident precursors (10,11). Depending on their tissue localization

macrophages also play different roles in homeostasis (12). For example, Kupffer cells in the liver phagocytize erythrocytes and recycle hemoglobin together with red pulp macrophages of the spleen to maintain iron homeostasis (13,14), whereas microglia cells in the brain aid in synaptic pruning (15,16). Alveolar macrophages in the lung clear mucus and take up surfactant lipids (17). These are only some examples of various macrophage types with distinct specialized functions (18). Thus, macrophages are defined by their outstanding diversity and plasticity reacting to environmental cues like cytokines, microbial antigens or damaged cells finally emerging into different functional states.

Importantly, despite their common monocytic origin tissue resident macrophages and monocytic-derived circulating macrophages show substantial differences. Tissue macrophages for example display a general immunomodulatory phenotype, while infiltrating recruited monocytes are activated into a pro- or anti-inflammatory phenotype depending on the tissue conditions (19). In steady state conditions the discrimination between tissue-resident and monocyte-derived macrophages is not possible, but alternatively activated tissue macrophages revealed different transcriptional and phenotypic profiles from alternatively activated monocyte-derived macrophages (20).

For many years, the classical paradigm of macrophage polarization into two major subtypes, classically activated (M1) and alternatively activated (M2) macrophages was established in the scientific community (7). In this model, the *in vitro* stimulation of bone-marrow derived murine macrophages with microbial agonists or cytokines mimicking *in vivo* situations allowed the characterization of two major functional and phenotypic profiles mirroring the Th1-Th2 polarization of T cells. Recently, computational methods analyzing at least 28 distinct human macrophage activation programs while utilizing 299 highly standardized microarray transcriptome data sets expanded the current view of macrophage polarization (21). Reverse network engineering and the integration of database-stored knowledge allowed not only the identification of a common macrophage program but also distinct activation stimulus specific transcriptional changes, suggesting a “spectrum model” with at least nine significantly different macrophage activation programs. The “spectrum model” better accommodates for the *in vivo* situation of macrophages, since a plethora of stimuli are present in the host microenvironment, generating various possible distinct activation states with functional differences, reflecting the fundamental changes on transcriptional level.

Nevertheless, previous publications describe in detail important mechanisms and involved factors in macrophage activation, which demonstrated valuable insights into macrophages independently of the recent findings in macrophage plasticity. Classically activated inflammatory macrophages M^{IFN γ} (formerly known as M1), targets of extensive effort to elucidate their activation and functional implications are generated *in vitro* by the stimulation

with TLR ligands, IFN γ or TNF- α (22). These IFN γ / TNF- α associated macrophages produce pro-inflammatory cytokines like IL-12p40, TNF- α , IL-6 or IL-23, also promoting T helper 1 and T helper 17 responses (23), whereas only low levels of IL-10 are expressed. The production of pro-inflammatory chemokines and chemokine receptors like CCL2, CCL3, CCL4, CCL5, IL-8 and CCR7 attract NK and T helper 1 cells to sites of inflammation (24). On the molecular level, various signaling molecules, transcription factors (TFs) and other transcriptional regulators (TRs) form the IFN γ / TNF- α macrophage-associated activation profile. Canonical IRF/STAT signaling pathways are activated by IFNs or TLR signaling via STAT1 (25). IRF5 successfully induces the expression of IL-12 and IL-23 (23). TLR4 stimulation of macrophages induces NF- κ B signaling subsequently expressing the SOCS3 protein, which can inhibit IL4 activated macrophage programs by the suppression of STAT3. Interestingly, NF- κ B signaling also activates a transcriptional program for the resolution of inflammation, demonstrating a negative feedback loop for the tight control of inflammatory processes (26,27). Thus, NF- κ B is not an exclusive TF for IFN γ / TNF- α associated processes, but also involved in anti-inflammatory functions. IFN γ / TNF- α associated macrophages phagocytose microorganisms and matrix debris, mainly important in early healing phases. In the disease background, IFN γ / TNF- α macrophages are amongst others associated with diminished metastasis and increased patient survival in colon carcinomas and are even able to kill tumor cells *in vitro* (28-30). However, in septic patients the production of pro-inflammatory IFN γ / TNF- α macrophage associated cytokines accompanies high mortality rates (31).

Macrophages with anti-inflammatory functions M^{L4} (formerly known as M2) are *in vitro* generated by IL-4/IL-13, immune complexes or by the stimulation with glucocorticoids. They express anti-inflammatory mediators like IL-4, IL-10 and TGF- β (32,33) and are characterized by low or no expression of IL-12 and the expression of surface receptors like CD23, MCR1 or scavenger receptors (7). On the molecular level, IL-4/IL-13 promotes the activation of canonical IRF/STAT signaling pathways via STAT6 (25), while IL-10 signals via STAT3 and allows the subsequent expression of genes like IL10, TGF β 1 or MRC1 (34,35). Other important promoters of IL4 activated macrophages are amongst others SOCS1, KLF4, PPAR γ or PPAR δ inhibiting the STAT1/NF- κ B associated IFN γ / TNF- α macrophage gene activation (36-39). It is important to note that differences between murine and human macrophages in their expression profiles were found, despite quite similar functions for example by the recruitment of the same cell types to sites of inflammation (40). Macrophages in murine models express Ym1, Fizz1, Arginase I and the chemokines CCL2 and CCL7 recruiting eosinophils and basophils, whereas human macrophages express the chemokines CCL13, 14 and 17. IL4 activated macrophages promote angiogenesis, tissue remodeling and repair (41). Additionally, increasing evidence suggests a major role for macrophages not only

in peritoneal fibrosis (42) but also in tumor formation with poor prognosis and disease progression in breast carcinoma (43,44).

The limitations of the M1/M2 nomenclature of macrophages is easily shown by macrophages from chronic inflamed tissues which belong neither to M1 nor to the M2 axis, since they display distinct transcriptomic changes in comparison to other activated macrophages (21). Chronic inflammation associated macrophages promote anti-inflammatory effects and strongly inhibit T cell proliferation. These macrophages are primed with GM-CSF and *in vitro* become activated by a combination of TNF- α , PGE₂, and P3C (TPP) and were therefore described as TPP stimulated macrophages (M^{TPP}). TNF- α , PGE₂ and TLR activation was linked to chronic granulomatous inflammation such as in granulomatous listeriosis or tuberculosis (45-47). These M^{TPP} macrophages express anti-inflammatory mediators like CD25, COX-2, IL10 and indoleamine 2,3-dioxygenase (IDO) which was also detected in human macrophages imbedded in human granulomatous structures (46,48). Furthermore, M^{TPP} macrophages highly expressed CD14, CD23, CXCR7 and CD197 in comparison to IFN γ or IL-4 activated macrophages and TFs like STAT4 or soluble effector molecules like IL-1 α were exclusively highly induced in M^{TPP} macrophages. Additional differences on miRNA expression levels in comparison to M^{IFN γ} or M^{IL4} activation further suggested the phenotypical and functional difference of M^{TPP} (21).

In general, markers for the differentiation between stimulus-activated macrophages are still a matter of debate. Modern computational approaches analyzing the transcriptome of differentially activated macrophages may profoundly change and extend the system of macrophage classification to mirror more and more the complex *in vivo* situation and allow a better understanding of these innate immune cells, responsible for many functions in immune defense, development and even disease.

1.2.1 PU.1 - A pioneer transcription factor in macrophage differentiation

TRs and particularly transcription factors (TFs) play an essential role in the establishment of transcriptional programs. During different phases of cell development, lineage commitment and activation different classes of TFs are involved. Pioneer TFs like FOXA1 or GATA family member proteins determine the cell lineage, are one of the first factors to be expressed and feed forward loops further increase their presence during early development (49-52). Polarizing TFs are expressed due to cell-extrinsic environmental signals and further fine tune the regulatory landscape, thus defining the cell type. Ultimately, effector TFs mediate the

activation into different functional cellular subtypes caused by transient changes in the microenvironment.

Pioneer TFs determine cell lineage and identity by the establishment of a core transcriptional program shaping the genome structure itself (53). They can bind in contrast to most other regulatory proteins DNA sequences even in a silent or native chromatin state and are crucial to make chromatin accessible for the binding of other TFs (54-56). The pivotal function of pioneer factors recruiting nucleosome remodeling and chromatin modifying complexes allows further adaption of the chromatin state at bound sites. Collectively, pioneer factors are proposed to act as global organizers, especially for the formation and maintenance of *cis*-regulatory elements like enhancers (57).

Enhancers are defined as DNA sequences influencing transcription in an orientation-independent manner with variable distances to corresponding genes (58,59). They are not only located at intergenic genomic regions but also at intronic sites (60,61) and can contain multiple accessible binding sites for TFs important in development and differentiation, controlling cell type- and tissue-specific gene expression (57,62,63). Furthermore, specific coactivator proteins like mediator complexes are located at enhancers (64-66). The mediator protein is associated with cohesion, is essential for the interaction of DNA bound TFs and the RNA polymerase II enzyme and acts as a bridge between distant regulatory sites and promoters mediating DNA looping (67). Furthermore, other important transcriptional coactivators like p300 or CBP histone acetyltransferases were also associated with enhancers (68-70).

The PU.1 pioneer TF, an essential regulator in the hematopoietic cell lineage, establishes the recruitment of other TFs to regulatory enhancer sites in murine macrophages (71-73). PU.1-deficient fibroblasts establish macrophage-specific nucleosome-depletion and gain open chromatin modifications going along with active transcription at putative macrophage specific promoter and enhancer sites upon PU.1 expression (71). Importantly, mice deficient for PU.1 do not only show alterations in B and T lymphocyte formation but moreover completely lack monocytes and macrophages (74,75). ChIP-seq binding experiments for PU.1 and other important TFs like C/EBP β or p300 correlated to different histone modification (HM) signals illustrates the chromatin remodeling capabilities of PU.1 in mouse macrophages and the gain and loss of specific enhancer marks during knockout experiments with simultaneous changes on nucleosome positioning (71,72). A study even intrinsically linked PU.1 binding and nucleosome positioning in macrophages (76). Non-active PU.1 bound sites in macrophages and PU.1 bound sites in cell types without PU.1 expression exhibit high nucleosomal occupancy. In contrast, in macrophages PU.1 binds a high fraction of putative active or poised enhancer sites (77). Interestingly, out of 600.000 predicted high affinity PU.1

binding sites only 10% are actually occupied by the PU.1 protein. These results were obtained from PU.1 ChIP-seq experiments performed with several human cell lines like Jurkat or HL60 cells. Obviously, another higher level of regulation seems to select actual PU.1 binding sites. A functional cooperation between pioneer TFs and secondary pioneer TFs or a direct DNA interaction by forming 3D structures might also be possible (78). Partner TFs could help access specific target sites or allow the pioneer TF to occupy DNA binding sites with suboptimal affinities, demonstrated by the interaction of PU.1 with IRF TFs binding ETS/IRF composite sites (79,80). Furthermore, partner TFs could additionally remodel chromatin states of pioneer factor bound sites by the recruitment of nucleosome remodeling enzyme complexes (81). Another issue could be the differential expression rates for PU.1 between cell types. Lower PU.1 expression in hematopoietic progenitors induced a B cell-like phenotype, while higher PU.1 expression leads to a reprogramming into macrophages (75). PU.1 expression itself is regulated by many factors. On the transcriptional level, high IRF8 or PAX5 expression can down-regulate PU.1 transcription and promote B cell formation (82,83). A prolonged cell cycle is another mechanism to enhance PU.1 expression in macrophages in contrast to B cells, where shorter cell cycles and decreased Spi1 gene transcription leads to less PU.1 protein (84).

Pivotal for transcriptionally active genomic sites are open chromatin positions without dense heterochromatin structures and only sparse nucleosome densities to enable the binding of regulatory factors, important parts of the transcriptional machinery. This process is amongst others mediated by pioneer TFs and only one aspect of cellular “epigenetic” regulation.

1.3 Epigenetics

The former meaning of “epigenetics” combined the words “epigenesis” and “genetics” to describe processes of cell fate commitment and lineage specification during development (85,86). Today, the term “epigenome” describes DNA sequence independent processes that modulate gene expression, amongst others DNA methylation, posttranslational modification of histone proteins or RNA-based mechanisms (87,88). By organizing the chromosomal conformation in the nucleus, chromatin is organized into sites with variable transcriptional activity by modulating the accessibility for the binding of TRs like TFs or RNA polymerases. Most importantly, epigenetics does not only augment the cellular genomic sequence with an additional dimension, thus increasing complexity in an order of magnitude, but also preserves memory on different levels particularly of past transcriptional events (89-91). The cellular epigenome is intrinsically linked to the structural organization of the genome and reflects the global patterns of transcriptional activity (92).

1.3.1 DNA methylation

The methylation of cellular DNA represents one major aspect of epigenetic regulation. A high fraction of the human genome is methylated (60-80%) typically at CpG dinucleotides by enzymes called DNA methyltransferases (93). DNA methylation is essential for developmental processes and the depletion of DNA methyltransferases leads to neonatal lethality in mice (94). Previous experiments suggest enrichments of DNA methylation especially at sites with low transcriptional activity. Next generation sequencing technology revealed that methylation marks are less stable than anticipated, and can also exist at non-CpG sites and are partially enriched at actively transcribed genes (95). Furthermore, DNA methylation is tightly connected to histone modifications and involved in the regulation of transcriptional processes (96).

1.3.2 DNA structure

Another level of epigenetic regulation addresses the basic structural subunit of DNA called “nucleosome”. Exactly 147 bp of DNA are wrapped around the nucleosome. Nucleosome subunits are connected by linker DNA fragments. ATP-dependent chromatin-remodeling enzymes act by sliding along the DNA, triggering conformational changes of the nucleosomes that enable subsequent positional changes (97). The nucleosomal positioning inhibits or promotes the binding of regulatory proteins by the occupation of sequence specific sites or inhibits the activity of RNA polymerase II (98). Histone tails of nucleosomes can also serve as binding sites for TRs and associated proteins. Furthermore, the genomic position of nucleosomes is not a random process and occurs more frequently at specific regions of DNA, depicting the active role of nucleosome positioning in regulatory processes (99). In addition, nucleosomal DNA may also support synergistic TF binding leading more efficiently to the displacement of nucleosomes in comparison to the presence of only one TF (100).

The nucleosome structure is only one level of cellular architecture in eukaryotes. Higher order 3D structures of chromatin like the 30 nm fiber of packed nucleosomes, stabilized by the H1 histone protein or 700 nm condensed sections of chromosomes, enable distant loci on the linear DNA to interact with each other and change transcriptional activity of adjacent genes (101). For example, distant enhancers with hundreds of kilobases distance can increase beta-globin gene transcription by the interaction with its promoter established by chromatin looping (102). In addition, chromatin close to the lamina of the inner nuclear membrane seems to exist in an inactive heterochromatic state (103). This further

demonstrates the importance of chromatin 3D structure onto the regulation of transcriptional activity for eukaryotes genes.

1.3.3 RNA mediated processes

Besides the structural chromatin regulation, a complex system of different RNA mediated epigenetic mechanisms further ensures the tight control of cellular transcription e.g. for the regulation of development (104), environmental stress (105) and disease relevant processes (106). The RNA polymerase II enzyme is a major subject of regulation mediating the transcription of eukaryotic DNA into messenger RNA precursors (mRNA) or different non-coding RNA species like infrastructural (ribosomal RNA, small nuclear RNA) or regulatory RNA (microRNA, long non-coding RNA and others) (107). The RNA polymerase II enzyme consists of 12 protein subunits (RPBs) highly interacting with each other and mediating different functions like DNA binding, stabilization and regulation of transcription (108). Importantly, the RPB1 protein possesses a specialized region called CTD (carboxy terminal domain) with up to 52 tandem heptapeptide repeats in vertebrates and is subject of extensive posttranslational modification (109,110). These modifications define complex patterns culminating in a CTD code tightly correlating the state of CTD modification to the actual function and activity of RNA polymerase II (111). For example, antibodies in macrophage ChIP-seq experiments targeting phosphorylated CTDs at the serine 5 residue recognized RNA polymerase II enzymes during their initiation actively transcribing DNA into RNA species (112), while subsequent phosphorylations at the serine 2 occur later in the transcriptional process, during elongation (113). After reading through the transcribed gene a phosphatase enzyme removes the phosphorylation when arriving at the 3' end of the protein coding region. The presence and modification status of RNA polymerase II enzymes allows the assessment of transcriptional regulation and activity at important genomic sites. Transcription does not only occur at protein coding genomic sites, but also in large amounts outside of these regions and can lead to different classes of extragenic transcripts (112,114,115).

Long non-coding RNAs (lncRNA) are defined by their transcript length of 200 nucleotides, their nuclear localization in very low amounts and their sequence conservation (116). They can originate from various intergenic genomic locations but were also found at intronic sites (117). lncRNAs are involved amongst others in functions like the regulation of allelic expression (118) or in developmental processes (119,120). Even disease associated functions were identified mainly in cancer related processes (121,122), whereas the analysis of the functional relevance for the majority of identified lncRNAs is still far from complete (123). Many possible mechanisms how lncRNAs regulate their targets include for example

the interaction with chromatin-modifying complexes (124), the modulation of DNA methylation (125) or direct effects on transcription by competing with TF binding (126).

The relatively new concept of enhancer RNA (eRNA), a special type of lncRNAs, transcribed at active *cis*-regulatory sites arises the question, if enhancers do not only influence transcription by their direct interaction with TFs bound to promoters forming DNA loops, but also regulate transcription by RNA-mediated processes (112,127,128). These highly unstable transcripts are produced either unidirectional as short (< 2 kb) and non-polyadenylated RNAs (1d-eRNA) or bidirectional in 5' to 3' and 3' to 5' direction as longer (> 4 kb) and polyadenylated eRNA species (2d-eRNA) (128,129). Three functional consequences of enhancer transcription are supported by an increasing amount of scientific data. First, random collisions of RNA polymerase II with accessible enhancer sites could cause eRNA transcription without functional consequences, thus defined as noise (130). Second, the process of eRNA transcription can alter the accessibility of the chromatin template due to the RNA polymerase II movement and its enzymatic activities altering the chromatin structure (131,132). Third, transcripts can cause functional changes by direct *in cis* or *in trans* interactions with DNA recruiting protein complexes (e.g. chromatin remodelers) or effector proteins (133,134).

Another class of extragenic RNA transcripts with lengths of approximately 22 nucleotides were defined as microRNAs (miRNAs). RNA polymerase II transcribes primary miRNA, which is subsequently processed by the Drosha and DGCR8 protein complex into a hairpin structure called precursor miRNA (135-138). Ran-GTP and exportin 5 transports the precursor miRNA into the cytoplasm, where the Dicer cleavage produces mature double-stranded miRNA. The passenger strand is degraded, while the guide strand is integrated into the RISC complex (RNA-induced silencing complex) and specifically targets mRNA transcripts, resulting in silencing of the corresponding target gene (139-142).

Amongst already specified epigenetic mechanisms, the posttranslational modification of histone proteins, also called histone modification (HM), plays a key role for the establishment and maintenance of cellular transcriptional programs.

1.3.4 Histone modifications

The fundamental repeated subunit of chromatin called “nucleosome” consists of chromosomal DNA wrapped around a histone octamer with protein subunits H2A, H2B, H3 and H4. Posttranslational modifications at the N-terminal tails of histone proteins and the existence of histone isoforms display a high possible variability for nucleosomes. Especially

covalent modifications like methylation, acetylation and phosphorylation are well studied examples of over 130 possible posttranslational HMs (Figure 1) (143,144). Different classes of chromatin-modifying enzymes accomplish the posttranslational modification of histone tails (97). The “histone code” hypothesis correlates an interaction of posttranslational HMs read by effector proteins to functions like DNA replication, recombination and transcription (145,146). For example, modified histone tails can form binding sites for specific classes of regulatory molecules. Proteins containing bromodomains bind to histone tails with acetylated lysine residues, while chromodomain-containing proteins bind to methylated histone lysine residues. These histone binding proteins could link covalently modified histone tails to transcriptional regulation (147). Posttranslational HMs play also a major role in developmental processes and disease, e.g. in cancer progression.

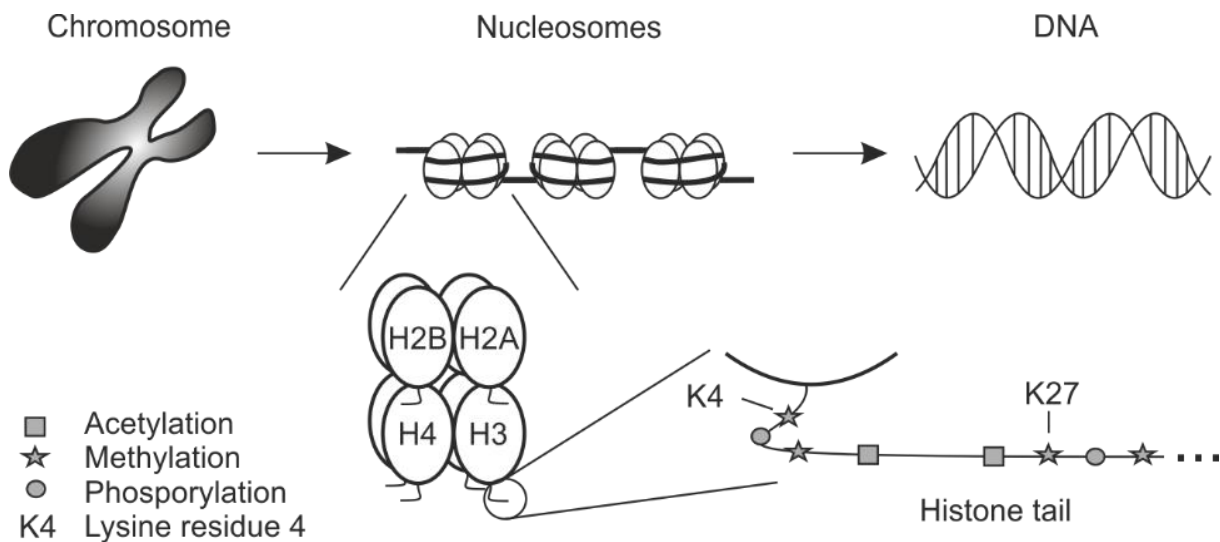


Figure 1. Schema of DNA/chromatin structure and site of histone modification

DNA consists of basic structural units called nucleosomes and is further packed into higher chromatin structures and may form chromosomes. The four histone protein subunits (H2A, H2B, H3, H4) of the nucleosome octamer can be target of covalent HMs. Various modifications like acetylation, methylation, phosphorylation or others can be located at specific N-terminal amino acid residues e.g. on lysine 4 (K4) or lysine 27 (K27).

In fact, several HMs could be linked to functional relevant regulatory sites in the genome and moreover assess their transcriptional activity. The trimethylation of histone H3 on lysine residue 4 (H3K4me3) was found to be mainly enriched at promoter sites in yeast, mouse and human cells, where transcriptional initiation takes place and important TRs like RNA polymerase II or TFs can bind (148-150). The simultaneous methylation of lysine residue 27 of histone protein H3 (H3K27me3) mediated by the polycomp repressor protein complex and of lysine 4 (H3K4me3) allowed the identification of bivalently marked poised promoters,

silenced in their transcriptional activity and important for a proper and robust differentiation (151,152). The polycomb repressor protein family can form different multiprotein complexes e.g. with co-repressor BCL6 or E2F6 proteins, which modify histone or other proteins to silence transcription (153-155). In contrast, H3K4me3 enriched sites with a simultaneous acetylation of histone protein H3 on lysine residue 27 (H3K27Ac) strongly correlated with accessible and active promoters at cell type specific promoters, important for the induction and maintenance of cellular transcriptional programs (156). Enhancers, another class of regulatory sites can further promote gene transcription and are located up- or downstream of transcriptional start sites (TSS) with variable distances up to 1 Mb (157). Different states of enhancers were previously defined with specific HMs and correlated to the transcriptional activity of adjacent genes. First of all, H3K4me1 (monomethylation of lysine residue 4 of histone protein H3) a general histone enhancer mark is solely located at weak enhancers regulating transcription of moderately active genes (158,159). Active and poised enhancer states can be discriminated with additional H3K27Ac or H3K27me3 HMs (68,160,161). Active enhancers increase transcriptional rates by the interaction of bound TRs with promoter sites or with non-coding enhancer RNA species (eRNAs) stabilizing these interactions (102,162). Poised enhancers, previously in an active state during development, were silenced in differentiated cells and correlate well with inactive genes. These combinations of HMs were established in various studies as standard strategy for the genome wide identification of promoter and enhancer states (163-166).

In summary, histone modifications can control transcription in an activating or repressive manner by changing the DNA accessibility and serve as targets for TRs (167). Additionally, genome wide epigenetic maps allow the delineation of unknown genes encoding for example for regulatory RNA species or new functional *cis*-regulatory sites.

1.3.5 Super enhancers – A distinct class of *cis*-regulatory sites

A special type of *cis*-regulatory regions were described recently amongst others in embryonic stem cells (168) and murine tissue macrophages (169). These super enhancer (SE) sites are not only characterized by a strong enrichment for active regulatory marks and TFs but also by a monomethylation of the histone protein 3 at the lysine 4 residue (H3K4me1) (170,171). Different approaches were used for their identification based on either mediator coactivator complex enrichments, essential for embryonic development and maintenance of the ESC state (67), or on the HM H3K27Ac present at active regulatory sites (169). Single enhancer positions with exceptionally high H3K27Ac signals and with distances less than 12.5 kb to each other are integrated into large SE regions. Due to the broad nature of SEs found at

genomic regions spanning tens of thousands of bases, new algorithms for their analysis were established and integrated into present ChIP-seq data analyzers like the HOMER program (72). These genomic regions were found at prominent genes encoding for key regulators required for cell identity in different cell types like embryonic stem cells or murine T-cells (172,173). In addition, SE also seem to be associated with oncogenic driver genes influencing the development and progression of cancer (171). Furthermore, at SE regions TRs are cooperatively bound in high numbers in comparison to standard enhancers. Due to this synergistic effects of TF binding, SEs can lose activity more rapidly with lower levels of regulators bound to the SE locus (174,175). The higher sensitivity of SEs for perturbation than typical enhancers, by the loss of key TFs like BRD4, may promote a targeted down-regulation of adjacent genes and can represent an efficient mechanism to profoundly influence oncogenes. SEs play also a major role for activation-induced cytidine deaminase (AID) processes, responsible for the initiation of somatic hypermutations during antibody maturation and DNA breakage during antibody class switch recombinations (176). As an integral part of the epigenetic landscape, SEs regulate diverse cellular functions in many cell types. Additionally, they may also reflect the extracellular environment and their formation is influenced by extracellular signals, demonstrated in heterogeneous murine tissue macrophages showing common macrophage SEs but also activation state specific sites (169), whereas the role for SEs during the establishment of different activity states in human macrophages still needs to be elucidated.

1.3.6 Epigenetic regulation in macrophages

Macrophages participate in a wide spectrum of biological processes and diseases, like neurological dysfunctions or cancer and are capable of rapid and dramatic alterations in their transcriptional programs adapting to environmental stimuli (92). Interestingly, despite the same genome and the expression of nearly the same repertoire of TFs, macrophages differentiate under various conditions and fulfill various functions after their specialization into distinct activation states (177-179). Such specialized programs seem to be based on specific combinations of a small number of expressed and repressed gene subsets. Epigenetic changes could hold the key to balance these “substantial similarities” with “specialized differences” mediating transcriptome variations, which can change cellular functions and identity. If epigenetic mechanisms alone or in combination with secondary control mechanisms at accessible loci are responsible for certain macrophage activation states is still unknown.

Different aspects of complex epigenetic regulatory mechanisms were already subject of detailed scientific investigations. First of all, *cis*-regulatory enhancer regions are established and maintained in macrophages mainly by the binding of the lineage determining TF PU.1, essential for the macrophage cell type (75,178). Between 35.000 and 45.000 enhancers could be identified in murine macrophages (71,72). They are used by signal depending TFs to integrate a diverse array of signals with gene transcription. TFs like AP-1 or C/EBP strongly correlate with PU.1 bound sites in thioglycolate-elicited mouse macrophages and cooperatively displace nucleosome barriers, prior to the establishment of functional enhancers (76,180). The identity of TFs cooperating with PU.1 to mediate development or survival of macrophages depends on the microenvironment and corresponding macrophage function. Tissue macrophages for example display slightly different TF repertoires. Whereas microglia located in the central nervous system depend on the TFs IRF8 and SMAD mediating TGF- β signaling (181-183), peritoneal macrophages require GATA6 for their survival (184). However, different murine spleen macrophage subtypes seem to depend on the ETS family TF SPI-C or LXR α (13,185).

Transcripts of active enhancer sites (eRNA) may also play an important role in the direct regulation of the macrophage transcriptome. For instance, Rev-Erb nuclear receptors repress the transcription of MMP9 and CX3CR1 genes in mouse macrophages by the direct binding of corresponding enhancer sites and mediate the down-regulation of eRNA transcripts adjacent to these genes (186). Other regulatory molecules like miRNA-155 can down-regulate M^{L4} genes in human macrophages by targeting the IL-13R α 1 subunit (187). Additionally, identified genomic SE regions in murine macrophages are concomitantly enriched for macrophage key TFs like C/EBP and seem to regulate genes responsible for cell identity and macrophage specific functions (168).

Furthermore, recently identified latent or *de novo* enhancers in mouse macrophages gain enhancer associated HMs (H3K4me1) upon macrophage activation and allow the binding of regulatory proteins like TFs (73,188). TLR4 stimulation leads to NF- κ B activation and finally results in PU.1/C/EBP binding, which recruits NF- κ B p65 to the enhancer site. Histone acetyl transferase enzymes (HATs) cause the acetylation of histones and the recruitment of RNA Polymerase II mediating in at least some cases the transcription of eRNAs. Histone methyl transferases (HMTs) MLL1-4 finally lead to the stabilization of the enhancer characterized by the methylation of lysine 4 on histone protein 3. Importantly, latent enhancers keep their corresponding H3K4me1 HM signal after loss of the stimulus and upon re-stimulation, a faster and stronger induction of macrophages could be detected, thus providing an epigenetic memory of the initial stimulation.

The histone modification enzyme (HME) repertoire itself can identify different activation states of macrophages and are associated with specific macrophage functions. MLL enzymes do not only mediate the *de novo* formation of latent enhancers, but are also known to increase the expression of CXCL10 in human and murine macrophages after IFN γ /LPS stimulation. CXCL10, an important chemokine in M^{IFN γ} macrophages attracts T helper 1 cells during inflammatory processes (7,189). Other HMEs, especially JMJD3 (Jumonji Domain Containing 3) and HDAC3 (histone deacetylase 3) play central roles in macrophage differentiation and activation (190). JMJD3 is not exclusively associated to only one macrophage population, is essential for the transcription of pro-inflammatory human and murine macrophage genes and is crucial for the activation of IL4 macrophages in helminth infections and responses to chitin (191-193). Both demethylases JMJD3 and KIAA1718 cooperate to resolve H3K27me3 repressive marks at genes poised for activation, allow RNA Polymerase II travelling and the activation of gene transcription (194). Overall, JMJD3 allows the response to different external stimuli in human and murine macrophages. In contrast, HDAC3 inhibits the activation of anti-inflammatory IL4 macrophages by the de-acetylation of IL-4 induced regulatory regions and promotes M^{IFN γ} pro-inflammatory responses in murine models (195). In HDAC3 deficient phenotypes, macrophages show M^{IL4}-like properties in the absence of external stimuli and are hyper responsive to IL-4. Furthermore, HDAC3 is required for the activation of STAT1-dependent inflammatory genes, assessed by defects in IFN β signaling in HDAC3 knockout mouse models (196). The expression and activity of HMEs is affected by environmental stimuli and requires the utilization of metabolites. The metabolism of M^{IFN γ} switches to glycolysis, while M^{IL4} macrophages perform enhanced fatty acid oxidation and mitochondrial oxidative phosphorylation (197-199). Interestingly, the inhibition of these metabolic pathways can impair M^{IFN γ} or M^{IL4} activation. Collectively, HMEs remodel the epigenetic landscape of macrophages upon environmental stimuli, thus may play an important role as linkers between environment, metabolism and macrophage activation (200).

How far in detail epigenetic processes are involved in the alteration and maintenance of transcriptional programs, specifically during the generation of primary human macrophage subsets still needs to be elucidated.

1.4 Next generation sequencing technology

Next generation sequencing (NGS) technology, an essential prerequisite for the modern high throughput epigenetic analysis of living cells, was developed in the early 2000s to combine the principle of DNA sequencing technology with a high throughput approach to massively

parallel sequence short DNA reads using solid phase sequencing by reversible terminators. In contrast to the ChIP-chip array technology, where ChIP-DNA fragments were identified by the hybridization to a microarray, in ChIP-seq experiments the ChIP-DNA is directly sequenced with higher resolution, fewer artefacts and greater coverage in a massively parallel fashion (201). The exponential increase of deep sequencing data set numbers generated with quite variable techniques required the establishment of standardized rules and optimized protocols to enable the integration and comparison of different data sets and to increase their overall accuracy and quality. For this purpose, the ENCODE and Roadmap Epigenomics Projects (202-204) did not only release tight quality protocols to ensure high quality sequencing data sets but also contributed gigantic data repositories to distribute thousands of NGS experiments to literally everyone. This approach allows the integration of data on regulatory processes to generate piece by piece a broader picture of the cellular regulatory landscape. Especially for cell line data sets this is already feasible, due to their highly abundant data availability, but it is also more and more possible for primary cells finally correlating epigenetic, transcriptional and translational data.

1.4.1 Next generation sequencing method variations

High throughput sequencing experiments were initially designed to identify enriched genomic regions for a protein of interest (ChIP-seq). The creative usage of digestion enzymes or different separation protocols for the isolation of specific RNA or DNA molecules multiplied the potential areas of application. Many aspects of cellular regulation can now be subject of detailed high throughput experiments, like epigenetics or transcriptional regulation.

Three NGS methods are specialized on the analysis of cellular DNA methylation and utilize different working principles (205,206). (I) In MRE-seq experiments (Methylation-sensitive Restriction Enzyme sequencing) methyl-sensitive restriction enzymes digest genomic DNA and finally lead to the identification of CpG methylated genomic sites. A disadvantage of this method is its incomplete enzymatic digestion that can lead to less accuracy and potentially biased sequencing results. Other more cost efficient approaches use affinity-based enrichment assays to select for methylated fragments in sonicated chromatin with an antibody or a methyl-binding domain. These sequencing methods are called (II) MeDIP-seq (Methylated DNA Immunoprecipitation sequencing) or MBD-seq (Methyl-CpG Binding Domain sequencing) (207,208). Other approaches like (III) Bisulfite sequencing experiments use a chemical conversion step to distinguish methylated from unmethylated cytosines (MethylC-seq) (209). After the discovery of four cytosine variants including intermediate states (210,211) classical MethylC-seq experiments were adapted due to their inability to

distinguish between the four cytosines (212). Enhanced OxBS-seq (Oxidative Bisulfite sequencing) or TAB-seq (TET Assisted Bisulfite sequencing) protocols can now discriminate between cytosine variants and produce accurate single-base resolution maps, despite concerns like DNA damage due to chemical treatments and their dependency on the enzymatic activity of TET (213,214)

The genome in eukaryotes is structured into basic nucleosome units and eventually forms chromosomes through a series of higher order structures. Genome-wide nucleosome positioning can be assessed by the enzymatic micrococcal nuclease digestion of chromatin (MNase-seq) which is afterwards used as input for following NGS. DNA wrapped around nucleosomes or occupied by TFs is protected from digestion and - after a purification step and the sequencing of 150 bp nucleosomal DNA segments - allows the identification of nucleosome protected regions. Genomic loci vary in their accessibility due to their packaging status which is mediated amongst others by nucleosomes and therefore promote or interfere with their capabilities to be bound by proteins. Transcriptionally active regulatory regions like promoters, enhancers, insulators and silencers are tightly correlated to open chromatin positions and are targets of specific transcriptional regulatory proteins. In DNase-seq (DNase I hypersensitive sites sequencing) experiments open regulatory regions are identified due to their vulnerability to DNase I enzymatic digestions. These experiments finally result after DNase I enzymatic treatment in nucleosome depleted open chromatin fragments (215). FAIRE-seq (Formaldehyde-assisted isolation of Regulatory elements), another approach to analyze chromatin accessibility uses formaldehyde cross-linked chromatin as input and an ultrasound mediated DNA shearing is performed (216). Open regulatory sites are more sensitive to ultrasonic shearing and thus can be sequenced after DNA purification. Another approach uses transposase enzymes (ATAC-seq), which shear open chromatin regions and integrate primer sequences into the cleaved genomic DNA (217). The biggest advantage for the analysis of nucleosome-free and open positions is the global identification of regulatory elements and protein binding sites without prior knowledge of TRs. Interestingly, specific nucleosome-free regions (NFR) e.g. at promoter sites can also be identified by a bioinformatics approach using HM data sets like H3K4me3 (tri-methylation of lysine 4 on histone protein 3) independently of previously listed methods with tools like the HOMER program to analyze the specific HM enrichment patterns (72,218).

For mapping higher order chromatin architectures, a variety of protocols was established. E.g. in all 3C (Chromosome Conformation Capture) methods formaldehyde cross-linked chromatin is digested with a restriction enzyme. An intramolecular ligation of DNA ends in close proximity to each other further allows the identification of ligation frequencies between two restriction fragments (one-to-one) and is finally measured by qPCR with specific primer

sequences (219). Variations of the 3C protocol measure the genome wide interaction frequency of one target site with 4C-seq experiments (Circular Chromosome Conformation Capture sequencing) (one-to-all), while many anchor and bait primers can identify thousands of interactions between targets in the 5C-seq method (Chromosome Conformation Capture Carbon Copy sequencing) (many-to-many) (220,221). Hi-C sequencing, the newest breakthrough in chromatin conformation measurements can identify possible interaction frequencies from an entire genome with itself in all combinations (all-to-all) (222). Chimeric DNA ligation junctions are selected with a biotin labeled nucleotide incorporated at the ligation junction and are subsequently used for NGS. The ChIA-PET (Chromatin Interaction Analysis by Paired-End Tag sequencing) protocol even combines the Hi-C method with the analysis of proteins like TFs or RNA polymerases interacting with the nuclear organization (223). Interestingly, Hi-C experiments identified cell type and species specific 1 Mb regions with high local interaction frequencies separated by boundary regions and were termed “topologically associated domains” (TADs) (224). TADS consist of regulatory elements like promoters or enhancers interacting with each other by DNA looping and can even separate chromosomes into higher structured compartments called chromosome territories (222).

NGS methods were also extended to the analysis of RNA molecules (225) to quantify known and unknown transcripts with corresponding splice junctions (226,227), alternative splicing (228,229) and even single nucleotide polymorphisms (SNPs) (230). Transcriptome analysis by massively parallel sequencing outperforms previous techniques like Sanger sequencing or microarray-based methods in resolution and accuracy. In RNA-seq experiments cellular total RNA, mRNA with poly-A tails, even micro RNA or other small non-coding RNA species (miRNA-seq) serve as input material for the complementary cDNA construction followed by RNA library preparation and high throughput sequencing. Standard RNA-seq protocols do not preserve the information about which strand was originally transcribed. Newer strand specific RNA-seq protocols can discriminate between RNA originated from + or – strand DNA (231). This enhanced RNA-seq protocol enables the accurate identification of regulatory antisense transcripts (232) or the correct measurement of expression levels for coding and non-coding transcripts located at overlapping genomic positions. Derivatives of RNA-seq experiments also allow the analysis of nascent RNA products measuring rates of transcription instead of steady state RNA levels (GRO-seq, Global Run-On sequencing) (232). During a chemically induced transcriptional pause, labeled nucleotides are added to isolated nuclei and transcription is briefly resumed. Therefore, the quantity and location of newly transcribed RNA molecules can be identified by sequencing the labeled RNA products. Furthermore, RNA-seq protocols were adapted to closely monitor translational processes (Ribo-seq, Ribosomal sequencing) (233). Only ribosome protected mRNA is isolated and used for sequencing experiments to detect global ribosomal densities along mRNA

transcripts (234) and translational start sites (235). RNA-seq experiments are the ideal tool to elucidate transcriptional and translational processes on a global cellular scale. Nevertheless, challenging bioinformatics data analysis requires correction and normalization steps to correctly transform experimental data into a profound biological interpretation (236).

Chromatin immunoprecipitation following deep sequencing (ChIP-seq) identifies genome wide binding patterns of proteins directly or indirectly bound to chromatin or covalent posttranslational modifications of histone proteins. In this method, formaldehyde cross-linked chromatin is sheared and DNA-protein complexes are immunoprecipitated with specific antibodies against a protein of interest. After the purification of ChIP-DNA, known adapter sequences are ligated to both ends and PCR reactions multiply the DNA amounts. Finally, the ChIP-seq library is sequenced and resulting sequence reads are computationally aligned to a reference genome for the corresponding species to find peak positions enriched for the protein of interest. Antibody quality plays an exceptionally high role for valid and accurate sequencing results and greatly influences unspecific and background signal intensities (204). The amount of starting material can also impact sequencing results. Especially for primary cells ChIP-seq experiments can be challenging and the success is not guaranteed while using less than 5 million cells for TF associated experiments and less than 1 million cells for detecting HM enrichments. Low cell protocols like ChIP-nano were adapted to overcome at least partially the cell number requirement hurdle using 50.000 cells or less by reducing sample loss, adapting shearing conditions and washing steps (237,238). Another protocol (ChIP-exo, exonuclease sequencing) increases footprinting resolution to nucleotide resolution by the 5'-to-3' exonuclease digestion of sonicated ChIP DNA ends up to the cross-linked protein. In general, the ChIP-seq method, can identify the genomic positions of histones and HMs, as well as TFs, DNA or histone modifying enzymes and other chromatin associated proteins. Careful experimental design and data analysis is required to minimize false-positive and background signals negatively affecting the correct identification of sites enriched by a protein of interest in ChIP-seq experiments.

1.4.2 Technical considerations for ChIP-seq experiments

ChIP-seq experiments can provide detailed genomic data on many regulatory levels in living cells and their design depends on the scientific question and chosen experimental protocol. Nevertheless, due to the complexity of experimental design, sequencing and data analysis multiple sources of variance need to be considered to allow a precise and correct biological data interpretation. First of all, the chromatin structure itself influences the performance of shearing experiments creating sites of variable fragility which depends on the hetero- or

euchromatin status, thus rendering euchromatic DNA more vulnerable to shearing (239). Under consistent shearing conditions using a highly efficient and standardized ultrasonic shearing device e.g. from Covaris with the same shearing protocols for target and control samples, the effects of this hurdle can be sufficiently abrogated. Incomplete protein digestion with the proteinase K enzyme could lead to biases due to the differential solubility of nucleic acids and proteins during phenol-chloroform extraction, which is still a widely used method during ChIP-seq experiments (240). Moreover, PCR over-amplifications due to inefficient amounts of immunoprecipitated ChIP-DNA is especially a problem while using low-cell protocols and can further create false-positive signals (241).

The choice between two different ChIP-seq approaches can additionally introduce variance. The X-Chip method uses formaldehyde to tightly fix proteins to corresponding genomic positions, which allows the study of non-histone proteins directly or even indirectly bound to DNA (242,243). A possible disadvantage is the inefficient antibody mediated precipitation due to potential masking of antibody epitopes during fixation and may require higher numbers of PCR cycles for sufficient DNA amplification. One positive aspect of X-ChIP experiments is the minimal chromatin rearrangement during ChIP-seq procedures. However, the N-ChIP method uses chromatin in its natural state without a fixation step and is specifically designed for proteins with strong DNA interactions e.g. histone proteins (244). In this protocol the chromatin shearing can be for example performed by an enzymatic digestion using the MNase enzyme (Micrococcal nuclease) (245). N-ChIP is normally not suitable for non-histone proteins, shows more pronounced sequencing bias than sonication and protein rearrangements could occur during chromatin preparation. (246). An over-digestion of chromatin can be controlled by thoroughly assessing optimal temperature and duration of MNase treatments. Positive aspects of N-ChIP experiments are e.g. the increased mapping resolution after MNase digestion by removing linker DNA more efficiently compared to an ultrasound mediated shearing. Furthermore, the generation of specific antibodies is normally performed with not-cross-linked peptides or proteins and is therefore more predictable for natural chromatin and can lead to a very efficient immunoprecipitation.

Mentioned examples of technical and biological variances create different kinds of false-positive ChIP-seq tag (short sequence reads) signals and need to be clearly discriminated during peak calling from biological relevant and specific ChIP-seq peak data (Figure 2). The antibody performance is critically important for successful ChIP-seq experiments and methods like immunoblotting or ChIP-string can assess their specificity and affinity (204,247,248). The amount of background noise and non-specific peak signals are inevitably connected to antibody quality. They can bind non-specifically with the constant FC region to DNA and beads or bind specifically by a cross-reaction to irrelevant epitopes introduced

during immunization. Especially for experiments with low enrichments of TF proteins the proper identification of significant peaks remains difficult (249). Another type of false-positive sites called “hyper ChIPable” regions were recently described in yeast and are characterized by huge amounts of non-specific ChIP enrichments for proteins like RNA polymerase II or various TFs, irrespective of protein origin (250). Even artificial proteins like GFP were bound to these regions in high amounts. These transcriptional active euchromatic sites are also vulnerable to DNase I cleavage. Hyper-ChIPable regions seem to display a general characteristic of the ChIP procedure per se. DNA from highly transcribed and open regions without a densely packed nucleosomal structure could promote stronger electrostatic interactions with antibodies or beads used during ChIP experiments. The ChIP antibody could also non-specifically interact with RNA polymerase II or III bound at these sites.

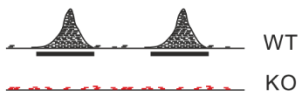
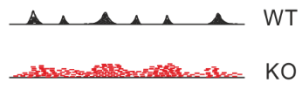
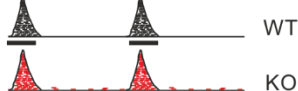
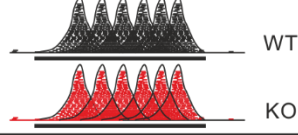
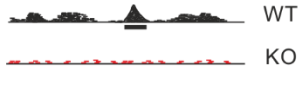
Variations in ChIP-seq data	Source of variance	Categories of signals	Peak calling
biological variance: • biologically relevant & specific peaks	experimentally induced variance: • stimulus induced • disease induced • cell-type associated signals	True positive	 WT KO
background signals	• unspecific epitope binding of ChIP Ab	False positive	 WT KO
non-specific peaks: • peaks identified in WT & KO samples	• specific epitope binding of ChIP Ab to other protein most likely present during immunization	False positive	 WT KO
Hyper ChIP-able regions	• open chromatin with protein complexes • strong protein cross-linking	False positive	 WT KO
other technical variances	• insufficient shearing • strong cross-linking • sequencing artifacts	False positive	 WT KO

Figure 2. Schematic overview of variances influencing ChIP-seq data files.

Sources for true- or false-positive ChIP-seq signals with corresponding graphical schema of raw tag signals and called peak positions marked with black bars. Figure has been previously published as part of manuscript “Optimization of transcription factor binding map accuracy utilizing knockout-mouse models” by myself as first shared author (Krebs, W.*, Schmidt, S.V.*, et al.).

True-positive and biological relevant ChIP-seq signals can be identified and separated from false-positive signals like background, non-specific and hyper-ChIPable signals or signals created by other technical variances, by the usage of ChIP-seq control data sets (204,251). For the first type of possible control sheared “input” chromatin is used, which cannot display the ChIP enrichments for specific binding sites of a protein of interest. Input chromatin allows an estimation of chromatin distribution and bias, but lacks the integration of variances occurring during the chromatin immunoprecipitation (ChIP). In a “mock” ChIP reaction, also called “IgG” control, an unspecific control antibody with the same immunoglobulin class than the target ChIP antibody reacts with irrelevant antigens on sheared chromatin. IgG controls mimic a ChIP experiment more closely than Input DNA but could recover not enough DNA amounts to create ChIP-seq libraries with sufficient complexity, thus introducing significant amounts of bias. A third type of control utilizing protein knockout (KO) experiments was until now only sparsely used in ChIP-seq experiments, combines the positive effects of input and IgG controls and simultaneously compensates for some of their negative properties. KO controls undergo the same experimental procedures as the ChIP-seq sample, thus representing the same bias. KO associated immunoprecipitations also normally result in sufficient amounts of ChIP DNA in comparison to the IgG control to produce control libraries with enough complexity to finally discriminate true- and false-positive sites without introducing additional bias. More importantly, KO controls could identify hyper-ChIPable regions in contrast to input controls and could exclude these highly expressed false positive sites from the analysis (250).

2. Objectives

The next generation sequencing technology, a very useful but complex method to gain insights into a plethora of scientific biological questions evolved in an astonishing way the last couple of years concerning data analysis strategies as well as in experimental methods. Thus, specific guidelines were recently published to ensure the comparability and reproducibility of ChIP-seq experiments (204). The ENCODE consortium did not specifically focus on knockout (KO) models and sparsely performed previous experiments concentrated more on the aspect of induced biological changes after the KO of a protein of interest in comparison to the WT situation. Therefore, it was speculated that KO experiments could be used as ChIP-seq control experiments, especially due to the known disadvantages of described IgG and Input chromatin controls to improve ChIP-seq data analysis (251).

These observations raised several questions: Can the standard ChIP-seq data analysis benefit from knockout controls used in the KO Implemented Normalization method (KOIN) method? In which extent influences KOIN false-positive signal ratios and more importantly the biological interpretation of ChIP-seq data? These questions were addressed with the establishment of KOIN utilizing six ChIP-seq data sets from mice, either generated with different antibodies or under different stimulatory conditions and with the assessment of the results in comparison to the standard ChIP-seq analysis.

In addition, in respect to the newly described spectrum model of human macrophage activation (21), the underlying complex mechanisms for their differential transcriptional programs need to be further elucidated. Previous observations in murine and human macrophages connecting epigenetic changes and transcriptional regulation already revealed useful insights into macrophage biology (62,157,161,252). An adapted approach for differentially activated primary human macrophages could further expand the knowledge of the underlying regulatory mechanisms for macrophage specialization.

Different questions arose in this matter: Is there a common macrophage core signature also visible on the epigenetic level? How extensive are the epigenetic differences for activated human macrophages? Do super-enhancers play a role in macrophage activation? Are important TRs differentially regulated on the epigenetics level? Four differentially activated primary human macrophages were generated to detect their epigenetic activity states at promoter, enhancer or super-enhancer positions with the analysis of ChIP-seq data sets for four histone modifications (HM) in different combinations (H3K4me1, H3K4me3, H3K27Ac, H3K27me3). The HM data was subsequently correlated to RNA-seq expression data to find answers for the previously described questions.

3. Materials

3.1. Chemicals and Reagents

Agencourt AMPure XP beads	Beckman Coulter, Brea, USA
BSA	New England Biolabs, UK
CFSE	Sigma-Aldrich, München, DE
Complete protease inhibitors	Hoffmann-La Roche, Basel, CH
dATP	New England Biolabs, UK
Dimethylsulfoxid	Sigma-Aldrich, München, DE
dNTP solution	New England Biolabs, UK
Dynabeads protein G	Invitrogen Life Technologies, Karlsruhe, DE
Dynabeads® M-450	InvitrogenLifeTechnologies,Karlsruhe, DE
EDTA	Calbiochem, San Diego, USA
EGTA	Calbiochem, San Diego, USA
Ethanol	Roth, Karlsruhe, DE
Ethylendiamintetraacetat	Sigma, St Louis, USA
Fetal calf serum	Invitrogen LifeTechnologies,Karlsruhe, DE
Formaldehyde 16% (wt/vol)	Thermo Scientific
Glutamax	Invitrogen Life Technologies, Karlsruhe, DE
Glycerol	Calbiochem, San Diego, USA
Glycine	Calbiochem, San Diego, USA
Glycogen	Roche Diagnostics GmbH, Mannheim, DE
HEPES buffer	Calbiochem, San Diego, USA
Ionomycin	Sigma Aldrich, Taufkirchen, DE
Klenow enzyme	New England Biolabs, UK
KOH solution	Sigma Aldrich, Taufkirchen, DE
Lithium chloride	Sigma Aldrich, Taufkirchen, DE
MicroBeads CD14 ⁺	MiltenyiBiotect, Bergisch Gladbach, DE
MNase enzyme	Sigma Aldrich, Taufkirchen, DE
NEB buffer 2	New England Biolabs, UK
N-Lauryl-sarcosine	Sigma Aldrich, Taufkirchen, DE
NP-40	Calbiochem, San Diego, USA
Pancoll	PAA Laboratories GmbH, Pasching, AT
PBS	PAA Laboratories GmbH, Pasching, AT
PEG-8,000	Sigma Aldrich, Taufkirchen, DE

Pfu Ultra buffer	Agilent Technologies, USA
Pfu Ultra II fusion	Agilent Technologies, USA
Propidium Iodide	Sigma Aldrich, Taufkirchen, DE
Prostaglandin E2	Sigma Aldrich, Taufkirchen, DE
Proteinase K	Invitrogen LifeTechnologies,Karlsruhe, DE
QIAzol®	Qiagen, Hilden, DE
Quick ligase	New England Biolabs, UK
Quick ligation buffer	New England Biolabs, UK
Rnase A	Hoffmann-La Roche, Basel, CH
RPMI 1640	PAA Laboratories GmbH, Pasching, AT
SDS	Calbiochem, San Diego, USA
Sodium acetate	Roth, Karlsruhe, DE
Sodium chloride	Roth, Karlsruhe, DE
Sodium deoxycholic acid	Calbiochem, San Diego, USA
Sodium hydroxide	Merck, Darmstadt, DE
T4 DNA polymerase	New England Biolabs, UK
T4 ligase buffer	New England Biolabs, UK
T4 polynucleotide kinase	New England Biolabs, UK
Trichloroaceticacid	Merck, Darmstadt, DE
TRIS (hydroxymethyl)-aminomethane	Roth, Karlsruhe, DE
Tris buffer	Calbiochem, San Diego, USA
Triton X-100	Calbiochem, San Diego, USA
Trypanblue	Merck, Darmstadt, DE
Tween-20	Calbiochem, San Diego, USA
Water, molecular biology grade	Sigma Aldrich, Taufkirchen, DE

3.2. Cytokines

rh GM-CSF	Immunotools, Friesoythe, DE
rh IL-4	Immunotools, Friesoythe, DE
rh IFN γ	Immunotools, Friesoythe, DE
rh TNF- α	Immunotools, Friesoythe, DE
Pam3Cysk4	Invivogen, San Francisco, USA
PGE2	Sigma Aldrich, Taufkirchen, DE

3.3. Antibodies

ChIP antibodies

Antigen	Host	IgG Type	Company
H3K27Ac	Rabbit	Polyclonal	Abcam, Cambridge, UK
H3K27me3	Rabbit	Polyclonal	Merck Millipore, Billerica, USA
H3K4me1	Rabbit	Polyclonal	Abcam, Cambridge, UK
H3K4me3	Rabbit	Monoclonal	Merck Millipore, Billerica, USA
PU.1	Rabbit	Polyclonal	Santa Cruz Biotechnology, Dallas, USA
RNA-Pol II	Rabbit	Polyclonal	Santa Cruz Biotechnology, Dallas, USA

Antibodies for flow cytometry

Antigen	Fluorophore	Company
CD11c	PE	BD Biosciences, Heidelberg, DE
CD14	APC	BD Biosciences, Heidelberg, DE
CD14	Pacific Blue	Biolegend, San Diego, USA
CD19	FITC	BD Biosciences, Heidelberg, DE
CD206	APC	BD Biosciences, Heidelberg, DE
CD23	FITC	BD Biosciences, Heidelberg, DE
CD25	FITC	BD Biosciences, Heidelberg, DE
CD3	FITC	Biolegend, San Diego, USA
CD56	PE	BD Biosciences, Heidelberg, DE
CD86	APC	Biolegend, San Diego, USA

3.4. RT-PCR primers

Primers were designed with the Beacon Designer software and oligonucleotides were ordered from Sigma-Aldrich (Taufkirchen, DE). Semi-quantitative real time PCR experiments were performed with listed primers to validate performed ChIP-seq experiments. Primers were either designed to target sites with known associated HM enrichments as positive controls (pos.) or negative sites without corresponding target HM (neg.).

Target Gene	Sequence 5' to 3'	Purpose
GAPDH_1	tactagcggttttacgggcg	H3K4me3 pos. / H3K27me3 neg.
	tcgaacaggaggagcagagagcga	H3K4me3 pos. / H3K27me3 neg.
Bmp2	ctcagcactccgcatttg	H3K27me3 pos. / H3K4me3 neg.
	ctccatccaacgcttag	H3K27me3 pos. / H3K4me3 neg.
GAPDH_2	gccttgctcttgctactc	H3K4me1 pos.
	gcctgcctggtgataatc	H3K4me1 pos.
GAPDH_3	atctcagtcgttcccaaagtc	H3K27Ac pos.
	gtgatcgggtgctggttcc	H3K27Ac pos.
Chr5_Intergenic_region	gaacaactggatgggacaaac	H3K4me1 neg / H3K27Ac neg.
	acacgaatggaaccttatctg	H3K4me1 neg / H3K27Ac neg.

3.5. Plastic ware

0,2 - 2 ml Eppendorf tubes	Eppendorf GmbH, Hamburg, DE
1.7 ml siliconized tubes	Sigma Aldrich, Taufkirchen, DE
Falcon 15 ml	Greiner bio-one, Frickenhausen, DE
Falcon 250 ml	Corning Life Sciences, Tewksbury, USA
Falcon 50 ml	Greiner bio-one, Frickenhausen, DE
LS columns	Miltenyi Biotech, Bergisch Gladbach, DE
Microtube, AFA fiber	Covaris, Woburn USA
Nunclon™ 24-well tissue culture plate	Thermo Scientific, Rockford, US
Nunclon™ 48-well tissue culture plate	Thermo Scientific, Rockford, US
Nunclon™ 6-well tissue culture plate	Thermo Scientific, Rockford, US
Parafilm	Pechiney, Chicago, US
Pipette filter tips, 10, 200, 1000 µl	Starlab GmbH, Hamburg, DE
Pipette tips, 10, 200, 1000 µl	Greiner bio-one, Frickenhausen, DE
Pipettes 2, 5, 10 and 25 ml	Greiner bio-one, Frickenhausen, DE
Pre-Separation Filters	MiltenyiBiotech, Bergisch-Gladbach, DE
Slide-A-Lyzer 3.5K dialysis cassette	Thermo Scientific, Rockford, USA
Sterile filter 22 µm	Sartorius, Hannover, DE
Syringe 50 ml	Braun, Melsungen, DE

3.6. Equipment

Auto MACS pro separator	MiltenyiBiotech, Bergisch Gladbach, DE
Bandelin Sono Plus	Bandelin, Berlin, DE
BD LSR II Flow cytometer	BD Biosciences, Heidelberg, DE
Bioanalyzer	Agilent Technologies, Santa Clara, USA
BioDoc Analyze	Biometra, Jena, DE
cBot	Illumina, Eindhoven, NL
Centrifuge Type 5415	Eppendorf GmbH, Hamburg, DE
Centrifuge Type 5424	Eppendorf GmbH, Hamburg, DE
Centrifuge Type 5810R	Eppendorf GmbH, Hamburg, DE
Covaris S220	LGC Genomics, Berlin, DE
Eppendorf Concentrator Plus	Eppendorf GmbH, Hamburg, DE
HiScanSQ system	Illumina, Eindhoven, NL
HiSeq 1500	Illumina, Eindhoven, NL
Incubator Binder B series	Binder, Tuttlingen, DE
Incubator Binder C series	Binder, Tuttlingen, DE
LightCycler 480 PCR system	Roche diagnostics, Basel, Switzerland
Magnet MACS Multi Stand	MiltenyiBiotech, Bergisch Gladbach, DE
Magnet MPC-S	DynalBiotech, Oslo, NO
Mikroskope SM-LUX	Leitz, Wetzlar, DE
NanoDrop 2000	Thermo Scientific, Rockford, USA
Neubauer chamber	Carl Roth Karlsruhe, DE
Odyssey® Infrared Imaging System	LI-COR Biosciences, Bad Homburg, DE
Orbital shaking incubator	Stuart, Staffordshire, UK
pH-meter	Knick, Berlin DE
Pipette boy	IBS Integra Biosciences, CH
PowerPac HC Power Supply	Bio-Rad Laboratories, München, DE
Roller Mixer SRT 1	Stuart, Staffordshire, UK
Shaker (type 3011)	Mettler-Toledo, Zwingenberg, DE
Tapestation	GFL, Burgwedel, DE
Water bath	Agilent Technologies, Santa Clara, USA
	Memmert, Schwabach, DE

3.7. Buffers and media

A-base addition buffer (XChIP)

6 μ l 10x NEB buffer 2, 0.1 μ l dATP (100 mM), 10.9 μ l nuclease free water

Adapter ligation buffer (XChIP)

29 μ l 2x DNA quick ligase buffer

End Repair Buffer (XChIP)

6.7 μ l 10x T4 DNA ligase buffer, 0.67 μ l BSA (10 mg/ml), 0.67 μ l dNTP mix (10 mM), 16.96 μ l nuclease free water

LiCl buffer (NChIP)

0.25 M LiCl, 0.5 % NP-40, 0.5 % sodium deoxycholate

MACS buffer

1 x PBS supplemented with 0.5 % BSA, 2 mM EDTA, pH=7.2 sterile-filtered

MNase digestion buffer (NChIP)

50 mM Tris-Cl (pH 7.6), 1 mM CaCl₂, 0.2% Triton X-100, 5 mM sodium butyrate, protease inhibitor cocktail (working concentration), 0.5 mM PMSF

MNase stop buffer (NChIP)

10 mM Tris-Cl (pH 7.6), 10 mM EDTA

PBS (NChIP)

137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄

RIPA buffer (NChIP)

10 mM Tris-Cl (pH 7.6), 1 mM EDTA, 0.1 % SDS, 0.1 % sodium deoxycholate, 1 % Triton X-100

RIPA-NaCl buffer (NChIP)

0.3 M NaCl added to RIPA buffer

TE buffer (NChIP)

10 mM Tris (pH 8.0), 1 mM EDTA

TE/Triton X-100 buffer (NChIP)

0.2 % Triton X-100 added to TE-buffer

3.8. Kits

Bioanalyzer high sensitivity DNA kit	Agilent Technologies, Santa Clara, USA
ChIP-seq sample preparation kit	Illumina, Eindhoven, NL
KAPA SYBR FAST Roche LightCycler 480 2X qPCR Master Mix	KAPABiosystems, Wilmington, USA
Maxima SYBR Green qPCR Master Mix	Fermentas, GmbH, DE
Multiplex sequencing primer kit	Illumina, Eindhoven, NL
NextFlex ChIP-Seq Barcodes - 48	Bioo Scientific, Austin, USA

QIAquick PCR Purification Kit	Qiagen, Hilden, DE
QIAquick Gel Extraction Kit	Qiagen, Hilden, DE
Tapestation DNA high sensitivity kit	Agilent Technologies, Santa Clara, USA
Truseq DNA sample preparation kit v1/v3	Illumina, Eindhoven, NL
Truseq RNA sample preparation kit v3	Illumina, Eindhoven, NL
Truseq SBS Kit v1/v3	Illumina, Eindhoven, NL

3.9. Software

Beacon Designer	PREMIER Biosoft, Palo Alto, USA
Bowtie 0.12.8	Ben Langmead, Cole Trapnell, USA
Circos 0.65	Martin Krzywinski, Canada's Michael Smith Genome Sciences Centre, CA
Cluster 3.0	Michael Eisen, Stanford University, USA
CorelDRAW X6	Corel Corporation, Ontario, CA
DESeq2 package (R-package)	Michael I Love, Harvard School of Public Health, USA
Eland 1.8 (NGS Alignment)	Illumina, Eindhoven, NL
Endnote X5	Thomson Reuters, Carlsbad, USA
FACSDiva	BD Biosciences, Heidelberg, GER
Flowjo 7.6.2	Tree Star, Ashland, USA
GenomeStudio	Illumina, Eindhoven, NL
HOMER 4.3	Chris Benner, UCSD/Salk Institute, USA
Integrative Genomics Viewer	Broad Institute, Cambridge, USA
Java Treeview 3.0	Alok Saldanha, GPLv2 License
LightCycler 480 SW 1.5	Roche applied sciences, Basel, CH
MACS 1.4.2	Yong Zhang, Tao Liu, Xiaole Shirley Liu Lab, USA
Microsoft Office	Microsoft GmbH, Unterschleissheim, DE
Partek genomics suite	Partek, St Louis, Missouri, USA
Python 2.6.9	The Python Software Foundation, Delaware, USA
R 2.15.0 - Programming language	Ross Ihaka, Robert Gentleman, GNU General Public License, USA
Samtools 0.1.18	Heng Li, BSD License/MIT License
SigmaPlot 12.0	Systat Software GmbH, San José, USA

TopHat2

Daehwan Kim, University of Maryland,
USA

3.10. Next generation sequencing data sets

All utilized next generation sequencing data sets are available for download from various sources and are listed in Table 1.

Table 1. Utilized next generation sequencing data sets

ChIP-seq data set name	Source	GEO ID
GATA3 (WT)	http://www.ncbi.nlm.nih.gov/geo/	GSM523224
GATA3 (KO)	http://www.ncbi.nlm.nih.gov/geo/	GSM742022
SRF (WT)	http://homer.salk.edu/homer/data/index.html	-
SRF (KO)	http://homer.salk.edu/homer/data/index.html	-
PU.1 (WT)	http://www.ncbi.nlm.nih.gov/geo/	GSM538003
PU.1 (KO)	http://www.ncbi.nlm.nih.gov/geo/	GSM537999
ATF3 ^{Unstim} (WT)	http://www.ncbi.nlm.nih.gov/geo/	GSM1334036
ATF3 ^{Unstim} (KO)	http://www.ncbi.nlm.nih.gov/geo/	GSM1334037
ATF3 ^{HDL} (WT)	http://www.ncbi.nlm.nih.gov/geo/	GSM1334038
ATF3 ^{HDL} (KO)	http://www.ncbi.nlm.nih.gov/geo/	GSM1334039
ATF3 ^{HDL+CpG} (WT)	http://www.ncbi.nlm.nih.gov/geo/	GSM1334040
ATF3 ^{HDL+CpG} (KO)	http://www.ncbi.nlm.nih.gov/geo/	GSM1334041
ATF3 ^{LPS 0-120min.} (WT)	http://www.ncbi.nlm.nih.gov/geo/	GSE36104
HM H3K4me3 (M ^b)	http://www.ncbi.nlm.nih.gov/geo/	GSM1146439
HM H3K4me3 (M ^{IFNy})	http://www.ncbi.nlm.nih.gov/geo/	GSM1146440
HM H3K4me3 (M ^{IL4})	http://www.ncbi.nlm.nih.gov/geo/	GSM1146441
HM H3K4me3 (M ^{TPP})	http://www.ncbi.nlm.nih.gov/geo/	GSM1146442
HM H3K27me3 (M ^b)	http://www.ncbi.nlm.nih.gov/geo/	GSE66595
HM H3K27me3 (M ^{IFNy})	http://www.ncbi.nlm.nih.gov/geo/	GSE66595
HM H3K27me3 (M ^{IL4})	http://www.ncbi.nlm.nih.gov/geo/	GSE66595
HM H3K27me3 (M ^{TPP})	http://www.ncbi.nlm.nih.gov/geo/	GSE66595
HM H3K27Ac (M ^b)	http://www.ncbi.nlm.nih.gov/geo/	GSE66595
HM H3K27Ac (M ^{IFNy})	http://www.ncbi.nlm.nih.gov/geo/	GSE66595
HM H3K27Ac (M ^{IL4})	http://www.ncbi.nlm.nih.gov/geo/	GSE66595
HM H3K27Ac (M ^{TPP})	http://www.ncbi.nlm.nih.gov/geo/	GSE66595
HM H3K4me1 (M ^b)	http://www.ncbi.nlm.nih.gov/geo/	GSE66595
HM H3K4me1 (M ^{IFNy})	http://www.ncbi.nlm.nih.gov/geo/	GSE66595
HM H3K4me1 (M ^{IL4})	http://www.ncbi.nlm.nih.gov/geo/	GSE66595
HM H3K4me1 (M ^{TPP})	http://www.ncbi.nlm.nih.gov/geo/	GSE66595
HM murine tissue M Φ	http://www.ncbi.nlm.nih.gov/geo/	GSE63339
RNA-Seq (M ^b)	http://www.ncbi.nlm.nih.gov/geo/	GSE66595
RNA-Seq (M ^{IFNy})	http://www.ncbi.nlm.nih.gov/geo/	GSE36952
RNA-Seq (M ^{IL4})	http://www.ncbi.nlm.nih.gov/geo/	GSE36952
RNA-Seq (M ^{TPP})	http://www.ncbi.nlm.nih.gov/geo/	GSE66595

Datasets for the analysis of 7 human tissues (5.2.7) are contained within Release 9 of the Human Epigenome Atlas (<http://www.epigenomeatlas.org>). In addition, 7 murine tissue macrophage datasets (5.2.8) are available under the main GEO accession number GSE63341, with the subseries accession numbers GSE63339 and GSE63340 for ChIP-seq and RNAseq experiments.

Data was either used in the Fastq file format or as SRA files (GEO). SRA files required an additional conversion step into the FastQ file format (see Methods section 4.8.1).

4. Methods

Buffy coats from healthy donors were obtained from the Institute for Experimental Hematology and Transfusion Medicine of the University Hospital Bonn following protocols accepted by the institutional review board at the University of Bonn (local ethics vote no. 045/09). Informed written consent was provided for each specimen according to the Declaration of Helsinki.

4.1. Isolation of monocytes

Peripheral blood mononuclear cells (PBMCs) were isolated in a Pancol density gradient by centrifugation at 250g for 25 minutes at room temperature. The interface containing the PBMCs was collected, washed twice with PBS and resuspended in 3 ml of MACS buffer. The fraction of white blood cells was incubated for 20 minutes in 200 μ l CD14⁺ magnetic beads to isolate CD14 positive monocytes with the magnetic assorted cell sorting technique (MACS) according to the manufacturer's protocol. LS columns were prepared in a MACS separation magnet and pre-separation filters were additionally used to exclude cell aggregates or fat pellets from downstream experiments. The white cell fraction was washed after bead incubation and transferred into the separation filter on top of the LS column. With the application of three times 3 ml of MACS buffer CD14⁺ cells moved through the LS column and were held in the magnetic field, while CD14 negative cells were washed through the LS column into a collection tube. After magnet removal, collected CD14⁺ cells were counted and their purity was assessed. Stainings with CD14, CD11c, CD19 (B-cells), CD56 (NK-cells) and CD3 (Lymphocytes) antibodies were performed to identify impurities with CD14 negative cells following fluorescence activated cell sorting (FACS) analysis. Successful monocyte isolations with purities above 95% were used for further experiments.

4.2. Generation and activation of human monocyte derived macrophages

Isolated monocytes were cultured in RPMI1640 medium supplemented with 10 % fetal calf serum (FCS) and 500 U/ml of recombinant human granulocyte macrophage colony-stimulating factor (rh GM-CSF). The cell concentration was adjusted to 2×10^6 cells/ml and maintained in 3 ml volume per well in 6-well plates for 72 h at 37°C and 5 % CO₂. Unstimulated macrophages (also called baseline macrophages, M^b) were harvested and aliquots were seeded out in RPMI/10% FCS medium with the addition of 500 U/ml rh GM-CSF. Treatments with different cytokines and chemicals for additional 72 h generated three macrophage activation states activated under different conditions described in Table 2 and Figure 3.

Table 2. Cytokines and chemicals used for macrophage activation

Macrophage activation state	Stimulation	Dose
M ^{IL4}	rhIL-4	500 U/ml
M ^{IFNγ}	rhIFN γ	200 U/ml
M ^{TPP}	rhTNF- α	800 U/ml
	Pam ₃ CSK4	1 μ g/ml
	PGE ₂	1 μ g/ml

Additionally, the expression of typical activation state specific surface proteins was validated by FACS analysis using CD14, CD23, CD25 and CD86 antibodies as described previously (21,253). For further details, see Methods section 4.4. Activated macrophages were further processed for CHIP-seq and RNA-seq experiments for 3 independent donors.

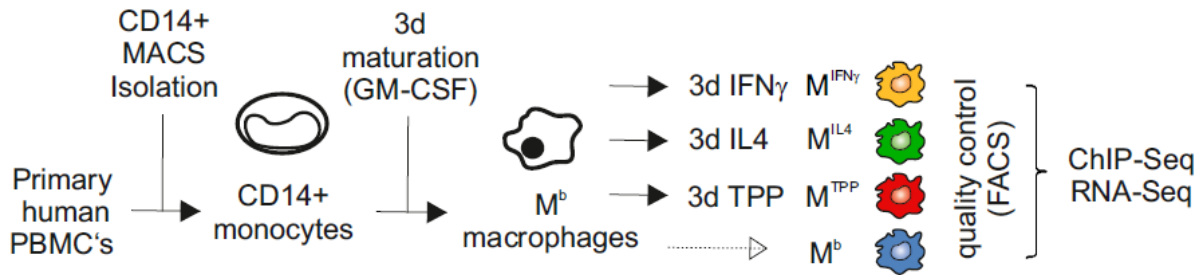


Figure 3. Workflow for primary human macrophage generation and differentiation.

Isolated PBMCs were positively selected for CD14 with the MACS technique. Monocytes were cultured with rh GM-CSF for 72 h to generate unstimulated baseline macrophages (M^b). These macrophages were further *in vitro* activated under different stimulatory conditions into classically ($M^{IFN\gamma}$) or alternatively activated (M^{IL4}) macrophages in addition to macrophages induced under chronic inflammatory conditions (M^{TPP}) by TNF- α , prostaglandin E2, and the TLR2 ligand P3C (TPP). Figure has been previously published as part of manuscript “The transcriptional regulator network of human inflammatory macrophages is defined by open chromatin” by myself as first shared author (Schmidt, S.V.*, Krebs, W.*, Ulas, T.* et al.

4.3. Generation of murine bone-marrow derived macrophages

Bone marrow-derived macrophages (BMDMs) involved in ATF3 ChIP-seq experiments were generated by Dominic De Nardo and Larisa Labzin (Institute of Innate Immunity, University Hospitals, University of Bonn, Germany). The detailed method and analysis was described previously (254). In brief, BMDMs from 6- to 8-week old wildtype (WT) C57BL/6 and ATF3 deficient knockout (KO) mice were obtained by culturing bone marrow cells for 6 days in DMEM supplemented with 40 ng/ml rh M-CSF, 10% (vol/vol) FCS and 10 μ g/ml Ciprobay-500 (R&D Systems). BMDMs were stimulated with 2 mg/ml HDL for 6 h (HDL condition) or with 2 mg/ml HDL for 6 h with additional treatment of 100nM CpG for 4 h (HDL + CpG condition). A third aliquot of cells was left untreated in medium (Unstim condition). Generated BMDM cells were subsequently used to perform cross-linked ATF3 ChIP experiments following deep sequencing.

4.4. Flow cytometry

For the quantitative and qualitative analysis of surface marker expression on macrophages, between 1×10^5 and 5×10^5 cells were resuspended in 100 μ l blocking buffer (PBS + 20 %

FCS) incubated for 10 min. at 4°C and stained with 1-5µl of antibody (used antibodies see Materials section 3.3) coupled to fluorochromes for 20 min. at 4°C. Next, cells were washed twice with 2 ml of PBS and pelleted by centrifugation at 300g for 8 minutes. After the supernatant was discarded, the cell volume was adjusted to approximately 300µl with PBS and measured on a flow cytometer FACS LSR II (Becton Dickinson, USA). Signal events were gated according to the expected size and granularity of living macrophages. Between 10.000 and 50.000 events in the final gates were recorded. Cell viability was assessed by the addition of propidium iodide (PI; 1µg/ml) to the cells directly before measurement to detect DNA not protected by an intact cell membrane, thus allowing the estimation of cell death ratios. The frequency emission overlap of fluorochromes in a multi-color setup made a compensation step necessary. Single stainings for each fluorochrome used with anti-IgG1k coated polystyrene beads made the clear assignment of fluorochrome light signals in the multi-color setup possible. Finally, raw data was exported as .fcs files and analyzed with the FlowJo Software.

4.5. Semi-quantitative real time PCR

The semi-quantitative real time PCR method was applied as validation for successful ChIP experiments using the Maxima SYBR Green/Fluorescein qPCR Master Mix kit (Fermentas) with volumes according to manufacturer's instructions listed in Table 3.

Table 3. qRT-PCR reaction mix

Type	Amount
Maxima SYBR Green qPCR MM (2x)	12.5 µl
Forward primer (10µM)	0.75 µl
Reverse primer (10µM)	0.75 µl
Water	10 µl
ChIP DNA (Diluted 1:5)	1 µl

Each qPCR reaction mix was loaded at least as duplicates into a 96 well plate and results were generated with the LightCycler 480 II System (Roche) with the standard temperature/time program recommended by the manufacturer listed in Table 4.

Table 4. Standard RT-qPCR program

Type	Temperature, °C	Time	Cycles
Initial denaturation	95	10 min.	1
Denaturation	95	15 sec.	40
Annealing	60	30 sec.	
Extension	72	30 sec.	

The relative PCR enrichment of positive (Pos) against negative (Neg) target sites were calculated according to the $\Delta\Delta Ct$ method with the following formula and ChIP validation results were visualized as the ratios of specific ChIP antibody values (S) to unspecific IgG antibody values (IgG) also called as relative enrichment of ChIP antibody against unspecific IgG antibody:

$$\Delta Ct = Ct_{(S)} - Ct_{(IgG)}$$

$$\Delta\Delta Ct = Ct_{(Pos)} - Ct_{(Neg)}$$

$$Relative\ enrichment = 2^{-\Delta\Delta Ct}$$

4.6. Isolation of RNA following RNA-sequencing

RNA isolation procedure and necessary downstream steps were performed by Michael Kraut and Laura Bohmann. RNA isolation, quality tests and RNA sequencing experiments were carried out according to previously published protocols (255).

In brief, 5×10^6 cultured macrophages were lysed in TRIZOL (Invitrogen). Total RNA was subsequently extracted according to the manufacturers' protocol (RNeasy kit, Qiagen) and quality assessments were performed. Ratios of the absorbance at 260 nm and 280 nm were detected using a Nanodrop 2000 Spectrometer (Thermo Scientific). Furthermore, the integrity of the 18/28 S RNA bands was visualized with a denaturing agarose gel. Only total RNA meeting the quality criteria with low degradation of 18/28 S RNA bands was used to perform the library construction with the Illumina TruSeq RNA sample preparation kit according to manufacturer's descriptions.

In brief, purified mRNA originating from 5-10 μ g of total RNA purified by the usage of poly-T oligo-attached magnetic beads was fragmented with divalent cations in fragmentation buffer (Illumina). Random oligonucleotides were used as a starting point bound to mRNA for the reverse transcriptase enzyme SuperScript II to conduct first strand cDNA synthesis. The second strand cDNA synthesis was realized by the DNA Polymerase I and RNase H enzyme.

Newly created cDNA overhangs were repaired into blunt ends with exonuclease/polymerase enzymes and an A-base was added to DNA 3' ends. Afterwards, Illumina PE adapter oligonucleotides were ligated to both ends of the cDNA molecules. A library size selection for preferentially 200 bp cDNA fragments was performed with 2 % (w/v) agarose gels and fragments were purified using the QIAquick gel extraction kit (Qiagen). A 15 cycles PCR reaction amplified the adapter ligated cDNA molecules with PE1.0 and PE2.0 PCR primers (Illumina). PCR products were purified with the QIAquick PCR purification kit (Qiagen). The measurements of quality and quantity for constructed cDNA library was carried out on a Bioanalyzer 2100 (Agilent) with the Agilent high sensitivity DNA assay kit. The validated cDNA library pieces were immobilized on a glass flow cell (Illumina) and clusters were generated out of each single strand cDNA molecule within the cBot machine (Illumina).

Finally, the constructed library was sequenced in a paired-end 100 bp sequencing run for in total 208 cycles on the HiScan SQ system (Illumina) with reagents from the TruSeq SBS kit v3 (Illumina). With the CASAVA v1.8 software (Illumina) base calling and data conversion was performed by Jil Sander. TopHat2 (256) was utilized for the alignment against the hg18 reference genome (257) and the transcript information for each gene was extracted using Partek Genomics Suite v6.6. Subsequently, the annotated data was normalized with the DESeq2 package (<http://dx.doi.org/10.1101/002832>), a plugin for the statistical software R v3.0.2 (258). For visualization purposes, normalized read counts below 1 were set to 1 (flooring) to allow the usage of logarithmic scales during the comparison of differentially activated macrophages. Genes with normalized RNA-seq values ≥ 10 were defined as expressed. Independent sequencing samples were created for 3 donors.

4.7. ChIP following deep sequencing

Chromatin immunoprecipitation following next generation sequencing (ChIP-seq) is a powerful tool to identify the location of DNA binding proteins or HMs associated with specific regions of the genome. This method combines a chromatin immunoprecipitation using specific antibodies against a protein of interest or a HM to isolate the corresponding bound DNA pieces with a next generation high throughput sequencing approach to finally identify the base pair sequences of unknown ChIP DNA fragments (201).

ChIP experiments were performed with two different approaches, either using chromatin fixated with formaldehyde to tightly bind proteins to corresponding genomic positions (X-ChIP) or chromatin in its natural state without a fixation step (N-ChIP). Each method possesses its own advantages and disadvantages described in the introduction (1.4.2).

ChIP-seq experiments were conducted for three independent donors and pooled prior to peak calling and downstream data analysis. HM data sets yielded in average 28 million ChIP-seq reads from 12 to 56 million reads for individual macrophage activation experiments

4.7.1. Native ChIP (N-ChIP)

The generation of H3K4me3 HM data sets was performed by adapting a previously published native ChIP protocol (244) and is published on the Gene Expression Omnibus (GEO) platform under accession number GSE47188 (21). In brief, 2×10^7 primary human macrophage cells were used from each activated macrophage population for native ChIP-seq experiments. Cell lysates were sheared with 0.3 units MNase enzyme (Sigma Aldrich) for 10 minutes at 37°C into even multiple pieces of 180 bp (mono-, di-, trinucleosome DNA). After an additional ultrasonic step (Bandelin), analysis of the resulting DNA fragment length was verified in an agarose gel after the shearing of the chromatin. A dialysis was executed with dialysis cassettes (Thermo scientific) exchanging the MNase shearing buffer with RIPA buffer. During the ChIP experiments rabbit monoclonal anti-trimethyl histone antibodies (Millipore) bound to protein G magnetic beads (Life Technologies) or unspecific isotype controls (Millipore) were used to isolate chromatin pieces modified with three methyl groups added to lysine 4 on histone protein 3 (H3K4me3). After several washing steps with low and high salt buffers, the de-cross-linking and protein digestion was carried out overnight. Finally, the DNA was purified with silica-membrane-based spin columns included in the Qiagen MinElute kit (Qiagen). A qPCR validation step was performed to calculate the DNA enrichment signals for target relative to control ChIP DNA using specific or unspecific isotype control antibodies (described in section 4.5)

The ChIP-Seq Sample Preparation Kit (Illumina) and the Multiplexing Sample Preparation Oligonucleotide Kit (Illumina) were used with 10 ng ChIP-DNA following the manufacturer's instructions to create multiplex sequencing libraries. Briefly, purified DNA ends were repaired using PNK and Klenow enzyme, followed by treatment with Klenow exo minus polymerase to add an A base to the 3' end used for the adaptor ligation. With a size selection step by an agarose gel, DNA fragments with approximately 220 bp size were excised and eluted with the QIAquick Gel Extraction kit (Qiagen). After subsequent adapter ligation, an amplification was performed for 5 cycles with PCR primers 1.1 and 2.1 (Illumina). An additional 13 cycles PCR further amplified DNA and integrated an index sequence for each library using specific multiplex PCR primers (Illumina). In a last step, library DNA was purified and 8 pM of single stranded library was loaded onto a glass Flow Cell (Illumina).

Finally, a 57 bp single read multiplex sequencing run was performed on the Illumina HiScanSQ/HiSeq 1000 systems resulting in DNA sequences with the length of 50 base pairs. These sequences were demultiplexed with CASAVA software (Illumina) and further processed with bioinformatics methods described in section 4.8.2.

4.7.2. Cross-linked ChIP (X-ChIP)

A previously published high-throughput low cell X-ChIP protocol was adapted to create ChIP-seq data sets for the HMs H3K4me1, H3K27ac, H3K27me3, also including ChIP-seq library construction procedures (238). All experiments were performed in 96-well plate format with multichannel pipettes to minimize technical errors and maximize speed for optimal protein stability. In brief, chromatin was in the first step cross-linked for ChIP reactions to bound proteins with 1 % formaldehyde (Sigma Aldrich) for 10 minutes. Cell nuclei from 0.5×10^6 lysed cells were isolated and chromatin was subsequently sheared for 45 min. into approximately 200 bp pieces with the Covaris S220 ultrasound system (Covaris Ltd.). Chromatin size verification was performed with the Bioanalyzer 2100 or TapeStation system (Agilent). Polyclonal rabbit ChIP antibodies were bound for 1 hour at RT to protein G magnetic beads (Life Technologies) for the subsequent detection of three different HMs H3K4me1 (Abcam), H3K27Ac (Abcam) or H3K27me3 (Millipore). Additionally, unspecific polyclonal rabbit IgG antibody (Millipore) was used to create IgG control data sets for the assessment of successful ChIP experiments. Antibody-bead-complexes were subsequently bound to chromatin in an overnight reaction. Washing steps with different salt buffers removed unspecific antibody-protein complexes. Chromatin de-cross-linking and RNA/protein digestions for 4 hours allowed the DNA purification with AMPure XP SPRI beads (Beckman Coulter). Isolated ChIP-DNA was measured with the Bioanalyzer 2100/TapeStation (Agilent) to record concentration and size. An additional qPCR validation measuring specific ChIP DNA signals relative to unspecific IgG control signals allowed the assessment of successfully performed ChIP experiments with high relative enrichments for target against control samples.

ChIP-seq library construction was conducted with 0.5 ng ChIP DNA or sheared chromatin prior to ChIP experiments (Input control) as starting material. ChIP DNA ends were repaired to blunt ends and an A-base was added to their 3' ends. Sample volumes were evaporated to 4 μ l to optimize efficacy of the ligation step, while using less DNA ligase enzyme in comparison to the described standard protocol. The ligation was performed with Illumina sequencing compatible NEXTflex adapter oligonucleotides (Bioo Scientific) in concentrations of 0.6 μ M and incubated for 15 min. at 25 °C. A SPRI beads cleanup (Beckman Coulter)

removed unbound adapter oligonucleotides and a 14 cycles PCR amplification step further increased the ChIP-seq library DNA amount. Another SPRI beads cleanup did not only exclude remaining PCR primers, but also allowed a size selection and resulted in ChIP-seq DNA library sizes between 100 and 500 bp. The double sided SPRI beads size selection was conducted with different volumes of water to PEG containing SPRI buffer, which enables the binding of different DNA sizes to the AMPure XP beads, thus mediating a size selection.

Concentration and molarity of purified library DNA was assessed with the KAPA qPCR system for Roche LightCycler 480 (Kapabiosystems) and the Bioanalyzer/Tapestation (Agilent Technologies). Approximately 11 pM of single stranded library DNA was loaded onto the Hiseq 1000 sequencer (Illumina). Multiplex single read runs for 51 cycles and additional 6 index sequence cycles were performed with the TruSeq SBS reagents kit v3 and the Truseq cluster generation kit v3 (Illumina). Generated 50 bp sequence data was demultiplexed with Casava (Illumina) to separate data according to the used samples and was further processed with bioinformatics methods described in section 4.8.2.

4.8. Bioinformatics data analysis

4.8.1. Conversion of published data sets

Data sets were downloaded as SRA files from the GEO database (see Materials section 3.10). This file format does not only allow an efficient storage of ChIP-seq data concerning file-sizes, the data can also be reconverted into multiple file formats like fasta, fastq or sam as needed (<http://www.ncbi.nlm.nih.gov/books/NBK158900>). The SRA Toolkit program was downloaded and used according to the manufacturer's recommendations to perform the conversion process (<http://www.ncbi.nlm.nih.gov/Traces/sra/?view=software>). The following command was used for the conversion of SRA files into FastQ files:

```
fastq-dump <downloaded file> > <filename.fastq>
```

4.8.2. Alignment to reference genome assembly

During the alignment step of ChIP-seq data, DNA sequences were mapped with an aligner algorithm to the recent reference genome to identify ChIP protein enriched genomic regions. The alignments were performed with the program Bowtie with the following command (259):

```
./bowtie -t -q -e 70 -l 28 -n 2 --best --maxbts 125 -S -p 8 <reference genome> <ChIP-seq-file.fastq> <aligned-ChIP-seq-file.sam>
```

With this command, bowtie uses the first 28 base calls for the alignment and accepts up to two mismatches with a Phred score sum of smaller than 70. Bowtie also reports not only the uniquely alignable sequences fitting to exact one genomic position, but also the best fitting genomic site for short sequences matching to more than one genomic position.

The input FastQ files do not only contain sequences of the ChIP-seq reads, but also list the quality values for each called base. With this information, the aligner tries to find the genomic position where the unknown DNA sequence belong to and thus the ChIP protein of interest was bound with a high probability (201,260). All ChIP-seq FastQ files either created from the author (see Methods section 4.7), converted from SRA files (see Methods section 4.8.1) or downloaded from other websites (see Material section 3.10) were aligned to the UCSC mouse (mm9) or UCSC human (hg18) reference genomes. TF data sets used in section 5.1 were aligned to the mouse mm9 reference genome, while all HM data sets used in section 5.2 were aligned to the human hg18 genome.

4.8.3. Peak Calling

Subsequently to the alignment of ChIP-seq data to the reference genome, regions with significant enrichment for ChIP-seq tags (short sequences reads) were identified by a peak calling step. In general, a peak caller takes several properties into account to separate background signals from true-positive signals. First of all, the ratio of treatment to control ChIP-seq tag counts at enriched regions is an important value for peak calling. Statistical calculations like binomial or poisson distribution models further improve the probability of finding true-positive peak positions, due to their correction for regional bias in tag density, copy number or amplification variations (201,261,262).

The peak calling for results depicted in section 5.1 were performed with MACS (Model-based Analysis of ChIP-Seq) v1.4.0/v2.0.1 (15), specialized on the identification of sharp peaks in standard ChIP-seq experiments with the following options: `-g 1.87e9 -s 51 --bw 150 -w --single-profile -p 1e-4 --on-auto`. In brief, WT peaks used in the standard method (described in section 5.1.1) were called with aligned '.bam' files as input using MACS v1.4.0. Genomic positions with bimodal signal enrichments were defined as peak regions, with strand and anti-strand tags enriched upstream and downstream of TF binding. MACS shifted sense and anti-sense tags to a peak midpoint and used shifted tags as significant peak position signals under a poisson distribution model with the parameter ($\lambda_{\text{local WT}}$).

$$\lambda_{\text{local WT}} = \max(\lambda_{\text{BG WT}}, \lambda_{5\text{kWT}}, \lambda_{10\text{kWT}})$$

MACS determined the maximum background signals $\lambda_{\text{local WT}}$ either from the whole WT data set ($\lambda_{\text{BG WT}}$) or from 5 or 10kb regions centered to the peak midpoint ($\lambda_{5\text{k}}$, $\lambda_{10\text{k}}$). Called WT peaks fitted to the poisson distribution model with p -values smaller than 0.04.

During KOIN calculations, MACS used aligned '.bam' files from WT experiments as "treatment" file and aligned '.bam' files from KO experiments as separate "control" to calculate background signals. Similar peak calling steps were performed according to the standard method described before. The sum of called WT peaks with p -values smaller than 0.04 (poisson model) are based on the parameter $\lambda_{\text{local KO}}$.

$$\lambda_{\text{local KO}} = \max(\lambda_{\text{BG KO}}, \lambda_{1\text{kKO}}, \lambda_{5\text{kKO}}, \lambda_{10\text{kKO}})$$

MACS estimated the maximum background signals with the parameter $\lambda_{\text{local KO}}$ either from the whole KO data set ($\lambda_{\text{BG KO}}$) or from 1, 5 or 10kbp regions centered to the peak middle points ($\lambda_{1\text{kKO}}$, $\lambda_{5\text{kKO}}$, $\lambda_{10\text{kKO}}$). Next, "treatment" and "control" files were swapped in a second peak calling to identify "negative" KO peaks present in the KO data set. KO peaks with poisson p -values smaller than 0.04 were described by the parameter $\lambda_{\text{local WT}}$. $\lambda_{\text{local WT}}$ was defined as estimated maximum background signal in the WT data set. This is calculated by MACS either from the whole data set ($\lambda_{\text{BG WT}}$) or at different regions with 1, 5 or 10kb length around the peak summit ($\lambda_{1\text{kWT}}$, $\lambda_{5\text{kWT}}$, $\lambda_{10\text{kWT}}$).

$$\lambda_{\text{local WT}} = \max(\lambda_{\text{BG WT}}, \lambda_{1\text{kWT}}, \lambda_{5\text{kWT}}, \lambda_{10\text{kWT}})$$

Non-specific peaks which were simultaneously present in the groups of WT and KO peaks were excluded for further analysis in the first two steps of the KOIN method. Finally, the sum of corrected peaks after the KOIN method only contained significant WT peaks as described in the following formula:

$$\text{Corrected peaks (KOIN method)} = \text{WT peaks} - (\text{KO peaks} \cap \text{WT peaks})$$

Peaks with fold changes < 2 for normalized WT tag counts compared to KO tag counts were filtered out to increase peak specificity. With the module "annotatePeaks.pl" included in the HOMER program, ChIP-seq tags were finally counted and total tag count normalized to WT and KO data sets with the -d option (command: *annotatePeaks peak-file.bed mm9 -size given -d peak-file-tag-directory/*). Fold changes (FC) between normalized tag counts for WT compared to KO data sets were subsequently calculated. P -values were adapted from the standard value 0.05 to 0.04 to reduce the number of enriched regions to a level allowing

statistical analysis with the GREAT tool. A command line-based description file for linux systems is available online to perform the KOIN correction during peak calling with the MACS program (<https://github.com/LIMES-NGS/KOIN-pipeline>).

In parallel, peak calling for primary human macrophage data sets for HM data was performed with comparable steps using the HOMER program (72). In HOMER the histone style option was used with peak sizes of 1000 bp and with minimal peak distances of 2500 bp for an optimal HM analysis, especially for broad ChIP-seq peaks. Input samples were utilized as control files during peak detection. Significantly called peaks were defined by a four-fold enrichment in treatment data sets over input tag counts and passed a false discovery rate of $1e^{-3}$ or better under a statistical Poisson model.

4.8.4. Peak annotation and tag distributions

MACS or HOMER called peak positions were annotated, concerning their genomic location on the reference genome mm9/hg18. The perl script “annotatePeaks.pl” included in the HOMER program identifies the center of each peak region and finds the closest TSS of a known annotated Refseq gene and finally connects this gene name to the identified peak position. Annotation data also allowed the generation of global peak distributions, describing the exact location for called peaks in relation to the known genomic function like intronic, exonic, intergenic and others.

Heatmaps centered to peaks of interest were created with HOMER using the “annotatePeaks.pl” script with the heatmap option (-ghist). During KOIN analysis (section 5.1), 10 bp sliding windows were utilized to calculate ChIP-seq raw signals 2 kb up- and downstream of peak middle points. For the second results part (section 5.2) 500 bp windows were used for the assessment of ChIP-seq signals and 12 kb regions were depicted and centered to each peak site. For both approaches ChIP-seq signals were normalized to 10^7 total tag numbers. Subsequently, the program Java-Treeview (v1.1.6.r4) visualized the HOMER output data as heatmaps.

For the visualization of genomic example sites with corresponding ChIP-seq raw data signals the integrative genomics viewer (IGV) was utilized (263). The normalization of ChIP-seq tag counts to 10^7 total tags allowed the direct comparison of WT and KO or activated macrophage data sets, independently of total tag sizes generated for each experiment.

4.8.5. *De novo* motif enrichment analysis

The *de novo* motif discovery was performed by HOMER with the "findMotifsGenome.pl" script. For TF peaks in the first part of the herein presented results (section 5.1), sequences 100 bp up- and downstream of peak centers were used as input, while for HM enriched sites (results section 5.2) sequences 500 bp up- and downstream of peak centers were utilized. Approximately 50000 random background sequences were used as controls with the same GC% content as the target sequences. An auto-normalization step removed imbalances in the short oligos between target and background sequences. Next, enriched sequences were identified in a global search with up to two allowed mismatches and most enriched motifs were optimized in their motif probability with a sensitive local optimization algorithm by scoring each oligo in the data to the probability matrix. Finally, enriched motifs with lengths of 8, 10 and 12 bp were identified and scored according to their cumulative binomial p -values. Corresponding positional weight matrices (PWMs) were visualized for each respective experiment together with the found motif.

Statistical p -values were converted into positive integer numbers with the formula $-\log_{10}(\text{motif } p\text{-value})$ and displayed as heatmaps with Partek genomics suite (v6.6) (results section 5.1). The percentages of target sequences with corresponding binding motif, determined by the standard method or the KOIN method were compared and depicted with horizontal aligned bar plots. For each motif, the percentage defines the number of ChIP-seq tags with a positive match for the particular motif among all tags used for the motif predictions. Multiple motif hits at a given tag site are possible.

Motif ratios for top 5 enriched motifs were called with HOMER, independently for WT and KO data sets. Numbers of motif occurrences in WT data were divided by the numbers found in KO data sets, after their normalization to total peak counts in respective WT or KO data sets. Higher motif abundance in WT compared to KO data sets are defined as positive FC values. Vice versa, higher ratios in KO compared to WT data sets are depicted with negative values.

4.8.6. Gene ontology enrichments

Peak sites were associated to the nearest TSS of known RefSeq genes using HOMER (4.8.4). For the first results section (5.1), gene names associated to peaks within ± 1000 bp of known TSS or inside gene bodies were used as input for the Cytoscape program (v2.8.3) (264) Gene set enrichment analysis (GOEA) was performed with the BiNGO plugin (265). For statistical significant results, the false discovery rate threshold (FDR) was set to 0.001. Output files for WT and KO data sets were visualized as differential network with the Cytoscape plugin Enrichment Map (v1.2). For the detailed options a Jaccard coefficient cutoff

was set to 0.001 and a FDR corrected Q-value of 0.01 was used. GO-terms in clusters were marked and named according to their main functions, categorized by the Cytoscape plugin Word cloud. Only subnetworks with more than four elements were visualized. A GOEA analysis for gene groups with and without KOIN correction including distal loci was performed with the Genomic Regions Enrichment of Annotations (GREAT) tool (v2.0.2) (266) at default parameters. Within this analysis, functions for sets of non-coding *cis*-regulatory genomic regions and TSS-associated regions were assessed. Results were listed in the respective figure according to their $(-\log_{10})$ binomial p -values, simultaneously statistical significant for the hypergeometric test as well.

In results section 5.2, genes with H3K4me3 HM peaks within ± 2500 bp around known RefSeq TSS were used for the assessment of epigenetic promoter states, while genes with associated H3K4me1 enhancer peaks with distances between 2.5 kb and 100 kb away from the next RefSeq TSS were utilized for the identification of their epigenetic enhancer status. These genes were used as input for the Cytoscape program upgraded by the plugins BiNGO, Enrichment map and Word cloud to perform a differential GOEA analysis. For the final resulting GO-terms statistical filters were set to 0.001 for the FDR corrected q-value and subnetworks with at least 2 GO-terms were shown.

4.8.7. Identification of “Hyper-ChIPable regions”

Top 25 TF peak positions with extraordinary high ChIP-seq signals in comparison to all called peaks were identified and listed according to their normalized tag counts. Only peak sites located in at least two data sets were used for further inspection with the UCSC genome browser (<http://genome.ucsc.edu>). Loci were defined as “hyper-ChIPable”, if the following criteria were visible: (1) Putative regions display additional enrichments for unrelated DNA binding proteins currently available at the ENCODE database (described below), (2) characterized as DNase I hypersensitive sites, (3) showed high RNA polymerase II binding and (4) showed enrichments for various HMs including H3K4me1, H3K4me3 and H3K27me3. Using the UCSC genome browser, putative “hyper-ChIPable regions” were visually analyzed for all six data sets (<https://genome.ucsc.edu/ENCODE>). In case the investigated regions showed high signals for all above mentioned criteria beyond the viewable default scaling, the region was defined as “hyper-ChIPable”. The following data sets were used: TF binding sites (Caltech TFBS; LICR TFBS; PSU TFBS; Stan/Yale TFBS), Histone modifications (Caltech Histone; LICR Histone, PSU Histone; Stan/Yale Histone), DNase I hypersensitive cleavage sites (PSU DNaseI HS; UW DNaseI DGF; UW DNaseI HS).

4.8.8. Epigenetic promoter and enhancer class identification

Promoter analyses were focused on H3K4me3 enriched regions 2500 bp up- and downstream of known RefSeq TSS. Accessible promoters (Pa) displayed both H3K4me3 and H3K27Ac signals, while poised promoters (Pp) gained repressive H3K27me3 HM marks. The integration of different HM data sets was performed with the HOMER program using the “mergePeaks” module. Promoter sites were defined as positively enriched for H3K27Ac or H3K27me3 signals, if a direct overlap between enriched genomic positions to H3K4me3 peaks was detected. Macrophage promoters were defined as activation specific, if H3K4me3 peak sites were exclusively found in only one macrophage activation state or in case the H3K4me3 HM signals were at least two times higher in one activation state than the H3K4me3 tag counts in the others.

The identification of putative enhancer sites according to H3K4me1 enrichments required the initial removal of promoter sites enabled by a coordinate data file, listing sites \pm 2500 bp up- and downstream of known RefSeq TSS. The intersection of promoter related positions in the coordinate data file with H3K4me1 enriched sites were excluded as promoter sites from the enhancer analysis. Strong enhancers (Es) defined by H3K4me1 and H3K27Ac signals lacked repressive H3K27me3 signals. Weak enhancers (Ew) were only enriched for H3K4me1 signals, whereas poised enhancers (Ep) gained additional repressive H3K27me3 marks. Specific enhancers were exclusively enriched for H3K4me1 in only one macrophage activation state or showed at least two times more H3K4me1 signals in one macrophage activation state compared to the others.

The super enhancer (SE) identification was performed according to methods described in the first SE publication (168) with HOMER, applying the SE option (-style super) within the peak calling module “findPeaks”. First, HOMER calls H3K27Ac peaks according to standard procedures and stitches peak sites with distances less than 12.5 kb together into SE regions. Second, H3K27Ac signals for each SE region are determined by the total normalized ChIP-seq reads in target minus input data sets. Finally, SE regions are sorted by their signal scores (y-axis) after normalization to the highest signals and to the number of putative total enhancer regions (x-axis). The results are plotted as line graph. SE regions are located past the point, where the slope of the line graph is greater than 1 (see example plot Figure 35B).

Subsequently, peak positions for different promoter and enhancer classes with corresponding normalized ChIP-seq raw data were displayed as heatmaps (detailed description in methods 4.8.4)

4.8.9. Correlation of gene expression to enhancer and promoter classification

Promoter and enhancer peak positions were annotated to known RefSeq gene symbols as described in the Method section 4.8.4. In addition, average RNA-seq (n=3) or array-based (n=3) expression data was subsequently connected to the same gene symbols, allowing a correlation of peak positions and expression data. Only enhancers with peak distances between 2.5 kb and 100 kb to the nearest TSS were used for the analysis. If more than one enhancer or promoter class correlated to one gene, for example due to multiple enhancers located in the proximity of a gene, the following weighting ensured a clear classification to only one epigenetic promoter and enhancer state:

Promoters: $P_a > P_p$

Enhancers: $E_s > E_w > E_p$

If not stated otherwise all genes with expression values above the following cut offs were defined as expressed:

RNA-seq > 10 (linear scale) or > 2.3 (e^x scale) or > 3.3 (log2 scale)

Array-based > 150 (linear scale) or > 5 (e^x scale) or > 7.22 (log2 scale)

Dashed lines in the figures correspond to these cut off values. Statistical significance between expression values for activated macrophages was assessed by the Wilcoxon rank-sum test.

4.8.10. Generation of co-regulation networks for transcriptional regulators and macrophage core signature genes

The generation of co-regulation networks for macrophage core signature genes was described in a previous publication (21). In brief, transcriptional information was used to perform a reverse network approach on GM-CSF stimulated baseline macrophages further stimulated with 28 different conditions with the ARACNe software to compute the interactions of important macrophage genes. 10% of resulting common macrophage genes with the highest values of predicted interactions (869 genes) were used for the correlation to established promoter and enhancer classifications. The network visualization was created with the Cytoscape program.

A list of TR (transcriptional regulator) genes for human and murine transcriptional regulators encompassing TFs, co-factors, RNA-binding proteins (RNBP) and chromatin remodelers was created originating from the TFCat database (267). Subsequent co-regulation networks based on variable Pearson's correlation coefficients were created for these important TR genes using expression data as input for the BioLayout Express3D program, described previously (21,268). For the visualization of TR-TR pairs and their correlation coefficients the Cytoscape program was utilized using a force-directed layout. Additional information like transcriptional fold changes calculated against a group mean value or corresponding epigenetic information for corresponding TR genes was mapped onto the co-regulation network. Networks were generated not only for differentially activated human macrophages, but also for human tissues and murine tissue macrophages to identify corresponding important TRs.

All presented co-regulation networks are also accessible online at VisuTranscript (http://www.s-khb.de/visutrans/content/index_eng.html).

5. Results

This thesis was designed to elucidate in its first chapter (5.1) the importance of false-positive binding signals falsifying the analysis and interpretation of next generation sequencing data sets for transcriptional regulators and transcription factors (204,250,251). A novel bioinformatics strategy called KOIN (knockout implement normalization) utilizes protein knockout experiments during the ChIP-seq data analysis to significantly and efficiently correct for false-positive signals derived from unspecific binding of ChIP-seq antibodies. The analyses of different ChIP-seq data sets were generated and online available data sets from other scientific publications with various bioinformatics tools were utilized to verify the improvement of KOIN onto the biological interpretation of TF ChIP-seq data.

In the second part of this thesis (chapter 5.2) various ChIP-seq experiments analyzing HMs in different macrophages were performed to gain insights into the epigenetic status of primary human macrophages activated under different stimulatory conditions. As HMs allow the classification of important genomic regulatory sites like promoters and *cis*-regulatory elements (enhancers) into different activation states describing the present chromatin status of these loci (156,160,167). The chromatin structure itself can be separated into condensed and inactive or open and active sites. This dynamic change of chromatin status at specific genomic sites is one of the major regulatory mechanisms for the transcription of cellular genes. ChIP-seq and RNA-seq data were merged into correlation networks of central TRs to get information about the co-occurrence of certain HM states or the presence of SE markings to the transcriptional activity of stimulus specific and common macrophage genes. Additionally, RNA-seq and HM ChIP-seq data sets from murine tissue macrophages under homeostatic conditions or cells originating from different tissues were used to perform epigenetic classifications of gene loci and comparisons of expression data for important TR genes in detailed co-regulation networks with four activated macrophage states analyzed in detail.

5.1. Knockout mouse models - A tool for the optimization of TF ChIP-seq experiments

Chromatin immunoprecipitation experiments following deep sequencing (ChIP-seq) require tight quality controls and outstanding antibody binding properties concerning binding specificity and efficiency to their protein targets to allow a precise data analysis and biological interpretation, while minimizing false-positive signals (204). Two different types of controls

were established in the past for ChIP-seq experiments (204,251), sheared chromatin (Input control) or chromatin eluted from antibodies with unspecific binding properties with a similar immunoglobulin classification as the specific ChIP antibody (IgG control). Both types of control data sets possess positive and negative aspects discussed in the introduction (section 1.4.2) (204). However, knockout experiments were underestimated in its value as essential ChIP-seq control until now. In fact, knock out experiments could better differentiate between false- and true-positive ChIP-seq signals and at the same time minimize negative factors of previously mentioned controls. During the analysis and normalization of ChIP-seq data sets knockout data were implemented. This correction step during ChIP-seq peak calling was termed “Knockout Implemented Normalization” (KOIN) and was published within the scope of this thesis (269). Figures shown in section 5.1 are part of this scientific publication.

5.1.1. Bioinformatic processing during “Knockout implemented normalization method” (KOIN)

Specific computational steps were necessary to utilize protein knockout (KO) experiments for the classical analysis (204) of ChIP-seq data sets (Figure 4). In the first step, next generation sequencing data sets from wild-type experiments (WT) and KO experiments for the corresponding WT protein were aligned to the human reference genome to find the target genomic position for each generated ChIP-seq sequence. Next, peak sites significantly enriched for WT ChIP-seq signals were identified during peak calling with the peak caller MACS. During the peak identification, the KO data set enabled the differentiation of false- and true-positive peaks. False-positive corrected peak files were afterwards further enriched for highly specific peaks with normalized tag counts at least two fold higher in WT compared to KO data sets. Finally, downstream data analysis was performed with different approaches like binding motif predictions or gene ontology enrichments. Details for conducted bioinformatics calculations are described in the methods section (4.8).

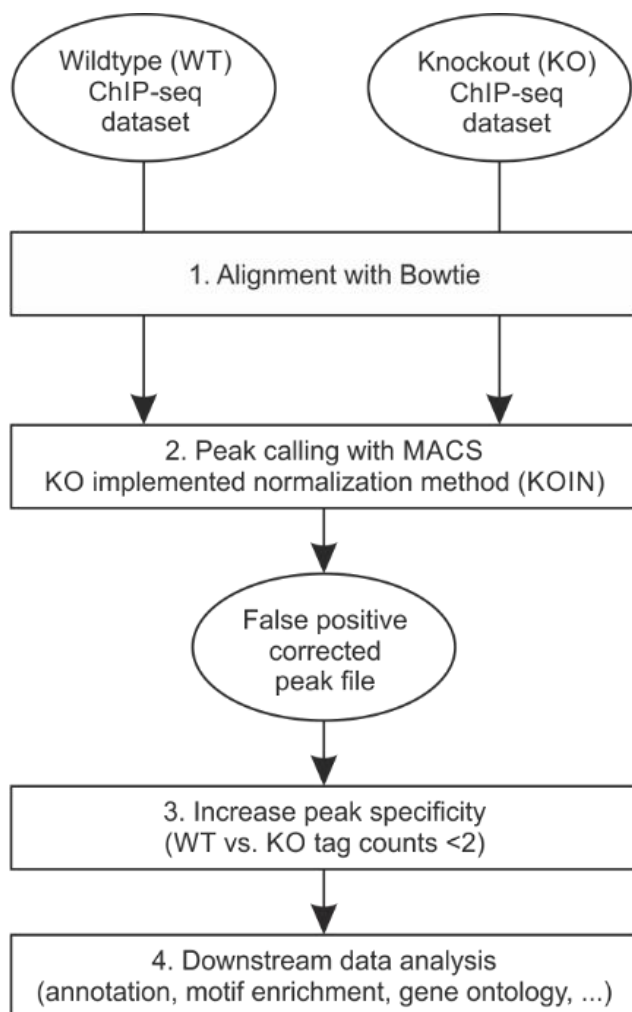


Figure 4. Flowchart depicting the generation of false-positive corrected peak files (KOIN method).

All necessary calculation steps and input/output files are displayed to generate false-positive corrected KOIN peak files. The aligned WT and KO data sets (step 1) were used as input to perform peak calling with MACS (step 2). In this step false-positive peaks were filtered out. Peak files were further filtered to increase peak specificity (step 3) using only peaks with at least two fold higher normalized tag counts for WT against KO experiments. In step 4 possible downstream data analysis options are mentioned ranging from peak annotation, connection of identified gene names to biological functions or to binding motif enrichment calculations.

For the evaluation of KOIN advantages in comparison to the standard analysis six ChIP-seq data sets were utilized to compute the global binding properties of the ATF3, GATA3, SRF and PU.1 proteins (72,254,255,270) (Figure 5, 3.10). The alignment was performed in both approaches (standard or KOIN) with Bowtie following the same protocols, whereas during the peak calling with MACS non-specific signals were identified and filtered out, following the KOIN protocol. Non-specific peak regions included high ChIP-seq tag counts in WT and KO data sets similarly, which possibly demonstrated unspecific binding capabilities of the utilized ChIP antibody (204). Improvements of KOIN in ChIP-seq analysis were compared in parallel to commonly used downstream data analysis methods using HOMER or Cytoscape.

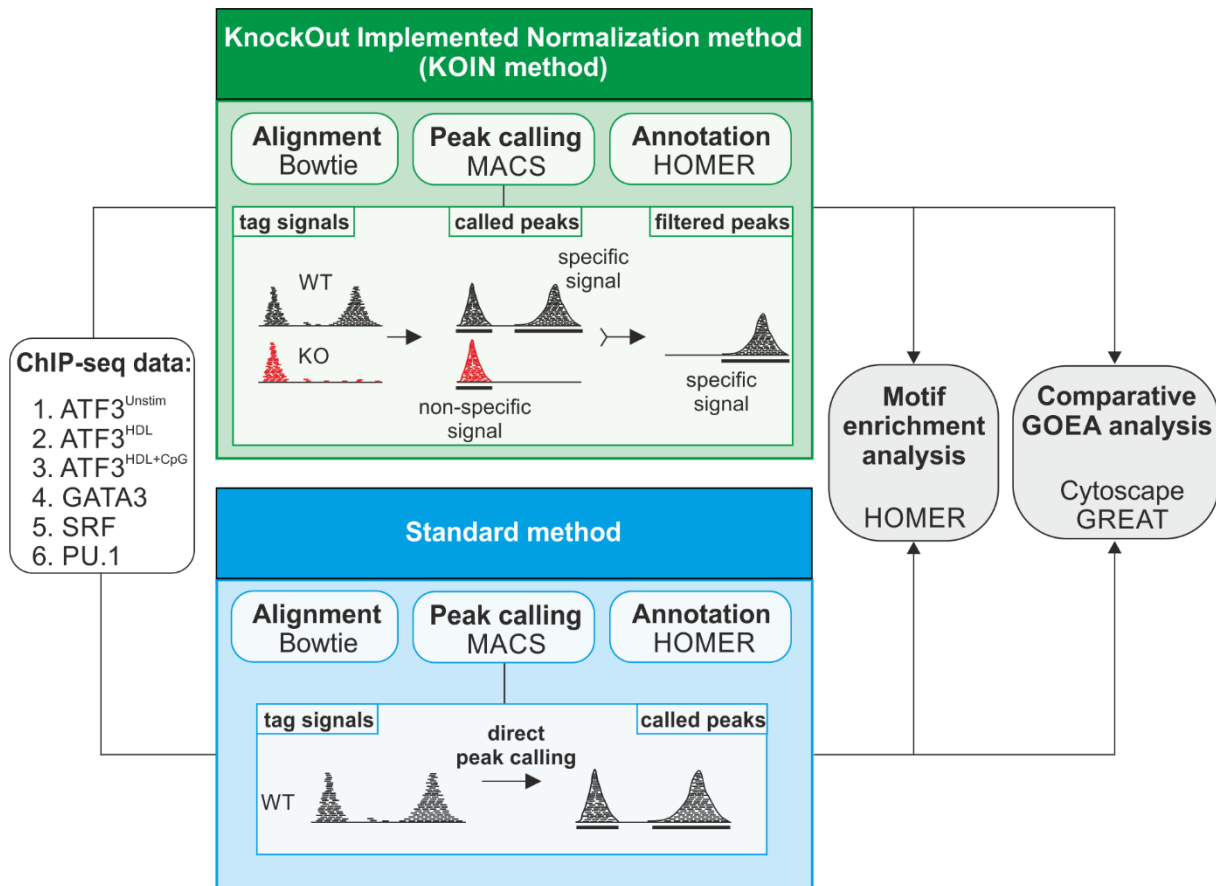


Figure 5. Schema of ChIP-seq data analysis using either the KOIN method or the standard method.

Six different data sets were used to perform the KOIN method (green) with wild type (WT) and knockout (KO) data sets, whereas for the standard method only WT data sets were utilized. In both approaches next generation sequencing data sets were aligned to the reference genome and peak calling was performed with MACS. KOIN allowed a discrimination between specific and non-specific signals, in contrast to the standard peak calling. ChIP-seq signals enriched peak positions were annotated using the HOMER program. For downstream data processing motif enrichment and a comparative gene ontology enrichment analysis was performed for both methods. The corresponding results were compared.

In conclusion, standard ChIP-seq analysis and the KOIN method differed only in one step facilitating the implementation of KOIN to the standard ChIP-seq data analysis. During the peak calling procedure non-specific peak sites could be excluded from downstream analysis utilizing KO ChIP-seq data sets.

5.1.2. KOIN identifies and excludes false-positive ChIP-seq signals

False-positive binding signals are inevitable in ChIP-seq experiments (204). Due to the binding properties of antibodies, not only specific bindings with high affinities to target antigens occur, but also non-specific bindings against irrelevant antigens or unspecific binding to the constant FC region. For the illustration of the amount of non-specific antibody binding, dot plot visualizations were created, plotting normalized tag counts for WT and KO data sets of uncorrected (grey color) and KOIN corrected (red color) peak positions found for the four analyzed TFs (Figure 6). The KOIN method identified significant numbers of false-positive peaks especially in the data sets for SRF and GATA3 proteins with respective percentages of 80 % and 78 % of false positive signals. Surprisingly, the correction rates for the ATF3 protein were dependent on the stimulatory conditions (unstimulated, stimulated with HDL, stimulated with HDL and CpG) ranging from 43 % to 79 %. In contrast, the KOIN correction slightly increased the already high peak numbers of the PU.1 data set (see section 5.1.3).

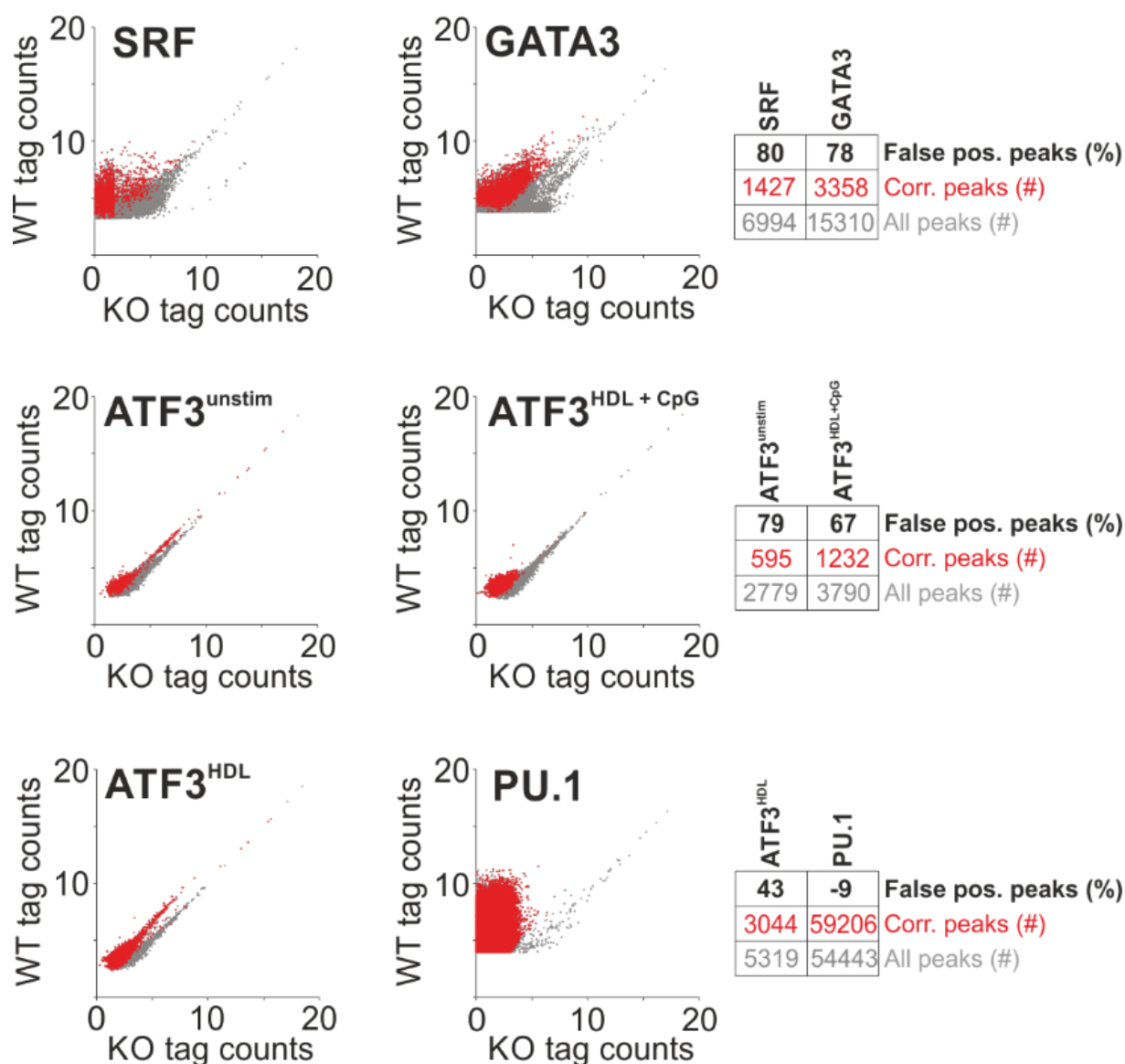


Figure 6. Independent ChIP-seq experiments reveal false-positive signals after KOIN calculations.

ChIP-seq tag counts for SRF, GATA3, PU.1 and ATF3 WT and KO data sets were visualized as dotplots with \log_2 scale normalized to 10^7 total tag counts. ATF3 data sets were stimulated under different stimulatory conditions as indicated. Peaks identified with the standard method were plotted in grey color whereas KOIN corrected peaks were overlaid as red dots into the same plots. Jitter was added to prevent excessive overlaps of data points and to improve the visualization. KOIN corrected (red) and uncorrected standard peak numbers (grey) were plotted additionally to the false-positive peak rates (in percent) to the right.

Illuminating the question, whether the false-positive peak signals were enriched at specific functional genomic sites, annotations of peaks to the reference genome were performed, thus enabling a functional categorization of peaks into intergenic, intronic and promoter

regions (Figure 7). SRF and GATA3 experiments showed similar fractions of KOIN corrected peaks between 18 % and 24 % irrespective of the genomic peak position. As mentioned before, KOIN corrected peak numbers differed in the four ATF3 data sets in dependency on the stimulatory conditions. A high specificity of promoter peaks was detected with 50 % and 63 % for KOIN corrected peaks in ATF3 control experiments without stimulation or stimulation with HDL and CpG, while intergenic and intronic sites displayed similar rates of false-positive signals. Monocytes stimulated with HDL displayed only minor changes between intergenic (53%), intronic (63%) and promoter (69%) peak numbers after KOIN correction. KOIN correction increased PU.1 enriched sites for true-positive peak numbers up to 15 % irrespective of the analyzed intergenic, intronic, or promoter sites (see section 5.1.3).

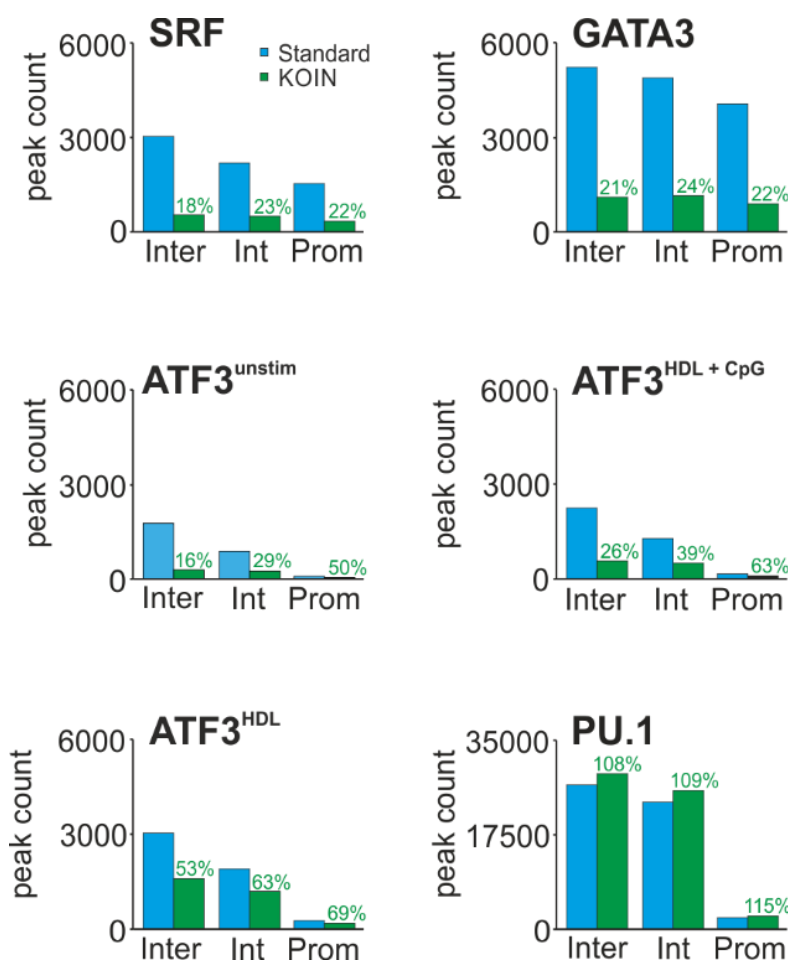


Figure 7. Genomic annotation of KOIN corrected and uncorrected peaks reveal a high distribution of false-positive peaks to various genomic positions.

A comparison of peak numbers between standard peaks (blue) and KOIN corrected peaks (green) was performed according to their genomic positions. Statistics for peaks located at promoter (Prom), intronic (Int) and intergenic (Inter) positions are shown as bar charts. The percentages of true-positive peaks which remained after KOIN correction were included as green numbers.

In addition to visualizing global numbers of KOIN corrected peaks, pile-up heatmap visualizations relative to peak centres were created to evaluate the false-positive signal strength before (standard method) and after the KOIN correction process (Figure 8a). WT and KO data sets were depicted ± 2000 bp from peak centres. As an example, the HDL and

CpG stimulated ATF3 data set showed distinct false-positive signals after the standard method in the KO data set (blue box), whereas these signals could be completely abrogated with KOIN correction (green box). Even in ChIP-seq experiments with high specificity like the PU.1 data set, KOIN successfully removed false-positive peak positions completely (Figure 8b).

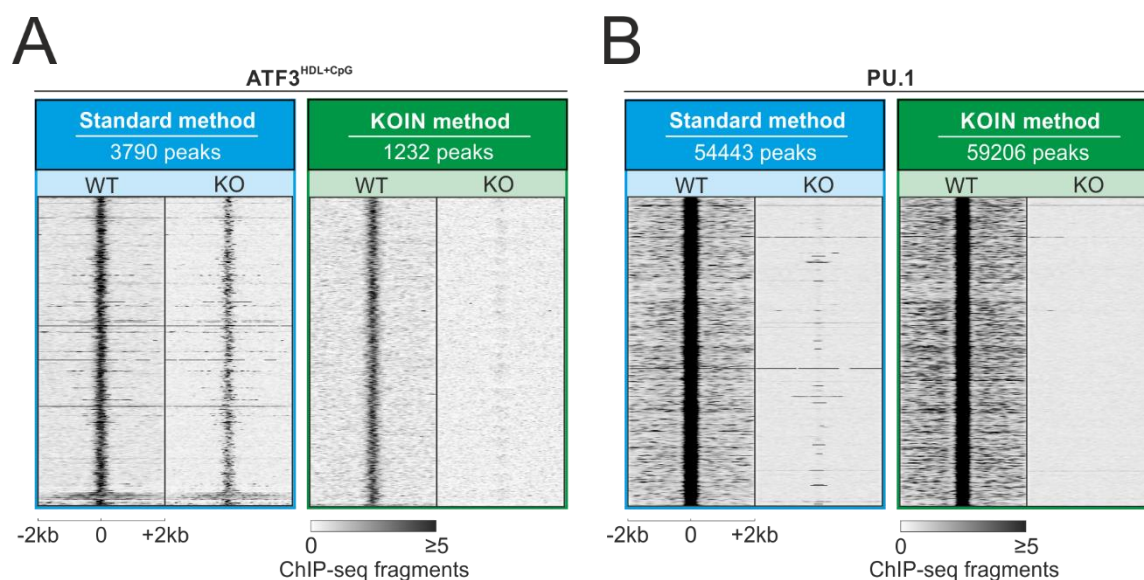


Figure 8. False-positive signals are depleted by KOIN correction in TF ChIP-seq data sets.

Pile-up heatmaps of ChIP-seq tag densities were created using 10bp sliding windows for total tag normalized data (normalized to 10^7 total tag counts). Tag densities are centred ± 2 kb to peak midpoints and are displayed for **(A)** ATF3 under the stimulatory influence of HDL and CpG as well as for **(B)** PU.1 visualized for both WT and KO samples before (blue) and after (green) KOIN correction.

An additional validation for the positive removal of false positive ChIP-seq signals was performed with a publicly available ChIP-seq data set, which was generated with the same ATF3 antibody (247). Two genomic sites (CD36 and CDK8) were visualized as examples (Figure 9). The validation data set (GSE36104) (3.10) called significant ATF3 peaks, marked by black boxes. In contrast, after KOIN correction significant ATF3 peaks remained at the CD36 locus but not at the CDK8 gene, which was simultaneously enriched for WT and KO signals. KOIN correction obviously allowed the differentiation of putative false-positive and true-positive peaks.

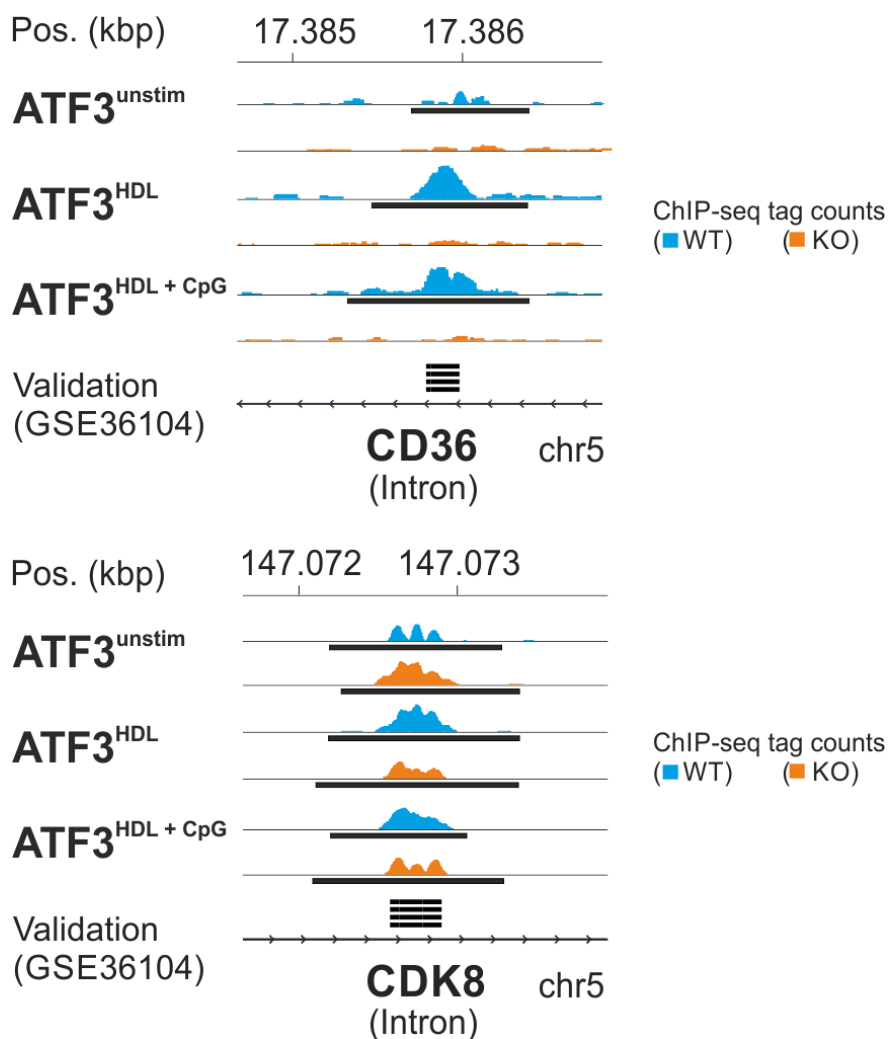


Figure 9. Representative genomic positions for true-positive and false-positive ChIP-seq peaks differ in KO signal intensities.

Representative ChIP-seq reads in the introns of true- and false-positive ATF3 binding sites. Black bars indicate significant peaks identified by MACS with p -values $\leq 10^{-4}$. An independently generated ATF3 data set (247) was used to visualize previously called significant peaks (black bars) called by the standard method.

Collectively, ChIP-seq data sets demonstrated variable numbers of false-positive peak signals. Some data sets contained high numbers of false-positive signals, whereas others like the PU.1 data set exhibited only marginal unspecific sites. In all cases, KOIN successfully identified and completely removed these non-specific signals.

5.1.3. Altered signal-to-noise ratios can increase peak numbers

In most cases, the KOIN method reduced the number of significant peaks in comparison to standard peak calling strategies correcting for non-specific tag signals. Surprisingly, KOIN also increased peak counts (9 %) in the PU.1 data set in comparison to the standard method and called 6037 new peaks (Figure 6, Figure 7 and Figure 8b). This effect of KOIN was visible due to lowered background signals, since KOIN concomitantly lowered the peak calling threshold. In highly specific data sets like PU.1 with only low numbers of false-positive peaks, decreased peak calling thresholds also lead to increased numbers of significant

peaks with lower tag counts. In sum, more new peaks with slightly lower tag counts were called as significant, compared to false-positive peaks which were lost during KOIN. Reduced background signals also improved statistical p -values determined during peak calling. WT and KOIN corrected p -values were visualized as dot plot (Figure 10). The apparent shift of KOIN/standard p -value ratios for called peaks to the upper left corner demonstrated the improvement of p -values during KOIN. Furthermore, the orange dots emphasized the newly called peaks with lower tag counts and higher but still significant p -values, whereas few grey dots visualized the KOIN excluded regions.

Taken together, KOIN even improved data analysis for highly specific antibodies by the reduction of noise, thereby identifying more significant binding regions, previously excluded as false-negative sites.

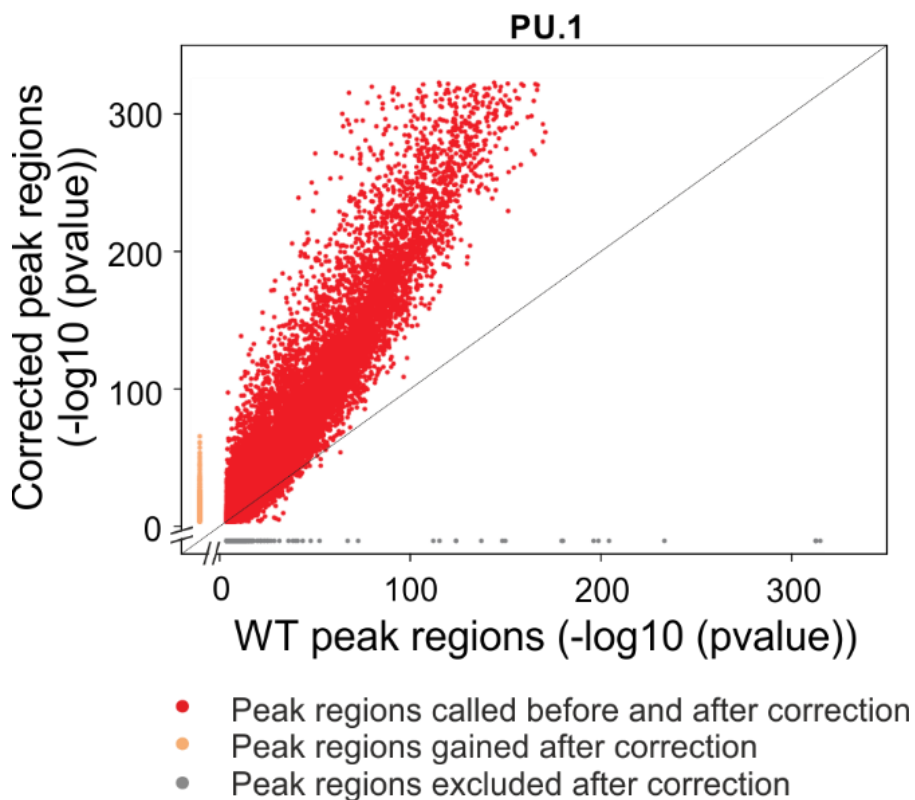


Figure 10. KOIN correction improves signal-to-noise ratios and corresponding statistical significance for KOIN peaks.

P -values for PU.1 peaks called with the standard method by MACS (x-axis, $-\log_{10}$ (p -value)) are plotted against the p -values for corresponding KOIN corrected peak positions (y-axis, $-\log_{10}$ (p -value)). Significantly called peaks marked in red were called with and

without KOIN correction. False-positive peaks lost after KOIN correction are marked in grey and orange peaks were exclusively called after application of KOIN correction.

5.1.4. Effect of KOIN on “hyper-ChIPable regions”

Recently, a special type of false-positive sites was described in yeast known as “hyper-ChIPable regions” (250). These highly expressed euchromatic regions are not only vulnerable to DNase I cleavage due to the absence of densely packed DNA to nucleosomes, but are also highly enriched for RNA Polymerase II allowing extremely high transcriptional activity. Other unrelated DNA-binding proteins or certain HMs can also be present at these sites. The potential of “hyper-ChIPable regions” to occur in other species as well influencing the ChIP-seq data analysis and interpretation justified a detailed analysis. For this task, loci with exceptionally high normalized ChIP-seq tag counts were identified, which were found in at least two out of six data sets. These sites were afterwards filtered concerning the above mentioned criteria using multiple previously published high quality ENCODE data sets to identify “hyper-ChIPable regions”. In fact, up to 18 „hyper-ChIPable regions“ were detected in the six analyzed ChIP-seq data sets highlighting the point that they could occur in most ChIP-seq data sets (Figure 11). More importantly, detected “hyper-ChIPable regions” were excluded by KOIN correction. To confirm the successful removal, the top 25 called and KOIN corrected peaks did not reveal “hyper-ChIPable” criteria.

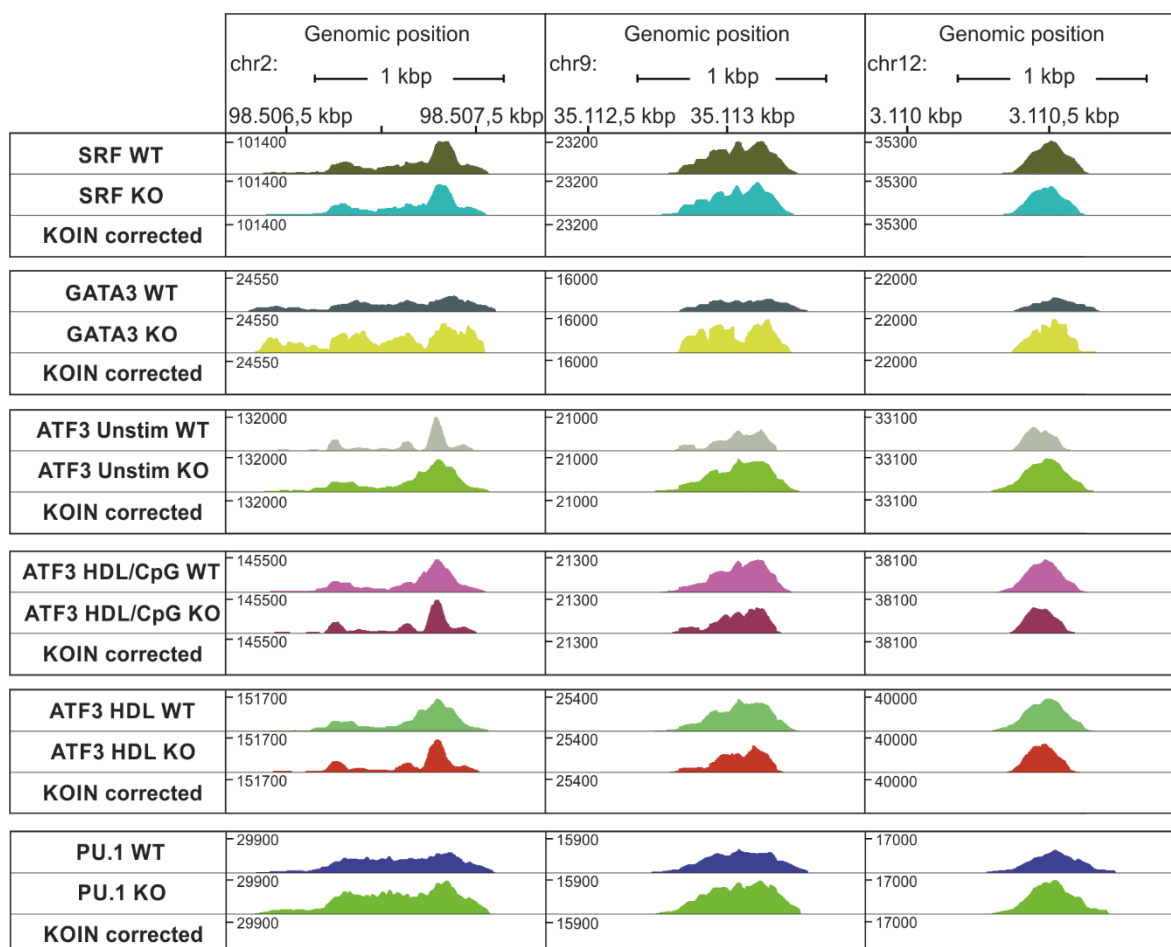


Figure 11. Examples of Hyper-ChIPable regions with highly enriched ChIP-seq signals.

Three examples for “hyper-ChIPable example regions” were visualized for six representative ChIP-seq TF data sets (SRF, GATA3, PU.1 and ATF3 generated under different stimulatory conditions). Histograms of WT and KO ChIP-seq signals are depicted in different colors. In KOIN corrected data sets “hyper-ChIPable regions” are absent.

The herein presented findings suggested that “hyper-ChIPable regions” not only exist in yeast but also in other species. More importantly, the KOIN correction curated TF ChIP-seq data from “hyper-ChIPable regions” beyond the correction for tag signals from other unspecific binding events.

5.1.5. Increased identification of ChIP protein specific binding motifs after KOIN correction

The analysis of TF-binding motifs is a possible approach to evaluate the accuracy of ChIP-seq data (204,271). In highly specific TF ChIP-seq data a significant enrichment for TF binding sites is observed. In ChIP-seq data with high fractions of false-positive binding sites the binding motif for the specific target protein is not enriched or can only be found with low

statistical significance. The effect of KOIN was determined by the measurement of binomial p -values and counts for motif enrichments in WT, KO and KOIN corrected data sets (Figure 12). First of all, HOMER calculated a high and statistical significant enrichment of the SRF, GATA3 and PU.1 binding motifs in WT and KO data sets. Additionally, Jun-AP1 and AP1 motifs were highly enriched in ATF3 data sets. As both TFs build heterodimers together with the ATF3 protein, they are also enriched at the same DNA regions as ATF3 (272). The amount of PU.1 positive peaks remained unchanged, possibly due to the high specificity of the PU.1 data, whereas in all other ChIP-seq data sets the KOIN correction strongly improved relative numbers of important binding motifs. Especially in SRF data KOIN improved the SRF binding motif occurrence from 15 % to 45 %.

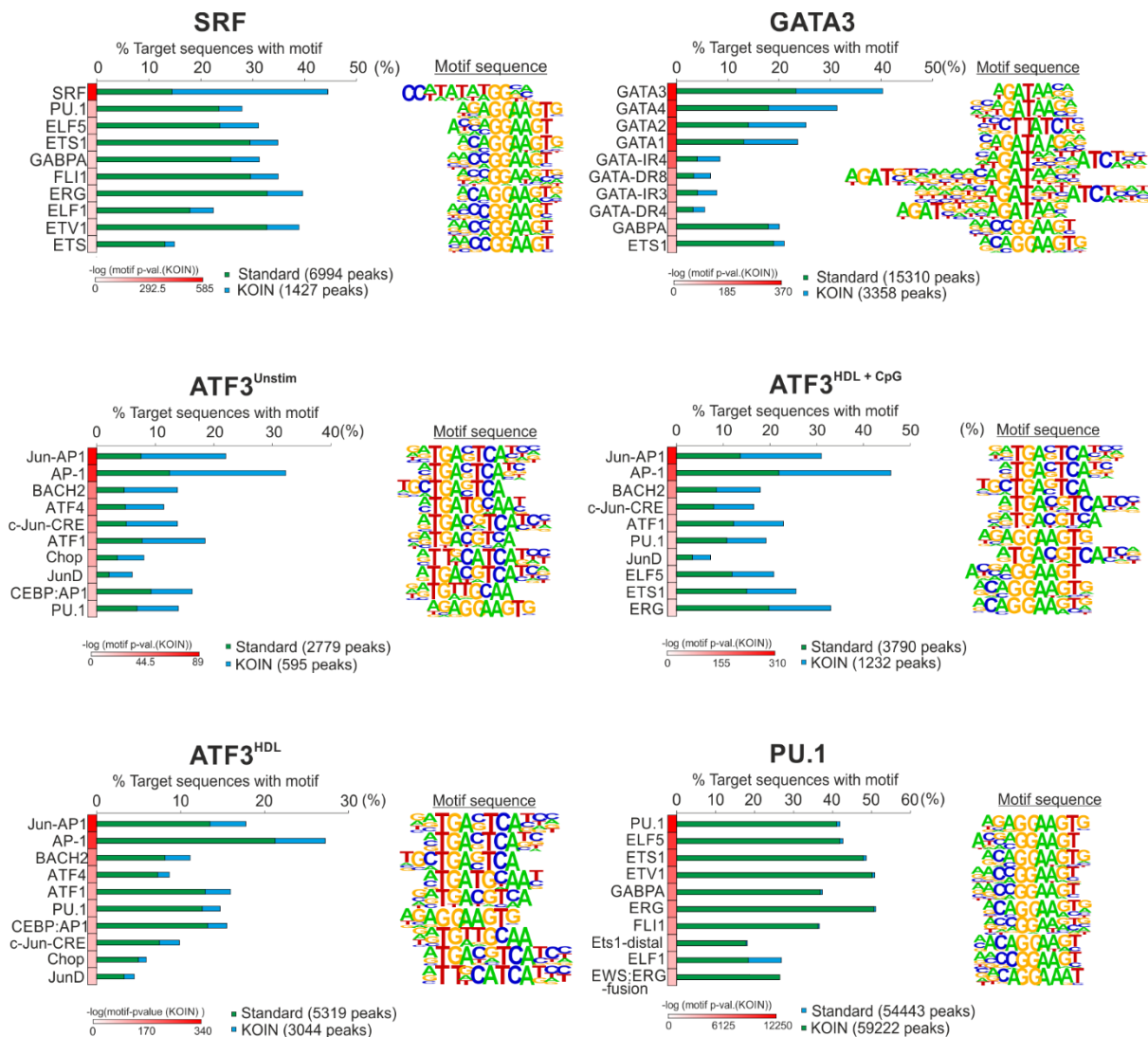


Figure 12. TF motif analysis is significantly improved after KOIN correction.

Top 10 TF binding motifs which were significantly enriched in SRF, PU.1, GATA3 and ATF3 data sets after KOIN correction were visualized and sorted according to their significance p -values. The

horizontally aligned bar plots describe the percentage of target sequences with corresponding TF motifs in each data set after KOIN correction (blue) or without KOIN correction (green). Graphical versions of positional weight matrices (PWM) for each motif are plotted to the right and KOIN motif p -values are depicted in heatmaps on the left side of each bar (white to red color).

Independent motif enrichments were also conducted solely with available ChIP-seq KO data and visualized as horizontal bar plots (Figure 13). In general, poorly enriched KO motifs lacked the main motifs of the respective TF as described before. Despite the sporadic enrichment of motifs at KO called peaks like the AR-halfsite motif in the GATA3 data set, the white colored heatmaps clearly illustrated the poor statistical significance of resulting enrichment scores suggesting putative false-positively called binding motifs.

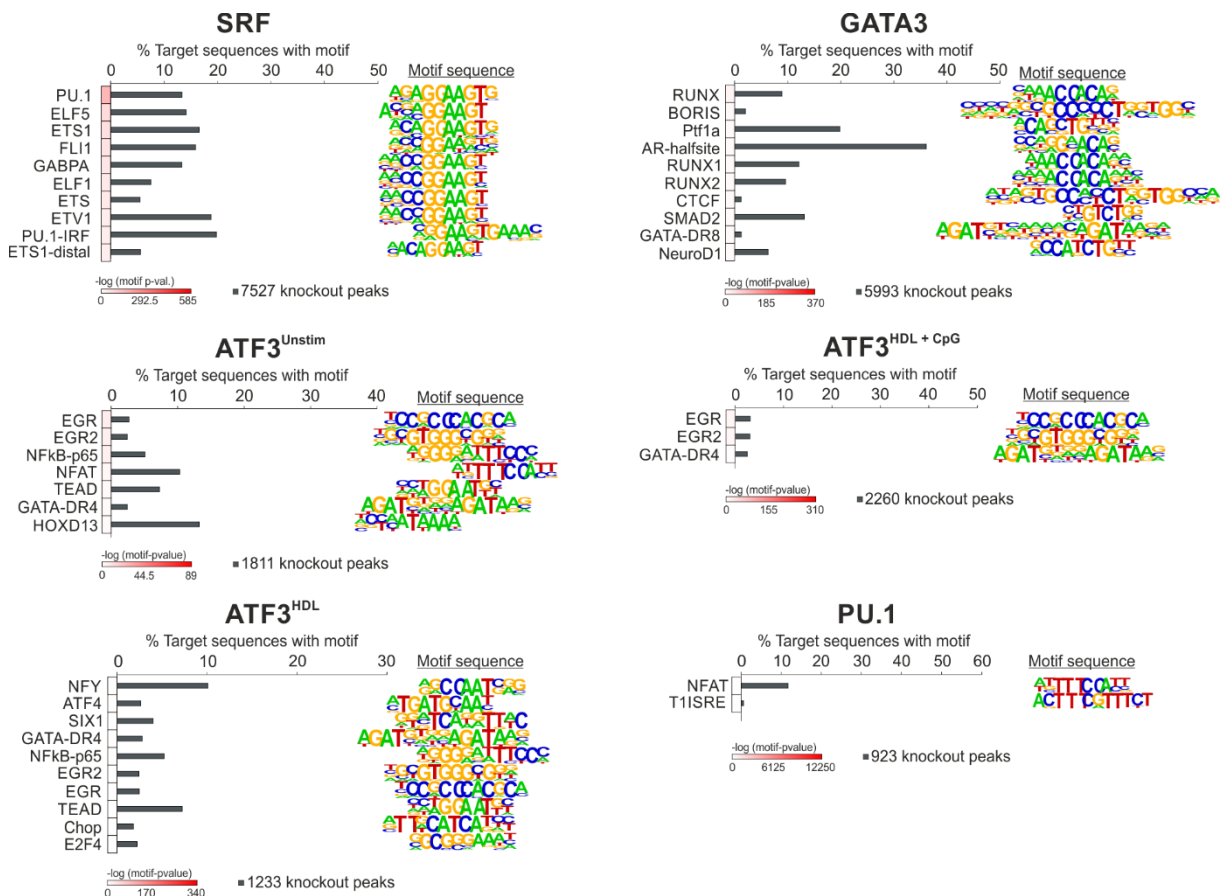


Figure 13. Absent enrichment for specific TF motifs in uncorrected KO data sets.

Percentages of target sequences with the corresponding top 10 enriched motifs exclusively at KO peak sites were ranked according to their enrichment p -values for six representative ChIP-seq data sets and depicted as horizontally aligned bar plots. Scales for p -values and percentages were used according to scales in Figure 12.

For further validation of the performed motif predictions, ratios of main motifs for independently analyzed WT and KO ChIP-seq data sets were plotted as bar charts (Figure 14). Fold change ratios were calculated for the top 5 motif counts in WT against KO peaks. In SRF and PU.1 data sets corresponding main motifs were found in the WT data set and enriched with high fold changes in comparison to KO data. Not only GATA3 but also other GATA TF family members were enriched with high fold changes exclusively in WT and not in KO data, due to high similarities in their core binding sequence (motif sequences, Figure 12). For ATF3 data sets the similarities in core sequences and the previously mentioned heterodimer formation with ATF3 caused an enrichment of Jun-AP1, AP1 and Bach2 ratios in WT data sets.

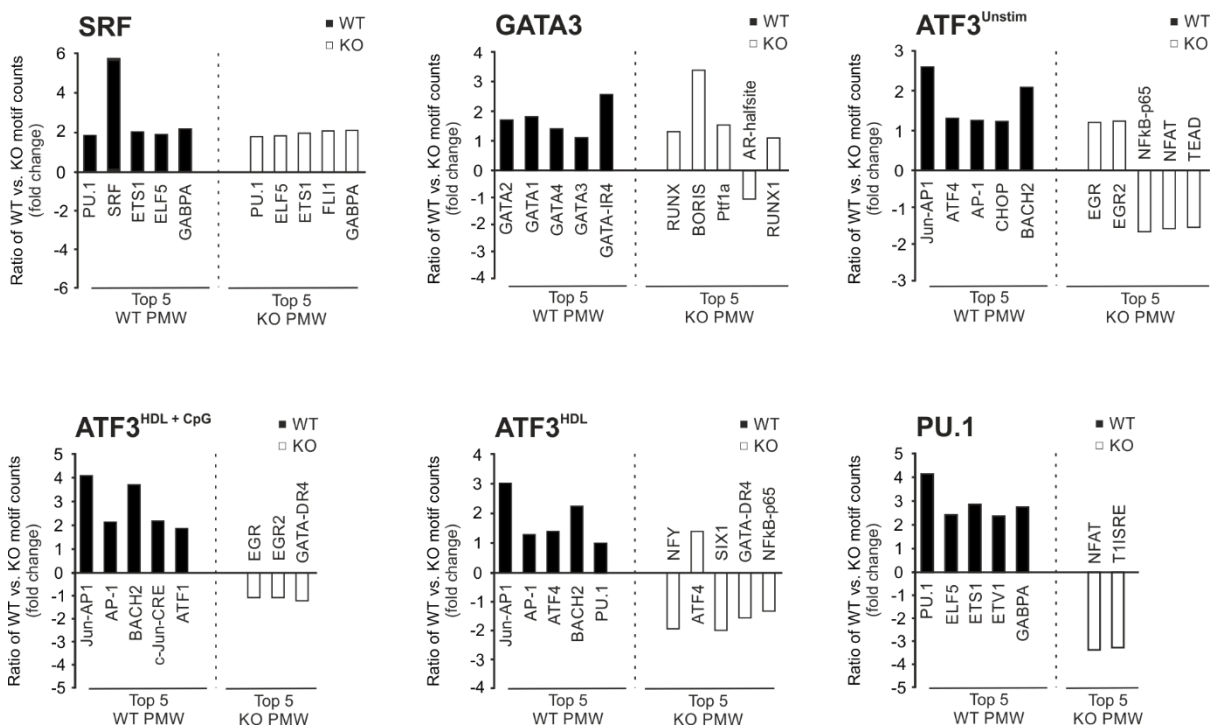


Figure 14. Enriched WT motifs are exclusively found in WT data sets compared to KO results.

Motif count ratios for the top 5 enriched motifs were depicted in WT (black) against KO (white) peak files. Motif counts were normalized to total peak counts and the ratio of WT against KO motif counts was calculated as fold changes. WT specific motifs exhibit positive ratio values and motifs with higher abundance in KO data sets show negative ratios.

In summary, motif binding predictions were altered by false-positive peaks in uncorrected WT data sets. The removal of false-positive sites drastically increased the percentages of peaks positive for the corresponding main TF motif. Additionally, comparisons for the top 5 enriched

motifs in WT against KO data sets illustrated the false-positive binding capabilities of ChIP antibodies and the enrichment of unrelated motifs without the target protein.

5.1.6. KOIN significantly enhances biological interpretation of TF ChIP-seq data

KOIN identified and significantly reduces false-positive binding signals as demonstrated in previous results (269) and in this thesis. The effects of the correction process onto the biological interpretation of ChIP-seq data was evaluated with a Gene Ontology Enrichment Analysis (GOEA) which connects gene groups with ChIP-seq enriched peak positions to known biological functions. The SRF data set was used as an example. GOEA analysis before and after KOIN correction was performed to investigate the improvement of biological interpretation by KOIN. Genes enriched for strong SRF binding within a 2 kb window centered to TSS or within the gene locus itself were utilized to create a differential network visualization for enriched GO-terms using BiNGO and EnrichmentMap plugins of Cytoscape (Figure 15). From 13 major subnetworks which were enriched in the WT SRF data set, two subnetworks were excluded (s7, s12) due to false-positive binding signals (blue edge and black border colors). Furthermore, only 5 (s16, s18, s20, s27, s28) out of 15 minor subnetworks remained after the KOIN correction process for false-positive curated genomic positions (green edge and red node colors). Therefore, KOIN correction curated the data set associated biological processes from irrelevant functions for the SRF protein and enriched significantly highly relevant biological processes the SRF protein is known for to participate in (255,273).

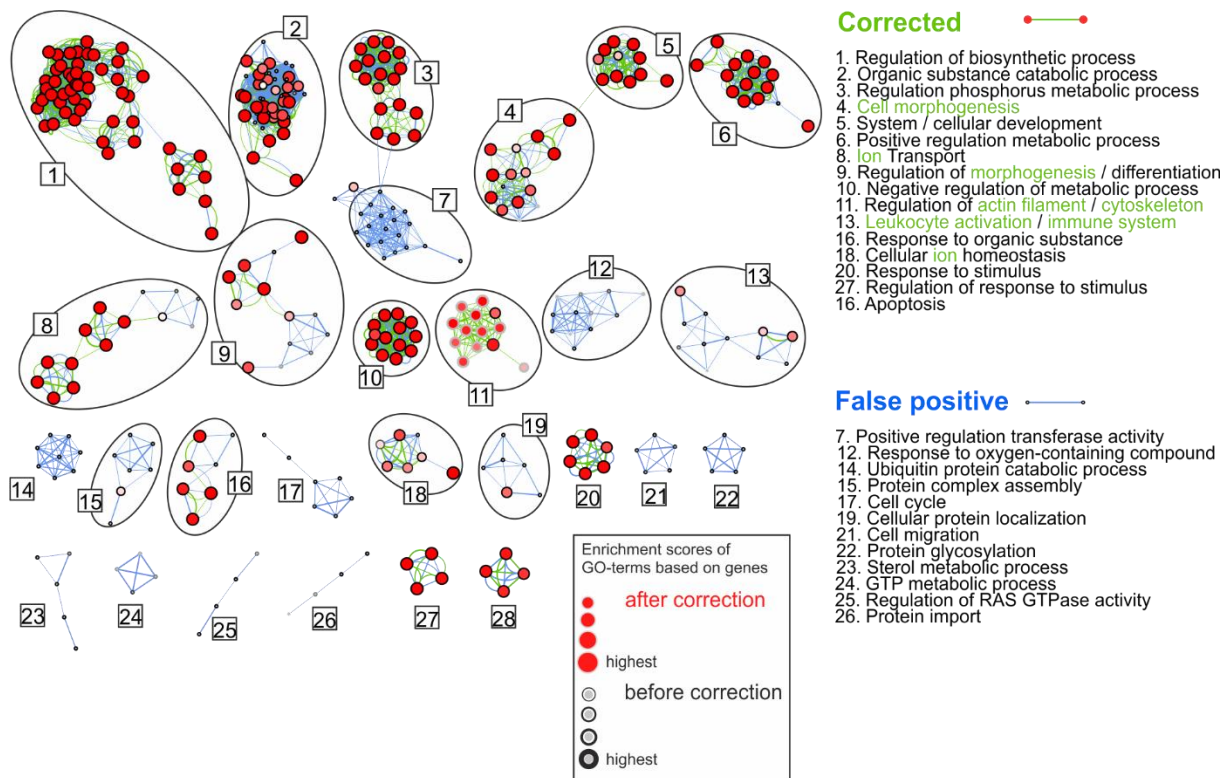


Figure 15. GO-term enrichment analysis for promoter regions is significantly improved after KOIN correction.

Gene Ontology Enrichment Analysis (GOEA) was performed for genes marked by SRF protein binding signals 1kb up- and downstream from their TSS. Results are visualized as networks for 1510 genes before KOIN correction (black node borders: GO-terms, blue edges: GO-term relations) and 327 genes after the correction process (red nodes: GO-terms, green edges: GO-term relations). The remaining true-positive GO-terms for SRF bound genomic regions after the KOIN correction are depicted as red nodes with black borders. The binomial FDR corrected p -value cutoff was set to <0.001 . The analysis was performed with Cytoscape and the plugins BiNGO and Enrichment Map. In green color marked corrected GO-terms were also found in GREAT GOEA analysis in Figure 16.

The improvement of GOEA was validated by the GREAT tool (Genomic Regions Enrichment of Annotations Tool) which allows the integration of known GO-terms also for distant regulatory sites far away from any promoter region into the biological interpretation of ChIP-seq peak signals (Figure 16) (266). GREAT utilizes stringent filter criteria by the application of not only hypergeometric but also binomial statistical models to predict meaningful biological processes for proximal and distal regulatory sites. In agreement with previous results found for GO-terms enriched at promoter regions (Figure 15) GREAT identified known SRF functions like “actin cytoskeleton dependent processes” (274) or “leukocyte activation in

immune response” (255). Other not related SRF functions like “mast-cell activation” or “cellular response to lithium ion” were excluded by KOIN.

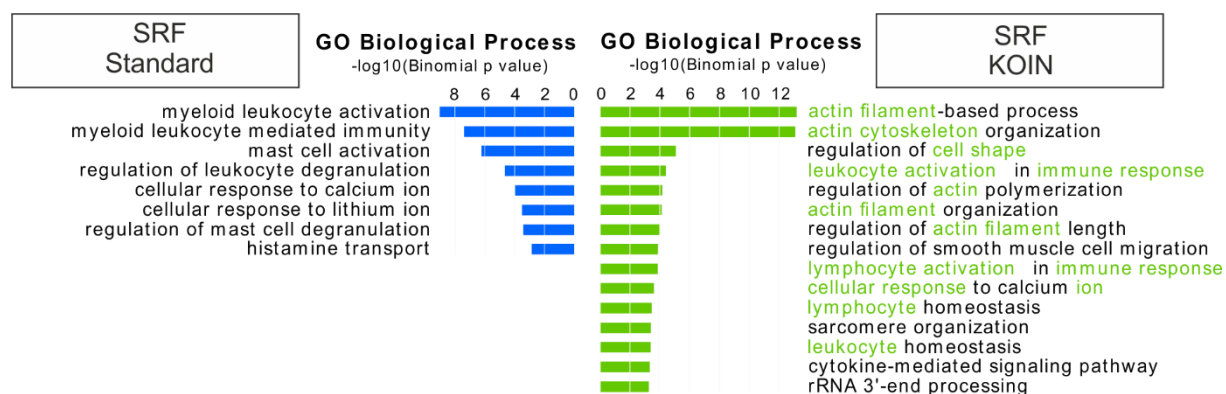


Figure 16. KOIN also improves GOEA analysis at *cis*-regulatory genomic regions.

Enriched GO-terms for *cis*-regulatory and non-coding genomic regions were calculated with the GREAT tool. Statistical stringency for significantly enriched terms was set to a FDR corrected *p*-value threshold of 0.05 using binomial and hypergeometric tests simultaneously. For the visualization the binomial *p*-value ($-\log_{10}(p\text{-value})$) of enriched GO-terms with (green) or without (blue) KOIN correction was used.

In summary, the herein presented results show that two independent tools for connecting ChIP-seq enriched genomic positions to known biological functions highly benefit from KOIN corrected data sets. Functional SRF relevant biological processes were correctly identified for KOIN corrected data, while false-positive GO-terms were excluded. Uncorrected data dilutes the true interpretation of ChIP-seq bound regions and could lead to overestimations concerning irrelevant biological functions.

5.2. Epigenetic changes during activation of primary human macrophages

The activation model for human monocyte-derived macrophages was recently extended beyond the established M1/M2 axis model to a multi-dimensional model, integrating a multitude of specific macrophage activation states equivalent to their multiple origins and employed stimulatory cues (21). The described macrophage plasticity with the establishment of different activation states displaying specialized functions, requires profound changes on the transcriptional level. Complex regulatory networks ensure the efficient and tight control of transcriptional processes. Amongst others, the epigenetic regulation of HMs can profoundly

alter the chromatin landscape and may also play a critical role as platform for the recruitment and binding of TR proteins, like TFs (167). Specific combinations of HMs can not only identify important regulatory regions like promoters or *cis*-regulatory sites in human and mouse, but can also predict the transcriptional activity at certain loci and their chromatin packaging status (68,148,151,156,158). Four different activation states of monocyte-derived primary human macrophages were used as an *in vitro* model for pro-inflammatory immune activation (21) to assess their chromatin landscapes. Four histone modifications were assessed (H3K4me1, H3K4me3, H3K27Ac, H3K27me3) in respect to the histone code hypothesis (145,146) and previously described adapted methods to identify regulatory regions of interest, including promoter and enhancer sites in up to three activity states for each defined genomic site. This epigenetic information correlated to transcriptional data from RNA-seq or microarray expression data demonstrated in detail the importance of epigenetic regulatory processes at promoter and enhancer sites during the activation of primary human macrophages and further elucidate underlying molecular mechanisms of their functional commonalities and differences.

The herein presented results (section 5.2) are also in most instances part of a scientific publication (275).

5.2.1. ChIP-seq validation procedures

In preparation of ChIP-seq experiments human monocytes were differentiated into baseline macrophages (M^b) and were further activated into $M^{IFN\gamma}$, M^{IL4} and M^{TPP} macrophages. GM-CSF was used as primary differentiation stimulus in the *in vitro* model to mimic an inflammatory milieu, often characterized by elevated systemic levels of GM-CSF (276). The complete activation of different macrophage populations was validated by the expression of typical macrophage marker proteins, like CD14 for M^b and other activation specific surface proteins like CD86 ($M^{IFN\gamma}$), CD23 (M^{IL4}) and CD25 (M^{TPP}) described previously (Figure 17) (21,253).

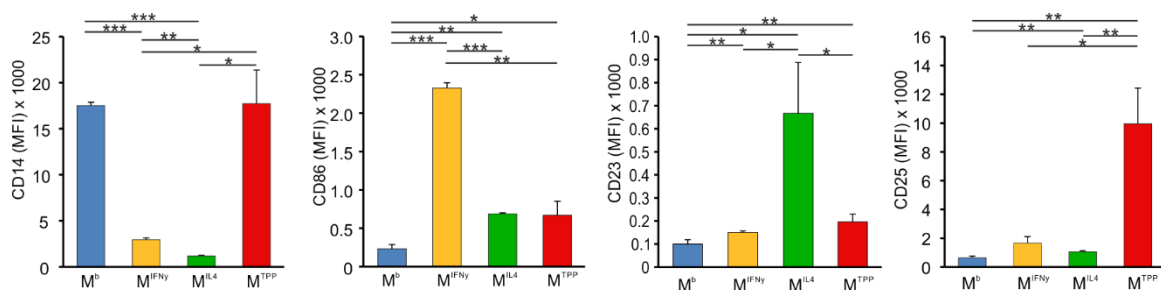


Figure 17. FACS analysis of surface marker expression in activated macrophages.

Typical macrophage surface markers used for the discrimination between certain activation states (CD14, CD86, CD23, CD25) were quantified by flow cytometry for unstimulated (baseline), IL4, IFN γ and TPP stimulated macrophages. Results are depicted as mean fluorescence intensity (MFI). * $p < 0.05$ (Student's t-test)

Chromatin of each macrophage population was sheared in an ultrasonic Covaris S220 into approximately 150 bp fragments and the exact length distributions were measured with the Agilent Bioanalyzer (Figure 18A). Suitable DNA fragments between 100 and 500 bp were purified by paramagnetic beads (Agencourt) during the library construction, while fragments with differing sizes were excluded. ChIP-seq experiments were performed as described in the methods section. The relative enrichments of known HM marked positions were quantified via semi-quantitative real time PCR for target experiments in comparison to DNA fragments enriched unspecifically with control IgG antibodies. Relative enrichments between 256 and 1024 folds for target against control DNA were detected for HM positive sites, whereas negative control regions displayed only minimal enrichments (Figure 18B).

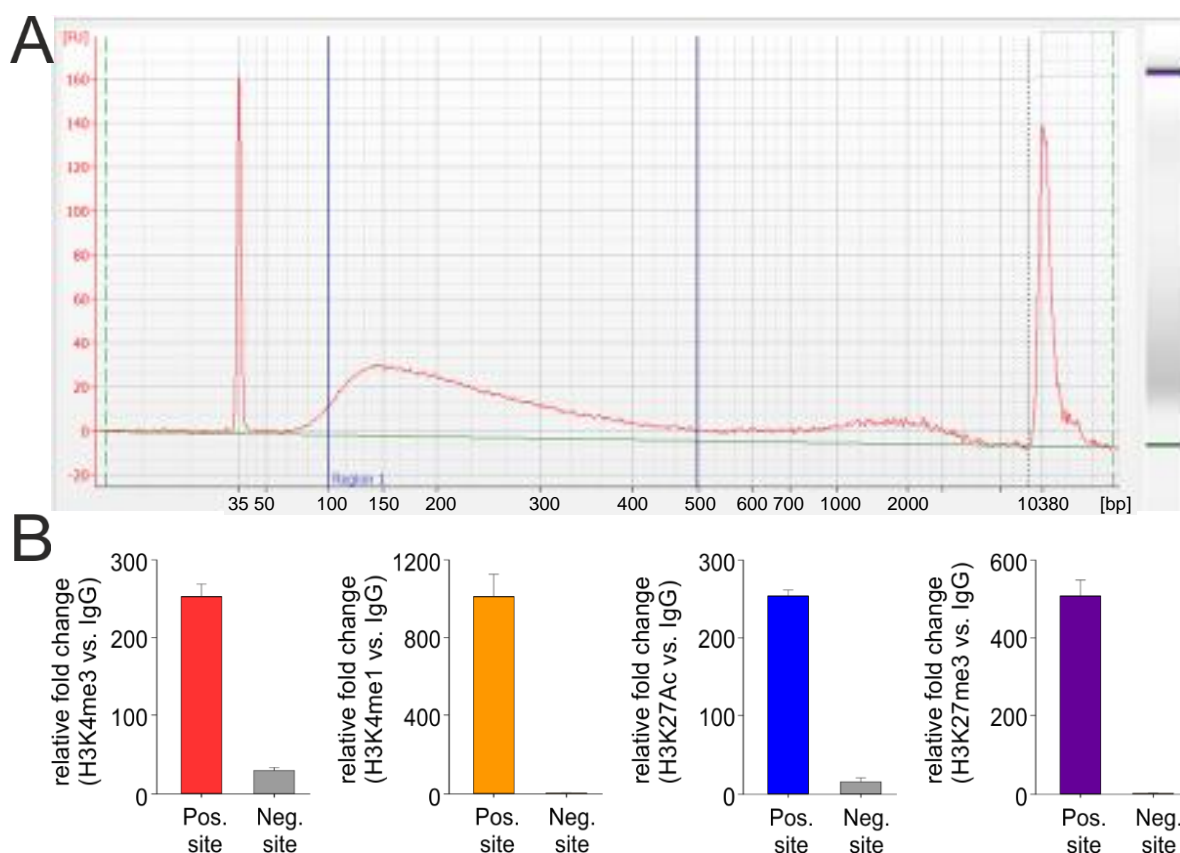


Figure 18. Quality controls for successful ChIP experiments.

(A) Representative gel electrophoresis result for cross-linked chromatin sheared for 45 minutes with chromatin sizes depicted in bp on the x-axis and fluorescence units on the y-axis. Additionally, optimal size for ChIP-seq library construction was marked with blue lines (Region 1). Lower and upper DNA markers are located at 35 and 10500 bp. **(B)** Relative qPCR enrichments of histone enriched known positive and negative sites are depicted as fold changes against background controls (IgG). Results represent the mean \pm standard deviation of at least three independent experiments and were performed for four different ChIP-seq data sets: H3K4me3 (red), H3K4me1 (orange), H3K27Ac (blue), H3K27me3 (violet).

5.2.2. Optimization of next generation sequencing library construction

Prior to ChIP-seq library construction experiments with primary human macrophages, jurkat cell DNA (Human T cell lymphoblast-like cell line, Clone E6-1) was isolated and sheared to assess the minimal necessary amount for efficient ChIP-seq library construction. These experiments were especially necessary due to the limited amount of isolated CD14 positive human monocytes per blood donor to establish the different macrophage activation states. For the library construction, between 10pg and 10ng chromatin was used as input starting material to construct 7 multiplex ChIP-Seq libraries. Constructed libraries were sequenced on a HiScan SQ machine (Illumina). Computational steps like demultiplexing and alignment to the reference genome were performed. Finally, sequencing results in form of tag counts normalized to total counts for each experiment were compared in circos plots for each chromosome to assess the optimal chromatin amount to cover the brought variety of genomic positions during ChIP-seq library construction and subsequent sequencing (Figure 19).

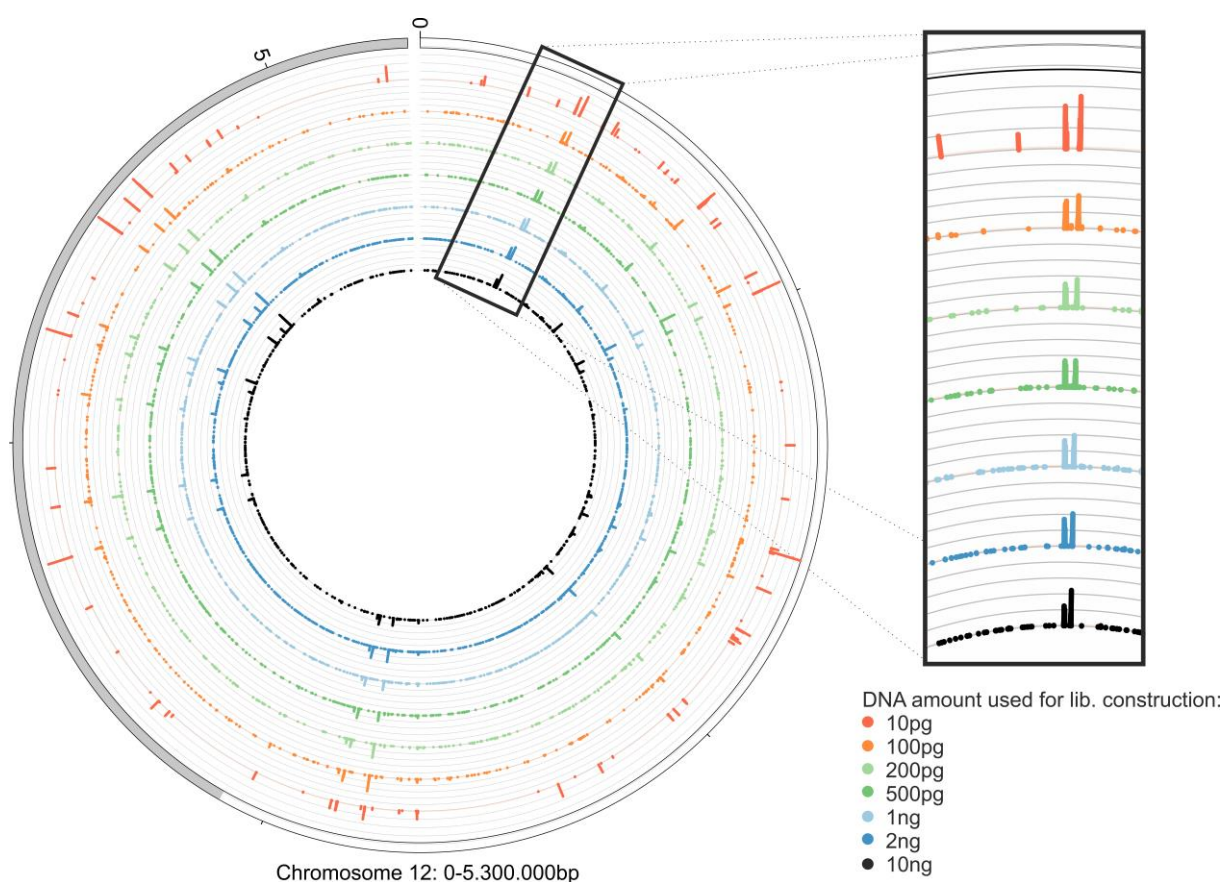


Figure 19. Sequencing results for ChIP-seq libraries constructed with variable amounts of DNA input material.

Normalized tag counts for 7 multiplex ChIP-seq libraries were visualized as circos plots and sorted according to their library construction starting material (10, 2, 1, 0.5, 0.2, 0.1, 0.01 ng DNA from Jurkat cells). Circos plot depicts normalized tag counts on chromosome 12 from position 0 to 5.3 million bp.

Despite manufacturer's specifications (Illumina) to use 1 μ g DNA as starting material for ChIP sequencing, 0.5-1ng of DNA for the library construction already displayed a broad distribution of tag signals for the example position on chromosome 12. Lower amounts of DNA lead to tremendously decreased sequencing signal resolutions, because of absent ChIP DNA pieces during library construction for various loci. In addition, over amplifications of ChIP-seq signals for specific DNA regions on chromosome 12 were observed for libraries constructed with 10pg input DNA. These signals are inevitable, if only low amounts of DNA are available and can falsify subsequent analysis steps.

Based on the above described optimization, 0.5-1 ng of sheared chromatin isolated from human macrophages was utilized for all following ChIP-seq library constructions described in the methods section. In addition, representative size distributions of constructed libraries

indicated correct sizes between 250 and 450 bp as recommended from the manufacturer (Illumina) for the optimal performance of next generation sequencing experiments (Figure 20).

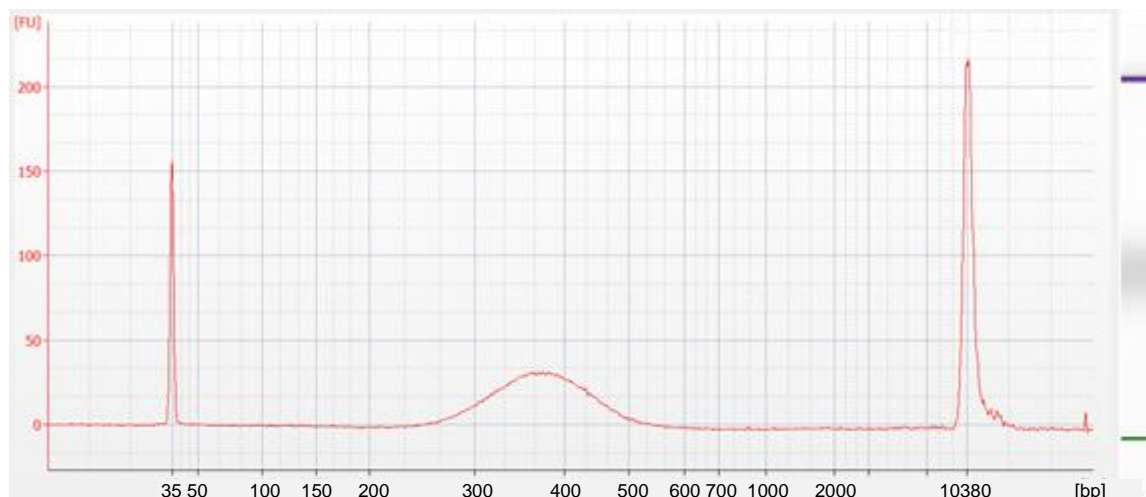


Figure 20. Representative size distribution of macrophage ChIP-seq library.

Representative DNA size distributions of successfully constructed ChIP-seq libraries (Agilent Bioanalyzer). The x-axis represents the chromatin size in bp, while the y-axis depicts fluorescence units of measured DNA fragments labelled with an intercalating dye. Lower and upper DNA markers are located at 35 and 10500 bp

Applied pipelines for the following downstream analysis of next generation sequencing data included the separation of data sets (demultiplexing), the alignment of unknown sequenced pieces to the human reference genome and the identification of regions significantly enriched for ChIP-seq signals (peak calling) finally determining locations of specific HM signals. Furthermore, the usage of 0.5-1 ng of sheared chromatin ensured the correct ChIP-seq library construction.

5.2.3. Epigenetic classification of specific promoter and enhancer states during human macrophage activation

Transcriptional processes and their regulation are tightly connected to changes on epigenetics level. Exogenous signals can regulate transcriptional processes by the modification of histone proteins not only at promoter but also at enhancer sites (73,277). Histone proteins, targets of different posttranslational modifications play a major role in these

changes and act as linkers to integrate amongst others chromatin organization, TF binding and transcription. In which extend HMs lead in primary human macrophages to their specialization adapting to environmental cues still needs to be elucidated, especially in primary human macrophages.

Previous publications lead to the genome wide identification of transcriptional regulatory sites like promoters or enhancers and the definition of activity states by the combination of different posttranslational HMs (Figure 22) (148,158,159). Accessible and poised promoters (Pa/Pp) as well as strong, weak and poised enhancers (Es, Ew, Ep) could shape the transcriptome of differentially activated primary human macrophages.

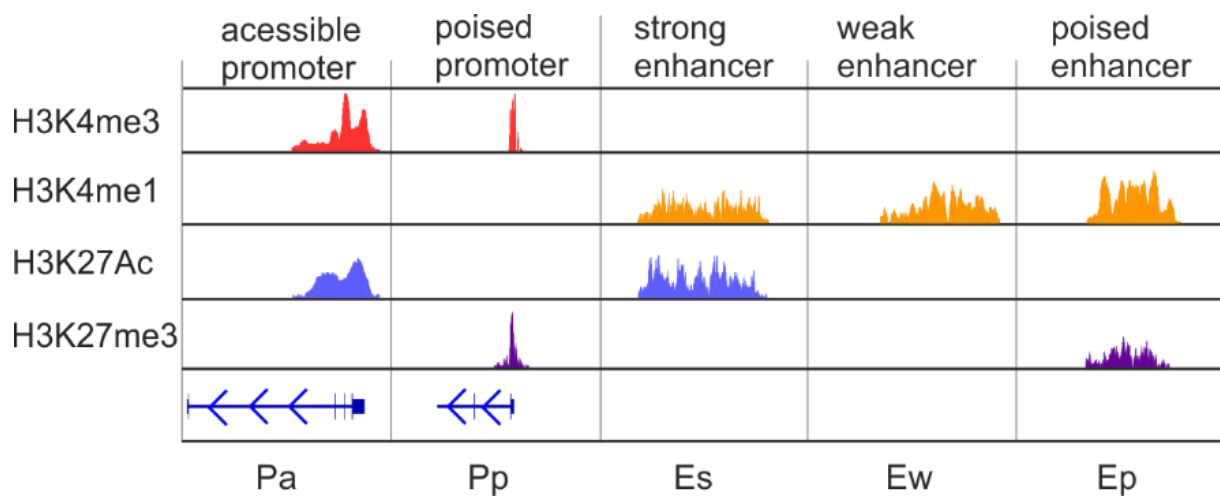


Figure 21. Definition of epigenetic states at promoter and enhancer sites.

Combinations of four HMs (H3K4me3, H3K4me1, H3K27Ac, H3K27me3) define the activity states of promoters (accessible Pa, poised Pp) and enhancer regions (strong Es, weak Ew, poised Ep).

Activation-signal-specific changes of the epigenetic landscape in human inflammatory macrophages were defined following two rules: First, specific loci were at least two times higher enriched for H3K4me3 or H3K4me1 total tag counts than the input control datasets. Second, tag signals for a specific H3K4me3 or H3K4me1 locus were increased at least two times in the respective activation state in comparison to the other activated macrophages. 430 activation state specific accessible promoters were identified (Figure 22A) by following the defined criteria during global analysis. Heatmaps clearly visualized the concomitant enrichments for H3K4me3 and H3K27Ac HM signals for Pa sites (Figure 22B). Pa numbers reflected the activation state of differentially activated macrophages with low numbers in M^b and increasing numbers in macrophages stimulated with different molecules (M^{IFN γ} , M^{IL4},

M^{TTP}) with highest numbers in pro-inflammatory $M^{IFN\gamma}$. In contrast to Pa sites, in total 459 Pp were marked with H3K4me3 and repressive H3K27me3 signals, while lacking the active H3K27Ac HM modification. Highest numbers of Pp in M^{IL4} could suggest the importance of promoter inactivation, necessary for the IL4-induced transcriptional program. M^{TTP} macrophages showed almost equal numbers for Pa and Pp, different to $M^{IFN\gamma}$ with higher numbers for Pa. Another interesting observation was the mutually exclusive enrichment of active H3K27Ac marks and repressive H3K27me3 HM at accessible and poised promoters.

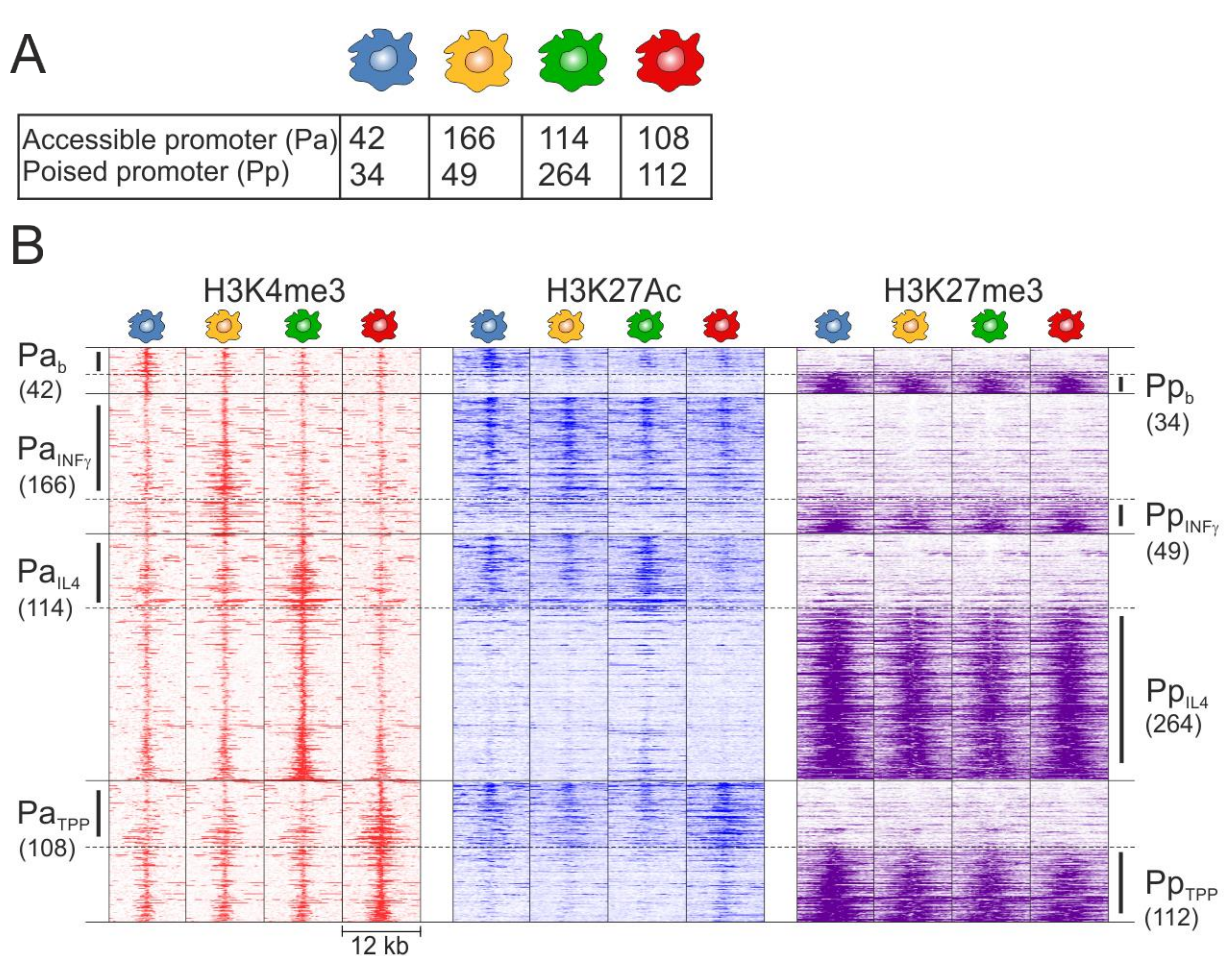


Figure 22. Activation state specific epigenetic promoter states in human macrophages.

(A) Numbers of macrophage activation specific promoter sites. **(B)** Heatmaps of normalized histone ChIP-seq tag counts at macrophage activation specific genomic positions for defined promoter states. Pa sites show simultaneous H3K4me3 and H3K27Ac signals, while Pp are characterized by H3K4me3 and H3K27me3 HM signals. Heatmaps were centred to H3K4me3 peaks and ChIP-seq signals were determined in 500 bp windows 6 kb up- and downstream of peak middle points. (M^b /blue, $M^{IFN\gamma}$ /yellow, M^{IL4} /green, M^{TTP} /red)

Promoter or enhancer sites need to be accessible for TRs and in an open chromatin state to allow actual gene transcription. In macrophages, the pioneer TF PU.1 is essential for the opening of closed inactive chromatin sites and mediates with the co-binding of other secondary TFs thereby enabling the transcription of corresponding genes (75,278,279). Previous publications assessed PU.1 binding mainly at *cis*-regulatory enhancer sites in macrophages, but PU.1 binding motifs are also present at promoter sites. In fact, in differentially activated macrophages approximately half of the Pa sites contained the PU.1 consensus motif (Figure 23A), except for Pa sites in M^b which showed even higher PU.1 motif rates of 71%. This finding could hint to the hypothesis that PU.1 binding is not a prerequisite for the activity of stimulus-specific macrophage promoters. Additionally, genes with Pa sites revealed gene ontology (GO) terms like “immune response”, “chemotaxis” or “inflammatory response”, correlating well with known functions as exemplified for M^{TPP}.

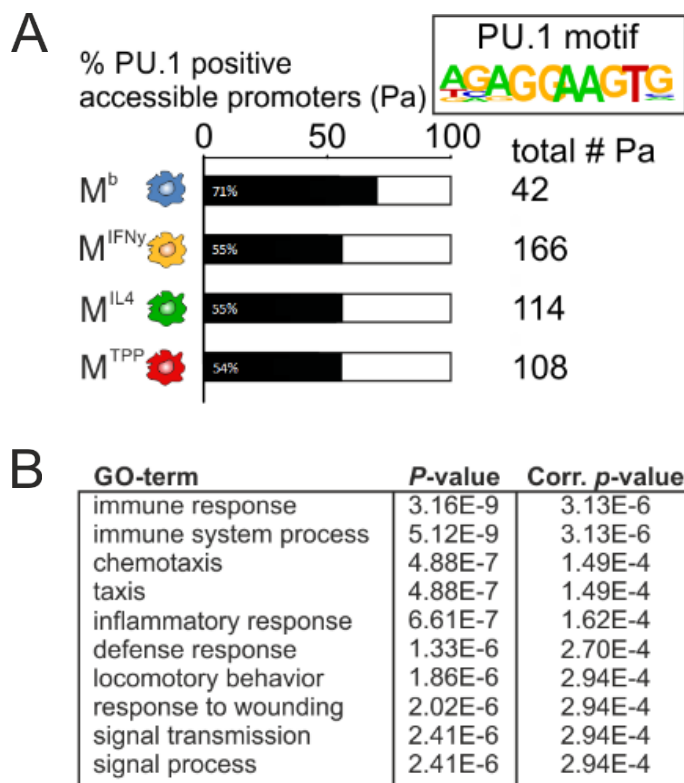


Figure 23. PU.1 frequencies at stimulus specific accessible M^{TPP} promoters and GO-terms related to typical macrophage functions.

(A) Percentages of accessible promoters enriched for the PU.1 consensus motif in differentially activated macrophages. Positional weight matrix of PU.1 motif is depicted in the upper right corner. **(B)** Gene set enrichment analysis (GOEA) for 108 genes marked as accessible promoters (Pa) in TPP activated macrophages (M^{TPP}). Corresponding *p*-values and corrected *p*-values were listed to the right (hypergeometric test, Benjamini & Hochberg correction).

The correlation of epigenetic promoter states to RNA-seq data allowed the assessment, if HM promoter markings have an effect on transcriptional activity of genes in primary human macrophages. Indeed, a significant increase in transcriptional activity was detected in all macrophages activated by the indicated stimuli, when genes with accessible or poised promoters were compared (Figure 24A). Furthermore, genes with specific accessible or

poised promoter marks were specific for the corresponding macrophage activation state (Figure 24B). M^{TPP} macrophage genes marked with Pa were significantly increased in their expression values in M^{TPP} in comparison to other macrophage activation states and showed stronger enrichments for H3K4me3 and H3K27Ac in comparison to decreased levels in other macrophages, especially for the HM H3K4me3 (Figure 22B). Poised promoter genes were in general drastically reduced in their expression values in comparison to accessible promoters, already stated previously. However, the less pronounced decrease in the expression rates for Pp in M^{TPP} in comparison to other activation states could be present due to the higher H3K4me3 signals in M^{TPP} almost diminished in the other macrophages. Comparable specific results were detected for the other three activated macrophages (data not shown).

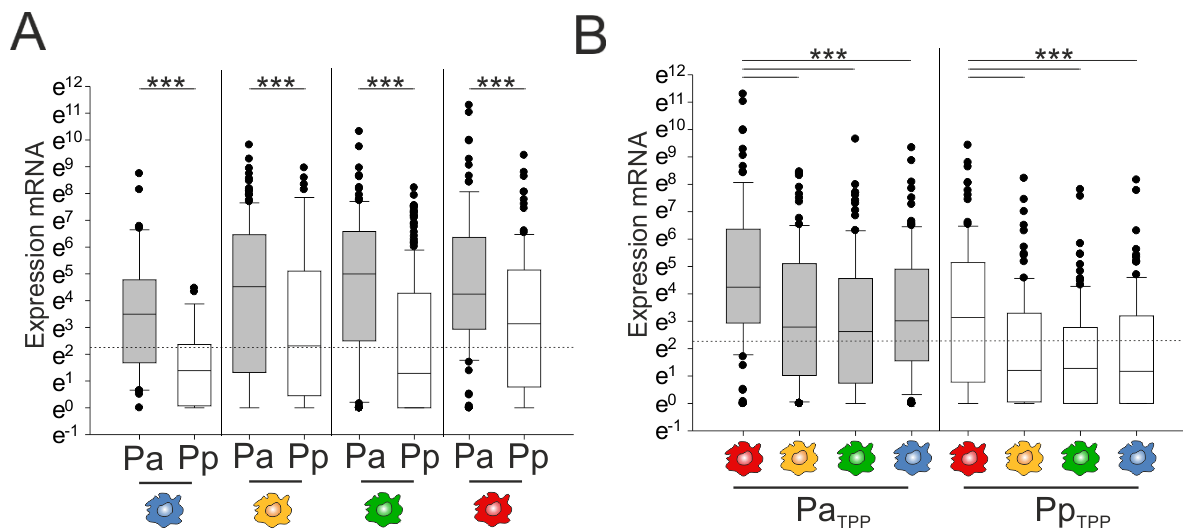


Figure 24. Boxplots of expression values for genes marked with stimulus specific accessible and poised promoters.

(A) Boxplots of RNA-seq expression values for genes marked with macrophage activation state specific accessible or poised promoters. Dashed line represents cut-off values for genes without gene expression. **(B)** RNA-seq expression values for M^{TPP} genes associated with specific accessible or poised promoters in all four macrophage activation states. (M^b /blue, $M^{IFN\gamma}$ /yellow, M^{IL4} /green, M^{TPP} /red). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (wilcoxon rank sum test)

Cis-regulatory enhancer sites, crucial for transcriptional regulation and maintenance act as platforms bound by TFs and other proteins and can mediate DNA looping to subsequently stimulate promoter activation and gene transcription (58,59,67). In addition, recently defined “latent enhancers” in murine macrophages suggested that exogenous signals can shape the epigenetic landscape (249). Therefore, three activity states of enhancers (strong, weak and

poised enhancers) were defined by the combination of three HM signal enrichments (Figure 21). In total 2024 Es, 15754 Ew and 2390 Ep sites specific for the macrophage activation states were identified (Figure 25A). Heatmaps illustrate HM signal combinations for Es with H3K4me1 and H3K27Ac signal enrichments, solely H3K4me1 for weak enhancers and a combination of H3K4me1 and H3K27me3 for poised enhancer sites (Figure 25B). Interestingly, Ew represented in numbers the largest group of defined enhancer states and thus could suggest their importance in shaping the activation associated transcriptome to induce stimulus-specific macrophage functions (68,160,280). Especially the M^{IFN γ} macrophage with the highest numbers of Es could also demonstrate the strong impact of dominant inflammatory signals on the activity of regulatory sites and their chromatin organization. Lower Es numbers in M^{TPP}, despite the inflammatory signals during activation could once more reveal negative feedback loops, due to the chronic inflammatory setting. M^{IL4} macrophages showed increased numbers of Ep sites, which is in line with previously described high numbers for Pp, emphasizing the importance of repressive mechanisms for anti-inflammatory macrophages. For answering the question if stimulus-specific enhancers require PU.1 presence for their activity PU.1 binding predictions were performed. 12 % to 54 % of described Es sites were positively marked with the PU.1 consensus motif, while Ew sites were slightly reduced in PU.1 frequencies (12 % - 32 %) (Figure 25C). Previously published results for mouse macrophages stated comparable results for the enrichment of the PU.1 motifs at H3K4me1 positive enhancer sites (157). These findings suggest - in line with previous results (281) - that the presence of PU.1 is not always required for stimulus-specific enhancers and that additional regulators are involved. TF binding predictions revealed significant enrichments for motifs like PU.1 in M^b, IRF1 in M^{IFN γ} , STAT6 in M^{IL4} (data not shown) and FOXL2 in M^{TPP}. Exemplarily, the top five enriched motifs for M^{TPP} were shown including enrichment p-values for expressed TFs at M^{TPP}-specific enhancer regions (Figure 25D).

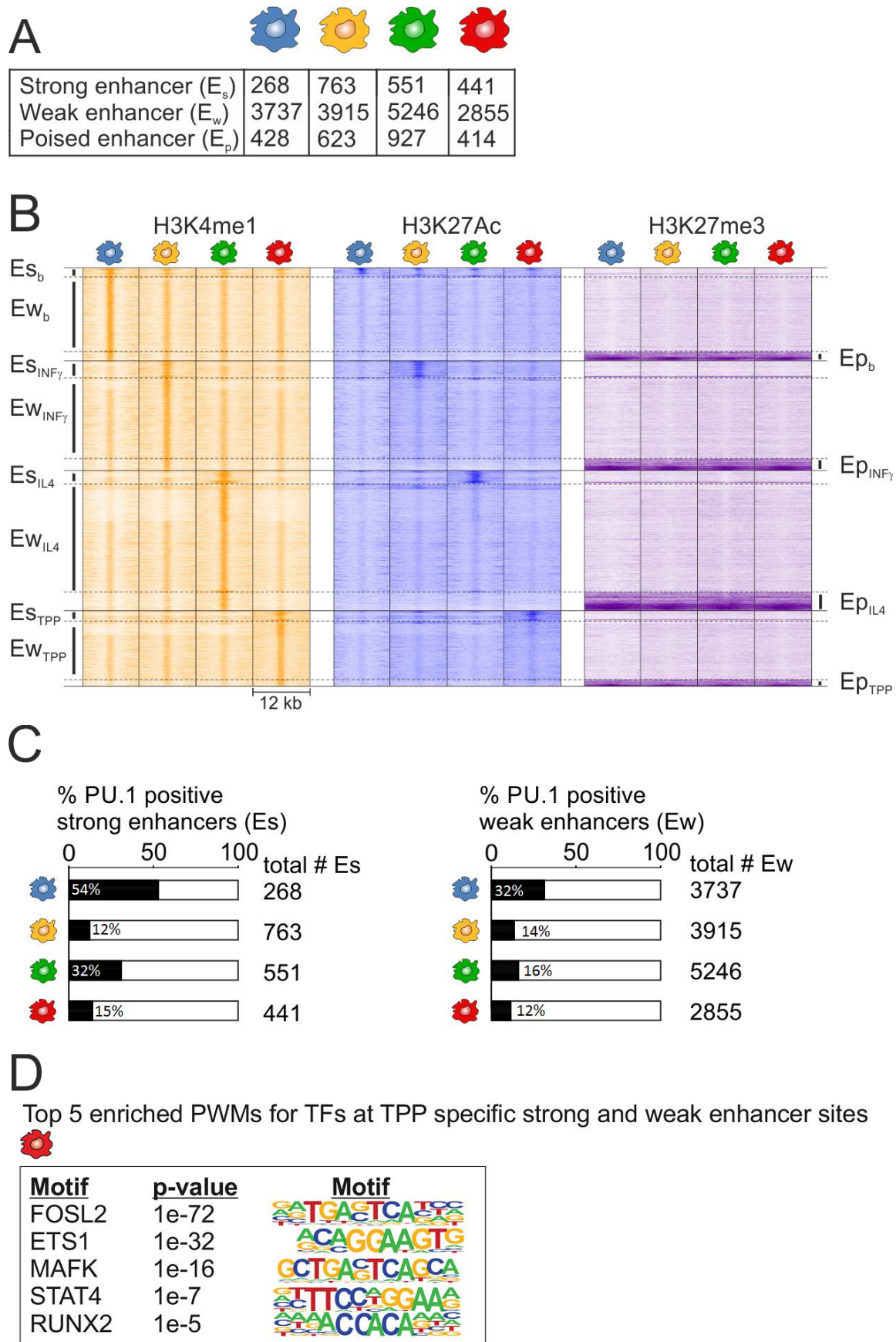


Figure 25. Identification of specific epigenetic enhancer states and corresponding PU.1 frequencies in four macrophage activation states

(A) Numbers of specific enhancer sites in differentially activated macrophages. **(B)** Normalized ChIP-seq tag counts of genomic positions for strong (Es), weak (Ew) and poised (Ep) enhancers plotted as heatmaps. All enhancers share the H3K4me1 signals, while Es are marked with additional H3K27Ac and poised enhancers with H3K27me3, respectively. Heatmaps were centred on H3K4me1 peak signals and assessed in 500 bp windows 6 kb up- and downstream of the peak midpoints **(C)** Frequencies for PU.1 positive enhancer sites at Es and Ew depicted for all four macrophage activation states in percent. Absolute gene numbers were plotted to the right. (M^b/blue, M^{IFN γ} /yellow, M^{IL4}/green, M^{TPP}/red)

To assess a correlation of the epigenetic enhancer status at a genomic locus to the actual transcriptional activity of adjacent genes, RNA-seq data were combined to enhancer states of known genes (Figure 26). Genes in the vicinity of Es sites demonstrated high expression levels, while genes with Ew were in comparison to Es significantly down-regulated in M^{IFN γ} and M^{IL4} and showed comparable tendencies in the other activation states. In contrast, Ep enriched sites were strongly correlated to genes with low transcriptional activity.

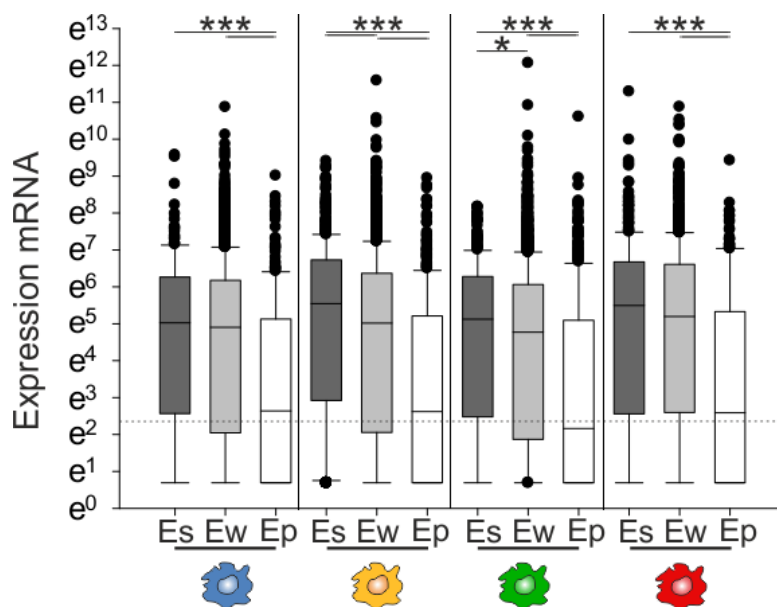


Figure 26. RNA-seq expression values of genes marked with activation specific enhancers.

Boxplots of RNA-seq expression values for genes marked with macrophage activation specific enhancers with strong (Es), weak (Ew) or poised (Ep) status. Dashed line represents the cut off for expressed genes. (M^b/blue, M^{IFN γ} /yellow, M^{IL4}/green, M^{TPP}/red). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (wilcoxon rank sum test)

Two examples for genomic sites with M^{TPP} specific Pa or Es sites and high expression values in RNA-seq data were visualized (Figure 27). Normalized HM tag counts were explicitly enriched at marked positions with at least two times higher tag counts in M^{TPP} compared to other macrophage activation states. The example genes IL2RA and CXCL1 neighboring specific M^{TPP} promoter and enhancer regions were highly expressed, specifically in M^{TPP}

macrophages and were also previously linked to chronic inflammation in macrophages (21,282).

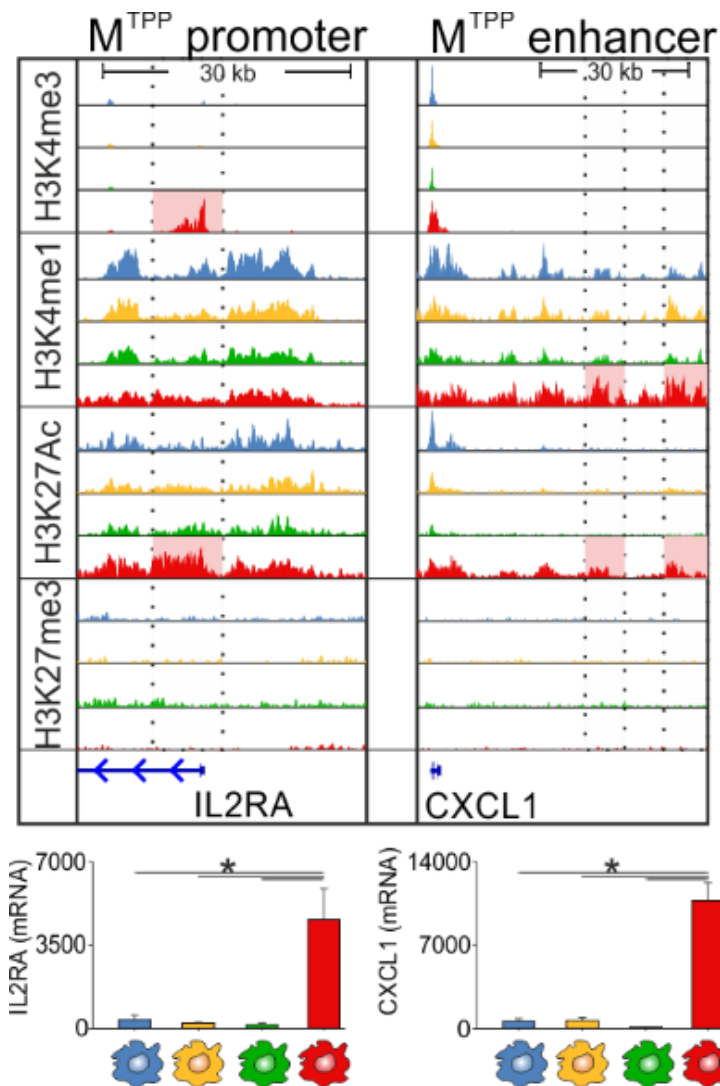


Figure 27. Visualization for representative sites in M^{TPP} with activation specific accessible promoter and strong enhancers.

Normalized ChIP-seq signals for different histone modification data sets (H3K4me3, H3K4me1, H3K27Ac and H3K27me3) are shown for an accessible promoter (left panel, IL2RA) or a strong enhancer (right panel, CXCL1). Locations of enriched ChIP-seq signals defining corresponding promoter or enhancer states are marked with red boxes. Corresponding RNA-seq expression values for M^{TPP} signature genes are depicted as box plots below the genomic visualization window. * p < 0.05 (wilcoxon rank sum test). (M^b/blue, M^{IFN γ} /yellow, M^{L4}/green, M^{TPP}/red)

Ongoing gene transcription is determined by the combination of promoter and enhancer activity. Hypergeometric probability tests, comparing pairs of gene groups marked by different epigenetic classes of promoters and enhancers, allowed the estimation of their relationships. Hypergeometric *p*-values smaller than 0.05 define a positive correlation between gene groups (Figure 28A). A strong correlation of Pa to Es and Ew and of Pp to Ep was detected especially for M^{L4} and M^{TPP}. In M^{IFN γ} , positive correlations between mentioned sites were less pronounced but still significantly detectable. Low correlations of Pa and Pp in M^b and Pp in M^{IFN γ} to corresponding enhancer sites could be caused by the relatively small

numbers of specific promoter sites (< 50) (Figure 22A). To assess the influence of combinations for promoter and enhancer markings at corresponding macrophage activation specific genes onto transcriptional activity, expression values for Pa genes with and without Es or Ew marks were compared (Figure 28B). Genes with additional marks for Es or Ew displayed elevated median expression levels in all macrophage activation states. Accessible and active promoters with already expressed genes (expression above dashed line) were especially increased by numbers in activated macrophages ($M^{IFN\gamma}$, M^{IL4} , M^{TPP}) due to the putative interactions with Es.

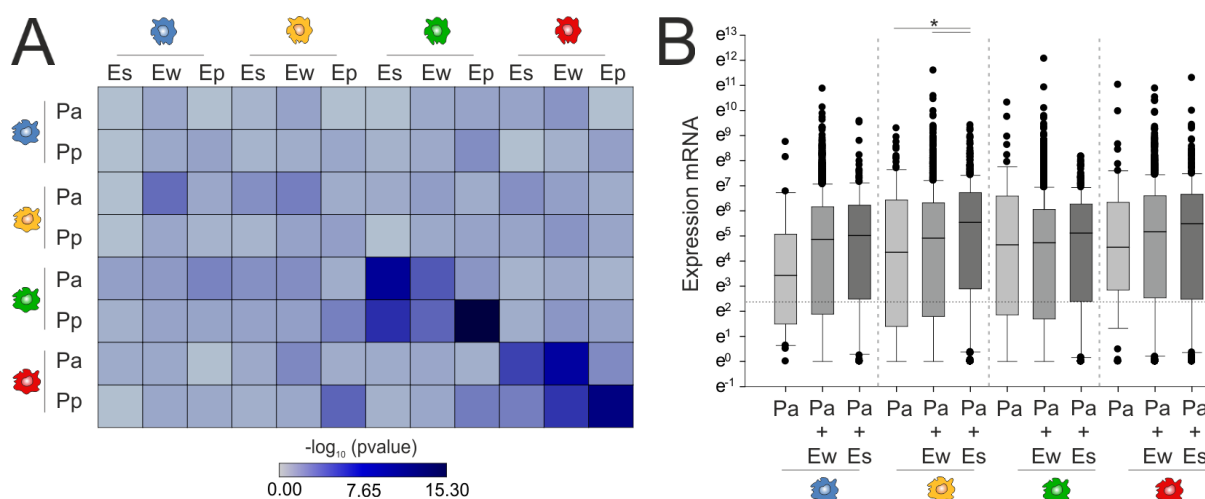


Figure 28. Correlation analysis of promoters and enhancers with different epigenetic activity markings.

(A) Pairwise correlations (Hypergeometric test) between gene groups marked with macrophage activation specific promoters (accessible Pa, poised Pp) and enhancers with different epigenetic activity states (strong Es, weak Ew, poised Ep). **(B)** RNA-seq expression values for genes marked with an Pa alone or with simultaneous Es or Ew markings are depicted as box plots for four macrophage activation states. (M^b /blue, $M^{IFN\gamma}$ /yellow, M^{IL4} /green, M^{TPP} /red). * $p < 0.05$ (wilcoxon rank sum test)

Collectively, macrophage activation specific genes with different epigenetic promoter and enhancer states were identified. These genes may shape the transcriptome of activated macrophages and mediate their specialization to fulfill different functions upon stimulation. The expression values of genes marked with Pp or Ep were significantly attenuated, while Pa genes with or without additional enhancer markings revealed high transcriptional rates in most cases. Mentioned exceptions could implicate additional regulatory mechanisms by fine tuning the activation-specific macrophage transcriptomes.

5.2.4. Epigenetic core program in differentially activated macrophages

Despite the specific differences of activated macrophages on epigenetic level, a high fraction of similarities was previously found for macrophages with different origins and stimulated under different conditions (21). Recent computational reverse engineering approaches utilizing 28 differentially activated human macrophages allowed the definition of a so called “macrophage core signature”, describing actively transcribed macrophage genes important for their identity.

Adapting the idea of a “macrophage core signature” for the regulation on epigenetic level, the HM status of promoter and *cis*-regulatory sites for macrophage activation states was assessed and sorted for genomic sites with common activity states. 7427 common accessible promoters (Pa_{common}) and 1247 common poised promoters (Pp_{common}) could be identified and were visualized as heatmaps (Figure 29A). Furthermore, a group of common enhancers was identified in all four macrophage activation states and subdivided into 3731 strong enhancers (Es_{common}) and 467 poised enhancers (Ep_{common}) (Figure 29B). Additionally, 3110 weak enhancers (Ew_{common}) were found. Common Pa and Es were defined as “common histone core signature” (CHCS), found in activated macrophages. Previous results comparing murine macrophages, neutrophils and monocytes corroborate the demonstrated strong overlap in the usage of accessible promoters in contrast to strong enhancers (157). The CHCS seems to be predominantly characterized by a vast number of promoters and to a lesser extent by *cis*-regulatory sites thereby capable of the establishment and maintenance of a common macrophage program.

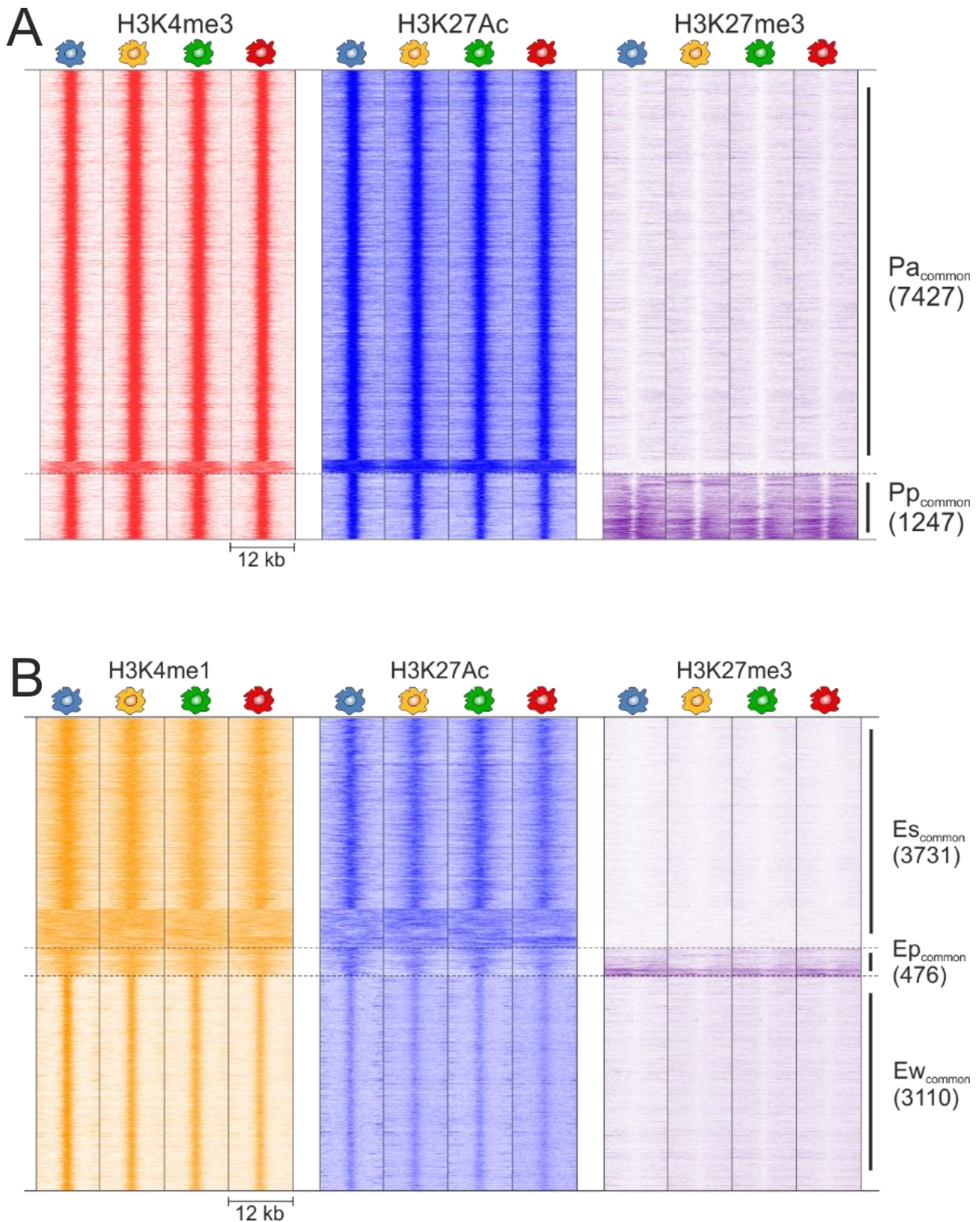


Figure 29. Common accessible and poised promoters in activated macrophages.

Heatmaps of normalized ChIP-seq tag counts for genomic positions with common **(A)** accessible ($P_{a_{\text{common}}}$) and poised ($P_{p_{\text{common}}}$) promoters or common **(B)** strong ($E_{s_{\text{common}}}$) and weak ($E_{w_{\text{common}}}$) enhancers in four macrophage stimulatory conditions. Heatmaps were centred on (A) H3K4me3 or (B) H3K4me1 peak signals and assessed in 500 bp windows 6 kb up- and downstream of peak middle points. (M^b /blue, $M^{IFN\gamma}$ /yellow, M^{L4} /green, M^{TPP} /red)

To verify, that genes with accessible or poised promoter state are expressed similarly in all investigated macrophage activation states or are absent, RNA-seq experiments were performed for activated macrophages. Expression values of genes with common Pa are plotted in comparison to genes with common Pp marks as boxplots for each macrophage activation state separately (Figure 30A). Promoters with common active HM signals were significantly higher expressed compared to genes with poised promoter sites. Comparable calculations were performed for genes with common enhancer marks found in all macrophage activation states (Figure 30B). Genes adjacent to *cis*-regulatory enhancer sites with Es correlated to increased expression values, while genes marked with Ew and Ep were significantly down-regulated. Interestingly, genes with poised promoter states were in a lower degree down-regulated in their expression in comparison to genes with a poised enhancer in their vicinity. In case CHCS genes are particularly important and moreover specific for macrophages, higher frequencies for PU.1 motifs at CHCS sites should be visible at corresponding promoter and enhancer sites, since PU.1 as macrophage lineage determining factor would initiate chromatin remodeling processes and permit the accessibility of gene loci (Figure 30C). Indeed, 89 % of Pa sites and 96 % of strong enhancers contained the PU.1 consensus motif, indicating a macrophage specific transcriptional regulation for the corresponding genes.

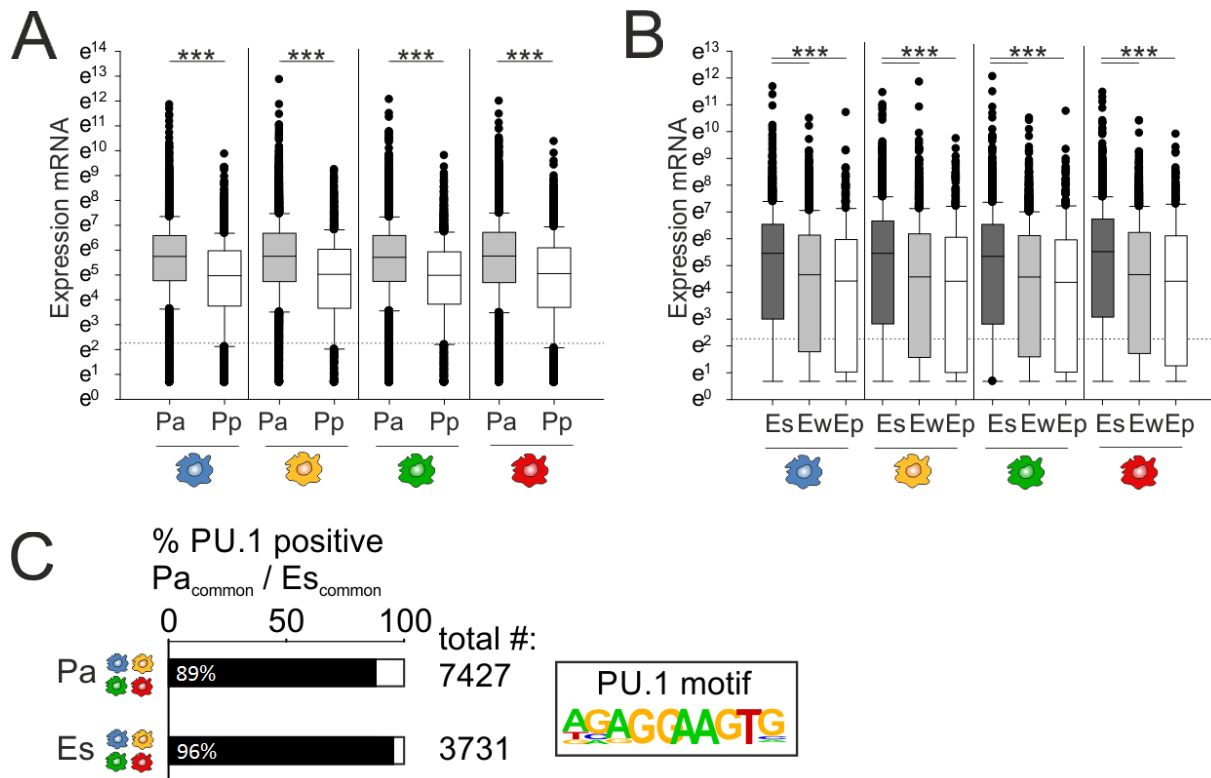


Figure 30. Expression values and PU.1 frequencies of common macrophage promoter and enhancer sites.

RNA-seq expression values depicted as boxplots for genes associated with common **(A)** Pa and Pp or with common **(B)** Es and Ep sites in all four macrophage activation states. *** $p < 0.001$ (wilcoxon rank sum test). **(C)** Percentages of PU.1 motif positive accessible promoter (Pa) or strong enhancer (Es) sites common in all four macrophage activation states. Absolute gene numbers and PU.1 consensus motif were plotted to the right (M^b /blue, $M^{IFN\gamma}$ /yellow, M^{IL4} /green, M^{TPP} /red)

Furthermore, a GO-term enrichment analysis of CHCS genes with common Pa sites and common Es marks was combined into one network with promoter data visualized in network nodes, while enhancer data was depicted in corresponding node borders (Figure 31). GO-terms were grouped according to their relationship of biological functions and resulted in 28 sub-clusters. The network revealed a significant correlation of CHCS gene functions to “housekeeping” functions, like “metabolic process”, “cell cycle” or “macromolecule catabolic process”.

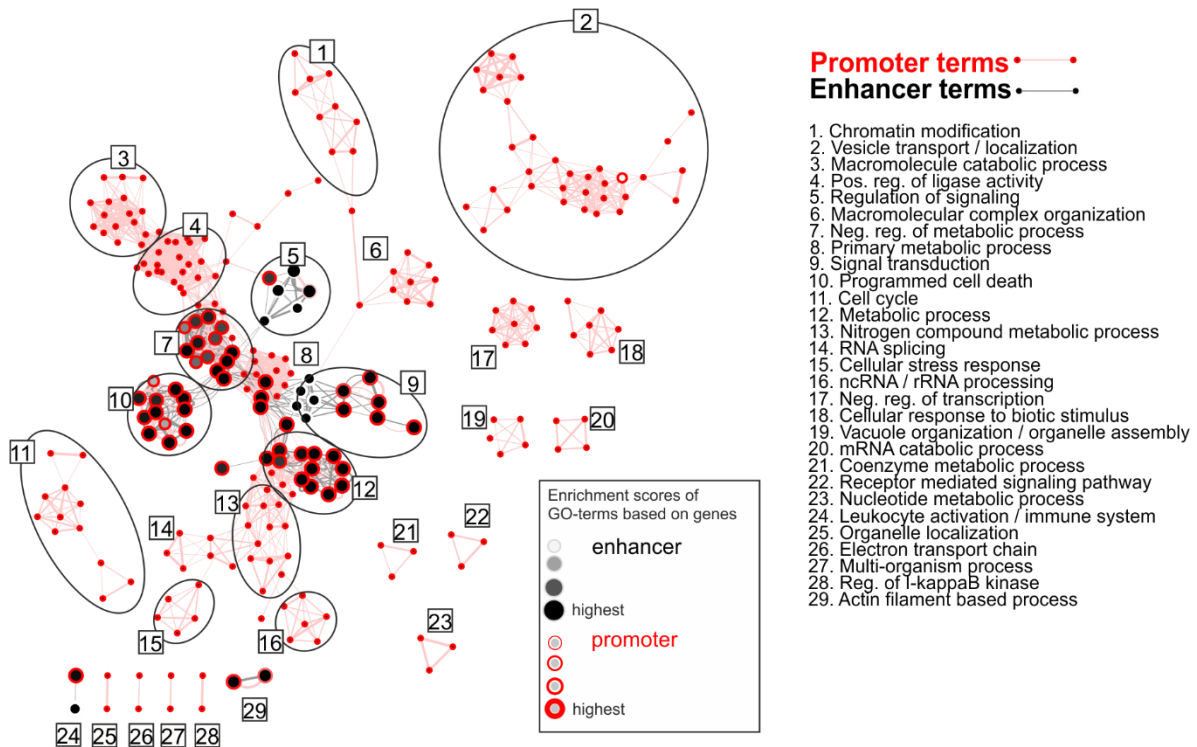


Figure 31. Differential Gene Ontology Enrichment Analysis of CHCS genes marked with common Pa or common Es.

GO-term associated functions of common macrophage genes marked with an accessible promoter (red border color) or strong enhancer (black nodes) are depicted as networks. Specific terms for network modules with strong correlations to each other are marked with numbers and listed to the right.

For a detailed analysis of the macrophage specificity for the defined CHCS gene signature, microarray data sets of other immune cells were utilized. CHCS expression values for all cell types were plotted as boxplots and compared to expression values in M^b . Statistical significant differences were found not only for M^b genes with accessible promoters (Figure 32A) but also for M^b genes in the proximity of Es compared to non-myeloid cell types (Figure 32B). Interestingly, M^b showed no statistical significant difference in CHCS expression values to monocytes or the other three activated macrophage groups ($M^{IFN\gamma}$, M^{IL4} , M^{TPP}), probably due to their strong relationship to each other as macrophage precursors or different M^b activation states.

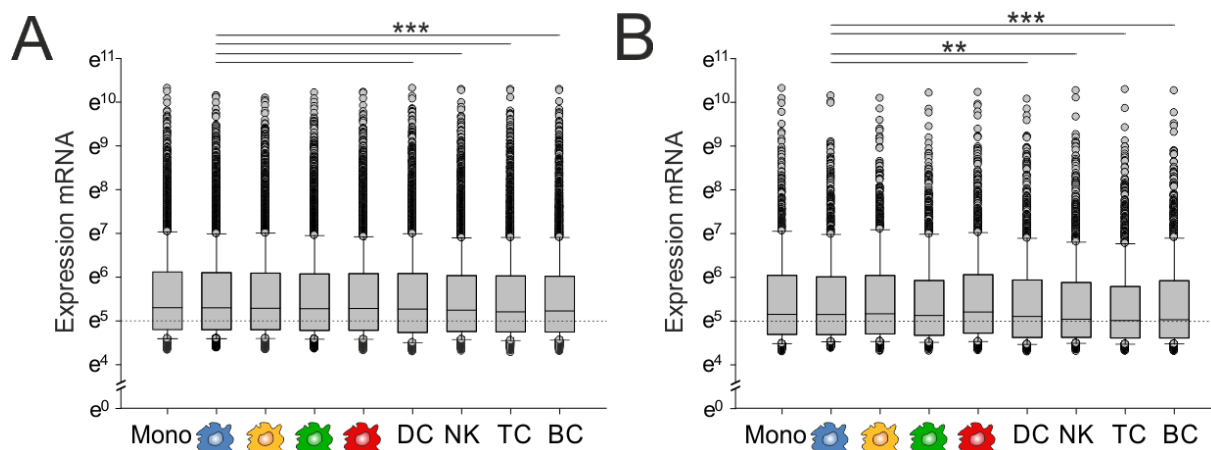


Figure 32. Expression values of common macrophage genes with similar epigenetic status at regulatory sites compared to other cell types of the immune system.

Boxplots of RNA-seq expression values for common macrophage genes associated with **(A)** accessible promoters and **(B)** strong enhancers in various primary cell types from peripheral blood. ** $p < 0.01$, *** $p < 0.001$ (wilcoxon rank sum test). (M^b/blue, M^IF^{Ny}/yellow, M^IL⁴/green, M^TPP/red, Mono=monocytes, DC=immature dendritic cells, NK=natural killer cells, TC=T cells, BC=B cells)

The establishment of a general macrophage phenotype depends on the activity of DNA binding proteins like TFs. A motif binding prediction analysis revealed the enrichment of DNA binding proteins located at accessible promoters (Figure 33A) or strong enhancer sites (Figure 33B) for CHCS genes. Top 10 motifs for proteins expressed in all macrophage activation states included the macrophage pioneer factor PU.1 overrepresented not only at common Pa but also at common Es sites (71,72). These results corroborated the performed PU.1 enrichment analysis described above (Figure 30C). Beyond PU.1, other motifs for TFs like ELK4, GABPA and ELK1 were significantly enriched at common Pa sites. These TFs were described to cooperate with the serum response factor (SRF) to regulate cytoskeletal gene expression in macrophages (255,274). Additionally, ETS family members like FLI1, ELF1 and ETS1 were also enriched and can interact with PU.1 to regulate macrophage development (283,284). At sites with marks for common Es, TFs like FRA1, FOSL2 (FRA2) and AP-1 were identified. Interestingly, these TFs play important roles in putative negative feedback loops suppressing exaggerated inflammatory responses, demonstrated previously for example for the TF BCL6 (285-287).

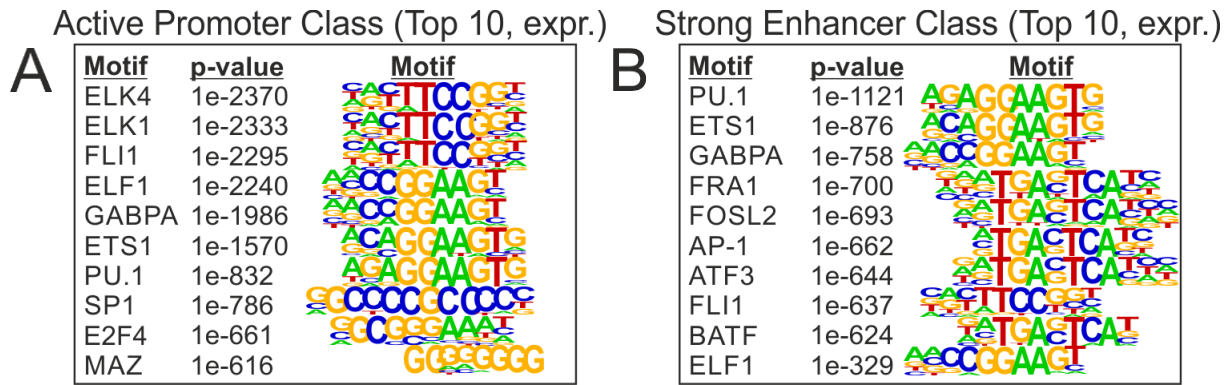


Figure 33. Motif enrichment analysis of common Pa and Es sites found in four activated macrophages.

Motif enrichment analysis for **(A)** accessible promoters or **(B)** strong enhancer sites found in all four macrophage activation states. Corresponding p -values (hypergeometric test) and positional weight matrices (PWMs) for top 10 enriched motifs were additionally plotted.

With the hypothesis that commonly expressed CHCS genes are best characterized by accessible promoters, previously defined major hub genes of the multi-dimensional model of human macrophage activation (21) should also be characterized by permissive histone modifications at their promoters. Therefore, a network representation of previously defined macrophage core signature genes (869 top 10 % highly connected major hub genes) was overlaid with CHCS information of Pa and Es/Ew enrichments (Figure 34). More than 94 % of major hub genes were tagged with a common accessible promoter, while only 27 % of gene names overlapped with genes adjacent to common strong or weak enhancers. The strong overlap for common Pa (CHCS data) with central macrophage genes identified by independently performed *in silico* predictions further demonstrates the importance of epigenetic data for the identification of central genes of a common macrophage program. Lower overlaps for CHCS enhancer data could indicate a minor role for *cis*-regulatory elements in common macrophage programs described before. Intriguingly, the previously built common macrophage network was based on 28 different macrophage stimulatory conditions, but epigenetic data of only four macrophage activation states already matched to a major extent to common macrophage genes.

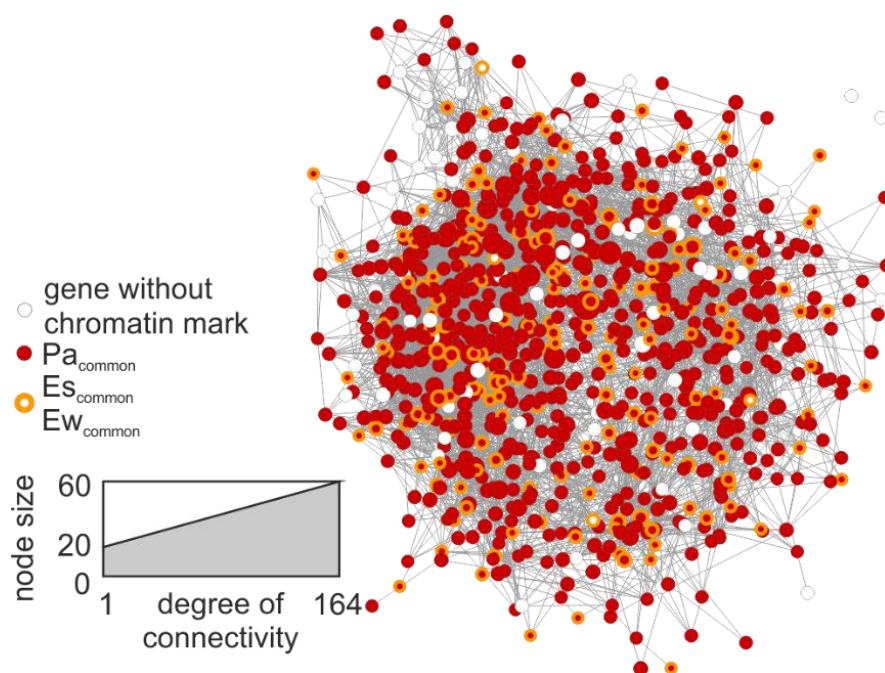


Figure 34. Network visualization for the epigenetic status of regulatory sites in common macrophage core signature genes.

The epigenetic status of promoter and enhancer sites is depicted as network visualization for 869 important regulatory macrophage genes: Pa (red node colour)/ Es/Ew (orange border colour)/ no mark (white border or node colour). The node

size specifies the degree of connectivity between depicted common macrophage genes.

In conclusion, accessible and poised promoters as well as enhancers common for all investigated macrophage activation states were identified and especially Pa correlated well with previously described macrophage core signature genes. Common promoter and enhancer sites were enriched for the macrophage pioneer TF PU.1 and other TFs related to important macrophage functions. CHCS genes were expressed in all differentially activated macrophages and participated in gene ontology terms for general housekeeping functions but more importantly for specific biological processes associated to macrophages. The collected data supports the hypothesis that the epigenetic regulation mediated by the common histone core signature is fundamentally linked to the macrophage phenotype.

5.2.5. Influence of super enhancers onto macrophage activation

Recently, special regulatory genomic regions called “super enhancers” (SE) were described in immune cells like murine macrophages or embryonic stem cells (169,170). Their main characteristics are defined by very strong enrichments for activating H3K27Ac HM signals and remarkably high numbers of TF binding sites. Additionally, SE regions consist of multiple numbers of classically defined enhancer sites with distances less than 12.5 kb to each other (Figure 35A). They were proposed to play important roles in disease initiation and

progression and regulate cell type-specific gene transcription and thus establishing cell identity.

In activated human macrophages between 417 and 806 SE regions could be identified (Figure 35B). As exemplified for M^{TTP} cells, 417 SEs were discriminated from classical enhancers by their disproportionately high H3K27Ac signals. Genomic M^{TTP} sites at SE regions were first sorted according to their total region size, starting with the biggest SE regions with 134.853 bp length and then visualized as heatmaps depicting their normalized H3K27Ac tag counts (Figure 36C). Broad H3K27Ac signals distributed over all SE areas further validated their identity in comparison to surrounding regions with significantly lower H3K27Ac levels. SEs in other macrophages (M^b , $M^{IFN\gamma}$, M^{IL4}) were identified and validated accordingly.

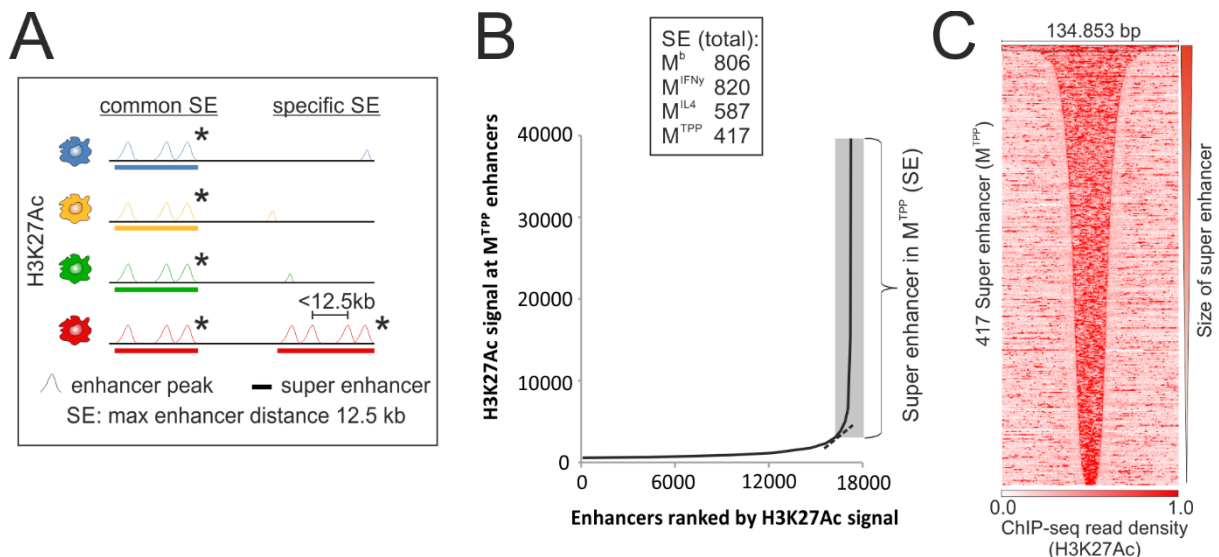


Figure 35. Identification of super enhancers in activated macrophages.

(A) Schematic overview of common and stimulus specific super enhancers (SE). **(B)** Identification of SE sites (grey box) at enhancers with extraordinary high H3K27Ac HM signals, exemplified for SEs found in M^{TTP} . Total SE counts for four activated macrophage activation states are listed in upper box. **(C)** Heatmap of normalized H3K27Ac tag densities at genomic SE sites found in M^{TTP} sorted according to their total region size in bp (M^b /blue, $M^{IFN\gamma}$ /yellow, M^{IL4} /green, M^{TTP} /red)

In differentially activated macrophages, common and specific SEs could be located. 200 common SE regions were identified in all four macrophage activation states (Figure 36A). Histograms of normalized ChIP-seq raw data were visualized for one example of a common SE with high gene expression values in all four investigated macrophage activation states

(Figure 36B). This example of a gene locus defined by high H3K27Ac signals showed additionally strong H3K4me1 HM signals defining it as enhancer site. The SE adjacent gene CCR1, which encodes a chemokine receptor is critical for macrophage immune functions and was expressed in all four macrophages (288). Unique SE sites present in only one macrophage activation state demonstrated a high variability in their numbers (Figure 36A). Highest numbers of activation specific SE were especially identified in M^b and $M^{IFN\gamma}$. A correlation to mRNA expression of genes with specific SE further demonstrated the capability of SE to promote transcription of adjacent genes beyond the expression levels of genes with Es or Ew in their proximity. For example, M^{TPP} genes with SE markings in their vicinity displayed significantly higher expression values in comparison to genes marked with Es or Ew, (Figure 36C). These results state a strong impact of SE on overall macrophage gene expression.

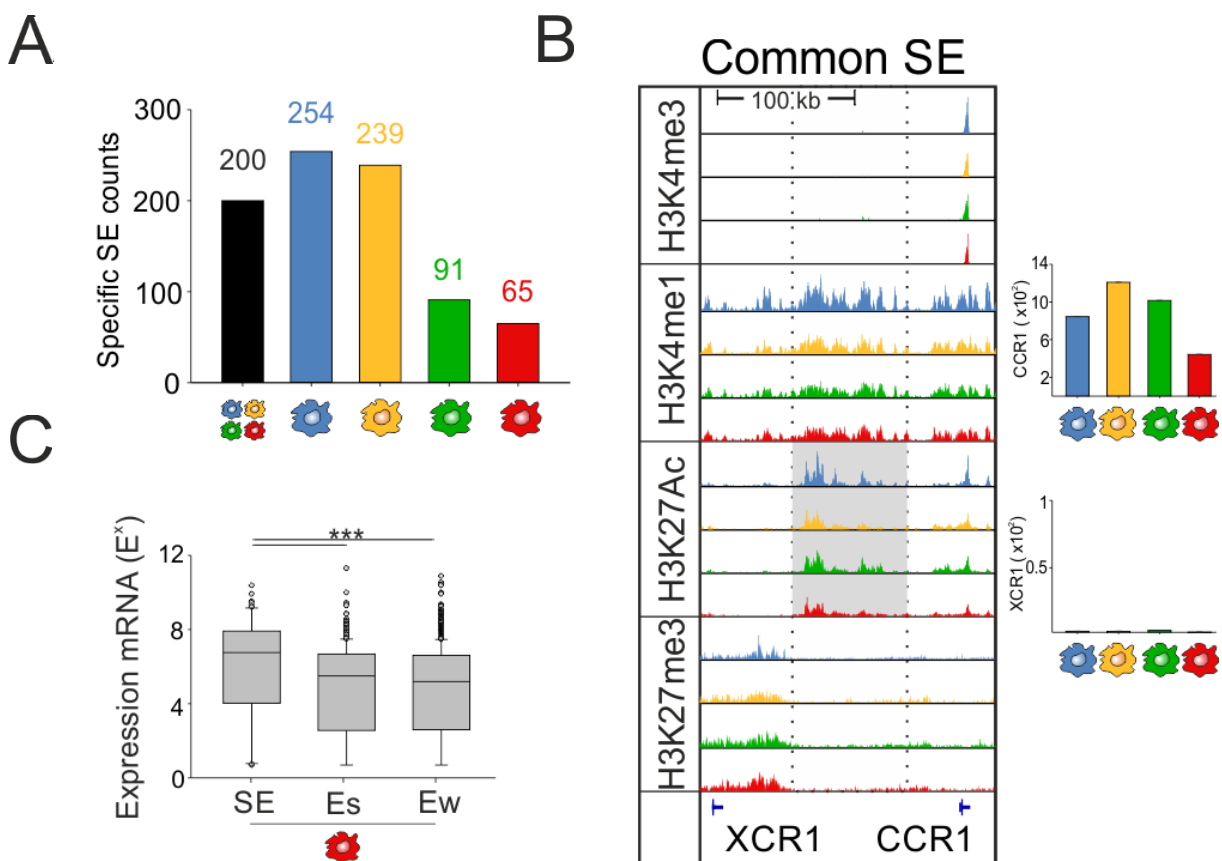


Figure 36. Super enhancer characterization.

(A) Bar chart for numbers of common and activation specific SE found in differentially activated macrophages. **(B)** Example for a genomic site identified as common macrophage SE visualized as histograms with normalized HM signals for four different HM data sets (H3K4me3, H3K4me1, H3K27Ac, H3K27me3). Expression values for the XCR1 and the super enhancer associated CCR1 gene in human macrophages are plotted to the right. **(C)** Boxplots for expression values of genes

marked by the different enhancer types (SE, Es, Ew) in M^{TTP} . *** $p < 0.001$ (wilcoxon rank sum test). (M^b /blue, $M^{IFN\gamma}$ /yellow, M^{IL4} /green, M^{TTP} /red)

Performing a GO analysis on the group of SE marked genes in M^{TTP} , GO-terms for macrophage relevant functions like “response to cytokine stimulus”, “wound healing” and “phagocytosis” were enriched (Figure 37A). Focusing on the smaller groups of genomic sites marked with activation state specific SEs occurred at least partially near genes which contribute to activation specific macrophage functions (Figure 37B). For example, IFI30 is constitutively expressed in antigen-presenting cells mediating lysosomal functions (289). NMI is induced by IFN γ stimulation and is responsible for STAT-mediated transcription, LIPA mediates lipolysis and the alternative activation of macrophages (290) and CXCL1 promotes inflammatory processes (291). Yet GO-term analysis on activation specific SE could be limited in its interpretations due to the low number of activation specific SE (Figure 36A).

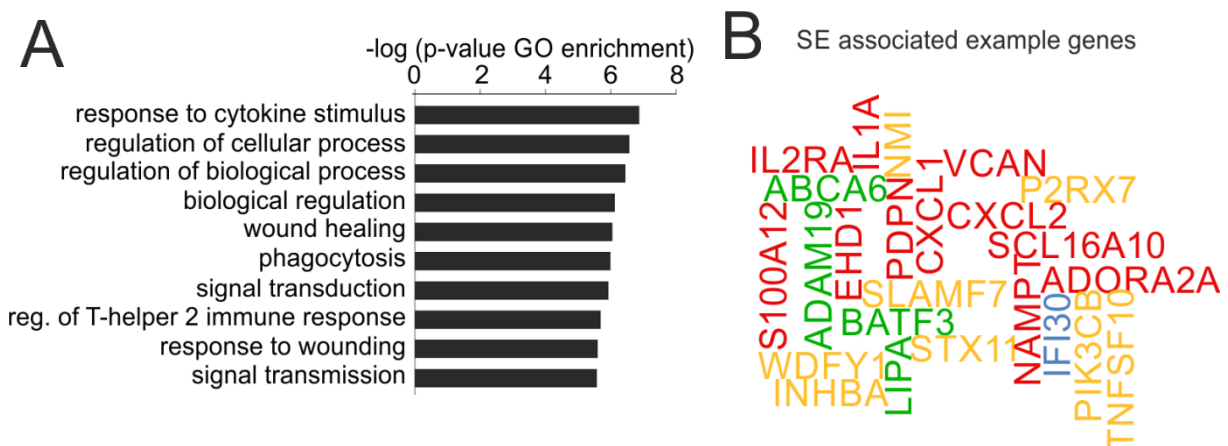


Figure 37. Correlation of gene ontology terms to genes with SE marks.

(A) Top 10 enriched GO-term functions for genes marked with SEs in M^{TTP} . **(B)** Word cloud visualization for exemplified macrophage related genes associated to SEs. (M^b /blue, $M^{IFN\gamma}$ /yellow, M^{IL4} /green, M^{TTP} /red)

All shown examples of expressed activation specific macrophage genes possessed an activation specific SE (Figure 38). In some cases, specific SE sites over spanned not only a single gen locus (LIPA) but were also located over several genes (IFI30) and displayed a broad enrichment for H3K27Ac signals. Moreover, a simultaneous enrichment for H3K4me1 signals was observed thus confirming their enhancer character.

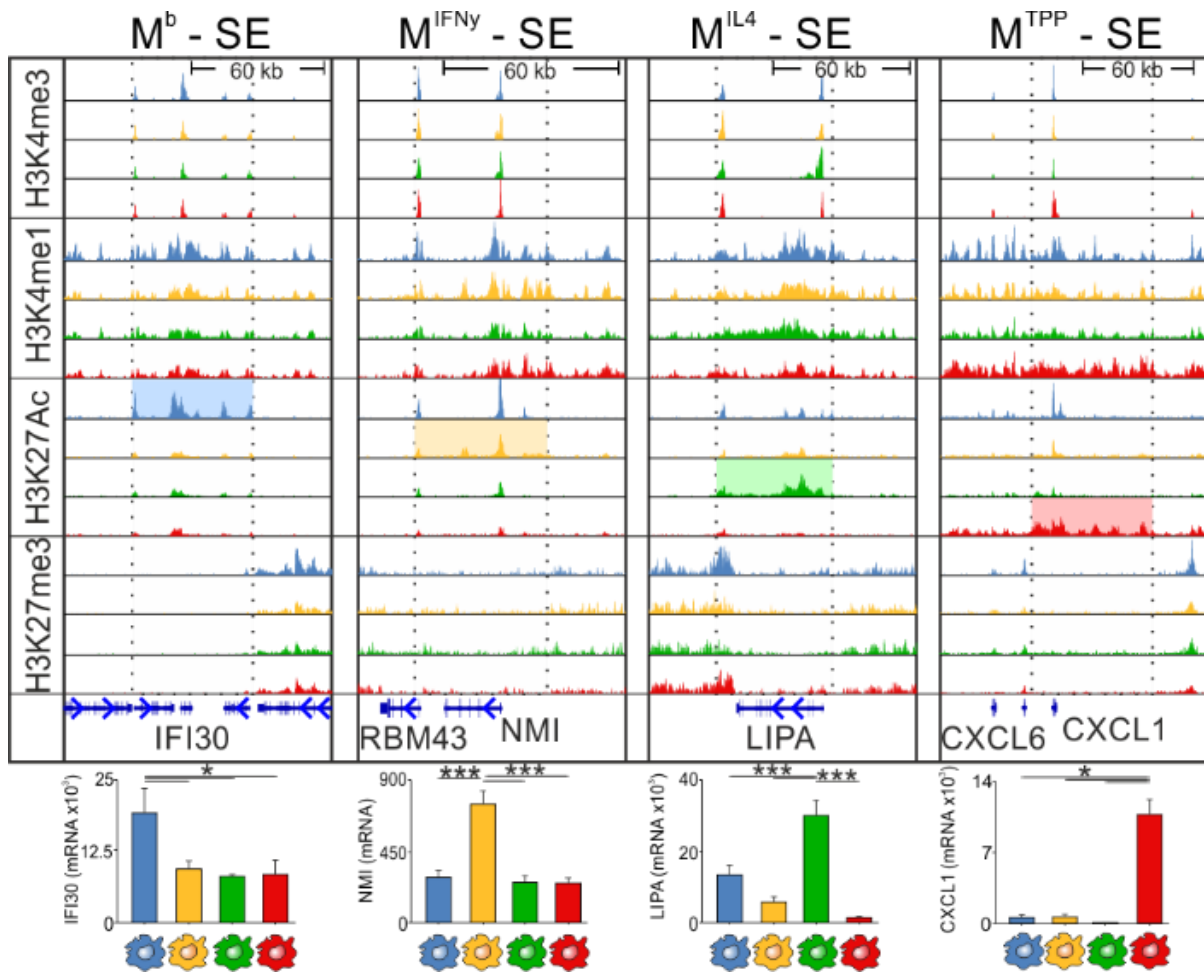


Figure 38. Visualization of genomic loci for macrophage activation state specific SE loci.

Normalized HM signals are visualized for four different data sets (H3K4me3, H3K4me1, H3K27Ac, H3K27me3) at macrophage activation specific SE sites (IFI30/ M^b , NMI/ $M^{IFN\gamma}$, LIPA/ M^{IL4} , CXCL1/ M^{TPP}). Corresponding H3K27Ac signals are marked for each macrophage population with coloured boxes. Expression values for genes at macrophage specific SE sites are depicted at the bottom. * $p < 0.05$, *** $p < 0.001$ (wilcoxon rank sum test). (M^b /blue, $M^{IFN\gamma}$ /yellow, M^{IL4} /green, M^{TPP} /red)

In summary, a special type of regulatory region called “super enhancer” was identified in activated human macrophages. Genes in proximity to common and specific SE sites displayed strongly increased expression values. Furthermore, a partial overlap with known macrophage functions in M^{TPP} and markings at macrophage associated genes in all four activation states were detected. These results suggested the involvement of SE in the input-signal specific induction of gene expression in inflammatory macrophages.

5.2.6. Epigenetic regulation of central macrophage transcriptional regulators in human

For the establishment and specialization of macrophage specific transcriptional programs, TRs are essential key components. Different classes of TRs mediate various functions like chromatin remodelers (e.g. histone methyl-transferases), pioneer factors (e.g. PU.1), transcriptional inhibitors mediating protein complex formation (e.g. NCOR2) or TFs. The expression of these regulatory proteins is influenced by epigenetic modifications, thus the analysis of HM signals at central TR loci could shed light on the regulation of TR expression during macrophage activation.

A previously published multi-dimensional model of macrophage activation allowed the definition of a distinct set of TR genes (21). In brief, GM-CSF derived macrophages (M^{ϕ}) were activated by 28 different stimulatory conditions to mimic inflammatory conditions (Figure 39A). 485 macrophage TRs were identified to be expressed at least in one of the 29 stimulatory conditions. Co-regulation network analysis allowed the creation of visualizations for 297 highly interconnected TR genes, which are to be considered as central key factors for macrophage identity and were termed as “macrophage activation TR network”. In this network each node represents one specific TR with connections to other co-regulated TRs. Surprisingly, almost all promoters (93%) of central TRs participating in macrophage activation were marked by an accessible promoter state already in untreated M^{ϕ} (Figure 39B, C). In addition, the vicinity of central TRs was occupied for more than 69% by strong or weak enhancers (Figure 39D). Concomitantly, TRs with accessible promoters and/or strong/weak enhancer markings showed expression levels above background level, while only a small subset of these TRs showed relatively low expression levels (Figure 39E, F). These results could support constitutively permissive histone modifications at TR loci independently of their transcriptional activity in M^{ϕ} .

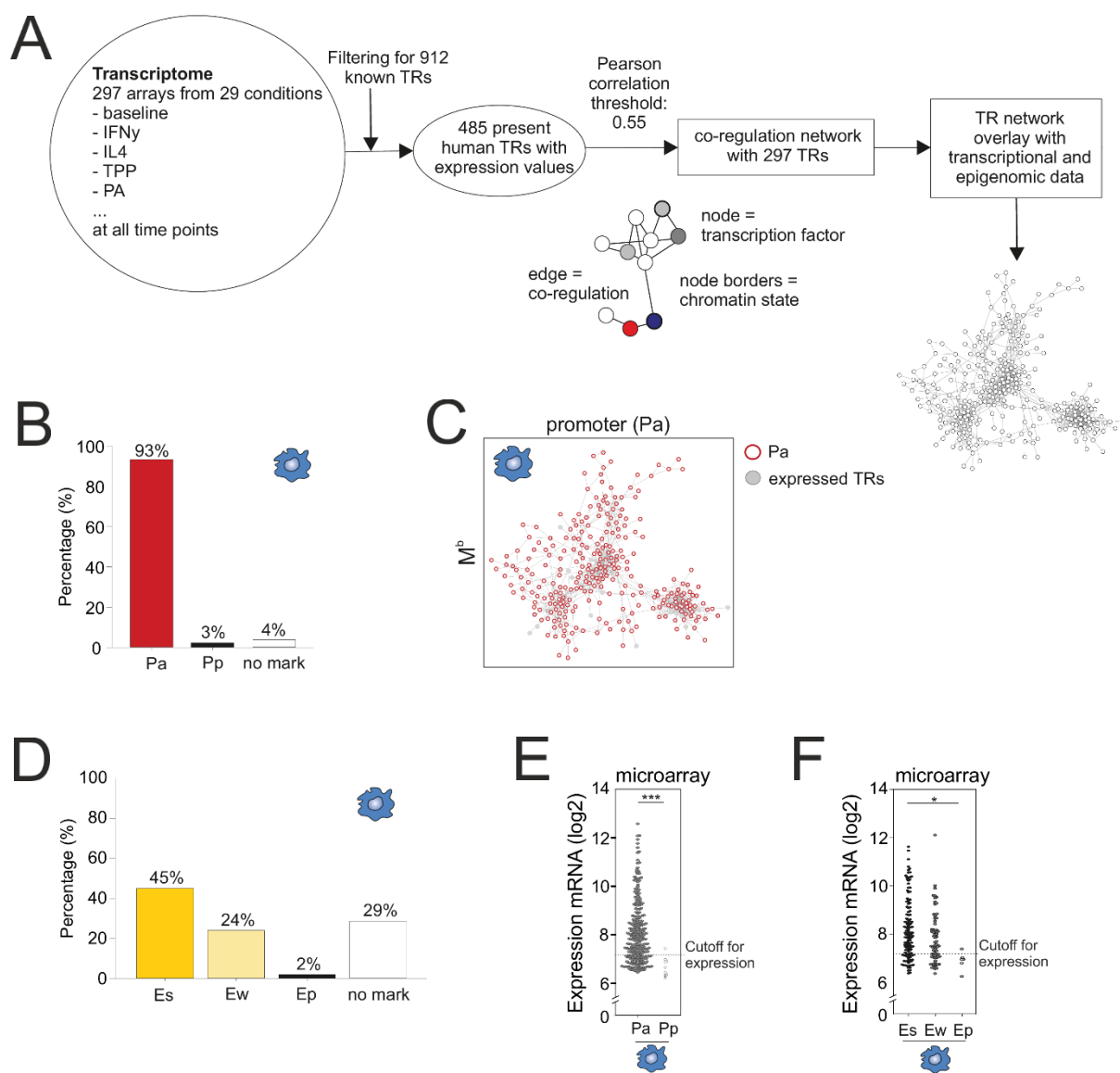


Figure 39. Epigenetic landscape of transcriptional regulators in baseline macrophages.

(A) Schema for the computational generation of co-regulation networks including 297 important activated macrophage TRs (transcriptional regulators) (21). Distribution of **(B)** accessible (Pa) and poised (Pp) promoter marks or **(C)** strong (Es), weak (Ew) or poised (Ep) enhancers in baseline macrophages for TRs being part of the generated macrophage activation TR network. **(D)** TR network visualization for baseline macrophages (M^b) with overlaid epigenetic information for accessible promoter states (Pa). **(E, F)** Microarray expression data for 287 TRs depicted as scatter plots in baseline macrophages (M^b) and sorted for (E) accessible (Pa) and poised (Pp) promoter states or (F) for strong (Es), weak (Ew) or poised (Ep) enhancer states. Mann-Whitney rank sum test (E) or Kruskal-Wallis One Way ANOVA on ranks with pairwise multiple comparisons according to Dunn's Method (F) were used for statistical calculations (* $p < 0.05$, *** $p < 0.001$, cutoff for expression: 7.2).

To answer the next question, if the activation-specific regulation of TR expression is influenced by specific changes of histone modifications at TR loci, three differentially activated macrophages were used. As observed for previous M^b results, a large majority of TR promoters for all activation conditions showed an accessible state (>93%, Figure 40A) with simultaneous strong/weak enhancer markings for a high fraction of TR loci (>69%, Figure 40B). The expression levels for these macrophage activation associated TR genes with accessible promoters or strong enhancers were once more significantly higher than for TR genes with poised promoters or weak/poised enhancer markings (Figure 40C, D). Permissive histone modifications seem to dominate not only the epigenetic TR landscape of M^b but also play an important role in activated macrophages under inflammatory conditions independently of their transcriptional activity.

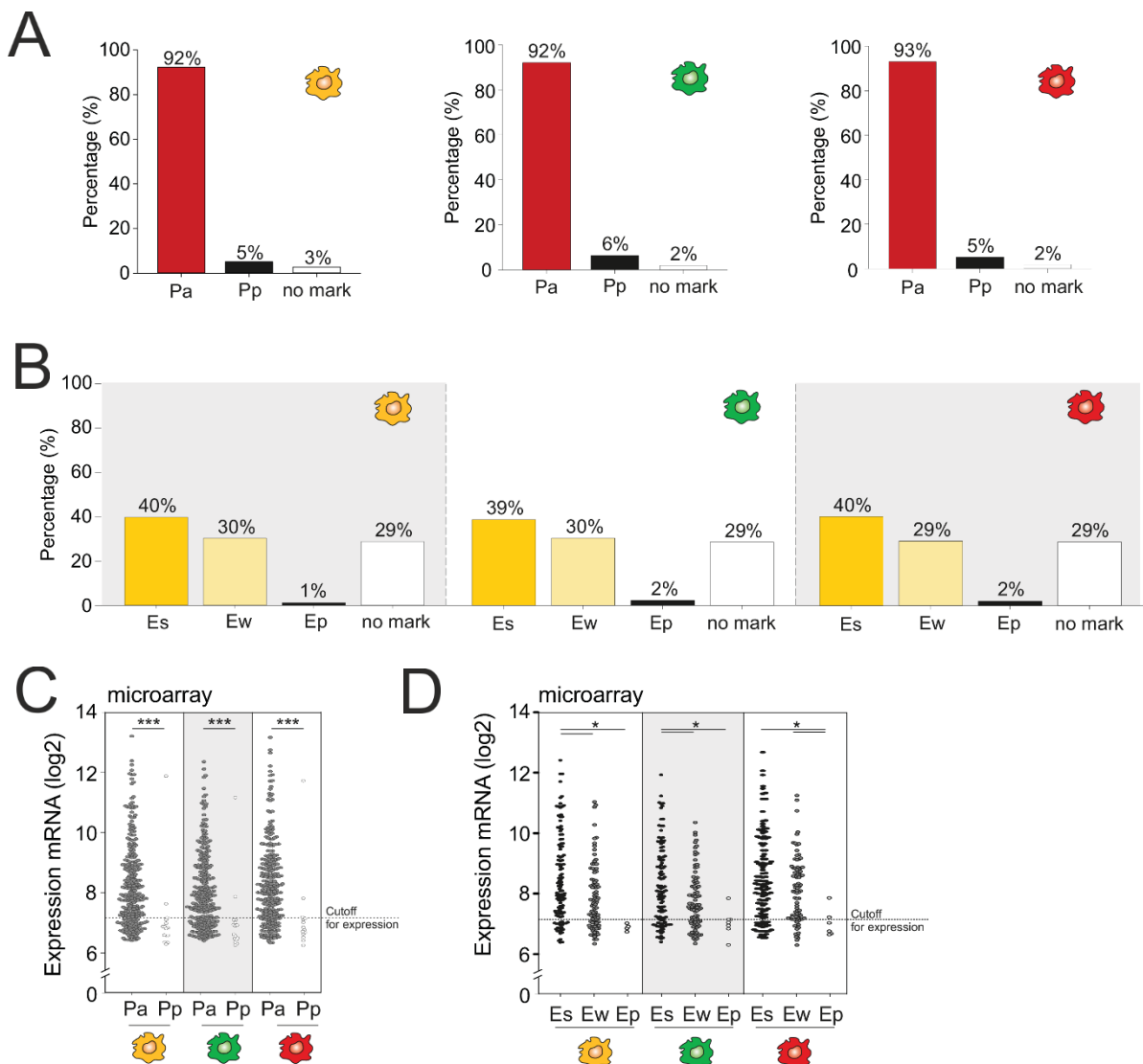
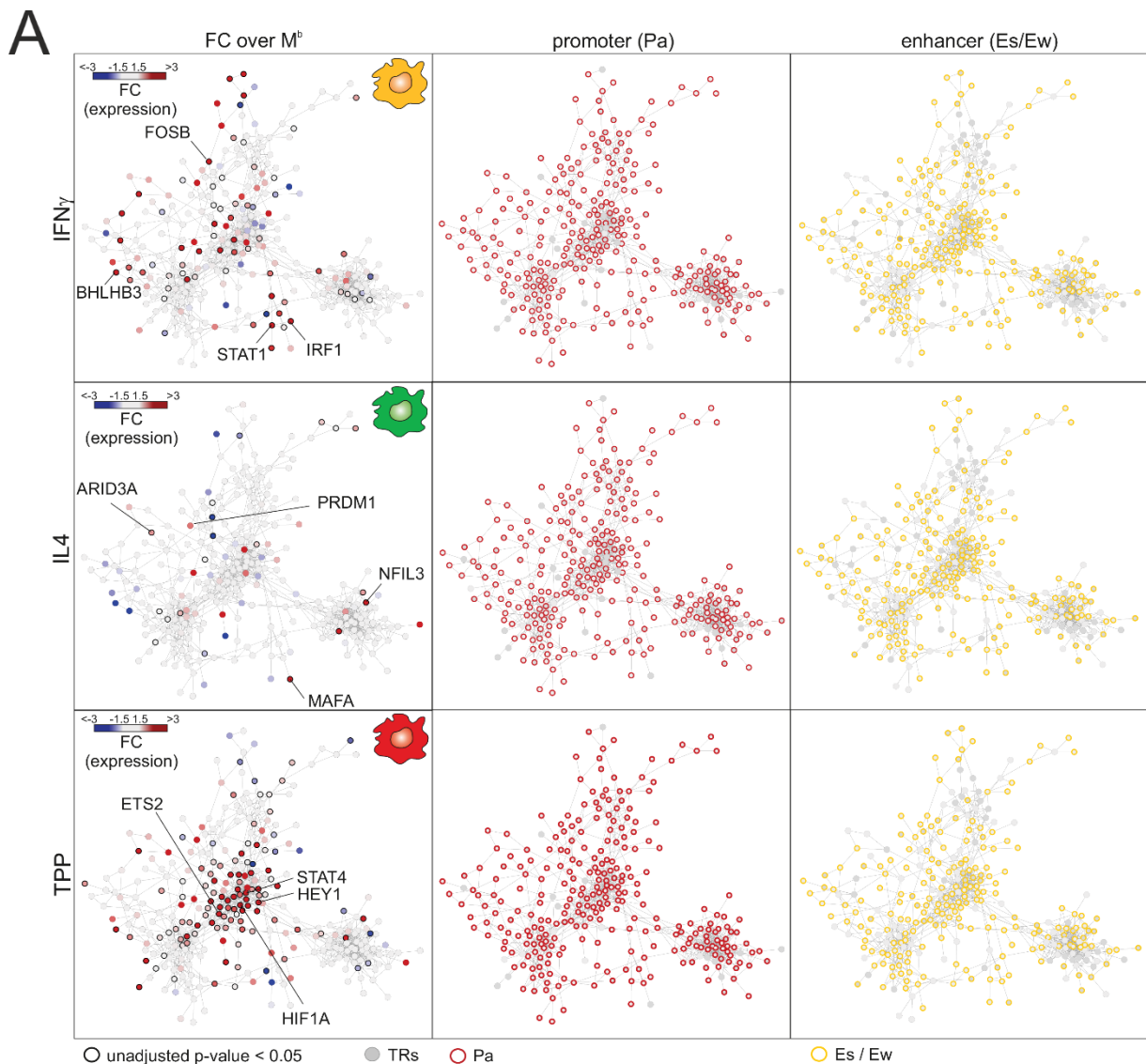


Figure 40. Promoter and enhancer mark distributions for transcriptional regulators in activated macrophages.

Distributions of **(A)** accessible (Pa) and poised (Pp) promoter marks or **(B)** strong (Es), weak (Ew) and poised (Ep) enhancer marks visualized as bar charts for TRs being part of the macrophage activation TR network in IFN γ (yellow), IL4 (green) and TPP (red) activated macrophages. **(C, D)** Microarray expression data for 287 TRs depicted as scatter plots sorted for (C) accessible (Pa) and poised (Pp) promoter states or (D) for strong (Es), weak (Ew) or poised (Ep) enhancer states in IFN γ (yellow), IL4 (green) and TPP (red) activated macrophages. Mann-Whitney rank sum test (C) or Kruskal-Wallis One Way ANOVA on ranks with pairwise multiple comparisons according to Dunn's Method (D) were used for statistical calculations (* $p < 0.05$, *** $p < 0.001$, cutoff for expression: 7.2).

A network visualization of transcriptional regulators allowed not only the direct comparison of transcriptional changes for activated macrophages in comparison to M^b but also the assessment of the epigenetic status for their promoter and enhancer sites (Figure 41A). For all three activated macrophages (M^{IFN γ} , M^{IL4}, M^{TPP}) TRs were transcriptionally regulated in a distinct and specific manner (Figure 41A left panel). These results demonstrate an enormous transcriptional regulation of TRs in an input-signal specific fashion in human macrophages. Compared to the noticeable changes for TRs on transcriptional level, the patterns for promoter and enhancer marks within the network were quite uniform (Figure 41A, middle and right panel) with a majority of promoters marked as accessible and enhancers with strong or weak HM marks. Interestingly, TRs not expressed in any of the 29 stimulatory conditions showed only a minority of accessible promoters (12-15%) or strong enhancer markings (12-17%) (data not shown). To confirm the hypothesis, that the permissive histone landscape is a unique feature for macrophage-activation associated TRs in monocyte-derived macrophages, a hypergeometric statistical test was performed for different gene sets to analyze the significance of accessible promoter occurrences (Figure 41). Both gene groups either for the central nervous system (CNS) or randomly chosen genes showed fewer accessible promoters and thus lower statistical significances compared to the TRs in activated human macrophages. To further identify master regulator genes in each stimulatory condition, TFs with significant enriched expression in M^{IFN γ} , M^{IL4} or M^{TPP} were determined and were predicted to bind to gene loci with an accessible promoter and strong or weak enhancer markings within the network. Important TF genes like STAT1 for M^{IFN γ} , IRF4 for M^{IL4} or STAT4 for M^{TPP} were found and previously connected to corresponding macrophage functions (193,292,293) (Figure 41C).



B

GO term related to:	comPa (n=7283)	Hypergeometric probability
CNS (GO0007417; n=405)	107	3.88e-02
TRs (n=925)	356	8.85e-14
Random genes (permuted; n=925)	-	3.35e-02
Total number of possible genes: 26.354		

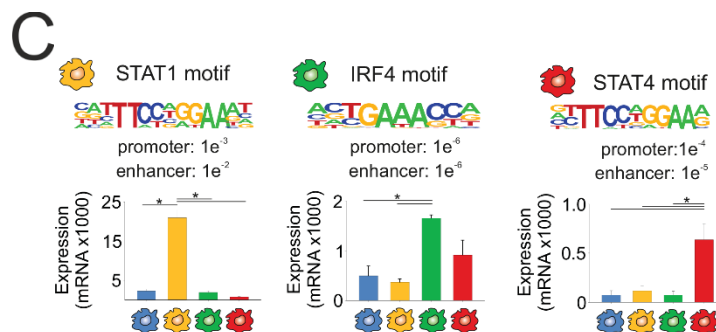


Figure 41. Epigenetic landscape of transcriptional regulators in activated macrophages.

(A) Network visualization of transcriptional regulators in activated macrophages (network generation described in Figure 39A). Left panel: RNA-seq expression values were overlaid as fold-changes (FC) compared to the expression values in M^b (FC with an unadjusted p-value <0.05 are marked by wider black border colour). Mid panel: Overlay with epigenetic data on accessible (Pa) promoters (red border colour). Right panel: Overlay with epigenetic data on strong (Es) or weak (Ew) enhancer marks (orange border colour). **(B)** Hypergeometric probability test for the enrichment of accessible promoters

(Pa) in GO-term gene sets related to the CNS (central nervous system), activated macrophages TR genes or a set of random genes. **(C)** Binomial p -values for motif enrichment calculations at promoter and enhancer positions, positional weight matrices and corresponding expression values were depicted for predicted master regulators in M^b (blue), M^{IFN γ} (yellow), M^{IL4} (green) and M^{T PP} (red) in all macrophage conditions. One Way ANOVA tests with pairwise multiple comparison according to Tukey were used for statistical analysis of expression values (* $p < 0.05$, $n=3$). RNA-seq and ChIP-seq data was derived from three independent experiments.

Overall, the established networks of active TRs during human macrophage activation showed in general an open chromatin state at promoter loci with concomitant strong or weak enhancer markings, while significant transcriptional differences between used stimulatory macrophage conditions were visible. These results could hint to a model where the quick adaption capabilities of macrophages to environmental signals are tightly connected to these epigenetic open TR loci. Transcriptional control mechanisms like the binding of transcription factors, co-repressors or –activators or non-coding RNAs which restrain TR mRNA transcription could regulate the predominantly accessible promoters and subsequently adapt the macrophage transcriptome.

5.2.7. Epigenetic and transcriptional regulation in human tissue-defined TR networks

Since human macrophage activation TR networks are characterized by an overall open chromatin state the question arises, if this epigenetic pattern is an exclusive feature in these cells. To answer this question similar histone modification networks were created for five human tissues (intestine, lung, ovary, heart, and muscle) (294) (Figure 43A). These datasets showed comparable RNA-seq statistics (data not shown), but up to 3-fold differences in their read numbers (Figure 42A) and called peaks (Figure 42B), probably due to their differences in function and location in the human body. Nevertheless, the histone distribution across the genome in correlation to important genomic functions demonstrated almost equal results, which was not only visible for the intestine and lung datasets (Figure 42C) but for all five human tissue datasets (data not shown).

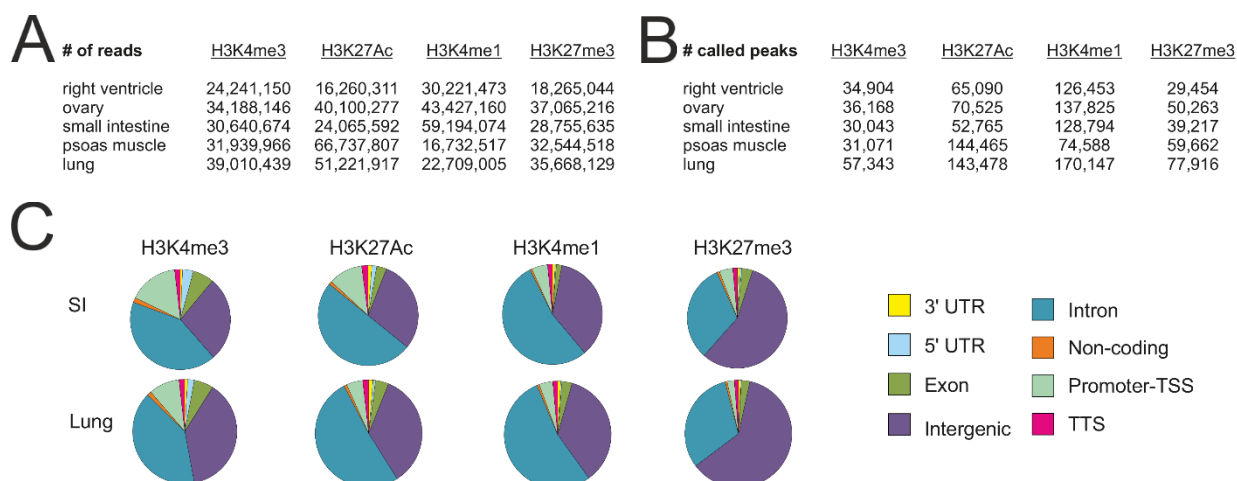


Figure 42. Main features of ChIP-seq datasets for five different human tissues

(A, B) Main features of consolidated histone ChIP-seq data ($n=1$) from the Road map consortium (294) describing (A) read counts and (B) called peak positions for corresponding histone datasets in five human tissues. **(C)** Histone modification peak distributions depicted by their localization to specific genomic regions for small intestine (SI) and lung tissue (Lung) datasets.

After the creation of a TR network for these 5 human tissues (Figure 43A) a similar approach introduced for the human macrophage activation TR network was performed. In the following results only two example datasets for small intestine and lung tissues were visualized but the other three datasets showed similar results. Transcriptional differences plotted onto the network revealed a specific upregulation of genes in only one subcluster for each of the five tissues (Figure 43B, left panel). Upregulated small intestine genes e.g. CDX2 or CREB3L3 are mainly located in the central subcluster, while important upregulated lung genes like TBX4 or NKX2-1 are located in the lower right network subcluster. TRs being not expressed in the particular tissue were located within the other not-tissue-specific subclusters in the network that marked the remaining tissues (Figure 43B, right panel, green nodes). Interestingly, genes not expressed in a particular tissue lacked the open promoter markings. This pattern strongly supports the hypothesis that transcriptional differences of TRs between tissues are epigenetically regulated and thus following previously published models(167). Indeed, the correlation of epigenetic promoter status and gene expression of TRs clearly demonstrated a dichotomous distribution with expressed TRs being defined by accessible promoters, while not expressed TR genes were characterized by the absence of accessible promoters (Figure 43C). Raw histone modification read counts on a genomic level also support these results (Figure 43D). For example, the CDX2 gene is not only specifically

expressed in small intestine cells but also shows specific open chromatin marks like H3K4me3 and H3K27Ac only in the small intestine dataset.

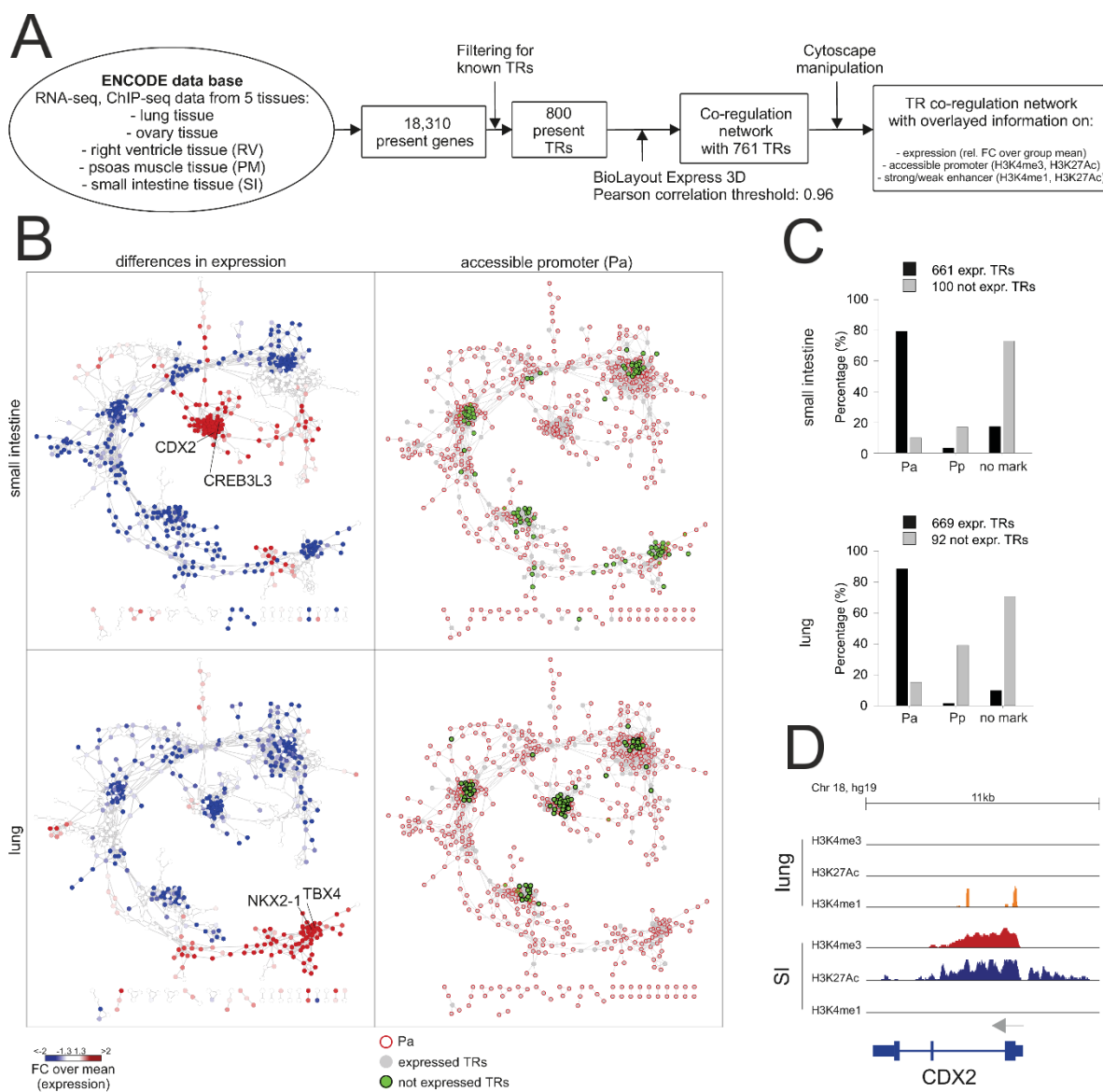


Figure 43. Epigenetic control of transcriptional regulators in human tissues

(A) Bioinformatics workflow for the generation of human tissue-related TR networks. **(B)** Network visualization of human-tissue related transcriptional regulators in small intestine and lung tissue datasets. Left Panel: Overlay with expression values depicted as fold-changes (FC) over the mean expression values of all five tissues included in the analysis. Right Panel: TR networks are overlaid with epigenetic data on accessible promoters (Pa) marked in red. Additionally, expressed TRs are marked in grey (RNA-seq expression values >10), while not expressed TRs are marked in green with black borders (RNA-seq expression values <10). **(C)** Human tissue related TRs were grouped into expressed (black bars) and not expressed (grey bars) genes and their promoter states were visualized as bar charts (accessible promoter (Pa), poised promoter (Pp), no H3K4me3 promoter mark (no

mark)). **(D)** Representative example position at CDX2 gene locus for a TR with tissue specific expression and promoter marks.

In conclusion, network visualizations clearly showed that highly upregulated genes in tissue specific clusters were simultaneously marked with open promoters, while not expressed genes lacked these markings. This finding suggests that tissue-associated TRs are unlike human macrophages coordinately regulated on the epigenetic and transcriptional level.

5.2.8. Transcriptional regulator networks in tissue macrophages

Previously depicted data in activated human macrophages and different tissues resulting in quite differential pictures about epigenetic and transcriptional regulation raised the question, if the described open chromatin state in monocyte-derived human macrophage TR networks is already a feature of macrophages during tissue homeostasis. To answer this question previously published data of seven murine tissue macrophage populations was used to generate TR networks (Figure 44), since there is not sufficient transcriptome and epigenome data available for human tissue macrophages (157).

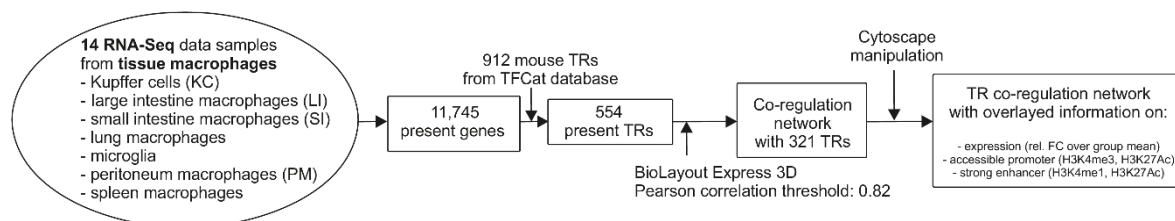


Figure 44. Workflow for the generation of TR networks in murine tissue macrophages.

Bioinformatics workflow for the generation of human tissue-related TR networks.

First of all, 554 expressed TRs were identified in at least one of the 7 populations, while in total 321 expressed TRs were highly interconnected (Figure 45A, left panel). Interestingly, after the plotting of differential gene expression onto the network, each of the 6 subclusters was associated with highly upregulated genes for a particular tissue macrophage population. Previously described specific tissue TFs were located in the described tissue specific subclusters within the network, for example Runx3, which was previously connected to ileal and colonic macrophages was present in the gut subcluster (157). This was also visible for

other TFs like Mef2c (microglia), Gata6 (peritoneal macrophages), Rxra (Kupffer cells), or Spic (spleen macrophages). In general, 2/3rd of all TRs are not specifically expressed in one tissue macrophage population but in any given population suggesting the importance of transcriptional changes in only a small number of TRs for the tissue specific transcriptional program. After adding epigenetic promoter information onto the network 77 to 90% of the specifically expressed TRs in the respective macrophage population were marked with an accessible promoter while only 19 to 41% of TRs being not expressed showed an accessible promoter state (Figure 45A, right panel, Figure 45B). Similar results were found for all 554 TRs expressed in at least one of the seven macrophage populations (data not shown).

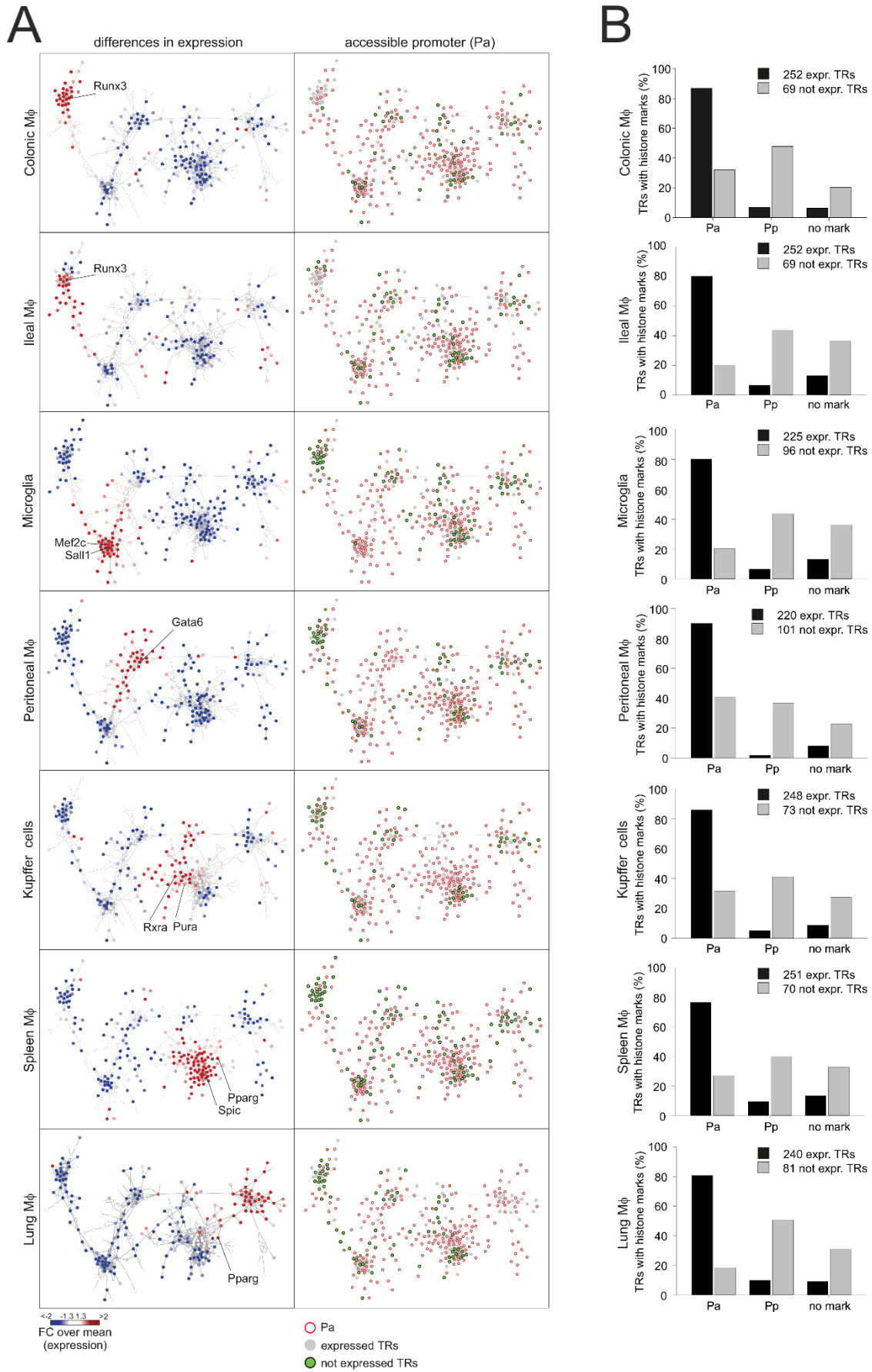


Figure 45. Epigenetic regulation of TRs in murine tissue macrophages.

(A) Network visualizations for transcriptional regulators in seven murine tissue macrophages (network generation described in Figure 44A). Left Panel: Overlay with expression values depicted as fold-changes (FC) over the mean expression of all seven tissue macrophage datasets. Right Panel: Epigenetic data for accessible promoters (Pa) was overlaid in red border colours. Expressed TRs (RNA-seq expression values >10) were visualized in grey colour, while not expressed TR genes were depicted in green with black border colour. **(B)** Distributions of expressed (black bars) and not expressed (grey bars) TRs belonging to the murine tissue macrophage TR network were depicted as bar charts and were also grouped by their epigenetic promoter states (accessible promoter (Pa), poised promoter (Pp), no H3K4me3 promoter mark (no mark)). ChIP-seq and RNA-seq data originated from two independent experiments.

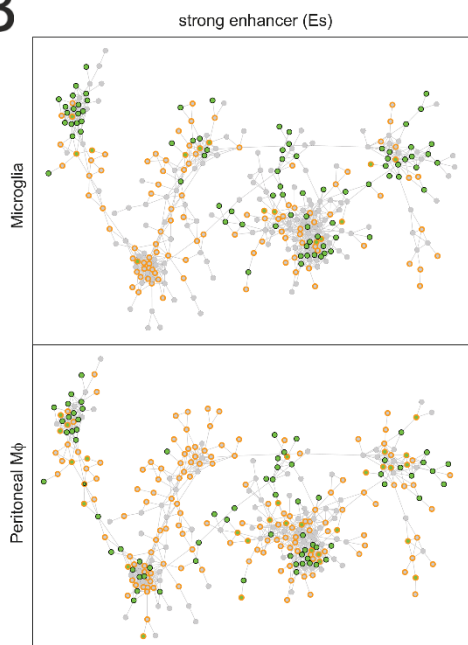
In contrast, TRs not included in the network ($n=358$) showed clearly lower percentages with 5-8% accessible promoter ratios (Figure 46A, upper numbers). In addition to epigenetic promoter state information two examples of strong enhancer datasets for microglia and peritoneal macrophages were also plotted to the network (Figure 46B). The highest percentages of strong enhancer enrichments were detected for expressed TRs within the network, followed by not expressed network TRs, while TRs not being part of the network showed the lowest enrichments for strong enhancers (Figure 46A, lower numbers). The genomic loci for two specific tissue TRs (Gata6 and Sall1) were chosen to further visualize the epigenetic regulation in murine tissue macrophages. Histone modification signals for H3K4me3, H3K27me3 and H3K4me1 were plotted for all tissue macrophages and depict the signal specificity of Gata6 for peritoneal macrophages and Sall1 for microglia (Figure 46C).

A

Histone modification state of 358 TRs being not present in murine macrophages:

Colonic	not expr.	Ileal	not expr.	Microgl.	not expr.	Periton.	not expr.	Kupffer	not expr.	Spleen	not expr.	Lung	not expr.
Pa	7 %	Pa	5 %	Pa	6 %	Pa	8 %	Pa	8 %	Pa	6 %	Pa	5 %
Pp	39 %	Pp	25 %	Pp	50 %	Pp	53 %	Pp	44 %	Pp	43 %	Pp	42 %
no mark	54 %	no mark	70 %	no mark	44 %	no mark	39 %	no mark	48 %	no mark	51 %	no mark	53 %
Es	9 %	Es	0 %	Es	4 %	Es	13 %	Es	9 %	Es	31 %	Es	0 %
Ep	24 %	Ep	0 %	Ep	12 %	Ep	28 %	Ep	31 %	Ep	3 %	Ep	0 %
no mark	67 %	no mark	100 %	no mark	84 %	no mark	59 %	no mark	60 %	no mark	66 %	no mark	100 %

B



C

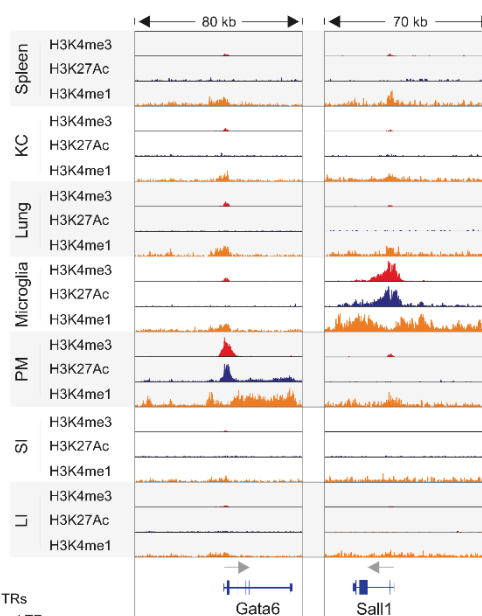


Figure 46. Epigenetic landscape of TR in murine tissue macrophages.

(A) Promoter and enhancer state distributions for 358 TRs not expressed in five murine tissue macrophage datasets (accessible promoter (Pa), poised promoter (Pp), strong enhancer (Es), weak enhancer (Ew), poised enhancer (Ep), no H3K4me3 (promoter) / H3K4me1 (enhancer) mark (no mark)). **(B)** Network visualizations of murine tissue macrophage related transcriptional regulators for microglia and colonic macrophages overlaid with epigenetic data on strong enhancers (orange border colour). Expressed genes are marked in grey colour, while not expressed TR genes are marked in green with black border colour. **(C)** Example positions for tissue macrophage-specific promoter (H3K4me3) and enhancer (H3K4me1) marks at the TF gene loci Gata6 and Sall1. ChIP-seq and RNA-seq data originated from two independent experiments.

In summary, the TR network approach revealed that tissue macrophages are defined by a whole cluster of co-regulated TRs and not only by the upregulation of a single TR. In addition, while activated human macrophage TR networks were characterized by a completely open promoter landscape, tissue macrophages showed a more tissue specific and coherent regulation of epigenetic and transcriptional activity.

5.2.9. Model of epigenetic regulation in human macrophages in an inflammatory model

The integration of herein generated human macrophage data on epigenetic HM level during activation and the correlation to transcriptional level can be summarized into a working model of several layers of epigenetic regulation beyond the activity status of promoter or *cis*-regulatory sites (Figure 47).

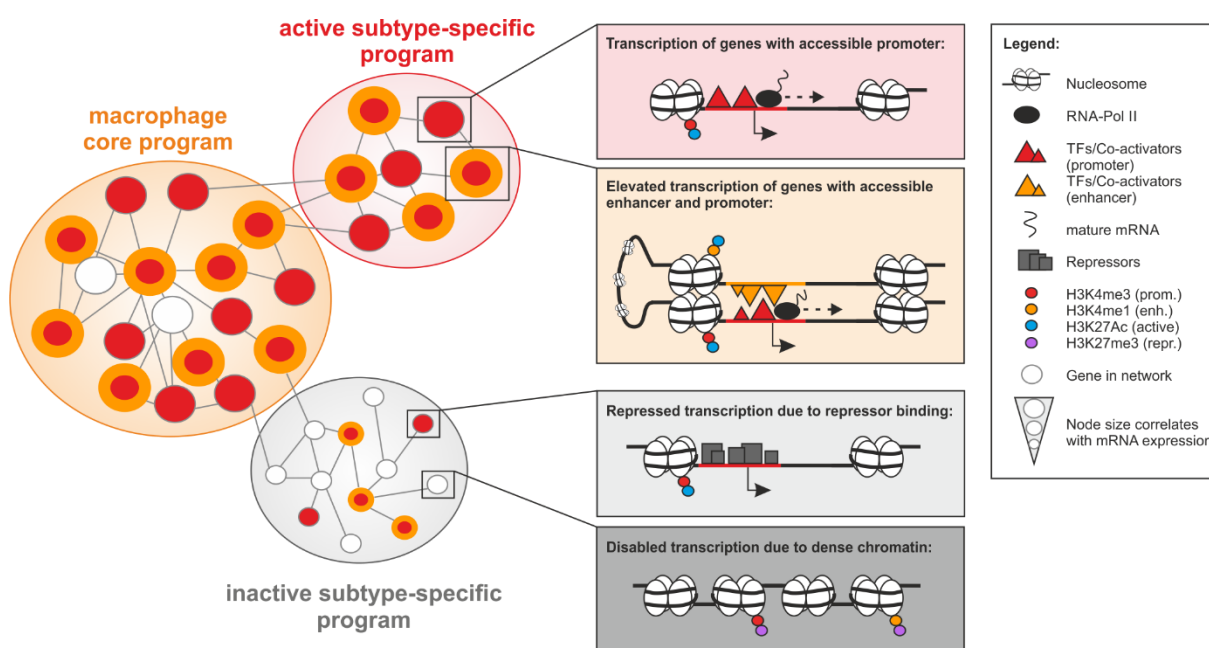


Figure 47. Model for epigenetic regulation in differentially activated macrophages.

Schema of central transcriptional regulator networks in human macrophage activation. Differentially activated human macrophages share a common core program with accessible promoter and enhancer sites marked with H3K4me3 and H3K4me1 and additional H3K27Ac HM activating signals. Activation state specific accessible and inactive sites are able to allow fast transcriptional regulation in case of environmental signals which are the basics for macrophage plasticity. Such regulatory mechanisms are established by several layers of epigenetic control mechanisms and an interplay between chromatin remodellers, repressors and transcriptional activators that allow a transcriptional specialization of activated macrophages.

In summary, this work describes a macrophage core program which shapes the general macrophage transcriptome and could be responsible for the regulation of basic cellular functions, while specific programs could drive macrophage specialization into different activation states with various functions. Interestingly, besides closed sites without functions in

the macrophage cell type, a high abundance of accessible regulatory promoter and enhancer sites could keep many genomic TR sites, which control the macrophage transcriptome in a general open chromatin state. Thus, these highly adaptive cells can quickly adapt to stimulatory signals, induced by changes in the microenvironment, independently of the slower chromatin remodeling process. This transcriptional fine-tuning for macrophage specialization could be regulated on other epigenetic levels beyond chromatin remodeling, possibly amongst others by the binding of transcriptional repressors.

6. Discussion

6.1. False-positive and negative signals in ChIP-seq experiments

ChIP-seq experiments inevitably contain false-positive signals introduced at multiple steps of experimental procedures (204). The herein presented method to reduce false positive ChIP-seq signals named KOIN in the first paragraph of this thesis, revealed false-positive ChIP-seq signals in up to 80 % of peak signals in six independent data sets. However, one data set for the important macrophage pioneer factor PU.1 displayed outstanding quality with minimal amounts of false-positive signals. Intriguingly, even in this high-quality data set the KOIN method increased signal-to-noise ratios and identified more significant peaks in comparison to peak calling without utilizing the KOIN method (Figure 10). Irrespective of the high variability of data set quality with various ratios of false- to true-positive signals, the KOIN method successfully identified and removed false-positive ChIP-seq signals (Figure 6, Figure 8, Figure 9) and recovered false-negative peak signals.

Interestingly, variable ratios of non-specific to true-positive signals between differentially stimulated ATF3 data sets were detected (Figure 6). Percentages of false-positive peaks were almost reduced to 50% after HDL stimulation in comparison to unstimulated macrophages. It still needs to be elucidated, how and to which extent the stimulation of cells can influence the amount of non-specific antibody binding, resulting in false-positive peak signals. Hypothetically, due to the strong impact of stimulation onto the chromatin landscape, demonstrated in mouse macrophages upon LPS stimulation, many new promoter or enhancer regions are in an open chromatin state and thus potentially introducing more ATF3 binding sites leading to more specific signals (73). Additionally, increased ATF3 expression upon HDL stimulation (254) could further present valid antibody targets for specific ChIP-seq binding signals and thus increase specific to unspecific signal ratios.

Particularly, the antibody quality has a strong impact on false-positive binding signals during ChIP-seq experiments. One approach called “ChIP-string” was designed to screen for effective antibodies against chromatin regulator proteins (248). Multiplexed ChIP-seq experiments with low sequencing depth were performed for every antibody and enrichments at approximately 500 representative loci were detected and correlated to IgG controls and already known chromatin states to assess the antibody quality. This approach or other antibody validation procedures, like immunoblotting suggested by the ENCODE consortium

(204) can initially help to evaluate antibody quality, aid the optimization of ChIP-seq experiments but cannot completely abolish unspecific antibody binding. Besides, KOIN even improved ChIP-seq data sets performed with high quality antibodies and is especially important for the identification of “hyper-ChIPable” regions, which is not possible with described antibody validation methods.

Overall, the complexity of the next generation sequencing (NGS) technology requires not only considerate experimental procedures but also highly specific and adaptive data processing methods to account for a data analysis as accurate as possible with the aim to correct biological downstream interpretation. The increasing availability of NGS technology due to greatly improved cost efficiency resulted in a rapid progression of data analysis strategies. These efforts allowed further improvements of NGS data analysis on different stages of the ChIP-seq analysis pipeline, which could be in addition to KOIN beneficial for future ChIP-seq experiments and decreasing false-positive ChIP-seq signals. For example, during the first major step aligning short DNA sequences (ChIP-seq tags) to the reference genome to discover locations of enriched ChIP protein binding, unaligned data is normally excluded as false-negative sequences from downstream peak calling (295). A remapping procedure of not alignable sequences with the short read aligner SHRIMP demonstrated the enhancement of gene regulatory grids and networks for the human TAL1 gene. Other approaches try to utilize “multi-mappable” reads. These sequences are additionally excluded from ChIP-seq analysis and can be introduced for example by excessive PCR amplification of ChIP DNA. They match to multiple locations on the reference genome and could therefore falsify the peak calling procedure (271). Yet, “not-uniquely” alignable reads can contain biological relevant information for proteins binding to highly repetitive sequences at different locations on the reference genome and can further increase sequencing depths (296).

It was demonstrated, that false-positive peak ratios can occur with high variability and are a common problem especially amongst “low-quality” ChIP-seq data sets. A systematic approach to assess and compare ChIP-seq data quality was performed previously (251). Between 20 % and 45 % of analyzed data sets were defined as low or intermediate quality. But, rigorous exclusion of all experiments failing very stringent and rigid criteria could be the wrong approach and could lead to the loss of important biological discoveries. KOIN could greatly improve the interpretation of low-quality ChIP-seq experiments and extract biologically meaningful data, especially when antibodies with lower binding specificities were used. Newly performed ChIP-seq experiments as well as already archived data could benefit from reanalysis with the KOIN method, as long as knockout data sets are available.

6.2. Correct biological interpretation requires KOIN correction

One type of biological interpretation for ChIP-seq binding signals correlates for example signals associated to specific genes to known gene groups involved in biological functions included in a gene ontology (GO) database (264,265,297). This type of analysis allows a general evaluation of putative functions associated to the analyzed TF and its binding patterns. The exclusion of false-positive signals during the KOIN method had a profound impact not only on the enrichments of specific TF binding motifs (Figure 12), often used as ChIP-seq quality criterion but also on the biological interpretation of ChIP-seq data sets, herein exemplified for the SRF protein (Figure 15, Figure 16). After false-positive correction by KOIN, SRF relevant GO-terms were identified, while unspecific SRF functions were lost.

Another type of biological interpretation utilizes the motif binding enrichments and predicts the cooperation of DNA binding proteins by forming putative multi-factor proteins, important for the corresponding analyzed cell type. The concomitant enrichment for the important pioneer TF PU.1 in the SRF data set at SRF binding sites demonstrated such a cooperative binding of TFs in the transcriptional regulation of macrophages (Figure 12) (255). Another example for the cooperation of PU.1 with the TF RUNX1 during myeloid differentiation at genomic sites like the granulocyte macrophage colony-stimulating factor receptor (GMCSFR) were described in murine macrophages (298). Interestingly, without a functional RUNX1 protein, PU.1 interacts with a corepressor complex and its histone deacetylase (HDAC) activity can subsequently decrease GMCSFR expression. It was demonstrated in this work that KOIN further sharpens the results for the identification of highly specific motifs cooperating with corresponding ChIP proteins, mediating important biological functions. Additionally, the low statistical significance of motif enrichments in KO data sets (Figure 13, Figure 14) showed the absence of specifically enriched motifs associated with KO peak sites and rather depicts randomly distributed motifs over the whole human genome, detected by unspecific antibody bindings to irrelevant target proteins.

Of note, in ATF3 data sets other TF motifs were identified (Jun, Jun-AP1) but not specifically the ATF3 motif (Figure 12). This might be caused by variations in TF binding strength based on modifications of the motif sequence, as it was described for the early stages of *Drosophila* development in embryonic cells (299). Another reason could be the hetero-dimerization of ATF3 with the Jun or AP1 protein to stabilize the ATF3 DNA binding followed by the induction of transcription for corresponding target sites (272).

In summary, the described observations like GO-term enrichments as well as TF cooperation predictions or other biological interpretations lead to the conclusion that the KOIN method performed during ChIP-seq peak calling significantly benefits bioinformatics data analysis.

6.3. Knockout data sets - Optimal control for TF ChIP-seq experiments

Already established controls reduce false-positive signals to a certain extent. Still, these established controls display several disadvantages. Input controls accurately describe the chromatin landscape of used chromatin material, but do not reflect bias introduced during the ChIP procedure. Whereas IgG controls exactly match the experimental steps of the ChIP samples they suffer from high variability, over-amplification and overestimation of background signals, leading to decreased signal-to-noise ratios and false-negative signals (251). In contrast, knockout samples combine both mentioned advantages – complete ChIP protocol bias description and the accurate display of the chromatin landscape - without the disadvantages of input or IgG controls.

The ChIP-seq experimental procedures themselves can introduce variation and potential bias due to different complex enzymatic reactions following DNA purification. “Hyper-ChIPable” regions are a special type of introduced ChIP-seq bias and are targets of massive protein binding with open chromatin states (250). These naturally occurring genomic sites are targets of non-specific ChIP enrichments, irrespective of the used antibody, potentially due to their open chromatin state allowing electrostatic or binding interactions of RNA polymerases or DNA to antibodies. KOIN identified 18 “hyper-ChIPable” regions in all six data sets and successfully removed these special type of false-positive signals (Figure 11). Importantly, sites with extraordinary high ChIP protein binding enrichments are of special interest for the interpretation of ChIP-seq results and the analysis of TF binding patterns. Without the KOIN method, all sites with high ChIP-seq signals are considered for data analysis, irrespective of their potential false-positive character.

Another advantage of KOIN is the possible integration into other ChIP-seq pipelines, utilizing different peak calling algorithms. In parallel to the herein presented protocol for ChIP-seq analysis, KOIN was tested with a second peak calling algorithm called SICER (269,300). Despite the detected improvements on data analysis with SICER utilizing KO data sets, MACS outperformed SICER. Different peak calling algorithms influence the identification of ChIP-protein enriched regions and thus can have an impact on the downstream biological data interpretation (269,301,302). Until now, no gold standard was defined for peak calling

algorithms. Every peak caller has its own beneficial properties and drawbacks strongly depending on the type of ChIP-seq data used as input for the data analysis. For example, the used MACS algorithm (262) showed a good performance not only in classical ChIP-seq but also in DNase-seq experiments (302) and is one of the best established algorithms until now.

One limitation of the KOIN method is the requirement for ChIP protein knockout cells. In murine models, the relatively new CRISPR/Cas-mediated genome engineering approach allows a fast and efficient generation of knockout mouse models and is also easily adaptable for the usage in other mammalian cells (303,304). This technique utilizes the CRISPR (clustered regulatory interspaced short palindromic repeat) and Cas (CRISPR-associated) proteins, originally found in bacteria and archaea mediating a RNA-based adaptive immunity (305,306). A single guide RNA (sgRNA) is artificially constructed to direct the Cas9 protein to a specific genomic DNA sequence to edit the genome in an efficient and simple manner, introducing for example reporter constructs or mediating gene knockouts in mammalian cells. The generation of KO control cells via the CRISPR/Cas technique might be a way to establish a convincing control for ChIP-seq data sets.

The positive effects on ChIP-seq data analysis utilizing KOIN peak calling demonstrated the profound influence and value of KOIN to reduce false-positive peak calls. Future studies of novel TFs, chromatin regulators or the recently introduced “occupied regions of genomes from affinity purified naturally isolated chromatin “(ORGANIC) method could greatly benefit from taking KOIN into account (307). The ORGANIC method uses native chromatin to identify direct TF chromatin interactions with high resolution, specificity and sensitivity in comparison to standard protocols using cross-linked chromatin. KOIN could further minimize bias and maximize specificity in combination with the advantages of native chromatin preparations, circumventing epitope masking and allowing minimal chromatin amounts as input.

In conclusion, KOIN should be considered as gold standard for the analysis of TF ChIP-seq data, optimizing not only false-positive signal reduction amongst others by the exclusion of “hyper-ChIPable” regions but also significantly improving the biological interpretation of TF binding data.

6.4. A common epigenetic core program defines the basic transcriptional landscape in macrophages

In the second part of this thesis, the epigenetic changes during activation of a human somatic cell type, namely monocyte-derived macrophages, were studied. Macrophages exist

upon stimulatory cues in different stimulus specific activation states and can thereby mediate various functions in inflammation, immune response modulation and tissue repair (33,308). Strong overlaps were found in this study for the epigenetic states of many regulatory promoter and enhancer sites in all investigated macrophages (Figure 29, Figure 30). The described common regions establish a macrophage specific transcriptional program (Figure 32) with corresponding general “housekeeping” functions (Figure 31). These high similarities were previously visible in different murine myeloid cells with 82 % overlap for H3K4me3 identified promoter regions comparing macrophages, monocytes and neutrophils (157). Common enhancer regions were less abundant (27 %) but still present. In terms of the similar origin of these cells, this common signature is not surprising, while cells only differ in their activation stimulus. In line with these results, even mouse tissue macrophages originated from different organs like lung or brain seem to manifest a common macrophage signature (177). Additionally, PU.1 but also other TFs like ETS1, GABPA or FLI1 seem to maintain the common epigenetic landscape of human macrophages (Figure 33). In previous publications, the cooperation of PU.1 with secondary TFs of the ETS family was described in mouse models amongst others in macrophage cells to initiate and maintain their transcriptional program (255,274,283,284). Furthermore, with a previous computational reverse engineering approach (21) commonly found promoter and enhancer sites could be verified by a high overlap of previous common macrophage core signature genes with active epigenetic promoter and enhancer markings rendering these genes as highly transcriptionally active, thus demonstrating their general importance (Figure 30, Figure 34).

6.5. Profound changes at promoter or *cis*-regulatory sites alter the transcriptional program in activated human macrophages

In this study, 430 differentially regulated H3K4me3 promoters were found while comparing four differentially activated primary human macrophages, activated under different stimulatory conditions in a pro-inflammatory background (Figure 22). These sites are involved in macrophage specific biological processes, exemplified for M^{TPP} and positively correlated with the transcriptional activity of corresponding genes (Figure 23, Figure 24). In comparison, 873 differential H3K4me3 promoters were previously found in murine macrophages compared to monocyte and neutrophils which are in line with found numbers in this study (157). These changes reflect the induced transcriptional changes on promoter level during the macrophage activation process. Interestingly, at least two times higher numbers of repressed promoters were detected in human M^{IL4} macrophages in comparison

to the other activation states. This could hint to a necessary repression of genes to prevent their transcriptional activity and finally to the establishment of the M^{IL4} phenotype with corresponding functions. In mouse models, comparable repressive cascades were described upon IL-4 stimulation e.g. for STAT6, KLF4 or PPAR proteins, inhibiting important mechanisms to finally repress a pro-inflammatory macrophage program (7).

In addition, a huge number of *cis*-regulatory enhancer regions with approximately 20.000 stimulus specific enhancer sites could be identified in the four investigated macrophage activation states (Figure 25). Their activity HM classification also positively correlated to the transcriptional activity of adjacent genes (Figure 26) and they outnumber the stimulus specific promoter regions determined before (Figure 22). The vast increase in activation state specific enhancer numbers indicates their extreme variability and complexity and also display their important role during the activation process of human macrophages as switch for transcriptional regulation and binding sites for signal-dependent TFs (71,72,188,309,310).

Furthermore, comparable numbers of super enhancers (SE) as determined here for human macrophages were found in previously described murine macrophages (169) (Figure 35). SEs were located in the vicinity of activation state relevant macrophage genes with corresponding functions and substantially increased the transcriptional activity of these genes (Figure 36, Figure 37). SEs could play an important role in the regulation of macrophage genes, since their role in the determination of cell fate and lineage in murine tissue macrophages or stem cells was previously verified (169,170,311).

This study further demonstrated the repertoire of *cis*-regulatory elements is shaped by different external and internal stimulatory events cells may undergo. Some of these important regulatory sites are established during cell differentiation and are directly controlled by the enhancer-organizing activities of lineage-determining TFs, which was demonstrated for PU.1 in macrophages (57,71,72). Other enhancers are formed by the functional cooperation between stimulus-activated TFs and lineage-restricted TFs (73). These stimulus dependent “latent enhancers” were demonstrated in murine macrophages to form an epigenetic memory after stimulatory signals and to change cellular reactions upon future stimulatory cues. Latent enhancers, defined by a *de novo* formation of H3K4me1 signals upon stimulatory signals are still present once the stimulatory signal, which established them, is terminated. It can be assumed, that the largest group of differential enhancer sites, identified as specific weak enhancers with solely H3K4me1 signals (Figure 25) partially consist of latent enhancers with no direct effect on the actual transcriptional program, but are important for putative reactions upon future stimulatory signals. Another argument for their “latent” enhancer character could be the loss of PU.1 binding (Figure 25) at many of the activation specific enhancer sites.

Epigenetic memory was also demonstrated for other HMs like H3K27me3 in *arabidopsis* (312), H3K9me3 in murine fibroblasts and pluripotent cells (313) or H3K4me1/2 and H3K9me3 in *Caenorhabditis elegans* (314). Interestingly, not all HMs seem to be able to maintain long-lasting marks at corresponding genomic locations (315). Mainly histone methylation, like H3K4me3 at promoter or H3K4me1 at enhancer regions, was suggested to act as a stable mark to establish longer lasting transcriptional changes with the potential to establish a mechanism for epigenetic memory.

The correlation of promoters with or without *cis*-regulatory elements located in their vicinity resulted in an increase of transcriptional activity (Figure 28B). Enhancers can increase gene expression by direct interactions forming chromatin loops or by indirect effects mediated by eRNAs (67,112,127,128). The strong overlaps for promoter and enhancer correlated gene names, especially for M^{IL4} and M^{TPP} but also for M^{IFN γ} in a less prominent but still significant manner, demonstrated the influence and correlation of promoter sites and enhancer presence in their surroundings (Figure 28A). Interestingly, the activation of human macrophages seems to influence the promoter-enhancer correlation. In unstimulated baseline macrophages (M^b), the overlaps are less specific in contrast to the other macrophage populations. Despite the defined activation state specificity for described promoter and enhancer sites, the results could indicate the usage of M^b enhancers not only for M^b promoters but also for other macrophage activation states. This postulate would also explain the strong correlation of Ep sites in M^{IFN γ} also with Pp sites in M^{TPP}, potentially reflecting their common inflammatory background and the partially shared character of these enhancer sites. More importantly, the method of promoter-enhancer pair identification could be partially responsible for these results as a potential source of variation.

The identification of regulatory pairs for promoter and corresponding enhancer sites on a global scale still remains difficult. Interactions between enhancers and their target genes can be orientation-independent, characterized by distances with up to 1 megabase pairs and enhancers can even be located within the coding region of the corresponding gene (58-60). Furthermore, enhancers and promoters can be targets of multiple interactions, rendering the proper identification of promoter-enhancer pairings even more complex. Until now, the common approach to annotate *cis*-regulatory elements to corresponding regulated genes is established by the identification of the nearest known TSS relative to the center of the corresponding enhancer site. This approach is a fast and efficient solution, especially for a high number of enhancer regions in the proximity of TSS for known genes. But due to previously mentioned enhancer properties, this approach is not an optimal solution, leading to false-positive or negative results. For single *cis*-regulatory elements, the classical luciferase assay is still a valid method to identify corresponding target genes, but for

hundreds or thousands of enhancer regions a global approach with the luciferase assay technique is not feasible. Future genome wide solutions, which are already utilized for a better and more precise prediction of promoter and enhancer cooperation could be the Hi-C seq method (222,316). This technique allows the prediction of 3D chromatin structures in a genome wide manner, thus identifying promoter to enhancer interactions by forming chromatin loops, allowing a direct interaction of factors involved in the regulation and transcription of the corresponding gene locus. Another revolutionary approach to identify promoter-enhancer pairs could be guanine-rich nucleic acid sequences capable of forming a four-stranded structure through Hoogsteen hydrogen bonding, also called “G-quadruplexes (317). These G-rich patterns found at the boundaries of DNase I hypersensitive promoters and enhancers could facilitate the formation of G-quadruplexes, thus promoting the formation of chromatin loops which mediate the interaction of promoter and enhancer sites (67). Another relatively new approach to identify actively transcribed enhancers in a broad variety of cell types e.g. human myeloid cells was performed by the FANTOM5 project by the detection of the bidirectional promoter activity of active enhancers with the CAGE (Cap analysis of gene expression) method and the correlation to histone modifications H3K4me1 and H3K27Ac (318-320).

In conclusion, several identified epigenetic promoter and enhancer differences for investigated macrophage activation states influence the transcriptional landscapes of these cells. A more complex analysis specifically for *cis*-regulatory elements is required to reveal more insights how these important sites impact the establishment of different macrophage activation programs by the regulation of corresponding promoter sites and the establishment of epigenetic memory.

6.6. Specialized transcriptional control of TRs in activated human macrophages

Due to the general differences found in epigenetic promoter and enhancer states in activated human macrophages, a detailed correlation analysis was performed for important TR proteins. Transcriptional regulators (TRs) like transcription factors (TFs), chromatin remodelers or co-factors are crucial elements responsible for changes on transcriptional levels and may act as cornerstones, establishing diverse macrophage programs upon inflammation-associated stimulatory signals in an input signal-specific manner. Astonishingly, co-regulation networks of TRs in differentially activated macrophages resulted in an overwhelming fraction of accessible sites with minimal differences between macrophage populations, despite clear transcriptional alterations (Figure 39, Figure 40, Figure 41). This

pre-defined open chromatin landscape in inflammatory macrophages could suggest that gene expression is solely guided by transcriptional regulatory processes mediated by input-specific master TRs. In contrast, differential observations were made for histone modification states in human tissues or murine tissue macrophages originating from different tissues. These cells display a tight regulation of TR on epigenetic level with an accessible chromatin landscape for highly expressed genes and vice versa inaccessible chromatin for non-expressed genes following a previously described general model (Figure 43, Figure 45, Figure 46) (167). Interestingly, peritoneal and colonic macrophages showed the highest numbers of genes not transcribed yet with accessible promoter markings in tissue macrophages, demonstrating a closer relation to the special epigenetic regulatory mechanisms in monocytes than other tissue macrophages. This closer connection to human monocyte-derived macrophages could be explained by the previously found replacement of these cells by monocyte-derived cells in the adulthood (321). In summary, these results suggest that the accessibility of the vast majority of TR loci is a special feature of TRs in activated monocyte-derived macrophages. This observed uncoupling of epigenetic and transcriptional regulation of TRs could not be found in tissue macrophages.

Only one other example of the uncoupling of epigenetic and transcriptional control was very recently found in stem cells and their progenies within intestinal crypts, where cell plasticity was associated with broadly permissive chromatin (322). The expression of one specific TF was responsible to switch cell phenotypes and functions from enterocytes to secretory cells and vice versa, despite the similarly open chromatin landscape. Furthermore, the enormous plasticity of gene expression of human macrophage activation demonstrated in the previously introduced multi-dimensional model (21) could be a direct consequence of expression-dependent regulation within this specialized TR network. Other cell types with functional plasticity may also contain similar activation TR networks, while the majority of other cell types without a broad functional plasticity are precisely regulated by epigenetic mechanisms.

Aside from macrophage plasticity, macrophages also need to react fast and efficient to environmental cues in the organism to mediate proper immune responses and other critical functions. It is speculated that the posttranslational HM, one important aspect of epigenetic regulation is a relatively slow process with multiple necessary steps until the transcriptional activity of corresponding genomic loci are altered (315,323). Pioneer factors open the chromatin landscape with or without secondary factors and recruit HM enzymes, which covalently modify histone proteins, leading to transcriptional initiation and elongation mediated by RNA polymerase II (73,324-326). Interestingly, amongst putative posttranslational modifications especially the methylation of histone proteins was described

to be a relatively slow and stable process (315). This observation could support the overall open promoter landscape in activated macrophages at TR regions, implicating other regulatory levels besides the posttranslational modification of histone residues.

First of all, specific repressor proteins like BCL6 (B cell leukemia) or nuclear receptors recruiting co-repressor complexes with additional transcriptional regulatory functions can restrict the expression of TRs (154,287,327,328). Other repressive HM marks, which are not part of this study like H3K9me3 or H4K20me3 could down-regulate the transcriptional activity of adjacent genes (329-331). Importantly, a cross-talk between HM was demonstrated for example by the stronger interaction of PHF8 to H3K4me3 with simultaneous H3K9 and H3K27 acetylation in human cells (332). Not only the methylation of histone proteins can repress gene transcription, but also the ubiquitinylation of H2AK119 inhibits the elongation of RNA polymerase II enzymes (333). Furthermore, eRNA transcripts could greatly influence the TR gene expression rates by RNA-mediated repression, altering for example the formation of DNA loopings between enhancer and promoter sites or changing the chromatin structure level itself (112,127,128). Also nuclear receptors can alter eRNA transcription and thus down-regulate the transcription of target genes, as it has been demonstrated for the Rev-Erb α and Rev-Erb- β proteins in murine macrophages (186,334). In addition, a disruption of the interaction between acetylated histone proteins and the bromodomain and extra terminal domain (BET) “epigenetic reader” proteins could additionally regulate transcription by the elongation of RNA polymerase II. This mechanism was described to selectively block the expression of a subset of TLR4 induced genes in mouse models of endotoxic shock or for pro-inflammatory cytokine expression (335-337). BET family members are indeed expressed in the utilized human macrophage model in all investigated populations, thus underlining their importance for a fast transcriptional response in macrophages (data not shown). Another possible regulatory level could be the extent of differential TR expression between differentially activated macrophages. In case of PU.1, differences in protein expression lead to the differentiation of hematopoietic progenitors into a B cell-like phenotype or into a macrophage cell type (75). In both cases, PU.1 is expressed and thus an open promoter state with corresponding *cis*-regulatory enhancer markings could be assumed. The same mechanism could be true for many TR genes, allowing for rapid changes in expression values upon stimulatory signals, due to already open promoter markings, rendering slow HM changes unnecessary.

Eventually, post-translational HMs are deposited as a consequence of TF binding. Multiple TFs bind in combinations or as multi-factor proteins to promoter and enhancer positions and play key regulatory roles in their activation (327). TF binding seems to be far more dynamic than the turnover of chromatin marks, which thus can extend and stabilize cellular activation

signals (247). In recent publications, the importance and abundance of these genomic loci co-occupied by multiple TFs was demonstrated in multiple cell types, mediating a complex regulation of transcriptional processes by their combinatorial binding (319,338-340). Possible candidates for important regulatory TFs in macrophages could be especially STAT TFs, previously demonstrated as essential proteins for the establishment of transcriptional programs in differentially activated macrophages. First of all, STAT1 was strongly associated to pro-inflammatory human or mouse macrophages important for IFN γ responsive genes (196,293,341), while STAT6 is induced by IL-4 stimulation in human and mouse macrophages and critically regulates PPAR γ function (25,342,343). Moreover, the motif enrichment analysis of Pa and Es revealed the enrichment of STAT6 in M^{IL4} macrophages (Figure 25). High STAT4 expression and relevance was previously demonstrated in M^{TPP} cells (21). STAT TFs were highly expressed in corresponding macrophage populations and were marked with Pa and Es markings (Figure 41C). These TFs could act in combination with other important regulatory proteins like PU.1, NF-kB or C/EBP with putative lineage determining functions in macrophages (92,338,344,345). In fact, multiple publications described the synergistic cooperation of STAT proteins with PU.1, C/EBP, NF-kB, AP-1 and other proteins (341,346-349). PU.1 and other lineage determining TFs may provide macrophages with a basic and plastic epigenomic architecture with many open and accessible sites, upon combination with other signals or TFs, a fine-tuning mechanism resulting in an optimal transcriptional output adapted to external and internal stimulatory signals (92).

Collectively, the epigenetic and transcriptional regulation in macrophage activation TR networks is uncoupled with an overall accessible epigenetic promoter and enhancer state, in contrast to other tissue cells or tissue macrophages without a functional and cellular plasticity, where a classical model of tight epigenetic and transcriptional correlation has emerged. Other regulatory mechanisms like eRNAs or TF binding could fine-tune the signal integration-dependent transcriptional regulation of network TRs.

6.7. Epigenetics - the “next generation” tool for the identification of regulatory genomic sites

The exciting identification of a “histone code”, connecting specific posttranslational histone modifications read by effector proteins to specific biological functions in virtually all known cell types, concomitantly with the development of the next generation sequencing technology allowed the location of regulatory sites and the discrimination of their activity status with unprecedented precision in a genome wide manner (145,146). This was demonstrated not

only in multiple cell types and species for the identification of *cis*-regulatory elements (62,160,161,350), but also for the assessment of disease relevant aspects (351). For the first time, differentially activated human macrophages were in detail analyzed in this study, correlating epigenetic activity states of regulatory promoter or enhancer sites with transcriptional RNA-seq data. Previous chapters described fascinating common but also differential properties of activated human macrophages and shed light especially to the regulatory processes of TRs, visualized in co-regulation networks.

The integration of collected data allowed the generation of a model, which described the epigenetic modification of important macrophage genes during the differential activation process (Figure 47). A macrophage core program defines sets of active and inactive genes important for general macrophage functions like the SP1 gene encoding for the TF PU.1 or PRDM1 (Figure 39B), while a stimulus specific active program promotes the specialization of the corresponding activated macrophage, like IL2RA for M^{TPP} (Figure 27). In contrast, a co-regulation network of important TR genes revealed an uncoupling of transcriptional and epigenetic regulation with a mainly accessible epigenetic landscape, which suggests other regulatory mechanisms in concert with accessible epigenetic HM marks in monocyte-derived human macrophages (Figure 41), which was discussed previously.

Additional future approaches are needed to elucidate this interesting point to extend the gathered information on HM changes in macrophages and link new knowledge to loose ends of macrophage biology. Within this study, cells in the used macrophage models were stimulated for 72 hours, harvested and were subsequently used for ChIP-seq experiments. Between the first minutes of stimulation and at later time points, various cascades of coordinate regulatory mechanisms display concerted actions in macrophage differentiation. Future time kinetic experiments could exactly specify the time points when epigenetic changes occur and thus, shed more light onto their functions mediating early responses at genomic regions for important TFs like TNF or NFkBIA enabling late responses for the maintenance of the corresponding macrophage activation state (352-354). Furthermore, previously demonstrated transient epigenetic effects in liver development can cause stable changes on transcriptional level (355). Comparable transient changes on epigenetic level during early phases of the activation process could be lost after the complete activation of macrophages and can thus only be targeted with the implementation of multi-time point experiments (68,280). Until now, it can only be speculated, how many reversible epigenetic marks are established during the differential activation of human macrophages in comparison to stable HM marks, visible after the completion of macrophage activation.

Other experimental approaches could further collect information about complete activation of human macrophages on the RNA polymerase II level. Global Run-On sequencing (GRO-

seq) experiments, already performed for example in mouse macrophages and cancer cells (162,186,356) could detect nascent RNA transcripts still associated with the RNA polymerase II enzyme in differentially activated macrophages. These experiments could further specify regulatory regions with high RNA polymerase II activity in combination with epigenetic HM signals at promoter or enhancer locations. Additionally, resulting precise catalogues of eRNA transcripts could assess the relevance of these non-coding and relatively unstable RNA species (186) in the regulation of human macrophage genes. Thus, especially genes in the depicted transcriptional co-regulation networks could be regulated by eRNAs, positively correlating with stabilized promoter-enhancer interactions and higher gene expression for corresponding adjacent genes and vice versa leading to decreased expression values after eRNA suppression.

In addition to future experiments which aim to gather further invaluable information about human macrophage biology, a few general issues should be kept in mind. For the characterization of human macrophages, the direct translation of data obtained from mouse models or from other species is a quite valuable approach, but should not be used in an unquestionable manner, without the consideration of important facts concerning murine and human macrophage biology and utilized experimental approaches (308). First, human inflammatory monocyte-derived macrophages are in general generated from blood monocytes, while murine macrophages are mostly generated from bone marrow *in vitro* or directly isolated from tissues. Moreover, differences in the production of iNOS and nitrite oxygen (NO) with antimicrobial functions seem to show fundamental differences in experiments performed with human or mouse macrophages (357-361). A similar scenario was described for the Arg1 enzyme. In humans it is only expressed in neutrophils, while it was found in both murine neutrophils and macrophages (362-364). Due to these differences, protein markers for specific murine macrophage states could not be validated in human macrophages and vice versa (365). Recent studies (21,366,367) tried to elucidate the similarities and differences of macrophages in human and mouse macrophage models to understand their core similarities but also highlight obvious evolutionary differences influencing their immune responses, like the IFN-regulated IRG protein with two members in human and approximately 18 in the mouse (368).

In addition to the mouse-human comparability, the actual *in vivo* situation for human macrophages should always stay in focus of future research. In this study, primary human macrophages were used in an inflammatory model using different stimuli to create functional activation states to approach the *in vivo* situation (Figure 3). In contrast to immortalized cell lines, where various artificial alterations to their genetic background could render results generated and biological interpretations at least partially questionable. The ENCODE project

(204) uses computational approaches (21,156,338) and other similar efforts to standardize and collect NGS data sets and thus makes the correlation of different data sets on various levels (RNA, promoter, enhancer, histones) possible to get with individual experiments more and more insights into complex biological systems. In this study generated data sets elucidated macrophage biology and added more aspects of information for the activation of primary human macrophages and their epigenetic regulation to the increasing amount of NGS databases. Further epigenetic experiments with human tissue macrophages in comparison to monocytes and monocyte-derived inflammatory macrophages cannot only shed more light on their regulatory processes, commonalities and differences but can also help to study epigenetic influences in macrophages in context of human diseases. Finally, single cell experiments concerning the genetic variability of macrophage regulation in different individuals can be important for personalized medicine-approaches for future patients.

7. References

1. Medzhitov, R. (2013) Pattern recognition theory and the launch of modern innate immunity. *J Immunol*, **191**, 4473-4474.
2. Janeway, C.A., Jr. (1989) Approaching the asymptote? Evolution and revolution in immunology. *Cold Spring Harbor symposia on quantitative biology*, **54 Pt 1**, 1-13.
3. Iwasaki, A. and Medzhitov, R. (2010) Regulation of adaptive immunity by the innate immune system. *Science*, **327**, 291-295.
4. Medzhitov, R. and Janeway, C., Jr. (2000) Innate immune recognition: mechanisms and pathways. *Immunological reviews*, **173**, 89-97.
5. Medzhitov, R. and Janeway, C., Jr. (2000) The Toll receptor family and microbial recognition. *Trends in microbiology*, **8**, 452-456.
6. Akira, K., Amano, M., Okajima, F., Hashimoto, T. and Oikawa, S. (2006) Inhibitory effects of amlodipine and fluvastatin on the deposition of advanced glycation end products in aortic wall of cholesterol and fructose-fed rabbits. *Biological & pharmaceutical bulletin*, **29**, 75-81.
7. Sica, A. and Mantovani, A. (2012) Macrophage plasticity and polarization: in vivo veritas. *The Journal of clinical investigation*, **122**, 787-795.
8. Gordon, S. and Martinez, F.O. (2010) Alternative activation of macrophages: mechanism and functions. *Immunity*, **32**, 593-604.
9. Biswas, S.K. and Mantovani, A. (2010) Macrophage plasticity and interaction with lymphocyte subsets: cancer as a paradigm. *Nature immunology*, **11**, 889-896.
10. Chow, A., Brown, B.D. and Merad, M. (2011) Studying the mononuclear phagocyte system in the molecular age. *Nature reviews. Immunology*, **11**, 788-798.
11. Geissmann, F., Manz, M.G., Jung, S., Sieweke, M.H., Merad, M. and Ley, K. (2010) Development of monocytes, macrophages, and dendritic cells. *Science*, **327**, 656-661.
12. Davies, L.C., Jenkins, S.J., Allen, J.E. and Taylor, P.R. (2013) Tissue-resident macrophages. *Nature immunology*, **14**, 986-995.
13. Kohyama, M., Ise, W., Edelson, B.T., Wilker, P.R., Hildner, K., Mejia, C., Frazier, W.A., Murphy, T.L. and Murphy, K.M. (2009) Role for Spi-C in the development of red pulp macrophages and splenic iron homeostasis. *Nature*, **457**, 318-321.
14. Ganz, T. (2012) Macrophages and systemic iron homeostasis. *Journal of innate immunity*, **4**, 446-453.
15. Paolicelli, R.C., Bolasco, G., Pagani, F., Maggi, L., Scianni, M., Panzanelli, P., Giustetto, M., Ferreira, T.A., Guiducci, E., Dumas, L. *et al.* (2011) Synaptic pruning by microglia is necessary for normal brain development. *Science*, **333**, 1456-1458.
16. Schafer, D.P., Lehrman, E.K., Kautzman, A.G., Koyama, R., Mardinly, A.R., Yamasaki, R., Ransohoff, R.M., Greenberg, M.E., Barres, B.A. and Stevens, B. (2012) Microglia sculpt postnatal neural circuits in an activity and complement-dependent manner. *Neuron*, **74**, 691-705.
17. Wright, J.R. (1990) Clearance and recycling of pulmonary surfactant. *The American journal of physiology*, **259**, L1-12.
18. Hume, D.A. (2008) Differentiation and heterogeneity in the mononuclear phagocyte system. *Mucosal immunology*, **1**, 432-441.
19. Italiani, P. and Boraschi, D. (2014) From Monocytes to M1/M2 Macrophages: Phenotypical vs. Functional Differentiation. *Frontiers in immunology*, **5**, 514.
20. Gundra, U.M., Girgis, N.M., Ruckerl, D., Jenkins, S., Ward, L.N., Kurtz, Z.D., Wiens, K.E., Tang, M.S., Basu-Roy, U., Mansukhani, A. *et al.* (2014) Alternatively activated

- macrophages derived from monocytes and tissue macrophages are phenotypically and functionally distinct. *Blood*, **123**, e110-122.
21. Xue, J., Schmidt, S.V., Sander, J., Draffehn, A., Krebs, W., Quester, I., De Nardo, D., Gohel, T.D., Emde, M., Schmidleithner, L. *et al.* (2014) Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. *Immunity*, **40**, 274-288.
 22. Mosser, D.M. and Edwards, J.P. (2008) Exploring the full spectrum of macrophage activation. *Nature reviews. Immunology*, **8**, 958-969.
 23. Krausgruber, T., Blazek, K., Smallie, T., Alzabin, S., Lockstone, H., Sahgal, N., Hussell, T., Feldmann, M. and Udalova, I.A. (2011) IRF5 promotes inflammatory macrophage polarization and TH1-TH17 responses. *Nature immunology*, **12**, 231-238.
 24. Mantovani, A., Sica, A., Sozzani, S., Allavena, P., Vecchi, A. and Locati, M. (2004) The chemokine system in diverse forms of macrophage activation and polarization. *Trends in immunology*, **25**, 677-686.
 25. Sica, A. and Bronte, V. (2007) Altered macrophage differentiation and immune dysfunction in tumor development. *The Journal of clinical investigation*, **117**, 1155-1166.
 26. Hoesel, B. and Schmid, J.A. (2013) The complexity of NF-kappaB signaling in inflammation and cancer. *Molecular cancer*, **12**, 86.
 27. Lawrence, T. and Gilroy, D.W. (2007) Chronic inflammation: a failure of resolution? *International journal of experimental pathology*, **88**, 85-94.
 28. Bogels, M., Braster, R., Nijland, P.G., Gul, N., van de Luitgaarden, W., Fijneman, R.J., Meijer, G.A., Jimenez, C.R., Beelen, R.H. and van Egmond, M. (2012) Carcinoma origin dictates differential skewing of monocyte function. *Oncoimmunology*, **1**, 798-809.
 29. Braster, R., O'Toole, T. and van Egmond, M. (2014) Myeloid cells as effector cells for monoclonal antibody therapy of cancer. *Methods*, **65**, 28-37.
 30. Allavena, P., Sica, A., Garlanda, C. and Mantovani, A. (2008) The Yin-Yang of tumor-associated macrophages in neoplastic progression and immune surveillance. *Immunological reviews*, **222**, 155-161.
 31. Bozza, F.A., Salluh, J.I., Japiassu, A.M., Soares, M., Assis, E.F., Gomes, R.N., Bozza, M.T., Castro-Faria-Neto, H.C. and Bozza, P.T. (2007) Cytokine profiles as markers of disease severity in sepsis: a multiplex analysis. *Crit Care*, **11**, R49.
 32. Ambarus, C.A., Krausz, S., van Eijk, M., Hamann, J., Radstake, T.R., Reedquist, K.A., Tak, P.P. and Baeten, D.L. (2012) Systematic validation of specific phenotypic markers for in vitro polarized human macrophages. *Journal of immunological methods*, **375**, 196-206.
 33. Martinez, F.O., Sica, A., Mantovani, A. and Locati, M. (2008) Macrophage activation and polarization. *Frontiers in bioscience : a journal and virtual library*, **13**, 453-461.
 34. Lang, R., Patel, D., Morris, J.J., Rutschman, R.L. and Murray, P.J. (2002) Shaping gene expression in activated and resting primary macrophages by IL-10. *J Immunol*, **169**, 2253-2263.
 35. Gordon, S. (2003) Alternative activation of macrophages. *Nature reviews. Immunology*, **3**, 23-35.
 36. Whyte, C.S., Bishop, E.T., Ruckerl, D., Gaspar-Pereira, S., Barker, R.N., Allen, J.E., Rees, A.J. and Wilson, H.M. (2011) Suppressor of cytokine signaling (SOCS)1 is a key determinant of differential macrophage activation and function. *Journal of leukocyte biology*, **90**, 845-854.
 37. Odegaard, J.I., Ricardo-Gonzalez, R.R., Goforth, M.H., Morel, C.R., Subramanian, V., Mukundan, L., Red Eagle, A., Vats, D., Brombacher, F., Ferrante, A.W. *et al.* (2007) Macrophage-specific PPARgamma controls alternative activation and improves insulin resistance. *Nature*, **447**, 1116-1120.
 38. Odegaard, J.I., Ricardo-Gonzalez, R.R., Red Eagle, A., Vats, D., Morel, C.R., Goforth, M.H., Subramanian, V., Mukundan, L., Ferrante, A.W. and Chawla, A. (2008)

- Alternative M2 activation of Kupffer cells by PPARdelta ameliorates obesity-induced insulin resistance. *Cell metabolism*, **7**, 496-507.
39. Liao, X., Sharma, N., Kapadia, F., Zhou, G., Lu, Y., Hong, H., Paruchuri, K., Mahabeleshwar, G.H., Dalmas, E., Venteclef, N. *et al.* (2011) Kruppel-like factor 4 regulates macrophage polarization. *The Journal of clinical investigation*, **121**, 2736-2749.
 40. Martinez, F.O., Helming, L. and Gordon, S. (2009) Alternative activation of macrophages: an immunologic functional perspective. *Annual review of immunology*, **27**, 451-483.
 41. Mantovani, A., Schioppa, T., Biswas, S.K., Marchesi, F., Allavena, P. and Sica, A. (2003) Tumor-associated macrophages and dendritic cells as prototypic type II polarized myeloid populations. *Tumori*, **89**, 459-468.
 42. Wang, J., Jiang, Z.P., Su, N., Fan, J.J., Ruan, Y.P., Peng, W.X., Li, Y.F. and Yu, X.Q. (2013) The role of peritoneal alternatively activated macrophages in the process of peritoneal fibrosis related to peritoneal dialysis. *International journal of molecular sciences*, **14**, 10369-10382.
 43. Zhang, B., Yao, G., Zhang, Y., Gao, J., Yang, B. and Rao, Z. (2011) M2-polarized tumor-associated macrophages are associated with poor prognoses resulting from accelerated lymphangiogenesis in lung adenocarcinoma. *Clinics (Sao Paulo)*, **66**, 1879-1886.
 44. Lewis, C.E. and Pollard, J.W. (2006) Distinct role of macrophages in different tumor microenvironments. *Cancer research*, **66**, 605-612.
 45. Marino, M.W., Dunn, A., Grail, D., Inglese, M., Noguchi, Y., Richards, E., Jungbluth, A., Wada, H., Moore, M., Williamson, B. *et al.* (1997) Characterization of tumor necrosis factor-deficient mice. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 8093-8098.
 46. Popov, A., Abdullah, Z., Wickenhauser, C., Saric, T., Driesen, J., Hanisch, F.G., Domann, E., Raven, E.L., Dehus, O., Hermann, C. *et al.* (2006) Indoleamine 2,3-dioxygenase-expressing dendritic cells form suppurative granulomas following *Listeria monocytogenes* infection. *The Journal of clinical investigation*, **116**, 3160-3170.
 47. Shay, J.E. and Celeste Simon, M. (2012) Hypoxia-inducible factors: crosstalk between inflammation and metabolism. *Seminars in cell & developmental biology*, **23**, 389-394.
 48. Popov, A., Driesen, J., Abdullah, Z., Wickenhauser, C., Beyer, M., Debey-Pascher, S., Saric, T., Kummer, S., Takikawa, O., Domann, E. *et al.* (2008) Infection of myeloid dendritic cells with *Listeria monocytogenes* leads to the suppression of T cell function by multiple inhibitory mechanisms. *J Immunol*, **181**, 4976-4988.
 49. Naik, S.H., Perie, L., Swart, E., Gerlach, C., van Rooij, N., de Boer, R.J. and Schumacher, T.N. (2013) Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature*, **496**, 229-232.
 50. Staber, P.B., Zhang, P., Ye, M., Welner, R.S., Nombela-Arrieta, C., Bach, C., Kerenyi, M., Bartholdy, B.A., Zhang, H., Alberich-Jorda, M. *et al.* (2013) Sustained PU.1 levels balance cell-cycle regulators to prevent exhaustion of adult hematopoietic stem cells. *Molecular cell*, **49**, 934-946.
 51. Serandour, A.A., Avner, S., Percevault, F., Demay, F., Bizot, M., Lucchetti-Miganeh, C., Barloy-Hubler, F., Brown, M., Lupien, M., Metivier, R. *et al.* (2011) Epigenetic switch involved in activation of pioneer factor FOXA1-dependent enhancers. *Genome research*, **21**, 555-565.
 52. Zheng, R. and Blobel, G.A. (2010) GATA Transcription Factors and Cancer. *Genes & cancer*, **1**, 1178-1188.
 53. Ostuni, R. and Natoli, G. (2013) Lineages, cell types and functional states: a genomic view. *Current opinion in cell biology*, **25**, 759-764.
 54. Zaret, K.S. and Carroll, J.S. (2011) Pioneer transcription factors: establishing competence for gene expression. *Genes & development*, **25**, 2227-2241.

55. Soufi, A., Donahue, G. and Zaret, K.S. (2012) Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell*, **151**, 994-1004.
56. Magnani, L., Eeckhoutte, J. and Lupien, M. (2011) Pioneer factors: directing transcriptional regulators within the chromatin environment. *Trends in genetics : TIG*, **27**, 465-474.
57. Natoli, G. (2010) Maintaining cell identity through global control of genomic organization. *Immunity*, **33**, 12-24.
58. Bulger, M. and Groudine, M. (2011) Functional and mechanistic diversity of distal transcription enhancers. *Cell*, **144**, 327-339.
59. Levine, M. (2010) Transcriptional enhancers in animal development and evolution. *Current biology : CB*, **20**, R754-763.
60. Roy, A.L., Sen, R. and Roeder, R.G. (2011) Enhancer-promoter communication and transcriptional regulation of Igh. *Trends in immunology*, **32**, 532-539.
61. Capone, M., Watrin, F., Fernex, C., Horvat, B., Krippel, B., Wu, L., Scollay, R. and Ferrier, P. (1993) TCR beta and TCR alpha gene enhancers confer tissue- and stage-specificity on V(D)J recombination events. *The EMBO journal*, **12**, 4335-4346.
62. Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W. *et al.* (2009) Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*, **459**, 108-112.
63. Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., 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.
64. Taatjes, D.J. (2010) The human Mediator complex: a versatile, genome-wide regulator of transcription. *Trends in biochemical sciences*, **35**, 315-322.
65. Malik, S. and Roeder, R.G. (2010) The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nature reviews. Genetics*, **11**, 761-772.
66. Francetic, T., Le May, M., Hamed, M., Mach, H., Meyers, D., Cole, P.A., Chen, J. and Li, Q. (2012) Regulation of Early Enhancer by Histone Acetyltransferase p300 during Stem Cell Differentiation. *Molecular biology*, **1**.
67. Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S. *et al.* (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature*, **467**, 430-435.
68. Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A. *et al.* (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature genetics*, **39**, 311-318.
69. Korzus, E., Torchia, J., Rose, D.W., Xu, L., Kurokawa, R., McInerney, E.M., Mullen, T.M., Glass, C.K. and Rosenfeld, M.G. (1998) Transcription factor-specific requirements for coactivators and their acetyltransferase functions. *Science*, **279**, 703-707.
70. Kwok, R.P., Lundblad, J.R., Chrivia, J.C., Richards, J.P., Bachinger, H.P., Brennan, R.G., Roberts, S.G., Green, M.R. and Goodman, R.H. (1994) Nuclear protein CBP is a coactivator for the transcription factor CREB. *Nature*, **370**, 223-226.
71. Ghisletti, S., Barozzi, I., Mietton, F., Polletti, S., De Santa, F., Venturini, E., Gregory, L., Lonie, L., Chew, A., Wei, C.L. *et al.* (2010) Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity*, **32**, 317-328.
72. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H. and Glass, C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell*, **38**, 576-589.
73. Ostuni, R., Piccolo, V., Barozzi, I., Polletti, S., Termanini, A., Bonifacio, S., Curina, A., Prosperini, E., Ghisletti, S. and Natoli, G. (2013) Latent enhancers activated by stimulation in differentiated cells. *Cell*, **152**, 157-171.

74. McKercher, S.R., Torbett, B.E., Anderson, K.L., Henkel, G.W., Vestal, D.J., Baribault, H., Klemsz, M., Feeney, A.J., Wu, G.E., Paige, C.J. *et al.* (1996) Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. *The EMBO journal*, **15**, 5647-5658.
75. DeKoter, R.P., Walsh, J.C. and Singh, H. (1998) PU.1 regulates both cytokine-dependent proliferation and differentiation of granulocyte/macrophage progenitors. *The EMBO journal*, **17**, 4456-4468.
76. Barozzi, I., Simonatto, M., Bonifacio, S., Yang, L., Rohs, R., Ghisletti, S. and Natoli, G. (2014) Coregulation of transcription factor binding and nucleosome occupancy through DNA features of mammalian enhancers. *Molecular cell*, **54**, 844-857.
77. Wei, G.H., Badis, G., Berger, M.F., Kivioja, T., Palin, K., Enge, M., Bonke, M., Jolma, A., Varjosalo, M., Gehrke, A.R. *et al.* (2010) Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *The EMBO journal*, **29**, 2147-2160.
78. Rohs, R., West, S.M., Sosinsky, A., Liu, P., Mann, R.S. and Honig, B. (2009) The role of DNA shape in protein-DNA recognition. *Nature*, **461**, 1248-1253.
79. Spitz, F. and Furlong, E.E. (2012) Transcription factors: from enhancer binding to developmental control. *Nature reviews. Genetics*, **13**, 613-626.
80. Martinez, G.J. and Rao, A. (2012) Immunology. Cooperative transcription factor complexes in control. *Science*, **338**, 891-892.
81. Biddie, S.C., John, S., Sabo, P.J., Thurman, R.E., Johnson, T.A., Schiltz, R.L., Miranda, T.B., Sung, M.H., Trump, S., Lightman, S.L. *et al.* (2011) Transcription factor AP1 potentiates chromatin accessibility and glucocorticoid receptor binding. *Molecular cell*, **43**, 145-155.
82. Wang, H., Lee, C.H., Qi, C., Taylor, P., Feng, J., Abbasi, S., Atsumi, T. and Morse, H.C., 3rd. (2008) IRF8 regulates B-cell lineage specification, commitment, and differentiation. *Blood*, **112**, 4028-4038.
83. Nutt, S.L., Heavey, B., Rolink, A.G. and Busslinger, M. (1999) Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature*, **401**, 556-562.
84. Kueh, H.Y., Champhekar, A., Nutt, S.L., Elowitz, M.B. and Rothenberg, E.V. (2013) Positive feedback between PU.1 and the cell cycle controls myeloid differentiation. *Science*, **341**, 670-673.
85. Holliday, R. (1990) Mechanisms for the control of gene activity during development. *Biological reviews of the Cambridge Philosophical Society*, **65**, 431-471.
86. Waddington, C.H. (1959) Canalization of development and genetic assimilation of acquired characters. *Nature*, **183**, 1654-1655.
87. Bernstein, B.E., Meissner, A. and Lander, E.S. (2007) The mammalian epigenome. *Cell*, **128**, 669-681.
88. Rivera, C.M. and Ren, B. (2013) Mapping human epigenomes. *Cell*, **155**, 39-55.
89. Turner, B.M. (2002) Cellular memory and the histone code. *Cell*, **111**, 285-291.
90. Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes & development*, **16**, 6-21.
91. Gialitakis, M., Arampatzi, P., Makatounakis, T. and Papamatheakis, J. (2010) Gamma interferon-dependent transcriptional memory via relocalization of a gene locus to PML nuclear bodies. *Molecular and cellular biology*, **30**, 2046-2056.
92. Gosselin, D. and Glass, C.K. (2014) Epigenomics of macrophages. *Immunological reviews*, **262**, 96-112.
93. Ziller, M.J., Gu, H., Muller, F., Donaghey, J., Tsai, L.T., Kohlbacher, O., De Jager, P.L., Rosen, E.D., Bennett, D.A., Bernstein, B.E. *et al.* (2013) Charting a dynamic DNA methylation landscape of the human genome. *Nature*, **500**, 477-481.
94. Li, E., Bestor, T.H. and Jaenisch, R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, **69**, 915-926.
95. Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M. *et al.* (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, **462**, 315-322.

96. Schmidl, C., Klug, M., Boeld, T.J., Andreesen, R., Hoffmann, P., Edinger, M. and Rehli, M. (2009) Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity. *Genome research*, **19**, 1165-1174.
97. Fry, C.J. and Peterson, C.L. (2001) Chromatin remodeling enzymes: who's on first? *Current biology : CB*, **11**, R185-197.
98. Bell, O., Tiwari, V.K., Thoma, N.H. and Schubeler, D. (2011) Determinants and dynamics of genome accessibility. *Nature reviews. Genetics*, **12**, 554-564.
99. Valouev, A., Johnson, S.M., Boyd, S.D., Smith, C.L., Fire, A.Z. and Sidow, A. (2011) Determinants of nucleosome organization in primary human cells. *Nature*, **474**, 516-520.
100. Mirny, L.A. (2010) Nucleosome-mediated cooperativity between transcription factors. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 22534-22539.
101. Felsenfeld, G. and Groudine, M. (2003) Controlling the double helix. *Nature*, **421**, 448-453.
102. Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F. and de Laat, W. (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. *Molecular cell*, **10**, 1453-1465.
103. Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W. *et al.* (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*, **453**, 948-951.
104. Golding, M.C., Magri, L.S., Zhang, L., Lalone, S.A., Higgins, M.J. and Mann, M.R. (2011) Depletion of Kcnq1ot1 non-coding RNA does not affect imprinting maintenance in stem cells. *Development*, **138**, 3667-3678.
105. Ferguson, L.R. (2011) RNA silencing: Mechanism, biology and responses to environmental stress. *Mutation research*, **714**, 93-94.
106. Yang, F., Yi, F., Zheng, Z., Ling, Z., Ding, J., Guo, J., Mao, W., Wang, X., Ding, X., Liang, Z. *et al.* (2012) Characterization of a carcinogenesis-associated long non-coding RNA. *RNA biology*, **9**, 110-116.
107. Lorenzen, J.M., Martino, F. and Thum, T. (2012) Epigenetic modifications in cardiovascular disease. *Basic research in cardiology*, **107**, 245.
108. Myer, V.E. and Young, R.A. (1998) RNA polymerase II holoenzymes and subcomplexes. *The Journal of biological chemistry*, **273**, 27757-27760.
109. Napolitano, G., Lania, L. and Majello, B. (2014) RNA polymerase II CTD modifications: how many tales from a single tail. *Journal of cellular physiology*, **229**, 538-544.
110. Hsin, J.P. and Manley, J.L. (2012) The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes & development*, **26**, 2119-2137.
111. Egloff, S. and Murphy, S. (2008) Cracking the RNA polymerase II CTD code. *Trends in genetics : TIG*, **24**, 280-288.
112. De Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B.K., Muller, H., Ragoussis, J., Wei, C.L. and Natoli, G. (2010) A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS biology*, **8**, e1000384.
113. Buratowski, S. (2009) Progression through the RNA polymerase II CTD cycle. *Molecular cell*, **36**, 541-546.
114. Prasanth, K.V. and Spector, D.L. (2007) Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. *Genes & development*, **21**, 11-42.
115. Mercer, T.R., Dinger, M.E. and Mattick, J.S. (2009) Long non-coding RNAs: insights into functions. *Nature reviews. Genetics*, **10**, 155-159.
116. Ponting, C.P., Oliver, P.L. and Reik, W. (2009) Evolution and functions of long noncoding RNAs. *Cell*, **136**, 629-641.
117. Guil, S., Soler, M., Portela, A., Carrere, J., Fonalleras, E., Gomez, A., Villanueva, A. and Esteller, M. (2012) Intronic RNAs mediate EZH2 regulation of epigenetic targets. *Nature structural & molecular biology*, **19**, 664-670.

118. Zhao, J., Sun, B.K., Erwin, J.A., Song, J.J. and Lee, J.T. (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science*, **322**, 750-756.
119. Deuve, J.L. and Avner, P. (2011) The coupling of X-chromosome inactivation to pluripotency. *Annual review of cell and developmental biology*, **27**, 611-629.
120. Bertani, S., Sauer, S., Bolotin, E. and Sauer, F. (2011) The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. *Molecular cell*, **43**, 1040-1046.
121. Prensner, J.R., Iyer, M.K., Balbin, O.A., Dhanasekaran, S.M., Cao, Q., Brenner, J.C., Laxman, B., Asangani, I.A., Grasso, C.S., Kominsky, H.D. *et al.* (2011) Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. *Nature biotechnology*, **29**, 742-749.
122. Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.C., Hung, T., Argani, P., Rinn, J.L. *et al.* (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*, **464**, 1071-1076.
123. Kung, J.T., Colognori, D. and Lee, J.T. (2013) Long noncoding RNAs: past, present, and future. *Genetics*, **193**, 651-669.
124. Zhao, J., Ohsumi, T.K., Kung, J.T., Ogawa, Y., Grau, D.J., Sarma, K., Song, J.J., Kingston, R.E., Borowsky, M. and Lee, J.T. (2010) Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Molecular cell*, **40**, 939-953.
125. Law, J.A. and Jacobsen, S.E. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature reviews. Genetics*, **11**, 204-220.
126. Hung, T., Wang, Y., Lin, M.F., Koegel, A.K., Kotake, Y., Grant, G.D., Horlings, H.M., Shah, N., Umbricht, C., Wang, P. *et al.* (2011) Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nature genetics*, **43**, 621-629.
127. Natoli, G. and Andrau, J.C. (2012) Noncoding transcription at enhancers: general principles and functional models. *Annual review of genetics*, **46**, 1-19.
128. Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S. *et al.* (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature*, **465**, 182-187.
129. Koch, F., Fenouil, R., Gut, M., Cauchy, P., Albert, T.K., Zacarias-Cabeza, J., Spicuglia, S., de la Chapelle, A.L., Heidemann, M., Hintermair, C. *et al.* (2011) Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nature structural & molecular biology*, **18**, 956-963.
130. Struhl, K. (2007) Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nature structural & molecular biology*, **14**, 103-105.
131. Cho, H., Orphanides, G., Sun, X., Yang, X.J., Ogryzko, V., Lees, E., Nakatani, Y. and Reinberg, D. (1998) A human RNA polymerase II complex containing factors that modify chromatin structure. *Molecular and cellular biology*, **18**, 5355-5363.
132. Wilson, C.J., Chao, D.M., Imbalzano, A.N., Schnitzler, G.R., Kingston, R.E. and Young, R.A. (1996) RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodeling. *Cell*, **84**, 235-244.
133. Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M.G., Glass, C.K. and Kurokawa, R. (2008) Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature*, **454**, 126-130.
134. Feng, J., Bi, C., Clark, B.S., Mady, R., Shah, P. and Kohtz, J.D. (2006) The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes & development*, **20**, 1470-1484.
135. Cai, X., Hagedorn, C.H. and Cullen, B.R. (2004) Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA*, **10**, 1957-1966.
136. Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H. and Kim, V.N. (2004) MicroRNA genes are transcribed by RNA polymerase II. *The EMBO journal*, **23**, 4051-4060.

137. Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N. and Shiekhattar, R. (2004) The Microprocessor complex mediates the genesis of microRNAs. *Nature*, **432**, 235-240.
138. Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S. *et al.* (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature*, **425**, 415-419.
139. Valencia-Sanchez, M.A., Liu, J., Hannon, G.J. and Parker, R. (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes & development*, **20**, 515-524.
140. Ambros, V. (2004) The functions of animal microRNAs. *Nature*, **431**, 350-355.
141. Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281-297.
142. Guo, H., Ingolia, N.T., Weissman, J.S. and Bartel, D.P. (2010) Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature*, **466**, 835-840.
143. Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N. *et al.* (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell*, **146**, 1016-1028.
144. Tian, Z., Tolic, N., Zhao, R., Moore, R.J., Hengel, S.M., Robinson, E.W., Stenoien, D.L., Wu, S., Smith, R.D. and Pasa-Tolic, L. (2012) Enhanced top-down characterization of histone post-translational modifications. *Genome biology*, **13**, R86.
145. Jenuwein, T. and Allis, C.D. (2001) Translating the histone code. *Science*, **293**, 1074-1080.
146. Strahl, B.D. and Allis, C.D. (2000) The language of covalent histone modifications. *Nature*, **403**, 41-45.
147. Marmorstein, R. (2001) Protein modules that manipulate histone tails for chromatin regulation. *Nature reviews. Molecular cell biology*, **2**, 422-432.
148. Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., 3rd, Gingeras, T.R. *et al.* (2005) Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell*, **120**, 169-181.
149. Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D. and Ren, B. (2005) A high-resolution map of active promoters in the human genome. *Nature*, **436**, 876-880.
150. Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E. *et al.* (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell*, **122**, 517-527.
151. Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., 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.
152. Voigt, P., Tee, W.W. and Reinberg, D. (2013) A double take on bivalent promoters. *Genes & development*, **27**, 1318-1338.
153. Di Croce, L. and Helin, K. (2013) Transcriptional regulation by Polycomb group proteins. *Nature structural & molecular biology*, **20**, 1147-1155.
154. Gearhart, M.D., Corcoran, C.M., Wamstad, J.A. and Bardwell, V.J. (2006) Polycomb group and SCF ubiquitin ligases are found in a novel BCOR complex that is recruited to BCL6 targets. *Molecular and cellular biology*, **26**, 6880-6889.
155. Attwooll, C., Oddi, S., Cartwright, P., Prosperini, E., Agger, K., Steensgaard, P., Wagener, C., Sardet, C., Moroni, M.C. and Helin, K. (2005) A novel repressive E2F6 complex containing the polycomb group protein, EPC1, that interacts with EZH2 in a proliferation-specific manner. *The Journal of biological chemistry*, **280**, 1199-1208.
156. Saeed, S., Quintin, J., Kerstens, H.H., Rao, N.A., Aghajani-Refah, A., Matarese, F., Cheng, S.C., Ratter, J., Berentsen, K., van der Ent, M.A. *et al.* (2014) Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. *Science*, **345**, 1251086.

157. Lavin, Y., Winter, D., Blecher-Gonen, R., David, E., Keren-Shaul, H., Merad, M., Jung, S. and Amit, I. (2014) Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. *Cell*, **159**, 1312-1326.
158. Ernst, J., Kheradpour, P., Mikkelson, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., 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.
159. Zentner, G.E., Tesar, P.J. and Scacheri, P.C. (2011) Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome research*, **21**, 1273-1283.
160. Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A. *et al.* (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 21931-21936.
161. Rada-Iglesias, A., Bajpai, R., Swigut, T., Bruggmann, S.A., Flynn, R.A. and Wysocka, J. (2011) A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*, **470**, 279-283.
162. Wang, D., Garcia-Bassets, I., Benner, C., Li, W., Su, X., Zhou, Y., Qiu, J., Liu, W., Kaikkonen, M.U., Ohgi, K.A. *et al.* (2011) Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature*, **474**, 390-394.
163. Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, **129**, 823-837.
164. May, D., Blow, M.J., Kaplan, T., McCulley, D.J., Jensen, B.C., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C. *et al.* (2012) Large-scale discovery of enhancers from human heart tissue. *Nature genetics*, **44**, 89-93.
165. Mercer, E.M., Lin, Y.C., Benner, C., Jhunjhunwala, S., Dutkowski, J., Flores, M., Sigvardsson, M., Ideker, T., Glass, C.K. and Murre, C. (2011) Multilineage priming of enhancer repertoires precedes commitment to the B and myeloid cell lineages in hematopoietic progenitors. *Immunity*, **35**, 413-425.
166. Cui, K., Zang, C., Roh, T.Y., Schones, D.E., Childs, R.W., Peng, W. and Zhao, K. (2009) Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. *Cell stem cell*, **4**, 80-93.
167. Kouzarides, T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693-705.
168. Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I. and Young, R.A. (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, **153**, 307-319.
169. Gosselin, D., Link, V.M., Romanoski, C.E., Fonseca, G.J., Eichenfield, D.Z., Spann, N.J., Stender, J.D., Chun, H.B., Garner, H., Geissmann, F. *et al.* (2014) Environment drives selection and function of enhancers controlling tissue-specific macrophage identities. *Cell*, **159**, 1327-1340.
170. Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-Andre, V., Sigova, A.A., Hoke, H.A. and Young, R.A. (2013) Super-enhancers in the control of cell identity and disease. *Cell*, **155**, 934-947.
171. Loven, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I. and Young, R.A. (2013) Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell*, **153**, 320-334.
172. Ding, J., Huang, X., Shao, N., Zhou, H., Lee, D.F., Faiola, F., Fidalgo, M., Guallar, D., Saunders, A., Shliaha, P.V. *et al.* (2015) Tex10 Coordinates Epigenetic Control of Super-Enhancer Activity in Pluripotency and Reprogramming. *Cell stem cell*, **16**, 653-668.
173. Vahedi, G., Kanno, Y., Furumoto, Y., Jiang, K., Parker, S.C., Erdos, M.R., Davis, S.R., Roychoudhuri, R., Restifo, N.P., Gadina, M. *et al.* (2015) Super-enhancers delineate disease-associated regulatory nodes in T cells. *Nature*, **520**, 558-562.

174. Giniger, E. and Ptashne, M. (1988) Cooperative DNA binding of the yeast transcriptional activator GAL4. *Proceedings of the National Academy of Sciences of the United States of America*, **85**, 382-386.
175. Griggs, D.W. and Johnston, M. (1991) Regulated expression of the GAL4 activator gene in yeast provides a sensitive genetic switch for glucose repression. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 8597-8601.
176. Meng, F.L., Du, Z., Federation, A., Hu, J., Wang, Q., Kieffer-Kwon, K.R., Meyers, R.M., Amor, C., Wasserman, C.R., Neubergh, D. *et al.* (2014) Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. *Cell*, **159**, 1538-1548.
177. Gautier, E.L., Shay, T., Miller, J., Greter, M., Jakubzick, C., Ivanov, S., Helft, J., Chow, A., Elpek, K.G., Gordonov, S. *et al.* (2012) Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. *Nature immunology*, **13**, 1118-1128.
178. Schulz, C., Gomez Perdiguero, E., Chorro, L., Szabo-Rogers, H., Cagnard, N., Kierdorf, K., Prinz, M., Wu, B., Jacobsen, S.E., Pollard, J.W. *et al.* (2012) A lineage of myeloid cells independent of Myb and hematopoietic stem cells. *Science*, **336**, 86-90.
179. Ginhoux, F., Greter, M., Leboeuf, M., Nandi, S., See, P., Gokhan, S., Mehler, M.F., Conway, S.J., Ng, L.G., Stanley, E.R. *et al.* (2010) Fate mapping analysis reveals that adult microglia derive from primitive macrophages. *Science*, **330**, 841-845.
180. Heinz, S., Romanoski, C.E., Benner, C., Allison, K.A., Kaikkonen, M.U., Orozco, L.D. and Glass, C.K. (2013) Effect of natural genetic variation on enhancer selection and function. *Nature*, **503**, 487-492.
181. Minten, C., Terry, R., Deffrasnes, C., King, N.J. and Campbell, I.L. (2012) IFN regulatory factor 8 is a key constitutive determinant of the morphological and molecular properties of microglia in the CNS. *PLoS one*, **7**, e49851.
182. Horiuchi, M., Wakayama, K., Itoh, A., Kawai, K., Pleasure, D., Ozato, K. and Itoh, T. (2012) Interferon regulatory factor 8/interferon consensus sequence binding protein is a critical transcription factor for the physiological phenotype of microglia. *Journal of neuroinflammation*, **9**, 227.
183. Butovsky, O., Jedrychowski, M.P., Moore, C.S., Cialic, R., Lanser, A.J., Gabriely, G., Koeglspenger, T., Dake, B., Wu, P.M., Doykan, C.E. *et al.* (2014) Identification of a unique TGF-beta-dependent molecular and functional signature in microglia. *Nature neuroscience*, **17**, 131-143.
184. Gautier, E.L., Ivanov, S., Williams, J.W., Huang, S.C., Marcelin, G., Fairfax, K., Wang, P.L., Francis, J.S., Leone, P., Wilson, D.B. *et al.* (2014) Gata6 regulates aspartoacylase expression in resident peritoneal macrophages and controls their survival. *The Journal of experimental medicine*, **211**, 1525-1531.
185. N, A.G., Guillen, J.A., Gallardo, G., Diaz, M., de la Rosa, J.V., Hernandez, I.H., Casanova-Acebes, M., Lopez, F., Tabraue, C., Beceiro, S. *et al.* (2013) The nuclear receptor LXRalpha controls the functional specialization of splenic macrophages. *Nature immunology*, **14**, 831-839.
186. Lam, M.T., Cho, H., Lesch, H.P., Gosselin, D., Heinz, S., Tanaka-Oishi, Y., Benner, C., Kaikkonen, M.U., Kim, A.S., Kosaka, M. *et al.* (2013) Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature*, **498**, 511-515.
187. Martinez-Nunez, R.T., Louafi, F. and Sanchez-Elsner, T. (2011) The interleukin 13 (IL-13) pathway in human macrophages is modulated by microRNA-155 via direct targeting of interleukin 13 receptor alpha1 (IL13Ralpha1). *The Journal of biological chemistry*, **286**, 1786-1794.
188. Kaikkonen, M.U., Spann, N.J., Heinz, S., Romanoski, C.E., Allison, K.A., Stender, J.D., Chun, H.B., Tough, D.F., Prinjha, R.K., Benner, C. *et al.* (2013) Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Molecular cell*, **51**, 310-325.

189. Cassol, E., Cassetta, L., Rizzi, C., Alfano, M. and Poli, G. (2009) M1 and M2a polarization of human monocyte-derived macrophages inhibits HIV-1 replication by distinct mechanisms. *J Immunol*, **182**, 6237-6246.
190. Van den Bossche, J., Neele, A.E., Hoeksema, M.A. and de Winther, M.P. (2014) Macrophage polarization: the epigenetic point of view. *Current opinion in lipidology*, **25**, 367-373.
191. De Santa, F., Narang, V., Yap, Z.H., Tusi, B.K., 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. *The EMBO journal*, **28**, 3341-3352.
192. Ishii, M., Wen, H., Corsa, C.A., Liu, T., Coelho, A.L., Allen, R.M., Carson, W.F.t., Cavassani, K.A., Li, X., Lukacs, N.W. *et al.* (2009) Epigenetic regulation of the alternatively activated macrophage phenotype. *Blood*, **114**, 3244-3254.
193. Satoh, T., Takeuchi, O., Vandenbon, A., Yasuda, K., Tanaka, Y., Kumagai, Y., Miyake, T., Matsushita, K., Okazaki, T., Saitoh, T. *et al.* (2010) The Jmjd3-Irf4 axis regulates M2 macrophage polarization and host responses against helminth infection. *Nature immunology*, **11**, 936-944.
194. Chen, S., Ma, J., Wu, F., Xiong, L.J., Ma, H., Xu, W., Lv, R., Li, X., Villen, J., Gygi, S.P. *et al.* (2012) The histone H3 Lys 27 demethylase JMJD3 regulates gene expression by impacting transcriptional elongation. *Genes & development*, **26**, 1364-1375.
195. Mullican, S.E., Gaddis, C.A., Alenghat, T., Nair, M.G., Giacomini, P.R., Everett, L.J., Feng, D., Steger, D.J., Schug, J., Artis, D. *et al.* (2011) Histone deacetylase 3 is an epigenomic brake in macrophage alternative activation. *Genes & development*, **25**, 2480-2488.
196. Chen, X., Barozzi, I., Termanini, A., Prosperini, E., Recchiuti, A., Dalli, J., Mietton, F., Matteoli, G., Hiebert, S. and Natoli, G. (2012) Requirement for the histone deacetylase Hdac3 for the inflammatory gene expression program in macrophages. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, E2865-2874.
197. Pearce, E.L. and Pearce, E.J. (2013) Metabolic pathways in immune cell activation and quiescence. *Immunity*, **38**, 633-643.
198. Tannahill, G.M., Curtis, A.M., Adamik, J., Palsson-McDermott, E.M., McGettrick, A.F., Goel, G., Frezza, C., Bernard, N.J., Kelly, B., Foley, N.H. *et al.* (2013) Succinate is an inflammatory signal that induces IL-1beta through HIF-1alpha. *Nature*, **496**, 238-242.
199. Vats, D., Mukundan, L., Odegaard, J.I., Zhang, L., Smith, K.L., Morel, C.R., Wagner, R.A., Greaves, D.R., Murray, P.J. and Chawla, A. (2006) Oxidative metabolism and PGC-1beta attenuate macrophage-mediated inflammation. *Cell metabolism*, **4**, 13-24.
200. Kaelin, W.G., Jr. and McKnight, S.L. (2013) Influence of metabolism on epigenetics and disease. *Cell*, **153**, 56-69.
201. Park, P.J. (2009) ChIP-seq: advantages and challenges of a maturing technology. *Nature reviews. Genetics*, **10**, 669-680.
202. Bernstein, B.E., Stamatoyannopoulos, J.A., Costello, J.F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M.A., Beaudet, A.L., Ecker, J.R. *et al.* (2010) The NIH Roadmap Epigenomics Mapping Consortium. *Nature biotechnology*, **28**, 1045-1048.
203. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57-74.
204. Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B.E., Bickel, P., Brown, J.B., Cayting, P. *et al.* (2012) ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome research*, **22**, 1813-1831.
205. Bock, C. (2012) Analysing and interpreting DNA methylation data. *Nature reviews. Genetics*, **13**, 705-719.
206. Laird, P.W. (2010) Principles and challenges of genomewide DNA methylation analysis. *Nature reviews. Genetics*, **11**, 191-203.

207. Down, T.A., Rakyan, V.K., Turner, D.J., Flicek, P., Li, H., Kulesha, E., Graf, S., Johnson, N., Herrero, J., Tomazou, E.M. *et al.* (2008) A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nature biotechnology*, **26**, 779-785.
208. Serre, D., Lee, B.H. and Ting, A.H. (2010) MBD-isolated Genome Sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. *Nucleic acids research*, **38**, 391-399.
209. Lister, R., O'Malley, R.C., Tonti-Filippini, J., Gregory, B.D., Berry, C.C., Millar, A.H. and Ecker, J.R. (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell*, **133**, 523-536.
210. Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L. *et al.* (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, **324**, 930-935.
211. Pastor, W.A., Aravind, L. and Rao, A. (2013) TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nature reviews. Molecular cell biology*, **14**, 341-356.
212. Huang, Y., Pastor, W.A., Shen, Y., Tahiliani, M., Liu, D.R. and Rao, A. (2010) The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PloS one*, **5**, e8888.
213. Booth, M.J., Branco, M.R., Ficz, G., Oxley, D., Krueger, F., Reik, W. and Balasubramanian, S. (2012) Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science*, **336**, 934-937.
214. Yu, M., Hon, G.C., Szulwach, K.E., Song, C.X., Zhang, L., Kim, A., Li, X., Dai, Q., Shen, Y., Park, B. *et al.* (2012) Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell*, **149**, 1368-1380.
215. Boyle, A.P., Davis, S., Shulha, H.P., Meltzer, P., Margulies, E.H., Weng, Z., Furey, T.S. and Crawford, G.E. (2008) High-resolution mapping and characterization of open chromatin across the genome. *Cell*, **132**, 311-322.
216. Giresi, P.G., Kim, J., McDaniell, R.M., Iyer, V.R. and Lieb, J.D. (2007) FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome research*, **17**, 877-885.
217. Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y. and Greenleaf, W.J. (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature methods*, **10**, 1213-1218.
218. Hon, G.C., Hawkins, R.D. and Ren, B. (2009) Predictive chromatin signatures in the mammalian genome. *Human molecular genetics*, **18**, R195-201.
219. Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002) Capturing chromosome conformation. *Science*, **295**, 1306-1311.
220. van Steensel, B. and Dekker, J. (2010) Genomics tools for unraveling chromosome architecture. *Nature biotechnology*, **28**, 1089-1095.
221. de Wit, E. and de Laat, W. (2012) A decade of 3C technologies: insights into nuclear organization. *Genes & development*, **26**, 11-24.
222. Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O. *et al.* (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, **326**, 289-293.
223. Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H. *et al.* (2009) An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature*, **462**, 58-64.
224. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S. and Ren, B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, **485**, 376-380.
225. Wang, Z., Gerstein, M. and Snyder, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics*, **10**, 57-63.

226. Wilhelm, B.T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., Penkett, C.J., Rogers, J. and Bahler, J. (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature*, **453**, 1239-1243.
227. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods*, **5**, 621-628.
228. Pan, Q., Shai, O., Lee, L.J., Frey, B.J. and Blencowe, B.J. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature genetics*, **40**, 1413-1415.
229. Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J. and Pachter, L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology*, **28**, 511-515.
230. Piskol, R., Ramaswami, G. and Li, J.B. (2013) Reliable identification of genomic variants from RNA-seq data. *American journal of human genetics*, **93**, 641-651.
231. Levin, J.Z., Yassour, M., Adiconis, X., Nusbaum, C., Thompson, D.A., Friedman, N., Gnirke, A. and Regev, A. (2010) Comprehensive comparative analysis of strand-specific RNA sequencing methods. *Nature methods*, **7**, 709-715.
232. Core, L.J., Waterfall, J.J. and Lis, J.T. (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science*, **322**, 1845-1848.
233. Ingolia, N.T., Ghaemmighami, S., Newman, J.R. and Weissman, J.S. (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*, **324**, 218-223.
234. Ingolia, N.T., Lareau, L.F. and Weissman, J.S. (2011) Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell*, **147**, 789-802.
235. Weiss, R.B. and Atkins, J.F. (2011) Molecular biology. Translation goes global. *Science*, **334**, 1509-1510.
236. Williams, A.G., Thomas, S., Wyman, S.K. and Holloway, A.K. (2014) RNA-seq Data: Challenges in and Recommendations for Experimental Design and Analysis. *Current protocols in human genetics / editorial board, Jonathan L. Haines ... [et al.]*, **83**, 11 13 11-11 13 20.
237. Adli, M. and Bernstein, B.E. (2011) Whole-genome chromatin profiling from limited numbers of cells using nano-ChIP-seq. *Nature protocols*, **6**, 1656-1668.
238. Blecher-Gonen, R., Barnett-Itzhaki, Z., Jaitin, D., Amann-Zalcenstein, D., Lara-Astiaso, D. and Amit, I. (2013) High-throughput chromatin immunoprecipitation for genome-wide mapping of in vivo protein-DNA interactions and epigenomic states. *Nature protocols*, **8**, 539-554.
239. Teytelman, L., Ozaydin, B., Zill, O., Lefrancois, P., Snyder, M., Rine, J. and Eisen, M.B. (2009) Impact of chromatin structures on DNA processing for genomic analyses. *PloS one*, **4**, e6700.
240. van Heesch, S., Mokry, M., Boskova, V., Junker, W., Mehon, R., Toonen, P., de Bruijn, E., Shull, J.D., Aitman, T.J., Cuppen, E. *et al.* (2013) Systematic biases in DNA copy number originate from isolation procedures. *Genome biology*, **14**, R33.
241. Gilfillan, G.D., Hughes, T., Sheng, Y., Hjorthaug, H.S., Straub, T., Gervin, K., Harris, J.R., Undlien, D.E. and Lyle, R. (2012) Limitations and possibilities of low cell number ChIP-seq. *BMC genomics*, **13**, 645.
242. Turner, B. (2001), *Mapping Protein/DNA Interactions by Cross-Linking*, Paris.
243. Orlando, V. (2000) Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends in biochemical sciences*, **25**, 99-104.
244. Cuddapah, S., Barski, A., Cui, K., Schones, D.E., Wang, Z., Wei, G. and Zhao, K. (2009) Native chromatin preparation and Illumina/Solexa library construction. *Cold Spring Harbor protocols*, **2009**, pdb prot5237.

245. Telford, D.J. and Stewart, B.W. (1989) Micrococcal nuclease: its specificity and use for chromatin analysis. *The International journal of biochemistry*, **21**, 127-137.
246. Chung, H.R., Dunkel, I., Heise, F., Linke, C., Krobitsch, S., Ehrenhofer-Murray, A.E., Sperling, S.R. and Vingron, M. (2010) The effect of micrococcal nuclease digestion on nucleosome positioning data. *PloS one*, **5**, e15754.
247. Garber, M., Yosef, N., Goren, A., Raychowdhury, R., Thielke, A., Guttman, M., Robinson, J., Minie, B., Chevrier, N., Itzhaki, Z. *et al.* (2012) A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Molecular cell*, **47**, 810-822.
248. Ram, O., Goren, A., Amit, I., Shores, N., Yosef, N., Ernst, J., Kellis, M., Gymrek, M., Issner, R., Coyne, M. *et al.* (2011) Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells. *Cell*, **147**, 1628-1639.
249. Bradley, R.K., Li, X.Y., Trapnell, C., Davidson, S., Pachter, L., Chu, H.C., Tonkin, L.A., Biggin, M.D. and Eisen, M.B. (2010) Binding site turnover produces pervasive quantitative changes in transcription factor binding between closely related *Drosophila* species. *PLoS biology*, **8**, e1000343.
250. Teytelman, L., Thurtle, D.M., Rine, J. and van Oudenaarden, A. (2013) Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 18602-18607.
251. Marinov, G.K., Kundaje, A., Park, P.J. and Wold, B.J. (2014) Large-scale quality analysis of published ChIP-seq data. *G3 (Bethesda)*, **4**, 209-223.
252. Pham, T.H., Benner, C., Lichtinger, M., Schwarzfischer, L., Hu, Y., Andreesen, R., Chen, W. and Rehli, M. (2012) Dynamic epigenetic enhancer signatures reveal key transcription factors associated with monocytic differentiation states. *Blood*, **119**, e161-171.
253. Beyer, M., Mallmann, M.R., Xue, J., Staratschek-Jox, A., Vorholt, D., Krebs, W., Sommer, D., Sander, J., Mertens, C., Nino-Castro, A. *et al.* (2012) High-resolution transcriptome of human macrophages. *PloS one*, **7**, e45466.
254. De Nardo, D., Labzin, L.I., Kono, H., Seki, R., Schmidt, S.V., Beyer, M., Xu, D., Zimmer, S., Lahrmann, C., Schildberg, F.A. *et al.* (2014) High-density lipoprotein mediates anti-inflammatory reprogramming of macrophages via the transcriptional regulator ATF3. *Nature immunology*, **15**, 152-160.
255. Sullivan, A.L., Benner, C., Heinz, S., Huang, W., Xie, L., Miano, J.M. and Glass, C.K. (2011) Serum response factor utilizes distinct promoter- and enhancer-based mechanisms to regulate cytoskeletal gene expression in macrophages. *Molecular and cellular biology*, **31**, 861-875.
256. Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology*, **14**, R36.
257. Pruitt, K.D., Brown, G.R., Hiatt, S.M., Thibaud-Nissen, F., Astashyn, A., Ermolaeva, O., Farrell, C.M., Hart, J., Landrum, M.J., McGarvey, K.M. *et al.* (2014) RefSeq: an update on mammalian reference sequences. *Nucleic acids research*, **42**, D756-763.
258. R Core Team, R.F.f.S.C. (2014) R: A Language and Environment for Statistical Computing.
259. Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology*, **10**, R25.
260. Trapnell, C. and Salzberg, S.L. (2009) How to map billions of short reads onto genomes. *Nature biotechnology*, **27**, 455-457.
261. Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P. *et al.* (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, **448**, 553-560.
262. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome biology*, **9**, R137.

263. Thorvaldsdottir, H., Robinson, J.T. and Mesirov, J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics*, **14**, 178-192.
264. Saito, R., Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P.L., Lotia, S., Pico, A.R., Bader, G.D. and Ideker, T. (2012) A travel guide to Cytoscape plugins. *Nature methods*, **9**, 1069-1076.
265. Maere, S., Heymans, K. and Kuiper, M. (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, **21**, 3448-3449.
266. McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M. and Bejerano, G. (2010) GREAT improves functional interpretation of cis-regulatory regions. *Nature biotechnology*, **28**, 495-501.
267. Fulton, D.L., Sundararajan, S., Badis, G., Hughes, T.R., Wasserman, W.W., Roach, J.C. and Sladek, R. (2009) TFCat: the curated catalog of mouse and human transcription factors. *Genome biology*, **10**, R29.
268. Theocharidis, A., van Dongen, S., Enright, A.J. and Freeman, T.C. (2009) Network visualization and analysis of gene expression data using BioLayout Express(3D). *Nature protocols*, **4**, 1535-1550.
269. Krebs, W., Schmidt, S.V., Goren, A., De Nardo, D., Labzin, L., Bovier, A., Ulas, T., Theis, H., Kraut, M., Latz, E. *et al.* (2014) Optimization of transcription factor binding map accuracy utilizing knockout-mouse models. *Nucleic acids research*, **42**, 13051-13060.
270. Wei, G., Abraham, B.J., Yagi, R., Jothi, R., Cui, K., Sharma, S., Narlikar, L., Northrup, D.L., Tang, Q., Paul, W.E. *et al.* (2011) Genome-wide analyses of transcription factor GATA3-mediated gene regulation in distinct T cell types. *Immunity*, **35**, 299-311.
271. Bailey, T., Krajewski, P., Ladunga, I., Lefebvre, C., Li, Q., Liu, T., Madrigal, P., Taslim, C. and Zhang, J. (2013) Practical guidelines for the comprehensive analysis of ChIP-seq data. *PLoS computational biology*, **9**, e1003326.
272. Liang, G., Wolfgang, C.D., Chen, B.P., Chen, T.H. and Hai, T. (1996) ATF3 gene. Genomic organization, promoter, and regulation. *The Journal of biological chemistry*, **271**, 1695-1701.
273. Mylona, A., Nicolas, R., Maurice, D., Sargent, M., Tuil, D., Daegelen, D., Treisman, R. and Costello, P. (2011) The essential function for serum response factor in T-cell development reflects its specific coupling to extracellular signal-regulated kinase signaling. *Molecular and cellular biology*, **31**, 267-276.
274. Xie, L. (2014) MKL1/2 and ELK4 co-regulate distinct serum response factor (SRF) transcription programs in macrophages. *BMC genomics*, **15**, 301.
275. Schmidt, S.V., Krebs, W., Ulas, T., Xue, J., Bassler, K., Gunther, P., Hardt, A.L., Schultze, H., Sander, J., Klee, K. *et al.* (2016) The transcriptional regulator network of human inflammatory macrophages is defined by open chromatin. *Cell research*.
276. Metcalf, D. and Nicola, N. (1995) *The hemopoietic colony-stimulating factors : from biology to clinical applications*. Cambridge University Press, Cambridge ; New York, NY, USA.
277. Alvarez-Errico, D., Vento-Tormo, R., Sieweke, M. and Ballestar, E. (2015) Epigenetic control of myeloid cell differentiation, identity and function. *Nature reviews. Immunology*, **15**, 7-17.
278. Krysinska, H., Hoogenkamp, M., Ingram, R., Wilson, N., Tagoh, H., Laslo, P., Singh, H. and Bonifer, C. (2007) A two-step, PU.1-dependent mechanism for developmentally regulated chromatin remodeling and transcription of the c-fms gene. *Molecular and cellular biology*, **27**, 878-887.
279. Scott, E.W., Simon, M.C., Anastasi, J. and Singh, H. (1994) Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science*, **265**, 1573-1577.
280. Lara-Astiaso, D., Weiner, A., Lorenzo-Vivas, E., Zaretzky, I., Jaitin, D.A., David, E., Keren-Shaul, H., Mildner, A., Winter, D., Jung, S. *et al.* (2014) Immunogenetics. Chromatin state dynamics during blood formation. *Science*, **345**, 943-949.

281. Mancino, A., Termanini, A., Barozzi, I., Ghisletti, S., Ostuni, R., Prosperini, E., Ozato, K. and Natoli, G. (2015) A dual cis-regulatory code links IRF8 to constitutive and inducible gene expression in macrophages. *Genes & development*, **29**, 394-408.
282. Dhawan, P. and Richmond, A. (2002) Role of CXCL1 in tumorigenesis of melanoma. *Journal of leukocyte biology*, **72**, 9-18.
283. Yordy, J.S. and Muise-Helmericks, R.C. (2000) Signal transduction and the Ets family of transcription factors. *Oncogene*, **19**, 6503-6513.
284. Davis, J.N. and Roussel, M.F. (1996) Cloning and expression of the murine Elf-1 cDNA. *Gene*, **171**, 265-269.
285. Morishita, H., Saito, F., Kayama, H., Atarashi, K., Kuwata, H., Yamamoto, M. and Takeda, K. (2009) Fra-1 negatively regulates lipopolysaccharide-mediated inflammatory responses. *International immunology*, **21**, 457-465.
286. Takada, Y., Gresh, L., Bozec, A., Ikeda, E., Kamiya, K., Watanabe, M., Kobayashi, K., Asano, K., Toyama, Y., Wagner, E.F. *et al.* (2011) Interstitial lung disease induced by gefitinib and toll-like receptor ligands is mediated by Fra-1. *Oncogene*, **30**, 3821-3832.
287. Barish, G.D., Yu, R.T., Karunasiri, M., Ocampo, C.B., Dixon, J., Benner, C., Dent, A.L., Tangirala, R.K. and Evans, R.M. (2010) Bcl-6 and NF-kappaB cistromes mediate opposing regulation of the innate immune response. *Genes & development*, **24**, 2760-2765.
288. Kaufmann, A., Salentin, R., Gemsa, D. and Sprenger, H. (2001) Increase of CCR1 and CCR5 expression and enhanced functional response to MIP-1 alpha during differentiation of human monocytes to macrophages. *Journal of leukocyte biology*, **69**, 248-252.
289. Arunachalam, B., Phan, U.T., Geuze, H.J. and Cresswell, P. (2000) Enzymatic reduction of disulfide bonds in lysosomes: characterization of a gamma-interferon-inducible lysosomal thiol reductase (GILT). *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 745-750.
290. Martinez, F.O. and Gordon, S. (2014) The M1 and M2 paradigm of macrophage activation: time for reassessment. *F1000prime reports*, **6**, 13.
291. Huang, S.C., Everts, B., Ivanova, Y., O'Sullivan, D., Nascimento, M., Smith, A.M., Beatty, W., Love-Gregory, L., Lam, W.Y., O'Neill, C.M. *et al.* (2014) Cell-intrinsic lysosomal lipolysis is essential for alternative activation of macrophages. *Nature immunology*, **15**, 846-855.
292. Frucht, D.M., Aringer, M., Galon, J., Danning, C., Brown, M., Fan, S., Centola, M., Wu, C.Y., Yamada, N., El Gabalawy, H. *et al.* (2000) Stat4 is expressed in activated peripheral blood monocytes, dendritic cells, and macrophages at sites of Th1-mediated inflammation. *J Immunol*, **164**, 4659-4664.
293. Coccia, E.M., Del Russo, N., Stellacci, E., Testa, U., Marziali, G. and Battistini, A. (1999) STAT1 activation during monocyte to macrophage maturation: role of adhesion molecules. *International immunology*, **11**, 1075-1083.
294. Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M.J. *et al.* (2015) Integrative analysis of 111 reference human epigenomes. *Nature*, **518**, 317-330.
295. Ouma, W.Z., Mejia-Guerra, M.K., Yilmaz, A., Pareja-Tobes, P., Li, W., Doseff, A.I. and Grotewold, E. (2015) Important biological information uncovered in previously unaligned reads from chromatin immunoprecipitation experiments (ChIP-Seq). *Scientific reports*, **5**, 8635.
296. Chung, D., Kuan, P.F., Li, B., Sanalkumar, R., Liang, K., Bresnick, E.H., Dewey, C. and Keles, S. (2011) Discovering transcription factor binding sites in highly repetitive regions of genomes with multi-read analysis of ChIP-Seq data. *PLoS computational biology*, **7**, e1002111.
297. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T. *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics*, **25**, 25-29.

298. Hu, Z., Gu, X., Baraoidan, K., Ibanez, V., Sharma, A., Kadkol, S., Munker, R., Ackerman, S., Nucifora, G. and Sauntharajah, Y. (2011) RUNX1 regulates corepressor interactions of PU.1. *Blood*, **117**, 6498-6508.
299. Cheng, Q., Kazemian, M., Pham, H., Blatti, C., Celniker, S.E., Wolfe, S.A., Brodsky, M.H. and Sinha, S. (2013) Computational identification of diverse mechanisms underlying transcription factor-DNA occupancy. *PLoS genetics*, **9**, e1003571.
300. Zang, C., Schones, D.E., Zeng, C., Cui, K., Zhao, K. and Peng, W. (2009) A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics*, **25**, 1952-1958.
301. Chen, Y., Negre, N., Li, Q., Mieczkowska, J.O., Slattery, M., Liu, T., Zhang, Y., Kim, T.K., He, H.H., Zieba, J. *et al.* (2012) Systematic evaluation of factors influencing ChIP-seq fidelity. *Nature methods*, **9**, 609-614.
302. Koohy, H., Down, T.A., Spivakov, M. and Hubbard, T. (2014) A comparison of peak callers used for DNase-Seq data. *PloS one*, **9**, e96303.
303. Sommer, D., Peters, A., Wirtz, T., Mai, M., Ackermann, J., Thabet, Y., Schmidt, J., Weighardt, H., Wunderlich, F.T., Degen, J. *et al.* (2014) Efficient genome engineering by targeted homologous recombination in mouse embryos using transcription activator-like effector nucleases. *Nature communications*, **5**, 3045.
304. Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L. and Jaenisch, R. (2013) One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*, **154**, 1370-1379.
305. Horvath, P. and Barrangou, R. (2010) CRISPR/Cas, the immune system of bacteria and archaea. *Science*, **327**, 167-170.
306. Wiedenheft, B., Sternberg, S.H. and Doudna, J.A. (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, **482**, 331-338.
307. Kasinathan, S., Orsi, G.A., Zentner, G.E., Ahmad, K. and Henikoff, S. (2014) High-resolution mapping of transcription factor binding sites on native chromatin. *Nature methods*, **11**, 203-209.
308. Murray, P.J., Allen, J.E., Biswas, S.K., Fisher, E.A., Gilroy, D.W., Goerdt, S., Gordon, S., Hamilton, J.A., Ivashkiv, L.B., Lawrence, T. *et al.* (2014) Macrophage activation and polarization: nomenclature and experimental guidelines. *Immunity*, **41**, 14-20.
309. Pennacchio, L.A., Ahituv, N., Moses, A.M., Prabhakar, S., Nobrega, M.A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A., Lewis, K.D. *et al.* (2006) In vivo enhancer analysis of human conserved non-coding sequences. *Nature*, **444**, 499-502.
310. Lefterova, M.I., Steger, D.J., Zhuo, D., Qatanani, M., Mullican, S.E., Tuteja, G., Manduchi, E., Grant, G.R. and Lazar, M.A. (2010) Cell-specific determinants of peroxisome proliferator-activated receptor gamma function in adipocytes and macrophages. *Molecular and cellular biology*, **30**, 2078-2089.
311. Adam, R.C., Yang, H., Rockowitz, S., Larsen, S.B., Nikolova, M., Oristian, D.S., Polak, L., Kadaja, M., Asare, A., Zheng, D. *et al.* (2015) Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. *Nature*.
312. Angel, A., Song, J., Dean, C. and Howard, M. (2011) A Polycomb-based switch underlying quantitative epigenetic memory. *Nature*, **476**, 105-108.
313. Hathaway, N.A., Bell, O., Hodges, C., Miller, E.L., Neel, D.S. and Crabtree, G.R. (2012) Dynamics and memory of heterochromatin in living cells. *Cell*, **149**, 1447-1460.
314. Greer, E.L., Beese-Sims, S.E., Brookes, E., Spadafora, R., Zhu, Y., Rothbart, S.B., Aristizabal-Corrales, D., Chen, S., Badeaux, A.I., Jin, Q. *et al.* (2014) A histone methylation network regulates transgenerational epigenetic memory in *C. elegans*. *Cell reports*, **7**, 113-126.
315. Barth, T.K. and Imhof, A. (2010) Fast signals and slow marks: the dynamics of histone modifications. *Trends in biochemical sciences*, **35**, 618-626.
316. Hwang, Y., Lin, C., Valladares, O., Malamon, J., Kuksa, P.P., Zheng, Q., Gregory, B.D. and Wang, L. (2014) HIPPIE: a high-throughput identification pipeline for promoter interacting enhancer elements. *Bioinformatics*.

317. Hegyi, H. (2015) Enhancer-promoter interaction facilitated by transiently forming G-quadruplexes. *Scientific reports*, **5**, 9165.
318. Joshi, A., Pooley, C., Freeman, T.C., Lennartsson, A., Babina, M., Schmidl, C., Geijtenbeek, T., Michoel, T., Severin, J., Itoh, M. *et al.* (2015) Transcription factor, promoter, and enhancer utilization in human myeloid cells. *Journal of leukocyte biology*.
319. Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T. *et al.* (2014) An atlas of active enhancers across human cell types and tissues. *Nature*, **507**, 455-461.
320. Schmidl, C., Renner, K., Peter, K., Eder, R., Lassmann, T., Balwiercz, P.J., Itoh, M., Nagao-Sato, S., Kawaji, H., Carninci, P. *et al.* (2014) Transcription and enhancer profiling in human monocyte subsets. *Blood*, **123**, e90-99.
321. Ginhoux, F. and Jung, S. (2014) Monocytes and macrophages: developmental pathways and tissue homeostasis. *Nature reviews. Immunology*, **14**, 392-404.
322. Kim, T.H., Li, F., Ferreira-Neira, I., Ho, L.L., Luyten, A., Nalapareddy, K., Long, H., Verzi, M. and Shivdasani, R.A. (2014) Broadly permissive intestinal chromatin underlies lateral inhibition and cell plasticity. *Nature*, **506**, 511-515.
323. Scharf, A.N., Barth, T.K. and Imhof, A. (2009) Establishment of histone modifications after chromatin assembly. *Nucleic acids research*, **37**, 5032-5040.
324. Li, B., Carey, M. and Workman, J.L. (2007) The role of chromatin during transcription. *Cell*, **128**, 707-719.
325. Berger, S.L. (2007) The complex language of chromatin regulation during transcription. *Nature*, **447**, 407-412.
326. Saunders, A., Core, L.J. and Lis, J.T. (2006) Breaking barriers to transcription elongation. *Nature reviews. Molecular cell biology*, **7**, 557-567.
327. Ivashkiv, L.B. (2013) Epigenetic regulation of macrophage polarization and function. *Trends in immunology*, **34**, 216-223.
328. Glass, C.K. and Saijo, K. (2010) Nuclear receptor transrepression pathways that regulate inflammation in macrophages and T cells. *Nature reviews. Immunology*, **10**, 365-376.
329. Stender, J.D., Pascual, G., Liu, W., Kaikkonen, M.U., Do, K., Spann, N.J., Boutros, M., Perrimon, N., Rosenfeld, M.G. and Glass, C.K. (2012) Control of proinflammatory gene programs by regulated trimethylation and demethylation of histone H4K20. *Molecular cell*, **48**, 28-38.
330. van Essen, D., Zhu, Y. and Sacconi, S. (2010) A feed-forward circuit controlling inducible NF-kappaB target gene activation by promoter histone demethylation. *Molecular cell*, **39**, 750-760.
331. Zhu, Y., van Essen, D. and Sacconi, S. (2012) Cell-type-specific control of enhancer activity by H3K9 trimethylation. *Molecular cell*, **46**, 408-423.
332. Vermeulen, M., Eberl, H.C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K.K., Olsen, J.V., Hyman, A.A., Stunnenberg, H.G. *et al.* (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell*, **142**, 967-980.
333. Zhou, W., Zhu, P., Wang, J., Pascual, G., Ohgi, K.A., Lozach, J., Glass, C.K. and Rosenfeld, M.G. (2008) Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. *Molecular cell*, **29**, 69-80.
334. Vandevyver, S., Dejager, L., Tuckermann, J. and Libert, C. (2013) New insights into the anti-inflammatory mechanisms of glucocorticoids: an emerging role for glucocorticoid-receptor-mediated transactivation. *Endocrinology*, **154**, 993-1007.
335. Nicodeme, E., Jeffrey, K.L., Schaefer, U., Beinke, S., Dewell, S., Chung, C.W., Chandwani, R., Marazzi, I., Wilson, P., Coste, H. *et al.* (2010) Suppression of inflammation by a synthetic histone mimic. *Nature*, **468**, 1119-1123.
336. Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D. and Patel, D.J. (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nature structural & molecular biology*, **14**, 1025-1040.

337. Belkina, A.C., Nikolajczyk, B.S. and Denis, G.V. (2013) BET protein function is required for inflammation: Brd2 genetic disruption and BET inhibitor JQ1 impair mouse macrophage inflammatory responses. *J Immunol*, **190**, 3670-3678.
338. Griffon, A., Barbier, Q., Dalino, J., van Helden, J., Spicuglia, S. and Ballester, B. (2015) Integrative analysis of public ChIP-seq experiments reveals a complex multi-cell regulatory landscape. *Nucleic acids research*, **43**, e27.
339. Negre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., Sealfon, R. *et al.* (2011) A cis-regulatory map of the *Drosophila* genome. *Nature*, **471**, 527-531.
340. Chapuy, B., McKeown, M.R., Lin, C.Y., Monti, S., Roemer, M.G., Qi, J., Rahl, P.B., Sun, H.H., Yeda, K.T., Doench, J.G. *et al.* (2013) Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer cell*, **24**, 777-790.
341. Ohmori, Y. and Hamilton, T.A. (2001) Requirement for STAT1 in LPS-induced gene expression in macrophages. *Journal of leukocyte biology*, **69**, 598-604.
342. Szanto, A., Balint, B.L., Nagy, Z.S., Barta, E., Dezso, B., Pap, A., Szeles, L., Poliska, S., Oros, M., Evans, R.M. *et al.* (2010) STAT6 transcription factor is a facilitator of the nuclear receptor PPAR γ -regulated gene expression in macrophages and dendritic cells. *Immunity*, **33**, 699-712.
343. Goenka, S. and Kaplan, M.H. (2011) Transcriptional regulation by STAT6. *Immunologic research*, **50**, 87-96.
344. Dahl, R., Walsh, J.C., Lancki, D., Laslo, P., Iyer, S.R., Singh, H. and Simon, M.C. (2003) Regulation of macrophage and neutrophil cell fates by the PU.1:C/EBP α ratio and granulocyte colony-stimulating factor. *Nature immunology*, **4**, 1029-1036.
345. Zhang, L. and Friedman, A.D. (2011) SHP2 tyrosine phosphatase stimulates CEBPA gene expression to mediate cytokine-dependent granulopoiesis. *Blood*, **118**, 2266-2274.
346. Stutz, A.M. and Woisetschlager, M. (1999) Functional synergism of STAT6 with either NF-kappa B or PU.1 to mediate IL-4-induced activation of IgE germline gene transcription. *J Immunol*, **163**, 4383-4391.
347. Delphin, S. and Stavnezer, J. (1995) Characterization of an interleukin 4 (IL-4) responsive region in the immunoglobulin heavy chain germline epsilon promoter: regulation by NF-IL-4, a C/EBP family member and NF-kappa B/p50. *The Journal of experimental medicine*, **181**, 181-192.
348. Shen, C.H. and Stavnezer, J. (2001) Activation of the mouse Ig germline epsilon promoter by IL-4 is dependent on AP-1 transcription factors. *J Immunol*, **166**, 411-423.
349. Aittomaki, S., Yang, J., Scott, E.W., Simon, M.C. and Silvennoinen, O. (2004) Molecular basis of Stat1 and PU.1 cooperation in cytokine-induced Fc γ receptor I promoter activation. *International immunology*, **16**, 265-274.
350. Arnold, C.D., Gerlach, D., Spies, D., Matts, J.A., Sytnikova, Y.A., Pagani, M., Lau, N.C. and Stark, A. (2014) Quantitative genome-wide enhancer activity maps for five *Drosophila* species show functional enhancer conservation and turnover during cis-regulatory evolution. *Nature genetics*, **46**, 685-692.
351. Mathelier, A., Shi, W. and Wasserman, W.W. (2015) Identification of altered cis-regulatory elements in human disease. *Trends in genetics : TIG*, **31**, 67-76.
352. Buxade, M., Lunazzi, G., Minguillon, J., Iborra, S., Berga-Bolanos, R., Del Val, M., Aramburu, J. and Lopez-Rodriguez, C. (2012) Gene expression induced by Toll-like receptors in macrophages requires the transcription factor NFAT5. *The Journal of experimental medicine*, **209**, 379-393.
353. Saccani, S., Pantano, S. and Natoli, G. (2001) Two waves of nuclear factor kappaB recruitment to target promoters. *The Journal of experimental medicine*, **193**, 1351-1359.
354. Hargreaves, D.C., Horng, T. and Medzhitov, R. (2009) Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell*, **138**, 129-145.

355. Cairo, S. and Buendia, M.A. (2012) How transient becomes stable: an epigenetic switch linking liver inflammation and tumorigenesis. *Journal of hepatology*, **57**, 910-912.
356. Hah, N., Danko, C.G., Core, L., Waterfall, J.J., Siepel, A., Lis, J.T. and Kraus, W.L. (2011) A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell*, **145**, 622-634.
357. Schneemann, M. and Schoeden, G. (2007) Macrophage biology and immunology: man is not a mouse. *Journal of leukocyte biology*, **81**, 579; discussion 580.
358. Schneemann, M. and Schoeden, G. (2002) Species differences in macrophage NO production are important. *Nature immunology*, **3**, 102.
359. Bogdan, C. (2001) Nitric oxide and the immune response. *Nature immunology*, **2**, 907-916.
360. Fang, F.C. and Nathan, C.F. (2007) Man is not a mouse: reply. *Journal of leukocyte biology*, **81**, 580.
361. Nathan, C. (2006) Role of iNOS in human host defense. *Science*, **312**, 1874-1875; author reply 1874-1875.
362. Munder, M. (2009) Arginase: an emerging key player in the mammalian immune system. *British journal of pharmacology*, **158**, 638-651.
363. Munder, M., Mollinedo, F., Calafat, J., Canchado, J., Gil-Lamainere, C., Fuentes, J.M., Luckner, C., Doschko, G., Soler, G., Eichmann, K. *et al.* (2005) Arginase I is constitutively expressed in human granulocytes and participates in fungicidal activity. *Blood*, **105**, 2549-2556.
364. Munder, M., Schneider, H., Luckner, C., Giese, T., Langhans, C.D., Fuentes, J.M., Kropf, P., Mueller, I., Kolb, A., Modolell, M. *et al.* (2006) Suppression of T-cell functions by human granulocyte arginase. *Blood*, **108**, 1627-1634.
365. Murray, P.J. and Wynn, T.A. (2011) Obstacles and opportunities for understanding macrophage polarization. *Journal of leukocyte biology*, **89**, 557-563.
366. Schroder, K., Irvine, K.M., Taylor, M.S., Bokil, N.J., Le Cao, K.A., Masterman, K.A., Labzin, L.I., Semple, C.A., Kapetanovic, R., Fairbairn, L. *et al.* (2012) Conservation and divergence in Toll-like receptor 4-regulated gene expression in primary human versus mouse macrophages. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, E944-953.
367. Martinez, F.O., Helming, L., Milde, R., Varin, A., Melgert, B.N., Draijer, C., Thomas, B., Fabbri, M., Crawshaw, A., Ho, L.P. *et al.* (2013) Genetic programs expressed in resting and IL-4 alternatively activated mouse and human macrophages: similarities and differences. *Blood*, **121**, e57-69.
368. Coers, J., Starnbach, M.N. and Howard, J.C. (2009) Modeling infectious disease in mice: co-adaptation and the role of host-specific IFN γ responses. *PLoS pathogens*, **5**, e1000333.

Zusammenfassung

Makrophagen, integrale Bestandteile des angeborenen Immunsystems sind an wichtigen Funktionen wie z.B. bei der Immunantwort, an metabolischen Prozessen, der Gewebe-Homöostase und an der Entwicklung beteiligt. Diese extrem heterogenen Zellen können ihre transkriptionelle Aktivität einer Vielzahl von stimulatorischen Signalen anpassen, um jeweils ihre spezialisierte Funktion erfüllen zu können. Das transkriptionelle Programm von aktivierten Makrophagen wird unter anderem durch transkriptionelle Regulatoren (TR) oder durch komplexe Veränderungen auf epigenetischer Ebene beeinflusst. Die DNS Sequenzieretechnologie der nächsten Generation ermöglicht die Bereitstellung von genomweiten Analysemethoden, um mit „ChIP-seq“ oder „RNA-seq“ Experimenten die DNS-Bindungsfähigkeiten von Proteinen im Genom oder die transkriptionelle Aktivität von unterschiedlich aktivierten menschlichen Makrophagen zu untersuchen. In dieser Arbeit wurden zwei wichtige Aspekte erarbeitet, zum einen eine methodische Verbesserung bei der Auswertung von ChIP-seq Daten für genomweite Transkriptionsfaktorbindungsstudien. Zum anderen wurde die epigenetische Regulation der Aktivierung humaner monozyten-abgeleiteter Makrophagen erstmalig beschrieben.

Für die Reduzierung der falsch-positiven Detektion von Proteinbindungsstellen („Peak-calling“) im Genom bei ChIP-seq Datensätzen wurde eine neue Methode entwickelt (Knockout implemented normalization method = KOIN). Bei KOIN werden ebenfalls Daten berücksichtigt, die aus Zellen stammen, die eine Nullmutante (Knockout = KO) für das jeweilige Protein darstellen. Mit Hilfe von KOIN konnte der größte Teil falsch-positiver Signale eliminiert werden, während das Signal-Stör-Verhältnis unabhängig der Datensatzqualität erhöht wurde. Außerdem wurden nicht nur kürzlich entdeckte „hyper-ChIPable“ Regionen, die mit falsch-positive Signalen assoziiert worden sind, erfolgreich ermittelt und entfernt, sondern die darauf folgende biologische Interpretation durch „KOIN“ maßgeblich verbessert. Insgesamt erzielte die „KOIN“ Methode durchweg bessere Ergebnisse als bisher verwendete „Input“ und „IgG“ Kontrollen, und wird hier deshalb als mögliche optimale „ChIP-seq“ Kontrolle für zukünftige ChIP-seq Experimente für Transkriptionsfaktoren – zumindest im murinen System - empfohlen.

Bei der Aktivierung von humanen Monozyten-abgeleiteten Makrophagen wurde ein Modell zur Charakterisierung von Histon-Modifizierungen nach Aktivierung erstellt. Mit Hilfe dieser Histon-Modifizierungen kann der Aktivierungszustand von wichtigen Regler-Elementen wie Promotoren oder Sequenzen die die Transkription eines Gens verstärken („Enhancer“) erkannt werden. Diese sind für die Etablierung und Aufrechterhaltung der Transkriptionslandschaft von entscheidender Bedeutung. In einem Modell, das teilweise mit

Faktoren aus Entzündungsreaktionen etabliert wurde, sind vier verschiedene primäre humane Makrophagen-Subpopulationen generiert worden. Zunächst wurde ein gemeinsames epigenetisches Makrophagen-Kernprogramm determiniert, was vor allem durch eine Vielzahl von gemeinsamen Promotern bestimmt wird. Außerdem konnten Subtyp-spezifische epigenetische Unterschiede an Promotoren, Super-„Enhancern“ und vor allem an einer Vielzahl von „Enhancern“ definiert werden, die auf Grund von stimulatorischen Signalen und einer darauf folgenden Spezialisierung der Makrophagen auftraten. Überraschenderweise wurden trotz transkriptioneller Unterschiede eine überraschend große Zahl an genomischen Bereichen mit aktiven und offenen Promotoren und partieller „Enhancer“ Markierung in aktivierten Makrophagen gefunden. Mit Hilfe sogenannter Koregulations-Netzwerke für transkriptionelle Regulatoren wurde gezeigt, dass für diese Klasse von Genen eine Entkoppelung von epigenetischer und transkriptioneller Kontrolle in Monozyten-abgeleiteten Makrophagen besteht. Die fehlende Notwendigkeit der epigenetischen Regulation dieser wichtigen Klasse von Genen könnte ein entscheidender Mechanismus sein, der der hohen Plastizität dieser Zellen als Antwort auf unterschiedlichste Signale aus dem Mikromilieu zugrunde liegt. Zusätzliche Mechanismen der transkriptionellen Regulierung wie „eRNA“, repressorische Proteine oder Wechselwirkungen zwischen Histon-Proteinen könnten für die Feinregulierung der Transkription in Makrophagen verantwortlich sein. Die kooperative Bindung zwischen Pionier-Transkriptions-Faktoren wie PU.1 mit sekundären Transkriptions-Faktoren wie z.B. STAT Proteinen könnte, neben der Regulation durch Histon-Protein-Modifikationen ein wichtiges Element für die Steuerung der Transkription für die offenen Positionen im Genom von aktivierten Makrophagen sein.