- Supplemental Information, Data S1

EXTENDED EXPERIMENTAL PROCEDURES

Ethics Statement

Buffy coats from healthy donors were obtained after written consent 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.

Primary cell isolation and macrophage generation

Cell isolation and generation was carried out following a previously published procedure [1]. In brief, peripheral blood mononuclear cells (PBMCs) were isolated from buffy coats by a density centrifugation step using Pancoll (PAN-Biotech) and were subsequently enriched for $CD14^+$ monocytes using CD14-specific MACS beads (Miltenyi Biotec), according to manufacturer's protocol. Isolated $CD14^+$ monocytes were cultured for 72h with granulocyte/macrophage colony stimulating factor (GM-CSF; 500U/ml) in RPMI1640 medium containing 10 % fetal calf serum (FCS) and 1% Penicillin/Streptomycin to generate unstimulated baseline macrophages (M^b). For further activation, M^b were stimulated for 72h with either IFN-γ (200U/ml), IL-4 (500U/ml) or a combination of TNF (800U/ml), PGE_2 (1 µg/ml) and Pam_3CSK4 (1 µg/ml). Distinct macrophage activation conditions were phenotypically assessed by flow cytometry using cell lineage and activation markers [2].

Flow Cytometry

Following FcR blockade with 20% FCS in PBS for 10 minutes at 4 ℃, cells were washed twice in ice-cold FACS buffer (10% FCS in PBS) and incubated with following antibodies for 30 min: CD14, CD25, CD23, CD86 (from BD or BioLegend). A LSR II cytometer (BD) was used to record data and detailed analysis was done with FlowJo software (Tree Star).

RNA isolation followed by deep sequencing and data processing

RNA isolation, quality tests and RNA sequencing experiments were performed as described earlier [1]. In brief, $5x10^6$ cultured macrophages were lysed in TRIZOL (Invitrogen). Total RNA was extracted according to the manufacturers' protocol. Only RNA with a A260/280 ratio between 1.8 and 2.1 and visible bands for 18S and 28S ribosomal RNA were converted to cDNA (Illumina TruSeq RNA Sample Preparation Kit) and used for high throughput sequencing on the HiScan SQ system (TruSeq SBS and cluster generation kit; Illumina). In brief, the mRNA purified from 5-10 µg of total RNA using poly-T oligo-attached magnetic beads was fragmented with divalent cations in fragmentation buffer (Illumina). The first strand cDNA synthesis was performed with the help of random oligonucleotides and the reverse transcriptase enzyme SuperScript II following second strand cDNA synthesis with the DNA Polymerase I and RNase H enzymes. cDNA overhangs were repaired to blunt ends with exonuclease/polymerase enzymes. An A-base was added to the 3' ends of cDNA fragments and Illumina PE adapter oligonucleotides were ligated to perform the following hybridization step. A library size selection for preferentially 200 bp cDNA fragments was conducted with 2 % (w/v) agarose gels. DNA fragments within the excised gel-fraction were purified using the QIAquick gel extraction kit (Qiagen). After the specific enrichment of fragments with ligated adapter molecules using Illumina PCR primer PE1.0 and PE2.0 in a 15 cycle PCR reaction, PCR products were purified using QIAquick PCR purification kit. Next, the assessments for quality and quantity of purified libraries were carried out on a Bioanalzyer 2100 (Agilent) with the Agilent high sensitivity DNA assay kit. After cluster generation of single strand DNA library pieces immobilized on a glass flow cell (Illumina), a paired-end 100 bp sequencing run for 208 cycles was performed.

RNA-seq libraries were sequenced and the data were demultiplexed using CASAVA v1.8 software (Illumina). Alignments for human macrophage samples were performed with TopHat2 (v2.0.3) [3] against the hg18 reference genome (RefSeq database release 64 [4]). Data are available via GEO (accession numbers GSE36952 and GSE66593). RNA-seq data for human tissues (GSE16256, [5]) were aligned against the hg19, RNA-seq data for murine tissue macrophages (GSE63341, [6]) were aligned against the mm10 reference genome. Gene and transcript information was extracted with Partek

Genomics Suite (v6.6). Annotated data were normalized using statistical software R (v3.0.2) [7] using the DESeq2 package [8]. Normalized read counts lower than 1 were set to the floor value 1. Genes were defined as expressed with normalized RNA-seq group mean values ≥ 10.

Chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) and data processing

Generation of H3K4me3 histone modification (HM) datasets were previously described and published on GEO under accession number GSE47188 [2]. ChIP-seq datasets for HMs (H3K4me1, H3K4me2, H3K27Ac, H3K27me3) and PU.1 were generated following a previously published cross-linking chromatin immunoprecipitation (X-ChIP) protocol [9]. Briefly, in case of HM ChIP chromatin from $0.5x10^6$ macrophages was cross-linked for ChIP reactions with 1 % formaldehyde for 10 minutes. For the PU.1 ChIP $6x10^6$ macrophages were cross-linked as describe before. Nuclei were lysed and chromatin was sheared into approximately 200 bp pieces with the Covaris S220 ultrasound system. For ChIP antibody reactions polyclonal rabbit antibodies were used (H3K4me1 (Abcam), H3K27Ac (Abcam), H3K27me3 (Millipore), PU.1 (SantaCruz, sc-352)). Additionally, polyclonal rabbit IgG antibody (Millipore) was used as control for validation of successful histone and PU.1 ChIP experiments. Washing steps and DNA purification as well as size selection were performed with AMPure XP SPRI beads (Beckman Coulter). ChIP-seq DNA concentration was measured with the Bioanalyzer 2100 or a Tapestation system (Agilent). Library construction was performed with 0.5 - 1 ng ChIP DNA. 5' and 3' ends of DNA were repaired to blunt ends and an A-base was added to the 3' ends. Afterwards, sample volumes were evaporated to approximately 4 µl to increase efficacy of the following ligation step using less total amount of DNA ligase enzyme in comparison to the standard protocol. For the ligation of Illumina sequencing compatible NEXTflex adapter oligonucleotides (Bioo Scientific) concentrations of 0.6 µM were added to the reaction mix and incubated for 15 min at 25 ℃. With a SPRI bead cleanup (Beckman Coulter) free adapter oligonucleotides were removed and the library was PCR amplified (14 cycles). Final DNA purification with additional size selection between 100 and 500 bp was performed with SPRI beads. ChIP-seq DNA library concentration, molarity, and average base pair length was assessed by Bioanalyzer or Tapestation analysis (Agilent Technologies) and the KAPA qPCR System (Kapabiosystems) on a Roche LightCycler 480 II system. ChIP-seq

libraries were sequenced on HiScanSQ and HiSeq 1000 sequencer (Illumina). Single read runs for 57 cycles were performed with the TruSeq SBS reagents kit (Illumina). Three biological replicates were performed for ChIP-seq experiments.

Fastq files were aligned with Bowtie [10] to the human reference genome hg18 (human macrophages, GSE66594), hg19 (human tissue, GSE16256 [5]) or mm9 (murine tissue macrophages, GSE63339 [6]) using best match parameters (bowtie -t -q -e 70 -l 28 -n 2 --best --maxbts 125 -S). Annotated SAM files were converted to tag directories using HOMERs´ makeTagDirectory module (makeTagDirectory /HistonDestinationFolder/ SAM-file.sam -genome). These directories are further used for peak calling (findPeaks /HistonDestinationFolder/ -i /Input_HistonDestinationFolder/ -style histone > peak_file.txt) or converted to the bedGraph (pos2bed.pl peak_file.txt > bed_file.bed) format with HOMER (makeUCSCfile module) normalized to 10^6 total tag counts. Visualization of normalized tag counts was performed with the Integrative Genomics Viewer (IGV) [11]. Statistics on tag counts and number of peaks for ChIP-seq data of the human activated macrophages are available in **Table S6A**. Peaks were identified with HOMER v.4.3 using the findPeaks module and the respective style option [12]. The "histone" style option (peak sizes of 1000 bp with minimal peak distances of 2500 bp) was used to optimize the analysis. Input samples were utilized as control files during peak detection for significantly called histone marks. For the identification of PU.1 binding sites, peaks were identified using the style "factor" option with the IgG control files during peak correction. Peaks were annotated using the same reference genomes as before.

Quantitative real-time PCR

Quantitative real-time PCR (qPCR) was performed using the Maxima SYBR Green/Fluorescein qPCR Master Mix (Fermentas) following manufacturer's protocol on a LightCycler 480 II system (Roche). The relative enrichment of positive but not negative control target sites in histone and PU.1 ChIP-DNA samples relative to IgG controls were calculated using the $2^{-\Delta\Delta CT}$ method. Primer sequences were designed using Beacon Designer 7 software. Primers for a genomic region in the house keeping gene GAPDH served as positive control for H3K4me3 ChIP and negative control for H3K27me3 ChIP: "tactagcggttttacgggcg" and "tcgaacaggaggagcagagagcga". Primers for a genomic position in the

BMP2 gene served as positive control for H3K27me3 ChIP and negative control for H3K4me3 ChIP: “ctcagcactccgcatttg” and “ctcccatccaacgcttag”. Primers for the positive control of H3K4me1 ChIP were “gccttgctcttgctactc” and “gcctgcctggtgataatc” (GAPDH), for H3K27Ac ChIPs “gcctgcctggtgataatc“ and “gtgatcggtgctggttcc“ (GAPDH), and primer in an intergenic region on chromosome 5 served as negative control for H3K4me1, H3K27Ac and PU.1 ChIP: “gaacaactggatgggacaaac” and “acacgaatggaaccttatatctg“. As positive control for the PU.1 ChIP a primer pair was designed against a genomic region on chromosome 2: “tccctatcccgcagatgg” and “gaggactgaggaagtggac”.

Identification and classification of promoter and enhancer sites

H3K4me3 positive promoters were further sorted into gene groups with accessible promoters (Pa) with H3K4me3 and H3K27Ac HM or poised promoters (Pp) lacking the H3K27Ac HM but encompassing the repressive H3K27me3 HM mark. Macrophage activation-specific sites matched at least one of two criteria: (1) H3K4me3 HM signals were exclusively present in only one subtype independently of signal strength, (2) HM signals were at least 2-fold higher in one macrophage type compared to the other macrophages.

Identification of putative enhancer sites with H3K4me1 HM required the exclusion of promoter sites, covering genomic regions ± 2500 bp up- and down-stream of TSS, which was enabled by HOMER. To exclude false positive enhancer sites from further analysis, positions with simultaneous H3K4me1 and H3K4me3 signal were excluded. Enhancers were separated into three groups of strong enhancer (Es) with both H3K4me1 and H3K27Ac HM, weak enhancer (Ew) with only H3K4me1 signals and poised enhancer (Ep) with simultaneous HM for H3K4me1 and H3K27me3. Macrophage specific enhancer sites were enriched at least two times more for H3K4me1 tag counts in one macrophage activation condition compared to the others or showed a significant H3K4me1 HM peak only in one subtype but not in the others.

Common accessible promoters were defined as H3K4me3 and H3K27Ac positive sites without H3K27me3 HM signals in all macrophages. At common poised promoters additionally H3K27me3 signals were present. Common strong enhancers enriched for H3K4me1 and H3K27Ac HMs lacked

the H3K27me3, whereas the poised common enhancers additionally showed H3K27me3 HM signals in all differentially activated macrophages.

Correlation of promoter and enhancer classes to expression

Biological replicates of RNA-seq data derived from macrophage activation conditions were averaged with R [7]. Statistical significance between gene expression levels was assessed by the Wilcoxon rank-sum test. Annotation was performed with HOMER using the annotatePeaks.pl (annotatePeaks.pl peak_file.txt genome > peaks_file_annotated.txt) module to connect positional peak information to gene names. For each promoter or enhancer class nearest genomic TSS was defined and corresponding gene name was linked to expression values. Enhancer sites with distances greater than 100 kb to the next TSS were excluded from the analysis.

Network of top ten percent interconnected macrophage core genes

The previously described network of highly interconnected or hub genes based on 29 macrophage activation states (299 transcriptomes) was generated by ARACNe [2, 13] and visualized with Cytoscape (v3.2.0).

***De novo* motif enrichment**

Discovery of *de novo* motifs was performed by HOMER findMotifsGenome.pl function (findMotifsGenome.pl peak_file.bed genome /DestinationFolder/ -size 1000 -nomotif -p 12). HM enriched sites in a region of 500 bp up- and downstream of nucleosome free regions within called peak positions were used to isolate sequences as input for motif analysis. Nucleosome free regions maintaining an euchromatic DNA state were identified with HOMER during peak calling (findPeaks –nfr option) to detect sites with high TF binding probability. HOMER identified enriched motifs according their corresponding positional weight matrices (PWM). The names and binomial *p*-values were calculated for significantly enriched motifs.

VisuTranscript

The principles of our WebPortal Visu Transcript will be described in more detail elsewhere (H. Schultze, unpublished results). Briefly, VisuTranscript is a HTML5 based, modular web application with responsive design and a cloud based backend. It makes use of a content management system, standard libraries like Javascript and jQuery for asynchronous content retrieval and a custom Python App on Googles highly scalable Plattform as-a-Service (PaaS) - App Engine – using NoSQL schemaless datastores and its scalability as the backend. Visualization of Boxplots make use of Google Charts API (https://developers.google.com/chart/), Heatmaps are visualized with jHeatmap [14] and networks are visualized with Cytoscape.js (http://js.cytoscape.org/)

ABREVIATIONS

Bp	base pairs
ChIP	chromatin immuno precipitation
comEp	common poised enhancer
comEs	common strong enhancer
comEw	common weak enhancer
comPa	common accessible promoter
comPp	common poised promoter
Ep	poised enhancer
eRNAs	enhancer-derived RNAs
Es	strong enhancer
Ew	weak enhancer
FC	fold change
FCS	fetal calf serum
GO	gene ontology
GOEA	gene ontology enrichment analysis
GSEA	gene set enrichment analysis
HM	histone modification
LI	large intestine (colon)
ln	logarithm naturalis
MFI	mean fluoresence intensity
Pa	accessible promoter
PBMCs	peripheral blood mononuclear cells
Pp	poised promoter
PWM	positional weight matrices
qPCR	quantitative real-time PCR
RNBPs	RNA-binding proteins
SE	super enhancer
SI	small intestine (ileus)
TF	transcription factor
TLR	Toll-like receptor
TR	transcriptional and epigenetic regulator
TSS	transcription start site

SUPPLEMENTAL REFERENCES

1 Beyer M, Mallmann MR, Xue J *et al.* High-resolution transcriptome of human macrophages. *PLoS One* 2012; **7**:e45466.

2 Xue J, Schmidt SV, Sander J *et al.* Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. *Immunity* 2014; **40**:274-288.

3 Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 2013; **14**:R36.

4 Pruitt KD, Brown GR, Hiatt SM *et al.* RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res* 2014; **42**:D756-763.

5 Roadmap Epigenomics C, Kundaje A, Meuleman W *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* 2015; **518**:317-330.

6 Lavin Y, Winter D, Blecher-Gonen R *et al.* Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. *Cell* 2014; **159**:1312-1326.

7 R Core Team RFfSC. R: A Language and Environment for Statistical Computing. 2014.

8 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; **15**:550.

9 Blecher-Gonen R, Barnett-Itzhaki Z, Jaitin D, Amann-Zalcenstein D, Lara-Astiaso D, Amit I. High-throughput chromatin immunoprecipitation for genome-wide mapping of in vivo protein-DNA interactions and epigenomic states. *Nat Protoc* 2013; **8**:539-554.

10 Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009; **10**:R25.

11 Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* 2013; **14**:178-192.

12 Heinz S, Benner C, Spann N *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010; **38**:576-589.

13 Margolin AA, Wang K, Lim WK, Kustagi M, Nemenman I, Califano A. Reverse engineering cellular networks. *Nat Protoc* 2006; **1**:662-671.

14 Deu-Pons J, Schroeder MP, Lopez-Bigas N. jHeatmap: an interactive heatmap viewer for the web. *Bioinformatics* 2014; **30**:1757-1758.