Feature Review

CellPress

Exploiting genomics and natural genetic variation to decode macrophage enhancers

Casey E. Romanoski[1], Verena M. Link[1,2], Sven Heinz[3], and Christopher K. Glass[1]

[1] Department of Cellular and Molecular Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0651, USA
[2] Faculty of Biology, Department II, Ludwig-Maximilians Universität München, Planegg-Martinsried 85152, Germany
[3] Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA

The mammalian genome contains on the order of a million enhancer-like regions that are required to establish the identities and functions of specific cell types. Here, we review recent studies in immune cells that have provided insight into the mechanisms that selectively activate certain enhancers in response to cell lineage and environmental signals. We describe a working model wherein distinct classes of transcription factors define the repertoire of active enhancers in macrophages through collaborative and hierarchical interactions, and discuss important challenges to this model, specifically providing examples from T cells. We conclude by discussing the use of natural genetic variation as a powerful approach for decoding transcription factor combinations that play dominant roles in establishing the enhancer landscapes, and the potential that these insights have for advancing our understanding of the molecular causes of human disease.

Exploiting macrophages to understand enhancer biology and enhancer biology to understand macrophages

Macrophages (see Glossary) are phagocytic cells of the innate immune system that reside in all tissues of the body and play key roles in responding to infection and injury through signaling downstream of pattern recognition receptors [1–3]. In addition to these general roles that operate throughout the body, each tissue-resident population of macrophages performs specific effector functions that contribute to the homeostasis of that tissue [2,4]. Some of the diverse roles that macrophages have *in vivo* that are unique to their tissue environments include neuronal synaptic pruning by microglia in the brain [5], bone resorption and remodeling by osteoclasts [6], control of insulin sensitivity and adaptive thermogenesis in adipose tissue [7,8], and surfactant recycling by lung alveolar macrophages [9]. Although the diverse functions of macrophages are normally adaptive, they can be co-opted to drive tissue pathology, particularly in the setting of chronic inflammatory diseases and cancer. For example, functions of macrophages that are important for pathogen recognition and initiation of inflammation play key roles the development and clinical complications of atherosclerosis [10,11]. Conversely, functions of macrophages that are important for wound repair contribute to tumor growth and metastasis [12]. Understanding the mechanisms by which various macrophage populations achieve their tissue-specific functions and determining whether these functions can be modulated for therapeutic purposes remain largely unmet goals.

Distinct macrophage phenotypic polarization states have been characterized *in vitro* by studying responses to various ligands that result in alternative gene expression programs [13]. Two extensively characterized *in vitro* polarization programs are broadly categorized as classically activated/proinflammatory M1 macrophages or alternately activated/anti-inflammatory M2 macrophages [14,15]. Treatment of macrophages with lipopolysaccharide (LPS), a component of Gram-negative bacteria, drives M1 polarization through Toll-like receptor (TLR)4-dependent activation of members of the nuclear factor (NF)-κB, activator protein 1 (AP-1) and interferon (IFN) regulated families of transcription factors (Figure 1) [16,17]. These factors induce the expression of hundreds of genes, many of which play key roles in innate immunity, inflammation, and initiation of adaptive immune responses. In contrast, treatment of macrophages with interleukin 4 (IL-4) drives M2 polarization through activation of signal transducer and activator of transcription 6 (STAT6), which induces a program of gene expression linked to immunity directed against parasitic infection (Figure 1) [14,18].

Recent studies indicate that tissue macrophages exhibit distinct gene expression programs that underlie their tissue-specific functions [19–21]. Furthermore, tissue environment has been shown to be a significant determinant of the gene expression patterns and the underlying transcriptional regulatory elements that are characteristic of a particular macrophage subtype [20,21]. The specific signals that dictate tissue-specific programs of macrophage gene expression are for the most part unknown. Some recently identified examples include transforming growth factor (TGFβ), which is essential for maintenance of microglia phenotypes [22]

Corresponding author: Glass, C.K. (ckg@ucsd.edu).

1471-4906/
 http://dx.doi.org/10.1016/j.it.2015.07.006

Glossary

C/EBP: a family of basic-leucine zipper (bZIP) transcription factors that bind DNA and form homo-and heterodimer interactions. C/EBPα and C/EBPβ are LDTFs in macrophages.
ChIP-Seq: chromatin immunoprecipitation followed by high-throughput sequencing. This assay identifies the genomic location and frequency with which a particular protein or histone modification associates with DNA.
Chromatin: DNA that is wrapped around nucleosomes. Chromatin compaction is dynamic with spatiotemporal patterns dependent on the cell cycle, developmental state, and chromosomal location. Chromatin provides a regulatory barrier between DNA and DNA-interacting proteins.
***cis*-eQTL:** An eQTL where the SNP and the gene locus for the associated transcript are close in linear genomic space (usually <1 megabase). *cis*-eQTLs typically quantify effects of genetic variation in the coding sequence, promoter or enhancer for the given gene.
***De novo* enhancer/latent enhancer:** an enhancer that transitions from a closed chromatin state to an open and active state by interactions involving SDTFs and LDTFs.
DNase-Seq: high-throughput DNA sequencing of accessible, often regulatory, regions of chromatin that result from chromatin digestion with DNase I.
E2A: encoded by the TCF3 locus, binds to E-box sequences in DNA and forms homo- and heterodimers with other transcription factors. E2A is a critical LDTF for B cell development.
EBF: a transcription factor expressed exclusively in the B cell lineage and directs B cell fate.
Enhancer: a region of DNA that can amplify RNA Pol II transcription at associated promoters. Enhancers are largely cell-type-specific and are bound by sequence-specific transcription factors.
Epigenetic marks: includes methylation of DNA as well as modifications such as acetylation, methylation, phosphorylation, and ubiquitinylation of amino acids on histone tails. Certain patterns of these epigenetic marks provide information about the function of the associated DNA.
eQTL: expression quantitative trait locus, which results when the abundance of a transcript associates to the genotypes at a given genetic variant (usually at SNPs). eQTL studies require measuring transcripts across many individuals and can be classified as *cis*- or *trans*- (also in Glossary).
GATA6: a member of the GATA family of zinc finger transcription factors that is induced by retinoic acid signaling in LPMs.
GWAS: genome wide association study. In most common form, uses specific SNPs (alleles) to link genomic loci to disease risk using cohorts of individuals with disease and healthy controls.
H3K27ac: acetylation of the lysine at position 27 of the histone tail of histone H3. H3K27ac marks active enhancers and promoters.
H3K4me1: monomethylation of the lysine at position 4 of the histone tail of histone H3. H3K4me1 marks primed and active enhancers.
H3K4me2: dimethylation of the lysine at position 4 of the histone tail of histone H3. H3K4me2 marks primed and active enhancers and promoters.
H3K4me3: trimethylation of the lysine at position 4 of the histone tail of histone H3. H3K4me3 marks promoters.
Histone tails: peptides that extend from histones that can be modified with epigenetic marks.
LDTF: lineage-determining transcription factor: a transcription factor required for the development of a specific cell type. LDTFs typically have the ability to select enhancers in concert with other LDTFs or collaborative factors. Examples include PU.1, C/EBPs in macrophages; E2A in B cells.
LPM: large peritoneal macrophage – a macrophage population resident in the peritoneal cavity that is dependent on retinoic acid.
LXRs: SDTF nuclear receptors that translocate to the nucleus and bind DNA in response to endogenous oxysterols to regulate cholesterol efflux and biosynthetic genes.
Macrophage: innate immune phagocytic cells of myeloid origin that reside in every organ of the body and perform diverse functions in health and disease.
MG: microglia – major resident macrophage population of the brain that is dependent on TGFβ.
NF-κB: SDTF complex that binds DNA upon TLR signaling to activate inflammatory genes in macrophages. During B and T cell differentiation, NF-κB has an important function more in line with an LDTF role.
Nucleosome: unit of chromatin that is composed of two copies each of histone proteins H2A, H2B, H3, and H4; 147 base pairs of DNA wrap around one nucleosome.
Promoter: region of DNA that contains binding sequences necessary to assemble the minimal transcriptional machinery and ultimately load RNA Pol II at gene start sites.
PU.1: encoded by the SPI1 human locus (Sfpi1 in *Mus musculus*), is a member of the ETS family of sequence-specific transcription factors that is a LDTF critical for macrophage and B cell differentiation.
RNA-Seq: high-throughput sequencing of RNA that is used to measure gene expression genome-wide.
SDTF: signal-dependent transcription factor: transcription factor that becomes active in response to an internal or external signal. Examples: NF-κB and LXRs.
SNP: single nucleotide polymorphism.
Super-enhancer: clusters of enhancers that are densely occupied by master regulators and the Mediator co-regulator complex, which frequently occur at genes that define the identity of a given cell. Super-enhancers can alternatively be defined by tracking epigenetic marks indicative of enhancer activity, such as H3K27ac.
***trans*-eQTL:** An eQTL where the SNP and the gene locus for the associated transcript are far in linear genomic space (usually >1 megabase). *trans*-eQTLs typically quantify effects of the genetic variant on the associated gene through an intermediate product, such as by altered expression of a transcription factor or signaling molecule that perpetuate expression differences on the target gene.

(Figure 1); retinoic acid, which is required for development/maintenance of large peritoneal macrophages [23] (Figure 1); IL-4, which is required for maintenance of homeostatic beige adipose tissue macrophages (Figure 1) [8]; and receptor activator of nuclear factor kappa-B ligand (RANKL), which is required for the development of osteoclasts [24]. Importantly, while these molecules are established to be important for the phenotypic characteristics of particular macrophage subsets *in vivo*, they represent only a part of the total spectrum of signals sensed by the macrophage within each tissue environment (Figure 1). Furthermore, as discussed below, signal-dependent transcription factors such as NF-κB primarily regulate gene expression by acting on pre-existing enhancers, which have recently been shown to differ among tissue macrophage subsets [20,21]. The implication of these findings is that the quantitative and qualitative responses of different tissue macrophages to the same signal, such as LPS, are likely to vary in a tissue-specific manner. Therefore, while *in vitro* studies of M1 and M2 macrophage activation provide powerful models to investigate mechanisms of signal-dependent gene expression, studies of the enhancer and promoter landscapes of macrophages *in vivo* provide important insights into how complex environmental signals regulate their development and function in distinct tissues.

In this review, we briefly introduce the current state of enhancer biology along with advances in the field of macrophage gene regulation. We begin with the characteristics of enhancers and underscore their dynamic behavior in cell lineage specification and environmental signaling contexts. We describe a working model in which distinct classes of sequence-specific transcription factors, referred to as lineage-determining and signal-dependent transcription factors (LDTFs and SDTFs), define the repertoire of active enhancers in macrophages through collaborative and hierarchical interactions. We note some important challenges to this model, specifically providing examples from T cells. Next, we highlight how natural genetic variation can be leveraged as a powerful tool to identify sets of collaborating transcription factors that establish enhancers in different cells. To this end, studies utilizing genetic variation in tissue-resident subsets of macrophages are discussed to exemplify the importance of tissue environment on enhancer selection. We conclude with a discussion of how these findings, which combine genetic variation and enhancer function, are highly informative for interrogating the molecular causes of human disease.