THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PhD)

Bence Dániel

Investigation of the function of Retionid X Receptor in mouse bone marrow-derived macrophages using novel genomic approaches

Supervisor: Prof. Dr. László Nagy

UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF MOLECULAR CELL AND IMMUNE BIOLOGY

DEBRECEN, 2014

Table of contents

Abbreviations

RNAPII	RNA Polymerase II
TFBS	Transcription Factor Binding Site
ChIP	Chromatin Immunoprecitation
NGS	Next-Generation Sequencing
STAT	Signal Transducer and Activator of Transcription
3C	Chromosome Conformation Capture
ChIA-PET	Chromatin Interaction Analysis by Paired-End Tag sequencing
GRO-seq	Global Run-On sequencing
TSS	Transcription Start Site
FISH	Fluorescent In Situ Hybridization
ENCODE	Encyclopedia of DNA Elements
DHS	DNase Hypersensitive Sites
CTCF	CCCTC-binding factor
eRNA	Enhancer RNA
LDTF	Lineage Determining Transcription Factor
LBD	Ligand-Binding Domain
DBD	DNA-Binding Domain
LBP	Ligand-Binding Pocket
PPAR	Peroxisome Proliferatior-Activated Receptor
RXR	Retinoid X Receptor

Introduction

Gene expression regulation

In 1956, a new field was born in molecular biology called gene expression regulation with the famous PaJaMa (designed and carried out by Art **Pa**rdee, Francois **Ja**cob and Jacque **M**onod) experiments [1]. These investigations shed light on the existence of a trans-acting repressor (lacI) which is able to inhibit the expression of a nearby gene called lacZ and supported the first evidence that trans-acting factors might possess the ability to regulate gene expression in bacteria. Importantly, these discoveries not only impacted the field of gene regulation but also dominated ideas about gene expression and the function of mRNA. Based on their experiments the most fundamental fact is that genes are controlled at the level of transcription by other genes encoding the regulators. These regulators act through specific DNA sequences near the genes they control and according to Jacob and Monod these are always having repressor functions.

With the help of these experiments they constructed the well-known Operon model, which describes how bacterial genes are often regulated [2]. What is an operon? A tightly regulated gene cluster under the control of one promoter and a regulator (Figure 1). In general this scheme applies to bacteria but largely inapplicable in eukaryotes, where the regulators can bind separate regions in the intergenic or even intragenic part of each gene they control. In addition their model explains not only how bacterial cells respond to environmental stimuli, but also in principle covers, how multi-cellular organisms develop through complex regulatory mechanisms.

The ideas of Jacob and Monod are in the center of our understanding not only of development, but also of how evolution remodeled the regulatory element landscape of animals to reach such diversity and to evolve various species. Francois Jacob in a Science paper entitled “Evolution and Tinkering” published in 1977 claimed that evolution is the result of changes in gene expression pattern, using the available repertoire of proteins encoded in the genome rather than designing new [3]. As he noted: “Biochemical changes do not seem, therefore, to be a main driving force in the diversification of living organisms. The really creative part of biochemistry must have occurred very early.” Instead he asserted that: “It seems likely that divergence and

specialization of mammals, for instance, resulted from mutations altering regulatory circuits rather than chemical structures. Small changes modifying the distribution in time and space of the same structures are sufficient to affect deeply the form, the functioning, and the behavior of the final product."

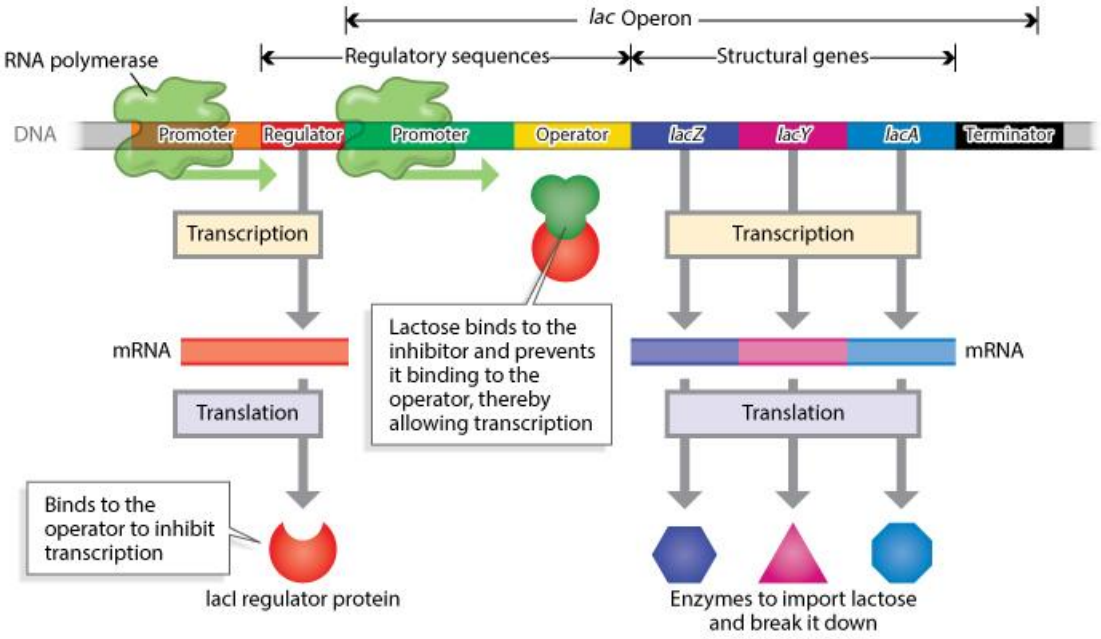

Figure 1. The lac operon in E. coli. *The lac operon contains three genes that encode enzymes involved in lactose metabolism. These are indicated as lacZ, lacY and lacA. These are the so-called structural genes which are encoding the beta-galactosidase, permease and transacetylase enzymes responsible for transporting and breaking down lactose to a usable food source. Upstream from the structural genes one promoter region is responsible for the recognition of the transcription machinery of the RNA polymerase complex. Other regulatory sequences are also crucial parts of this unit and responsible for the proper functioning of the operon. The operator between the promoter and the structural genes is able to shut down the entire operon by recruiting the lacI repressor protein in the absence of lactose. The terminator sequence serves as a stop signal to the transcription machinery. 2013 Nature Education Adapted from Pierce, Benjamin. Genetics: A Conceptual Approach, 2nd ed.*

Prokaryotic gene expression

In prokaryotes as shown by Jacob and Monod the genes are usually fall into the operons in which they are under the control of a regulator protein having repressor function on the members of the gene unit and the promoter region capable of binding the RNAPII (RNA polymerase II) and the repressor. In the absence of the activator molecule, the repressor constitutively expressed and binds to a special part of the promoter region called operator, thus preventing the initiation of transcription. In the presence of the activator the repressor cannot bind the operator region of the operon, thus transcription initiation takes place and upon elongation, a long so called polycistronic RNA is produced which made up by the members of the operon. These RNAs are then serve as a template for translation and encodes a set of genes usually involved in similar biological functions (Figure 1). According to these it seems that prokaryotic gene expression regulation is mainly maintained through regulatory elements located in the close proximity (so-called proximal regulatory elements) of the regulated genes [2].

Eukaryotic gene expression

The mammalian genome carries the potential to create more than 200 cell types from the very same genetic material. What kind of driving forces are responsible for this enormous diversity? Each cell type has its own characteristic pattern of gene expression which is the result of the collaborative action of various regulatory elements/factors either as a cis-acting or trans-acting factor on the DNA. Cis-acting elements (enhancers, silencers, insulators, locus control regions) are encoded in the DNA and possess the ability to bind the so-called trans-acting factors (transcription factors). Importantly the location of the cis-acting element repertoire relative to the regulated gene set can also be cis or trans depending on their distance or chromosomal location. Regulatory elements that are located in cis covers smaller distances, while those located in trans might be embedded into a different chromosome. This is in stark contrast with the prokaryotic gene expression regulation, where genes are commonly coordinated over proximal regulatory regions. Another obvious difference is the appearance of

monocistronic RNA which made up only by one gene and containing the intronic sequences which than later excluded by the process known as splicing (Figure 2).

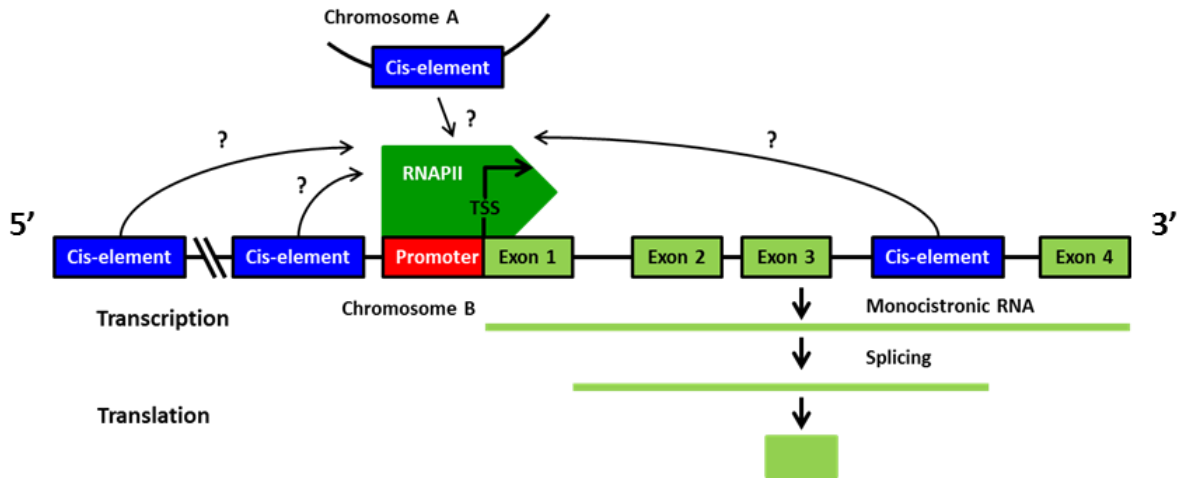

Figure 2. *Eukaryotic genes can be regulated by multiple cis-acting elements located in the non-coding part of the genome. The inner character of these elements is that they can be located far away from their regulated genes. It is also conceivable that these are part of another chromosome, thus it is challenging to pair them with their regulated genes. RNA synthesized from eukaryotic genes is monocistronic and undergoes the process called splicing, in which the intronic regions are excluded from the nascent transcript before translation [66].*

The difficulties of investigating the regulatory element toolset of mammalian genomes

The fact that mammalian genes produce monocistronic RNA molecules suggests that each gene has its own regulatory circuit**.** Most of the genes have an individual cis/trans-acting element toolkit, which might be altered throughout differentiation or upon an external stimulus. Importantly, these elements can be located far away from the regulated genes [4], hence it is very difficult to pair them up with the affected genes. Before the genomic era it was impossible to search for these, unless one was satisfied with the closest ones, the promoters. It is worth noting that many studies appeared in the literature using the so called “promoter bashing” approach in which a given DNA

element was under investigation [5]. In these experiments the element is sliced into pieces and the activities of these are assayed linked to reporter genes in various cell based systems. Essentially with the help of this technology more and more response element was described, but probably by examining only the promoter region many important regulatory elements and the factors act on it remained elusive. The discovery of the enhancer elements and the fact that their location has no major effect on the activity of the enhancer unequivocally showed that the assumption was correct so the identification of far way enhancers is like finding the needle in a haystack [4]. Shortly the weaknesses of these systems were recognized, but without the technical advances, it was almost impossible to deepen our insights into the location, function and collaborative action of distant cis-acting elements.

Another important aspect of investigating gene expression regulation is the expansion of the analysis providing a global view at the level of the genome. The first high throughput method which is able to recognize the changing transcripts was first used in 1982 aimed at detecting the level of 382 transcripts between normal and tumor tissues [6]. Actually this was a filter paper spotted technology which is known as the ancestor of the well-known microarray methodology. Although the miniaturized microarray technology was first used to measure the expression of 45 Arabidopsis genes [7]. These studies were the very first attempts to enlarge our view on gene expression and its regulation.

The impact of the genome programs and the technological advances

The result of the genome programs and the technological revolution coming with it, major breakthroughs helped scientists working in the field of mammalian gene expression regulation. As a key example, the profiling of transcription factor binding sites (TFBSs) and histone modification patterns along with nucleosome positions have been determined by linking Chromatin Immunoprecipitation (ChIP) to microarray technology [8]. This allowed the identification of histone modification and/or TFBSs in a close to genome-wide fashion and showed how old molecular biology techniques (i.e. immunoprecipitation) can be effectively revolutionized by coupling them to genome-based approaches. During the last several years, sequencing technologies have greatly

evolved and almost fully replaced the microarray-based methods. With the ability to sequence tens of millions of reads (DNA sequences) in a parallel fashion, more applications, which could be only imagined before, became a reality. This rapid evolution of sequencing, now called Next-Generation Sequencing (NGS), completely revolutionized the field of gene expression research, along with other areas of research. The linking of NGS with simple molecular biology approaches proved to be very useful and effective and led to the immediate generation of genome-wide data in a number of areas including whole genome sequencing, mRNA sequencing and revealing MNase (Micrococcal Nuclease) and DNase I hypersensitive sites [8]. ChIP has been one of the earliest applications linked to NGS leading to the determination of the typical histone acetylation and methylation patterns of gene promoters, enhancers, insulators and repressed chromatin territories [9] (Figure 3). ChIP-seq also allows the determination of the “cistrome” of any transcription factor meaning - all the binding sites in a given cell type under given circumstances, which was first carried out for Signal Transducer and Activator of Transcription (STAT) 1 [10]. These results supported the first set of evidence in a genome-wide manner that cis-acting elements are likely to be widespread in the mammalian genomes and are mostly located in the non-coding part of the genome, covering large distances relative to the putative regulated genes.

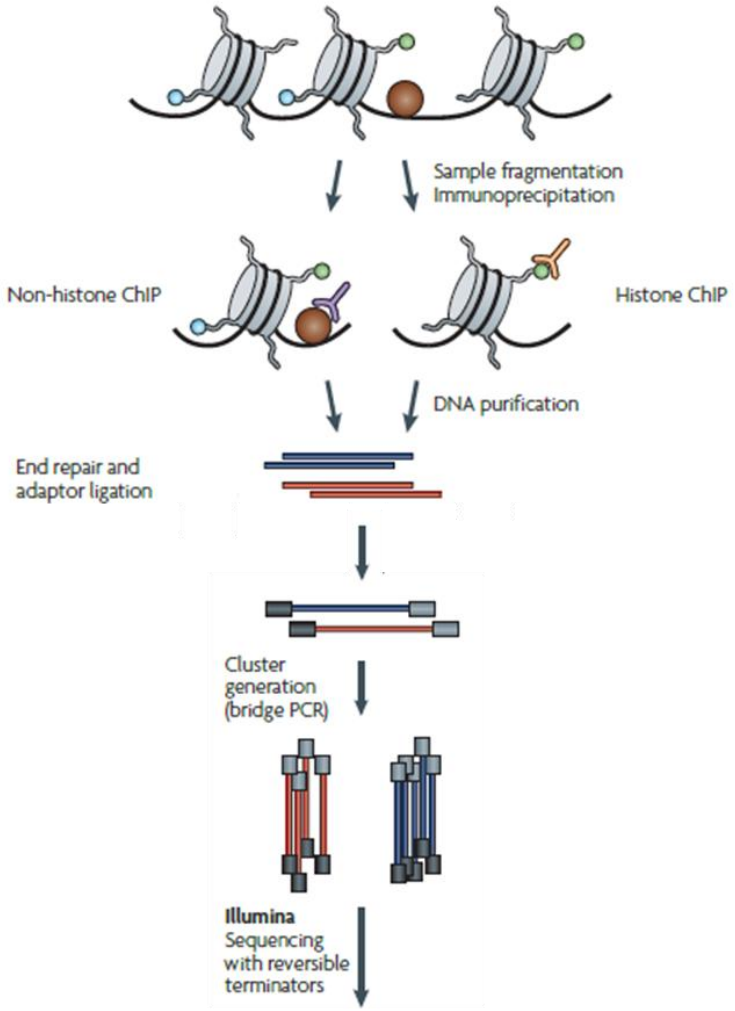

Figure 3. Overview of the ChIP-seq method. *After crosslinking with formaldehyde, chromatin is fragmented by sonication and immunoprecipitated by specific antibodies. The enriched DNA fragments then end repaired and ligated with adapter sequences indispensable for efficient cluster generation and discrimination between different samples. Histone and non-histone ChIPs are represented on the figure [8].*

Our capabilities to detect intra- and interchromosomal interactions have also expanded quite a bit. The availability of technologies, such as 3C (Chromosome Conformation

Capture), greatly increased our opportunities to construct the interaction map of cis-elements beyond their interactions with the regulated genes, which has been shown first in yeast [11]. The 3C method is based on the fixation of chromatin loops mediated by protein complexes bound to DNA. After making a snapshot about the chromatin interactions, a restriction enzyme is used to cut the genome into smaller pieces containing the compatible sticky ends. Then the chromatin is subjected to ligation in a highly diluted fashion which favors intramolecular ligation events, thus capable of connecting those genomic regions residing in the proximity of each other (Figure 4.) [11].

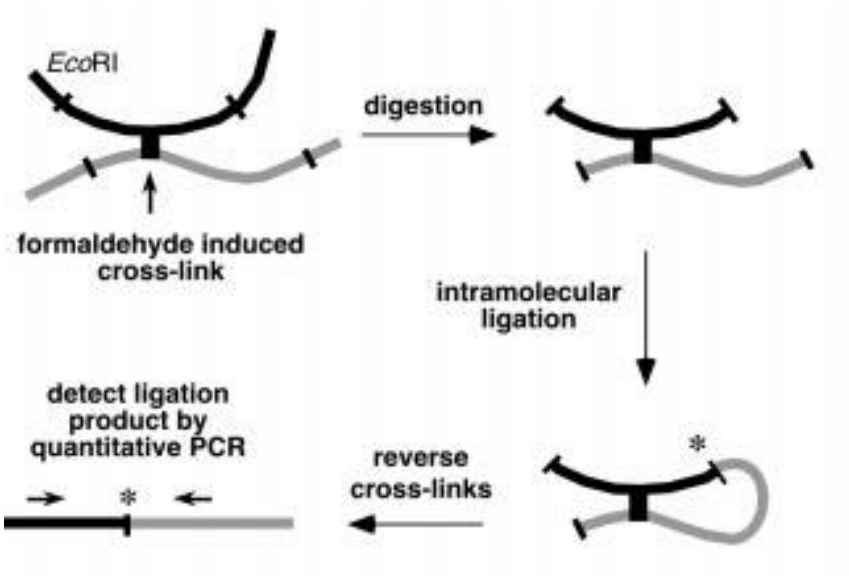

Figure 4. Overview of the chromosome conformation capture method.
Chromatin is cross-linked with formaldehyde and digested with a restriction enzyme (in this case EcoRI). After these steps using T4 DNA ligase intramolecular ligation takes place and the newly formed restriction sites will be detected by PCR reactions [11].

In the last couple of years 3C has been also linked to NGS and also combined with previously mentioned ChIP which resulted in the following technologies (3C-seq, 4C-seq, 5C, ChIA-PET, Hi-C). Depending on the biological question, one can choose these methods to reveal the chromatin interactions of the genome or a particular locus. The 4C assay can be very useful if a singular genomic element and its interactions are examined. Initially, libraries obtained from 4C experiments were hybridized to microarrays to get insights into chromatin interactions, actually this why the method is

called 4C referring to Chromosome conformation capture-on-chip [12]. 3C-sequencing basically holds the same advantage as the 4C experiments and it can be used for detecting one to all interactions [13]. 5C (Chromosome Conformation Capture Carbon Copy) overcomes the disadvantage of 4C and 3C-sequencing by offering the opportunity to map all the interactions in a large chromosomal territory on a megabase scale [14]. However, Hi-C has been the most powerful amongst these by offering the unbiased detection of all the interactions genome-wide [15]. ChIA-PET (Chromatin interaction analysis by paired-end tag sequencing) [16] is the genome-wide version of the method called ChIP-loop [17], in which the combination of ChIP and 3C is applied to reveal the interactions between any two loci bound by the protein of interest. These technological improvements definitely offered the possibility to explore the interaction maps about the identified cis-acting elements and their promoters. In addition, these can inform one about the organization of higher-order chromatin structure in the nucleus. These technologies are quite robust, provide a lot of data, but are prone to artifacts and require significant bioinformatics efforts to analyze and interpret the data.

Finding the link between the regulator and the regulated

The combination of molecular biology, NGS and bioinformatics applications and methods continue to provide a large amount of useful information about the location of cis-acting elements and the gene expression profile of a given cell type. However, assigning the cis-elements to the affected genes remained very cumbersome, if not impossible. The strategy to simply link TFBSs to the closest regulated gene is not reliable in case of larger distances between the putative, predicted enhancers and the regulated genes. Experimental evidence is required for such interactions, by methods such as Chromosome Conformation Capture (3C). 3C coupled to quantitative PCR reaction, is suitable for measuring the interaction frequency between any two loci in the genome. However, prior information (e.g. TFBSs or histone modification patterns) of the given loci is critical for the experimental design and the control template preparation is also time consuming [18]. The development of Hi-C overcomes these problems [15]. This method is based on 3C, of which ligation products are sequenced on an NGS platform detecting all the genome-wide interactions. However, its resolution is far from

ideal to map enhancer-promoter interactions because the restriction enzyme of choice dictates the resolution of the experiments. Fullwood et al. reported a new technology called ChIA-PET in 2009 as previously mentioned. The method harnesses the combination of ChIP and 3C, which was first applied to map the chromatin interactions influenced by the estrogen receptor alpha in the human genome [16]. The usage of this technology provided a tool to understand how a transcription factor can act from long distances, and implied that chromatin interactions are one of the driving molecular mechanisms for regulating gene expression in the mammalian genomes. Development of Global Run-On sequencing (GRO-seq) was an additional breakthrough (Figure 5.). The procedure is based on the classical nuclear run-on transcription assay, which has been used in molecular biology for decades, providing a snapshot about the level of *in vivo* (at least in an intact nucleus) synthesized nascent RNAs [19]. If done as a time course, the dynamics of the induced/changing transcription can be assessed and even quantitated. GRO-seq was first utilized to map the amount, position and orientation of the transcriptionally engaged RNA polymerases in primary human lung fibroblasts [20]. The obtained results were very striking, showing that about 30% of the human genes are occupied by active polymerases, genes are transcribed beyond the 3' end of the annotated regions, and surprisingly, most promoters possess engaged polymerases in the opposite orientation to the annotated gene. This divergent transcription is associated with active genes, but usually is not elongated efficiently to the upstream regions relative to the TSS [20]. The technical advance made it possible to go forward using these approaches to reveal the complex regulation of the genes in a genome-wide manner, however, it became clear that the combination of the approaches is necessary to understand and explain the most exciting findings.

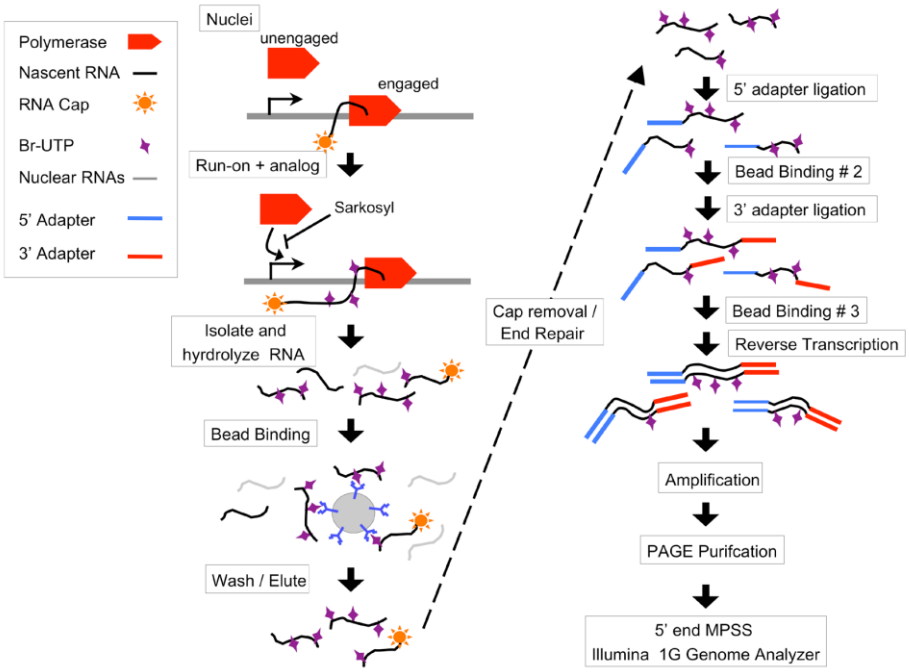

Figure 5. Overview of the GRO-seq protocol. *Run-on reaction is allowed to proceed, while RNA polymerases transcribe 100 bases in intact, isolated nuclei in the presence of sarkosyl and Br-UTP. Nuclear run-on RNA is then base hydrolyzed and isolated using α-BrdUTP antibody coated agarose beads. RNA is then further modified by removing the cap and repairing the ends for adapter ligation. Adapter containing RNAs are then reverse transcribed and amplified followed by sequencing [20].*

Enhancing gene expression by well-defined cis-elements, termed enhancers

In complex genomes, cis-acting elements are dispersed and can be located over several hundred kilobases away from their targeted genes [21]. Developments regarding NGS rapidly expanded our view about the putative location of regulatory elements and shed light on the problematic nature of linking cis-elements to the regulated genes. However, it has been shown in one of the earliest studies that their mutations may result in

congenital diseases, meaning that their proper action is indispensable for gene expression [22].

According to these data, the precise annotation of cis-elements has become an important and widely studied issue in the field of transcriptional regulation, and an absolute requirement for appropriate annotation of genome function. Enhancers are DNA sequences with the ability to recruit various types of transcription factors for the interaction with the mediator complex, as well as with the members of the (pre-)initiation complex. By looping mechanisms, the complexes assembled on the DNA can facilitate RNA-polymerase II (RNAPII) binding to the promoter, thus the initiation of gene transcription [23]. It has been also shown that transcription factors bound to enhancers are able to remodel the surrounding chromatin structure to establish NFRs (Nucleosome-free regions) by recruiting ATP-dependent chromatin remodeling enzyme complexes, thus facilitating transcription factor binding, transcription initiation and elongation [24]. The appearance of MNase-, DNase I-, and ChIP-seq as main tools to investigate open chromatin, transcription factor binding and epigenetically marked histone landscapes has greatly impacted and improved our understanding about the characteristic features of enhancers. Results from genome-wide studies mapping nucleosome occupancy indicate that at cis-regulatory elements, histone replacement is more enhanced than at other genomic locations not harboring such enhancer-like properties [25]. Active promoters/TSSs are barely occupied by nucleosomes as well, thus these form also NFRs. These results suggested that nucleosome stability contributes to gene regulation [26]. Later, it has been shown that the two alternative, minor histone variants, H3.3 and H2A.Z, are enriched near NFRs [27]. High-resolution co-activator CREB-binding protein (CBP) and P300 ChIP-seq experiments provided further insights into the chromatin signatures of enhancers. These proteins interact with various transcription factors and possess histone acetyl-transferase activity, which makes them capable to modify histones [28]. Several studies showed that these factors are good indicators of enhancer function in a tissue specific manner [29]. These results suggest that CBP/P300 co-factors are key functional components of the enhancer binding complexes. Certain histone modifications also participate in cis-element function and it has been described that the main characteristic features of active promoters are

characterized by the residence of RNAPII and TBP-associated factor 1 (TAF1), marked by NFRs flanking with trimethylated histone H3 at lysine 4 (H3K4me3), acetylated H3 (H3ac) and TFIID [30]. On the other hand, as previously mentioned P300 is one of the well-documented and accepted active enhancer marks along with the enriched H3K4me1, H3K4me2 and H3K27ac [31]. These observations were confirmed in different cellular model systems leading to the identification of enhancer landscapes of a given cell type [32] [33]. Based on these studies, the identification of enhancers and their characteristic features are relatively straightforward, however, the annotation process to the affected genes remained largely elusive. Utilizing the features of RNA-seq and GRO-seq, another useful feature has been described, namely that active enhancers transcribe the so-called enhancer RNAs [34] [35]. The observation added another layer to enhancer features and it turned out that this might be the most reliable indicator that the enhancer actively participates in gene regulation [35].

Enhancers and the “position effect”

As far as the main genomic characters of enhancers are recognized, their identification has been more efficient using enhancer prediction methods based on both evolutionary conservation [36] and ChIP-seq results [37]. Based on transcription factor/cofactor occupancy and histone modification landscapes, the enhancers can be identified, but their targeted genes are more difficult to find. In the absence of a better method, most studies in the field utilized simply proximity based, predictive approaches to link enhancers to the regulated genes [38] [39]. As it has been previously mentioned the development of 3C methodology and its combination with NGS technology rapidly changed our view about genome structure [40]. Traditionally, nuclear organization was evaluated by microscopy based methods. Since then, different 3C related methods have been emerged leading to the 3D determination of chromatin structure at various gene loci [40]. The advantage of these 3C related methods over microscopy is their higher resolution and the ability to allow the investigation of a single gene and its interaction profile with the surrounding chromatin environment. It is important to mention that all the 3C-based approaches will always need a helping hand from the side of microscopy to fully uncover the shape of the genome and to identify the most reliable interactions [40] .

The most evident example is the comparison of gene expression in active and inactive chromatin regions. Chromatin segregation into active and inactive regions raised the question whether positioning into these regions affects gene expression. Later on it was clearly documented using fluorescent in situ hybridization (FISH), that certain genomic regions change their nuclear position upon the activating stimulus [41]. Probably, some of these changes at the level of gene expression are attributable to cis-acting elements such as enhancers. The correlation between nuclear position and gene expression has been shown in several studies [40]. Silent genes are localized closer to nuclear lamina than their active counterparts and supported the basis of the so-called ‘‘position effect’’, which describes the behavior and action of cis-regulatory elements in the context of higher-order chromatin structure. Another intriguing concept is the existence of subnuclear compartments enriched in transcription factors and RNAPII, called transcriptional factories [41]. These would be highly relevant for enhancer activity and might even suggest that enhancers and promoters co-localize in such subnuclear regions.

The relevance of annotating the human regulatory element landscape, ENCODE project

Humans are estimated to have about 20000 protein-coding genes, but this covers only 1.5% of the entire genome. The other 98% accounts for intronic and intergenic regions, non-coding RNAs and short or long interspersed elements [42]. ENCODE is a project launched in 2003 aimed at revealing all the cis-acting elements in the human genome via the collaboration of several research groups as part of this specialized consortium [43]. This initiative is the continuation of the Human Genome Project, but in this particular case the final aim was the delineation of functional DNA sequences that act at the protein and RNA levels in a given cell type [44]. The functional DNA sequences defined as distinct genomic sections encoding a specified output for instance, protein product or non-coding RNA. Another main character of these elements is the reproducible biochemical trademark exemplified as protein binding or specific chromatin structure. Using 1640 data sets from 147 different cell types complemented with all

ENCODE data regarding candidate regions from genome-wide association studies (GWAS) and evolutionarily constrained genomic territories revealed crucial components about the function and organization of the human genome. The main conclusions made by the ENCODE consortium were the following: ‘‘(1) More than 80% of the human genome is associated with at least one biochemical signature in one particular cell type. (2) Classification of the genome into functionally different chromatin states implies an initial set of approximately 400000 enhancer-like regions and more than 70000 promoter-like elements. (3) Quantitative correlation of RNA production versus chromatin marks and transcription factor binding on the promoter regions indicate that RNA expression is mostly dependent on the functionality of the promoter. (4) ENCODE annotated at least as much functional non-coding DNA sequences as the protein-coding genes. (5) Single nucleotide polymorphisms (SNPs) associated with disease phenotypes determined by GWAS enriched in non-coding functional elements annotated by the consortium’’ [45]. Based on these considerations, it is evident that the non-coding part of the genome is full of functional and disease associated cis-regulatory regions. The basis to connect these to the distal target genes remained unexplored. Gene promoters and their cis-acting elements can participate in looping that is involved in gene regulation [46]. In order to link genes and their putative cis-regulators, chromosome conformation capture carbon copy (5C) is carried out and interaction maps are generated from three different cell lines and these results were integrated with the ENCODE data. Combining these datasets revealed more than 1000 interactions between enhancers, CCCTC-binding factor (CTCF) bound sites and promoters in each cell lines. Significant correlations are observed between gene expression and the existence of promoter-enhancer interaction and the presence of transcripts originating from enhancers (enhancer RNAs, eRNAs). Surprisingly, unlike the reported functions of CTCF as an insulator, long-range interactions are not blocked by CTCF/cohesin co-bound sites, demonstrating that many of these sites are not demarcated physically insulated gene domains. The fact that only 7% of the loops are detected with the closest gene suggests that genomic proximity is not necessarily a good indicator for long-range interactions [47]. This study clearly demonstrates that cis-acting elements communicate with their targeted promoters via looping; nevertheless, if one would like to reveal the entire

interaction map between these elements, one needs a more robust method capable of detecting all the chromatin loops. Although 5C is a very powerful method to map the interactions, it is limited to a single locus. The usage of ChIA-PET solved this issue. In order to link the regulatory elements to their targets, DNase I hypersensitive sites (DHS) are identified leading to the determination of open chromatin landscape of several cell lines representing the human genome. These regions were then aligned with RNAPII ChIA-PET results gathering all the participants of RNAPII dependent open chromatin interactions. This large-scale interaction analysis confirmed that cooperation between DHS sites and promoters are markedly enriched. Surprisingly, this kind of integration revealed that approximately half of the DHS sites are detectable in the close proximity of more than one promoter. These results suggest that the human cis-acting element network is more complex than expected [48].

Taken together, ENCODE identified an enormous amount of functional elements in the human genome and provides a beneficial resource for the field. On the other hand, the data presented has greatly enlarged our understanding about the functionality of the human genome directing us toward new challenges regarding cis-acting element annotation and how these act genome-wide. Although the functional significance of the detected chromatin interactions is not known at all.

Fine tuners of gene expression regulation, pioneering, bookmarking and the role of higher order chromatin structure

In the previous sections the knowledge regarding the features of cis-acting elements and how one can recognize them in the very complex mammalian genome were summarized. Each of the genes encoded in the human or mammalian genetic material is likely to have many individual cis-acting elements distributed across tens to hundreds of kilobases. These regulatory sequences act in collaboration to fine-tune gene expression in a highly tissue and signal specific manner. What kind of factors dictate the functionality of such elements in the genome leading to differential gene expression patterns in various cell types?

From a functional point of view there are at least three categories of enhancers: (1) Potential enhancers not binding its cognate transcription factor. (2) Non-active enhancers, binding a particular transcription factor, but not participating in enhancement of transcription and (3) Active enhancers, binding the required transcription factor and activating transcription. Molecularly, the diversity is likely to be greater. These need to be identified, sorted out, linked to regulated gene(s) and functionally validated. From a mechanistic point of view enhancer elements must be prepared before activation. It has been shown that the forkhead transcription factor A (FoxA) binds to inactive genomic regions and by remodeling the nucleosome and histone modification pattern, it is able to recruit other transcription factors leading to enhancer activation. This observation leads to the establishment of the pioneering factor concept. The first two pioneer factors (FoxA, GATA) have been described in the liver differentiation program. By definition, pioneer factors hold the property to bind nucleosomes and compact chromatin, and remain bound during mitosis. FoxA is a typical pioneer factor, which is capable of transforming the actual enhancer to a state called ‘‘poised’’ for activation [49]. This state of the enhancers makes them ready for rapid activation once the specific signal or its downstream effector appears. The most accepted view is that during differentiation pioneering factors establish the active cis-regulatory element repertoire, thus contribute to the acquisition of cell identity. More precisely, this concept describes that the regulatory landscape evolved through the differentiation program determines the sites where transcriptional regulation occurs [50]. Interestingly, in 2013 Ostuni et al. reported the existence of the so-called cryptic or latent enhancers in macrophages. These regulatory regions are not bound by the lineage-specific transcription factor PU.1 in terminally differentiated macrophages and do not show the main histone profiles of active enhancers. Importantly, upon stimulation by an activating signal (IL-4, INFg, TGFb), their histone profile can suddenly change due to the binding of the downstream effectors and leads to the subsequent binding of PU.1. In this study the authors took a closer look at the STAT6 (Signal transducer and activator of transcription) activated latent enhancers. STAT6 upon activation by IL-4 gets phosphorylated, form homodimers and translocate to the nucleus, thus regulates transcription. Activated STAT6 established the formation of latent enhancers but importantly, after washing out the

activating signal, most of them do not return to the latent state. Instead they remain marked by H3K4me1 and upon restimulation, they mediate a faster and stronger response, thus provides an epigenetic memory to the cell (Figure 6) [51]. It remains to be seen how wide spread is this mechanism among different cell types.

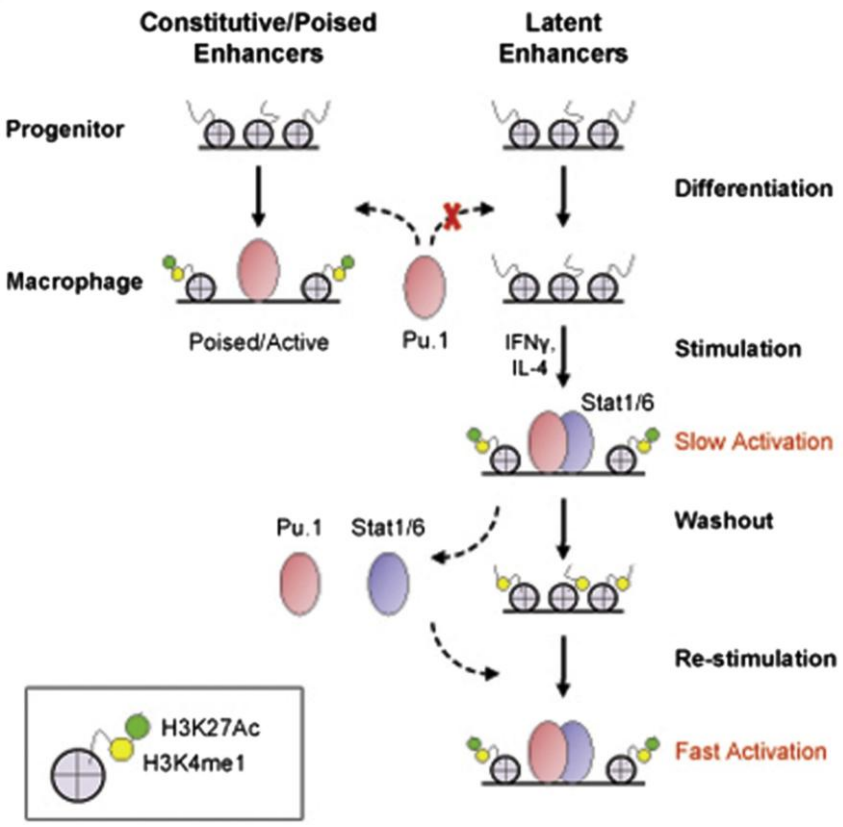

Figure 6. The mechanism of latent enhancer formation in macrophages. *Both constitutive/ poised enhancers and latent enhancers are closed by nucleosomes at the progenitor stage. Upon differentiation constitutive/poised enhancers are opened up by the myeloid specific factor PU.1, but not the latent enhancers. Once the activating stimuli is present (INFγ, IL-4), signal specific transcription factors STAT1/6 are phosphorylated, translocated to the nucleus and bind to these closed chromatin regions, initiating the establishment of functional enhancers. Latent enhancers are then marked by the well-known active histone signature H3K27ac and the nucleosome free cis-regulatory element signature H3K4me1. Washing out the activating molecule leads to the releasing of STAT factors, PU.1 and the active enhancer mark H3K27ac, but not H3K4me1. Upon restimulation these sites are capable of triggering a more pronounced effect on the target genes [51].*

Altogether, the picture seems quite complex. Cell type specific pioneer factors exist and at least in part shape the cis-acting element landscape across cell types. Other factors affecting the architecture of the genome have been also shown to be important in establishing the functional regulatory elements through contact insulation [52]. At the same time it is also known that not all lineage specific transcription factors have chromatin remodeling activities associated with. Therefore, these are more appropriately called bookmarking factors. It remains to be discovered how these factors establish cellular memory mechanistically.

It has been shown that each chromosome has its own physical location in the nucleus [40]. Chromosomal territories are functionally different and spatially separated, but what are the determinants of this higher-order chromatin structure, which is implicated in the regulation of gene expression and also responsible for cell autonomous transcriptomes?

This line of investigations started when insulator sequences were discovered in vertebrates due to their ability to block enhancer function [53]. Later on, it has been also described that CTCF allows these insulator sequences to function as separating enhancer/promoter interactions and also active/passive gene domains [54] [52]. CTCF is a transcription factor possessing eleven zinc-fingers and is ubiquitously expressed in higher eukaryotes. The function of CTCF was further clarified, showing that it functions with the multiprotein cohesin complex containing the following subunits: SMC1, SMC3, RAD21, SA1, SA2 [55]. Cohesin has a ring-like shape with a diameter of approximately 60 nm. This important feature of cohesin makes it suitable to handle the chromatin fiber as shown in sister chromatid cohesion [56]. According to the described features of these proteins, it is conceivable that they are crucial components of shaping the higher-order chromatin structure.

CTCF had long been thought to contribute to the structural organization of the genome, but its long-range interaction mediating effect has remained elusive until it has been linked to cohesin on the mouse *Infg* locus [57]. This study was the very first to show that both CTCF and cohesin are indispensable for genomic interactions. Others also reported the crucial function of these proteins in mediating chromatin interactions. Two independent studies showed that cohesin depletion leads to diminished

promoter/enhancer interactions in embryonic stem cells [58] and in thymocytes [59]. Importantly, series of genome-wide studies show that CTCF and cohesin co-occupy regions in the genome [60] [61]. The extensive interaction between these factors may explain how CTCF separates functionally different domains. Recently, contact mapping of chromosomes determined by Hi-C methods revealed the topological domain structure of the genome [15]. These domains contain multiple genes and show differential gene expression profile and epigenetic pattern. Presumably, these domains serve as fundamental building blocks that support active and passive chromosomal architectures. It has been shown that the anchoring points of chromatin loops, organizing the domain structure, are enriched for CTCF and cohesin binding sites [62]. Other studies have reported that CTCF/cohesion co-bound regions mediate the looping events surrounding promoter enhancer elements, while those regions occupied only by cohesin are responsible for enhancer/promoter interactions [58]. Based on these results, several studies showed evidence that CTCF and cohesin are required to maintain topological domain structures, interchromosomal interactions and enhancer/promoter interactions [63] [64] [65]. Perturbation of the cohesin complex has been shown to affect gene expression involving not only the cohesin bound genes, but also those that are devoid of cohesin suggesting its function in maintaining topological domain structure [64]. Based on the above mentioned results, CTCF and cohesin appear to have crucial roles in proper gene regulation, although their genomic binding sites suggest that their effects on gene expression are not cell type specific because the CTCF cistrome is largely invariant between cell types. Surprisingly, a study as part of the ENCODE project, compared CTCF binding sites from 19 different human primary and immortal cell lines and show that there is plasticity in CTCF binding across cell types indicative of strong cell-selective regulation of CTCF binding. Using massively parallel bisulfite sequencing, the authors could show that approximately 40% of variable CTCF binding is due to differential methylation states at two specific points of the binding motif of the transcription factor. Strikingly, they could demonstrate also that CTCF binding is greatly different between primary and immortal cell lines. The latter harbors widespread disruption of CTCF sites associated with increased methylation [63].

Taken together, CTCF/cohesin co-bound sites appear to be responsible, at least in part, for the configuration of topological domain structure. By shaping the genome architecture these factors significantly contribute to the regulation of gene expression in collaboration with the cell type specific pioneer transcription factors and also the signal specific ones (Figure 7).

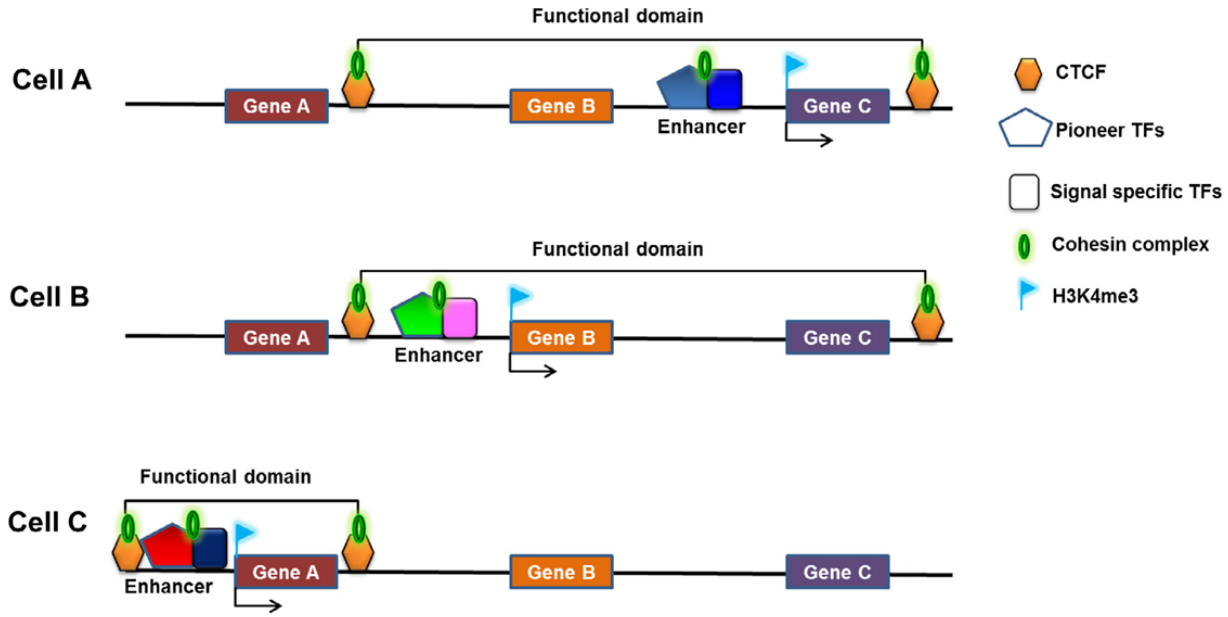

Figure 7. The major mechanisms by which cell type specific gene expression is executed. *Cell type specific gene expression program is established by the appearance of various cell type specific transcription factors (pioneering/bookmarking factors), which can shape the functional cis-acting element landscape during differentiation. Higher order chromatin structure is also involved in the formation of active gene domains. In cell A, the functional gene domain is demarked by CTCF/cohesin co-bound regions and the active enhancer element is occupied by the pioneering factor, thus support an entry point to the signal dependent transcription factor. In cell B, the same gene domain is active, but because of the presence of a different pioneering factor it uses a different regulatory element, which in turn lead to a differential gene expression output, while in cell C the chromatin conformation of the locus is reshaped due to the differential binding of CTCF/cohesin, hence a new enhancer element gets into the close proximity of Gene A, possessing another cell type specific pioneering factor with the ability to open up the enhancer waiting for activation by the signal specific transcription factor [66].*

Enhancer RNAs, functionally relevant or a byproduct?

Recent advances in genomic technologies made the surprising finding that active enhancers are transcribed into RNA molecules, called enhancer RNAs (eRNA). The first results describing the existence of eRNAs originated from the locus control region (LCR) of the beta-globin gene clusters [67]. The fact that there is pervasive transcription on enhancer elements came with the advent of total RNA sequencing, showing that in neuronal activity regulated and T-cell specific enhancers are transcribed [34] [68]. More and more studies were published recently in the field using GRO-seq in various cells and species, which have supported clear evidence that enhancers are transcribed to eRNAs and to a given stimuli, the activation dynamics of these elements are similar to their targeted genes [35] [69].

There is a great debate on the field about the functionality of eRNAs. The most important question regarding these molecules originated from enhancers is: Are these functionally relevant in gene expression regulation or eRNAs are just merely byproducts of gene transcription? Recently, several studies published using novel methods to test the functionality of enhancer-derived transcripts. Specific degradation of eRNAs using either RNA interference or antisense oligonucleotides demonstrated that the expression of the adjacent gene targeted by the enhancer is reduced [70] [71] [72] [73]. Two out of the four studies also performed an eRNA tethering assay connected to a reporter system. Interestingly, eRNA linked to either the enhancer [70] or the promoter [72] was capable of increasing the expression of the reporter gene. As further evidence, Lam et al. showed that by cloning various sizes of genomic regions from an enhancer into a reporter vector could differentially affect the activity of the reporter gene. If the core sequence containing only the TFBS was cloned, they detected increased reporter activity compared to a reporter plasmid harboring random DNA sequence. Surprisingly, the reporter consisting of the core and the full length eRNA-coding sequence had the highest transcriptional readout. Inverting the eRNA-coding region supported the most striking result because changing the sequence of the eRNA diminished its enhancing effect, suggesting that the sequence of these short RNAs can be very important in the context of gene regulation [71]. According to these results, it seems that eRNAs possess

regulatory function, but the question is how. What is the molecular mechanism through which these small molecules reach the heart of transcription?

It has been demonstrated that enhancers actively participating in looping with their targeted promoters transcribe higher level of eRNAs [74]. Based on these studies, the predicted model must be that eRNAs somehow secure the contact between enhancers and their corresponding promoters. Nuclear receptors like estrogen receptor alpha (ERa) have been shown to bind enhancer elements and in the presence of the agonist, oestrogen, these cis-acting elements are anchored at target gene promoters through long-range chromatin interactions determined by ChIA-PET [16]. Recently, a very important finding came to light showing that knockdown of eRNAs immediately next to ERa bound enhancers reduced enhancer/promoter interactions and resulted in a reduced expression level of the corresponding gene. Potentially, these ERa-mediated eRNAs are taking part in the modulation of looping, which was further supported by the fact that eRNAs could pull-down the subunits of the cohesin complex. On the other hand, RNA immunoprecipitation (RIP) assays performed against RAD21 showing that eRNAs could enrich in the cohesin complex [70], which has been shown to control enhancer/promoter interactions [58]. Furthermore, targeted degradation of eRNAs by RNA interference led to the inhibition of oestrogen dependent RAD21 recruitment at several ERa bound enhancers. Strikingly, knockdown of RAD21 almost fully diminished the interaction on the gene loci NRIP1 and GREB1 between the enhancers and their corresponding promoters. In addition, knockdown of SMC1, another component of the cohesin complex, almost completely abolished the oestrogen mediated gene activation program [70]. Thus, eRNAs may participate in the process of looping by initiating or stabilizing the interactions of enhancer/promoter pairs (Figure 8).

Overall these studies imply that at least in some cases eRNAs possess regulatory function and contributes to gene expression regulation. Clearly, further studies are needed to clarify their precise role in gene expression regulation.

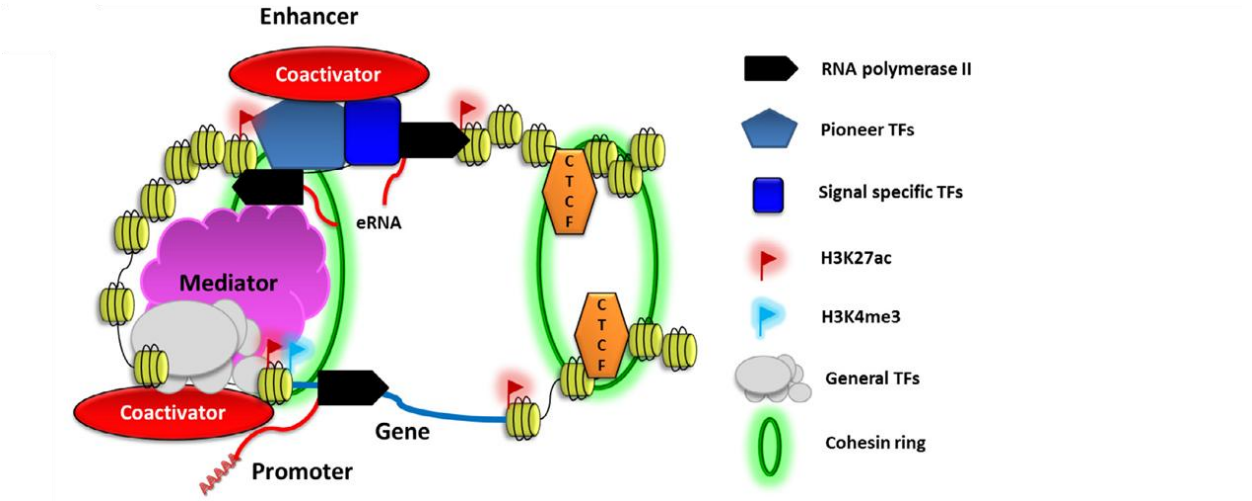

Figure 8. The proposed action and features of active enhancers. *Enhancers and their target genes are located in the same topological domain shaped by the CTCF/cohesin factors. Enhancers upon activation by the stimuli get acetylated at H3K27 by various coactivators then loop into the close proximity of the promoter region of the targeted gene potentially by an eRNA-dependent mechanism, which also relies on the cohesin complex. After these steps the enhancers can initiate or boost the expression of the target gene by assisting the assembly of the general transcription factors and the mediator complex along with the deposition of the active transcription start site marker H3K4me3 [66].*

Enhancers and disease

Based on the fact that enhancers are the main drivers of gene expression it is obvious that their proper action is required for gene transcription. There are more and more studies in the literature, which connect single nucleotide polymorphisms (SNPs) with cis-acting elements supporting the notion that SNPs in the regulatory regions may cause disease phenotypes and thus contributes to evolution. Disruptions in chromosomal regions not harboring coding genes provided the first evidence that mutations in the non-coding part of the genome may contribute to disease development.

Mutations in the cis-acting elements can be easily annotated to a given disease phenotype if they fulfill one or more of the following criteria's: (1) Genetic evidence is present to link the phenotype to a known disease locus. (2) Validated chromosomal

anomaly (deletion, amplification, rearrangement) can be annotated to a known disease gene. (3) Resulting phenotype is very similar to the phenotypic change caused by a mutation in the coding region of a known disease gene. (4) Disease associated variation accounts for all or a significant fraction of disease risk [75].

One of the earliest studies mentioned the regulatory element mutations of the PAX6 locus involved in Aniridia. Aniridia is characterized by the absence of iris and is mostly due to mutations occurred in the coding sequence of the PAX6 gene [76]. However, a fraction of the cases are not caused by mutations in the coding sequence. It has been shown that the downstream genomic region is full of rearrangements. The most distal point harboring the mutation lies approximately 125 kb from the last exon of PAX6 and fall in the intronic region of the ubiquitously expressed gene ELP4, although haploinsufficiency for ELP4 has been shown not to contribute to the disease phenotype. YAC-based transgene experiments in mice revealed an approximately 80 kb long genomic region containing series of DHSs. Later on, it has been described that these cis-acting elements are required for PAX6 expression [77].

Another striking example has been demonstrated with the POU3F4 gene. Mutations occurred in the coding region responsible for X-linked deafness type 3, however a smaller group of cases was identified that lack the gene variation. Interestingly, approximately 900 kb upstream from the gene's TSS, a very important 2 kb element has been described to overlap with an otic vesicle enhancer, likely regulating POU3F4 [78].

These experiments have shown that regulatory elements function over long distances, they can reside in other transcriptional units and their mutations can cause disease phenotypes. The listed diseases are typically inherited in a Mendelian manner, although mutations present in the non-coding regions of the genome where they associated with non-Mendelian diseases.

The rapid evolution of genotyping technologies has resulted in GWAS data that generally imply a powerful role for regulatory variation in common genetic disorders [93]. Recently, meta-analysis of approximately 1200 SNPs representing the most significant association with disease phenotypes has been done. Surprisingly 40% of these falls into the non-coding part of the genome, suggesting that disease causing mutations may act on enhancer elements [79].

Several studies successfully demonstrated roles for non-coding mutations in disease risk. One of the earliest was identified in the intronic region of the RET gene contributing to Hirschprung disease [80]. Recently, it has been shown that a mutation residing in the non-coding region belonging to the IRF6 gene is a risk factor of cleft lip associated with Van der Woude Syndrome [81].

Overall, the vast amount of data unequivocally support the notion that cis-acting elements are crucial components of proper gene regulation and their mutations give rise to various pathological conditions. Importantly, until now no one could investigate the effect of these genetic variations on the functional cis-acting element toolkit in a special cell type.

Recently, Heinz et al. used different mouse strains looking for differential binding of transcription factors caused by natural genetic variation. In this study they could delineate the strain specific differences in functional enhancer usage in macrophages. Most of the cells express hundreds of transcription factors to control the non-coding part of the genome, thus shaping the cell specific transcriptome. In macrophages, transcription factors responsible for marking regulatory elements are PU.1, C/EBPs and the AP-1 family members. They showed that in macrophages these lineage determining transcription factors (LDTF) collaboratively occupy 70% of the active enhancer elements. Surprisingly, if a SNP can be detected in the PU.1 motif leading to diminished PU.1 binding, it is negatively affected by the binding of the other two LDTFs. Conversely, if the binding motif of C/EBP or AP-1 harbors the mutation, PU.1/AP-1 and PU.1/C/EBP binding also diminished. In addition, H3K4me2 and H3K27ac markers of active enhancers are also abolished, meaning that these enhancers were no longer functional. These results provided a definitive answer to the question: How enhancer function and transcription factor binding is lost where there are no mutations in its binding motif? Based on these, the answer is the requirement for collaborative binding. Interestingly, further examination of the strains in the context of a signal specific transcription factor, NF-κB, led to the observation that mutations occurred in the LDTF motifs are approximately three times more likely to result in decreased NF-κB binding, than mutations occurred in the NF-κB motif. This observation showed the importance of LDTFs in chromatin priming/remodeling, as previously mentioned [82].

The authors claim that this collaborative binding model can be very useful if one determines the LDTFs in a given cell type and merge these binding sites with the annotated genetic variations to pinpoint the potential disease-causing variants. The future challenges will be to expand these studies and use them in different model systems to understand disease-causing natural genetic variations [82].

Nuclear receptors

Nuclear receptors are in the centrum of gene regulatory processes. Their precise action is known to implicate various developmental progresses, regulation of metabolic pathways and also the proper functioning of the endocrine system. The relevance of this protein family is enormous due to the fact that these transcription factors contribute to normal human physiological processes, but also on the other hand implicated in the manifestation of many pathological conditions. It is obvious that a detailed understanding of these regulatory proteins is needed in order to develop new drug treatments.

Their activity on transcription is tightly regulated by small lipophilic molecules which can dock into the ligand-binding domain (LBD). The first receptors were described close to 50 years ago as intracellular proteins capable of sensing steroids [83] [84].

More than 20 years were needed to clarify that these receptors are part of a large superfamily of metazoan transcription factors and that share conserved domain structure consisting of separate DNA-binding (DBD) and ligand-binding domain. Nearly all nuclear receptors possess this domain architecture. The DBD is located to the N-terminal part of the protein, while the LBD is part of the C-terminus. The DBD is highly conserved and it has a nuclear localization signal at its carboxyl-terminus, required for proper nuclear translocation. Two zinc-binding domains can be also found on the DBD, these are responsible to maintain the architecture of the DBD, and function as one single unit. The LBD is more diverse part of the receptor and interacts with the ligands and mediates the transcriptional response in a ligand dependent fashion. The binding of the ligands determine the recruitment of various transcriptional coregulators which then lead to the repression or induction of the target gene. Among the coregulators, coactivators like P300 (histone acetyltransferase p300), SRC (steroid receptor coactivator) and corepressors like SMRT (silencing mediator for retinoid and thyroid receptor) or N-CoR

(nuclear corepressor) are described. Taken together, the LBD is a multifunctional domain which not only serves to bind the ligand, but also responsible for coregulator recruitment and implicated in the dimerization processes. Based on crystal structure studies it has been shown that the LBD has an apolar part named as the ligand-binding pocket (LBP) which is a conserved domain within the nuclear receptor family, but importantly it can be very distinct between receptors. For instance, the peroxisome proliferator activated receptors (PPARs) has one of the biggest LBP and it results in the various ligand-binding ability of the receptor. Actually this is one good reason why the PPAR receptors are in the main focus of the pharmaceutical investigations. The other reason is that their key functions in the regulation of metabolic processes. Going back to the LBP, it has another crucial part which is responsible for the stabilization of the ligand bound LBP, namely the Helix 12 (H12). H12 plays an indispensable role in the molecular switch, which describes how the receptors can change coregulators in the presence of the unique activator ligand. In general, nuclear receptors bind corepressors in the absence of the agonist. Once the activating molecule bound to the LBP it causes conformational changes in the structure of the receptor, which makes it capable to release the corepressors and bind the coactivator proteins. This coregulator exchange will lead to the modification of the surrounding chromatin structure and enhances the initiation of transcription [85]

.

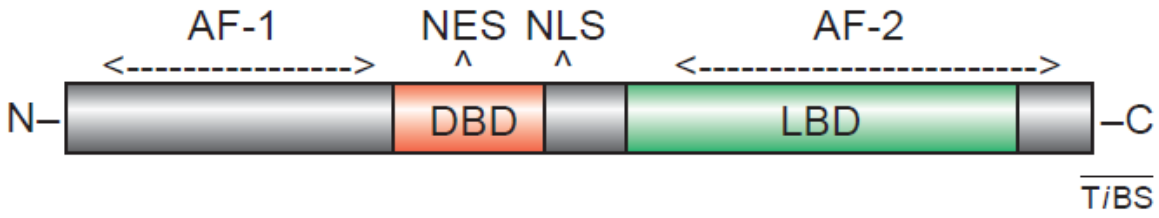

Figure 9. Domain structure of nuclear receptors. *The N and C termini are indicated. Abbreviations: AF-1, Activation function-1responsible for ligand independent function, AF-2, Activation function-2 responsible for ligand dependent function. The DBD and LBD are also indicated along with the nuclear export signal (NES) and nuclear localization signal (NLS) [86].*

The classification of nuclear receptors

Nuclear receptors can function through homo- or heterodimer formation. These protein dimers can bind to specific response elements in the DNA, these sites can exist as half-sites separated by various spacer nucleotides. According to their direction the half-sites can form direct or inverted half-site repeats. Mangelsdorf et al. suggested four categories of nuclear receptor classes: 1, Steroid hormone receptors which can act by forming homodimers binding to inverted half-site repeats. 2, RXR heterodimers and the last two 3, and 4, including the orphan receptors for which the activator ligand is not known. The difference between the two groups is that the 3, class can bind to the DNA via direct repeats forming homodimers, while the 4, class receptors can bind to the DNA as monomers [85].

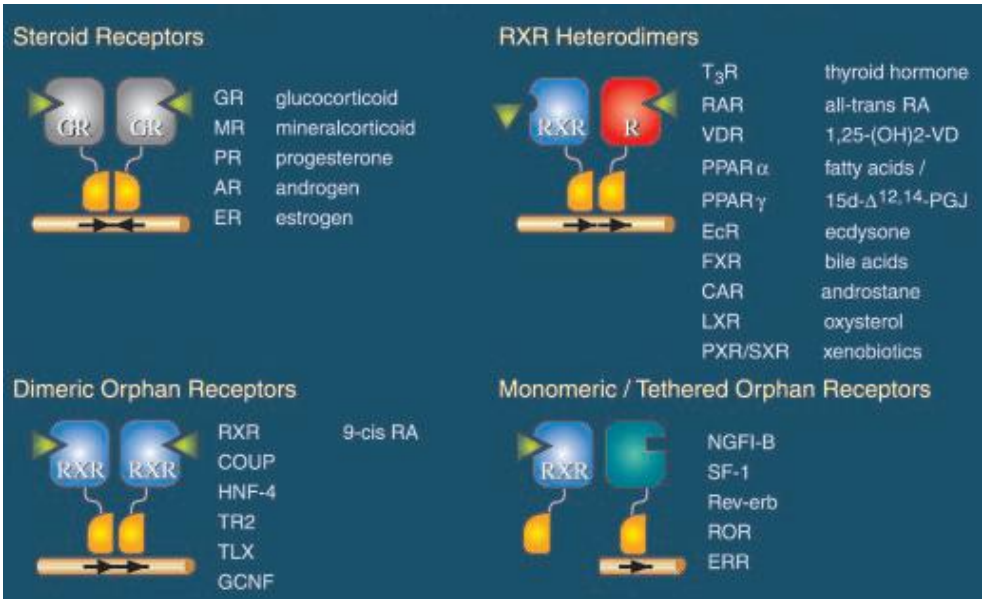

Figure 10. Classification of nuclear receptors based on their dimerization, DNA- and ligand binding properties *[85].*

Macrophages as biologically relevant model system in nuclear receptor biology

More than 100 years ago Elie Metchnikoff shed light on the existence of macrophages and designated them as big phagocytes [87]. Today, the picture is quite complex. It

seems that macrophages represent a highly specialized and heterogenous cell population [88]. These cells are taking part in various steps of innate and adaptive immune responses, fundamental components of the antigen presentation system, but also involved in wound healing, tumor progression, bacterial elimination and cytokine production. In addition, more and more study supports their contribution to the pathogenesis of metabolic diseases for instance, atherosclerosis, diabetes and metabolic syndrome [89] [90]. It is not surprising based on the above mentioned pathological disorders that nuclear receptors have obvious and indispensable functions in macrophages due to their evidenced effect on metabolism. The key question is whether there is a new nuclear receptor related function exist waiting to be explored.

The action of nuclear receptors in macrophages

LXRs

Liver X receptor has been described in 1994. The receptor is involved in cholesterol- and bile-acid related metabolic processes. Two isoforms were identified: LXRα which is highly present in the liver, but also can be found in adipose tissue, intestine, lung and in macrophages and LXRβ which is ubiquitously expressed [93] [94] . LXRs have been identified as sensors of oxysterols and it has been shown that 24(*S*)-hydroxycholesterol, 25-hydroxycholesterol, 22(*R*)-hydroxycholesterol, and 24(*S*),25-epoxycholesterol can bind and activate the receptor. The main functions of LXR in macrophages are: regulation of lipid metabolism, modulation of immune functions and the inhibition of the apoptotic pathway [95].

PPARs

Three members of the PPAR subfamily have been identified. PPARα is predominantly expressed in the liver and regulates the oxidation of fatty acids [96]. PPARγ regulates the uptake and storage of lipids in adipose tissue, but also in immune cells like macrophages and dendritic cells. PPARγ has been also linked to the regulation of immune responses in macrophages [97] [98]. The main roles of PPARβ are linked to fatty acid metabolism, mitochondrial respiration, thermogenesis and muscle lipid

metabolism [99]. All the PPAR receptors bind various nonesterified-, polyunsaturated fatty acids, eicosanoids and also prostanoids, thus regulate lipid metabolism. Interestingly, all of the PPAR isotypes are under very intensive investigation due to their positive effects on insulin sensitization and muscle regeneration [100].

The retinoic acid receptors, RARs and RXRs

The Retinoic acid receptor has unequivocal effect on differentiation, development and various aspects of metabolism. The discovery of the retinoic acid hormone as a morphogen contributed to the understanding of how RARs exert their multiple effects in the body [101]. Two families of retinoid receptors has been identified, both RARs Retinoic acid receptors and Retionid X receptors are capable of binding the retionoid molecules and thus regulate the transcription of their target genes [102]. RARs have been shown to be indispensable factors in embryonic and myeloid development, wound healing and also in the buildup of the nervous system [103] [104].

RXR

RXR is a unique nuclear receptor because it functions as an obligate heterodimerization partner of other receptors. It is expressed in every cell type and is required for post-natal life in mice [105] [106]. Its presumed main molecular function is to regulate the activity of a dozen or so nuclear receptors. There is also evidence that it can form homodimers and/or have heterodimer-independent signaling capacity [107] [108]. A key concept regarding RXR signaling is the permissiveness, non-permissiveness mutually exclusive dual paradigm. According to this, in certain heterodimers such as RXR:PPAR, RXR:LXR ligand activation of RXR results in transcriptional activation, hence these are permissive heterodimers, whilst in other heterodimers such as RAR:RXR, TR:RXR and VDR:RXR RXR is suppressed or “subordinated” and therefore these so-called non-permissive heterodimers cannot be activated from the RXR side [109]. Therefore the activation of all permissive heterodimers present in a particular cell type might lead to pleiotropic gene activation and engagement of potentially conflicting pathways.

The existence of pleiotropy, the role of RXR activation of permissive heterodimers and the presence and activity of RXR homodimers have been debated and remain largely unresolved.

The fact that certain natural lipids such as 9-cis RA, docosahexanoic acid and phytanic acid are able to activate RXR gives support to the biological role of RXR activation *in vivo* [92]. There are also potent and selective synthetic compounds such as Bexarotone (LG10069) and LG100268 (LG268) [110] [111] that have been used to dissect the role of the receptor in various biological systems and/or used in therapies. Ligand activation of RXR is believed to regulate distinct and coherent gene expression in various cell types and it has been shown to have beneficial pharmacological effects in various animal models such as increasing insulin sensitivity in diabetic animals [112] and more recently selective activation of RXR in microglia and astrocytes has been shown to clear β-amyloid and reverse deficits in an AD mouse model [113] although the extent of the improvement was recently questioned [114]. These data support that ligand activation of RXR is biologically and therapeutically relevant. Therefore understanding the liganded receptor's genomic and cistromic activity is essential to further map its biological role and therapeutic potential.

In macrophages there are several heterodimeric receptors with key cellular roles such as PPARγ regulating oxLDL uptake and processing, LXR regulating cholesterol efflux and immune function, NR4A1 (NUR77) regulating inflammatory response. These heterodimeric receptors have been linked to the development of atherosclerosis and also immune function and provide means to reprogram macrophages [95] [115]. Therefore it is biologically important to understand how activation of RXR contributes to these pathways and potentially to novel ones, and regulates gene expression in macrophages.

Macrophage activation

In the field of immunology, macrophage activation becomes a very important area. Diverse terms have been applied to characterize different macrophage subtypes. These

macrophage types are basically the result of a given stimulus, most probably a cytokine which converts the gene expression program of the cell. Due to the fact that macrophage activation is involved in the outcome and progression of several diseases including metabolic diseases, allergic disorders, autoimmune diseases, cancer, bacterial, parasitic, fungal, and viral infections it is indispensable to investigate their activation state. Recently, in order to clarify the nomenclature and the differentiation protocols regarding macrophage activation, leading scientists on the field set the gold standards of macrophage polarization. They claimed that the two most popular differentiation methods are straightforward. Therefore macrophages stimulated with INF-γ are the so-called M1 type (classically activated) macrophages and those growing in the presence of IL-4 are the M2 (alternatively activated) ones (Figure 11). In our studies we followed these instructions [91]. In one of our earlier studies we found that in M2 macrophages the IL-4 activated STAT6 can potentiate the PPARγ response, which supported the fact that STAT6 signaling and nuclear receptor signaling meet at the level of gene expression regulation through direct interaction [92]. According to these results it is worth to further investigate the collaboration between STAT6 the "alternative activator of macrophages" and nuclear receptors.

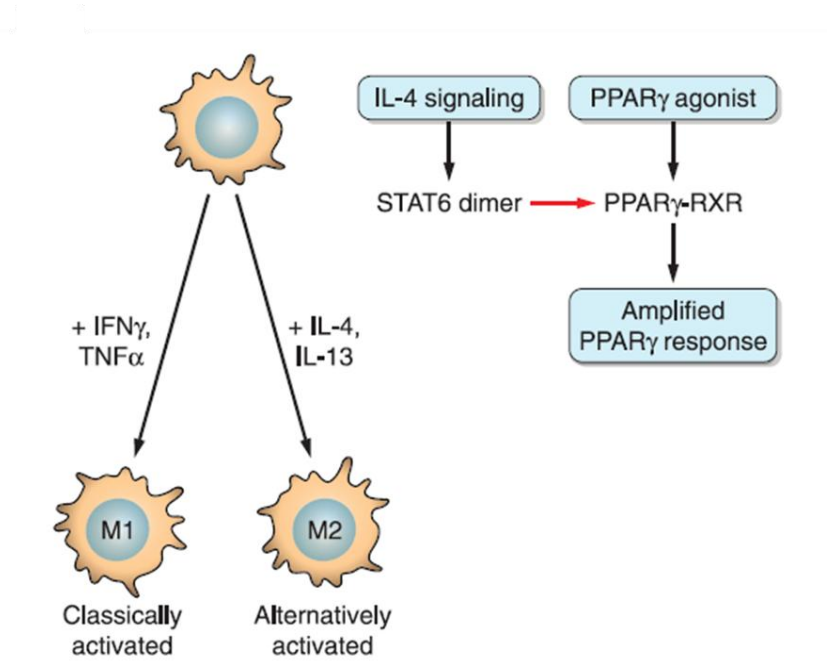

Figure 11. *The two main type of macrophage activation (left). Scheme about the collaboration of the STAT6 and PPARγ signaling pathway (right) [95].*

The importance of lysosomes in macrophages

Lysosomes are characterized by acidic pH, various characteristic proteins and lipids, and the activity of multiple acidic hydrolases capable of catalyzing the degradation of substances that are taken up by the cell through fluid-phase endocytosis, phagocytosis and autophagy [148] [149]. Abnormalities regarding lysosomal functions are leading to atherosclerosis and neurodegenerative diseases (Alzheimer's disease) [150]. On the other hand lysosomal overproduction can lead to extracellular matrix degradation, thus metastasis, emphysema, atherosclerosis, osteoporosis. Almost all of these pathological conditions have been linked to macrophages, thus it is important to investigate the molecular mechanisms driving the expression of lysosomal genes in macrophages.

The transcriptional regulation of lysosomal gene expression has been studied mainly during autophagy and a few transcription factors are shown to be involved in their regulation [151] [152].

Based on the fact that macrophages are one of the most powerful phagocytes in our body it seems that lysosomal function has to be important in order to fulfill their purifier activity, however how lysosomal genes are regulated in these cells during differentiation, activation or stress is not investigated in detail.

Hypotheses

IL-4 activated STAT6 directly regulates the expression of a set of lysosomal genes.

RXRs are likely to have a distinct impact on gene expression in macrophages.

The unique effect of RXR on gene regulation leads to a specialized biological program.

IL-4 activated STAT6 alters the genomic distribution of RXR, thus reshapes the nuclear receptor signaling during alternative macrophage activation.

Aims

1. Validation of the binding of STAT6 by ChIP-qPCR on the enhancer elements of lysosomal genes in macrophages.
2. Determination of the effect of RXR activation at the level of the macrophage transcriptome.
3. Identifiaction of the RXR bound genomic regions in macrophages by ChIP-seq.
4. Determination of the primarily RXR activated genes and the corresponding enhancer elements by GRO-seq.
5. Validation of enhacer/promoter interactions at the chromatin level by 3C methods.
6. Investigating the possible unique effect of RXR on gene expression regulation.

Methods

Materials

Ligands: LG268, LG1208 gifts from M. Leibowitz (Ligand Pharmaceuticals), RSG (Sigma), GW3965 was a gift from T. M Wilson (GlaxoSmithKline), AM580 (Sigma).

Differentiation of bone marrow derived macrophages

Bone-marrow was flushed from the femur of wild-type and RXRα/β DKO C57Bl6/J male animals. Cells were purified through a Ficoll-Paque gradient (Amersham Biosciences, Arlington Heights, IL) and cultured in DMEM containing 20% endotoxin-reduced fetal bovine serum and 30% L929 conditioned medium for 5 days.

RNA-seq

RNA-seq library was prepared from two biological replicates by using TruSeq RNA Sample Preparation Kit (Illumina) according to manufacturer protocol. Briefly, 2.5µg total RNA was used for the library preparation. In the first step poly-A tailed RNA molecules (mRNA) were purified with poly-T oligo-attached magnetic beads. Following the purification mRNA is fragmented using divalent cations at 85 ºC, then first strand cDNA was generated using random primers and SuperScript II reverse transcriptase (Invitrogen, Life Technologies). This was followed by the second strand cDNA synthesis, then double stranded cDNA fragments went through an end repair process, the addition of a single ‘A’ base and then barcode indexed adapter ligation. Adapter-ligated products were enriched with adapter specific PCR to create cDNA library. Agarose gel electrophoresis was performed on E-Gel EX 2% agarose gel (Invitrogen, Life Technologies) and the library was purified from the gel using QIAquick Gel Extraction Kit (Qiagen). Fragment size and molar concentration were checked on Agilent BioAnalyzer using DNA1000 chip (Agilent Technologies). Libraries were sequenced with Illumina HiScanSQ sequencer. RNA-seq experiments have been done by Szilárd Póliska.

RNA-seq analysis

The TopHat-Cufflinks-CummeRbund toolkit trio [117] was used for mapping spliced reads, making transcript assemblies, and getting, sorting and visualizing gene expression data. Expressional values were in FPKM format: Fragments Per Kilobase of exon per Million fragments mapped. Interaction network of the top genes (FDR < 0.1, FC < 0.5 or FC > 2 in at least one condition) was made by GeneSpring 12.5. RNA-seq analysis has been done by Gergely Nagy and Attila Horváth.

Real-Time Quantitative PCR

RNA was isolated with Trizol Reagent (Molecular Research Center). For mRNA measurements, Tetro Reverse Transcriptase (Bioline) Kit was used and transcript quantification was performed by Quantitative Real-Time PCR reaction using SYBR green dye and Universal Mastermix (Diagenode). For eTranscript measurements RNA was reverse transcribed using High-Capacity cDNA Reverse Transcription Kit (Life Technologies) and quantification was performed with LightCycler® 480 SYBR Green I Master (Roche). Transcript levels were normalized to Ppia or Rplp0 in the case of eTranscript measurements.

ChIP-seq

ChIP-seq experiments were performed based on the protocol of our laboratory [118]. Cells were crosslinked with DSG (Sigma) for 30 minutes and then with formaldehyde (Sigma) for 10 minutes. After fixation chromatin were sonicated with Diagenode Bioruptor to generate 200-1000 bp fragments. Chromatin was immunoprecipitated with antibodies against pre-immune IgG (Millipore, 12-370), RXR (sc-774), P300 (sc-585), PU.1 (sc-352), CTCF (Millipore, 07-729), RAD21 (ab992), H4ac (Millipore, 06-866), H3K27ac (ab4729), H3K4me2 (Upstate, 07-030) and H3K4me3 (ab8580). Chromatin antibody complexes were precipitated with Protein A coated paramagnetic beads (Life Technologies). After 6 washing steps complexes were eluted and reverse crosslinked. DNA fragments were column purified (Qiagen, MinElute). The amount of

immunoprecipitated DNA was quantified with Qubit fluorometer (Invitrogen). DNA was applied for qPCR analysis or library preparation.

ChIP library preparation for sequencing

ChIP-seq library was prepared with Ovation Ultralow Library Systems (NuGen) at least from two biological replicates according to manufacturer intructions. Briefly, 1 ng immunoprecipitated DNA was submitted to end repair reaction. Adaptors were ligated to end repaired DNA fragments. Library was amplified with adaptor specific primers in 16 PCR cycles. Libraries were gel-purified with E-Gel® systems (Life Technologies) to remove needless primers. Libraries were quantified by Qubit fluorometer and the quality was assessed with Agilent 1000 DNA Chip. Libraries were sequenced with Illumina HiScanSQ sequencer.

ChIP-seq analysis

Primary analysis of the ChIP-seq raw reads has been carried out using the ChIP-seq_analyze command line pipeline (Barta, 2011). Alignment to mm9 genome assembly was made by the Burrows-Wheeler Alignment (BWA) Tool. Genome coverage files (bedgraph) were made by makeUCSCfile.pl (Homer) [119] and used for visualization with IGV 2 [120]. Peaks were predicted by MACS2 [121], and artifacts were predicted based on their presence in all samples in the same position (both in factor and histone ChIP-seq results, together 19 samples), and then eliminated from peaksets.

Two parallels of control and LG268-treated RXR samples were analyzed by DiffBind v1.0.9 [122] with an input control: consensus peaks were formed from peaks predicted from two of four samples; peaks with significantly changing ‘binding affinity’ were defined using the ‘full library size’ parameter. As the used 1h treatments did not change the location of the RXR peaks, agonist and antagonist effects were also examined on the previous ‘consensus peakset’.

Peaks of non-RXR samples were selected manually based on MACS2 scores. Fold changes of peak sizes were determined based on the normalized read counts of vehicle and LG268 treated samples by intersectBed (BEDtools) [123] and further command line

programs. Histograms of reads centered to peak summits were made by annotatePeaks.pl (Homer).

Summit +/-50 bases of top 1000 peaks (based on peak score) were used for prediction of motif enrichment by findMotifsGenome.pl (Homer) (repeat masking was set, and both unmasked and final target region numbers are marked on figures). On Figure 16 and 24 insert, target percent refers to the ratio of the peaks having the given motif, and background (Bg) percent shows the motif enrichment of a large random sequence collection with 100-base lengths as the target sequences. P-values were calculated based on the comparison of these two enrichments.

As RXR binds to multifarious sites and Homer usually determines (now determined) only the merge of these (=half sites), repeat element enrichments had to be predicted one-by-one based on their RGGTCANnRGGTCA consensus with the optimization (-opt) function. Under annotated RXR peaks, the remaining „halfsite" (motif score > 6) +/-11 base sequences were got by homerTools extract (Homer), and DRn (RGKKSANnRGKKSA), ERn (TSMMCYNnRGKKSA) and IRn (RGKKSANnTSMMCY) elements (where n was between 0 and 5, and 1 mismatch was allowed) were searched by fuzznuc (EMBOSS). Peakset overlaps were defined by intersectBed (BEDTools) and visualized by VennMaster-0.37.5 [124]. ChIP-seq analysis has been done by Gergely Nagy.

GRO-seq

Approximately 10 million of nuclei were isolated from bone marrow-derived macrophages and Run-On reactions were performed for 5 minutes. RNA was then isolated by acid:phenol-chloroform extraction and base hydrolysis was performed for 20 minutes in the presence of 1N NaOH followed by DNase I treatment. Nucelar run-on RNA was immunoprecipitated with Br-dUTP antibody coated agarose beads and then washed to remove nonspecifically bound substances. Specifically bound RNA was eluted from the beads and end repair was performed in the presence of Tobacco acid pyrophosphatase and the in the presence of Polynucleotide Kinase followed by RNA

extraction with acid:phenol-chloroform. For adapter ligation the RNA was incubated for 4 hours at room tmperature in the presence of the 5'- or the 3'- adapter oligo. RNA enrichment was performed with the above mentioned anti-deoxy Br-UTP agarose beads. The eluted RNAs were then reverse transcribed and the remaining RNAs were degraded by the addition of a RNase mixture. Sequencing libraries were amplified with Phusion high fidelity DNA polymerase using the primers according to Illumina. Libraries were run on a non-denaturing 1xTBE, 8% acrylamid gel and fragments larger than 90 nucleotides were extracted from the gel and then libraries were then extracted, quantified and sequenced.

GRO-seq analysis

Primary analysis of the raw sequencing data has been carried out similarly as detailed for ChIP-seq. Alignment to the mm9 genome assembly was made by the Burrows-Wheeler Alignment Tool after clipping the reads from adaptors using Fastx tools. A pool of 2x4 macrophage sample reads was used for transcript prediction and annotation. Genome coverage files (bedgraph) were generated using makeUCSCfile.pl (Homer) with the following parameters for both strands: -fsize 5e8 -fragLength 120 -noadj -style chipseq. Based on the pooled coverage file (and a bed file containing mm9 chromosomes), all 'subpeaks' which had higher signals compared to the neighboring regions and the background, were detected by PeakSplitter (EBI Bertone Group Software). These regions were limited in +/-250 bases from their summit, and read numbers were calculated for each region from the alignment (bam) file by intersectBed (BEDTools). Divergent sites were defined from region pairs with different direction to each other when both contained more than 15 reads, and more than 10 reads per 300 bases, and after a 150-base shift to 5' direction these overlapped. Regions with the highest read numbers were selected in the case of multiple overlap. Transcripts were built from regions closer than 600 bases on the same strand (BEDTools and command line programs). As several "alignment gaps" could be found in introns due to the genetic differences between the reference genome and the strain used, these gaps in transcribed regions were filled in according to the Ensembl reference genome annotation. Transcript positions were downloaded from Ensembl by BioMart: GRCm37

(mm9). The short noncoding genes overlapping in sense direction with coding transcripts were omitted from the further annotation steps.

Control ChIP-seq data of H3K4me3 histone modification was used to separate transcription start tites (TSSs) from proximal and intronic enhancers. H3K4me3 subpeaks were predicted in a very similar manner as GRO-seq subpeaks. Subpeaks exceeded score 10 were limited in +/-1000 bases from their summit (scores were determined by PeakSplitter). The most intensive 5' GRO-seq regions were assigned to each H3K4me3 subpeaks: primarily the divergent sites above 30 reads and then the other single initiation regions above 50 reads. All possible transcripts were created from the previous predictions based on these potential TSSs. Longer transcriptional events were categorized by the following criteria: Transcripts were annotated as a known gene when the predicted TSS was closer to a known TSS than 1500 bases (Group1). 3' overhangs were collected separately from transcript bodies where it was possible. The remaining H3K4me3-started transcript predictions were separated to intergenic (longer than 1000 bases) and antisense (longer than 3000 bases because of the intronic enhancers) transcripts (Group2). All potential transcript variants were collected from Group1 and Group2 predictions. Potential transcripts (longer than 1000 bases) with divergent or single start site were collected from the remaining regions without H3K4me3 site (Group3). The rest of the H3K4me3 and divergent sites overlapping with known TSS were collected as "full pausing" sites (Group4). Unknown transcript predictions were re-annotated, potential short sense overlapping genes were collected, and a known and unknown transcript category was formed.

All transcript bodies were collected and because of the about 45 base / sec polymerase speed, up to 50 kb 5' fragments were used for gene expression analyses, without 3' overhangs, "alignment gaps" and any divergent sites, TSSs and intronic enhancers which show different expression. Calculating unique read numbers for all samples was done on these fragments if their joint length was longer than 500 bases, and RPKM-like gene expression values (Reads Per Kilobase of transcript fragments per mapped read number of control) were determined. GRO-seq analysis has been done by Gergely Nagy.

Domain predictions based on the CTCF and RAD21 peaks

CTCF/RAD21 co-peaks (the focal points of the chromatin-chromatin interactions) were designated if the overlapping CTCF and RAD21 peak scores were similarly high (over score 15, and with less than 3 fold difference). The closest co-peak pairs in 1Mb range were assigned as loop borders if their overall scores showed less than 5/3 fold difference. Active loops were united if these were closer than 100 kb to predict the active topological domains. The remaining regions between form the inactive topological domains. Predicted loops were annotated to the regulated genes by intersectBed (BEDTools). Differential binding analysis of RAD21 on the insulator (CTCF/RAD21 co-peaks) and the GRO +/- RXR enhancer regions was done similarly as described above. Domain predictions have been done by Gergely Nagy.

Chromosome conformation capture

Cells were fixed with 2% formaldehyde for 10 minutes. Nuclei were isolated in buffer containing 10mM Tris-HCl pH7.5, 10mM NaCl, 0.2% NP40 (Sigma), and protease inhibitor tablets (Roche). Chromatin was digested with 400U of HindIII (Fermentas) restriction enzyme at 37 ºC for 16 hours and for an additional 1 hour with 100U. Chromatin fragments were ligated with 100U of T4 DNA ligase (Fermentas) at 16 ºC for 4 hours. After ligation chromatin was de-crosslinked overnight at 65 ºC. Ligation products were column purified (Roche, High Pure PCR Template Preparation Kit) and DNA concentration was determined by Nanodrop. DNA fragments were submitted to qPCR reactions using TaqMan probes designed to the assayed enhancer region. Tandem primers were designed in the close proximity of the restriction enzyme cutting sites. BAC control DNA pools were used to determine primer efficiency in each analyzed genomic region and GAPDH was used as a loading control.

3C-sequencing

After the first digestion and ligation the 3C DNA pool was purified with phenol/chloroform/ isoamyl alcohol (25:24:1) (Sigma). Second restriction digestion was performed by using DpnII (NEB) for 16 hours per manufacturer's instruction. Second ligation was performed at 16C for 6 hours with 200U of T4 DNA ligase. DNA was then purified again with phenol/chloroform/isoamyl alcohol (25:24:1) followed by QIAquick gel purification column (Qiagen) purification. Bait specific inverse PCRs were performed using primers coupled to Universal Illumina adapters and Barcode sequences. Reactions were purified by QIAquick gel purification columns. Amplicon libraries were quantified and qualified by Agilent using DNA 7500 chip cartridge. Amplicon libraries were sequenced on Illumina MiSeq and HiSeq2000 sequencer. Two biological replicates were sequenced twice (technical replicates).

3C-seq analysis

Samples were de-multiplexed by Fastx-tools based on their index than bait sequences ending with 3' HindIII site, and the 68 to 82 bases long fragments starting with 5' HindIII site were aligned onto the mm9 genome assembly by BWA. Mapped reads were counted in 500, 1000, 2000 and 4000 bases windows of the genome, and the read numbers were normalized to 1000 reads for the comparison. Target coverages of 1 Mb fragments covering the whole mouse genome were determined as thousandths of all interactions. Fragments of the inactive and active topological domains predicted as described above, and the RXR regulated and independent fragments of the latter were separated and distribution of their interchromosomal interaction frequencies were plotted for the baits. 3C-seq analysis has been done by Gergely Nagy.

Reporter construct preparation

Enhancer sequences were PCR amplified from BACs or genomic DNA with overhangs on the forward (atataagctt-) and reverse (tataggatcc-) primers, digested with BamHI and

HindIII and cloned into pUC18 HSV TK-LUC. BACs were ordered from BACPAC Resources Center.

Transient transfection

COS1 cells were transfected with reporter constructs along with plasmids encoding β-galactosidase and full-length receptors in triplicates using poliethylenimine. Six hours after transfection, cells were exposed to ligands. After 48 hours of incubation luciferase activity was determined by Luciferase Assay System (Promega) and normalized to β-galactosidase activity.

ELISA

ELISA experiments were carried out based on manufacturer instructions (R&D Systems). Cells from three biological replicates were counted and plated. After ligand exposure, supernatants were collected and centrifuged. Clear supernatants were applied to the ELISA plates. Unbound substances were removed by washing and the specific polyclonal antibody was added to the wells. Unbound complexes were eliminated by washing, then substrates were added to the wells and enzyme activity was determined.

Chorioallantoic membrane (CAM) assay

Fertilized white leghorn chicken eggs (Wyverkens, Halle, Belgium) were incubated for 3 days at 37°C before removing 3ml of albumen to detach the shell from the developing CAM. After making a window in the eggshell to expose the CAM, it was covered with cellophane. Eggs were then further incubated until day 9, when the test conditions were applied. Sterile absorbable gelatin sponges (1-2 mm^3; Hospithera, Brussels, Belgium) were impregnated with $3x10^4$ BMDM treated with 100nM LG268 for 6 hours and then placed on the CAM. LG268 alone was used as a negative control and recombinant hVEGF-A_{165} was used as a positive control. The windows were again covered and the incubation continued until day 13, when angiogenesis was assessed. Membranes were

fixed with 4% paraformaldehyde at room temperature for 2h. A large area around the contact or insertion sites was removed, placed on a glass slide, and photographed under using a Zeiss Lumar V.12 stereomicroscope. Digital images were captured using an AxioCam MRc5 and processed with Axiovision 4.5 Software (Zeiss). To determine the number of blood vessels, a grid containing three concentric circles with diameters of 4, 5, and 6 mm was positioned on the surface of the CAM and all vessels radiating from the sample spot and intersecting the circles were blindly counted under a stereomicroscope. CAM assay has been done by Conny A. Gysemanns and Jiri Keirsse.

Statistical tests

qPCR, ELISA and CAM assay were presented as means +/-SD. We made at least three biological replicates and we performed paired (two-tailed) t-tests and results were considered significant with $p < 0.05$. Interaction frequencies of the chosen genomic regions were considered significantly different with $p < 0.0001$ according to the unpaired (two-tailed) t-test.

Results

STAT6 regulates the expression of many lysosomal genes during alternative macrophage activation

In this study we used gene expression correlation analyses to identify transcription factors which might regulate lysosomal function. This *in silico* method is based on the observation that the expression of transcription factors and their target genes are often positively related [126]. By using this approach we found STAT6 as the most significant hit, which is a signal dependent transcription factor activated by IL-4 (Interleukine-4) involved in the regulation of immune system and also has crucial functions during alternative activation of macrophages. The background of such analysis is that if there is a transcriptional regulator which coordinated action is responsible for the expression of a group of lysosomal genes, it might be possible to identify such a factor by using correlation analyses across a great number of microarray data. Clustering of all known lysosomal genes led to the determination of several subgroups whose expression appears to be coordinated. Importantly, STAT6 was correlated with the biggest lysosomal cluster consisting of known acidic hydrolases and vacuolar H+ ATPase subunits (Figure 12).

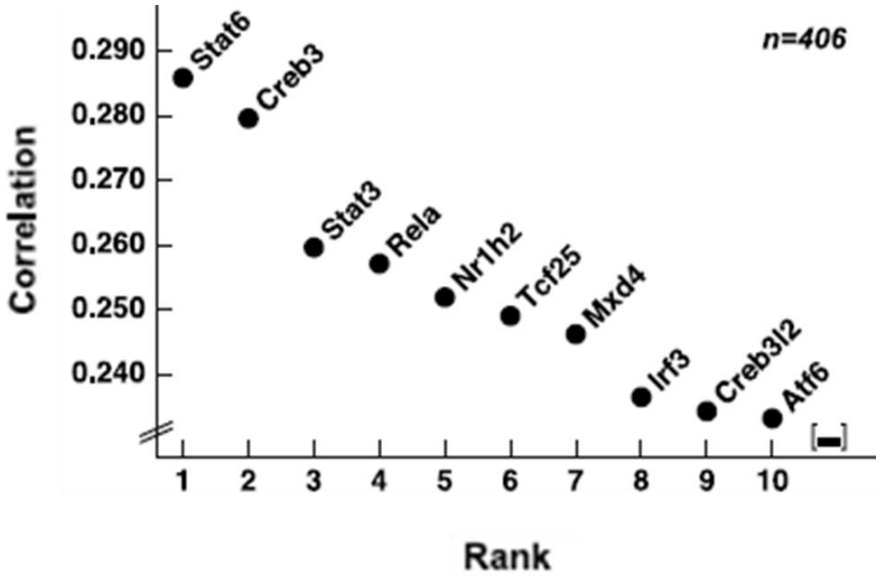

Figure 12. Correlation of lysosomal genes (n=406) with transcriptional regulators.

Importantly, 14 out of the 15 subunits building up the vacuolar H+ ATPase were shown to be regulated by STAT6 and three subunits (*Atp6v0a1*, *Atp6v0d2* and *Atp6v1b2*) were among the most strongly induced lysosomal genes by IL-4/STAT6 (Figure 13).

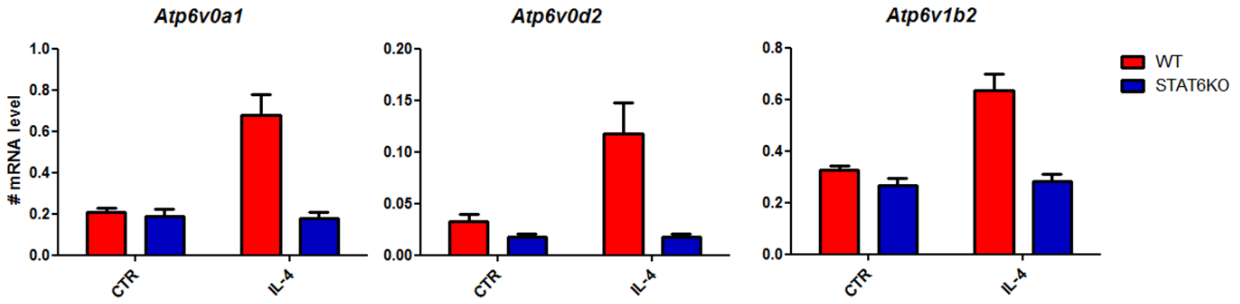

Figure 13. Gene expression measurements of lysosomal genes in the presence or absence of IL-4 (20ng/ul) for 6 hours in wild type and STAT6 knock out bone marrow-derived macrophages. *The mean and +/- SD of three technical replicates are shown.*

In order to find the regulatory elements from where STAT6 regulates the lysosomal target genes, we used publicly available ChIP-seq results in alternatively activated macrophages. Global analysis around lysosomal genes revealed that STAT6 bound genomic regions are overrepresented in the close proximity of this gene set (Figure 14). We could successfully validate these binding events on the above mentioned three genes by ChIP-qPCR (Figure 15).

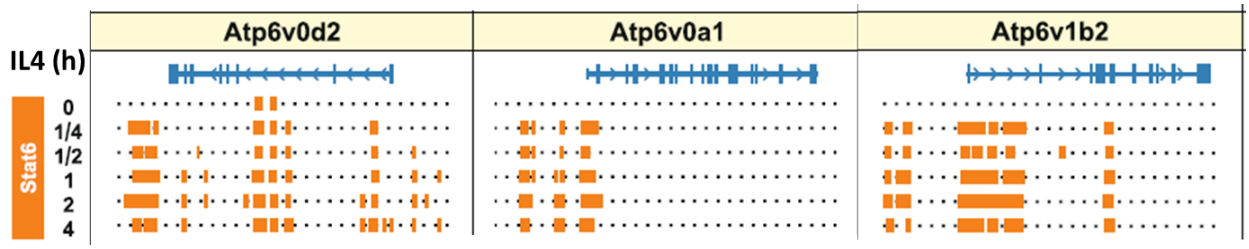

Figure 14. STAT6 binding is overrepresented around lysosomal genes. *Analysis is based on publicly available STAT6 ChIP-seq results in bone marrow-derived macrophages deposited by Ostuni et al. [51]. STAT6 bound regions are highlighted with orange rectangles on three representative examples treated with IL-4 for the indicated period of time.*

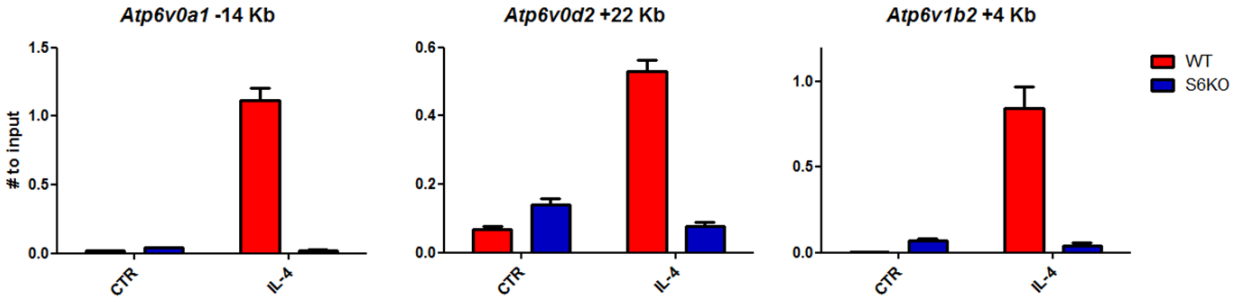

Figure 15. ChIP-qPCR measurements against STAT6 on the enhancer regions of three lysosomal genes in BMDMs. *Cells were treated for 30 minutes with IL-4 (20ng/ml) and ChIP was carried out in wild type and STAT6 knock out bone marrow-derived macrophages. The mean and +/- SD of three technical replicates are shown.*

Taken together, in this study we assigned a new transcription factor to the lysosomal gene panel in macrophages which can fine tune their expression. The data presented here identified a novel pathway that is implicated in the lysosomal function during alternative macrophage activation. In our next study we made an attempt on revealing the global, genome-wide effect of a signal specific nuclear receptor (RXR) in the regulation of macrophage specific gene expression.

The transcriptional readout of the activated RXR in murine macrophages

In order to study the effect of the RXR activation in macrophages we set out to systematically dissect the genomic events following RXR liganding in BMDMs (Figure 16). The synthetic and selective RXR agonist, LG268 was used throughout the studies in 100 nM concentration.

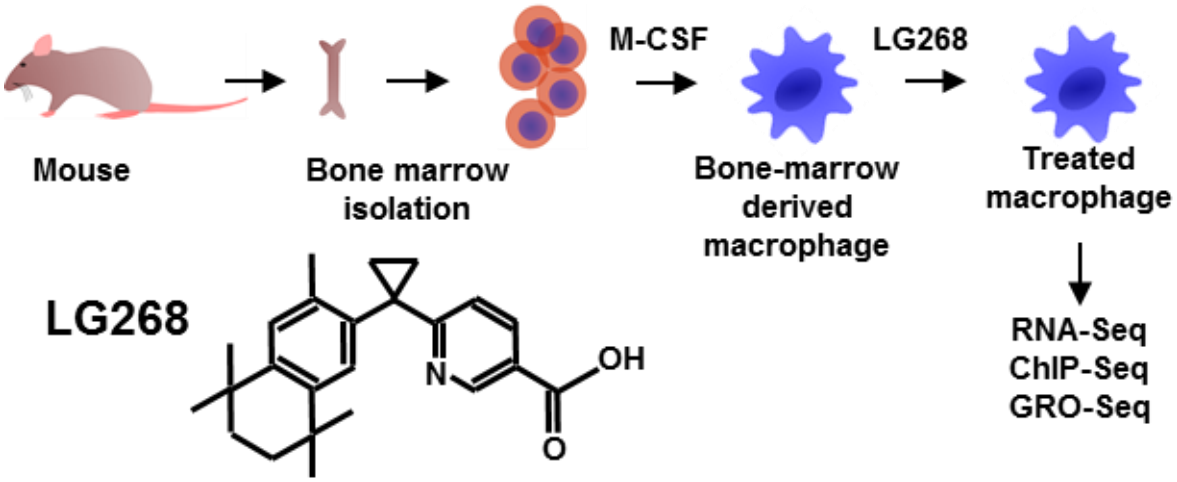

Figure 16. The scheme of bone marrow-derive macrophage differentiation. *Cells from the bone marrow were differentiated in the presence of M-CSF and then stimulated with LG268 as a synthetic activator of the RXR receptor. The chemical structure of LG268 is shown.*

First we identified the variable transcripts by RNA-seq and found that selective activation of RXR affects hundreds of genes at the steady state mRNA levels. Figure 17 shows the hierarchical clustering of the top 200 changing genes as a time course.

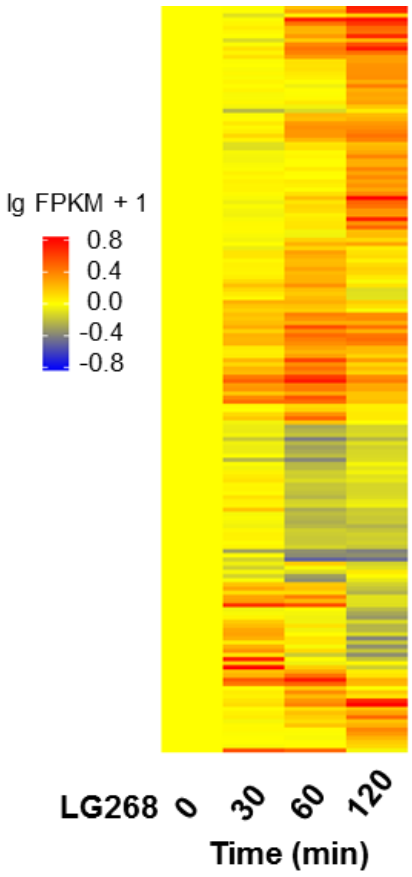

Figure 17. *Heat map representation of the changing transcripts obtained from RNA-seq experiments in the presence of LG268. The presented gene set was filtered at FDR < 0.1 and a more than twofold change.*

In order to identify the immediate early responder targets to RXR activation we wanted to analyze and determine in detail both the genomic binding events of RXR (RXR cistrome) and the enhancers where the receptor acts and responsible for the altered transcription.

The liganding of RXR results in enhanced RXR binding on preformed PU.1 marked cis-regulatory elements and leads to P300 recruitment and modification of surrounding chromatin structure

Performing ChIP-seq experiments we revealed the cistromes of RXR in the absence and presence of the activator ligand LG268, and its relationship to the lineage specific transcription factor PU.1 and the co-factor P300 along with a marker of transcription initiation H3K4me3 and active histone marks H3K27ac, H4ac and H3K4me2. We found

circa 5,200 RXR bound genomic regions. Importantly, the distribution of RXR cistrome does not changed significantly by ligand treatment in 60 minutes, although the peaks gained around 30% more reads suggesting that RXR binding is enhanced (Figure 18). We confirmed these observations by ChIP-qPCR on binding regions of known directly regulated genes [127] [128] (Figure 18).

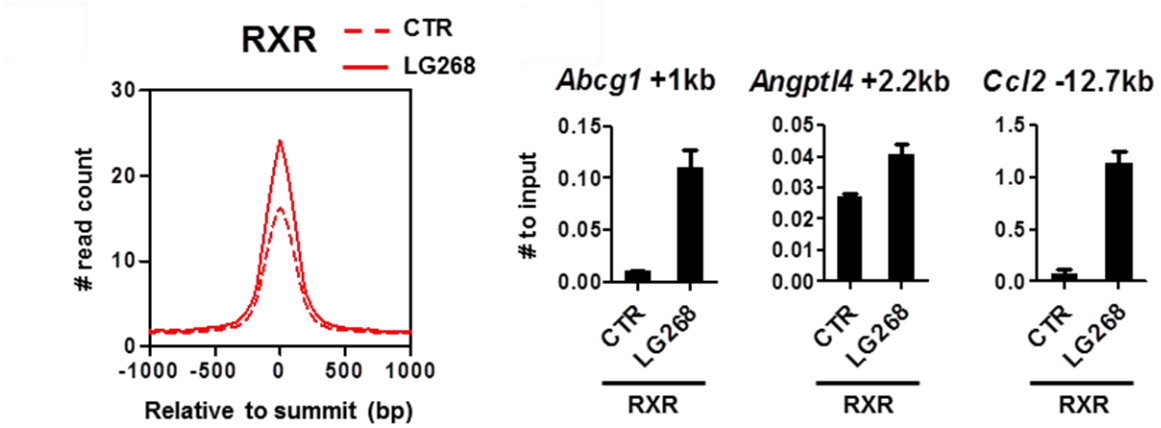

Figure 18. *Histogram representation of the average read distribution of RXR ChIP-seq relative to the peak summit in the absence (dashed line) and presence (continuous line) of LG268 (left). (#) means normalized. RXR binding on the indicated individual enhancers confirmed using ChIP-qPCR. Macrophages were treated with 100 nM LG268 for 1 h (right). The mean and +/- SD of three technical replicates are shown.*

The motif rank order on the detected RXR peaks revealed that PU.1 had the most significantly enriched motif followed by various combinations of nuclear receptor repeats, including DR1 (motif capable of binding the PPAR/RXR dimer) and DR4 (motif capable of binding the LXR/RXR dimer), of the nuclear receptor binding (half) site. Furthermore, however with lower abundance two other macrophage-associated transcription factor motifs were also observed, C/EBP and AP-1 (Figure 19).

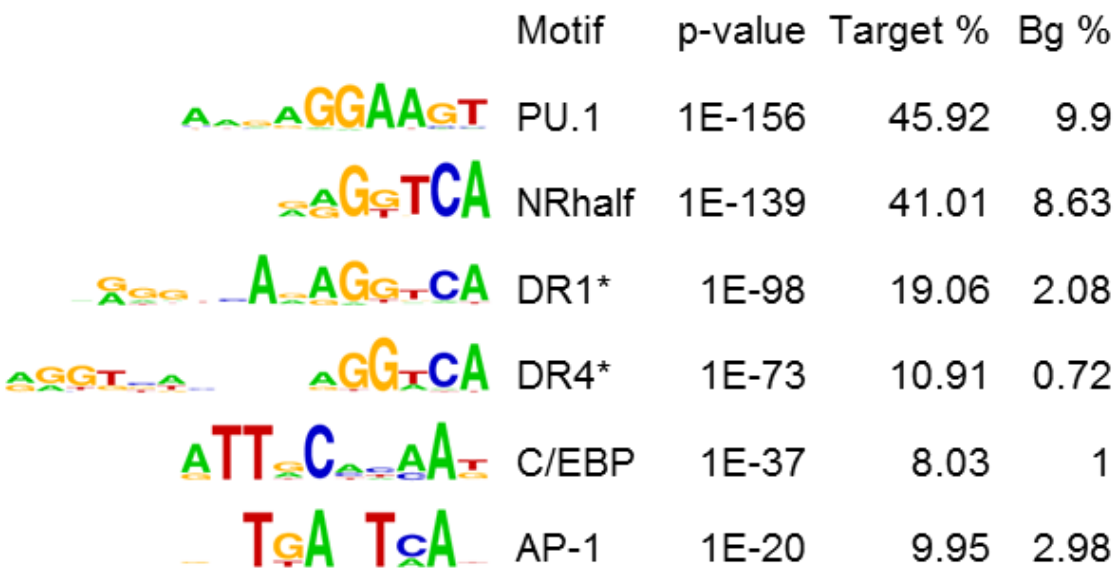

Figure 19. *Identification of the motifs under the detected RXR binding regions by de novo or targeted (marked by asterisk) motif search using Homer. "Target %" refers to the ratio of the peaks having the given motif, and "Bg %" refers to the ratio of a random background as described in the methods section.*

As far as the genomic distribution of the detected peaks is concerned, almost 90% were found outside of the potentially directly regulated 823 genes identified by RNA-seq experiments i.e. not within 10 kb of their transcription start sites (TSS) (Figure 20) suggesting that linking binding sites to regulated genes simply based on proximity will be difficult and not reliable.

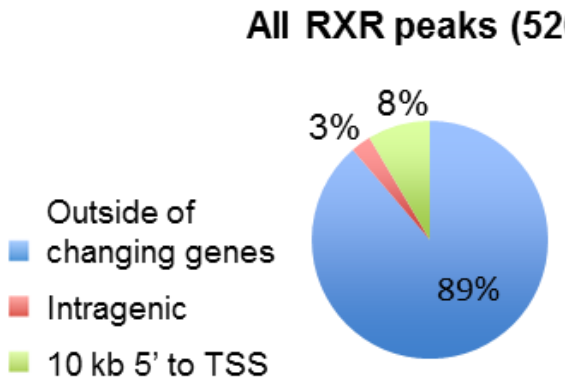

Figure 20. *Genome-wide localization of the detected RXR binding sites relative to the transcription start site (TSS) of the closest regulated genes determined by RNA-seq (P < 0.05; n = 823).*

ChIP-seq against PU.1 revealed around 30,000 binding sites, with only a minimal rearrangement upon ligand exposure. Intriguingly, P300 binding follows the genome-wide RXR enrichment, suggesting that P300 is likely to be recruited to the genome upon RXR activation. Global analyses show that RXR activation leads to P300 recruitment which could be validated by ChIP-qPCR at individual genomic regions (Figure 21).

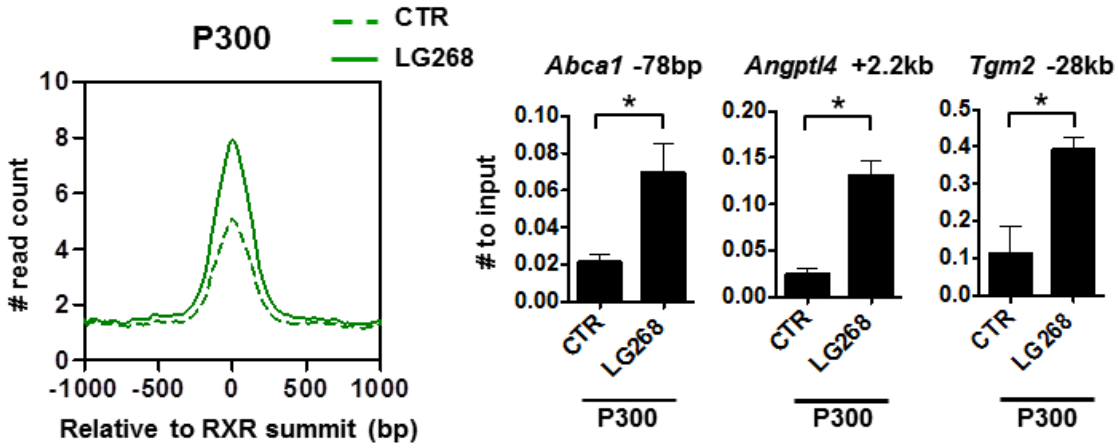

Figure 21. *Histogram representation of the average read distribution of P300 ChIP-seq on the RXR peaks relative to the peak summit in the absence (dashed line) and presence (continuous line) of LG268 (left). (#) means normalized. P300 binding on the indicated individual enhancers confirmed using ChIP-qPCR. Macrophages were treated with 100 nM LG268 for 1 h (right). The mean and +/- SD of three biological replicates are shown. Asterisk represents significant difference at P < 0.05; n = 3.*

Next we determined the overlap between the cistromes of PU.1 and RXR in the ligand-activated state which revealed that more than 2/3 of the RXR binding regions co-bound by the myeloid lineage determining factor PU.1 (Figure 22).

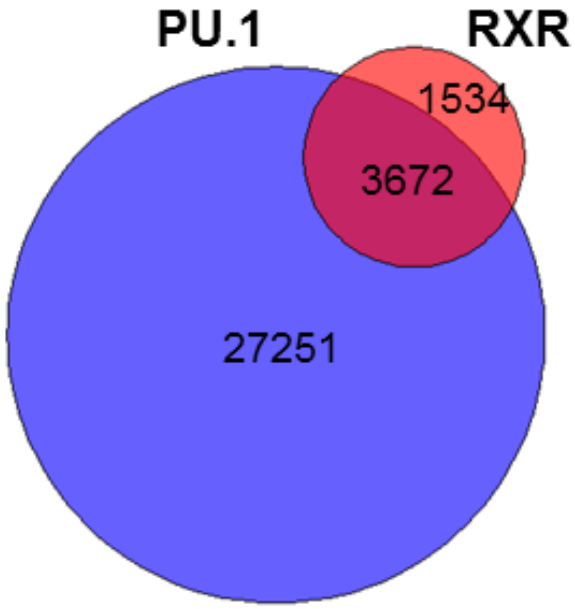

Figure 22. Venn-diagram representation of the intersection between the RXR and PU.1 cistromes.

As a next step we analyzed the RXR cistrome with regard to P300 and histone modifications known to mark active enhancer regions. Using the consensus RXR peak set (~5200) we observed enhanced RXR, P300 and H4 acetylation in the presence of the activator ligand, whilst H3K4 dimethylation and H3K27 acetylation remained unaltered (Figure 23).

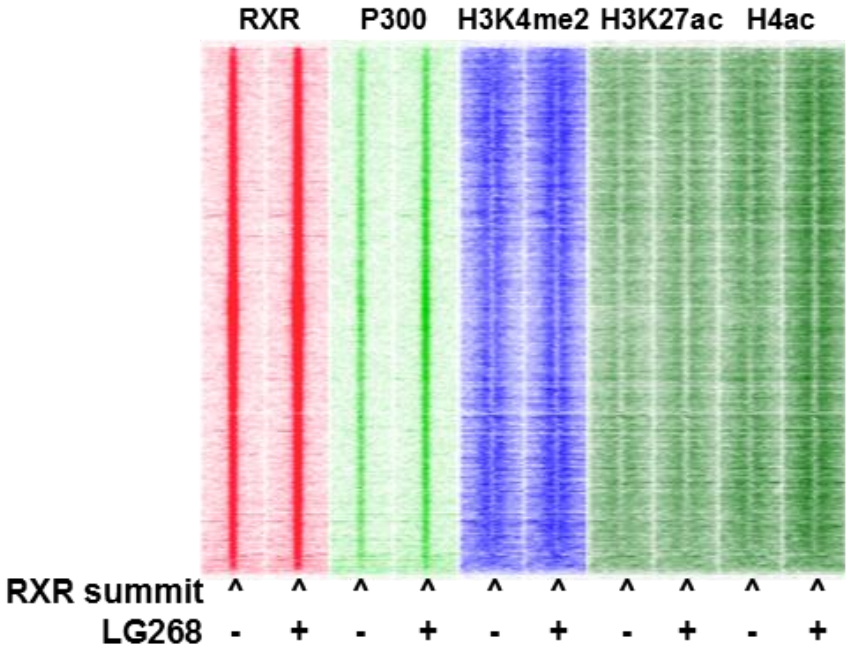

Figure 23. Read distribution plot of RXR, P300, H3K4me2, H3K27ac, and H4ac occupancies in 3-kb windows around the summit of the RXR peaks in the presence or absence of LG268. *Read distribution was determined by HOMER, clustering was done by Gene Cluster 3.0 using centered correlation similarity metric with single linkage clustering method, and heat maps were created by Java TreeView in log2 scale.*

Determination of the immediate early response of liganded RXR on the living genome

As we presented before the merging of the ChIP-seq and RNA-seq datasets proved to be insufficient to reliably identify the network of primary regulated genes and their enhancers. This is due to the fact that the enhancer elements can be located very far from the regulated genes and thus in many cases make the analysis prone to mistakes. Almost 90% of the RXR peaks are located outside of the 10 kb region upstream of the TSS of the closest regulated gene (Figure 20) which is in good correlation with others' observations [48].

We reasoned that in order to link a liganded RXR occupied enhancer to the corresponding regulated gene, one needs a high resolution method, which can provide dynamic temporal information about transcription. The recently recognized fact that

nascent RNA production can be detected both on the regulated gene and on the enhancer element offered a plausible solution [35].

Consequently we decided to measure the immediate early response of RXR at the level of nascent RNA production using Global Run-On Sequencing (GRO-seq). We carried out a time-course experiment to determine the sites and the dynamics of nascent RNA production in the presence of LG268 (Figure 24A, B).

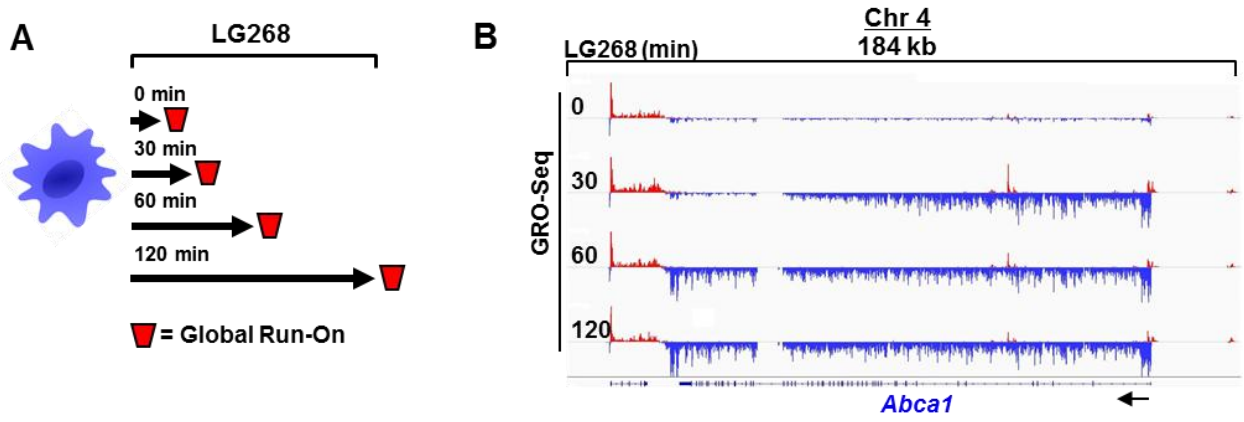

Figure 24. A, The scheme of GRO-seq experiments. *Cells were treated for the indicated periods of time and then nuclei were isolated, and run-on was performed followed by sequencing.*

B, Nascent RNA production measured by GRO-seq. *The induction profile of Abca1 gene is shown. Strand-specific coverage is represented by red and blue. Black arrow shows the direction of transcription.*

Importantly, we observed and verified the presence and induction of nascent RNA transcripts at sites of enhancers and also short divergent transcripts [20]. Similarly to others we call these enhancer RNAs (eRNA). Strikingly, if one aligns the detected GRO-seq activities with cistromic data, it is clear that short transcripts produced from the close proximity of transcription factor bound regions. It is spectacular how PU.1, RXR and P300 co-bound regions transcribe enhancer RNAs (Figure 25).

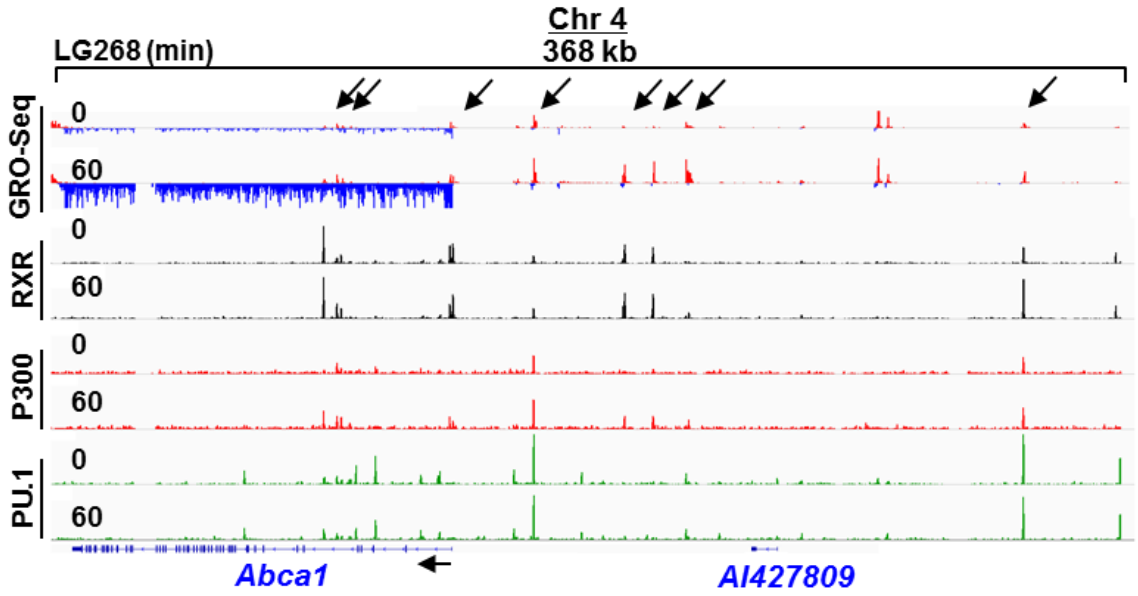

Figure 25. IGV (Integrative genomics viewer) screenshot around the Abca1 locus. *GRO-seq and ChIP-seq against RXR (black), P300 (red) and PU.1 (green) are shown. Black arrows on the top are showing the sites where enhancer transcription induced upon LG268 treatment and overlaps with cistromic data.*

Using this technology we determined changing (induced/repressed) gene transcripts and eRNAs as well. If one plots the changing levels of nascent RNA production of already established direct target genes, the dynamics of the changes is indicative of an immediate induction (Figure 26), showing that this approach also allows easy and reliable identification of new direct targets.

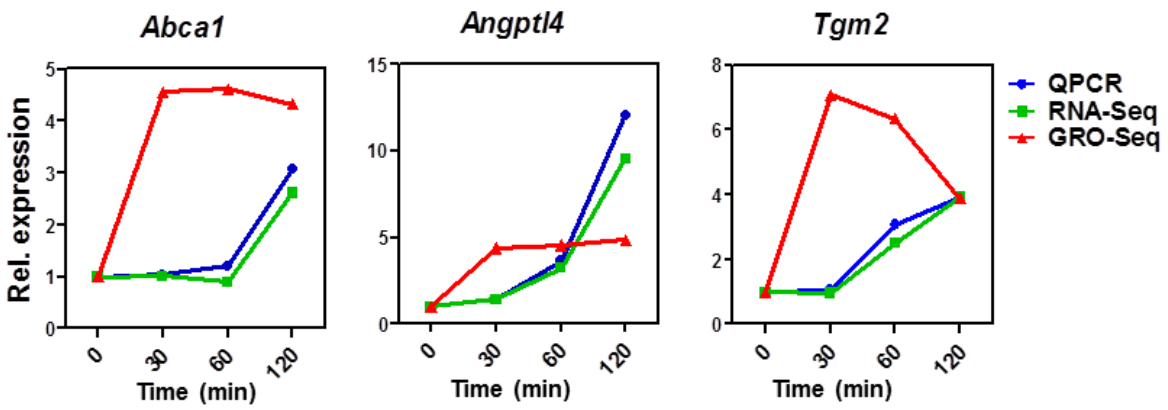

Figure 26. Gene transcription dynamics assayed at the nascent (GRO-seq) and at the steady-state RNA level (RNA-seq, qPCR). *Macrophages were treated with LG268 for the indicated time period, and specific gene expression was detected with the indicated method. A representative set of experiments is shown. Gene expression was normalized to the 0 time point.*

These data further reinforced the notion that the dynamically changing nascent RNA landscape provides clues about direct transcriptional regulation specific for RXR liganding.

As far as the active enhancers are revealed, we classified the identified circa 5,200 RXR binding sites into two categories using an algorithm (described in the methods section): 1. Enhancers that overlap with a GRO-seq positive region (divergent site) (2,781), and 2. Enhancers that do not overlap with GRO-seq activity (2,425). This classification showed an enrichment of RXR peaks within the proximity of the closest TSS (<10 kb from the TSS), suggesting that our assumption was likely correct and we were identifying the functional/ transcriptionally active binding sites and efficiently separating those from the silent/non-active/parking ones. Interestingly, the distribution of GRO-seq positive and negative RXR bound regions relative to the nearest TSS shows a bias towards upstream regions in cases of the actively transcribed ones. Most interestingly, the enriched motif distribution under these peaks also showed characteristic differences. RXR peaks with GRO-seq activity showed an enrichment for DR1 and DR4 (binding sites of RXR:PPAR

and RXR:LXR heterodimers, respectively) as well as AP-1 sites as opposed to the GRO-seq negative sites (Figure 27).

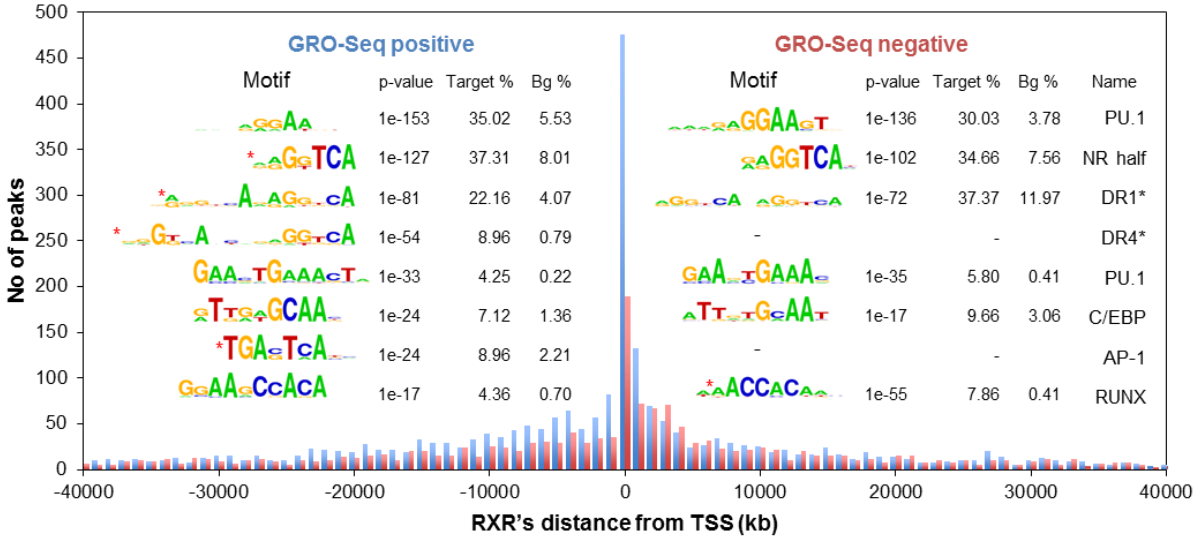

Figure 27. Distribution of GRO-seq positive (blue) and GRO-seq negative (red) RXR peaks compared with the TSSs of the closest expressed transcripts. *Columns represent the peak number of the 1kb distance bins. (Insert) De novo and targeted (black asterisks) identification of motifs under GRO-seq-positive (left side) and GRO-seq-negative (right side) RXR peaks determined using HOMER. Differentially enriched motifs are marked by red asterisks. "Target %" refers to the ratio of the peaks having the given motif, and "Bg %" refers to the ratio of a random background as described in the Materials and methods section.*

Next we further analyzed the GRO-seq positive RXR binding sites and we found that these characterized by an increased H3K27ac, H4ac and also H3K4me2 histone marks if compared to negative ones (Figure 28). These observations supported further evidence to the notion that these sites are indeed functionally distinct and probably represent active enhancers.

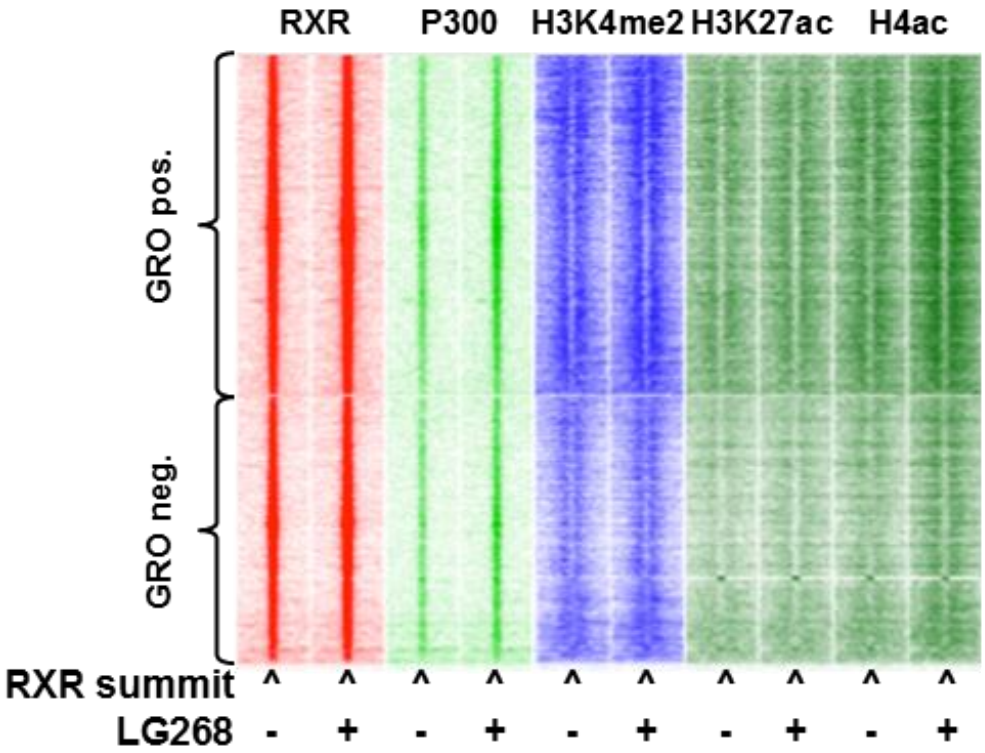

Figure 28. *Read distribution plot of RXR, P300, H3K4me2, H3K27ac, and H4ac occupancies in 3-kb windows around the summit of the GRO-seq-positive and GRO-seq-negative RXR peaks in the presence or absence of LG268. Read distribution was determined by HOMER, clustering was done by Gene Cluster 3.0 using centered correlation similarity metric with single linkage clustering method, and heat maps were created by Java TreeView in log2 scale.*

Next, we matched up all the divergent GRO-seq sites with RXR bound genomic regions. The identified 51,657 divergent GRO-seq sites likely contain all TSSs and eRNAs. We found more than 3,300 of this change upon RXR ligand activation and 718 overlap with RXR binding sites as well. Finally, we identified 252 regulated genes to which we could assign 414 RXR bound genomic regions on which eRNA production increased (387 enhancers) or decreased (27 silencers) using the algorithm developed by us (described in the Materials and methods section (Figure 29). These results show that RXR is predominantly a transcriptional activator rather than repressor.

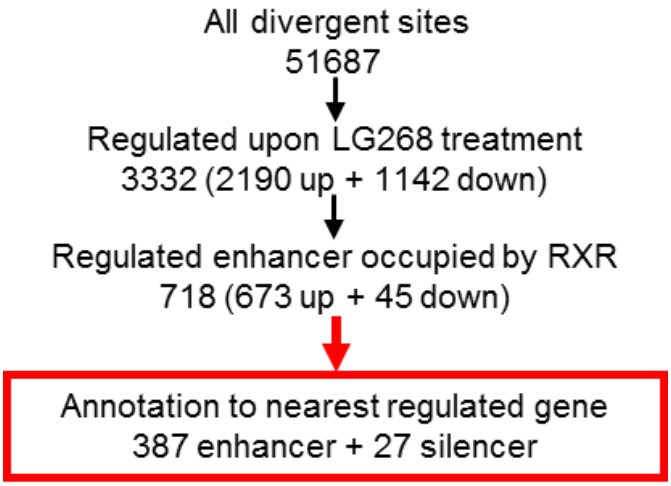

Figure 29. Flow of annotation of GRO-seq positive RXR-binding regions with changing enhancer transcription to the closest regulated gene.

Using the active RXR enhancer network (387) we could show that these genomic regions have the most pronounced increase in RXR occupancy, P300 recruitment and H4ac, whilst H3K4me2 and H3K27ac are not different upon ligand treatment which is in agreement with the reported acetyltransferase selectivity of P300 on H4 histones (Figure 30).

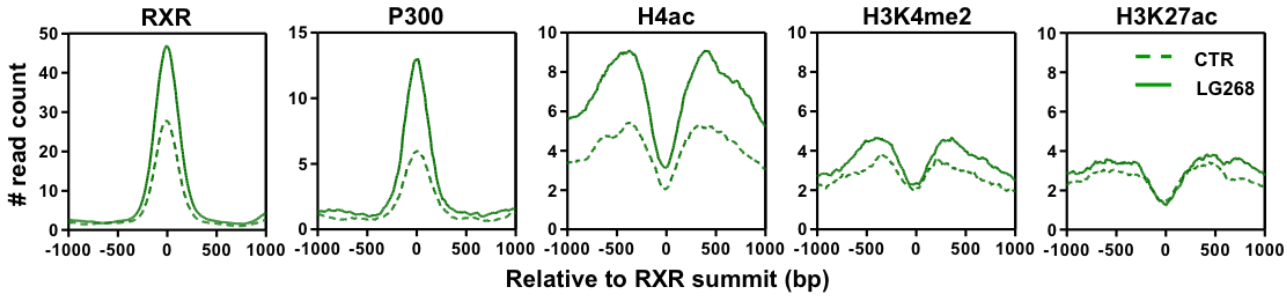

Figure 30. *Histograms representing the Read distribution of the indicated proteins on the identified 387 enhancer elements in the presence (continuous line) or absence (dotted line) of LG268 relative to the RXR summit. # indicates normalized.*

Taken together, we developed a novel algorithm through which we shed light on the core active enhancer network operated by liganded RXR. We could efficiently separate the transcriptionally active sites from non-functional, silent/parking RXR binding sites to which we successfully assign the regulated genes.

Characterization and functional validation of newly identified distant and long-range enhancers

Before we embarked on molecular validation of the enhancers we intended to assess the contribution of the activated partner receptors to the induction of various RXR regulated genes and eRNAs also. Therefore we determined the induction of steady state RNA levels of selected genes. We found that *Vegfa* and *Hbegf* were induced primarily by RXR liganding with some activity by RAR, *Tgm2*, *Ccl6* by both RXR and RAR liganding, *Abcg1* was induced by RXR and LXR, and *Angptl4* by the RXR and PPARγ ligands (Figure 31).

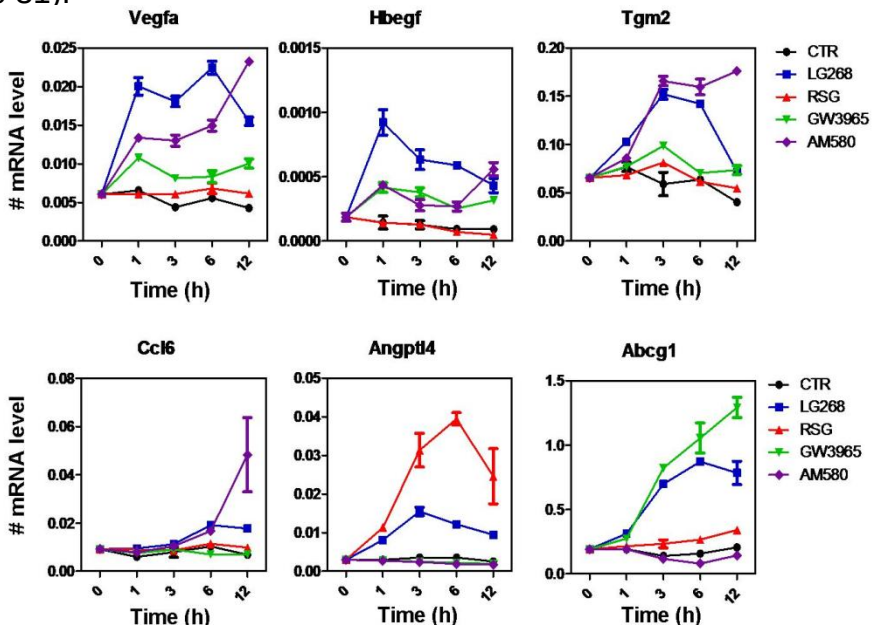

Figure 31. *Comparative induction profile of the indicated RXR target genes in the presence of various nuclear receptor ligands. Cells were treated with 1 μM of the indicated ligand or with 100 nM LG268 for the indicated period of time and specific mRNA levels were determined using qPCR. The mean and +/- SD of triplicate determinations is shown. # indicates normalized.*

Interestingly, primarily RXR induced eRNA production could be detected on an enhancer assigned to *Vegfa* or *Tgm2*, whilst an enhancer of *Abcg1* also showed robust LXR ligand activation as expected (Figure 32). These data suggested that the eRNAs can be easily validated and show ligand induction similarly to the regulated genes which means that the enhancers and their target genes are functionally linked and therefore most likely are part of the same regulatory unit.

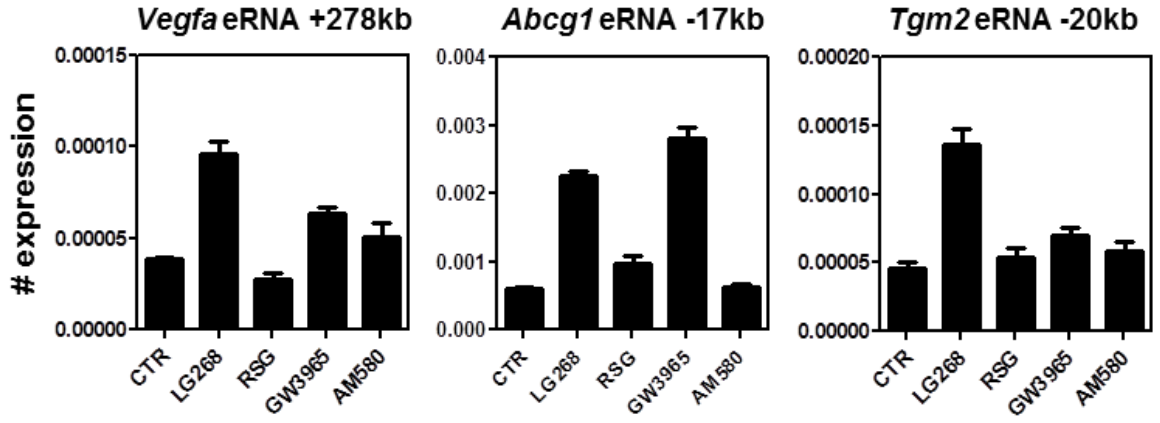

Figure 32. *QPCR analysis of Vegfa, Abcg1, and Tgm2 related eRNAs. Macrophages were treated with LG268, RSG, GW3965, and AM580 for 1 hour, and eRNAs were measured. Data represent mean and +/- SD, with expression normalized to Rplp0.*

For functional validation and characterization we chose 45 newly identified RXR occupied regulatory regions linked to 23 genes. We selected 30 distant enhancer elements (>10 kb from TSS), which earlier studies performing traditional enhancer trap or promoter bashing approaches [129] would likely have missed including one of ours [130]. In addition we aimed to measure the activity of 15 more proximal enhancers (<10 kb from TSS).

According to our aims we inserted 45 putative RXR regulated *cis*-regulatory elements (35 enhancers and 10 silencers) in front of a luciferase reporter gene and measured their transcriptional rates in COS1 (monkey kidney fibroblast cell line) cells in various receptor combinations and ligands. We clustered the enhancers based on the induction patterns and responsibility obtained in the transient transfection based reporter system

(Figure 33). First we determined whether RXR binds to these elements using a VP16-RXR construct. Most reporters showed such binding, but interestingly 12 out of 45 bind the receptor only in the presence of LG268 (Cluster A). These enhancers typically cannot be activated efficiently by other ligands, even if partners are added. Five of the twelve cis-regulatory elements are silencers, which might explain this behavior. Alternatively, these sequences might represent highly macrophage specific regulatory elements. We classified further those enhancers, which bind readily the receptor in the absence of its ligand. Cluster B consists of those regulatory regions, which are activated only if, besides RXR, another receptor is expressed at sufficiently high levels and a suitable ligand (LG268, GW3965, AM580) is present. These enhancers appear to be sensitive to heterodimer expression levels. This cluster has four silencers. Cluster C is representing the most interesting ones, because these are activated selectively by RXR and its ligand, under conditions when other ligands can activate heterodimers on other elements (compare to Cluster D - only RXR). However if the partners are overexpressed, they also responded to PPARγ, LXR, or RAR activation, showing sensitivity to receptor levels. The remaining two clusters (Cluster D1, D2) can be activated by RXR and LXR heterodimers, while transfecting RAR makes cluster D2 responsive for AM580 as well.

These experiments supported evidence that the identified and tested enhancers are RXR-specific enhancers, but also revealed the complex and very versatile activation pattern allowing context-dependent regulation of the genes by distinct dimers.

Based on these results it seems that the main driving force can be the expression profiles of the receptors and probably also the presence of the lineage-specific transcription factors along with selectively bound co-factors. To sum up, probably each single gene has its own enhancer repertoire depending on the cell type which is marked by the lineage determining transcription factor. Once the enhancers are ready to bind the signal specific transcription factors, their spatiotemporal character will determine which signals exert the final effect on the actual element leading to gene regulation.

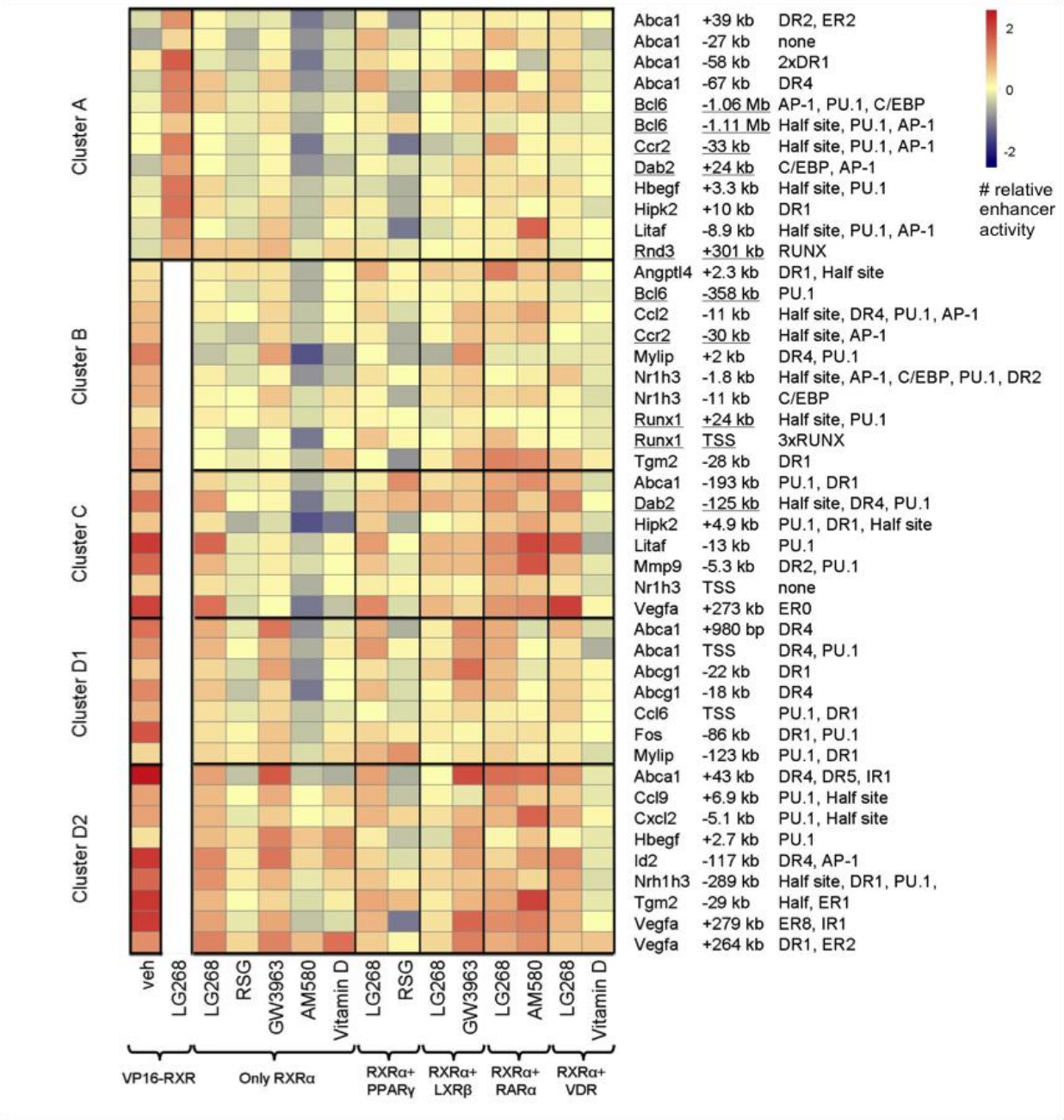

Figure 33. *Heat map representation of the activity of 45 cis-regulatory elements in luciferase reporter assays in the presence of the indicated ligands in COS-1 cells. Heat map intensities represent normalized luciferase values relative to control measurements. Silencers are marked by underlining.*

RXR regulated gene expression in the context of higher order chromatin structure

Once we determined the RXR regulated genes and linked, at least a subset of them, to their enhancers (some of them are long-range ones) we wanted to get some insight if these regulatory units are confined to known features of genome architecture.

Therefore we decided to explore the genome architectural context of RXR signaling asking the question if RXR signaling is confined by known structural features of genome organization and/or contributes to reorganization of these. We determined the cistromes of CTCF and a member of the cohesin complex (RAD21) in the absence or presence of ligand activated RXR. We have found around 30,000 peaks for CTCF and around 24,500 for RAD21, the latter showed some increase upon ligand activation. Next we determined an overlap of around 12,660 peaks representing similarly high CTCF and RAD21 occupancy (with peaks having a MACS2 score higher than 15 and having less than 3-fold difference). We considered these as boundaries of functional domains as suggested by others [64] [131]. We identified 10,204 such functional domains by pairing the CTCF and RAD21 co-peaks with the closest ones with similar scores.

Next we validated molecularly the interactions in such functional domains and the impact of ligand on these for long-range enhancers of *Abcg1*, *Vegfa* and *Tgm2* using qPCR based 3C (Figure 34). These data clearly documented that the enhancers identified by the combination of RNA-, ChIP- and GRO-seq act in functional domains, loop to the promoter and can be readily functionally validated using transient transfection and 3C as well.

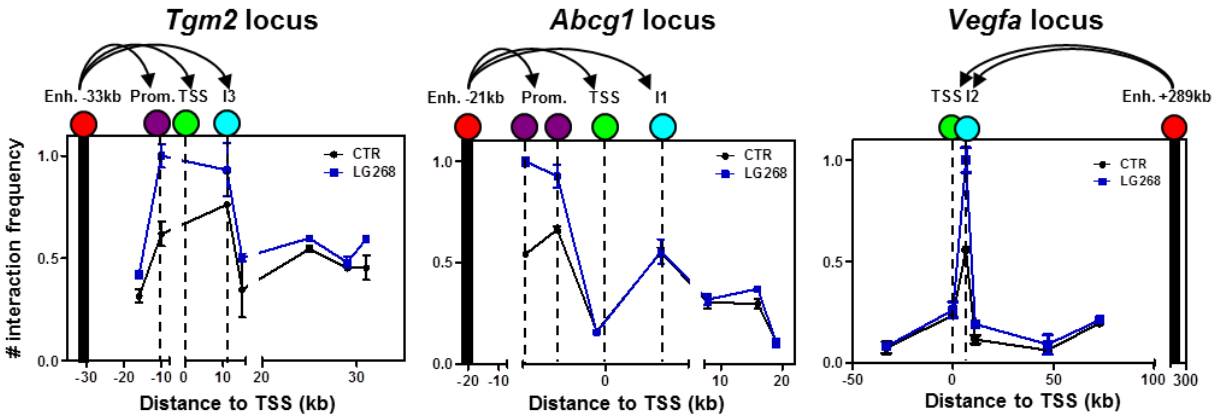

Figure 34. *3C-qPCR measurements in the presence or absence of LG268 on the Tgm2, Abcg1, and Vegfa loci. Constant primers giving the anchor points of the analysis were designed to the enhancers and are represented by black columns. Red circles indicate the enhancers, purple ones are the promoter regions, TSSs are depicted with green, and intronic regions are highlighted by cyan circles. The mean and +/- SD of triplicate determinations are shown.*

Finally we asked the question if any of these enhancers communicate with other cis-regulatory elements in the genome. Thus we carried out 3C-seq using pairs of baits located in or close to these regions. We could detect proximal and also long-range interactions. At the *Vegfa* locus we could detect interactions between the distant enhancer and the neighboring enhancers as well as the intronic region with remarkable specificity (Figure 35). Similar results were obtained in the case of the *Tgm2* and *Abcg1* loci.

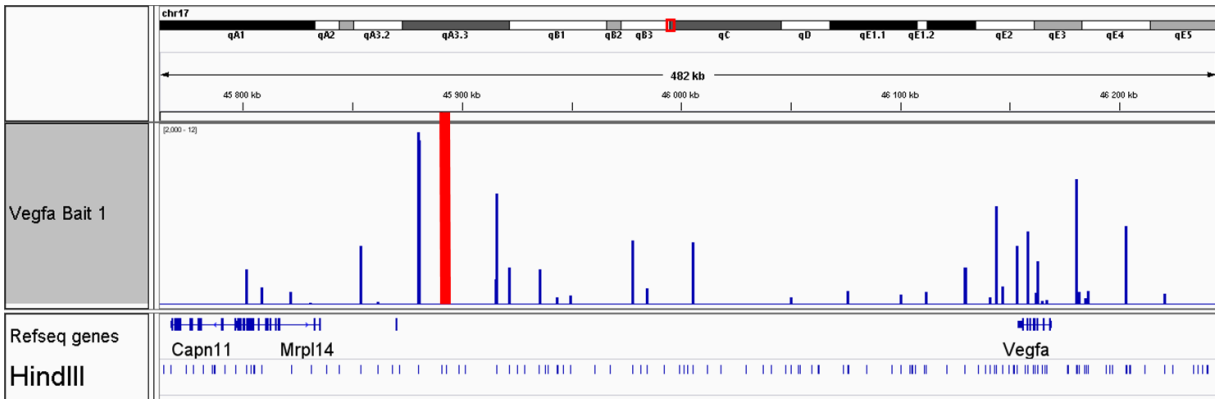

Figure 35. *IGV representation of 3C-seq interactions on the Vegfa locus. Bait sequence (starting point of the analysis) is highlighted with red column. HindIII cutting sites are also marked with blue columns at the bottom.*

Importantly, we also detected interchromosomal interactions with much less frequency (at least 50 fold less) though. In order to provide statistical context to these findings we compared the interaction frequency of given bait with the previously predicted functional domains in which we analyzed the RXR regulated enhancer gene pairs. Therefore we first separated the transcriptionally inactive domains from those that are RXR signaling negative but transcriptionally, on the other hand we determined also the active domains with RXR-regulated regions. With the help of this analysis we successfully show that there is significant difference between the interaction frequencies of interchromosomal interactions among functional and non-functional chromosomal territories in the context of RXR signaling regarding the *Vegfa* enhancer (Figure 36). Similar results were obtained with additional enhancers of the *Abcg1* and the *Tgm2* gene..

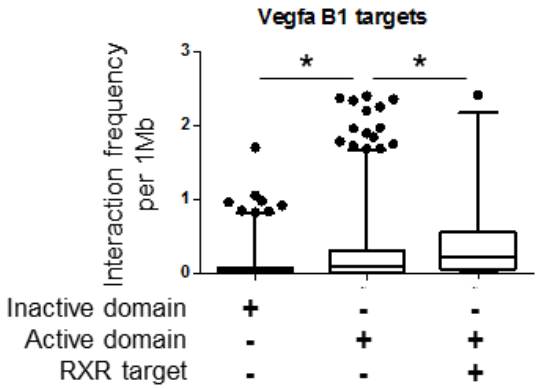

Figure 36. *Box plot representation of the distribution of interaction frequency of Vegfa B1 enhancer region assayed by 3C-seq. The detected interactions were mapped onto 1-Mb fragments covering the mouse genome. Interaction frequency was determined by expressing the read number per 1 Mb normalized to 1000 reads. Asterisks represent significant difference at P < 0.0001.*

This suggests that active RXR enhancers interact with other active genomic regions and with an even higher likelihood with other RXR regulated ones. These formally suggest that the active RXR enhancers form an interchromosomal hub or network.

Determination of the angiogenic capacity of RXR programed macrophages

Finally we wanted to see if some of the identified novel transcriptional pathways could be validated biologically. The functional annotation of the genes controlled by liganded-RXR enhancers assigned a large number of genes into the angiogenesis category (Figure 37). These include *Vegfa*, *Hbegf*, *Litaf*, *and Hipk2*. Therefore we decided to functionally test this activity.

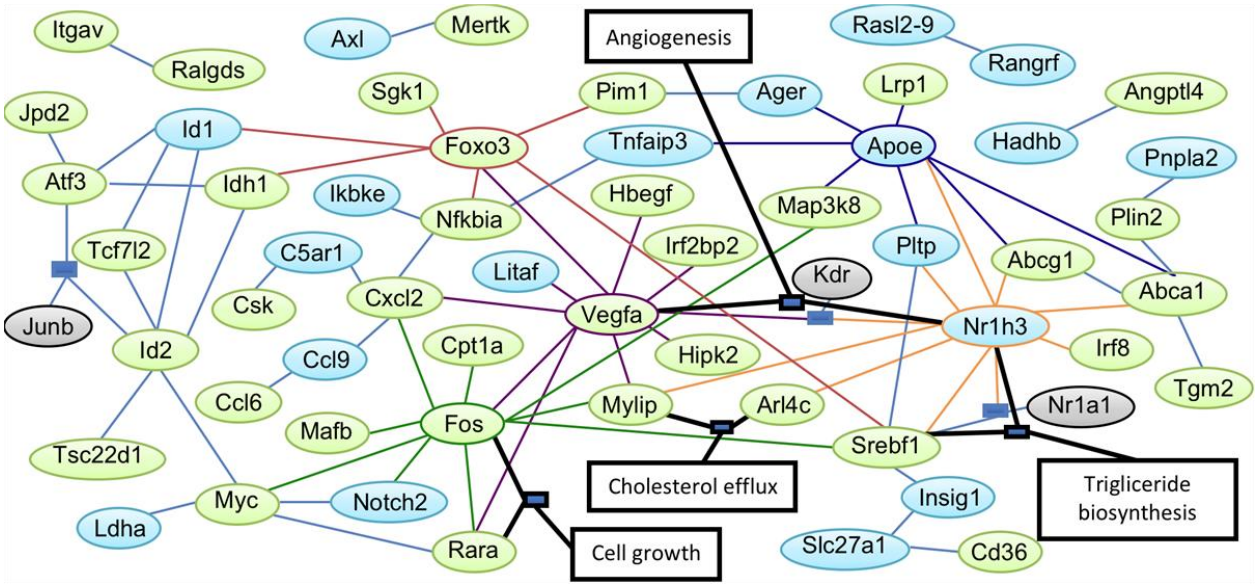

Figure 37. *Interaction network showing the significantly regulated genes upon ligand treatment. Genes which are significantly regulated only in GRO-seq (cyan), only RNA-seq (grey) or both (green) are shown.*

First we determined the excreted VEGFα protein levels from cell supernatants (Figure 38 A) using different ligands and could show that RXR induced this protein as was suggested by our gene expression measurements as well (Figure 38 B). Using qPCRs we could show the RXR requirement using RXRα/β double KO macrophages (Figure 38 B).

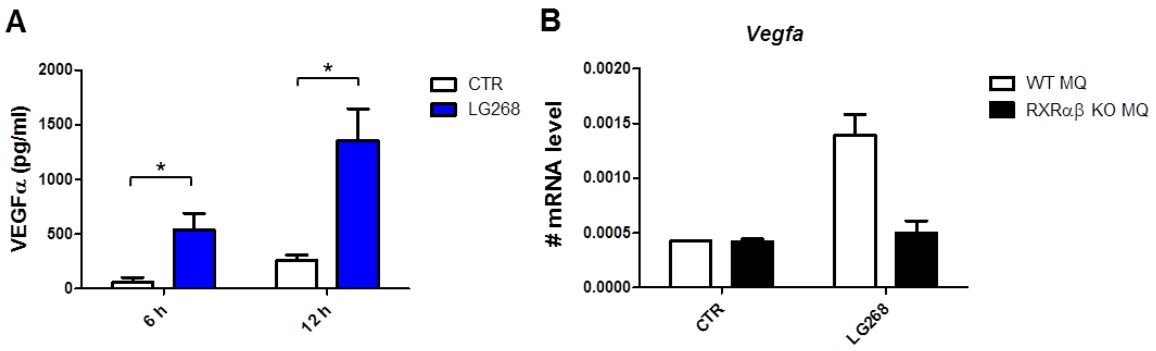

Figure 38. RXR programed macrophages are pro-angiogenic.

A, VEGFa protein levels determined from supernatants of macrophages treated with LG268 for the indicated periods of time by ELISA. The mean and +/- SD of triplicate determinations are shown. Asterisk represents significant difference at P < 0.05; n = 3. B, Vegfa mRNA levels determined using qPCR in wild-type (WT) and RXRα/β knockout (KO) macrophages. The mean and +/- SD of duplicate determinations are shown.

To test if activation of RXR strengthens angiogenesis in an *in vivo* relevant setting we carried out a ChorioAllantoic Membrane (CAM) assay using LG268 pretreated macrophages. Importantly, these cells show significantly improved vessel formation, which was not detectable in RXRα/β double KO macrophages (Figure 39).

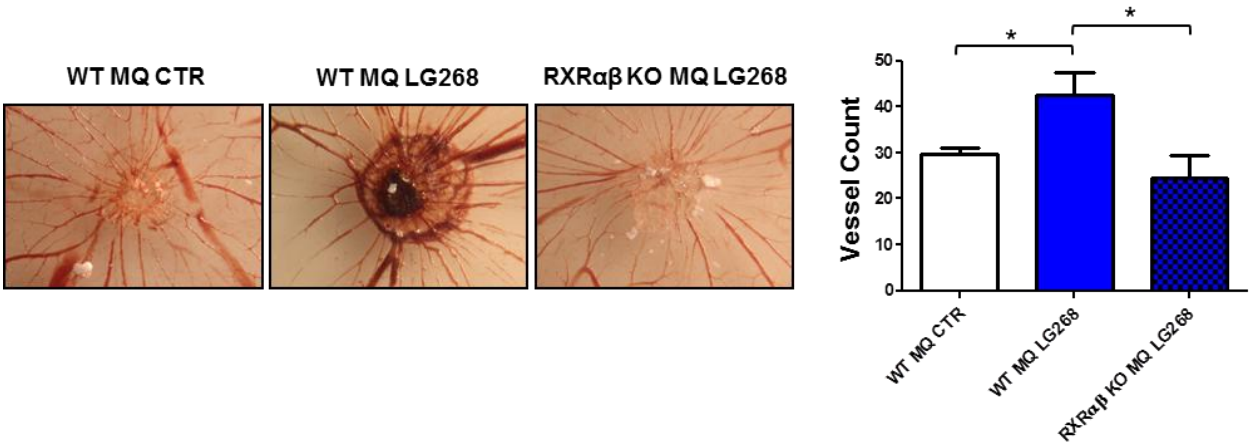

Figure 39. *Angiogenic activity of macrophages was determined using CAM assay in the presence or absence of LG268 using wild-type and RXRα/β knockout macrophages. Representative images (left) and bar chart of quantification (right) are shown. The mean and +/- SD of three biological replicates are shown. Asterisk represents significant difference at P < 0.05; n = 3.*

These data suggest that RXR activation can program macrophages toward a distinct cell autonomous angiogenic phenotype.

The effect of the IL-4/STAT6 signaling pathway on the RXR cistrome during alternative macrophage activation

Our preliminary results show that the RXR binding pattern in alternatively activated macrophages is greatly impacted. We could show that more than 4000 RXR binding regions are significantly more occupied in response to IL-4 mediated differentiation (Figure 40). It is very interesting, especially if we are considering the fact that the RXR specific ligand could induce the receptor occupancy at around 800 genomic regions only.

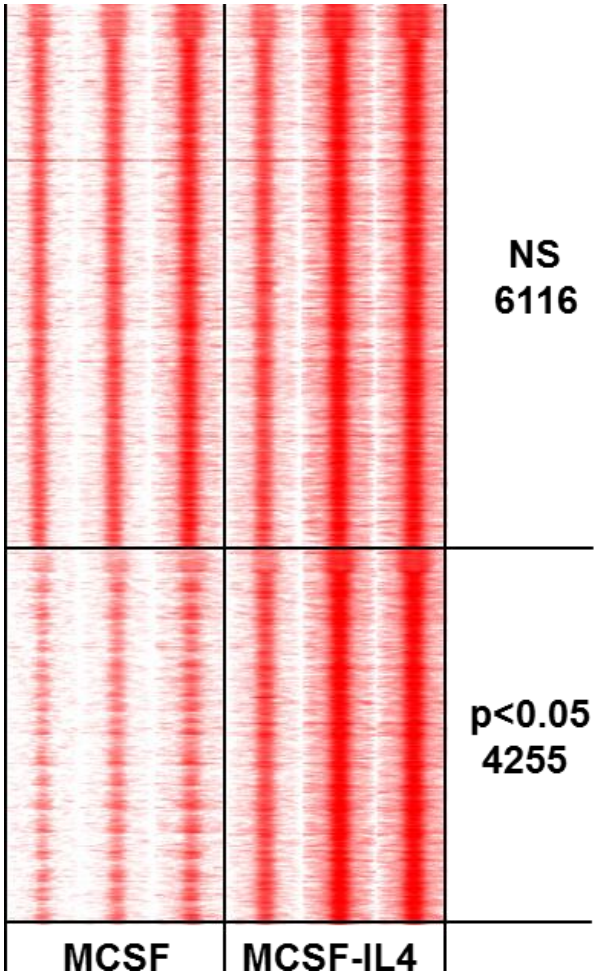

Figure 40. *Heat map representation of RXR occupancy in 3 kb windows around the summit of the RXR peaks in MCSF/IL4 differentiated macrophages from three biological replicates. NS means non-significant difference. Read distribution was determined by Homer, clustering was done by Gene Cluster 3.0 using Centered Correlation Similarity Metric with Single Linkage Clustering Method, and heat maps were created by Java TreeView in log2 scale.*

As a next step we chose those RXR binding sites which showed the biggest change in response to IL-4 differentiation compared to those that are differentiated in the absence of IL-4. We carried out *de novo* motif analysis on these which revealed that under these binding regions DR1 (capable of bind the PPARγ/RXR and possibly RAR/RXR

heterodimers), AP-1 and EGR2 motifs are more enriched compared to their IL-4 absent counterparts (Figure 41).

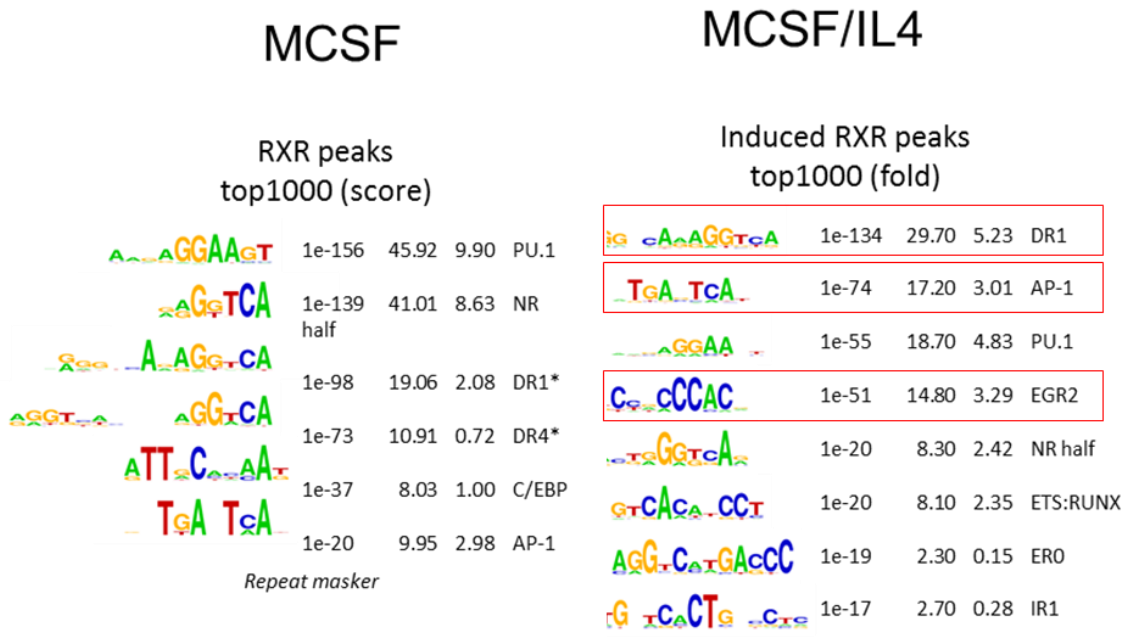

Figure 41. *De novo identification of motifs under RXR peaks from ChIP-seq data obtained from MCSF and MCSF/IL4 differentiated macrophages using HOMER. Peaks showing the highest scores (MCSF) and the highest fold induction (MCSF/IL4) were used in the analysis. Target % refers to the ratio of the peaks having the given motif, and Bg % refers to the ratio of a random background.*

Interestingly, we could show that the genes (*Rara, Pparg, Egr2, Batf3* member of the AP-1 complex) encoding the proteins that are capable of bind to these elements are highly induced at the mRNA level in response to IL-4 in a STAT6 dependent manner (Figure 42). These results suggest that these factors might be crucial for the altered RXR cistrome and dynamically contribute to the rearrangement of the receptor.

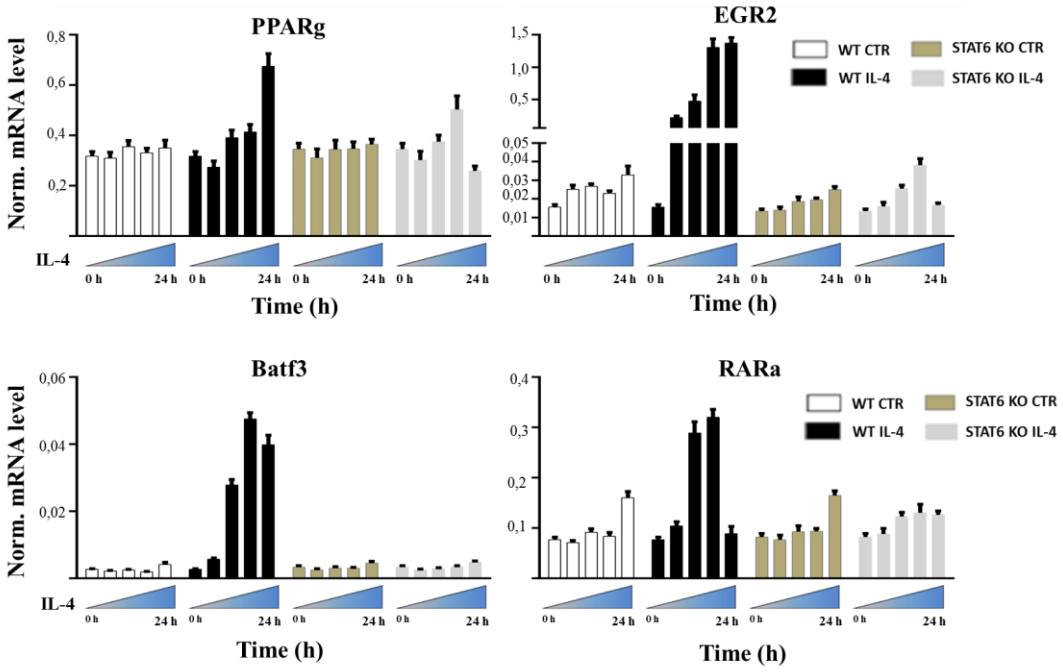

Figure 42. *Gene expression profile of Pparg, Egr2, Batf3 and Rara in the presence of IL-4 from wild type and STAT6 KO macrophages. The mean and +/- SD of three replicates are shown.*

Surprisingly, we could also show that *Vegfa* follows different gene expression pattern in response to RXR activation in IL-4 differentiated macrophages which leads to a more robust *Vegfa* expression and importantly secretion also, measured by ELISA (Figure 43). These data indicate that RXR signaling is impacted by IL-4 and probably leads to an enhanced pro-angiogeinc phenotype in macrophages.

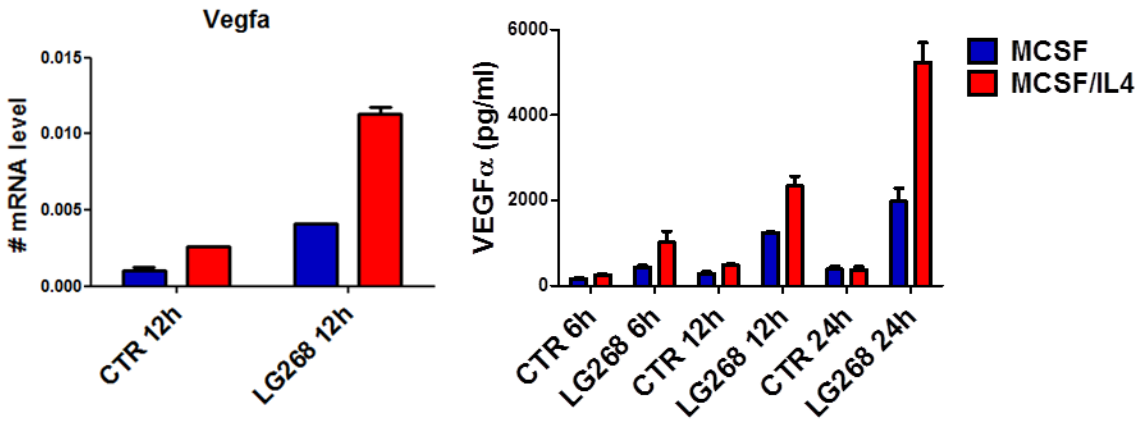

Figure 43. *Gene expression profile of Vegfa in MCSF and MCSF/IL4 differentiated macrophages in the presence of the RXR agonist (LG268) (left) treated with LG268 for 12 hours. ELISA assay directed against VEGFA in MCSF and MCSF/IL4 differentiated macrophages treated with LG268 for the indicated period of time (right). The mean and +/- SD of three replicates are shown.*

Discussion

The fundamental aim of studying the function of a signal specific transcription factor is to identify its primary biological action in the genome of a given cell type. One way to do this can be the determination of the cistrome and pairing the binding sites with the regulated genes. Using in silico methods one can identify the potential RXR binding sites on the scale of several x 100,000 in the mouse genome. Combining all the RXR ChIP-seq data available from various cell types or tissues like fat, liver and now in macrophages the joined number of genomic regions bound by the receptor is a few x 10,000. Still in macrophages the number of genomic regions is around 5200, but the number of regulated genes by liganding the receptor is a few x 100. This is even more complicated if we are considering the fact that primary targets are difficult to reliably identify by simply measuring steady state mRNA levels.

The question is given: How can one find the immediate early responder genes and the related RXR bound enhancers?

Integration based genomics identifies the active RXR enhancer landscape

The most important part of this work is that genome-wide ChIP-seq experiments can be used in combination with data from GRO-seq revealing nascent RNA levels. These data together hold the potential to identify not only the primarily regulated genes but also the relevant enhancer elements. This relies on the observation that *in vitro* (nuclear run-on) determination of the activity of RNA polymerases (RNAPs) is a reliable indicator of transcriptional activation. The usage of GRO-seq for the determination of direct transcriptional responses has two major advantages: (1) the dynamics of the nascent RNA levels depends only on the rate of RNAPs activity and therefore it is matching the expected time course of a directly regulated gene. Our data is clearly demonstrating this, because known and established direct target genes (*Abca1, Abcg1, Angptl4, Tgm2*) as well as newly identified ones (i.e.: *Vegfa, Hbegf*) show an immediate induction with GRO-seq rather than a complex, partly delayed (due to the RNA maturation processes) one with RNA-seq or qPCR. (2) Enhancers occupied by engaged RNAPs and their

activity can be registered by GRO-seq. These are the so-called enhancer RNAs characterized as a typically few hundred base pair long divergent transcript. The existence of such transcripts has been shown before by RNA-seq and later on also by GRO-seq [34], [35]. Despite the fact that their mechanism of production and role in transcription regulation is far from being understood, they represent active enhancers. Additionally, if the level of nascent RNA production changes (increase or decrease) at these sites and this is in accordance with the stimulating signal, it is very likely to be the consequence of the activation of the signal specific transcription factor acts on it. Today's technological limitation is the lack of high throughput validation of the identified enhancers. However, our high rate of success in validating such enhancers with low throughput enhancer trap approaches and 3C suggests that our integration based enhancer annotation method can be used to filter active enhancers for a given signal dependent transcription factor from the general pool of cis-regulatory regions. Importantly, this also means that the combination of these methodologies can serve as *proof-of-concept* and a model to solve similar problems with other signal dependent transcription factors. Our approach has a remarkable feature, namely that it allows the recognition of active enhancers in spite of the fact that many of them are very long-range enhancers with high probability and accuracy. It is also interesting that the enhancers' distribution relative to the TSS appears to be symmetrical instead of being enriched in the upstream regions suggesting that one needs to look in both directions and also far away to identify the true enhancer controlling a particular gene in a given cellular context, however the transcriptionally active ones are shifted toward the upstream regions. A cautionary conclusion of this study is that it is very likely that many of the enhancers identified with less comprehensive methods need to be revisited.

RXR binds together with the myeloid specific PU.1 and recruits P300 as a cofactor to its binding sites

Our results provided significant insights into the transcriptional regulation of liganded RXR. We showed that RXR has approximately 5,200 binding regions in macrophages. Due to the fact that RXRα is the dominant isotype in macrophages, these genomic

regions are most probably bound by RXRα, though the used antibody is pan-RXR specific and would recognize all three subtypes. Not surprisingly most of these sites contain nuclear receptor half sites (AGGTCA) suggesting that RXR mainly bind to the chromatin through direct DNA binding in a nuclear receptor motif dependent manner. In addition, two 18 independent lines of evidence suggest that there is an intimate relationship between the lineage determining factor PU.1 and RXR: (1.) 45 % of the RXR peaks contain PU.1 sites and (2.) there is a significant overlap between RXR and PU.1 cistromes, as more than 2/3 of RXR binding regions overlap with PU.1 peaks. This is in agreement with the literature, where it has been reported that binding regions for PPARγ, a heterodimerization partner of RXR, colocalizes with PU.1 [132], [133]. This finding further support the pioneering or bookmarking factor theory which says that cell type specific transcription factors shape the cis-regulatory element landscape during differentiation, thus contribute to the proper functioning of other signal specific transcription factors [134]. Our results are fully compatible with this concept and interestingly show that there is only 14 % overlap between our RXR binding regions and the ones found in 3T3L1 cells (Adipo-L1) differentiating into adipocytes [135] and even less if we compared them to the binding regions of the receptor in liver (Figure 44) [136].

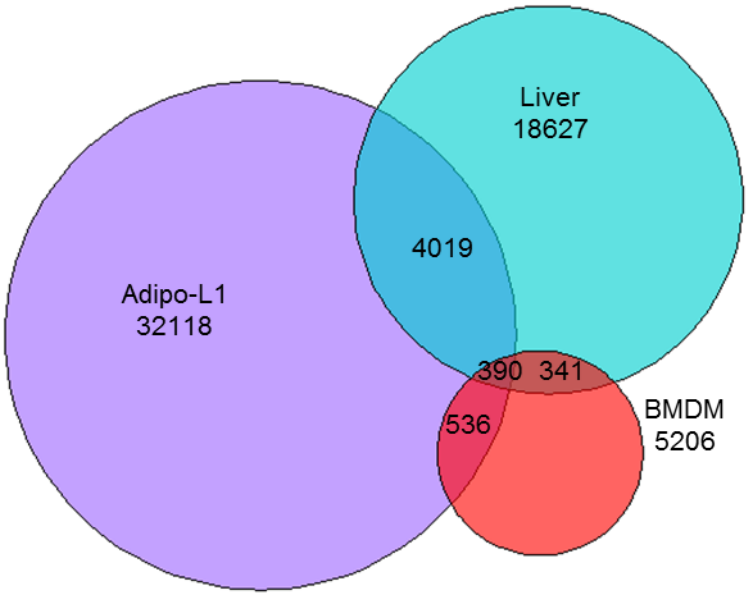

Figure 44. *Overlapping RXR binding regions of 3T3-L1, liver and our BMDM cells represented as Venn diagram.*

Our work suggests that many if not all RXR heterodimers, colocalize with PU.1 in macrophages and that upon ligand treatment, PU.1 is at least partially released and therefore might not be required for further binding or transcriptional activity. This data might also implicate that PU.1 dependent RXR recruitment to the genome is not likely to be a major mechanism in macrophages, although it might contribute to the establishment of the RXR landscape to a small degree. Another unexpected aspect of the cistromic interactions is that the P300 coactivator is dynamically recruited to the RXR sites upon ligand treatment. Importantly, P300 occupancy is increasing in a more pronounced way on GRO-seq positive sites compared to the transcriptionally inactive ones. The main conclusions which can be drawn from these findings is that P300 seems to be the major coactivator through which activated RXR modifies the surrounding chromatin structure and activates gene expression. This is accordance with the initially proposed role for this protein in nuclear receptor signaling [137] and with the more recent findings [138], but also suggests that P300 is likely to be specific for this signaling pathway in macrophages. The key question regarding to signal specific transcription factors is whether the activating stimulus contributes to the recruitment of rearrangement of the transcription factor of interest. On the nuclear receptor field the most popular view regarding RXR localization is that the receptor binds to the genome both in the absence and presence of ligand [136]. Our findings are compatible with the conventional view and now by adding genome-wide perspective show that the only effect of the RXR ligand is that the enrichment on the pre-formed binding sites increases. These data obviously has to be interpreted within the context of the experiment, which in chromatin immunoprecipitation meaning that the sum of all binding is detected in all cells in the timeframe of the 40 minutes crosslinking used. Therefore, our conclusion is that RXRs genome-wide localization is determined by its own DNA binding capacity, which allows it to find preformed sites in the genome where it can bind to and/or this is facilitated by additional factors such as a lineage specific factor i.e. PU.1. The explanation for increased occupancy that is observed is either the reflection of more cells being involved in the response or a higher affinity for the binding site. However, the experimental

approaches used are not able to determine the relative contribution of these two mechanisms.

Characteristics of the liganded RXR operated enhancer network

Our data show, as expected, that the main permissive heterodimers present in bone marrow-derived macrophages are RXR:PPAR and RXR:LXR as shown by the induction profile of their established target genes *Angptl4, Abca1* and *Abcg1* and the motifs revealed under their RXR peaks are RXR:PPAR specific (DR1) and RXR:LXR specific ones (DR4). Surprisingly, our experiments also shed light on a set of primarily "RXR-only" regulated genes including *Tgm2* and novel ones, such as *Vegfa* and *Hbegf*, which could be induced in a more RXR selective way. However, ligands for RAR:RXR, or LXR:RXR heterodimers have also shown to affect the expression of these but unequivocally to a lesser extent. These results suggest that an RXR selective activation program is an inner character of the enhancer network regulating these genes but also propose that permissive RAR:RXR or a combinatorial enhancer network is implicated in the expression of *Tgm2, Vegfa* and *Hbegf*.

This, along with the facts that using our most stringent criteria we identified only 200 regulated genes and only 10% of the identified binding sites as active enhancers suggest that activation of RXR is tightly controlled in macrophages and leads to distinct, selective and not to pleiotropic gene expression. This was further supported by the identification of an entire gene expression module supporting angiogenesis. Unexpectedly, a network of enhancers could be linked to a set of genes including *Vegfa, Hbegf, Hipk2, Litaf, Ccl2, Cxcl2* and *Foxo3,* which have been clearly linked to angiogenesis, was also revealed. Several lines of evidence suggest that *Vegfa, Hbegf, Hipk2, Litaf, Cxcl2* and *Foxo3* have important roles in angiogenesis [139], [140], [141], [142], [143], [144]. It has been revealed that CXCL2 dose-dependently promotes angiogenesis and tumor growth of CXCR-2 expressing colon cancer cells, most probably by overexpression of VEGF [139]. HB-EGF is also a potent inducer of tumor growth and angiogenesis [140]. It has been also shown that LITAF plays key roles in

cytokine (CCL2, TNF) and *Vegfa* regulation in complex with STAT6/B [142]. Regarding CCL2 it is confirmed that different kinds of tumor cells along with TAMs release this cytokine to promote tumor angiogenesis, tumor growth through monocyte recruitment and TAM differentiation. VEGFa is almost the best known and the most potent angiogenic stimulator produced by tumor cells, macrophages and T-cells in many human tumors. It has been shown that VEGF is also chemotactic for monocytes via the VEGFR. There are several studies to show that VEGFa expression level is in good correlation with macrophage content of several tumors. Based on these pieces of evidence macrophages appear to be very important for metastatic capacity and survival of tumor cells [145] and our results having in mind with all of these information suggest that RXR programed macrophages are pro-angiogenic.

The regulation of *Vegfa* by RXR is very interesting, because it uses a set of very long-range enhancers (270 kb downstream) and shows RXR specificity, including RXR ligand induced looping. However, RARa activation could also induce the expression of the gene and also the enhancer transcript. Importantly, considering the permissive heterodimeric receptors only LXR signaling seems to have a small effect on the expression of Vegfa as reported by the Tontonoz group in peritoneal macrophages [146]. The presence of CTCF sites outside of this regulatory unit and the involvement of the Mediator complex (MED12) further support the role of this novel long-range enhancer in the RXR-regulated expression of *Vegfa*. A similar regulatory mechanism was uncovered in the case of *Tgm2*.

We presented here a novel integrated approach to identify enhancers with very high probability and link those to primarily regulated genes for a signal dependent transcriptional factor, liganded RXR. This integrated approach revealed that ligand stimulation of RXR activates only a small fraction of the DNA bound molecules and leads to a distinct gene expression programs likely including angiogenesis.

RXR programed macrophages are pro-angiogenic

Myeloid cells are key components in tumor progression. In particular, macrophages are known to be the most abundantly infiltrating immune cell type to the sites of tumors.

These cells are termed Tumor Associated Macrophages (TAMs). TAMs are originated from monocytes, but there is evidence to demonstrate the local proliferation of these cells in the tumor microenvironment. Chemokines are able to dictate the recruitment of immune cells to the sites of inflammation, development, hemostasis and also tumorigenesis [145].

According to the results presented above, it is reasonable to investigate the RXR programed macrophages in an *in vivo* relevant model system. Considering the fact that RXR induces an angiogenesis related gene panel we wanted to clarify whether the appearance of these might shift the phenotype of macrophages toward angiogenesis. The ChorioAllantoic Membrane assay revealed their facilitated angiogenic effect on the vascularization of the CAM. Interestingly, using ligands for the key permissive partner (LXR, PPARγ) could not induce significantly the angiogenic capacity of macrophages, not even the RARα agonist which was shown to activate *Vegfa*. Importantly, in the absence of the receptor macrophages could not induce the formation of new vessels. These results evidenced that activating RXR by LG268 affects the expression of several angiogenesis related genes and this will lead to improved vascularization in a receptor dependent manner. Strikingly, the effect of RXR activation on agiogenesis cannot be recapitulated by the activation of the most relevant permissive and non-permissive heterodmieric partners. Altogether, RXR is responsible for the regulation of a unique gene set in macrophages which resulted in angiogenesis.

Alternatively activated macrophages might possess an enhanced pro-angiogenic phenotype

Several studies showed that IL-4 differentiated, alternatively activated macrophages are crucial players in tumor development and regenerative processes in our body [147]. Our results showed that RXR activation programs macrophages toward the direction of angiogenesis. We also demonstrated that liganding the receptor does not lead to the rearrangement of the receptor meaning that RXR sits on “ready to use” enhancers poised for activation. It is intriguing that upon IL-4 differentiation the RXR cistrome undergoes through enormous changes. At least 4000 new sites are established by the IL-4 triggered molecular mechanisms. Surprisingly, the enriched motifs under the newly

appearing binding regions are not STAT6 related sequences. The new binding regions are characterized by the presence of DR1, AP-1 and EGR2 motifs. Measuring their expression revealed that IL-4 directly regulates the *Pparg* (capable of bind the DR1 element), *Rara* (might be capable of bind the DR1), *Batf3* (capable of bind the AP-1 motif) and *Egr2* genes in a STAT6 dependent manner. The main conclusion that can be drawn from these is that STAT6 induces the expression of various transcription factors which then facilitate the binding of RXR either to newly established latent enhancers or ready to use enhancers.

Another striking result is that in IL-4 differentiated macrophages the basal expression of *Vegfa* is higher, furthermore the activation of RXR triggers a more pronounced induction, which could be detected also at the protein level. How can IL-4 mediate such response? Probably we should search for the answer in the changed RXR cistrome. Considering the fact that STAT6 induces the expression of several transcription factors at the mRNA level for which the motifs can be found under the newly developed RXR sites, it is conceivable that these will lead to the formation of new enhancer elements in the proximity of *Vegfa*, thus leading to improved gene expression.

Taken together, these preliminary data show that a signal specific transcription factor which is dynamically recruited to the genome upon IL-4 stimulation can act as a redistribution signal to another signal specific one, RXR, which is sitting on the genome on its preformed sites. Importantly, this redistribution might lead to an enhanced pro-angiogenic phenotype, which might be in the background of the well-recognized angiogenic potential of alternatively activated macrophages.

Summary

In our study we focused on the action of the activated RXR nuclear receptor in bone marrow-derived macrophages. Taking advantage of the technological developments we determined the action of the receptor in a genome-wide manner.

First, we identified the regulated gene network by using RNA-sequencing, then we performed chromatin immunoprecipitation followed by deep sequencing in order to find the receptor bound genomic loci. Using these methods we found approximately 800 regulated genes and 5200 binding sites for the receptor. Interestingly, almost 90% of the RXR sites are located to intergenic regions, often more than 100kb far from the closest genes thus their annotation to the regulated genes remained elusive. By using global run-on sequencing we analyzed the amount of the freshly synthesized nascent RNA on gene bodies and as reported previously also the transcripts originated from cis-elements known as enhancer RNAs. Based on the fact that enhancer RNAs are induced in a very similar manner as their targeted genes we assigned them to their potential targets, yielding the transcriptionally active RXR cis-element/gene pair network. The combination of the aforementioned methods let us to identify 252 target genes to which we could assign 414 cis-acting RXR elements. To our surprise, these genes showed a convergent angiogenic network consisting well-known agiogenesis inducers, for example: *Vegfa*, *Hbegf*, *Litaf*. In order to characterize the RXR programed macrophages in the context of angiogenesis we carried out the well-accepted chorioallantoic membrane assay, which unequivocally shed light on the improved angiogenic capacity of macrophages harboring the activated RXR program. Taken together, these results show that the RXR regulated gene/cis-element network promotes the establishment of a pro-angiogenic macrophage phenotype.

Our highly integrated method is capable of identifying the regulated genes and their cis-element network for any kind of signal specific transcription factor thus able to reveal new biological functions in a transcription factor and cell type dependent manner.

Összefoglalás

A kísérleteink során az aktivált RXR magreceptor hatásának molekuláris mechanizmusát és funkcióját vizsgálatunk egér csontvelői eredetű makrofágokban. Kihasználva a technológiai fejlődés adta metodikai újításokat, teljes genom szinten határoztuk meg a receptor működését új-generációs szekvenáláshoz kapcsolt módszerekkel.

A receptor által szabályozott génhálózat azonosításához RNS szekvenálást használtunk. Kromatin immunprecipitációhoz kapcsolt szekvenálással meghatároztuk az RXR genomi kötőhelyeit. Ezekből a kíséreltekből kiderült, hogy a receptor hozzávetőleg 800 gént szabályoz, ami 5200 kötőhelyről valósulhat meg. A receptor kötőhelyei majdnem 90%-ban intergenikus régiókban találhatók, sokszor több száz kilobázisos távolságra a génektől, így ezek génekhez való annotálása nem megbízható. A globális run-on technológia felhasználásával képesek voltunk a frissen keletkezett naszcens RNS mennyiségét detektálni teljes genom szinten. Irodalmi adatok igazolják, hogy ez a metódus képes a transzkripciós szabályozásban aktívan részt vevő cisz-elemek kimutatására, mivel ezekről úgynevezett enhanszer transzkriptek keletkeznek, amelyek az aktiváló stimulus hatására a célgénen történő transzkripció mértékét követik. Ezt kihasználva meghatároztuk azokat az RXR kötőhelyeket, amelyek részt vesznek a gének transzkripciós szabályozásában. A fent említett módszerek integrációjával 414 RXR által szabályozott cisz-elemet kapcsoltunk össze 252 célgénnel. Érdekes módon az RXR által szabályozott gén/cisz-elem hálózat több olyan gént is tartalmazott, amelyek az angiogenezis indukciójában játszanak jelentős szerepet, például: *Vegfa*, *Hbegf, Litaf*. Az RXR aktivátor által programozott makrofágokat ezután chorioallantois membránon végzett angiogenezis kísérletekben teszteltük. Az eredményeink egyértelműen bebizonyították, hogy az RXR által szabályozott gén/cisz-elem hálózat egy érképződést elősegítő makrofág fenotípus kialakításáért felelős.

Az általunk kifejlesztett, genomikai módszerek kombinációján alapuló metódus alkalmas a különböző szignálok hatására bekapcsolódó transzkripciós faktorok biológiai funkcióinak feltárásához, gyakorlatilag bármilyen sejttípusban.

List of Keywords

genomics, nuclear receptor, enhancer, macrophage, cistrome, RXR, ChIP-seq, GRO-seq, 3C-seq

Kulcsszavak

genomika, magreceptor, enhanszer, makrofág, cisztrom, RXR, ChIP-seq, GRO-seq, 3C-seq

Acknowledgements

First of all I would like to thank the continuous support and training received during my PhD studentship from my supervisor, Laszlo Nagy. He was an excellent mentor whose main goal was to lead me towards becoming a real scientist.

I would like to thank Prof. László Fésüs and Prof. József Tőzsér the former and recent leaders of the Department of Biohemistry and Molecular Biology for the opportunity to work in a well-equipped environment with highly skilled colleagues.

Special thanks to Gergely Nagy who basically performed the bioinformatics of this study and the members of the Nagy Laboratory for their help and support during my PhD.

I also thank Edit Hathy, Tímea Cseh, Márta Béládi, Beáta Szalka and Ibolya Fürtös for their excellent technical assistance.

I am very grateful to my family, especially my wife who was very sympathetic with me during my PhD.

References

1. Pardee, A.B., *PaJaMas in Paris.* Trends Genet, 2002. 18(11): p. 585-7.
2. Jacob, F., et al., *[Operon: a group of genes with the expression coordinated by an operator].* C R Hebd Seances Acad Sci, 1960. 250: p. 1727-9.
3. Jacob, F., *Evolution and tinkering.* Science, 1977. 196(4295): p. 1161-6.
4. Banerji, J., S. Rusconi, and W. Schaffner, *Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences.* Cell, 1981. 27(2 Pt 1): p. 299-308.
5. Stanford, W.L., J.B. Cohn, and S.P. Cordes, *Gene-trap mutagenesis: past, present and beyond.* Nat Rev Genet, 2001. 2(10): p. 756-68.
6. Augenlicht, L.H. and D. Kobrin, *Cloning and screening of sequences expressed in a mouse colon tumor.* Cancer Res, 1982. 42(3): p. 1088-93.
7. Schena, M., et al., *Quantitative monitoring of gene expression patterns with a complementary DNA microarray.* Science, 1995. 270(5235): p. 467-70.
8. Park, P.J., *ChIP-seq: advantages and challenges of a maturing technology.* Nat Rev Genet, 2009. 10(10): p. 669-80.
9. Barski, A., et al., *High-resolution profiling of histone methylations in the human genome.* Cell, 2007. 129(4): p. 823-37.
10. Robertson, G., et al., *Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing.* Nat Methods, 2007. 4(8): p. 651-7.
11. Dekker, J., et al., *Capturing chromosome conformation.* Science, 2002. 295(5558): p. 1306-11.
12. Simonis, M., et al., *Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C).* Nat Genet, 2006. 38(11): p. 1348-54.
13. Stadhouders, R., et al., *Multiplexed chromosome conformation capture sequencing for rapid genome-scale high-resolution detection of long-range chromatin interactions.* Nat Protoc, 2013. 8(3): p. 509-24.

14. Dostie, J., et al., *Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements.* Genome Res, 2006. 16(10): p. 1299-309.
15. Lieberman-Aiden, E., et al., *Comprehensive mapping of long-range interactions reveals folding principles of the human genome.* Science, 2009. 326(5950): p. 289-93.
16. Fullwood, M.J., et al., *An oestrogen-receptor-alpha-bound human chromatin interactome.* Nature, 2009. 462(7269): p. 58-64.
17. Horike, S., et al., *Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome.* Nat Genet, 2005. 37(1): p. 31-40.
18. Hagege, H., et al., *Quantitative analysis of chromosome conformation capture assays (3C-qPCR).* Nat Protoc, 2007. 2(7): p. 1722-33.
19. Weber, J., W. Jelinek, and J.E. Darnell, Jr., *The definition of a large viral transcription unit late in Ad2 infection of HeLa cells: mapping of nascent RNA molecules labeled in isolated nuclei.* Cell, 1977. 10(4): p. 611-6.
20. Core, L.J., J.J. Waterfall, and J.T. Lis, *Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters.* Science, 2008. 322(5909): p. 1845-8.
21. Nobrega, M.A., et al., *Scanning human gene deserts for long-range enhancers.* Science, 2003. 302(5644): p. 413.
22. Lettice, L.A., et al., *A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly.* Hum Mol Genet, 2003. 12(14): p. 1725-35.
23. Malik, S. and R.G. Roeder, *The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation.* Nat Rev Genet, 2010. 11(11): p. 761-72.
24. Clapier, C.R. and B.R. Cairns, *The biology of chromatin remodeling complexes.* Annu Rev Biochem, 2009. 78: p. 273-304.
25. Mito, Y., J.G. Henikoff, and S. Henikoff, *Histone replacement marks the boundaries of cis-regulatory domains.* Science, 2007. 315(5817): p. 1408-11.

26. Heintzman, N.D., et al., *Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome.* Nat Genet, 2007. 39(3): p. 311-8.
27. Jin, C., et al., *H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions.* Nat Genet, 2009. 41(8): p. 941-5.
28. Bedford, D.C., et al., *Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases.* Epigenetics, 2010. 5(1): p. 9-15.
29. Lam, M.T., et al., *Enhancer RNAs and regulated transcriptional programs.* Trends Biochem Sci, 2014. 39(4): p. 170-82.
30. Kim, T.H., et al., *A high-resolution map of active promoters in the human genome.* Nature, 2005. 436(7052): p. 876-80.
31. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. 128(4): p. 693-705.
32. Lupien, M., et al., *FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription.* Cell, 2008. 132(6): p. 958-70.
33. Lin, Y.C., et al., *A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate.* Nat Immunol, 2010. 11(7): p. 635-43.
34. Kim, T.K., et al., *Widespread transcription at neuronal activity-regulated enhancers.* Nature, 2010. 465(7295): p. 182-7.
35. Wang, D., et al., *Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA.* Nature, 2011. 474(7351): p. 390-4.
36. Pennacchio, L.A., et al., *In vivo enhancer analysis of human conserved non-coding sequences.* Nature, 2006. 444(7118): p. 499-502.
37. Visel, A., et al., *ChIP-seq accurately predicts tissue-specific activity of enhancers.* Nature, 2009. 457(7231): p. 854-8.
38. Carroll, J.S., et al., *Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1.* Cell, 2005. 122(1): p. 33-43.

39. Barish, G.D., et al., *The Bcl6-SMRT/NCoR cistrome represses inflammation to attenuate atherosclerosis.* Cell Metab, 2012. 15(4): p. 554-62.
40. de Wit, E. and W. de Laat, *A decade of 3C technologies: insights into nuclear organization.* Genes Dev, 2012. 26(1): p. 11-24.
41. West, A.G. and P. Fraser, *Remote control of gene transcription.* Hum Mol Genet, 2005. 14 Spec No 1: p. R101-11.
42. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. 409(6822): p. 860-921.
43. Consortium, E.P., *The ENCODE (ENCyclopedia Of DNA Elements) Project.* Science, 2004. 306(5696): p. 636-40.
44. Maher, B., *ENCODE: The human encyclopaedia.* Nature, 2012. 489(7414): p. 46-8.
45. Consortium, E.P., *An integrated encyclopedia of DNA elements in the human genome.* Nature, 2012. 489(7414): p. 57-74.
46. Hah, N., et al., *Enhancer transcripts mark active estrogen receptor binding sites.* Genome Res, 2013. 23(8): p. 1210-23.
47. Sanyal, A., et al., *The long-range interaction landscape of gene promoters.* Nature, 2012. 489(7414): p. 109-13.
48. Thurman, R.E., et al., *The accessible chromatin landscape of the human genome.* Nature, 2012. 489(7414): p. 75-82.
49. Watts, J.A., et al., *Study of FoxA pioneer factor at silent genes reveals Rfx-repressed enhancer at Cdx2 and a potential indicator of esophageal adenocarcinoma development.* PLoS Genet, 2011. 7(9): p. e1002277.
50. Magnani, L., J. Eeckhoute, and M. Lupien, *Pioneer factors: directing transcriptional regulators within the chromatin environment.* Trends Genet, 2011. 27(11): p. 465-74.
51. Ostuni, R., et al., *Latent enhancers activated by stimulation in differentiated cells.* Cell, 2013. 152(1-2): p. 157-71.
52. Ong, C.T. and V.G. Corces, *CTCF: an architectural protein bridging genome topology and function.* Nat Rev Genet, 2014. 15(4): p. 234-46.

53. Chung, J.H., M. Whiteley, and G. Felsenfeld, *A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila.* Cell, 1993. 74(3): p. 505-14.
54. Bell, A.C., A.G. West, and G. Felsenfeld, *The protein CTCF is required for the enhancer blocking activity of vertebrate insulators.* Cell, 1999. 98(3): p. 387-96.
55. Xiao, T., J. Wallace, and G. Felsenfeld, *Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity.* Mol Cell Biol, 2011. 31(11): p. 2174-83.
56. Nasmyth, K. and C.H. Haering, *Cohesin: its roles and mechanisms.* Annu Rev Genet, 2009. 43: p. 525-58.
57. Hadjur, S., et al., *Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus.* Nature, 2009. 460(7253): p. 410-3.
58. Kagey, M.H., et al., *Mediator and cohesin connect gene expression and chromatin architecture.* Nature, 2010. 467(7314): p. 430-5.
59. Seitan, V.C., et al., *A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation.* Nature, 2011. 476(7361): p. 467-71.
60. Parelho, V., et al., *Cohesins functionally associate with CTCF on mammalian chromosome arms.* Cell, 2008. 132(3): p. 422-33.
61. Rubio, E.D., et al., *CTCF physically links cohesin to chromatin.* Proc Natl Acad Sci U S A, 2008. 105(24): p. 8309-14.
62. Dixon, J.R., et al., *Topological domains in mammalian genomes identified by analysis of chromatin interactions.* Nature, 2012. 485(7398): p. 376-80.
63. Wang, H., et al., *Widespread plasticity in CTCF occupancy linked to DNA methylation.* Genome Res, 2012. 22(9): p. 1680-8.
64. Sofueva, S., et al., *Cohesin-mediated interactions organize chromosomal domain architecture.* EMBO J, 2013. 32(24): p. 3119-29.
65. Zuin, J., et al., *Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells.* Proc Natl Acad Sci U S A, 2014. 111(3): p. 996-1001.

66. Daniel, B., G. Nagy, and L. Nagy, *The intriguing complexities of mammalian gene regulation: how to link enhancers to regulated genes. Are we there yet?* FEBS Lett, 2014. 588(15): p. 2379-91.
67. Collis, P., M. Antoniou, and F. Grosveld, *Definition of the minimal requirements within the human beta-globin gene and the dominant control region for high level expression.* EMBO J, 1990. 9(1): p. 233-40.
68. Koch, F., et al., *Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters.* Nat Struct Mol Biol, 2011. 18(8): p. 956-63.
69. Hah, N., et al., *A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells.* Cell, 2011. 145(4): p. 622-34.
70. Li, W., et al., *Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation.* Nature, 2013. 498(7455): p. 516-20.
71. Lam, M.T., et al., *Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription.* Nature, 2013. 498(7455): p. 511-5.
72. Melo, C.A., et al., *eRNAs are required for p53-dependent enhancer activity and gene transcription.* Mol Cell, 2013. 49(3): p. 524-35.
73. Mousavi, K., et al., *eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci.* Mol Cell, 2013. 51(5): p. 606-17.
74. Lin, Y.C., et al., *Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate.* Nat Immunol, 2012. 13(12): p. 1196-204.
75. Noonan, J.P. and A.S. McCallion, *Genomics of long-range regulatory elements.* Annu Rev Genomics Hum Genet, 2010. 11: p. 1-23.
76. van Heyningen, V. and K.A. Williamson, *PAX6 in sensory development.* Hum Mol Genet, 2002. 11(10): p. 1161-7.
77. Kleinjan, D.A., et al., *Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6.* Hum Mol Genet, 2001. 10(19): p. 2049-59.
78. de Kok, Y.J., et al., *Identification of a hot spot for microdeletions in patients with X-linked deafness type 3 (DFN3) 900 kb proximal to the DFN3 gene POU3F4.* Hum Mol Genet, 1996. 5(9): p. 1229-35.

79. Visel, A., E.M. Rubin, and L.A. Pennacchio, *Genomic views of distant-acting enhancers.* Nature, 2009. 461(7261): p. 199-205.
80. Emison, E.S., et al., *A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk.* Nature, 2005. 434(7035): p. 857-63.
81. Rahimov, F., et al., *Disruption of an AP-2alpha binding site in an IRF6 enhancer is associated with cleft lip.* Nat Genet, 2008. 40(11): p. 1341-7.
82. Heinz, S., et al., *Effect of natural genetic variation on enhancer selection and function.* Nature, 2013. 503(7477): p. 487-92.
83. Toft, D., G. Shyamala, and J. Gorski, *A receptor molecule for estrogens: studies using a cell-free system.* Proc Natl Acad Sci U S A, 1967. 57(6): p. 1740-3.
84. Jensen, E.V. and E.R. DeSombre, *Mechanism of action of the female sex hormones.* Annu Rev Biochem, 1972. 41: p. 203-30.
85. Olefsky, J.M., *Nuclear receptor minireview series.* J Biol Chem, 2001. 276(40): p. 36863-4.
86. Nagy, L. and J.W. Schwabe, *Mechanism of the nuclear receptor molecular switch.* Trends Biochem Sci, 2004. 29(6): p. 317-24.
87. Kaufmann, S.H., *Immunology's foundation: the 100-year anniversary of the Nobel Prize to Paul Ehrlich and Elie Metchnikoff.* Nat Immunol, 2008. 9(7): p. 705-12.
88. Geissmann, F., et al., *Unravelling mononuclear phagocyte heterogeneity.* Nat Rev Immunol, 2010. 10(6): p. 453-60.
89. Glass, C.K. and J.L. Witztum, *Atherosclerosis. the road ahead.* Cell, 2001. 104(4): p. 503-16.
90. Olefsky, J.M. and C.K. Glass, *Macrophages, inflammation, and insulin resistance.* Annu Rev Physiol, 2010. 72: p. 219-46.
91. Murray, P.J., et al., *Macrophage activation and polarization: nomenclature and experimental guidelines.* Immunity, 2014. 41(1): p. 14-20.
92. Szanto, A., et al., *STAT6 transcription factor is a facilitator of the nuclear receptor PPARgamma-regulated gene expression in macrophages and dendritic cells.* Immunity, 2010. 33(5): p. 699-712.
93. Willy, P.J., et al., *LXR, a nuclear receptor that defines a distinct retinoid response pathway.* Genes Dev, 1995. 9(9): p. 1033-45.

94. Lehmann, J.M., et al., *Activation of the nuclear receptor LXR by oxysterols defines a new hormone response pathway.* J Biol Chem, 1997. 272(6): p. 3137-40.
95. Nagy, L., et al., *Nuclear hormone receptors enable macrophages and dendritic cells to sense their lipid environment and shape their immune response.* Physiol Rev, 2012. 92(2): p. 739-89.
96. Desvergne, B., L. Michalik, and W. Wahli, *Transcriptional regulation of metabolism.* Physiol Rev, 2006. 86(2): p. 465-514.
97. Szatmari, I. and L. Nagy, *Nuclear receptor signalling in dendritic cells connects lipids, the genome and immune function.* EMBO J, 2008. 27(18): p. 2353-62.
98. Tontonoz, P. and B.M. Spiegelman, *Fat and beyond: the diverse biology of PPARgamma.* Annu Rev Biochem, 2008. 77: p. 289-312.
99. Di Croce, L., et al., *Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor.* Science, 2002. 295(5557): p. 1079-82.
100. Kersten, S., B. Desvergne, and W. Wahli, *Roles of PPARs in health and disease.* Nature, 2000. 405(6785): p. 421-4.
101. Giguere, V., et al., *Identification of a receptor for the morphogen retinoic acid.* Nature, 1987. 330(6149): p. 624-9.
102. De Luca, L.M., *Retinoids and their receptors in differentiation, embryogenesis, and neoplasia.* FASEB J, 1991. 5(14): p. 2924-33.
103. Mendelsohn, C., et al., *Function of the retinoic acid receptors (RARs) during development (II). Multiple abnormalities at various stages of organogenesis in RAR double mutants.* Development, 1994. 120(10): p. 2749-71.
104. Kastner, P. and S. Chan, *Function of RARalpha during the maturation of neutrophils.* Oncogene, 2001. 20(49): p. 7178-85.
105. Mangelsdorf, D.J., et al., *Characterization of three RXR genes that mediate the action of 9-cis retinoic acid.* Genes Dev, 1992. 6(3): p. 329-44.
106. Szanto, A., et al., *Retinoid X receptors: X-ploring their (patho)physiological functions.* Cell Death Differ, 2004. 11 Suppl 2: p. S126-43.

107. Zhang, X.K., et al., *Homodimer formation of retinoid X receptor induced by 9-cis retinoic acid.* Nature, 1992. 358(6387): p. 587-91.
108. Szeles, L., et al., *Research resource: transcriptome profiling of genes regulated by RXR and its permissive and nonpermissive partners in differentiating monocyte-derived dendritic cells.* Mol Endocrinol, 2010. 24(11): p. 2218-31.
109. Germain, P., et al., *Co-regulator recruitment and the mechanism of retinoic acid receptor synergy.* Nature, 2002. 415(6868): p. 187-92.
110. Boehm, M.F., et al., *Synthesis and structure-activity relationships of novel retinoid X receptor-selective retinoids.* J Med Chem, 1994. 37(18): p. 2930-41.
111. Boehm, M.F., et al., *Design and synthesis of potent retinoid X receptor selective ligands that induce apoptosis in leukemia cells.* J Med Chem, 1995. 38(16): p. 3146-55.
112. Mukherjee, R., et al., *Sensitization of diabetic and obese mice to insulin by retinoid X receptor agonists.* Nature, 1997. 386(6623): p. 407-10.
113. Cramer, P.E., et al., *ApoE-directed therapeutics rapidly clear beta-amyloid and reverse deficits in AD mouse models.* Science, 2012. 335(6075): p. 1503-6.
114. Fitz, N.F., et al., *Comment on "ApoE-directed therapeutics rapidly clear beta-amyloid and reverse deficits in AD mouse models".* Science, 2013. 340(6135): p. 924-c.
115. Calkin, A.C. and P. Tontonoz, *Transcriptional integration of metabolism by the nuclear sterol-activated receptors LXR and FXR.* Nat Rev Mol Cell Biol, 2012. 13(4): p. 213-24.
116. Barish, G.D., et al., *A Nuclear Receptor Atlas: macrophage activation.* Mol Endocrinol, 2005. 19(10): p. 2466-77.
117. Trapnell, C., et al., *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.* Nat Protoc, 2012. 7(3): p. 562-78.
118. Daniel, B., et al., *Mapping the genomic binding sites of the activated retinoid x receptor in murine bone marrow-derived macrophages using chromatin immunoprecipitation sequencing.* Methods Mol Biol, 2014. 1204: p. 15-24.

119. Heinz, S., et al., *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.* Mol Cell, 2010. 38(4): p. 576-89.
120. Thorvaldsdottir, H., J.T. Robinson, and J.P. Mesirov, *Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration.* Brief Bioinform, 2013. 14(2): p. 178-92.
121. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS).* Genome Biol, 2008. 9(9): p. R137.
122. Ross-Innes, C.S., et al., *Differential oestrogen receptor binding is associated with clinical outcome in breast cancer.* Nature, 2012. 481(7381): p. 389-93.
123. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics, 2010. 26(6): p. 841-2.
124. Kestler, H.A., et al., *Generalized Venn diagrams: a new method of visualizing complex genetic set relations.* Bioinformatics, 2005. 21(8): p. 1592-5.
125. Movahedi, B., et al., *Pancreatic duct cells in human islet cell preparations are a source of angiogenic cytokines interleukin-8 and vascular endothelial growth factor.* Diabetes, 2008. 57(8): p. 2128-36.
126. Segal, E., et al., *Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data.* Nat Genet, 2003. 34(2): p. 166-76.
127. Kersten, S., et al., *Characterization of the fasting-induced adipose factor FIAF, a novel peroxisome proliferator-activated receptor target gene.* J Biol Chem, 2000. 275(37): p. 28488-93.
128. Repa, J.J., et al., *Regulation of absorption and ABC1-mediated efflux of cholesterol by RXR heterodimers.* Science, 2000. 289(5484): p. 1524-9.
129. Schwartz, K., R.M. Lawn, and D.P. Wade, *ABC1 gene expression and ApoA-I-mediated cholesterol efflux are regulated by LXR.* Biochem Biophys Res Commun, 2000. 274(3): p. 794-802.
130. Nagy, L., et al., *Identification and characterization of a versatile retinoid response element (retinoic acid receptor response element-retinoid X receptor response*

element) in the mouse tissue transglutaminase gene promoter. J Biol Chem, 1996. 271(8): p. 4355-65.

131. Merkenschlager, M. and D.T. Odom, *CTCF and cohesin: linking gene regulatory elements with their targets.* Cell, 2013. 152(6): p. 1285-97.
132. Lefterova, M.I., et al., *Cell-specific determinants of peroxisome proliferator-activated receptor gamma function in adipocytes and macrophages.* Mol Cell Biol, 2010. 30(9): p. 2078-89.
133. Pott, S., et al., *PPARG binding landscapes in macrophages suggest a genome-wide contribution of PU.1 to divergent PPARG binding in human and mouse.* PLoS One, 2012. 7(10): p. e48102.
134. Zentner, G.E. and P.C. Scacheri, *The chromatin fingerprint of gene enhancer elements.* J Biol Chem, 2012. 287(37): p. 30888-96.
135. Nielsen, R., et al., *Genome-wide profiling of PPARgamma:RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis.* Genes Dev, 2008. 22(21): p. 2953-67.
136. Boergesen, M., et al., *Genome-wide profiling of liver X receptor, retinoid X receptor, and peroxisome proliferator-activated receptor alpha in mouse liver reveals extensive sharing of binding sites.* Mol Cell Biol, 2012. 32(4): p. 852-67.
137. Chakravarti, D., et al., *Role of CBP/P300 in nuclear receptor signalling.* Nature, 1996. 383(6595): p. 99-103.
138. Jin, Q., et al., *Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation.* EMBO J, 2011. 30(2): p. 249-62.
139. Hoeben, A., et al., *Vascular endothelial growth factor and angiogenesis.* Pharmacol Rev, 2004. 56(4): p. 549-80.
140. Ongusaha, P.P., et al., *HB-EGF is a potent inducer of tumor growth and angiogenesis.* Cancer Res, 2004. 64(15): p. 5283-90.
141. Aikawa, Y., et al., *Roles of HIPK1 and HIPK2 in AML1- and p300-dependent transcription, hematopoiesis and blood vessel formation.* EMBO J, 2006. 25(17): p. 3955-65.

142. Tang, X., et al., *Novel transcriptional regulation of VEGF in inflammatory processes.* J Cell Mol Med, 2013. 17(3): p. 386-97.
143. Kollmar, O., et al., *Macrophage inflammatory protein-2 promotes angiogenesis, cell migration, and tumor growth in hepatic metastasis.* Ann Surg Oncol, 2006. 13(2): p. 263-75.
144. Park, J.H., et al., *Loss of Mel-18 induces tumor angiogenesis through enhancing the activity and expression of HIF-1alpha mediated by the PTEN/PI3K/Akt pathway.* Oncogene, 2011. 30(45): p. 4578-89.
145. Bingle, L., N.J. Brown, and C.E. Lewis, *The role of tumour-associated macrophages in tumour progression: implications for new anticancer therapies.* J Pathol, 2002. 196(3): p. 254-65.
146. Walczak, R., et al., *Transcription of the vascular endothelial growth factor gene in macrophages is regulated by liver X receptors.* J Biol Chem, 2004. 279(11): p. 9905-11.
147. Qian, B.Z. and J.W. Pollard, *Macrophage diversity enhances tumor progression and metastasis.* Cell, 2010. 141(1): p. 39-51.
148. de Duve, C., *The lysosome turns fifty.* Nat Cell Biol, 2005. **7**(9): p. 847-9.
149. Saftig, P. and J. Klumperman, *Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function.* Nat Rev Mol Cell Biol, 2009. **10**(9): p. 623-35.
150. Klein, A.D. and A.H. Futerman, *Lysosomal storage disorders: old diseases, present and future challenges.* Pediatr Endocrinol Rev, 2013. **11 Suppl 1**: p. 59-63..
151. Kang, Y.A., et al., *Autophagy driven by a master regulator of hematopoiesis.* Mol Cell Biol, 2012. **32**(1): p. 226-39
152. Pena-Llopis, S., et al., *Regulation of TFEB and V-ATPases by mTORC1.* EMBO J, 2011. **30**(16): p. 3242-58.

Publications related to dissertation

UNIVERSITY OF DEBRECEN
UNIVERSITY AND NATIONAL LIBRARY
PUBLICATIONS

Register number: DEENKÉTK/261/2014.
Item number:
Subject: Ph.D. List of Publications

Candidate: Bence Dániel
Neptun ID: FVHSAG
Doctoral School: Doctoral School of Molecular Cell and Immune Biology

List of publications related to the dissertation

1. **Dániel, B.**, Nagy, G., Nagy, L.: The intriguing complexities of mammalian gene regulation: How to link enhancers to regulated genes. Are we there yet?
   FEBS Lett. 588 (15), 2379-2391, 2014.
   DOI: http://dx.doi.org/10.1016/j.febslet.2014.05.041
   IF:3.341 (2013)

2. **Dániel, B.**, Nagy, G., Hah, N., Horváth, A., Czimmerer, Z., Póliska, S., Gyuris, T., Keirsse, J., Gysemans, C., Van Ginderachter, J.A., Bálint, B.L., Evans, R.M., Barta, E., Nagy, L.: The active enhancer network operated by liganded RXR supports angiogenic activity in macrophages.
   Genes Dev. 28 (14), 1562-1577, 2014.
   DOI: http://dx.doi.org/10.1101/gad.242685.114
   IF:12.639 (2013)

3. Brignull, L.M., Czimmerer, Z., Saidi, H., **Dániel, B.**, Villela, I., Bartlett, N.W., Johnston, S.L., Meira, L.B., Nagy, L., Nohturfft, A.: Reprogramming of lysosomal gene expression by interleukin-4 and Stat6.
   BMC Genomics. 14 (1), 1-20, 2013.
   DOI: http://dx.doi.org/10.1186/1471-2164-14-853
   IF:4.041

List of other publications

List of other publications

4. **Dániel, B.**, Bálint, B.L., S. Nagy, Z., Nagy, L.: Mapping the genomic binding sites of the activated retinoid x receptor in murine bone marrow-derived macrophages using chromatin immunoprecipitation sequencing.
Methods Mol. Biol. 1204, 15-24, 2014.
DOI: http://dx.doi.org/10.1007/978-1-4939-1346-6_2

5. Brázda, P., Krieger, J., **Dániel, B.**, Jonás, D., Szekeres, T., Langowski, J., Tóth, K., Nagy, L., Vámosi, G.: Ligand binding shifts highly mobile RXR to chromatin-bound state in a coactivator-dependent manner as revealed by single cell imaging.
Mol. Cell. Biol. "accepted by publisher" (2014)
DOI: http://dx.doi.org/10.1128/MCB.01097-13
IF:5.036 (2013)

6. Nagy, G., **Dániel, B.**, Jonás, D., Nagy, L., Barta, E.: A novel method to predict regulatory regions based on histone mark landscapes in macrophages.
Immunobiology. 218 (11), 1416-1427, 2013.
DOI: http://dx.doi.org/10.1016/j.imbio.2013.07.006
IF:3.18

Total IF of journals (all publications): 28,237
Total IF of journals (publications related to the dissertation): 20,021

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of Web of Science, Scopus and Journal Citation Report (Impact Factor) databases.

11 September, 2014

Oral and poster presentations

2010 iGEM (International Genetically Engineered Machine competition) International synthetic biology competition - Cambridge, Boston, Massachusetts, USA – MIT (presentation and poster)

2011 EMDS (European Macrophage and Dendritic Cell Society) Conference Debrecen

Presentation: The global view on RXR regulated transcription events in mouse bone marrow derived macrophages
Poster: The global view on RXR regulated transcription events in mouse bone marrow derived macrophages

2013 FEBS Advanced lecture course on nuclear receptor signaling in physiology and disease

Presentation: Cistromic features of liganded RXR’s action in macrophages define a distinct, partly hidden enhancer network
Poster: Cistromic features of liganded RXR’s action in macrophages define a distinct, partly hidden enhancer network

2013 Semmelweis Symposium

Poster: Cistromic features of liganded RXR's action in macrophages define a distinct, partly hidden enhancer network

Appendix

Brignull, L.M., et al., Reprogramming of lysosomal gene expression by interleukin-4 and Stat6. BMC Genomics, 2013. 14: p. 853.

Daniel, B., et al., The active enhancer network operated by liganded RXR supports angiogenic activity in macrophages. Genes Dev, 2014. 28(14): p. 1562-77.

Daniel, B., G. Nagy, and L. Nagy, The intriguing complexities of mammalian gene regulation: how to link enhancers to regulated genes. Are we there yet? FEBS Lett, 2014. 588(15): p. 2379-91.