From the DEPARTMENT OF MICROBIOLOY, TUMOR AND CELL BIOLOGY

Karolinska Institutet, Stockholm, Sweden

TISSUE- AND SITE-SPECIFIC EFFECTS OF PPARγ ACTIVATION AND ITS ROLE IN CHRONIC INFLAMMATION

Sebastian Pott

Karolinska Institutet

Stockholm 2011

Published by Karolinska Institutet. Printed by Larserics Digital Print AB.

ISBN 978-91-7457-343-5

Abstract

Developmental programming, metabolism, immune function and tissue homeostasis in multicellular organisms are regulated by a plethora of stimuli. Transcription factors (TFs) convert incoming signals to appropriate transcriptional responses of the genome. Nuclear receptors (NRs), a particular family of TFs, are especially well suited for this task given their capacity to influence gene regulation in multiple tissues and conditions. In this work, I addressed how environmental cues affect the expression of NRs and how gene regulation is mediated in different tissues.

My thesis comprises a set of four separate manuscripts of which the first paper established how an important environmental factor, the gut microbiome, modulates NRs, including the peroxisome proliferator-activated receptor gamma (PPARγ), in vivo. The ligand-activated TF PPARγ is a key regulator of adipogenesis and glucose homeostasis, and possesses profound anti-inflammatory properties. In the subsequent manuscripts (paper II-IV), I have used PPARγ as a model to gain insights into the mechanisms that guide tissue-specific activity of NRs on a genome-wide level. To this end I have identified PPARγ binding sites in a genome-wide manner using chromatin immunoprecipitation combined with high-throughput sequencing (ChIP-seq) in adipocytes, macrophages and intestinal epithelium. These data were superimposed with information from gene-expression profiling to facilitate identification of direct PPARγ target genes.

Specifically, our comprehensive genomic analysis of PPARγ binding during adipogenesis confirmed the role of PPARγ as key regulator of this process and, potentially, revealed novel adipogenic target genes of PPARγ (paper II). In addition, an inter-species comparison of PPARγ binding sites in human and mouse macrophages identified a core set of conserved PPARγ targets. In this study we also identified PU.1 as a co-factor necessary for PPARγ function in macrophages (paper III). In the final manuscript, I mapped the genomic landscape of PPARγ-DNA interactions in intestinal epithelial cells (paper IV). These studies revealed that PPARγ antagonizes the WNT/TCF4 signaling pathways potentially identifying a mechanism by which activation of PPARγ affects cell fate of intestinal epithelial cells.

My PhD work yielded important novel insights into general mechanisms related to PPARγ-dependent gene regulation. While in the tissues studied, PPARγ activation seemed to always induce a core set of lipid metabolic genes, tissue-specific utilization of PPARγ binding sites appeared to be dependent on cell type restricted transcription factors which may determine binding site accessibility at the chromatin level.

My studies further suggest that the regulation of lipid metabolism is the evolutionary most conserved function of PPARγ and additional functions might have developed later, representing adaptations to changing metabolic needs and environmental challenges. While not studied in depth, our data on tissue-specific mechanisms of TF binding might also have implications for the interpretation of population-wide genetic studies.

In conclusion, my work has revealed common principles that guide PPARγ activation in a tissue-dependent and -independent manner and has laid the fundament for further detailed molecular studies of NRs in general and PPARγ in particular.

List of Publications

Paper I

Lundin A, Bok CM, Aronsson L, Björkholm B, Gustafsson JA, **Pott S**, Arulampalam V, Hibberd M, Rafter J, Pettersson S.
Gut flora, Toll-like receptors and nuclear receptors: a tripartite communication that tunes innate immunity in large intestine. Cell Microbiol (2008) 10(5):1093-103.

Paper II

Hamza MS, **Pott S***, Vega VB, Thomsen JS, Kandhadayar GS, Ng PW, Chiu KP, Pettersson S, Wei CL, Ruan Y, Liu ET.
De-novo identification of PPARamma/RXR binding sites and direct targets during adipogenesis. PLoS One (2009) 4(3):e4907.

Paper III

Pott S, Kamrani NK, Bourque G, Pettersson S, Liu ET
Interplay between PPARG and PU.1 during evolution shape divergent PPARG binding landscapes in human and mouse.
Manuscript

Paper IV

Pott S, Kamrani NK, Liu ET, Pettersson S.
Integrative genomic analysis of PPARγ activity suggests broad antagonism of WNT/β-catenin signaling in intestinal epithelial cells.
Manuscript

*shared first authorship

Table of content

1. Introduction……1

1.1 Chromatin structure……2
1.2 Transcription factors and the regulation of gene expression……3
1.3 Nuclear receptors……5
1.3.1 Co-activators and co-repressors……7

1.4 Peroxisome proliferator-activated receptor gamma (PPARγ)……7

1.4.1 PPARγ in the intestinal epithelium……9
1.4.2 PPARγ in adipocytes……10
1.4.3 PPARγ in myeloid cells……11

2. Aims……13

3. Methodology……14

3.1 Genome-wide identification of TF binding sites using ChIP-seq ……14
3.2 Microarray-based gene expression analysis……15

4. Results and discussion……17

4.1 Paper I: Gut flora, Toll-like receptors and nuclear receptors: a tripartite communication that tunes innate immunity in large intestin……17

4.2 Genome-wide exploration of PPARγ activity across different tissues……19
4.2.1 Paper II: De-novo identification of PPARgamma/RXR binding sites and direct target during Adipogenesis……19
4.2.2 Paper III: Interplay between PPARG and PU.1 during evolution shape divergent PPARG binding landscapes in human and mouse……21
4.2.3 Paper IV: Integrative genomic analysis of PPARγ activity suggests broad antagonism of WNT/β-catenin signaling in intestinal epithelial cells……24
4.2.4 Insights from genome-wide mapping of PPARγ binding in different tissues……27

5. Concluding remarks and perspectives……31

6. Acknowledgements ……34

7. References……35

List of Abbreviations

AF	Activation function
BAT	Brown adipose tissue
CTD	Carboxy-terminal domain
ChIP	Chromatin immunoprecipitation
DBD	DNA binding domain
DC	Dendritic cell
DSS	Dextran sulfate sodium
GF	Germ free
GTF	General transcription factor
LBD	Ligand binding domain
NR	Nuclear receptor
PCR	Polymerase chain reaction
PIC	Pre-initiation complex
PolII	RNA polymerase II
PPAR	Peroxisome proliferator-activated receptor
PTM	Posttranslational modification
RSG	Rosiglitazone
RXR	Retinoid X receptor
SPF	Specific pathogen free
T2D	Type 2 diabetes
TF	Transcription factor
TLR	Toll-like receptor
TSS	Transcription start site
TZD	Thiazolidinedione
WAT	White adipose tissue

1. Introduction

The accurate interpretation of intrinsic and extrinsic signals is crucial for development, reproduction, survival and overall well-being of all multicellular organisms. While the genome sequence holds the information necessary for appropriate physiological responses, its correct utilization depends on the interaction between sophisticated cellular machinery and the genome in order to change expression of specific genes. Transcriptional regulation is one of several essential tuning components that maintain homeostasis. If inappropriately guided, dysregulation of transcriptional processes might lead to chronically altered gene expression and ultimately manifest in disease. It is therefore vital and important to decipher the complex mechanisms that regulate how the information encoded in the genome produces the appropriate transcriptional and phenotypic outputs in order to understand the pathophysiological pathways underlying disease development and progression. Central to this endeavor is the understanding of how factors interact with the genome in response to a variety of stimuli. Such studies have become feasible with the availability of full mammalian genome sequences [1-4] and the development of powerful high-throughput sequencing technologies [5, 6] which have enabled us to functionally interrogate genomes across species, tissues and under different stimuli.

Accessibility to regulatory options, rather than the number of genes, underlies cellular diversity

The last two decades have witnessed dramatic developments in biology. The publication of the human and other genome sequences allowed for the first time to directly assess the critical components of genomes. One of the major findings in these initial analyses was that the number of genes encoded within the human genome was around ~21.000, a much smaller number than the anticipated 50.000-100.000 genes [7]. The number of genes turned out to be similar in mouse and rat and even in more distantly related species. Moreover, most of these genes were conserved across mammalian species. These data strongly indicated that variation in the protein-coding regions alone was not enough to explain the phenotypic differences between species. Therefore, variation in the regulation of transcription shifted into focus and it became clear that a significant proportion of the non-coding regions of the genome showed some degree of conservation (~6%) indicating their functional relevance [8]. Functional characterization of the human genome confirmed that an even larger portion of the human genome appeared to fulfill some regulatory roles [9]. Thus, the non-coding regions of

genomes, once considered of little functional use, are now regarded as central elements in the regulation of gene expression. Evolutionary change therefore does not depend entirely on the 'invention' of novel functional parts (i.e. proteins) but rather uses changes in regulatory regions to fine tune and re-write the instructions of where and when these parts would be used in cells. The contribution of regulatory regions to diversity is further underscored by studies aiming to characterize the genetic diversity within human populations. These studies strongly indicate that phenotypic variation between humans is in part based on sequence variation in regulatory regions which can results in subtle variation in gene expression and differential responses to extrinsic stimuli [10-12].

1.1 Chromatin structure

Genomes of multi-cellular organisms are able to specify cell types with vastly different characteristics (e.g. neuron, muscle cell, epithelium). Such diversity is achieved through tightly controlled gene expression programs. Not all genes are expressed all the time, most genes are rather not or lowly expressed while some are highly expressed either representing genes with essential functions in one or few tissues (e.g. Oct4, MyoD, Insulin) or in most tissues ('housekeeping genes'; e.g. Actin). Control of gene expression is critically dependent on several interconnected regulatory layers. In eukaryotic cells the genomic DNA is stored in the form of chromatin, a nucleoprotein complex consisting of DNA and histone proteins [13]. Nucleosomes are the fundamental unit of chromatin and are formed by approximately 147 bp DNA wrapped around an octamer of histone proteins [14]. Histones can be posttranslationally modified (e.g. phosphorylation, methylation, acetylation). It has been suggested that the combinatorial patterns of different histone modifications serve as a 'histone-code' guiding the recognition and interpretation of specific genomic regions [15]. Such a regulatory 'code' can for example integrate the information stored in the genome sequence with more transient, environmental signals represented by different posttranslational modifications (PTMs) on the histones [16].

The advent of high-throughput sequencing methods allowed the systematic interrogation of histone modifications and the investigation of combinatorial patterns across the genome in great detail. It is now clear that the combination of different histone modifications specify different chromatin states that are associated with distinct modes of activation. Genome-wide mapping of histone modification in different cell types provided a high resolution picture of

histone marks associated with different regulatory regions and functions [17-19]. These studies showed that the transcription start site (TSS) of actively transcribed genes are characteristically marked by histone 3 lysin 4 trimethyl (H3K4Me3) and histone 3 lysine 9 acetyl (H3K9Ac), while repressed genes preferentially carry H3K9Me3 and H3K27Me3. In contrast, distal regulatory regions (enhancers) were demarcated by strong enrichment of H3K4Me1 in the absence of H3K4Me3 [19] and the gene bodies of actively transcribed regions show enrichment of H3K36Me3 [20]. In addition, direct methylation of DNA serves as a stable mark and is associated with repressive and inactive regions [21, 22]. Comprehensive chromatin maps in humans [23, 24], Drosophila [25-27] and C. elegans [28] revealed a functional compartmentalization based on the combinatorial association of a large number of different chromatin-associated proteins. These compartments, referred to as chromatin domains, correlate with functional states such as transcriptional activation or repression and are thought to facilitate the propagation and maintenance of transcriptional states in cell-type-specific manner.

However, while chromatin marks participate in establishing stable transcriptional states and facilitate the efficient read-out of the genome, it is important to keep in mind that sequence-specific transcription factors (TFs) are required to initially set and later change chromatin states in a sequence-specific manner. The fact that over-expression of few TFs is sufficient to reprogram cells into pluripotent stem cells (iPSC) [29] provides compelling evidence that TFs are both necessary and sufficient for the establishment of tissue-specific chromatin states, DNA methylation and gene expression patterns. Such reprogramming experiments demonstrated the plasticity of the chromatin environment [30]. In summary, the primary genome sequence is embedded in a regulatory scaffold that restricts and facilitates transcriptional regulation in a tissue- and signal-specific manner.

1.2 Transcription factors and the regulation of gene expression

The fact that the primary DNA sequence in eukaryotic cells is packaged into chromatin is important for the understanding of the basic mechanism that guide transcription. Protein-coding genes are transcribed into RNA by the RNA polymerase II (PolII) (reviewed in: [31, 32]). Transcription is initiated at regions immediately surrounding the TSS, the promoters, which often contain common recognition sequences such as TATA boxes or initiator sequences to engage the basal transcription machinery. While transcription is induced by

different stimuli and controlled in a tissue-specific manner, the sequence of events which controls transcription factor binding eventually leading to active transcription mediated by PolII is highly conserved and includes both specific recognition of DNA by TFs and remodeling or modifying of the chromatin structure by co-factors (Fig. 1).

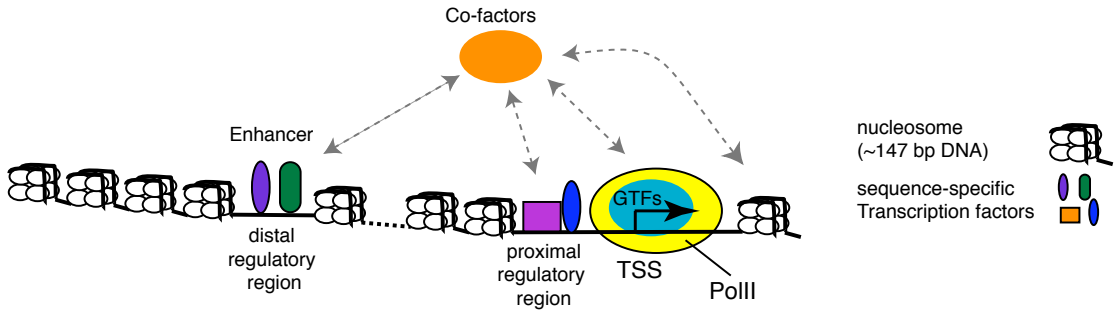

Figure 1. Schematic representation of the interaction between TFs binding to distal and proximal regulatory elements with co-factors and PolII. *Adapted and modified from [33].*

Binding of sequence-specific TFs (e.g. NFκB, p53) to their recognition sites located at the promoter, in its proximity or at distal enhancer elements initiates the events. In eukaryotes, these TFs recruit co-activators which remodel chromatin, interact with and modify histones and make contact with components of the pre-initiation complex (PIC). The PIC is formed by general transcription factors (GTFs) and PolII [34]. To proceed with transcription, the carboxy-terminal domain (CTD) is phosphorylated at serine 5 and single-stranded DNA is loaded onto PolII. Yet, after the formation of the transcription bubble and transcription of a few base pairs, PolII pauses, until phosphorylation of serine 2 in the CTD by P-TEFb allows recruitment of mRNA processing complexes and facilitates elongation. Recently, several studies revealed that the transition from pausing to transcriptional elongation is also controlled by transcription factors (e.g. MYC, NFκB) [35-37], illustrating that TFs might have multiple roles during these processes, depending on signals and promoters. Given the conserved basal transcriptional machinery and the similarity in its activation across different tissues, the crucial role of sequence-specific TFs in the generation of specific responses to the various cellular signals becomes clear. Of note, TFs might also act as repressors by inhibiting activators or through the recruitment of histone-modifying enzymes that deposit repressive histone marks. In addition, sequence-specific DNA binding factors might serve as insulators by separating active chromatin domains from inactive domains (e.g. CTCF [38]), further broadening the spectrum of regulatory options. Given their central function it is not surprising that about 1400 TFs [39] are encoded by the human genome, as predicted by sequence

homology. Some of these TFs are widely expressed but most of the TFs have a restricted expression pattern, limited to few or a single tissue.

1.3 Nuclear receptors: signal-specific ligand induced transcriptional regulators

Nuclear receptors are a class of ligand inducible TFs which have evolved to translate environmental stimuli into transcriptional responses. NRs derive their name from the fact that many of them act as receptors for small lipophilic compounds such as hormones, vitamins and metabolites. NRs can translocate to the nucleus where they directly bind to DNA and activate or repress transcription. In humans 48 receptors of the NR superfamily have been identified on the basis of sequence homology (Reviewed in: [40]) and many of them exert fundamental biological functions in reproduction, development and metabolism**.** NRs are expressed throughout the body, with some NRs being ubiquitously expressed (e.g. Retinoid X receptor alpha (RXRα), liver X receptor beta (LXRβ)), and others showing more restricted expression patterns, in agreement with their respective biological functions (e.g. constitutive androstane receptor (CAR), pregnane X receptor (PXR), farnesoid X receptor (NR1H4)) [41]. Recognition of specific binding sites is mediated through two highly conserved zinc fingers located in the N-terminal DNA binding domain (DBD). In most cases, the DNA recognition site is a variation of the hexameric 5'-AGGTCA consensus sequence [42], referred to as hormone-response element (HRE) or half-site. NRs might bind to DNA as monomers, homo- or hetero-dimers depending on the factor. In some cases, a NR can bind in different configurations, depending on its binding partner and binding sites [43]. The first NRs to be identified were receptors for steroid hormones (such as glucocorticoid receptor (GR) [44] and androgen receptor [45]) which bind DNA as homodimers on half-sites arranged in a palindromic configuration. Subsequent studies identified receptors which were activated by non-steroid hormones, vitamins or other lipophilic compounds such as thyroid hormone, retinoid acid, Vitamin D, prostaglandins [40]. These receptors predominantly form heterodimers with members of the RXR family (RXRα, RXRβ, RXRγ) and bind a broad spectrum of different recognition motifs. In general RXR and its heterodimerisation partner bind to direct repeats (DRs). The spacing between the two half sites varies depending on the binding partners. Some NRs are referred to as 'orphans' because no specific ligand has yet been identified. In fact a number of them might act independently of ligands as constitutive activators or repressors of transcription. Monomeric binding of receptors has also been demonstrated.

In addition to the DBD, the N-terminus of NRs generally contains an activation domain (AF-1), which contributes to the transcriptional activation in a ligand-independent fashion. Part of the interaction between the partners in a dimer is also mediated via the DBD, most of it, however is mediated through the ligand-binding domain (LBD). The LBD is connected to the DBD via a flexible 'hinge' region (Fig. 2).

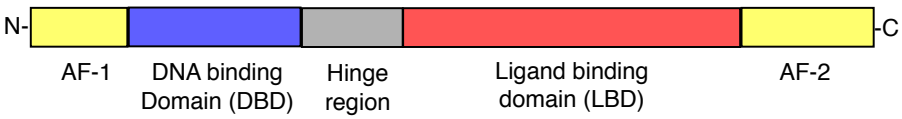

Figure 2. **Schematic representation of the domains of NRs**. *AF-1 and -2 domains are located and the N- and C-terminus, respectively. The DBD is located in the N-terminal part and connected to the LBD through a flexible hinge region.*

Ligand recognition and interaction with co-activators and repressors is mediated via the LBD which also contains a second activation domain (AF-2). LBDs are relatively conserved between NRs in terms of structure, but vary to some degree in their sequence and show a remarkable degree of ligand specificity. Binding of the ligand induces conformational changes in the LBD which together with the AF-2 forms an interaction surface that is recognized by the LXXLL motif, which is common to many co-activators [46, 47]. In addition, the activity of most, if not all, NRs can be further modulated by post-translational modifications (PTMs) (such as phosphorylation, ubiquitylation, acetylation, sumoylation [48-50]). PTMs facilitate context-specific modulation of NR activity by taking into account input from other signaling pathways (e.g. MAPK [51, 52]). While molecular features and functional domains are often shared by many NRs, the combination of DNA-recognition, ligand specificity, posttranslational modifications and expression levels of the NR itself ensure a high level of specificity and restricted access to target gene regulation.

The high specificity of gene regulation by NRs and their direct amenability to pharmacological interventions have shifted the focus of NR research in many disease-relevant areas. For example, ERα antagonists are used to treat breast cancer and glucocorticoid receptor and members of the peroxisome proliferator-activated receptors are selectively exploited in inflammatory disorders and to manage insulin resistance and type 2 diabetes (T2D) [53], respectively. Given the many functions of NRs in crucial processes of

development and metabolism, it is immediately clear that NRs are promising targets in a variety of disorders.

1.3.1 Co-activators and co-repressors

Nuclear receptors, like other TFs, require co-activators or co-repressors in order to control transcription and guide their mode of action on the DNA template. While some NRs have been shown to directly interact with the basal transcription machinery [54, 55], in most cases additional factors are recruited to assist in transcriptional activation. Most of these co-activators directly interact with NRs via their LXXLL motif which recognizes a surface formed by helices of the ligand-bound LBD/AF-2 [47]. Co-activators represent a diverse group of factors which act either by directly interacting with the transcriptional machinery or through modification of histones and chromatin remodeling. Many cofactors facilitate acetylation or methylation of histone tails (e.g. CBP/p300, p160, P/CAF [56, 57]). These factors therefore directly influence the chromatin state and potentially set activation marks at enhancers and/or TSS. Other co-activators such as the SWI/SNF complex are involved in ATP-dependent chromatin remodeling and thus in removing nucleosomes away from the TSS facilitating access for PolII, while yet other factors seem to act as bridges or mediators between the NR and the transcriptional machinery. Co-repressors usually bind NRs in the unliganded state as their common motif recognizes the LBD/AF-2 in the unliganded conformation. Co-repressors include histone deacetylases (HDACs) such as SMRT and NCoR. Of note, some co-repressors do bind to NRs in a ligand-dependent manner and therefore mediate repression of transcription after ligand activation, providing potential negative-feedback regulation. Generally, removal of the co-repressor complex (containing SMRT and NCoR among others) bound to the unliganded NR precedes the recruitment of co-activators upon ligand stimulation. Elaborate molecular analysis of NR co-activator complexes after activation of the receptor suggest that co-activator and -repressor recruitment and replacement are complex cyclical multistep processes [57, 58].

1.4 Peroxisome proliferator-activated receptor gamma (PPARγ)

Lifestyle-related diseases like obesity, T2D and metabolic syndrome constitute a major health problem in a large part of the world [59]. Many of these disorders share several features,

including dyslipidemia and, often, insulin resistance. Of note, all these conditions appear to have chronic inflammatory components [60]. While it is likely that the rapid increase in incidences of metabolic disorders and inflammatory conditions is causally related to lifestyle changes such as nutritional oversupply and a generally sedentary lifestyle, more research is required to fully understand the pathophysiology underlying this rapidly growing group of diseases. Interestingly, the composition of the microbial flora in the gut, representing an environmental factor, was demonstrated to differ between lean and obese individuals [61, 62] and suggested to affect diabetes onset [63]. Given the role of nuclear receptors as sensors for a wide variety of endogenous and exogenous ligands they have been considered prime targets in some conditions. Notably, activation of Peroxisome proliferator-activated receptors (PPARs) was shown to modulate beneficial effects in such metabolic and inflammatory conditions. One in particular was PPARγ, identified as direct mediator of the insulin-sensitizing effects of the thiazolidinediones (TZD) class of drugs [53].

PPARγ, like PPARα and PPARβ/δ, forms constitutive heterodimers with RXRs. These heterodimers bind to direct repeats of two half-sites separated by 1bp (DR1) in which PPARγ occupies the 5' half of the motif [64]. While the PPARs recognize similar DR1 motifs, they show distinct specificities for ligands and vary substantially in their expression across different tissues. In addition to TZDs (such as Rosiglitazone (RSG)), PPARγ is activated by polyunsaturated fatty acids [65], prostanoids such as 15-deoxy-delta12,14-prostaglandin J2 [66], and components of oxidized low-density lipoproteins (oxLDL) such as 9-HODE and 13-HODE [67]. The different members of the PPAR family were originally identified on the basis of homology in *Xenopus* [68], and the murine *Pparg* gene was later cloned as *ARF6*. Two isoforms of PPARγ are reported which differ in their AF-1 domain, with PPARγ2 extending an additional 30 amino acids at its N-terminal compared to PPARγ1. Expression of PPARγ2 is almost exclusively restricted to the adipose tissue where it serves as a key regulator of adipogenesis and of lipid metabolism in general. In addition, PPARγ is expressed in myeloid cells (macrophages and dendritic cells (DCs)), the colonic epithelium, as well as in the lung and heart [41]. In addition to the insulin-sensitizing effects of TZDs, PPARγ ligands possess anti-inflammatory properties and attenuate the development of atherosclerosis. While the response to treatment with TZD might be similar across different organs, one has to keep in mind that the cellular environment and the supposed tissue-specific function of PPARγ might differ vastly between organs in vivo. While the intestinal epithelium is exposed to, and occasionally awash in, both dietary lipids and bacteria derivatives (e.g. short-chain fatty acids such as butyrate) the cellular environment for adipocytes is considerably different. So are the

metabolic functions of the two organs, with the epithelium mediating uptake and adipose tissue the storage of lipids. A third situation is exemplified by macrophages which do not exert obvious direct metabolic roles but whose dysregulation is frequently observed in association with metabolic disorders. Here too, ligands for PPARγ are derived from their specific cellular environment as phagocytosed material such as oxLDL contains PPARγ ligands and contributes to activation of PPARγ [67].

1.4.1 PPARγ in the intestinal epithelium

The single layer of intestinal epithelium that separates the outside from the underlying tissues has to facilitate the efficient uptake of nutrients and simultaneously serve as a physical barrier for the content of the gut lumen. This is especially important as gut contains a significant amount of commensal bacteria. Therefore integration of nutritional signals and bacterial cues is crucial to maintain gut homeostasis. It had been observed that PPARγ expression levels are high in the gut epithelium, especially in the colon [69] and this observation led to intensive studies on the role of PPARγ in the intestinal epithelium. To date, a number of effects mediated by PPARγ or its ligands in the intestine have been reported. Early studies focused on the role of PPARγ in the development of colonic tumors and provided evidence that PPARγ suppresses tumor growth [69, 70]. In addition several studies showed decreased proliferation of intestinal cells after TZD treatment in vitro [71]. It is therefore generally assumed that PPARγ has anti-proliferative effects in the intestine. However, data from the murine tumor model (APCmin) suggest that PPARγ activation might facilitate tumor progression under certain conditions [72, 73]. In vivo treatment of tumors with TZD showed no beneficial effect but it has been suggested that TZDs might work in a combinatorial fashion with other treatments, for example in combination with cisplatin based drugs [74].

PPARγ activation in the intestine, as in most other cell types, has been shown to attenuate inflammatory gene expression. Mice with dextran sulfate sodium (DSS) induced-colitis benefited from treatment with TZDs and epithelium-specific loss of PPARγ expression led to an exacerbated inflammatory response to DSS treatment [75]. These data are complemented by clinical reports that show decreased levels of PPARγ in the gut of patients with ulcerative colitis compared to healthy subjects [76].

Further evidence on the role of PPARγ as negative regulator of inflammatory processes was obtained through findings that the commensal bacterium *Bacteroides thetaiotaomicron* but

not the pathogen *Salmonella enteridites* attenuated NFκB-dependent inflammatory responses via PPARγ [77]. The molecular mechanism for this differential anti-inflammatory activity has not been characterized in detail, but it appears to involve translocation of PPARγ from the nucleus and possibly the sequestration of NFκB to the cytoplasm. The cross-talk between PPARγ and components of the intestinal flora is further emphasized by findings that the bacterial flora in the gut directly affects the activation status of PPARγ. Exposure of human colonic cells to *Enterococcus feacalis* increased PPARγ activity through modulation of its phosphorylation status [78].

Additional effects of PPARγ in the intestine have been documented. Most notably, the phenotype of the murine model for cystic fibrosis (Cftr -/-) which is characterized by mucus accumulation and colonic obstruction, was in part linked to a defect in PPARγ signaling [79]. Treatment of Cftr -/- mice RSG ameliorated bowel obstruction and increased survival through the PPARγ-dependent induction of genes involved in bicarbonate secretion.

1.4.2 PPARγ in adipocytes

White adipose tissue (WAT) serves both as a depot for storage of surplus energy in the form of lipids and as an endocrine organ involved in the regulation of energy homeostasis. Early studies demonstrated that the absence of PPARγ inhibited adipocyte differentiation in vitro and in vivo and established PPARγ as a key regulator in adipocytes. PPARγ induces a range of target genes in adipocytes most of which are implicated in the regulation of lipid metabolism and glucose homeostasis. Over-expression of PPARγ in fibroblast induces many of these targets and leads to lipid deposition. Most studies point to a major role of adipocyte PPARγ in mediating the insulin-sensitizing effects of TZDs. Improved insulin sensitivity in response to TZDs can be in part attributed to lower blood lipid levels resulting from the redistribution of lipids from the muscle and liver to the adipose tissue, thus storing excess lipids in a tissue where their physiological consequences are less severe [80].

In addition to the direct benefit of increased lipid uptake by adipocytes, PPARγ regulates the expression of several cytokines secreted from the adipose tissue, aptly named ‘adipokines’. These adipokines might act on other target tissues and therefore mediate insulin sensitivity in an endocrine manner. Notably, adiponectin is regulated by PPARγ in adipocytes [81] and its secretion is increased by TZDs [82, 83]. Adiponectin levels are inversely correlated with

insulin-resistance, while the level of another adipokine, resistin, which is reduced by TZD treatment, correlates with the degree of insulin-resistance [84].

PPARγ is expressed in brown adipose tissue (BAT) as well [85]. BAT serves an important function by regulating body temperature through catabolizing stored fatty acids [86]. These effects are in part mediated through activation of BAT-specific PPARg target genes (e.g. UCP-1) and it has been shown that the differential activation of target genes is dependent on the presence of the co-activator PCG-1α [87].

1.4.3 PPARγ in myeloid cells

Myeloid cells, namely macrophages and dendritic cells (DCs) are part of the innate immune system and are activated by a host of stimuli including pro-inflammatory cytokines and microbial components. Macrophages serve as effector cells and modulators of the inflammatory response. DCs serve as antigen presenting cells and thus connect the innate arm of the immune response to the adaptive arm. Activity of PPARγ in macrophages and DCs, has been intensively studied. Initial demonstration of PPARγ expression in foam cell of atherosclerotic lesions raised concerns about the safety of TZDs. These observations initiated a number of studies on PPARγ activity in macrophages, revealing a broad spectrum of PPARγ activity in macrophages and DCs. PPARγ activation leads to the up regulation of the scavenger receptor CD36 which facilitates the uptake of fatty acids and oxidized LDL [88]. Contrary to initial assumptions and despite the increased uptake of LDL, activation of PPARγ in foam cells proved to be beneficial in murine atherosclerosis models [89]. This effect appeared to be due to the fact that PPARγ simultaneously activates the cholesterol efflux pathway.

In addition to improving metabolic parameters (i.e. blood glucose levels), activation of PPARγ was shown to repress expression of pro-inflammatory genes in macrophages. The effects of PPARγ activation in myeloid cells in vivo are not fully understood, but several mechanisms have been suggested through in vitro and in vivo studies. The induction of pro-inflammatory genes (e.g. iNOS) by LPS was repressed by RSG in PPARγ+/+ but not in PPARγ-/- macrophages [90]. Subsequent molecular studies revealed that PPARγ represses the activation of AP-1 and NFκB by stabilizing the co-repressor containing complex and therefore inhibiting activation [49, 91, 92]. It was elegantly demonstrated that the repression of LPS-induced gene activation by PPARγ was mediated via ligand-dependent SUMOylation

of lysine 365 in the LBD of PPARγ [49]. The SUMOylated form of PPARγ is recruited to promoters occupied by NCoR-bound NFκB and in turn represses Ubc5-dependent proteosomal degradation of NCoR. This study showed, that PPARγ activation inhibits the exchange of the NFκB co-factor complex after activation by LPS and therefore suppresses gene activation. This mechanism has been shown in other cell types and by other factors [93, 94, 95], indicating that NR-mediated transrepression is a general mechanism employed by different NRs in a variety of cell types.

Interestingly, recent studies suggest additional models for the anti-inflammatory effects observed for PPARγ in vivo. PPARγ appears to exert strong anti-inflammatory effects by influencing the differentiation of different macrophage subtypes. Based on the expression of key molecular markers, macrophages are classified according to their function. M1 macrophages represent classical macrophages which, after stimulation with LPS or TNFa, secrete pro-inflammatory cytokines and mount an anti-microbial response. M2 macrophages are associated with regulatory functions and characterized by the expression of anti-inflammatory cytokines such as IL-10. Differentiation or polarization of macrophages towards the M2 phenotype is mediated by STAT6 in the presence of IL-4/IL-13 [96]. M1 and M2 macrophages are derived from the same progenitors in vivo, but their differentiation is not well understood yet. Mice with PPARγ-deficient macrophages (macPPARg-/-) showed increased insulin-resistance under high-fat diet compared to their wild type littermates [97, 98]. Treatment with RSG only partially restored insulin sensitivity, strongly suggesting that PPARγ plays a role in TZD-mediated insulin sensitization in macrophages. Characterization of the macrophage population revealed the depletion of a population resembling M2 cells in macPPARγ-/- mice. Besides, treatment of human individuals with RSG let to increase in monocytes with M2 characteristics in peripheral blood [99]. This anti-inflammatory cell type is therefore clearly depending on PPARγ activity. However, unlike through direct trans-repression, these cells typically express a cassette of genes involved in lipid metabolism [100]. Similarly, activation of PPARγ during differentiation of human DCs primarily led to the induction of genes involved in lipid metabolism [101]. This was later followed by an increase in the level of expression of CD1 molecules, which facilitated lipid antigen presentation [102]. These PPARγ-dependent effects appear to be indirectly mediated through induction of the retinoid acid signaling pathway. Together these studies suggest that the anti-inflammatory effects of PPARγ observed in myeloid cells are at least in part a result of its metabolic activity.

2. Aims of this thesis

The work presented in this thesis aimed to first establish how expression and activity of NRs were subject to regulation by environmental cues. With this information at hand the remaining work aimed to explore, on a genome-wide scale, the mechanism by which the NR PPARγ affects gene expression in different tissues and across species.

Specific Aims

Aim 1:

To characterize effects of exogenous signals such as the normal intestinal microflora on NRs and TLRs by combining expression analysis with the use of specific-pathogen free and germ free animals. (Paper I)

Aim 2:

To identify DNA binding sites and direct target genes of PPARγ in adipocytes, macrophages and intestinal cells. (Paper II, Paper III and Paper IV, respectively)

Aim 3:

To explore evolutionary changes in PPARγ-dependent gene regulatory networks in human and mouse macrophages. (Paper III)

Aim 4:

To characterize the properties of ligand induced activation of PPARγ in the intestinal epithelium and its consequence on cell growth on a genome-wide scale. (Paper IV)

3. Methodology

3.1 Genome-wide identification of TF binding sites using ChIP-seq

The studies presented in papers II-IV are based on genome-wide analysis of TF binding site locations and gene-expression profiles. Genome-wide identification of novel TF binding sites was facilitated through ChIP-seq which stands for chromatin immunoprecipitation followed by high-throughput massively parallel sequencing [103]. Due to its specificity and sensitivity ChIP has become a standard methods in molecular biology. It is used to test and quantify the binding of transcription factors or other DNA associated factors (e.g. histones) to specific genomic loci [104]. While protocols may vary in the details, generally, the Protein-DNA interaction is captured in vivo through cross-linking after which chromatin is prepared by either mechanic or enzymatic fragmentation of the DNA. Since the protein-DNA interaction is preserved, immunoprecipitaion with specific antibodies will enrich for DNA bound by the factor of interest. After isolation, the enrichment of specific genomic loci can be assessed by PCR a method usually limited to relatively few loci which have to be known in advance. Next-generation sequencing provides the means to identify binding loci in an unbiased manner as DNA enrichment can be simultaneously assessed throughout the genome. Several sequencing platforms are currently used for ChIP-seq with Illumina (Solexa) being the most often used. Depending on the platform and specific application each sequencing run produces between 15 and 30 million reads of 35-100bp in length per lane.

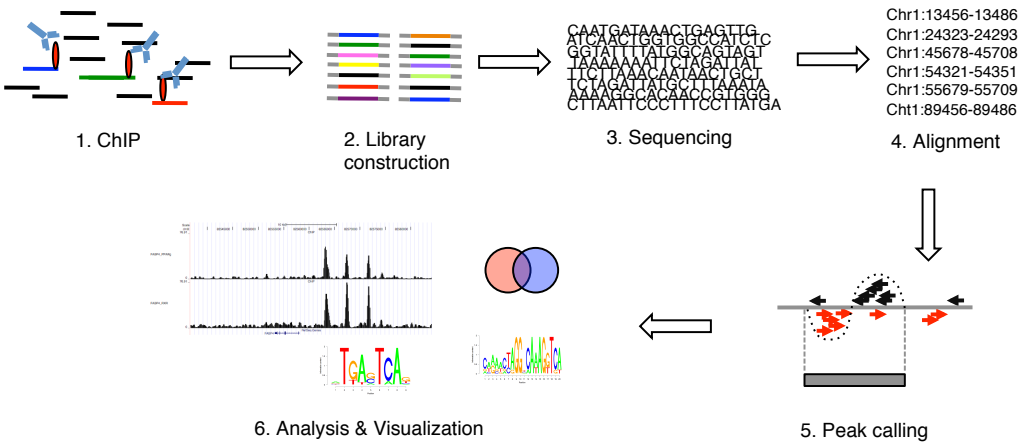

Figure 3. Schematic representation of the ChIP-seq procedure.

Binding regions can be identified by algorithms specifically that purpose such as MACS, SISSRS, CCAT [105-107]. These 'peak-callers' generally work by comparing the number of sequencing reads at a specific genomic location in the library obtained from ChIP material to a control library (e.g. non-enriched input material) in order to take into account background enrichment. The analysis of binding site locations provides a means to identify specific binding locations in a genome-wide manner and thus identify potential direct target genes of a given TF. Furthermore, the fact that often several thousands binding sites are identified for a single factor allows inferring general features that would go unnoticed otherwise. In this manner it is possible to identify the precise binding motif for the factor of interest and often also additional co-motifs, thus enabling the identification of associated binding factors (e.g. FoxA1 [108], CEBPα [109]). ChIP-seq provides a powerful tool for the unbiased interrogation of cis-regulatory regions across the entire genome because prior knowledge of binding locations for a particular factor is not needed. A drawback of this and similar ChIP-based methods lies in their dependency on highly specific antibodies that work well under the given conditions. Furthermore, the correct interpretation of these datasets provides challenges in the analysis and requires bioinformatics resources. Despite these small drawbacks ChIP-seq has proven to be extremely useful approach for genomic studies, witnessed by numerous genome-wide location analyses of TFs, histone modifications and co-factors have been published in the brief period following the first reports in 2007 [110, 111]. It is of note that the approach used in Paper II represents a variant of sequencing-based approaches for binding site identification (ChIP-PET [112]), the details of the sequencing and analysis process vary, but the principles are the same.

3.2 Microarray-based gene expression analysis

Microarray-based gene expression analysis enables the simultaneous interrogation of all annotated transcripts in a given genome. A variety of methods from different companies are available but most of the principles are similar. In a given sample, the amount of RNA (or DNA) sequences representing a specific gene is quantified by hybridization to an array containing probes complementary to all or a subset of relevant genes in the particular genome [113]. To facilitate detection, the oligonucleotides in the sample are usually labeled with fluorescent dyes. The amount of RNA is proportional to the measured intensity. To compare intensities between different samples and to identify transcripts with significantly altered expression between samples several processing steps are required. First, the intensities need

to be normalized to account for differences between samples due to background variation (e.g. scanning intensity). Many different procedures have been described and the right choice depends on the system that was used and the actual distribution of the intensity values. Second, the expression levels of transcripts in different samples are compared and significance of the difference is tested. The appropriate method depends on the experimental design and the exact aim of the study (reviewed in: [114]). Due to the inherent experimental noise and to obtain reliable estimates each biological condition should be represented by several replicates. Genome-wide studies of gene-expression using micro arrays have provided deep insights into gene regulation and have been instrumental in the detection of aberrant gene expression patters that are associated with diseases [115]. However, akin to the development of ChIP-seq, direct sequencing of the whole transcriptome (called RNA-seq) provides a powerful alternative to hybridization-based methods [116-118]. RNA-seq has several advantages; most importantly it is unbiased in the detection of transcripts. While microarrays need to be predesigned, direct sequencing enables the identification of novel transcripts and it accurately detects splice-isoforms. In addition, deep sequencing enables the detection of nucleotide variants, a fact that is especially valuable in the study of transcriptomes from cancer samples. At the current time costs for sequencing are still relatively high and array applications are therefore useful especially in screen expression patterns on a large number of samples. However, it seems likely that most microarray-based methods will be replaced with next-generation sequencing approaches in the future.

4. Results and Discussion

This thesis consists of four papers all attempting to elucidate mechanisms by which NRs, and PPARγ in particular, respond to environmental cues and regulate signal-specific gene-expression in different tissues. Paper I investigates the role of NRs as second messengers in the communication between the intestinal epithelium and the gut microbiome. The subsequent Papers II-IV are specifically focused on the identification PPARγ binding site selection and target genes in different tissues, namely in adipocytes, macrophages and intestinal epithelium.

4.1 Paper I: Gut flora, Toll-like receptors and nuclear receptors: a tripartite communication that tunes innate immunity in large intestine

The intestinal epithelium represents a particularly instructive system to study the complex mechanisms that facilitate the integration of different signals. The intestinal lumen contains about 10^{14} bacteria, about ten times more bacteria than cells present in the human body. Under normal condition these 'commensal' bacteria pose no threat to the host and some were shown to be beneficial provided they are kept in check and within the lumen of the gut. Therefore modulating the immune response in a way to tolerate commensal bacteria while at the same time recognizing and eliminating pathogenic bacteria provides a formidable challenge to the epithelium and cells of the innate and adaptive immune system [119]. Cells of both the adaptive and innate arm of the immune system are important in mediating the correct responses [120] and are thought to play an important role in establishing oral tolerance. More recently it was recognized that the epithelium acts as a first line of innate immune cells and senses the presence of bacterial components or antigens. In order to sense these molecules, epithelial cells appear to use at least in part the same pathways documented in adaptive immune cells. Many of these receptors including members of the family of Toll-like receptors recognize generic antigens shared by both beneficial and pathogenic bacteria. Interestingly, signaling through these pathways appears to be important for the homeostasis of the intestinal epithelium and innate immunity. In paper I, we investigated if the presence of gut microflora affected expression levels and therefore potentially modulated the activation threshold of nuclear receptors. In a previous study, PPARγ was shown to attenuate NFκB activation in response to the commensal *Bacteroides thetaiotaomicron* but not in response to the pathogen *Salmonella enteriditis* [77]. In addition, several NRs exhibit anti-inflammatory properties in a variety of cell-types. These data made NRs promising candidates to mediate

some of the signaling important for gut homeostasis in the presence of normal microflora. Therefore, we set out to characterize the expression of NRs and TLRs in mice with normal flora (specific-pathogen free; SPF) and germ-free mice (Paper I).

Expression of 5 NRs in the colonic epithelium is affected by the resident gut flora

Using realtime qPCR with primers designed against all 49 NRs present in the mouse genome we detected the expression of 37 of them in the colonic epithelium in both SPF and GF animals. Interestingly, we found that the expression levels of five of the expressed receptors were modulated by the resident flora in the colonic epithelium. Nur77 and GCNF expression was increased in SPF mice, while expression of RORγ, LRXα and CAR was decreased in these mice compared to animals raised under germ-free conditions. TLR expression remained relatively unaffected by the intestinal flora with the exception of TLR2 and TLR5 which were strongly induced under SPF conditions. Since TLR2 showed strong differential expression we asked if TLR2 might mediate some of the gut flora associated effects observed on NR expression. Using TLR2-/- mice to test this hypothesis we found that CAR was significantly up regulated compared to wt controls. These data suggested that the down-regulation of CAR in SPF mice may possibly be mediated via a TLR2-dependent mechanism.

Nuclear receptor expression change after inflammatory challenges

Having characterized the expression levels of NRs and TLRs in the presence and absence of normal bacterial flora we then asked if the expression of NRs and TLRs might be altered in the presence of pathogenic and inflammatory stimuli. Exposure of SFP mice to the pathogen *Yersinia pseudotuberculosis* (YP) resulted in expression changes in 11 out of the 37 expressed NRs (PPARγ, GR, Rev-ErbA, PPARα, and RORα among others). In contrast, only the expression of CAR was altered in GF mice upon YP exposure, CAR expression in SPF mice remained unaffected. In addition, the expression of TLR2 and TLR5 was significantly repressed in SPF mice upon YP exposure.

Together, these data demonstrate that the normal flora influences the expression level of several NRs as well as TLRs and establishes a potential link between colonialization and gut-specific metabolic functions (e.g. via LXRα and CAR). Similar to humans, mice are born with a virtually germ-free gastrointestinal tract, followed by rapid post-natal colonization. The maturation of this organ in vivo is therefore closely linked to its first encounters with the microflora. It is tempting to speculate that the presence of normal flora serves as an environmental cue for the intestinal epithelium (potentially through pathways of the innate

immune system) and helps to establish correct expression levels of receptor molecules such as NRs and TLRs. How the expression changes are induced by the bacteria is presently not clear, but the correlated expression of CAR and TLR2 indicates that TLRs are potential candidates to relay the microbial signal to NRs. Additional factors, such as changes in metabolite concentrations or secondary effects of the gut development are likely to play a role as well. Here, we addressed one layer of the complex interplay between the gut flora and cellular sensors, namely the expression levels of NRs and TLRs. Given that many NRs are ligand-activated it is conceivable that their activity is modulated through microflora-induced changes in ligand concentration and composition.

4.2 Genome-wide exploration of PPARγ activity across different tissues

In the papers II-IV PPARγ was used to explore genome-wide binding patterns and target genes across different tissues. These studies cover a broad spectrum of PPARγ activity reported in human and mouse. As described in the introduction chromatin environment and expression level of TFs differ between tissues factors that might have strong influence on TF binding. Furthermore, in Paper I we directly demonstrated that the tissue environment might influence expression levels of NRs. Genome-wide binding analysis of PPARγ in a single tissue would therefore only results in very limited insights into PPARγ biology, potentially restricted to the particular tissue. However, combining studies in multiple tissues has the potential to reveal common rules of PPARγ-dependent gene regulation and to identify mechanisms that facilitate tissue-specific activities. In paper II, we investigated PPARγ binding sites and target genes in murine adipocytes. In paper III, the focus shifted towards PPARγ in macrophages. Finally, in paper IV we investigated the PPARγ DNA binding landscape in intestinal epithelial cells. Common to all these studies was our approach to combine PPARγ binding site data and expression analysis to identify direct target genes in all three tissues

4.2.1 Paper II: De-novo identification of PPARgamma/RXR binding sites and direct targets during Adipogenesis

Despite the well-documented role of PPARγ during adipogenesis and in mature adipocytes, relatively few direct PPARγ targets have been identified in adipocytes. Therefore, in Paper II we set out to identify PPARγ binding sites in a genome-wide manner and to combine these

data with gene-expression analysis to comprehensively identify PPARγ target genes during adipogenesis.

Identification of differentially expressed genes during adipogenesis in 3T3-L1 cells

We used the pre-adipogenic 3T3-L1 cell line as it represents a well established model system of adipogenesis. Adipogenesis was chemically induced in 3T3-L1 cells and differentiation was confirmed by OilRed-O staining and by western blot for adipocytes specific genes (FABP4, Apo-A1). To obtain genes whose expression levels changed during adipogenesis, we collected mRNA from 3T3-L1 cells treated PPARγ-specific and control siRNA over the course of adipocyte differentiation. Genes with significantly altered expression between the two conditions were identified using microarrays. In total we found about 1700 genes that were differentially expressed during adipogenesis, 40% of which were repressed and 60% induced after PPARγ siRNA treatment. Regulated genes were associated with biological processes annotated in the PANTHER database [121, 122] and significantly enriched processes were obtained for both groups. Genes down-regulated after PPARγ knock-down (KD) were strongly associated with lipid and fatty acid metabolism, while genes induced by PPARγ KD were significantly associated with processes such as development and cell division.

Genome-wide mapping of PPARγ and RXR binding sites

The gene expression analysis revealed potential PPARγ target genes but since ablation of PPARγ expression inhibits adipogenesis entirely, a substantial proportion of regulated genes might not be directly targeted by PPARγ. To identify direct PPARγ target genes we sought to map all PPARγ and RXR binding sites occupied in mouse adipocytes. We obtained chromatin from ChIP against PPARγ and RXR, in 3T3-L1 cells on day 6 after induction of adipogenesis. After sequencing we identified a total of 7700 binding sites (at FDR $\leq$0.05) for PPARγ and its binding partner RXR. The analysis and validation of PPARγ and RXR binding data confirmed the constitutive dimerization of PPARγ with RXR since RXR was detectable by ChIP-qPCR at virtually all PPARγ binding sites. PPARγ/RXR binding sites were scattered throughout the genome, but showed significant enriched towards transcriptional start sites, as had previously been observed for other factors [123]. Binding of PPARγ/RXR was strongly associated with genes that were normally induced during adipogenesis, but not with repressed genes. Using luciferase reporter constructs we demonstrated that several novel PPARγ/RXR regions were indeed able to activate transcription in a PPARγ-dependent manner.

Novel direct PPARγ targets are potentially important during adipogenesis.

Direct PPARγ target genes are enriched for fatty acid and lipid metabolism confirming the role of PPARγ as the transcriptional master regulator of adipogenesis. Interestingly, PPARγ seemed to broadly regulate adipocyte metabolism by directly inducing enzymes with functions in lipid and fatty acid metabolism, rather than activating secondary TFs. We reasoned that direct target genes of PPARγ that had not been characterized or at least not been investigated in the context of adipogenesis might play an as of yet unrecognized role during adipocyte differentiation. To test this hypothesis we selected 20 candidate targets and used specific siRNA to assess their role during adipogenesis in 3T3-L1 cells. Indeed, 6 out of the 20 genes showed significant reduction in lipid accumulation, suggesting that they might be functional during adipogenesis. Knock-down of Smaf1 (Adipogenin), Pim3, Mnk2, Agt, Fsp27, and Pdzm3 led to significantly reduced OilRed-O staining. It will be of interest to further decipher mechanisms by which PPARγ regulate adipogenesis by investigating the exact roles of selected downstream target genes. It remains to be established whether the observed effect on lipid accumulation is a primary effect or whether the decreased lipid accumulation is a downstream effect of differentiation processes. Finally, it will be important to investigate if these genes or genes with similar roles in lipid metabolism are regulated in other tissues with high PPARγ activity as well. These questions are higly relevant in order to understand PPARγ activity on a systemic level and these data provide a starting point for future studies.

4.2.2 Paper III: Interplay between PPARG and PU.1 during evolution shape divergent PPARG binding landscapes in human and mouse

The study presented in Paper II focused on the central role of PPARγ in murine adipocytes and yielded relevant insights into the role of PPARγ during adipogenesis. However it was not clear how binding sites vary between different tissues and species. We were particularly interested to address these questions in macrophages because its anti-inflammatory, anti-atherogenic and TZD-sensitizing effects that had been demonstrated in these cells [49, 89, 91, 97-99]. These data provided a strong rational to study PPARγ in macrophages and similar to Paper II, we set out to study PPARγ activity in a comprehensive manner by combining PPARγ binding and expression data.

Direct interpretation of binding and gene expression data is often complicated by the fact that many, if not most, binding sites appear to be non-functional. By comparing data from different species one would be potentially able to filter out binding sites and target genes that are less relevant and therefore evolutionarily not retained. Initial cross-species comparisons on few TFs revealed surprisingly little similarities in the binding locations across different mammalian [124-126] or vertebrate species [127]. Such comparisons help to understand the forces that shape binding landscapes during evolution and might guide the identification of functional binding sites. In Paper III we reported a genome-wide localization analysis of PPARγ, RXR and PU.1 in in the well characterized human myeloid cell line THP-1 and presented an interspecies analysis of PPARγ binding and gene regulation in human and mouse macrophages. In our case such an approach was facilitated by the fact that PPARγ binding data in murine macrophages was publicly available [128].

Exceedingly low PPARγ binding site conservation contrasts a high degree of functional conservation

Using ChIP-seq we obtained genome-wide binding profiles for PPARγ and RXR in differentiated THP-1 cells. In total, we identified 2133 loci which showed simultaneous enrichment for PPARγ and RXR. In order to assess the proportion of sites shared between both species we compared the human PPARγ binding sites to PPARγ sites reported in mouse macrophages [128]. We found that only a very small proportion (5%) of the human binding sites also showed PPARγ binding at homologous loci in mice which is in agreement with previous studies on other TFs [127].

Previous genome-wide cross-species comparisons suggested that gene-regulatory networks might be conserved, despite an apparently high divergence on the level of binding sites [124, 129, 130]. In light of the extremely low overlap between PPARγ binding sites in human and mouse we therefore tested if the overlap was higher on the level of putative PPARγ target genes. We found that about 21% of human PPARγ target genes were also targeted in mouse and therefore shared between the two species. As expected, most of these targets were 'indirectly' shared (i.e. the adjacent binding site in both species are not conserved) contributing about 15% to the shared target genes while 6% of human targets were adjacent to at least in conserved binding site. These data suggest that the massive changes in the binding site landscape might be in part compensated by the emergence of evolutionary novel binding sites to preserve the regulatory network. Not surprisingly, we found that directly shared targets were much more likely to be regulated by RSG (> 30%) as compared to human

specific targets (around 12%). Importantly, indirectly shared target genes showed enrichment for RSG responsive genes compared to human specific genes as well.

Preserved macrophage-specific configuration of cis-regulatory modules

Genome-wide surveys of TF binding revealed combinatorial binding patterns for many factors[131, 132]. More specifically, PPARγ was reported to bind together with CEBPα in adipocytes [109] and recently, PU.1 was identified as determinant for PPARγ binding in mouse macrophages [128]. We therefore asked if PU.1 acts as a sequence-specific determinant for PPARγ binding in human macrophages as well. We found strong enrichment of the known PU.1 binding motif at PPARγ binding sites and subsequent PU.1 ChIP-seq confirmed the presence of PU.1 at 60% of all PPARγ sites in human macrophages. These data suggest that the composition of the cis-regulatory modules is maintained in human and mouse [132, 133]. These data are in agreement with other studies that demonstrated how dramatic changes in regulatory regions of the genome are contrasted by small changes in the cellular machinery that reads-out the genome [134].

PU.1 potentially facilitates PPARγ binding site turnover during evolution

Given the high degree of co-occupancy between PPARγ and PU.1 in both species and the apparent role of PU.1 as a determinant of PPARγ binding in macrophages, we wondered if the emergence of novel and functionally relevant PPARγ binding sites during evolution was associated with PU.1. By comparing PPARγ-PU.1 modules at shared genes and human-specific genes we found indeed evidence for a role of PU.1 in PPARγ binding site turnover. Based on the assumption that indirectly shared genes evolved from an ancestral state with a single PPARγ binding site, the evolutionary history of indirectly shared targets genes would includes a state in which (at least) two PPARγ binding sites are found adjacent to the particular gene in one of the species.

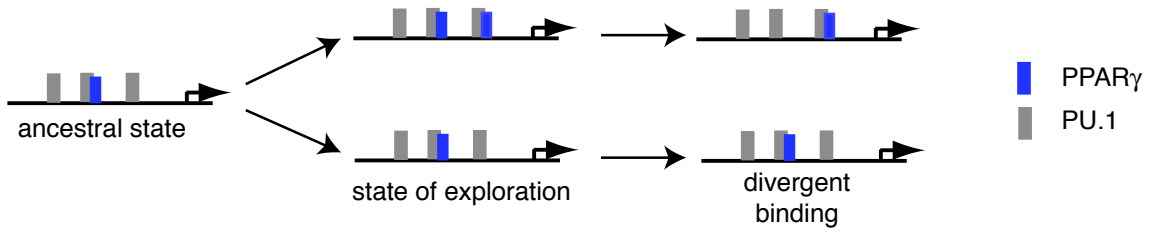

Figure 4. Model for PU.1 assisted PPARγ binding site turnover

A role of PU.1 in the exploration of novel sites would be indicated by a higher proportion of conserved/old PU.1 sites in PPARγ-PU.1 modules at indirectly shared genes. Not

surprisingly, human-specific PPARγ-PU.1 sites showed contained around 20% of conserved PU.1 sites, comparable to the genome-wide average of PU.1 binding site conservation. In contrast PPARγ-PU.1 binding sites where the PPARγ binding site was shared between human and mouse usually also contained a shared PU.1 binding sites (>80%). Importantly, in agreement with our hypothesis, PPARγ-PU.1 binding sites adjacent to indirectly shared targets showed significant increase in conserved PU.1 binding sites. It is of note however, that directly conserved target genes appear to be the functionally most relevant group likely representing a small core regulatory network which is conserved between human and mouse macrophages, mostly contains genes involved in lipid metabolism. This is particularly interesting since alternatively activated macrophages or M2 macrophages are characterized by the expression of cassette of genes involved in lipid metabolism and this metabolic activity appears to be important for the anti-inflammatory potential of these macrophages [100]. It is tempting to speculate that the cassette of conserved PPARγ target genes might be linked to this function. Together these findings suggest that genome-wide cross-species comparisons of TF binding data might provide a useful approach to identify biologically important loci and gene targets. By combining H3K27Ac data from human and mouse during adipogenesis such a strategy helped to identify novel regulators of this process [125].

4.2.3 Paper IV: Integrative genomic analysis of PPARγ activity suggests broad antagonism of WNT/β-catenin signaling in intestinal epithelial cells

Strong expression of PPARγ in the intestinal epithelium, especially of the colon, motivated research into the role for PPARγ in the gut [135]. Subsequent studies found that PPARγ ligands inhibited tumor growth and that PPARγ acted as tumor suppressor with anti-proliferative effects in colon cancer [69, 70]. Treatment of human colonic tumor explants with PPARγ ligand in nude mice showed a significant decrease in tumor size. However, the molecular mechanisms underlying these effects remain poorly understood and furthermore some studies added controversy by showing that treatment of mice genetically susceptible to colon cancer (APCmin-/-) with RSG led to a significant increase in colonic tumors [72]. In addition PPARγ was reported to attenuate inflammation in the gut [75]. To investigate the mechanisms by which PPARγ regulates genes expression in the gut epithelium and to identify direct target genes, we combined gene-expression profiling with genome-wide binding site mapping for PPARγ in the human intestinal cell line HT-29.

PPARγ ligand induces a variety of genes and suppresses cell growth in HT-29 cells

We found that treatment with RSG affected the expression of 749 genes, suggesting strong PPARγ activity in HT-29 cells. A considerable proportion of the 408 induced genes was associated with lipid metabolism while repressed genes appeared to be enriched for functions associated with cell proliferation and differentiation. In agreement with the gene expression data we found that treatment with RSG inhibited cell growth in these cells, confirming observations in vivo and on some cell lines. These data demonstrated, that PPARγ controls a relatively large set of genes and that exposure to ligand affected cell growth in these cells, HT-29 cells therefore provide a good model to study PPARγ activity in the intestinal epithelium.

PPARγ directly binds to the majority of induced genes

Gene-expression studies on their own are limited by the fact that distinction between direct and indirect targets is not possible. Furthermore, to infer the molecular mechanisms involved in gene activation or repression mediated through the particular factor knowledge of the precise binding location for a given TF is imperative. To complement the gene-expression data we therefore obtained PPARγ and RXR binding sites in HT-29 cells using ChIP-seq and identified a set of 6608 shared PPARγ/RXR sites. These sites demarcated most of the genes induced by PPARγ ligands (~70%) and almost all of the strongly induced genes. Similar to the effects in other tissues (Paper II and III), PPARγ binding showed significantly less association with genes repressed by RSG.

PPARγ and TCF4 binding sites coincide on a genome-wide scale

Genome-wide binding studies for PPARγ and other TFs identified additional determinants of binding through analysis of enriched motifs at binding sites (Paper III). Scanning all PPARγ binding regions in HT29 cells we found strong enrichment of binding motifs for AP1, SP1, ETS and KLF4. Remarkably, we also found that the motif for TCF4, a key downstream effector of WNT/β-catenin signaling was significantly enriched at PPARγ binding sites. Furthermore, by using publicly available data for TCF4 binding in intestinal epithelial cells [136] we confirmed the physical presence of this factor at approximately 25% of PPARγ binding sites.

RSG and β-catenin regulate a set of genes in opposing fashion.

The co-occurrence of PPARγ and TCF4 on the level of binding sites prompted us to directly compare genes regulated by RSG and genes regulated after knock-down of β-catenin[136]. When taking into account the directionality of the expression changes in both groups we found that PPARγ and WNT/β-Catenin/TCF4 signaling affected the expression of the majority of shared target genes in opposite directions. Most of the shared target genes were induced by RSG and repressed by β-catenin. These targets included genes associated with epithelial differentiation (KRT20, CDH1) and potential tumor supressors (Rhob1 and TOB1). Of note, among RSG-repressed targets that appear to be induced by β-catenin we found IL-8 and MMP7 both of which are implicated in the development of colon cancer and IL-8 also plays an important role in inflammatory diseases of the intestine.

Lastly, we sought to investigate the nature of the relationship between PPARγ and TCF4 at shared binding sites. We wondered if co-binding was a consequence of accessible chromatin with little direct functional implications or if the co-binding has effects on transcriptional activity of either factor. Notably, the frequency of PPARγ/TCF4 co-binding in adjacency of RSG repressed genes significantly increases to 37% compared to the genome-wide average of approximately 25%. While our data do not allow us to infer molecular mechanisms it is tempting to speculate that the increased co-binding of PPARγ and TCF4 points towards a regulatory mechanism involving inhibition or competition between the two factors.

This manuscript highlights several aspects of PPARγ activity in the intestine. PPARγ appears to directly target several genes with anti-growth and tumor suppressor activity. It will be of interest to study some of theses target genes in more detail. In particular it will be important to investigate if the mechanisms by which PPARγ and β-catenin signaling antagonize each other are mediate through direct interaction or by indirect activation of a shared set of genes. In addition, the regulatory role of PPARγ in glucose homeostasis and lipid metabolism in the intestinal epithelium deserves more attention. PPARγ induced an array of genes associated with lipid metabolism, many of which are ligand responsive in other tissues as well. In addition, NR1H4(FXRα) whose activity was previously shown to lower triglyceride levels [137] was strongly induced by RSG suggesting that different layers of metabolic control are linked through PPARγ.

4.2.4 Insights from genome-wide mapping of PPARγ binding in different tissues

The studies presented in this thesis all aimed to further our understanding on the activity of NRs on a genome-wide level. In Paper I it was demonstrated that exogenous factors (i.e. gut microflora) are able to influence expression of NRs and, in a broader sense, potentially modulate the activity thresholds for NRs. In the subsequent three papers, we explored how activation PPARγ, relates to gene regulation in different tissues. Our results yielded novel insights in the activity of PPARγ in adipocyte, macrophages and the intestinal epithelium. In addition these studies also provides useful insights into general principles underlying PPARγ binding site selection and gene activation. Here, I discuss the findings common to all studies and insights that were gained from combined analyses of these studies. On a basic level, we confirmed biochemical data by showing that PPARγ exclusively forms heterodimers with RXR that bind to the canonical DR1 element. Besides, PPARγ binding seems to be predominantly associated with ligand-dependent (in this case RSG) transcriptional activation rather than with repression. However, since I did not simultaneously activate additional transcription factors (such as AP-1 and NFκB) this observation is potentially limited to this particular experimental system and transrepression effects are likely precluded. It is of note that I did not observe transrepression of NFκB activation in THP-1 cells in a set of preliminary experiments, potentially as a consequence of the rapid decrease of PPARγ mRNA and protein levels in response to LPS stimulation.

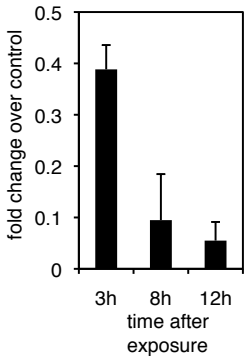

Figure 5: PPARγ expression decrease dramatically after LPS exposure in THP-1 cells. *PPARG mRNA levels were compared between LPS-treated cells and cells treated with vehicle.*

The down regulation of PPARγ following LPS stimulation is in concordance with observation by others [138], however transrepression by PPARγ has been clearly demonstrated in different cellular contexts [49, 90]. Two considerations might be important to understand the apparently contrasting effects. Firstly, subsequent studies demonstrated transrepression of NFκB and AP-1 by additional NRs targeting co-repressor checkpoints in a similar manner. Secondly, since the molecular mechanism for PPARγ-dependent transrepression was

demonstrated by over-expression systems [49], one might speculate that the in vivo mechanism might deviate. It is therefore entirely possible that additional regulation of PPARγ levels under the same conditions restrict this effect to a smaller set of conditions in vivo.

In the three cell types studied, PPARγ appeared to mainly regulate gene sets associated with lipid metabolism. Compared to simpler organisms (e.g. Drosophila) metabolic functions are highly compartmentalized in mammals (fat body vs. adipocytes, liver and macrophages), the shared expression cassette between adipocytes, macrophages and intestinal epithelium might indicate some overlapping functions pertinent to lipid metabolism between these organs. It is possible that this co-regulation of a gene cassette of lipid metabolic genes by PPARγ reflects the evolutionary relationship between these organs and demonstrated the ancestral function of PPARγ. Many genes are targeted by PPARγ in all three tissues and both in human and mouse macrophages. PDK4, ANGPTL4 (FIAF), ADFP are examples of this core set of PPARγ target genes. Given the strong conservation of metabolic functions among PPARγ targets it is possible that many of the systemic features of TZDs, including anti-inflammatory effects, are mediated through the primary metabolic role orchestrated by PPARγ. In the case of macrophages this was demonstrated directly [100] and similar mechanisms might be in place in other tissues as well. Interestingly, additional genes with metabolic roles might be included as target genes in a tissue-specific manner. NR1H4 appear to be targeted by PPARγ only in the intestinal epithelium, and activation of NR1H4 has been associated with lowered triglyceride levels and increased insulin sensitivity. It is interesting to speculate that different layers of metabolic control are linked through PPARγ. Activation of PPARγ in the gut (e.g. through ligands in dietary lipids or secreted from bacteria) might generally increase transcription of genes involved in metabolic processes such as NR1H4 which in turn control regulatory circuits of their own (bile acid metabolism in the case of NR1H4). One might speculate that during evolution a set of core targets can be amended in tissue-specific manner to include (or exclude) target genes that are additionally required.

The mechanisms by which tissue-specific inclusion or exclusion of target genes into the PPARγ-dependent gene regulatory network might work were brought to light through comparing binding profiles of PPARγ across different tissues. This comparison showed that most binding sites are tissue specific and tissue-specific binding sites are often associated with additional TFs the expression of which might be regulated in a tissue-specific manner. Earlier studies identified CEBPα as an important determinant of PPARγ binding in adipocytes [128, 139].

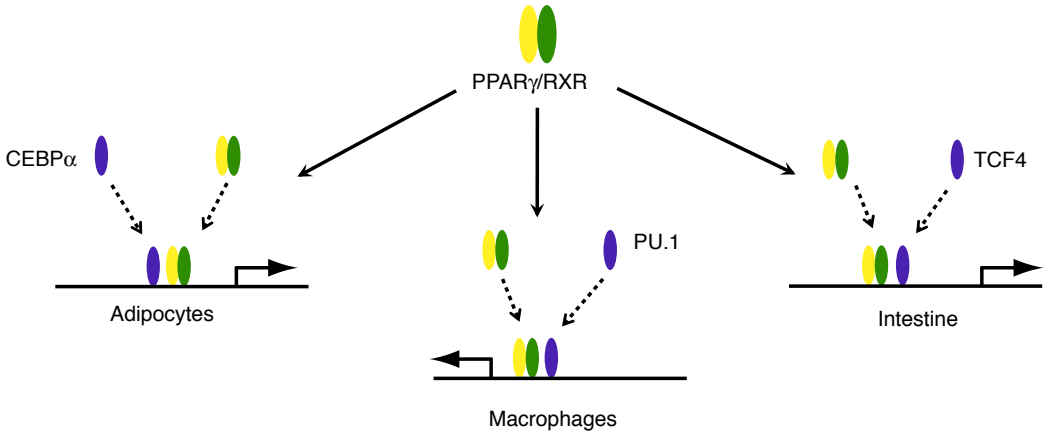

Figure 6: Schematic of the tissue-specific PPARγ/RXR binding configurations. *Illustration of tissue-specific PPARγ binding site configurations.*

In macrophages, our studies in humans and studies in mice [128] identified PU.1 as a determinant for PPARγ binding in macrophages, and PU.1 expression is almost exclusive to the hematopoietic lineage. PPARγ binding sites in the intestinal epithelium showed enrichment for the TCF motif and I showed that TCF4 co-occupied a substantial proportion of PPARγ binding sites. According to the enrichment of TF binding motifs at PPARγ binding sites, several additional factors are likely present at PPARγ binding sites and important for binding in different tissues. In contrast to the tissue-specific co-occupants AP-1 appears to be present at most PPARγ binding sites as suggested by the presence of the AP-1 motif. Our data, together with other studies suggest a model in which general and tissue-specific co-factors engage in a combinatorial manner to establish (tissue-specific) enhancer regions [132]. Since each cell and tissue in any given individual organism contains the same genome, differential binding of PPARγ cannot depend on the presence or absence of a PPARγ binding motif. However, the presence of additional motifs, potentially bound by factors only present in a specific tissue would provide a basis for tissue-specific accessibility of chromatin. Such combinatorial codes are likely to be identified in more detail by comprehensive ChIP-seq studies in different tissues and species [23]. Apart from the presence or absence of sequence-specific TFs several studies also demonstrate the ability of co-factors to direct tissue-specific gene expression programs for PPARγ. In BAT PCG-1α is required for activation of the BAT-specific induction of UCP-1 by PPARγ [87], while Hic-5 was shown to direct PPARγ-dependent gene expression towards a program of epithelial differentiation in the gut [140]. Understanding the specificity of these interactions is not straightforward as the co-factors

themselves do not bind specific DNA sequences and should therefore not discriminate between different PPARγ binding loci in a similar manner as associated sequence-specific TFs. However, a possible explanation is that such co-factors recognize PPARγ in combination with a second TF, i.e. recognize tissue-specific binding modules (e.g. PPARγ-PU.1 or PPARγ-TCF4). Besides, a potentially very important determinant for such interactions is the spatial relation between different parts of the genome. The complicated three-dimensional organization of the genome has recently been inferred for few cell lines from sequencing data describing the points of interaction between genomic loci. These data suggest a highly organized structure that shows tissue-specific features [141]. On a finer scale, focusing on the NR ERα, a genome-wide map of long-range interactions showed extensive interactions between loci in cis- and trans [142]. Association of co-regulated genes in 'transcription factories' might be an additional mechanism by which co-factors (such as Hic-5 and PCG-1α) enrich at certain promoters while not at others [143].

5. Concluding remarks and perspectives

The availability of genome sequences from humans and other mammals species in combination with the enormous advances in sequencing technology have transformed many aspects of modern biology. As a consequence, many areas of molecular biology evolved from focusing on individual genes to developing novel concepts on a genome-wide scale. The work presented in this thesis might provide an entry point to study PPARγ activity in a systemic manner by describing a genome-wide view of PPARγ binding sites and target genes across different tissues. To date, most studies on chromatin structure and TF binding are based on cell lines or model systems, however improvements in ChIP-seq technology enable the use of minimal DNA amounts e.g. derived from in vivo material [144, 145]. Using similar approaches PPARγ binding site distribution and gene regulation could be assessed in a variety of different cell-types derived from human blood or tissue biopsies. Such an approach would overcome many of the limitations of the use of cell lines.

Following the observation that PPARγ seems to activate a core set of metabolic genes in different cell types after RSG treatment, the next step would be to monitor this activity in vivo. It would be important to assess if these genes are controlled in a similar manner or if there may be tissue- and ligand-specific differences. While it is currently difficult to address the contribution of specific ligands in vivo, additional characterization of the chromatin landscape and associated co-factors might provide insight into tissue-specific gene regulation.

Monitoring PPARγ binding, chromatin landscapes and co-factor recruitment in response to different ligands at different time-points will provide insights into ligand- and gene-specific dynamics at a previously unachieved resolution. Using ERα, meticulous experiments on a single locus level suggested that cycling of ERα binding occurs after ligand-dependent activation and that the recruitment and replacement of different co-factors during such cycles plays an important role during gene activation [58]. In addition it will be important to address the role of PTM on the NR (as well as at co-factors). Biochemical as well as epidemiological data show that PTMs are key regulatory features for specific states [146].

An important finding of the studies presented here and other genome-wide TF binding site studies is that the number of binding sites usually exceeds by far the number of regulated genes. Binding sites show some enrichment for TSS and gene body but are in general scattered across the genome. This raises the question of how many of these binding sites are actually functionally relevant in regard to gene activation. Indeed, establishing the link

between binding sites and their target genes is a major challenge for the interpretation of these data. It is likely that a considerable proportion of binding sites are merely a consequence of accessible chromatin and the casual occurrence of binding motifs. Many studies only consider binding sites in proximity to TSS and therefore increase the likelihood of functional association between TF binding and gene regulation. However, even distant binding events were shown to be involved in transcriptional regulation [147, 148]. Novel techniques developed to capture long-range interactions on a genome-wide scale will be helpful in closing the gap between the mapping of binding sites and the identification of direct target genes [141, 142]. In theory these techniques would also allow the detection of interactions in trans. Another way to enrich for functional binding sites was demonstrated in Paper II where evolutionary retention of binding sites was applied as a filter to separate functional binding sites from sites that were more likely to represent noise. This approach is likely to identify a set of high-confidence functional binding sites and present a useful method to characterize biological systems. However, such methods are insensitive to functionally important binding sites that are not conserved.

A novel approach that combines population genetic approaches with data on TF binding and chromatin structure has the potential to become a very important method as one might be able to overcome some problems of large-scale population studies. Genome-wide association studies (GWAS) have identified hundreds of loci associated with a variety of diseases and traits (e.g. Obesity, Diabetes, Blood lipid levels [149, 150]), however they are often limited by different factors. One major limitation is that the sample size has to be very high (tens of thousands of samples) to discover low frequency variants or variants associated with small risks, making these studies very costly and in many cases unfeasible. Another point is that even if sequence variants are found which strongly associate with a given trait it is often not clear whether the variant itself (or which of the variants in a given locus) is causally related to the trait. These complications could in part be overcome by including TF binding data or chromatin maps into the analysis of GWAS. Several studies indicate that single-nucleotide polymorphisms (SNPs) associated with traits in a given tissue are often located in accessible and active regulatory regions in that particular tissue, but not in others [23]. Indeed, preliminary analysis of PPARγ binding sites data in the colonic epithelium showed PPARγ binding sites coinciding with several loci associated with predisposition to colon cancer [151, 152]. In contrast PPARγ binding in macrophages was enriched adjacent to genes implicated by GWAS in the control of blood lipid levels [149, 150, 153-155] (Fig. 7).

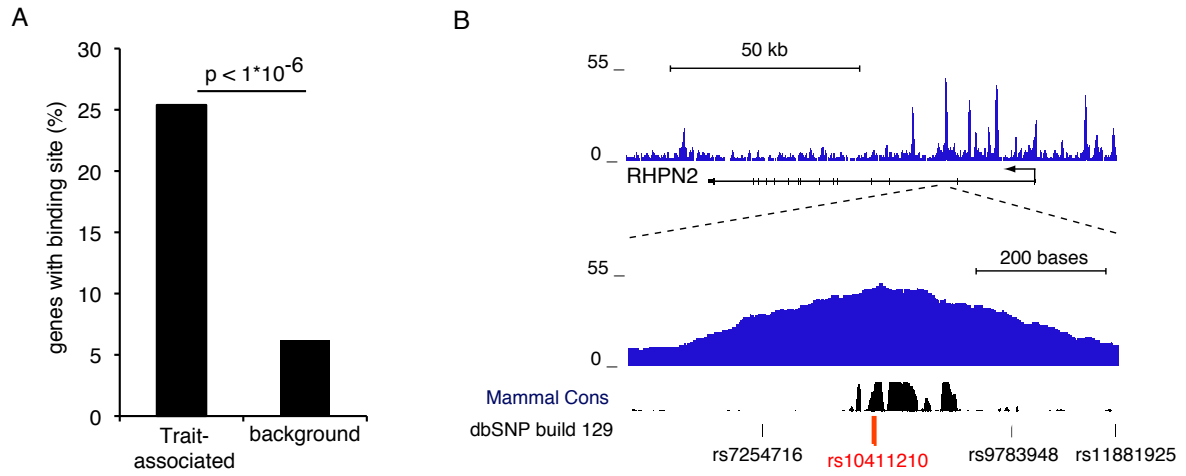

Figure 7. PPARγ binding associates with SNPs that are associated with different traits in a tissue-specific manner. *A) Genes implicated in at least one of five GWAS for blood lipid levels [149, 150, 153-155] are significantly more likely to contain a PPARγ binding site in human macrophages (THP-1 cells) compared to all genes in the genome (background). B) PPARγ binding in intestinal epithelial cells (HT29 cells) is shown across the RHPN2 locus, magnification shows strong PPARγ binding in vicinity of rs10411210. This SNP is strongly associated with colorectal cancer [151]. Plotted is the density of sequencing reads detected in the PPARg ChIP-seq library in the specific region.*

PPARγ and NRs in general provide a unique therapeutic opportunity. However, the benefits of such therapies crucially depend on a proper understanding of the principles that underlie NR-dependent gene regulation in a systemic manner. PPARγ itself provides a sobering example as the clinical use of RSG in diabetes has been recently discouraged because of its association with increased heart failure [156]. While only at the beginning, combining sequencing based approaches and clinical samples might provide a powerful tool to better understand the precise actions of the ligands on a systemic level and avoid adverse side-effects in future applications. The challenge is to recognize and identify how the many regulatory layers, ranging from primary sequence over chromatin structure and interacting TFs to ligands and PTMs of NRs are all integrated in a genome-wide and systemic level. The work presented here comprehensively assessed binding of PPARγ across three organs and between species and demonstrated how general principles for gene regulation can be inferred from these data that were previously not accessible. While still descriptive in their nature these data provide the foundation for mechanistic and functional studies on PPARγ in different contexts.

6. Acknowledgements

During the five years of my PhD numerous colleagues and friends have shared their ideas, excitement and time with me. All of you made this journey a rewarding one!

I am especially grateful to my supervisor Sven Pettersson for giving me the opportunity to pursue my PhD in his lab and for letting me explore and develop my own ideas and interests. This was truly a great experience- I am not sure if we ever fully agreed on something- but I sure hope you agree that our discussions were stimulating to you too.

I am thankful to my co-supervisor Ed Liu, who took me in as a temporary visitor, and allowed me to stay a little(!) longer. Your encouragement to take full advantage of the opportunities at hand (and there were quite a few!) and the constant challenge you provided ('..to not get lost like a kid in the candy store') have in many ways shaped the way I address problems.

I have benefited tremendously from the patience and support of Vega, Mikael and Guillaume who shared their knowledge (and thankfully also their scripts) with me- Your encouragement and willingness to discuss my questions (often at length) made this experience most rewarding for me, thank you!

I am most grateful to the lab members in Singapore, especially Kartiki and Jane, for providing such a relaxed environment.

The dear friends and colleagues without who work would have been rather dull- Gaye, Nima, Leena, Faranak and Francesca (and many others); I will miss your lunch company, the coffee breaks and the smaller breaks in between those.

I would also like to thank the swedish part of the lab- it has always been great to be back! Linda, Ing-Marie, Agne, thanks for welcoming me to the group- back then! Thank you, Kerstin and Gesan for remaining my reliable contacts once I moved to Singapore, it's pretty safe to say that things would have been messy without you!

It goes without saying (rather, almost without) that I am indebted to friends and family in ways these few lines can do no justice. The encouragement and support I get from you is more than I could have asked for – nothing I do would matter without you.

Tanya, thank you for your patience and everything else- it's only beginning…

7. References

1. Lander, E.S., et al., *Initial sequencing and analysis of the human genome*. Nature, 2001. 409(6822): p. 860-921.
2. Venter, J.C., et al., *The sequence of the human genome*. Science, 2001. 291(5507): p. 1304-51.
3. Waterston, R.H., et al., *Initial sequencing and comparative analysis of the mouse genome*. Nature, 2002. 420(6915): p. 520-62.
4. Gibbs, R.A., et al., *Genome sequence of the Brown Norway rat yields insights into mammalian evolution*. Nature, 2004. 428(6982): p. 493-521.
5. Mardis, E.R., *Next-generation DNA sequencing methods*. Annu Rev Genomics Hum Genet, 2008. 9: p. 387-402.
6. Margulies, M., et al., *Genome sequencing in microfabricated high-density picolitre reactors*. Nature, 2005. 437(7057): p. 376-80.
7. Clamp, M., et al., *Distinguishing protein-coding and noncoding genes in the human genome*. Proc Natl Acad Sci U S A, 2007. 104(49): p. 19428-33.
8. Siepel, A., et al., *Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes*. Genome Res, 2005. 15(8): p. 1034-50.
9. Birney, E., et al., *Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project*. Nature, 2007. 447(7146): p. 799-816.
10. Kasowski, M., et al., *Variation in transcription factor binding among humans*. Science, 2010. 328(5975): p. 232-5.
11. Dimas, A.S., et al., *Common regulatory variation impacts gene expression in a cell type-dependent manner*. Science, 2009. 325(5945): p. 1246-50.
12. McDaniell, R., et al., *Heritable individual-specific and allele-specific chromatin signatures in humans*. Science, 2010. 328(5975): p. 235-9.
13. Kornberg, R.D., *Chromatin structure: a repeating unit of histones and DNA*. Science, 1974. 184(139): p. 868-71.
14. Davey, C.A., et al., *Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution*. J Mol Biol, 2002. 319(5): p. 1097-113.
15. Jenuwein, T. and C.D. Allis, *Translating the histone code*. Science, 2001. 293(5532): p. 1074-80.
16. Foster, S.L., D.C. Hargreaves, and R. Medzhitov, *Gene-specific control of inflammation by TLR-induced chromatin modifications*. Nature, 2007. 447(7147): p. 972-8.
17. Wang, Z., et al., *Combinatorial patterns of histone acetylations and methylations in the human genome*. Nat Genet, 2008. 40(7): p. 897-903.
18. Barski, A., et al., *High-resolution profiling of histone methylations in the human genome*. Cell, 2007. 129(4): p. 823-37.
19. 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.
20. Guenther, M.G., et al., *A chromatin landmark and transcription initiation at most promoters in human cells*. Cell, 2007. 130(1): p. 77-88.
21. Meissner, A., et al., *Genome-scale DNA methylation maps of pluripotent and differentiated cells*. Nature, 2008. 454(7205): p. 766-70.
22. Lister, R., et al., *Human DNA methylomes at base resolution show widespread epigenomic differences*. Nature, 2009. 462(7271): p. 315-22.
23. Ernst, J., et al., *Mapping and analysis of chromatin state dynamics in nine human cell types*. Nature, 2011.
24. Hon, G., B. Ren, and W. Wang, *ChromaSig: a probabilistic approach to finding common chromatin signatures in the human genome*. PLoS Comput Biol, 2008. 4(10): p. e1000201.
25. Negre, N., et al., *A cis-regulatory map of the Drosophila genome*. Nature, 2011. 471(7339): p. 527-31.
26. Roy, S., et al., *Identification of functional elements and regulatory circuits by Drosophila modENCODE*. Science, 2010. 330(6012): p. 1787-97.
27. Filion, G.J., et al., *Systematic protein location mapping reveals five principal chromatin types in Drosophila cells*. Cell, 2010. 143(2): p. 212-24.

28. Gerstein, M.B., et al., *Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project*. Science, 2010. 330(6012): p. 1775-87.
29. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors*. Cell, 2006. 126(4): p. 663-76.
30. Hawkins, R.D., et al., *Distinct epigenomic landscapes of pluripotent and lineage-committed human cells*. Cell Stem Cell, 2010. 6(5): p. 479-91.
31. Lee, T.I. and R.A. Young, *Transcription of eukaryotic protein-coding genes*. Annu Rev Genet, 2000. 34: p. 77-137.
32. Farnham, P.J., *Insights from genomic profiling of transcription factors*. Nat Rev Genet, 2009. 10(9): p. 605-16.
33. Fuda, N.J., M.B. Ardehali, and J.T. Lis, *Defining mechanisms that regulate RNA polymerase II transcription in vivo*. Nature, 2009. 461(7261): p. 186-92.
34. Roeder, R.G., *The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly*. Trends Biochem Sci, 1991. 16(11): p. 402-8.
35. Hargreaves, D.C., T. Horng, and R. Medzhitov, *Control of inducible gene expression by signal-dependent transcriptional elongation*. Cell, 2009. 138(1): p. 129-45.
36. Rahl, P.B., et al., *c-Myc regulates transcriptional pause release*. Cell, 2010. 141(3): p. 432-45.
37. Ramirez-Carrozzi, V.R., et al., *A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling*. Cell, 2009. 138(1): p. 114-28.
38. Kim, T.H., et al., *Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome*. Cell, 2007. 128(6): p. 1231-45.
39. Vaquerizas, J.M., et al., *A census of human transcription factors: function, expression and evolution*. Nat Rev Genet, 2009. 10(4): p. 252-63.
40. Mangelsdorf, D.J., et al., *The nuclear receptor superfamily: the second decade*. Cell, 1995. 83(6): p. 835-9.
41. Bookout, A.L., et al., *Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network*. Cell, 2006. 126(4): p. 789-99.
42. Naar, A.M., et al., *The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors*. Cell, 1991. 65(7): p. 1267-79.
43. Mangelsdorf, D.J. and R.M. Evans, *The RXR heterodimers and orphan receptors*. Cell, 1995. 83(6): p. 841-50.
44. Hollenberg, S.M., et al., *Primary structure and expression of a functional human glucocorticoid receptor cDNA*. Nature, 1985. 318(6047): p. 635-41.
45. Lubahn, D.B., et al., *Cloning of human androgen receptor complementary DNA and localization to the X chromosome*. Science, 1988. 240(4850): p. 327-30.
46. Heery, D.M., et al., *A signature motif in transcriptional co-activators mediates binding to nuclear receptors*. Nature, 1997. 387(6634): p. 733-6.
47. McInerney, E.M., et al., *Determinants of coactivator LXXLL motif specificity in nuclear receptor transcriptional activation*. Genes Dev, 1998. 12(21): p. 3357-68.
48. Wilson, B.J., et al., *An acetylation switch modulates the transcriptional activity of estrogen-related receptor alpha*. Mol Endocrinol, 2010. 24(7): p. 1349-58.
49. Pascual, G., et al., *A SUMOylation-dependent pathway mediates transrepression of inflammatory response genes by PPAR-gamma*. Nature, 2005. 437(7059): p. 759-63.
50. Faus, H. and B. Haendler, *Post-translational modifications of steroid receptors*. Biomed Pharmacother, 2006. 60(9): p. 520-8.
51. Barger, P.M., et al., *p38 mitogen-activated protein kinase activates peroxisome proliferator-activated receptor alpha: a potential role in the cardiac metabolic stress response*. J Biol Chem, 2001. 276(48): p. 44495-501.
52. Hu, E., et al., *Inhibition of adipogenesis through MAP kinase-mediated phosphorylation of PPARgamma*. Science, 1996. 274(5295): p. 2100-3.
53. Nolan, J.J., et al., *Improvement in glucose tolerance and insulin resistance in obese subjects treated with troglitazone*. N Engl J Med, 1994. 331(18): p. 1188-93.
54. Baniahmad, A., et al., *Interaction of human thyroid hormone receptor beta with transcription factor TFIIB may mediate target gene derepression and activation by thyroid hormone*. Proc Natl Acad Sci U S A, 1993. 90(19): p. 8832-6.

55. Schulman, I.G., et al., *Interactions between the retinoid X receptor and a conserved region of the TATA-binding protein mediate hormone-dependent transactivation*. Proc Natl Acad Sci U S A, 1995. 92(18): p. 8288-92.
56. Vo, N. and R.H. Goodman, *CREB-binding protein and p300 in transcriptional regulation*. J Biol Chem, 2001. 276(17): p. 13505-8.
57. Perissi, V. and M.G. Rosenfeld, *Controlling nuclear receptors: the circular logic of cofactor cycles*. Nat Rev Mol Cell Biol, 2005. 6(7): p. 542-54.
58. Metivier, R., et al., *Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter*. Cell, 2003. 115(6): p. 751-63.
59. Hu, F.B., et al., *Diet, lifestyle, and the risk of type 2 diabetes mellitus in women*. N Engl J Med, 2001. 345(11): p. 790-7.
60. Hotamisligil, G.S., *Inflammation and metabolic disorders*. Nature, 2006. 444(7121): p. 860-7.
61. Ley, R.E., et al., *Obesity alters gut microbial ecology*. Proc Natl Acad Sci U S A, 2005. 102(31): p. 11070-5.
62. Ley, R.E., et al., *Microbial ecology: human gut microbes associated with obesity*. Nature, 2006. 444(7122): p. 1022-3.
63. Wen, L., et al., *Innate immunity and intestinal microbiota in the development of Type 1 diabetes*. Nature, 2008. 455(7216): p. 1109-13.
64. Ijpenberg, A., et al., *Polarity and specific sequence requirements of peroxisome proliferator-activated receptor (PPAR)/retinoid X receptor heterodimer binding to DNA. A functional analysis of the malic enzyme gene PPAR response element*. J Biol Chem, 1997. 272(32): p. 20108-17.
65. Kliewer, S.A., et al., *Fatty acids and eicosanoids regulate gene expression through direct interactions with peroxisome proliferator-activated receptors alpha and gamma*. Proc Natl Acad Sci U S A, 1997. 94(9): p. 4318-23.
66. Kliewer, S.A., et al., *A prostaglandin J2 metabolite binds peroxisome proliferator-activated receptor gamma and promotes adipocyte differentiation*. Cell, 1995. 83(5): p. 813-9.
67. Nagy, L., et al., *Oxidized LDL regulates macrophage gene expression through ligand activation of PPARgamma*. Cell, 1998. 93(2): p. 229-40.
68. Dreyer, C., et al., *Control of the peroxisomal beta-oxidation pathway by a novel family of nuclear hormone receptors*. Cell, 1992. 68(5): p. 879-87.
69. Sarraf, P., et al., *Differentiation and reversal of malignant changes in colon cancer through PPARgamma*. Nat Med, 1998. 4(9): p. 1046-52.
70. Sarraf, P., et al., *Loss-of-function mutations in PPAR gamma associated with human colon cancer*. Mol Cell, 1999. 3(6): p. 799-804.
71. Bush, C.R., et al., *Functional genomic analysis reveals cross-talk between peroxisome proliferator-activated receptor gamma and calcium signaling in human colorectal cancer cells*. J Biol Chem, 2007. 282(32): p. 23387-401.
72. Lefebvre, A.M., et al., *Activation of the peroxisome proliferator-activated receptor gamma promotes the development of colon tumors in C57BL/6J-APCMin/+ mice*. Nat Med, 1998. 4(9): p. 1053-7.
73. Saez, E., et al., *Activators of the nuclear receptor PPARgamma enhance colon polyp formation*. Nat Med, 1998. 4(9): p. 1058-61.
74. Girnun, G.D., et al., *Synergy between PPARgamma ligands and platinum-based drugs in cancer*. Cancer Cell, 2007. 11(5): p. 395-406.
75. Adachi, M., et al., *Peroxisome proliferator activated receptor gamma in colonic epithelial cells protects against experimental inflammatory bowel disease*. Gut, 2006. 55(8): p. 1104-13.
76. Dubuquoy, L., et al., *Impaired expression of peroxisome proliferator-activated receptor gamma in ulcerative colitis*. Gastroenterology, 2003. 124(5): p. 1265-76.
77. Kelly, D., et al., *Commensal anaerobic gut bacteria attenuate inflammation by regulating nuclear-cytoplasmic shuttling of PPAR-gamma and RelA*. Nat Immunol, 2004. 5(1): p. 104-12.
78. Are, A., et al., *Enterococcus faecalis from newborn babies regulate endogenous PPARgamma activity and IL-10 levels in colonic epithelial cells*. Proc Natl Acad Sci U S A, 2008. 105(6): p. 1943-8.

79. Harmon, G.S., et al., *Pharmacological correction of a defect in PPAR-gamma signaling ameliorates disease severity in Cftr-deficient mice*. Nat Med, 2010. 16(3): p. 313-8.
80. Virtue, S. and A. Vidal-Puig, *It's not how fat you are, it's what you do with it that counts*. PLoS Biol, 2008. 6(9): p. e237.
81. Berg, A.H., T.P. Combs, and P.E. Scherer, *ACRP30/adiponectin: an adipokine regulating glucose and lipid metabolism*. Trends Endocrinol Metab, 2002. 13(2): p. 84-9.
82. Pajvani, U.B., et al., *Complex distribution, not absolute amount of adiponectin, correlates with thiazolidinedione-mediated improvement in insulin sensitivity*. J Biol Chem, 2004. 279(13): p. 12152-62.
83. Yu, J.G., et al., *The effect of thiazolidinediones on plasma adiponectin levels in normal, obese, and type 2 diabetic subjects*. Diabetes, 2002. 51(10): p. 2968-74.
84. Arner, P., *Insulin resistance in type 2 diabetes -- role of the adipokines*. Curr Mol Med, 2005. 5(3): p. 333-9.
85. Park, K.W., D.S. Halperin, and P. Tontonoz, *Before they were fat: adipocyte progenitors*. Cell Metab, 2008. 8(6): p. 454-7.
86. Cannon, B. and J. Nedergaard, *Metabolic consequences of the presence or absence of the thermogenic capacity of brown adipose tissue in mice (and probably in humans)*. Int J Obes (Lond), 2010. 34 Suppl 1: p. S7-16.
87. Puigserver, P., et al., *A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis*. Cell, 1998. 92(6): p. 829-39.
88. Endemann, G., et al., *CD36 is a receptor for oxidized low density lipoprotein*. J Biol Chem, 1993. 268(16): p. 11811-6.
89. Chawla, A., et al., *A PPAR gamma-LXR-ABCA1 pathway in macrophages is involved in cholesterol efflux and atherogenesis*. Mol Cell, 2001. 7(1): p. 161-71.
90. Welch, J.S., et al., *PPARgamma and PPARdelta negatively regulate specific subsets of lipopolysaccharide and IFN-gamma target genes in macrophages*. Proc Natl Acad Sci U S A, 2003. 100(11): p. 6712-7.
91. Ogawa, S., et al., *Molecular determinants of crosstalk between nuclear receptors and toll-like receptors*. Cell, 2005. 122(5): p. 707-21.
92. Ogawa, S., et al., *A nuclear receptor corepressor transcriptional checkpoint controlling activator protein 1-dependent gene networks required for macrophage activation*. Proc Natl Acad Sci U S A, 2004. 101(40): p. 14461-6.
93. Klotz, L., et al., *The nuclear receptor PPAR gamma selectively inhibits Th17 differentiation in a T cell-intrinsic fashion and suppresses CNS autoimmunity*. J Exp Med, 2009. 206(10): p. 2079-89.
94. Lee, J.H., et al., *Differential SUMOylation of LXRalpha and LXRbeta mediates transrepression of STAT1 inflammatory signaling in IFN-gamma-stimulated brain astrocytes*. Mol Cell, 2009. 35(6): p. 806-17.
95. Saijo, K., et al., *A Nurr1/CoREST pathway in microglia and astrocytes protects dopaminergic neurons from inflammation-induced death*. Cell, 2009. 137(1): p. 47-59.
96. Gordon, S., *Alternative activation of macrophages*. Nat Rev Immunol, 2003. 3(1): p. 23-35.
97. Hevener, A.L., et al., *Macrophage PPAR gamma is required for normal skeletal muscle and hepatic insulin sensitivity and full antidiabetic effects of thiazolidinediones*. J Clin Invest, 2007. 117(6): p. 1658-69.
98. Odegaard, J.I., et al., *Macrophage-specific PPARgamma controls alternative activation and improves insulin resistance*. Nature, 2007. 447(7148): p. 1116-20.
99. Bouhlel, M.A., et al., *PPARgamma activation primes human monocytes into alternative M2 macrophages with anti-inflammatory properties*. Cell Metab, 2007. 6(2): p. 137-43.
100. Vats, D., et al., *Oxidative metabolism and PGC-1beta attenuate macrophage-mediated inflammation*. Cell Metab, 2006. 4(1): p. 13-24.
101. Szatmari, I., et al., *PPARgamma regulates the function of human dendritic cells primarily by altering lipid metabolism*. Blood, 2007. 110(9): p. 3271-80.
102. Szatmari, I., et al., *PPARgamma controls CD1d expression by turning on retinoic acid synthesis in developing human dendritic cells*. J Exp Med, 2006. 203(10): p. 2351-62.
103. Park, P.J., *ChIP-seq: advantages and challenges of a maturing technology*. Nat Rev Genet, 2009. 10(10): p. 669-80.

104. Solomon, M.J., P.L. Larsen, and A. Varshavsky, *Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene*. Cell, 1988. 53(6): p. 937-47.
105. Xu, H., et al., *A signal-noise model for significance analysis of ChIP-seq with negative control*. Bioinformatics, 2010. 26(9): p. 1199-204.
106. Jothi, R., et al., *Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data*. Nucleic Acids Res, 2008. 36(16): p. 5221-31.
107. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS)*. Genome Biol, 2008. 9(9): p. R137.
108. Lupien, M., et al., *FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription*. Cell, 2008. 132(6): p. 958-70.
109. Lefterova, M.I., et al., *PPARgamma and C/EBP factors orchestrate adipocyte biology via adjacent binding on a genome-wide scale*. Genes Dev, 2008. 22(21): p. 2941-52.
110. 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.
111. Johnson, D.S., et al., *Genome-wide mapping of in vivo protein-DNA interactions*. Science, 2007. 316(5830): p. 1497-502.
112. Ng, P., et al., *Multiplex sequencing of paired-end ditags (MS-PET): a strategy for the ultra-high-throughput analysis of transcriptomes and genomes*. Nucleic Acids Res, 2006. 34(12): p. e84.
113. Schena, M., et al., *Quantitative monitoring of gene expression patterns with a complementary DNA microarray*. Science, 1995. 270(5235): p. 467-70.
114. Nadon, R. and J. Shoemaker, *Statistical issues with microarrays: processing and analysis*. Trends Genet, 2002. 18(5): p. 265-71.
115. Miller, L.D., et al., *An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival*. Proc Natl Acad Sci U S A, 2005. 102(38): p. 13550-5.
116. Wang, Z., M. Gerstein, and M. Snyder, *RNA-Seq: a revolutionary tool for transcriptomics*. Nat Rev Genet, 2009. 10(1): p. 57-63.
117. Guttman, M., et al., *Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs*. Nat Biotechnol, 2010. 28(5): p. 503-10.
118. Trapnell, C., et al., *Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation*. Nat Biotechnol, 2010. 28(5): p. 511-5.
119. Abreu, M.T., *Toll-like receptor signalling in the intestinal epithelium: how bacterial recognition shapes intestinal function*. Nat Rev Immunol, 2010. 10(2): p. 131-44.
120. Slack, E., et al., *Innate and adaptive immunity cooperate flexibly to maintain host-microbiota mutualism*. Science, 2009. 325(5940): p. 617-20.
121. Thomas, P.D., et al., *PANTHER: a library of protein families and subfamilies indexed by function*. Genome Res, 2003. 13(9): p. 2129-41.
122. Thomas, P.D., et al., *Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools*. Nucleic Acids Res, 2006. 34(Web Server issue): p. W645-50.
123. Lin, C.Y., et al., *Whole-genome cartography of estrogen receptor alpha binding sites*. PLoS Genet, 2007. 3(6): p. e87.
124. Kunarso, G., et al., *Transposable elements have rewired the core regulatory network of human embryonic stem cells*. Nat Genet, 2010. 42(7): p. 631-4.
125. Mikkelsen, T.S., et al., *Comparative epigenomic analysis of murine and human adipogenesis*. Cell, 2010. 143(1): p. 156-69.
126. Odom, D.T., et al., *Tissue-specific transcriptional regulation has diverged significantly between human and mouse*. Nat Genet, 2007. 39(6): p. 730-2.
127. Schmidt, D., et al., *Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding*. Science, 2010. 328(5981): p. 1036-40.

128. 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.
129. Boj, S.F., et al., *Functional targets of the monogenic diabetes transcription factors HNF-1alpha and HNF-4alpha are highly conserved between mice and humans*. Diabetes, 2009. 58(5): p. 1245-53.
130. Weirauch, M.T. and T.R. Hughes, *Conserved expression without conserved regulatory sequence: the more things change, the more they stay the same*. Trends Genet, 2010. 26(2): p. 66-74.
131. Chen, X., et al., *Integration of external signaling pathways with the core transcriptional network in embryonic stem cells*. Cell, 2008. 133(6): p. 1106-17.
132. 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.
133. Ghisletti, S., et al., *Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages*. Immunity, 2010. 32(3): p. 317-28.
134. Wilson, M.D., et al., *Species-specific transcription in mice carrying human chromosome 21*. Science, 2008. 322(5900): p. 434-8.
135. Lefebvre, M., et al., *Peroxisome proliferator-activated receptor gamma is induced during differentiation of colon epithelium cells*. J Endocrinol, 1999. 162(3): p. 331-40.
136. Mokry, M., et al., *Efficient double fragmentation ChIP-seq provides nucleotide resolution protein-DNA binding profiles*. PLoS One, 2010. 5(11): p. e15092.
137. Watanabe, M., et al., *Bile acids lower triglyceride levels via a pathway involving FXR, SHP, and SREBP-1c*. J Clin Invest, 2004. 113(10): p. 1408-18.
138. Necela, B.M., W. Su, and E.A. Thompson, *Toll-like receptor 4 mediates cross-talk between peroxisome proliferator-activated receptor gamma and nuclear factor-kappaB in macrophages*. Immunology, 2008. 125(3): p. 344-58.
139. 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.
140. Drori, S., et al., *Hic-5 regulates an epithelial program mediated by PPARgamma*. Genes Dev, 2005. 19(3): p. 362-75.
141. 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.
142. Fullwood, M.J., et al., *An oestrogen-receptor-alpha-bound human chromatin interactome*. Nature, 2009. 462(7269): p. 58-64.
143. Schoenfelder, S., et al., *Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells*. Nat Genet, 2009. 42(1): p. 53-61.
144. Goren, A., et al., *Chromatin profiling by directly sequencing small quantities of immunoprecipitated DNA*. Nat Methods, 2009. 7(1): p. 47-9.
145. Gaulton, K.J., et al., *A map of open chromatin in human pancreatic islets*. Nat Genet, 2010. 42(3): p. 255-9.
146. Heikkinen, S., et al., *The Pro12Ala PPARgamma2 variant determines metabolism at the gene-environment interface*. Cell Metab, 2009. 9(1): p. 88-98.
147. Antoniou, M., et al., *The human beta-globin gene contains multiple regulatory regions: identification of one promoter and two downstream enhancers*. EMBO J, 1988. 7(2): p. 377-84.
148. Dekker, J., et al., *Capturing chromosome conformation*. Science, 2002. 295(5558): p. 1306-11.
149. Kathiresan, S., et al., *Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans*. Nat Genet, 2008. 40(2): p. 189-97.
150. Kathiresan, S., et al., *Common variants at 30 loci contribute to polygenic dyslipidemia*. Nat Genet, 2009. 41(1): p. 56-65.
151. Houlston, R.S., et al., *Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer*. Nat Genet, 2008. 40(12): p. 1426-35.

152. von Holst, S., et al., *Association studies on 11 published colorectal cancer risk loci*. Br J Cancer, 2010. 103(4): p. 575-80.
153. Aulchenko, Y.S., et al., *Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts*. Nat Genet, 2009. 41(1): p. 47-55.
154. Willer, C.J., et al., *Newly identified loci that influence lipid concentrations and risk of coronary artery disease*. Nat Genet, 2008. 40(2): p. 161-9.
155. Sabatti, C., et al., *Genome-wide association analysis of metabolic traits in a birth cohort from a founder population*. Nat Genet, 2009. 41(1): p. 35-46.
156. Nissen, S.E. and K. Wolski, *Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes*. N Engl J Med, 2007. 356(24): p. 2457-71.