Review

A Genomic Perspective on Vitamin D Signaling

CARSTEN CARLBERG[1,2] and SABINE SEUTER[1]

[1]*Life Sciences Research Unit, University of Luxembourg, Luxembourg;*
[2]*Department of Biochemistry, University of Kuopio, Kuopio, Finland*

Abstract. *The vitamin D receptor (VDR) is a ligand-inducible transcription factor whose target genes play key roles in cellular metabolism, bone formation, cellular growth, differentiation and in controlling inflammation. Many of these VDR target genes are also involved in dysregulated pathways leading to common human diseases, such as cancer. The activation of VDR by natural and synthetic ligands may improve such pathological conditions. On a genomic level, these pathways converge on regulatory modules, some of which contain VDR-binding sites, so-called vitamin D response elements (VDREs). Transcriptome analysis, chromatin immunoprecipitation scans and in silico screening approaches have already identified many genomic targets of the VDR. Important cancer regulatory modules with VDREs should have a major impact on understanding the role and potential therapeutic value of VDR in cancer.*

Cellular proliferation and differentiation are controlled by transcriptional regulation of a large subset of the human genome. The transcriptomes of normal and tumor cells as revealed by microarray analysis show significant differences (1), suggesting that in cancer the precise transcriptional control is lost due to overactive oncogenes and loss of function of tumor suppressor genes, many of which code for transcription factors. For a molecular insight into cancer, the transcriptional regulation of probably thousands of genes has to be uncovered in detail by integrating expression array data with regulatory site location data (2). Although the understanding of the regulation of a few of key genes, such as the cyclin-dependent kinase (CDK) inhibitor *p21*$^{WAF1/CIP1}$ (3), is already quite advanced, for the majority of the cancer-associated genes such detailed analyses have not been performed. Even "big biology" projects, such as ENCODE (4), have thus far focused only on 1% of the human genome sequence, while other genome-wide scans, *e.g.* for histone modifications (5-7) or transcription factor binding (8, 9), had to concentrate on only a subset of modifications and factors under limited experimental conditions.

Each individual gene is under the control of a large set of transcription factors that can bind upstream and downstream of its transcription start site (TSS) (10). These sites typically arrange into collections of neighboring sites, so-called modules or enhancers. Modules of transcription factors that act on focused genomic regions have been shown to be far more effective than individual factors on isolated locations and can act from distances up to hundreds of thousands of base pairs. In an ideal case, such transcription factor modules can be identified by parallel and comparative analysis of their binding sites. Here bioinformatics approaches could be of great help, provided they can predict the actions of the transcription factors precisely enough (11).

The Nuclear Receptor Superfamily

Nuclear receptors represent an important transcription factor family (12). The 48 human members of this superfamily belong to the best-characterized genes of the approximately 3,000 mammalian genes that are involved in transcriptional regulation (13). Often triggered by small lipophilic ligands, nuclear receptors modulate genes that affect processes as diverse as reproduction, development, inflammation and general metabolism. This property has attracted a lot of

interest in the family members as possible therapeutical targets in the context of cancer. Nuclear receptors mostly regulate transcription by binding directly to specific response elements, either as monomers, homo- or heterodimers, but they can also indirectly influence transcription through other DNA-binding transcription factors.

Nuclear receptors have a modular structure, onto which certain functions can be ascribed. The amino-terminus is of variable length and sequence in the different family members. It contains a transactivation domain, termed AF-1, which is recognized by co-activator proteins and/or other transcription factors, often in a ligand-independent fashion. The central DNA-binding domain has two zinc-finger motifs that are common to the entire family. The carboxy-terminal ligand-binding domain, whose overall architecture is well conserved between the various family members, nonetheless diverges sufficiently to guarantee selective ligand recognition as well as to accommodate the broad spectrum of nuclear receptor ligand structures. The ligand-binding domain consists of 250-300 amino acids in 11-13 α-helices (14). Ligand binding causes a conformational change within the ligand-binding domain, whereby, at least in the case of endocrine nuclear receptors, helix 12, the most carboxy-terminal α-helix, closes the ligand-binding pocket via a "mouse-trap like" intramolecular folding event (15). The ligand-binding domain is also involved in a variety of interactions with nuclear proteins, such as other members of the nuclear receptor superfamily and co-regulator proteins.

The Vitamin D Receptor

The vitamin D receptor (VDR) is an endocrine member of the nuclear receptor superfamily (16), because it is the only nuclear protein that binds the nuclear hormone $1\alpha,25$-dihydroxyvitamin D_3 ($1\alpha,25(OH)_2D_3$) with high affinity (K_d=0.1 nM) (17, 18). Other members of the endocrine receptor subgroup of the nuclear receptor superfamily are the receptors for the nuclear hormones retinoic acid, thyroid hormone, estradiol, progesterone, testosterone, cortisol and aldosterol (19).

In mammals, the highest VDR expression is found in metabolic tissues, such as intestine and kidney, as well as in skin and the thyroid gland, but moderate expression is found in nearly all tissues (20). Moreover, the receptor is also expressed in many malignant tissues (21). Mice lacking a functional *VDR* gene develop alopecia (likewise found in many patients with mutations in the *VDR*) (22); these mice also exhibit a defect in epidermal differentiation. Moreover, *VDR*-null mice also show an increased susceptibility to tumor formation (23). VDR ligands are important for the regulation of normal cellular growth and differentiation and have a potential in the prevention and the treatment of hyper-proliferative malignancies, such as cancer and psoriasis (24,

25). However, frequently transformed cells get resistant to these actions of $1\alpha,25(OH)_2D_3$, making the clinical application of VDR ligands difficult. Epidemiological evidence supports the importance of adequate vitamin D_3 nutrition for several types of cancer (26).

VDR has been shown to form homodimers (27, 28) and heterodimers with the thyroid hormone receptor (29, 30) and the retinoic acid receptor (31), but by far the strongest binding partner of VDR is one of the three retinoid X receptors (RXRs) α, β and γ (32). Primary $1\alpha,25(OH)_2D_3$ responding genes contain a specific VDR-binding site in their regulatory regions, which is referred to as a vitamin D response element (VDRE). An essential prerequisite for the direct modulation of transcription by VDR ligands is the location of at least one activated VDR protein close to the TSS of the respective primary target gene. This is achieved, in most cases, through the specific binding of the DNA-binding domain of the VDR to the major groove of a hexameric DNA sequence, referred to as core binding motif, with the consensus sequence RGKTSA (R=A or G, K=G or T, S=C or G) (18). The trimeric complex of VDR, RXR and a VDRE can be considered as a molecular switch for primary $1\alpha,25(OH)_2D_3$ responding genes. VDR-RXR heterodimers bind to VDREs formed by a direct repeat (DR) of two hexameric core binding motifs with 3 intervening nucleotides (DR3-type) (27), but also to DR4-type REs along with other members of the nuclear receptor superfamily (33). It should also be noted that effective VDR binding has also been observed on everted repeat (ER)-type REs with 6 to 9 spacing nucleotides (ER6, ER7, ER8, ER9) (30, 34).

Chromatin and Co-factors

The major protein constituents of chromatin are the four different histones that form a nucleosome, around which DNA is wound. Covalent modifications of the lysines at the amino-terminal tails of these histone proteins neutralize their positive charge and thus their attraction for the negatively charged DNA backbone is diminished (35). As a consequence, the association between the histone and the DNA becomes less stable. This influences the degree of chromatin packaging and regulates the access of transcription factors to their potential binding sites. When nuclear receptors are bound to response elements in the regulatory regions of their target genes, they recruit positive and negative co-regulatory proteins, referred to as co-activators (36) and co-repressors (37), respectively. In a simplified view of nuclear receptor signaling, in the absence of ligand, the nuclear receptor interacts with co-repressor proteins, such as NCoR1, SMRT, hairless and Alien, which in turn associate with histone deacetylases (HDACs) leading to a locally increased chromatin packaging (38, 39). The binding of ligand induces the dissociation of the co-repressor and the

association of a co-activator of the p160-family, such as SRC-1, TIF2 or RAC3 (40). Some co-activators have histone acetylase activity or are complexed with proteins harboring such activity and this results in the net effect of local chromatin relaxation (41). In a subsequent step, ligand-activated nuclear receptors change rapidly from interacting with the co-activators of the p160-family to those of mediator complexes, such as Med1 (42). The mediator complexes, which consist of approximately 15-20 proteins, build a bridge to the basal transcriptional machinery (43). In this way ligand-activated nuclear receptors execute two tasks, the modification of chromatin and the regulation of transcription.

The anti-proliferative effects of HDAC inhibitors and $1\alpha,25(OH)_2D_3$ converge *via* the interaction of un-liganded VDR with co-repressors recruiting multi-protein complexes containing HDACs and *via* the induction of CDK inhibitor genes (44, 45). The HDAC inhibitor Trichostatin A induces the genes *p18* and *p19*, whereas the gene $p21^{WAF1/CIP1}$ is stimulated by $1\alpha,25(OH)_2D_3$ (46). The co-repressor NCoR1 and some HDAC family members complex un-liganded VDR and repress the basal level of CDK inhibitor genes, but their roles in regulating expression of these genes by Trichostatin A and $1\alpha,25(OH)_2D_3$ are contrary. HDAC3 and HDAC7 attenuate $1\alpha,25(OH)_2D_3$-dependent induction of the $p21^{WAF1/CIP1}$ gene, for which NCoR1 is essential. In contrast, Trichostatin A-mediated induction of the *p18* gene is dependent on HDAC3 and HDAC4, but is opposed by NCoR1 and un-liganded VDR (46). This suggests that the attenuation of the response to TSA or $1\alpha,25(OH)_2D_3$ by NCoR1 or HDACs, respectively, can be overcome by their combined application achieving maximal induction of anti-proliferative target genes.

Cell- and time-specific patterns of relative protein expression levels of some co-regulators can distinctly modulate nuclear receptor transcriptional activity. This aspect may have some diagnostic and therapeutic value in different types of cancer (47). Concerning skin cancer it was postulated that the stoichiometric ratio between co-activators of the p160 family and Med1 might regulate a $1\alpha,25(OH)_2D_3$-dependent balance between proliferation and differentiation of keratinocytes (48). However, the switch between gene repression and activation is more complex than a simple alternative recruitment of two different regulatory complexes (46). Most co-regulators are co-expressed in the same cell type at relatively similar levels, which raises the possibility of their concomitant recruitment to a specific promoter. This has been resolved by the mutually exclusive binding of co-activators and co-repressors to ligand-bound and -unbound nuclear receptors, respectively. Therefore, repression and activation is more likely to be achieved by a series of sequential multiple enzymatic reactions that are promoter and cell-type specific. Transcriptional regulation is

a highly dynamic event of rapid association and dissociation of proteins and their modifications, including proteolytic degradation and *de novo* synthesis. A pattern of recruitment and release of cohorts of co-regulatory complexes was demonstrated on a single region of the *trefoil factor-1* promoter in breast cancer cells (49). This study revealed detailed and coordinated patterns of co-regulator recruitment and preferential selectivity for factors that have similar enzymatic activities. Interestingly, similar cyclical behavior was also observed for the VDR (50).

Primary VDR Target Genes

The essential role of $1\alpha,25(OH)_2D_3$ for mineral homeostasis and skeletal integrity is due to its regulation of Ca^{2+}-transporting proteins and bone-forming enzymes, such as calbindin D_{9K}, osteocalcin and osteopontin (51). $1\alpha,25(OH)_2D_3$ inhibits cellular proliferation by blocking cells at the G_0/G_1 to S transition. This block is presumed to occur because of an up-regulation of the CDK inhibitors p21 (52) and p27. The $p21^{WAF1/CIP1}$ gene promoter carries three VDREs positioned at -2.3 kbp, -4.5 kbp and -6.9 kbp from the TSS with p53 binding co-localized in two of these regions (3). In contrast, p27 is a short-lived protein and its activity is regulated almost exclusively post-translationally by the ubiquitin-proteosome protein degradation system. Decreased p27 protein levels are common in many tumor types (53).

Another interesting $1\alpha,25(OH)_2D_3$-target gene is *cyclin C*. The cyclin C-CDK8 complex was found to be associated with the RNA polymerase II basal transcriptional machinery (54) and is considered as a functional part of those mediator protein complexes that are involved in gene repression (55). The fact that the *cyclin C* gene, being located in chromosome 6q21, is deleted in a subset of acute lymphoblastic leukemias, suggests its involvement in tumorigenesis (56).

The steady state mRNA expression levels of some VDR target genes, such as that of *24-hydroxylase* (*CYP24*), are very low in the absence of ligand, but are induced by up to 1,000-fold through stimulation with $1\alpha,25(OH)_2D_3$ (57). Most other known primary $1\alpha,25(OH)_2D_3$ target genes, such as *cyclin C* and $p21^{WAF1/CIP1}$, often show an inducibility of 2-fold or less after short-term treatment with $1\alpha,25(OH)_2D_3$ (58, 59). However, both genes have 10,000- to 100,000-fold higher basal expression levels compared to that of the *CYP24* gene. Therefore, when the relative levels are taken into account, 2- to 20-fold more *cyclin C* and $p21^{WAF1/CIP1}$ than *CYP24* mRNA molecules are produced after induction with $1\alpha,25(OH)_2D_3$.

Another interesting primary $1\alpha,25(OH)_2D_3$ target is the *peroxisome proliferator activated recepto*r (*PPAR*) δ gene, which carries a potent DR3-type VDRE in close proximity to its TSS (60). PPARδ and VDR proteins are widely

expressed and an apparent overlap in the physiological action of the two nuclear receptors is their involvement in the regulation of cellular growth, particularly in neoplasms. High *PPARδ* expression in tumors seems to be positive for the prognosis of the respective cancer (61). Thus, the up-regulation of *PPARδ* expression can be considered as a part of the mechanism of the anti-proliferative action of $1\alpha,25(OH)_2D_3$ and its synthetic analogs.

Expression profiling using microarray technology indicates that comparable numbers of genes are down-regulated by $1\alpha,25(OH)_2D_3$ as are up-regulated by the hormone (62). In general, the mechanisms of the down-regulation of genes by $1\alpha,25(OH)_2D_3$ are much less understood, but they also seem to require the binding of an agonist to the VDR. The down-regulation of the $25(OH)D_3$ *1a-hydroxylase* (*CYP27B1*) gene by $1\alpha,25(OH)_2D_3$ has been proposed to involve a negative VDRE located at position -0.5 kB, where VDR-RXR heterodimers do not bind directly, but *via* the transcription factor VDR interacting repressor (63). In addition, two positive VDREs are located -2.6 and -3.2 kbp upstream from the TSS and modulate the cell-specific activity of the negative VDRE (64).

In addition, secondary $1\alpha,25(OH)_2D_3$-responding genes contribute to the physiological effects of $1\alpha,25(OH)_2D_3$, but their induction is delayed by a few hours or even days and is probably mediated by primary $1\alpha,25(OH)_2D_3$ responding gene products, such as transcription factors or co-regulator proteins (59).

VDR Target Gene Analysis

There are a number of modern methods for the identification and characterization of VDR target genes. Often the first step is the treatment of the cellular model (either an established cell line or primary cells) or *in vivo* model (mostly rodents) with $1\alpha,25(OH)_2D_3$ or its analogs. When the focus is on the identification of primary VDR target genes, the stimulation times are short (2-6 hours), but when the overall physiological effects are the center of the study, longer treatment times are used (24 to 72 hours). For a limited number of putative VDR target genes, quantitative PCR can be applied, but for a whole genome perspective on VDR signaling, microarrays have to be used. A few years ago cDNA arrays with an incomplete number of genes were mostly used and rather short lists of VDR target genes from colon (59), prostate (47, 65-67), breast (58), osteoblasts (68, 69) and squamous cell carcinoma (70) were published. Interestingly, the number of overlapping VDR target genes in these lists was low. Since the setups of these microarray analyses were different in treatment times and probe sets, this suggests that most VDR target genes respond to $1\alpha,25(OH)_2D_3$ in a very tissue-specific fashion and may have only a rather transient response. However, on the basis

of these results, the total number of convincing primary $1\alpha,25(OH)_2D_3$ target genes is in the order of 250. For a more detailed meta-analysis of VDR target genes, standardized microarray procedures performed on whole genome chips from Affymetrix, Illumina or other commercial suppliers are essential. Results from such approaches will be published soon.

For a detailed analysis of the regulatory regions of primary VDR target genes, the method of chromatin immunoprecipitation (ChIP) has become the golden standard over the last couple of years. For the genes *CYP24* (57), *CYP27B1* (64), *cyclin C* (71) and $p21^{WAF1/CIP1}$ (3), 7.1 to 8.4 kbp of their promoter regions were investigated by using in each case a set of 20-25 overlapping genomic regions. The spatio-temporal, $1\alpha,25(OH)_2D_3$-dependent chromatin changes in the four gene promoters were studied by ChIP assays with antibodies against acetylated histone 4, VDR, RXR and RNA polymerase II. Promising promoter regions were then screened *in silico* for putative VDREs, whose functionality was analyzed sequentially with gel shift, reporter gene and re-ChIP assays. This approach identified four VDREs for both the CYP24 and *cyclin C* genes, three in the $p21^{WAF1/CIP1}$ promoter and two in the *CYP27B1* gene. However, most of them are simultaneously under the control of other transcription factors, such as p53 in case of the $p21^{WAF1/CIP1}$ gene (3), and therefore possess significant basal levels of transcription. Consequently, the fold induction afforded by $1\alpha,25(OH)_2D_3$ stimulation is much less than for the *CYP24* gene, despite the large numbers of new mRNA molecules, which may be produced.

An alternative approach to the identification of primary $1\alpha,25(OH)_2D_3$ target genes was performed with the six members of the *insulin-like growth factor-binding protein* (*IGFBP*) gene family. Initially, an *in silico* screen was performed, which was then followed by the analysis of candidate $1\alpha,25(OH)_2D_3$-responsive sequences by gel shift, reporter gene and re-ChIP assays (72). Induction of gene expression was confirmed independently using quantitative PCR. By using this approach, the genes *IGFBP1, 3* and *5* were demonstrated to be primary $1\alpha,25(OH)_2D_3$ target genes. The *in silico* screening of the 174 kB of genomic sequence surrounding all six *IGFBP* genes identified 15 candidate VDREs, ten of which were shown to be functional in ChIP assays. Importantly, the *in silico* screening approach was not restricted to regulatory regions that comprise only maximal 2 kbp of sequence up- and downstream of the TSS, as in a recent whole genome screen for regulatory elements (73), but involved up to 10 kbp of flanking sequences as well as intronic and intergenic sequences. In a similar approach the *5-lipoxygenase* (*5-LO*) gene was analyzed and confirmed to be a primary $1\alpha,25(OH)_2D_3$ target gene. From the 22 putative VDREs identified in the whole *5-LO* gene sequence (-10 kB - $+74$ kbp) by *in silico* screening, at least two have

been validated to be functional *in vitro* and in living cells. One of these VDREs is located far downstream of the TSS (+42 kbp) and is one of the strongest known VDREs of the human genome (74). No functional VDRE had been previously reported for *5-LO*, since preceding studies had been restricted to the proximal promoter region (75, 76). Therefore, this approach revealed candidate VDREs that are located more than 30 kbp distant from their target gene's TSS. Based on the present understanding of enhancers, DNA looping and chromatin units being flanked by insulators or matrix attachment sites, these distances are not limiting (77).

The combination of ChIP assays with hybridization of the resulting chromatin fragments on microarrays, the so-called ChIP-chip assays, provides an additional step for a larger scale analysis of VDR target genes. The ChIP-chip technology has been applied for the analysis of the *VDR* gene itself (78), the intestinal calcium ion channel gene *TRPV6* (79), the Wnt signaling co-regulator *LRP5* (80) and the TNF-like factor *RankL* that promotes the formation of calcium resorbing osteoclasts (81). For all those genes, a number of VDR-associated chromatin regions were identified, some of which were far upstream of the gene's TSS. These studies confirmed that many, if not all, VDR target genes have multiple VDR-associated regions. However, not all of these VDREs may be functional, *i.e.* they may not contact the gene's TSS *via* DNA looping. Therefore, it is necessary to apply an additional method, the so-called chromatin-conformation-capture (3C) assay. 3C assays have confirmed the functionality of the VDREs in the *CYP27B1* (64) and the *p21*$^{WAF1/CIP1}$ (46) genes.

The next step in genome-wide association studies will be massive parallel sequencing of genomic fragments obtained after ChIP assays, also referred to as ChIP-Seq, with antibodies against VDR and its partner proteins. Results are to be published soon.

Methods for *In Silico* Screening of Transcription Factor Binding Sites

The specificity of VDR for its DNA-binding sites allows construction of a model to describe the VDRE properties that can be used to predict potential binding sites in genomic sequences. For this the VDR-binding preference, often expressed as position weight matrix, has to be described on the basis of experimental data, such as series of gel shift assays with a large number of natural binding sites (82-85). However, VDR-RXR heterodimers do not only recognize a pair of the nuclear receptor consensus binding motifs AGGTCA, but also a number of variations to it. Dependent on the individual position weight matrix description, this leads to a prediction of VDREs every 1,000 to 1,0000 bp of genomic sequence. This probably contains many false-positive predictions, which is mainly due to scoring

methodology and the limitations that are imposed by the available experimental data. Wang *et al.* combined microarray analysis and *in silico* genome-wide screens for DR3- and ER6-type VDREs (70). This approach identified several novel VDREs and VDR target genes, but most of the VDREs await a confirmation by ChIP and 3C assays.

In a position frequency matrix, the quantitative characteristics of a transcription factor, *i.e.* its relative binding strength to a number of different binding sites, is neglected, since the total number of observations of each nucleotide is simply recorded for each position. Moreover, in the past there was a positional bias of transcription factor binding sites upstream in close vicinity to the TSS. This would be apparent from the collection of identified VDREs (18), but is in contrast with a multi-genome comparison of nuclear receptor binding site distribution (73) and other reports on wide-range associations of distal regulatory sites (7).

Internet-based software tools, such as TRANSFAC (86), screen DNA sequences with databases of matrix models. The accuracy of such methods can be improved by taking the evolutionary conservation of the binding site and that of the flanking genomic region into account. Moreover, cooperative interactions between transcription factors, *i.e.* regulatory modules, can be taken into account by screening for binding site clusters. The combination of phylogenetic footprinting and position weight matrix searches applied to orthologous human and mouse gene sequences reduces the rate of false predictions by an order of magnitude, but leads to some reduction in sensitivity (87). Recent studies suggest that a surprisingly large fraction of regulatory sites may not be conserved but yet are functional, which suggests that sequence conservation revealed by alignments may not capture some relevant regulatory regions (88).

The recently published classifier method for the *in silico* screening of transcription factor binding sites (89) showed at the example of PPAR-RXR heterodimers, how a set of *in vitro* binding preferences of the three PPAR subtypes can be used as an experimental data set. Single nucleotide variants were sorted into three classes, where in class I the PPAR subtypes are able to bind the sequence with a strength of 75±15% of that of the consensus PPAR response element (PPRE), in class II with 45±15% and in class III with 15±15%. An additional 130 PPREs were sorted on the basis of counting an increasing number of variations from the consensus and taking into account the single nucleotide variant binding strength. Those variants that alone decrease the binding only modestly (class I) could be combined with even three deviations from consensus still resulting in more than 20% binding relative to consensus. Other combinations resulted in faster loss of binding detailed in 11 categories, where such combinations still resulted in more than 1% relative binding. The main advantage, when comparing the classifier to position weight matrix methods, is a clear

separation between weak PPREs and those of medium and strong strength (89). With this method the gene-dense human chromosome 19 (63.8 Mbp, 1,445 known genes) and its syntenic mouse regions (956 genes have known orthologs) were screened. Twenty percent of genes of chromosome 19 were found to contain a strong PPRE and an additional 4% have more than two medium PPREs or one proximal medium PPRE. These numbers suggest a total of 4,000 to 5,000 targets for PPARs in the human genome. Presently, the same approach is used for a genome-wide screening of VDREs and the results will be published soon. Initial results already indicated that the number of putative VDR target genes is in the same order as that of possible PPAR targets. Certainly not all sites will be accessible and the human genome also contains weak binding sites that could gain function *via* interaction with other transcription factors.

In effect, these approaches and tools are still insufficient and there has to be a focus on the creation of bioinformatics resources that include more directly the biochemical restraints to regulate gene transcription. One important aspect is that most putative transcription factor binding sites are covered by nucleosomes, so that they are not accessible to the transcription factor. This repressive environment is found in particular for those sequences that are either contained within interspersed sequences, are located isolated from transcription factor modules, or lie outside of insulator sequences marking the border of chromatin loops (90). This perspective strongly discourages the idea that isolated, simple VDREs, may be functional *in vivo*. In turn, this idea implies that the more transcription factor binding sites a given promoter region contains and the more of these transcription factors are expressed, the higher is the chance that this area of the promoter becomes locally decondensed.

The PAZAR information mall (91) is a tertiary database that is built on the resource of a multitude of secondary databases and provides a computing infrastructure for the creation, maintenance and dissemination of regulatory sequence annotation. The unambiguous identification of the chromosome location for any given transcription factor binding site using genomic coordinates allows to link the results from "big biology" projects, such as ENCODE (4), and other whole genome scans for histone modification and transcription factor association.

Conclusion

Databases, such as Oncomine (1) for gene expression data and the UCSC genome browser (92) for visualization of genome-wide ChIP data and transcription factor binding site location data, allow the combination of data from various projects. Together, these data resources may provide sufficient insight into understanding the regulation of an individual gene in a complex disease state, such as cancer.

In addition, efforts to improve bioinformatics methods predicting the binding and interaction of transcription factors together with more extensive experimental datasets will fill important gaps (87). The identification of genes showing a primary response to VDR and its ligands, the so-called VDR regulome, can be used as a prediction of their therapeutic potential as well as their possible side-effects. Methods incorporating both experimental- and informatics-derived evidence to arrive at a more reliable prediction of VDR targets and binding modules can bring all available data together with the aim of predicting the outcome in a specific context. It is envisioned that in the future the emphasis will shift from target genes to target regulatory modules to alter a physiological response, and from individual gene to whole genome response. Therefore, a much larger challenge lies ahead when we will be confronted with the higher order of regulated networks of genes, where the sum of the effects of ligand treatments may reveal itself. In an effort to study this, systems biology is being applied to the field of nuclear receptor biology, through an EU-funded Marie Curie Research Training Network, NUCSYS (www.uku.fi/nucsys).

Acknowledgements

References

1 Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Varambally R, Yu J, Briggs BB, Barrette TR, Anstet MJ, Kincead-Beal C, Kulkarni P, Varambally S, Ghosh D and Chinnaiyan AM: Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. Neoplasia *9*: 166-180, 2007.

2 Rhodes DR and Chinnaiyan AM: Integrative analysis of the cancer transcriptome. Nature Genetics *37(Suppl)*: S31-37, 2005.

3 Saramäki A, Banwell CM, Campbell MJ and Carlberg C: Regulation of the human *p21(waf1/cip1)* gene promoter *via* multiple binding sites for p53 and the vitamin D_3 receptor. Nucleic Acid Res *34*: 543-554, 2006.

4 ENCODE-Consortium: The ENCODE (ENCyclopedia Of DNA Elements) Project. Science *306*: 636-640, 2004.

5 Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES and Bernstein BE: Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*: 553-560, 2007.

6 Barrera LO, Li Z, Smith AD, Arden KC, Cavenee WK, Zhang MQ, Green RD and Ren B: Genome-wide mapping and analysis of active promoters in mouse embryonic stem cells and adult organs. Genome Res *18*: 46-59, 2008.

7 Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I and Zhao K: High-resolution profiling of histone methylations in the human genome. Cell *129*: 823-837, 2007.

8 Hallikas O, Palin K, Sinjushina N, Rautiainen R, Partanen J, Ukkonen E and Taipale J: Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. Cell *124*: 47-59, 2006.

9 Johnson DS, Mortazavi A, Myers RM and Wold B: Genome-wide mapping of *in vivo* protein-DNA interactions. Science *316*: 1497-1502, 2007.

10 Sperling S: Transcriptional regulation at a glance. BMC bioinformatics *8(Suppl 6)*: S2, 2007.

11 Jones SJ: Prediction of genomic functional elements. Ann Rev Genom Hum Gen *7*: 315-338, 2006.

12 Chawla A, Repa JJ, Evans RM and Mangelsdorf DJ: Nuclear receptors and lipid physiology: opening the X-files. Science *294*: 1866-1870, 2001.

13 Maglich JM, Sluder A, Guan X, Shi Y, McKee DD, Carrick K, Kamdar K, Willson TM and Moore JT: Comparison of complete nuclear receptor sets from the human, *Caenorhabditis elegans* and *Drosophila* genomes. Genome Biol *2*: R29, 2001.

14 Rochel N, Wurtz JM, Mitschler A, Klaholz B and Moras D: Crystal structure of the nuclear receptor for vitamin D bound to its natural ligand. Mol Cell *5*: 173-179, 2000.

15 Moras D and Gronemeyer H: The nuclear receptor ligand-binding domain: structure and function. Curr Opin Cell Biol *10*: 384-391, 1998.

16 Nuclear Receptor Committee: A unified nomenclature system for the nuclear receptor superfamily. Cell *97*: 161-163, 1999.

17 Bikle DD, Gee E and Pillai S: Regulation of keratinocyte growth, differentiation, and vitamin D metabolism by analogs of 1,25-dihydroxyvitamin D. J Invest Dermatol *101*: 713-718, 1993.

18 Carlberg C and Polly P: Gene regulation by vitamin D_3. Crit Rev Eukaryot Gene Expr *8*: 19-42, 1998.

19 Carlberg C: Mechanisms of nuclear signalling by vitamin D_3. Interplay with retinoid and thyroid hormone signalling. Eur J Biochem *231*: 517-527, 1995.

20 Bookout AL, Jeong Y, Downes M, Yu RT, Evans RM and Mangelsdorf DJ: Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network. Cell *126*: 789-799, 2006.

21 Norman AW: Receptors for $1\alpha,25(OH)_2D_3$: past, present, and future. J Bone Mineral Res *13*: 1360-1369, 1998.

22 Kato S, Takeyama K-i, Kitanaka S, Murayama A, Sekine K and Yoshizawa T: *In vivo* function of VDR in gene expression-VDR knock-out mice. J Ster Biochem Mol Biol *69*: 247-251, 1999.

23 Welsh J: Vitamin D and breast cancer: insights from animal models. Am J Clin Nutr *80*: 1721S-1724S, 2004.

24 Carlberg C and Saurat JH: Vitamin D-retinoid association: molecular basis and clinical applications. J Investig Dermatol Symp Proc *1*: 82-86, 1996.

25 Carlberg C: The antiproliferative effect of vitamin D_3 analogues. Dermatology *192*: 195-197, 1996.

26 Schwartz GG: Vitamin D and the epidemiology of prostate cancer. Semin Dial *18*: 276-289, 2005.

27 Carlberg C, Bendik I, Wyss A, Meier E, Sturzenbecker LJ, Grippo JF and Hunziker W: Two nuclear signalling pathways for vitamin D. Nature *361*: 657-660, 1993.

28 Cheskis B and Freedman LP: Ligand modulates the conversion of DNA-bound vitamin D^3 receptor (VDR) homodimers into VDR-retinoid X receptor heterodimers. Mol Cell Biol *14*: 3329-3338, 1994.

29 Schräder M and Carlberg C: Thyroid hormone and retinoic acid receptors form heterodimers with retinoid X receptors on direct repeats, palindromes, and inverted palindromes. DNA Cell Biol *13*: 333-341, 1994.

30 Schräder M, Müller KM, Nayeri S, Kahlen JP and Carlberg C: VDR-T3R receptor heterodimer polarity directs ligand sensitivity of transactivation. Nature *370*: 382-386, 1994.

31 Schräder M, Bendik I, Becker-Andre M and Carlberg C: Interaction between retinoic acid and vitamin D signaling pathways. J Biol Chem *268*: 17830-17836, 1993.

32 Mangelsdorf DJ and Evans RM: The RXR heterodimers and orphan receptors. Cell *83*: 841-850, 1995.

33 Quack M and Carlberg C: Ligand-triggered stabilization of vitamin D receptor/retinoid X receptor heterodimer conformations on DR4-type response elements. J Mol Biol *296*: 743-756, 2000.

34 Schräder M, Nayeri S, Kahlen JP, Müller KM and Carlberg C: Natural vitamin D_3 response elements formed by inverted palindromes: polarity-directed ligand sensitivity of vitamin D_3 receptor-retinoid X receptor heterodimer-mediated transactivation. Mol Cell Biol *15*: 1154-1161, 1995.

35 Jenuwein T and Allis CD: Translating the histone code. Science *293*: 1074-1080, 2001.

36 Aranda A and Pascual A: Nuclear hormone receptors and gene expression. Physiol Rev *81*: 1269-1304, 2001.

37 Burke LJ and Baniahmad A: Co-repressors 2000. FASEB J *14*: 1876-1888, 2000.

38 Polly P, Herdick M, Moehren U, Baniahmad A, Heinzel T and Carlberg C: VDR-Alien: a novel, DNA-selective vitamin D_3 receptor-corepressor partnership. FASEB J *14*: 1455-1463, 2000.

39 Privalsky ML: The role of corepressors in transcriptional regulation by nuclear hormone receptors. Annu Rev Physiol *66*: 315-360, 2004.

40 Leo C and Chen JD: The SRC family of nuclear receptor coactivators. Gene *245*: 1-11, 2000.

41 Castillo AI, Jimenez-Lara AM, Tolon RM and Aranda A: Synergistic activation of the prolactin promoter by vitamin D receptor and GHF-1: role of coactivators, CREB-binding protein and steroid hormone receptor coactivator-1 (SRC-1). Mol Endocrinol *13*: 1141-1154, 1999.

42 Rachez C, Suldan Z, Ward J, Chang C-P, Burakov D, Erdjument-Bromage H, Tempst P and Freedman LP: A novel protein complex that interacts with the vitamin D_3 receptor in a ligand-dependent manner and enhances transactivation in a cell-free system. Genes Develop *12*: 1787-1800, 1998.

43 Rachez C, Lemon BD, Suldan Z, Bromleigh V, Gamble M, Näär AM, Erdjument-Bromage H, Tempst P and Freedman LP: Ligand-dependent transcription activation by nuclear receptors requires the DRIP complex. Nature *398*: 824-828, 1999.

44 Liu M, Lee M-H, Cohen M, Bommakanti M and Freedman LP: Transcriptional activation of the Cdk inhibitor p21 by vitamin D_3 leads to the induced differentiation of the myelomonocytic cell line U937. Genes Develop *10*: 142-153, 1996.

45 Warrener R, Beamish H, Burgess A, Waterhouse NJ, Giles N, Fairlie D and Gabrielli B: Tumor cell-selective cytotoxicity by targeting cell cycle checkpoints. FASEB J *17*: 1550-1552, 2003.

46 Malinen M, Saramäki A, Ropponen A, Degenhardt T, Väisänen S and Carlberg C: Distinct HDACs regulate the transcriptional response of human cyclin-dependent kinase inhibitor genes to Trichostatin A and $1\alpha,25$-dihydroxyvitamin D_3. Nucleic Acid Res *36*: 121-132, 2008.

47 Khanim FL, Gommersall LM, Wood VH, Smith KL, Montalvo L, O'Neill LP, Xu Y, Peehl DM, Stewart PM, Turner BM and Campbell MJ: Altered SMRT levels disrupt vitamin D₃ receptor signalling in prostate cancer cells. Oncogene *23*: 6712-6725, 2004.

48 Oda Y, Sihlbom C, Chalkley RJ, Huang L, Rachez C, Chang CP, Burlingame AL, Freedman LP and Bikle DD: Two distinct coactivators, DRIP/mediator and SRC/p160, are differentially involved in vitamin D receptor transactivation during keratinocyte differentiation. Mol Endocrinol *17*: 2329-2339, 2003.

49 Metivier R, Penot G, Hubner MR, Reid G, Brand H, Kos M and Gannon F: Estrogen receptor α directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. Cell *115*: 751-763, 2003.

50 Kim S, Shevde NK and Pike JW: 1,25-Dihydroxyvitamin D3 stimulates cyclic vitamin D receptor/retinoid X receptor DNA-binding, co-activator recruitment, and histone acetylation in intact osteoblasts. J Bone Miner Res *20*: 305-317, 2005.

51 Sutton AL and MacDonald PN: Vitamin D: more than a "bone-a-fide" hormone. Mol Endocrinol *17*: 777-791, 2003.

52 Jiang H, Lin J, Su Z-z, Collart FR, Huberman E and Fisher PB: Induction of differentiation in human promyelotic HL-60 leukemia cells activates p21, WAF1/CIP1, expression in the absence of p53. Oncogene *9*: 3397-3406, 1994.

53 Kaldis P: Another piece of the p27^{Kip1} puzzle. Cell *128*: 241-244, 2007.

54 Rickert P, Seghezzi W, Shanahan F, Cho H and Lees E: Cyclin C/CDK8 is a novel CTD kinase associated with RNA polymerase II. Oncogene *12*: 2631-2640, 1996.

55 Bourbon HM, Aguilera A, Ansari AZ, Asturias FJ, Berk AJ, Björklund S, Blackwell TK, Borggrefe T, Carey M, Carlson M, Conaway JW, Conaway RC, Emmons SW, Fondell JD, Freedman LP, Fukasawa T, Gustafsson CM, Han M, He X, Herman PK, Hinnebusch AG, Holmberg S, Holstege FC, Jaehning JA, Kim YJ, Kuras L, Leutz A, Lis JT, Meisterernest M, Näär AM, Nasmyth K, Parvin JD, Ptashne M, Reinberg D, Ronne H, Sadowski I, Sakurai H, Sipiczki M, Sternberg PW, Stillman DJ, Strich R, Struhl K, Svejstrup JQ, Tuck S, Winston F, Roeder RG and Kornberg RD: A unified nomenclature for protein subunits of mediator complexes linking transcriptional regulators to RNA polymerase II. Mol Cell *14*: 553-557, 2004.

56 Li H, Lahti JM, Valentine M, Saito M, Reed SI, Look AT and Kidd VJ: Molecular cloning and chromosomal localization of the human cyclin C (*CCNC*) and cyclin E (*CCNE*) genes: deletion of the *CCNC* gene in human tumors. Genomics *32*: 253-259, 1996.

57 Väisänen S, Dunlop TW, Sinkkonen L, Frank C and Carlberg C: Spatio-temporal activation of chromatin on the human *CYP24* gene promoter in the presence of 1α,25-dihydroxyvitamin D₃. J Mol Biol *350*: 65-77, 2005.

58 Swami S, Raghavachari N, Muller UR, Bao YP and Feldman D: Vitamin D growth inhibition of breast cancer cells: gene expression patterns assessed by cDNA microarray. Breast Cancer Res Treat *80*: 49-62, 2003.

59 Palmer HG, Sanchez-Carbayo M, Ordonez-Moran P, Larriba MJ, Cordon-Cardo C and Munoz A: Genetic signatures of differentiation induced by 1α,25-dihydroxyvitamin D₃ in human colon cancer cells. Cancer Res *63*: 7799-7806, 2003.

60 Dunlop TW, Väisänen S, Frank C, Molnar F, Sinkkonen L and Carlberg C: The human peroxisome proliferator-activated receptor d gene is a primary target of 1α,25-dihydroxyvitamin D₃ and its nuclear receptor. J Mol Biol *349*: 248-260, 2005.

61 Reed KR, Sansom OJ, Hayes AJ, Gescher AJ, Winton DJ, Peters JM and Clarke AR: PPARδ status and Apc-mediated tumouri-genesis in the mouse intestine. Oncogene *23*: 8992-8996, 2004.

62 White JH: Profiling 1,25-dihydroxyvitamin D₃-regulated gene expression by microarray analysis. J Ster Biochem Mol Biol *89-90*: 239-244, 2004.

63 Murayama A, Kim MS, Yanagisawa J, Takeyama K and Kato S: Transrepression by a liganded nuclear receptor *via* a bHLH activator through co-regulator switching. Embo J *23*: 1598-1608, 2004.

64 Turunen MM, Dunlop TW, Carlberg C and Väisänen S: Selective use of multiple vitamin D response elements underlies the 1α,25-dihydroxyvitamin D₃-mediated negative regulation of the human *CYP27B1* gene. Nucleic Acids Res *35*: 2734-2747, 2007.

65 Peehl DM, Shinghal R, Nonn L, Seto E, Krishnan AV, Brooks JD and Feldman D: Molecular activity of 1,25-dihydroxyvitamin D₃ in primary cultures of human prostatic epithelial cells revealed by cDNA microarray analysis. J Ster Biochem Mol Biol *92*: 131-141, 2004.

66 Krishnan AV, Shinghal R, Raghavachari N, Brooks JD, Peehl DM and Feldman D: Analysis of vitamin D-regulated gene expression in LNCaP human prostate cancer cells using cDNA microarrays. Prostate *59*: 243-251, 2004.

67 Ikezoe T, Gery S, Yin D, O'Kelly J, Binderup L, Lemp N, Taguchi H and Koeffler HP: CCAAT/enhancer-binding protein delta: a molecular target of 1,25-dihydroxyvitamin D₃ in androgen-responsive prostate cancer LNCaP cells. Cancer Res *65*: 4762-4768, 2005.

68 Eelen G, Verlinden L, van Camp M, van Hummelen P, Marchal K, de Moor B, Mathieu C, Carmeliet G, Bouillon R and Verstuyf A: The effects of 1α,25-dihydroxyvitamin D₃ on the expression of DNA replication genes. J Bone Miner Res *19*: 133-146, 2004.

69 Eelen G, Verlinden L, Van Camp M, Mathieu C, Carmeliet G, Bouillon R and Verstuyf A: Microarray analysis of 1α,25-dihydroxyvitamin D₃-treated MC3T3-E1 cells. J Ster Biochem Mol Biol *89-90*: 405-407, 2004.

70 Wang TT, Tavera-Mendoza LE, Laperriere D, Libby E, MacLeod NB, Nagai Y, Bourdeau V, Konstorum A, Lallemant B, Zhang R, Mader S and White JH: Large-scale *in silico* and microarray-based identification of direct 1,25-dihydroxyvitamin D₃ target genes. Mol Endocrinol *19*: 2685-2695, 2005.

71 Sinkkonen L, Malinen M, Saavalainen K, Väisänen S and Carlberg C: Regulation of the human cyclin C gene *via* multiple vitamin D₃-responsive regions in its promoter. Nucleic Acids Res *33*: 2440-2451, 2005.

72 Matilainen M, Malinen M, Saavalainen K and Carlberg C: Regulation of multiple insulin-like growth factor binding protein genes by 1α,25-dihydroxyvitamin D₃. Nucleic Acids Res *33*: 5521-5532, 2005.

73 Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, Lindblad-Toh K, Lander ES and Kellis M: Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. Nature *434*: 338-345, 2005.

74 Seuter S, Väisänen S, Radmark O, Carlberg C and Steinhilber D: Functional characterization of vitamin D responding regions in the human 5-lipoxygenase gene. Biochim Biophys Acta *1771*: 864-872, 2007.

75 Klan N, Seuter S, Schnur N, Jung M and Steinhilber D: Trichostatin A and structurally related histone deacetylase inhibitors induce 5-lipoxygenase promoter activity. Biol Chem *384*: 777-785, 2003.

76 Sorg BL, Klan N, Seuter S, Dishart D, Radmark O, Habenicht A, Carlberg C, Werz O and Steinhilber D: Analysis of the 5-lipoxygenase promoter and characterization of a vitamin D receptor-binding site. Biochim Biophys Acta *1761*: 686-697, 2006.

77 Ogata K, Sato K and Tahirov TH: Eukaryotic transcriptional regulatory complexes: cooperativity from near and afar. Curr Opin Struct Biol *13*: 40-48, 2003.

78 Zella LA, Kim S, Shevde NK and Pike JW: Enhancers located within two introns of the vitamin D receptor gene mediate transcriptional autoregulation by 1,25-dihydroxyvitamin D_3. Mol Endocrinol *20*: 1231-1247, 2006.

79 Meyer MB, Watanuki M, Kim S, Shevde NK and Pike JW: The human transient receptor potential vanilloid type 6 distal promoter contains multiple vitamin D receptor binding sites that mediate activation by 1,25-dihydroxyvitamin D_3 in intestinal cells. Mol Endocrinol *20*: 1447-1461, 2006.

80 Fretz JA, Zella LA, Kim S, Shevde NK and Pike JW: 1,25-Dihydroxyvitamin D_3 induces expression of the Wnt signaling co-regulator LRP5 *via* regulatory elements located significantly downstream of the gene's transcriptional start site. J Ster Biochem Molec Biol *103*: 440-445, 2007.

81 Kim S, Yamazaki M, Zella LA, Shevde NK and Pike JW: Activation of receptor activator of NF-kappaB ligand gene expression by 1,25-dihydroxyvitamin D_3 is mediated through multiple long-range enhancers. Mol Cell Biol *26*: 6469-6486, 2006.

82 Mader S, Leroy P, Chen J-Y and Chambon P: Multiple parameters control the selectivity of nuclear receptors for their response elements. J Biol Chem *268*: 591-600, 1993.

83 Schräder M, Müller KM, Becker-André M and Carlberg C: Response element selectivity for heterodimerization of vitamin D receptors with retinoic acid and retinoid X receptors. J Mol Endocrinol *12*: 327-339, 1994.

84 Schräder M, Müller KM and Carlberg C: Specificity and flexibility of vitamin D signaling. Modulation of the activation of natural vitamin D response elements by thyroid hormone. J Biol Chem *269*: 5501-5504, 1994.

85 Schräder M, Becker-Andre M and Carlberg C: Thyroid hormone receptor functions as monomeric ligand-induced transcription factor on octameric half-sites. Consequences also for dimerization. J Biol Chem *269*: 6444-6449, 1994.

86 Matys V, Fricke E, Geffers R, Gossling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV, Kloos DU, Land S, Lewicki-Potapov B, Michael H, Munch R, Reuter I, Rotert S, Saxel H, Scheer M, Thiele S and Wingender E: TRANSFAC: transcriptional regulation, from patterns to profiles. Nucleic Acids Res *31*: 374-378, 2003.

87 Wasserman WW and Sandelin A: Applied bioinformatics for the identification of regulatory elements. Nat Rev Genet *5*: 276-287, 2004.

88 Odom DT, Dowell RD, Jacobsen ES, Gordon W, Danford TW, MacIsaac KD, Rolfe PA, Conboy CM, Gifford DK and Fraenkel E: Tissue-specific transcriptional regulation has diverged significantly between human and mouse. Nature Genetics *39*: 730-732, 2007.

89 Heinäniemi M, Uski JO, Degenhardt T and Carlberg C: Meta-analysis of primary target genes of peroxisome proliferator-activated receptors. Genome Biol *8*: R147, 2007.

90 Burns JL, Jackson DA and Hassan AB: A view through the clouds of imprinting. FASEB J *15*: 1694-1703, 2001.

91 Portales-Casamar E, Kirov S, Lim J, Lithwick S, Swanson MI, Ticoll A, Snoddy J and Wasserman WW: PAZAR: a framework for collection and dissemination of *cis*-regulatory sequence annotation. Genome Biol *8*: R207, 2007.

92 Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM and Haussler D: The human genome browser at UCSC. Genome Res *12*: 996-1006, 2002.