Nuclear hormone 1α,25-dihydroxyvitamin D₃ elicits a genome-wide shift in the locations of VDR chromatin occupancy

Sami Heikkinen[1], Sami Väisänen[1], Petri Pehkonen[1], Sabine Seuter[2], Vladimir Benes[3] and Carsten Carlberg[1,2,*]

[1]Department of Biosciences, University of Eastern Finland, FIN-70210 Kuopio, Finland, [2]Life Sciences Research Unit, University of Luxembourg, L-1511 Luxembourg, Luxembourg and [3]Genomics Core Facility, EMBL, D-69117 Heidelberg, Germany

ABSTRACT

A global understanding of the actions of the nuclear hormone 1α,25-dihydroxyvitamin D₃ (1α,25(OH)₂D₃) and its vitamin D receptor (VDR) requires a genome-wide analysis of VDR binding sites. In THP-1 human monocytic leukemia cells we identified by ChIP-seq 2340 VDR binding locations, of which 1171 and 520 occurred uniquely with and without 1α,25(OH)₂D₃ treatment, respectively, while 649 were common. *De novo* identified direct repeat spaced by 3 nucleotides (DR3)-type response elements (REs) were strongly associated with the ligand-responsiveness of VDR occupation. Only 20% of the VDR peaks diminishing most after ligand treatment have a DR3-type RE, in contrast to 90% for the most growing peaks. Ligand treatment revealed 638 1α,25(OH)₂D₃ target genes enriched in gene ontology categories associated with immunity and signaling. From the 408 upregulated genes, 72% showed VDR binding within 400 kb of their transcription start sites (TSSs), while this applied only for 43% of the 230 downregulated genes. The VDR loci showed considerable variation in gene regulatory scenarios ranging from a single VDR location near the target gene TSS to very complex clusters of multiple VDR locations and target genes. In conclusion, ligand binding shifts the locations of VDR occupation to DR3-type REs that surround its target genes and occur in a large variety of regulatory constellations.

BACKGROUND

The vitamin D receptor (VDR) is an endocrine member of the nuclear receptor superfamily, since it is activated already by sub-nanomolar concentrations of its natural ligand 1α,25-dihydroxyvitamin D₃ (1α,25(OH)₂D₃) (1). The classical, physiological role of 1α,25(OH)₂D₃ is the regulation of calcium and phosphate homeostasis and bone mineralization (2), but there is both epidemiological and pre-clinical evidence that VDR ligands have also anti-proliferative and immuno-modulatory actions (3,4). VDR is nearly ubiquitously expressed (5) and many cell types [for example, bone, skin and monocytes (6)] are capable of metabolizing the main circulating form of vitamin D, 25-hydroxyvitamin D₃ (25(OH)₂D₃), to 1α,25(OH)₂D₃ by the enzyme 25(OH)D₃-1α-hydroxylase, encoded by the gene *CYP27B1*, i.e. 1α,25(OH)₂D₃ can have in these tissues both auto- and paracrine roles.

In recent years, an increasing amount of data has shown a physiological function of 1α,25(OH)₂D₃ in the immune system, both in defense and in maintenance of self-tolerance (7). VDR ligands affect the immune system in many different ways, but the most important seems to be the enhanced capacity of innate anti-bacterial defense and a more tolerogenic profile toward autoimmune phenomena (4). One important target cell of 1α,25(OH)₂D₃ is the monocyte/macrophage, where the nuclear hormone promotes the anti-bacterial defense system by inducing the secretion of anti-bacterial proteins, such as cathelicidin anti-microbial peptide (CAMP) (8). Importantly, decreased anti-bacterial and anti-viral defenses are obvious in vitamin D deficiency, and restoration of the vitamin D-sufficient state restores immune responses (4).

As a member of the nuclear receptor superfamily VDR acts as a transcription factor that binds to specific vitamin D response elements (VDREs) within the regulatory regions of its primary target genes (1). VDR forms a heterodimer with retinoid X receptor (RXR), another member of the nuclear receptor superfamily, and binds preferentially to VDREs that are formed as direct repeat

*To whom correspondence should be addressed. Tel: +358 40 355 3062; Fax: +358 17 2811510; Email: carsten.carlberg@uef.fi

of two hexameric core binding motifs with 3 nucleotides spacing, so-called DR3-type VDREs (9). However, also other types of VDREs, such as direct repeats with 4 intervening nucleotides (DR4) (10) or everted repeats (ERs) with a spacer of 6–9 nucleotides (ER6-9) (11,12) have been reported. Single gene analyses suggest that VDR target genes often contain multiple functional VDREs (13–16).

VDR belongs to those nuclear receptors that bind to their genomic targets already in the absence of ligand. Under this constellation VDR has been shown to associate via co-repressor proteins with histone deacetylases and to repress its target genes (17). Ligand binding induces a conformational change that leads to dissociation of co-repressors and to the association of different types of co-activator proteins that on one hand lead to chromatin opening and on the other hand are part of the mediator complex, which forms a bridge to the basal transcription machinery associated with the transcription start sites (TSSs) of primary VDR target genes (18).

Thus far, both single gene and microarray analyses in various tissues and cells from different species have suggested a long list of VDR target genes. However, the actions of VDR seem to be highly cell-specific, since there is only very little overlap of VDR targets between the different tissues (19–22). Moreover, such studies have often been performed at such late time points that they did not allow distinguishing primary VDR targets from secondary ones. Recently, Ramagopalan *et al.* (23) reported the first genome-wide view on VDR locations in human lymphoblastoid cell lines suggesting 2776 genomic VDR locations after ligand treatment and 229 target genes. However, also these data were obtained from a very late time point (36 h after ligand stimulation) and therefore may not fully represent the primary actions of VDR.

In this study, we used undifferentiated THP-1 human monocytic leukemia cells as a model system, since they reflect reasonably well the $1\alpha,25(OH)_2D_3$ response of primary human monocytes (24,25). To gain a genome-wide view of the primary actions of VDR in these cells, we determined its genomic binding sites after 40 min stimulation with $1\alpha,25(OH)_2D_3$ and in the unstimulated state and compared it with the VDR target genes after 4 h ligand treatment. We chose human monocytes as cellular model out of interest in the biological role of $1\alpha,25(OH)_2D_3$ signaling in the initial events of immune response. However, the main focus of this study is of mechanistic nature showing that VDR is associated with chromatin even in the absence of ligand, but that the balance of VDR binding shifts from locations that are gene-centered and lacking classical VDREs to locations that are more distal to the genes and contain one or more DR3-type response elements (REs).

MATERIAL AND METHODS

Cell culture

THP-1 human monocytic leukemia cells were grown in RPMI 1640 medium supplemented with 10% fetal calf serum, 2 mM L-glutamine, 0.1 mg/ml streptomycin and 100 U/ml penicillin and the cells were kept at 37°C in a humidified 95% air/5% CO_2 incubator. Prior to mRNA or chromatin extraction, cells were grown overnight in phenol red-free medium supplemented with charcoal-stripped fetal bovine serum. Then, for RNA extractions cells were treated for 4 h with $1\alpha,25(OH)_2D_3$ (Sigma-Aldrich) or with solvent (ethanol, finally 0.1%) or for chromatin immunoprecipitation (ChIP) assays for 40 min with $1\alpha,25(OH)_2D_3$ or left unstimulated.

RNA extraction and cDNA synthesis

Total RNA was extracted using the Mini RNA Isolation II Kit (Zymo Research, HiSS Diagnostics, Freiburg, Germany). cDNA synthesis was performed for 60 min at 37°C using 1 μg of total RNA as a template, 100 pmol oligo(dT)$_{18}$ primers (Eurogentec, Seraing, Belgium), 500 μM dNTPs, 40 U Ribolock Ribonuclease Inhibitor and 40 U MMuLV reverse transcriptase (Fermentas, St Leon-Rot, Germany). Prior to real-time quantitative PCR (qPCR) reaction, the cDNA was diluted 10-fold.

qPCR

qPCR was performed using an ABI 7500 Fast System (Applied Biosystems, Lennik, Belgium). The reactions were performed using 250 nM of reverse and forward primers (for sequence see Supplementary Table S1), 4 μl 1/10 diluted cDNA template and the ABsolute Q-PCR SYBR Green Low ROX Mix (Abgene, Westburg, Leusden, The Netherlands) in a total volume of 20 μl. In the PCR reaction Hotstart Taq polymerase was activated for 15 min at 95°C, followed by 40 amplification cycles of 30 s denaturation at 95°C, 30 s annealing at primer-specific temperatures (Supplementary Table S1) and 40 s elongation at 72°C and a final elongation for 5 min at 72°C. PCR product specificity was monitored using post-PCR melt curve analysis. Relative expression levels were determined with the comparative delta threshold cycle (ΔC_t) method. Amplification efficiencies of the primer pairs were taken into account and have been determined on the basis of a standard curve of a cDNA dilution series (with two exceptions, all PCR efficiencies were between 80% and 120%). Relative expression levels of the target genes were normalized to the three most stable of 10 tested internal control genes (*b2M*, *GAPDH*, *HPRT1*). The expression stability of the control genes was determined by the geNorm algorithm (26). Briefly, the arithmetic mean of replicated C_t values for each gene is transformed to a relative quantity (setting the sample with the highest expression as calibrator to 1), using the ΔC_t formula $Q = E^{\Delta Ct} = E^{(\text{calibratorCt} - \text{sampleCt})}$ (Q = quantity sample relative to calibrator sample; E = amplification efficiency). For normalization, the relative quantities were divided by the normalization factor being the geometric mean of the three reference genes.

Microarray analysis

Total RNA was checked for RNA integrity using a Biorad Experion electrophoresis station (Biorad, Nazareth, Belgium) and analyzed on Sentrix Human-6 v2 Expression BeadChips from Illumina (San Diego, CA, USA) at the

Finnish Microarray Centre (Turku, Finland) using protocols recommended by the manufacturer. Raw microarray data is available at Gene Expression Omnibus (GEO) under accession number GSE27270 (SubSeries of GSE27438). Microarray data analysis was performed using R statistical software version 2.11 (R Dev Core) with associated libraries from Bioconductor project version 2.6 (27). Data was normalized using VST transformation and RSN normalization used as standard approach for Illumina arrays with the lumi package (28). Probe sets that were not linked to any known or predicted human gene were filtered out. Linear Models for Microarray Data (limma) package (29) using linear model fitting for statistical testing with empirical Bayes variance smoothing procedure was applied to detection of differentially expressed genes. Obtained *P*-values were corrected for multiple testing using Benjamini–Hochberg false discovery rate (FDR) procedure (30). To identify the biological processes targeted by the $1\alpha,25(OH)_2D_3$-treatment, we used GeneTrail (31) to analyze the enrichment of up and downregulated VDR target genes into GO categories using default settings.

ChIP-seq

After ligand or no treatment of THP-1 cells, nuclear proteins were cross-linked to DNA by adding formaldehyde directly to the medium to a final concentration of 1% and incubating at room temperature for 15 min on a rocking platform. Cross-linking was stopped by adding glycine to a final concentration of 0.125 M and incubating at room temperature for 5 min on a rocking platform. The cells were collected by centrifugation and washed twice with ice cold PBS (140 mM NaCl, 2.7 mM KCl, 1.5 mM KH_2PO_4, 8.1 mM $Na_2HPO_4\cdot2H_2O$). The cell pellets were resuspended in 1 ml of lysis buffer (1% SDS, 10 mM EDTA, protease inhibitors, 50 mM Tris–HCl, pH 8.1) and the lysates were sonicated to result in DNA fragments of 200–400 bp. Cellular debris was removed by centrifugation. Aliquots of 100 µl of the lysate were diluted 1:10 in ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 167 mM NaCl, protease inhibitors, 16.7 mM Tris–HCl, pH 8.1) and 1 µg of anti-VDR antibody (ab3508-100, Abcam, Cambridge, UK) or non-specific IgG (12-370, Millipore, Temecula, CA, USA) were added and the samples were incubated for overnight at 4°C on a rotating platform. The immunocomplexes were collected using 20 µl of Magna ChIP protein A magnetic beads (Millipore) at 4°C for 1 h with rotation. The beads were washed sequentially for 5 min in rotating platform with 1 ml of the following buffers: low salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris–HCl, pH 8.1), high salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 500 mM NaCl, 20 mM Tris–HCl, pH 8.1) and LiCl wash buffer (0.25 M LiCl, 1% Nonidet P-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris–HCl, pH 8.1). Finally, the beads were washed twice with 1 ml TE buffer (1 mM EDTA, 10 mM Tris–HCl, pH 8.0) and the immune complexes were eluted twice using 250 µl elution buffer (1% SDS, 100 mM $NaHCO_3$) at room temperature for 15 min with rotation. The supernatants were combined and the immune complexes were reverse cross-linked at 64°C overnight in the presence of proteinase K (Fermentas) in a final concentration of 40 µg/ml. DNA was extracted with phenol/chloroform/isoamyl alcohol (25/24/1) and precipitated with 0.1 volumes of 3 M sodium acetate, pH 5.2 and 2 volumes of ethanol using glycogen as carrier. For ChIP-seq five individual VDR ChIP samples were pooled and concentrated using a SpeedVac. The ChIP templates were sequenced using a Solexa Gene Analyzer II platform (Illumina) at 36 bp read length using standard manufacturer protocols at the Genomics Core Facility in Heidelberg, Germany. For the validation of the ChIP-seq data, selected genomic regions containing VDR peaks were analyzed by quantitative ChIP analysis of independent samples using Maxima^TM Sybr Green qPCR Master Mix (Fermentas) and the genomic primers listed in Supplementary Table S2. In addition to VDR and IgG antibodies, immunoprecipitation was performed using 1 µg of anti-RXR antibody (sc-553, Santa Cruz Biotechnologies, Santa Cruz, CA, USA). The qPCR reactions were performed with a LightCycler 480 apparatus (Roche) using the following PCR profile: 10 min at 95°C, followed by 50 cycles of 30 s at 95°C, 30 s at 58°C and 30 s at 72°C. The results were normalized with respect to input. Non-specific IgG results were subtracted by using the formula $2^{-(\Delta Cp)}*100_{(\text{specific antibody})} - 2^{-(\Delta Cp)}*100_{(\text{non-specific IgG})}$, where ΔCp is $Cp_{(\text{immunoprecipitated DNA})} - Cp_{(\text{input})}$ and Cp is the fractional cycle number. Statistical significance was calculated in reference to non-treated sample using two-tailed Student's *t*-test.

ChIP-seq data analysis

Alignment of sequence reads against the human reference genome version hg19 was done using Bowtie software version 0.12.2. (32) with the following command line arguments: `bowtie -n 1 -m 1 -e 70 -l 28 -k 1 -t -p 8 -q -S --best hg19 input_file_name output_file_name`. Of note, one Solexa run with 1, 2 and 1 lanes of IgG, unstimulated and ligand-treated VDR ChIP sample sequencing, respectively, suffered from a minor post-sequencing processing error at the sequencing platform that had resulted in an omission of the 16th nucleotide from the large majority of reads. This was corrected for by inserting an N at the missing position in each read and using two modified settings in the Bowtie alignment (`-n 2 -m 3`). The short read data publically available under accession number GSE27437 (SubSeries of GSE27438) contains the corrected reads. Aligned reads from multiple lanes were combined per sample. In the case of ChIP-seq analysis, the common IgG background read set was scaled so that an individual background was generated for each VDR ChIP-seq sample by randomly removing IgG reads such that the read ratio IgG/VDR became the same for each VDR ChIP-seq sample. The scaled IgG and un-scaled VDR reads were then converted to sorted BAM format using SAM tools (33) and subsequently also to TDF format using igvtools to allow visualization in the IGV genome browser (34).

For ChIP-seq analysis statistically significant peaks were identified using MACS program version 1.3.7.1 (35) with the following command line arguments: `macs --pvalue=1e-2 --nomodel --wig -t input_sample_file_name -c input_control_file_name --tsize=36 --format=BAM --name=analysis_topic --mfold=13 --shiftsize=250 --bw=250 --verbose=3`. A loose initial *P*-value cutoff was purposeful used to enable catching all true signals with less regard to possible false positive signals at this early phase of analysis. Subsequent refinement of MACS peaks was done using PeakSplitter program version 1.0 (36) with the following command line arguments: `PeakSplitter -p peak_folder_name -w aligned_read_wig_-folder_name -o output_folder -c 5 -v 0.6 -f`. An in-house R script was used to calculate fold enrichments (FEs) over IgG, *P*-values and FDR estimates for the subpeaks in an identical manner used internally in MACS 1.3.7.1. Since splitting the original peaks into several subpeaks may produce large sets of weaker flanking residual peaks, thus biasing for example the genomic and FE distributions of the peaks, we kept for further analysis only those peaks that either had FDR <1% or were the best subpeak (by FDR) within their parental MACS peak.

Further data analysis was conducted using an in-house R pipeline containing tools for identifying peaks overlapping in two samples, the analysis of genomic distribution of peaks and the integration of peak and gene expression data sets. Of note, since there are genomic positions that cannot be uniquely assigned to the selected 10 types of genomic elements used in the analysis of peak distributions therein, a prioritization scheme was employed, where the peaks were uniquely overlapped to the elements in a step-wise exclusive scheme, starting from coding elements (UTRs and coding exons), and then moving to introns and outward from the gene.

Detection of transcription factor binding sites

For the *de novo* identification of the VDR binding consensus sequence within ±100 bp of the summits of all 2340 FDR <1% VDR peaks we employed the rGADEM R-package with the following settings: `GADEM (Sequences, genome=Hsapiens, weightType =1, populationSize=500, numGeneration=20)`. Of note, the default settings did not identify the DR3-like sequence in the full peak set. The resulting candidate position weight matrices (PWMs) were then compared to the set of known PWMs in the JASPAR database using the MotIV R-package, and their locations mapped back to the original peaks using in-house R-scripts. To scan the peak sequences for PWMs representing different types of DRs and ERs with varying spacer lengths, we used 'matrix-scan (quick)' that is one of the regulatory sequence analysis tools [RSAT, available at http://rsat.ulb.ac.be/rsat/ (37)]. For the RSAT scans, background was modeled at Markov order 2 from the 'upstream-noorf' subset of the human genome and a *P*-value cutoff of 1×10^{-4} was used.

RESULTS

Identification of VDR target genes

To identify VDR target genes, THP-1 cells were treated for 4 h with either vehicle (ethanol) or $1\alpha,25(OH)_2D_3$ and collected for the extraction of total RNA that was subjected to gene expression microarray analysis. We chose the 4 h time point as a compromise to obtain reasonable mRNA up or downregulation of VDR target genes, the majority of which still can be considered as primary targets. Statistically significant effect on gene expression levels (adjusted $P < 0.05$) was detected for transcripts representing 638 separate Ensembl genes (604 separate RefSeq gene symbols), out of which 408 (384) were upregulated and 230 (221) downregulated (Supplementary Table S3). The broad expression range of these 638 genes is indicated in Supplementary Figure S1A. In addition to the larger number of affected genes, also the magnitude of change was higher for the upregulated set of VDR target genes (maximal upregulation 6.65-fold), compared to the downregulated set (maximal downregulation 1.48-fold). Validation of 43 selected genes on independent samples by qPCR showed a high correlation ($r = 0.889$, $P < 1.8 \times 10^{-14}$) to the microarray results across a wide range of fold changes (Supplementary FigureS1B). The validation results of five upregulated immunity-related genes *CAMP*, *CD14*, *SP100*, *LRRCA8* and *CD93* and a downregulated immunity-related gene *HSH2D* are shown in Supplementary Figure S1C.

Gene ontology (GO) enrichment analysis of the 408 upregulated VDR target genes using GeneTrail (31) highlighted the role of $1\alpha,25(OH)_2D_3$ in the regulation of immune response and signaling cascades and in the positive regulation of higher-level processes (Table 1). In particular, a positive effect of $1\alpha,25(OH)_2D_3$ on the production and secretion of the cytokine interleukin (IL) 1 was implicated. The 230 downregulated genes showed association with the negative regulation of many similar higher-level processes as the upregulated genes. In contrast to upregulated genes, GO terms enriched for downregulated genes lacked categories related to immunity and signaling, but interestingly included a category for transcription regulator activity (Table 1). Combining the up and downregulated genes for the enrichment analysis did not point to additional functions (data not shown).

In summary, we identified in THP-1 cells a set of 638 genes as primary targets of regulation by $1\alpha,25(OH)_2 D_3$, and thus by VDR, that were enriched for genes that are associated with the GO terms related to immunity and signaling.

Genome-wide mapping of VDR binding sites

To identify genomic VDR binding sites, THP-1 cells were treated for 40 min with $1\alpha,25(OH)_2D_3$ or remained unstimulated and ChIP assays were performed with antibodies against VDR and non-specific IgGs. Purified chromatin samples were subjected to deep sequencing using a Solexa GAII platform. Analysis of the resulting

Table 1. GO analysis of VDR target genes

GO ID	Subcategory	Genes		FDR
		Expected	Observed	
Upregulated $1\alpha,25(OH)_2D_3$ targets				
GO:0007242	Intracellular signaling cascade	29.9	55	0.0062
GO:0048518	Positive regulation of biological process	40.7	70	0.0062
GO:0048522	Positive regulation of cellular process	37.2	63	0.0111
GO:0023052	Signaling	80.0	113	0.0125
GO:0005829	Cytosol	26.0	47	0.0130
GO:0009966	Regulation of signal transduction	18.7	37	0.0130
GO:0023051	Regulation of signaling process	18.7	37	0.0130
GO:0050718	Positive regulation of interleukin-1 beta secretion	0.2	4	0.0130
GO:0080134	Regulation of response to stress	5.5	17	0.0130
GO:0050706	Regulation of interleukin-1 beta secretion	0.3	4	0.0151
GO:0050716	Positive regulation of interleukin-1 secretion	0.3	4	0.0151
GO:0007229	Integrin-mediated signaling pathway	1.1	7	0.0153
GO:0010646	Regulation of cell communication	21.6	40	0.0153
GO:0032731	Positive regulation of interleukin-1 beta production	0.3	4	0.0153
GO:0050704	Regulation of interleukin-1 secretion	0.3	4	0.0153
GO:0002376	Immune system process	21.8	40	0.0171
GO:0006950	Response to stress	35.2	57	0.0171
GO:0050702	Interleukin-1 beta secretion	0.3	4	0.0171
GO:0007275	Multicellular organismal development	57.9	84	0.0171
GO:0048583	Regulation of response to stimulus	9.4	22	0.0179
GO:0032732	Positive regulation of interleukin-1 production	0.3	4	0.0192
Downregulated $1\alpha,25(OH)_2D_3$ targets				
GO:0051093	Negative regulation of developmental process	2.8	14	0.0011
GO:0048519	Negative regulation of biological process	20.0	41	0.0023
GO:0048523	Negative regulation of cellular process	18.3	39	0.0023
GO:0050793	Regulation of developmental process	7.3	21	0.0040
GO:0030528	Transcription regulator activity	16.2	33	0.0132
GO:0045595	Regulation of cell differentiation	5.2	16	0.0132
GO:0045165	Cell fate commitment	1.4	8	0.0137
GO:0045596	Negative regulation of cell differentiation	2.3	10	0.0141
GO:0050794	Regulation of cellular process	74.3	100	0.0141

The GO terms enriched for either up or downregulated VDR genes were identified using GeneTrail. All sub-categories with $P < 0.02$ are listed.

sequences, using IgG-precipitated sample as a negative control, revealed the genome-wide landscape of the chromatin occupation of VDR in THP-1 cells. The reads at various stages of the analysis are summarized in Supplementary Table S4. In the $1\alpha,25(OH)_2D_3$-treated sample the number of high-confidence (FDR <1%) peaks indicating genomic VDR binding sites was 1820, in contrast to 1169 sites in the unstimulated sample (Figure 1A). The total number of VDR binding sites was 2340, with 649 (27.7%) sites common in both unstimulated and $1\alpha,25(OH)_2D_3$-treated samples and 520 (22.2%) and 1171 (50.0%) unique in unstimulated and $1\alpha,25(OH)_2D_3$-treated samples, respectively (Figure 1A). The genomic locations of the 2340 binding sites are listed in Supplementary Table S5.

VDR binding locations, grouped to unique and common as in the Venn diagram in Figure 1A, showed differences in their distribution to various genomic elements (Figure 1B). Within both the unstimulated and $1\alpha,25(OH)_2D_3$-treated peak sets, the high-confidence peaks, compared to peaks with lower confidence, were enriched to genomic regions representing proximal promoter and exons, and were depleted from distal regions and introns other than the first (Supplementary Figure S2). However, comparing the distributions of

unstimulated and $1\alpha,25(OH)_2D_3$-treated peak sets revealed that upon ligand treatment VDR favors significantly more such binding locations that are distal to genes (>1 kb upstream or >10 kb downstream) and disfavors those representing coding sequences (Figure 1B). The shift of VDR binding from proximal to distal locations upon ligand treatment was also highlighted by the VDR peak proximity to TSS (Figure 1C). The density of VDR binding proximal to the TSS diminished with decreasing peak confidence, i.e. with increasing FDR (peaks with FDR >10% may be considered random-like), for peaks that were either unique in the unstimulated state (left panel) or common in both treatments (middle panel). In contrast, such enrichment was nearly completely lacking in the $1\alpha,25(OH)_2D_3$-treated cells (right panel). Together with Figure 1B, this suggests a non-random enrichment of VDR binding locations to gene-dense regions which, at first glance, appears weaker for the $1\alpha,25(OH)_2D_3$-treated peak set but may in fact indicate enrichment to a subset of genes, such as regulated genes.

Taken together, we have identified in THP-1 cells 2340 genomic VDR binding locations, a surprisingly large proportion of which display at least some binding also in the unstimulated state and not only upon $1\alpha,25(OH)_2D_3$-treatment.

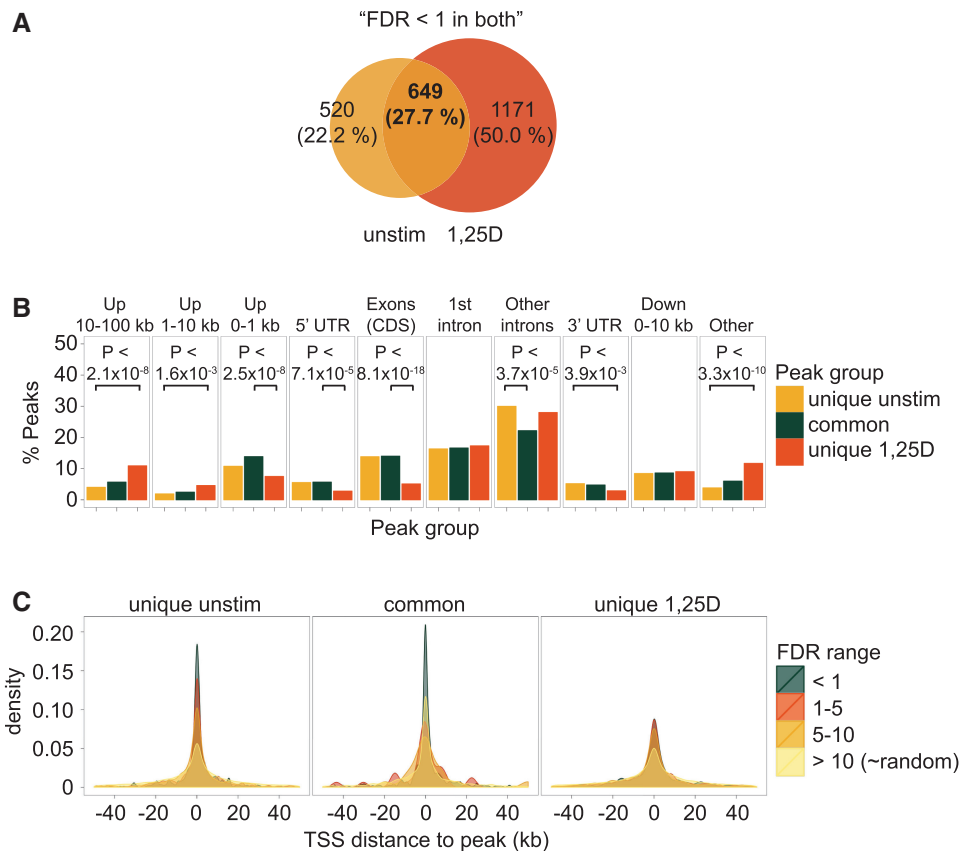**Figure 1.** VDR binding locations and genomic distribution in unstimulated and 1α,25(OH)$_2$D$_3$-treated THP-1 cells. ChIP-Seq was performed from chromatin templates extracted from THP-1 cells that were either unstimulated or treated for 40 min with 10 nM 1α,25(OH)$_2$D$_3$. (**A**) Peak counts (percentages in parenthesis) are given for locations in the unique and common categories. The 'common' category represents locations where both samples had an FDR <1% peak. Peaks were considered to be at the same location when the narrower peak overlapped the wider by >50%. (**B**) Distribution of VDR peaks into genomic elements. For clarity, only the most significant *P*-value from the binomial test is given for each element. Only *P*-values <0.005 are shown. (**C**) Distance to the TSS of the closest Ensembl transcript. The whole MACS peak set was divided into the indicated FDR ranges, unique and common peak locations were defined within each range and the densities of distances between each peak and the nearest TSS plotted by peak group and FDR range. Decrease in the density at 0 kb from high-confidence peaks (FDR <1%) to nearly random peaks (FDR >10%) indicates non-random proximity to TSS. Unstim, unstimulated.

Transcription factor binding sequences at genomic VDR locations

De novo binding site search using rGADEM and MotIV R/Bioconductor packages identified the classical DR3-type RE consensus sequence for VDR–RXR heterodimers as the most highly enriched within ±100 bp of the VDR peak summits (*E*-value 2.65×10^{-12} for the JASPAR consensus matrix MA0074.1 for RXRA::VDR). Of note, while the JASPAR consensus has near-exclusive T at positions 4, 12 and 13, more variability is allowed at these positions in the *de novo* DR3-type REs actually occupied by VDR in THP-1 cells (Figure 2A). However, only 31.7% (742) of all VDR peak summits included one or more *de novo* DR3-type RE. The quality of the identified *de novo* DR3-type REs appears high, since the more each of the *de novo* DR3-type REs resembled their common consensus, i.e. had lower *E*-value, the higher was the FE of the respective VDR peak (Pearson *r* = −0.217, *P* < 3×10^{-15}, Figure 2B). Hoping to identify any atypical VDR binding consensus sequences for the remaining 68.3% (1598) of VDR peaks, we re-ran the *de novo*

analysis on them only, but even the most frequent consensus nrGGrGswGGrrn, corresponding best to a SP1 binding site in the JASPAR database (MA0079.2, *E*-value 1.03×10^{-7}), was found in 538 peaks (23.0%). Another consensus sequence, nsAGGAAGn, was even less frequent with 274 peaks (11.7%). The latter had high similarity to the consensus for ETS transcription factor family member SPI1/PU.1 (MA0080.2; *E*-value 8.50×10^{-7}). *De novo* analysis of even more focused subgroups, like the peaks unique to unstimulated cells, did not reveal additional consensus sequences (data not shown). Using a set of DR- and ER-type VDREs where the half sites were based on the identified *de novo* DR3-type RE and which covered all spacer lengths from 0 to 16 nucleotides, VDR peak sequences lacking the *de novo* DR3-type RE were scanned using RSAT (37). This scan failed to identify any VDRE with statistically significant enrichment to ±100 bp of the peak summit versus the two 150 bp flanks of that summit region (data not shown). It should be noted that we found no genome-wide evidence that VDR binds to any other type of
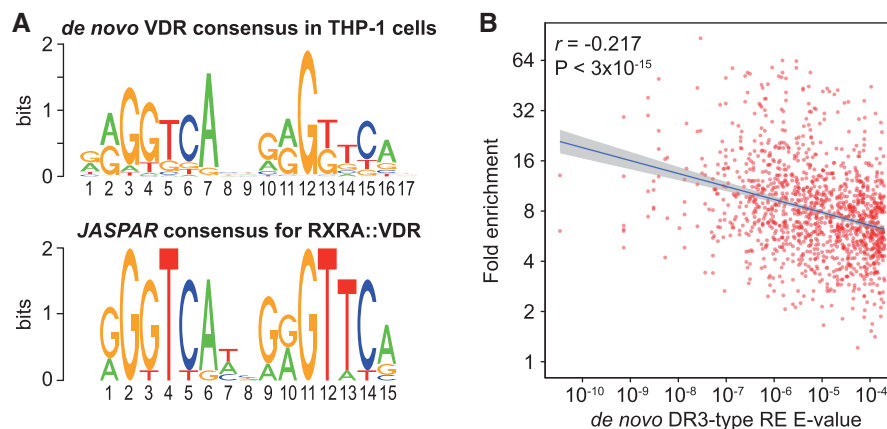
Figure 2. *De novo* identified DR3-type REs. *De novo* analysis was performed on sequences ±100 bp of the FDR <1% VDR peak summits. (**A**) The consensus DR3-type RE identified by the rGADEM *de novo* search is displayed in comparison to the public RXRA::VDR heterodimer consensus sequence in the JASPAR database. Note that the *de novo* DR3-type RE allows more variability at certain positions than the JASPAR DR3-type sequence. (**B**) Correlation between the *de novo* DR3-type RE *E*-value and the peak FE. *E*-value describes how perfectly each individual element matches the common consensus (*x-axis*). For peaks with multiple *de novo* DR3-type REs, only the element with the lowest *E*-value was selected. The blue line indicates the linear fit with 95% confidence intervals in gray, with associated Pearson correlation *r* and *P*-values given in the upper left corner of the graph.

VDRE than DR3. However, the possibility can not be excluded that a few individual non-DR3 sites exist that bind VDR.

In summary, only about a third of the genomic VDR binding sites in THP-1 cells contained a DR3-type RE and no common element was found to consistently explain the presence of VDR at locations lacking the classical VDRE.

Ligand-induced genome-wide shift in DR3-type RE usage by VDR

While the *de novo* DR3-type RE was, at first look, apparently relatively infrequent at sites of VDR chromatin occupation in THP-1 cells (Figure 2), on closer inspection its presence strikingly correlated with the $1\alpha,25(OH)_2D_3$-dependent effects on VDR binding at each location (Figure 3A). The *de novo* DR3-type RE was present in about 20% of genomic VDR locations where the VDR binding in the unstimulated sample exceeded that in the $1\alpha,25(OH)_2D_3$-treated sample by >10 FE. Conversely, when the VDR binding after a short 40 min $1\alpha,25(OH)_2D_3$ treatment superseded that of the unstimulated state, the percentage of peak locations with a DR3-type RE climbed to >90%. Furthermore, the number of *de novo* DR3-type REs near the peak summit tended to associate with increased peak FE within the common and $1\alpha,25(OH)_2D_3$-exclusive peaks, but not within the unstimulated-exclusive peaks (Figure 3B). The presence or absence of either the *de novo* SPI1/PU.1 or SP1-like consensus sequences, irrespective of coinciding *de novo* DR3-type REs, did not affect the FE of VDR peaks (data not shown).

Taken together, even without the ligand VDR occupies many chromatin sites, but most of these do not contain a DR3-type RE. In contrast, even a short ligand-treatment focuses the bulk of VDR binding to sites that do contain at least one DR3-type RE.

Proximity of VDR binding sites to $1\alpha,25(OH)_2D_3$ target genes

Combining gene expression data with the genomic locations of VDR offered an avenue to explore the mechanisms of target gene regulation by VDR. First we mapped the VDR binding locations into the neighborhoods of each gene regulated by the 4 h $1\alpha,25(OH)_2D_3$ treatment. For the 408 upregulated genes the bulk of VDR regulation seems to occur via sites within only 30 kb from the target gene TSS. However, a considerable proportion of regulation appears to be orchestrated by VDR binding sites up to 400 kb away, especially when the sites are unique to the $1\alpha,25(OH)_2D_3$ treatment (Figure 4A, top panel and Supplementary Figure S3A, top right panel). In contrast, the 230 downregulated genes had only slight enrichment for common and $1\alpha,25(OH)_2D_3$-unique peaks near the TSS (Figures 4A, Supplementary Figures S3A and S3B, bottom panels). The frequency of VDR peaks around the TSS was also higher for upregulated than for downregulated genes (2.04 ± 2.54 versus 1.05 ± 2.16 peaks per gene, Wilcoxon rank sum test *P*-value 5.01×10^{-13}). The strong enrichment of $1\alpha,25(OH)_2D_3$-treated peaks to the $1\alpha,25(OH)_2D_3$-regulated gene neighborhoods confirms regulated genes as the likeliest source to the apparent depletion of these peaks from the proximity of any TSS seen in Figure 1B and 1C. Thus, VDR seems to shift from sites in the vicinity of the TSS of non-regulated genes to the proximity of VDR-regulated genes.

Since the presence of the *de novo* DR3-type REs strongly associated with ligand-dependency of the VDR binding (Figure 3), we investigated whether this is also reflected by the VDR target gene regulation (Figure 4B). We identified for each regulated gene the strongest VDR binding location in $1\alpha,25(OH)_2D_3$-treated cells within 400 kb of the gene TSS and defined them as the most likely regulating VDR locations (Supplementary Table S6). These sites were further split to those in proximal
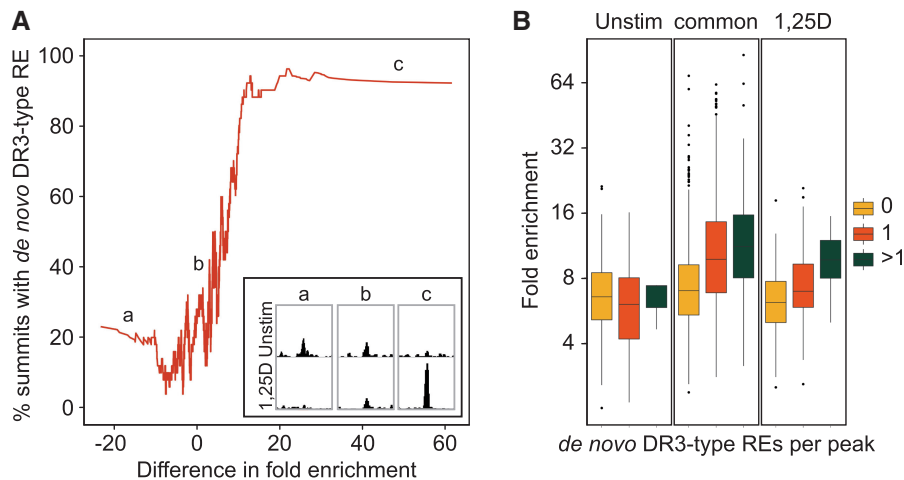
Figure 3. Ligand-dependent VDR occupation on the *de novo* DR3-type REs. The usage of the *de novo* DR3-type RE is strongly ligand dependent. (**A**) The running mean of the difference in peak FE at VDR binding locations (*x-axis*; 1α,25(OH)$_2$D$_3$ minus unstimulated, window size 50 locations) was plotted against the percentage of locations with the *de novo* DR3-type RE (*y-axis*). For unique peaks, the minimum FE of the whole 2340 peak set was used as the value for the 'missing' sample, to represent the lower detection limit. The inset shows 3 examples of peak pairs with different ligand effects on peak FE (a, decrease; b, no change; and c, increase). (**B**) The effect of the number of *de novo* DR3-type REs on the peak FE. Unstim, unstimulated.

(0 to ± 30 kb of the TSS) and distal regions (±30 to ±400 kb). If no VDR peak was present within the 400 kb limit, we used from the outlying 'desert' the closest peak of any FE to represent a potential very distant regulator. Genes were divided, separately for both up and downregulated sets, into equally sized tertiles according to their stimulated fold change [4 h 1α,25(OH)$_2$D$_3$ versus vehicle treatment]. The most likely regulating VDR locations were also divided at median FE to equally sized 'low' and 'high' VDR peak sets and further to peaks with or without any DR3-type RE near the peak summit. Gene counts in these 12 sub-categories suggest that the higher the VDR peak is, the more likely it is to strongly upregulate its target gene, especially when it contains a DR3-type RE. As already seen in Figure 4A and Supplementary Figure S3, strong regulation by VDR is possible from both proximal and distal sites, but less likely from sites beyond ± 400 kb. Also among the genes for which the most likely regulating VDRE was in the 'low' category, the proportion of DR3-type RE containing peaks versus peaks without such sites increased toward the high fold change genes. In contrast to upregulated genes, the VDR peak FE or the presence of DR3-type REs associated much less with the fold change of downregulated target genes.

Taken together, 72% of the 408 upregulated 1α,25(OH)$_2$D$_3$ target genes showed VDR binding within 400 kb of their TSS, while this applied only for 43% of the 230 downregulated genes.

Examples of VDR target genes

Examples of previously unidentified VDR target genes that were upregulated by a 4 h 1α,25(OH)$_2$D$_3$ treatment, and for which the most likely regulating VDRE could be easily proposed, are shown in Figure 5. *SP100*, a gene encoding *SP100* nuclear antigen that has been identified as a co-repressor for transcription factors, such as ETS1

(38), is characterized by a single, strong and ligand-dependent VDR binding region with a single *de novo* DR3-type RE located in the first intron at + 0.5 kb of the gene's TSS (a category 4 peak, Figure 5A).

Also for *ELL*, a gene encoding for an elongation factor of RNA polymerase II that has also been identified as a transcriptional co-factor for transcription factors, such as HIF-1 (39), the single dominant VDRE is located in the first intron at +19.3 kb of the gene's TSS, is slightly ligand-dependent although with considerable basal binding and contains a DR3-type RE (Figure 5B). The locus carries also a distal 1α,25(OH)$_2$D$_3$ target gene, *LRRC25* encoding the protein leucine rich repeat containing 25 of unknown function, for which the *ELL* VDR binding region, located 105 kb upstream of the *LRRC25* TSS, appears the most likely regulator. Therefore, for the *ELL* gene the VDR peak is of category 4, while for the *LRRC25* gene it is of category 5.

The third example is a complex locus containing the three 1α,25(OH)$_2$D$_3$ target genes thrombomodulin (*THBD*), encoding for the integral membrane protein *THBD* (also called CD141) (40), *CD93*, encoding for the C-type transmembrane receptor CD93 (41), and *GZF1*, encoding for the GDNF-inducible zinc finger protein 1 (42) (Figure 5C). However, the microarray result for *GZF1* could not be validated by qPCR. The *THBD* gene responds most effectively to 1α,25(OH)$_2$D$_3$ treatment, while the strongest ligand-dependent VDR binding region of the genomic locus is 285 kb upstream of the *THBD* TSS and contains eight distinct DR3-type REs and, i.e. it is a category 5 peak. There is also one other good genomic VDR location at 176 kb upstream of the *THBD* TSS, but this lacks a DR3-type RE. The locus has still two additional, but weaker VDR binding regions, one with a DR3-type RE and one without, at 77 and 107 kb upstream of the *THBD* TSS, respectively. Interestingly, there is also an unstimulated-unique VDR
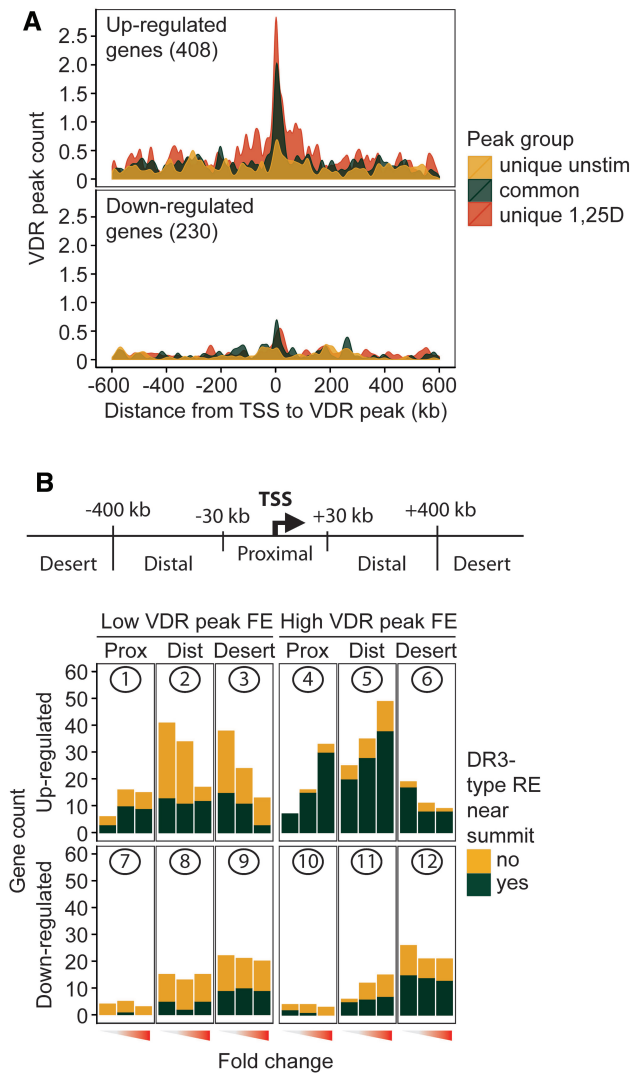
Figure 4. VDR peak height, location at <400 kb from the TSS, and the presence of *de novo* DR3-type REs predicts strong upregulation of $1\alpha,25(OH)_2D_3$ target genes. (**A**) Common and treatment-unique VDR binding sites in the neighborhoods of up- (upper panel) and downregulated genes (lower panel). (**B**) The diagram on top describes the distance categories used for selecting the most likely regulating VDR binding site as follows (distances not to scale). For each differentially expressed gene, the most likely regulating FDR <1% VDR peak was assumed to be the largest peak (by FE) that was within ±400 kb of the gene TSS. The most likely regulators were further divided into proximal and distal at ±30 kb to identify distance-related differences. If no peak was within ±400 kb, then the closest one of any FE was taken to represent a potential regulating site. See Supplementary Figure S3 for the basis of the selected distance limits. The bar graph below shows that the most likely gene regulatory VDR binding sites for up but not downregulated genes are enriched to within 400 kb of the target gene TSS and preferentially use DR3-type REs. Here, up- and downregulated genes were separately split by fold change into tertiles, indicated by increasing intensity of red in the triangle below the *x-axis*. Peaks were split into two equal sized halves by FE (left and right panels). Gene counts are plotted for up- (upper panel) and downregulated genes (lower panel). The gene counts are further split according to the presence or absence of a *de novo* DR3-type RE near the summit of the most likely regulating VDR peak, as indicated in the legend. The likely regulatory scenarios are indicated by the circled numbers 1–12, and used as peak labels in Figure 5 and Supplementary Figure S4. Unstim, unstimulated; Prox, proximal; Dist, distal.

binding site within the short, single exon *THBD* gene. It is tempting to speculate that this VDR binding site could have a repressive role that could, together with the strong ligand-dependent VDR binding site in the vicinity, facilitate the relatively strong upregulation of *THBD*.

For all three examples the VDR binding at the sites indicated with an arrow was validated from independent chromatin samples of unstimulated and 40 min $1\alpha,25(OH)_2D_3$-treated THP-1 cells by qPCR (Figure 5D). In addition to VDR, also the presence of the heterodimerizing partner RXR was verified from the same chromatin samples. In all cases the $1\alpha,25(OH)_2D_3$ treatment significantly increased the amount of both VDR and RXR at the respective VDR binding region. In addition, the upregulation by $1\alpha,25(OH)_2D_3$ of *SP100*, *ELL* and *THBD* mRNA was confirmed on independent samples (Figure 5D).

Additional examples of VDR target gene and VDR binding site scenarios are shown in Supplementary Figure S4. A strong double VDR peak with each a DR3-type VDRE was found very close to the TSS of the uncharacterized VDR target gene in chromosome 20 called open reading frame 197 (*C20orf197*), which also seems to serve as regulatory site for the VDR target genes *C20orf177* and *PPP1R3D* (Supplementary Figure S4A). The gene ArfGAP with SH3 domain, ankyrin repeat and PH domain (*ASAP2*) is regulated by three VDR binding sites, of which one is proximal and two are distal (Supplementary Figure S4B). The location of the myosin IXB (*MYO9B*) gene represents a similar complex locus as that of the *THBD* gene, as the *MYO9B* gene seems to be regulated together with the genes *OCEL1* and *ABHD8* by two proximal VDR binding sites that both contain a DR3-type RE and four proximal sites (Supplementary Figure S4C). As examples of downregulated VDR target genes the genes glucosaminyl (N-acetyl) transferase 1, core 2 (*GCNT1*) (Supplementary Figure S4D) and SUMO1 pseudogene 1 (*SUMO1P1*) (Supplementary Figure S4E) are shown having their next VDR binding site nearly 1 Mb and more than 300 kb distant from their TSS regions, respectively. Interestingly, the genomic region of *SUMO1P1* also contains the 24-hydroxylase (*CYP24A1*) gene, which is in most cellular systems a strong VDR target gene with a number of ChIP-confirmed VDR binding locations (13,43).

In summary, the eight exemplary genomic loci showed the variation of gene regulatory scenarios ranging from a single VDR location close to the $1\alpha,25(OH)_2D_3$ target gene TSS, more complex ones with one VDR location distal to the TSSs of one or more $1\alpha,25(OH)_2D_3$ target genes and very complex ones with up to six VDR locations clustering with multiple $1\alpha,25(OH)_2D_3$ target genes.

DISCUSSION

For a general understanding of the mechanisms of $1\alpha,25(OH)_2D_3$ signaling it is essential to monitor the genome-wide location of VDR in relation to primary $1\alpha,25(OH)_2D_3$ target genes. In this study, we investigated the sites of VDR chromatin occupancy in THP-1 human

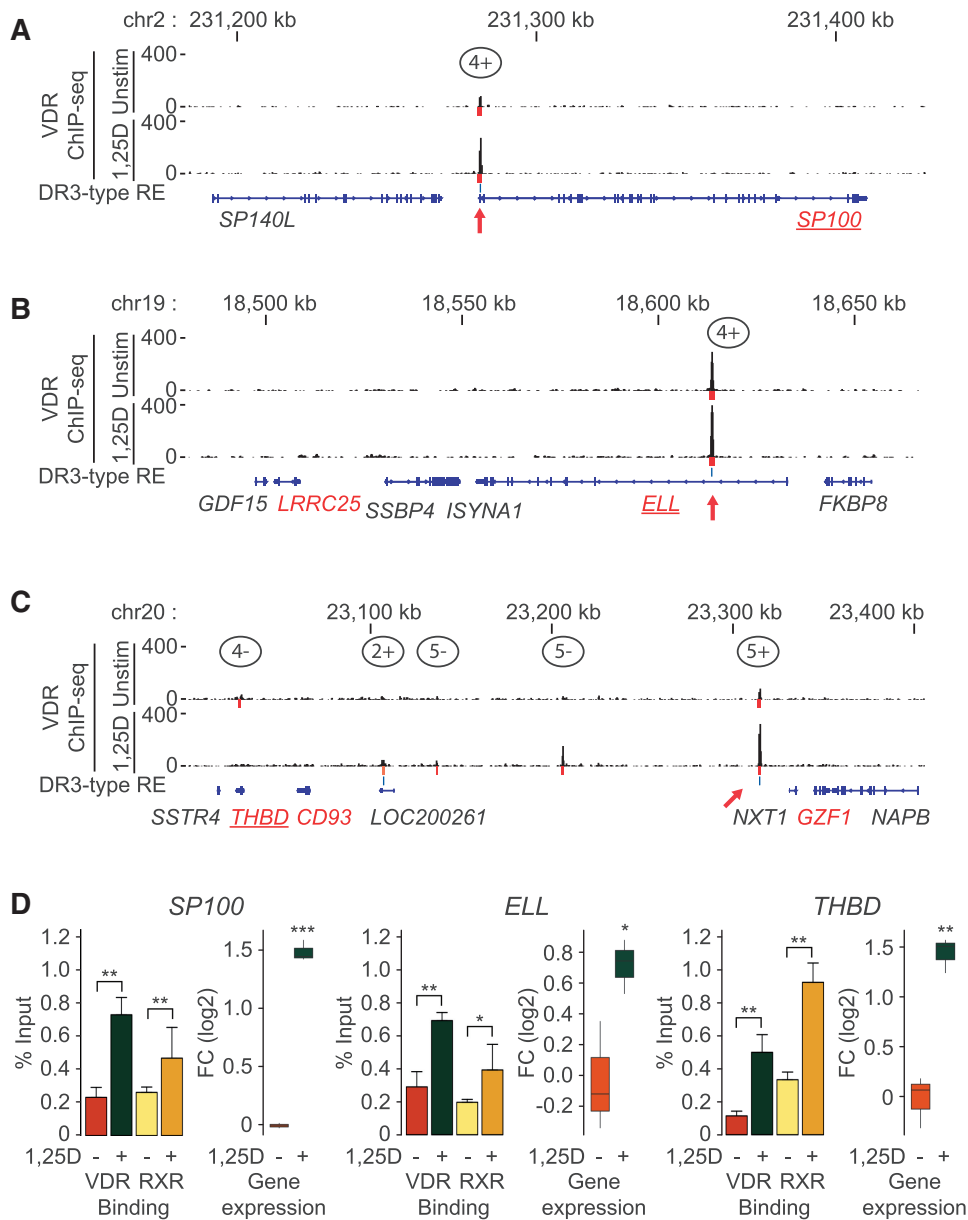**Figure 5.** VDR ChIP-seq analysis of genomic loci for the $1\alpha,25(OH)_2D_3$ target genes *SP100*, *ELL* and *THBD*. In (**A–C**), the peak tracks show the unstimulated and $1\alpha,25(OH)_2D_3$-treated VDR peaks. The high-confidence peaks are indicated as yellow-to-orange-to-red boxes, color scale signifying FE (yellow, low FE; red, high FE), below each peak track. DR3-type REs within ±100 bp of VDR peak summits are indicated by blue vertical lines below the ligand-treated VDR peak track. The gene structures are shown in blue at the bottom. Gene names in red color indicate upregulated genes (no downregulated genes were present in the loci shown). The red arrows above the peak tracks indicate the location of the manually validated VDR binding locations. The circled numbers above the VDR peaks indicate their category in reference to the underlined gene (see Figure 4B for details), plus symbol and minus symbol indicating the presence or absence of a DR3-type VDRE near the peak summit. Unstim, Unstimulated. (**D**). ChIP validation of VDR as well as RXR binding at sites indicated by the red arrows in A–C near the exemplary target genes *SP100*, *ELL* and *THBD* (bar graphs, mean ±SD), and the qPCR validation of their target gene status (box plots). The 40 min (ChIP) or 4 h (qPCR) experiments with unstimulated/vehicle (minus symbol) or $1\alpha,25(OH)_2D_3$ (plus symbol) treatments are indicated below the graphs.

monocytes after 40 min treatment with $1\alpha,25(OH)_2D_3$ and in the unstimulated state using ChIP-seq, and related these to the VDR target genes identified using microarrays from the same cellular model 4 h after ligand treatment. As expected, GO term analysis of the 638 primary $1\alpha,25(OH)_2D_3$ target genes that we identified in THP-1 cells showed a clear enrichment for immune-related genes. In particular, the positive effect of $1\alpha,25(OH)_2D_3$ on IL1 production and secretion was highlighted, which

fits well with previous reports (44,45). However, we did not find *IL1A* or *IL1B* within our 638 $1\alpha,25(OH)_2D_3$ target genes, but only the gene of the IL1 receptor antagonist (*IL1RN*) downregulated. The latter suggests that the regulation of IL1 activity by $1\alpha,25(OH)_2D_3$ is indirect via the block of the IL1 receptor. In addition, Kaler *et al.* (46) suggested that in THP-1 cells $1\alpha,25(OH)_2D_3$ downregulates the transcription factor STAT1, which is central in IL1 regulation.

Since the only nuclear target of $1\alpha,25(OH)_2D_3$ is VDR (47) and the 4 h treatment is a reasonably short time, most of the genes that we found in our microarray analysis are assumed to be primary VDR target genes, i.e. VDR should bind a regulatory chromatin region that loops to the target gene's TSS region. The majority (64%) of the $1\alpha,25(OH)_2D_3$ target genes were upregulated, which is typical for most VDR target tissues (48,49).

Through the choice of early time points both in microarray and ChIP-seq experiments we aimed to detect mainly primary VDR effects. This is in contrast to the recent study of Ramagopalan *et al.* (23), who reported VDR ChIP-seq data in HapMap lymphoblastoid cells after a long (36 h) $1\alpha,25(OH)_2D_3$ treatment, focusing more on the association of VDR binding sites with disease SNPs and aspects of human evolution.

In this study, we identified 2340 genomic sites of VDR binding, 520 of which are exclusive to the unstimulated THP-1 cells. By size, our set of 1820 $1\alpha,25(OH)_2D_3$-treated genomic VDR binding sites in THP-1 cells is comparable with the 2776 VDR binding sites in $1\alpha,25(OH)_2D_3$-treated lymphoblastoids (23), but the overlap is only 18.2%. On the level of $1\alpha,25(OH)_2D_3$ target genes identified by microarray analysis the overlap between the two cellular models is even lower at 5.6%. Although the poor overlap may be partly explained by differences in cutoffs when considering high-confidence peaks and target genes as well as the rather different duration of $1\alpha,25(OH)_2D_3$ treatment (40 min or 4 h versus 36 h), it also suggests that VDR utilizes considerably variant binding sites in different cells. *In silico* screening for VDR binding showed an approximately 100-fold higher number of putative REs than actually used (50). The main regulator for the access to these sites is the organization of chromatin, which is very cell specific. Moreover, in cells of primary origin there is a significant inter-individual variation in chromatin organization and gene expression. Therefore, it is not surprising that monocytes and lymphoblasts have a rather small overlap of VDR binding sites. This is fully in line with the diversity of VDR target gene sets between different tissues (19–22).

In the absence of ligand, VDR has been shown to actively repress its target genes, likely via a mechanism involving an interaction with co-repressor proteins (17,51). Interestingly, only 14% of the 520 genomic VDR binding locations that occur uniquely in the absence of ligand contain a DR3-type VDRE. In contrast, up to 90% of the best $1\alpha,25(OH)_2D_3$-dependent VDR binding regions contain at least one DR3-type RE. This means that upon ligand treatment the VDR occupancy is strongly shifted from non-DR3 locations to DR3-type RE locations. It is possible that the non-DR3 locations may serve as a nuclear store of VDR to be utilized rapidly upon the introduction of the ligand, partly substituting for the need to transport VDR into the nucleus from outside. This mechanism may also apply to retinoic acid and thyroid hormone receptors, but not for the classical steroid receptors, since they do not bind to DNA in the absence of ligand.

Interestingly, for 300 (73.5%) of the 408 upregulated $1\alpha,25(OH)_2D_3$ target genes we found VDR binding within 400 kb of their TSS, while this applied only for 104 (45.2%) of the 230 downregulated genes. This suggests that the mechanisms of downregulation of VDR target genes are rather complex and/or varied, and may require gene-specific investigations as demonstrated for the *CYP27B1* gene (52). In this case the repressive function of VDR is known to result from indirect interaction with the chromatin, via transcription factor 3, also known as VDR interacting repressor (53). An additional question is whether for $1\alpha,25(OH)_2D_3$ target genes there is a ligand-induced derepression mechanism as described for few other members of the nuclear receptor superfamily (54). From the 408 $1\alpha,25(OH)_2D_3$ target genes that we can explain via a VDR peak within 400 kb of their TSS only 11 (2.7%) meet the derepression criteria that they have a VDR peak in the unstimulated sample and no peak in the $1\alpha,25(OH)_2D_3$-treated sample. Additional 32 genes (7.9%) can be called dominantly derepressed, since their main peak shows only in the unstimulated sample. This indicates that for some 10% of all $1\alpha,25(OH)_2D_3$ target genes a derepression mechanism may apply.

The fact that DR3-type REs are detected by our *de novo* motif search as the most dominating sequence in the genomic VDR peak regions confirms earlier reports (9,55) and the general belief in the field that DR3-type REs are the main type of VDREs. Nevertheless, only 31.7% of the identified VDR binding regions contain at least one DR3-type RE, i.e. the majority of genomic VDR binding sites do not use this type of VDRE. Since the most dominant ligand-induced VDR binding sites are highly enriched for DR3-type REs that, moreover, preferentially locate to the neighborhoods of upregulated $1\alpha,25(OH)_2D_3$ target genes, it is likely that DR3-type REs are primarily used in gene activation. However, this leaves the question about the nature of VDR binding at those 68.3% of locations that lack a DR3-type VDRE. Although we could not identify any non-DR3-type VDRE that would explain any significant proportion of the VDR binding, it does not exclude the possibility that a low number of such sites exist. Indeed, a case of repressive gene regulation via an ER9-type VDRE has been demonstrated by Polly *et al.* (17). Other possible explanations include sites where VDR binds indirectly to genomic DNA via other transcription factors, such as SP1, although again our *de novo* motif analysis did not identify specific enrichment of other transcription factor binding sites in VDR binding locations lacking a DR3-type VDRE compared to those that had it. A probably more likely explanation for both the VDR-peak-less target genes as well as the target gene-less VDR peaks comes from the recent demonstration that gene regulation by VDR is a very dynamic process with rapid changes of VDR binding site occupancy and, consequently, on the mRNA accumulation of the $1\alpha,25(OH)_2D_3$ target genes (13,56). Therefore, the time points chosen in this study represent only snap-shots of the early actions of $1\alpha,25(OH)_2D_3$ and its receptor VDR. Due to the fast dynamics of VDR binding site occupancy and resulting fluctuations in target gene mRNA amounts, it is likely that without time-course data a

considerable proportion of all transient VDR binding sites remains unknown.

Although most of the upregulated genes may be explained by a more unified mechanism of one or multiple VDREs, as already suggested for various VDR target genes, such as *CDKN1A* (57), *CYP24A1* (13,43), *ALOX5* (16) and *CCNC* (14), we demonstrate here with three exemplary genomic regions that there is quite a variation in the gene regulatory scenarios of upregulated genes. Genome-wide there are only about 20 genes that show, as in the case of the *SP100* gene, a single VDR location close to the target gene TSS. More common are situations where either one target gene has multiple VDR binding sites in its regulatory region or, as shown for *ELL* and *LRRC25*, a pair of closely located VDR target genes share one or more VDR binding sites. The first constellation we have already demonstrated for the above mentioned VDR target genes, while the second scenario applies for the members of the *IGFBP* gene family (15). However, the most complex regulatory situations we found in the clusters around the VDR target genes *THBD* and *MYO9B*, which each contains five or six VDR binding sites of different characteristics.

In conclusion, our genome-wide characterization of early-responding VDR binding sites close to primary $1\alpha,25(OH)_2D_3$ target genes identified a clear shift of VDR binding from non-DR3 sites to DR3-type REs upon ligand treatment. Furthermore, our data suggest regulatory explanations for a large majority of especially upregulated $1\alpha,25(OH)_2D_3$ target genes in THP-1 cells.

ACCESSION NUMBERS

The gene expression microarray and sequence data have been submitted collectively to the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) under accession number GSE27438.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Carlberg,C. and Polly,P. (1998) Gene regulation by vitamin D_3. *Crit. Rev. Eukaryot. Gene Expr.*, **8**, 19–42.
2. DeLuca,H.F. (2004) Overview of general physiologic features and functions of vitamin D. *Am. J. Clin. Nutr.*, **80**, 1689S–1696S.
3. Ingraham,B.A., Bragdon,B. and Nohe,A. (2008) Molecular basis of the potential of vitamin D to prevent cancer. *Curr. Med. Res. Opin.*, **24**, 139–149.
4. Verstuyf,A., Carmeliet,G., Bouillon,R. and Mathieu,C. (2010) Vitamin D: a pleiotropic hormone. *Kidney Int.*, **78**, 140–145.
5. Bouillon,R., Carmeliet,G., Verlinden,L., van Etten,E., Verstuyf,A., Luderer,H.F., Lieben,L., Mathieu,C. and Demay,M. (2008) Vitamin D and human health: lessons from vitamin D receptor null mice. *Endocr. Rev.*, **29**, 726–776.
6. Stoffels,K., Overbergh,L., Giulietti,A., Verlinden,L., Bouillon,R. and Mathieu,C. (2006) Immune regulation of 25-hydroxyvitamin-D_3-1a-hydroxylase in human monocytes. *J. Bone Miner Res.*, **21**, 37–47.
7. Baeke,F., Etten,E.V., Overbergh,L. and Mathieu,C. (2007) Vitamin D3 and the immune system: maintaining the balance in health and disease. *Nutr. Res. Rev.*, **20**, 106–118.
8. Gombart,A.F., Borregaard,N. and Koeffler,H.P. (2005) Human cathelicidin antimicrobial peptide (CAMP) gene is a direct target of the vitamin D receptor and is strongly up-regulated in myeloid cells by 1,25-dihydroxyvitamin D_3. *FASEB J.*, **19**, 1067–1077.
9. Carlberg,C., Bendik,I., Wyss,A., Meier,E., Sturzenbecker,L.J., Grippo,J.F. and Hunziker,W. (1993) Two nuclear signalling pathways for vitamin D. *Nature*, **361**, 657–660.
10. Quack,M., Frank,C. and Carlberg,C. (2002) Differential nuclear receptor signalling from DR4-type response elements. *J. Cell. Biochem.*, **86**, 601–612.
11. Schräder,M., Nayeri,S., Kahlen,J.P., Müller,K.M. and Carlberg,C. (1995) 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.
12. Tavera-Mendoza,L., Wang,T.T., Lallemant,B., Zhang,R., Nagai,Y., Bourdeau,V., Ramirez-Calderon,M., Desbarats,J., Mader,S. and White,J.H. (2006) Convergence of vitamin D and retinoic acid signalling at a common hormone response element. *EMBO Rep.*, **7**, 180–185.
13. Väisänen,S., Dunlop,T.W., Sinkkonen,L., Frank,C. and Carlberg,C. (2005) Spatio-temporal activation of chromatin on the human CYP24 gene promoter in the presence of 1a,25-dihydroxyvitamin D_3. *J. Mol. Biol.*, **350**, 65–77.
14. Sinkkonen,L., Malinen,M., Saavalainen,K., Väisänen,S. and Carlberg,C. (2005) Regulation of the human cyclin C gene via multiple vitamin D_3-responsive regions in its promoter. *Nucleic Acids Res.*, **33**, 2440–2451.
15. Matilainen,M., Malinen,M., Saavalainen,K. and Carlberg,C. (2005) Regulation of multiple insulin-like growth factor binding protein genes by 1a,25-dihydroxyvitamin D_3. *Nucleic Acids Res.*, **33**, 5521–5532.
16. Seuter,S., Väisänen,S., Radmark,O., Carlberg,C. and Steinhilber,D. (2007) Functional characterization of vitamin D responding regions in the human 5-lipoxygenase gene. *Biochim Biophys Acta*, **1771**, 864–872.
17. Polly,P., Herdick,M., Moehren,U., Baniahmad,A., Heinzel,T. and Carlberg,C. (2000) VDR-Alien: a novel, DNA-selective vitamin D_3 receptor-corepressor partnership. *FASEB J.*, **14**, 1455–1463.
18. Carlberg,C. and Molnar,F. (2006) Detailed molecular understanding of agonistic and antagonistic vitamin D receptor ligands. *Curr. Top. Med. Chem.*, **6**, 1243–1253.
19. Palmer,H.G., Sanchez-Carbayo,M., Ordonez-Moran,P., Larriba,M.J., Cordon-Cardo,C. and Munoz,A. (2003) Genetic signatures of differentiation induced by 1a,25-dihydroxyvitamin D_3 in human colon cancer cells. *Cancer Res.*, **63**, 7799–7806.
20. Wang,T.T., Tavera-Mendoza,L.E., Laperriere,D., Libby,E., MacLeod,N.B., Nagai,Y., Bourdeau,V., Konstorum,A., Lallemant,B., Zhang,R. et al. (2005) Large-scale in silico and microarray-based identification of direct 1,25-dihydroxyvitamin D_3 target genes. *Mol. Endocrinol.*, **19**, 2685–2695.

21. Suzuki,T., Tazoe,H., Taguchi,K., Koyama,Y., Ichikawa,H., Hayakawa,S., Munakata,H. and Isemura,M. (2006) DNA microarray analysis of changes in gene expression induced by 1,25-dihydroxyvitamin D_3 in human promyelocytic leukemia HL-60 cells. *Biomed. Res.*, **27**, 99–109.

22. White,S.L., Belov,L., Barber,N., Hodgkin,P.D. and Christopherson,R.I. (2005) Immunophenotypic changes induced on human HL60 leukaemia cells by 1a,25-dihydroxyvitamin D_3 and 12-O-tetradecanoyl phorbol-13-acetate. *Leuk. Res.*, **29**, 1141–1151.

23. Ramagopalan,S.V., Heger,A., Berlanga,A.J., Maugeri,N.J., Lincoln,M.R., Burrell,A., Handunnetthi,L., Handel,A.E., Disanto,G., Orton,S.M. *et al.* (2010) A ChIP-seq defined genome-wide map of vitamin D receptor binding: associations with disease and evolution. *Genome Res.*, **20**, 1352–1360.

24. Matilainen,J.M., Husso,T., Toropainen,S., Seuter,S., Turunen,M.P., Gynther,P., Yla-Herttuala,S., Carlberg,C. and Vaisanen,S. (2010) Primary effect of 1a,25(OH)$_2$D$_3$ on IL-10 expression in monocytes is short-term downregulation. *Biochim. Biophys. Acta*, **1803**, 1276–1286.

25. Gynther,P., Toropainen,S., Matilainen,J.M., Seuter,S., Carlberg,C. and Vaisanen,S. (2011) Mechanism of 1a,25-dihydroxyvitamin D_3-dependent repression of interleukin-12B. *Biochim. Biophys. Acta*, **1813**, 810–818.

26. van Antwerp,D.J., Martin,S.J., Kafri,T., Green,D.R. and Verma,I.M. (1996) Suppression of TNF-a-induced apoptosis by NF-kB. *Science*, **274**, 787–789.

27. Gentleman,R.C., Carey,V.J., Bates,D.M., Bolstad,B., Dettling,M., Dudoit,S., Ellis,B., Gautier,L., Ge,Y., Gentry,J. *et al.* (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.*, **5**, R80.

28. Du,P., Kibbe,W.A. and Lin,S.M. (2008) lumi: a pipeline for processing Illumina microarray. *Bioinformatics*, **24**, 1547–1548.

29. Smyth,G.K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet Mol. Biol.*, **3**, Article3.

30. Klipper-Aurbach,Y., Wasserman,M., Braunspiegel-Weintrob,N., Borstein,D., Peleg,S., Assa,S., Karp,M., Benjamini,Y., Hochberg,Y. and Laron,Z. (1995) Mathematical formulae for the prediction of the residual beta cell function during the first two years of disease in children and adolescents with insulin-dependent diabetes mellitus. *Med. Hypotheses*, **45**, 486–490.

31. Keller,A., Backes,C., Al-Awadhi,M., Gerasch,A., Kuntzer,J., Kohlbacher,O., Kaufmann,M. and Lenhof,H.P. (2008) GeneTrailExpress: a web-based pipeline for the statistical evaluation of microarray experiments. *BMC Bioinformatics*, **9**, 552.

32. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.

33. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G. and Durbin,R. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

34. Robinson,J.T., Thorvaldsdottir,H., Winckler,W., Guttman,M., Lander,E.S., Getz,G. and Mesirov,J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.*, **29**, 24–26.

35. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.

36. Salmon-Divon,M., Dvinge,H., Tammoja,K. and Bertone,P. (2010) PeakAnalyzer: genome-wide annotation of chromatin binding and modification loci. *BMC Bioinformatics*, **11**, 415.

37. Thomas-Chollier,M., Sand,O., Turatsinze,J.V., Janky,R., Defrance,M., Vervisch,E., Brohee,S. and van Helden,J. (2008) RSAT: regulatory sequence analysis tools. *Nucleic Acids Res.*, **36**, W119–W127.

38. Yordy,J.S., Moussa,O., Pei,H., Chaussabel,D., Li,R. and Watson,D.K. (2005) SP100 inhibits ETS1 activity in primary endothelial cells. *Oncogene*, **24**, 916–931.

39. Liu,L., Ai,J., Xiao,W., Liu,J., Wang,Y., Xin,D., He,Z., Guo,Y. and Wang,Z. (2010) ELL is an HIF-1a partner that regulates and responds to hypoxia response in PC3 cells. *Prostate*, **70**, 797–805.

40. Koutsi,A., Papapanagiotou,A. and Papavassiliou,A.G. (2008) Thrombomodulin: from haemostasis to inflammation and tumourigenesis. *Int. J. Biochem. Cell. Biol.*, **40**, 1669–1673.

41. Greenlee,M.C., Sullivan,S.A. and Bohlson,S.S. (2008) CD93 and related family members: their role in innate immunity. *Curr. Drug Targets*, **9**, 130–138.

42. Fukuda,N., Ichihara,M., Morinaga,T., Kawai,K., Hayashi,H., Murakumo,Y., Matsuo,S. and Takahashi,M. (2003) Identification of a novel glial cell line-derived neurotrophic factor-inducible gene required for renal branching morphogenesis. *J. Biol. Chem.*, **278**, 50386–50392.

43. Meyer,M.B., Goetsch,P.D. and Pike,J.W. (2010) A downstream intergenic cluster of regulatory enhancers contributes to the induction of CYP24A1 Expression by 1a,25-dihydroxyvitamin D_3. *J. Biol. Chem.*, **285**, 15599–15610.

44. Kong,J., Grando,S.A. and Li,Y.C. (2006) Regulation of IL-1 family cytokines IL-1a, IL-1 receptor antagonist, and IL-18 by 1,25-dihydroxyvitamin D_3 in primary keratinocytes. *J. Immunol.*, **176**, 3780–3787.

45. Liu,P.T., Schenk,M., Walker,V.P., Dempsey,P.W., Kanchanapoomi,M., Wheelwright,M., Vazirnia,A., Zhang,X., Steinmeyer,A., Zügel,U. *et al.* (2009) Convergence of IL-1b and VDR activation pathways in human TLR2/1-induced antimicrobial responses. *PLoS One*, **4**, e5810.

46. Kaler,P., Augenlicht,L. and Klampfer,L. (2009) Macrophage-derived IL-1b stimulates Wnt signaling and growth of colon cancer cells: a crosstalk interrupted by vitamin D_3. *Oncogene*, **28**, 3892–3902.

47. Haussler,M.R., Haussler,C.A., Jurutka,P.W., Thompson,P.D., Hsieh,J.C., Remus,L.S., Selznick,S.H. and Whitfield,G.K. (1997) The vitamin D hormone and its nuclear receptor: molecular actions and disease states. *J. Endocrinol.*, **154(Suppl.)**, S57–S73.

48. Pike,J.W., Zella,L.A., Meyer,M.B., Fretz,J.A. and Kim,S. (2007) Molecular actions of 1,25-dihydroxyvitamin D_3 on genes involved in calcium homeostasis. *J. Bone Miner Res.*, **22(Suppl. 2)**, V16–V19.

49. White,J.H. (2004) Profiling 1,25-dihydroxyvitamin D_3-regulated gene expression by microarray analysis. *J. Steroid Biochem. Mol. Biol.*, **89–90**, 239–244.

50. Carlberg,C., Dunlop,T.W., Saramäki,A., Sinkkonen,L., Matilainen,M. and Väisänen,S. (2007) Controlling the chromatin organization of vitamin D target genes by multiple vitamin D receptor binding sites. *J. Steroid. Biochem. Mol. Biol.*, **103**, 338–343.

51. Sanchez-Martinez,R., Zambrano,A., Castillo,A.I. and Aranda,A. (2008) Vitamin D-dependent recruitment of corepressors to vitamin D/retinoid X receptor heterodimers. *Mol. Cell. Biol.*, **28**, 3817–3829.

52. Turunen,M.M., Dunlop,T.W., Carlberg,C. and Väisänen,S. (2007) Selective use of multiple vitamin D response elements underlies the 1a,25-dihydroxyvitamin D_3-mediated negative regulation of the human CYP27B1 gene. *Nucleic Acids Res.*, **35**, 2734–2747.

53. Murayama,A., Kim,M.S., Yanagisawa,J., Takeyama,K. and Kato,S. (2004) Transrepression by a liganded nuclear receptor via a bHLH activator through co-regulator switching. *EMBO J.*, **23**, 1598–1608.

54. Baniahmad,A., Ha,I., Reinberg,D., Tsai,S., Tsai,M.-J. and O'Malley,B.W. (1993) Interaction of human thyroid hormone receptor b with transcription factor TFIIB may mediate target gene derepression and activation by thyroid hormone. *Proc. Natl Acad. Sci. USA*, **90**, 8832–8836.

55. Umesono,K., Murakami,K.K., Thompson,C.C. and Evans,R.M. (1991) Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D_3 receptors. *Cell*, **65**, 1255–1266.

56. Saramäki,A., Diermeier,S., Kellner,R., Laitinen,H., Väisänen,S. and Carlberg,C. (2009) Cyclical chromatin looping and transcription factor association on the regulatory regions of the p21 (CDKN1A) gene in response to 1a,25-dihydroxyvitamin D_3. *J. Biol. Chem.*, **284**, 8073–8082.

57. Saramäki,A., Banwell,C.M., Campbell,M.J. and Carlberg,C. (2006) Regulation of the human p21[(waf1/cip1)] gene promoter via multiple binding sites for p53 and the vitamin D_3 receptor. *Nucleic Acids Res.*, **34**, 543–554.