

PROCEEDINGS OF THE
5TH INTERNATIONAL SYMPOSIUM ON
VITAMIN D AND ANALOGS IN CANCER
PREVENTION AND THERAPY

*2-3 May, 2014
Krefeld, Germany*

Local Organizing Committee

J. Reichrath¹

M. Friedrich²

Th. Vogt¹

*¹Department of Dermatology, Venerology and Allergology,
The Saarland University Hospital, Homburg/Saar, Germany;*

²Department of Obstetrics and Gynecology, HELIOS Hospital, Krefeld, Germany

Review

What Do We Learn from the Genome-wide Perspective on Vitamin D₃?

CARSTEN CARLBERG

School of Medicine, Institute of Biomedicine, University of Eastern Finland, Kuopio, Finland

Abstract. *Vitamin D₃ insufficiency is associated with a number of diseases, such as cancer and autoimmune disorders. This important medical problem leads to the question, whether an insight into the genome-wide actions of the transcription factor vitamin D receptor (VDR) and its high affinity ligand 1 α ,25-dihydroxyvitamin D₃ (1,25(OH)₂D₃) can help in a more global appreciation of the physiological impact of vitamin D₃. Chromatin immunoprecipitation sequencing (ChIP-seq) studies in 6 human cell culture models demonstrated 1,000 to 10,000 genomic VDR binding sites per*

cell type that sum-up to more than 23,000 non-overlapping loci of the receptor. After ligand stimulation VDR associates with many new binding loci, of which the most important have a higher rate of DR3-type VDR binding sequences than average sites. On the majority of latter VDR interacts directly or indirectly with genomic DNA in a presently uncharacterized fashion. Formaldehyde-assisted isolation of regulatory elements sequencing (FAIRE-seq) monitors the dynamically opening chromatin regions after 1,25(OH)₂D₃ stimulation. The integration of ChIP-seq and FAIRE-seq data combined with a screening for DR3-type sequences facilitates the identification of key VDR binding sites and primary 1,25(OH)₂D₃ target genes. Recent results of the FANTOM5 project strongly suggest a shift from in vitro cell culture experiments to primary human cells stimulated in vivo. First results suggest that both the number of genome-wide VDR binding sites and the expression of VDR target genes correlate with vitamin D status of the studied human individuals. In conclusion, a genome-wide overview provides a broader basis for addressing vitamin D's role in health and disease.

Abbreviations: 1,25(OH)₂D₃: 1 α ,25-dihydroxyvitamin D₃; 25(OH)D₃: 25-hydroxyvitamin D₃; ASAP2: ArfGAP with SH3 domain, ankyrin repeat and PH domain 2; CBS: cystathionine β -synthase; CCNC: cyclin C; CD14: CD14 molecule; CDKN1A: cyclin-dependent kinase inhibitor 1A; ChIA-PET: chromatin interaction analysis by paired-end tag sequencing; ChIP: chromatin immunoprecipitation; ChIP-seq: ChIP sequencing; CTCF: CCCTC-binding factor; CYP: cytochrome P450; DR3: direct repeat spaced by 3 nucleotides; DUSP10: dual specificity phosphatase 10; FAIRE-seq: formaldehyde-assisted isolation of regulatory elements sequencing; FANTOM: functional annotation of the mammalian genome HOMER: hypergeometric optimization of motif enrichment IGV: Integrative Genomics Viewer; LPS: lipopolysaccharide; LRP5: low density lipoprotein receptor-related protein 5; MACS: model-based analysis of ChIP-seq data; MYC: v-myc avian myelocytomatosis viral oncogene homolog; NRIP1: nuclear receptor interacting protein 1; PBMC: peripheral blood mononuclear cell; RXR: retinoid X receptor; TBP: TATA box binding protein; TNFSF11: tumor necrosis factor (ligand) superfamily, member 11; TRPV6: transient receptor potential cation channel, subfamily V, member 6; TSS: transcription start site; VDR: vitamin D receptor.

Correspondence to: Prof. Carsten Carlberg, School of Medicine, Institute of Biomedicine, University of Eastern Finland, POB 1627, FIN-70210 Kuopio. Tel: +358 403553062, e-mail: carsten.carlberg@uef.fi

Key Words: Vitamin D insufficiency, genome-wide research, vitamin D receptor, 1,25(OH)₂D₃, ChIP-seq, review

Vitamin D₃ is a pleiotropic signaling molecule, which is involved in the regulation of a large number of physiological processes, such as bone formation, immune function and cellular growth and differentiation (6, 8, 26, 54). Via binding to and activation of the transcription factor VDR the biologically active form of vitamin D₃, 1,25(OH)₂D₃, has a direct effect on the transcriptional regulation of the genome. Since VDR is the only protein that binds 1,25(OH)₂D₃ with high-affinity (17), the genomic effects of vitamin D are a subset of those exerted by its nuclear receptor. VDR is one of some 1,900 human transcription factors (55), but within these it plays a special role, since it belongs to the few dozens proteins that are specifically activated by lipophilic molecules directly reaching the cell nucleus (4). Therefore, the signal transduction by vitamin D₃/1,25(OH)₂D₃ is in contrast to that of hydrophilic signaling molecules, such as peptide hormones, growth factors and cytokines, a rather straightforward process. In addition, VDR is expressed in

Table I. Number of VDR ChIP-seq peaks in human cellular models. The raw data of six publically available VDR ChIP-seq experiments had been re-analyzed under identical settings (49). The resulting number of significant VDR peaks is reported in bold, while the originally reported peaks is indicated in brackets. ND, Not determined.

Publication	Year	Human cell Type	Name	Unstimulated VDR peaks	Ligand-stimulated VDR peaks
Ramagopalan <i>et al.</i> (36)	2010	B cells	GM10855	3,144 (623)	6,172 (2,776)
			GM10861	4,072 (623)	12,448 (2,776)
Heikkinen <i>et al.</i> (19)	2011	Monocytes	THP-1	613 (1,169)	774 (1,820)
Meyer <i>et al.</i> (29)	2012	Colon carcinoma	LS180	165 (262)	3,777 (2,209)
Ding <i>et al.</i> (9)	2013	Hepatic stellate	LX2	1,474 (ND)	1,532 (6,281)
Tuoresmäki <i>et al.</i> (49)	2014	Macrophages	LPS-THP-1	529	955

most human tissues and cell types (57). This suggests that understanding of the actions of VDR on a genome-wide level will lead to a comprehensive insight on the genomic effects of vitamin D in health and disease.

This review discusses how the understanding of vitamin D signaling has changed and improved during the last years based on the knowledge of the genome-wide locations of VDR in a number of key human cell types.

Monitoring Genomic VDR Loci

ChIP is a method that is used since more than 10 years (33), in order to monitor the genomic location of transcription factors. The core of the technique is a physical cross-linking between nuclear proteins and genomic DNA and a sonication of chromatin into small fragments in the size of 200-400 bp. A precipitation step with an antibody specific for a nuclear protein of choice (for example against VDR) and a reverse cross-linking allows the enrichment for those regions of the genome that, at the time of the crosslinking, had been in complex with the chosen nuclear protein (27).

The specific enrichment of a chosen genomic region in reference to a negative control region can be monitored by quantitative PCR (ChIP-qPCR). This approach had been successfully used in the past for the study of the regulatory regions of known primary VDR target genes, such as *CYP24A1* (52), *CYP27B1* (50), *CCNC* (46) and *CDKN1A* (40, 41). For a short while also the use of tiled microarrays, so-called “chips”, was popular for the identification of enriched chromatin fragments (ChIP-chip) and allowed for identification of VDR binding sites of the mouse genes *Vdr* (60), *Trpv6* (30), *Lrp5* (12), *Tnfrsf11* (also known as *Rankl*) (24), *Cyp24a1* (28) and *Cbs* (25). However, PCR and microarrays are based on nucleic acid hybridization, which could cause a bias for certain genomic regions, since the efficiency of these methods varies with the investigated sequence.

This possible bias does not apply to the so-called “next generation sequencing” methods, such as ChIP-seq. These techniques often provide only small stretches of 35-50

nucleotides, so-called “sequence tags”, of the investigated genomic regions. However, the length of these tags is sufficient to locate them back to their origin in the genome of the species, in which the experiment had been originally performed. The accumulation of sequence tags to a genomic region looks like a peak, when it is visualized *via* a genome browser, such as the Integrative Genome Viewer (IGV). So-called “peak calling software”, such as MACS, is then applied to identify those regions within the genome, where the peaks are significantly above the signal of a control, such as the input of the ChIP reaction (Figure 1). The specific peaks mark in this way the genomic loci of the investigated nuclear protein (14, 35). It should be noted that ChIP-seq peaks represent in most cases the average of a signal from millions of cells. Therefore, the ChIP signal of an individual cell can differ from the often heterogeneous average.

VDR belongs to those nuclear receptors, for which the complex formation with genomic DNA depends on the absence or presence of their specific ligand. Therefore, the VDR ChIP-seq signal from 1,25(OH)₂D₃-treated cells differs from those that were not stimulated. Figure 1 illustrates three scenarios, where 1,25(OH)₂D₃ stimulation either leads to i) a strong induction of VDR binding, such as illustrated for the locus of the known VDR target gene *ASAP2* (45) (Figure 1A), ii) no significant differences, such as those demonstrated for the transcription start site (TSS) of the *TBP* gene (Figure 1B) or iii) a down-regulation of VDR binding, as shown for the VDR site of the *MYC* gene (Figure 1C). The latter example demonstrates that VDR also acts in the absence of its ligand and, in fact, the *MYC* gene is down-regulated by 1,25(OH)₂D₃ (48).

At present (August 2014) VDR ChIP-seq data are publically available for 6 human cellular models. These are the immortalized lymphoblastoid cell lines GM10855 and GM10861 (36), THP-1 monocyte-like cells (19), lipopolysaccharide (LPS)-polarized THP-1 macrophage-like cells (49), LS180 colorectal cancer cells (29) and LX2 hepatic stellate cells (9). For a direct comparison the raw datasets were re-analyzed under identical settings (49)

(Table I). The original publications report between 1,820 and 6,281 VDR binding sites in stimulated samples and 262 to 1,161 VDR peaks in unstimulated cells, *i.e.* they agree on that ligand activation of VDR increases the number of its genomic binding sites. The harmonized re-analysis of the datasets came to the same general conclusion, but reported for the B cells higher and for the other cellular models lower VDR peak numbers ranging from 774 to 12,448 for stimulated cells and from 165 to 4,072 for unstimulated samples.

The overlap between the VDR peaks sets of ligand-treated and untreated cells varied between 31% and 77% (49). This suggests that a reasonable number of VDR loci are occupied in the presence and absence of ligand (Figure 1B and C). In total 23,409 non-overlapping VDR binding sites were identified when allowing a distance of 100 bp between the peak summits (49). Interestingly, some 70% of these VDR loci are unique for one of the six cellular models. This indicates that the genomic binding of VDR is largely cell-specific. In contrast, within related cellular models, such as between B cells or between monocytes and macrophages, 53 to 73% of the VDR peaks were the same (49). However, within all six VDR ChIP-seq datasets only 43 loci were identical. This number increases to 60, when larger overlap distances were allowed (49). Nevertheless, only at a very limited number of genomic loci VDR is found in all cellular systems.

In summary, ChIP-seq studies demonstrated that there are 1,000 to 10,000 genomic VDR binding loci per human cell type. Most of these sites are cell-specific, but the closer cells are developmentally related, the higher is their overlap in VDR binding. Ligand stimulation changes the genome-wide VDR binding profile: the receptor disappears from some sites, stays on approximately half of all loci, but in particular finds new binding locations.

Mechanistic Implications of Genome-wide VDR Binding

The genome-wide view on VDR binding suggests some reconsideration of the models of vitamin D signaling. In the past only a few dozens VDR binding sites were known, which were located in rather close distance to the TSS regions of physiologically relevant 1,25(OH)₂D₃ target genes (18). In contrast, today one is facing thousands of VDR binding sites in the vicinity of genes, out of which most had not been related to vitamin D before. This may suggest that far greater number of genes are vitamin D targets than previously assumed. In fact, transcriptome-wide studies suggest that in every cellular system hundreds of genes are up- or down-regulated, when stimulated with 1,25(OH)₂D₃ (1). For example, in THP-1 cells 408 genes were found to be statistically significantly up-regulated after 4-h treatment

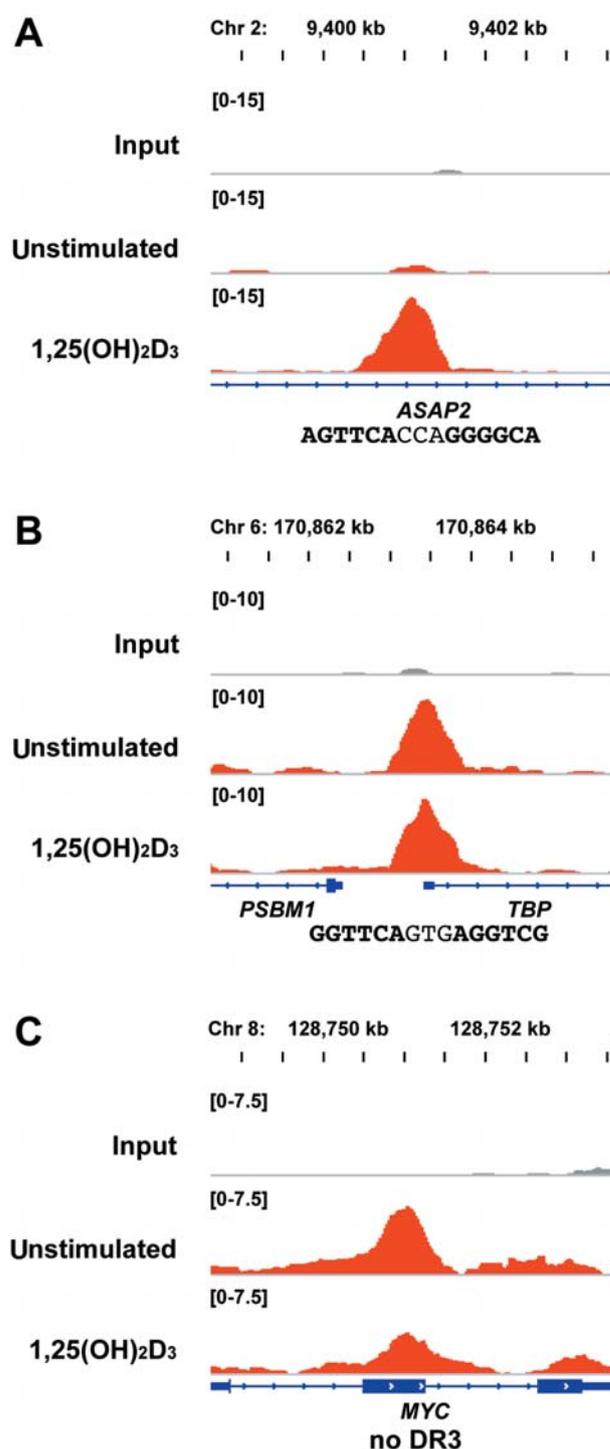


Figure 1. Different types of VDR binding scenarios. The Integrative Genomics Viewer (IGV) browser (38) was used to visualize a ligand-inducible VDR binding site within the ASAP2 gene (A, 54 kb downstream of the TSS), a ligand-independent VDR site at the TSS of the TBP gene (B) and a ligand-repressed VDR locus within the MYC gene (C, 2.7 kb downstream of the TSS). The peak tracks display data from input controls (grey) or VDR ChIP-seq datasets (red) from THP-1 human monocytic leukemia cells (19). The gene structures are shown in blue and the sequence of the DR3-type VDR binding motifs is indicated below the respective peaks.

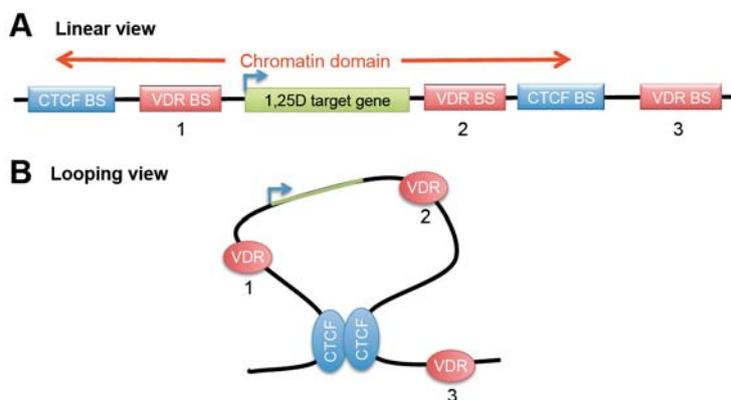


Figure 2. Chromatin domain containing VDR binding sites. The linear view (A) and the looping view (B) of a hypothetical 1,25(OH)₂D₃ target gene is shown. The chromatin domain of the gene is defined by upstream and downstream CTCF binding sites (BS). VDR BS 1 and 2 are within the chromatin domain and therefore able to regulate the gene, while VDR BS 3 is located outside of the loop and therefore not implicated in the regulation.

with 1,25(OH)₂D₃ (19). However, only 67 of these genes were more than 1.5-fold induced. Interestingly, only a few of the latter genes are close to a conserved VDR binding site, *i.e.* these sites may have a more general function than controlling genes in their vicinity.

In this context it is important to define the term “vicinity” more precisely. Chromatin is organized into loops of in average a few hundred kb in size (23). These loops are mostly separated by insulator regions (53), which are often characterized by the binding of the highly conserved transcription factor CCCTC-binding factor (CTCF) (42). Genes of a given chromatin loop, which is often also referred as chromatin domain, are preferentially regulated by transcription factors of the same loop (Figure 2). This means that a primary 1,25(OH)₂D₃ target gene should have a VDR binding site within the borders of CTCF binding loci. The recently developed method chromatin interaction analysis by paired end-tag sequencing (ChIA-PET) (13) is able to monitor genome-wide CTCF-mediated chromatin loops. Since the genomic binding of CTCF is exceptionally highly conserved, ChIA-PET data, which the ENCODE project delivered for K562 monocytic leukemia cells and MCF-7 breast cancer cells (10), allows a good estimation of chromatin domain borders (Figure 2). By extrapolating K562 CTCF ChIA-PET data to THP-1 cells, some 1,600 chromatin domains were defined that contain at least one VDR binding site (43). The average size of these loops was some 200 kb, but some chromatin domains are larger than 1 Mb. Similar numbers can be assumed for other 1,25(OH)₂D₃ responding tissues and cell types.

Since more than 20 years VDR is known to form heterodimeric complexes with the retinoid X receptor (RXR) on sequences that are direct repeats of hexameric motifs with three intervening nucleotides (DR3) (2, 51). By performing so-called “*de novo* motif searches” all VDR ChIP-seq studies

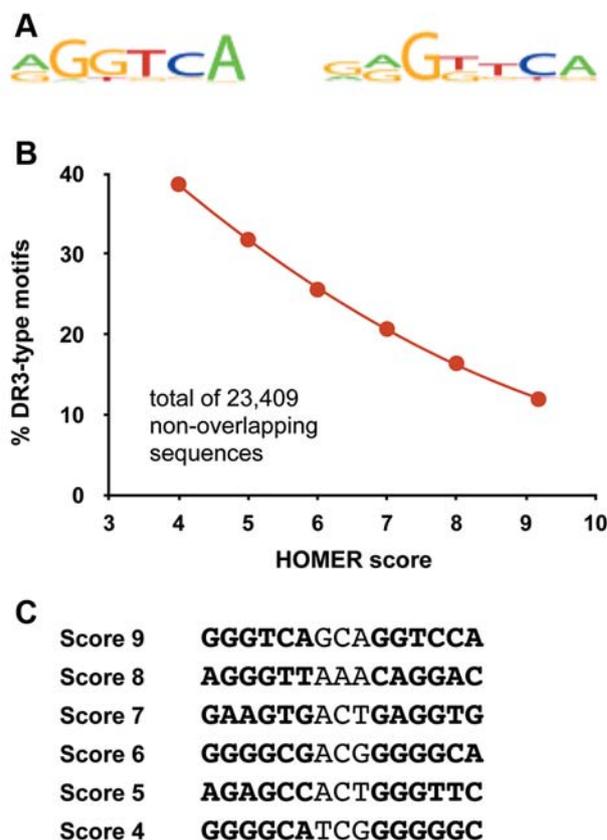


Figure 3. DR3-type VDR binding motifs. All ChIP-seq studies identified by *de novo* motif screening DR3-type binding sites as the predominant sequence motif below the summits of VDR peaks (A). The transcription factor motif screening software HOMER was used in different settings (referred to as “scores”), in order to determine the percentage of DR3-type motifs in the total set of 23,409 non-overlapping VDR binding sites (B). Random examples of DR3-type motifs demonstrate that lower HOMER scores represent more degenerate sequences (C).

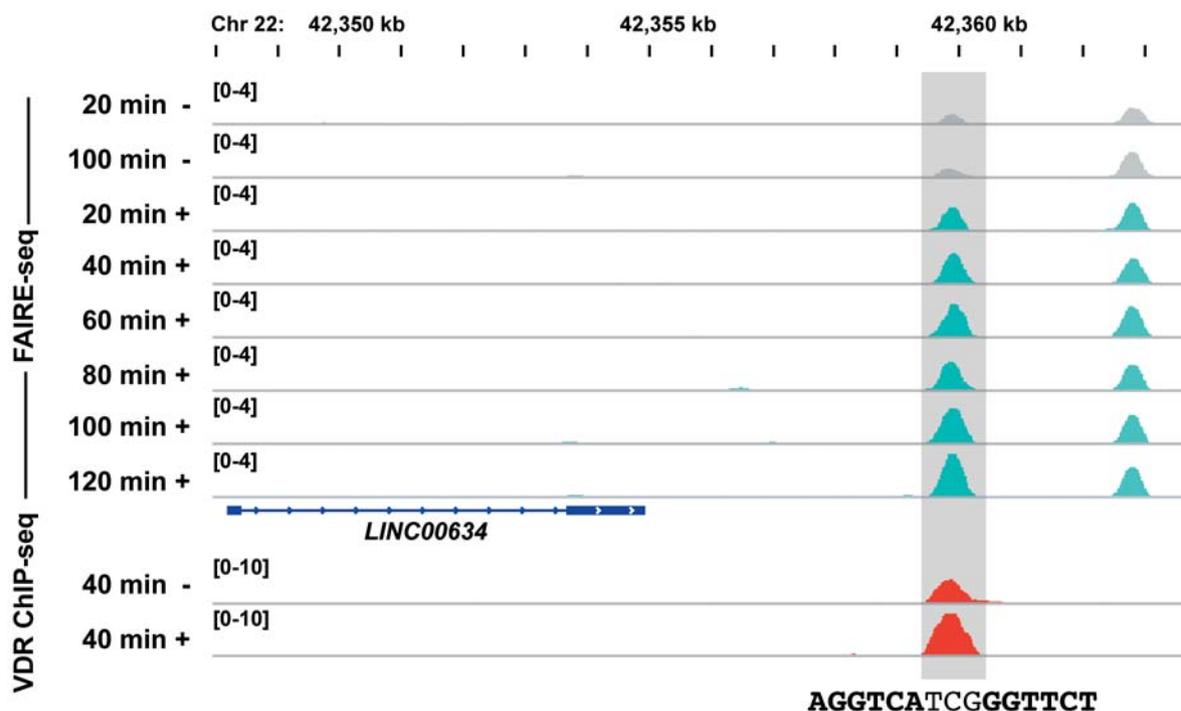


Figure 4. Vitamin D effects on chromatin. The IGV browser was used to visualize $1,25(\text{OH})_2\text{D}_3$ -dependent chromatin opening determined by FAIRE-seq (44) (grey for solvent-treated (-) and light blue for $1,25(\text{OH})_2\text{D}_3$ -stimulated (+) samples) and ligand-dependent VDR binding measured by ChIP-seq (19) (red) at a locus 10 kb downstream of the TSS of the long non-coding RNA gene LINC00634 in THP-1 cells. The gene structure is shown in blue and the sequence of the DR3-type VDR binding motif is indicated below the peak.

confirmed the preferential occurrence of DR3-type sequences below the summits (± 100 bp) of VDR peaks (Figure 3A). However, by far not below all VDR peaks a DR3-type motif could be identified. Transcription factor binding site screening algorithms, such as HOMER (20), allow different threshold settings, so-called “scores”, for detecting more or less deviations from the consensus sequence. The screening of the total set of 23,409 non-overlapping VDR binding sites with HOMER scores between 4 and 9 (Figure 3B) demonstrates that even with a score of 4, which represents rather degenerated sequence motifs (Figure 3C), only some 40% of all VDR peaks have DR3-type sequence below their summit. A more detailed analysis of the VDR binding sites of six cellular models and a differentiation between $1,25(\text{OH})_2\text{D}_3$ -treated cells and unstimulated samples (49) showed that ligand-stimulated samples have a higher rate of DR3-type sequences than that of non-treated cells. Moreover, cellular models, such as monocytes, for which a lower total number of VDR peaks have been identified, show a higher DR3 percentage than B cells, for which a large number of peaks have been reported. Importantly, the top 200 VDR sites of all six ChIP-seq datasets have a DR3 rate of more than 60%. This suggests that the most prominent VDR binding sites, which are often also the most

ligand responsive loci (Figure 1A), play an important role in the response of the respective cellular system to vitamin D. Moreover, this implies that the total number of VDR peaks is less important than the quality of the most responsive sites.

The results of the transcription factor binding site screening (Figure 3B) suggest that at least 60% of all presently known VDR binding sites do not contain a DR3-type sequence. This indicates that at these loci VDR uses a different mode of interaction with genomic DNA and implies that it partners with an alternative protein at these sites. Together with these partners VDR may recognize different binding sequences or it may even bind “backpack” of a DNA-binding transcription factor (3). No prominent alternative VDR partner has yet been characterized, but HOMER searches indicate that the receptor may use a divergent group of proteins, such as the transcription factors PU.1 (also called SPI1), ESRRB (also called NR3B2) and GABPA (49). A partnership of VDR with the well-known pioneer factor PU.1 (59) has already been suggested in the context of monocytic differentiation (32).

Taken together, in each vitamin D-responsive tissue or cell type there are more than 1,000 chromatin domains that combine VDR with its target genes. The more prominent and

ligand-responsive VDR binding sites have a higher rate of DR3-type sequences. In contrast, on the majority of its binding loci VDR interacts directly or indirectly with genomic DNA in a presently uncharacterized fashion.

Chromatin Effects of Vitamin D

Due to the intrinsic repressive nature of chromatin, accessible genomic DNA is an essential condition for the binding of a regular, so-called “settler” transcription factor, such as VDR (37). The recently developed method FAIRE-seq is a rather simple and straightforward approach to detect genome-wide accessible, *i.e.* not protein-bound, DNA within chromatin (15). Like in ChIP-seq, the technique involves crosslinking and sonication of chromatin but replaces the antibody immunoprecipitation by a simple phenol extraction. A FAIRE-seq peak represents open chromatin, while regions without a peak are within non-accessible closed chromatin. In average a cell has some 50,000 to 100,000 FAIRE-seq peaks, *i.e.* this number of accessible chromatin regions (10, 44).

FAIRE-seq was used to monitor the dynamic response of chromatin after a stimulation of THP-1 cells with $1,25(\text{OH})_2\text{D}_3$ through measurements every 20 min over a time period of 2 h (43, 44). The vast majority of all VDR binding sites (87% in THP-1 cells) are associated with open chromatin but only at some 200 genomic regions a significant increase of accessible chromatin was detected. Figure 4 displays the example of a ligand-inducible VDR binding locus (comparable to that of the *ASAP2* gene shown in Figure 1A) located 10 kb downstream of the TSS of the long non-coding RNA gene *LINC00634*, at which significant opening of chromatin can be observed. Interestingly, VDR binding sites at dynamically opening chromatin regions contain a far higher rate of DR3-type sequences below the peak summits (66% in THP-1 cells) as an average site (20% in THP-1 cells) (44). This suggests that the association with $1,25(\text{OH})_2\text{D}_3$ -triggered chromatin opening is another indication a prominent VDR binding site.

In summary, FAIRE-seq is a technically simpler but more comprehensive approach than performing VDR ChIP-seq. The method allows for identification of the most prominent VDR binding loci *via* monitoring dynamically opening chromatin regions after $1,25(\text{OH})_2\text{D}_3$ stimulation.

Shifting from Cultured Cells to Primary Human Tissues and Cell Types

Experiments in cultured cells of the past 25 years discovered many of the principles of gene regulation by $1,25(\text{OH})_2\text{D}_3$. In particular, the above described genome-wide insight from ChIP-seq, ChIA-PET and FAIRE-seq assays provided a major advance in understanding. However, it should be noted that most of these experiments were performed with cancer cells under conditions that do not reflect the reality of vitamin D

endocrinology. For example, in most assays the culture medium was depleted from lipophilic molecules, such as vitamin D, and a pharmacologically high dose of $1,25(\text{OH})_2\text{D}_3$ (10 to 100 nM) was used, in order to activate VDR within a few hours. However, in reality the levels of vitamin D_3 and its metabolites 25-hydroxyvitamin D_3 ($25(\text{OH})\text{D}_3$) and $1,25(\text{OH})_2\text{D}_3$ do not undergo rapid changes (8, 31). Serum $25(\text{OH})\text{D}_3$ concentrations, which are the widely accepted indicator of the vitamin D_3 status of the human body (21), change only in the order of weeks and months due to seasonal variations in sun exposure (56). Nevertheless, the vitamin D status of human individuals varies widely due to differences in diet, sun exposure, age, level of adiposity and (epi)genetic polymorphisms (11, 34, 47). As a consequence, worldwide billions of people have a serum $25(\text{OH})\text{D}_3$ concentrations below 50 nM and are vitamin D-deficient (22).

The genome-wide impact of different serum $25(\text{OH})\text{D}_3$ levels was investigated first by VDR ChIP-seq with primary T-cells, which were isolated from nine human individuals showing a variant vitamin D status (16). Interestingly, the number of observed VDR peaks ranged from 200 for a vitamin D-deficient individual to more than 7,000 for a person with high circulating $25(\text{OH})\text{D}_3$ concentrations. Unfortunately, the raw data of this study is not available, in order to perform a harmonized re-analysis in comparison with the six VDR ChIP-seq dataset from cell culture models. Nevertheless, for the sum of the nine individuals 14,044 unique VDR peaks were reported, from which HOMER analysis identified only 442 (3.1%) to carry a DR3-type sequence within its summit area. This represents a 6.7-times lower DR3 rate than observed with cultured cells under identical settings (HOMER score 7) (Figure 3B). This may be due to the fact that *in vivo* ChIP-seq is technically more challenging and may have led to a sub-optimal data quality of this study (16). However, it could also mean that gene regulation by vitamin D is *in vivo* less complex than *in vitro*. The latter possibility was impressively demonstrated by the FANTOM5 project (7) that compared gene expression in 750 primary human samples with that of 250 cell lines. In general, the FANTOM5 data suggest that the transcriptome-wide gene expression profile in primary human tissues and cell types differs significantly from their cell culture surrogates. This also implies that in future experiments should be preferably performed *in vivo*, since the culture of primary cells *in vitro* is inducing changes in their epigenome.

Human *in vivo* experiments are an ethically-difficult issue. However, vitamin D_3 is a save micronutrient that is taken daily as a supplement by millions of people and endogenously produced by everyone after sun exposure. This provides numerous possibilities for the investigation of gene regulatory effects of vitamin D_3 . One of these was a study that used samples donated by 71 elderly, pre-diabetic individuals that participated in a 5-month vitamin D_3 intervention trial (VitDmet) during Finnish winter (5). Peripheral blood

mononuclear cells (PBMCs) (5, 58) and adipose tissue biopsies (39) were obtained both at the start and the end of the study and were investigated for changes in mRNA expression of a series of primary VDR target genes, such as *CD14*, *NR1P1* and *DUSP10*. These gene expression changes were correlated with alterations in serum 25(OH)D₃ levels during the 5-month intervention and allowed a classification of human individuals based on their responsiveness to vitamin D₃. Interestingly, only for some 60% of the individuals a significant correlation between gene expression and vitamin D status could be obtained. This suggests that, at least on the level of gene expression, vitamin D₃ supplementation was unnecessary for the remaining studied subjects (5).

Taken together, results of the FANTOM5 project strongly suggest shifting the study of gene regulation by vitamin D from *in vitro* cell culture experiments to primary cells stimulated *in vivo*. First results suggest that both the number of genome-wide VDR binding sites and the expression of VDR target genes correlate with the vitamin D status of the studied individuals.

Conclusion

The genome-wide view on VDR locations *via* ChIP-seq or on 1,25(OH)₂D₃-induced chromatin opening *via* FAIRE-seq have significantly broadened the understanding of gene regulation by vitamin D. Some previous knowledge, such as the preferential binding of VDR to DR3-type sequences, had been confirmed but also challenged at the same time, since only a minority of all VDR loci contain these sequences. The large number of 23,000 non-overlapping VDR sites known so far from six human cellular models or the 14,000 VDR loci identified in primary T-cells from nine human individuals is initially overwhelming. However, not all these VDR sites seem to be equally important. In cell culture models only some 10% to 30% of the VDR loci carry a DR3-type sequence (translating to 300 to 1,000 sites), some 200 VDR sites are at 1,25(OH)₂D₃-triggered chromatin regions and only at some 50 genomic positions sites VDR is found in all tested cellular models. This suggests that in a given cell type only a few hundred VDR sites play a critical role, while many of the other VDR loci may rather represent “noise”, than having a specific function. Interestingly, this number is in the same order as the sum of primary 1,25(OH)₂D₃ target genes. Therefore, a combination of genome-wide assessment of VDR loci by ChIP-seq, monitoring of 1,25(OH)₂D₃-inducible chromatin sites and screening for DR3-type motifs below the peak summits may be presently the best experimental approach to understand the effects of vitamin D on the genome of a given cell type. Applying these methods to primary human tissues and cell types that have been stimulated *in vivo* may be the best way to evaluate the responsiveness to, and needs for, vitamin D of a human individual on the molecular level.

Acknowledgements

The Author thanks the Academy of Finland and the Juselius Foundation for their support.

References

- Campbell MJ: Vitamin D and the RNA transcriptome: more than mRNA regulation. *Front Physiol* 5: 181, 2014.
- 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.
- Carlberg C and Campbell MJ: Vitamin D receptor signaling mechanisms: Integrated actions of a well-defined transcription factor. *Steroids* 78: 127-136, 2013.
- Carlberg C and Molnár F: Current status of vitamin D signaling and its therapeutic applications. *Curr Top Med Chem* 12: 528-547, 2012.
- Carlberg C, Seuter S, de Mello VD, Schwab U, Voutilainen S, Pulkki K, Nurmi T, Virtanen J, Tuomainen TP and Uusitupa M: Primary vitamin D target genes allow a categorization of possible benefits of vitamin D₃ supplementation. *PLoS One* 8: e71042, 2013.
- Chun RF, Liu PT, Modlin RL, Adams JS and Hewison M: Impact of vitamin D on immune function: lessons learned from genome-wide analysis. *Front Physiol* 5: 151, 2014.
- FANTOM consortium: A promoter-level mammalian expression atlas. *Nature* 507: 462-470, 2014.
- DeLuca HF: Overview of general physiologic features and functions of vitamin D. *Am J Clin Nutr* 80: 1689S-1696S, 2004.
- Ding N, Yu RT, Subramaniam N, Sherman MH, Wilson C, Rao R, Leblanc M, Coulter S, He M, Scott C, Lau SL, Atkins AR, Barish GD, Gunton JE, Liddle C, Downes M and Evans RM: A vitamin D receptor/SMAD genomic circuit gates hepatic fibrotic response. *Cell* 153: 601-613, 2013.
- ENCODE-Project-Consortium, Bernstein BE, Birney E, Dunham I, Green ED, Gunter C and Snyder M: An integrated encyclopedia of DNA elements in the human genome. *Nature* 489: 57-74, 2012.
- Engelman CD, Fingerlin TE, Langefeld CD, Hicks PJ, Rich SS, Wagenknecht LE, Bowden DW and Norris JM: Genetic and environmental determinants of 25-hydroxyvitamin D and 1,25-dihydroxyvitamin D levels in Hispanic and African Americans. *J Clin Endocrinol Metab* 93: 3381-3388, 2008.
- Fretz JA, Zella LA, Kim S, Shevde NK and Pike JW: 1,25-Dihydroxyvitamin D₃ induces expression of the Wnt signaling co-regulator LRP5 *via* regulatory elements located significantly downstream of the gene's transcriptional start site. *J Steroid Biochem Mol Biol* 103: 440-445, 2007.
- Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EG, Huang PY, Welboren WJ, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY, Wansa KD, Zhao B, Lim KS, Leow SC, Yow JS, Joseph R, Li H, Desai KV, Thomsen JS, Lee YK, Karuturi RK, Herve T, Bourque G, Stunnenberg HG, Ruan X, Cacheux-Rataboul V, Sung WK, Liu ET, Wei C, Cheung E, and Ruan Y: An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature* 462: 58-64, 2009.
- Furey TS: ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions. *Nat Rev Genet* 13: 840-852, 2012.

- 15 Giresi PG, Kim J, McDaniell RM, Iyer VR, and Lieb JD: FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res* 17: 877-885, 2007.
- 16 Handel AE, Sandve GK, Disanto G, Berlanga-Taylor AJ, Gallone G, Hanwell H, Drablos F, Giovannoni G, Ebers GC, and Ramagopalan SV: Vitamin D receptor ChIP-seq in primary CD4+ cells: relationship to serum 25-hydroxyvitamin D levels and autoimmune disease. *BMC Med* 11: 163, 2013.
- 17 Haussler MR, Haussler CA, Jurutka PW, Thompson PD, Hsieh JC, Remus LS, Selznick SH, and Whitfield GK: The vitamin D hormone and its nuclear receptor: molecular actions and disease states. *J Endocrinol* 154 Suppl: S57-73, 1997.
- 18 Haussler MR, Whitfield GK, Kaneko I, Haussler CA, Hsieh D, Hsieh J-C, and Jurutka PW: Molecular mechanisms of vitamin D action. *Calcif Tissue Int* 92: 77-98, 2013.
- 19 Heikkinen S, Väisänen S, Pehkonen P, Seuter S, Benes V and Carlberg C: Nuclear hormone $1\alpha,25$ -dihydroxyvitamin D_3 elicits a genome-wide shift in the locations of VDR chromatin occupancy. *Nucleic Acids Res* 39: 9181-9193, 2011.
- 20 Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H and Glass CK: Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38: 576-589, 2010.
- 21 Hollis BW: Circulating 25-hydroxyvitamin D levels indicative of vitamin D sufficiency: implications for establishing a new effective dietary intake recommendation for vitamin D. *J Nutr* 135: 317-322, 2005.
- 22 Institute-of-Medicine: Dietary reference intakes for calcium and vitamin D. Washington, DC: National Academies Press 2011.
- 23 Kadauke S and Blobel GA: Chromatin loops in gene regulation. *Biochim Biophys Acta* 1789: 17-25, 2009.
- 24 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.
- 25 Kriebitzsch C, Verlinden L, Eelen G, van Schoor NM, Swart K, Lips P, Meyer MB, Pike JW, Boonen S, Carlberg C, Vitvitsky V, Bouillon R, Banerjee R and Verstuyf A: $1,25$ -dihydroxyvitamin D_3 influences cellular homocysteine levels in murine preosteoblastic MC3T3-E1 cells by direct regulation of cystathionine beta-synthase. *J Bone Miner Res* 26: 2991-3000, 2011.
- 26 Larriba MJ, Gonzalez-Sancho JM, Bonilla F and Munoz A: Interaction of vitamin D with membrane-based signaling pathways. *Front Physiol* 5: 60, 2014.
- 27 Maston GA, Landt SG, Snyder M and Green MR: Characterization of enhancer function from genome-wide analyses. *Annu Rev Genomics Hum Genet* 13: 29-57, 2012.
- 28 Meyer MB, Goetsch PD and Pike JW: A downstream intergenic cluster of regulatory enhancers contributes to the induction of CYP24A1 expression by $1\alpha,25$ -dihydroxyvitamin D_3 . *J Biol Chem* 285: 15599-15610, 2010.
- 29 Meyer MB, Goetsch PD and Pike JW: VDR/RXR and TCF4/beta-catenin cisromes in colonic cells of colorectal tumor origin: impact on c-FOS and c-MYC gene expression. *Mol Endocrinol* 26: 37-51, 2012.
- 30 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.
- 31 Norman AW: From vitamin D to hormone D: fundamentals of the vitamin D endocrine system essential for good health. *Am J Clin Nutr* 88: 491S-499S, 2008.
- 32 Novershtern N, Subramanian A, Lawton LN, Mak RH, Haining WN, McConkey ME, Habib N, Yosef N, Chang CY, Shay T, Frampton GM, Drake AC, Leskov I, Nilsson B, Preffer F, Dombkowski D, Evans JW, Liefeld T, Smutko JS, Chen J, Friedman N, Young RA, Golub TR, Regev A and Ebert BL: Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* 144: 296-309, 2011.
- 33 Orlando V: Mapping chromosomal proteins *in vivo* by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends Biochem Sci* 25: 99-104, 2000.
- 34 Orton SM, Morris AP, Herrera BM, Ramagopalan SV, Lincoln MR, Chao MJ, Vieth R, Sadovnick AD and Ebers GC: Evidence for genetic regulation of vitamin D status in twins with multiple sclerosis. *Am J Clin Nutr* 88: 441-447, 2008.
- 35 Park PJ: ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet* 10: 669-680, 2009.
- 36 Ramagopalan SV, Heger A, Berlanga AJ, Maugeri NJ, Lincoln MR, Burrell A, Handunnetthi L, Handel AE, Disanto G, Orton SM, Watson CT, Morahan JM, Giovannoni G, Ponting CP, Ebers GC and Knight JC: A ChIP-seq defined genome-wide map of vitamin D receptor binding: associations with disease and evolution. *Genome Res* 20: 1352-1360, 2010.
- 37 Razin A: CpG methylation, chromatin structure and gene silencing—a three-way connection. *Embo J* 17: 4905-4908, 1998.
- 38 Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G and Mesirov JP: Integrative genomics viewer. *Nat Biotechnol* 29: 24-26, 2011.
- 39 Ryyänen J, Neme A, Tuomainen TP, Virtanen JK, Voutilainen S, Nurmi T, de Mello VD, Uusitupa M and Carlberg C: Changes in vitamin D target gene expression in adipose tissue monitor the vitamin D response of human individuals. *Mol Nutr Food Res* 2014.
- 40 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 Acids Res* 34: 543-554, 2006.
- 41 Saramäki A, Diermeier S, Kellner R, Laitinen H, Väisänen S and Carlberg C: Cyclical chromatin looping and transcription factor association on the regulatory regions of the p21 (CDKN1A) gene in response to $1\alpha,25$ -dihydroxyvitamin D_3 . *J Biol Chem* 284: 8073-8082, 2009.
- 42 Schmidt D, Schwalie PC, Wilson MD, Ballester B, Gonçalves A, Kutter C, Brown GD, Marshall A, Flicek P and Odom DT: Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell* 148: 335-348, 2012.
- 43 Seuter S, Neme A and Carlberg C: Characterization of genomic vitamin D receptor binding sites through chromatin looping and opening. *PLoS ONE* 9: e96184, 2014.
- 44 Seuter S, Pehkonen P, Heikkinen S and Carlberg C: Dynamics of $1\alpha,25$ -dihydroxyvitamin D-dependent chromatin accessibility of early vitamin D receptor target genes. *Biochim Biophys Acta* 1829: 1266-1275, 2013.
- 45 Seuter S, Ryyänen J and Carlberg C: The ASAP2 gene is a primary target of $1,25$ -dihydroxyvitamin D in human monocytes and macrophages. *J Steroid Biochem Mol Biol* 144: 12-18, 2014.

- 46 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.
- 47 Snellman G, Melhus H, Gedeberg R, Olofsson S, Wolk A, Pedersen NL and Michaelsson K: Seasonal genetic influence on serum 25-hydroxyvitamin D levels: a twin study. *PLoS One* 4: e7747, 2009.
- 48 Toropainen S, Väisänen S, Heikkinen S and Carlberg C: The down-regulation of the human MYC gene by the nuclear hormone 1 α ,25-dihydroxyvitamin D₃ is associated with cycling of corepressors and histone deacetylases. *J Mol Biol* 400: 284-294, 2010.
- 49 Tuoesmäki P, Väisänen S, Neme A, Heikkinen S and Carlberg C: Patterns of genome-wide VDR locations. *PLoS ONE* 9: e96105, 2014.
- 50 Turunen MM, Dunlop TW, Carlberg C and Väisänen S: Selective use of multiple vitamin D response elements underlies the 1 α ,25-dihydroxyvitamin D₃-mediated negative regulation of the human CYP27B1 gene. *Nucleic Acids Res* 35: 2734-2747, 2007.
- 51 Umesono K, Murakami KK, Thompson CC and Evans RM: Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D₃ receptors. *Cell* 65: 1255-1266, 1991.
- 52 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 1 α ,25-dihydroxyvitamin D₃. *J Mol Biol* 350: 65-77, 2005.
- 53 Van Bortle K and Corces VG: The role of chromatin insulators in nuclear architecture and genome function. *Curr Opin Genet Dev* 23: 212-218, 2013.
- 54 van de Peppel J and van Leeuwen JP: Vitamin D and gene networks in human osteoblasts. *Front Physiol* 5: 137, 2014.
- 55 Vaquerizas JM, Kummerfeld SK, Teichmann SA and Luscombe NM: A census of human transcription factors: function, expression and evolution. *Nat Rev Genet* 10: 252-263, 2009.
- 56 Virtanen JK, Nurmi T, Voutilainen S, Mursu J, and Tuomainen TP: Association of serum 25-hydroxyvitamin D with the risk of death in a general older population in Finland. *Eur J Nutr* 50: 305-312, 2011.
- 57 Wang Y, Zhu J and DeLuca HF: Where is the vitamin D receptor? *Arch Biochem Biophys* 523: 123-133, 2012.
- 58 Wilfinger J, Seuter S, Tuomainen T-P, Virtanen JK, Voutilainen S, Nurmi T, de Mello VDF, Uusitupa M, and Carlberg C: Primary vitamin D receptor target genes as biomarkers for the vitamin D₃ status in the hematopoietic system. *J Nutr Biochem* 25: 875-884, 2014.
- 59 Zaret KS and Carroll JS: Pioneer transcription factors: establishing competence for gene expression. *Gen Dev* 25: 2227-2241, 2011.
- 60 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 D₃. *Mol Endocrinol* 20: 1231-1247, 2006.

Received August 30, 2014

Revised October 2, 2014

Accepted October 9, 2014