Vitamin D Receptor Gene Polymorphisms and Bone

Acknowledgements

The work presented in this thesis was conducted at the Department of Internal Medicine of the Erasmus Medical Center, Rotterdam, The Netherlands.

The studies described in this thesis were supported the Netherlands Organization for Scientific Research grant (NWO 903-46-178) and the European Commission under grant "GENOMOS" (OLK6-CT-2002-02629).

The Rotterdam Study is supported by the Erasmus Medical Center and Erasmus University Rotterdam, the Netherlands Organization for Scientific Research (NWO), the Netherlands Organization for Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry of Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam.

The contributions of the general practitioners and pharmacists of the Ommoord district to the Rotterdam Study are greatly acknowledged.

Financial support by the Departments of Internal Medicine of the Erasmus Medical Center, the Dutch Society for Calcium and Bone Metabolism (NVCB) is gratefully acknowledged.

Cover design: Yue Fang & André Uitterlinden

Front flap: Conservation of the human and mouse genomic VDR gene sequences.
Linkage disequillibrium (LD) structure of the VDR gene for Caucasian, Asian and African American populations. Figure legends are in page 23, 42 and 43.

Layout: Yue Fang
Printed by: Optima Grafische Communicatie, Rotterdam
ISBN: 90-8559-127-9

© Yue Fang, 2005
No part of this book may be reproduced, stored in a retrieval system or transmitted in any form or by any means, without permission of the author or, when appropriate, of the scientific journal in which parts of this book have been published.

Vitamin D Receptor Gene Polymorphisms and Bone

Polymorfismen van het vitamine D receptor gen en bot

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus
Prof.dr. S.W.J. Lamberts
and in accordance with the decision of the Doctorate Board

The public defence shall be held on
Wednesday 21 December 2005 at 9:45 hrs

by

Yue Fang
born in Wuhan, Hubei province, China

Doctoral Committee

Promotor:	Prof.dr. H.A.P. Pols
Other members:	Prof.dr.ir. T.J. Visser Prof.dr. P. Lips Prof.dr.ir. C. M. van Duijn
Copromotor:	Dr. A.G. Uitterlinden

献给：
我的父母、姐姐方真及家人

Dedicated to
my parents,
sister Fang Zhen
and her family

CONTENTS

List of abbreviations

1,25-(OH)2D3	1,25-dihydroxyvitamin D3
25-(OH)D3	25-hydroxyvitamin D3
ANCOVA	analysis of covariance
ANOVA	analysis of variance
BMD	bone mineral density
BMI	body mass index
CDCV	common disease/common variant
COL1A1	collagen type 1 alpha 1
COL2A1	collagen type 2 alpha 1
DBP	vitamin D binding protein
DBP-MAF	DBP-macrophage activating factor
DXA	dual energy X-ray absorptiometry
DNA	deoxyribonucleic acid
ddNTP	dideoxyribonucleic triphosphate
dNTP	deoxyribonucleic triphosphate
EMSA	electrophoretic mobility shift assay
EST	expressed sequence tag
FISH	Fluorescent in Situ Hybridization
Gc	group-specific component
HR (95% CI)	hazard ratio (95% confidence interval)
HVDRR	hereditary vitamin D-resistant rickets
HWE	Hardy-Weinberg equilibrium
htSNPs	haplotype tagging SNPs
LD	linkage disequilibrium
MAF	minor allele frequency
mRNA	messenger ribonucleic acid
PAC	P1 derived artificial chromosome vector
PAR	population attributable risk
PCR	polymerase chain reaction
RR	relative risk
RFLP	restriction fragment length polymorphism
SNP	single nucleotide polymorphism
TFBS	transcription factor binding site
TRPV	transient receptor potential channel of vanilloid receptor
UTR	untranslated region
VDR	vitamin D receptor
VNTR	variable number of tandem repeats
WGA	whole genome association

Manuscripts based on the studies described in this thesis

Chapter 2, 3, 5.1, & 7

Fang, Y., van Meurs, J.B., d'Alesio, A., Jhamai, M., Zhao, H., Rivadeneira, F., Hofman, A., van Leeuwen, J.P., Jehan, F., Pols, H.A. Uitterlinden AG (2005) Promoter and 3'-Untranslated-Region Haplotypes in the Vitamin D Receptor Gene Predispose to Osteoporotic Fracture: The Rotterdam Study. Am J Hum Genet, 77, 807-23.

Chapter 4

Fang Y, van Meurs JB, Bergink AP, Hofman A, van Duijn CM, van Leeuwen JP, Pols HA, Uitterlinden AG (2003) Cdx-2 polymorphism in the promoter region of the human vitamin D receptor gene determines susceptibility to fracture in the elderly. J Bone Miner Res 18:1632-1641

Chapter 5.2

Fang Y, Rivadeneira F., van Meurs JB, Pols HA, Ioannidis J., Uitterlinden AG Vitamin D Receptor Gene Bsm I Polymorphism and Fracture Risk: A Meta-analysis (Submitted)

Chapter 6

Fang Y, van Meurs JB, Rivadeneira F, van Schoor N, van Leeuwen JP, Lips P, Pols HA, Uitterlinden AG. Vitamin D receptor gene haplotype is associated with body height and bone size (To be submitted)

Chapter 8

Fang Y, Arp P, Meurs JB, van Leeuwen JP, Pols HA, Uitterlinden AG. Association of Vitamin D Binding Protein (DBP) Haplotype with Serum Vitamin D level and Osteoporosis (To be submitted)

Chapter 1

General Introduction

1.1 Function of vitamin D endocrine system

The Vitamin D endocrine system, consisting of the metabolites of vitamin D protein (involved in their synthesis and metabolism) and the vitamin D receptor (VDR) as well as its cofactors, plays an essential role in calcium and phosphate homeostasis and skeletal metabolism. This classical biological action includes intestinal calcium and phosphate absorption, bone modeling and remodeling, and renal calcium and phosphate reabsorption. The neoclassical actions of the system have an important role in hematopoietic, immune, nervous and endocrine systems[1,2]. 1 α, 25-dihydroxyvitamin D_3 ($1,25(OH)_2D_3$, the active form of vitamin D) suppresses proliferation and induces differentiation of mouse myeloid leukemia cells into macrophages[3], and induces differentiation of human promyelocytic leukemia cell into macrophages[4]. The immunosuppressive actions of $1,25(OH)_2D_3$ have been applied to control autoimmune diseases such as rheumatoid arthritis, systemic lupus erythematosus, type I diabetes, and transplant rejection. Vitamin D also contributes to hair growth and skin development, and $1,25(OH)_2D_3$ or its analogs have potential therapeutic effects on psoriasis[5,6], alopecia[7] and cancer[8,9]. In this thesis, we will only focus on the classical action of the system: skeletal and calcium metabolism.

The vitamin D receptor (VDR) mediates the pleiotropic effects of the vitamin D endocrine system. Inter-individual differences in aspects of the vitamin D endocrine system have been well documented. For example, circulating vitamin D levels vary by the season and latitude but also vary between individuals[10,11]. Furthermore, it is known that subjects respond differently when vitamin D analogs are administered[12-14]. Such inter-individual variability can also contribute to the differences in risk for diseases in which vitamin D plays an important role. They could be caused by genetic differences in important proteins in the vitamin D endocrine system, such as the VDR. The VDR mediates most effects of its ligand, $1,25(OH)_2D_3$ (see Fig. 1). Thus, one approach to understand inter-individual differences in the vitamin D endocrine system is to study the influence of variations in the DNA sequence of important proteins of this system. For example, Mutations in the VDR gene, such as in the DNA binding domain[15], the ligand-bind domain[16], or in splice sites[17], cause the hereditary vitamin D-resistant rickets (HVDRR), which is characterized by growth retardation and small body size. More subtle sequence variations are "polymorphisms" that are defined as sequence variations that occur at a frequency of > 1% in the population. Their influence on the vitamin D endocrine system is currently under scrutiny in relation to a number of so-called complex diseases and traits, such as aspects of bone health and osteoporosis in particular on which we focus in this thesis.

1.2 Structure of the VDR gene

After the cloning of the human VDR cDNA in 1988 by Baker et al.[18], it took almost 10 years before major parts of the genomic structure of the human VDR gene became clear as described by Miyamoto et al.[19]. Simultaneously, Croft et al.[20]

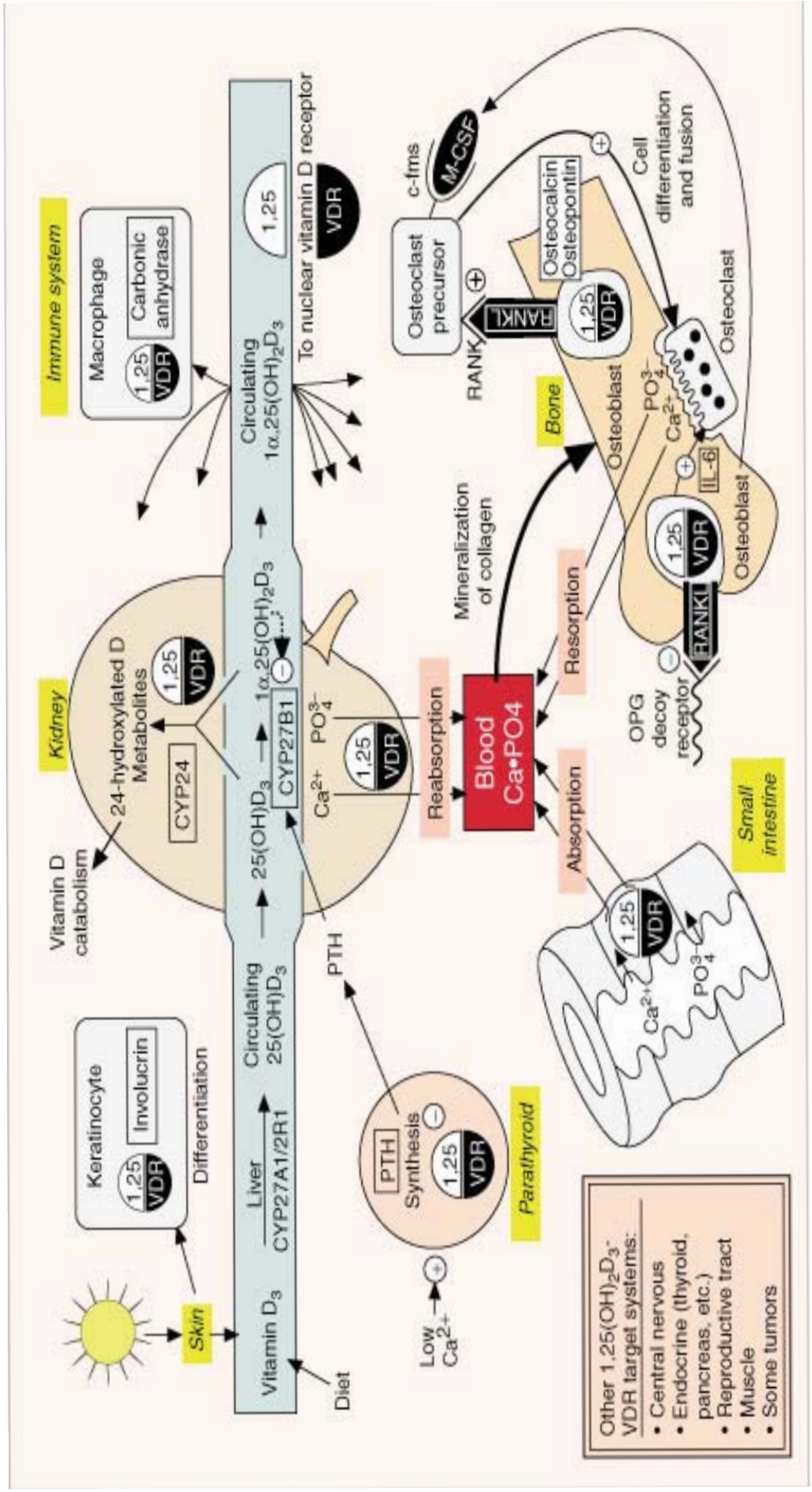


Figure 1 Vitamin D metabolism and its biologic actions via the nuclear VDR. Depicted at the upper left are the sources of vitamin D_3, which are either dietary or via a UV light-mediated photolysis of 7-dehydrocholesterol in the epidermis. The central bar depicts the enzymatic conversion of vitamin D_3 to the 25-hydroxy and then to the 1α, 25-dihydroxy derivative (1, $25(OH)_2D_3$), which is the active, hormonal form. These and subsequent hydroxylation reactions are catalyzed by CYPs. The 1, $25(OH)_2D_3$ hormonal form (shown as 1, 25 inside a white semicircle) then circulates to the various target tissues (selected target tissue identified *in italics*), where it binds to the nuclear VDR (depicted in white letters inside a black semicircle). (Adapted from Whitfield GK et al. [61])

demonstrated multiple tissue-specific transcripts (exon 1d to 1f) that differed at the 5'-end of VDR gene. Although much of this became known well before the Human Genome Project was finished, the data from Croft et al. is still not integrated in the Celera or NCBI gene databases. The size of the VDR gene and the exact location of the 5'-end exons were still unknown (Fig. 2). The location of the VDR gene on the physical map of chromosome 12 was indicated first by linkage mapping by Labuda et al.[21] and later on somewhat more refined by Fluorescent in Situ Hybridization (FISH) and radiation hybrid mapping by Taymans et al.[22]. However, these studies defined the position of the VDR gene in very general terms and this is insufficient for understanding the role of VDR polymorphisms in disease. The major reason to understand the exact gene structure and its size is coming from genetic association studies. Since mostly anonymous polymorphisms have been used in association studies, one explanation of an association observed was that not only the VDR gene but also another gene nearby could explain the association. It is therefore important to know which genes are also present in the area and how close they are to the VDR gene, and how likely it is that these other genes play a role in explaining the associations.

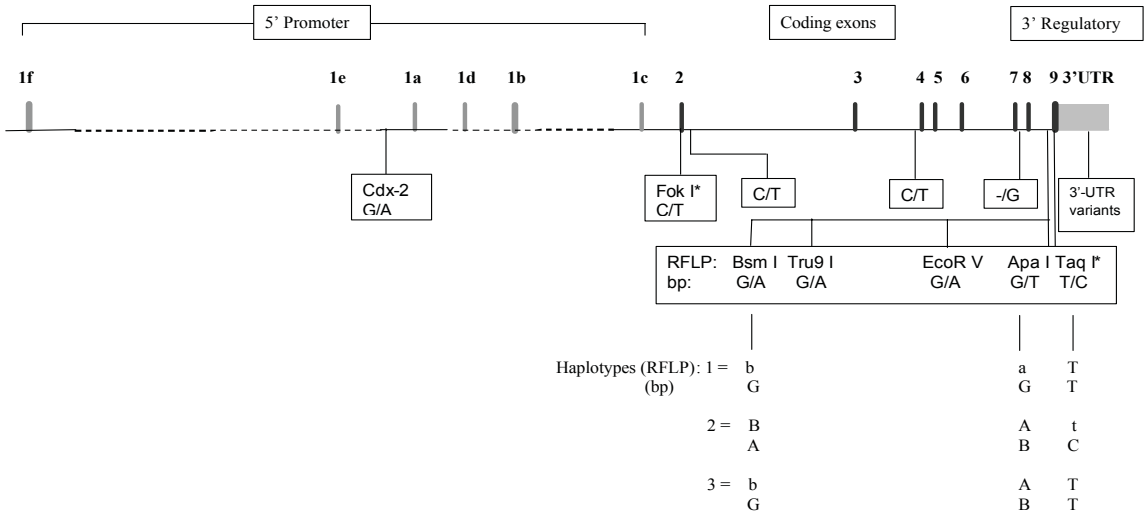


Figure 2 Genomic structure of the VDR gene and position of known polymorphisms before our study. The dashed line between exons indicates the distance between the two exons is uncertain. * indicates that these polymorphisms are in the coding sequence.

1.3 VDR Polymorphisms

Studies of VDR polymorphisms were initially focused on the 3'-end of the gene, ever since an association of the VDR polymorphisms with bone turnover and bone mineral density (BMD) was reported by Morrison[23,24]. A partial correction of the data on the BMD association appeared a few years later in which the association was much diminished[25]. In the lower portion of Fig.2, a number of the VDR polymorphisms are depicted that were known before the start of this

project. In the region of intron 8 and exon 9, Apa I[26], EcoR V[23], Bsm I[23], Taq I[24] and Tru9 I[27] restriction fragment length polymorphisms (RFLPs) were discovered and used in association studies. For the 3'-untranslated region (UTR) of the VDR gene, Morrison et al.[24] sequenced two individuals, who were homozygous for the most frequent Bsm-Apa-Taq haplotypes: BAt-BAt and baT-baT, and reported 13 distinct polymorphic sites including a poly(A)-tract with a varying number of adenosines. Durrin et al. expanded this approach and sequenced the 3'-UTR region in eight subjects and identified seven polymorphisms, of which four were common and three were rare in the eight annalysed subjects[28]. However, the number of subjects analysed in these studies is still limited and they were highly selected, so it is likely that more as yet uncovered polymorphisms exist in the complete 3'-UTR.

So far only two polymorphisms were reported in the coding exons of the VDR gene. One is the Taq I RFLP which is located in exon 9 but does not change the amino acid sequence of the VDR protein. Another is the Fok I RFLP. On comparison of the original Baker sequence of the VDR cDNA[18], two potential translation initiation start sites (ATG) were observed and subsequent sequence comparisons have shown that a T to C polymorphism exists (ATG to ACG) at the first potential start site[29,30]. This polymorphism, also referred to as the Start Codon Polymorphism (SCP), was later on defined using the FokI restriction enzyme in an RFLP test[31]. Thus, two protein variants can exist corresponding to the two available start sites: a long version of the VDR protein (the T nucleotide allele detected as the "f" allele, also referred to as the M1 form, i.e., the methionine at first position) and a protein shortened by three amino acids (the C nucleotide allele detected as the "F" allele, also referred to as the M4 form, i.e., the methionine at fourth position). This is the only known protein polymorphism in the VDR gene so far. Brown et al. sequenced the VDR coding region in 59 parathyroid tumors to find mutations[32]. Apart from the previously reported Taq I and Fok I polymorphisms, they reported no polymorphism in the coding region and found two intronic polymorphisms near exon 2 and 8.

Another VDR polymorphism was found through sequence analysis of a targeted area in the promoter region of VDR gene. After Yamamoto et al. found a binding site for Cdx-2, an intestinal-specific transcription factor, in the VDR 1a promoter region (based on the genomic structure of VDR from Miyamoto et al.[19]), Arai and colleagues reported a G to A sequence variation in this binding site among Japanese women[33].

1.4 Linkage Disequilibrium and Haplotypes of VDR Gene

Linkage Disequilibrium (LD) measures describe the association (or co-occurrence) of alleles of adjacent polymorphisms with each other[34]. This means in practice that the genotype of one polymorphism can predict the other adjacent "linked" one because very little recombination has occurred between them over the time of evolution and population history. High levels of LD in a certain area

will coincide with a limited number of "haplotypes" in that area. Haplotypes are blocks of linked alleles of adjacent polymorphisms, whereby the length of such a block coincides with the strength of LD aross the area. In practice, this means that relatively few polymorphisms have to be genotyped to "cover" the variance in a certain area. Therefore, a massive effort is currently under way to determine a haplotype map of the human genome[34,35].

It follows that the LD (or haplotype) structure of a certain candidate gene, such as the VDR, is important for association analyses to understand how the polymorphic variation in such a gene can contribute to risk of disease and population variance of certain phenotypes of interest. When a certain allele of one polymorphism has been found to be associated with "risk" of fracture (e.g.), it follows that this association might be explained by the effect of a particular allele, or by one (or more) other alleles that happen to be linked to this allele within the haplotype block, because of LD and the haplotype structure. Once we know which haplotype carries this risk allele, we can determine by cell biological and molecular biological functional analyses which of the variants on that haplotype allele truly cause this effect.

Based on some of the known polymorphisms shown in Fig. 1, several studies have analyzed the extent of LD especially at the 3'-end of the VDR gene. Since these analyses have used only a small number of polymorphisms, accurate information on LD and haplotypes has been very limited so far. Nevertheless, strong LD at this region has been observed for the Bsm I, Apa I, EcoR V, and Taq I RFLPs as well as the poly(A) variable number of tandem repeats (VNTR) in the 3'-UTR[23,24,28,36,37]. Thus, an LD map using a high density of polymorphisms across the VDR gene is necessary to be determined, to use the haplotype information in association studies.

1.5 Osteoporosis and VDR Polymorphisms

Osteoporosis is defined as a systemic skeletal disease characterized by low bone mass and microarchitectural deterioration of bone tissue, with a consequent increase in bone fragility and susceptibility to fracture[38]. It is a complex genetic disorder, which involves interaction between environmental and genetic factors. Because of the importance of the vitamin D endocrine system for a proper bone homeostasis, it quickly became a target for genetic association studies. Many genetic association studies have since demonstrated a relationship between the polymorphisms of the VDR gene (depicted in Fig. 2) with decreased bone mineral density (BMD), and increased fracture risk, but there are also many negative studies.

BMD is one of the most important predictors of osteoporosis, especially for postmenopausal women. Low BMD in this population results from a low peak BMD or/and faster bone loss with aging, and is a strong determinant of fracture risk. Morrison et al. first demonstrated that the Bsm I RFLP in the last intron of the VDR gene was related to serum osteocalcin concentration, which is an important biochemical marker of bone turnover[23]. They subsequently found the Bsm I RFLP

to be associated with differences in BMD in a twin study and in postmenopausal women[24]. Although the initial observations on the twin study (but not those on postmenopausal women) have been withdrawn[25], in the following years dozens of papers were published analyzing the same RFLP in relation to BMD. However, also controversial observations of the relationship between the commonly used RFLP and BMD were reported. Three studies[39-41] summarized association studies between the 3' and Fok I VDR polymorphisms and BMD, but still conflicting conclusions were drawn and some questions remained from those analyses.

The recently described G to A sequence variation in the Cdx-2 binding site just upstream of exon 1a has also been found to be associated with BMD[33]. This site is suggested to play an important role in intestinal-specific transcription of the VDR gene. As the intestine is the site where the calcium absorption predominantly takes place, the Cdx-2 site is thought to influence the vitamin D regulation of calcium absorption. The A-allele has been demonstrated to have higher affinity than the G-allele for binding to the Cdx-2 transcription factor, and thus having more transcriptional activity [31]. With more VDR expression in the intestine, the A-allele thereby can increase the transcription of VDR downstream genes, especially, calcium transport proteins such as TRPV5, TRPV6, calbindin-D_{9k}, and calbindin-D_{28k}. This could enhance the intestinal absorption of calcium and result in increased BMD. Indeed, this increased BMD has been demonstrated for Japanese postmenopausal women who carry the A-allele. However, because the population analysed encompassed only 55 postmenopausal women, the power of this study was low. In addition, this study did not analyse the relationship with fracture, the clinically most relevant endpoint in osteoporosis. The association between this polymorphism with other endpoints of osteoporosis such as fracture and bone geometry parameters is therefore interesting to be investigated in other preferably large populations, including those of different ethnic background than Japanese.

Fracture is an import clinical endpoint to assess in genetic studies of osteoporosis but also has some drawbacks. The etiology of different types of fracture (e.g., vertebral fracture versus hip fracture versus wrist fracture) is likely to be different. For example, the mean age at which fractures occur at the hip and the wrist is around 80 years and 65 years, respectively[42]. In addition, fractures are not determined by bone strength alone. Also falling risk plays a role, as well as cognitive and neuromuscular fitness contribute to the overall risk of suffering a fracture. Finally, the few studies that have analysed the heritability of fracture risk showed low heritability ranging from 20 – 45%[43-45]. This low heritability could be explained by the factors discussed above, which might also involve a large contribution of environmental factors. Polymorphisms in the 3'-end of the VDR gene, especially Bsm I and Taq I RFLPs, are most commonly used in VDR gene association studies, but they are highly linked to polymorphisms in the 3'-UTR, which is thought to be the potentially functional region involving the stability of VDR mRNA. Eighteen association studies (written in English) have so far analysed the relation-

ship between VDR polymorphisms and fracture risk but with conflicting results. The reason of this controversy could be small study size, diverse study designs and use of different study populations, but also involve heterogeneity in types of fractures studied. The Fok I RFLP has been analysed in relation to fracture in two studies[46,47], but no evidence of an association between Fok I and fracture risk was observed. So far, no meta-analysis has systematically reviewed the published data of this polymorphism.

Body height is another skeletal phenotype with a strong genetic background. Several twin studies indicate a heritability of up to 90%[48-50]. Bone size is an important determinant of bone strength, and thus, a risk factor for osteoporosis. Vitamin D is known to regulate the proliferation, differentiation, and maturation of cells responsible for skeletal growth, bone modeling and remodeling. Hence, genetic variations in the VDR gene might contribute to inter-individual differences in bone dimensions, growth and skeletal size characteristics, expressed as differences in height/stature, vertebral area, or femur shaft diameter. Not only deleterious mutations in the VDR gene cause HVDRR (which is characterized by growth retardation and small body size), but VDR polymorphisms also seem to be associated with differences in body height[51-54], but not consistently[55]. Altogether, the data indicate that several polymorphisms of the VDR gene might be associated with anthropometric differences. A meta-analysis is necessary to systematically review the data, and more polymorphisms across the VDR gene are interesting to be investigated in different populations to test the consistency of the association.

1.6 Functionality of VDR Polymorphisms

The interpretation of association studies using VDR polymorphisms is severely hindered by the fact that most of the polymorphisms used are anonymous, i.e., have an unknown functional effect. The likely explanation for any observed association is then to assume the presence of one or more truly functional sequence variations elsewhere in the gene which is – to a certain extent – in linkage with an allele of the anonymous (or marker) polymorphism used. As can be understood from the complex organization of the VDR gene (Fig. 1), the identification of these functional polymorphisms in the VDR gene is a challenging enterprise.

To identify functional sequence variations in the VDR gene most investigators have focused on the 3' regulatory region because this is close to the anonymous markers used mostly in association studies. While the Bsm I, Apa I, and Taq I RFLPs are located in the 3'-end of the gene, and LD extends into the 3'-UTR which contains several polymorphisms[24]. The 3'-UTR of genes is known to be involved in regulation of gene expression, especially through regulation of mRNA stability. Morrison et al.[23] earlier provided evidence of different luciferase activity in monkey COS-7 and rat ROS 17/2.8 cell lines for the two 3'-UTRs that are linked to the two most frequent haplotypes "baT" and "BAt" according to Bsm I, Apa I and Taq I RFLPs (see Fig. 2). Durrin et al. demonstrated that certain parts of the 3'-UTR,

so-called destabilizing elements, are involved in determining stability of the VDR-mRNA[28]. However, the UTRs linked to the "baT" and "BAt" haplotypes were not found by them to differ with respect to mRNA stability. Furthermore, heterologous constructs (human VDR-UTR sequences coupled to a rabbit beta-globin gene) and cell types (mouse NIH3T3 cells) were used to test for functionality. Especially, since it is known that 3'-UTRs display cell type specific effects on mRNA stability, this could be important in demonstrating functionality of sequence variations in the UTR.

Several studies[31,56-59] have been carried out to find evidence for the functionality of the non-anonymous polymorphism, detected as an Fok I RFLP, in the initial coding region of the VDR gene. From those studies, it appears that the Fok I polymorphism seems to be functional and the 424 aa VDR variant (F-allele) is somewhat more active than the 427 aa variant (f-allele) in terms of its transactivation capacity as a transcription factor. However, no consistent association studies support those functionality findings so far. The Cdx-2 polymorphism, which is located in the VDR promoter region, has been well-characterized by studies of Yamamoto et al. and Arai et al.[33,60]. Although the functionality of this polymorphism has indeed been convincingly demonstrated and the association of the polymorphism and BMD is also described in a small Japanese population, the exact mechanism whereby the polymorphism would relate to the risk for fracture has not been elucidated yet and requires further study.

1.7 Aim and scope of the study

The main purpose of this study is to characterize functional sequence variations across genes of the vitamin D endocrine system, in particular the VDR gene, and to study association of such polymorphisms to aspects of osteoporosis. To this end we performed studies to find polymorphisms, to analyse the patterns of LD across the gene area, and we analysed functional aspects of several combinations of polymorphisms. This research involves several approaches including molecular genetics, genetic epidemiology, bio-statistics, bioinformatics and cell biology.

1.8 Description of Chapters

Chapter 2 *Genome structure, homology analysis and re-sequencing of the VDR gene*: In this chapter, we first describe the updated genome structure of the human VDR gene according to the integrated data from publications and web resources. The result of homology analysis between human and mouse VDR gene genomic DNA is presented for the determination of potentially functional regions across the human VDR gene. We present sequence variations that we found in a re-sequencing study of those regions across the VDR gene.

Chapter 3 *LD structure of the VDR gene*: This chapter presents high resolution LD maps and race-specific tagging SNPs to describe the haplotype block structure of the VDR gene. This information forms the basis for furthre association studies.

Chapter 4 *Cdx-2 polymorphism and susceptibility for fracture risk:* This section demonstrates an ecological study of the correlation between the Cdx-2 genotype with hip fracture risk for different ethnic populations, and the association of the Cdx-2 polymorphism and fracture risk in a subset of the Rotterdam Study.

Chapter 5 *VDR haplotypes and susceptibility for fracture risk*: In **Section 5.1** we demonstrate the relationship of haplotypes across the VDR gene (according to the genotyping of haplotype tagging SNPs) and risk of different types of fractures in the complete Rotterdam Study population. **Section 5.2** presents a meta-analysis of published data on the relationship between the VDR Bsm I polymorphism and fracture risk.

Chapter 6 *VDR haplotypes and body height*: This chapter presents association studies of VDR haplotypes and body height, and bone geometry parameters in two Dutch populations. A meta-analysis of published data on the relationship between VDR Bsm I polymorphism and body height was performed.

Chapter 7 *Functionality studies of VDR polymorphisms*: In this chapter we present functionality studies of EMSA and transactivation experiments for SNPs in the promoter region, and studies of VDR mRNA level and stability by 3'-UTR haplotypes in different cell lines.

Chapter 8 *Interaction of VDR and DBP haplotypes on fracture*: This chapter presents the LD map of the DBP gene using bioinformatics. The correlation of DBP haplotypes and serum vitamin D level, interactions between DBP haplotype with dietary calcium intake and VDR haplotype on fracture risk are demonstrated.

Chapter 9 *General discussions and conclusions*: We finally summarize the overall results of the thesis and pitfalls in the process, and the general discussion ends with suggestions for further research.

References

1. Brown, A.J., Dusso, A. & Slatopolsky, E. Vitamin D. *Am J Physiol* **277**, F157-75 (1999).
2. Haussler, M.R. et al. The nuclear vitamin D receptor: biological and molecular regulatory properties revealed. *J Bone Miner Res* **13**, 325-49 (1998).
3. Abe, E. et al. Differentiation of mouse myeloid leukemia cells induced by 1 alpha,25-dihydroxyvitamin D3. *Proc Natl Acad Sci U S A* **78**, 4990-4 (1981).
4. Mangelsdorf, D.J., Koeffler, H.P., Donaldson, C.A., Pike, J.W. & Haussler, M.R. 1,25-Dihydroxyvitamin D3-induced differentiation in a human promyelocytic leukemia cell line (HL-60): receptor-mediated maturation to macrophage-like cells. *J Cell Biol* **98**, 391-8 (1984).
5. Smith, E.L., Pincus, S.H., Donovan, L. & Holick, M.F. A novel approach for the evaluation and treatment of psoriasis. Oral or topical use of 1,25-dihydroxyvitamin D3 can be a safe and effective therapy for psoriasis. *J Am Acad Dermatol* **19**, 516-28 (1988).
6. Feliciani, C. et al. Nail Psoriasis: Combined Therapy with Systemic Cyclosporin and Topical Calcipotriol. *J Cutan Med Surg* (2004).
7. Vegesna, V. et al. Vitamin D3 analogs stimulate hair growth in nude mice. *Endocrinology* **143**, 4389-96 (2002).
8. Simboli-Campbell, M., Narvaez, C.J., Tenniswood, M. & Welsh, J. 1,25-Dihydroxyvitamin D3 induces morphological and biochemical markers of apoptosis in

MCF-7 breast cancer cells. *J Steroid Biochem Mol Biol* **58**, 367-76 (1996).
9. Narvaez, C.J., Zinser, G. & Welsh, J. Functions of 1alpha,25-dihydroxyvitamin D(3) in mammary gland: from normal development to breast cancer. *Steroids* **66**, 301-8 (2001).
10. Webb, A.R., Kline, L. & Holick, M.F. Influence of season and latitude on the cutaneous synthesis of vitamin D3: exposure to winter sunlight in Boston and Edmonton will not promote vitamin D3 synthesis in human skin. *J Clin Endocrinol Metab* **67**, 373-8 (1988).
11. Lips, P. Which circulating level of 25-hydroxyvitamin D is appropriate? *J Steroid Biochem Mol Biol* **89-90**, 611-4 (2004).
12. Wieder, R. et al. Pharmacokinetics and safety of ILX23-7553, a non-calcemic-vitamin D3 analogue, in a phase I study of patients with advanced malignancies. *Invest New Drugs* **21**, 445-52 (2003).
13. Koeffler, H.P., Aslanian, N. & O'Kelly, J. Vitamin D(2) analog (Paricalcitol; Zemplar) for treatment of myelodysplastic syndrome. *Leuk Res* **29**, 1259-62 (2005).
14. Ringe, J.D., Faber, H., Fahramand, P. & Schacht, E. Alfacalcidol versus plain vitamin D in the treatment of glucocorticoid/inflammation-induced osteoporosis. *J Rheumatol Suppl* **76**, 33-40 (2005).
15. Hughes, M.R. et al. Point mutations in the human vitamin D receptor gene associated with hypocalcemic rickets. *Science* **242**, 1702-5 (1988).
16. Kristjansson, K., Rut, A.R., Hewison, M., O'Riordan, J.L. & Hughes, M.R. Two mutations in the hormone binding domain of the vitamin D receptor cause tissue resistance to 1,25 dihydroxyvitamin D3. *J Clin Invest* **92**, 12-6 (1993).
17. Cockerill, F.J. et al. Mutations in the vitamin D receptor gene in three kindreds associated with hereditary vitamin D resistant rickets. *J Clin Endocrinol Metab* **82**, 3156-60 (1997).
18. Baker, A.R. et al. Cloning and expression of full-length cDNA encoding human vitamin D receptor. *Proc Natl Acad Sci U S A* **85**, 3294-8 (1988).
19. Miyamoto, K. et al. Structural organization of the human vitamin D receptor chromosomal gene and its promoter. *Mol Endocrinol* **11**, 1165-79 (1997).
20. Crofts, L.A., Hancock, M.S., Morrison, N.A. & Eisman, J.A. Multiple promoters direct the tissue-specific expression of novel N-terminal variant human vitamin D receptor gene transcripts. *Proc Natl Acad Sci U S A* **95**, 10529-34 (1998).
21. Labuda, M. et al. Two hereditary defects related to vitamin D metabolism map to the same region of human chromosome 12q13-14. *J Bone Miner Res* **7**, 1447-53 (1992).
22. Taymans, S.E. et al. The human vitamin D receptor gene (VDR) is localized to region 12cen-q12 by fluorescent in situ hybridization and radiation hybrid mapping: genetic and physical VDR map. *J Bone Miner Res* **14**, 1163-6 (1999).
23. Morrison, N.A., Yeoman, R., Kelly, P.J. & Eisman, J.A. Contribution of trans-acting factor alleles to normal physiological variability: vitamin D receptor gene polymorphism and circulating osteocalcin. *Proc Natl Acad Sci U S A* **89**, 6665-9 (1992).
24. Morrison, N.A. et al. Prediction of bone density from vitamin D receptor alleles. *Nature* **367**, 284-7 (1994).
25. Morrison, N.A. et al. Prediction of bone density from vitamin D receptor alleles [correction]. *Nature* **387**, 106 (1997).
26. Faraco, J.H., Morrison, N.A., Baker, A., Shine, J. & Frossard, P.M. ApaI dimorphism at the human vitamin D receptor gene locus. *Nucleic Acids Res* **17**, 2150 (1989).
27. Ye, W.Z., Reis, A.F. & Velho, G. Identification of a novel Tru9 I polymorphism in the human vitamin D receptor gene. *J Hum Genet* **45**, 56-7 (2000).
28. Durrin, L.K., Haile, R.W., Ingles, S.A. & Coetzee, G.A. Vitamin D receptor 3'-untranslated region polymorphisms: lack of effect on mRNA stability. *Biochim Biophys Acta* **1453**, 311-20 (1999).
29. Saijo, T. et al. A unique mutation in the vitamin D receptor gene in three Japanese patients with vitamin D-dependent rickets type II: utility of single-strand conformation polymorphism analysis for heterozygous carrier detection. *Am J Hum Genet* **49**, 668-73 (1991).
30. Gross, C. et al. The presence of a polymorphism at the translation initiation site of the vitamin D receptor gene is associated with low bone mineral density in postmenopausal

Mexican-American women. *J Bone Miner Res* **11**, 1850-5 (1996).
31. Arai, H. et al. A vitamin D receptor gene polymorphism in the translation initiation codon: effect on protein activity and relation to bone mineral density in Japanese women. *J Bone Miner Res* **12**, 915-21 (1997).
32. Brown, S.B. et al. Vitamin D receptor as a candidate tumor-suppressor gene in severe hyperparathyroidism of uremia. *J Clin Endocrinol Metab* **85**, 868-72 (2000).
33. Arai, H. et al. The polymorphism in the caudal-related homeodomain protein Cdx-2 binding element in the human vitamin D receptor gene. *J Bone Miner Res* **16**, 1256-64 (2001).
34. Wall, J.D. & Pritchard, J.K. Haplotype blocks and linkage disequilibrium in the human genome. *Nat Rev Genet* **4**, 587-97 (2003).
35. Gabriel, S.B. et al. The structure of haplotype blocks in the human genome. *Science* **296**, 2225-9 (2002).
36. Uitterlinden, A.G. et al. A large-scale population-based study of the association of vitamin D receptor gene polymorphisms with bone mineral density. *J Bone Miner Res* **11**, 1241-8 (1996).
37. Ingles, S.A. et al. Strength of linkage disequilibrium between two vitamin D receptor markers in five ethnic groups: implications for association studies. *Cancer Epidemiol Biomarkers Prev* **6**, 93-8 (1997).
38. Consensus development conference: diagnosis, prophylaxis, and treatment of osteoporosis. *Am J Med* **94**, 646-50 (1993).
39. Cooper, G.S. & Umbach, D.M. Are vitamin D receptor polymorphisms associated with bone mineral density? A meta-analysis. *J Bone Miner Res* **11**, 1841-9 (1996).
40. Gong, G. et al. The association of bone mineral density with vitamin D receptor gene polymorphisms. *Osteoporos Int* **9**, 55-64 (1999).
41. Thakkinstian, A., D'Este, C., Eisman, J., Nguyen, T. & Attia, J. Meta-analysis of molecular association studies: vitamin D receptor gene polymorphisms and BMD as a case study. *J Bone Miner Res* **19**, 419-28 (2004).
42. Schuit, S.C. et al. Fracture incidence and association with bone mineral density in elderly men and women: the Rotterdam Study. *Bone* **34**, 195-202 (2004).
43. Deng, H.W. et al. Genetic determination of Colles' fracture and differential bone mass in women with and without Colles' fracture. *J Bone Miner Res* **15**, 1243-52 (2000).
44. Keen, R.W., Hart, D.J., Arden, N.K., Doyle, D.V. & Spector, T.D. Family history of appendicular fracture and risk of osteoporosis: a population-based study. *Osteoporos Int* **10**, 161-6 (1999).
45. Andrew, T., Antioniades, L., Scurrah, K.J., Macgregor, A.J. & Spector, T.D. Risk of wrist fracture in women is heritable and is influenced by genes that are largely independent of those influencing BMD. *J Bone Miner Res* **20**, 67-74 (2005).
46. Langdahl, B.L., Gravholt, C.H., Brixen, K. & Eriksen, E.F. Polymorphisms in the vitamin D receptor gene and bone mass, bone turnover and osteoporotic fractures. *Eur J Clin Invest* **30**, 608-17 (2000).
47. Gennari, L. et al. FokI polymorphism at translation initiation site of the vitamin D receptor gene predicts bone mineral density and vertebral fractures in postmenopausal Italian women. *J Bone Miner Res* **14**, 1379-86 (1999).
48. Silventoinen, K., Kaprio, J. & Lahelma, E. Genetic and environmental contributions to the association between body height and educational attainment: a study of adult Finnish twins. *Behav Genet* **30**, 477-85 (2000).
49. Roberts, D.F., Billewicz, W.Z. & McGregor, I.A. Heritability of stature in a West African population. *Ann Hum Genet* **42**, 15-24 (1978).
50. Hirschhorn, J.N. et al. Genomewide linkage analysis of stature in multiple populations reveals several regions with evidence of linkage to adult height. *Am J Hum Genet* **69**, 106-16 (2001).
51. Tao, C. et al. Vitamin D receptor alleles predict growth and bone density in girls. *Arch Dis Child* **79**, 488-93; discussion 493-4 (1998).
52. Suarez, F., Rossignol, C. & Garabedian, M. Interactive effect of estradiol and vitamin D

receptor gene polymorphisms as a possible determinant of growth in male and female infants. *J Clin Endocrinol Metab* **83**, 3563-8 (1998).
53. Lorentzon, M., Lorentzon, R. & Nordstrom, P. Vitamin D receptor gene polymorphism is associated with birth height, growth to adolescence, and adult stature in healthy caucasian men: a cross-sectional and longitudinal study. *J Clin Endocrinol Metab* **85**, 1666-70 (2000).
54. Minamitani, K., Takahashi, Y., Minagawa, M., Yasuda, T. & Niimi, H. Difference in height associated with a translation start site polymorphism in the vitamin D receptor gene. *Pediatr Res* **44**, 628-32 (1998).
55. Baroncelli, G.I. et al. Vitamin-D receptor genotype does not predict bone mineral density, bone turnover, and growth in prepubertal children. *Horm Res* **51**, 150-6 (1999).
56. Gross, C. et al. The vitamin D receptor gene start codon polymorphism: a functional analysis of FokI variants. *J Bone Miner Res* **13**, 1691-9 (1998).
57. Jurutka, P.W. et al. The polymorphic N terminus in human vitamin D receptor isoforms influences transcriptional activity by modulating interaction with transcription factor IIB. *Mol Endocrinol* **14**, 401-20 (2000).
58. Whitfield, G.K. et al. Functionally relevant polymorphisms in the human nuclear vitamin D receptor gene. *Mol Cell Endocrinol* **177**, 145-59 (2001).
59. Colin, E.M. et al. Consequences of vitamin D receptor gene polymorphisms for growth inhibition of cultured human peripheral blood mononuclear cells by 1, 25-dihydroxyvitamin D3. *Clin Endocrinol (Oxf)* **52**, 211-6 (2000).
60. Yamamoto, H. et al. The caudal-related homeodomain protein Cdx-2 regulates vitamin D receptor gene expression in the small intestine. *J Bone Miner Res* **14**, 240-7 (1999).
61. Whitfield, G.K. et al. Nuclear Vitamin D Receptor: Structure-Function, Molecular Control of Gene Transcription, and Novel Bioactions. in *Vitamin D*, Vol. 1 (ed. David Feldman, J.W.P., Francis H. Glorieux) 219-61 (Elsevier, Burlington, 2005).

Chapter 2

Genomic Structure, Homology Analysis and Re-sequencing of the Vitamin D Receptor (VDR) Gene

This chapter represents a part of the manuscript:
Fang Y, van Meurs JB, d'Alesio A, Jhamai M, Zhao H, Rivadeneira F, Hofman A, van Leeuwen JP, Jehan F, Pols HA, Uitterlinden AG
Promoter and 3'-Untranslated-Region Haplotypes in the Vitamin D Receptor Gene Predispose to Osteoporotic Fracture: The Rotterdam Study
Am J Hum Genet (2005) 77:807-823

ABSTRACT

Introduction: Polymorphisms of the vitamin D receptor (VDR) gene have been reportedly associated with several complex diseases, including osteoporosis, but by unknown mechanisms and with inconsistent results. Most studies have used anonymous polymorphisms at the 3'-end and one in the initial coding region of the gene. However, no functional effect of the polymorphisms has been described.

Materials and Methods: We therefore searched the complete VDR gene for additional polymorphisms. We first corrected the genomic structure of the human VDR according to bioinformatics analysis of different databases. We sequenced 22 kb in the regions displaying high homology between human and mouse genomic VDR sequence, including the 3.2 kb 3'-untranslated region (3'-UTR), 4.1 kb of all coding exons and flanking introns and 14.7 kb of the 6 promoter exons 1a - 1f, in 15 Caucasians.

Results: We determined the structural organisation of the VDR promoter region by aligning our sequencing results and the Celera database, resolved a 500 bp gap in front of exon 1b. We identified 62 polymorphisms, including 55 SNPs and 7 tandem repeats. 22 SNPs (40%) were new and not contained in the NCBI and Celera databases. In the VDR promoter region, 14 polymorphisms change the putative recognition sequences of transcription factors, while 4 SNPs are located in destabilizing elements (DE) in 3'UTR.

Conclusions: We identified additional polymorphisms in potentially functional regions across the VDR gene. The result is a fundamental step for our next study of linkage disequilibrium (LD) analysis across the VDR gene.

INTRODUCTION

The vitamin D receptor (VDR, 12q13, MIM 601769) is a steroid receptor acting as a transcription factor responding to the biologically active form of the secosteroid vitamin D hormone. The vitamin D endocrine system is pleiotropic and plays an important role in skeletal metabolism, including intestinal calcium absorption and regulation of osteoblast differentiation, but has also been shown to modulate the immune response, insulin secretion, the renin/angiotension system, and growth of cancer cells[1].

Although Baker et al.[2] reported the human VDR cDNA in 1988, Miyamoto et al.[3] and Croft et al.[4] described the genomic structure and multiple tissue-specific transcripts at the 5'-end of VDR gene in 1998, but the complete genomic structure of the VDR gene is still not correctly described. The data from Croft et al. is still not integrated in the Celera or NCBI gene databases. The size of the VDR gene and the exact location of the 5'-end exons were still unknown (see Fig. 2 of Chapter 1).

As shown in chapter 1 of the thesis, many studies have been carried out to investigate the association of VDR polymorphisms with complex diseases, including osteoporosis, but conflicting results were observed. One of the reason of that is anonymous polymorphisms were used in most studies. To identify polymorphisms in potentially functional region of the VDR gene, we currently determined an accurate structure of the gene, and sequenced functional region across the VDR gene.

MATERIALS AND METHODS

Subjects

We sequenced genomic DNA from 15 young Caucasian individuals encompassing five homozygotes for each major VDR 3'-UTR haplotype, i.e., 11, 22, and 33, as defined in a previous study[5] based on Bsm I, Apa I and Taq I RFLPs.

Homology analysis

Human and mouse VDR genomic sequences (>105 kb from Celera database) were analyzed with the Vista program[6] to visualise the pair-wise percentage identity as calculated for every 100 bp.

Resequencing and sequence analysis

We sequenced 37 overlapping PCR fragments covering 22 kb of the VDR gene (shown in the upper-part of Figure 3) in 30 chromosomes including the 3.2 kb 3'-UTR, coding exons 2 to 9 and flanking introns (4.1 kb), and 14.7 kb promoter area containing 6 exons (1a-1f). PCR primers were designed according to published sequences (J03258[2]; PAC clone, AC004466[3]; AF080454-AF080456[4,7]) and the Celera database. Melting temperatures of primers were calculated with DNAMAN

Table 1 Primers and PCR conditions for the sequence analysis of the VDR gene

Primer name	Sequence of primer (5'→3')	$MgCl_2$ (nM)	Annealing Temp. (°C)	Size of PCR product (bp)	GenBank** Access No.
Promoter region					
1f-23h For	GAACCAGGATCATGTTTTGGA	1.5*	57	655	BV210414
1f-23h Rev	ACAATCCAAGTCTGTGGATAGATG				
1f-2kb For	CAAGTTTTGCCAGATTCACC	1.5*	58	932	BV210415
1f-2kb Rev	CCGCAGCACTATACAGAGGT				
1f-1kb For	CAAGTCCATATCTCCTTCTCCAG	1.5*	59	855	BV210416
1f-1kb Rev	AAGTGGTTTCCCTACCATCAAT				
1f-4h For	AGATACCTCTGGTTCCTCTAATGC	1.5*	59	631	BV210417
1f-4h Rev	CACACTTGTTCACCTCCACAC				
VDR1a-40h For	CTCAAGCATAGTGGCATGATCA	1.5	60	549	BV210418
VDR1a-40h Rev	TTCCTGATACAAAAGATGTTCTACAATG				
VDR Cdx-2 G For	AGGATAGAGAAAATAATAGAAAACATT	1.5	54	297	BV210419
VDR Cdx-2 A Rev	ACGTTAAGTTCAGAAAGATTAATTC				
VDR 1a-36h1 For	GAGTCATCCCTGATCCTTTTTGT	1.5	60	2201	BV210420
VDR 1a-36h1 Rev	TTCTGAGAAACAGATGAAGTGCC				BV210421
VDR 1a-36h2 For	CTGGCAGAGAGGTCAAGAGAC	1.5	60	1119	BV210422
VDR 1a-36h2 Rev	GCTGAGGAGTATCAGAGCTACTG				BV210423
VDR 1a-36h3 For	GAATGGTGCTTGTCATCGAGG	1.5	60	651	BV210424
VDR 1a-36h2 Rev	GCTGAGGAGTATCAGAGCTACTG				
VDR 1a-36h4 For#	AAGACCAGAGAATTGACAGTTCCTT	1.5	56	352	BV210425
VDR 1a-36h2 Rev	GCTGAGGAGTATCAGAGCTACTG				
VDR 1a-36h2 For	CTGGCAGAGAGGTCAAGAGAC	1.5	60	768	BV210421
VDR 1a-36h4 Rev	AGGAACTGTCAATTCTCTGGTCTT				
VDR 1a-14h For	GGGGATCCTTCCTATTACTTTCATTACA	1.5*	60	612	BV210426
VDR 1a-14h Rev	CTTAGACTCACTGTGCAGTGGAGATG				
VDR 1a-10h For	GGAGGTCATCGACTGCTGGA	1.5	60	535	BV210427

VDR 1a-10h Rev	CTAGCTCCGACGAATGGGAAA				
VDR 1a-5h For	ATCTGTGGGATCAGGCTGAGC	1.5*	60	550	BV210428
VDR 1a-5h Rev	CGCCTTTTGACAAGCAGAGACA				
VDR 1ad For	TGGTTGATTCCAAGTCAAGATGG	1.5*	60	604	BV210429
VDR 1ad Rev	CTCCAGCAGTTCTGAGCACCA				
VDR1d+7h For	CGCTAACACAGTGCTTAGCACTT	1.5	60	710	BV210430
VDR1d+7h Rev	AGTAAAATGCCTGCCCAACTG				
1gA-For	TCCTGTCTGGGCTCAGTTG	1.5*	60	879	BV210431
1gA-Rev	TAGCTTGGTAAGGGTCCAAGTC				
1gB-For	AATTCACAGTCTATGCTCTGGCT	1.5*	59	896	BV210432
1gB-Rev	GATCCTCTCAGAACTGGACAATAGT				
1gC-For	TGAGAGATAGCAGGAAGCAGAAC	1.5*	60	967	BV210433
1gC-Rev	CTGGCTGCTTACCTGCTTTAC				
1b-2k For	GACCACACTATCCCACAGAAAGT	1.5*	59	842	BV210434
1b-2k Rev	GAATGTCCCACCTTGCATAAC				
1b-gap For	AAGCGGAGTCTGTGCAGAC	1.5*	59	835	BV210435
1b-gap Rev	AGGTGGCTGCATCTCTTTAGA				
1c-2kb For	AAGGACCGTCCACTATTGGA	1.5*	59	875	BV210436
1c-2kb Rev	AAGTCCTTCCCAGCTGACC				
1c-12h For	CTGAGGAATCAATAAGGCCAG	1.5*	59	622	BV210413
1c-12h Rev	TGGATCTTAGCAGCTGGCT				
1c-8h For	TTGTCTATGTCTGCAGGTGGA	1.5*	59	816	BV210437
1c-8h Rev	CTTTTATCTAAGGCGGAGCG				
VDR E1c-For	CACTTCTGTTTGCAGTCACTGA	1.5	60	489	BV210438
VDR E1c-Rev	GGCTGGATAGGAAACATCAGA				
Exons					
VDR E2-For	ATGCTCTGAGCCAGCTATGTAG	1.5	60	571	BV210439
VDR E2-Rev	GAGAGTCAGAGGAACATCTGGA				
VDR E3-For	TGTCTTCTGTTGGAGAAATGGA	1.5	60	495	BV210440
VDR E3-Rev	AGTGCATCTGACCCTGGACT				

VDR E4-For	TTTCTTCACACAGTGGAGTGG	1.5	60	526	BV210441
VDR E4-Rev	GGCTTTGAGGAAGGTCTACAG				
VDR E5-For	ATCCTGAACAGAACTGGGGTA	1.5	60	519	BV210442
VDR E5-Rev	TGATGCTACACAGCTGGAATC				
VDR E6-For	TACTGCCTTATGCTGCTGAAA	1.5	60	557	BV210443
VDR E6-Rev	GAGAATCGCTTGAACCTAGGA				
VDR E78-For	CAGCAGGTGTATACCTGTCAAAG	1.5	60	634	BV210444
VDR E78-Rev	AGCAGGTCTTTGTCCTTCATACT				
VDR E9-For	CTAGGTCTGGATCCTAAATGCA	1.5	60	628	BV210445
VDR E9-Rev	TTAGGTTGGACAGGAGAGAGAA				
3'-UTR					
VDRU1-For	TGAGTGCAGCATGAAGCTAAC	1.5	59	580	BV210446
VDRU1-Rev	ATATAACCAGGGCAATGGGAT				
VDRU2-For	TCCTGCCTTACTCACGATAAATAA	1.5	60	612	BV210447
VDRU2-Rev	CTAGCTCTTAGCCCTGTGGTG				
VDRU3-For	AAGAATTTTCAGACCCCAGC	1.5	60	621	BV210448
VDRU3-Rev	TTTCCACCTGAAGAATTCTGAG				
VDRU4-For	GAGGAATCAGACTTCACACTGC	1.5	61	592	BV210449
VDRU4-Rev	AGAGACAGGGTTTCTCCATGTT				
VDRU5-For	GTAGGTGGATCACCTGAGGTC	1.5	61	656	BV210450
VDRU5-Rev	AGTAACTGATATTTCAGGAGTTCCC				
VDRU5-For	GTAGGTGGATCACCTGAGGTC	1.5	60	403	BV210451
VDRU5-PoA-Rev	CAGATCCAGACTTGGCTCTTT				
VDRU6-For	ATGCTGTTGCCTCATCTATAACA	1.5	60	680	BV210452
VDRU6-Rev	ATAATGATTCATCTCCCATAAGGTC				

#: Semi-nested PCR using the fragment generated with primer set of VDR 1a-36h2 For and VDR 1a-36h2 Rev as template

*: Add 5× Q-Solution (Qiagen) in PCR reaction system

**: Accession numbers of small sequence fragments in the GenBank database, and accession numbers of four long sequence fragments (2.5 to 9.6 kb, I to IV in Fig. 3) are AY827085 - AY827088.

(v4.0, Lynnon BioSoft). Primer sequence and PCR conditions are described in Table 1. The 50 μl PCR contained 20-80 ng of genomic DNA, 3-5 pmol primers, and 2 units Taq DNA polymerase (SUPER TAQ, HT Biotechnology LTD. UK, TP05c; or Promega, M1665) and was done 25 - 38 cycles in a Thermal Cycler 480 (Perkin Elmer) or a GeneAmp PCR System 9700 (Applied Biosystems, ABI). Taq PCR Core Kit (5 × Q-Solution from QIAGEN Ltd.) was used to amplify some GC-rich PCR fragments in promoter regions (Table 1). Sequencing extension reaction system (20 μl) consists of 11 μl purified PCR product, 4 μl of Terminator Mixture (ABI PRISM® Big-Dye Terminator cycles sequencing Ready Reaction Kit), 4 μl of 5 × buffer (400 mM Tris-Cl pH 9.0, 10 mM $MgCl_2$) and 1 μl of 10 pM primer,. We purified PCR and sequencing products by Quantum Prep® PCR Kleen Spin Columns and Micro Bio-spin® Chromatography Columns (BIO-RAD 732-6301 and 732-6224). Sequencing products were analyzed on an ABI PRISM® 310, or 3100 automated capillary Genetic Analyzer. All sequencing data was analyzed and aligned by Navigator (Perkin Elmer) and Sequencher (Gene Codes Corporation) sequence analysis software.

We used the TFSEARCH program[8] from the TRANSFAC databases developed by GBF-Braunschweig, Germany and MatInspector[9] from Genomatix software GmbH, to search for potential transcription factor binding sites (TFBS) around promoter polymorphisms.

RESULTS

Genomic structure of the VDR gene

The physical organisation of the human VDR gene region on chromosome 12q12-14 (kbp 46950 - 47350; Fig. 1a) and the VDR gene structure (Fig. 1b) is based on PAC clones (PAC228P16 and PAC1057I20), two published data[3,4], the Celera database, and our re-sequencing efforts. The VDR gene encompasses at least 105 kb: the large 5'-region of non-coding exons 1f to 1c is 60 kb, with exon 1f being 35 kb upstream of exon 1e, while exon 1e is 2 kb upstream of exon 1a. The nearest genes are COL2A1 at 20 kb distance upstream, with a small gene-like structure (MGC5576) close to the COL2A1 gene, and HDAC7A at 10 kb distance downstream.

Homology analysis

The overall percentage-identity between the entire human and mouse genomic sequence of the VDR gene is 28.8% (Fig. 2). While coding exons are found in highly conserved region (from 86.5 to 92.6% identity), except for exon 5 (61.2%). We also observed high homology regions in intronic areas, i.e., in total 14 kb of regions between 1f and 1e, 5kb of regions between exon 1b and 1c, and a small 500 bp region between exon 2 and 3. Interestingly, there is a lack of homology of

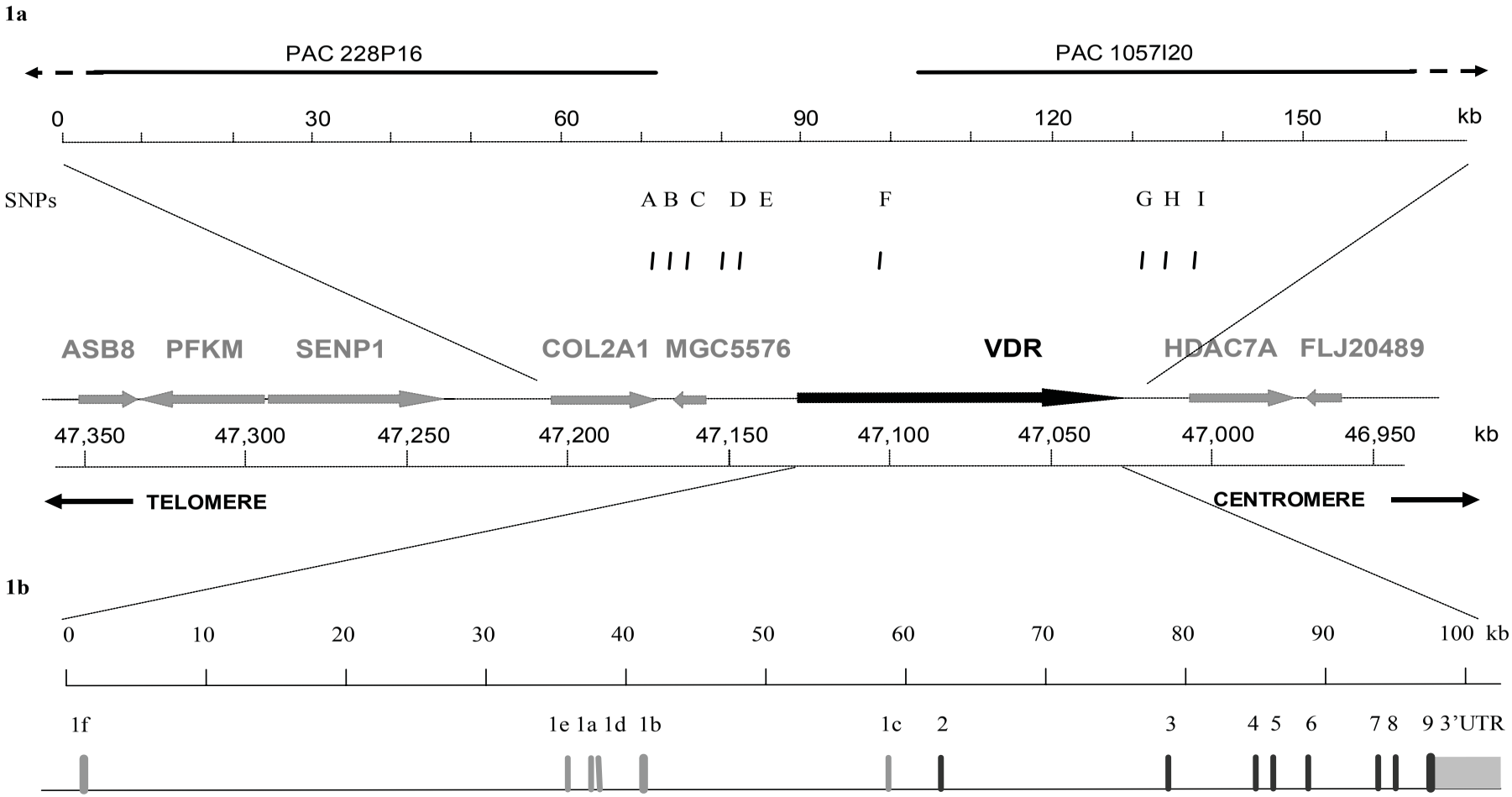


Figure 1 Genomic structure of the human VDR gene. **1a** Physical organization of the 12q12 area containing the VDR gene mostly based on the Celera database (47032 – 47145 kb at chromosome 12q12). The arrows for each gene indicate the transcription direction (distance in kb). **1b** The genomic structure of the human VDR gene. Dark bars indicate the coding exons of the VDR gene, the grey bars and box indicate 5' exons and 3'-UTR.

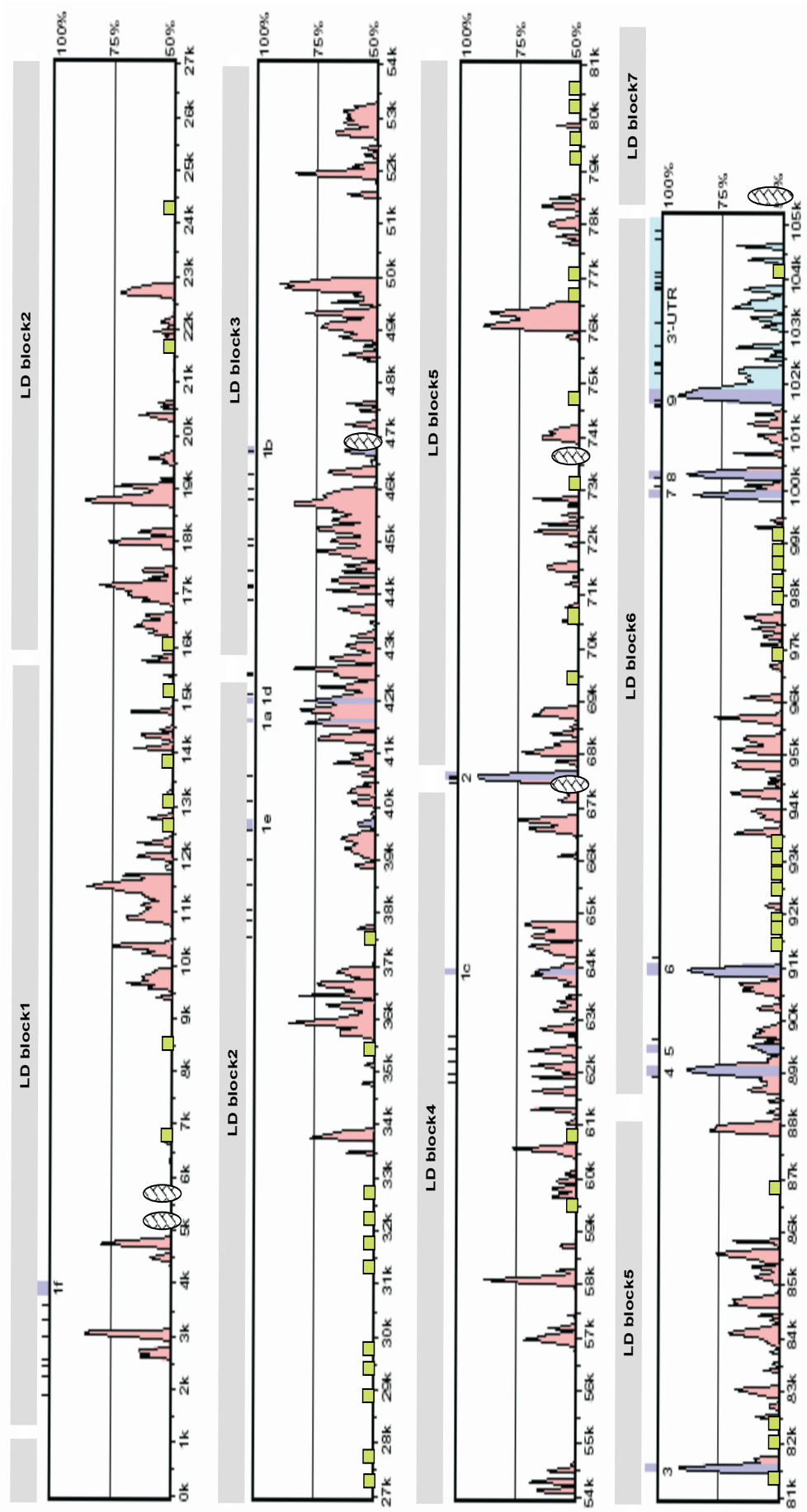


Figure 2 Conservation of the human and mouse genomic VDR gene sequence. Y-axis is the homology rate between the human and mouse, X-axis is the physical distance of the human VDR gene (in kilobasepairs). A color version of this figure is presented at the front flap page. All exons are indicated in purple, 3'-UTR in light blue and the conserved non-coding region in red. The small black bars on the top of each frame indicate the polymorphisms we observed by resequencing, the grey blocks on top indicate the LD blocks of the VDR gene (described in chapter 3), green boxes indicate Alu-repeats and elliptical boxes indicate Line-1 insertions.

the human exons 1b, 1e, and 1f with the mouse gene, while 1a, 1d and 1c are well conserved. Overall, much of the lack of homology can be explained by presence of Alu-repeats in the human VDR gene. We detected 62 Alu-like sequences mostly located in introns and between exons 1f and 1e, while one is located in the 3'-UTR (data not shown).

Resequencing of the VDR gene

We found 62 polymorphisms (Table 2 and Fig. 3) including 57 SNPs and 5 variable number tandem repeat (VNTR) polymorphisms, of which 18 polymorphisms (15 SNPs) were not present in the Celera database (chromosome 12, location from 47033000 to 47139000, April 2005 freeze/ NCBI Human Genome Build 35.1), NCBI dbSNP (dbSNP build 123), the Human Gene Mutation Database (HGMD), the HGBASE database and/or in publications. Nine (out of 62) polymorphisms locate in highly conserved non-coding regions (identity rate > 70%) in the 3'-UTR and in the promoter region. The average frequency of variation across 22 kb was 1/355 bp, which ranged from 1/247 bp (13/3.2 kb) in the 3'-UTR, 1/315 (13/4.1 kb) in exons (2 - 9) and surrounding introns, to 1/408 (36/14.7 kb) across the promoter region. 50 SNPs, with a minor allele frequency (MAF) ≥ 5% in the 15 subjects, were identified as "informative SNPs".

Three informative polymorphisms were located in the so-called "destabilizing elements (DE; described before[10]) in the 3'-UTR of the VDR gene (Fig. 3.). Four informative SNPs were observed in coding exons 2-9: a previously known C to T substitution (E2-C4T; detected as a FokI RFLP), E2-C59T and E8-C2T are synonymous substitutions; and a previously known synonymous SNP (E9-T32C; detected as a Taq I RFLP). In the promoter region exons 1a-1f, we found 36 sequence variations, including 5 VNTRs and 31 SNPs. One SNP in exon 1b (1b-C25A) would change the predicted amino acid sequence by threonine to lysine (1b-Thr8Lys). TRANSFAC analysis of the 5' upstream regions of promoter exons 1a-1f indicated that 14 polymorphisms change the core recognition sequence of potential transcription factors (Table 2).

DISCUSSION

The VDR gene is an important candidate gene of osteoporosis, and it has been studied for more than 17 years in genetic association analyses. However, the genomic structure of the gene described in the NCBI and the Celera databases still does not cover the whole length of the gene. In our current study, we determined a more accurate genomic structure of the human VDR gene, especially at the 5'-end of the gene. The length of the gene by our data is one third longer as compared to other databases. The borders of the VDR gene seem indicated by the homology and linkage analyses.

Table 2 Sequence variations of the VDR gene based on sequence analysis of 30 Caucasian chromosomes

No.	Polymorphism Nomenclature[A]	Minor allele (frequency by %)	Polymorphism and flanking sequence[B]	Position[C]	Characteristics[D]	Reference[C]	dbSNP SS# (build 124)
	Promoter region						
	Exon 1f						
1	1f-G-1904A	A (37)	GTGTGTGTGC**G/A**CCTGTGTGGG	47136401		Celera (hCV26069889)	32465633
2	1f-1889 $(CATGTGTGTG)_{2/1}$	$(CATGTGTGTG)_1$ (20)	TGGGTGTG**(CATGTGTGTG)$_{2/1}$**TGGATGTGG	47136219	USF (-)	Novel	32465817
3	1f-C-1570T	T (43)	CATATTTATC**C/T**TTTTATAC/TTTCTTTCAAG	47136067	Evi-1 (-)	Celera (hCV2880884)	32465634
4	1f-C-1562T	T (30)	CATATTTATCC/TTTTTATA**C/T**TTCTTTCAAG	47136059		Celera (hCV26069888)	32465635
5	1f-G-1344A	A (37)	TTAAACAGAT**G/A**TATCATCATC	47135841	GATA (+)	Celera (hCV2880883)	32465636
6	1f-T-1198G	G (20)	CAGGCTCAGC**T/G**GCCCTTTGAC	47135695	AP4 (+)	dbSNP (ss5759311, rs4237856), Celera (hCV27909097)	32465637
7	1f-G-777A	A (13)	CCAGGTAGGC**G/A**TGGTGCCCCC	47135107		ss20420964*	32465638
8	1f-G-465C	C (37)	TCCTCCCTGG**G/C**CAAGCCATCT	47134962		dbSNP (ss5452989, rs4073726), Celera (hCV27537612), ss20420962*	32465639
9	1f-C-217T	T (30)	CCCCCCTGCT**C/T**CCTCATGGCC	47134714		dbSNP (ss6198081, rs4073729), Celera (hCV27909106), ss20420961*	32465640
	Exon 1e						
10	1e-C-2090T	T (23)	ACTGGGATTA**C/T**AGGCCTGAGC	47100603	Cdx-A (+)	Novel	32465641
11	1e-A-2054G	G (7)	GATTTTTAAT**A/G**CTGTATTTTG	47100567		Novel	32465642
12	1e-G-1739A	A (37)	CTAGGTCACA**G/A**TAAAAACTTA	47100296	Cdx-A (-)	Cdx-2 [7], Celera (hCV2880808), ss20420955*	32465643
13	1e-A-1553G	G (3)	AGGAAATATA**A/G**ATGAATTAAT	47100066	Cdx-A (-)	Novel	32465644
14	1e-C-1092T	T (23)	TGAGGACAGG**C/T**TGCAGTCCTG	47099649		Celera (hCV2880807)	32465645

15	1e-C-577A	A (23)	CTCTGTAGTG**C/A**AGGAATGGTG	47099134		Celera (hCV2880806), ss20421064*	32465646
16	1e-19 $(T)_{9\text{-}11}$	$(T)_9$ (20) $(T)_{10}$ (50) $(T)_{11}$ (30)	TGCATGGAAA**(T)$_{9\text{-}11}$**GTTTTTTAGA	47098532		Novel	32465647
	Exon 1a						
17	1a-G-1521C	C (23)	GCTTTCCCAC**G/C**ATGCTTTGGG	47098041	GATA-1 (+)	ss20420954*	32465648
18	1a-A-1012G	G (23)	GAATAGCAATA**/G**TCTTCCCTGG	47097576	GATA-1, 2, 3 (-)	Celera (hCV2880805), ss20420953*	32465649
	Exon 1d						
19	1d+77 $(T)_{23\text{-}28}$	$(T)_{23}$ (13) $(T)_{24}$ (40) $(T)_{25}$ (33) $(T)_{26}$ (7) $(T)_{27}$ (3) $(T)_{28}$ (3)	TTTTAAGTAA**(T)$_{23\text{-}28}$**ACTTTTCTTT	47095974		Novel	32465818
20	1d+425 $(TCC)_{3/2}$	$(TCC)_2$ (13)	TTCTCACAGT**(TCC)$_{3/2}$**T$(TCC)_{7/8}$TCTTGTTCCT	47095624		Novel	32465819
21	1d+435 $(TCC)_{7/8}$	$(TCC)_8$ (47)	TTCTCACAGT$(TCC)_{3/2}$T**(TCC)$_{7/8}$**TCTTGTTCCT	47095612		Novel	32465650
	Exon 1b						
22	1b-T-2746C	C (23)	AACTCGCTGT**T/C**GCCTGTGCAG	47094187		dbSNP (ss6566285, rs4760658), Celera (hCV27904554)	32465651
23	1b-G-2528A	A (37)	GCCCTCTGAG**G/A**GATAGCAGGA	47093969		ss20420951*	32465652
24	1b-C-2481A	A (13)	GACTCCAAGA**C/A**AGTTTGAGCC	47093922		Novel	32465653
25	1b-A-2225G	G (37)	GATGCCATGC**A/G**TGGTTATACC	47093666		Celera (hCV2880804), ss20421063*	32465654
26	1b-T-1748A	A (40)	TACAGCCTCC**T/A**CATTTTGCTC	47093189		Celera (hCV2880803)	32465655
27	1b-G-1641C	C (3)	GGCAAACCTC**G/C**CCCACTCTGA	47093082	MZF1 (-)	Novel	32465656
28	1b-G-886A	A (37)	CAGCTGCACC**G/A**GCGGGAAAGC	47092327		Celera (hCV26014291), ss20421061*	32465657
29	1b-C-673T	T (30)	CAGCTACTTA**C/T**TGAGCACTCA	47092114		Novel	32465658
30	1b-T-391C	C (40)	ATTATACACC**T/C**GCAGTAAATG ATTATACACC**T/C**GCAGTAAATG ATTATACACC**T/C**GCAGTAAATG	47091832	E47 (-) deltaEF1 (+) USF (+)	ss20421060*	32465659
31	1b-C25A (Thr/Lys)	A (23)	CCCAGCTGGA**C/A**GGAGAAATGG	47091409		ss20421059*	32465660

	Exon 1c						
32	1c-C-2040T	T (23)	CCTGCCATCC**C/T**TTGGCCTGGG	47076284		ss20421046*	32465661
33	1c-T-1930C	C (27)	AAGGTGTCCA**T/C**GGCTTAGGGT	47076174		dbSNP (ss4040170, rs2853564), Celera (hCV15823889), ss20420934*	32465662
34	1c-G-1633C	C (27)	GCCTAGCTGT**G/C**GGACCCTGGG	47075877		dbSNP (ss3193006, rs2238135), Celera (hCV16031778), ss20420933*	32465663
35	1c-C-1453T	T (27)	TGGTTGTCTA**C/T**CTGGATGTCA TGGTTGTCTA**C/T**CTGGATGTCA	47075558	GATA-1, 2 (-) AREB6 (+)	dbSNP (ss3193007, rs1989969), Celera (hCV12060044), ss20420932*	32465664
36	1c-G-1156A	A (27)	CCCAGCTTAG**G/A**TTATCTTGGC	47075261		dbSNP (ss3193008, rs2238136), Celera (hCV3290655), ss20421045*	32465665
	Coding exons and flanking intron region						
37	E2-G-51A	A (3)	ATGCCAGCTG**G/A**CCCTGGCACT	47070636		Novel	32465666
38	E2-C4T (codon 1)	T (40)	CTTACAGGGA**C/T**GGAGGCAATG	47070443		dbSNP (ss16359158, rs2228570), Fok I (Saijo, 1991), HGMD (CM972826), Celera (hCV12060045), ss20420928*	32465667
39	E2-C59T (codon 19)	T (7)	TTGACCGGAA**C/T**GTGCCCCGGA	47070527		Exon 2, C57T (Brown S., 1999), ss20420927*	32465668
40	E4-C-71T	T (27)	ACCTTTACCC**C/T**CAACCGCAA/GGAGGAA	47049093		Celera (hCV3290629)	32465669
41	E4-A-62G	G (33)	TTACCCC/TCAACCGCA**A/G**GAGGAAGGTT	47049084		Celera (hCV3290628), ss20421022*	32465670
42	E5-A+118T	T (7)	AACTTACATA**A/T**ATACTGTGCG	47048463		ss20420902*	32465671
43	E6-G+154A	A (33)	GTGCAGTGGC**G/A**CGATCTCGGC	47046925		Novel	32465672
44	E7-D+75G	G (13)	TGGGGTTTGG**-/G**CTCCAATCAG	47038024		Intron 7, G-insertion (Brown S., 1999)	32465673

45	E8-C2T (codon 303)	T (3)	CTCTCACAGC**C/T**GGACACAGCC	47037891		ss20420899*	32465674
#	E8-G+284A (Bsm I)	A (41)	ACAGGCCTGC**G/A**CATTCCCAAT	47038290		rs1544410	-
46	E9-G-111C	C (33)	TAGAGGGGTG**G/C**CCTAGGGGGT	47036557		Novel	32465675
47	E9-G-94A	A (13)	GGGTGCTGCC**G/A**TTGAGTGTCT	47036540		Novel	32465676
48	E9-T-48G (Apa I)	G (33)	AGCAGTGAGG**T/G**GCCCAGCTGA	47036493		dbSNP (ss20420896, rs7975232), Apa I [11], ss20420896*	32465677
49	E9-T32C (Codon 352, Taq I)	C (33)	CCGCGCTGAT**T/C**GAGGCCATCC	47036308		dbSNP (ss20420895, rs731236), Taq I [12], HGBASE (SNP000008184), Celera (hCV2404008), ss20420895*	32465678
	3'-UTR						
50	U-A311C	C (33)	CCACCGCTGC**A/C**TAAGTGGCTG	47035772		dbSNP (ss87853, rs739837), C1588A [10], Celera (hCV2404007), ss20420894*	32465679
51	U-C440G	G (17)	TCTCCCTCTC**C/G**TGCCTACTCA	47035748		ss20420893*	32465680
52	U-G464T	T (33)	TAAATAATCG**G/T**CCCACAGCTC	47035619		dbSNP (ss87854, rs3847987), Celera (hCV2404006), ss20420892*	32465681
53	U-D796T	T (30)	GTCCTCCCCC**-/T**GCCAGTGCCT	47035391		T2074del [10], ss20421003*	32465682
54	U-T1229C	C (3)	AAGTGCATGC**T/C**CTCTGCAGCC	47034958	in destabilizing elements	ss20421001*	32465683
55	U-1238(CCAGC)$_{3/4}$	(CCAGCC)$_4$ (13)	GCATGCT/CCTCTGCAG**(CCAGCC)$_{3/4}$** TGGTGGGAAG	47034947	in destabilizing elements	AGCCC2517del [10]	32465820
56	U-G1868A	A (7)	TTCAGTGGGA**G/A**AAAACACTTG	47034316		Novel	32465684
57	U-A1909C	C (30)	TCCCCTCATT**A/C**AGGAAAACTG	47034174		C3185A [10], Celera (hCV8716058), HGBASE (SNP000016363), ss20420891*	32465685

58	U-G1982C	C (7)	AGGACAGGCC**G/C**GGCGCGGTGG	47034101		Celera (hCV3290620), ss20420890*	32465686
59	U-2119$(A)_{13-24}$	$(A)_{13}$ (20) $(A)_{16}$ (17) $(A)_{18}$ (13) $(A)_{19}$ (17) $(A)_{23}$ (27) $(A)_{24}$ (7)	TTAAAAATAC**$(A)_{13-24}$**TAGCCGGGCA	47033958		Poly (A) (Ingles, 1997), Celera (hCV26075087),	32465687
60	U-T2147A	A (30)	TGGTGGCGCA**T/A**GCCTGTAATC	47034035		A3424T [10], ss20420889*	32465688
61	U-G2795A	A (13)	GGGAGAAAAG**G/A**TCATCATCGA	47033289	in destabilizing elements	dbSNP (ss4040169, rs2853563), A4107G [10], Celera (hCV15823846), ss20420888*	32465689
62	U-A2978T	T (13)	AATCCAAGCG**A/T**GGTCAACAGA	47033201	in destabilizing elements	ss20420887*	32465690

Note:

A. The nomenclature of polymorphism is based on requencing. The first position of the polymorphism ID is designated as exon (coding exons in capital) or 3'-UTR (U), the major nucleotide allele (for Caucasian) is following a dash, the nucleotide location is designated as plus (+) or minus (-) relative to the first or last base of the nearest exon, the location without (+) or (-) is in the exon, the last position of the polymorphism ID is the minor nucleotide allele (for Caucasian).

B. Bold letter indicates the variation, and underlined nucleotides indicate a potential transcription factor consensus sequence. For multi-allelic polymorphisms, we list all observed alleles

C. The location and access codes of SNPs are based on the April 2005 freeze of the Celera database and National Center for Biotechnology Information (NCBI) dbSNP build 123, *: SNPs were reported by Nejentsev et al., (Hum. Mol Genet., 2004)

D. The potential transcription factors are described, (+) indicates the sense recognition sequence and (-) indicates the antisense recognition sequence

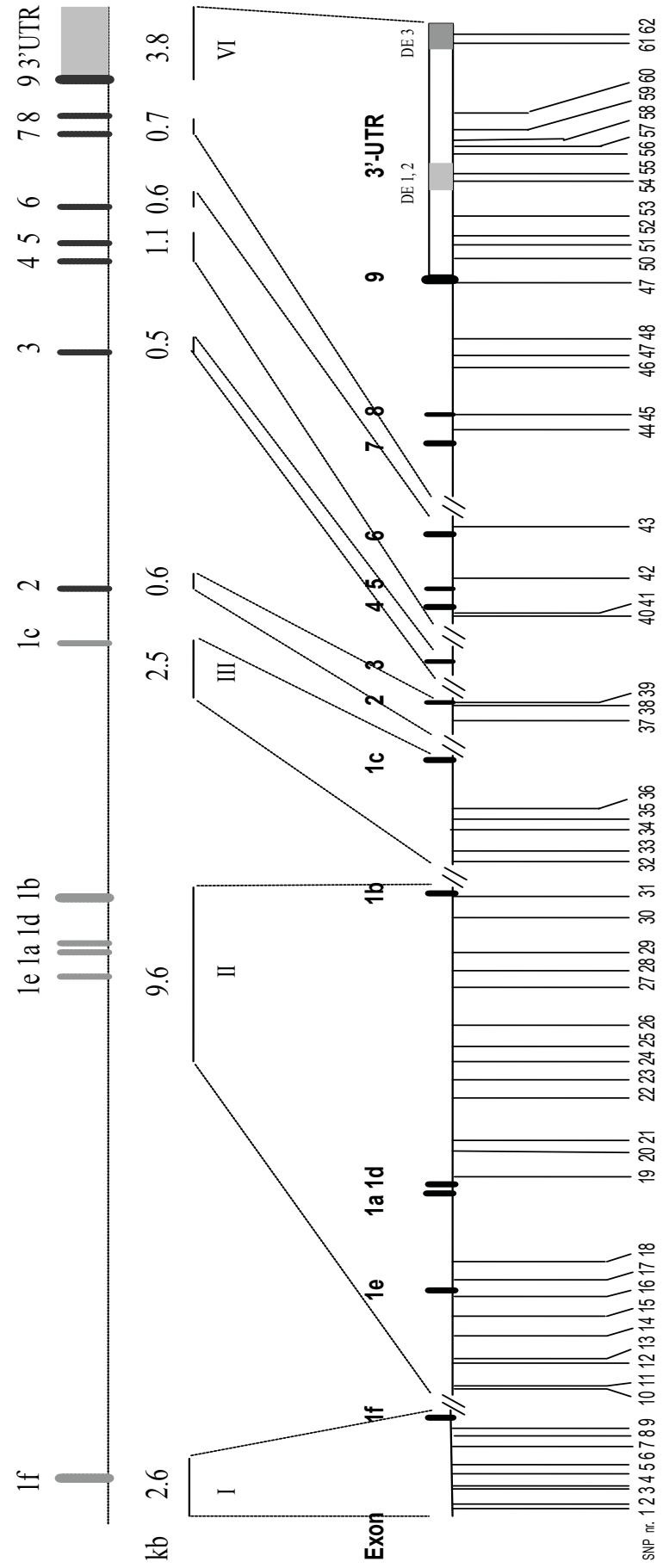


Figure 3 Sequenced areas and position of the 62 variations. Grey boxes in the 3'-UTR indicate destabilizing elements (DE1, 2 &3; Durrin et al.[10]). Numbers of the sequence variations refer to Table 2.

Since the human VDR gene is a complex (with six extra exons in 5'-end) and large gene (at least 105 kb), it was a time and money consuming design to sequence the complete VDR gene for functional variations. Especially in the beginning of this study, no high-throughput sequencing equipment was available. The homology analysis was therefore used to identify potentially functional regions of the VDR gene, we therefore sequenced in total 22 kb (instead of 105 kb) of the VDR gene. Those fragments covered most potentially functional region: all coding exons, the complete 3'-UTR and 2 kb promoter region of all extra exons in 5'-end.

One limitation of our homology analysis is we only compared the VDR sequence between human and mouse, because in the beginning of our study (1999), only human and mouse VDR sequences were available. Now more species genomic sequences are available for the homology analysis to identify commonly conserved fragments among those species.

The analyses in this chapter (determining the VDR structure, homology analysis and sequencing analysis) are time-consuming (cost 2.5 years) but they are the fundamental steps for our following studies.

ACKNOWLEDGEMENTS

We thank Drs. John Eisman and Linda Crofts for helpful sharing of unpublished sequence information, Dr. Mark Haussler for kindly providing VDR cDNA.

This project was funded by the Dutch Research Organisation (NWO 903-46-178, 925-01-010, 014-90-001 and 911-03-012).

REFERENCES

1. Haussler, M.R. et al. The nuclear vitamin D receptor: biological and molecular regulatory properties revealed. *J Bone Miner Res* **13**, 325-49 (1998).
2. Baker, A.R. et al. Cloning and expression of full-length cDNA encoding human vitamin D receptor. *Proc Natl Acad Sci U S A* **85**, 3294-8 (1988).
3. Miyamoto, K. et al. Structural organization of the human vitamin D receptor chromosomal gene and its promoter. *Mol Endocrinol* **11**, 1165-79 (1997).
4. Crofts, L.A., Hancock, M.S., Morrison, N.A. & Eisman, J.A. Multiple promoters direct the tissue-specific expression of novel N-terminal variant human vitamin D receptor gene transcripts. *Proc Natl Acad Sci U S A* **95**, 10529-34 (1998).
5. Uitterlinden, A.G. et al. A large-scale population-based study of the association of vitamin D receptor gene polymorphisms with bone mineral density. *J Bone Miner Res* **11**, 1241-8 (1996).
6. Vista. http://www-gsd.lbl.gov/VISTA/index.html.
7. Yamamoto, H. et al. The caudal-related homeodomain protein Cdx-2 regulates vitamin D receptor gene expression in the small intestine. *J Bone Miner Res* **14**, 240-7 (1999).
8. Tfsearch. http://www.cbrc.jp/research/db/TFSEARCH.html.
9. Genomatix. http://www.genomatix.de/cgi-bin/./matinspector_prof/mat_fam.pl.
10. Durrin, L.K., Haile, R.W., Ingles, S.A. & Coetzee, G.A. Vitamin D receptor 3'-untranslated region polymorphisms: lack of effect on mRNA stability. *Biochim Biophys Acta* **1453**,

311-20 (1999).
11. Morrison, N.A., Yeoman, R., Kelly, P.J. & Eisman, J.A. Contribution of trans-acting factor alleles to normal physiological variability: vitamin D receptor gene polymorphism and circulating osteocalcin. *Proc Natl Acad Sci U S A* **89**, 6665-9 (1992).
12. Morrison, N.A. et al. Prediction of bone density from vitamin D receptor alleles. *Nature* **367**, 284-7 (1994).

Chapter 3

Linkage Disequilibrium (LD) Structure of the Vitamin D Receptor (VDR) Gene

This chapter represents a part of the manuscript:
Fang Y, van Meurs JB, d'Alesio A, Jhamai M, Zhao H, Rivadeneira F, Hofman A, van Leeuwen JP, Jehan F, Pols HA, Uitterlinden AG
Promoter and 3'-Untranslated-Region Haplotypes in the Vitamin D Receptor Gene Predispose to Osteoporotic Fracture: The Rotterdam Study
Am J Hum Genet (2005) 77:807-823

ABSTRACT

Introduction: After 62 polymorphisms across the VDR gene were detected, the next questions were: a. what is the allelic association between them [i.e., linkage disequilibrium (LD)], b. how many functional polymorphisms were among them, and c. how to identify them. Haplotype and LD analysis is a powerful tool to examine the relationship between polymorphisms in a gene. This analysis can cluster polymorphisms as haplotype blocks according to linkage level, and determine haplotype tagging polymorphisms for association and functionality studies to identify potentially functional regions or polymorphisms in a gene.

Materials and Methods: We genotyped 47 single nucleotide polymorphisms (SNPs) in 234 Caucasians, 107 Asians and 58 African Americans. We analysed several databases and determined the race-specific haplotype structure and haplotype tagging SNPs (htSNPs) across the whole VDR gene for different ethnic population.

Results: LD analysis of common SNPs (frequency $\geq$ 5 %) revealed 4 – 8 high linkage haplotype blocks (D' > 0.8) which are conserved among Caucasians and Asians, but more fragmented in Africans. The haplotype frequencies differ extensively among the ethnic groups. Fifteen htSNPs are diagnostic for the VDR haplotype blocks for large-scale association studies in Caucasians, 10 htSNPs for studies in Asians and 28 for studies in Africans.

Conclusions: These htSNPs can be used for further studies to investigate the relationship between VDR haplotypes and clinical outcomes in different ethnic populations.

INTRODUCTION

The interpretation of polymorphic variations in the VDR gene is severely hindered by the fact that until now only few polymorphisms in this large gene have been studied, and that most of these are anonymous (non-functional) polymorphisms. To explain the associations observed with complex diseases they should be in linkage disequilibrium (LD) with truly functional polymorphisms. Haplotype-based methods offer a powerful approach to study association of genetic variation with complex disease[1]. Following a recent comprehensive study of VDR polymorphisms[2] we here present a detailed description of the genomic organisation of the VDR gene region, the identification of 62 polymorphisms across relevant areas of the gene, and analysis of LD and haplotype diversity of VDR variations in different ethnic groups. We also compared our haplotype data with data from Nejentsev et al. and data in several databases including Perlegen, HapMap and SNPbrowser.

MATERIALS AND METHODS

Subjects

We studied LD in 234 random blood Caucasian bank donors, and in DNA from 107 Asian and 58 African individuals (the Coriell Institute, Camden, NJ. USA: 90 Chinese Han (HD100), 9 Chinese (HD02), 8 Japanese (HD07), 47 African Americans [HD04 and HD50) and 9 Africans from south of the Sahara (HD12)].

Genotyping

We genotyped 47 SNPs (Table 1) in 3 ethnic groups with the high throughput TaqMan allelic discrimination assays. The Assay Mixes (including unlabelled PCR primers, FAM™ and VIC® dye-labelled TaqMan MGB probes) of three Assays-on-Demand℠ and 36 Assays-by-Design™ were designed and provided by ABI. The reaction system contained 1 - 5 ng of dried genomic DNA, 2.5 μl of TaqMan® Universal PCR Master Mix, No AmpErase® UNG (2 ×), 0.125 μl (40 ×) or 0.0625 μl (80 ×) of Assay Mix, and adjusted with Milli-Q H_2O to a total volume of 5 μl. The genotyping results were analyzed with an endpoint reading in the ABI Prism® 7900HT independently by two operators, and 5% random samples were independently repeated to confirm genotyping results.

Linkage disequilibrium (LD) and haplotype analyzes

We first analyzed LD among all 62 polymorphisms based on the 15 sequenced Caucasian samples to select 37 SNPs, and also nine additional SNPs (A to I in Table 1) from the Celera database, which are flanking the VDR gene, and the E8-G+284A (Bsm I RFLP[3]). Selection criteria for VDR SNPs in the LD analysis is based on: A. Minor allele frequency (MAF) > 10 % (3/30 alleles) from our sequence analysis;

Table 1 Allele frequencies of 47 SNPs in different ethnic groups

Code [I]	SNPs ID [II]	Allele	Minor allele frequency (%) [III]		
			Caucasian (# chr = 468)	Asian (# chr = 214)	African (# chr = 116)
	Intergenic region of COL 2A1 - VDR (5 SNPs)				
A	hCV2615323 (G/A)	A	10	38	4
B	hCV8724908 (G/C)	C	44	15	27
C	hCV2626648 (G/A)	A	11	33	4
D	hCV8724903 (C/G)	G	7	1	5
E	hCV8724902 (C/T)	T	12	34	28
	Promoter (24 SNPs)				
1	1f-G-1904A	A	46	5	10
3	1f-C-1570T	T	27	78	66*
5	1f-G-1344A	A	46*	5*	10*
6	1f-T-1198G	G	28*	17*	12*
7	1f-G-777A	A	13*	29*	23*
F	1e-T-3743C	C	17	48	76
10	1e-C-2090T	T	43	3	6*
12	1e-G-1739A (Cdx-2)	A	17*	49*	75*
15	1e-C-577A	A	42	3	10*
17	1a-G-1521C	C	43	3	6
18	1a-A-1012G (GATA)	G	43*	3	6*
22	1b-T-2746C	C	34	3	13*
23	1b-G-2528A	A	28	66	75*
24	1b-C-2481A	A	10*	0.0	3
25	1b-A-2225G	G	27	64	74
26	1b-T-1748A	A	38	31	10
28	1b-G-886A	A	25*	44	9*
29	1b-C-673T	T	10*	0	2
30	1b-T-391C	C	39	31*	10*
31	1b-C25A (Thr/Lys)	A	32*	3	12*
33	1c-T-1930C	C	40	34	14
34	1c-G-1633C	C	28*	20*	30*
35	1c-C-1453T	T	41*	33*	51*
36	1c-G-1156A	A	29	17	11*
	Coding (8 SNPs)				
38	E2-C4T (Fok I)	T	34*	51*	21*
41	E4-A-62G	G	42	75	35*
44	E7-D+75G	G	2	1	22*
#	E8-G+284A (Bsm I)	A	42	6	36*
46	E9-G-111C	C	13	17	10
47	E9-G-94A	A	2	1	14*
48	E9-T-48G (Apa I)	G	44	76	26
49	E9-T32C (Taq I)	C	43	8	31*
	UTR (7 SNPs)				
50	U-A311C	C	45*	76*	31*
51	U-C440G	G	2	0.5	1
52	U-G464T	T	13*	17*	10*
53	U-D796T	T	44	24	59*
57	U-A1909C	C	44	75	29*
61	U-G2795A	A	2	1	16*
62	U-A2978T	T	2*	0.5	1
	Intergenic region of VDR HDAC7A (3 SNPs)				
G	hCV3290614 (A/T)	T	35	75	23
H	hCV3290610 (T/C)	C	46	64	42
I	hCV16253844 (T/C)	T	46	36	33

I: The SNP codes from A to I indicate SNPs which are selected from the Celera database, SNP A is a mis-sense mutation in exon 53 of COL 2α1; # indicates the RFLP SNP, E8-G+284A (Bsm I)[3]. The number is according to Table 2 in chapter 2.
II: The SNP ID is according to the Celera database and the resequencing result in Table 2 in chapter 2.
III: Minor alleles are based on the frequency in the Caucasian population.
*: VDR Tagging SNPs for association study in different ethnic groups

B. In potential promoter TFBS or destabilising element in 3'-UTR; C. in a highly conserved region; D. a tagging SNP (based on the 15 sequenced subjects).

We then genotyped 47 SNPs in Caucasians, Han-Chinese and African-Americans to calculate allele frequencies (Table 1). We determined the race-specific SNPs, whose MAF is > 3% in either of the ethnic study populations, and identified 42 SNPs for Caucasian, 33 for Asian and 41 for African American. We constructed haplotype structure by the PHASE program[4], then used PHASE outcome to calculate the pair- wise standardized disequilibrium coefficient (D') with the "haploxt"[5] to estimate the linkage magnitude between two SNPs, and depicted the graphic overviews of LD by the GOLD program[6]. We identified haplotype blocks and calculated haplotype frequency in each block by HaploBlockFinder program[7]. Selection of the htSNPs in each ethnic group was based on: A. the minimal combination of htSNPs in each haplotype block to represent ≥ 95% of the haplotypes; B. potential functionality in 3'-UTR or promoter; C. SNP is unlinked to any block e.g., E2-C4T (Fok I).

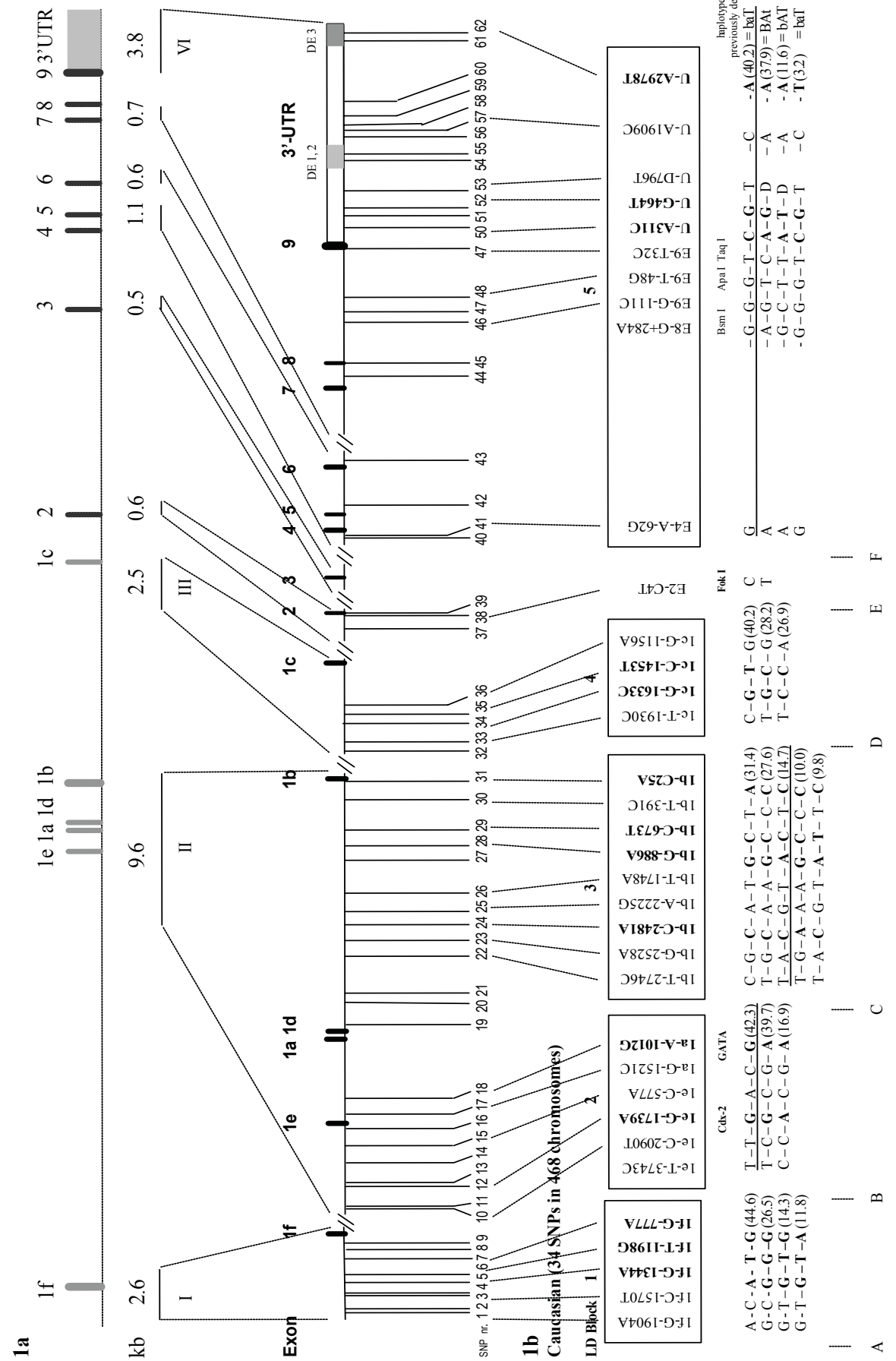
LD data comparison

We compared our data with the data from Nejentsev et al.[2] (including 68 SNPs for Caucasians), and data from Perlegen[8] (including 48 SNPs for European American), HapMap[9] (Haploview version 3.0, including 33 SNPs) and SNPbrowser[10] (version 3.0, including 11 SNPs)

RESULTS

Linkage disequilibrium (LD) and haplotype analyzes

Genomic structure and polymorphisms of the VDR gene were presented in Figure 1a. In Caucasians, we identified 5 blocks with high LD (coded 1 to 5 in Fig. 1b and 2a) which range in size from 2 - 17 kb. SNPs in such a block are in strong LD with the other SNPs inside the block but show very little LD with SNPs outside the block. Four blocks (1-4) were found in the promoter region, while the largest block 5 is encompassing 17 kb and includes exons 4-9 and the 3'-UTR. Block 2 and 3, encompassing the 1b-1e promoter region, could not be considered as one LD block in our analyses even when D' > 0.50 is taken as cut-off to define a block. There are also clear areas of very low (or absent) LD, i.e., between exons 1e and 1f (block 1 and 2), between exons 2 and 3 (block 4 and 5), and at the end of the VDR gene (3' of block 5). The E2-C4T SNP (detected as a Fok I RFLP) has no LD with any of the other SNPs and cannot be assigned to any of the blocks. The most distal 3' haplotype block 5 shows no LD with SNPs after the VDR gene and the most proximal 5' block 1 shows only weak LD with more 5' SNPs near and in the COL2A1 gene. At least 7 blocks can be identified when we compared different sources on VDR LD block structures (Fig. 2b). It is difficult to define the exact boundaries of these


1a
1f
1e 1a 1d 1b
1c 2
3 4 5 6 7 8 9 3'UTR
kb 2.6 9.6 2.5 0.6 0.5 1.1 0.6 0.7 3.8
I II III VI
Exon 1f 1e 1a 1d 1b 1c 2 3 4 5 6 7 8 9 3'-UTR
DE 1, 2
DE 3
SNP nr. 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 50 51 52 53 54 55 56 57 58 59 60 61 62
1b
Caucasian (34 SNPs in 468 chromosomes)
LD Block 1 2 3 4 5
1f-G-1904A
1f-C-1570T
1f-G-1344A
1f-T-1198G
1f-G-777A
1e-T-3743C
1e-C-2090T
1e-G-1739A
1e-C-577A
1a-G-152IC
1a-A-1012G
1b-T-2746C
1b-G-2528A
1b-C-2481A
1b-A-2225G
1b-T-1748A
1b-G-886A
1b-C-673T
1b-T-391C
1b-C25A
1c-T-1930C
1c-G-1633C
1c-C-1453T
1c-G-1156A
E2-C4T
E4-A-62G
E8-G+284A
E9-G-111C
E9-T-48G
E9-T32C
U-A3111C
U-G4647T
U-D796T
U-A1909C
U-A2978T
Cdx-2
GATA
Fok I
Bsm I Apa I Taq I
haplotype allele previously defined[11]
A-C-A-T-G (44.6) T–T–G–A–C–G (42.3) C–G–C–A–T–G–C–T–A (31.4) C–G–T– G (40.2) C – G – G – T – C – A – G – D – C – A (40.2) = baT
G-C-G-G-G (26.5) T–C–G–C–G–A (39.7) T–G–C–A–A–G–C–C–C (27.6) T–G–C– G (28.2) T – G – C – A – G – D – A – A (37.9) = BAt
G-T-G-T-G (14.3) C–C–A–C–G–A (16.9) T–A–C–G–T–A–C–T–C (14.7) T–C–C– A (26.9) G – C – T – T – A – T – D – A – A (11.6) = bAT
T–G–A–A–A–G–C–C–C (10.0) G – G – G – T – C – G – T – C – C – T (3.2) = baT
T–A–C–G–T–A–T– T–C (9.8)
A B C D E F
A

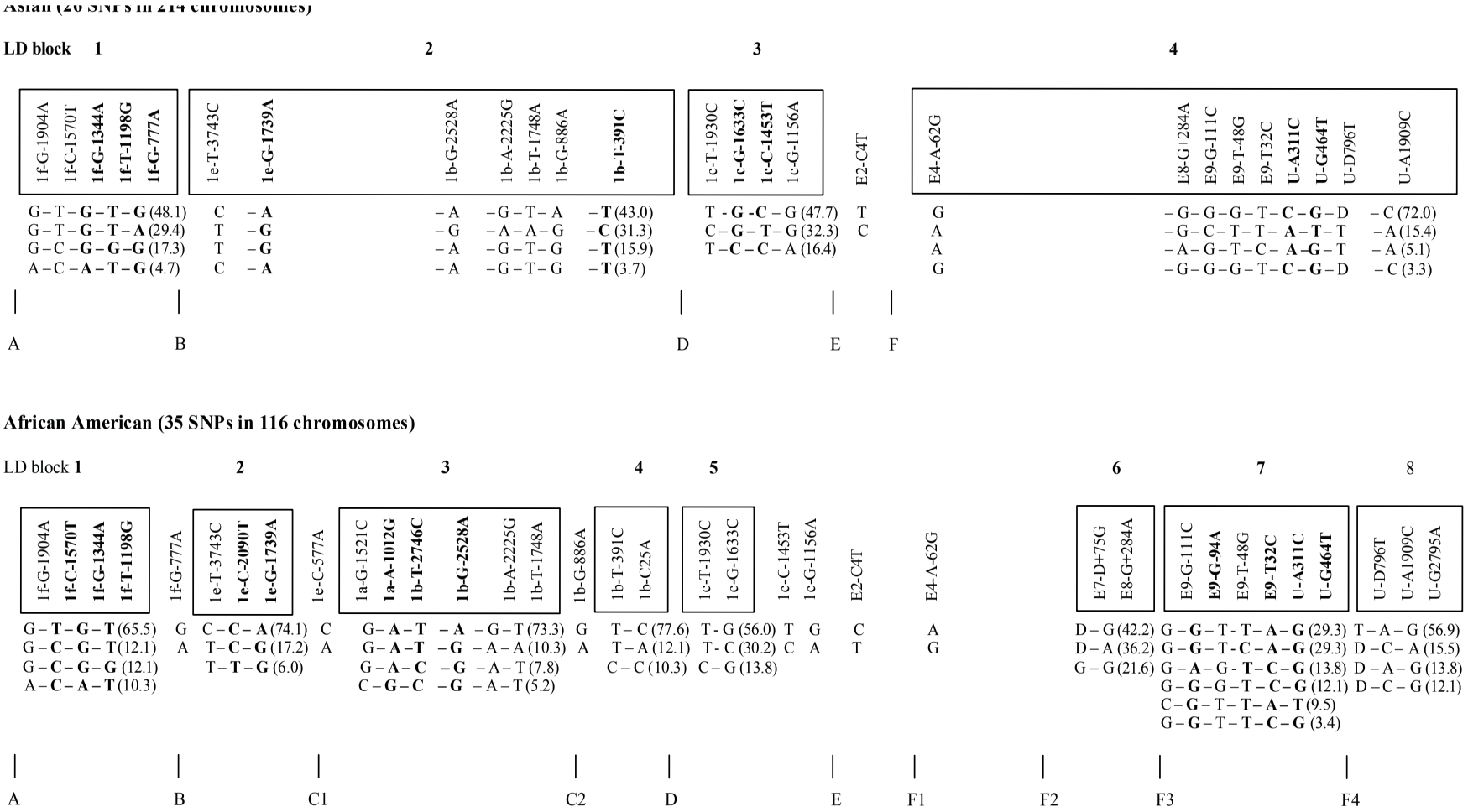


Figure 1 LD map of the human VDR gene **1a** Sequenced areas and position of the 62 variations. Grey boxes in the 3'-UTR indicate destabilizing elements (DE1, 2 &3; Durrin et al.[13]). Numbers of the sequence variations refer to Table 2 in Chapter 2. **1b** Haplotype map of the VDR gene in Caucasians, Asians and African-Americans, based on SNPs with a MAF of ≥ 5% in each of the different ethnic populations. Common haplotype alleles in each block with a frequency >3% are presented below the blocks. Bold SNPs and alleles indicate the haplotype tagging SNPs (htSNPs). Fracture risk haplotype alleles (described in chapter 5.1) are underlined. Correspondence to Caucasians for our previous Bsm-Apa-Taq haplotype allele definition[11] in block 5 is shown.

low LD areas because not all studies analyzed a high density of SNPs across the VDR gene. Interestingly, we observed LINE-1 repetitive elements to be localized between the boundaries of the LD blocks of the VDR gene (Fig. 2b).

The LD map of Han Chinese is similar to that of Caucasians (Fig. 1b and 3a), but with only 4 haplotype blocks and with areas of low LD co-localizing with those in Caucasians. For Asians we observed a large haplotype block in the promoter region (from 1e-T-3743C to 1b-T-391C) which corresponds to the combined Block 2 and 3 in Caucasians. This larger promoter block was also observed in Caucasians when we used SNPs with MAF ≤ 5% (data not shown). The LD map of African-Americans is more fragmented and substantially different compared with that of Caucasians and Asians (Fig. 1b and 3b). There are eight small haplotype blocks and seven unlinked SNPs. Again, the E2-C4T SNP (the Fok I RFLP) was found to be independent of any other haplotype block. Some areas of low LD are co-localizing in all three ethnic groups, indicated as A in Figure. 1b (separating exon 1f from the upstream region into COL2A1), B (separating exon 1e-1b from exon 1f), D (separating exon 1c from 1e-1b), E (separating exon 2 from exon 3-9/3'-UTR), and F (separating the exon 4 - 3'-UTR from the area downstream of VDR), while C (separating exon 1f from exon 1e-1b) is shared between Caucasians and African-Americans, but is absent in Han Chinese. In African-Americans two additional areas of low LD (C1 and C2) can be distinguished.

We reconstructed haplotype alleles and analyzed diversity and frequencies across ethnic groups (Fig. 1b; only haplotypes with MAF > 3% are shown). The number of haplotype alleles in each of the LD blocks increases from Han Chinese to Caucasians and is the greatest in African Americans. Only for relatively "conserved" LD blocks across ethnic groups comparison is possible and this shows substantial differences in haplotype allele frequencies. For example, in block 1 (around exon 1f) the most common allele in Caucasians has a frequency of 44.6% but this same allele is 4.7% in Han Chinese and 10.3% in African-Americans. Similarly, the most common allele in block 4 (around exon 1c) is 40.2% in Caucasians, 32.3% in Han Chinese, and 13.8% in African Americans.

While for Caucasians 15 tagging SNPs (14 htSNPs and E2-C4T) are required to cover the common genetic diversity across the VDR gene, this number is only 10 for Han Chinese but increases to 28 for African Americans. When only the tagging SNPs were used to genotype a population (6,148 subjects form the Rotterdam Study), a similar LD pattern was obtained compared to using the 34 original SNPs with MAF > 5% in the smaller sample of 234 bloodbank donors, indicating that the tagging SNPs are effectively predicting the other SNPs within the blocks.

We observed that the haplotypes constructed from the well-known Bsm I, Apa I, and Taq I RFLPs, as described previously by our group[11], could correctly predict all the common haplotypes at the 3'-UTR region in Caucasians as defined using the 9 SNPs with MAF > 5% (Fig. 1b). For Han Chinese such correspondence is lower due to the U-D796T polymorphism which has changed phase, while for African

Americans the haplotype structure of this area is much different.

DISCUSSION

We recently identified 62 polymorphisms in potentially functional areas of the VDR gene (chapter 2), and now characterised the LD structures across the whole VDR gene for three different ethnic populations. For Caucasian population we identified 15 tagging SNPs to represent the common haplotypes of these blocks in potentially functional areas of the VDR gene, we also combined several studies and databases[2,8-10,12] to conclude that there are 7 haplotype blocks across the VDR gene in Caucasians.

We analysed the LD and haplotype structure of the VDR gene using 49 SNPs with MAF > 5% in 234 subjects, and identified 5 haplotype blocks across the gene for Caucasian population (Fig 1b), haplotype pattern does not change when we selected 50 random subjects in the same analysis. Nejentsev et al.[2], analysed 68 SNPs with MAF > 10% in 916 subjects. Their analysis has the similar LD map of the gene in Caucasians, although they described three haplotype block including subset blocks. However, they missed one block structure around exon 1b, because only one SNP was selected in the region; and since not analysing the intron2 and 3, we missed the block(s) in that region. Perlegen[8] database defines three blocks with high density SNPs (1 kb per SNP on average) in 23 Caucasians. HapMap[9] and SNPbrowser[10] databases show some small haplotype fragments in the gene, partially describe the LD map of the VDR gene. Combining all data together, we conclude that there are seven haplotype blocks across the VDR gene. To determine the reliable haplotype structure for the association study, one has to analyse high density of informative SNPs (1 – 2 kb per SNP and MAF > 5%) in the relatively big study population (> 50).

Using the D' threshold of 0.8, we gained five clear border blocks with a few common haplotypes (MAF > 5%) which cover > 95% haplotypes in the association population. Using a few common haplotypes in association study, we limited the number of analysis against the multiple testing.

We constructed the LD map and identified htSNPs of the VDR gene for three major ethnic populations. There are several further approaches to identify the (set of) responsible functional SNPs in the LD areas where we identified risk haplotype alleles. Since African Americans have in general smaller LD blocks, replicating the associations in African Americans could "zoom in" on a smaller region with concomitantly less SNPs to test functionally. Subsequently, the functionality test we will show (Fig. 2 and 5 in chapter 7) can be repeated but now with constructs of promoter and 3'-UTR variants that differ at less and down to only 1 position.

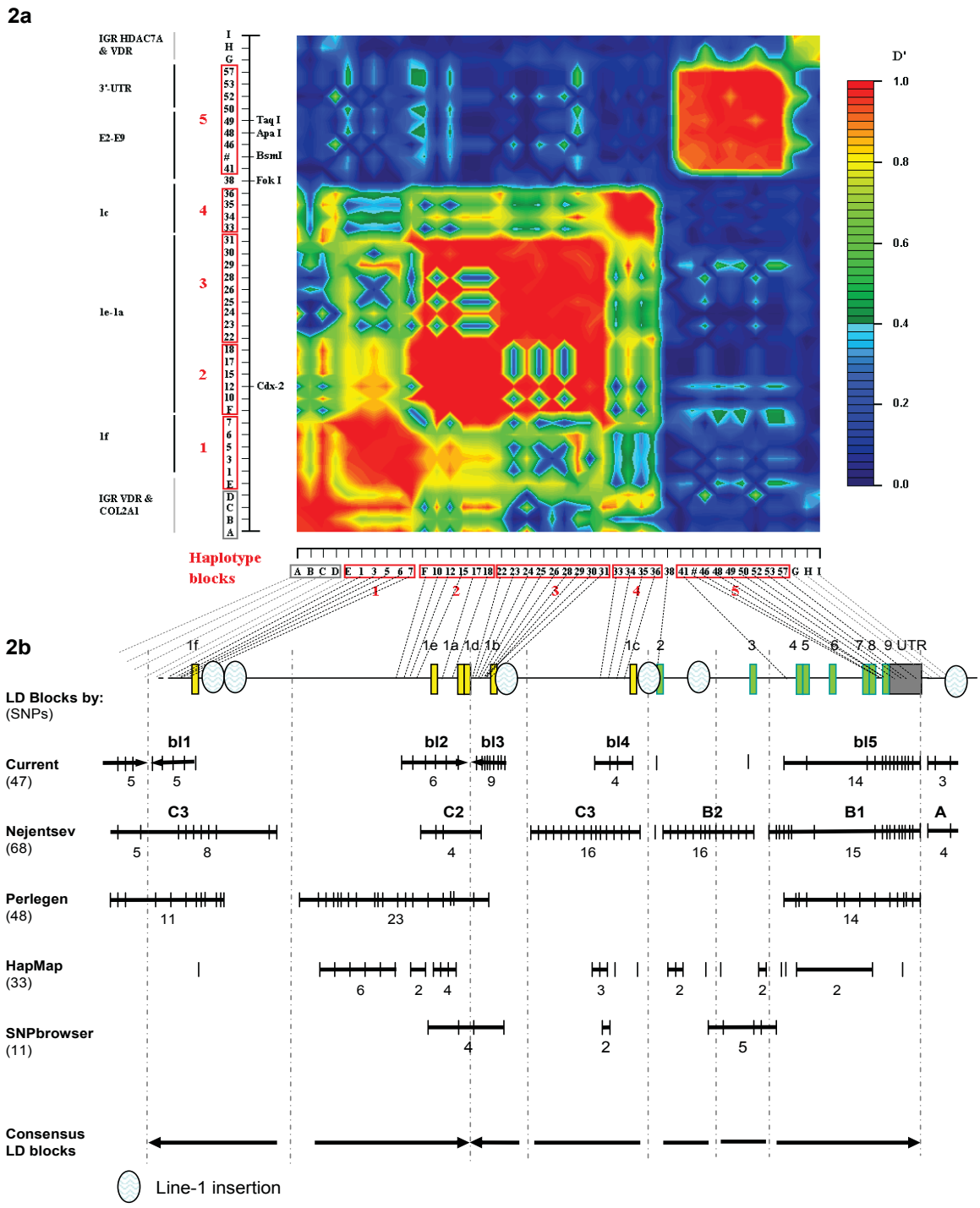
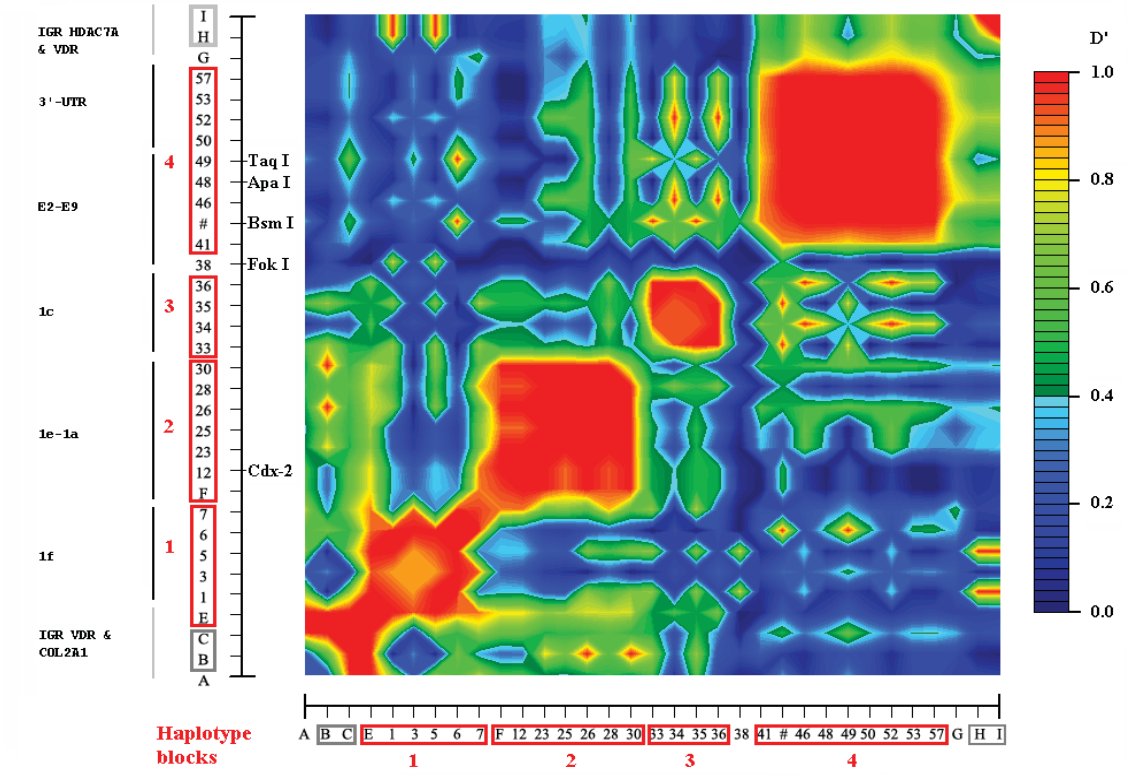


Figure 2 The linkage disequilibrium (LD) structure of the VDR gene in Caucasians. A color version of this figure is presented at the front flap page. **2a** Blocks with pair-wise D'-values with D'>0.8 are numbered 1-5. The analyzed SNPs (Table 1) include: five SNPs in the COL2 α 1 and VDR intergenic region (IGR VDR & COL 2A1), 39 VDR SNPs and three SNPs in the VDR and HDAC7A intergenic region (IGR HDAC7A & VDR). SNP ID is based on Table 1 in Chapter 2; Figure 1a and Table 1 in the chapter. The red boxes on the X and Y-axes indicate the high LD blocks used to define haplotype alleles. Physical organization of the VDR gene is represented with vertical lines on the Y-axis (Fig. 2a, 3a and 3b). **2b** Aligned LD analyzes from different sources and estimated consensus LD structure of the VDR gene. Total number of SNPs analyzed in each study is indicated between brackets. Thick lines indicate haplotype blocks with the number of analyzed SNPs below the line and the name of the block above the line.

3a



3b

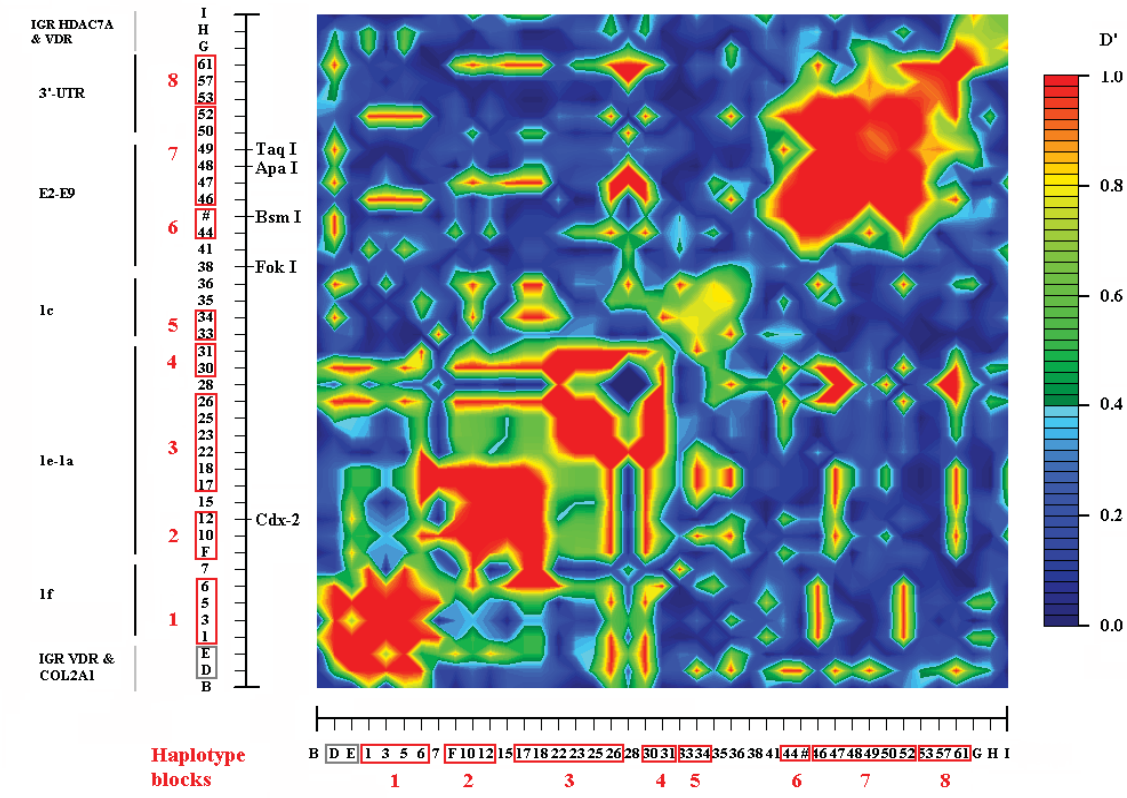


Figure 3 LD map of the VDR gene in different ethnic groups. A color version of this figure is presented at the front flap page. **3a** LD map of 33 SNPs in 107 Asians (214 chromosomes). **3b** LD map of 41 SNPs in 58 African-Americans (116 chromosomes).

REFERENCES

1. Gabriel, S.B. et al. The structure of haplotype blocks in the human genome. *Science* **296**, 2225-9 (2002).
2. Nejentsev, S. et al. Comparative high-resolution analysis of linkage disequilibrium and tag single nucleotide polymorphisms between populations in the vitamin D receptor gene. *Hum Mol Genet* **13**, 1633-9 (2004).
3. Morrison, N.A., Yeoman, R., Kelly, P.J. & Eisman, J.A. Contribution of trans-acting factor alleles to normal physiological variability: vitamin D receptor gene polymorphism and circulating osteocalcin. *Proc Natl Acad Sci U S A* **89**, 6665-9 (1992).
4. Phase. http://archimedes.well.ox.ac.uk/pise/PHASE.html.
5. Haploxt. http://archimedes.well.ox.ac.uk/pise/haploxt.html.
6. Gold. http://www.sph.umich.edu/csg/abecasis/GOLD.
7. HaploBlockFinder. http://cgi.uc.edu/cgi-bin/kzhang/haploBlockFinder.cgi.
8. Perlegen. http://genome.perlegen.com/browser/index.html.
9. HapMap. http://www.hapmap.org/cgi-perl/gbrowse/gbrowse/hapmap.
10. SNPbrowser. http://events-na.appliedbiosystems.com/mk/get/snpb_landing?isource=fr_E_RD_www_allsnps_com_snpbrowser.
11. Uitterlinden, A.G. et al. A large-scale population-based study of the association of vitamin D receptor gene polymorphisms with bone mineral density. *J Bone Miner Res* **11**, 1241-8 (1996).
12. Hinds, D.A. et al. Whole-genome patterns of common DNA variation in three human populations. *Science* **307**, 1072-9 (2005).
13. Durrin, L.K., Haile, R.W., Ingles, S.A. & Coetzee, G.A. Vitamin D receptor 3'-untranslated region polymorphisms: lack of effect on mRNA stability. *Biochim Biophys Acta* **1453**, 311-20 (1999).

Chapter 4

Cdx-2 Polymorphism in the Promoter Region of the Human Vitamin D Receptor Gene Determines Susceptibility to Fracture in the Elderly

Yue Fang, Joyce B. J. van Meurs, Arjan P. Bergink, Albert Hofman, Cornelia M. van Duijn, Johannes P. T. M. van Leeuwen, Huibert A. P. Pols, André G. Uitterlinden

Journal of Bone Mineral Research. 2003 18(9), 1632-41

ABSTRACT

Introduction: A single nucleotide polymorphism (SNP) within a binding site of the intestinal-specific transcription factor Cdx-2 in the promoter region of the human Vitamin D Receptor (VDR) gene was previously reported. It was found to modulate the transcription of the hVDR gene and to be associated with decreased BMD in a small group of postmenopausal Japanese women. In this study we investigated the relationship between the VDR Cdx-2 genotype and risk of fracture.

Materials and Methods: We first determined the location of this SNP in the VDR gene by sequencing analysis, and developed an allele-specific multiplex PCR test to determine the Cdx-2 genotype. We then performed an ecological study in 8 ethnic groups and an association analysis in a large epidemiological cohort of 2848 Dutch Caucasian men and women, aged 55 years or older.

Results: The location of the G to A substitution was found in the promoter region of exon 1e (1e-G1760A) of the VDR gene. By comparing the frequency of the A-allele in eight different ethnic groups, we observed a negative correlation between prevalence of the A-allele and published hip fracture incidence rates in these ethnic groups ($p = 0.006$ for men and $p = 0.02$ for women), suggesting a protective effect of this allele on fracture risk. Subsequently, in the association study the A-allele (population frequency 19%) was observed to have a protective effect on occurrence of osteoporotic fractures, especially for non-vertebral fracture in women (Relative Risk of AA vs. GG genotype is 0.2, 95%CI is 0.05-0.8). This effect remained after adjustment for age, weight and bone mineral density (BMD).

Conclusions: The A-allele of the VDR Cdx-2 polymorphism is present in Caucasians, albeit at low frequency, and has a protective effect on risk of fracture.

Keywords: osteoporosis, ethnic groups, genetics, allele-specific multiplex PCR

INTRODUCTION

Polymorphisms of the vitamin D receptor (VDR) gene have been found to be associated with many clinical endpoints, such as osteoporotic fracture[1], osteoarthritis[2], diabetes[3], breast cancer[4], prostate cancer[5] and low BMD[6]. Most of these studies involved the analysis of polymorphisms which are located at the 3' end of the VDR gene, such as the Bsm I, Apa I and Taq I restriction fragment length polymorphisms (RFLPs) or poly (A) repeat. However, these polymorphisms are not likely to be functional by themselves, either because they are in an intron or they do not change the sequence of amino acids, while the potential functional effect of poly (A) repeat polymorphism is still unclear. Another commonly studied polymorphism in the VDR gene is the Fok I RFLP which detects a translation initiation codon polymorphism[7]. Although it changes the structure of the VDR protein and is a potential functional polymorphism[8], so far there is no strong evidence of association of this polymorphism with any clinical endpoint including fracture[9,10].

Recently, Yamamoto et al.[11] described a functional binding site for the intestinal-specific transcription factor Cdx-2 in the 1a promoter region of VDR gene. Subsequently, Arai et al.[12] described a G to A substitution polymorphism at this Cdx-2 site which was found to modulate the intestine-specific transcription of the VDR gene. Functional differences between these two alleles could be demonstrated whereby the A-allele had increased binding to the Cdx-2 protein *in vitro* and showed in the human colonic carcinoma cell line Caco-2, increased transcription activity of the VDR promoter compared with the G-allele. Furthermore, in a small group of Japanese postmenopausal women (n = 55) these authors found the A-allele to be associated with increased bone mineral density (BMD). Together, these data suggest the Cdx-2 polymorphism to be a functional polymorphism in the VDR gene. However, data on the occurrence of this polymorphism in other ethnic groups and from large genetic epidemiological study-cohorts are currently not available. The Cdx-2 polymorphism is a single nucleotide polymorphism (SNP) which was detected by direct sequencing[12]. However, this technique is too laborious and time consuming for screening purposes in large populations. In this study we therefore developed a so-called allele-specific multiplex PCR (ASM-PCR), to detect the Cdx-2 polymorphism. We first analysed this polymorphism in a panel of several ethnic groups. In addition, we performed a large-scale association study in a prospective cohort study of Caucasian men and women, aged 55 years and older. We analysed this polymorphism in relation to fracture and BMD.

MATERIALS AND METHODS

SUBJECTS

Panel of ethnic groups

We genotyped a panel of DNA from 88 subjects of different ethnicity. The panel was obtained from the Coriell Institute (Camden, New Jersey, USA). It consists of DNA from 10 African Americans (HD04), 9 Africans from south of the Sahara (HD12); 10 Chinese (HD02), 10 Japanese (HD07), 10 Southeast Asians (excluding Chinese and Japanese, HD13); 10 Northern Europeans (HD01), 9 Indo Pakistanis (HD03), 10 Middle Easterns (HD05) and 10 Mexicans (HD08). We grouped these 88 subjects into three major human race groups defined as: African (19 subjects, HD04 and HD 12), Mongoloid (30 subjects, HD02, HD07 and HD13) and Caucasian (39 subjects, HD01, HD03, HD05 and HD08)

Data on the incidence rates of hip fracture in different ethnic groups were collected from published studies[13-25], and are presented in Table 1. To enable comparisons among different published studies, we standardised the age-adjusted incidence of hip fracture in men and women aged ≥ 55 years according to the US population distribution in 1990[26,27]. The direct standardisation method was used, every age-specific rate was multiplied by the corresponding US population in the same age category to get the number of hip fractures for each age category. The sum of fractures was then divided by the total US population to get age-adjusted incidence of hip fracture. For each ethnic group, we computed the average incidence rates, which were weighted by the size of the study population, to compare incidence rates of hip fracture among these eight ethnic groups.

Table 1 Overview of studies used to determine the age-adjusted* incidence of hip fracture in men and women of age 50 years or older

Ethnicity	Source population	Author (ref.)	Incidence (per 100 000)	
			Men	Women
Chinese	Chinese	Lau et al.[13], Koh et al.[14], Schwartz et al.[15], Xu et al.[16], Yan et al.[17], Zhang et al.[18].	118	158
African	African American	Bauer et al.[19]	16	182
Mexican	Mexican American	Bauer et al.[19]	95	183
Japanese	Japanese	Hashimoto et al.[20], Iga et al.[21].	144	254
Southeast Asian	Singapore, Malaysia, Thailand	Lau et al.[13], Koh et al.[14]	111	284
Middle Eastern	Kuwaiti	Memon et al.[22]	200	295
Indo Pakistani	Singapore Indian	Koh et al.[14]	128	361
Northern European	Norwegian	Schwartz et al.[15], Lofthus et al.[23], Luthje et al.[24], Flach at el.[25]	243	419

*: Adjusted according to the US population distribution in 1985 or 1990[26,27]

Study population

The study population sample of Caucasian elderly was derived from the Rotterdam Study (the source population), a single centre prospective population-

based cohort study including 7983 individuals with 3105 men (38.9%) and 4878 women (61.1%) to analyse determinants and prognosis of chronic and disabling diseases in the elderly[28]. The baseline measurements were performed between 1990 and 1993. The third follow-up examination phase took place from 1997 to 1999; the mean follow-up period was 6.6 years (range 5.3-10.2). Baseline measurements of BMD were available for 5931 independently living subjects from the study. For the current study 1453 of these were excluded on the basis of age (>80), use of a walking aid, diabetes mellitus or use of estrogen, thyroid hormone or cytostatic drug therapy. From the 4478 remaining subjects, a random sample of 2848 subjects (the study population) with 1131 men (39.7%) and 1717 women (60.3%) was drawn at baseline, comprising independently living participants aged 55 to 80 years. All subjects have records of incident non-vertebral fracture, while 1915 subjects (67.2% of study population) survived and had follow-up radiograph records to assess incident vertebral fracture. Data on dietary intake was available for 2536 subjects (90.0% of study population).

Measurements

Anthropometric variables

Information about medical history, dietary habits, age at menopause and smoking was obtained with a computerized questionnaire during a home interview. Intakes of calcium and total energy were calculated by food frequency questionair (based on all food and drinks consumed in one month) with the use of Dutch food composition tables. Dietary calcium intake was adjusted for total energy intake. Anthropometric measurements of participants were obtained at the research centre. Body Mass Index (BMI) was calculated as weight (kg) divided by the height square (m^2). Bone mineral density (BMD in g/cm^2) was determined by dual energy X-ray absorptiometry (Lunar DPX-L densitometer, Lunar Corp., Madison, WI, USA) at the femoral neck and lumbar spine (vertebral L2-L4) as described before[29].

Definition of fracture

Non-vertebral fractures (including hip, wrist and other fractures, but excluding head, foot, hand and pathological fracture) were recorded by general practitioners (GPs) who covered 80% of the population. Research physicians confirmed follow-up information by checking GPs' patient records and collected the data of the remaining 20% of the population. Discharge reports and letters from medical specialists were additionally used to verify the hospitalized non-vertebral fracture patients. All fractures were coded independently by two research physicians according to the International Classification of Diseases, 10^{th} revision (ICD-10). If there was disagreement, consensus was reached in a separate session. A medical expert in the field reviewed all coded events for a final classification. The incidence of non-vertebral fracture was defined as all new cases occurring during the follow-

up period.

To assess presence of vertebral fracture, lateral radiographs of the spine from the fourth thoracic to the fifth lumbar vertebrae were obtained and analysed morphometrically by McCloskey-Kanis method as described previously[30]. All vertebral fractures were confirmed through visual interpretation by an expert in the field. When a vertebra was determined to be normal at baseline and any of the three vertebral heights (anterior, central or posterior) showed a minimum decrease of at least 4.6 mm and 15% in absolute height of the later film, then it was considered an incident vertebral fracture. The incidence of vertebral fracture was defined as all new cases occurring during the follow-up period.

Genotyping

Genomic DNA was isolated form peripheral venous blood samples according to standard proteinase K digestion and phenol-chloroform extraction. The position of the Cdx-2 (also denoted 1e-G1760A) polymorphism is based on our sequence analysis of the 1e promoter region of the VDR gene (manuscript in preparation). Two sets of primers were designed for the ASM-PCR test:

G-For: 5'-AGGATAGAGAAAATAATAGAAAACATT-3'
G-Rev: 5'-AACCCATAATAAGAAATAAGTTTTTA**C**-3'
A-For: 5'-TCCTGAGTAAACTAGGTCACA**A**-3'
A-Rev: 5'-ACGTTAAGTTCAGAAAGATTAATTC-3'

A schematic representation of the method and localization of the allele-specific primer sets is shown in Figure 1a. G-Rev and A-For are allele-specific primers. The primer A-For is designed from 5' to 3' of the sense strand (+ strand), and the last base is "A" at the site of the polymorphism. The primer G-Rev is from 5' to 3' of the antisense strand (- strand), and stops at "C" (the complement base of "G") at the polymorphic site. These four primers generate three PCR fragments: primer set G-For and G-Rev specifically amplifies the G-allele with a size of 110 basepair (bp), A-For and A-Rev specifically amplify the A-allele with a size of 235 bp, the out-primer pair (G-For and A-Rev) amplifies the internal control PCR fragment with a size of 297 bp. A schematic representation of the gel electrophoresis pattern is shown in Figure 1b. To verify the first result of genotype calls, a random set of 5% of the samples were genotyped again.

PCR and gel electrophoresis

The PCR amplification was carried out in a GeneAmp PCR system 9700 (Applied Biosystems) with MicroAmp Optical 96-well Reaction Plate (Applied Biosystems) and in a PTC-225 DNA Engine Tetrad (MJ research) with polypropylene 96-well and 384-well thin wall microplates. 10 µl of the PCR reaction system consisted of 1.0 µl 10× PCR Buffer (1× Buffer = 10 mM Tris-Cl, pH 8.3; 50 mM KCl; 1.25 mM $MgCl_2$), 1.0 µl 10× dNTPs (2mM), 0.4 pmol G-For, 0.6 pmol G-Rev, 0.6 pmol A-For, 0.4 pmol A-Rev, 0.5 units Super Taq (HT Biotechnology LTD) and

10 ng genomic DNA. PCR was performed with an initial denaturation at 96°C for 5 minutes, followed by 28 cycles of denaturation at 94°C for 45 seconds, annealing at 56°C for 45 seconds and extension at 72°C for 45 seconds. The final extension was at 72°C for 5 minutes. PCR products were size-separated on a 2.5% agarose gel at 125 volts for one hour. The 100-bp DNA ladder of GIBCOBRL™ was used to determine the size of fragments.

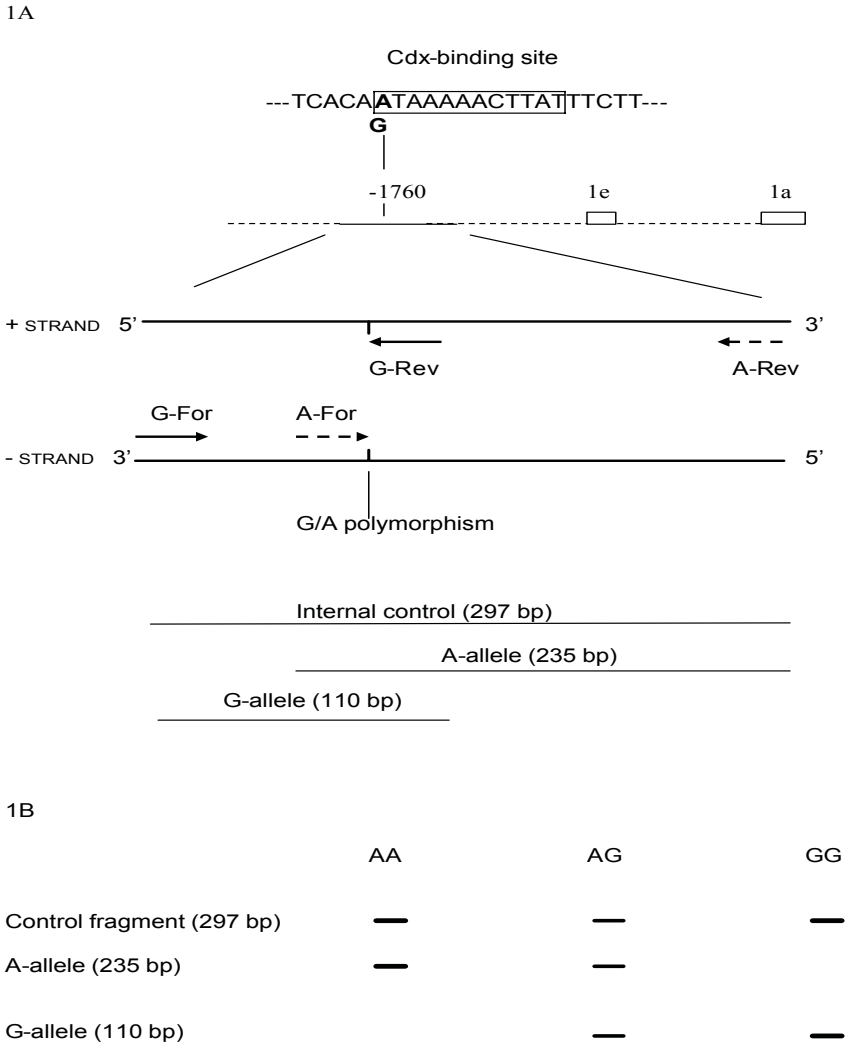


Figure 1 (A) Schematic representation of ASM-PCR for detecting the G to A SNP at the Cdx-2 binding site in the VDR 1e promoter region (1e-G1760A). Primers A-For and G-Rev are located flanking the SNP in opposite directions. The last base is at the SNP site, the last base of A-For is "A" to match the A-allele and the last base of G-Rev is "C" (complement of "G") to match the G-allele. Primer set G-For and G-Rev generates a 110 bp fragment when the G-allele is present. Primer set A-For and A-Rev generates a 235-bp fragment when the A-allele is present. The primer set G-For and A-Rev generates a 297-bp fragment which is an internal control and independent of the presence of either the G or A-allele. (B) Schematic representation of the gel pattern obtained with the ASM PCR test for Cdx-2 genotyping. The largest PCR fragment (297 bp) is the internal control fragment which can be seen in any successful PCR. The 235-bp fragment is the A-allele specific PCR product, and the 110-bp fragment is the G-allele specific fragment.

Sequencing

We sequenced the entire region of 6.5 kilo base pairs (kb) before the exon 1a of VDR gene in DNA from 15 young female Caucasians as a part of our sequence analysis of the VDR gene[31]. To sequence the region around the Cdx-2 polymorphism, the 297-bp PCR product, which was amplified with the out-primer set of G-For and A-Rev, was purified by using Quantum Prep® PCR Kleen Spin Columns (BIO-RAD). We performed direct sequence analysis of PCR products by cycle sequencing using an ABI PRISM bigdye® kit from Applied Biosystems. The sequencing reaction (20 μl) included 4 μl of 5× buffer (400 mM Tris-Cl pH 9.0, 10 mM $MgCl_2$), 1.0 nM of G-For or A-Rev, 4 μl of Terminator Ready Reaction Mixture (containing BigDye®, Applied Biosystems) and 11 μl of PCR product. The sequencing reaction was performed in GeneAmp® PCR system 9700 (Applied Biosystems) with the following 25 cycles: 96°C for 10 seconds, 50°C for 5 seconds and 60°C for 4 minutes. The sequencing product was purified by using Micro Bio-Spin® chromatography Columns (BIO-RAD). Forward and reverse sequences of PCR products were produced with an ABI PRISM 310 or 3100 Genetic Analyzer (Applied Biosystems).

Statistical Analysis

For the ecological study Spearman's Correlation test was used to test for correlation between VDR Cdx-2 genotype distribution and the eight ethnic groups ranked by gender-specific hip fracture incidence. To compare the differences of the major anthropometric characteristics and clinical endpoints between the source population and the study population for the epidemiological study, we performed independent sample t-tests for continuous variables and the Pearson chi-square analysis for categorical variables. Hardy-Weinberg equilibrium was tested for the Cdx-2 genotype by a chi-square goodness of fit test. We grouped subjects by their genotype for the Cdx-2 polymorphism as GG, GA and AA. The GG genotype was defined as reference group because it was the most frequent genotype in our study group. Differences in anthropometric characteristics by genotype were evaluated by analysis of variance (ANOVA) for continuous variables and Pearson chi-square p-value for categorical variables. Three possible genetic models were allowed to explain differences between groups, i.e., an allele dose effect, a dominant effect or a recessive effect. Allele dose was defined as the number of copies of a certain allele in the genotype, and genotype was treated as a continuous variable. For the A-allele the genotype is expressed as a numeric variable with 0 = GG, 1 = GA and 2 = AA. In case of a consistent trend reflected as an allele dose effect, a linear regression analysis was performed and a "trend" p-value was calculated to quantify the association. In case of a recessive or dominant effect of the test allele, a two by two chi-square test or an independent sample T-test was performed to test for differences between two genotype groups. For recessive effects homozygous subjects for the test group (e.g., AA) were compared to the combined group of

heterozygous carriers (e.g., GA) and non-carriers (e.g., GG). For dominant alleles we compared the test allele carriers (e.g., GA and AA) versus non-carriers (e.g., GG).

We first analysed risk of any fracture, which includes vertebral and non-vertebral fracture, by the Cdx-2 genotype. This analysis was limited to the group for which data on vertebral fracture was available (n=1915). We then stratified the analysis by type of fracture for vertebral vs. non-vertebral fracture. For analysis of non-vertebral fracture the complete study sample (n=2848) could be included. 84 individuals who had both vertebral and non-vertebral fracture during the follow-up time were included in all analysis. Differences in the non-vertebral fracture frequency by genotype were compared using Pearson chi-square tests in general and a linear regression model for the allele-dose effect. Relative Risk (RR) and 95% confidence intervals (95% CI) were calculated for the relationship between the Cdx-2 genotype and fracture by using Cox regression models. The same models were used to estimate the RR adjusted for potential confounders, such as age, gender, BMI and BMD.

To analyse effects of potential confounders and modification by dietary calcium intake, we also stratified by gender; percentiles and particular cut-off level (less or more 600 mg/day) of dietary calcium intake. Differences in BMD by genotype were adjusted for age and BMI by a general linear model. All statistical analyses (except Hardy-Weinberg equilibrium) were carried out with the SPSS software package (version 9.0).

RESULTS

The 1e promoter region

We determined the sequence of the promoter region of 6.5 kb in front of the VDR gene exon 1a using a sequence walking strategy. We used reference sequence information from Miyamoto et al.[32], Yamamoto et al.[11], NCBI (http://www.ncbi.nlm.nih.gov/) and the Celera database. We found exon 1e (accession number of the NCBI genomic database is AH006427 from Crofts et al.[33]) was only 2 kb in front of exon 1a (see Figure 1A). The G to A substitution in the Cdx-2 binding site is located 1760 basepairs in front of exon 1e. Therefore the Cdx-2 polymorphism is referred to as 1e-G1760A.

Genotyping

The ASM-PCR method to determine Cdx-2 genotype was applied to genotype 88 subjects from different ethnicity and 2848 men and women from the Rotterdam Study. Figure 2 shows a representative gel separation pattern of ASM-PCR genotype analysis for 95 samples and one negative control (no genomic DNA). In order to confirm the genotype result, we re-genotyped 5% random samples from those

subjects and found no discrepancy. In addition, 15 random DNA samples were directly sequenced after PCR to evaluate the Cdx-2 genotype. Both genotype results were completely identical. No other polymorphism was found in this 297-bp region around Cdx-2 polymorphism in the 1e promoter region of the VDR gene according to the sequencing result of these 15 random samples. The genotype distribution in the sample from the Rotterdam Study obeyed Hardy-Weinberg equilibrium (p= 0.17). The frequency of the G-allele and A-allele in the large Dutch Caucasian population is 81% (4614/5696) and 19% (1082/5696), respectively. Allele frequencies were not different in men and women (p = 0.93), did not vary by age in men (p = 0.42) nor in women (p = 0.17), and did also not vary in the different subsets drawn from the Rotterdam Study (p = 0.86).

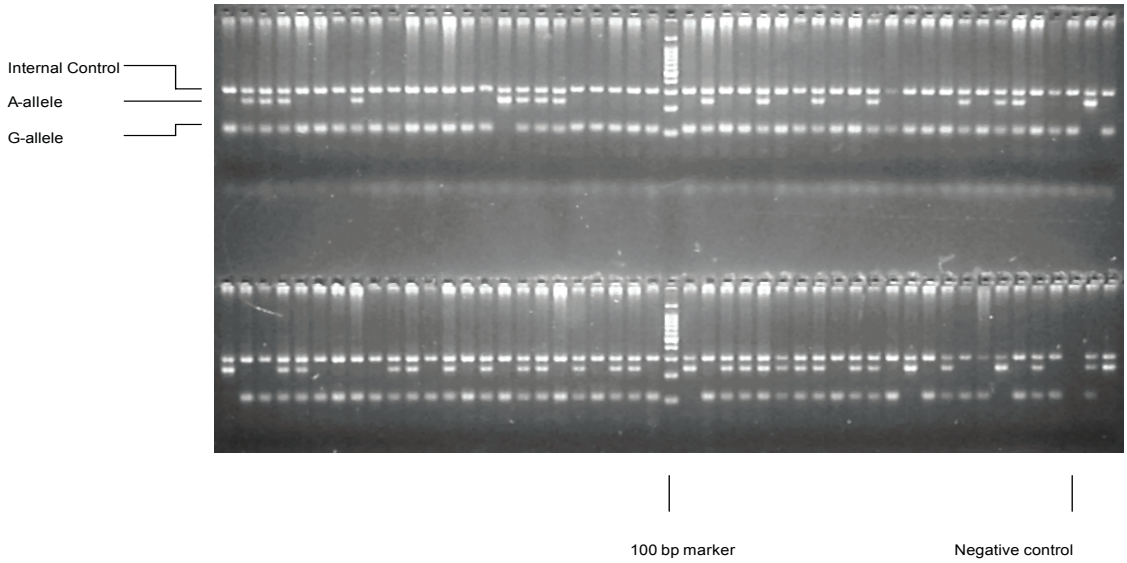


Figure 2 Gel electrophoretic separation pattern of ASM-PCR analysis of 93 samples of our study population and 3 controls for the Cdx-2 polymorphism

Ecological study of the Cdx-2 polymorphism

The frequencies of the Cdx-2 alleles and the genotype distribution in different ethnic groups are presented in Table 2. Genotype distribution and allele frequency differed substantially among ethnic groups. When we combined individual Coriell panel groups into the three major races we found the A-allele to have highest frequency in Africans (74%, 95% CI 60-88%), intermediate in Asian (43%, 95%CI 39-47%) and lowest in Caucasian (19%, 95% CI 18-20%). We observed an inverse relation between Cdx-2 A-allele frequency and the incidence rate of hip fractures, both for men and women (Spearman's correlation test r = -0.87 p = 0.006 for men; and r = -0.79 p = 0.02 for women). In Figure 3 we plotted the age-adjusted incidence rates of hip fracture in women of age 50 years or older, against the frequencies of the VDR Cdx-2 A-allele in the eight ethnic groups. We excluded HD12 (Africans

from south of the Sahara) in the figure, because no hip fracture incidence data are available for this group.

Table 2 The VDR 1e promoter Cdx-2 polymorphism in different ethnic groups

Ethnic group	Coriell code	Number of subjects	Genotype frequency (No.) GG	GA	AA	Allele frequency (No.) G	A
African		**19**	**0.05 (1)**	**0.42 (8)**	**0.53 (10)**	**0.26 (10) [95 % CI]#**	**0.74 (28) [0.60-0.88]**
South Sahara	HD12	9	0	0.44 (4)	0.56 (5)	0.22 (4)	0.78 (14)
African American	HD04	10	0.10 (1)	0.40 (4)	0.50 (5)	0.30 (6)	0.70 (14)
Mongoloid		**291**	**0.32 (92)**	**0.50 (147)**	**0.18 (52)**	**0.57 (331) [95 % CI]#**	**0.43 (251) [0.39-0.47]**
Southeast Asian*	HD13	10	0.10 (1)	0.80 (8)	0.10 (1)	0.50 (10)	0.50 (10)
Chinese	HD02	10	0.30 (3)	0.50 (5)	0.20 (2)	0.55 (11)	0.45 (9)
Japanese	HD07	10	0.60 (6)	0.30 (3)	0.10 (1)	0.75 (15)	0.25 (5)
Japanese (Arai)		261	0.31 (82)	0.50 (131)	0.18 (48)	0.56 (295)	0.44 (227)
Japanese (total)		271	0.32 (88)	0.49 (134)	0.18 (49)	0.57 (310)	0.43 (232)
Caucasian		**2887**	**0.67 (1921)**	**0.29 (848)**	**0.04 (118)**	**0.81 (4690) [95 % CI]#**	**0.19 (1096) [0.18-0.20]**
Mexican	HD08	10	0.60 (6)	0.30 (3)	0.10 (1)	0.75 (15)	0.25 (5)
Indo Pakistani	HD03	9	0.67 (6)	0.22 (2)	0.11 (1)	0.78 (14)	0.22 (4)
Middle Eastern	HD05	10	0.60 (6)	0.40 (4)	0	0.80 (16)	0.20 (4)
Northern European	HD01	10	0.90 (9)	0.10 (1)	0	0.95 (19)	0.05 (1)
Dutch (current study)		2848	0.67 (1894)	0.29 (838)	0.04 (116)	0.81 (4626)	0.19 (1082)
Northern European (total)		2858	0.67 (1903)	0.29 (839)	0.04 (116)	0.81 (4645)	0.19 (1083)

#: 95% confidence interval of A-allele frequency
*: Excluding Chinese and Japanese

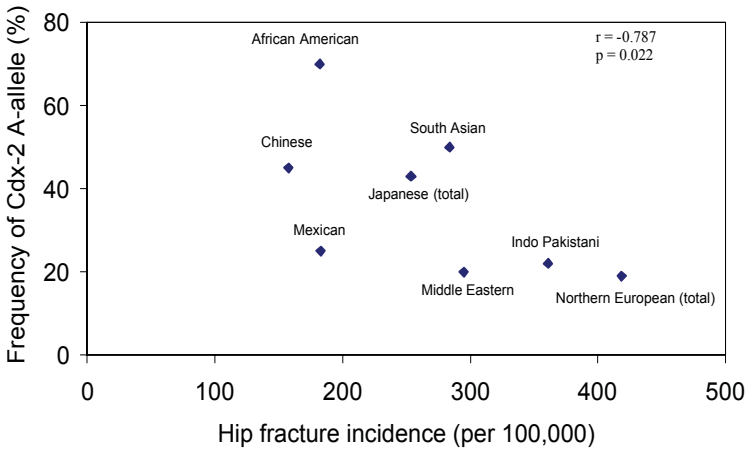


Figure3 Age-adjusted incidence rates of hip fracture in women of different ethnic groups according to the frequency of the VDR Cdx-2 A-allele. Incidence rates of each group were standardised according to US population in 1985 or 1990, and we calculated the female-specific incidence rates for subjects aged 50 or older.

Association study of Cdx-2 polymorphism with fracture

We went on to analyse the relation between fracture risk and the VDR Cdx-2 polymorphism in a large Caucasian study population of Dutch men and women. We first compared baseline characteristics, including age, gender, height, weight, BMI, age of menopause, dietary calcium intake and current smoking, between the source population and our study population. The subjects of our study group were 4 years younger, 0.6 cm shorter and had 0.2 (kg/m2) lower BMI on average than the source population. In our study population we observed that all fracture risk increased by increasing age (the linear regression $p < 0.001$ for all kinds of fractures) and the incidences of all fractures in women were significantly higher than in men (chi-square $p < 0.001$ for all). Baseline characteristics of men and women in the study population were not found to be significantly different when stratified by VDR

Cdx-2 genotype (data not shown). The genotype distribution was not significantly different in men and women as well as 5-year age categories.

Table 3 Risk of fracture according to the VDR Cdx-2 genotype in men and women

	Total	By Cdx-2 genotype			
		GG	GA	AA	p-value#
Case/total (%)					
Any fracture	381/1915 (19.9)	268/1270 (21.1)	103/567 (18.2)	10/78 (12.8)	0.06
Vertebral fracture	217/1915 (11.3)	156/1270 (12.3)	56/567 (9.9)	5/78 (6.4)	0.04
Non-vertebral fracture	248/2848 (8.8)	173/1894 (9.1)	70/838 (8.4)	5/116 (4.3)	0.12
RR (95%CI)					Per copy of A-allele
Any fracture	Crude	1	0.8 (0.6-1.1)	0.7 (0.4-1.3)	0.8 (0.7-1.0)
	*Adjusted**	1	0.8 (0.6-1.1)	0.7 (0.4-1.3)	0.8 (0.7-1.0)
Vertebral fracture	Crude	1	0.8 (0.6-1.1)	0.5 (0.2-1.2)	0.8 (0.6-1.0)
	*Adjusted**	1	0.8 (0.6-1.1)	0.5 (0.2-1.3)	0.8 (0.6-1.0)
Non-vertebral fracture	Crude	1	0.9 (0.7-1.3)	0.5 (0.2-1.2)	0.9 (0.7-1.1)
	*Adjusted**	1	0.9 (0.7-1.2)	0.5 (0.2-1.2)	0.8 (0.6-1.1)

#: Calculated by linear regression model
*: RR was adjusted for age, gender, weight and femoral neck BMD

We compared by Cdx-2 genotype the incidence of any fracture (including vertebral and non-vertebral fracture), and then separately by vertebral fracture and by non-vertebral fracture during the 6.6 years follow-up period (Table3). We found subjects carrying the A-allele to have fewer fractures with evidence for an allele dose effect. The association was borderline significant for vertebral fracture ($p = 0.04$) and any fractures ($p = 0.06$), while a similar trend was found for non-vertebral fracture ($p = 0.12$). When stratifying the same analysis by gender, the protective effect of the A-allele was similar in women and men. Because of the higher incidence of non-vertebral fracture in women (13.6%) than in men (6.2%) and the low frequency of the Cdx-2 A-allele in our Caucasian population, we further investigated the effect of Cdx-2 genotype on non-vertebral fracture (including hip fracture) in women separately (Table 4). Compared to the whole study population we observed a similar protective effect of the Cdx-2 genotype in the women which was borderline significant ($p=0.05$). When Relative Risks (RR) were calculated we found presence of the Cdx-2 A-allele to be associated with reduced risk for fracture with evidence for an allele dose effect. The RR estimate did not essentially alter after adjustment for age, weight as well as for femoral neck BMD. We did not observe the calcium intake to modify the relation between VDR Cdx-2 genotype and fracture (data not shown).

Association of Cdx-2 genotype with BMD

We also analysed the effect of Cdx-2 genotype on BMD but could not find an association of Cdx-2 genotype with femoral neck BMD or lumbar spine BMD.

Our data showed that BMD at femoral neck and lumbar spine were correlated to calcium intake and age ($p < 0.001$ for both by linear regression analysis). In general, individuals with a high calcium intake had higher BMD than individuals with a low calcium intake. The dietary calcium intake was not different by Cdx-2 genotype ($p = 0.57$, which was adjusted for age and gender). We then went on to analyse the influence of dietary calcium intake on the relation between Cdx-2 genotype and BMD.

In quartiles of dietary calcium intake, no difference in BMD was observed by VDR Cdx-2 genotype. There were only 119 subjects (4% of the study population) with a calcium intake less than 600 mg/day. In this group, a trend could be observed towards A-allele having increased BMD but probably because we did not have sufficient subjects in the group, this failed to reach significance ($p = 0.47$ for femoral neck BMD and $p = 0.56$ for lumbar spine BMD). In the group with high calcium intake (>600 mg/day) we did not see differences in BMD among VDR Cdx-2 genotype groups.

Table 4 Relative Risk of non-vertebral fracture in 1717 women according to the VDR Cdx-2 genotype

Genotype	Case/total (%)	Crude RR (95% CI)	Adjusted RR (95% CI)*	Adjusted RR (95% CI)**
GG	139/1139 (12.2)	1	1	1
GA	57/505 (11.3)	1.0 (0.7-1.3)	1.0 (0.7-1.3)	1.0 (0.7-1.3)
AA	2/73 (2.7)	0.2 (0.06-0.9)	0.2 (0.05-0.8)	0.2 (0.05-0.8)
Per copy of A-allele		0.8 (0.6-1.0)	0.8 (0.6-1.0)	0.7 (0.5-1.0)
p-value	0.05			

*: RR was adjusted for age and weight
**: RR was adjusted for age, weight and femoral neck BMD

DISCUSSION

Location of Cdx-2 polymorphism and genotyping method

From our analysis of the physical map of the region in front of exon 1a and 1e of the VDR gene, we found that exon 1e was only about 2 kb in front of exon 1a. The Cdx-2 polymorphism is therefore positioned in a promoter region in front of exon 1e and not of exon 1a. Yamamoto et al.11 determined that the region around the Cdx-2 polymorphism is important for expression of the VDR in intestinal cell lines. However, the constructs used in that study contained either only exon 1a or both 1e and 1a. The effect of the promoter of exon 1e on the expression of the VDR gene was not analysed separately in that study. Crofts et al.33 analysed the tissue-specific mRNA expression of multiple promoters of the VDR gene, and found exon 1a

expressed in all tissues or cell lines in that study. Transcripts containing exon 1f and 1e were found to be expressed in kidney and two tumour cells, but exon 1e was not analysed separately. Cdx-2 protein is an intestine-specific transcription factor which could regulate the expression of VDR in the same tissue, and influence calcium homeostasis consequently. Therefore, to confirm whether 1e is a tissue-specifically expressed exon, further analysis of tissue specific expression of VDR (especially in intestinal cells) is necessary to elucidate the underlying molecular mechanism of the association study.

We introduced the ASM-PCR genotyping method for detecting the Cdx-2 polymorphism in our study. The Cdx-2 polymorphism does not create a natural or artificial restriction site by itself. In the ASM-PCR method we designed four primers, to generate three potential PCR fragments, which have about 100 bp difference in size, so they can be readily distinguished by normal agarose gel electrophoresis. In previous allele-specific PCR systems34, detection of each allele required one pair of primers, so two PCR reactions had to be performed for two alleles. This made it difficult to distinguish whether the relevant allele is truly absent or whether this is due to PCR failure. In the ASM-PCR system described here, genotypes can be determined in a single PCR without any additional treatment before agarose gel electrophoresis. The reliability of the method was proven by reproducible results obtained by sequencing and by repeated genotyping.

Ecological study of A-allele frequency and hip fracture incidence

The frequency of the A-allele in this Dutch Caucasian population (19%) was much lower than that reported previously for the Japanese population (43%12). Our analysis of the ethnic panel also suggests large differences in A-allele frequency of this polymorphism ranging from 19% in North European Caucasian to 74% in African subjects. The A-allele frequency in the small sample (10) Japanese subjects used for the ecological study (25%) differed from that found in the relatively large study (261 subjects) by Arai et al. (43%), which can be due to differences in power between these two samples. The previously reported hip fracture incidence rates13-25 appeared to be highest in subjects of northern European extraction and lowest in those of Asian and African origin. By comparing incidence rates with the frequency of the A-allele among these ethnic groups, we observed an apparent inverse relationship between the A-allele frequency and hip fracture rate (Figure 3). However, results from ecological studies have to be interpreted with caution because we cannot rule out alternative explanations for the observed relation ("ecological fallacy"). Yet, the protective effect on fracture of the Cdx-2 A-allele suggested by this comparison is in line with the results from our epidemiological study in the Caucasian population of elderly subjects, especially for non-vertebral fracture in elderly women (table 4). In the previous Japanese study12, Arai reported a significant association between A-allele and an increased BMD in postmenopausal women, and this would predict the A-allele to be associated with a decreased risk

of fracture as observed in our current study. Nevertheless, separate epidemiological studies of the relation between the VDR Cdx-2 polymorphism and fracture risk have to be performed in the different ethic groups, to determine if the protective effect of the A-allele on fracture risk is true and consistent.

Epidemiological study in Caucasians

In the population-based cohort study of Caucasian elderly subjects we performed a large-scale association analysis to investigate the relationship between the Cdx-2 polymorphism and fracture risk. The genotype distribution followed the Hardy-Weinberg equilibrium, suggesting absence of selection bias. The associations between Cdx-2 genotype and fractures were observed and independent of age, gender, weight and BMD. We found a trend of decreasing frequency of fracture by increased number of A-alleles, suggesting an allele dose effect. Because we did not have sufficient statistical power, in particular for the AA genotype group, the associations were borderline significant, which makes it difficult to distinguish an allele-dose effect from a recessive effect.

The association of the VDR Cdx-2 A-allele with reduced fracture risk is consistent with the results of functional studies of this polymorphism that were previously reported 11,12. In these studies the A-allele was found to bind more efficiently the Cdx-2 protein and showed increased transcription level of the VDR gene. The Cdx-2 transcription factor plays an important role in intestine-specific gene transcription35,36. As a transcription factor, Cdx-2 could mediate the transcription of the VDR gene via the special cis-element in the 1e promoter region of the VDR gene, thereby affecting the expression of the VDR in the intestine. Thus, the VDR content of intestinal cells of the GA and AA genotype may be higher than that of the GG genotype. The VDR is a transcription factor, and regulates the transcription of other down stream genes in many tissues. Interestingly, by using VDR- knockout mice the expression of two intestinal calcium channels, epithelial calcium channel (ECaC) and calcium transport protein type 1 (CaT1), was shown to be strongly vitamin D-dependent 37. The VDR mediates the effect of vitamin D. It can therefore be hypothesized that VDR Cdx-2 A-allele carriers could have higher intestinal calcium absorption, because of the elevated expression of these intestinal calcium channel proteins. Increased calcium absorption in turn could increase the BMD, and might, thus, contribute to a decreased fracture risk.

In order to investigate the potential mechanism underlying the association between fracture and the VDR Cdx-2 genotype, we further analysed the relationship between BMD and Cdx-2 genotype. However, in our study of Dutch Caucasian elderly we could not observe VDR Cdx-2 genotype to be associated with BMD, which is in contrast with the results reported for the Japanese population. Arai et al. found the VDR Cdx-2 A-allele to be associated with increased BMD at the lumbar spine in a Japanese population of postmenopausal women12. However, several environmental factors involved in bone metabolism, could be different between

the Japanese and the Dutch population, including dietary calcium intake and serum vitamin D level. For example, the dietary calcium intake of this Dutch study population is higher than that of the Japanese population (average intake of 1117 vs. 600 mg/day38). We therefore went on to analyse the influence of dietary Ca-intake on the relationship between Cdx-2 genotype and BMD in our study population.

In our study group increased dietary calcium intake was associated with increased BMD, but the association was independent of VDR Cdx-2 genotype. Yet, when we analysed subjects with a dietary calcium intake of less than 600 mg/day (n=119) which is the similar to that of the Japanese population, we observed a trend towards the A-allele being associated with increased BMD but this failed to reach significance, probably due to low statistical power. Thus, to draw conclusions in this respect it is necessary to analyse VDR Cdx-2 genotype in relation to BMD and fracture in a population with relatively low calcium intake and, for Caucasians, in an even larger sample size than that of the current study population.

In conclusion, in the present study we introduced a simple and specific genotyping method for association analysis of the Cdx-2 polymorphism in the 1e promoter region of VDR gene. In an ecological study we found a strong correlation between frequency of Cdx-2 A-allele and the incidence rates of hip fracture from different ethnic groups. In an epidemiological study we demonstrated that the A-allele has a protective effect on the risk of fracture in Caucasian elderly, especially for women. The association seems not to be directly explained by differences in BMD. Our results prompt the further association analysis of this polymorphism in relation to fracture risk and environmental factors, particularly in Asian and African populations.

ACKNOWLEDGEMENTS

We gratefully acknowledge statistical and meta-analysis advice from Prof. Theo Stijnen, (Department of Epidemiology and Biostatistics, Medical Centre of Erasmus University Rotterdam) and Dr. Chris de Laet (Institute for Medical Technology Assessment, Medical Centre of Erasmus University Rotterdam) for help with the ecological study. We gratefully acknowledge help of the DXA and radiograph technicians L. Buist and H.W.M. Mathot. We thank the participating general practitioners and the many field workers at the research center in Ommoord. This study was sponsored by grants 925-01-010, 903-46-178 and 014-90-001 from the Netherlands Organisation for Scientific Research (NWO).

REFERENCES

1. Uitterlinden, A.G. et al. Interaction between the vitamin D receptor gene and collagen type Ialpha1 gene in susceptibility for fracture. *J Bone Miner Res* **16**, 379-85. (2001).
2. Uitterlinden, A.G. et al. Vitamin D receptor genotype is associated with radiographic osteoarthritis at the knee. *J Clin Invest* **100**, 259-63. (1997).
3. Pani, M.A. et al. Vitamin D receptor allele combinations influence genetic susceptibility to type 1 diabetes in Germans. *Diabetes* **49**, 504-7. (2000).
4. Ruggiero, M. et al. Vitamin D receptor gene polymorphism is associated with metastatic breast cancer. *Oncol Res* **10**, 43-6. (1998).
5. Taylor, J.A. et al. Association of prostate cancer with vitamin D receptor gene polymorphism. *Cancer Res* **56**, 4108-10. (1996).
6. Cooper, G.S. & Umbach, D.M. Are vitamin D receptor polymorphisms associated with bone mineral density? A meta-analysis. *J Bone Miner Res* **11**, 1841-9. (1996).
7. Miyamoto, K. et al. A novel polymorphism in the vitamin D receptor gene and bone mineral density: study of vitamin D receptor expression and function in COS-7 cells. *J Bone Miner Res* **11**, 85-90. (1996).
8. Colin, E.M. et al. Consequences of vitamin D receptor gene polymorphisms for growth inhibition of cultured human peripheral blood mononuclear cells by 1, 25-dihydroxyvitamin D3. *Clin Endocrinol (Oxf)* **52**, 211-6. (2000).
9. Langdahl, B.L., Gravholt, C.H., Brixen, K. & Eriksen, E.F. Polymorphisms in the vitamin D receptor gene and bone mass, bone turnover and osteoporotic fractures. *Eur J Clin Invest* **30**, 608-17. (2000).
10. Bell, N.H., Morrison, N.A., Nguyen, T.V., Eisman, J. & Hollis, B.W. ApaI polymorphisms of the vitamin D receptor predict bone density of the lumbar spine and not racial difference in bone density in young men. *J Lab Clin Med* **137**, 133-40. (2001).
11. Yamamoto, H. et al. The caudal-related homeodomain protein Cdx-2 regulates vitamin D receptor gene expression in the small intestine. *J Bone Miner Res* **14**, 240-7. (1999).
12. Arai, H. et al. The polymorphism in the caudal-related homeodomain protein Cdx-2 binding element in the human vitamin D receptor gene. *J Bone Miner Res* **16**, 1256-64. (2001).
13. Lau, E.M. et al. The incidence of hip fracture in four Asian countries: the Asian Osteoporosis Study (AOS). *Osteoporos Int* **12**, 239-43. (2001).
14. Koh, L.K., Saw, S.M., Lee, J.J., Leong, K.H. & Lee, J. Hip fracture incidence rates in Singapore 1991-1998. *Osteoporos Int* **12**, 311-8. (2001).
15. Schwartz, A.V. et al. International variation in the incidence of hip fractures: cross-national project on osteoporosis for the World Health Organization Program for Research on Aging. *Osteoporos Int* **9**, 242-53. (1999).
16. Xu, L., Lu, A., Zhao, X., Chen, X. & Cummings, S.R. Very low rates of hip fracture in Beijing, People's Republic of China the Beijing Osteoporosis Project. *Am J Epidemiol* **144**, 901-7. (1996).
17. Yan, L., Zhou, B., Prentice, A., Wang, X. & Golden, M.H. Epidemiological study of hip fracture in Shenyang, People's Republic of China. *Bone* **24**, 151-5. (1999).
18. Zhang, L. et al. Epidemiology of cervical and trochanteric fractures of the proximal femur in 1994 in Tangshan, China. *J Bone Miner Metab* **18**, 84-8. (2000).
19. Bauer, R.L. Ethnic differences in hip fracture: a reduced incidence in Mexican Americans. *Am J Epidemiol* **127**, 145-9. (1988).
20. Hashimoto, T., Sakata, K. & Yoshimura, N. Epidemiology of osteoporosis in Japan. *Osteoporos Int* **7**, S99-102. (1997).
21. Iga, T., Dohmae, Y., Endo, N. & Takahashi, H.E. Increase in the incidence of cervical and trochanteric fractures of the proximal femur in Niigata Prefecture, Japan. *J Bone Miner Metab* **17**, 224-31. (1999).
22. Memon, A. et al. Incidence of hip fracture in Kuwait. Int J Epidemiol **27**, 860-5. (1998).
23. Lofthus, C.M. et al. Epidemiology of hip fractures in Oslo, Norway. *Bone* **29**, 413-8.

(2001).
24. Luthje, P. et al. Hip fractures in urban population in Finland. *Ann Chir Gynaecol* **81**, 316-21. (1992).
25. Falch, J.A., Kaastad, T.S., Bohler, G., Espeland, J. & Sundsvold, O.J. Secular increase and geographical differences in hip fracture incidence in Norway. *Bone* **14**, 643-5. (1993).
26. US population estimates by Age, Sex, Race and Hispanic Origin: 1980 to1991, (US Department of Commerce, Economics and Statistics Administration, Bureau of the Census, 1993).
27. National and State Population Estimates: 1990 to 1994 (age and sex), (US Department of Commerce, Economics and Statistics Administration, Bureau of the Census, 1995).
28. Hofman, A., Grobbee, D.E., de Jong, P.T. & van den Ouweland, F.A. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur J Epidemiol* **7**, 403-22. (1991).
29. Burger, H. et al. The association between age and bone mineral density in men and women aged 55 years and over: the Rotterdam Study. *J Bone Miner Res* **25**, 1-13. (1994).
30. McCloskey, E.V. et al. The assessment of vertebral deformity: a method for use in population studies and clinical trials. *Osteoporos Int* **3**, 138-47. (1993).
31. Fang, Y., van Meurs, J.B.J., van Leeuwen, J., Pols, H.A.P. & Uitterlinden, A.G. sequence variation of the vitamin D receptor gene. *Calcif Tissue Int* **67**, 485 (2000).
32. Miyamoto, K. et al. Structural organization of the human vitamin D receptor chromosomal gene and its promoter. *Mol Endocrinol* **11**, 1165-79. (1997).
33. Crofts, L.A., Hancock, M.S., Morrison, N.A. & Eisman, J.A. Multiple promoters direct the tissue-specific expression of novel N-terminal variant human vitamin D receptor gene transcripts. *Proc Natl Acad Sci U S A* **95**, 10529-34. (1998).
34. Kirschbaum, N.E. & Foster, P.A. The polymerase chain reaction with sequence specific primers for the detection of the factor V mutation associated with activated protein C resistance. *Thromb Haemost* **74**, 874-8. (1995).
35. Suh, E., Chen, L., Taylor, J. & Traber, P.G. A homeodomain protein related to caudal regulates intestine-specific gene transcription. *Mol Cell Biol* **14**, 7340-51. (1994).
36. Jin, T. & Drucker, D.J. Activation of proglucagon gene transcription through a novel promoter element by the caudal-related homeodomain protein cdx-2/3. *Mol Cell Biol* **16**, 19-28. (1996).
37. Van Cromphaut, S.J. et al. Duodenal calcium absorption in vitamin D receptor-knockout mice: functional and molecular aspects. *Proc Natl Acad Sci U S A* **98**, 13324-9. (2001).
38. Tsuchida, K., Mizushima, S., Takahashi, H., Misugi, N. & Soda, K. Relationship between dietary calcium and bone mineral density before menopause. *Nippon Koshu Eisei Zasshi 45*, 121-8. (1998).

Chapter 5

Vitamin D Receptor (VDR) Gene Haplotypes and Susceptibility of Fracture Risk

Chapter 5.1

Vitamin D Receptor (VDR) Gene Haplotypes and Fracture Risk: the Rotterdam Study

This chapter represents a part of the manuscript:
Fang Y, van Meurs JB, d'Alesio A, Jhamai M, Zhao H, Rivadeneira F, Hofman A, van Leeuwen JP, Jehan F, Pols HA, Uitterlinden AG
Promoter and 3'-Untranslated-Region Haplotypes in the Vitamin D Receptor Gene Predispose to Osteoporotic Fracture: The Rotterdam Study
Am J Hum Genet (2005) 77:807-823

ABSTRACT

Introduction: Most genetic association studies for the VDR focused on the analysis of individual polymorphisms of the vitamin D receptor (VDR) gene and complex diseases, but with conflicting results. We therefore recently determined sequence variation across the major relevant parts of the VDR gene, including construction of linkage disequilibrium (LD) blocks and identification of "haplotype tagging" single nucleotide polymorphisms (htSNPs).

Materials and Methods: We here analyzed Fok I, Bsm I, Apa I, Taq I RFLPs and 14 htSNPs in relation to 937 clinical fractures and 335 vertebral fractures recorded in 6,148 elderly Caucasians over a follow-up period of 7.4 years.

Results: Haplotype alleles of the 5' 1a/1e (block 2-hap1), 1b (block 3-hap3) promoter region and of the 3' untranslated region (UTR, block 5-hap1) were associated with increased fracture risk of 15% ($p = 0.06$), 74% ($p = 0.002$) and 23% ($p = 0.004$), respectively. For the 16% of subjects who had risk genotypes at all three regions, risk increased 48% for clinical fractures ($p=0.0002$), independent of age, gender, height, weight, and BMD. The population attributable risk varied between 1% - 12% for each LD block, and was 4% for the combined VDR risk genotypes.

Conclusions: This comprehensive candidate gene analysis demonstrates small but consistent effect of risk alleles of multiple VDR LD blocks on fracture risk.

INTRODUCTION

Rare deleterious mutations in the VDR gene cause the well known 1,25-dihydroxyvitamin D resistant rickets (rickets type II), a rare monogenetic disease characterised by osteomalacia, alopecia, and increased 1,25-$(OH)_2$ D_3 levels. Some isolated, more common VDR single nucleotide polymorphisms (SNPs) have previously been associated with several complex diseases and traits, such as osteoporosis[1-4]. Yet, most studies have used polymorphisms at the 3'end or in exon 2, but information on other polymorphism across the VDR gene is scarce. The relationship between disease and haplotype alleles across the VDR gene has not been systematically analyzed.

We recently sequenced 22 kb genomic VDR sequence, including the 3'untranslated region (3'UTR), all coding exons and the 6 promoter exons 1f – 1c (chapter 2), and determined linkage disequilibrium (LD; pair-wise D') between single nucleotide polymorphisms (SNPs) across and flanking the VDR gene in different ethnic populations, and also identified haplotype tagging SNPs (htSNPs) of the VDR gene (chapter 3). In the current study, we use this information in a large-scale association analysis of htSNPs in relation to osteoporosis in a group of 6,148 Caucasian elderly men and women.

MATERIALS AND METHODS

Subjects

The association of VDR genotype with fracture risk was analyzed in the Rotterdam Study with 7,983 subjects[5], of which 6,580 DNA were collected, and 6,148 DNA samples were available and succeeded for genotyping of all SNPs. Genomic DNA was isolated from blood according to standard procedures.

Genotyping

We genotyped Fok I, Bsm I, Apa I, Taq I RFLPs and 14 haplotype tagging SNPs (htSNPs) in 6,148 Caucasian subjects from the Rotterdam Study with the high throughput TaqMan allelic discrimination assays. The detail of genotyping assays are provided in the Materials and Methods section of chapter 3. Five percent random samples were independently repeated to confirm genotyping results. The disagreement rate varied from 0.3-1.2 % for five htSNPs, while genotype results of all other htSNPs were completely consistent.

Epidemiological analysis

Genotype distribution was tested for Hardy-Weinberg equilibrium. Anthropometric measurements, body mass index (BMI, kg/m^2), BMD (g/cm^2;

dual energy X-ray absorptiometry, Lunar DPX-L densitometer; Madison, WI, USA) at the femoral neck and lumbar spine (L2-L4) and other variables were measured as described before[6,7]. BMD measurements were available for 5,027 (82% of the cohort with genotypes) subjects. Gender-specific T-scores were calculated from the femoral neck BMD using the NHANES reference population[8] of Caucasian males and females aged 20 to 29 years. For assessing incidence of fracture, follow-up ended before 1st of January 2002 (mean follow-up period of 7.4 years ± 3.3 years). We defined "clinical fracture" as all fractures confirmed by general practitioners or hospitals as described before[4,7], and excluded head, foot, hand and pathological, post-procedural, skull and face fractures. Presence of vertebral fracture was analyzed as described previously[9,10].

We applied the Pearson χ^2 test, and calculated the Relative Risk and 95% confidence interval (RR, 95% CI) by logistic regression models, and calculated Hazard Ratio (HR, 95% CI) for incidence of clinical, wrist and hip fractures by Cox regression model. Both logistic regression and Cox models were adjusted for potential confounders, such as age, gender, height, weight, BMD and bone loss. The population attributable risk (PAR) was calculated for genetic and other markers for clinical fracture risk. All statistical analyses were done with the SPSS (11.0). The PAR was calculated as: $P \times (HR-1) / [P \times (HR - 1) + 1]$, where P is the proportion of the study population that is exposed to the risk factor for fracture and HR is the Hazard Ratio for the risk factor.

RESULTS

The 15 tagging SNPs were in Hardy-Weinberg Equilibrium. Haplotype alleles in block 2, 3, and 5 had consistent, albeit not always significant, effects on clinical fractures (Fig 1), with similar effects in men vs. women and for vertebral, hip and wrist fractures (data not shown). Subjects homozygous for the block 2 -hap1 allele, or homozygous for the block 3-hap3 allele, or carrying the block 5-hap1 allele had an increased risk for clinical fracture of 15% ($p = 0.06$), 74% ($p = 0.002$) and 23% ($p = 0.004$), respectively (Table 1). We went on to investigate the combined effect of promoter and 3'-UTR risk genotypes on fracture risk, and found that subjects carrying two or three risk genotypes across the VDR gene have 48% increased risk for clinical fracture ($p = 0.0002$).

Baseline characteristics of 6,148 Caucasian elderly of the Rotterdam Study are presented in Table 2. While age was borderline significantly different by combined VDR genotype, all associations with fracture are independent of age, gender, height, weight, bone loss (data not shown). We did not observe BMD to influence the association of individual or combined blocks with fracture risk.

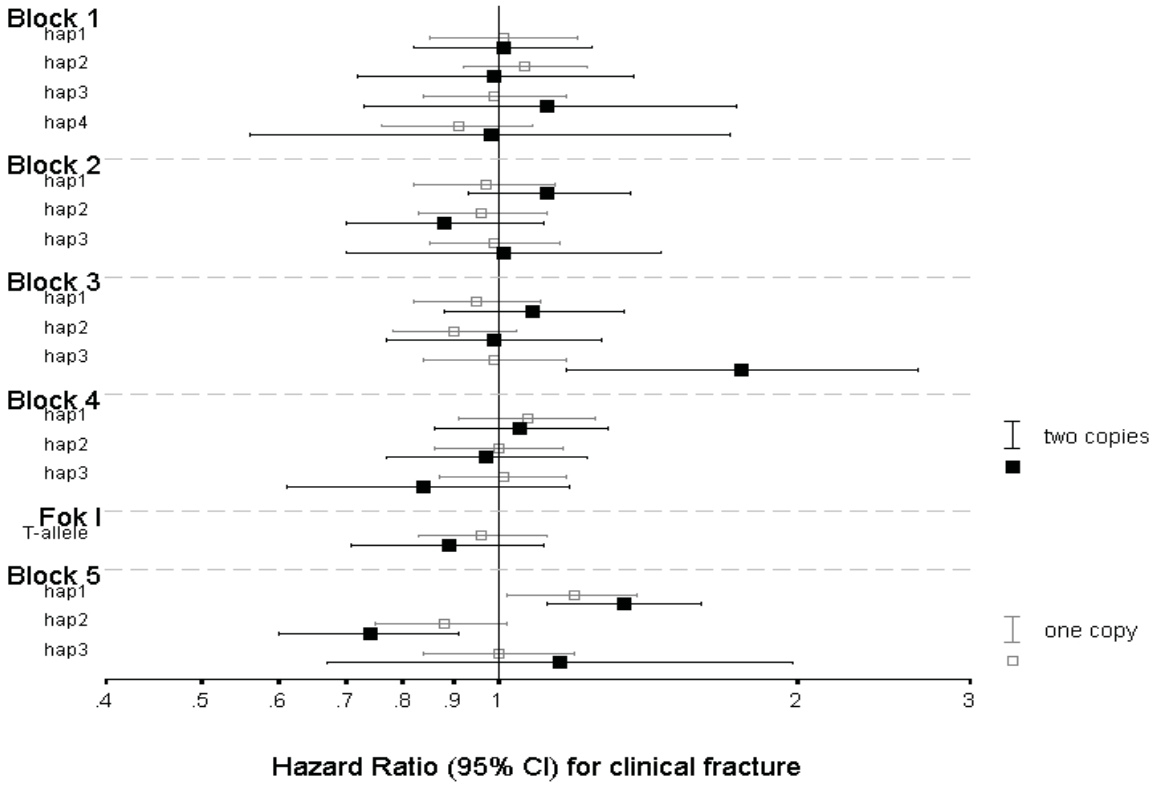


Figure 1 Hazard Ratio (HR) for clinical fracture by VDR genotypes based on haplotype alleles in five haplotype blocks (1-5) and the Fok I RFLP. The HR point estimate and the surrounding 95% Confidence Interval (95% CI) are represented with squares and lines. The HR for one copy versus no-copy of the test allele is represented with open squares and grey lines, the HR for two copies versus no-copy of the allele is represented with solid squares and black lines. The logarithmic HR is plotted for the frequent haplotype alleles (frequency > 3%) in all haplotype blocks (see Fig. 1b for Caucasians in Chapter 3) and the Fok I RFLP.

The population attributable risk (PAR) of the VDR genetic markers (Table 3) was 1% (frequency of subjects homozygous for the block 3-hap3 allele = 2%), and 2% (frequency of subjects homozygous for the block 2-hap1 = 21%) for the promoter region, but increased to 12% for the 3' region (including the 3'-UTR, frequency of the block 5 hap1 allele-carriers = 68%). This was higher than the 1% PAR of the best validated genetic marker for osteoporosis so far (COL1A1 Sp1[11,12] with a frequency of 3% of subjects homozygous for the Sp1 T-allele in our study population), and similar to PAR's for smoking and use of a walking aid.

DISCUSSION

A major limitation so far of association studies using VDR polymorphisms in relation to complex disease endpoints has been the small number of analyzed polymorphisms and, thus, the lack of knowledge about influence of and relation between other polymorphisms in the gene. In addition, the lack of statistical

Table 1 Incidence of fractures in 6,148 men and women from the Rotterdam Study by VDR genotype defined by haplotype allele status in blocks 2, 3 and 5

	Total	Genotype groups*		p-value
Clinical fx. by individual LD blocks:				
Block 2-hap1		**Non-Homozygous**	**Homozygous**	
Case/total (%)	906/6148 (14.7)	695/4844 (14.3)	211/1304 (16.2)	0.06
Crude HR (95% CI)		1	1.15 (0.97-1.35)	0.06
Block 3-hap3		**Non-Homozygous**	**Homozygous**	
Case/total (%)		876/6019 (14.6)	30/129 (23.3)	**0.006**
Crude HR (95% CI)		1	**1.74 (1.22-2.50)**	**0.002**
Block 5-hap1		**Non-Carrier**	**Carrier**	
Case/total (%)		256/1988 (12.9)	650/4160 (15.6)	**0.004**
Crude HR (95% CI)		1	**1.23 (1.07-1.42)**	**0.004**

Fractures by combined LD blocks:		**Number of risk genotypes in blocks 2, 3 & 5#**			Trend p-value
		Zero	**One**	**Two or Three**	
Case/Total (%)					
- Clinical fx.	906/6148 (14.7)	207/1586 (13.1)	516/3560 (14.5)	183/1002 (18.3)	**0.001**
- Vertebral fx.**	335/3055 (11.0)	78/814 (9.6)	186/1748 (10.6)	71/493 (14.4)	**0.01**
- Hip fx.	261/6148 (4.2)	61/1586 (3.8)	144/3560 (4.0)	56/1002 (5.6)	**0.05**
- Wrist fx.	257/6148 (4.2)	58/1586 (3.7)	145/3560 (4.1)	54/1002 (5.4)	**0.04**
Crude HR (95% CI)					
- Clinical fx.		1	1.14 (0.97-1.33)	**1.48 (1.21-1.80)**	**0.0002**
- Vertebral fx.		1	1.12 (0.85-1.48)	**1.59 (1.13-2.24)**	**0.01**
- Hip fx.		1	1.06 (0.79-1.43)	**1.49 (1.04-2.14)**	**0.04**
- Wrist fx.		1	1.13 (0.83-1.53)	**1.53 (1.06-2.21)**	**0.03**

*: "Non-homozygous" includes the genotype groups without the risk allele and the heterozygotes. "Homozygous" is the genotype group homozygous for the risk allele. "Non-Carrier" is the genotype group without the risk allele, "Carrier" includes genotypes heterozygous and homozygous for the risk allele
#: "Zero" indicates the "Non-Homozygous": genotype groups for blocks 2 and 3, and the "Non-Carrier" genotype groups for block 5
"One" indicates either the "Homozygous" genotype group for blocks 2 or 3 or the "Carrier" genotype group for block 5
"Two and three" indicates two or three of the risk genotype groups, i.e., the "Homozygous" genotype group for blocks 2 and 3, and "Carrier" for block 5
**: Vertebral fracture was diagnosed by the X-ray

Table 2 Characteristics of the study population

Characteristic*	a. Total Cohort	b. by combined promoter and 3'-UTR VDR genotype#			p-value
		0	1	2 & 3	
Number (%)	6148	1586 (25.8)	3560 (57.9)	1002 (16.3)	
Female (%)	3649 (59.4)	932 (58.8)	2102 (59.0)	612 (61.1)	0.49
Age (years)	69.5 ± 9.1	69.1 ± 9.1	69.5 ± 9.1	70.0 ± 9.4	0.03
Height (cm)	166.8 ± 9.5	167.0 ± 9.5	166.7 ± 9.4	166.6 ± 9.8	0.16
Weight (kg)	73.1 ± 12.0	73.3 ± 12.2	73.1 ± 12.0	72.9 ± 11.7	0.55
FN BMD (g/cm^2)	0.867 ± 0.142	0.872 ± 0.143	0.864 ± 0.142	0.868 ± 0.145	0.40
LS BMD (g/cm^2)	1.090 ± 0.198	1.098 ± 0.197	1.089 ± 0.198	1.082 ± 0.199	0.25
FN BMD change (10^{-3} g/cm^2/year)	-2.1 ± 9.4	-1.6 ± 8.6	-1.9 ± 9.5	-3.2 ± 10.0	0.02

*: The following adjustments were applied:
- Age: adjusted for gender
- Height: adjusted for age and gender
- Weight: adjusted for age, gender, and body height
- BMD: subset n = 5027, adjusted for age, gender, height, and weight
- BMD change: subset n = 2391, 7.4 years follow-up, adjusted for age, gender, and clinical fracture
#: Definition of the combined genotype is shown in Table 1.

Table 3 Population attributable risk of independent risk factors for clinical fracture in 6148 men and women from the Rotterdam Study

	Frequency (%) at baseline	HR (95% CI)*	PAR (%, 95% CI)*
Age > 75 years	32	2.3 (2.1-2.6)	29 (25-34)
T-score< -2.5	16	2.7 (2.3-3.1)	21 (17-25)
Current Smoking	23	1.3 (1.1-1.5)	6 (2-10)
Use of walking aid	12	1.3 (1.1-1.6)	3 (1-4)
Genetic Markers			
ColIA1 Sp1 T-allele homozygotes	3	1.4 (1.0-1.9)	1 (0-3)
VDR risk genotypes			
Block 2 hap 1	21	1.1 (0.9-1.3)	2 (0-5)
Block 3 hap 3	2	1.6 (1.2-2.4)	1 (0-3)
Block 5 hap 1	68	1.2 (1.1-1.4)	12 (4-21)
Carrier of two or three risk genotypes	16	1.3 (1.1-1.5)	4 (2-6)

* HR = Hazard Ratio, 95% CI = 95% confidence Interval; PAR = population attributable risk; all HR and PARs were adjusted for age and gender.

power of most studies to detect the expected subtle effects and misconceptions about how such small biological effects could be translated to the risk of disease has led to a number of controversies in the field. In our recent studies we identified 62 polymorphisms in potentially functional areas of the VDR gene, characterised the LD structure and identified htSNPs, and used 15 tagging SNPs to represent

the common haplotypes for five of LD blocks in potentially functional areas of the VDR gene. We here examined the association between those VDR haplotypes and fracture risk in a large-scale population.

A large association study with the htSNPs in a population-based cohort (6,148 elderly subjects) identified haplotype alleles in 3 LD blocks that confer risk for fracture, and we present evidence for intragenic additive effects. These relationships were found in both men and women and for different types of fractures (including vertebral, hip and wrist fractures), and were independent of age, height, weight, BMD and BMD change. In our study population (15% prevalence of fracture and 6,148 subjects), we have 80% statistical power to detect a 25% increased fracture risk for a polymorphism/haplotype with a frequency > 10%. We analyzed haplotypes instead of individual SNPs, which limited the number of tests. For block 2 and 5, we did not adjust the significant p-value by the conservative Bonferroni correction for multiple comparisons, because in our previous studies[3,4] we observed two fracture risk alleles from these blocks in a subset of our current study, and we confirmed the association in this study, but now for the haplotype alleles encompassing the previously observed risk alleles. However, we used $p = 0.005$ (0.05/11), based on a Bonferroni correction for analyzes in the other three blocks and Fok I, only the association between block 3-hap 3 and fracture risk remains border line significant ($p = 0.002$). Thus, we cannot reject the possibility that some of the associations we observed are false positives, even in this very large population.

We previously reported haplotype 1 of the 3'-end variants, defined by only Bsm-Apa-Taq RFLPs, to be associated with increased fracture risk in 1004 postmenopausal women[3], which is a subgroup of the current study population. We here confirm this association between Bsm-Apa-Taq haplotype 1 and fracture risk in the complete study population ($p = 0.03$, data not shown). Yet, in our more detailed haplotype analysis based on more SNPs we defined a sub-type of the Bsm-Apa-Taq haplotype 1, i.e., block 5-hap1 (40.2%, Fig. 1b for Caucasians in chapter 3), which shows a stronger and more significant association with fracture risk. These subtle differences in exact definition of the risk allele could contribute to heterogeneity in association results observed for different studies. While we here focussed on the more common (haplotype) alleles in the population, we cannot exclude (many) less frequent risk alleles to contribute to VDR genotype-dependent fracture risk. Further studies will be necessary to assess their contribution. We show that effects are modest (about 20-70% increased risk) as can be expected for common variants in relation to complex disease. The VDR risk haplotype alleles therefore have a modest influence on individual risk of fracture but make a substantial contribution at the population level (PAR = 4 – 12%) in comparison to other genetic markers we previously identified in this population, such as COL1A1 Sp1[12] and ESR1 Pvu II and Xba I polymorphisms[13]. The most prominent genetic effect on fracture risk according to the PAR analyzes is the block 5-hap1 risk allele. Many previous conflicting association studies on VDR and BMD as well as fracture,

analyzed usually (very) small study populations and used the Bsm I, Apa I or Taq I polymorphisms in this block, but mostly analyzed separately. We therefore suppose that this controversy can partly be explained by a lack of statistical power as a result of small sample size, and failure to use haplotypes. In addition, population stratification, such as mixed ethnic groups with different allele frequencies, population-specific differences of some environmental factors, such as (dietary) calcium intake, (dietary) vitamin D intake, sunlight-exposure, and other characteristics of the study population pertinent to bone metabolism and fracture risk could result in heterogeneity of association observed across different study populations. Based on a meta-analysis of published data on the relationship between the VDR Bsm I RFLP and fracture risk we have some evidence that this could be an important factor (chapter 5.2). We have data on serum vitamin D and dietary calcium intake in a subset of our study population (n=1,312 and 4,747, respectively), but we don't have reliable dietary vitamin D intake data. We note that our study population has a very high dietary calcium intake (1,120 mg/day). For the VDR-fracture relationship we repeated the analyses but now stratified by the median, tertiles and quartiles of vitamin D level or dietary calcium intake. However, the association did not differ in these strata, and no interaction was observed (data not shown).

VDR as a transcription factor influences the expression of down-stream genes, such as TRPV5, TRPV6, and Calbindin, which are involved in calcium absorption, which could therefore impact BMD and thus fracture risk. However, we found that the association between VDR genotype and fracture risk was BMD-independent, which suggests other mechanisms (such as bone micro-architecture, bone quality and bone strength) to determine fracture risk. Alternatively, we can hypothesize that the bone tissue of the subjects carrying a risk allele is thought to have a somewhat lower sensitivity to vitamin D since the expression of VDR is lower. Therefore, the osteoblast activity could be lower, and the bone formation rate (in the bone remodelling balance) could be decreased. The age-related expansion of the outer diameter of long bone is associated with a marked increment of bone strength. We recently have observed that the fracture risk haplotype alleles were also associated with decreased bone size (chapter 6) in the same population. This genotype-related bone geometry difference reflects that the bone gain resulting in outer bone diameters expanding is smaller than the bone loss resulting in inner bone diameter expanding. This bone geometry difference leads to a decrease in bone strength for the risk allele carriers, possibly increasing the fracture risk.

In conclusion, basing on our recent sequencing and LD analysis of the VDR gene, we demonstrated polymorphisms in the 5' promoter region and the 3'-UTR of the VDR gene to contribute to fracture risk in a large population. This "whole gene" analysis demonstrates intragenic interaction of VDR polymorphisms in the promoter region and in the 3'UTR to significantly contribute to fracture risk. Our finding suggests these risk alleles to encode a "less-sensitive" VDR, probably by encoding lower numbers of VDRs on target cells.

ACKNOWLEDGEMENTS

We thank all participants of the Rotterdam Study and the general practitioners, pharmacists, and the many field workers at the research center in Ommoord, Rotterdam, The Netherlands.

This project was funded by the Dutch Research Organisation (NWO 903-46-178, 925-01-010, 014-90-001 and 911-03-012) and the European Commission under grant QLK6-CT-2002-02629 ("GENOMOS").

REFERENCES

1. Morrison, N.A. et al. Prediction of bone density from vitamin D receptor alleles. *Nature* **367**, 284-7 (1994).
2. Young, R.P., Lau, E.M., Birjandi, Z., Critchley, J.A. & Woo, J. Interethnic differences in hip fracture rate and the vitamin D receptor polymorphism. *Lancet* **348**, 688-9 (1996).
3. Uitterlinden, A.G. et al. Interaction between the vitamin D receptor gene and collagen type Ialpha1 gene in susceptibility for fracture. *J Bone Miner Res* **16**, 379-85 (2001).
4. Fang, Y. et al. Cdx-2 polymorphism in the promoter region of the human vitamin D receptor gene determines susceptibility to fracture in the elderly. *J Bone Miner Res* **18**, 1632-41 (2003).
5. Hofman, A., Grobbee, D.E., de Jong, P.T. & van den Ouweland, F.A. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur J Epidemiol* **7**, 403-22 (1991).
6. Burger, H. et al. The association between age and bone mineral density in men and women aged 55 years and over: the Rotterdam Study. *Bone Miner* **25**, 1-13 (1994).
7. Schuit, S.C. et al. Fracture incidence and association with bone mineral density in elderly men and women: the Rotterdam Study. *Bone* **34**, 195-202 (2004).
8. Looker, A.C. et al. Updated data on proximal femur bone mineral levels of US adults. *Osteoporos Int* **8**, 468-89 (1998).
9. McCloskey, E.V. et al. The assessment of vertebral deformity: a method for use in population studies and clinical trials. *Osteoporos Int* **3**, 138-47 (1993).
10. Van der Klift, M., De Laet, C.E., McCloskey, E.V., Hofman, A. & Pols, H.A. The incidence of vertebral fractures in men and women: the Rotterdam Study. *J Bone Miner Res* **17**, 1051-6 (2002).
11. Grant, S.F. et al. Reduced bone density and osteoporosis associated with a polymorphic Sp1 binding site in the collagen type I alpha 1 gene. *Nat Genet* **14**, 203-5 (1996).
12. Uitterlinden, A.G. et al. Relation of alleles of the collagen type Ialpha1 gene to bone density and the risk of osteoporotic fractures in postmenopausal women. *N Engl J Med* **338**, 1016-21 (1998).
13. van Meurs, J.B. et al. Association of 5' estrogen receptor alpha gene polymorphisms with bone mineral density, vertebral bone area and fracture risk. *Hum Mol Genet* **12**, 1745-54 (2003).

Chapter 5.2

Vitamin D Receptor (VDR) Gene Bsm I Polymorphism and Fracture Risk: A Meta-analysis

Yue Fang, Fernando Rivadeneira, Joyce B. J. van Meurs, Huib A. P. Pols, John P. A. Ioannidis, André G. Uitterlinden

Submitted

ABSTRACT

Introduction: Fracture is the major clinical outcome of osteoporosis. The vitamin D receptor (VDR) gene is thought to be a candidate gene for osteoporosis. Many genetic studies have suggested an association of VDR polymorphisms and osteoporosis, but evidence remains conflicting.

Materials and Methods: We searched published studies from 1996 to September 2005 through PubMed, and evaluated the genetic effect of Bsm I polymorphism of VDR on fracture risk in a meta-analysis. Thirteen studies with a total of 20 eligible comparisons (1,632 fracture cases and 5,203 controls) were analyzed with fixed and random effects models.

Results: No evidence of relationship between VDR Bsm I polymorphism and fracture risk was observed with any genetic model. The odds ratio (95% confidence interval) of b-allele versus B-allele was 0.98 (0.86 – 1.12) with random effects calculations. There was significant between-study heterogeneity. Small studies did not differ significantly from larger ones.

Conclusions: No relationship of VDR Bsm I polymorphism and fracture risk was found in the meta-analysis of published data.

Keywords: VDR; Bsm I genotype; fracture; meta-anlaysis

INTRODUCTION

Osteoporosis is a common complex genetic disease, characterized by decreased bone mineral density (BMD), increased bone fragility and fracture risk. Vitamin D plays a crucial role in calcium and phosphate homeostasis and skeletal metabolism. The vitamin D receptor gene (VDR) mediates the action of its ligand, and results in normal bone mineralization and remodeling, therefore VDR has been considered an important candidate gene of osteoporosis. Since 1994 when Morrison et al.[1] reported the relationship between VDR polymorphisms and BMD, about a hundred subsequent studies have been carried out on the association between VDR polymorphisms and diverse outcomes of osteoporosis, including BMD and fracture. At least three systematic reviews[2-4] have tried to summarize association studies between VDR polymorphisms and BMD. However, analyzing the association between VDR polymorphisms and BMD may be not sufficient to estimate the role of VDR in osteoporosis. Bone fragility also depends on the morphology, micro-architecture, and remodeling of bone. Some studies[5,6] also showed that the relationship of VDR polymorphisms and fracture risk was independent of BMD.

Fracture is a relevant clinical outcome to assess in genetic studies of osteoporosis. Polymorphisms in the 3'-end of the VDR gene, especially Bsm I and Taq I RFLPs, have most commonly been used in association studies to date. These are highly linked to polymorphisms in 3'-UTR, which is thought to be the potential functional region involving the stability of VDR mRNA. Several association studies have shown conflicting results of the relationship between VDR polymorphisms and fracture risk. The reason of the controversy could be due to small study size, diverse study design and different populations. So far, no meta-analysis has systematically reviewed the published data of on VDR polymorphisms and the risk of osteoporotic fracture. In the current study, we performed a meta-analysis of all studies relating Bsm I or Taq I polymorphisms of VDR gene with fracture risk.

MATERIALS AND METHODS

Identification of studies

We searched publications on PubMed (1996 to September 2005) with the combination of keywords: VDR AND (polymorphism OR genotype) AND fracture. References from retrieved publications were checked for additional studies. This search tracked 40 studies, of which 18 studies (all in English) addressed the association between fractures and VDR genotype. Others were reviews (n = 4), or association and meta-analysis studies of VDR polymorphisms with other outcomes such as BMD, bone turnover, muscle strength, hormone therapy and vitamin D supplementation (n = 18), but no original fracture data per VDR genotype were available. We also reviewed 108 references from three meta-analysis studies[2-4] of

VDR polymorphisms and BMD, including a meta-analysis on other bone phenotypes (Gong et al.[3]). These 108 references included 59 published original association studies, and 49 abstracts. Five references contained fracture data by VDR genotypes, but they were already included in the 18 studies of fractures.

Of the 18 potentially eligible studies, further scrutiny suggested that 14 could be included in the meta-analysis. This included eight studies for Bsm I restriction fragment length polymorphisms (RFLPs)[7-14], one for Taq I RFLP[15], two for Bsm-Apa-Taq (B-A-T) haplotype[5,16], and three studies investigated more than one polymorphisms (Fok I, Bsm I, Apa I, Taq I or B-A-T haplotype) of the VDR for fracture risk[17-19]. Conversely, we excluded from the meta-analysis two studies that investigated the Cdx-2[6] and Fok I[20] polymorphisms in relation to fracture risk, respectively, and two ecological studies[21,22] of Bsm I and Taq I. Of the 14 remaining reports, two reports presented overlapping data from the same population[5,16], and we retained for the meta-analysis only the study with the most extensive data. Thus, 13 studies were finally included in our meta-analysis (Table 1). The study of Ramalho et al.[9,14] had two control groups, which we combined in one.

Data Extraction

Data were extracted according to agreement of two reviewers (Y. Fang and F. Rivadeneira), and the third investigator (A. Uitterlinden) arbitrated the disagreement by discussion and consensus. We extracted information on authors, published year, country of origin, study design, mean or range of age, gender, site of fracture (e.g. vertebral, non-vertebral, hip, forearm, low-energy and any fracture), number of cases or controls and frequency of B-allele in cases, controls, and the whole population for each study (Table 1). We also examined whether the genotype distributions of study populations followed Hardy-Weinberg equilibrium (HWE) proportions using a χ^2-test[23].

Data Analysis

Three polymorphisms, Bsm I, Apa I and Taq I, have been commonly analyzed in published association studies, and Bsm I is most consistently reported. They are highly linked to each other with pair-wise D's from 0.95 to 0.98, according to our recent study of 6,500 subjects[24]. With a direct haplotyping test, the "B"-allele of Bsm I has been shown to be 99% linked to the "t"-allele of Taq I, but only 41% linked to the "A"-allele of Apa I in a large-scale Caucasian population[25]. We therefore considered Bsm I and Taq I as interchangeable genetic markers of the VDR gene (at least in the Caucasian population). In the current study, we focused on analyzing the relationship of Bsm I and fracture. When Bsm I genotype was absent, Taq I genotype was used to infer the Bsm I genotype.

For the meta-analyses we used the Cochrane Review Manager (RevMan Version4.2 program, the Cochrane Collaboration, Oxford, UK). The association between the Bsm I polymorphism and fracture was assessed under the following

Table 1 Studies of VDR Bsm I (or Taq I or B-A-T haplotype) genotype and fracture risk included in meta-analysis

Author, year	Country	Study design	Mean age (range), yrs	Gender	Fracture type	Eligible subjects (n)		Studied Polymorphism	B-allele freq. (%)*	p-value (HWE)*
						Case	Control			
Berg, 1996	Norway	Cohot	65	Women	Any	19	30	Bsm I	46	0.15
Houston, 1996	Scotland	Case-control	60 (45-88)	Women	Vertebral	44	44	Bsm I	45	0.44
Ramalho, 1998	Brazilian	Case-control	32–78	Women and men	Hip	73	148	Bsm I	34	0.78
Feskanich, 1998	USA	Case-control	43-69	Women	Hip	54	108	Bsm I	39	0.77
				Women	Forearm	163	163		43	0.15
Gomez, 1999	Spain	Cohort	67± 8	Men	Any	39	114	Bsm I	41	0.59
				Women		37	122		37	0.24
Ensrud, 1999	USA	Cohort	72 (65-)	Women	Hip	181	693	Apa I, Taq I	44	-
				Women	Vertebral	127	482		43	-
				Women	Other	223	368		44	-
Aerssens, 2000	Belgium	Case-control	77 (60-95)	Women	Hip	135	239	Bsm I	48	0.46
Langdahl, 2000	Denmark	Case-control	51± 16	Men	Vertebral	30	73	Fok I, Bsm I, Apa I, Taq I	40	0.67
				Women	Vertebral	80	80		52	0.18
Valimaki, 2001	Finland	Case-control	53 (50-56)	Women	Low-energy	57	315	Bsm I	43	0.53
			89 (85-98)	Women	Hip	64	108		34	0.25
Uitterlinden, 2001	Netherlands	Cohort	67 (55-80)	Women	Any	97	905	Haplotype (B-A-T)	41	0.09
Alvarez-Hernandez, 2003	Spain	Cohort	64 ± 9	Men	Vertebral	20	134	Bsm I, Apa I, Taq I, haplotype	41	0.69
Garnero, 2005	France	Cohort	61± 8	Women	Vertebral	34	589#	Bsm I	40	0.71
				Women	Non-vertebral	86	589#		40	0.71
Nguyen, 2005	Australia	Cohort	70 ± 7	Women	Hip	69	608	Taq I	40	0.74
TOTAL						1632	5203			

*: Frequency of B-allele in entire population of cohort studies or control group of case-control studies
#: The same control group was used in the original study, and is counted as one control group in the table

genetic models, which were treated as a dichotomous variable: (a) b-allele versus B-allele for the allele level comparison; (b) bb versus Bb + BB for recessive model of b-allele, (c) bb + Bb versus BB for dominant model of b-allele, and (d) bb versus BB for extreme genotypes. The study of Ensrud et al. only presented fracture data by b or B alleles, but not by genotype in the original report, hence it was only included in the analysis of model (a). We calculated fracture type-specific and overall odds ratios (ORs) and 95% confidence intervals (95% CIs) for all pair-wise comparisons. The p-value cutoff level, adjusted for multiple comparison, was considered significant when lower than 0.0125 (Bonferroni corrected: $p = 0.05/4$, to consider pair-wise comparisons in the four genetic models).

Assessment of heterogeneity and bias diagnostics

Between-study heterogeneity for difference of fracture risk was tested using the χ^2 – distributed Q statistic (with p-value considered significant at a level of 0.10[26]), and the inconsistency index I^2 (suggesting inconsistency among the studies' results with values of 50%, or higher and large heterogeneity for values of 75% or higher[27] estimated by the RevMan program. ORs (95% CIs) were estimated by both fixed effect and random effects models. Fixed effects assume that the genetic factor shows a similar effect on fracture across all investigated study groups, and the observed variation among study groups is caused by chance alone. Random effects assume that there may be substantial diversity among different study groups, and assesses both within-study sampling error and between-study variance. When there is no heterogeneity between study groups, the two models are similar; otherwise, random effects model usually give wider CIs than the fixed effect model. Random effects are preferable in the presence of significant between-study heterogeneity. In order to clarify the source of potential heterogeneity, we also performed subgroup analyses by type of fracture (vertebral, hip, any fractures), type of study (cohort and case-control studies) design and gender.

We also evaluated whether smaller studies gave different results from larger studies and using: the Begg-Mazumdar test[28], which is based on the Kendall's tau rank correlation coefficient between the natural logarithm of OR [ln(OR)] and its variance; and its regression equivalent[29], which is a linear regression of ln(OR) on its standard error (SE). We also performed recursive cumulative meta-analysis to evaluate whether the summary odds ratio changed in the same direction over time and appraised whether the first published study gave different results from subsequent ones[30-32].

RESULTS

Eligible data

The 13 eligible studies included 1,632 fracture cases and 5,203 controls with genotype data (Table 1). Some studies considered separate data for different types of fractures or different gender, thus there were 20 comparisons available. Two comparisons from the Feskanich et al.[10] study (for hip and forearm fractures, respectively) have no overlap in case and control groups, since first-reported fractures were defined as cases and independent controls for hip and forearm fractures were selected. However, there is some overlap in the control groups of Ensrud et al.[17] who studied different type of fractures, because control groups were selected from the same pool of a cohort. Two comparisons (vertebral vs. non-vertebral) from the Garnero et al.[14] study used the same cohort, and cases could suffer more than two times (or types) of a fracture. We cannot clarify the exact extent of overlap according to the original data of these studies[14,17]. The mean age for each study was 51 – 89 years. Eleven of the comparisons originated from cohort studies and nine from case-control designs, and sample size per comparison varied from 49 to 1,002. The frequency of the B-allele in the entire population of cohort studies or control group of case-control studies ranged from 34 – 52%, and we observed highly significant differences across all study groups ($p < 0.001$). All populations were of Caucasian descent and most of them were peri- or postmenopausal women. Three studies[11,18,19] contained separate data for men (mean age from 51 – 67 years), and one study[9,14] did not show separate data by gender. The genotype frequencies of all studies did not deviate significantly from HWE proportions.

Data synthesis-main analysis

We first compared the b-allele frequency difference in fracture cases and controls (Fig. 1). There was significant between-study heterogeneity when all comparisons were considered (Q statistic χ^2-test $p = 0.003$, $I^2 = 54\%$). No relationship between the b-allele and different types of fracture was observed in any analyses. The summary OR (95%CI) was 0.98 (0.86-1.12) for random effects model. Summary ORs (95%CI) of fixed and random effects models with various genetic models are shown in Table 2. The study of Ensrud et al., which only presented fracture data by b-allele versus B-allele, but not by Bsm I genotype from the original study, was only included in the analysis by allele. There was no evidence for the effect of Bsm I polymorphism of the VDR gene and fracture risk in any of the genetic models.

Subgroup analyses

When we stratified fracture as vertebral, hip and any type of fractures [non-vertebral fracture data from Garnero et al.[14] was included in the analysis of any fracture] according to the original report, no evidence of heterogeneity was found for vertebral fracture, but modest heterogeneity still remained in hip and any type

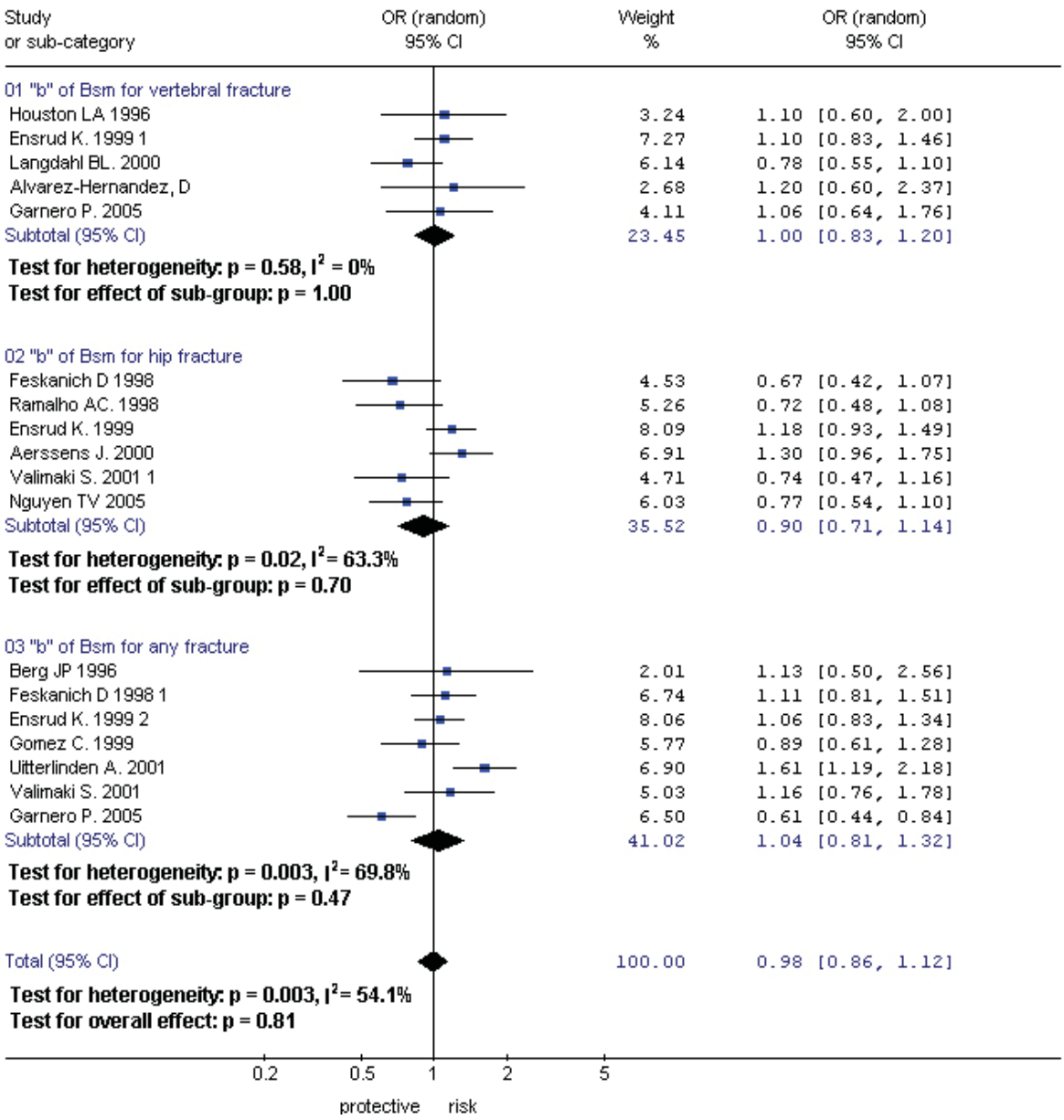


Figure 1 Meta-analysis of relationship between the b-allele of the VDR Bsm RFLP and fracture risk stratified by type of fractures. Point estimates and 95% confidence intervals (CI) are shown for the odds ratio (OR) in each study group. Summary estimates of OR and their 95% CI (diamonds), test for heterogeneity and effect are given by random effects models by vertebral, hip and any fractures, as well as for the overall database.

of fracture sub-group analyses.

When we stratified the analyses by type of fracture in women, similar results were observed in each sub-group (Table 3) as compared to the results presented in Figure 1. No significant heterogeneity was found in men, which only contained three study groups. We also stratified the meta-analysis by cohort and case-control designs. Eleven comparisons containing 4,857 subjects (with 932 fractures) were derived from cohort association studies, and significant heterogeneity was observed among those studies ($p = 0.005$, $I^2 = 62\%$). Nine comparisons containing 1,978 subjects (with 700 fracture cases) were derived from case-control studies, and some

heterogeneity was found also among those studies (p = 0.09, I^2 = 44%).

Table 2 Summary odds and 95% confidence intervals for various comparisons

Contrast	Study groups	Sample size	Summary OR (95% CI)		Heterogeneity	
	(n)	(n)	Fixed effect	Random effects	p-value	I^2 (%)
b vs. B (alleles)	18	13670	1.01 (0.93-1.10)	0.98 (0.86-1.12)	0.003	54
bb vs. Bb + BB*	15	4761	0.98 (0.84- 1.14)	0.97 (0.79-1.20)	0.03	45
bb + Bb vs. BB*	15	4761	0.90 (0.75-1.08)	0.86 (0.66-1.11)	0.03	45
bb vs. BB*	15	1934	0.91 (0.74-1.11)	0.87 (0.62-1.21)	0.003	57

*: The data of Ensrud et al. was not included (no fracture data by Bsm I genotype were available).

Table 3 Subgroup analyses for "b" vs. "B" comparison

Strata	Comparisons	Sample size	Summary OR (95% CI)		Heterogeneity	
	(n)	(n)	Fixed effect	Random effects	p-value	I^2 (%)
Type of fracture for women*						
Vertebral	4	2720	1.07 (0.88-1.31)	1.07 (0.88-1.31)	0.99	0
Hip	5	4515	0.97 (0.85- 1.12)	0.94 (0.72-1.22)	0.02	65
Any	7	6111	1.04 (0.92-1.19)	1.01 (0.78-1.32)	0.002	72
Summary	16	13346	1.04 (0.95-1.13)	1.00 (0.88-1.15)	0.005	55
All fracture for men*	3	820	0.92 (0.66-1.30)	0.92 (0.58-1.44)	0.18	41
Type of study						
Cohort	11	9714	1.04 (0.94-1.16)	1.02 (0.85-1.22)	0.005	62
Case-contol	9	3956	0.96 (0.84-1.10)	0.94 (0.78-1.13)	0.09	44

In the analysis by gender stratification, Ramalho et al. was excluded (data could not be separated by gender)

Other bias and heterogeneity diagnostics

In our meta-analysis of allele model, there is no evidence of significant differences in the genetic effects of large versus small studies based on tests of rank correlation (p = 0.78) and regression on the SE (p = 0.86). Two studies, Uitterlinden et al. and Garnero et al. showed statistically significant effects on their own, but these were in the opposite direction (Figure 1). We then carried out a cumulative meta-analysis (Figure 2). Starting with the first study [7], we added stepwise the subsequent studies, and observed that no cumulative step showed statistical significance until the end of the meta-analyses and the summary effect changes were rather small at all steps.

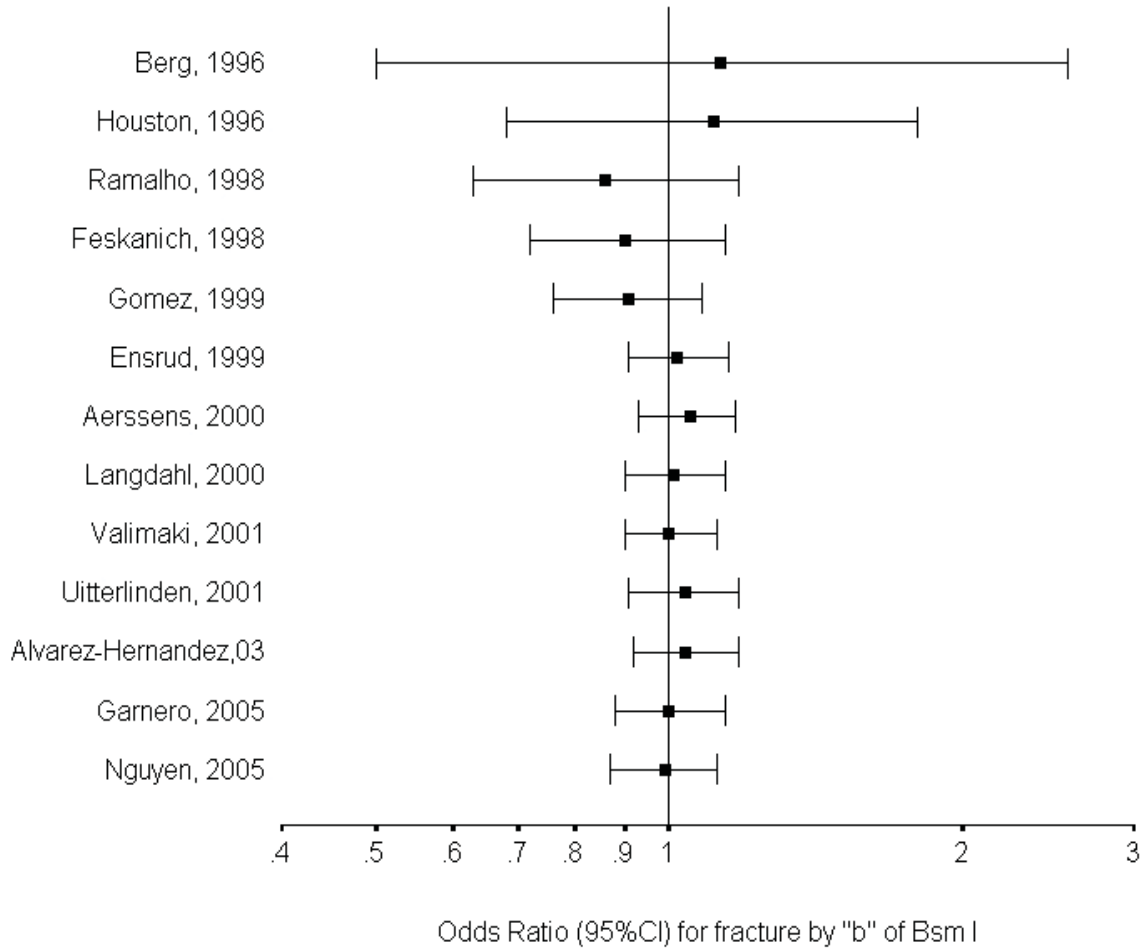


Figure 2 Cumulative meta-analysis. By each study, the cumulative odds ratio point estimates and the surrounding 95% CI for fracture by "b" allele of Bsm I of the VDR gene are represented with squares and lines. The cumulative estimate is calculated with random effects model.

DISCUSSION

Association studies are a powerful tool to identify genetic factors conferring susceptibility to common disorders like osteoporosis[33]. Yet, most reported association studies are underpowered to detect the modest genetic effects underlying the genetic susceptibility to develop common diseases[34]. This has resulted in the continuous reporting of inconsistent findings in the literature that could be due to false positive studies, false negative studies or true variability in associations among different populations[34]. A systematic meta-analytic approach may permit the estimation of population-wide effects and the identification of sources of variability of genetic risk factors in human disease [31]. In the current study, we combined the data of 6,835 subjects from published studies to evaluate the genetic association of the most commonly used polymorphism of the VDR gene, Bsm I, to fracture

risk. The results of our meta-analyses show there is no evidence supporting an association between the VDR Bsm I polymorphism and the risk of osteoporotic fracture. Moreover, the tight CIs exclude a 15% increase or decrease in the relative risk of fractures.

We have found considerable between-study heterogeneity which is highlighted by the fact that two studies have provided significant associations in the opposite direction[35]. The definition of osteoporotic fracture is difficult to standardize and the definition can be influenced in several different ways depending on diagnosis (clinical versus X-ray confirmed), type of trauma (low-energy versus high-impact), history of fracture (prevalent versus incident) and location (vertebral, non-vertebral, fragility, hip or wrist). However, most studies included in our meta-analysis did not provide sufficient information to address all those issues. When we stratified our meta-analyses by type of fracture, we observed that heterogeneity was more prominent for the studies of hip and any fractures, while data on vertebral fracture were more consistent across studies.

Another potential source of heterogeneity is differences in study design. However, there was no strong evidence that cohort studies had less or more heterogeneity than case-control studies. Also, we did not find strong evidence to conclude that gender is a source of heterogeneity in our meta-analysis: there was no observed association in either gender. Nevertheless, the number of studies in men was limited. Other possible sources of heterogeneity are differences in the study population characteristics including differences in age, ethnicity, serum vitamin D level, Ca^{++}-intake, smoking and physical activity among others.

Thakkinstrian et al.[4] had indicated that differences in the study setting (population-based and non-population-based) might be an important source of heterogeneity affecting the strength of the detected association between Bsm I polymorphism and BMD. However, full dissection of heterogeneity would require both large-scale evidence and detailed data on individual subjects. The GENOMOS consortium[36], the largest case collection yet reported for osteoporosis, is an example of a collaborative effort where such reliable answers in this field may be pursued[37].

Genetic heterogeneity in the form of differences in allele frequencies of analyzed polymorphisms and the presence of admixture within the study populations could also be a source of heterogeneity influencing the conclusion of our meta-analysis. Another potential explanation for differences in allele frequency is genotyping error, caused by one SNP present in the reverse primer sequence of Bsm I genotyping PCR[38]. This could have contributed to the modest variation in allele frequencies we observed across studies.

The VDR gene is a large and complex gene extending a size of 105 kilo base pairs, encompassing more than 120 single nucleotide polymorphisms (SNPs) according to Celera and NCBI databases, and seven haplotype blocks across the gene have been determined[24]. Thus the Bsm I polymorphism might not completely represent the genetic effect of the whole VDR gene on fracture risk. The genetic

effect of "whole gene" haplotype tagging SNPs on fracture need to be carried out in large-scale populations[24] or consortia[36].

Previous meta-analyses have focused on analyzing the genetic effect of the Bsm I polymorphism on BMD. Three studies[2-4] have summarized the relationship, but conflicting conclusions were drawn and some questions have remained from those analyses. Cooper et al.[2] reported that the BB genotype of the Bsm I was associated with low BMD only at the hip, and that younger women with BB genotype had borderline significant lower BMD compared to women with the bb genotype. Nevertheless, this effect disappeared after excluding the data with clearly documented genotyping error[1]. That study only analyzed one pair-wise comparison, BB versus bb, but did not involve other genetic models. A model of comparing extreme genotypes perhaps has insufficient statistical power to estimate a small genetic effect in relatively small study populations. In addition, different ethnic groups were mixed in the analysis, and frequencies of the B-allele and BB genotype differed across studies (1.4 – 31.3% of BB). Furthermore, three included small studies of northern Europe Caucasian population had low BB genotype frequencies (7.1-12.2%), while, in our current meta-analysis the frequencies of the BB genotype from the same region have been between 18.0 – 26.1% across studies[7,13,18].

The meta-analysis of Gong et al.[3] summarized 75 studies on the association between Bsm I, Apa I, Taq I and Fok I RFLPs of VDR gene with BMD as well as other related skeletal phenotypes. The authors grouped all studies as "positive" and "non-positive", where the "positive" study was defined as studies with association between either one of the b, a, T, or F alleles and increased BMD or other phenotypes (see below). This approach has some drawbacks. It simply counts the "positive rate", but it ignores the study size and magnitude of the genotype effect on BMD. Furthermore, the inclusion of all polymorphisms cannot be justified based on current knowledge. According to our recent data[24], Fok I RFLP is clearly not linked to Bsm I, Apa I and Taq I polymorphisms. The latter three polymorphisms are in the same haplotype block of the VDR gene. Bsm I and Taq I are highly linked to each other, but Apa I has relatively low linkage to the other two. Therefore, association studies of Bsm I and Taq I can be merged as one genotype group, but studies on Fok I and Apa I polymorphisms should be analyzed as different genotype groups, or employ Bsm-Apa-Taq haplotype in the association analyses. Some phenotypes in the study of Gong et al. were not always in line with BMD, such as, bone geometry variables, response to vitamin D treatment, bone turnover and fracture, and they should not be considered as the same phenotype group as BMD.

In another study, Thankkinstian et al.[4] observed that the BB genotype had lower spine BMD than Bb + bb genotypes but only in studies of postmenopausal women with a big range of B-allele frequency (29 – 53%). Moreover, there was considerable heterogeneity in this analysis. So far, almost one hundred association studies and meta-analyses, which adjusted for or stratified by age, gender, menopausal status, different measured instruments and different locations, show that Bsm I genotype

is not consistently associated with BMD, and more recent studies[39,40] on genome-wide linkage scans of BMD have not found the VDR locus (12q12) linked to BMD. Alltogether, this indicates that there is also no clear evidence of a genetic effect of Bsm I RFLP on BMD either.

In the current meta-analysis, we only included published English-language association studies of Bsm I and fracture. Perhaps we missed some studies, especially with negative results, or studies on BMD but that did not report on fracture data. Missing studies with negative results, however, would probably not change our conclusions. Such data would most likely reinforce our observations about the lack of an effect.

In conclusion, no evidence of a relationship between VDR Bsm I polymorphism and fracture risk was observed in the meta-analysis of published data. Further research should examine the effect of other haplotype tagging SNPs in large-scale populations or consortia.

ACKNOWLEDGEMENTS

This project was funded by the Dutch Research Organisation (NWO 903-46-178, 925-01-010, 014-90-001 and 911-03-012) and the European Commission under grant QLK6-CT-2002-02629 ("GENOMOS").

We gratefully acknowledge Dr. Fakhredin Ali Sayed Tabatabaei for initial help with the analysis and helpful comments.

REFERENCE

1. Morrison, N.A. et al. Prediction of bone density from vitamin D receptor alleles. *Nature* **367**, 284-7 (1994).
2. Cooper, G.S. & Umbach, D.M. Are vitamin D receptor polymorphisms associated with bone mineral density? A meta-analysis. *J Bone Miner Res* **11**, 1841-9 (1996).
3. Gong, G. et al. The association of bone mineral density with vitamin D receptor gene polymorphisms. *Osteoporos Int* **9**, 55-64 (1999).
4. Thakkinstian, A., D'Este, C., Eisman, J., Nguyen, T. & Attia, J. Meta-analysis of molecular association studies: vitamin D receptor gene polymorphisms and BMD as a case study. *J Bone Miner Res* **19**, 419-28 (2004).
5. Uitterlinden, A.G. et al. Interaction between the vitamin D receptor gene and collagen type Ialpha1 gene in susceptibility for fracture. *J Bone Miner Res* **16**, 379-85 (2001).
6. Fang, Y. et al. Cdx-2 polymorphism in the promoter region of the human vitamin D receptor gene determines susceptibility to fracture in the elderly. *J Bone Miner Res* **18**, 1632-41 (2003).
7. Berg, J.P., Falch, J.A. & Haug, E. Fracture rate, pre- and postmenopausal bone mass and early and late postmenopausal bone loss are not associated with vitamin D receptor genotype in a high-endemic area of osteoporosis. *Eur J Endocrinol* **135**, 96-100 (1996).
8. Houston, L.A., Grant, S.F., Reid, D.M. & Ralston, S.H. Vitamin D receptor polymorphism, bone mineral density, and osteoporotic vertebral fracture: studies in a UK population. *Bone* **18**, 249-52 (1996).

9. Ramalho, A.C. et al. Fractures of the proximal femur: correlation with vitamin D receptor gene polymorphism. *Braz J Med Biol Res* **31**, 921-7 (1998).
10. Feskanich, D. et al. Vitamin D receptor genotype and the risk of bone fractures in women. *Epidemiology* **9**, 535-9 (1998).
11. Gomez, C. et al. Vitamin D receptor gene polymorphisms, bone mass, bone loss and prevalence of vertebral fracture: differences in postmenopausal women and men. *Osteoporos Int* **10**, 175-82 (1999).
12. Aerssens, J. et al. Polymorphisms of the VDR, ER and COLIA1 genes and osteoporotic hip fracture in elderly postmenopausal women. *Osteoporos Int* **11**, 583-91 (2000).
13. Valimaki, S. et al. Relation of collagen type I alpha 1 (COLIA 1) and vitamin D receptor genotypes to bone mass, turnover, and fractures in early postmenopausal women and to hip fractures in elderly people. *Eur J Intern Med* **12**, 48-56 (2001).
14. Garnero, P., Munoz, F., Borel, O., Sornay-Rendu, E. & Delmas, P.D. Vitamin D Receptor gene polymorphisms are associated with the risk of fractures in postmenopausal women, independently of bone mineral density. The OFELY study. *J Clin Endocrinol Metab* (2005).
15. Nguyen, T.V. et al. Contribution of the Collagen I Alpha 1 and Vitamin D Receptor Genes to the Risk of Hip Fracture in Elderly Women. *J Clin Endocrinol Metab* (2005).
16. Colin, E.M. et al. Interaction between vitamin D receptor genotype and estrogen receptor alpha genotype influences vertebral fracture risk. *J Clin Endocrinol Metab* **88**, 3777-84 (2003).
17. Ensrud, K.E. et al. Vitamin D receptor gene polymorphisms and the risk of fractures in older women. For the Study of Osteoporotic Fractures Research Group. *J Bone Miner Res* **14**, 1637-45 (1999).
18. Langdahl, B.L., Gravholt, C.H., Brixen, K. & Eriksen, E.F. Polymorphisms in the vitamin D receptor gene and bone mass, bone turnover and osteoporotic fractures. *Eur J Clin Invest* **30**, 608-17 (2000).
19. Alvarez-Hernandez, D. et al. Influence of polymorphisms in VDR and COLIA1 genes on the risk of osteoporotic fractures in aged men. *Kidney Int Suppl*, S14-8 (2003).
20. Gennari, L. et al. FokI polymorphism at translation initiation site of the vitamin D receptor gene predicts bone mineral density and vertebral fractures in postmenopausal Italian women. *J Bone Miner Res* **14**, 1379-86 (1999).
21. Beavan, S., Prentice, A., Yan, L., Dibba, B. & Ralston, S. Differences in vitamin D receptor genotype and geopgraphical variation in osteoporosis. *Lancet* **348**, 136-7 (1996).
22. Young, R.P., Lau, E.M.C., Birjandi, Z., Critchley, J.A.J.I.I. & Woo, J. Interethnic differences in hip fracture rate and the vitamin D receptor polymorphism. *Lancet* **348**, 688-9 (1996).
23. Hernandez, J.L. & Weir, B.S. A disequilibrium coefficient approach to Hardy-Weinberg testing. *Biometrics* **45**, 53-70 (1989).
24. Fang, Y. et al. Promoter and 3'-Untranslated-Region Haplotypes in the Vitamin D Receptor Gene Predispose to Osteoporotic Fracture: The Rotterdam Study. *Am J Hum Genet* **77**, 807-23 (2005).
25. Uitterlinden, A.G. et al. A large-scale population-based study of the association of vitamin D receptor gene polymorphisms with bone mineral density. *J Bone Miner Res* **11**, 1241-8 (1996).
26. Lau, J., Ioannidis, J.P. & Schmid, C.H. Quantitative synthesis in systematic reviews. *Ann Intern Med* **127**, 820-6 (1997).
27. Higgins, J.P. & Thompson, S.G. Quantifying heterogeneity in a meta-analysis. *Stat Med* **21**, 1539-58 (2002).
28. Begg, C.B. & Mazumdar, M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* **50**, 1088-101 (1994).
29. Ioannidis, J.P., Trikalinos, T.A., Ntzani, E.E. & Contopoulos-Ioannidis, D.G. Genetic associations in large versus small studies: an empirical assessment. *Lancet* **361**, 567-71 (2003).
30. Ioannidis, J.P., Contopoulos-Ioannidis, D.G. & Lau, J. Recursive cumulative meta-

analysis: a diagnostic for the evolution of total randomized evidence from group and individual patient data. *J Clin Epidemiol* **52**, 281-91 (1999).
31. Ioannidis, J.P., Ntzani, E.E., Trikalinos, T.A. & Contopoulos-Ioannidis, D.G. Replication validity of genetic association studies. *Nat Genet* **29**, 306-9 (2001).
32. Ioannidis, J. & Lau, J. Evolution of treatment effects over time: empirical insight from recursive cumulative metaanalyses. *Proc Natl Acad Sci U S A* **98**, 831-6 (2001).
33. Risch, N.J. Searching for genetic determinants in the new millennium. *Nature* **405**, 847-56 (2000).
34. Lohmueller, K.E., Pearce, C.L., Pike, M., Lander, E.S. & Hirschhorn, J.N. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* **33**, 177-82 (2003).
35. Ioannidis, J.P. & Trikalinos, T.A. Early extreme contradictory estimates may appear in published research: the Proteus phenomenon in molecular genetics research and randomized trials. *J Clin Epidemiol* **58**, 543-9 (2005).
36. Ioannidis, J.P. et al. Differential genetic effects of ESR1 gene polymorphisms on osteoporosis outcomes. *JAMA* **292**, 2105-14 (2004).
37. Ioannidis, J.P. et al. A network of investigator networks in human genome epidemiology. *Am J Epidemiol* **162**, 302-4 (2005).
38. Zajickova, K., Krepelova, A. & Zofkova, I. A single nucleotide polymorphism under the reverse primer binding site may lead to BsmI mis-genotyping in the vitamin D receptor gene. *J Bone Miner Res* **18**, 1754-7 (2003).
39. Ralston, S.H. et al. Loci for regulation of bone mineral density in men and women identified by genome wide linkage scan: the FAMOS study. *Hum Mol Genet* **14**, 943-51 (2005).
40. Devoto, M. et al. Univariate and bivariate variance component linkage analysis of a whole-genome scan for loci contributing to bone mineral density. *Eur J Hum Genet* **13**, 781-8 (2005).

Chapter 6

Vitamin D Receptor Gene Haplotype Is Associated With Body Height and Bone Size

Yue Fang, Joyce B. J. van Meurs, Fernando Rivadeneira, Natasja M. van Schoor, Johannes P. T. van Leeuwen, Paul Lips, Huibert A. P. Pols, André G. Uitterlinden

To be submitted

ABSTRACT

Introduction: Adult stature is a complex genetic trait. We studied whether vitamin D receptor (VDR) gene variants contribute to the genetic variation in height. Recently, we resolved the linkage disequilibrium (LD) structure of the VDR gene, and identified promoter and 3'-UTR haplotype alleles in LD blocks 2, 3 and 5 respectively that influence VDR expression.

Materials and Methods: We studied VDR haplotype alleles and body height in two independent populations (total n = 7,198). In a subsequent meta-analysis (including 11,378 subjects from 19 studies and our current data) we evaluated the effect of a polymorphism in block 5 (Bsm I) which was studied previously.

Results: Haplotypes of LD block 3 and block 5 were associated with body height differences with evidence for additive effects in the Rotterdam study ($p = 0.00002$) and the LASA study ($p = 0.001$). Height differences between the extreme genotypes were 1.4 cm and 2.7 cm, respectively. The relationship was independent of age, gender, presence of vertebral fractures and age-related height loss. In the Rotterdam population we found the combined genotype to be associated with decreased vertebral area ($p = 0.03$), and femoral narrow neck width ($p = 0.002$). In the meta-analysis, subjects with the "BB" genotype were 0.6 cm (95% CI, 0.5 – 1.4 cm) taller than those with "bb" genotype ($p = 0.006$).

Conclusions: Alleles of the VDR gene are associated with differences in body height, as evidenced by our study and by a meta-analysis. The underlying mechanism of the association might involve slightly lower copy numbers of VDR protein in cells important for determining bone size.

INTRODUCTION

Human adult stature is one of the most heritable human traits (heritability up to 90%), as demonstrated in twin studies[1,2] and family studies[3,4], while also ethnicity influences body height. On the other hand, environmental changes (e.g., improved nutrition) have progressively increased height (as a secular trend) during the 20th century. Adult stature has been reported to be inversely associated with a number of common complex genetic diseases, such as cardiovascular diseases, diabetes, pulmonary diseases, cancer and osteoporosis, as well as disease-specific mortalities in different populations[5-11]. Short stature also seems to share some risk factors with these diseases, such as serum low density lipoprotein cholesterol level, blood pressure and socio-economic position[12,13]. The genetic contribution of the relationship between height and those risk factors has been described previously. In a case-control study[14], students with a paternal history of myocardial infarction had lower body height as compared to age- and gender- matched controls. In addition, a family study[12] showed familial clustering and negative correlation between height and low density lipoprotein cholesterol and triglycerides. Therefore, stature appears to be an interesting model complex trait because the understanding of the underlying genetic causes of height determination, may provide insight into mechanisms of other related diseases.

The vitamin D endocrine system has been shown to have pleiotropic effects on a number of endocrine pathways, e.g., related to immune-modulation, regulation of cell proliferation and differentiation, and on skeletal metabolism[15]. Long term vitamin D deficiency results in rickets in children and osteomalacia in adults. The vitamin D receptor (VDR) gene is a central regulator in this endocrine system and therefore, it is an interesting candidate gene for genetic studies of stature. Mutations in the VDR gene, such as in the DNA binding domain[16], the ligand-bind domain[17], or in splice sites[18], cause hereditary vitamin D-resistant rickets (HVDRR). Like other rickets, HVDRR is commonly associated with clinical manifestations of impaired longitudinal growth resulting in short stature and/or regular bone deformities. Adult height was found to be correlated with severeness in 13 HVDRR patients[19], while growth retardation nearly invariably accompanies vitamin D resistant rickets[20]. Also a deletion of the DNA binding domain VDR null mice show growth retardation, in addition to rickets, secondary hyperparathyroidism and alopecia[21]. The phenotype of rickets or osteomalacia in both human and animal models show the same skeletal abnormalities associated with defective mineralization in the growing skeleton, small body size and weaker bones. Recent genome-wide linkage analyses[22,23] showed evidence of linkage between stature and a region on chromosome 12 (12p11.2-q14), where the VDR gene is located. Therefore, sequence variation of the VDR gene might influence body height differences in normal population. Although several association studies have investigated the relationship between VDR polymorphisms and body height in the past, conflicting results were found,

possibly because of small sample sizes, variations in study design, heterogeneous populations, and because mostly anonymous VDR polymorphisms were used so far.

Recently, we and others have resolved the linkage disequilibrium (LD) structure of the VDR gene[24,25]. Haplotypes consisting of several polymorphisms might provide more power over individual polymorphisms in genetic association studies. However, evolutionary structure exists in the haplotype variation on either side of the recombinational hotspot[26]. Recombination events decrease the linkage between polymorphisms, and markedly damage the power of study. So-called cladistic analysis provides a way to estimate evolutionary history of the haplotype variation in a statistical design for an association study[26]. Haplotype alleles within clusters of the cladistic analysis are expected to share more-common ancestry than haplotypes between clusters. Therefore, the combination of cladistic and haplotype analysis can further increase power for an association study.

In functional studies we have identified promoter and 3'-UTR haplotype alleles in block 2 and 5, respectively, that influence VDR expression[24]. In a subsequent association study we observed these risk alleles in block 2, block 3 and block 5 to be associated with increased risk to develop an osteoporostic fracture. The hypothesis of the current study was that the VDR fracture risk alleles might be associated with weak bone possibly due to small bone size, and subsequently with a decreased body height. We investigated the relationship of haplotypes (including the fracture risk haplotypes and others) across the VDR gene with height as well as bone geometry in a population-based population (the Rotterdam Study[27]), and replicated the analysis of height-related haplotypes in another independent populations, the LASA study. Finally, we compared our results with those of previous studies on VDR polymorphisms and height in a meta-analysis of published data. These studies focused on the anonymous Bsm I and /or Taq I polymorphisms in the 3'-end of the VDR gene. We have, however, demonstrated that these SNPs can be used to identify the most common haplotypes in that area[24,25,28]. Thus, we assessed the genetic effect of VDR Bsm I or Taq I polymorphism on body height with a meta-analysis of published literature and including our current studies.

MATERIALS AND METHODS

Subjects

The Rotterdam Study population: The Rotterdam Study is a single center prospective population-based cohort study, and includes 7,983 individuals, with 3,105 men (38.9%) and 4,878 women (61.1%), designed to analyse determinants and prognosis of chronic and disabling diseases in the elderly (> 55 years) Caucasians[27]. The baseline measurements were performed between 1990 and 1993. The latest follow-up period ended 1st of January 2002, and the mean follow-up period was

7.4 years for clinical fracture. For the current study, 6,580 DNA samples (82.4% of the whole cohort population) were available and 6,031 DNAs were successfully genotyped. For the association analysis, we included 5,931 subjects with baseline measurement data of BMD, 5,933 subjects with clinical data of vertebral fracture, 3,114 subjects with X-ray confirmed data of vertebral fracture data, 5,165 subjects with vertebral area data, and 4,230 subjects with hip bone geometry data.

LASA population: The Longitudinal Aging Study Amsterdam (LASA) is an ongoing cohort study of older persons aged 55-85 years. The sampling and data collection procedures have been described in detail elsewhere[29]. In summary, a random sample stratified by age, sex, and expected 5-years mortality rate was drawn from the population registers of 11 municipalities in three regions of the Netherlands. In total, 3,107 persons were enrolled in the baseline examination in 1992-1993. For the current study, persons who participated in the medical interview, and were born in or before 1930 (aged 65 years and older as of January 1, 1996), were selected (n=1509). Of these, 1,352 blood samples were obtained, and 911 DNAs were successfully genotyped for all SNPs.

DNA isolation and genotyping

Genomic DNA was isolated from peripheral venous blood specimens according to standard protocols. We genotyped 14 haplotype tagging single nucleotide polymorphisms (htSNPs) according to our recent haplotype analysis across the entire VDR gene[24] and the Fok I polymorphism in 6,580 Caucasian subjects from the Rotterdam Study with the high throughput TaqMan allelic discrimination assays. We successfully genotyped 6,031 DNA samples for all 15 SNPs. The genotype results were analysed independently by two operators. 332 (5%) random samples from the Rotterdam Study were independently repeated to confirm the genotyping results for all SNPs, the disagreement rate of the genotype results were 0.3-1.2 % for five htSNPs, and the genotype results of other htSNPs were completely consistent. The Assay Mixes (including unlabelled PCR primers, FAM™ and VIC® dye-labelled TaqMan MGB probes) of three Assays-on-Demand℠ and 36 Assays-by-Design™ were designed and supported by ABI. The reaction system contained 2 ng of dried genomic DNA, 2.5 μl of TaqMan® Universal PCR Master Mix, No AmpErase® UNG (2 ×), 0.125 μl (40 ×) or 0.0625 μl (80 ×) of Assay Mix, and adjusted Milli-Q H_2O in a total volume of 2 μl. The reaction conditions consisted of an initial step at 95 °C for 10 minutes, followed by 40 cycles with 92 °C for 15 seconds and 60 °C for 60 seconds in a 384-well plate using the PCR machines, ABI Prism® 7900HT, or ABI GeneAmp® PCR system 9700 (ABI), or PTC-225 Peltier Thermal Cycler (MJ Research, INC.). The genotyping results were analysed with an endpoint reading in the ABI Prism® 7900HT.

Linkage disequilibrium (LD) haplotype analyses and cladogram construction

Our recent LD and haplotype analyses were based on 47 SNPs (identified by re-

sequencing) across the VDR gene and indicated that the whole VDR gene contains seven haplotype blocks[24]. In this study, we combined the nested cladistic method, pair-wise LD measurement and LD block definition together, constructed cladogram and haplotypes within LD blocks. We included nine SNPs from block 3 and ten SNPs from block 5 for LD and haplotype analyses with the PHASE program[30]. PHASE output was used to calculate the pair-wise standardized disequilibrium coefficient (D') with the "haploxt"[31] to estimate the linkage magnitude between two SNPs. We represent haplotype diversity by means of a cladogram, that was generated with a PHASE output linked to the "hapdist" program[32], and depicted the graphic overviews of the cladogram for each LD block with CLadogramer[33].

Clinical examination of association study

Anthropometric measurements of participants were obtained at the research center. Height and weight were measured in standing position wearing indoor clothing without shoes, and all height measurements were attained by a research assistant using a standard wall-mounted stadimeter. Height loss was calculated as the difference of height between the baseline and the ending of follow-up measurements. Body mass index (BMI) was calculated as weight (kg) divided by height squared (m^2). BMD (g/cm^2) was determined by dual energy X-ray absorptiometry (DXA, Lunar DPX-L densitometer, Lunar Radiation Corporation, Madison, WI, USA) at the femoral neck and lumbar spine (vertebral L2 - L4) as described before [34]. Vertebral body area (cm^2) was measured over L2 – L4 by postero-anterior scanning using the same DXA instrument. Hip structure analysis, such as the section modulus (Z, an index of bending strength), the buckling ratio (BR, in index of bone instability), were calculated as described in previous reports[35,36]. The narrow-neck (NN) width (cm) of the hip is across the narrowest point of the femoral neck. Clinical vertebral fracture was diagnosed according to a previous description[37], X-ray confirmed vertebral fractures at baseline were determined by the McCloskey-Kanis method[37,38]. The incidence of vertebral fracture was considered as new cases diagnosed during the follow-up period.

Statistical and association analysis

We applied one-way analysis of variance (ANOVA) to investigate the relationship between genotypes and age, height, bone geometry variables and other continuous outcomes. All genotyping results were tested for Hardy-Weinberg equilibrium. To clarify the possible confounding effect of potential factors on the association between VDR genotypes with height and bone geometry data, we employed adjusted analysis of covariance (ANCOVA) for age, gender and BMI, and we also stratified the analyses by gender, clinical or X-ray confirmed vertebral fracture. To test the association between three fracture risk haplotypes in block 2, 3 and 5 with height, normal significance cut-off p-value (0.05) was used, since we had a clear up-front test hypothesis for those three haplotypes. We adjusted the significance

cut-off p-value using the Bonferroni correction in the association analyses of other seven common haplotypes (MAF > 5%) from LD block 1 and 4, as well as Fok I polymorphism. This resulted in a cut-off p-value at 0.006 (0.05/8), considering that eight independent tests were performed.

Three possible genetic models were allowed to explain differences between genotype groups, i.e., an allele dose effect, a dominant effect or a recessive effect. Allele dose was defined as the number of copies of a certain allele in the genotype, and genotype was treated as a continuous variable. For a certain haplotype or combined genotype allele, the genotype is expressed as a numeric variable with 0 = non-carrier, 1 = heterozygous (one copy of the allele) and 2 = homozygous (two copies of the allele). In case of a consistent trend reflected as an allele dose effect, a linear regression analysis was performed and a "trend" p-value was calculated to quantify the association. In case of a recessive or dominant effect of the test allele, a two by two chi-square test for binary outcomes or an independent sample T-test for continuous outcomes was performed to test for differences between two genotype groups. For dominant alleles we compared the test group of "**Carrier**" (e.g., heterozygous and homozygous of block 3-hap 3) versus "**Non-Carrier**" (e.g., subjects without block 3-hap 3 allele). For recessive effects homozygous subjects for the test allele (e.g. homozygous block 3-hap 3), were compared to the combined group of heterozygous (e.g. heterozygous block 3-hap 3) and non-carriers (e.g. subjects without b_hp 3). The intragenic effect of combined VDR polymorphisms was defined as the combined genotypes of haplotype 3 (block 3-hap 3) from block 3 in the exon 1b promoter region, and haplotype 2 (block 5-hap 2) from block 5 in the 3'-untranslated region (UTR). According to the genetic model of individual haplotypes in our study populations, we combined genotype of the VDR gene with a dominant model for block 3-hap 3 and an allele dose model for block 5-hap 2. "**Zero**" indicates the "Non-Carrier" of block 3-hap 3, and the homozygous block 5-hap 2; "**One**" indicates either the "Carrier" of block 3-hap 3 and homozygous block 5-hap 2, or the "Non-carrier" of block 3-hap 3 and the heterozygous block 5-hap 2; "**Two**" indicates the "Non-carrier" of block 3-hap 3 and the non-carrier of block 5-hap 2, as well as the "Carrier" of block 3-hap 3 and the heterozygous of block 5-hap 2; "**Three**" indicates the "Carrier" of block 3-hap 3 and the non-carrier of block 5-hap 2.

All statistical analyses of the association study were carried out with the SPSS software package (version 11.0).

Meta-analysis

Identification of studies: We first searched publications on PubMed with a combination of keywords: [vitamin D receptor] and [polymorphism or genotype] and [height]. References from retrieved publications were checked for additional studies. In addition, we also reviewed 108 references from three meta-analysis

studies[39-41] of VDR polymorphisms and different bone phenotypes to detect data of height distribution by VDR genotypes. In total, we identified 27 studies[42-68] with genotyping of Bsm I or Taq I or Bsm-Apa-Taq haplotyping and body height from 1995 to May 2005. We also included our current association results in the meta-analysis. Of studies from the same population, we retained for the meta-analysis only the study with the most extensive data. We extracted information on authors, published year, country of origin, ethnicity, study design, mean or range of age, gender, number of subjects and frequency of the B-allele in the control or the complete follow-up population for each study (Table 5). We also examined whether the genotype distribution frequencies of the studies followed Hardy-Weinberg equilibrium (HWE) proportions.

Data Analysis: Three polymorphisms, Bsm I, Apa I and Taq I – restriction fragment length polymorphisms (RFLPs), are commonly analysed in association studies, and Bsm I is the most consistently reported. Bsm I, Taq I and haplotypes in block 5 (in current association study) are highly linked with each other with a pair-wise D' = (0.88 – 1.0) according to our recent study[24], "B" of Bsm I is linked to "t" of Taq I, and block 5-hap 2 in block 5. In the meta-analysis, we focused on analyzing relationship of Bsm I and height, when Bsm I genotype was absent, Taq I and block 5-hap 2 genotype were used to predict Bsm I genotype. We analysed the reliable data using the RevMan 4.2 program from the Cochrane Collaboration (www.cochrane.dk). According to our cross-sectional analysis, the relationship between VDR Bsm I genotype and height followed an allele-dose effect genetic model. We therefore only evaluated the comparison between homozygous genotypes, BB versus bb, in the meta-analysis. BB genotype group was absent or only one individual had BB genotype three studies from Asian population[54,57,65], we extracted height data by Bb versus bb. One study[67] only presented height data by B-carrier versus bb, we also used data by BB+Bb versus bb.

Sensitivity of meta-analysis: Between-study heterogeneity was evaluated by the chi-square distributed Q statistic with a cut-off p-value level at 0.10, and the inconsistency index I^2 (suggesting inconsistency among studies' results with values of 50% or higher, and larger heterogeneity for values of 75% or higher) as estimated from the RevMan program. The standardized mean difference (SMD) and its confidence interval (95% CI) were estimated by both fixed effect and random effects models. The fixed effect model assumes that the genetic factor shows a similar effect on height across all investigated comparisons, and the observed variation among comparisons is caused by chance alone. The random effects model assumes that there may be substantial diversity among different comparisons, and assesses both within-study sampling error and between-study variance. When there is no heterogeneity between comparisons, the outcomes of the two models are similar. Otherwise, the random effects model will give a wider CI than the fixed effect model. We also stratified studies by ethnicity and by age to attempt to clarify potential between-study heterogeneity. We carried out sensitivity analyses by

excluding one study with the largest sample size, or/and excluding studies deviating from HWE proportions. We also performed recursive cumulative meta-analysis to evaluate whether the summary SMD changed in the same direction over time and appraised whether the first published study gave different results from subsequent ones[69-71].

RESULTS

Cladogram of VDR haplotype blocks

Nine SNPs around exon 1b and ten SNPs from exon 4 to the end of 3'-UTR of the VDR gene were found to be highly linked in block 3 (D' is 0.93 – 1.00) and 5 (D' is 0.88 – 1.00), respectively (Fig. 1). Twelve haplotype alleles were predicted by PHASE in block 3, and 18 in block 5 for these Caucasian populations. The cladogram graphs showed that these haplotypes could be categorized as three main clusters (labled as I, II and III in Fig. 1) in the blocks with one or two common haplotype alleles (frequency > 10%). Some SNPs (with arrow below the figure) are specific for certain clusters, e.g. 1b-G-2528A and 1b-G-886A for cluster III in LD block 3; E9-T-48G, U-A311C and U-D796T for cluster I in LD block 5. We combined this feature with the potential functional background (possible transcriptional factor binding site) and a reliable genotyping method together to determine haplotype tagging SNPs (htSNPs[24]). We identified four and three htSNPs which can represent most (93 – 95%) of the estimated five common haplotype alleles (frequency > 3%) in block 3 (4 htSNPs) and 5 (3 htSNPs), respectively. Therefore these htSNPs were used in our association study.

VDR haplotypes and body height

We genotyped seven htSNPs for block 3 and 5 in the Rotterdam and LASA population. The genotype distribution of all htSNPs followed HWE proportion. The baseline characteristics of each study population are shown in Table 1. Data from the Rotterdam and LASA populations are presented by the combined VDR genotype (based on the block 3-hap 3 allele and the block 5-hap 2 allele, as described in the Materials and Methods section). No significant differences in baseline characteristics were found according to VDR genotype. The allele frequencies of block 3-hap 3 and block 5-hap 2 were similar between the Rotterdam and LASA populations.

We first investigated the relationship of body height in the Rotterdam population with the haplotype risk alleles for fracture (Table 2). Block 3-hap 3 carriers were found to have a 0.3 – 0.7% decreased body height, and the block 5-hap 2 homozygotes were found to have a 0.4 – 0.8% increased body height, with evidence for an allele-dose effect. The combined genotype of block 3-hap 3 and block 5-hap 2 had an additive effect on body height. The height difference between two

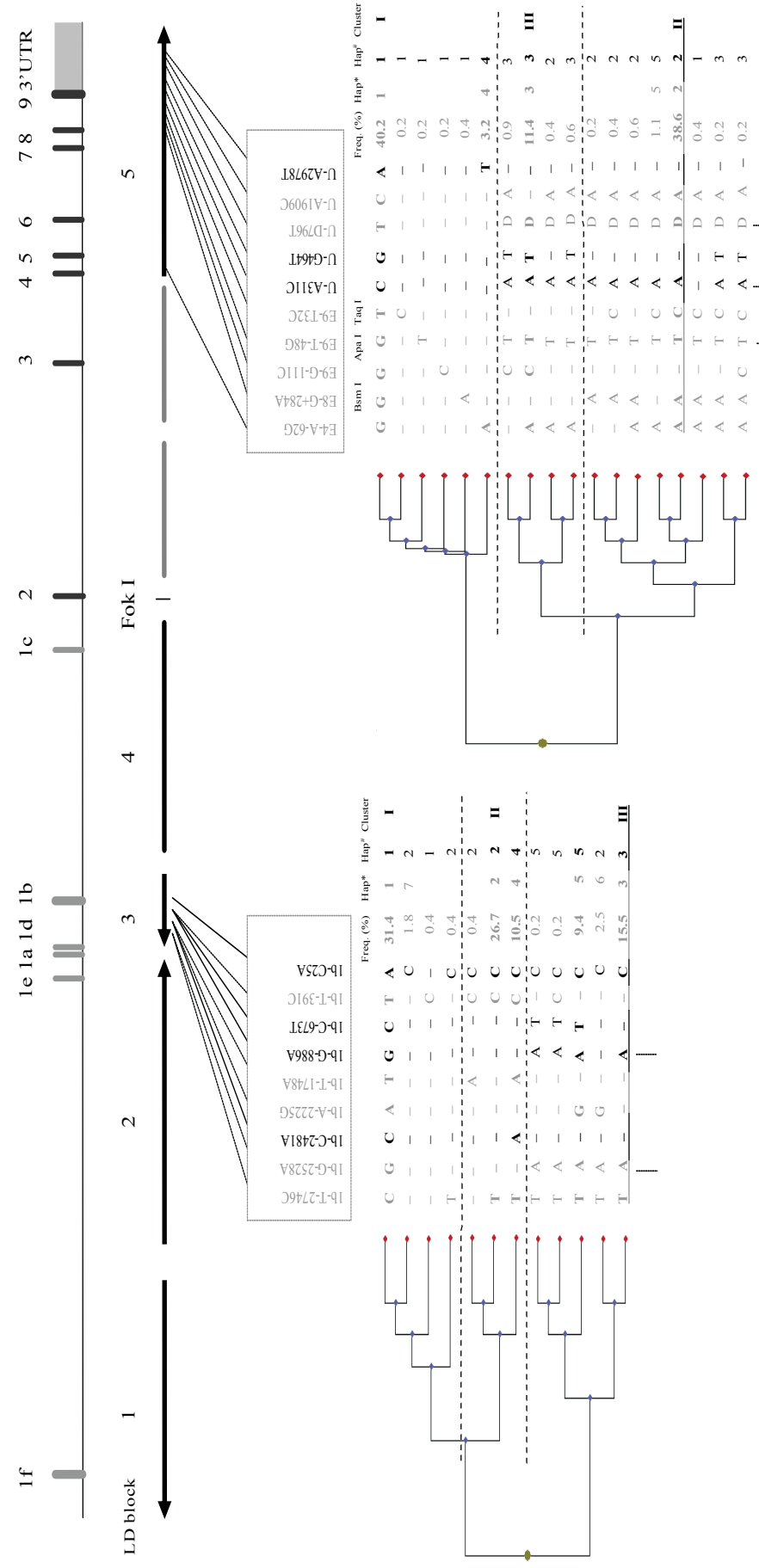


Figure 1. Haplotype constructs and cladograms of LD block 3 and 5 in the VDR gene. The genomic structure of the VDR gene is on the top of the figure. The LD map shows five determined haplotype blocks with black lines based on our data and two expected blocks with grey lines based on other data [24]. Analysed SNPs for block 3 and 5 are presented in vertical boxes, dark SNPs are haplotype tagging SNPs (htSNPs) of the blocks. Cladograms and haplotype constructs are at the lower panel of the figure. Three main haplotype clusters (I, II and III) are separated by dashed lines for each block. Two SNPs (with small bars on the bottom of the figure) are specific for the cluster III in block 2 and three SNPs for cluster I in block 5. The most frequent haplotype alleles are presented on the top of the LD block structures, dashes in the haplotype alleles indicate that the allele of one SNP is the same as the one on top. "Hap*" indicates haplotype code according to all analysed SNPs in the blocks, the most frequent haplotype allele was assigned as hap 1. "Hap#" indicates haplotype code according to htSNPs in the blocks. Haplotypes used in the association analyses are underlined.

Table 1 Characteristics by VDR genotype in different study populations

	Total	Combined VDR genotype#				p-value##
		Zero	**One**	**Two**	**Three**	
Rotterdam population						
Number (%)*	6276	881 (14.0)	2416 (38.5)	2323 (37.0)	656 (10.5)	-
Female (%)*	3732 (59.5)	531 (14.2)	1409 (37.8)	1387 (37.2)	405 (10.9)	-
Age (year)**	69.5 ± 9.1	69.2 ± 8.9	69.5 ± 9.1	69.6 ± 9.2	69.3 ± 9.2	0.70
BMI (kg/m^2)**	26.3 ± 3.7	26.4 ± 3.8	26.2 ± 3.6	26.3 ± 3.8	26.4 ± 3.7	0.49
LASA population						
Number (%)*	911	133 (14.6)	361 (39.6)	334 (36.7)	83 (9.1)	
Female (%)*	467 (51.3)	66 (14.1)	189 (40.5)	167 (35.8)	45 (9.6)	
Age (year)	75.6 ± 6.6	75.2 ± 6.8	75.3 ± 6.6	76.0 ± 6.5	76.2 ± 6.3	0.10
BMI (kg/m^2)	26.9 ± 4.2	27.2 ± 4.2	26.9 ± 4.3	26.9 ± 4.0	26.2 ± 4.4	0.19

#: Combined VDR genotype for risk genotypes in haplotype blocks 3 and 5 (see materials and methods)
##: ANOVA p-value is presented
*: Number (proportion in %)
**: Age was adjusted for gender, BMI was adjusted for age and/or gender

extreme genotype groups was 1.4 cm [0.8%; p = 0.00002, corresponding to 0.15 standard deviations (SD.)]. This effect was replicated in the LASA population (1.6% difference; p = 0.001) using the same genetic model. We observed that the VDR genotype dependent height difference by block 5-hap 2 genotype increased with increasing mean age of the study population. This increase was 0.9 cm (0.09SD.) with mean age 70 years in the Rotterdam Study and 1.4 cm (0.15SD.) with mean age 76 years in LASA.

An inverse correlation between body height and age was observed for both men and women in the Rotterdam Study (r = -0.31, p < 0.001) and the LASA cohort (r = -0.21, p < 0.001). We then analysed the relationship between VDR genotypes and body height by 5-year age strata to investigate the contribution of the VDR genotypes to age-related genotype dependent height differences (secular trends). In all the age categories of the two populations, VDR genotype was associated with increased body height. In addition, we analysed "height loss" longitudinally during a 7.4 years follow-up in the Rotterdam Study. Height loss was not significantly different by VDR genotype. We also examined the relationship between body height and VDR haplotype alleles from other LD blocks, but no other LD blocks showed evidence of association (data not shown).

Vertebral fracture is known to be an important confounder to influence associations regarding body height in elderly subjects. In the Rotterdam population, there were 98 cases (out of 6,031 analysed) with incident clinical vertebral fracture during the follow-up period, and 232 X-ray confirmed vertebral fractures (out of 3,114 screened) at baseline. In the LASA cohort, seven cases (out of 902 analysed) of incident clinical vertebral fracture were reported during the follow-up time. When

Table 2. Height by VDR genotype in different study populations

	Total	Height by VDR haplotype genotype*				Δ Height (%) #	p-value
block 3-hap 3		**Non-Carrier**	**Carrier**		Allele freq. (%)		
Rotterdam Study	166.8 ± 9.4 (6082)	166.9 ± 9.5 (4485)	166.4 ± 9.5 (1597)		14.2	0.5 (0.3%)	0.03
LASA	166.2 ± 9.2 (922)	166.5 ± 9.3 (691)	165.4 ± 8.8 (231)		13.3	1.1 (0.7%)	0.02
block 5-hap 2		**Homozygous**	**Heterozygous**	**Non-Carrier**	Allele freq. (%)		
Rotterdam Study	166.8 ± 9.5 (6154)	167.3 ± 9.5 (1106)	166.8 ± 9.3 (2850)	166.4 ± 9.6 (2198)	41.1	0.9 (0.5%)	0.0002**
LASA	166.3 ± 9.2 (919)	167.2 ± 9.8 (174)	166.2 ± 8.8 (435)	165.8 ± 9.5 (310)	42.6	1.4 (0.8%)	0.03**
Combined block 3 + 5		**Zero**	**One**	**Two**	**Three**		
Rotterdam Study	166.8 ± 9.5 (6031)	167.4 ± 9.4 (859)	166.8 ± 9.4 (2313)	166.6 ± 9.5 (2225)	166.0 ± 9.6 (634)	1.4 (0.8%)	0.00002**
LASA	166.2 ± 9.2 (905)	167.5 ± 10.3 (132)	166.4 ± 8.6 (360)	166.0 ± 9.1 (331)	164.8 ± 9.7 (82)	2.7 (1.6%)	0.001**

*: Data are presented as mean ± SD. (n) Height (cm) is adjusted for age and gender at baseline

#: Height difference (cm) was calculated between extreme genotype groups

**: Trend p-value was calculated by linear regression analysis

Table 3. The influence of vertebral fractures on the association between height and combined VDR genotypes in the Rotterdam Study

	Total	Height by combined VDR genotype*				Δ Height (%)#	p-value**
		Zero	One	Two	Three		
Whole population	166.8 ± 9.5 (6031)	167.4 ± 9.4 (859)	166.8 ± 9.4 (2313)	166.6 ± 9.5 (2225)	166.0 ± 9.6 (634)	1.4 (0.8%)	0.00002
Subjects without clinical vertebral fracture[a.]	166.8 ± 9.5 (5933)	167.5 ± 9.4 (845)	166.9 ± 9.5 (2280)	166.7 ± 9.5 (2182)	166.0 ± 9.6 (626)	1.5 (0.9%)	0.00004
Subjects with X-ray record[b.]	168.3 ± 9.2 (3114)	169.2 ± 9.3 (457)	168.4 ± 9.2 (1197)	168.2 ± 9.2 (1143)	167.7 ± 9.1 (317)	1.5 (0.9%)	0.001
Subjects without vertebral fracture (by X-ray)[b.]	168.4 ± 9.2 (2882)	169.4 ± 9.4 (427)	168.5 ± 9.2 (1112)	168.3 ± 9.2 (1046)	167.6 ± 9.3 (297)	1.8 (1.1%)	0.001

*: Data is presented as mean ± SD. (n). Height (cm) is adjusted for age and gender at baseline
#: Height difference (cm) was calculated between extreme genotype groups
**: Trend p-value was calculated by linear regression analysis
a.: Subset of subjects without incident clinical vertebral fractures during a 7.4-year follow-up period
b. Subset subjects without X-ray confirmed vertebral fractures at baseline

Table 4. Vertebral and hip bone geometry by combined VDR genotype in the Rotterdam Study

	Total	Combined VDR genotypes				Difference (%)	p-value
		Zero	One	Two	Three	Zero vs. Three	
Spine							
Vertebral area (cm^2)#	46.6 ± 6.6 (5165)	47.0 ± 6.8 (745)	46.6 ± 6.5 (1968)	46.7 ± 6.6 (1908)	46.2 ± 6.3 (544)	1.7	0.03
Hip structure analysis							
Femoral narrow neck width (cm)#	3.20 ± 0.34 (4230)	3.22 ± 0.34 (619)	3.19 ± 0.34 (1605)	3.20 ± 0.34 (1574)	3.16 ± 0.35 (432)	1.9	0.002
Inner diameter (cm)	2.93 ± 0.36 (4230)	2.95 ± 0.35 (619)	2.93 ± 0.35 (1605)	2.94 ± 0.36 (1574)	2.89 ± 0.38 (432)	2.0	0.007
section modulus(cm^3)	1.14 ± 0.36 (4230)	1.17 ± 0.38 (619)	1.14 ± 0.36 (1605)	1.14 ± 0.36 (1574)	1.12 ± 0.34 (432)	4.2 ↓	0.02
Femoral shaft width (cm)	3.10 ± 0.25 (837)	3.14 ± 023 (132)	3.11 ± 0.25 (312)	3.10 ± 0.22 (304)	3.06 ± 0.35 (89)	2.5	0.02*
inner diameter (cm)	2.32 ± 3.3 (837)	2.35 ± 0.27 (132)	2.32 ± 0.33 (312)	2.32 ± 0.33 (304)	2.29 ± 0.39 (89)	2.6	0.20

Data are presented as mean ± SD. (n). All analyses are adjusted for age and gender. Vertebral area is adjusted for age, gender and incident clinical or X-ray confirmed vertebral fracture
#: Vertebral area is of L2 – L4
*: is ANOVA p-value. Others are trend p-value calculated by the linear regression analysis

we only analysed subjects without clinical or X-ray confirmed vertebral fracture, the association of VDR genotype with height remained essentially the same (Table 3).

We also analysed bone geometry parameters available in the Rotterdam study (Table 4). The combined VDR genotype was associated with decreased vertebral area of L2 – L4, and decreased femoral size parameters (external and internal diameters) at the narrow-neck and at the shaft, and lower section modulus (index of bending strength) of femoral neck. Associations remained when we adjusted for age, gender, vertebral fracture, height and weight. We did not see the cortical thickness difference by the VDR genotype.

Meta-analysis of VDR genotype and body height

We identified 27 eligible published studies and the two current studies, in total reporting data from 35 different comparisons, which together included 14,157 subjects from Caucasian, Asian and African American populations (Table 5). Study design contained cohort, case-control, normal subject-based and hospital-based cross-section studies. The mean age for each study was 2 – 75 years for both male and female. The sample size of studies varied from 24 – 6,154. Twenty one comparisons presented Bsm I and height data, four comparisons had Taq I genotype data, while six studies had Bsm I and Apa I and Taq I, or Bsm-Apa-Taq haplotype data. Because of the high linkage disequilibrium, we assumed the "t"-allele (or the C nucleotide) of Taq I, the haplotype 2 allele of Bsm-Apa-Taq[56], or the block 5-hap 2 allele (in the current study) to represent the "B"-allele (or the A nucleotide) of Bsm I. The frequency of the B-allele ranged from 3 – 49%.

In our current association analysis, an allele-dose genetic model was observed for the block 5-hap 2 allele. In the meta-analysis, we therefore compared the extreme genotype groups, BB (2,447 subjects) versus bb (5,329 subjects), based on an allele dose genetic model. A significant between-study heterogeneity was observed ($p = 0.0002$, 52%). The random effects model was used to test the overall genetic effect of Bsm I genotype on body height from all comparisons (in the bottom of Fig. 2). Subjects with BB genotype were found to be 0.63 cm (95% CI: 0.17 to 1.08) higher compared to bb genotype carriers ($p = 0.006$).

To investigate the source of the heterogeneity, we first stratified the analyses in the Table 6 by age (< 20 years or > 20 years). Body height changes remarkably during puberty and the age of puberty differs between individuals. Evidence of strong heterogeneity was found in the young (< 20 years) group ($p = 0.002$, $I^2 = 76\%$, in the upper of Fig. 2), and no association was found. In the adult (> 20years) group, a significant but mild heterogeneity was found ($p = 0.003$, $I^2 = 47\%$), and a significant and strong association was observed ($p = 0.002$, height difference: 0.75 [0.34, 1.15]) compared to overall effect. Five comparisons for Asian populations[53,54,57,65], which had a significantly lower B-allele frequency (3-9%, $p < 0.001$), showed modest heterogeneity. Other studies with a similar higher B-allele

Table 5. Studies of VDR Bsm I genotype and body height included in the meta-analysis

Author, year	Country	Ethnicity	Study design	Age (yrs)	Population	Eligible subjects (n)	Genotype	**Bsm I B-allele frequency (%)***	**p-value (HWE)***
Garnero P. 1995	Lyon, France	Caucasian	Cross-section	31-57	Premenopausal	189	Bsm	42	0.50
Need A. 1996	Adelaide, Australia	Caucasian	cohort	20-83	Men	146	Bsm	41	0.32
Tsai K. 1996	Taipei, China Taiwan	Chinese	Cross-section	40-53	Women	113	Bsm, Apa, Taq	3	0.71
				22-88	Men	155		5	0.14
Garnero P. 1996	Lyon, France	Caucasian	cohort	50-70	Postmenopausal	268	Bsm	47	**< 0.001**
Vandevyver C. 1997	Limburg, Belgium	Caucasian	Case-control	75.5 ± 5.0	Women	698	Bsm, Apa, Taq	44	0.08
				66.6 ± 8.4	Women	86		43	0.08
Alahari K. 1997	Durham, US	Caucasian	Cross-section	18-45	Premenopausal	69	Bsm	42	0.56
Keen R. 1997	London, England	Caucasian	Cross-section	63-73	Women	65	Taq	41	0.26
Zmuda J. 1997	Pittsburgh, US	African-American	Cohort	71 ± 5	Women	87	Bsm, Apa, Taq	36	0.85
Sainz J. 1997	Los Angeles, US	Caucasian, Mexican descent	Cross-section	6.7-11.7	Girls	100	Bsm	35	0.44
Suarez F. 1997	Paris, France	Caucasian	Cross-section	22.9 ± 1.2	Girls	119	Bsm	39	0.83
		North Afr.		months	Boys	142	Bsm	36	0.80
Ferrari S. 1998	Geneva, Switzerland		Clinical trial	7-9	Prepubertal girls	101	Bsm	37	0.97
Hansen T. 1998	Odense, Denmark	Caucasian	Cohort	50.8 ± 2.8	Perimenopausal	200	Bsm	48	0.80
Tsuritani I. 1998	Loughborough, UK	Caucasian	Case-control	60-70	Postmenopausal	69	Bsm	38	0.55
Gennari L. 1998	Siena, Florence, Italy	Caucasian	Cross-section, hospital-based	57.7± 8.2	Postmenopausal Women	410	Bsm	44	0.09
Gomez C. 1999	Oviedo, Spain	Caucasian	Cohort	67 ± 8	Men	164	Bsm	40	0.75
				67 ± 8	Women	162	Bsm	39	0.44
Kikuchi R. 1999	Yokohama, Japan	Asian	Cross-section	45-78	Women	127	Bsm	9	0.24
Lorentzon M. 2000	Umea, Sweden	Caucasian	Cohort	19.3 ± 0.7	Boys	88	Bsm	36	0.53
Holmberg-Marttila D. 2000	Kangasal, Finland	Caucasian	Cohort	30.4 ± 4.9	Pregnant women	43	Bsm	49	0.17
Uitterlinden A. 2001	Rotterdam, Netherlands	Caucasian	cohort	67 (55-80)	Women	1004	haplotype	41	0.09
Laaksonen M. 2002	Helsinki, Finland	Caucasian	Case-control	22-45	Premenopausal	24	Bsm	27	0.80
				48-65	Postmenopausal	69	Bsm	36	**0.02**
Taverna M. 2002	Paris, France	Caucasian	Case-control	42.9 ± 13.5	Men ad women	99	Taq	35	0.78
Grundberg E. 2003	Uppsala, Sweden	Caucasian	Cross-section	20-39	Women	304	Bsm	40	0.85
Kim J. 2003	Seoul, Korea	Asian	Hospital-based	48-75	Postmenopausal	417	Bsm	7	0.54
Kurabayashi T. 2004	Niigata, Japan	Asian	Cross-section, hospital-based	49.5 ± 5.4	Women	81	Taq	8	0.47
Garnero P. 2005	Lyon, France	Caucasian	cohort	61 ± 8	Post-menopausal	589	Bsm	40	0.72
Nguyen, 2005	Dubbo, Australia	Caucasian	cohort	70 ± 7	Women	677	Taq	40	0.74
Remes T. 2005	Kuopio, Finland	Caucasian	cohort	50-60	Men	127	Taq	31	0.41
Current studies									
Rotterdam	Rotterdam, NL	Caucasian	cohort	69.5 ± 9.1	Women	3619	haplotype	41	-
					Men	2535		41	-
LASA	Amsterdam, NL	Caucasian	cohort	75.6 ± 6.6	Women	469	haplotype	43	-
					Men	450		42	
Total						14157			

*: HWE p-value was not calculated for two current studies, since Bsm I genotypes were deduced by linked haplotype analyses.

frequency (27-49%, including one study[62] of African Americans with a B-allele frequency of 36%) showed a borderline significant heterogeneity ($p = 0.05$, $I^2 = 34\%$). The point estimates of the height difference by genotype were similar between the random effects model and the fixed effect model, and the genetic effect of Bsm I genotype on height remained. After stratifying for gender, status of menopause (pre- vs. post-menopausal) and study design, we still could see heterogeneity in all strata (data not shown).

To test the sensitivity of the meta-analysis, we replaced the study with the largest sample size, the current complete Rotterdam Study population, with the previous published data[56] using a subset (1,004 subjects) of the same population. This did not impact on the conclusion of association and heterogeneity, but height difference by genotype became smaller [$p = 0.03$; mean difference (95% CI), 0.57 (0.04 – 1.09) cm in the random effects model]. When we excluded two studies which deviated from HWE (Garnero et al. 1996 and Laaksonen et al.), the association remained the same.

Table 6. Mean difference of body height in body height between BB versus bb genotype of Bsm I in the meta-analysis

Studies	Study group (n)	Subjects (n)	Mean (95% CI)		Heterogeneity#		Effect#
			Random effects	Fixed effect	p-value	I^2 (%)	p-value
Overall	35	14157	0.63 (0.17, 1.08)	0.85 (0.64, 1.06)	0.0002	52	0.006
Young (< 20 yrs)	5	550	-0.42 (-2.91, 2.08)	0.14 (-0.84, 1.13)	0.002	76	0.95
Adult (> 20 yrs)	30	13607	0.75 (0.34, 1.15)	0.88 (0.67, 1.10)	0.003	47	0.002
Low B-allele freq.*	5	893	0.99 (-0.08, 2.26)	0.97 (-0.06, 2.23)	0.04	61	0.10
High B-allele freq.*	25	12714	0.70 (0.16, 1.24)	0.71 (0.36, 1.06)	0.05	34	0.01

*: Low B-allele freq. indicates the frequency of B-allele is from 3 to 9%, and high B-allele freq. is from 31 to 49% in Table 5;
#: Heterogeneity and effect were tested by the Random effects model.

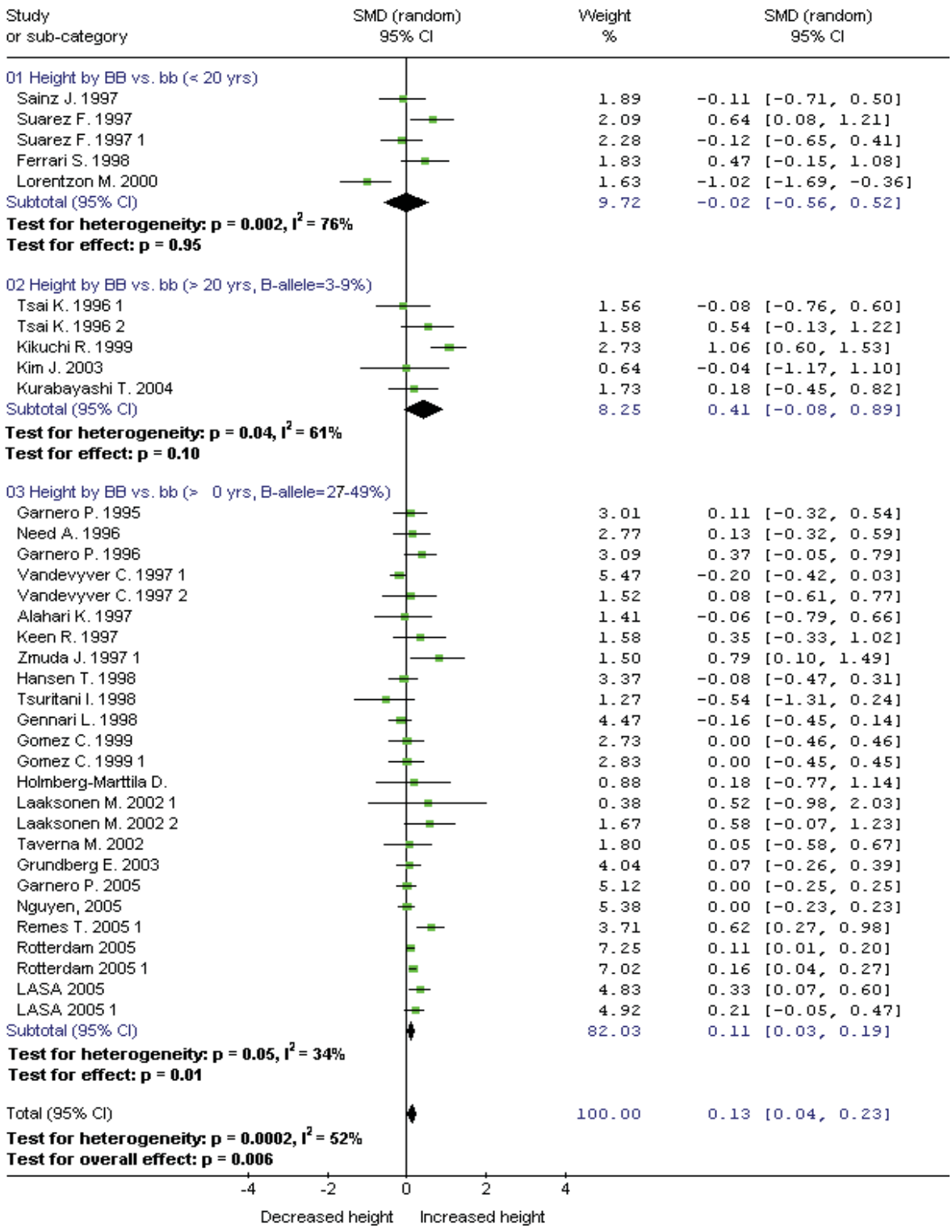


Figure 2. Meta-analysis of the relationship between BB versus bb genotypes of VDR Bsm I RFLP and body height. Point estimates of the standardized mean difference (SMD) of height between the genotypes and 95% confidence intervals (CI) are presented for each study group. Summary estimates of SMD and their 95% CI (diamonds) are given by random effects models in "young" (< 20 years) and "adult" (> 20 years) strata, as well as in the overall database.

DISCUSSION

Although several polymorphisms of the VDR gene have been assessed in the association analyses of body height[47,48,50,72,73], conflicting results were reported. One reason for this might be that non-functional or non-causal polymorphisms were used in the studies. Recently, we systematically analysed the linkage disequilibrium (LD) across the whole VDR gene, and determined five haplotype blocks and 15 htSNPs to represent all common haplotypes in each block[24]. The cladogram, enables inference of evolutionary history patterns and identifies recombination events[74]. "Old" alleles tend strongly to be common alleles, recombination is common but concentrates into a hotspot and recurrent mutations at multiple sites may have occurred[26]. We combined cladistic and haplotype analyses and identified a small number of common haplotype alleles (frequency >10%) in LD blocks, which contain polymorphisms after recombination. In an LD block, haplotype alleles within the clusters of the cladogram are more similar than haplotypes between clusters. According to the common disease/common variation hypothesis[75], a common risk-allele accounts for more contribution to disease in a population. In the association study we therefore, only focus on common haplotype alleles in each block. The homozygous block 3-hap 3 genotype and the block 5-hap 2 carriers were found to be associated with 74% increased and 23% decreased fracture risk, respectively. In the consequent functionality analysis, we found that *in vitro* assays of block 5-hap 2 showed a 15% higher VDR mRNA level compared to the block 5-hap 1 in five different cell lines (including one osteoblast cell line). We here demonstrate an association between these two alleles and adult body height in two independent large populations of elderly subjects.

Interestingly, the summarized genetic effect of Bsm I genotype on height from our meta-analysis was in line with our association results, although a heterogeneity among studies was found. The height difference between BB and bb carriers was 0.63 cm. In the sensitivity analysis, we found that the result of the Rotterdam Study did not substantially influence the significance of the association and heterogeneity in the meta-analysis, although the magnitude of the height difference decreased when we remove our current study from the meta-analysis. Taken together, this indicates that the VDR polymorphisms have a modest and constant genetic effect on the progressing and accumulated height different through lifetime to influence the body height.

To investigate the source of the between-study heterogeneity, we first stratified the analysis by age in young (< 20 years) or adult (> 20 years) groups. In young populations, large heterogeneity was found, because the direction of the effects is different between studies. Childhood and adolescence is the period of most rapid skeletal growth in an individual's lifetime. After this period, the stature tends to be relatively stable, and is more suitable for investigation of the relationship between adult body height and genetic effects of VDR genotypes. However, in the adult

group we still observed significant but mild heterogeneity. Further stratification by B-allele frequency was carried out in the adult group. In the low B-allele frequency group (freq. = 3-9%), a modest heterogeneity resulted from a different magnitude of genetic effects between studies, although a similar direction of the effect was found between studies. This probably is because of insufficient statistical power to detect the genetic effect in the small number of BB homozygote at this low allele frequency. In the high B-allele frequency group (freq. = 27-49%), a borderline significant heterogeneity was found, and other factors of heterogeneity need to be further analysed in this group. We did not stratify the analysis by ethnic groups, because according to our previous study[24], only Asian populations have a lower B-allele frequency (5%) compared to Caucasian and African populations (39% and 36%, respectively). Therefore, age and B-allele frequency variation both influence heterogeneity of this meta-analysis. Although heterogeneity between studies was observed, the genetic effect of Bsm I genotype on body height still exists, especially in the adult population.

As we demonstrated, adult body height (this study) and fracture risk[24] are two complex skeletal phenotypes related to the variations of the VDR gene. We demonstrated fracture risk alleles of the VDR gene to be associated with decreased body height, decreased bone size (e.g., vertebral area, femoral external and internal diameters) and lower bending strength (section modulus) at femoral neck. The effect we observed of VDR genotype on vertebral area could contribute to the VDR genotype effect we observed on stature. Hence, the correlations of VDR genotype with body height and bone size are in line with each other. Therefore, our study indicates the genetic effect of VDR genotype on related bone phenotypes, e.g., fracture risk, short stature and small bone size.

Another fracture-risk haplotype allele found in the Rotterdam Study population[24], block 2-hap1, was not found to be related to height difference in our elder and high calcium intake (1,117 mg/day) populations: the Rotterdam Study and the LASA. However, this haplotype was found to be associated with decreased body height in young and low calcium-intake (< 865 mg/day) French adolescent girls[76], and this haplotype also associated with decreased serum 25(OH)D_3 and decreased serum IGF-1 level in the same population. Since Cdx-2 polymorphism is one of the htSNPs which involve calcium absorption through intestine[77], the interaction between block 2-hap 1 with calcium-intake and/or serum vitamin D level on bone phenotype is interesting to be further investigated in other young and old populations.

In our functionality study of VDR alleles we observed the block 5-hap 2 allele to result in 15% higher mRNA level and stability compared to the block 5-hap 1 in several cell lines including an osteoblast cell line[24]. A possible underlying mechanism might be that for individuals with block 5-hap 2 haplotype, bone tissue (especially osteoblast cells) has a higher sensitivity to vitamin D ligands because of higher expression of VDR, and higher osteoblast activity. As a result, block 5-hap

2 carriers might have relatively higher bone formation to make bigger and stronger bones more resistant to fracture.

In addition, the intervertebral disc height and toughness might influence the body height especially in the elderly populations. The "t" allele of the Taq I polymorphism of the VDR gene was reported to be associated with decreased quantitative signal intensity (more degeneration) of thoracic and lumbar disc[78], increased risk of disc bulges[79], disc degeneration and herniation[80] in different populations. In our study population we don't have data on intervertebral disc to investigate disc changes by VDR genotype. The relationship between VDR polymorphisms and spine phenotypes is certainly interesting to be investigated in further studies.

In conclusion, we observed haplotype alleles in the promoter region and the 3'-end LD block of the VDR gene to be associated with decreased body height in two elderly populations. This association was independent of age, gender and presence of vertebral fracture. A meta-analysis of published data confirmed the relationship between the same allele in the 3'-end of the VDR gene and decreased "adult" height. The underlying mechanism of the association might involve slightly lower copy numbers of VDR protein (in carriers of this risk haplotype) in cells important for determining bone size and strength.

ACKNOWLEDGEMENTS

We thank all participants of the Rotterdam, LASA and EPOS studies. In particular, we acknowledge the general practitioners, pharmacists, and the many field workers at the research center in Ommoord, Rotterdam, the Netherlands for their help in collecting the data for the Rotterdam Study.

This project was funded by the Dutch Research Organisation (NWO 903-46-178, 925-01-010, 014-90-001 and 911-03-012), the European Commission under grant QLK6-CT-2002-02629 ("GENOMOS") and the Dutch Ministry of Health, Welfare and Sports.

REFERENCES

1. Silventoinen, K., Kaprio, J., Lahelma, E., Viken, R.J. & Rose, R.J. Assortative mating by body height and BMI: Finnish twins and their spouses. *Am J Hum Biol* **15**, 620-7 (2003).
2. Schousboe, K. et al. Twin study of genetic and environmental influences on adult body size, shape, and composition. *Int J Obes Relat Metab Disord* **28**, 39-48 (2004).
3. Leiberman, E. et al. Short stature in carriers of recessive mutation causing familial isolated growth hormone deficiency. *Am J Med Genet* **90**, 188-92 (2000).
4. Schwinger, E. et al. Short stature in a mother and daughter with terminal deletion of Xp22.3. *Am J Med Genet* **63**, 239-42 (1996).
5. Sorlie, P., Gordon, T. & Kannel, W.B. Body build and mortality. The Framingham study. *JAMA* **243**, 1828-31 (1980).
6. Hebert, P.R. et al. Height and incidence of cardiovascular disease in male physicians. *Circulation* **88**, 1437-43 (1993).

7. Rich-Edwards, J.W. et al. Height and the risk of cardiovascular disease in women. *Am J Epidemiol* **142**, 909-17 (1995).
8. Han, T.S., Feskens, E.J., Lean, M.E. & Seidell, J.C. Associations of body composition with type 2 diabetes mellitus. *Diabet Med* **15**, 129-35 (1998).
9. Albanes, D., Jones, D.Y., Schatzkin, A., Micozzi, M.S. & Taylor, P.R. Adult stature and risk of cancer. *Cancer Res* **48**, 1658-62 (1988).
10. Jousilahti, P., Tuomilehto, J., Vartiainen, E., Eriksson, J. & Puska, P. Relation of adult height to cause-specific and total mortality: a prospective follow-up study of 31,199 middle-aged men and women in Finland. *Am J Epidemiol* **151**, 1112-20 (2000).
11. Song, Y.M., Smith, G.D. & Sung, J. Adult height and cause-specific mortality: a large prospective study of South Korean men. *Am J Epidemiol* **158**, 479-85 (2003).
12. La Batide-Alanore, A. et al. Family study of the relationship between height and cardiovascular risk factors in the STANISLAS cohort. *Int J Epidemiol* **32**, 607-14 (2003).
13. Langenberg, C., Shipley, M.J., Batty, G.D. & Marmot, M.G. Adult socioeconomic position and the association between height and coronary heart disease mortality: findings from 33 years of follow-up in the Whitehall Study. *Am J Public Health* **95**, 628-32 (2005).
14. Kee, F. et al. Short stature and heart disease: nature or nurture? The EARS Group. *Int J Epidemiol* **26**, 748-56 (1997).
15. Uitterlinden, A.G., Fang, Y., Van Meurs, J.B., Pols, H.A. & Van Leeuwen, J.P. Genetics and biology of vitamin D receptor polymorphisms. *Gene* **338**, 143-56 (2004).
16. Hughes, M.R. et al. Point mutations in the human vitamin D receptor gene associated with hypocalcemic rickets. *Science* **242**, 1702-5 (1988).
17. Kristjansson, K., Rut, A.R., Hewison, M., O'Riordan, J.L. & Hughes, M.R. Two mutations in the hormone binding domain of the vitamin D receptor cause tissue resistance to 1,25 dihydroxyvitamin D3. *J Clin Invest* **92**, 12-6 (1993).
18. Cockerill, F.J. et al. Mutations in the vitamin D receptor gene in three kindreds associated with hereditary vitamin D resistant rickets. *J Clin Endocrinol Metab* **82**, 3156-60 (1997).
19. Herweijer, T.J. & Steendijk, R. The relation between attained adult height and the metaphyseal lesions in hypophosphataemic vitamin-D resistant rickets. *Acta Paediatr Scand* **74**, 196-200 (1985).
20. Chan, J.C. & Bartter, F.C. Hypophosphatemic rickets: effect of 1 alpha, 25-dihydroxyvitamin D3 on growth and mineral metabolism. *Pediatrics* **64**, 488-95 (1979).
21. Erben, R.G. et al. Deletion of deoxyribonucleic acid binding domain of the vitamin D receptor abrogates genomic and nongenomic functions of vitamin D. *Mol Endocrinol* **16**, 1524-37 (2002).
22. Hirschhorn, J.N. et al. Genomewide linkage analysis of stature in multiple populations reveals several regions with evidence of linkage to adult height. *Am J Hum Genet* **69**, 106-16 (2001).
23. Xu, J. et al. Major recessive gene(s) with considerable residual polygenic effect regulating adult height: confirmation of genomewide scan results for chromosomes 6, 9, and 12. *Am J Hum Genet* **71**, 646-50 (2002).
24. Fang, Y. et al. Promoter and 3'-Untranslated-Region Haplotypes in the Vitamin D Receptor Gene Predispose to Osteoporotic Fracture: The Rotterdam Study. *Am J Hum Genet* **77**, 807-23 (2005).
25. Nejentsev, S. et al. Comparative high-resolution analysis of linkage disequilibrium and tag single nucleotide polymorphisms between populations in the vitamin D receptor gene. *Hum Mol Genet* **13**, 1633-9 (2004).
26. Templeton, A.R., Weiss, K.M., Nickerson, D.A., Boerwinkle, E. & Sing, C.F. Cladistic structure within the human Lipoprotein lipase gene and its implications for phenotypic association studies. *Genetics* **156**, 1259-75 (2000).
27. Hofman, A., Grobbee, D.E., de Jong, P.T. & van den Ouweland, F.A. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur J Epidemiol* **7**, 403-22 (1991).
28. Uitterlinden, A.G. et al. A large-scale population-based study of the association of

vitamin D receptor gene polymorphisms with bone mineral density. *J Bone Miner Res* **11**, 1241-8 (1996).
29. Deeg, D.J.H., Beekman, A.T.F., Kriegsman, D.M.W. & Westendorp de Seriere, M. Autonomy and well-being in the Aging population I. Report from the Longitudinal Aging Study Amsterdam 1992-1993. *Amsterdam: VU University Press* **1998**, 9-22 (1994).
30. Phase. http://archimedes.well.ox.ac.uk/pise/PHASE.html.
31. Haploxt. http://archimedes.well.ox.ac.uk/pise/haploxt.html.
32. Hapdist. http://archimedes.well.ox.ac.uk/pise/hapdist.html.
33. Zhang, L., Heng, C.K. & Tan, T.W. Cladogramer: incorporating haplotype frequency into cladogram analysis. *Bioinformatics* **17**, 481-2 (2001).
34. Burger, H. et al. The association between age and bone mineral density in men and women aged 55 years and over: the Rotterdam Study. *Bone Miner* **25**, 1-13 (1994).
35. Beck, T.J., Looker, A.C., Ruff, C.B., Sievanen, H. & Wahner, H.W. Structural trends in the aging femoral neck and proximal shaft: analysis of the Third National Health and Nutrition Examination Survey dual-energy X-ray absorptiometry data. *J Bone Miner Res* **15**, 2297-304 (2000).
36. Rivadeneira, F. et al. The influence of an insulin-like growth factor I gene promoter polymorphism on hip bone geometry and the risk of nonvertebral fracture in the elderly: the Rotterdam Study. *J Bone Miner Res* **19**, 1280-90 (2004).
37. Fang, Y. et al. Cdx-2 polymorphism in the promoter region of the human vitamin D receptor gene determines susceptibility to fracture in the elderly. *J Bone Miner Res* **18**, 1632-41 (2003).
38. McCloskey, E.V. et al. The assessment of vertebral deformity: a method for use in population studies and clinical trials. *Osteoporos Int* **3**, 138-47 (1993).
39. Cooper, G.S. & Umbach, D.M. Are vitamin D receptor polymorphisms associated with bone mineral density? A meta-analysis. *J Bone Miner Res* **11**, 1841-9 (1996).
40. Gong, G. et al. The association of bone mineral density with vitamin D receptor gene polymorphisms. *Osteoporos Int* **9**, 55-64 (1999).
41. Thakkinstian, A., D'Este, C., Eisman, J., Nguyen, T. & Attia, J. Meta-analysis of molecular association studies: vitamin D receptor gene polymorphisms and BMD as a case study. *J Bone Miner Res* **19**, 419-28 (2004).
42. Garnero, P., Borel, O., Sornay-Rendu, E. & Delmas, P.D. Vitamin D receptor gene polymorphisms do not predict bone turnover and bone mass in healthy premenopausal women. *J Bone Miner Res* **10**, 1283-8 (1995).
43. Keen, R.W. et al. Polymorphisms of the vitamin D receptor, infant growth, and adult bone mass. *Calcif Tissue Int* **60**, 233-5 (1997).
44. Sainz, J. et al. Vitamin D-receptor gene polymorphisms and bone density in prepubertal American girls of Mexican descent. *N Engl J Med* **337**, 77-82 (1997).
45. Suarez, F., Zeghoud, F., Rossignol, C., Walrant, O. & Garabedian, M. Association between vitamin D receptor gene polymorphism and sex-dependent growth during the first two years of life. *J Clin Endocrinol Metab* **82**, 2966-70 (1997).
46. Ferrari, S.L., Rizzoli, R., Slosman, D.O. & Bonjour, J.P. Do dietary calcium and age explain the controversy surrounding the relationship between bone mineral density and vitamin D receptor gene polymorphisms? *J Bone Miner Res* **13**, 363-70 (1998).
47. Gennari, L. et al. Vitamin D and estrogen receptor allelic variants in Italian postmenopausal women: evidence of multiple gene contribution to bone mineral density. *J Clin Endocrinol Metab* **83**, 939-44 (1998).
48. Gomez, C. et al. Vitamin D receptor gene polymorphisms, bone mass, bone loss and prevalence of vertebral fracture: differences in postmenopausal women and men. *Osteoporos Int* **10**, 175-82 (1999).
49. Lorentzon, M., Lorentzon, R. & Nordstrom, P. Vitamin D receptor gene polymorphism is associated with birth height, growth to adolescence, and adult stature in healthy caucasian men: a cross-sectional and longitudinal study. *J Clin Endocrinol Metab* **85**, 1666-70 (2000).
50. Laaksonen, M. et al. Vitamin D receptor gene BsmI-polymorphism in Finnish premeno-

pausal and postmenopausal women: its association with bone mineral density, markers of bone turnover, and intestinal calcium absorption, with adjustment for lifestyle factors. *J Bone Miner Metab* **20**, 383-90 (2002).
51. Taverna, M.J. et al. Taq I polymorphism of the vitamin D receptor and risk of severe diabetic retinopathy. *Diabetologia* **45**, 436-42 (2002).
52. Grundberg, E. et al. A poly adenosine repeat in the human vitamin D receptor gene is associated with bone mineral density in young Swedish women. *Calcif Tissue Int* **73**, 455-62 (2003).
53. Kim, J.G. et al. Association between vitamin D receptor gene haplotypes and bone mass in postmenopausal Korean women. *Am J Obstet Gynecol* **189**, 1234-40 (2003).
54. Kurabayashi, T. et al. Association of vitamin D and estrogen receptor gene polymorphism with the effects of longterm hormone replacement therapy on bone mineral density. *J Bone Miner Metab* **22**, 241-7 (2004).
55. Garnero, P., Munoz, F., Borel, O., Sornay-Rendu, E. & Delmas, P.D. Vitamin D Receptor gene polymorphisms are associated with the risk of fractures in postmenopausal women, independently of bone mineral density. The OFELY study. *J Clin Endocrinol Metab* (2005).
56. Uitterlinden, A.G. et al. Interaction between the vitamin D receptor gene and collagen type Ialpha1 gene in susceptibility for fracture. *J Bone Miner Res* **16**, 379-85 (2001).
57. Tsai, K.S. et al. Bone mineral density and bone markers in relation to vitamin D receptor gene polymorphisms in Chinese men and women. *Bone* **19**, 513-8 (1996).
58. Need, A.G. et al. Vitamin D receptor genotypes are related to bone size and bone density in men. *Eur J Clin Invest* **26**, 793-6 (1996).
59. Garnero, P., Borel, O., Sornay-Rendu, E., Arlot, M.E. & Delmas, P.D. Vitamin D receptor gene polymorphisms are not related to bone turnover, rate of bone loss, and bone mass in postmenopausal women: the OFELY Study. *J Bone Miner Res* **11**, 827-34 (1996).
60. Vandevyver, C., Wylin, T., Cassiman, J.J., Raus, J. & Geusens, P. Influence of the vitamin D receptor gene alleles on bone mineral density in postmenopausal and osteoporotic women. *J Bone Miner Res* **12**, 241-7 (1997).
61. Alahari, K.D., Lobaugh, B. & Econs, M.J. Vitamin D receptor alleles do not correlate with bone mineral density in premenopausal Caucasian women from the southeastern United States. *Metabolism* **46**, 224-6 (1997).
62. Zmuda, J.M., Cauley, J.A., Danielson, M.E., Wolf, R.L. & Ferrell, R.E. Vitamin D receptor gene polymorphisms, bone turnover, and rates of bone loss in older African-American women. *J Bone Miner Res* **12**, 1446-52 (1997).
63. Hansen, T.S. et al. Vitamin D receptor alleles do not predict bone mineral density or bone loss in Danish perimenopausal women. *Bone* **22**, 571-5 (1998).
64. Tsuritani, I. et al. Does vitamin D receptor polymorphism influence the response of bone to brisk walking in postmenopausal women? *Horm Res* **50**, 315-9 (1998).
65. Kikuchi, R., Uemura, T., Gorai, I., Ohno, S. & Minaguchi, H. Early and late postmenopausal bone loss is associated with BsmI vitamin D receptor gene polymorphism in Japanese women. *Calcif Tissue Int* **64**, 102-6 (1999).
66. Holmberg-Marttila, D., Sievanen, H., Jarvinen, T.L. & Jarvinen, T.A. Vitamin D and estrogen receptor polymorphisms and bone mineral changes in postpartum women. *Calcif Tissue Int* **66**, 184-9 (2000).
67. Remes, T. et al. Bone mineral density, body height, and vitamin D receptor gene polymorphism in middle-aged men. *Ann Med* **37**, 383-92 (2005).
68. Nguyen, T.V. et al. Contribution of the Collagen I Alpha 1 and Vitamin D Receptor Genes to the Risk of Hip Fracture in Elderly Women. *J Clin Endocrinol Metab* (2005).
69. Ioannidis, J.P., Contopoulos-Ioannidis, D.G. & Lau, J. Recursive cumulative meta-analysis: a diagnostic for the evolution of total randomized evidence from group and individual patient data. *J Clin Epidemiol* **52**, 281-91 (1999).
70. Ioannidis, J.P., Ntzani, E.E., Trikalinos, T.A. & Contopoulos-Ioannidis, D.G. Replication validity of genetic association studies. *Nat Genet* **29**, 306-9 (2001).
71. Ioannidis, J. & Lau, J. Evolution of treatment effects over time: empirical insight from

recursive cumulative metaanalyses. *Proc Natl Acad Sci U S A* **98**, 831-6 (2001).
72. Fischer, P.R. et al. Vitamin D receptor polymorphisms and nutritional rickets in Nigerian children. *J Bone Miner Res* **15**, 2206-10 (2000).
73. Minamitani, K., Takahashi, Y., Minagawa, M., Yasuda, T. & Niimi, H. Difference in height associated with a translation start site polymorphism in the vitamin D receptor gene. *Pediatr Res* **44**, 628-32 (1998).
74. Kroon, L.P., Bakker, F.T., van den Bosch, G.B., Bonants, P.J. & Flier, W.G. Phylogenetic analysis of Phytophthora species based on mitochondrial and nuclear DNA sequences. *Fungal Genet Biol* **41**, 766-82 (2004).
75. Collins, F.S., Guyer, M.S. & Charkravarti, A. Variations on a theme: cataloging human DNA sequence variation. *Science* **278**, 1580-1 (1997).
76. d'Alesio, A. et al. Two new SNPs in the human vitamin D receptor promoter switch transcription factor binding, change promoter activity and are associated with growth in femal adolescents. *J Bone Miner Res* **20**, S184 (2005).
77. Arai, H. et al. The polymorphism in the caudal-related homeodomain protein Cdx-2 binding element in the human vitamin D receptor gene. *J Bone Miner Res* **16**, 1256-64 (2001).
78. Videman, T. et al. Intragenic polymorphisms of the vitamin D receptor gene associated with intervertebral disc degeneration. *Spine* **23**, 2477-85 (1998).
79. Videman, T. et al. The relative roles of intragenic polymorphisms of the vitamin d receptor gene in lumbar spine degeneration and bone density. *Spine* **26**, E7-E12 (2001).
80. Kawaguchi, Y. et al. The association of lumbar disc disease with vitamin-D receptor gene polymorphism. *J Bone Joint Surg Am* **84-A**, 2022-8 (2002).

Chapter 7

Functionality Studies of the Vitamin D Receptor (VDR) Gene Polymorphisms

This chapter represents a part of the manuscript:
Fang Y, van Meurs JB, d'Alesio A, Jhamai M, Zhao H, Rivadeneira F, Hofman A, van Leeuwen JP, Jehan F, Pols HA, Uitterlinden AG
Promoter and 3'-Untranslated-Region Haplotypes in the Vitamin D Receptor Gene Predispose to Osteoporotic Fracture: The Rotterdam Study
Am J Hum Genet (2005) 77:807-823

ABSTRACT

Introduction: Vitamin D receptor (VDR) gene polymorphisms are associated with complex diseases, including osteoporosis, but with inconsistent results. We previously analysed LD haplotype blocks in Caucasians across the VDR gene, and found haplotype alleles of the promoter and 3'untranslated region (UTR) associated with increased fracture risk and decreased body height. To investigate the underlying mechanism, we performed functional experiments for individual polymorphisms and haplotypes in both the VDR promoter and 3'UTR areas.

Material and Methods: To examine VDR promoter SNPs electrophoretic gel mobility shift assay (EMSA) were done using HEK293 or CaCo2 nuclear extracts, and luciferase reporter constructs were assayed in HEK293. We measured VDR mRNA level for the 3'-UTR variations by transfecting constructs with VDR cDNA and the entire 3'UTR of either the block 5-hap1 (risk) allele or the block 5-hap2 (protective) allele in osteoblast cell lines MG63 and SV-HFO, and in HEK293, CaCo2 and COS1. mRNA stability was determined in MG63 at 0-8-24 h after inhibiting transcription.

Results: For the 1e/1a promoter, we found that the "G" weaker binding (by EMSA) alleles at a known Cdx-2 site (1e-G-1739A) and a novel GATA binding site (1a-A-1012G) are both on the promoter risk block 2-hap1 haplotype allele. We transfected reporter constructs containing 2 kb 1a-promoter sequence, and observed 53% lower expression of the risk block 2-hap1 allele vs. block 2-hap2 or 3 ($p=5\times10^{-7}$, n = 51 exp.). For the 3'UTR variants the normalized VDR mRNA level of block 5-hap1 was 15% lower for all cell lines combined with similar patterns in all individual cell lines ($p=2\times10^{-6}$, n = 53). We observed a 30% faster mRNA decay of block 5-hap1 vs. block 5-hap2 in the MG63 ($p=0.02$, n = 9).

Conclusions: We demonstrate *in vitro* that risk haplotypes in both the VDR promoter and the 3'-UTR cause a reduction in mRNA level in potential target cells. We postulate these subtle functional effects to underlie the associations we reported for fracture and height, by resulting in a reduced sensitivity for vitamin D signalling.

INTRODUCTION

Vitamin D receptor (VDR) gene polymorphisms are associated with complex diseases, including osteoporosis, but with inconsistent results. One reason for the conflicting results might be the fact that non-functional polymorphisms were used in most of association studies. To explain the associations of those polymorphisms with complex diseases, linkage disequilibrium (LD) is assumed to truly functional polymorphisms. Functionality studies of some VDR polymorphisms have been performed but also with inconsistent results (see chapter 1.6). Reasons for this inconsistency could involve the use of non-functional polymorphisms, the use of different cell lines (from different species and different tissues), and the use of heterologous constructs.

We previously analysed sequence variations (chapter 2) and LD haplotype blocks (chapter 3) in Caucasians across the VDR gene. According to our sequence analysis, we observed 14 out of 35 polymorphisms to change the putative recognition sequences of transcription factors in the promoter region, while four out of 13 polymorphisms are located in destabilizing elements (DE) in the 3'-untranslated region (UTR) of the VDR gene. In our association studies, we found haplotype alleles of the promoter and 3'-UTR to be associated with increased fracture risk and decreased body height. We therefore carried out functionality experiments, especially for polymorphisms in the promoter region and in the 3'-UTR, which regulate the mRNA level to help understand the underlying mechanism of the associations we observed.

MATERIALS AND METHODS

Electrophoretic gel mobility shift assay (EMSA) and transactivation assay for promoter polymorphisms

Cell nuclear extracts from Caco2 (human colonic adenocarcinoma) and HEK293 (human embryonic kidney) were prepared as described previously[1]. Annealed oligonucleotides (Table 1) were [$\gamma-^{32}$P] ATP end-labeled (Amersham Biosciences) by the T3-polynucleotide kinase (Invitrogen Life Technologies) and purified on a non-denaturing 10% polyacrylamide gel prior to gel shift studies. Nuclear extracts (20 μg) were incubated at 4 °C for 30 min in the presence of 1 μg of double strand poly [dI-dC]-poly [dI-dC] (Amersham Biosciences) and 10 fmoles of purified labeled oligonucleotide in a final buffer containing 10 % glycerol, 5 mM Tris (pH 7.5) and 150 mM KCl. Samples were resolved on a 4% non-denaturing polyacrylamide gel in a low salt buffer (pH7.5) of Tris (6.7 mM), acetate (3.3 mM) and EDTA (1 mM). Supershift experiments were done similarly except for a 1 h incubation at 4 °C in the presence of 1 μl of monoclonal anti Cdx-2 antibody (MU392-UC, Biogenex, San Ramon, California, USA) prior to the addition of the oligonucleotide probe.

Table 1 Oligonucleotides for EMSA

Name	Sequence
Cdx-2	
1e-2090-C sense	5'-AAGTACTGGGATTA**C**AGGCCTGAGCCACT-3'
1e-2090-C antisense	5'-AGTGGCTCAGGCCT**G**TAATCCCAGTACTT-3'
1e-2090-T sense	5'-AAGTACTGGGATTA**T**AGGCCTGAGCCACT-3'
1e-2090-T antisense	5'-AGTGGCTCAGGCCT**A**TAATCCCAGTACTT-3'
1e-1739-G sense	5'-TAAACTAGGTCACA**G**TAAAAACTTATTTC-3'
1e-1739-G antisense	5'-GAAATAAGTTTTTA**C**TGTGACCTAGTTTA-3'
1e-1739-A sense	5'-TAAACTAGGTCACA**A**TAAAAACTTATTTC-3'
1e-1739-A antisense	5'-GAAATAAGTTTTTA**T**TGTGACCTAGTTTA-3'
SIF sense	5'-GAGGGTGCAATAAAACTTTATGAGTAGGT-3'
SIF antisense	5'-ACCTACTCATAAAGTTTTATTGCACCCTC-3'
GATA	
1a-1012-A sense	5'-AGGCGAATAGCAAT**A**TCTTCCCTGGCTAA-3'
1a-1012-A antisense	5'-TTAGCCAGGGAAGA**T**ATTGCTATTCGCCT-3'
1a-1012-G sense	5'-AGGCGAATAGCAAT**G**TCTTCCCTGGCTAA-3'
1a-1012-G antisense	5'-TTAGCCAGGGAAGA**C**ATTGCTATTCGCCT-3'
1a-1521-G sense	5'-GCTAGCTTTCCCAC**G**ATGCTTTGGGCAAG-3'
1a-1521-G antisense	5'-CTTGCCCAAAGCAT**C**GTGGGAAAGCTAGC-3'
1a-1521-C sense	5'-GCTAGCTTTCCCAC**C**ATGCTTTGGGCAAG-3'
1a-1521-C antisense	5'-CTTGCCCAAAGCAT**G**GTGGGAAAGCTAGC-3'

Three reporter constructs containing 2 kb 1a-promoter sequence with 2 SNPs (1a-G-1521C and 1a-A-1012G) were created. Site-directed mutagenesis at the SNP locations in the hVDRp was performed using the Gene Editor TM system (Promega Corp., Madison, WI, USA). The mutated oligonucleotides used were 5'-AGG CGA ATA GCA AT**G** TCT TCC CTG GCT AA-3'for -1012 SNP and 5'-GCT AGC TTT CCC AC**C** ATG CTT TGG GCA AG-3' for -1521 SNP (the mutated base is underlined). Each site-directed mutant was confirmed by sequencing.

Cells were in 6 seeded-well plates ($2x10^5$ cells/well) and grown in 2 ml medium for 24 h. Then, a mixture of 1 μg of plasmid DNA, containing 50 ng of pCMV-β plasmid (Clontech), 950 ng of luciferase reporter vector and 2 μl of Fugene-6® (Roche) was added (maintaining the presence of serum and antibiotics). No change of medium occurred during 72 h culture. Cells were grown for a further 48 h and harvested with 200 μl of 1x reporter lysis buffer (Promega Corp.). Luciferase activities were measured in 10 μl of cell extract with a LG Berthold Lumat LB 9507 and corrected over β-galactosidase activity determined by a standard colorimetric procedure using *o*-nitrophenyl-β-D-galactopyranoside (ONPG) as substrate.

2-D structure of RNA analysis and functionality experiments for 3'-UTR polymorphisms

RNA secondary structure of VDR 3'-UTR was predicted by the RNAfold

server[2], which offers prediction of secondary structure of RNA from a single sequence. The RNA secondary structure of the four most frequent haplotypes of the entire 3.2 kb VDR 3'-UTR were predicted based on free energy minimization[3].

Constructs encompassed VDR exons 2-9, linked to the entire 3'-UTR of the block 5-hap1 or block 5-hap2, were cloned into a pCI-Neo mammalian expression vector (Promega). The accuracy of all cloning was verified by direct sequencing of constructs.

Before transfection MG63 (human osteoblast), SV-HFO (human osteoblast), Caco2, HEK293 and Cos1 (Green Monkey kidney) cells were grown overnight in Dulbecco's modified Eagle's medium (DMEM, Gibco), supplied with 10% fetal calf serum (FCS, Gibco), 50 U/ml penicillin and 50 μg/ml streptomycin (Gibco), and transfected using Fugene transfection reagent (Roche) at a reagent to DNA ratio of 3:1. After 24 hours, cells were harvested and VDR mRNA level was measured. For the mRNA stability measurements cells were put on fresh medium with 7.5 μl/ml actinomycin D (Sigma), cells were harvested at time points 0, 8 and 24 hours after addition of actinomycin D.

Harvested cells were washed once with phosphate buffered saline (PBS), and total RNA was extracted with High pure RNA isolation kit (Roche). One microgram total RNA was reverse transcribed into cDNA using a cDNA synthesis kit according to the protocol of the manufacture (MBI Fermentas, St. Leon-Rot, Germany). Quantitative real-time PCR was performed using an ABI PRISM^{TM} 7700 sequence detector. 25 μl reaction system with qPCR core kit (Eurogentec, Seraing, Belgium) contained 20 ng cDNA, 5 mM $MgCl_2$, 200 μM dNTPs and 0.025 U/μl Hot GoldStar enzyme.Primers and probes were designed with the Primer Express program from ABI (v1.5). cDNA primers: VDR For (5'-CCTCCAGTTCGTGTGAATGATG-3') and VDR Rev (5'-TCATGTCTGAAGAGGTGATACA-3'), NEO For (5'-GCGCCCGGTTCTTTTTG-3') and NEO Rev (5'-CCTCGTCCTGCAGTTCATTCA-3') were used to amplify the VDR and neomycin cDNA respectively. Neomycin was used as an internal control at 0 hour time point. Reaction condition of real time PCR was followed: 50 ºC for 2 min., 95 ºC for 10 min., 40 cycles with 95 ºC for 15 sec. and 60 ºC for 1 min. All transfection experiments were carried out for six to 21 times. The mRNA levels were calculated by the equation of $2^{(20-Ct)}$.

RESULTS

Functionality studies for promoter polymorphisms

Among 14 polymorphisms which are located at potential transcription factor binding sites (TFBS) in the VDR promoter region, we performed EMSA for four common SNPs (MAF > 20%; 1e-C-2090T, 1e-G1739A, 1a-G1521C and 1a-A-1012G) in the promoter region of exon 1e and 1a using Caco2 or HEK293 nuclear extracts (Fig. 1). We identified a putative GATA-binding site for the A-allele of

the 1a-A-1012G SNP in exon 1a promoter region (AGATAT in reverse orientation) and, demonstrated that G-allele has markedly decreased binding to GATA compared with the A-allele.

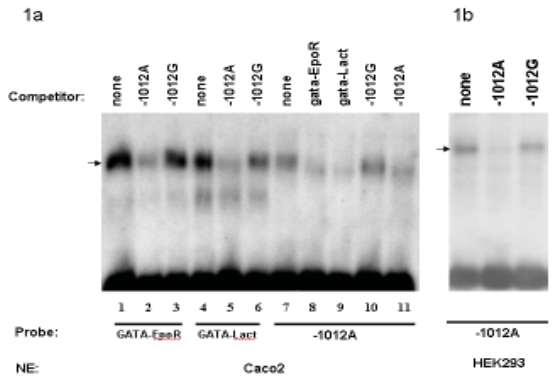


Figure 1 EMSA of 1a-A-1012G for GATA protein. **1a** GATA binding assay using Caco2 cell line nuclear extract. The binding of GATA to the 1a-1012A was analyzed in competition experiments for Caco2 nuclear extract using the well-characterized GATA sites of the EpoR and Lactase gene promoters[14,7]. These experiments showed similar binding characteristics of the complexes bound to the 1a-1012A site and to the other GATA sites (lanes 1 and 4 versus lane 7). In addition, the signal found on the 1a-1012A site was eliminated by a hundred-fold excess of unlabeled GATA sites of the EpoR and Lactase genes (lane 7-10). Inversely, a hundred-fold excess of the 1a-1012A site eliminated the binding of GATA on the EpoR and Lactase GATA sites (lanes 1 versus 2, and lanes 4 versus 5). **1b** GATA binding using HEK293 cell line nuclear extract. Competition experiments for HEK293 nuclear extract using the -1012 GATA site as probe revealed that the 1a-1012G variant were unable to compete the binding of the 1a-1012A variant.

Reporter constructs containing 2 kb 1a-promoter sequence with the 2 SNPs 1a-G-1521C and 1a-A-1012G in block 2 (Fig. 2) showed that in HEK293 cells, the normalised luciferase activity of the block 2-hapl allele was decreased 53% and 50% compared to the block 2-hap 2/3, and hap 4/5 alleles ($p = 5 \times 10^{-7}$, and 8×10^{-7}, respectively). This indicates that the G-allele of the htSNP, 1a-A-1012G, has a 2-fold lower transcription compared to the A-allele. The same results were observed in COS-7 cells (data not shown).

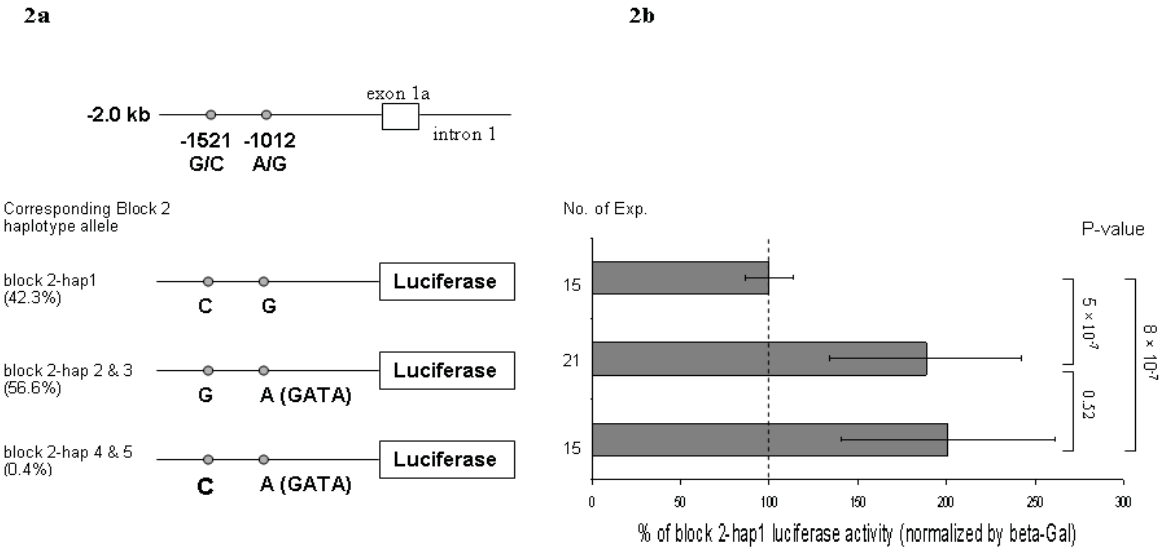


Figure 2 Relative luciferase activity in HEK293 cells of VDR exon 1a promoter activity including two SNPs. **2a** The three constructs containing the 2 kb 1a-promoter sequence with the SNPs: 1a-G-1521C and 1a-A-1012G. **2b** beta-Gal normalized luciferase activity for the three constructs. The block 2-hap1 allele is set at 100% as the reference group, p-value is calculated by independent t-test.

Surrounding sequence analyses showed that the 1e-C-2090T and 1e-G-1739A SNPs locate at potential Cdx-2 binding-sites (Fig. 3a). EMSA confirmed that the G-allele of 1e-G-1739A and the T-allele of the 1e-C-2090T have a relatively decreased

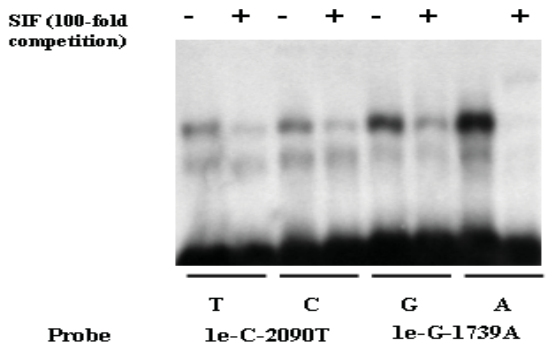
3a

Cdx2-binding sites	Promoter
AAACTTTATGAGT	Sucrose isomaltase a (SIF)
AAGTTTTATTGCA	Sucrose isomaltase b (SIF)
TGCATTTATGGGC	9 K cdxd (rat)
CACTTTTATGGCC	Hoxc 8 site a
TTGTTTTATGGTT	Hoxc 8 site d
GTATTTTACAACC	Lactase-phlorizin hydrolase
TAGGTTTATGCAT	Clusterin (Upstream element)
CTGGTTTATGGAT	Claudin
CCATTTTATATGC	Glucose 6 phosphatase
TCTGTTTATAGCT	Guanyl cyclase C
TCTCTTTATGGCT	human MUC2 mucin gene
AAACTTTATGAGC	RELM/FIZZ
ATGGTTTATGATT	RELM/FIZZ
TTTTTTTATGACA	UDP-Glucuronosyltransferase 1
CTGTTTTATGATG	Ileal bile acid-binding protein

A	4	5	4	1	0	0	0	**15**	0	3	6	2	2
C	4	3	2	3	0	0	0	0	1	0	1	6	5
G	1	1	6	4	0	0	0	0	0	**11**	7	4	1
T	6	6	3	7	**15**	**15**	**15**	0	**14**	1	1	3	7
Consensus:	t/c/a	t/c/a	n	t/n	**T**	**T**	**T**	**A**	**T**	G/a	a/g	n	c/t
1e-2090-**T**	T	G	G	G	A	T	T	A	**T**	A	G	G	C
1e-2090-**C**	T	G	G	G	A	T	T	A	**C**	A	G	G	C
1e-1739-**G**(*)	A	G	T	T	T	T	T	A	**C**	T	G	T	G
1e-1739-**A**(*)	A	G	T	T	T	T	T	A	**T**	T	G	T	G

(*) Reverse strand

3b



3C

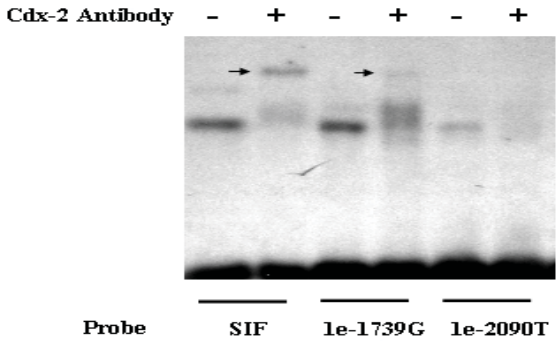


Figure 3 Two SNPs in the 1e promoter region of the human VDR gene are at a Cdx-2 binding site. **3a** 15 well-characterized Cdx-2 sites from mammalian gene promoters were lined up. Base usage was summarized in a table and compared to DNA sequences surrounding SNP 1e-C-2090T and 1e-G-1739A. DNA bases involved in SNPs are underlined. **3b** Double strand oligonucleotides either containing the Sucrose isomaltase (SIF) Cdx2 binding site as control or sequences encompassing SNPs 1e-C-2090T and 1e-G-1739A have been ^{32}P-labeled and purified on a 10% polyacrylamide gel. EMSA experiments were performed using Caco-2 cell nuclear extracts. The competition experiments with an oligonucleotide containing the SIF element. Gel shift experiments was performed in the absence (-) or the presence (+) of a hundred-fold excess of cold SIF probe as competitor, and resulted in relatively more elimination of these specific complexes for the A-allele of 1e-G-1739A and the T-allele of 1e-C-2090T compared with their counterparts. **3c** The antibody experiments with a monoclonal anti-Cdx-2. EMSA was performed with nuclear extract alone (-) or the presence (+) of monoclonal anti Cdx-2 antibody. Supershift complexes were identified by an arrow when present. A clear supershift was observed with SIF, but weakly in the 1e-1739G, and very low intensity of the complex was seen with 1e-2090. Comparison of the surrounding 1e-C-2090T and 1e-G-1739A sequences with the consensus Cdx-2 sequence (lower panel of Suppl. Fig. 3a.) evidenced the presence of a substitution (T/A) in the 1a-2090 sequence, corrupting somehow the Cdx-2 site, maybe explaining the lower intensity observed in EMSA compared to 1e-G-1739A.

binding to Cdx-2, compared with their allelic counterparts (Fig. 3b). The risk allele of the GATA and Cdx-2 promoter SNPs are contained in the block 2-hap1, which is the risk allele for fracture.

Functionality studies of 3'-UTR polymorphisms

Thirteen polymorphisms and four common haplotypes (frequency > 3%) were identified in the VDR 3'-UTR (see chapter 2 and 3). The block 5-hap1 allele and the block 5-hap2 allele in the 3'-UTR differ in only five polymorphisms (Fig. 4a). These result in differences in the RNA secondary structure (Figure 4b).

We found that 24 hours after transfection of reporter constructs containing the complete 3.2 kb 3'-UTR of the block 5-hap1 or hap2, the neomycin normalized VDR mRNA level of the block 5-hap1 transcript was 15% lower than that of the block 5-hap2 transcript when results of all tested cell lines were combined ($p = 2 \times 10^{-6}$, $n = 53$) while similar patterns were observed in individual cell lines (Fig. 5a). We then investigated the stability of VDR mRNAs transcribed from the block 5-hap1 and hap2, and observed that the decay rate of VDR mRNA for the block 5-hap1 was 30% faster than that for the hap2 in MG63 24 hours after inhibiting transcription ($p = 0.02$, $n = 9$; Fig. 5b).

DISCUSSION

According to results from our association studies of VDR SNPs and fracture risk, the genetic effect of the VDR polymorphisms on fracture risk is modest (15 – 48%), which corresponds to the modest difference we observed (15%) in the VDR mRNA level by genotype in our functionality analysis. For comparison, in one previous study[4], the "Ss" (or GT) genotype of COL1A1 Sp1 was found to have an 21% increased COL1A1/COL1A2 protein ratio compared to "SS" (or GG) as measured in osteoblasts. Another functional study of a 3'-UTR SNP [+1073C/T in the oxidised LDL receptor (OLR1) gene) that associates with increased risk for Alzheimer's disease (AD) demonstrated that C-allele carriers had 41% decreased OLR1 mRNA level compared to "TT" homozygotes[5]. These examples indicate that genetic effects of polymorphisms on gene expression and clinical phenotypes are modest.

By analysing functionality of the risk alleles *in vitro*, we demonstrate that the molecular mechanisms underlying these associations are likely to involve a lower expression of VDR mRNA. Some of promoter polymorphisms results in altered transcription factor binding for Cdx-2 and GATA. The previously reported Cdx-2 site at 1e-G-1739A[6,7] and the GATA site at 1a-A-1012G[8] are encompassed in the block 2-hap1, the risk allele of fracture. Our and previous[9] functionality experiments show that both these two weak-binding alleles together result in decreased transcription activity of this VDR promoter. Further research is needed to establish

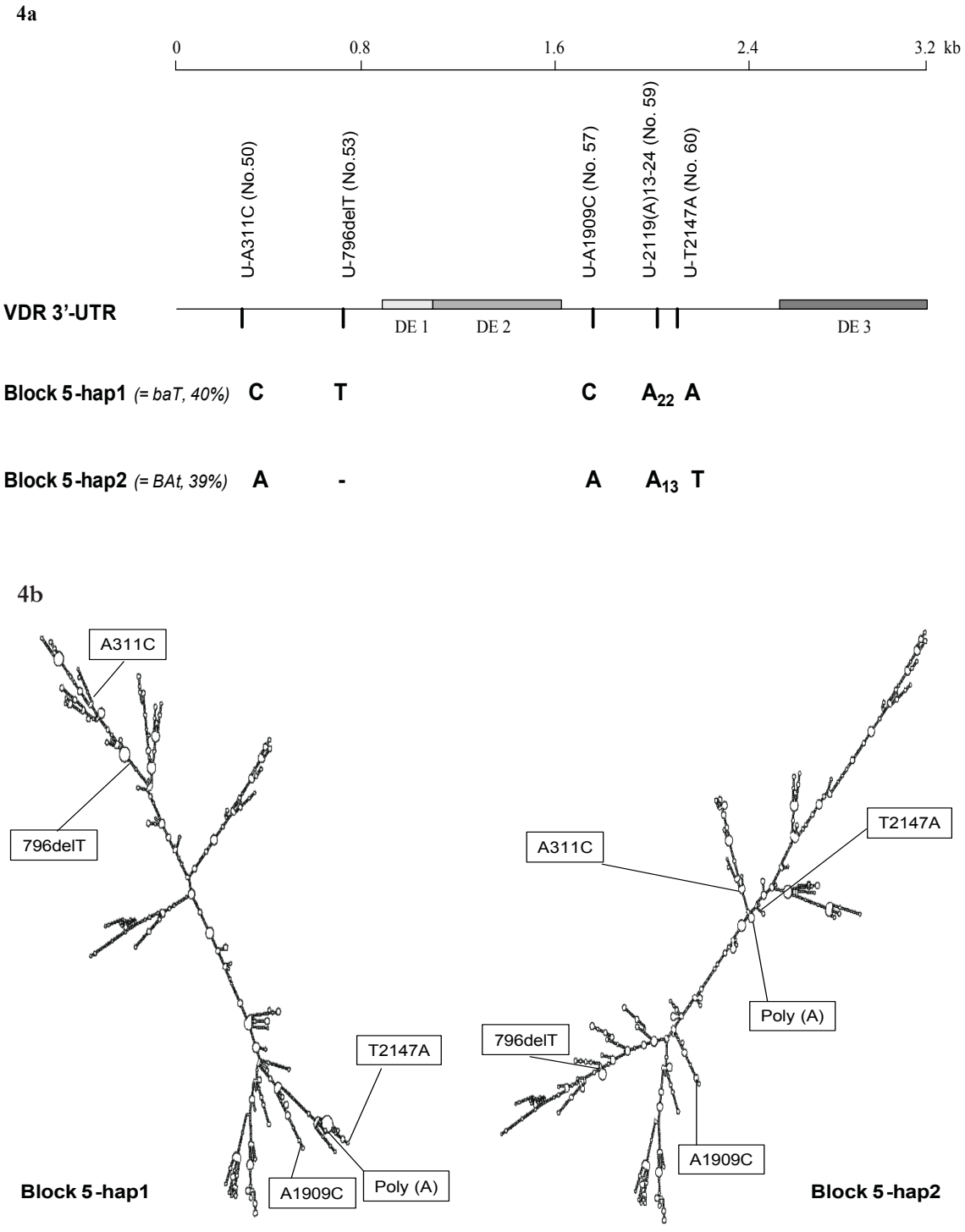


Figure 4 Sequence difference between the block 5-hap1 and hap2 in the VDR 3'-UTR. **4a** The VDR 3'-UTR with sequence variations which distinguish the block 5-hap1 from the block 5-hap2. "No." of SNP refers to the Table 2 in chapter 2. **4b** RNA secondary structure difference between the block 5-hap1 and hap2. Entire 3'-UTR single sequence of block 5-hap1 and hap2 were used to predict the RNA secondary structure. Five polymorphisms were labeled in the figure.

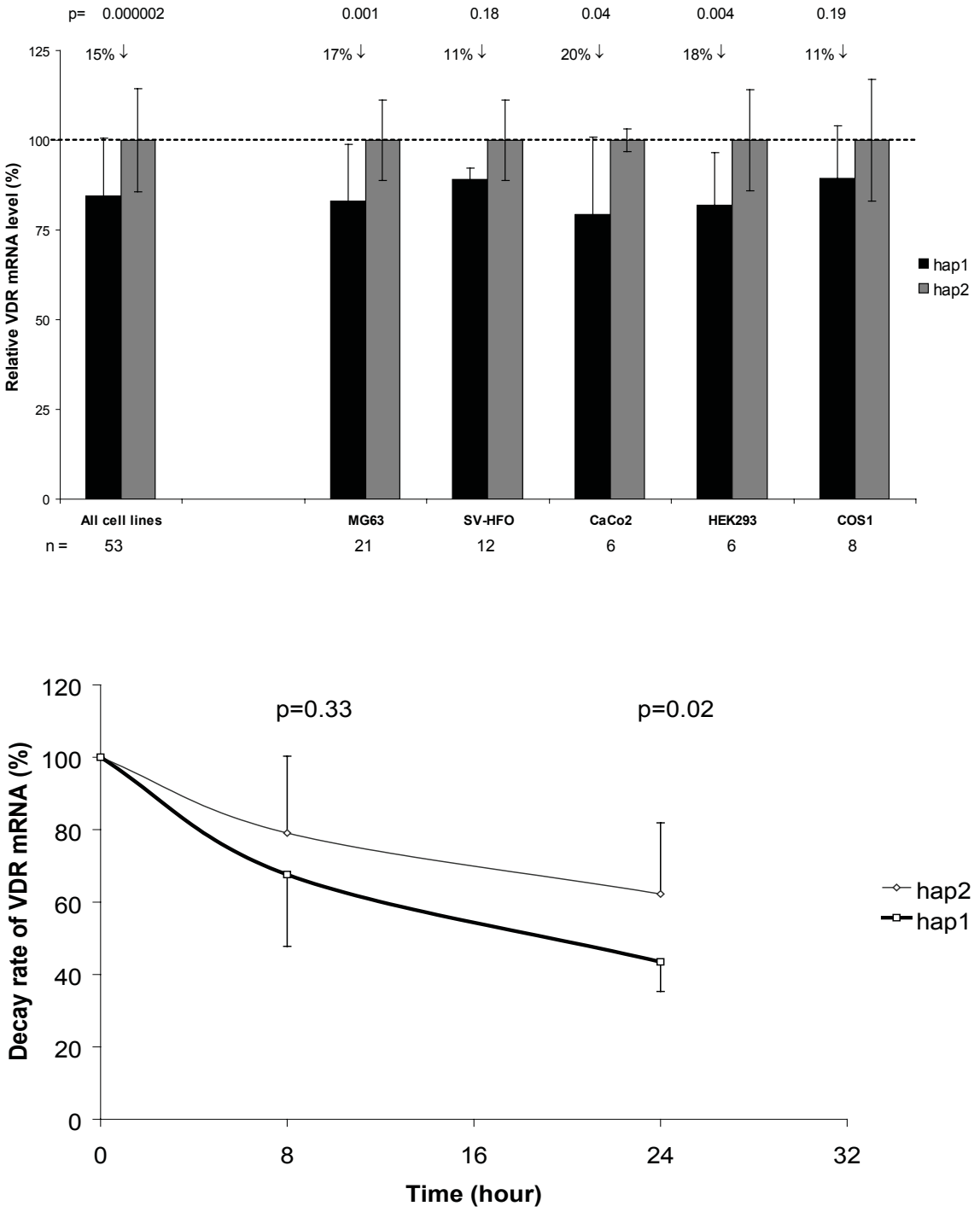


Figure 5 VDR mRNA expression level and stability analysis by 3'-UTR haplotypes in different cell lines. **5a** Neomycin normalized VDR mRNA expression level (mean ± SD) by VDR block 5-hap1 vs. hap2. Hap2 is set at 100% as the reference group, p-value is calculated by independent t-test, n = number of experiments for each cell line. **5b** Decay rate of VDR mRNA by block 5-hap1 vs. hap2 determined in MG63. Time point at 0 hour is defined as 100% mRNA level for both haplotypes, p-value of each time point is calculated by independent t-test.

in more detail, in which cells/tissues this promoter part is influencing VDR expression. While GATA is expressed in many tissues, Cdx-2 is expressed predominantly in the intestine. Thus, the hap1 allele might cause relatively lower VDR expression in target cells for vitamin D, including the intestine.

The 3'-UTR of genes is known to be involved in regulation of gene expression, especially through regulation of mRNA stability. The Bsm I, Apa I and Taq I SNPs are anonymous, and block 5 does not include polymorphisms beyond the 3'-UTR of the VDR gene. Therefore, SNPs in 3'-UTR are the most likely candidates for the truly functional sequence variations explaining the associations we observed. We identified differences in VDR mRNA expression level and stability between the block 5-hap1 and hap2 alleles, which differ at only five positions across the 3.2 kb 3'-UTR. The fracture risk allele, the block 5-hap1, causes 15% lower levels of mRNA expression compared to the hap2 in all tested cell lines. This is in line with the 30% faster decay of or lower of VDR mRNA stability we observed in MG63, an osteoblast cell line. This observation also corresponds to other studies *in vivo* and *in vitro*[10-13]. This is likely to also result in lower numbers of VDR protein being present in target cells for vitamin D, giving such target cells a decreased response to vitamin D.

We demonstrate that polymorphisms within the promoter area and the 3'-UTR area of a gene are having effects that can influence VDR gene function in certain cells and/or subjects. Thus, the 5' promoter and 3'-UTR polymorphisms together can determine how much of a given VDR mRNA will be expressed in a given target cell. The combined risk genotypes in promoter region and 3'-UTR represent a moderate genetic effect of the entire VDR gene on fracture risk. The vitamin D endocrine system has been implicated in several other complex diseases including osteoarthritis, diabetes, and cancer. Whether our findings have relevance for these other diseases needs to be tested in separate association studies using the LD and haplotype information we here provide.

The determination of RNA structure from sequence contains two levels of complexity: secondary structure and tertiary structure (i.e., the three-dimensional shape). It may be possible to infer tertiary structure from interactions of secondary structure elements. It is still unclear how three-dimensional structure influences the stability of mRNA. Although we here showed that only four SNPs and one tandem repeat polymorphisms results in the RNA secondary structure difference of block 5-hap1 and hap2 in the VDR 3'-UTR, the mechanism underlying the secondary structure difference for mRNA stability needs to be further investigated.

In conclusion, we demonstrate *in vitro* that risk haplotypes in both the VDR promoter region and the 3'-UTR cause a reduction in mRNA level in potential target cells. We postulate these subtle functional effects to underlie the associations we reported for fracture and height, by resulting in a reduced sensitivity for vitamin D signalling.

ACKNOWLEDGEMENTS

We thank Drs. John Eisman and Linda Crofts for helpful sharing of unpublished sequence information, Dr. Mark Haussler for kindly providing VDR cDNA, and Marco Eijken for initial help with the RT-PCR of 3'-UTR experiments.

This project was funded by the Dutch Research Organisation (NWO 903-46-178, 925-01-010, 014-90-001 and 911-03-012) and the European Commission under grant QLK6-CT-2002-02629 ("GENOMOS").

REFERENCES

1. Dame, M.C., Pierce, E.A. & DeLuca, H.F. Identification of the porcine intestinal 1,25-dihydroxyvitamin D3 receptor on sodium dodecyl sulfate/polyacrylamide gels by renaturation and immunoblotting. *Proc Natl Acad Sci U S A* **82**, 7825-9 (1985).
2. RNAfold. http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi.
3. Mathews, D.H., Sabina, J., Zuker, M. & Turner, D.H. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol* **288**, 911-40 (1999).
4. Mann, V. et al. A COL1A1 Sp1 binding site polymorphism predisposes to osteoporotic fracture by affecting bone density and quality. *J Clin Invest* **107**, 899-907 (2001).
5. Lambert, J.C. et al. Association of 3'-UTR polymorphisms of the oxidised LDL receptor 1 (OLR1) gene with Alzheimer's disease. *J Med Genet* **40**, 424-30 (2003).
6. Yamamoto, H. et al. The caudal-related homeodomain protein Cdx-2 regulates vitamin D receptor gene expression in the small intestine. *J Bone Miner Res* **14**, 240-7 (1999).
7. Fang, R., Olds, L.C., Santiago, N.A. & Sibley, E. GATA family transcription factors activate lactase gene promoter in intestinal Caco-2 cells. *Am J Physiol Gastrointest Liver Physiol* **280**, G58-67 (2001).
8. Halsall, J.A., Osborne, J.E., Pringle, J.H. & Hutchinson, P.E. Vitamin D receptor gene polymorphisms, particularly the novel A-1012G promoter polymorphism, are associated with vitamin D3 responsiveness and non-familial susceptibility in psoriasis. *Pharmacogenet Genomics* **15**, 349-55 (2005).
9. Arai, H. et al. The polymorphism in the caudal-related homeodomain protein Cdx-2 binding element in the human vitamin D receptor gene. *J Bone Miner Res* **16**, 1256-64 (2001).
10. Morrison, N.A., Yeoman, R., Kelly, P.J. & Eisman, J.A. Contribution of trans-acting factor alleles to normal physiological variability: vitamin D receptor gene polymorphism and circulating osteocalcin. *Proc Natl Acad Sci U S A* **89**, 6665-9 (1992).
11. Morrison, N.A. et al. Prediction of bone density from vitamin D receptor alleles. *Nature* **367**, 284-7 (1994).
12. Carling, T., Ridefelt, P., Hellman, P., Rastad, J. & Akerstrom, G. Vitamin D receptor polymorphisms correlate to parathyroid cell function in primary hyperparathyroidism. *J Clin Endocrinol Metab* **82**, 1772-5 (1997).
13. Ogunkolade, B.W. et al. Vitamin D receptor (VDR) mRNA and VDR protein levels in relation to vitamin D status, insulin secretory capacity, and VDR genotype in Bangladeshi Asians. *Diabetes* **51**, 2294-300 (2002).
14. Zon, L.I., Youssoufian, H., Mather, C., Lodish, H.F. & Orkin, S.H. Activation of the erythropoietin receptor promoter by transcription factor GATA-1. *Proc Natl Acad Sci U S A* **88**, 10638-41 (1991).

Chapter 8

Association of Vitamin D Binding Protein (DBP) Gene Haplotype with Serum Vitamin D level and Osteoporosis

Yue Fang, Pascal Arp, Joyce B. J. van Meurs, Johannes P. T. van Leeuwen, Albert Hofman, Huibert A. P. Pols, André G. Uitterlinden

To be submitted

ABSTRACT

Introduction: Vitamin D binding protein (DBP) is a polymorphic and multifunctional serum protein. DBP action is influenced by environmental factors since it binds to vitamin D and transports it to target tissues to maintain calcium homeostasis. However, DBP can also be converted to DBP-macrophage activating factor (DBP-MAF), which mediates bone resorption by activating osteoclasts. Therefore, the DBP gene is a candidate gene of osteoporosis.

Materials and Methods: We summarized the haplotype structure across the DBP gene according to Perlegen, HapMap and SNPbrowser datasets, and proposed consensus haplotype structures for different ethnic populations. We genotyped two non-synonymous SNPs (rs7041 and rs4588) in a population-based Caucasian population of 6,181 elderly. Haplotypes of these SNPs correspond to protein variations Gc1s (hap1), Gc2 (hap2) and Gc1f (hap3). We analised DBP genetic variation in relation to BMD and fracture risk and studied interaction with dietary calcium intake and vitamin D receptor (VDR) genotype.

Results and Conclusions: In a subgroup of 1,312 subjects, DBP genotype was found to be associated with differences in serum 25-$(OH)D_3$ for hap1 ($P = 3\times10^{-4}$), and for hap2 ($p = 3\times10^{-6}$). Similar associations were observed for 1,25-$(OH)_2D_3$. The DBP genotype was weakly and borderline significantly associated with fracture risk ($p = 0.13$) and the effect on fracture risk was likely to be independent of vitamin D level. The hazard ratio for clinical fracture risk of DBP hap1-homozygote versus non-carrier was 1.47 (95% confidence interval: 1.06–2.05) in a subgroup with dietary calcium intake < 1.09 g/day. In the total study population, we demonstrated interaction between DBP and VDR haplotypes on fracture risk. In the DBP hap1-carrier group, VDR risk genotype carriers had 33% increased fracture risk compared to non-carriers ($p = 0.005$). All associations were independent of age and gender.

INTRODUCTION

Osteoporosis is defined as a reduction in bone mass associated with disruption of bone microarchitecture, resulting in increased bone fragility and increased fracture risk[1]. It is a complex genetic disorder with interaction between environmental and genetic factors. The genes involved remain ill defined but candidate gene association studies seem a powerful tool to identify causative genes[2]. Indeed, several genetic association studies have demonstrated a relationship between polymorphisms of candidate genes with decreased bone mineral density (BMD), and increased fracture risk[3-8].

Vitamin D binding protein (DBP), was initially named as group-specific component (Gc), and is a polymorphic serum protein with different functions. Two functions of DBP involve skeletal metabolism.

First, DBP binds to vitamin D metabolites such as 25-hydroxyvitamin D_3 [$25(OH)D_3$], the major circulating metabolite, and 1,25-dihydroxyvitamin D_3 [$1,25(OH)_2D_3$], the most active form of vitamin D. Strong and positive correlations were found between those vitamin D metabolites and DBP levels in serum[9]. Vitamin D plays an essential role in regulating serum calcium and phosphate homeostasis as well as bone metabolism. A deficiency of vitamin D results in rickets in children and osteomalacia in adults and increased risk of osteoporosis[10]. DBP transports vitamin D metabolites to liver, kidney, bone and other target tissues. DBP also plays a crucial role in storing and prolonging the half-life of the circulating vitamin D metabolites[11].

Second, serum DBP can be converted to a DBP-macrophage activating factor (DBP-MAF) by deglycosylation of DBP. Though the mechanism is still unclear, some studies suggested that DBP-MAF upregulates oxidative metabolism in osteoclasts[12] and potently inhibits the sensitivity of osteoclast to extracellular calcium[13] which is a powerful anti-resorptive signal[14]. DBP-MAF plays a role in osteoclast differenciation from monocytes[15], and is a cytokine to mediate bone resorption by directly activating osteoclasts[16]. DBP-MAF treated osteopetrosis rats have increased number and activity of osteoclast, and decreased bone mass[12].

Hence, the contribution of the DBP to bone metabolism can not only be through assisting the vitamin D endocrine system, but also through directly influencing bone resorption. Therefore, the DBP gene, on chromosome 4q13.3, is a potential osteoporosis candidate gene. In addition, DBP has been observed to have several other functions. It is able to bind and clear gelsolin-actin (G-actin) from the circulation, and prevent the formation of filamentous actin (F-actin) and actin-mediated damage in the microcirculation after cellular trauma[17,18]. Moreover, DBP-MAF consequently activates macrophages[19] and enhances a number of specialized immune functions such as chemotaxis, phagocytosis and destruction of tumor cells[20].

Three most commonly studied protein variations (Gc1f, Gc1s and Gc2) can be identified by two DNA polymorphisms in exon 11: a G/A substitution in codon 416 (rs7041) leading to a Glu/Asp amino acid change and a C/A substitution in codon 420 (rs4588) leading to Thr/Lys amino acid change[21]. Haplotypes of the nucleotide changes result in the protein isoforms: Gc1f, Gc1s and Gc2 (see Fig. 2). Affinity of the DBP isoforms for vitamin D metabolites was reported to be different: Gc1f > Gc1s > Gc2[22]. A positive correlation was found between serum $1,25(OH)_2D_3$ and DBP level[23]. Recently, two studies demonstrated a relationship of a DBP haplotype (or protein isoform: Gc1) with decreased BMD and increased fracture risk in Japanese and Caucasian populations[24,25]. However, it remains necessary to replicate these associations in a large-scale population-based study.

The action of these genes is under the influence of several environmental factors, e.g., dietary vitamin D intake and calcium intake. Furthermore, osteoclasts can have a different response to some environmental factors, such as dietary calcium intake. The number of osteoclasts significantly increases when rats were at a calcium deficiency condition, and the number decreased to a similar level of control groups after calcium or phosphorus replenished[26,27]. No studies have investigated the interaction between DBP and dietary calcium intake on osteoporosis.

Another osteoporosis candidate gene, the vitamin D receptor (VDR) gene mediates the action of the vitamin D endocrine system in calcium homeostasis and bone metabolism. In a previous study of >6,000 elderly Dutch Caucasian subjects, we reported that the VDR block 5-haplotype 1 allele associated with increased fracture risk. In a functionality study, this haplotype allele was found to associate with 15% decreased VDR mRNA level comparing to the haplotype 2 allele, especially in osteoblast cells[28]. DBP stimulates the activity of osteoclasts, and influences the bone resorption[12]. Thus, VDR and DBP might influence the bone remodelling in two different pathways and further impact risk of osteoporosis. So far, no studies have addressed the interaction between VDR and DBP polymorphisms on osteoporosis.

In addition, it is necessary to systematically analyse the complete gene haplotype structure of the DBP gene in major ethnic populations to guide further genetic studies. In this study, we first determined the haplotype structure of the whole DBP gene according to existing databases. Subsequently, we investigated the relationship between the most common polymorphisms of the DBP gene with serum vitamin D level and bone phenotypes in a population-based large study population, and finally we tested the interaction between DBP haplotypes with VDR haplotypes and dietary calcium intake on fracture risk.

MATERIALS AND METHODS

Subjects:

The Rotterdam Study population: The Rotterdam Study is a single center prospective population-based cohort study, and includes 7,983 (> 55 years) individuals Caucasian, with 3,105 men (38.9%) and 4,878 women (61.1%), to analyze determinants and prognosis of chronic and disabling diseases in the elderly[29]. The baseline measurements were performed between 1990 and 1993. The latest follow-up period ended 1st of January 2002. For the current study, 6,580 DNA samples (82.4% of the whole cohort population) were available and 6,181 DNA was successfully genotyped. A subgroup with 4,747 subjects we had data on dietary calcium intake. 5,931 subjects had baseline measurement data of bone mineral density (BMD, g/cm^2). On baseline, for 1,317 subjects we measured serum vitamin D level data.

Panel of ethnic groups: We genotyped a panel of DNA from 107 Chinese Han and 68 African individuals [the Coriell Institute, Camden, NJ. USA: 98 Chinese Han (HD100), 9 Chinese Han (HD02), 59 African Americans (HD04 and HD50) and 9 Africans from south of the Sahara (HD12)].

DNA isolation and genotyping

Genomic DNA was isolated from peripheral venous blood specimens according to standard protocols. In the DBP gene we genotyped two single nucleotide polymorphisms (SNPs) in exon 11, rs7041 and rs4588, which change the amino acid sequence at codon 416 and 420, respectively. The genotype results were detected using the Taqman procedure. The Assay Mixes (including unlabelled PCR primers, FAM™ and VIC® dye-labelled TaqMan MGB probes) of Assays-by-Design™ were designed and supported by ABI. The reaction system contained 2 ng of dried genomic DNA, 2.5 μl of TaqMan® Universal PCR Master Mix, No AmpErase® UNG (2 ×), 0.125 μl (40 ×) or 0.0625 μl (80 ×) of Assay Mix, and adjusted Milli-Q H_2O in a total volume of 5 μl. The reaction conditions consisted of an initial step at 95 °C for 10 minutes, followed by 40 cycles with 92 °C for 15 seconds and 60 °C for 60 seconds in a 384-well plate using PCR machines, ABI Prism® 7900HT, or ABI GeneAmp® PCR system 9700 (ABI), or PTC-225 Peltier Thermal Cycler (MJ Research, INC.). The genotyping results were analysed with an endpoint reading in the ABI Prism® 7900HT. The genotype results were determined independently by two operators and 5% random samples were re-genotyped to check for genotyping errors. No inconsistencies were observed.

Linkage disequilibrium (LD) and haplotype analyses

We first extracted from the Perlegen database[30] genotype data of 40 SNPs across the DBP gene and one SNP out of 3'-untranslated region (UTR) of the DBP gene for different ethnic groups: 24 European Americans, 24 Han Chinese and 23 African Americans. We selected race-specific SNPs with a minor allele fre-

quency (MAF) > 10% in either of the ethnic study populations, and identified 16 SNPs for European American, 22 for Chinese Han and 25 for African American. These race-specific SNPs were used to determine haplotype structure of the DBP gene with the PHASE program [31]. We identified haplotype blocks (defined as, the standardized pair-wise disequilibrium coefficient between SNPs, D' > 0.8) and calculated haplotype frequency in each block by HaploBlockFinder program[32]. Other haplotype structure data of the DBP were included in the further analysis: A. the original Perlegen haplotype data, including 32 SNPs for European American, 31 for Chinese Han and 33 for African American; B. HapMap data (Haploview version3.2[33]), including 30 trios CEPH (Utah residents with ancestry from northern and western Europe), 45 singletons of Han Chinese in Beijing, 30 trios YRI (Yoruba in Ibadan, Nigeria); C. SNPbrowser data (version 3.0[34]), including 60 parent samples from 30 CEPH (Caucasian from HapMap data), 45 Han Chinese in Beijing and 45 African American from Coriell DNA samples. The definition of haplotype block was as follows: MAF of analysed SNPs > 10% and D' > 0.8, and haplotype block from original Perlegen was defined by the HAP program[35]. We extracted consensus race-specific LD maps of the DBP gene by comparing our haplotype structures (re-analysed Perlergen data) with other sources: Perlegen original data, HapMap, SNPbrowser and Ezura et al.[24].

The haplotypes of rs7041 and rs4588 SNPs in the Rotterdam study population were generated by the PHASE program.

Clinical data for the association study

Information on medical history, dietary habits, age at menopause, and smoking was obtained with a computerized questionnaire during a home interview at baseline. Intakes of calcium and total energy were calculated by food frequency questionnaire (based on all food and drinks consumed in 1 month) with the use of Dutch food composition tables. We calculated the total dietary energy intake adjusted dietary calcium intake for all individuals. Anthropometric measurements of participants were obtained at the research center. Height (cm) and weight (kg) were measured in standing position in indoor clothing without shoes, and all height measurements were attained by a research assistant using a standard wall-mounted station-meter. Body mass index (BMI) was calculated as weight (kg) divided by the height squared (m^2). BMD (g/cm^2) was determined by dual energy X-ray absorptiometry (DXA, Lunar DPX-L densitometer, Lunar Radiation Corporation, Madison, WI, USA) at the femoral neck and lumbar spine (vertebral L2 - L4) as described before[36]. The presence of a clinical vertebral fracture was diagnosed as described according to previously[6,37,38]. The incidence of fracture was considered as new cases diagnosed during the follow-up period (7.4 years ± 3.3 years).

50μl serum obtained at baseline was stored at -20 °C was used for the quantitative determination of 25-hydroxyvitamin D_3 [$25(OH)D_3$] and 1, 25-dihydroxyvitamin D_3 [$1, 25(OH)_2D_3$] level by using radioimmunoassay kits from IDS (Immunodianostic

systems limited, UK).

Statistical and association analysis

All genotyping results were tested for Hardy-Weinberg equilibrium. We applied one-way analysis of variance (ANOVA) to investigate the relationship between genotypes and age, height, weight, BMD, BMD loss, bone geometry variables, serum vitamin D level, dietary calcium intake, and other continuous outcomes. We employed the Pearson χ^2 test, and calculated the Hazard Ratio and 95% confidence interval (HR, 95% CI) to investigate the relationship between incidence of clinical fracture and haplotypes of the DBP gene by Cox regression model. Both analysis of covariance (ANCOVA) and Cox models were adjusted for potential confounders, such as age, gender, height, weight, BMD and dietary calcium intake. We also stratified the analyses by gender and high/low, tertiles and quartiles of dietary calcium intake in the study population. The interaction term of high/low calcium intake and DBP haplotype for facture risk was estimated in the Cox model with the adjustment for age and gender. All statistical analyses of the association study were carried out with the SPSS software package (version 11.0).

RESULTS

Haplotype structure of the DBP gene

The DBP gene structure is shown in Figure 1. It is at least 42 kb in length, contains 12 exons with a 4.3 kb 3'-untranslated region (UTR) and has at least 162 SNPs according to the dbSNP database (build 124[39]). Three genotype data sources, Perlegen[30], HapMap[33], SNPbrowser[34], and one publication[24] provided the haplotype structure of the DBP gene in different ethnic populations. The Perlegen database includes genotype data of the highest number of SNPs as compared to other databases, and it also uses the HAP program[35] to infer haplotypes. In order to summarise the haplotype structure of the DBP gene in Caucasian, Han Chinese and African American populations, we extracted genotype data from Perlegen, determined haplotype and LD block structure with the PHASE[31] and HaploBlockFinder[32] programs. We used the same criteria to define haplotype blocks for the three datasets, and aligned all blocks across the gene to describe the linkage disequilibrium (LD) structure of the DBP gene. The distribution of selected SNPs in the DBP gene based on each dataset is shown in the upper panel of Figure 1. Seven common SNPs across the DBP gene were used in these three databases. Three consensus haplotype blocks across the DBP gene were observed in the Caucasian population (A-C), four blocks in Han Chinese and eight blocks in African American populations. In each ethnic population, the selected SNPs and the size of the study populations were different between soures, and therefore the LD structures of different sources were observed to be different from each other.

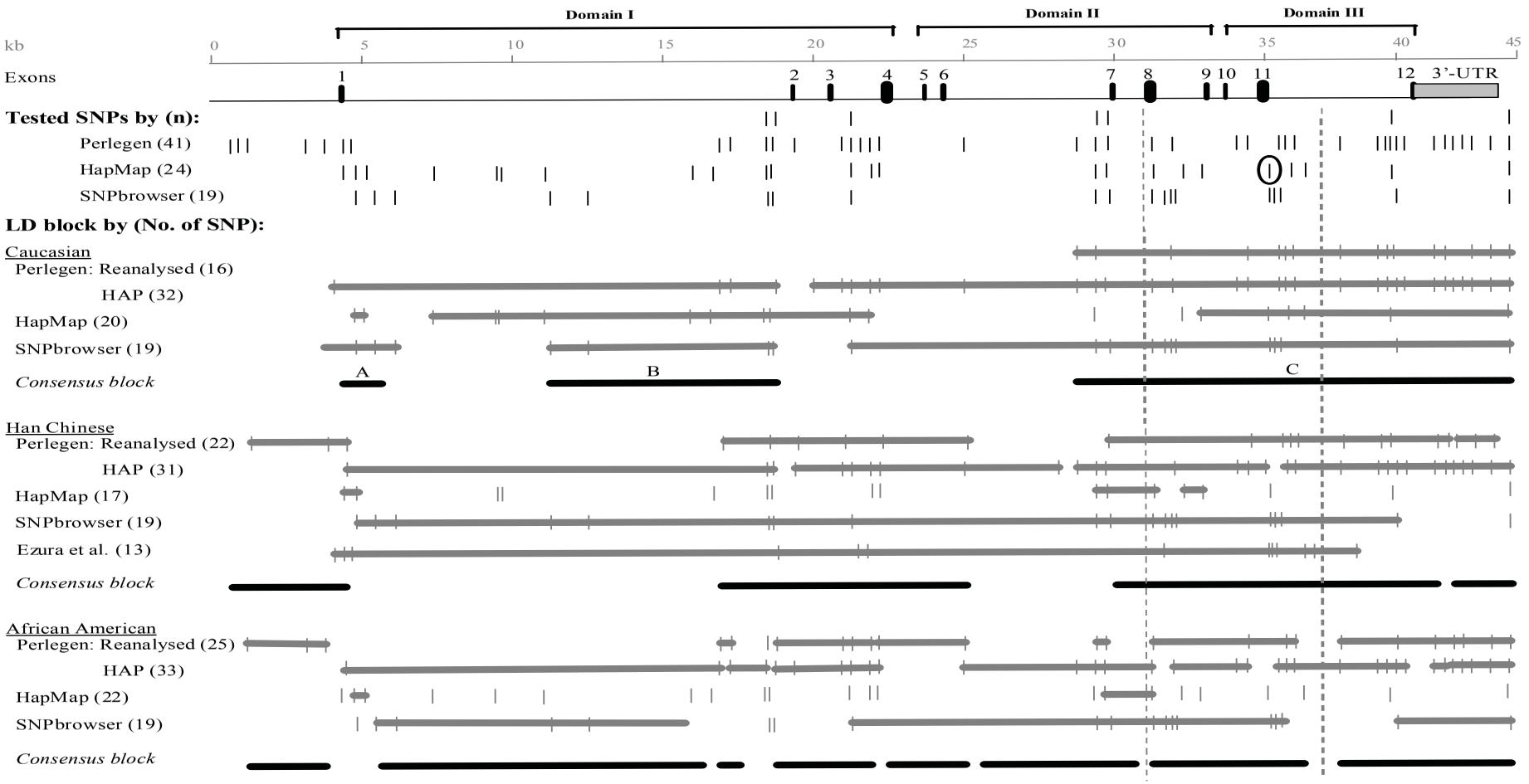
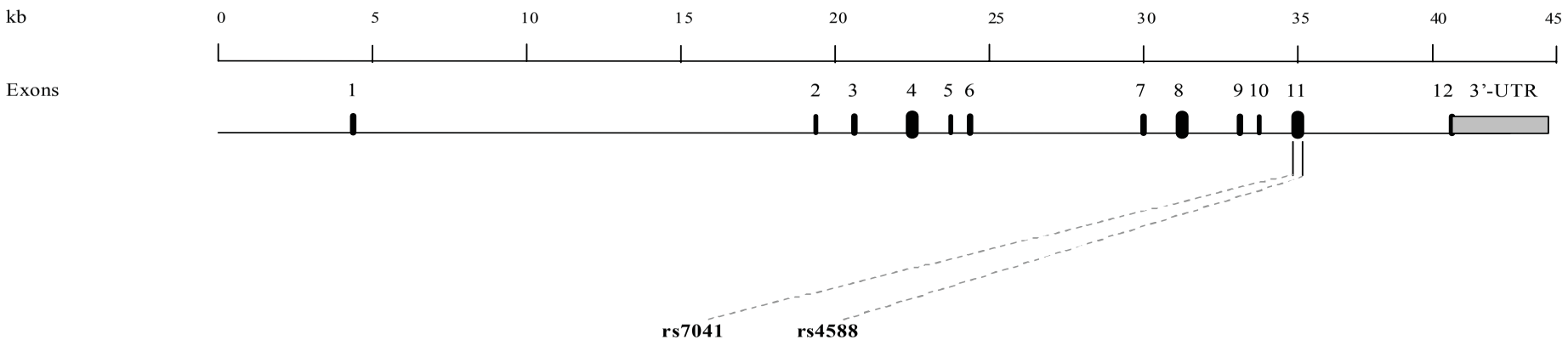


Figure 1 Haplotype structure across the vitamin D binding protein (DBP) gene in different ethnic populations. The triple domain structure of the protein and gene structure of DBP are indicated on the top. The numbers of analysed SNPs are in parentheses of each dataset. SNPs from Perlegen, HapMap and SNPbrowser databases are indicated as small bars below the DBP gene. The SNP with a circle in the HapMap database indicates the SNP rs7041. The grey lines and bars indicate the haplotype blocks and SNPs from our reanalysed LD structure according to reanalysed Perlegen genotype data, and according to original Perlegen data, HapMap, SNPbrowser as well as Ezura et al. [24]. Common SNPs across databases are indicated with arrows. Consensus blocks are shown in black lines below different datasets of each ethnic population. Three consensus haplotype blocks across the DBP gene in the Caucasian population is labelled with A-C. An overlapping region of haplotype blocks among three ethnic populations is indicated between vertical dash lines.



	rs7041	rs4588			
	E416D	T420K		Number	Freq. (%)
hap 1	**G**	**C**	**Caucasian**	6148	**56**
Gc1s	Glu	Thr	Han Chinese	107	32
			Japanese	384	18
			African American	68	15
hap 2	**T**	**A**	**Caucasian**		**27**
Gc2	Asp	Lys	Han Chinese		24
			Japanese		22
			African American		10
hap 3	**T**	**C**	**Caucasian**		**16**
Gc1f	Asp	Thr	Han Chinese		44
			Japanese		39
			African American		75

Figure 2 Structures and allele frequency comparison of haplotypes according to SNPs rs7041 and rs4588 of the DBP gene. Arrows below the DBP gene structure indicate the location of SNPs. Haplotype and protein variation alleles are in the left part of the lower panel. The comparisons of allele frequency in different ethnic populations are shown in the right part of the lower panel. "Caucasian population" means the Rotterdam study population. Han Chinese and African American were from the Coriell DNA panel (see Materials and Methods), Japanese data is from Ezura et al.[24].

Exon 11 contains two non-synonymous SNPs, rs7041 and rs4588, which are located in an overlapping region of the 3'-end haplotype blocks among the three ethnic populations (between two vertical lines in Figure 1). Because of high linkage between these two SNPs, only three instead of the possible four haplotypes were present in our population, corresponding to three Gc isoforms distinguished at the protein level. The allele frequency distribution of haplotypes in different study populations is shown in Figure 2.

Association study of DBP haplotypes

Serum vitamin D level: We measured serum vitamin D level in a sub-group (n = 1,317) of the study population. The vitamin D level in our study population [mean and standard deviation (SD)] of serum $25(OH)D_3$ level was 65.5 ± 27.3 nmol/l (Table 1). Haplotype 1 was found to be significantly associated with increased serum level of $25(OH)D_3$ ($p = 3 \times 10^{-4}$ by a linear regression test for trend) and 1, $25(OH)_2D_3$ ($p = 4 \times 10^{-4}$ for trend). In contrast, haplotype 2 was found to be highly significantly associated with decreased serum level of $25(OH)D_3$ ($p = 3 \times 10^{-6}$) and 1, $25(OH)_2D_3$ ($p = 5 \times 10^{-5}$). The difference in serum vitamin D level between the extreme genotype groups was 0.2 – 0.6 standard deviation. No significant difference in serum vitamin D level by haplotype 3 genotype was found. The associations were independent of age and gender.

Fracture risk and bone phenotypes: No significant difference of fracture incidence was observed by DBP haplotypes (Table 2), although we found a small increasing trend of fracture risk by the DBP haplotype 1. The baseline characteristics by haplotype 1 of the DBP gene are shown in Table 3. In our study population, we did not observe a relationship between DBP haplotype 1 genotype and age, gender, body height, weight, clinical fracture and BMD.

Influence of dietary calcium intake: In the subgroup with record of dietary calcium intake, we observed a similar trend (as seen in the whole study population) and borderline significant association of the DBP haplotype 1 and fracture risk ($p = 0.06$, Table 4). We further investigated the relationship between DBP haplotype 1 and clinical fracture risk after stratification for low or high dietary calcium intake. In the low dietary calcium intake group, we found that subjects homozygous for haplotype 1 had 47% increased risk of clinical fracture compared to non-carriers ($p = 0.02$). The association remained when we adjusted for age and gender. In the high dietary calcium intake group, no association was found. The interaction p-value between the haplotype 1 genotype and dietary calcium intake for clinical fracture was 0.06 after adjustment for age and gender. No effect of the interaction between DBP genotype and calcium intake on BMD and serum vitamin D level was found. We also stratified calcium intake as tertiles and quartiles, and similar patterns of increased fracture risk by DBP haplotype2 was observed in those with a relatively lower calcium intake (Fig. 3).

Interaction between DBP and VDR: We stratified the study population as non-

Table 1 Serum vitamin D level by DBP (Gc) haplotypes in 1,317 elderly men and women

	Total	By haplotypes of DBP Codon416-420			Non-carrier vs.	p-value#
		Non-carrier	Heterzygous	Homozygous	Homozygous (SD)	
25-OH-D_3 (nmol/l)	65.5 ± 27.3 (1312)					
Hap 1 (Gc1s)		61.1 ± 26.3 (269)	65.4 ± 26.9 (647)	68.5 ± 28.3 (396)	↑ 0.3	3×10^{-4}*
Hap 2 (Gc2)		74.2 ± 28.3 (665)	63.0 ± 25.6 (540)	58.5 ± 27.1 (107)	↓ 0.6	3×10^{-6}*
Hap 3 (Gc1f)		65.0 ± 27.2 (917)	66.7 ± 27.6 (360)	64.9 ± 27.1 (35)	↓ 0.004	0.57
1,25-$(OH)_2$-D_3 (pmol/l)	109.3 ± 30.3 (1317)					
Hap 1 (Gc1s)		103.7 ± 29.1 (270)	109.8 ± 30.8 (648)	112.4 ± 29.9 (399)	↑ 0.3	4×10^{-4}*
Hap 2 (Gc2)		113.0 ± 30.9 (668)	105.6 ± 29.2 (541)	105.7 ± 28.9 (108)	↓ 0.2	5×10^{-5}
Hap 3 (Gc1f)		108.9 ± 29.7 (922)	111.0 ± 31.5 (360)	102.0 ± 32.2 (35)	↓ 0.2	0.18

Data are presented as mean ± SD (number of subjects), Vitamin D level was adjusted for age and gender
#: p-value for trend (*, estimated by linear regression analysis) and ANOVA

Table 2 Risk of clinical fracture by DBP haplotypes in 6,181 elderly men and women

	Total	Haplotypes of DBP Codon416-420			p-value
		Non-carrier	Heterzygous	Homozygous	
Hap 1 (Gc1s)					
Case/total (%)	905/6181 (14.6)	155/1183 (13.1)	456/3068 (14.9)	294/1930 (15.2)	0.13*
Crude HR (95% CI)		1	1.16 (0.95-1.41)	1.19 (0.97-1.41)	0.13*
Adjusted HR (95% CI)#		1	1.13 (0.93-1.38)	1.20 (0.97-1.49)	0.13*
Hap 2 (Gc2)					
Case/total (%)		482/3221 (15.1)	348/2477 (14.0)	72/483 (14.9)	0.56
Crude HR (95% CI)		1	0.92 (0.80-1.07)	0.99 (0.76-1.29)	0.56
Adjusted HR (95% CI)		1	0.89 (0.77-1.04)	0.95 (0.72-1.25)	0.35
Hap 3 (Gc1f)					
Case/total (%)		649/4348 (14.9)	239/1686 (14.2)	17/147 (11.6)	0.24*
Crude HR (95% CI)		1	0.94 (0.80-1.10)	0.74 (0.45-1.24)	0.24*
Adjusted HR (95% CI)		1	0.95 (0.81-1.12)	0.84 (0.50-1.42)	0.42*

#: Hazard Ratio (HR) was adjusted for age and gender
*: Trend p-value estimated by linear regression analysis

Table 3 Characteristics of the study population of 6,181 elderly men and women by DBP haplotype 1 (Gc1s)

Characteristic*	Total Cohort	Haplotype 1 of DBP Codon416-420			p-value
		Non-carrier	Heterozygous	Homozygous	
Number (%)	6181	1183 (19.1)	3068 (49.6)	1930 (31.2)	
Female (%)	3689 (59.7)	703 (59.4)	1831 (59.7)	1155 (59.8)	0.97
Age (years)#	69.4 ± 9.1	69.3 ± 9.1	69.7 ± 9.3	69.1 ± 8.9	0.07
Height (cm) #	166.8 ± 9.5	166.7 ± 9.4	166.6 ± 9.5	167.0 ± 9.7	0.11
Weight (kg) #	73.2 ± 12.0	73.2 ± 12.3	73.2 ± 12.3	73.0 ± 11.5	0.68
Dietary Calcium intake (g/day) #	1.13 ± 0.36	1.13 ± 0.37	1.13 ± 0.36	1.12 ± 0.35	0.36
Femoral neck BMD (g/cm^2) #	0.84 ± 0.14	0.84 ± 0.13	0.84 ± 0.14	0.84 ± 0.14	0.90
Lumbar spine BMD (g/cm^2) #	1.09 ± 0.20	1.10 ± 0.20	1.09 ± 0.20	1.09 ± 0.19	0.82

*: The following adjustments were applied:
- Age: adjusted for gender
- Height: adjusted for age and gender
- Weight: adjusted for age, gender, and body height
- Dietary Calcium intake: subset n = 4747; adjusted for age, gender and dietary energy intake
- BMD: subset n = 5027, adjusted for age, gender, height, and weight

#: Data are presented as mean ± SD (number of subjects)

carriers of DBP haplotype 1 and carriers of DBP haplotype 1 (Table 5). In the group of non-DBP haplotype 1 carriers, fracture risk was not different by the VDR haplotype 1 ($p = 0.70$). However, the subjects who carried DBP haplotype 1 and were homozygous for VDR block 5-haplotype 1, had a 33% increased fracture risk ($p = 0.008$ by the Cox model). This association was independent of age and gender. The p-value of interaction between the DBP genotype and VDR genotype in relation to fracture risk was $p = 0.09$.

No interaction between DBP or VDR genotypes in relation to vitamin D level or BMD was observed in the study population (data not shown).

DISCUSSION

In this study we first compared the LD structure of the DBP gene according to different haplotype databases. The haplotype structure analysis may facilitate to identify functional SNPs (in both association and functionality studies) of the DBP gene for disorders such as osteoporosis. The spans and boundaries of haplotype blocks were found to be different between different datasets. The different selection of SNPs, determination of common SNPs for the LD analysis, definition of haplotype block (e.g., threshold of D') and the size of the study population can contribute to the LD map differences. We observed the most common haplotypes (haplotype 1 and haplotype 2 constructed by rs7041 and rs4588), to reside in an LD block "C" and to correspond to the common protein isoforms of the Gc protein: Gc1s and Gc2. The LD block spans > 20 kb from exon 7 to the 3' end of the gene, including the 3'-UTR. Thus, variants in this region and linked to the two SNPs we analysed, can contribute to explain any observed association.

The classical function of DBP is to store and prolong the half-life of circulating vitamin D metabolites. DBP binds 88% and 85% of serum $25(OH)D_3$ and 1, $25(OH)_2D_3$, respectively[20]. In our study, haplotype 1 (corresponding to Gc1s) was positively correlated with both $25(OH)D_3$ and 1, $25(OH)_2D_3$, while haplotype 2 (Gc2) was negatively correlated and haplotype 3 (Gc1f) was not associated with vitamin D level. Lauridsen et al.[40] demonstrated that Gc1 (consisting of 80% Gc1s and 20% Gc1f in our analysis) was associated with increased plasma vitamin D and DBP level, and the DBP Gc2 allele was negatively correlated to plasma vitamin D and DBP level. Thus, our finding of a relationship between haplotype 1 and vitamin D level indeed reflects the correlation between serum DBP level and vitamin D level, and this is in line with a previous report[23].

A relationship of Gc1 (associated with increased serum vitamin D level) and increased fracture risk was reported by Lauridsen et al.[25,40]. We also observed a similar trend in the complete study population (Table 2) and borderline significance in a sub-group population based on availability of dietary calcium intake data (Table 5). This relationship can not be explained by the effect of DBP on the

Table 4 Relationship between DBP haplotype 1 (Gc1s) and clinical fracture by high or low dietary Calcium intake

	Total	Haplotype 1 of DBP Codon416-420			p-value**
		Non-carrier	Heterzygous	Homozygous	
Total*					
Case/total (%)	651/4743 (13.7)	102/891 (11.4)	331/2342 (14.1)	218/1510 (14.4)	0.06
Crude HR (95% CI)		1	1.25 (1.00-1.56)	1.27 (1.00-1.60)	0.08
Adjusted HR (95% CI)#		1	1.19 (0..95-1.48)	1.23 (0.97-1.56)	0.11
Dietary Ca^{++}-intake < 1.09 g/day					
Case/total (%)	341/2371 (14.4)	49/439 (11.2)	167/1158 (14.4)	125/774 (16.1)	0.02
Crude HR (95% CI)		1	1.30 (0.94-1.78)	1.47 (1.06-2.05)	0.02
Adjusted HR (95% CI)#		1	1.22 (0.89-1.68)	1.42 (1.02-1.97)	0.03
Dietary Ca^{++}-intake > 1.09 g/day					
Case/total (%)	310/2372 (13.1)	53/452 (11.7)	164/1184 (13.9)	93/736 (12.6)	0.48
Crude HR (95% CI)		1	1.20 (0.88-1.64)	1.07 (0.77-1.50)	0.88
Adjusted HR (95% CI)		1	1.15 (0.85-1.57)	1.05 (0.75-1.47)	0.94

*: Subjects with calcium intake data
#: Hazard Ratio (HR) was adjusted for age and gender
**: Trend p-value estimated by linear regression analysis

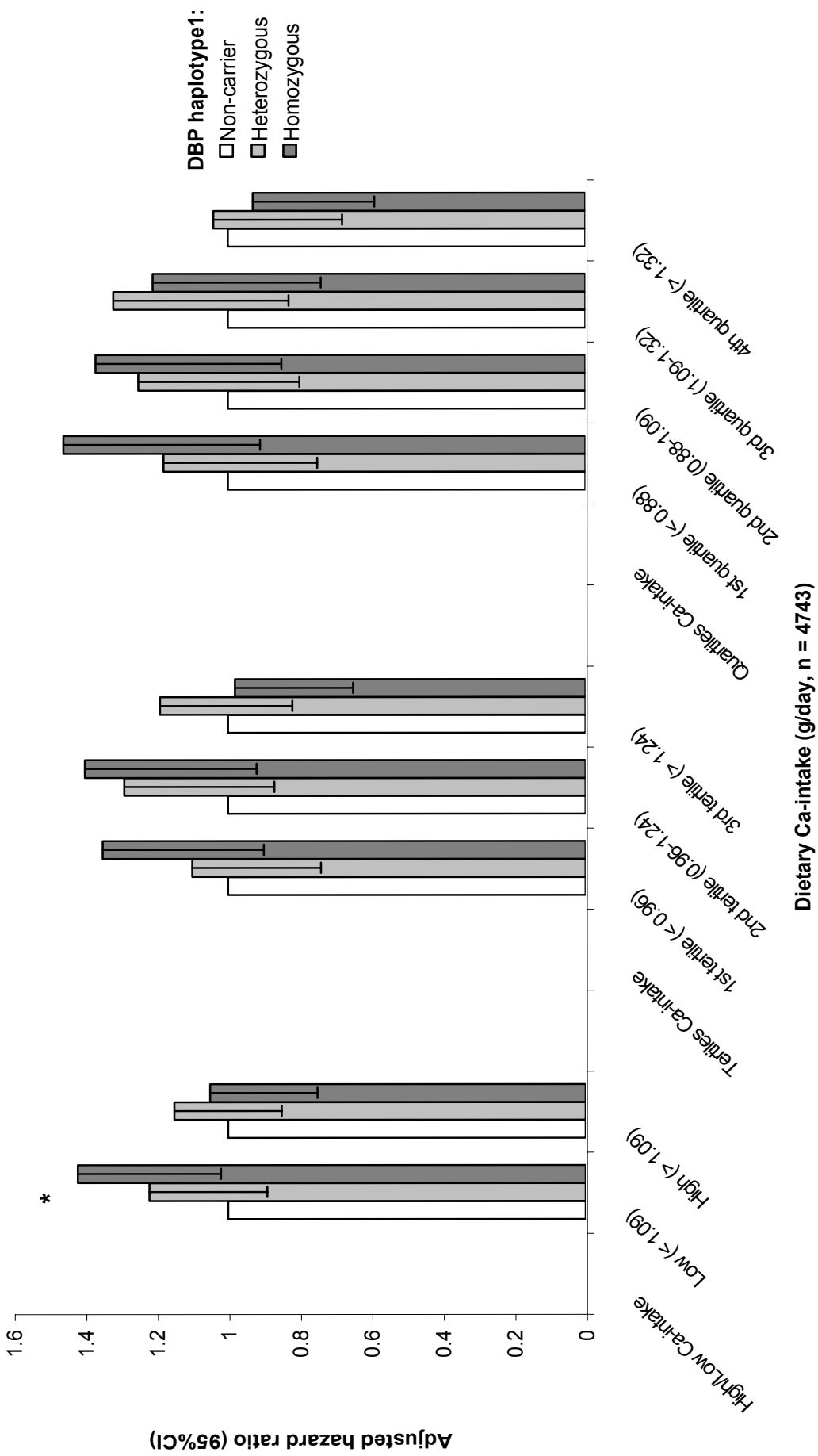


Figure 3 Relationship between DBP haplotype 1 and fracture risk stratified by dietary calcium intake (g/day) in a subgroup study population (n = 4743). Dietary calcium intake was stratified as high/low by the median intake, and as tertiles and quartiles. Hazard ratios (95% CI) were adjusted for age and gender and were calculated using the DBP haplotype 1 non-carriers as the reference group. The star (*) indicates p-vale < 0.05.

Table 5 Interaction between DBP haplotype 1 (Gc1s) and VDR block 5-haplotype 1 in relation to clinical fracture

	Total	Genotype of VDR block 5-haplotype 1			p-value**
		Non-carrier	Heterozygous	Homozygous	
Total study population					
Case/total (%)	881/6009 (14.7)	255/1930 (13.2)	428/2887 (14.8)	198/1192 (16.6)	0.009
Crude HR (95% CI)		1	1.14 (0.98-1.33)	1.29 (1.07-1.55)	0.007
Adjusted HR (95% CI)#		1	1.11 (0.95-1.30)	1.27 (1.06-1.53)	0.01
By VDR genotype					
Non-carrier of DBP hap 1					
Case/total (%)	153/1158 (13.2)	46/366 (12.6)	77/566 (13.6)	30/226 (13.3)	0.90
Crude HR (95% CI)		1	1.06 (0.74-1.53)	1.09 (0.69-1.72)	0.70
Adjusted HR (95% CI)#		1	1.02 (0.71-1.47)	1.08 (0.68-1.70)	0.77
Carrier of DBP hap 1					
Case/total (%)	728/4851 (15.0)	209/1564 (13.4)	351/2321 (15.1)	168/966 (17.4)	0.006
Crude HR (95% CI)		1	1.16 (0.98-1.38)	1.33 (1.09-1.63)	0.005
Adjusted HR (95% CI)		1	1.14 (0.96-1.35)	1.32 (1.07-1.61)	0.008

*: Subjects with DBP and VDR genotype data
#: Hazard Ratio (HR) was adjusted for age and gender
**: Trend p-value was estimated by Cox regression analysis

vitamin D endocrine system, since increased vitamin D level is not associated with increased fracture risk. Moreover, according to the "free hormone" hypothesis, the free hormone performs the biological activity rather than bound hormone[41], and the free vitamin D level is only <12-15% of total vitamin D level we measured.

The genomic and protein structures also reveal at least two different functions of the DBP. The vitamin D-binding region of the DBP is located between residues 35-49 in the N-terminal end of the DBP (domain I in Figure 1)[20]. The two non-sysnonymous coding SNPs are located in the C-terminal end (domain III, from residue 375 to the end) with a single glycosylation site nearby. Our LD structure analysis showed that these two domains are in different haplotype blocks for all ethnic populations. These two domains have a different function: domain I can bind to sterols, domains II and III are responsible for non-sterol binding activities of DBP[42]. The macrophage/osteoclast activating activity is also related to domain III[16,42]. The glycosylation of DBP in domain III is important for macrophage and osteoclast activation, while it has been shown that binding of vitamin D does not influence this activity[43]. This indicates that domain III plays an independent role of domain I in the function of stimulating osteoclast.

DBP haplotype 1 is associated with increased serum DBP level[40]. We hypothesize that the association of DBP haplotype with increased DBP levels is caused by SNPs in this block that result in increased mRNA levels, and thereby in increased DBP protein levels. The most likely position of these SNPs is in the 4.3 kb large 3'-UTR area. About 17 SNPs have been identified in this region and further functional studies will be necessary to determine if and which of these SNPs cause differential mRNA expression. We note that we have recently identified a similar situation for SNPs in the 3'-UTR of the VDR gene[28].

DBP-MAF is an activator of osteoclasts, while *in vivo* studies demonstrated that a calcium-deficient diet in rat studies could significantly increase the number of osteoclast in different sites of bone[26,27]. In our sub-group population of human subjects with record of dietary calcium intake, we found that dietary calcium intake influenced the strength of the association between DBP haplotype 1 (Gc1s) and fracture risk. The association between the DBP haplotype 1 and fracture risk was only seen in subjects with a relatively low dietary calcium intake (< 1.09 g/day). We note that Lauridsen et al.[25] showed a similar relationship between Gc1 and fracture risk in a population of Danish Caucasian with a dietary calcium intake at 0.8 g/day level[40]. Another study in a Japanese population[24] found the Glu-allele at codon 416 (corresponding to our haplotype 1) to be associated with decreased radial BMD. The dietary calcium intake of Japanese populations is considered to be lower [0.6 g/day[44]] than ours and the Danish study populations. Although this seems to be in line with our observations, the association between the DBP haplotype 1 (or Gc1) and fracture risk in ours and the Danish studies was independent of BMD. We therefore also analysed subjects with dietary calcium intake < 0.6 g/day for DBP related differences in BMD. No association was observed. However, since we

had only 147 subjects in this category, we were not sure whether there is really no association, or we did not have sufficient statistical power to detect the relationship between DBP haplotypes and BMD.

In this study we also demonstrated interaction of VDR and DBP polymorphisms on fracture risk. This gene-gene interaction on fracture risk is likely to involve two aspects of bone modelling and remodelling, which are active during the whole lifespan, and includes bone formation (by osteoblasts) and bone resorption (by osteoclasts) to maintain serum calcium homeostasis and bone metabolism. The block 5-haplotype 1, fracture risk allele, of the VDR gene results in a lower expression of VDR mRNA in osteoblasts[28]. This indicates that the osteoblast of block 5-haplotype 1 carriers could give a compromised response to vitamin D since lower numbers of VDR protein are present in the cells. Thus, this is likely to result in lower bone formation for the VDR block 5-haplotype 1 carriers. As mentioned above, DBP haplotype 1 is associated with increased DBP level, and this might result in increased osteoclast activity, and thus increased bone resorption. Taken together, carriers of both risk alleles might have a lower bone formation and higher bone resorption resulting in increased bone turnover which negatively influences bone quality and strength[45]. Thus, bone of both risk-allele carriers becomes weaker with time, and thus fracture risk consequently increases. This hypothesis needs to be confirmed by functionality studies.

In conclusion, we summarised the haplotype structure of the DBP gene according to three major haplotype datasets. The most commonly used haplotypes (or protein isoforms) of the DBP gene were found to be associated with serum vitamin D level, while an additive effect of VDR block 5-hap 1 and DBP haplotype 1 (Gc1s) on fracture risk was observed, and DBP haplotype 1 was found to be associated with fracture risk especially in a low calcium intake group. Further functional analysis of DBP 3'-UTR variants is necessary to investigate the underlying mechanism, while the relationship between DBP haplotypes in other LD blocks and osteoporosis remains to be studied.

ACKNOWLEDGEMENTS

We thank all participants of the Rotterdam study and the general practitioners, pharmacists, and the many field workers at the research center in the Netherlands.

This project was funded by the Dutch Research Organisation (NWO 903-46-178, 925-01-010, 014-90-001 and 911-03-012) and the European Commission under grant QLK6-CT-2002-02629 ("GENOMOS").

REFERENCES

1. Ringe, J. *Pathophysiology of postmenopausal osteoporosis*, 2 (Science Press Ltd, London, 2004).
2. Lohmueller, K.E., Pearce, C.L., Pike, M., Lander, E.S. & Hirschhorn, J.N. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* **33**, 177-82 (2003).
3. Uitterlinden, A.G. et al. A large-scale population-based study of the association of vitamin D receptor gene polymorphisms with bone mineral density. *J Bone Miner Res* **11**, 1241-8 (1996).
4. Uitterlinden, A.G. et al. Relation of alleles of the collagen type Ialpha1 gene to bone density and the risk of osteoporotic fractures in postmenopausal women. *N Engl J Med* **338**, 1016-21 (1998).
5. van Meurs, J.B. et al. Association of 5' estrogen receptor alpha gene polymorphisms with bone mineral density, vertebral bone area and fracture risk. *Hum Mol Genet* **12**, 1745-54 (2003).
6. Fang, Y. et al. Cdx-2 polymorphism in the promoter region of the human vitamin D receptor gene determines susceptibility to fracture in the elderly. *J Bone Miner Res* **18**, 1632-41 (2003).
7. Murray, R.E., McGuigan, F., Grant, S.F., Reid, D.M. & Ralston, S.H. Polymorphisms of the interleukin-6 gene are associated with bone mineral density. *Bone* **21**, 89-92 (1997).
8. Ioannidis, J.P. et al. Differential genetic effects of ESR1 gene polymorphisms on osteoporosis outcomes. *JAMA* **292**, 2105-14 (2004).
9. Grymonprez, A. et al. Vitamin D metabolites in childhood nephrotic syndrome. *Pediatr Nephrol* **9**, 278-81 (1995).
10. Parfitt, A.M. et al. Vitamin D and bone health in the elderly. *Am J Clin Nutr* **36**, 1014-31 (1982).
11. Safadi, F.F. et al. Osteopathy and resistance to vitamin D toxicity in mice null for vitamin D binding protein. *J Clin Invest* **103**, 239-51 (1999).
12. Schneider, G.B., Benis, K.A., Flay, N.W., Ireland, R.A. & Popoff, S.N. Effects of vitamin D binding protein-macrophage activating factor (DBP-MAF) infusion on bone resorption in two osteopetrotic mutations. *Bone* **16**, 657-62 (1995).
13. Adebanjo, O.A., Moonga, B.S., Haddad, J.G., Huang, C.L. & Zaidi, M. A possible new role for vitamin D-binding protein in osteoclast control: inhibition of extracellular Ca2+ sensing at low physiological concentrations. *Biochem Biophys Res Commun* **249**, 668-71 (1998).
14. Datta, H.K., MacIntyre, I. & Zaidi, M. Intracellular calcium in the control of osteoclast function. I. Voltage-insensitivity and lack of effects of nifedipine, BAYK8644 and diltiazem. *Biochem Biophys Res Commun* **167**, 183-8 (1990).
15. Yamamoto, N., Naraparaju, V.R. & Orchard, P.J. Defective lymphocyte glycosidases in the macrophage activation cascade of juvenile osteopetrosis. *Blood* **88**, 1473-8 (1996).
16. Haddad, J.G. Plasma vitamin D-binding protein (Gc-globulin): multiple tasks. *J Steroid Biochem Mol Biol* **53**, 579-82 (1995).
17. Lee, W.M. & Galbraith, R.M. The extracellular actin-scavenger system and actin toxicity. *N Engl J Med* **326**, 1335-41 (1992).
18. Dahl, B. et al. Plasma concentration of Gc-globulin is associated with organ dysfunction and sepsis after injury. *Crit Care Med* **31**, 152-6 (2003).
19. Yamamoto, N., Homma, S., Haddad, J.G. & Kowalski, M.A. Vitamin D3 binding protein required for in vitro activation of macrophages after alkylglycerol treatment of mouse peritoneal cells. *Immunology* **74**, 420-4 (1991).
20. White, P. & Cooke, N. The multifunctional properties and characteristics of vitamin D-binding protein. *Trends Endocrinol Metab* **11**, 320-7 (2000).
21. Braun, A., Bichlmaier, R. & Cleve, H. Molecular analysis of the gene for the human vitamin-D-binding protein (group-specific component): allelic differences of the common genetic GC types. *Hum Genet* **89**, 401-6 (1992).
22. Arnaud, J. & Constans, J. Affinity differences for vitamin D metabolites associated with the genetic isoforms of the human serum carrier protein (DBP). *Hum Genet* **92**, 183-8 (1993).
23. Bouillon, R., Van Assche, F.A., Van Baelen, H., Heyns, W. & De Moor, P. Influence of

the vitamin D-binding protein on the serum concentration of 1,25-dihydroxyvitamin D3. Significance of the free 1,25-dihydroxyvitamin D3 concentration. *J Clin Invest* **67**, 589-96 (1981).
24. Ezura, Y. et al. Association of molecular variants, haplotypes, and linkage disequilibrium within the human vitamin D-binding protein (DBP) gene with postmenopausal bone mineral density. *J Bone Miner Res* **18**, 1642-9 (2003).
25. Lauridsen, A.L. et al. Female premenopausal fracture risk is associated with gc phenotype. *J Bone Miner Res* **19**, 875-81 (2004).
26. Liu, C.C. & Baylink, D.J. Differential response in alveolar bone osteoclasts residing at two different bone sites. *Calcif Tissue Int* **36**, 182-8 (1984).
27. Wright, K.R. & McMillan, P.J. Osteoclast recruitment and modulation by calcium deficiency, fasting, and calcium supplementation in the rat. *Calcif Tissue Int* **54**, 62-8 (1994).
28. Fang, Y. et al. Promoter and 3'-Untranslated-Region Haplotypes in the Vitamin D Receptor Gene Predispose to Osteoporotic Fracture: The Rotterdam Study. *Am J Hum Genet* **77**, 807-23 (2005).
29. Hofman, A., Grobbee, D.E., de Jong, P.T. & van den Ouweland, F.A. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur J Epidemiol* **7**, 403-22 (1991).
30. Perlegen. http://genome.perlegen.com/browser/index.html.
31. Phase. http://archimedes.well.ox.ac.uk/pise/PHASE.html.
32. HaploBlockFinder. http://cgi.uc.edu/cgi-bin/kzhang/haploBlockFinder.cgi.
33. HapMap. http://www.hapmap.org/cgi-perl/gbrowse/gbrowse/hapmap.
34. SNPbrowser. http://events-na.appliedbiosystems.com/mk/get/snpb_landing?isource=fr_E_RD_www_allsnps_com_snpbrowser.
35. Halperin, E. & Eskin, E. Haplotype reconstruction from genotype data using Imperfect Phylogeny. *Bioinformatics* **20**, 1842-9 (2004).
36. Burger, H. et al. The association between age and bone mineral density in men and women aged 55 years and over: the Rotterdam Study. *Bone Miner* **25**, 1-13 (1994).
37. McCloskey, E.V. et al. The assessment of vertebral deformity: a method for use in population studies and clinical trials. *Osteoporos Int* **3**, 138-47 (1993).
38. Van der Klift, M., De Laet, C.E., McCloskey, E.V., Hofman, A. & Pols, H.A. The incidence of vertebral fractures in men and women: the Rotterdam Study. *J Bone Miner Res* **17**, 1051-6 (2002).
39. dbSNP. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=snp&cmd=search&term=human+Gc.
40. Lauridsen, A.L. et al. Plasma concentrations of 25-Hydroxy-Vitamin D and 1,25-Dihydroxy-Vitamin D are Related to the Phenotype of Gc (Vitamin D-Binding Protein): A Cross-sectional Study on 595 Early Postmenopausal Women. *Calcif Tissue Int* **77**, 15-22 (2005).
41. Mendel, C.M. The free hormone hypothesis: a physiologically based mathematical model. *Endocr Rev* **10**, 232-74 (1989).
42. Ray, R. Molecular recognition in vitamin D-binding protein. *Proc Soc Exp Biol Med* **212**, 305-12 (1996).
43. Swamy, N., Ghosh, S., Schneider, G.B. & Ray, R. Baculovirus-expressed vitamin D-binding protein-macrophage activating factor (DBP-maf) activates osteoclasts and binding of 25-hydroxyvitamin D(3) does not influence this activity. *J Cell Biochem* **81**, 535-46 (2001).
44. Tsuchida, K., Mizushima, S., Takahashi, H., Misugi, N. & Soda, K. [Relationship between dietary calcium and bone mineral density before menopause]. *Nippon Koshu Eisei Zasshi* **45**, 121-8 (1998).
45. Bauer, D.C. et al. Change in bone turnover and hip, non-spine, and vertebral fracture in alendronate-treated women: the fracture intervention trial. *J Bone Miner Res* **19**, 1250-8 (2004).

Chapter 9

General Discussion

In this thesis, we investigate the relationship between polymorphisms in the vitamin D endocrine pathway and bone phenotypes. The thesis consists of three major consecutive portions: molecular genetic studies, genetic association studies and functionality studies.

First, we integrated all resources from web databases and publications to determine the most accurate genomic structure of the VDR gene. This complete genomic sequence of the human VDR gene was compared with the mouse VDR sequence to identify high homology regions between these two species as potentially functional regions and as targets for sequencing. 62 polymorphisms were detected by re-sequencing those potentially functional regions, and used to determine the LD structure and tagging SNPs across the VDR gene for different ethnic groups. We then investigated the association between VDR haplotypes (defined by tagging SNPs) and bone phenotypes, including fracture, body height and bone geometry parameters in a large population-based study population and in an independent sample: the LASA population. Potentially functional haplotypes were found according to association and functionality studies. Gene-gene and gene-environment interactions in relation to fracture risk were briefly described in the study. Finally, we performed functionality studies to provide possible molecular mechanism that could explain the association results.

9.1 Usage of bioinformatics in our study

9.1.1 Structure and polymorphisms of VDR gene: Baker et al.[1] reported the DNA sequence of human VDR cDNA in 1988, while Miyomoto et al.[2] and Croft et al.[3] simultaneously described the major genomic structure and four tissue-specific exons at 5'-end of the VDR gene after 10 years. However, the genomic structure of the VDR gene as is reported in the NCBI database[4] so far does not include the results of the Croft study, and the size of the VDR gene is only about 65 kb. In our study, we integrated all published data and data from the Celera database, and described the genomic structure of the VDR gene with a size of > 105 kb (up to exon 1f). Our study and Nejentsev et al.[5] re-sequenced the VDR gene in 30 and 16 chromosomes, respectively, and thus polymorphisms with a minor allele frequency (MAF) of > 3% and 6% could be detected, respectively. Therefore, both studies described a higher quantity and quality of polymorphism data than the Celera[6] and NCBI[7] database (before these studies submitted their data to the dbSNP database). 40% of the polymorphisms we found were not in those databases or in any publications.

9.1.2 Haplotype structures from different resources: Our study and the study of Nejentsev[5] focused on analysing the haplotype structure of and around one single gene, e.g. the VDR gene. The initiatives of Perlegen[8], HapMap[9] and SNPbrowser[10] create a genome-wide haplotype map of SNPs. The studies on single

gene LD structure so far have high densities of selected SNPs. For example in our study we determined an LD map used on average one SNP per 0.54 kilo-basepairs (kb), as compared to one SNP per 1.7 kb by Nejentsev et al.[5]. The three databases generated LD maps of the VDR gene with a density of one SNP per 2.2 – 9.5 kb (including SNPs with MAF < 10%). With a high density of SNPs, a similar haplotype block definition and the same program (PHASE[11]), both our study and the study by Nejentsev et al. identified a similar and higher resolution LD map compared to the three databases. The consensus LD map of the VDR gene according to all studies and databases shows differences in haplotype block size. The LD block sizes determined in our study and Nejentsev et al. (< 17 kb) are smaller than that of the three databases (see Figure 2 in chapter 3).

These three databases analyse LD structure of the whole genome. Large numbers of SNPs are included to be analysed: Perlegen included 1.6 million SNPs in February 2005, HapMap have 2.25 million in its database (until August 2005), SNPbrowser contains 5 million till August 2005. The data of Perlegen and HapMap have been merged together, and by the end of 2005 Perlegen plans to test another 4.6 million SNPs in public databases and add them to the map. Although more and more SNPs have and will be included in those databases, the selected SNPs are not based on the gene structure, or based on functional domains of the gene. This contrasts somewhat with our approach as is shown in the Figure 2 of chapter 3 for the VDR gene and the Figure 1 of chapter 8 for the DBP gene. Reasons why these databases do not show consistent LD structures are not only because of the lack of common SNPs in all three databases (see selected SNPs in the Figure 2 of chapter 3 and Figure 1 of chapter 8), but also because of different definitions of haplotype blocks that were used. For example, Perlegen selects MAF of SNPs > 10% and uses the Hap algorithm[12] to infer haplotypes from their diploid genotype data[13]. HapMap and SNPbrowser use the PHASE algorithm[14] to estimate pair-wise LD of SNPs and keep the selection of MAF and haplotype block cutoff value open for users. MAF is known to affect LD structure and determination of tagging SNP, and MAF depends on sample size. Small sample size requires high MAF cutoff value for the LD analysis. Another pitfall of these databases is that they mostly use computational constraints on SNP selection, and ignore common variations and the functional aspect in the SNP selection criteria. For example, two non-anonymous SNPs, rs7041 and rs4588 SNPs, are close to each other and in exon 11 of the DBP gene. They define the most common DBP phenotypes, and two phenotypes were found to be significantly correlated with serum vitamin D level in our and other studies, and also have interaction with VDR haplotype for fracture risk. They are functional SNPs of the DBP gene. However, SNP rs7041 was not included in the LD analyses of all three databases, and was not determined as haplotype tagging SNP. Unlike most genome-wide studies, databases and DNA-array design, we provide a way to determine haplotype tagging SNPs (htSNPs) not only according to computer programs based LD analysis but also taking SNP that are frequently used

in other studies and taking functional aspects into account. For the LD analysis of a gene, we suggest to combine information from Perlegen, HapMap, SNPbrowser and other publications, and to use the same haplotype definition and program to achieve a relatively accurate haplotype structure of the gene.

9.2 Genetic studies

9.2.1 Linkage or candidate gene association study: The initial genetic study design for inherited disease dominantly focused on linkage analysis in pedigrees. This design is very powerful to identify rare high-risk disease alleles, such as in single-gene Mendelian diseases. The linkage design searches for the disease-causal alleles in relative large regions of the genome. Because of low frequent recombination events within most families, it is mostly difficult to narrow the interesting region below several megabases. Inversely, association design is more powerful to detect common alleles with modest disease risks[15]. For the association design it is easier to recruit large numbers of unrelated affected individuals than to collect large numbers of families, especially for diseases in the elderly, and it can investigate the relationship between single variations and disease. However, association analysis requires markedly higher density of genetic markers than linkage analysis. The combined study design, candidate gene analyses within linked regions, which have been greatly accelerated by the availability of the complete sequence of the human genome, provides a reasonable strategy to identify fine disease genes in suggestive-linkage regions[16]. SNPs are the most frequent sequence variations of the genome by now, more than eleven million SNPs have been identified in the human genome. Since SNPs are relatively easy to be analysed compared to multi-allele polymorphisms, they are commonly used in genetic association studies.

We can distinguish direct and indirect association studies, which are based on functional polymorphisms and "marker" polymorphisms respectively. A direct association study can be carried out for the putatively functional SNPs based on the evidence of functionality studies, e.g., the Cdx-2 polymorphism of the VDR gene. Alternatively, an indirect association between unknown functional SNPs and a disease requires to test a dense map of marker SNPs and to identify the LD map across the gene, e.g., other analysed SNPs across the VDR gene. Both direct and indirect association studies have proven to be effective methods to identify association of candidate genes that have been implicated in disease pathogenesis in our study.

9.2.2 Whole genome association (WGA) study: WGA study, a non-hypothesis study design and analysing association between sequence variations in whole genome and one phenotype, can theoretically analyse the entire genome at once in an unbiased fashion. This design intends to investigate highly dense polymorphic markers across the whole genome for diseases, and is a promising research direction

to identify the genetic basis of complex disorders. For WGA, high-density SNP genotyping assays (gene chips) are required to be used in a case-control design. For example, Perlegen offers a set of 1.6 million validated SNP assays to genotype pooled case or control DNAs, and Affymetrix is developing a 500,000 SNP-chip[17]. However, SNP selection strategies (like those based on haplotype tagging, or on potential function) differ for the different platforms, e.g., Affymetrix selects equi-distant SNPs, Perlegen focuses on tagging SNPs and Illumina selects coding and tagging SNPs.

9.2.3 Genetic meta-analysis: A retrospective meta-analysis (for instance, Mann et al.[18] for COLIA1 Sp1 SNP and our study in chapter 8) of the VDR Bsm I RFLP can systematically summarize and obtain sufficient statistical power to detect an overall modest genetic effect according to published association studies (with different study sizes). However, publication bias (positive results are likely to be published) and heterogeneity between studies (e.g. the meta-analysis of Bsm I genotype and fracture risk in chapter 5.1) are remarkably influencing the reliability of the conclusion from meta-analyses. The GENOMOS (Genetic Markers for Osteoporosis) project provides an example of another design of meta-analysis of genetic association studies: prospective meta-analysis. This design involves the study of several candidate gene polymorphisms in relation to osteoporosis-related outcomes in approximately 20,000 individuals drawn from eight European centers. Genotype data is obtained prospectively from all centers and all are included in the analyses to avoid the publication bias. Standardization of genotyping and analysing can avoid attributing to genuine genetic variability difference between study teams that might be simply due to analytical inconsistencies. This design detected the genetic effect of ESR1 gene polymorphisms on osteoporosis[19].

9.3 Individual polymorphism or haplotypes in association study

Initially most genetic association studies investigated relationships between individual polymorphisms and clinical endpoints. Nowadays, more and more studies are focusing on studies of haplotypes, as haplotypes combine genetic information of individual polymorphisms. Haplotype-based methods offer a powerful approach to disease gene mapping[20], and the Human Genome Project, the SNP project and the HapMap project make great efforts to provide haplotype data for association studies. In a study, we identified the VDR Cdx-2 polymorphism to be associated with fracture risk in the Rotterdam Study population, especially for non-vertebral fracture in women (chapter 4). In our more recent and more comprehensive study (1st section of chapter 5), we used the Cdx-2 and GATA polymorphisms as haplotype tagging SNPs (htSNPs) of the VDR LD block 2, and demonstrated that haplotype 1 defined by these two SNPs and containing the risk allele of the Cdx-2 SNP was associated with any fracture in the complete study population. Furthermore,

functionality studies (chapter 8) showed that GATA is another potential functional polymorphism (in addition to Cdx-2) in the haplotype block. This suggests that several functional SNPs in a haplotype block can "collaborate" to confer risk for a certain disease.

In the 3'-end of the VDR gene, most previous studies performed association studies for individual RFLPs: Bsm I, Apa I and Taq I which are in the same haplotype block as 3'-UTR SNPs. Bsm I and Apa I RFLPs are in intron 8, and although Taq I is in exon 9, it does not change the amino acid sequence of the VDR, so they are all anonymous (not functional) RFLPs. In our study (1st section of chapter 5), haplotype 1 defined by 3'-UTR SNPs was found to be associated with fracture risk, and similar associations were detected for haplotype 1 of Bsm-Apa-Taq and the "b-allele" of Bsm I RFLP. The function of 3'-UTR of a gene likely involves influencing the gene expression by regulating the stability of mRNA. Thus, we hypothesized that the causal polymorphisms to potentially explain those associations in the haplotype block were in the 3'-UTR. In the functional study (chapter 7), when we transfected the same coding sequence, as well as complete VDR 3'-UTR sequences of haplotype 1 and 2 (excluding Bsm I and Apa I RFLPs) into different cell lines, the 3'-UTR in haplotype 1 was indeed found to have 15% lower mRNA level compared to haplotype 2. The differences between haplotype 1 and haplotype 2 are limited to 25 sequence variations across 17 kb (LD block 5) in Caucasian population. When we assume the causal nucleotide variants to reside in the 3'-UTR, this is further limited to five polymorphisms. The question arises whether it is one of these variants that "drives" the association, or perhaps the combination of > 2 polymorphisms. This has to be investigated in further studies by analysing 3'-UTR constructs of different combination of nucleotide variants. As a parallel alternative study, one could also perform a similar association study, but now in African Americans. The LD blocks there are smaller, thus, we could "zoom in" on smaller region/haplotype blocks that show association. In any case, we can conclude from this study that VDR haplotypes represent the genetic effect of VDR variations on fracture risk in our population.

Polymorphisms within a gene are interacting with each other (intragenic interaction) to determine the gene functions in a give cell and/or subject. The promoter area regulates production of mRNA while the 3'-UTR is involved in stability/degradation of mRNA and their interaction/combined effects regulates the net availability of the mRNA for translation into the VDR protein. In our study we detected not only haplotypes in separate LD blocks that are related to bone phenotypes, but we also further investigated the combined genetic effect of haplotypes in the promoter region and in the 3'-UTR. Indeed, we observed additive effects of haplotypes in both regions on fracture risk and body height.

In addition to our study, Nejentsev et al.[5] also similarly re-sequenced the complete VDR gene, performed LD analysis, determined haplotype structure across the VDR gene for three major ethnic populations. However, in their subsequent as-

sociation study[21], they analysed 98 individual SNPs instead of haplotypes for type 1 diabetes risk in 3,763 families. They found two intronic SNPs and one 3'-UTR SNP with p = 0.02 – 0.05 (which no longer are significant when p-values are adjusted for multiple testing by the Bonferroni correction). When they tested those three SNPs in an independent case-control (n = 1,587 and 1,827, respectively) study, no significant p-value was found. Finally, they conclude that sequence variation in the VDR gene had no major effect on type 1 diabetes in their populations. Although the conclusion might be true, some questions and discussions remain after this study. First, the authors had LD structure and race-specific tagging SNPs of the whole VDR gene, but did not report the association of VDR haplotypes and diabetes. Although they did not find individual SNPs to be associated with type 1 diabetes, the relationship between VDR haplotypes and diabetes remains to be further investigated. Secondly, since there might be less than 20 common haplotype alleles (frequency > 5%) across the whole VDR gene (according to our study), the number of the analyses using haplotypes is much lower than the number of the analyses using 98 individual SNPs. Thus using haplotypes in the study will reduce the number of analyses thereby diminishing multiple testing. Thirdly, the MAF of one "significant" SNP for diabetes in the family study was 1%. Even in 3,763 subjects, the authors have not sufficient statistical power to detect the subtle genetic effect of the SNP on diabetes. Therefore, it is important to determine the cut off value of allele frequencies for selected SNPs or haplotypes according to the statistical power of the study population before carrying out the association study or LD analysis.

9.4 VDR polymorphisms and bone phenotypes

VDR mediates the endocrine function of vitamin D for calcium and phosphate homeostasis and bone metabolism. Over 1,000 of studies have investigated the relationship between VDR polymorphisms with osteoporosis (especially focusing on the most important clinical endpoints: BMD change and fracture risk), adult stature and bone geometry variables.

9.4.1 Association studies of VDR polymorphisms and BMD: Low BMD is an important risk factor of osteoporosis. As mentioned in chapter 1, many association studies have been performed to test the relationship of VDR polymorphisms and BMD. Furthermore, three studies[22-24] summarized association studies between VDR polymorphisms and BMD. Cooper et al.[22] reported that the BB genotype of the Bsm I was associated with decreased BMD only at the hip, and that younger women with BB genotype had border line significantly lower BMD compared to women with the bb genotype. Nevertheless, such effect disappeared after omitting the data[25] with genotyping error. The meta-analysis of Gong et al.[23] summarized 75 studies on the association between Bsm I, Apa I, Taq I and Fok I RFLPs of

VDR gene with BMD as well as other related skeletal phenotypes. These authors grouped all studies as "positive" and "non-positive", while the "positive" study was defined as studies with association between either one of the b, a, T, or F alleles and increased BMD or other phenotypes. However, this study has several drawbacks. Some phenotypes in the study were not always in line or even comparable with BMD, such as, bone geometry variables, response to vitamin D treatment, bone turnover and fracture, and they should not be counted together in the same phenotype group as BMD. Furthermore, according to what is known about linkage of the VDR gene polymorphisms, the Fok I RFLP is not linked to Bsm I, Apa I and Taq I polymorphisms, and can therefore not "explain" the association results of the Bsm I, Apa I and Taq I RFLPs. Studies on the Fok I RFLP should be analysed as a different genotype group in the meta-analysis. Thakkinstian et al.[24] observed that the BB genotype had lower spine BMD than Bb + bb genotypes but only in studies of postmenopausal women with a big range of B-allele frequency (29 – 53%, possibly due to ethnic admixture) and heterogeneity was generally present in their analyses. All together, this indicates that there is also no evidence of a genetic effect of Bsm I and Taq I RFLPs on BMD, and the genetic effect of Fok I and Apa I RFLPs on BMD is unclear.

A functional polymorphism, Cdx-2, in the promoter region of the VDR gene was implicated to influence vitamin D regulation of calcium absorption in the intestine[26]. Consequently, the A-allele of the Cdx-2 binding site was found to be associated with increased BMD in Japanese postmenopausal women[27]. However, we only see a borderline significant BMD difference by Cdx-2 (or haplotype 1 of VDR LD block 2, containing the G-allele of the Cdx-2 SNP) genotype in the Rotterdam Study population, and no association was found in the LASA population. One interpretation might be a difference in dietary calcium intake between Japanese and Dutch populations. The mean of dietary calcium intake in the Rotterdam Caucasian population is 1,117g/day, but is only about 600g/day for the Japanese Asian population[28]. We also analysed in the Dutch Caucasian the relationship between VDR Cdx-2 genotype and BMD in subjects with calcium intake < 600 or < 800 mg/day, but we did not have sufficient statistical power to detect the association, since the A-allele frequency in Caucasian population (19%) is lower than in Asian population (44%).

9.4.2 VDR polymorphisms predict multiple bone phenotypes in the Rotterdam Study: The Rotterdam study has collected multiple phenotypes and in this respect is likely to help us to understand the contribution of genetic factors to components of disease by looking for biological consistency in associations. In association studies of the thesis we demonstrated a relationship between VDR polymorphisms with several related bone phenotypes, including fracture risk, body height, bone size.

We have described associations between VDR polymorphisms and osteoporosis

in our previous studies[29-31]. Recently we extended the study size and investigated most potentially functional polymorphisms across the VDR gene. Our current findings not only confirmed the previous associations, but also presented an overview of genetic effects of VDR polymorphisms on osteoporosis. We first found that three haplotype alleles in the promoter and 3'-UTR were associated with increased fracture risk, while an additive effect of these risk alleles across the VDR gene was observed. The haplotype alleles in the 3'-UTR show the strongest effect on the fracture risk (chapter 5.1).

We secondly demonstrated the associations between two fracture-risk haplotypes and their combined genotype with decreased body height in the same population and another independent elderly population (the LASA study). These associations were further confirmed by a meta-analysis of VDR genotype (defined by the Bsm I RFLP) and body height according to published data. The same haplotype alleles were observed to be associated with decreased bone size and strength (chapter 6). Another fracture-risk haplotype allele found in our population, block 2-hap1, was not found to be related to height difference in our elder and high calcium intake (1,117 mg/day) populations: the Rotterdam Study and the LASA. However, this haplotype was found to be associated with decreased body height in young and low calcium-intake (< 865 mg/day) French adolescent girls[32]. This haplotype also associated with decreased serum 25(OH)D_3 and decreased serum IGF-1 level in the same French population. Since Cdx-2 polymorphism is one of the htSNPs which are related to calcium absorption through the intestine, the interaction between block 2-hap 1 with calcium-intake and/or serum vitamin D level on bone phenotype is interesting to be further investigated in other young and old populations.

We then described an interaction effect between VDR 3'-UTR haplotype and DBP haplotype on fracture risk (chapter 8). All those associations are independent of BMD. With the evidence of functionality study of VDR haplotypes (chapter 7), the underlying mechanism of the association studies seems to be that VDR fracture-risk allele carriers have a lower sensitivity to the vitamin D ligand because of lower expression of VDR. Perhaps this is due to lower osteoblast activity compared to non risk allele carriers. Thus, the bone gain decreases because of decreasing bone formation, consequently the bone loss is relatively increased, which is reflected in decrease of the bone size and strength, and bone shape change, finally height loss increases and fracture risk increases. On the other hand, DBP-macrophage activating factor (DBP-MAF) can mediate bone resorption by activating osteoclasts[33], and the fracture-risk haplotype allele of the DBP gene (according to Lauridsen et al.[34]) is associated with increased plasma DBP level[35]. Thus, the fracture-risk allele of the DBP carriers is likely to have higher osteoclast activity, and thus higher bone resorption. When carrying both risk alleles of VDR and DBP genes, the subject has low bone formation because of low osteoblast activity and high bone resorption because of high osteoclast activity. Bone strength in these subjects could therefore decrease and fracture risk could increase. However, those hypotheses need to be

confirmed by further studies.

9.4.3 Reason of conflicting association results: Many studies reported conflicting associations between VDR polymorphisms and fracture risk. Our meta-analysis does not show a genetic effect of Bsm I RFLP on fracture risk, and this seems to contradict our association results in the Rotterdam Study. Some factors could underlie the differences in association results. Our association study is based on a population-based and large-scale study design. All subjects of the study population share similar environmental factors and have the same ethnical background, and haplotypes of each of the blocks were used for the association analysis. Many previous conflicting association studies on VDR and fracture, analysed usually (very) small study populations and used the Bsm I, Apa I or Taq I polymorphisms in the VDR LD block 5, but mostly analysed separately. We therefore suppose that this controversy can partly be explained by a lack of statistical power as a result of small sample size, and failure to use haplotypes. The meta-analysis of the association between Bsm I and fracture risk showed a significant heterogeneity between different individual studies ($I^2 < 0.001$). The direction and strength of association results from different studies is shown different in figure 1 of chapter 5.2. The definition of phenotype, such as fracture, crucially affects the prospects of an association analysis, and it is difficult to standardize the definition of fracture among the individual published studies for the meta-analysis. The different type of study design is also another important source of heterogeneity (such as equal-size case-control, non equal-size case-control, hospital source case-control, population-based and non population-based cohorts). Such settings will determine study size and different numbers of fracture, such as incidence and prevalence of fracture in the study population. In addition, population stratification, such as mixed ethnic groups with different allele frequencies, population-specific differences of some environmental factors, such as (dietary) calcium intake, (dietary) vitamin D intake, sunlight-exposure, and other characteristics of the study population pertinent to bone metabolism and fracture risk could result in heterogeneity of associations observed across different study populations. Therefore, non-standardized phenotype definition, poorly characterized and non-normal distributed study populations, unregulated environmental exposures can essentially influence (and/or cover up) to identify modest genetic effects of sequence variations, because those factors bring genetic and environmental noise in the meta-analysis. Especially, for supposedly small effects such as for the VDR (see below), this can mask their detection in meta-analysis.

9.5 Magnitude of genetic effect of gene variation

The genetic effect of VDR polymorphisms on bone phenotypes is generally modest in our study. In the association studies, we demonstrated that two haplotype

alleles (in block 3 and 5) increased fracture risk by 15 – 74 %, and decreased 0.09 – 0.15 SD. of body height. Our results reflect the common disease/common variant (CDCV) hypothesis[36]. In a public health context, few high-risk alleles may have a large population attributable ratio (PAR, >50%), which is thought of as the fraction of the disease that would be eliminated if the risk factor were removed. Common modest-risk alleles may account for a greater PAR in common disease than do rare high-risk alleles[16]. The haplotype 3 homozygous in VDR LD block 3 increases fracture risk by 60%, but contributes only 1% PAR because of 2% genotype frequency in the population. With 68% genotype frequency, haplotype 1 carriers in block 5 increases fracture by 20%, being associated with 12% PAR. In a previous study we showed the homozygous "s-allele (or T-allele)" of the COL1A1 Sp1 polymorphism to be associated with 50% increased fracture risk and 0.23 and 0.35 SD. decreased femoral neck and lumbar spine BMD[37]. We now demonstrated it has 1% PAR in our complete study population, and genotype frequency of TT homozygotes is 3%. These small risks are in line with small effects detected in functional studies. Our functionality study showed that the difference of VDR mRNA level by 3'-UTR haplotype was 15%, and Mann et al.[18] presented the "Ss" (or GT) genotype of COL1A1 Sp1 was found to have a 21% increased COL1A1/COL1A2 protein ratio compared to "SS" (or GG) as measured in osteoblasts. Another functional study[38] of a 3'-UTR SNP [+1073C/T in the oxidised LDL receptor (OLR1) gene] that associates with increased risk for Alzheimer's disease (AD) demonstrated that C-allele carriers had 41% decreased OLR1 mRNA level compared to "TT" homozygotes. Therefore, common variants contribute modest genetic effect to common diseases, but they are the most interesting variations for genetic studies in the terms of public health.

9.6 Suggestions of future studies

9.6.1 Identification of functional variations of a candidate gene: Our studies in the thesis provide a study strategy, which combines bioinformatics, molecular genetics, genetic association and functionality aspects together, to identify functional polymorphisms for candidate gene studies. A procedure of genetic study for a candidate gene is described in Figure 1. The study begins with bioinformatics aspects: a. view previous studies (association, functional experiments *in vivo* or *in vitro*) to make a hypothesis for the current study; b. find genomic organization of the gene, collect all information of sequence variations of the gene (from genetic databases, e.g., NCBI, and publications); c. download genotype data from Perlegen, HapMap and SNPbrowser databases, compare LD maps of those resources (as well as published data) with the same haplotype definition criteria to generate a consensus LD map. If few common SNPs are present in all resources, and LD structures are much different among the resources, an assessment of an accurate LD pattern in a small set of control individuals (called the LD sample) is suggested

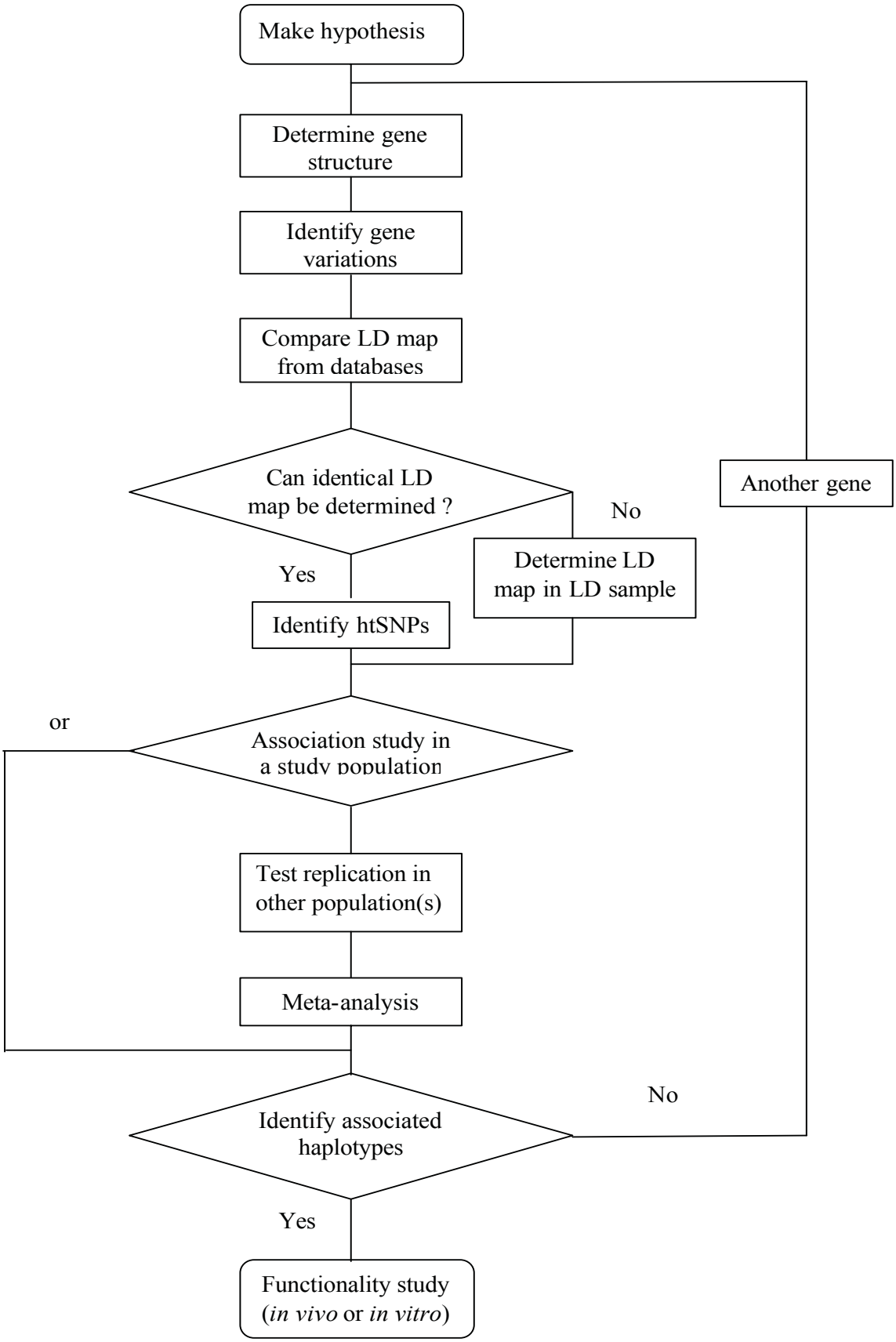


Figure 1. A procedure of genetic study for a candidate gene.

to be carried out. The consensus LD pattern is then intended to determine commonly used and/or potentially functional variations as tagging SNPs across the gene for the association study. Genotyping all tagging SNPs in the association study population, generates haplotypes to test the hypothesis of the association between haplotypes of the gene and phenotype. Analysing an association in one study population is followed by replication studies in other populations, and retrospective or prospective meta-analysis to systematically summarize the association studies. Finally, functionality studies are carried out if the associated haplotypes are identified.

9.6.2 WGA study: With the developing support of the SNP project and the HapMap project, high-throughput genotyping technology and statistics, WGA study becomes a critical tool for the identification of susceptibility genes for complex traits[15]. The study design of a WGA might briefly consist of following portions. The major study population of the WGA studies is the Rotterdam Study population, the replication of positive results will be tested in several independent populations, such as LASA, EPOS and so on. The phenotypes of the studies should be precisely defined, it will have greater chances of success and will increase the likelihood that the results will be replicated[39]. The tagging SNPs across the whole genome should be well selected according to systematical analysis of resources information on SNPs and LD maps from Perlegen, HapMap, SNPbrowser, other publications and LD-study samples. The selection criteria should consider allele frequencies of SNPs (following CDCV hypothesis), physical distance between SNPs, SNP distribution in the genome, potential function of SNPs and linkage status of SNPs. Common haplotypes in LD blocks are suggested to be used in the study to restrict multiple testing. The threshold of common haplotype frequency is determined by power calculation based on different type of study populations.

9.6.3 Pharmacogenetics: Complex genetic traits involve gene-environment and gene-gene interactions. Pharmacogenetics studies the role of DNA sequence variation in individual variation in drug response, and represents the field of gene-environment interactions. Functional variations in gene products that play essential roles in determining variability in drug responses are analysed for the correlations between the genetic variant and drug associated toxicity or therapeutic effect[40,41]. Difference in response to vitamin D supplementation for BMD change by VDR genotype was reported in a Caucasian population[42]. In this study, a placebo-controlled clinical trial was carried out in elderly women (age > 70 years), and the effect of vitamin D supplementation on BMD and fracture was examined after one and two years follow-up. The increase of BMD in the vitamin D group relative to the placebo group was significantly higher in BB and Bb genotypes of the VDR gene as compared to the bb genotype. This was the first study to demonstrate that the effect of vitamin D supplementation on BMD depends on the VDR genotype. More

and more polymorphisms and mutations are identified as prognostic markers, and first phase of "personalized medicine" is already under way with the genotyping required to prescribe drug for patients, although there is still a long way to go[43]. Pharmacogenetic studies will be an attractive field for scientists.

WGA studies followed by candidate gene studies provide the opportunity to investigate the pathway and network of gene-gene and gene-environmental interactions. Overall disease risk can be modeled as the product of risks at many independent genetic risk loci and environmental factors. It is interesting and a great challenge to determine the genetic contribution of variations of candidate genes and environmental contributions to a certain clinical endpoint, such as osteoporosis, with statistic models. In osteoporosis, more than 20 candidate genes have been analysed in more or less detail so far to determine their contribution. Studies on the Rotterdam study[30,37,44-46] and in particular the GENOMOS consortium[19] provide important research environments to dissect the genetics of osteoporosis and determine their contribution to the overall risk for this disabling disease.

References

1. Baker, A.R. et al. Cloning and expression of full-length cDNA encoding human vitamin D receptor. *Proc Natl Acad Sci U S A* **85**, 3294-8 (1988).
2. Miyamoto, K. et al. Structural organization of the human vitamin D receptor chromosomal gene and its promoter. *Mol Endocrinol* **11**, 1165-79 (1997).
3. Crofts, L.A., Hancock, M.S., Morrison, N.A. & Eisman, J.A. Multiple promoters direct the tissue-specific expression of novel N-terminal variant human vitamin D receptor gene transcripts. *Proc Natl Acad Sci U S A* **95**, 10529-34 (1998).
4. VDR-NCBI. http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=nucleotide&val=32891816.
5. Nejentsev, S. et al. Comparative high-resolution analysis of linkage disequilibrium and tag single nucleotide polymorphisms between populations in the vitamin D receptor gene. *Hum Mol Genet* **13**, 1633-9 (2004).
6. Celera. http://www.celera.com/celera.
7. dbSNP. http://www.ncbi.nlm.nih.gov/SNP.
8. Perlegen. http://genome.perlegen.com/browser/index.html.
9. HapMap. http://www.hapmap.org/cgi-perl/gbrowse/gbrowse/hapmap.
10. SNPbrowser. http://events-na.appliedbiosystems.com/mk/get/snpb_landing?isource=fr_E_RD_www_allsnps_com_snpbrowser.
11. Phase. http://archimedes.well.ox.ac.uk/pise/PHASE.html.
12. Halperin, E. & Eskin, E. Haplotype reconstruction from genotype data using Imperfect Phylogeny. *Bioinformatics* **20**, 1842-9 (2004).
13. Hinds, D.A. et al. Whole-genome patterns of common DNA variation in three human populations. *Science* **307**, 1072-9 (2005).
14. Stephens, M., Smith, N.J. & Donnelly, P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* **68**, 978-89 (2001).
15. Risch, N. & Merikangas, K. The future of genetic studies of complex human diseases. *Science* **273**, 1516-7 (1996).
16. Carlson, C.S., Eberle, M.A., Kruglyak, L. & Nickerson, D.A. Mapping complex disease loci in whole-genome association studies. *Nature* **429**, 446-52 (2004).
17. Gibson, N., Jawaid, A. & March, R. Novel technology and the development of pharmacogenetics within the pharmaceutical industry. *Pharmacogenomics* **6**, 339-56 (2005).

18. Mann, V. et al. A COL1A1 Sp1 binding site polymorphism predisposes to osteoporotic fracture by affecting bone density and quality. *J Clin Invest* **107**, 899-907 (2001).
19. Ioannidis, J.P. et al. Differential genetic effects of ESR1 gene polymorphisms on osteoporosis outcomes. *JAMA* **292**, 2105-14 (2004).
20. Gabriel, S.B. et al. The structure of haplotype blocks in the human genome. *Science* **296**, 2225-9 (2002).
21. Nejentsev, S. et al. Analysis of the vitamin D receptor gene sequence variants in type 1 diabetes. *Diabetes* **53**, 2709-12 (2004).
22. Cooper, G.S. & Umbach, D.M. Are vitamin D receptor polymorphisms associated with bone mineral density? A meta-analysis. *J Bone Miner Res* **11**, 1841-9 (1996).
23. Gong, G. et al. The association of bone mineral density with vitamin D receptor gene polymorphisms. *Osteoporos Int* **9**, 55-64 (1999).
24. Thakkinstian, A., D'Este, C., Eisman, J., Nguyen, T. & Attia, J. Meta-analysis of molecular association studies: vitamin D receptor gene polymorphisms and BMD as a case study. *J Bone Miner Res* **19**, 419-28 (2004).
25. Morrison, N.A. et al. Prediction of bone density from vitamin D receptor alleles. *Nature* **367**, 284-7 (1994).
26. Yamamoto, H. et al. The caudal-related homeodomain protein Cdx-2 regulates vitamin D receptor gene expression in the small intestine. *J Bone Miner Res* **14**, 240-7 (1999).
27. Arai, H. et al. The polymorphism in the caudal-related homeodomain protein Cdx-2 binding element in the human vitamin D receptor gene. *J Bone Miner Res* **16**, 1256-64 (2001).
28. Tsuchida, K., Mizushima, S., Takahashi, H., Misugi, N. & Soda, K. [Relationship between dietary calcium and bone mineral density before menopause]. *Nippon Koshu Eisei Zasshi* **45**, 121-8 (1998).
29. Uitterlinden, A.G. et al. A large-scale population-based study of the association of vitamin D receptor gene polymorphisms with bone mineral density. *J Bone Miner Res* **11**, 1241-8 (1996).
30. Uitterlinden, A.G. et al. Interaction between the vitamin D receptor gene and collagen type Ialpha1 gene in susceptibility for fracture. *J Bone Miner Res* **16**, 379-85 (2001).
31. Fang, Y. et al. Cdx-2 polymorphism in the promoter region of the human vitamin D receptor gene determines susceptibility to fracture in the elderly. *J Bone Miner Res* **18**, 1632-41 (2003).
32. d'Alesio, A. et al. Two new SNPs in the human vitamin D receptor promoter switch transcription factor binding, change promoter activity and are associated with growth in femal adolescents. *J Bone Miner Res* **20**, S184 (2005).
33. Haddad, J.G. Plasma vitamin D-binding protein (Gc-globulin): multiple tasks. *J Steroid Biochem Mol Biol* **53**, 579-82 (1995).
34. Lauridsen, A.L. et al. Female premenopausal fracture risk is associated with gc phenotype. *J Bone Miner Res* **19**, 875-81 (2004).
35. Lauridsen, A.L. et al. Plasma concentrations of 25-Hydroxy-Vitamin D and 1,25-Dihydroxy-Vitamin D are Related to the Phenotype of Gc (Vitamin D-Binding Protein): A Cross-sectional Study on 595 Early Postmenopausal Women. *Calcif Tissue Int* (2005).
36. Collins, F.S., Guyer, M.S. & Charkravarti, A. Variations on a theme: cataloging human DNA sequence variation. *Science* **278**, 1580-1 (1997).
37. Uitterlinden, A.G. et al. Relation of alleles of the collagen type Ialpha1 gene to bone density and the risk of osteoporotic fractures in postmenopausal women. *N Engl J Med* **338**, 1016-21 (1998).
38. Lambert, J.C. et al. Association of 3'-UTR polymorphisms of the oxidised LDL receptor 1 (OLR1) gene with Alzheimer's disease. *J Med Genet* **40**, 424-30 (2003).
39. Ehm, M.G., Nelson, M.R. & Spurr, N.K. Guidelines for conducting and reporting whole genome/large-scale association studies. *Hum Mol Genet* **14**, 2485-8 (2005).
40. Nebert, D.W. Pharmacogenetics and pharmacogenomics: why is this relevant to the clinical geneticist? *Clin Genet* **56**, 247-58 (1999).
41. Evans, W.E. & McLeod, H.L. Pharmacogenomics--drug disposition, drug targets, and

side effects. *N Engl J Med* **348**, 538-49 (2003).
42. Graafmans, W.C. et al. The effect of vitamin D supplementation on the bone mineral density of the femoral neck is associated with vitamin D receptor genotype. *J Bone Miner Res* **12**, 1241-5 (1997).
43. Lewis, R. An individual approach. *Nature* **436**, 746-7 (2005).
44. van Meurs, J.B. et al. Association of 5' estrogen receptor alpha gene polymorphisms with bone mineral density, vertebral bone area and fracture risk. *Hum Mol Genet* **12**, 1745-54 (2003).
45. van Meurs, J.B. et al. Homocysteine levels and the risk of osteoporotic fracture. *N Engl J Med* **350**, 2033-41 (2004).
46. Fang, Y. et al. Promoter and 3'-Untranslated-Region Haplotypes in the Vitamin D Receptor Gene Predispose to Osteoporotic Fracture: The Rotterdam Study. *Am J Hum Genet* **77**, 807-23 (2005).

Chapter 10

Summary

Chapter 10.1

Summary

Genetic variations of the vitamin D receptor (VDR) gene have been found associated with several complex diseases, including osteoporosis, but with controversial results. This can be due to small sample size of the association studies and non-functional sequence variations analysed. We performed the current study to find and characterize genetic variations across the whole VDR gene, investigate functionality of variations and determine association with bone phenotypes. This research involved several approaches including bioinformatics, molecular genetics, genetic epidemiology and cell biology.

In **Chapter 2,** we integrated information from different resources and updated the genomic organization of the human VDR gene. Homology analysis between human and mouse VDR genomic DNA sequences was used to determine the potentially functional regions of the human VDR gene for the re-sequencing study. We then sequenced those potentially functional regions (22 kb) across the VDR gene to find sequence variations (polymorphisms) in 15 Caucasian individuals. We observed 62 polymorphisms, and compared the findings with other genetic databases using bioinformatic methods and determined a more accurate genomic structure of the VDR gene.

In **Chapter 3,** we examined the linkage disequilibrium (LD) and haplotype structure across the VDR gene based on the sequencing results, defined the haplotype maps and haplotype tagging SNPs for association studies in Caucasians, Asians and Africans. Four to eight haplotype blocks and nine to 28 tagging SNPs were determined in those three ethnic groups which can be used in association studies.

In **Chapter 4**, we investigated the relationship between a functional SNP in a VDR promoter region and fracture risk. The Cdx-2 SNP is a G to A substitution and was first reported at a position in front of exon 1a by a Japanese group, but they only tested the relationship with BMD in a small Japanese female population. During our re-sequencing analysis, we also observed this SNP, and identified the location of the SNP is at 1012 basepair in front of exon 1e, and we resolved that the distance between exon 1e and 1a is 2kb. We then developed a simple and quick genotype method to analyse DNA samples from 8 different ethnic groups and a large sample of the Rotterdam Study. We found a correlation between the Cdx-2 genotype and prevalence of fracture in different ethnic populations, and identified a BMD-independent association between the G-allele and fracture risk in the Rotterdam Study population.

In **Chapter 5**, we focus on the analyses of VDR haplotypes and one 3'-end polymorphism in relation to fracture risk. An association study in the complete Rotterdam Study population and a meta-analysis of published data were carried

out.

In Chapter 5.1, we demonstrated a relationship between haplotypes defined by tagging SNPs across the VDR gene (in chapter 3), and fracture risk in > 6,500 subjects from the Rotterdam Study population. We observed an association of fracture risk with three haplotype alleles in the promoter region and the 3'-end of the VDR gene, and an additive effect of those haplotype alleles on fracture risk. These associations were observed for different kinds of fractures, such as wrist, hip, and vertebral fractures, and independent of age and gender. We also analyzed a possible relationship of VDR haplotypes with BMD and bone loss. One fracture risk haplotype allele in the promoter region was found to be associated with decreased BMD, and a risk allele at the 3'-end of the VDR gene was observed to be associated to bone loss. However, those effects were not explaining the relationship between VDR haplotypes and fracture risk.

In Chapter 5.2, we systematically reviewed all published association studies of VDR Bsm I RFLP and fracture risk and performed a meta-analysis. The Bsm I RFLP is the most widely studied polymorphism until now and so the meta-analsis had to be limited to this SNP. No evidence of a relationship between the VDR BsmI RFLP and fracture risk was observed. However, Bsm I RFLP is not a good marker of haplotype diversity in the 3' area of the VDR gene. Therefore, the meta-analysis result might not completely represent the genetic effect of other polymorphisms across the VDR gene on fracture risk.

In **Chapter 6,** we first demonstrated a relationship between VDR fracture-risk haplotype alleles and combined genotypes with body height in the Rotterdam Study and the LASA population. This association was found to be independent of age, gender and vertebral fracture. In addition, the combined promoter and 3'-UTR VDR genotype was found to be associated with decreased lumbar spine area and narrow neck width of the femoral neck in the Rotterdam population. We then carried out a meta-analysis according to published data, and found that the VDR Bsm I RFLP was associated with differences in body height in adult elderly populations, which is in line with results from our association analysis in the Rotterdam Study.

In **Chapter 7**, we performed functionality studies by EMSA and transactivation experiments for SNPs in the 1e-1a promoter region, and measured VDR mRNA level and stability at the two major 3'-UTR haplotypes in different cell lines. We found that fracture risk alleles in the promoter region that associate with increased fracture risk, had low binding affinity to transcriptional factors and resulted in lower expression of the VDR gene. The fracture risk haplotype allele of the 3'-UTR resulted in decreased VDR mRNA level, and had lower mRNA stability. The functionality results are in line with our association results on fracture and body height.

In **Chapter 8**, we described an LD map across the DBP gene by integrating DBP LD data from different resources. In a subset of the Rotterdam Study population we observed a significant correlation of common haplotypes of the DBP and serum vitamin D level. In the complete Rotterdam Study population, we also observed interactions between DBP haplotype with dietary calcium intake and VDR haplotypes in relation to fracture risk.

In **Chapter 9**, basing on the results of the thesis, we discuss the usage of bioinformatics in genetic studies, the comparison of linkage studies and candidate gene association studies, whole genome association studies, genetic meta-analysis and the use of single polymorphisms or haplotypes for association studies. We also discuss an overview of association of VDR polymorphisms and bone phenotypes, possible reasons of conflicting association results, the magnitude of genetic effects of common gene variation. Finally we provide suggestions for future studies following the conclusions of the thesis, including identification f functional variations in a candidate gene and applications of genetic association study in Pharmacogenetics.

Chapter 10.2

Samenvatting

Genetische variaties in het vitamine D receptor gen zijn in het verleden geassocieerd gevonden met verschillende complexe ziekten, waar osteoporose (botontkalking) er één van is. Deze associaties worden echter niet consistent gevonden en zijn daarom controversieel. Deze inconsistente bevindingen kunnen zowel door te kleine studiegrootte worden veroorzaakt als door de niet-functionele DNA-variaties die meestal in het VDR gen worden onderzocht.

Het hier beschreven onderzoek had tot doel om alle relevante variaties in het VDR gen te vinden en te karakteriseren, en de functionele gevolgen van deze variaties te onderzoeken. Tevens was het doel om de relatie te bestuderen tussen deze variaties en osteoporose. Tijdens het beschreven onderzoek is er gebruik gemaakt van verschillende soorten benaderingen, waaronder bio-informatica, moleculaire genetica, genetische epidemiologie en celbiologie.

In **hoofdstuk 2** hebben we de genomische organisatie van het VDR gen opgehelderd en verbeterd ten opzichte van eerder gepubliceerde gegevens, waarbij gebruik is gemaakt van verschillende informatiebronnen. Om de potentieel functionele gedeeltes van het VDR-gen te bepalen, werd er een homologie analyse uitgevoerd, waarbij het muizen VDR-gen werd vergeleken met het mensen VDR-gen. De gedeelten van het gen die grote gelijkenis vertonen (hoge homologie) werden vervolgens in detail bestudeerd in een aantal verschillende mensen. De basenpaarvolgorde (sequentie) van het VDR gen werd bepaald in 15 verschillende mensen, zodat alle veelvoorkomende DNA variaties (polymorfismen) in het VDR gen gevonden werden; dit waren er 62 in totaal.

In **hoofdstuk 3** is de correlatie tussen de verschillende variaties in het VDR-gen onderzocht. In het gen werden blokken van polymorfismen met hoge onderlinge correlatie gevonden, deze blokken worden ook wel "linkage disequilibrium" (LD) blokken genoemd. Binnen een dergelijk blok voorspellen de variaties elkaar goed, zodat niet alle polymorfismen bepaald dienen te worden om toch alle genetische variatie te kunnen voorspellen. Binnen elk LD-blok werden de variaties bepaald die nodig waren om alle andere variaties te kunnen voorspellen, de zogenaamde "tagging" polymorfismen. De correlatie van de DNA polymorfismen in het VDR-gen werd bestudeerd in 3 verschillende rassen: het Kaukasische, Aziatische, en Afrikaanse ras. Daarbij zagen we dat in de individuen van Afrikaanse afkomst meer en beduidend kleinere blokken van SNPs met hoge correlatie gevonden werden.

In **hoofdstuk 4** is de relatie tussen een functionele variant in het VDR gen en het risico op het krijgen van een botbreuk onderzocht. De variant die werd onderzocht zit in het gedeelte dat het VDR gen bestuurt, de zogenaamde promoter. De promoter bepaald hoe hard het gen "aan" staat. Variaties in dit gedeelte kunnen er dus voor zorgen dat er meer of minder VDR mRNA gemaakt wordt. De variatie die in dit hoofdstuk is onderzocht, het Cdx2-polymorfisme, werd voor het eerst

onderzocht door een Japanse groep in een kleine populatie van Japanse vrouwen. Tijdens onze eerdere sequentie-bepaling (hoofdstuk 2), zagen we dat dit polymorfisme ook in onze populatie voorkwam en konden we de exacte positie van deze variant bepalen. Vervolgens vonden we, in onze grote Rotterdam-populatie van ongeveer 7000 mensen, een correlatie tussen dit polymorfisme en het risico op het krijgen van een botbreuk. Deze correlatie was onafhankelijk van de botdichtheid.

In **hoofdstuk 5** wordt de relatie beschreven tussen de polymorfismen in het VDR gen en het risico op het krijgen van een botbreuk. In hoofdstuk 5.1 laten we zien dat er een relatie bestaat tussen bepaalde variaties in het VDR-gen en botbreukrisico. De beschreven studie werd uitgevoerd in meer dan 6500 individuen van de Rotterdam studie. Bepaalde polymorfismen in de promoter en in het uiteinde (het zogenaamde 3'einde) van het gen voorspelde samen een verhoogd risico voor het krijgen van een botbreuk. Deze relatie bleek heel consistent aanwezig te zijn en was onafhankelijk van geslacht, leeftijd en type botbreuk.

Er werd ook een relatie gevonden tussen de promoter variant en botdichtheid, daarnaast vonden we een relatie tussen de variatie in het uiteinde van het gen en botverlies. Beide relatie konden echter niet het verhoogde risico op botbreuken verklaren dat eerder was gevonden.

In hoofdstuk 5.2 is een analyse beschreven van alle eerder gepubliceerde artikelen (een meta-analyse) over een VDR polymorfisme en botbreuk risico: het BsmI polymorfisme. Deze DNA variant is gesitueerd in het 3'-uiteinde van het gen en is al door vele groepen onderzocht. Alle gepubliceerde artikelen met als onderwerp de relatie tussen dit BsmI polymorfisme en het risico van een botbreuk werden gebruikt voor de meta-analyse. Er werd geen relatie tussen dit polymorfisme en botbreuken gevonden. Onze eerder beschreven studies (hoofdstuk 2 en 3) laten echter zien dat het BsmI polymorfisme de totale genetische variabiliteit van het 3'uiteinde van het VDR-gen niet goed voorspelt. Dit zou kunnen verklaren waarom er geen relatie gevonden werd tussen het BsmI polymorfisme en botbreuken in alle verschillende studies tot nu toe.

In **hoofdstuk 6** laten we zien dat de variaties die eerder gerelateerd waren gevonden met botbreuken (hoofdstuk 5) ook samen hangen met verschillen in lengte van de individuen. Dit werd gevonden in zowel de Rotterdamse ERGO populatie als de Amsterdamse LASA Studie.

Vervolgens is er een meta-analyse uitgevoerd van alle gepubliceerde data voor het BsmI polymorfisme in relatie tot lengte. Deze meta-analyse had als conclusie dat het BsmI polymorfisme gerelateerd is met verschillen in lengte in volwassen populaties. Dit ondersteunt de eerder gevonden relatie in de Rotterdam studie en de LASA Studie.

In **hoofdstuk 7** is de mogelijke functionaliteit van een aantal polymorfismen in het promoter gedeelte van het VDR-gen onderzocht. Dit werd gedaan door het onderzoeken van de capaciteit van (verschillende varianten van) de promoter om een gen tot expressie te brengen ("aan" te zetten) . We vonden dat de variaties die eerder (hoofdstuk 5) gerelateerd waren gevonden met fractuurrisico, ook minder capaciteit hadden om een gen tot expressie te brengen. Waarschijnlijk leiden deze varianten dus tot een lagere expressie van het VDR-gen.

In dit hoofdstuk werden ook de polymorfismen in het 3'uiteinde van het gen onderzocht. Deze polymorfismen zijn gelegen in een gedeelte van het gen dat wordt gedacht de stabiliteit van het mRNA te beïnvloeden. Het bleek dat deze polymorfismen, welke al eerder bleken te relateren met botbreukrisico (hoofdstuk 5), ook samen hingen met een verlaagde stabiliteit van het mRNA. Dit leidt uiteindelijk tot minder mRNA van het VDR gen in de cel.

Beide resultaten van de functionaliteit experimenten zijn in lijn met de eerder gevonden relaties van de DNA varianten met verhoogd botbreuk risico.

In **hoofdstuk 8** wordt data van verschillende databanken gebruikt om de LD blokken van het vitamine D bindings eiwit (DBP) gen te bepalen. Het DBP is erg belangrijk voor het Vitamin D endocrien systeem, het zorgt voor transport van het Vitamine D naar de doelcellen toe. Daarnaast kan het DBP ook een rechtstreeks effect uitoefenen op de botafbrekende cellen, de osteoclasten. In hoofdstuk 8 zijn er een tweetal polymorfismen in het DBP-gen onderzocht in de Rotterdam studie. Er werd gevonden dat deze polymorfismen samenhangen met de vitamine D spiegel in het bloed. Bovendien zagen we een interactie van deze polymorfismen met calcium-inname en met de eerder beschreven VDR-polymorfismen in relatie tot botbreukrisico.

Gebaseerd op de beschreven resultaten in dit proefschrift, worden in **hoofdstuk 9** een 5-tal onderwerpen bediscussieerd: 1. het gebruik van bioinformatica in genetische studies, 2. de voor- en nadelen van linkage studies en associatie studies, 3. associatie studies van het hele genoom, 4. meta-analyse van genetische studies, 5. het gebruik van aparte polymorfismen of combinaties van polymorfismen (haplotypes). Daarnaast wordt er in dit hoofdstuk ook een overzicht gegeven van alle studies die zijn verschenen op het gebied van VDR polymorfismen in relatie tot osteoporose en worden mogelijke verklaringen gegeven voor de tegenstrijdig resultaten die in eerdere studies zijn gevonden. Geconcludeerd kan worden dat VDR polymorfismen een effect op botbreuk risico kunnen hebben in sommige populaties, maar niet noodzakelijkerwijs in alle populaties. Als laatste worden enkele suggesties voor vervolgstudies gedaan.

Chapter 10.3

总　结

通过以前的研究人们发现维生素D受体（ＶＤＲ）基因的变异与某些具有遗传背景的综合性疾病（其中包括骨质疏松）的发病相关，但研究结果并不一致。其原因可能是由于过小的样本研究数量不能提供可靠的结果，以及非功能性遗传变异被用于研究。通过本研究，我们着眼于发现并描述整个ＶＤＲ基因的遗传变异，通过研究基因变异对骨质疏松和人体身高等骨代谢特征的影响，确定ＶＤＲ基因的功能性变异。为将来研究ＶＤＲ基因和其它疾病（如心血管疾病、肿瘤、糖尿病、脱发和一些免疫性疾病）的遗传易感性提供可靠的研究线索和研究模式，并可为临床诊断以上疾病提供科学依据，该结果还可用于药物遗传学领域，着眼于今后开发具有高效、特异、付作用小的治疗药物，实现“个体治疗”，即根据个体的遗传特征选择不同种类、不同计量的药物进行治疗。本研究涉及到多方面的研究方法：生物信息学、分子遗传学、遗传流行病学和细胞生物学。

第一章为论文集的前言部分，其中主要描述了相关背景知识、当前对骨代谢和ＶＤＲ基因变异研究的现状和主要存在的问题，其中包括：维生素D内分泌系统的功能、当前对ＶＤＲ基因结构的了解、ＶＤＲ基因多肽性变异、ＶＤＲ基因变异的联系不平衡和单倍型、骨质疏松和ＶＤＲ基因多肽性变异、ＶＤＲ基因多肽性变异的功能性。同时本章也引出了本课题研究的目的、方法和该论文集各章节的内容。

第二章，利用生物信息学手段，我们综合多方面的信息资源，更新了人类ＶＤＲ基因的物理结构。通过比较人类和鼠类ＶＤＲ基因的基因组序列，我们确定了具有潜在功能的区域作为下一步脱氧核苷酸（ＤＮＡ）序列区域。我们接下来用了两年半的时间对那些潜在功能区（２，２０００碱基对）进行了ＤＮＡ测序分析以发现序列变异（多肽性），在３０条白种人的染色体中我们检测了３９个大小在２９７到2，２０１碱基对的ＤＮＡ序列，发现了６２个多肽性变异，填补了５００碱基对的序列空白区。通过运用生物信息方法将我们发现的序列变异和其它遗传学数据库进行了比较，确定了更为精确的ＶＤＲ基因的基因组结构，将ＶＤＲ基因的长度从以前描述的8，５００碱基对更正为１０，５００碱基对。

第三章，根据序列分析的结果，我们创立了一套新颖方法用于研究整个ＶＤＲ基因的连锁不均衡（ＬＤ）和单倍型结构，确认了白种人、亚洲人和非洲人所特异的单倍型图谱和单倍型标记单核苷酸多肽性（ＳＮＰ），在以上三个人种中我们确定了四到八个单倍型区和九到２８个标记ＳＮＰ，这些结果将用于流行病学的相关性研究。

第四章，我们对一个ＳＮＰ（它位于ＶＤＲ基因启动区，并具有功能性）和骨折的危险关系进行了调查。这个被称作Cdx-2的ＳＮＰ是一个G对A的变异，一个日本研究小组曾经报道它位于外显子1a的前方，但他们只在小范围的日本女性人群中调查了该ＳＮＰ与骨密度（ＢＭＤ）的关系。通过对ＤＮＡ的测序分析，我们也发现了该ＳＮＰ，并确认它定位于外显子1e前方１，０１２碱基对处，而且我们还确定外显子1a和1e之间的距离为２，０００碱基对。我们进而研发出一种简单快速

的基因分型方法，对八个不同人种和“鹿特丹研究”人群的ＤＮＡ样品进行了基因分型。我们发现Cdx-2的基因型别和不同人群的股骨骨折发病具有相关性，同时在“鹿特丹研究”的人群中我们也发现该变异的Ｇ－等位基因和骨折的发生有联系，ＢＭＤ不影响这种联系。

第五章，我们着重研究了整个ＶＤＲ基因的单倍型和一个位于3’端的多肽性变异同骨折的危险性关系。本章描述了一个基于整个“鹿特丹研究”人群的流行病学相关性研究和一个基于文献的荟萃分析（Meta-analysis）。

在5.1章节，根据对鹿特丹人群的６，５００对象的研究，我们揭示了由特征性ＳＮＰ定义并含盖整个ＶＤＲ基因的单倍型（基于第三章的结果）与骨折发病具有相互关联。我们发现发生骨折的危险性与两个位于ＶＤＲ基因启动区和一个位于3’端的单倍型有明显相关，并且这三个单倍型对骨折的危险性有相加效应。这种联系还见于不同类型的骨折，例如：手腕、股骨和脊椎骨折，并不受性别和年龄的影响。我们也分析了ＶＤＲ单倍型与ＢＭＤ或骨量减少的可能联系，一个位于启动区的骨折相关性单倍型被发现与ＢＭＤ的下降有关，另一个位于3’端的骨折相关性单倍型与骨量减少有关。但是这些ＢＭＤ或骨量减少并不能影响ＶＤＲ单倍型与发生骨折的危险性，这提示我们可能存在别的病理机制。

在5.2章节，我们系统地回顾了所有已发表的关于ＶＤＲ基因Bsm Ⅰ限制性内切酶多肽性（RFLP）和骨折危险性的流行病学研究，并进行了荟萃分析。到目前为止Bsm Ⅰ RFLP是被研究得最多的ＶＤＲ多肽性变异点，因此我们的荟萃分析着眼于该ＲＦＬＰ。我们的分析发现没有证据显示ＶＤＲ基因的Bsm Ⅰ RFLP与骨折的发病相关。然而，Bsm Ⅰ RFLP并不是一个好的遗传标志代表ＶＤＲ３’端的单倍型。因此，荟萃分析的结果并不能完全代表ＶＤＲ其它多肽性变异，特别是单倍型对骨折发生的遗传学作用。

第六章，根据对鹿特丹人群和另一个人群（LASA）研究，我们揭示了以上ＶＤＲ骨折相关性单倍型等位基因及其联合基因型与身高的联系，该联系不受年龄、性别和脊椎骨折地影响。另外在鹿特丹人群，该启动区和3’端的联合基因型同时被发现与腰椎的纵向骨面积及股骨颈狭窄处的宽度的减少相关。在接下来的荟萃分析中我们还发现ＶＤＲ的Bsm Ⅰ RFLP与成人的身高差异有关联，这一结果和我们的流行病学发现相吻合。

第七章，我们运用电泳迁移率变动分析（EMSA）和转录活性试验对ＶＤＲ基因1e-1a启动区的ＳＮＰ进行了功能性研究，同时也在不同的细胞系里对位于ＶＤＲ基因3’端非转录区（3’-UTR）的主要单倍型等位基因，进行了ＶＤＲ信息核糖核酸（mRNA）水平及其稳定性的检测。我们发现位于启动区的与骨折相关的等位基因和转录因子的结合力低，以至于ＶＤＲ基因的表达水平较低。在3’-UTR的与骨折相关的等位基因也表现出低的ＶＤＲ的mRNA表达水平及较低的mRNA稳定性。这些功能性实验的结果与我们关于骨折和身高的流行病学结论相符。

第八章，通过整合不同来源的有关维生素Ｄ结合蛋白（ＤＢＰ）基因的ＬＤ数据资料，我们描绘出关于整个ＤＢＰ基因的ＬＤ图谱。在部分鹿特丹人群中我们观

察到ＤＢＰ基因的单倍型与血清中的维生素Ｄ水平有显著性关联。我们同时在整个鹿特丹研究人群中发现，ＤＢＰ基因的单倍型与饮食中的钙摄入量的互相作用以及ＤＢＰ和ＶＤＲ基因的单倍型之间的互相作用对骨折的发生有显著影响。本研究也提示，ＤＢＰ对骨折的影响基理可能是通过对破骨细胞的直接活化作用，而不是通过维生素D内分泌系统发挥作用。

第九章，基于整个论文集的结果，我们讨论了生物信息学在遗传学研究上的应用，遗传学上连锁分析、相关性研究的比较，全基因组相关性分析，遗传流行病的荟萃分析，以及基因单独的多肽性变异点或单倍型分析在遗传流行病学上的应用。我们还总的评述了ＶＤＲ多肽性变异与不同的骨代谢表型的相互关系，造成不同相关性研究结果的可能原因，普通基因变异所表现的遗传性作用的强度。最后我们也提供了一些对今后研究工作的建议，包括：如何确定被选基因的功能性变异和遗传流行病学研究在药物遗传学上的应用。

第十章，论文集的英文、荷兰文和中文总结。

Epilogue

ACKNOWLEDGEMENT - 致谢

The thesis is finished now, but the contribution of those people, who helped me for years to make this landmark in my career and life, will forever stay in my heart. I feel very honoured to acknowledge them at the end of this work.

The first gratitude goes to my promoter, Prof. Dr. H.A.P. Pols. Your knowledge and experience in the research field and medical practice is enormous. I am honored to have a world famous scientist as my mentor. I still remember when you gave me great support and encouragement before and after my first oral presentation at an international conference in Davos. I learned a lot from you and I was impressed by your brilliant suggestions every time when we had discussions. Thank you so much to help me to realize my desire and also thanks for the guidance of the project, even for the more detailed procedures of my defense.

My co-promoter, Dr. A.G. Uitterlinden, is a miracle maker in my eye. Success needs hard work, but hard work does not always reach success. You showed me how hard work brings people success. Usually you are one of those people who go home latest. After you picked me up (holding a piece of paper with my name on it) in Schipol airport in 1997, only two of us struggled in the genetics of osteoporosis in our lab, now we have nine people working together. Your door is always open for my professional and personal advice. You told me once that my success was also one of yours, I believe this is also true for other people in our group. You provide opportunities and support to all of us to get achievement. We all are pleased to see that our group is becoming one of the leading ones over the world in our research field. You patiently guide me so much in my work. People know that you are my co-promoter, but I think we are friends, and we are also fans of classical guitar.

I would like to thank the members of my core and plenary doctoral committee. Prof. Theo Visser, who is the secretary of the core committee, collected all evaluations and comments from other members in the final phase while producing this thesis. I acknowledge the great opinion for the thesis from Prof. Paul Lips, Prof. Cornelia van Duijn, Prof. Ben Oostra, Dr. Albert Brinkmann. Especially thanks Prof. Hongwen Deng, who gave me so many inspiring and helpful comments. We also had many nice discussions during several science meeting. It is pity that you can not come to my defense, because you are moving and to be a director of the Orthopedic Surgery and Basic Medical Sciences department in University of Missouri.

I will never forget all my dear colleagues in our laboratory:

Dr. Joyce van Meurs, you are a successful scientist in our lab (at least in my mind). You started your post-doc project only half a year before my PhD project, and now you have so many publications, especially the one in the New England

Journal of Medicine. In addition, you have very important "products": Coen and Paula (two kids), as you joked once. You also are a co-author of my manuscripts and the Dutch translator of the summary, since you are so familiar with my work. We had so many delightful and helpful discussions in the last few years.

Dr. Fernando Rivadeneira and I have a similar background: we are foreigners in the Netherlands, we have a medical background, we have a NIHES Master of Science (MSc.) degree, we are carrying out genetic association studies in the bone field, we are also roommates during several bone conferences, and, finally, we are "team-player" of the ECTS football tournaments. I am impressed by your diligence and carefulness in your research, and inspiring attitude in life. As Prof. Pols evaluated at the end of your PhD ceremony, you survived in the Dutch environment.

One of my paranymphs, Pascal. You are one of the people who worked with me for the longest time in our lab. You are also one of my best friends in the Erasmus MC. We have shared so many happy times during our work. Besides me, you have also kindly helped many people from our lab, our department, other departments, other universities and even in other countries. Lisette is my other paranymph. Your excellent suggestions and help with the arrangement of the reception and party was a vital part of the day of my defense. Mila contributed very much work to the functionality assay of the VDR 3'-UTR *in vitro*. You were the first technician student who I met in 1997, you were only two days earlier than I to join the "vitamin D group". Thanks for your close friendship. Thanks also to Wendy, for your helpful introductions and suggestions of Dutch culture (either in English or in Dutch), I learned many interesting Dutch stories from you. Rowena is another bright technician in our lab, I was so surprised to see that you could learn how to analyse LD structure of the VDR so quickly with programs you had never touched. Marco (Medici), you also surprised me with your submission of a manuscript to JBMR in such a short time of your study. Of course I want to thank very much the visitor from Shanghai, Dr. Hongyan Zhao and my student, Angela, for your helps in genotyping VDR polymorphisms in DNA samples.

I would like to express my gratitude to Dr. Hans van Leeuwen and the people in our "bone group". Thanks Hans for your pertinent comments on my manuscripts, for your advice and encouragement of my presentations, and also for your academic support of my research career. Also thanks for the "werkbesprekingen" you organised every Tuesday. It extended my knowledge fundamental of bone biology with presentations by Marjolein (van Driel), Bram, Marco (Eijken), Irene, Sander, Holger, Justus, Martijn (van der Velde), Nadia, Yvonne, Marijke and Cok. Here I will specially thank Marco and Marijke for initial help with cell culture and the RT-PCR of our functional studies.

I also would like to thank Ronald van de Wal who gave me much help with computer and software support. We also worked together on the sequencing machine (ABI 310 and 3100), when I started this project. Carola Zillikens, thanks for

your suggestion of the treatment of my knee. Many thanks to other friends and colleagues from the department of Internal Medicine for making my stay at this group a happy one: Marielle (secretary of Prof. Pols), Edith, Monique, Ellen, Wim, Robin, Wendy, Jose, Jurgen, Frank, Lenie, Carine, Nel, Marianna, Axel, Jenny, Bas, Miriam, Anke (van Kerkwijk), Anke (Mclusky-Dankbar), Martijn, Marlies, Marlijn, Diana, Leo, Jan Willem, Henk, Pauline, Giovanni, Carlotta, Patrick, Michel, Peter, Jolanda, Saskia, Hans, Piet, Jonneke, Jeanine, Brigitta and Ineke.

I also acknowledge people from the department of Epidemiology & Biostatistics. Prof. Theo Stijnen not only taught me statistical courses during my Master of Science program, but also provided much statistical advice for the thesis and my manuscripts. Dr. Chris de Laet also provided help with the ecological study. Other people from the department, Marjolein (van der Klift), Stephanie, Mariette and Wietske, we have shared so many discussions in the past years.

Many acknowledgements go to Prof. John Ioannidis who much impressed me with helpful comments and suggestions on my meta-analyses. Those comments guided me to learn many new developments, methods and more accurate analyses in the field of meta-analysis. I was also very much impressed that he replied his comments on my manuscript so quickly (within only three hours!). I would like to thank Dr. Fakhredin Ali Sayed Tabatabaei who provided initially suggestions for doing a meta-analysis (especially for the RevMan program). We also had many helpful discussions during my analysis.

Drs. Frederic Jehan and Arnold d'Alesio from the collaborating group in Paris carried out several functional studies for the VDR promoter polymorphisms. Our happy and impressive collaboration creates two publications in high rank journals. We had a delightful collaboration with Prof. Paul Lips and Dr. Natasja van Schoor on the replication studies in the Rotterdam Study and in the LASA population.

Many thanks to all members of BV Erasmus, our badminton club. When we shared time together, I felt happy (as a king), strong (for my body), and relaxed (form hard work). Especially for members of our team 2, Pascal, Bart (also my partner of man's double championship in NSK 2005), Cathelijn, Jolanda and Linda, we won a championship of our pool in 2004.

Of course, I would like to thank my Chinese friends in the Netherlands.

当然我也要感谢那些在荷兰的中国朋友们。张大刚先生和张太太（黄颖玉女士），我是参加你们为留学生们提供的家庭式聚会最多的一员，十多年来你们无偿地付出大量时间和经历，以这种形式为无数的在荷中国留学生提供了每周一次的相互交流机会，你们是留学生们真正的朋友。朱银南医生和朱太太（朱丽芳女士），忠心感谢你们在我刚到荷兰那几年里对我以长辈的照顾。刘迎迎博士和马利宾博士，我终身难忘我们那些知心的交流和你们给予的最具建设性的帮

助。其他还有许多我先后“共同战斗”在荷兰朋友们：陈永平、陈可美、姜宇飞、林海祥、王琳、杜青、林涛、彦军、李颖、麻晓迁、罗崇德、李铁强、柴文霞、李雯、史保明、张桌力、梨德元、王蕾、刘帆、张学源、胡雅梅、梁国勇、李英、王勇、刘嘉玲、赵宇浩、苏番、朱泽、段芮、商鹏、侯君、周晨、焦涛和张玲，非常高兴和你们渡过了那么多的快乐时光，让我消去了学习和工作的疲劳和压力。

Finally, I would like to acknowledge my dear parents, and all family members:

最后我还要表达我对年迈父母的无限亲情。也是您们那不计回报的爱一直支持着我到现在的一切，特别是我出国以后，每周一次的越洋电话是您们给予我支持和教诲方式，您们虽然不懂我从事的专业，但您们每次的提醒还是让我受益匪浅，我从而调整那些不当的言行。姐姐方真和姐夫章锡平不仅要承担自己家里的家务，还完全照顾起了父母那里一切大大小小的事务，这是我能安心完成多年在国外学习和工作的可靠保障，父母能安享晚年的幸福生活完全是你们的辛劳结果，章霄堃也能在紧张的学习之余为视力不好的外公读报、交谈，并帮外婆干些活儿，让他们感到家庭生活的温馨。君芝姐和友德姐夫，我同时也要感谢你们多年来对我学习工作的指导和帮助，你们还不顾自己年事已高关心我父母的生活。我还要感谢其他那些在中国和海外的、关心过我的亲友们，感谢你们对我和家人的关怀和照顾。

Yue Fang

Rotterdam, November, 2005

CURRICULUM VITAE

1967, 05, 19	Born in Wuhan, Hubei province, China
1979, 09 – 1985, 06	Wuhan 12th Middle School, China
1985, 09 – 1990, 06	**Bachelor degree of Medicine** Department of Public Health TongJi Medical University in Wuhan, China (Reference: Prof. Dr. Wu Zhaide)
1990, 06 - 1995, 08	**Research trainer & physician trainer** Center for Disease Control and Prevention of Hubei Province, China (Reference: Prof. Dr. Huang Xianhao)
1995,09 – 1997,05	**Bachelor degree project of computer scienc**e for 1.5 years (not accomplished) Chinese Geological University (Wuhan), China (Reference: Mr. Zhou Biao)
1995, 09 - 1997, 05	**Research Assistant & Senior Physician** Center for Disease Control and Prevention of Hubei Province, China (Reference: Prof. Dr. Huang Xianhao)
1997, 08 - 1999, 06	**Master of Science Degree** in general epidemiology The Netherlands Institute for Health Sciences (NIHES) and Erasmus Medical Center Rotterdam The Netherlands (Reference: Dr. Andre Uitterlinden)
1999, 02 - 2000, 11	**Master of Science Degree** in genetic statistics & epidemiology Department of epidemiology Pavia University, Pavia, Italy (References: Prof. Luisa Bernardinelli, Prof. Marco Fraccaro & Prof. Roberto Schmid)
2000, 09 - 2005, 12	**Doctor of Philosophy (PhD.) Degree** in Molecular Genetics, Genetic Epidemiology Department of Internal Medicine Erasmus Medical Center Rotterdam The Netherlands (References: Dr. Andre G. Uitterlinden & Prof. Dr. Huibert A. P. Pols)

AWARDS AND GRANTS

2005, Nov.	Nominee (to represent Chinese students in the Netherlands) of 2005 **Chinese Government Award** for outstanding self-financed students abroad
2005, Jun.	**Young Investigator Award** at 2nd joint meeting of ECTS (European Calcified Tissue Society) and IBMS (International Bone and Mineral Society) in Geneva Switzerland
2004, Oct.	2004 Web Lee **Young Investigator Award** during 26 ASBMR (American Society for Bone and Mineral Research) annual meeting in Seattle, USA
2004, Jun.	2004 ECTS (European Calcified Tissue Society) **Young Investigator Award** at 31 ECTS annual meeting in Nice, France
2004, Jan.	**Young Presenter's Award** and **travel grant** from the Second Asian Regional IOF (International Osteoporosis Foundation) Conference on Osteoporosis in HongKong
2003, Jun.	**Travel grants** for the 1st Joint Meeting of the International Bone and Mineral Society and the Japanese Society for Bone Mineral Research in Osaka, Japan, from IBMS and NVCB (Nederlands Tijdschrift voor Calcium- en Botstofwisseling)
2003, May	2003 ECTS (European Calcified Tissue Society) **Young Investigator Award** at 30 ECTS annual meeting in Roma, Italy
2002, May	2002 ECTS (European Calcified Tissue Society) **Young Investigator Award** at 29 ECTS annual meeting in Zagreb, Croatia

LIST OF PUBLICATIONS

1. **Fang Y**, van Meurs JB, d'Alesio A, Jhamai M, Zhao HY, Rivadeneira F, Hofman A, van Leeuwen J, Jehan F, Pols HA, Uitterlinden AG. 2005 Promoter and 3'-UTR haplotypes in the vitamin D receptor (VDR) gene predispose to osteoporotic fracture: The Rotterdam Study. Am. J. Hum. Genet, 77:807-823.

2. Uitterlinden AG, **Fang Y**, van Meurs JB, Pols HA 2005 Genetic Vitamin D Receptor Polymorphisms and Risk of Disease. In: David Feldman JWP, Francis H. Glorieux (ed.) Vitamin D, 2nd edition, vol. 2. Elsevier, Burlington, pp 1121-57.

3. Claassen M, Nouwen J, **Fang Y**, Ott A, Verbrugh H, Hofman A, van Belkum A, Uitterlinden A 2005 Staphylococcus aureus nasal carriage is not associated with known polymorphism in the Vitamin D receptor gene. FEMS Immunol Med Microbiol 43(2):173-176.

4. **Fang Y**, Uitterlinden AG 2004 62 polymorphisms of the vitamin D receptor gene were published in dbSNP database (accession code: ss32465633 to ss32465690, and ss32465817to ss32465820) of NCBI build 124

5. **Fang Y,** Uitterlinden AG 2004 40 sequencing fragments of the vitamin D receptor gene were published in Genbank database (accession code: BV210413 to BV210452↑ and AY827085 to AY827088)

6. Uitterlinden AG, Arp PP, Paeper BW, Charmley P, Proll S, Rivadeneira F, **Fang Y**, van Meurs JB, Britschgi TB, Latham JA, Schatzman RC, Pols HA, Brunkow ME 2004 Polymorphisms in the Sclerosteosis/van Buchem Disease Gene (SOST) Region Are Associated with Bone-Mineral Density in Elderly Whites. Am J Hum Genet 75(6):1032-45.

7. Uitterlinden AG, **Fang Y**, Van Meurs JB, Pols HA, Van Leeuwen JP 2004 Genetics and biology of vitamin D receptor polymorphisms. Gene 338(2):143-56.

8. Uitterlinden AG, **Fang Y**, Van Meurs JB, Van Leeuwen H, Pols HA 2004 Vitamin D receptor gene polymorphisms in relation to Vitamin D related disease states. J Steroid Biochem Mol Biol 89-90:187-93.

9. Schuit SC, van Meurs JB, Bergink AP, van der Klift M, **Fang Y**, Leusink G, Hofman A, van Leeuwen JP, Uitterlinden AG, Pols HA 2004 Height in pre- and postmenopausal women is influenced by estrogen receptor alpha gene polymorphisms. J Clin Endocrinol Metab 89(1):303-9.

10. van Meurs JB, Schuit SC, Weel AE, van der Klift M, Bergink AP, Arp PP, Colin EM, **Fang Y**, Hofman A, van Duijn CM, van Leeuwen JP, Pols HA, Uitterlinden AG 2003 Association of 5' estrogen receptor alpha gene polymorphisms with bone mineral density, vertebral bone area and fracture risk. Hum Mol Genet 12(14):1745-54.

11. **Fang Y**, van Meurs JB, Bergink AP, Hofman A, van Duijn CM, van Leeuwen JP, Pols HA, Uitterlinden AG 2003 Cdx-2 polymorphism in the promoter region of the human vitamin D receptor gene determines susceptibility to fracture in the elderly. J Bone Miner Res 18(9):1632-41.

12. Colin EM, Uitterlinden AG, Meurs JB, Bergink AP, van de Klift M, **Fang Y**, Arp PP, Hofman A, van Leeuwen JP, Pols HA 2003 Interaction between vitamin D receptor genotype and estrogen receptor alpha genotype influences vertebral fracture risk. J Clin Endocrinol Metab 88(8):3777-84.

13. Bergink AP, van Meurs JB, Loughlin J, Arp PP, **Fang Y**, Hofman A, van Leeuwen JP, van Duijn CM, Uitterlinden AG, Pols HA 2003 Estrogen receptor alpha gene haplotype is associated with radiographic osteoarthritis of the knee in elderly men and women. Arthritis Rheum 48(7):1913-22.

14. Uitterlinden AG, **Fang Y**, Bergink AP, van Meurs JB, van Leeuwen HP, Pols HA 2002 The role of vitamin D receptor gene polymorphisms in bone biology. Mol Cell Endocrinol 197(1-2):15-21.

15. Kann P, Bergink AP, **Fang Y**, Van Daele PL, Hofman A, Van Leeuwen JP, Beyer J, Uitterlinden AG, Pols HA 2002 The collagen Ia1 SP1 polymorphism is associated with differences in ultrasound transmission velocity in the calcaneus in postmenopausal women. Calcif Tissue Int 70(6):450-6.

16. Havelka S, Uitterlinden AG, **Fang Y,** Arp PP, Pavelkova A, Vesela M, Halman L, Forejtova S, Ruzickova S, Pavelka K 2002 Collagen type I(alpha 1) and vitamin D receptor polymorphisms in diffuse idiopathic skeletal hyperostosis. Clin Rheumatol 21(4):347-8.

17. Uitterlinden AG, Weel AE, Burger H, **Fang Y**, van Duijn CM, Hofman A, van Leeuwen JP, Pols HA 2001 Interaction between the vitamin D receptor gene and collagen type Ialpha1 gene in susceptibility for fracture. J Bone Miner Res 16(2):379-85.

18. van Meurs JB, Arp PP, **Fang Y**, Slagboom PE, Meulenbelt I, van Leeuwen JP, Pols HA, Uitterlinden AG 2000 Detection of sequence variability of the collagen type IIalpha 1 3' variable number of tandem repeat. Electrophoresis 21(17):3571-7.

19. Uitterlinden AG, Burger H, Huang Q, **Yue F**, McGuigan FE, Grant SF, Hofman A, van Leeuwen JP, Pols HA, Ralston SH 1998 Relation of alleles of the collagen type Ialpha1 gene to bone density and the risk of osteoporotic fractures in postmenopausal women. N Engl J Med 338(15):1016-21.

20. **Fang Y**, Huang X, Liu W 1997 Serotyping of hemorrhagic fever with renal syndrome in Hubei province. Zhonghua Yu Fang Yi Xue Za Zhi 31(5):272-4.

21. **Fang Y**, et al. Analysis on Hemorrhagic Fever with Renal Syndrome (HFRS) Surveillance in Hubei Province 1991-1995. Chinese Journal of Disease Surveillance, 1997; 12(7):245-248.

22. **Fang Y**, et al. Trend of epidemic situation changes on HFRS in Early 1990s in Hubei Province. Chinese Journal of Vector Biology and Control, 1995; 6(2):111-113.

23. Huang X, **Fang Y,** et al. Epidemiological surveillance of Haemorrhagic Fever with Renal Syndrome in Hubei province. Chinese Journal of Epideniology, 1994; 15(4) suppl: 28.

24. **Fang Y**, et al. Sero-typing of acute hepatitis patients in Hubei province. Chinese Journal of Epidemiology, 1994; 15(4) suppl:1-4.

25. Cheng DL, **Fang Y,** et al. Surveillance of Haemorrhagic Fever with Renal Syndrome in Wuchang County in 1992. Chinese Journal of Vector Biology and Control, 1993; 4(3)suppl.

26. **Fang Y**, et al. Surveillance and analysis about Haemorrhagic Fever with Renal Syndrome (HFRS) in Hubei province in 1992. Chinese Journal of Vector Biology and Control, 1993; 4(3)suppl.