Advances in Biological Chemistry, 2015, 5, 24-33
Published Online February 2015 in SciRes. http://www.scirp.org/journal/abc
http://dx.doi.org/10.4236/abc.2015.51003

Scientific Research Publishing

The Polymorphisms in the Vitamin D Receptor Gene and Disease Severity in Sickle Cell Disease

E. Leila Jerome Clay[1*], Alison Motsinger-Reif[2], Janelle Hoskins[3], Lindsay Veit[4], Ali Calikoglu[5], Rupa Redding-Lallinger[6]

[1]Departments of Pediatrics and Internal Medicine/Division of Hematology and Oncology, Georgia Regents University, Augusta, GA, USA
[2]Department of Statistics, Bioinformatics Research Center, North Carolina State University, Raleigh, NC, USA
[3]Gentris Corporation, Morrisville, NC, USA
[4]Orthopedic Surgery Resident, University of Toledo Medical Center, Toledo, OH, USA
[5]Department of Pediatrics/Division of Endocrinology, University of North Carolina School of Medicine, Chapel Hill, NC, USA
[6]Departments of Pediatrics and Internal Medicine/Division of Hematology and Oncology, University of North Carolina School of Medicine, Chapel Hill, NC, USA
Email: *leila.clay@gru.edu

Received 31 December 2014; accepted 20 February 2015; published 27 February 2015

Abstract

Vitamin D is important in multiple aspects of health and its effects are mediated through the Vitamin D Receptor (VDR). We wanted to test the hypothesis that specific haplotypes of the VDR gene are associated with markers of disease severity, inflammation and bone health in Sickle Cell Disease (SCD). Genotyping was performed on DNA specimens from 1141 study participants in the NIH-funded Silent Infarct Transfusion (SIT) trial. We used the clinical and laboratory data to create separate endothelial dysfunction, vaso-occlusive severity scores and phenotype variables. Seventy-nine Single Nucleotide Polymorphisms (SNP) in the VDR gene and three associated genes—CYP27B1, VD binding protein, retinoid X receptor, were evaluated. The discovery cohort individuals had VDR haplotype information from a prior Genome-Wide Association Study (GWAS). The validation cohort was analyzed for SNPs that were significant in the discovery cohort. The phenotype data were obtained from the demographic and clinical information of the participants, and were used to create the severity scores, vaso-occlusive score, endothelial dysfunction severity, and overall severity score. Potential gene-gene interactions were analyzed for prediction of disease severity within each severity score. Two SNPs were associated with the overall severity score, 3

*Corresponding author.

How to cite this paper: Clay, E.L.J., Motsinger-Reif, A., Hoskins, J., Veit, L., Calikoglu, A. and Redding-Lallinger, R. (2015) The Polymorphisms in the Vitamin D Receptor Gene and Disease Severity in Sickle Cell Disease. *Advances in Biological Chemistry*, **5**, 24-33. http://dx.doi.org/10.4236/abc.2015.51003

SNPs with the endothelial dysfunction severity score and 4 SNPs with the vaso-occlusive severity score. After permutation testing to correct for multiple comparisons, only one of the associations remained significant. SNP rs7965281 was found to be associated with the endothelial dysfunction severity score and remained significant after correcting for multiple comparisons using permutation testing. In the validation cohort, that SNP was again tested for association with each of the severity scores. There was no association with the endothelial or the overall severity score but a trend towards association with the vaso-occlusive severity score (p = 0.02). None of the known functional polymorphisms in the VDR gene were found to have an association with severity in sickle cell disease. Further work analyzing for gene-gene interaction using the same significant SNPs remains to be done in association with inflammatory markers and measure of bone health. Those studies may provide insight on the contribution of VDR polymorphisms to sickle cell disease severity.

Keywords

Vitamin D (VD), Sickle Cell Disease (SCD)

1. Introduction

Sickle Cell Disease (SCD) is a hereditary disorder primarily affecting people of African ancestry [1]. SCD is a single gene disorder and the complications of this disease affect every organ of the body including the vascular and immune systems along with bone metabolism. Vitamin D deficiency is commonly seen in patients with sickle cell disease especially those with SCD-SS [2] [3]. Vitamin D and calcium are required for optimal bone health with close to 90% of required vitamin D synthesis coming from exposure to sunlight. Dark-skinned individuals usually require 5 - 10 times more exposure to sunlight to produce the same amount of vitamin D3 in their skin [4]. As common as vitamin D deficiency is in SCD, its contribution to disease manifestations is not yet known. Vitamin D is important in multiple aspects of health, including cardiovascular, immune and skeletal systems. The active metabolite of the pro-hormone vitamin D exerts its effect through interaction with the Vitamin D Receptor (VDR). A Single Nucleotide Polymorphism (SNP) is DNA sequence variation occurring commonly within a population in which a single nucleotide in the genome differs between members of a biological species or paired chromosome. The VDR is expressed in multiple cell types and there are specific VDR polymorphisms associated with diseases such as diabetes, hyperparathyroidism, and renal disease (to name a few) [5]. Some of the polymorphisms such as bsmI, fokI, taqI have been found to be associated with bone health, inflammation and vascular disease, respectively [6] [7]. Although vitamin D deficiency has been studied in sickle cell disease to some extent, the polymorphisms in the VDR gene have not been explored in sickle cell disease or in Africans or African Americans [8]. Because of the presence of the VDR on cells in endothelial, inflammatory and bone systems, polymorphisms in this gene may be associated with disease manifestations in these organ systems in SCD. Our hypothesis is that specific VDR polymorphisms are associated with disease severity in sickle cell disease. Potentially some of the variation in the severity of sickle cell disease can be explained by differing polymorphisms in the VDR gene.

2. Materials and Methods

After approval of the study was obtained from the Biomedical Institutional Review Board of the University of North Carolina, genotyping was performed on DNA specimens from 1141 study participants in the NIH-funded Silent Infarct Transfusion (SIT) trial (WU-04-60/PO29892B). In this multi-center international trial, the participants were children ages 4 through 13 with SCD who were screened for the presence of silent cerebral infarction and had demographic and clinical data collected, as well as samples for a biologic repository which included DNA. We used the clinical and laboratory data to create separate endothelial dysfunction and vaso-occlusive severity scores for phenotype variables, following the current thinking on the subtypes of sickle cell disease expression [9]. The initial 570 participants served as our discovery cohort, used to detect potential associations from all the variants tested. The subsequently enrolled 530 individuals formed our validation cohort to try to

replicate/validate statistically significant findings from the discovery cohort. We evaluated 79 Single Nucleotide Polymorphisms (SNPs) in the VDR gene, three associated genes: CYP27B1, VD binding protein, retinoid X receptor, and tagging SNPs from the African American population from Hap map (**Appendix 1**). The discovery cohort individuals had VDR haplotype information from a prior Genome-Wide Association Study (GWAS) study, and analysis for additional VDR-related SNPs was performed using a specifically designed Sequenom assay.

The validation cohort was analyzed for SNPs that were significant in the discovery cohort.

The phenotype data was obtained from the demographic and clinical information of the participants, and was used to create severity scores. The vaso-occlusive severity score includes: number of hospitalizations for pain, number of hospitalizations for acute chest syndrome and avascular necrosis. The endothelial dysfunction severity score includes priapism, Transient Ischemic Attacks (TIA), silent cerebral infarct, systolic and diastolic blood pressure, transcranial doppler velocity, white count and baseline hemoglobin. The overall severity score includes all of the endothelial dysfunction and vaso-occlusive severity variables (**Figures 1(a)-(c)**). To derive the scores, the variables were transformed into quartiles. Each individual subject was assigned values of 1, 2, 3, or 4 for each variable with 1 representing lowest severity and 4, the highest. In addition, in concert with prior analyses of the SIT data, the variable for number of hospitalizations for pain was used alone as a severity measure.

Before association analysis, genotypes underwent quality control filters. First, markers with missing data rates greater that 5% were removed. Second, individuals with missing data rates greater than 5% were removed. Next, markers were tested for deviation from Hardy-Weinberg proportions using a Fisher's exact test and a Bonferroni corrected alpha =0.05 for the number of SNPs tested and markers with significant deviations were removed prior to analysis. This QC resulted in no markers of individuals needing to be removed.

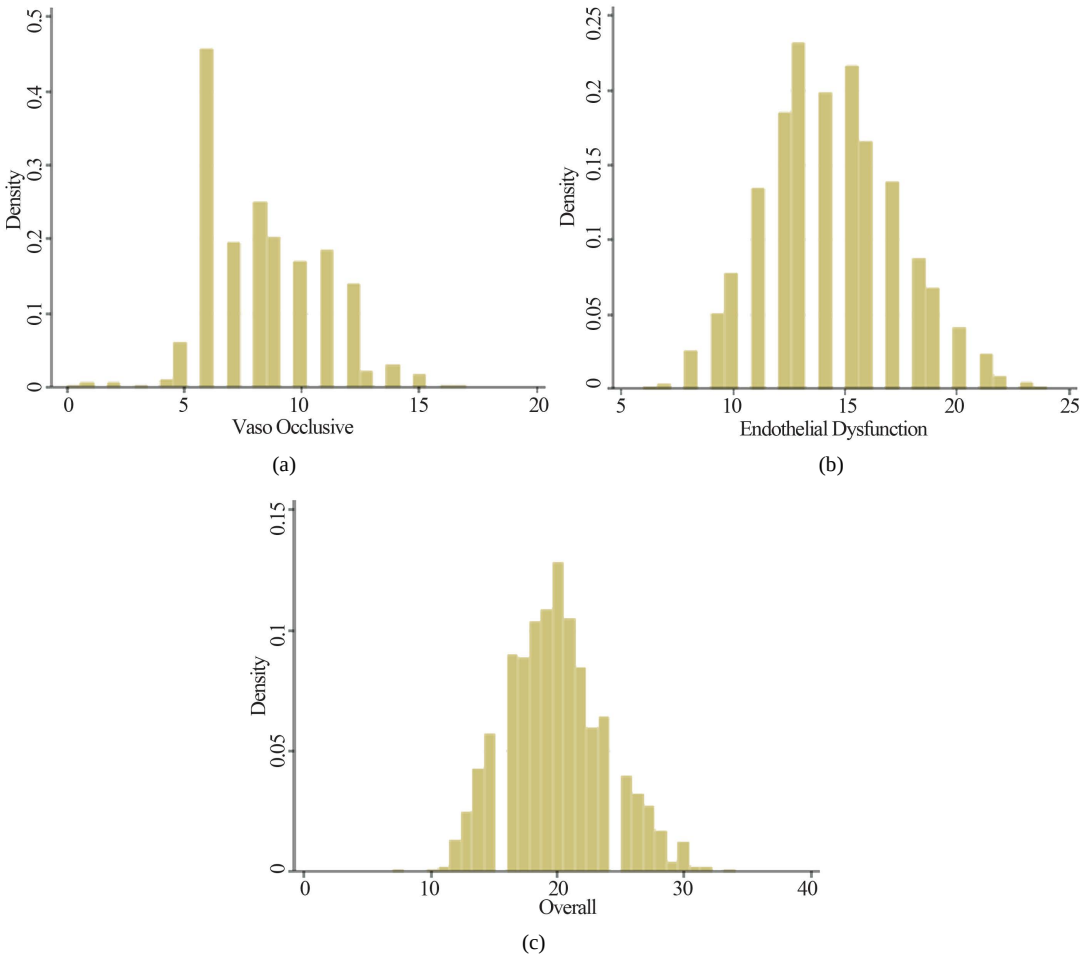

Figure 1. Histogram of severity score distributions in the discovery cohort. The relative density of the validation population for the (a) vaso-occlusive, (b) endothelial dysfunction and (c) overall severity score are shown.

After QC, genetic markers were tested for association with each of the severity scores using a two stage approach. Initially, univariate tests for association were performed. Because severity scores were not normally distributed and were not totally continuous distributions, the Kruskal-Wallis test was used in association analysis (treating each genotype as a category). To correct for multiple comparisons, permutation testing was performed. One thousand permuted datasets (permuting the risk score outcomes while preserving the genotype matrix) were created and the Kruskal-Wallis tests were performed for each of the variants. The lowest p-value from the analysis of each permuted dataset was used to build an empirical distribution of p-values expected by chance, and the p-values for the original analyses were determined by their hypothetical percentile rank in the empirical distribution. Such a permutation approach is a well validated approach for multiple testing correction that is less conservative than a Bonferroni correction when there is correlation between predictor variables [10] [11].

To look for complex genetic models, including potential gene-gene interactions for prediction of disease severity, the Generalized Multifactor Dimensionality Reduction (GMDR) method was utilized, with repeat analyses performed for each severity score. GMDR introduces the concept of a score statistic into the Multifactor Dimensionality Reduction (MDR) framework to obtain an appropriate statistic to classify multifactor contingency table cells into two different groups. The score value is based on a generalized linear model of the phenotype variable, and allows for the use of covariates in the score approximation. Details of GMDR have been previously described in detail [12].

In step one, data was divided into a training set and an independent testing set for cross validation. Five-fold cross-validation was used, with 4/5 of the data used for training and 1/5 for testing. A set of n genetic factors were then selected. These factors and their multiple classes were divided in n-dimensional space. In the traditional MDR approach, the ratios of cases to controls were then calculated within each multifactor class. Each multifactor cell class was labeled “high risk” if the ratio exceeded 1.0, or “low risk” if less than 1.0, thus reducing n-dimensional space to one dimension with two levels. In GMDR, the ratio of cases to controls in each cell is replaced by the score values. The null hypothesis assumes there are no effects of the putative factors or their interactions, so the score values are the same for all different factor classifications. In the third step of GMDR, the cumulative score value was calculated within each multifactor cell and in the fourth step, each multifactor cell was labeled either as high-risk if the average score meets or exceeded a pre-assigned threshold T (e.g., 0), or as low-risk if the threshold was not exceeded. The collection of these multifactor classes comprised the GMDR model. The result is a set of models, one for each model size considered. The final model was chosen that minimizes prediction error while maximizing Cross Validation Consistency (CVC). The statistical significance of the final best model was determined through permutation testing, which involved creating 1000 permuted datasets by randomizing the value of the phenotype variable. The entire procedure was repeated for each, generating a distribution of 1000 prediction errors that could be expected by chance alone. The significance of the final model was determined by comparing the prediction error of the final model to the distribution. A p-value was extracted for the model by its theoretical location in the permutation distribution.

In the present study, GMDR analysis was performed with the same outcome variables listed above as the phenotypes, using stratification when necessary and all other genetic markers as potential predictor variables. Analysis was performed with 5-fold cross-validation, and single-variable through four-variable interactions were evaluated. The permutation testing was used to empirically define the prediction accuracies that are significant with a family-wise type I error rate of 0.05.

The study was designed to have at least 80% power to detect odds ratios of 2.15 for the univariant association of variants with minor allele frequencies of at least 0.25; 58% of our candidate genes fell within this range. These power calculations were performed assuming an F distribution with a Bonferroni corrected alpha of 0.05. SNPs that were significantly associated with severity in the discovery cohort were tested for association in the validation cohort. Each SNP was directly tested for association using a Kruskal-Wallis test after QC checks. Since this replication analysis represents a single statistical hypothesis, no correction for multiple testing was performed.

3. Results

By univariate testing in the discovery cohort, 2 SNPs (rs1491710, rs11574114) were nominally associated with the overall severity score, 3 SNPs (rs11829917, rs2853563, rs11574138) with the endothelial dysfunction severity score and 4 (rs7855881, rs12348547, rs11574114, rs7965281) with the vaso-occlusive severity score. One of

the SNPs, rs11574114, was associated with both the overall and the vaso-occlusive severity scores.

Our severity scores, although based on current thinking about sickle cell disease, have not been validated in other work. The measure that has been validated as a measure severity in the SIT population (the source of the data for this study) is number of hospitalization for pain. When tested by univariate analysis, there was a nominal association with 2 of the SNIPS (rs7855881, rs34312136) with this variable. However, after permutation testing to correct for multiple comparisons, none of the associations remained significant.

Using MDR to test for significant gene-gene interactions in the discovery cohort, one single locus association between SNP rs7965281 and the endothelial dysfunction score was significant, and remained significant after permutation testing to correct for multiple comparisons. The mean endothelial dysfunction severity score for those homozygous for A and for G (the variant allele) was 15, and it is 13 for the heterozygotes (AG). The range for the endothelial dysfunction score is 7-23, with higher numbers corresponding to greater severity.

In the validation cohort, there was no univariate association with the one significant SNP from the discovery cohort, rs7965281. However, that SNP was significantly associated with the vaso-occlusive severity score in the validation cohort by MDR testing (p value = 0.02).

4. Discussion

Overall there does not appear to be a significant association between any of the currently described functional VDR genes, such as bsmI, fokI, taqI and severity in SCD. With the current study's sample size and power, only 58% of our candidate genes had a minor allele frequency sufficient for testing for an association with disease severity. The association between the VDR and the SCD severity could be further explored in a larger SCD population. In addition, a sickle cell disease population for which there are data on bone health, disease and inflammatory markers should be studied in relation to VDR haplotypes.

Concerning the single SNP that was found to be significantly associated with SCD severity, not much is known currently. In the literature, SNP rs7965281 is a tagging SNP whose significance is not fully clear. It is located in the regulatory region UTR + 3713: Chr position 45262658 as per dbSNP (HuRef). It was found to be associated with reduced risk for cutaneous melanoma in a large population [13]. In a marker report, a trend analysis was done with this SNP and an association with systolic blood pressure was found in a British population, with an adjusted p value of 0.029. More work is being currently done on this particular SNP (personal communication with Dr. Orlow [13]).

While our study did not show a relationship between the well-described VDR gene polymorphisms and the SCD severity, one SNP was found which may be of importance. In this study with a two-stage approach to association analysis, the SNP found was associated with disease severity in both the discovery and the validation cohorts, but it was not associated with the same severity scores in the two cohorts, making the results somewhat less precise that one would like. However, in the current conception of the clinical expression of sickle cell disease, the vaso-occlusive and the endothelial dysfunction disease manifestations do overlap, so the association of SNP rs7965281 with both of the scores does not invalidate its significance. Clearly, an understanding of the biologic function of SNP rs7965281 is necessary if the results of the current study are to be built upon, and these results should be evaluated in other studies of the VDR and sickle cell disease severity to confirm this association.

Acknowledgements

This project is an ancillary study of the SIT trial (NIH WU-04-60/PO29892B).

This manuscript was prepared during the corresponding author's training and was supported by the T32 NIH grant *PHS GRANT* 5*T*32 *HL* 7149-35.

This project was supported by a grant from North Carolina Translational and Clinical Sciences Institute.

We thank the staff, clinicians and patients for their participation in the Silent Infarct Transfusion (SIT) Trial study. We would like to thank the members of the SIT Executive Committee for their comments and suggestions during the planning and conduct of this study. The SIT study (PI: M. DeBaun, MD) is supported by the National Institute of Neurological Disorders and Stroke (NINDS) (NIH-NINDS 5U01-NS042804). We also thank the SIT Trial Biologic Repository (J. F. Casella, MD, E. Barron-Casella, PhD and K. Jones) for providing plating f genomic DNA. Genotyping services were provided by the Center for Inherited Disease Research (CIDR) and the Centers for Disease Control and Prevention (CDC). CIDR is fully funded through a federal contract from the

National Institutes of Health to The Johns Hopkins University (PI: J.F. Casella, MD) (CIDR contract #: HHSN268200782096C). Additional assistance with the GWAS data was provided by Drs. D.E. Arking and P. Bhatnagar (JHU) and W.C. Hooper and C.J. Bean (CDC).

We also would like to acknowledge David Barrow, DDS, for his contribution in the early stages and planning of this project.

References

[1] Fixler, J. and Styles, L. (2002) Sickle Cell Disease. *Pediatric Clinics of North America*, **49**, 1193-210, vi. http://dx.doi.org/10.1016/S0031-3955(02)00089-5

[2] Buison, A.M., Kawchak, D.A., Schall, J., Ohene-Frempong, K., *et al*. (2004) Low Vitamin D Status in Children with Sickle Cell Disease; Osteopenia and Vitamin D Deficiency in Children with Sickle Cell Disease. *Journal of Pediatrics*, **145**, 622-627. http://dx.doi.org/10.1016/j.jpeds.2004.06.055

[3] Reed, J.D., Redding-Lallinger, R. and Orringer, E.P. (1987) Nutrition and Sickle Cell Disease. *American Journal of Hematology*, **24**, 441-455. http://dx.doi.org/10.1002/ajh.2830240416

[4] Adewoye, A.H., Chen, T.C., Ma, Q., McMahon, L., Mathieu, J., Malabanan, A., *et al.* (2008) Sickle Cell Bone Disease: Response to Vitamin D and Calcium. *American Journal of Hematology*, **83**, 271-274. http://dx.doi.org/10.1002/ajh.21085

[5] Valdivielso, J.M. and Fernandez, E. (2006) Vitamin D Receptor Polymorphisms and Diseases. *Clinica Chimica Acta*, **371**, 1-12. http://dx.doi.org/10.1016/j.cca.2006.02.016

[6] Bouillon, R., Norman, A.W. and Lips, P. (2007) Vitamin D Deficiency. *New England Journal of Medicine*, **357**, 1980-1981. http://dx.doi.org/10.1056/NEJMc072359

[7] Rovner, A.J., Stallings, V.A., Kawchak, D.A., Schall, J.I., Ohene-Frempong, K. and Zemel, B.S. (2008) High Risk of Vitamin D Deficiency in Children with Sickle Cell Disease. *Journal of the American Dietetic Association*, **108**, 1512-1516. http://dx.doi.org/10.1016/j.jada.2008.06.433

[8] Uitterlinden, A.G., Fang, Y., Van Meurs, J.B., Pols, H.A. and Van Leeuwen, J.P. (2004) Genetics and Biology of Vitamin D Receptor Polymorphisms. *Gene*, **338**, 143-156. http://dx.doi.org/10.1016/j.gene.2004.05.014

[9] Kato, G.J., Gladwin, M.T. and Steinberg, M.H. (2007) Deconstruction Sickle Cell Disease: Reappraisal of the Role of Hemolysis in the Development of Clinical Subphenotypes. *Blood Reviews*, **21**, 37-47. http://dx.doi.org/10.1016/j.blre.2006.07.001

[10] Lou, X.Y., Chen, G.B., Yan, L., Ma, J.Z., Zhu, J., Elston, R.C., *et al.* (2007) A Generalized Combinatorial Approach for Detecting Gene-by-Gene and Gene-by-Environment Interactions with Application to Nicotine Dependence. *American Journal of Human Genetics*, **80**, 1125-1137. http://dx.doi.org/10.1086/518312

[11] Good, P. (2001) Extensions of the Concept of Exchangeability and Their Applications. *Journal of Modern Applied Statistical Methods*, **1**, 243-247.

[12] Lou, X.Y., Chen, G.B., Yan, L., Ma, J.Z., Zhu, J., Elston, R.C. and Li, M.D. (2007) A Generalized Combinatorial Approach for Detecting Gene by Gene and Gene by environment Interactions with Application to Nicotine Dependence. *American Journal of Human Genetics*, **80**, 1125-1137. http://dx.doi.org/10.1086/518312

[13] Orlow, I., Roy, P., Reiner, A.S., Yoo, S., Patel, H., Paine, S., *et al*. (2011) Vitamin D Receptor Polymorphisms in Patients with Cutaneous Melanoma. *International Journal of Cancer*, **130**, 405-418. http://dx.doi.org/10.1002/ijc.26023

Appendix 1

Original SNPs	Gene	Chromosome	Position	Reference allele	Other allele	African ancestry in South-west USA (HapMap III) other allele frequency	Yoruba in Ibadan, Nigeria (HapMap III) other allele frequency	Source	Genotyped SNPs	r^2	Genotype platform	Comment/ Alt. snp name
rs1048691	CYP27B1	12	58152948	C	T	0.33	0.394	Haploview Tagger	rs1048691		650Y	
rs10488854	GC	4	72624135	C	T	0.198	0.155	Haploview Tagger	rs10488854		Sequenom	
rs10735810 /rs2228570	VDR	12	48272895	A	G	0.802	0.808	PharmGKB	rs2228570		Sequenom	FokI
rs10877012	CYP27B1	12	58162085	G	T	0.16	0.088	Haploview Tagger	rs10877012		Sequenom	
rs10881578	RXRA	9	137232535	A	G	0.377	0.394	Haploview Tagger	rs10881578		Sequenom	
rs10881582	RXRA	9	137256078	G	A	0.623	0.73	Haploview Tagger	rs10881582		650Y	
rs11168266	VDR	12	48251533	C	T	0.528	0.673	Haploview Tagger	rs11168266		Sequenom	
rs11168268	VDR	12	48251812	G	A	0.566	0.708	Haploview Tagger	rs11168268		Sequenom	
rs11172327	CYP27B1	12	58161660	G	A	0.066	0.071	Haploview Tagger	rs8176353		Sequenom	
rs11185649	RXRA	9	137228370	T	C	0.585	0.643	Haploview Tagger	rs11185649		Sequenom	
rs11185659	RXRA	9	137243383	C	T	0.283	0.319	Haploview Tagger	rs11185659		650Y	
rs1155563	GC	4	72643488	T	C	0.057	0.049	Haploview Tagger	rs1155563		650Y	Ahn *et al.* 2010
rs11568820	VDR	12	48302545	C	T	0.717	0.978	PharmGKB	rs11568820		Sequenom	Cdx2
rs11574114	VDR	12	48238883	C	T	0.151	0.226	Haploview Tagger	rs11574114		Sequenom	
rs11574138	VDR	12	48235702	T	C	0.066	0.124	Haploview Tagger	rs11574138		Sequenom	
rs11574143	VDR	12	48234917	C	T	0.132	0.08	Haploview Tagger	rs11574143		650Y	
rs11829917	CYP27B1	12	58164455	G	A	0.123	0.124	Haploview Tagger	rs11829917		Sequenom	
rs12308082	VDR	12	48252139	G	A	0.066	0.054	Haploview Tagger	rs12308082		Sequenom	
rs12314197	VDR	12	48242722	A	G	0.245	0.217	Haploview Tagger	rs12314197		Sequenom	
rs12339187	RXRA	9	137229327	A	G	0.206	0.221	Haploview Tagger	rs12339187		650Y	
rs12348547	RXRA	9	137246262	T	C	0.123	0.168	Haploview Tagger	rs12348547		Sequenom	
rs12351315	RXRA	9	137225103	A	T	0.689	0.788	Haploview Tagger	rs11185647	0.965	650Y	LD with SNP on original list
rs12640179	GC	4	72612687	C	G	0.085	0.081	Haploview Tagger	rs12640179		Sequenom	

Continued

rs1352841	GC	4	72605170	T	C	0.16	0.124	Haploview Tagger	rs1352841		Sequenom	
rs1352844	GC	4	72647749	C	T	0.113	0.125	Haploview Tagger	rs1352844		650Y	
rs1352845	GC	4	72647762	A	G	0.34	0.403	Haploview Tagger	rs2298849	1	650Y	LD with SNP on original list
rs1491709	GC	4	72613566	G	A	0.075	0.08	Haploview Tagger	rs1491709		650Y	
rs1491710	GC	4	72609094	A	C	0.415	0.433	Haploview Tagger	rs1491710		Sequenom	
rs1491711	GC	4	72602252	C	G	0.358	0.354	Haploview Tagger	rs1491711		Sequenom	
rs1540339	VDR	12	48257326	C	T	0.245	0.204	PharmGKB	rs1540339		650Y	
rs1544410	VDR	12	48239835	C	T	0.217	0.279	PharmGKB	rs1544410		650Y	BsmI
rs16847024	GC	4	72650679	C	T	0.066	0.102	Haploview Tagger	rs16847024		Sequenom	
rs1873590	GC	4	72653480	A	G	0.142	0.15	Haploview Tagger	rs1873590		Sequenom	
rs188812	GC	4	72627684	A	T	0.189	0.137	Haploview Tagger	rs188812		Sequenom	
rs2107301	VDR	12	48255570	G	A	0.179	0.137	Haploview Tagger	rs2107301		650Y	
rs2189480	VDR	12	48263828	G	T	0.387	0.389	Haploview Tagger	rs2189480		650Y	
rs222014	GC	4	72632931	C	T	0.066	0.064	Haploview Tagger	rs222014		650Y	
rs222016	GC	4	72634975	G	A	0.462	0.389	Haploview Tagger	rs222016		650Y	
rs222047	GC	4	72610208	C	A	0.802	0.888	Haploview Tagger	rs222047		650Y	
rs222049	GC	4	72605955	G	C	0.154	0.177	Haploview Tagger	rs222046	0.869	650Y	LD with SNP on original list
rs2239179	VDR	12	48257766	T	C	0.311	0.292	PharmGKB	rs2239179		650Y	
rs2239182	VDR	12	48255411	T	C	0.519	0.611	Haploview Tagger	rs2239182		650Y	
rs2239184	VDR	12	48244583	G	A	0.594	0.634	Haploview Tagger	rs2239184		650Y	
rs2239185	VDR	12	48244559	G	A		0.542	PharmGKB	rs2239185		Sequenom	
rs2248098	VDR	12	48253356	A	G	0.423	0.593	Haploview Tagger	rs2248098		Sequenom	
rs2282678	GC	4	72608792	A	G	0.264	0.312	Haploview Tagger	rs16846912 and rs4752	0.959	650Y	LD with SNP on original list
rs2525045	VDR	12	48233020	G	A	0.132	0.168	Haploview Tagger	rs2525045		650Y	
rs2853563	VDR	12	48235738	C	T	0.17	0.177	Haploview Tagger	rs2853563		Sequenom	
rs34312136	RXRA	9	137269456					Haploview Tagger	rs34312136		Sequenom	
rs3733359	GC	4	72649774	G	A	0.217	0.283	Haploview Tagger	rs3733359		650Y, Sequenom	Failed 650Y, passed on sequenom

Continued

rs3737549	GC	4	72631668	G	A	0.208	0.219	Haploview Tagger	rs3737549		650Y	
rs3755967	GC	4	72609398	C	T	0.094	0.04	Haploview Tagger	rs3755967		Sequenom	In complete LD with rs2282679 (Wang *et al.* 2010)
rs3782905	VDR	12	48266167	G	C		0.217	PharmGKB	rs3782905		Sequenom	Failed Sequenom
rs3819545	VDR	12	48265006	A	G	0.302	0.221	Haploview Tagger	rs3819545		650Y	
rs41400444	RXRA	9	137231651	G	T	0.075	0.098	Haploview Tagger	No snps in LD			Failed Sequenom assay design
rs4364228	GC	4	72623347	A	G	0.396	0.468	Haploview Tagger	rs4364228		650Y	
rs4646536	CYP27B1	12	58157988	A	G	0.368	0.27	Haploview Tagger	rs703842	0.876	650Y	LD with SNP on original list
rs4646537	CYP27B1	12	58157281	T	G	0.075	0.097	Haploview Tagger	rs4646537		650Y	
rs4917352	RXRA	9	137232030	T	C	0.528	0.58	Haploview Tagger	rs4917352		Sequenom	
rs4917354	RXRA	9	137237661	C	T	0.453	0.5	Haploview Tagger	rs4917354		650Y	
rs6817912	GC	4	72653705	C	T	0.094	0.121	Haploview Tagger	rs6817912		Sequenom	
rs7039190	RXRA	9	137266704	A	C	0.104	0.17	Haploview Tagger	rs7039190		650Y	
rs7041	GC	4	72618334	A	C	0.179	0.088	Haploview Tagger	rs7041		650Y	Ahn *et al.* 2010
rs7048602	RXRA	9	137253447	G	A	0.387	0.527	Haploview Tagger	rs7048602		Sequenom	
rs705117	GC	4	72608115	C	T	0.298	0.155	Haploview Tagger	rs705117		650Y	
rs705119	GC	4	72613036	A	C	0.179	0.097	Haploview Tagger	rs705119		Sequenom	
rs705120	GC	4	72614140	A	C	0.547	0.496	Haploview Tagger	rs705120		Sequenom	
rs731236	VDR	12	48238757	A	G	0.236	0.288	PharmGKB	rs731236		Sequenom	TaqI
rs739837	VDR	12	48238221	G	T	0.557	0.588	Haploview Tagger	rs739837		650Y	
rs757343	VDR	12	48239675	C	T	0.113	0.062	Haploview Tagger	rs757343		650Y	
rs7853934	RXRA	9	137252760	C	T	0.292	0.288	Haploview Tagger	rs7853934		Sequenom	
rs7855881	RXRA	9	137220543	G	T	0.179	0.221	Haploview Tagger	rs7855881		Sequenom	
rs7954412	VDR	12	48237287	T	C	0.085	0.071	Haploview Tagger	rs7967673	1	650Y	LD with SNP on original list
rs7962898	VDR	12	48242837	C	T	0.34	0.336	Haploview Tagger	rs7962898		Sequenom	
rs7965281	VDR	12	48231610	A	G	0.358	0.385	Haploview Tagger	rs7965281		Sequenom	
rs7967152	VDR	12	48244184	A	C	0.623	0.615	Haploview Tagger	rs7967152		650Y	

Continued

rs7968585	VDR	12	48232093	C	T	0.6040.628	Haploview Tagger	rs7968585		650Y	
rs7975232	VDR	12	48238837	C	A	0.6320.628	PharmGKB	rs7967152	0.815	650Y	ApaI, LD with SNP on original list
rs8176348	CYP27B1	12	58157133	G	A	0.1470.146	Haploview Tagger	rs8176348		Sequenom	
rs872298	RXRA	9	137222137	G	A	0.3490.381	Haploview Tagger	rs872298		650Y	
rs886441	VDR	12	48262964	G	A	0.5850.606	Haploview Tagger	rs886441		650Y	