Strength of Linkage Disequilibrium between Two Vitamin D Receptor Markers in Five Ethnic Groups: Implications for Association Studies[1]

Sue Ann Ingles,[2] Robert W. Haile, Brian E. Henderson, Laurence N. Kolonel, Gail Nakaichi, Chen-Yang Shi, Mimi C. Yu, Ronald K. Ross, and Gerhard A. Coetzee

Departments of Preventive Medicine [S. A. I., R. W. H., B. E. H., G. N., M. C. Y., R. K. R.] and Urology [G. A. C.], Norris Comprehensive Cancer Center, University of Southern California School of Medicine, Los Angeles, California 90033; Cancer Research Center of Hawaii, University of Hawaii, Honolulu, Hawaii 96813 [L. N. K.]; and Department of Community, Occupational, and Family Medicine, National University of Singapore, Singapore 0511, Singapore [C-Y. S.]

Abstract

Markers in the 3′ end of the vitamin D receptor gene have recently been associated with prostate cancer risk. To evaluate the adequacy of the commonly used *Bsm*I restriction fragment length polymorphism as a marker of this locus, we genotyped 627 individuals from five ethnic groups for this marker, as well as for a polymorphic site in the 3′ untranslated region of this gene. At the latter site, we identified 12 alleles, A_{13} to A_{24}, of a poly(A) microsatellite. Allele size followed a bimodal distribution with distinct short (A_{13}–A_{17}) and long (A_{18}–A_{24}) allele populations. Poly(A) allele frequency differed by ethnicity, with the frequency of short alleles being highest in non-Hispanic whites (41%), intermediate in Hispanics and African-Americans (31 and 29%, respectively), and lowest in Japanese-Americans and Chinese (8 and 9%, respectively). In each of the ethnic groups, some degree of coupling was observed between *Bsm*I *B* and short poly(A) alleles and between *Bsm*I *b* and long poly(A) alleles. However, the strength of the linkage disequilibrium varied by ethnicity, with departures from complete disequilibrium producing disagreement between the *Bsm*I and poly(A) genotypes. Genotypic disagreement was lowest in Japanese-Americans and non-Hispanic whites (6 and 7%, respectively), intermediate in Chinese and Hispanics (11 and 19%, respectively), and highest among African-Americans (37%), indicating that *Bsm*I is not a good marker for the vitamin D receptor 3′ untranslated region genotype in all populations. This finding may explain contradictory results from recent association studies using the *Bsm*I marker.

Received 9/10/96; revised 10/30/96; accepted 10/30/96.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported by funds provided by the Breast Cancer Fund of the State of California through the Breast Cancer Research Program of the University of California, Grant 1IB-0353 and NIH Research Grants CA54281, CA17054, CA53890, CA65726, and UCLA CNRU CA42710.

[2] To whom requests for reprints should be addressed, at USC/Norris Cancer Center, 1441 Eastlake Ave., MS 44, Room 4415, Los Angeles, CA 90033.

Introduction

The vitamin D receptor, a member of the steroid receptor family, is a transcription factor involved in the regulation of a variety of genes in a wide range of target tissues (reviewed in Refs. 1 and 2). The classical target organs of vitamin D are bone, intestine, and kidney, where it transcriptionally regulates genes involved in calcium and phosphorus homeostasis. Vitamin D also regulates a number of cancer-related genes (reviewed in Ref. 3), including the gene for the cyclin-dependent kinase inhibitor p21 (4) and the metastasis-related fibronectin gene (5). VDR[3]-mediated regulation of these genes is probably responsible for the antiproliferation and prodifferentiation effects of vitamin D seen in a number of cell lines and probably also plays a role in cellular differentiation in tissues that express VDR, including the breast, uterus, ovary, testis, pancreas, thyroid, lung, and spleen. Given this role as a mediator of antiproliferative and prodifferentiative signals, VDR functional variants could possibly contribute to variation in cancer susceptibility.

Functional VDR alleles were recently identified (6) using restriction fragment length polymorphism markers (*Eco*RV, *Bsm*I, *Apa*I, and *Taq*I) that lie in the region from exon 7 to the 3′ UTR. Although these restriction fragment length polymorphisms do not change the VDR coding sequence, and thus are likely to be neutral polymorphisms, they are presumed to mark functional sequence elements that lie nearby, possibly in the VDR 3′ UTR. 3′ UTR regions are known to contain sequence elements that can destabilize mRNAs, and such elements can influence levels of gene expression (reviewed in Refs. 7 and 8). The 3′ UTRs associated with the two most common *Bsm*I/*Apa*I/*Taq*I haplotypes have been reported to result in substantial differences in *VDR* gene expression in a reporter gene assay (9); however, the responsible sequence variants have not yet been identified. The *Bsm*I marker continues to be widely used to mark the *VDR* locus.

At least 16 studies of the *Bsm*I marker and bone density have been reported (reviewed in Ref. 10). The original report that *Bsm*I was associated with bone mineral density among Australian women (9) was confirmed by studies in the United Kingdom (11), in Japan (12), and in one U. S. study of premenopausal women (13). However, a lack of association was reported in numerous other studies, in Sweden, France, Finland, and in at least six U. S. studies. One possible explanation of these discrepant results is that these studies have been done in populations with varying ethnicities and varying degrees of ethnic heterogeneity, and that the *Bsm*I polymorphism may not be an adequate marker of the relevant genotype in all populations. In the Australian, British, American, and Japanese stud-

[3] The abbreviations used are: VDR, vitamin D receptor; UTR, untranslated region; *B*, *Bsm*I uncut allele; *b*, *Bsm*I cut allele; *S*, short poly(A) allele (A_{13}–A_{17}); *L*, long poly(A) allele (A_{18}–A_{24}).

ies, the *Bsm*I *B* allele was associated with low bone density. It is possible, however, that in some ethnic groups the *Bsm*I *b* allele might be in linkage disequilibrium with the at-risk functional allele. If this is the case, then in an ethnically heterogeneous population, the *Bsm*I genotype would not be fully informative for the functional *VDR* locus.

We have recently described a polymorphic microsatellite located approximately 1 kb upstream from the 3′ end of the VDR 3′ UTR (14). The microsatellite consists of a string of adenosine residues with polymorphism in the length of the poly(A) string. In a pilot study, long poly(A) alleles (18 or more adenosines) were associated with prostate cancer in non-Hispanic white men (14). A second study found that a *Taq*I polymorphism, which is tightly linked to the *Bsm*I marker, is also associated with prostate cancer in white men (15). In the present study, we examine the distribution of two VDR markers, *Bsm*I and poly(A), in five ethnic groups: non-Hispanic whites, African-Americans, Hispanics, Japanese-Americans, and Chinese. If the *Bsm*I and poly(A) marker genotypes are not in agreement in one or more ethnic groups, then the utility of these markers for purposes of association studies must be questioned, and further studies will be necessary to evaluate which of these two markers is in disequilibrium with the relevant alleles of the *VDR* locus.

Patients and Methods

Study Subjects. Blood samples were obtained from 627 healthy subjects from three population-based epidemiological studies. From an ongoing multi-ethnic cohort study of environmental exposures and genetic susceptibility to cancer risk being conducted by the University of Hawaii and the University of Southern California (16), blood samples were obtained from 198 African-American (103 male, 95 female), 98 Hispanic (43 male, 55 female), and 66 Japanese-American (37 male, 29 female) subjects. Blood samples from 96 Chinese subjects belonging to the Cantonese dialect group (41 male, 55 female) were obtained from a similar cohort study being conducted in Singapore (Grant R35 CA53890). Blood samples also were available for 169 non-Hispanic white male subjects who comprised the control group of a population-based case-control study of bladder cancer being conducted in Los Angeles County (Grant R01 CA65726). Appropriate informed consent was obtained from subjects in all three of these studies.

VDR 3′ UTR Markers. An 825-bp region of genomic DNA containing the *Bsm*I polymorphic site in intron 8 was amplified and analyzed as described previously (9). The existence of the cut allele, *b*, is indicated by the formation of a 625-bp product.

The region surrounding the poly(A) microsatellite was amplified as described previously (14). The PCR products were separated on 6% polyacrylamide sequencing gels and autoradiographed. Length polymorphisms were scored from smallest to largest, and allele sizes were confirmed by re-running equal sized alleles side-by-side.

Sequencing of Poly(A) Alleles. PCR products from individuals whose alleles were typed previously regarding their relative PCR product sizes were cloned into the pCR II vector using the TA Cloning kit (Invitrogen, San Diego, CA). Forty-one independent clones (covering all sizes) were sequenced using the fmol DNA sequencing system (Promega Corp., Madison, WI). As expected, because the source DNA was initially obtained by PCR, clones of different poly(A) sizes were obtained from the same allele. Clones obtained from the same allele normally differed by one A only and were probably generated during PCR amplification by Taq polymerase. To assign poly(A) sizes to each of the relative allele sizes, we considered not only which poly(A) size was most common among the clones of a given relative size but also which poly(A) size was consistent with a pattern of increasing poly(A) sizes separated by single adenosine residues.

Statistical Methods. Tests for Hardy-Weinberg equilibrium were performed by comparing observed and expected genotype frequencies using a χ^2 test with, where indicated, an exact P as implemented in StatXact-3 (Cytel Software, Cambridge, MA).

Linkage disequilibrium was evaluated using standard likelihood methods (17, 18). The multinomial log likelihood is:

$$\Sigma_{ij} n_{ij} \log(f_{ij}) + n_{22} \log(f_{11} f_{22} + f_{12} f_{21})$$

where $i = 1,2$ and $j = 1,2$ index the two alleles at the *Bsm*I and poly = A loci, respectively, and the f_{ij} are the haplotype frequencies. Because haplotypes cannot be distinguished for doubly heterozygous individuals, the EM algorithm (19) was used to obtain maximum likelihood estimates of the haplotype frequencies and the disequilibrium coefficient. To facilitate comparison among ethnic groups, the standardized disequilibrium coefficient, D_{ij}', was calculated by dividing the raw disequilibrium coefficient, D_{ij}, by its maximum absolute value, D_{max}, where $D_{max} = \min\{p_i(1 - q_j),(1 - p_i)q_j\}$ when $D_{ij} > 0$, or $\min\{p_i q_j,(1 - p_i)(1 - q_j)\}$ when $D_{ij} \leq 0$ and where p_i and q_j are the frequencies of alleles i and j at the *Bsm*I and poly(A) markers, respectively (20).

Results

For the poly(A) microsatellite, 12 allele sizes were identified among the 627 subjects. Sequencing revealed that the alleles differed in the length of the poly(A) (Fig. 1). All alleles sizes between A_{13} and A_{24} were observed. The distribution of allele size differed by ethnicity, and some alleles were unique to a single ethnic group (Fig. 2). The longest alleles, A_{23} and A_{24}, were seen only in African-Americans. and the shortest allele, A_{13}, was seen only in Hispanics.

Because poly(A) allele size follows a bimodal distribution, we categorized the alleles as short (A_{13}–A_{17}) and long (A_{18}–A_{24}). Non-Hispanic whites had the highest frequency of short alleles (41%), and Asians had the lowest (8 and 9% in Japanese-Americans and Chinese, respectively). A similar pattern was seen for *Bsm*I alleles; non-Hispanic whites had the highest frequency of the *B* allele (43%), and Asians had the lowest (10 and 5% in Japanese-Americans and Chinese, respectively). In each of the five ethnic groups, *Bsm*I and poly(A) heterozygote frequencies agreed closely with expected Hardy-Weinberg equilibrium values (Table 1).

Estimated *Bsm*I/poly(A) haplotype frequencies are presented in Table 2 and Fig. 3. In all ethnic groups, coupling was observed between *Bsm*I *B* and short poly(A) alleles and between *Bsm*I *b* and long poly(A) alleles, as indicated by an excess of two haplotypes, *BS* and *bL*, compared to expected frequencies (Table 2). Linkage disequilibrium was nearly complete among non-Hispanic whites and Japanese-Americans, as indicated by *BS* and *bL* haplotype frequencies totaling 97% in each group and standardized disequilibrium coefficients close to 1. Greater departures from complete disequilibrium were seen in other ethnic groups and was most pronounced among African-Americans, with a standardized disequilibrium coefficient of 0.53, intermediate between 1 (complete linkage disequilibrium) and 0 (no linkage disequilibrium; Table 2).

Departure from complete linkage disequilibrium creates disagreement between the two marker genotypes (Table 3). Assuming that *Bsm*I *B* and *b* alleles are in disequilibrium with

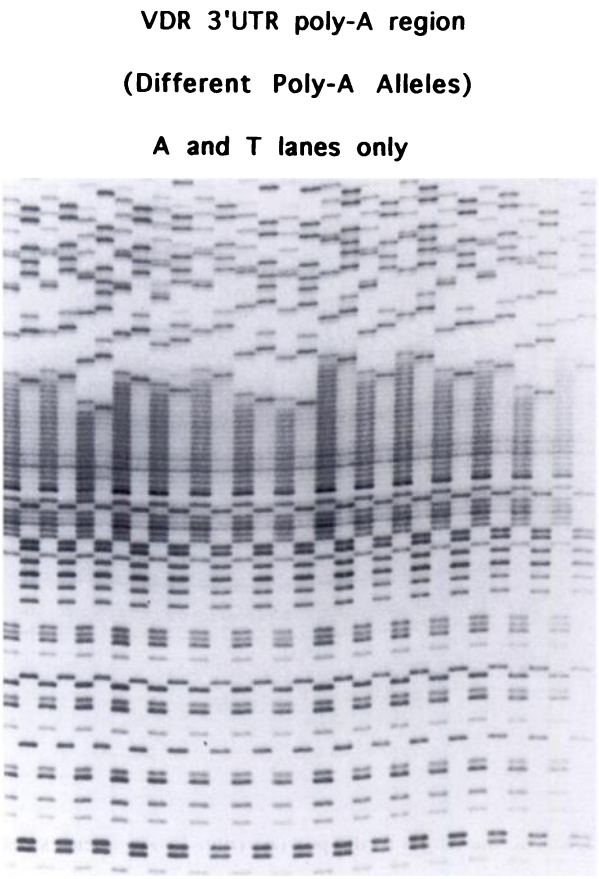

Fig. 1. Sequencing gel of 15 cloned VDR 3′ UTR alleles, A and T lanes only, showing alleles of lengths from A_{15} to A_{24}. All sizes from A_{13} to A_{24} were identified by similar analyses (data not shown).

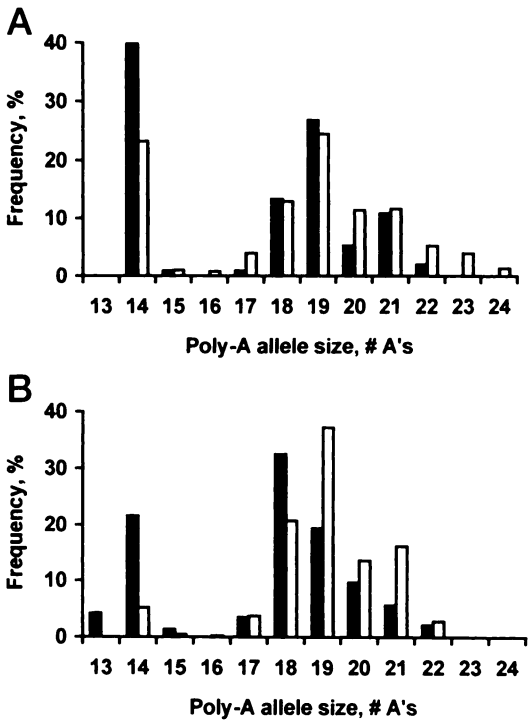

Fig. 2. Frequency distribution of VDR 3′ UTR poly(A) allele sizes by ethnicity. *A*, ■, non-Hispanic whites; □, African-Americans. *B*, ■, Hispanics; □, Asians (Japanese-Americans and Chinese).

the poly(A) short and long alleles, respectively, agreement was 94% for Japanese-Americans and 93% for non-Hispanic whites. Genotypic agreement for Chinese and Hispanic subjects was 89 and 81%, respectively; however, for African-Americans, agreement was only 63%. The nature of the genotypic disagreement in African-Americans is illustrated in Fig. 4. Homozygous *BB* individuals had a substantial frequency of very long alleles, and homozygous *bb* individuals had a non-trivial frequency of very short alleles (7% A_{14}), indicating that the *Bsm*I/poly(A) disagreement is not due to a poor choice of the long/short poly(A) cutpoint.

Discussion

The *Bsm*I *B* allele frequencies reported here (43, 32, 32, 10, and 5% for non-Hispanic whites, African Americans, Hispanics, Japanese-Americans, and Chinese, respectively) are in close agreement with values reported previously. Morrison *et al.* (6) reported a *B* allele frequency of 44% among 182 white subjects. Hustmyer *et al.* (21) reported *B* allele frequencies for whites, blacks, and Asians of 44, 21, and 6%, respectively, based on samples of 85, 19, and 16 subjects, respectively.

We have shown that the poly(A) site is a microsatellite with at least 12 alleles, varying from 13 to 24 adenosine residues. Microsatellites are DNA sequences of one to five nucleotides, which are repeated varying numbers of times and are thought to have arisen due to replication errors related to their repetitive nature. A general model for microsatellite evolution has been proposed that encompasses a bias toward an increase in repeat unit (22). For the poly(A) microsatellite reported here, such growth probably happened in two stages. Long alleles may have evolved from short alleles by a jump in allele size. Subsequent replication error could then produce single base pair differences, causing expansion of the short and long allele populations. Microsatellite expansion rate is related to the number of repeats, with replication error occurring more frequently in longer repeats, consistent with our observation of greater variability in the long allele population. Our choice to dichotomize the 12 poly(A) alleles as short and long is motivated by our belief that these represent two distinct allele populations, and that different alleles within each of the two populations merely reflect minor differences due to microsatellite expansion/contraction.

Only two *Bsm*I/poly(A) haplotypes, *BS* and *bL*, were observed to be common in non-Hispanic whites, and only one haplotype, *bL*, was common in Asians. Thus, the entire region, from the *Bsm*I site upstream of the 3′ UTR to the poly(A), which is approximately 1 kb from the 3′ end, appears to be inherited as a single unit in these populations. Recombination is expected to be a rare event in a chromosomal segment of this length (approximately 3000 bp), because markers that are 1 million bp apart cross-over in 1% of meioses on average.

In the African-American population, in contrast, all four possible haplotypes are commonly observed. One theory that would explain ethnic differences in haplotype frequencies is that all four haplotypes were present prior to the divergence of the races. All non-African populations are thought to have evolved "out-of-Africa" through a series of bottlenecks, with

Table 1 *Bsm*I and poly(A) alleles are in Hardy-Weinberg equilibrium

	Frequency		Heterozygosity		*P* for test of Hardy-Weinberg equilibrium
	Allele *B*	Short alleles	Observed	Expected[a]	
*Bsm*I					
Whites	0.43		0.47	0.49	0.84
African-Americans	0.32		0.49	0.44	0.22
Hispanics	0.32		0.49	0.44	0.42
Japanese-Americans	0.10		0.17	0.18	0.94[b]
Chinese	0.05		0.10	0.10	0.84[b]
Poly(A)					
Whites		0.41	0.45	0.48	0.64
African-Americans		0.29	0.44	0.41	0.65
Hispanics		0.31	0.46	0.43	0.78
Japanese-Americans		0.08	0.17	0.15	0.88[b]
Chinese		0.09	0.15	0.17	0.32[b]

[a] Expected heterozygosity under the assumption of Hardy-Weinberg equilibrium.
[b] Exact χ^2 test.

Table 2 *Bsm*I/poly(A) haplotype frequencies

	Haplotype frequencies[a]				Frequency of haplotypes BS + bL		Disequilibrium coefficient	
	bL	*BS*	*BL*	*bs*	Estimated[a]	Expected[b]	*D*[c]	$D' = D/D_{max}$
Whites	0.56	0.41	0.02	0.01	0.97	0.51	0.23	0.96
African-Americans	0.59	0.20	0.12	0.09	0.79	0.58	0.10	0.53
Hispanics	0.64	0.26	0.05	0.05	0.90	0.57	0.16	0.77
Japanese-Americans	0.89	0.08	0.02	0.01	0.91	0.80	0.07	0.90
Chinese	0.90	0.04	0.01	0.05	0.94	0.86	0.04	0.77

[a] Maximum likelihood estimates based on observed genotype frequencies.
[b] Expected frequencies under the assumption of no linkage disequilibrium.
[c] $D = D_{BS} = D_{bL} = -D_{BL} = -D_{bS}$.

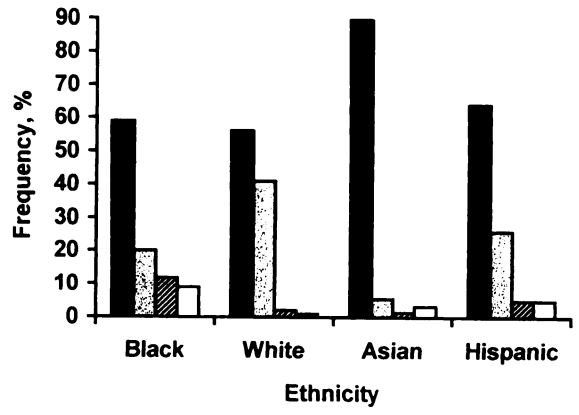

Fig. 3. Estimated *Bsm*I/poly(A) haplotype frequencies. ■, *bL*; □, *BS*; ▨, *BL*; □, *bS*.

only a subset of African alleles being represented in each new population (23). The predominance of two *Bsm*I/poly(A) haplotypes, *BS* and *bL*, in Caucasians and a single haplotype, *bL*, in Asians is consistent with the belief that the Asian populations evolved from early Europeans who, in turn, originated from Africa. Hispanics appear to be intermediate between whites and Asians, consistent with the known ethnic makeup of this population (*i.e.*, part white, part native-American).

The observation of four rather than two *Bsm*I/poly(A) haplotypes in human populations indicates that the two markers are not interchangeable for evaluating putative disease loci in association studies. Although alleles at a disease locus in this region might segregate with the *Bsm*I/poly(A) haplotype within a family or within a homogeneous population stratum, the at-risk haplotype may not be the same in all families. Thus, in an association study, several different at-risk haplotypes might be present in the study population.

As a hypothetical example, consider the case in which the poly(A) itself is the disease locus. If long poly(A) alleles confer increased risk of disease, then the *BL* and *bL* haplotypes confer increased risk, and the haplotypes *BS* and *bS* are "normal." If the *Bsm*I locus is chosen as a marker, misclassification of the disease locus occurs. The effect of nondifferential misclassification in association studies is to bias results toward the null hypothesis. No association or only a weak association may be observed between the *Bsm*I marker and the disease.

When the disease locus is not the poly(A) itself but is tightly linked to the poly(A), use of the *Bsm*I marker will produce similar misclassification and downward bias. Conversely, if the disease locus is tightly linked to the *Bsm*I marker, use of the poly(A) marker will result in misclassification and downward bias. Moreover, for any causal sequence element in this region, estimates of association using these two markers will not be in agreement, because disequilibrium between these two markers is not complete.

It is possible that the negative results in many studies of *Bsm*I genotype and bone density might be explained by atten-

Table 3 Agreement between *Bsm*I and poly(A) marker genotypes

Whites	*SS*	*SL*	*LL*
BB	30	3	0
Bb	2	72	5
bb	0	1	56

Agreement = 158/169 = 93%.
Expected agreement = 38%[a].

African-Americans	*SS*	*SL*	*LL*
BB	5	8	2
Bb	9	56	32
bb	0	23	63

Agreement = 124/198 = 63%.
Expected agreement = 42%[a].

Hispanics	*SS*	*SL*	*LL*
BB	5	2	0
Bb	3	37	8
bb	0	6	37

Agreement = 79/98 = 81%.
Expected agreement = 42%[a].

Japanese-Americans	*SS*	*SL*	*LL*
BB	0	1	0
Bb	0	9	2
bb	0	1	53

Agreement = 62/66 = 94%.
Expected agreement = 71%[a].

Chinese	*SS*	*SL*	*LL*
BB	0	0	0
Bb	1	7	2
bb	1	7	78

Agreement = 85/96 = 89%.
Expected agreement = 76%[a].

[a] Expected agreement under the assumption of no linkage disequilibrium.

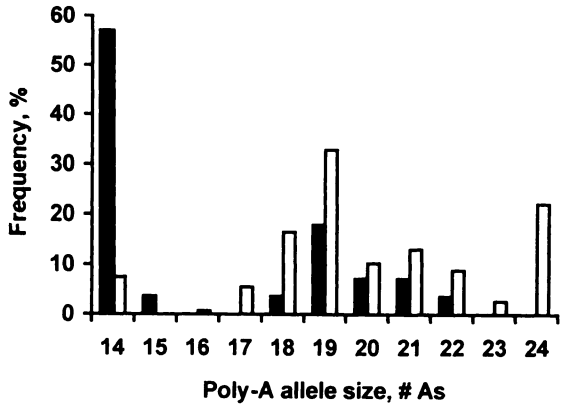

Fig. 4. Frequency distribution of VDR 3′ UTR poly(A) allele sizes in African-Americans by *Bsm*I genotype. ■, *Bsm*I *BB* homozygotes. Long alleles (≥18) are not in agreement with the *Bsm*I locus. □, *Bsm*I *bb* homozygotes. Short alleles (<18) are not in agreement with the *Bsm*I locus.

uation due to misclassification. Although the exact location of the putative at-risk sequence elements is unknown, they may lie within the 3′ UTR. The *Bsm*I site lies upstream of the 3′ UTR, and we have shown that it may not be a good marker of the 3′ UTR itself, as judged by the poly(A) site. Misclassification of 3′ UTR poly(A) alleles by the *Bsm*I marker is most severe in African Americans (37%), minimal among Japanese-Americans and non-Hispanic whites (6 and 7%, respectively), and intermediate (11 and 19%) among Chinese and Hispanics. Thus, even after stratification to control for ethnic differences, residual misclassification may remain and may have contributed to the failure to find an association between genotype and disease status in some, but not all, of the negative studies. Observing an association between genotype and disease status may require adequate exposure of the population to factors that modify the effect of the VDR genotype, such as sunlight or dietary vitamin D. It is possible that in the relatively ethnically homogeneous populations of Sweden and Finland, low sunlight exposure rather than genotype misclassification may have been the major reason for failing to observe an association between genotype and bone density. Resolution of these issues, however, awaits identification of the relevant VDR functional sequence elements. Until these sequence elements have been identified, associations between neutral markers and disease status must be interpreted with caution.

Acknowledgments

We gratefully acknowledge Lucy Xia, Kristine Monroe, and Wu Zhang for expert technical assistance.

References

1. Darwish, H., and DeLuca, H. F. Vitamin D-regulated gene expression. CRC Crit. Rev. Eukaryotic Gene Expression, *3:* 89–116, 1993.

2. Hannah, S. S., and Norman, A. W. 1α25(OH)$_2$ vitamin D$_3$-regulated expression of the eukaryotic genome. Nutr. Rev., *52:* 376–382, 1994.

3. Studzinski, G. P., McLane, J. A., and Uskokovic, M. R. Signaling pathways for vitamin D-induced differentiation: implications for therapy of proliferative and neoplastic diseases. CRC Crit. Rev. Eukaryotic Gene Expression, *3:* 279–312, 1993.

4. Liu, M., Mong-Hong, L., Cohen, M., Bommakanti, M., and Freedman, L. P. Transcriptional activation of the Cdk inhibitor p21 by vitamin D$_3$ leads to the induced differentiation of the myelomonocytic cell line U937. Genes Dev., *10:* 142–153, 1996.

5. Polly, P., Carlberg, C., Eisman, J. A., and Morrison, N. A. Identification of a vitamin D$_3$ response element in the fibronectin gene that is bound by a vitamin D$_3$ receptor homodimer. J. Cell. Biochem., *60:* 322–333, 1996.

6. Morrison, N. A., Yeoman, R., Kelly, P. J., and Eisman, J. A. Contribution of *trans*-acting factor alleles to normal physiological variability: vitamin D receptor gene polymorphisms and circulating osteocalcin. Proc. Natl. Acad. Sci. USA, *89:* 6665–6669, 1992.

7. Sachs, A. B. Messenger RNA degradation in eukaryotes. Cell, *74:* 413–21, 1993.

8. Beelman, C. A., and Parker, R. Degradation of mRNA in eukaryotes. Cell, *81:* 179–183, 1995.

9. Morrison, N. A., Qi, J. C., Tokita, A., Kelly, P. J., Crofts, L., Nguyen, T. V., Sambrook, P. N., and Eisman, J. A. Prediction of bone density from vitamin D receptor alleles. Nature (Lond.), *367:* 284–287, 1994.

10. Peacock, M. Vitamin D receptor gene alleles and osteoporosis: a contrasting view. J. Bone Mineral Res., *10:* 1294–1297, 1995.

11. Spector, T. D., Keen, R. W., Arden, N. K., Morrison, N. A., Major, P. J., Nguyen, T. V., Kelly, P. J., Baker, J. R., Sambrook, P. M., Lanchbury, P. S., and Eisman, J. A. Influence of vitamin D receptor genotype on bone mineral density in postmenopausal women: a twin study in Britain. Br. Med. J., *310:* 1357–1360, 1995.

12. Yamagata, Z., Miyamura, T., Iijima, S., Asaka, A., Sasaki, M., Kato, J., and Koizumi, K. Vitamin D receptor gene polymorphism and bone mineral density in healthy Japanese women. Lancet, *344:* 1027, 1994.

13. Fleet, J. C., Harris, S. S., Wood, R. J., and Dawson-Hughes, B. The *Bsm*I vitamin D receptor restriction fragment length polymorphism (BB) predicts low bone density in premenopausal black and white women. J. Bone Mineral Res., *10:* 985–990, 1995.

14. Ingles, S. A., Ross, R. K., Yu, M. C., Irvine, R. A., La Pera, G., Haile, R. W., and Coetzee, G. A. Association of prostate cancer risk with vitamin D receptor and androgen receptor genetic polymorphisms. J. Natl. Cancer Inst., *89(2):* in press, 1997.

15. Taylor, J. A., Hirvonen, A., Watson, M., Pittman, G., Mohler, J. L., and Bell, D. A. Association of prostate cancer with vitamin D receptor gene polymorphism. Cancer Res., *56:* 4108–4110, 1996.

16. Monroe, K. R., Yu, M. C., Kolonel, L. N., Coetzee, G. A., Wilkens, L. R., Ross, R. K., and Henderson, B. E. Evidence of an X-linked or recessive genetic component to prostate cancer risk. Nat. Med., *1:* 827–829, 1995.

17. Hill, W. G. Estimation of linkage disequilibrium in randomly mating populations. Heredity, *33:* 229–239, 1974.

18. Weir, B. S. Genetic Data Analysis, pp. 89–109. Sunderland, MA: Sinauer Associates, 1990.

19. Dempster, A. P., Laird, N., and Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. J. R. Statist. Soc. Ser. B, *39:* 1–38, 1977.

20. Lewontin, R. C. The interaction of selection and linkage. I. General considerations: heterotic models. Genetics, *49:* 49–67, 1964.

21. Hustmyer, F. G., DeLuca, H. F., and Peacock, M. *Apa*I, *Bsm*I, *Eco*RV and *Taq*I polymorphisms at the human vitamin D receptor gene locus in Caucasians, Blacks and Asians. Hum. Mol. Genet., *2:* 487, 1993.

22. Rubinsztein, D. C., Amos, W., Leggo, J., Goodburn, S., Jain, S., Li, S. H., Margolis, R. L., Ross, C. A., and Ferguson-Smith, M. A. Microsatellite evolution—evidence for directionality and variation in rate between species. Nat. Genet., *10:* 337–343, 1995.

23. Tishkoff, S. A., Dietzsch, E., Speed, W., Pakstis, A. J., Kidd, J. R., Cheung, K., Bonne-Tamir, B., Santachiara-Benerecetti, A. S., Moral, P., Krings, M., Paeabo, S., Watson, E., Risch, N., Jenkins, T., and Kidd, K. K. Global patterns of linkage disequilibrium at the *CD4* locus and modern human origins. Science (Washington DC), *271:* 1380–1387, 1996.

Strength of linkage disequilibrium between two vitamin D receptor markers in five ethnic groups: implications for association studies.

S A Ingles, R W Haile, B E Henderson, et al.

Cancer Epidemiol Biomarkers Prev 1997;6:93-98.