AMERICAN JOURNAL OF INDUSTRIAL MEDICINE 41:457–465 (2002)

# Racial Differences in Prevalence of a Supratypic HLA-Genetic Marker Immaterial to Pre-Employment Testing for Susceptibility to Chronic Beryllium Disease

**Ainsley Weston, PhD,[1]* James Ensey,[1] Kathleen Kreiss,[2] Channa Keshava,[1] and Erin McCanlies[1]**

**Background** *A beryllium materials manufacturer is conducting a limited pilot program that offers testing for HLA-DPβ1$^{E69}$ with genetic counseling through a third party to applicants for employment. An important consideration in this regard is the prevalence of this marker in the general population, and its consequent positive predictive value of disease susceptibility.*

**Methods** *Polymerase chain reaction and restriction fragment length polymorphism analyses were used to determine HLA-DPβ1$^{E69}$ population frequencies. Estimation of positive predictive values assumed a disease frequency among beryllium workers of either 5 or 15% and used an odds ratio for disease risk of 35 for the HLA-DPβ1$^{E69}$ marker.*

**Results** *Allelic/carrier frequencies were found to be 0.21/0.33, 0.24/0.40, 0.27/0.47, and 0.38/0.59 for Caucasians, African-Americans, Hispanics, and Chinese, respectively. Ranges of positive predictive values for a genetic test based on HLA-DPβ1$^{E69}$ in these populations were calculated to be 8.3–14.3% for carriers with an assumed disease frequency of 5%. For high risk subgroups with disease frequencies of 15%, the range of positive predictive values was found to span between 24.9–43.0%.*

**Conclusions** *These estimates suggest that using HLA-DPβ1$^{E69}$ genotyping for general pre-employment screening in the beryllium industry has a low positive predictive value, which varies little among racial groups where carrier frequencies differ significantly.* Am. J. Ind. Med. 41:457–465, 2002. *© 2002 Wiley-Liss, Inc.*

***KEY WORDS: HLA-DP; beryllium disease; RFLP; genetic testing***

## INTRODUCTION

A genetic polymorphism in the human leukocyte antigen (HLA), DPβ1 gene was first reported to be associated with chronic beryllium disease almost a decade ago [Richeldi et al., 1993]. The specific disease marker is called *HLA-DPβ1$^{E69}$* , where E69 refers to genetic variants with a glutamic acid (E) in the 69th position of the protein sequence or the 69th codon of the gene itself. Codon 69 of the *HLA-DPβ1* gene can code for one of the three amino acids, lysine (K), glutamic acid (E), or arginine (R). The lysine variants are found in 60 of the 96 known variants of this gene, glutamic acid is found in 31, while arginine is found in only 5 [Marsh, 1998, 2000; Steiner et al., 1999; Rozmuller et al., 2000; Varney and Tait, 2000; Voorter et al., 2000; Lui et al., 2001; McTernan et al., 2000]. Individuals with chronic beryllium disease are more likely to carry an *HLA-DPβ1* allele with glutamic acid in the 69th position

[1]Health Effects Laboratory Division, National Institute for Occupational Safety and Health, Morgantown, West Virginia
[2]Division of Respiratory Disease Studies, National Institute for Occupational Safety and Health, Morgantown, West Virginia
*Correspondence to: Ainsley Weston, NIOSH-CDC, MS-L3014, 1095 Willowdale Road, Morgantown, WV26505-2888. E-mail: AGW8@CDC.GOV

Accepted 30 January 2002
DOI 10.1002/ajim.10072. Published online in Wiley InterScience (www.interscience.wiley.com)

(*HLA-DPβ1*$^{E69}$) than either a lysine or arginine. Now, five published studies have reported an association of the genetic marker *HLA-DPβ1*$^{E69}$ with susceptibility to disease following occupational exposure to beryllium [Richeldi et al., 1993, 1997; Wang et al., 1999, 2001; Saltini et al., 2001]. Because the strongest association was found with inheritance of E69 as compared with other variant positions, such as position-36 or -55, *HLA-DPβ1*$^{E69}$ is often referred to as a supratypic marker [Richeldi et al., 1993]. These findings indicated that it designates multiple genetic variants that have in common the glutamic acid in the 69th position and that this residue most prominently signals elevated risk of chronic beryllium disease in exposed beryllium workers. A detailed description of *HLA* nomenclature can be found at www.ebi.ac.uk/imgt/allele.html and in Marsh [2000]. In the case of *HLA-DPβ1*, the reference sequence is *HLA-DPβ1*01011*. As new sequences are discovered, the second digit is advanced. These new designations are truncated at the fourth digit (e.g., *HLA-DPβ1*0401*), and the fourth and fifth digits are used to identify highly related sequences (e.g., *HLA-DPβ1*0401* and *HLA-DPβ1*0402* or *HLA-DPβ1*02012* and *HLA-DPβ1*02013*).

Many studies in the literature address genetic variation in *HLA-DP*, but true allele frequencies are difficult to ascertain [al-Daccak et al., 1991; Magzoub et al., 1992; Moonsamy et al., 1992; Wang et al., 1992; Cariappa et al., 1998; May et al., 1998; Poulton et al., 1998; Loudova et al., 1999; Ravikumar et al., 1999; MacHulla et al., 2000; Nishimaki et al., 2000; Rani et al., 1999]. To address this question, we have used a simple polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP)-based assay to determine the frequency and genotypic distribution of *HLA-DPβ1*$^{E69}$ in four different racial populations: African-Americans, Caucasians and Hispanics from the United States, and Chinese from mainland China. While we recognize that race definitions lack precision, especially in a multiracial society like the United States of America, they remain a useful reference for delineating the genetic underpinnings of certain phenotypes [Wood, 2000]. Also, because of the reported powerful association of this marker with risk of beryllium disease in conjunction with occupational beryllium exposure (OR range = 12–76; 95%CIs range = 2–322) (Table I) [Richeldi et al., 1993, 1997; Wang et al., 1999], some advocate its use as a genetic pre-screen to determine suitability for employment in the industry [Richeldi et al., 1993; Lang, 1994; Fields, 2001]. In fact, a pilot pre-employment screening project, based on a test for *HLA-DPβ1*$^{E69}$ carrier status (yes/no), is currently underway in a major beryllium manufacturer. The prudence of this course of action is discussed in light of calculations of the positive predictive value, sensitivity and specificity of this supratypic marker. The debate concerning the use of genetic tests in the employment arena is not new [Williams and Siegel, 1961; Zavon, 1962; Stokinger and Scheel, 1972; Holden, 1982]; however, the application of improved genetic technologies should help to guide progress towards non-prejudicial protection of all workers. Genetic research can support beneficial uses in contrast to testing, and such pertinent approaches are discussed also.

**TABLE I.** Summary of Literature Reports of the Frequency of *HLA-DPβ1*$^{E69}$ in Chronic Beryllium Disease Cases and Controls

| Study | KK[a] | | KE/EE | OR (95%CI) | CF[b] |
|---|---|---|---|---|---|
| Richeldi et al. [1993] | | | | | |
| Cases (33) | 1 | | 32 | 76 (18–322) | |
| Controls (44) | 31 | | 13 | | 0.30 |
| Richeldi et al. [1997] | | | | | |
| Cases (6) | 1 | | 5 | 12 (2–70) | |
| Controls (121) | 85 | | 36 | | 0.30 |
| | **KK** | **KE** | **EE**[c] | | |
| Wang et al. [1999] | | | | | |
| Cases (20) | 1 | 13 | 6 | 23 (5–108) | |
| Controls (75) | 41 | 33 | 1 | | 0.45 |
| Analysis of all published data[d] | **KK** | | **KE/EE** | | |
| Cases (59) | 3 | | 56 | 35 (15–82) | |
| Controls (240) | 157 | | 83 | | 0.35 |

[a]In this case, K represents either lysine (the common allele) or arginine (the rare allele).
[b]Carrier frequency (i.e., glutamic acid homozygotes and heterozygotes combined).
[c]Hardy–Weinberg equilibrium for this genotypic distribution: $\chi^2 = 3.96$, $P = 0.05$.
[d]Does not include Saltini et al. [2001]; n = 22 cases, 93 controls, OR = 3.7 (95%CI = 1.4–10.0).

## MATERIALS AND METHODS

### Human Subjects

Collections of blood-derived constitutive DNA samples from 100 unrelated African-Americans and 100 unrelated Caucasians were purchased from the Coriell Institute, Camden, NJ; these two collections were specifically developed to facilitate characterization of genetic polymorphisms. A collection of DNA from 100 unrelated Chinese was a gift from Dr. Tong-man Ong, NIOSH-CDC, Morgantown, WV; and a collection of DNA samples from 100 unrelated non-black Hispanic women, principally of Caribbean descent, enrolled in a New York City breast cancer case-control study, was also available to us through a collaboration with Dr. Mary Wolff (Mount Sinai Medical Center, NYC) [Weston et al., 1997; Wolff et al., 2000]. The protocol, including statement of purpose for obtaining these DNA samples, was given expedited review by the NIOSH Human Studies Review Board because no personal identifiers were involved. Its own Institutional Review Board approves the Coriell Institute for such studies. Approval of the Mount Sinai Medical Center Institutional Review Board was also obtained. Overall quality of DNA was assessed by determining: $OD^{260/280}$ (for purity, with an expected ratio between 1.6 and 1.9), agarose gel-electrophoresis characteristics (for high molecular weight, with a narrow band expected $>30$ kb) and facile amplification of a $>300$-bp fragment by PCR. Samples that did not meet these criteria were repurified; high molecular weight DNA was redissolved in Tris:EDTA (10 mM:1 mM) and stored at $-70^{\circ}$C; and working solutions for PCR amplification (10 μg/ml, containing Tris-Cl 10 mM and EDTA 1 mM) were prepared and stored at 0–4°C.

### Genetic Analysis of the Polymorphism at Amino Acid Residue 69 in Exon 2 of HLA-DPβ1

Analysis of white blood cell DNA samples from African-Americans, Caucasians, Chinese and Hispanics ($n=100$ for each group) was conducted to determine the *HLA-DPβ1* genotypes that code for the 69th amino acid (codon 69). The PCR was used to generate fragments of *HLA-DPβ1* that were 322 bp in length. Amplification of human constitutive, genomic DNA was performed in a PCR reaction mixture (50 μl) containing: primers (FDPI1-5′-GAGAGTGGCGCCTCC-3′ and RDPI2-5′-CCCAAAGCC CTCACTC-3′, 2 pM; these primers flank exon 2 of *HLA-DPβ1*), genomic DNA (100 ng), $MgCl_2$ (1.25 mM), dNTPs (0.8 mM), and Taq-polymerase (1 U, Perkin-Elmer Applied Biosystems, Foster City, CA). Amplification was allowed to proceed with an initial melt (95°C, 5 min), then 35 cycles of melting (95°C, 30 s), annealing (60°C, 35 s) and extension (72°C, 60 s), and a final extension at 72°C (10 min). The resulting PCR product (322 bp) was subjected to digestion with *BsrBI* (cleaves GAGCGG, where GAG codes for glutamic acid in position 69 [E69] and CGG (arginine) is invariant in position 70; does not cleave AAGCGG or AGGCGG, where AAG codes for lysine 69 [K69] and AGG codes for arginine 69 [R69]), and where glutamic acid at codon 68 is invariant (Fig. 1).

Adequacy of PCR was determined by electrophoretic analysis of a portion (20%) of the reaction mixture (2% agarose gels). Digestion of a further portion (20%) of the amplicon with *BsrBI* was conducted in a reaction mixture (16.4 μl) according to the manufacturer's instructions (New England Biolabs, Beverly, MA). Restriction fragments (322 bp uncut, K69 or R69, 88/234 bp cut or E69) were resolved by agarose gel (3%) electrophoresis.

### Statistical Analysis

Exact methods and the Chi-square test were used to compare allele frequencies for $HLA\text{-}DP\beta1^{E69}$ among four racial groups, African-Americans, Caucasians, Chinese and Hispanics. The Chi-square test was also used to determine if the genotypic distributions conformed to Hardy–Weinberg population laws. SAS statistical software was used to conduct all statistical analyses [SAS Institute, 1999]. Gene frequencies were estimated using standard methods, where the allelic frequency is the proportion of chromosomes in the population that harbor a specific gene variant [Russell, 1992].

Positive predictive values, sensitivity and specificity were estimated according to the method described by Khoury et al. [1985, 1993]. However, odds ratios based on analysis of previously published reports, were substituted for relative risks. First, an odds ratio of 35 for carriers was used to calculate positive predictive values for a range of $HLA\text{-}DP\beta1^{E69}$ carrier frequencies [Richeldi et al., 1993, 1997; Wang et al., 1999] (Table I). The calculations are also based on the highest estimate (5%) of beryllium disease among generally exposed workers (1–5%), and 15% for high risk subgroups (workers performing tasks such as machining and lapping) [Kreiss et al., 1996]. Caution should be used in interpretation of the calculations for the populations studied here because they assume disease frequency does not vary by race [Kreiss et al., 1996]. This strategy provides for finding the highest positive predictive values, that is, it is expected that the positive predictive value will be over-estimated. Second, to counter-balance this approach, a comparison of positive predictive values for an OR of 3, over the range of $HLA\text{-}DP\beta1^{E69}$ carrier frequencies, was also calculated. This strategy was adopted because obtaining an odds ratio as high as 35 is unusual in most studies; and while 3 might still be considered high, it might better reflect the odds ratios of other disease/gene associations. Summary of formulae: Positive predictive value $= Rp/1 +$

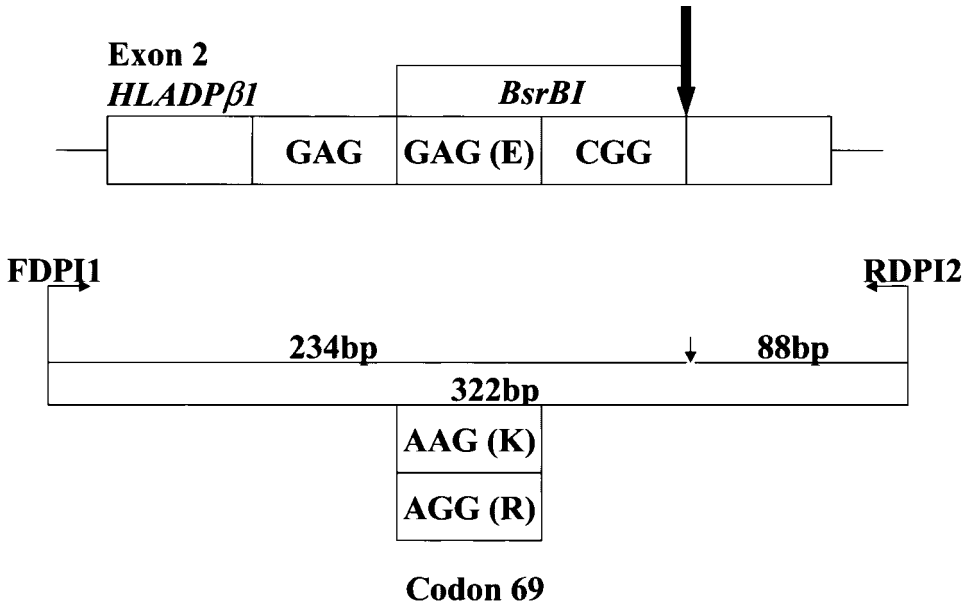


**FIGURE 1.** Depiction of exon 2 of *HLA-DPβ1* indicating the relative positions of codon 69, the restriction enzyme recognition site and polymerase chain reaction (PCR) primer reagents used to amplify the DNA segment. This PCR and restriction fragment length polymorphism (RFLP) analysis strategy was used to determine the frequency of the *HLA-DPβ1*$^{E69}$ supratypic marker in four racial populations. Exon 2 was amplified using primers (FDPI1 and RDPI2) that flank the exon, and which generated 322-bp fragments for each sample. Then *BsrBI* was used to interrogate each amplicon. Amplicon cleavage by *BsrBI* occurred for those samples containing the E69, resulting in fragments of 88 and 234 bp. Amplicons containing K or R69 remain uncleaved. Isoschizomers *AccBSI* and *MbiI* would be expected to produce a similar result.

$g(R-1)$, where R = odds ratio substituted for relative risk, p = disease frequency, and g = genetic trait frequency. Sensitivity $= Rg/1+g(R-1)$, and Specificity $= 1-(g \times [1-\{Rp\}/\{1+g\lceil R-1\rceil\} \div \{1-p\}])$ [Khoury et al., 1985, 1993]. Thus, positive predictive value is the proportion of individuals who develop disease among those who have the trait.

## RESULTS

Results for six representative DNA samples are shown in Figure 2. The enzyme, *BsrB1* cleaves the PCR product into two pieces (88 and 234 bp) if the sequence codes for E69. The 322-bp product remains intact if it codes for K69 or R69. Reference to the molecular weight marker (lane M)

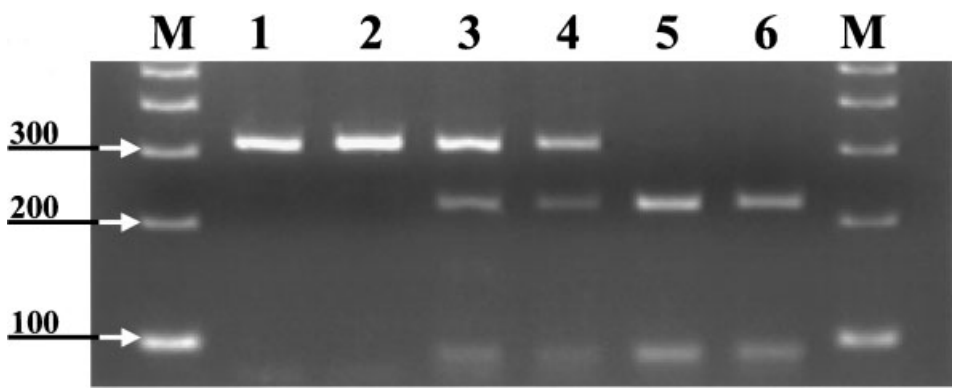


**FIGURE 2.** Typical results of PCR-RFLP analysis designed to determine the presence or absence of the *HLA-DPβ1*$^{E69}$ supratypic marker in human, constitutive, genomic DNA (according to the scheme described in the legend to Fig. 1 and the Materials and Methods section of the text). Analysis of six different samples, from four collections of 100 unrelated African-Americans, Caucasians, Chinese and Hispanics is shown. Individual samples were amplified with primers specific for DNA sequences flanking exon 2 of *HLA-DPβ1* (FDPI1 and RDPI2). Amplified materials were restricted with *BsrBI* and analyzed on agarose gels: lanes marked **M** are a molecular weight marker (100 base pair DNA ladder, as indicated), **lanes 1** and **2**, homozygous restriction site absent (KK, KR, or RR), **lanes 3** and **4**, heterozygous EK or ER, **lanes 5** and **6**, homozygous EE.

**TABLE II.** Genotypic Distribution of *HLA-DPβ1*$^{E69}$ in Four Racial Groups*

| Population | KK[a] | KE | EE | F[b] | CF[c] | HWE[d] *P*-value |
|---|---|---|---|---|---|---|
| African-American | 60 | 32 | 8 | 0.24 | 0.40 | 0.22 |
| Caucasian | 67 | 25 | 8 | 0.21 | 0.33 | 0.02 |
| Chinese | 41 | 43 | 16 | 0.38 | 0.59 | 0.41 |
| Hispanic | 53 | 40 | 7 | 0.27 | 0.47 | 0.88 |

*N = 100 per group.
[a]In this case, K represents either lysine (the common allele) or arginine (the rare allele).
[b]Frequency of the alleles containing a glutamic acid residue at codon 69 (exon 2) of *HLA-DPβ1*.
[c]Carrier frequency (i.e., glutamic acid homozygotes and heterozygotes combined).
[d]Hardy–Weinberg equilibrium ($\chi^2$ not shown).

shows two lysine/arginine homozygotes (lanes 1 and 2), two glutamic acid homozygotes (lanes 5 and 6), and two heterozygotes (lanes 3 and 4). This assay did not discriminate between the lysine and arginine alleles; however, the goal was to determine the genotypic distribution with respect to the glutamic acid variant. No unexpected bands were seen.

Genotyping data for each sample collection is shown in Table II. Among African-Americans, the allelic frequency for *HLA-DPβ1*$^{E69}$ was determined to be 0.24, with the expected number of heterozygotes and *HLA-DPβ1*$^{E69}$ homozygotes (Hardy–Weinberg $\chi^2 = 1.51$, $P = 0.22$). Similarly for Chinese (F = 0.38) and Hispanics (F = 0.27), there was no departure from Hardy–Weinberg population laws (Chinese $\chi^2 = 0.68$, $P = 0.41$, and Hispanics $\chi^2 = 0.02$, $P = 0.88$). Among Caucasians, where the allelic frequency was 0.21, an excess of *HLA-DPβ1*$^{E69}$ homozygotes was observed (0.08 vs. 0.04; $\chi^2 = 5.43$, $P = 0.02$).

When the frequency of *HLA-DPβ1*$^{E69}$ was compared across the ethnic groups, there was essentially no difference in the allelic frequency or genotypic distribution between African-Americans and Caucasians (Table III), with each group having a carrier frequency of 0.40 and 0.33, respectively. However, Chinese had a significantly higher *HLA-DPβ1*$^{E69}$ allelic frequency (0.38) and concomitantly, the expected higher carrier frequency (0.59) compared to both Caucasians and African-Americans. When this genotypic distribution was compared with that in Caucasians, the difference was found to be statistically significant ($\chi^2 = 13.69$, $P = 0.001$) (Table III). Similarly, the difference in allelic distribution between Chinese and African-Americans was statistically significant ($\chi^2 = 7.85$, $P = 0.02$). For Hispanics, in which the allelic frequency falls between Chinese and African-Americans, there was no significant difference (Table III). However, their carrier frequency was signifi-

**TABLE III.** Comparison of *HLA-DPβ1*$^{E69}$ Genotypic Distributions Among Four Racial Groups

| | | Genotype | | Carrier frequency | |
|---|---|---|---|---|---|
| Comparison groups | F[a] | $\chi^2$ | *P*-value[b] | $\chi^2$ | *P*-value[c] |
| African-American | 0.24 | | | | |
| Caucasian | 0.21 | 1.25 | 0.54 | 1.06 | 0.30 |
| African-American | 0.24 | | | | |
| Chinese | 0.38 | 7.85 | 0.02 | 7.22 | 0.01 |
| African-American | 0.24 | | | | |
| Hispanic | 0.27 | 1.39 | 0.50 | 1.00 | 0.32 |
| Caucasian | 0.21 | | | | |
| Chinese | 0.38 | 13.69 | 0.001 | 13.61 | 0.001 |
| Caucasian | 0.21 | | | | |
| Hispanic | 0.27 | 5.16 | 0.08 | 4.08 | 0.04 |
| Chinese | 0.38 | | | | |
| Hispanic | 0.27 | 5.16 | 0.08 | 2.89 | 0.09 |

[a]Frequency of the alleles containing a glutamic acid residue at codon 69 (exon 2) of *HLA-DPβ1*.
[b]df = 2.
[c]df = 1.

**TABLE IV.** Estimated Positive Predictive Value, Sensitivity, and Specificity of *HLA-DPβ1*$^{E69}$ for Chronic Beryllium Disease

| *HLA-DPβ1*$^{E69}$ carrier frequency | Disease prevalence = 5% | | | | Disease prevalence = 15% | | | |
|---|---|---|---|---|---|---|---|---|
| | PPV[a] | Spec. | Sens. | NPV | PPV | Spec. | Sens. | NPV |
| 1 | 100 | 100 | 26 | 96 | 100 | 100 | 26 | 86 |
| 10 | 40 | 94 | 80 | 99 | 100 | 100 | 80 | 94 |
| 30 | 16 | 73 | 94 | 99 | 47 | 81 | 94 | 98 |
| 40 | 12 | 63 | 96 | 100 | 36 | 70 | 96 | 99 |
| 60 | 8 | 42 | 98 | 100 | 25 | 49 | 98 | 100 |

[a]PPV, positive predictive value; Spec., specificity; Sens., sensitivity; NPV, negative predictive value (Spec. $\times$ [1 − p]/1 − g) [Khoury et al., 1993]. All values are percentages, calculated using an odds ratio of 35.

cantly different when compared to the frequency among Caucasians ($\chi^2 = 4.08$, $P = 0.04$) (Table III).

A range of carrier frequencies, encompassing those found for the four racial groups studied here, both low (5%) and high (15%) prevalence of disease, and data generated by others on beryllium disease cases and controls (Table I), were used to estimate the positive predictive value of *HLA-DPβ1*$^{E69}$ for development of beryllium disease (Table IV). In the case of each racial group, the positive predictive value of *HLA-DPβ1*$^{E69}$ for carriers among beryllium workers in general is quite low, falling between 8.3 and 14.3% (Fig. 3a and Table IV). In Chinese, especially, where the carrier frequency is highest (0.59), assuming an OR for disease of 35, the positive predictive value is only 8.3%, with a specificity of 43% indicating that the test would yield misleading information more often than not (57%). When the calculations were extended to a disease prevalence of 15% (such as might be associated with workers in high risk tasks, e.g., machining and lapping), the findings were still unimpressive. Thus, for the range of carrier frequencies found here, the range of positive predictive values was 24.9–43.0% (Table IV and Fig. 3a). Similarly, the specificity of the test was found to be quite low, again implying that many people who test positive will not develop disease. Correspondingly, the negative predictive value was found to be high (Table IV).

Evident in Figures IIIb and IIIc, is the relative lack of impact that a reduction in OR has on positive predictive value and the diminution of its effect with increasing population frequency of the genetic trait. This would also be the case had it been possible to more correctly use a relative risk estimate. For a 5% disease prevalence, at the point when the carrier frequency reaches 33% this impact, from a policy perspective, is almost negligible on the positive predictive value, which ranges between 9.0% for an OR of 3 and 14.3% for an OR of 35 (Fig. 3a,b). At a carrier frequency of 59%, the range is 6.9–8.3% for the same ORs and disease prevalence. Similarly, for a disease prevalence of 15%, these ranges are 27.0–43.0% and 20.6–24.9%, respectively (Fig. 3a,c).

## DISCUSSION

The rationale for this study centers on the ethical issues involved in the use of genetic markers of susceptibility for pre-employment screening [Richeldi et al., 1993; Lang, 1994; Fields, 2001]. Five published studies have previously implicated inheritance of HLA haplotypes characterized by the supratypic marker *HLA-DPβ1*$^{E69}$ as a significant risk factor for the development of a debilitating granulomatous lung disease in beryllium exposed workers [Richeldi et al., 1993, 1997; Wang et al., 1999, 2001; Saltini et al., 2001]. Specifically, pertaining to the case of *HLA-DPβ1*$^{E69}$ and beryllium disease, there are many deficiencies that weigh against its adoption for pre-employment screening. These include limitations of the published data, sensitivity, specificity, and poor positive predictive value.

Although high odds ratios for disease susceptibility were reported (OR range = 12 − 76; 95%CIs range = 2 − 108), only small numbers of cases were studied (33, 6, 20, and 22, respectively)[1] [Richeldi et al., 1993, 1997; Wang et al., 1999, 2001; Saltini et al., 2001]. Three of the studies did not distinguish between homozygotes and heterozygotes (copy number: two or one) [Richeldi et al., 1993, 1997; Saltini et al., 2001], whereas the most recent study that used detailed or allele-specific DNA sequence analysis suggested that homozygotes were 10–15 times more likely to develop disease compared to heterozygotes [OR > 200 estimated from Wang et al., 1999, 2001]. In this technically superior study of twenty cases (75 controls), the genotypic distribution of *HLA-DPβ1*$^{E69}$ among controls did not conform to Hardy–Weinberg laws, arising from a deficiency of homozygotes [$P < 0.05$, calculated from Wang et al., 1999]. This result suggests perhaps a winnowing of the high-risk alleles from the control population. Taken together, these factors suggest that control carrier frequencies from these studies [which were 0.30, 0.30, 0.45, and 0.40, reported by

[1] The report by Saltini et al., 2001 appeared while the present manuscript was being reviewed and the data have not been included in the analysis (Table I). Inclusion of this study would result in reducing the positive predictive values further.

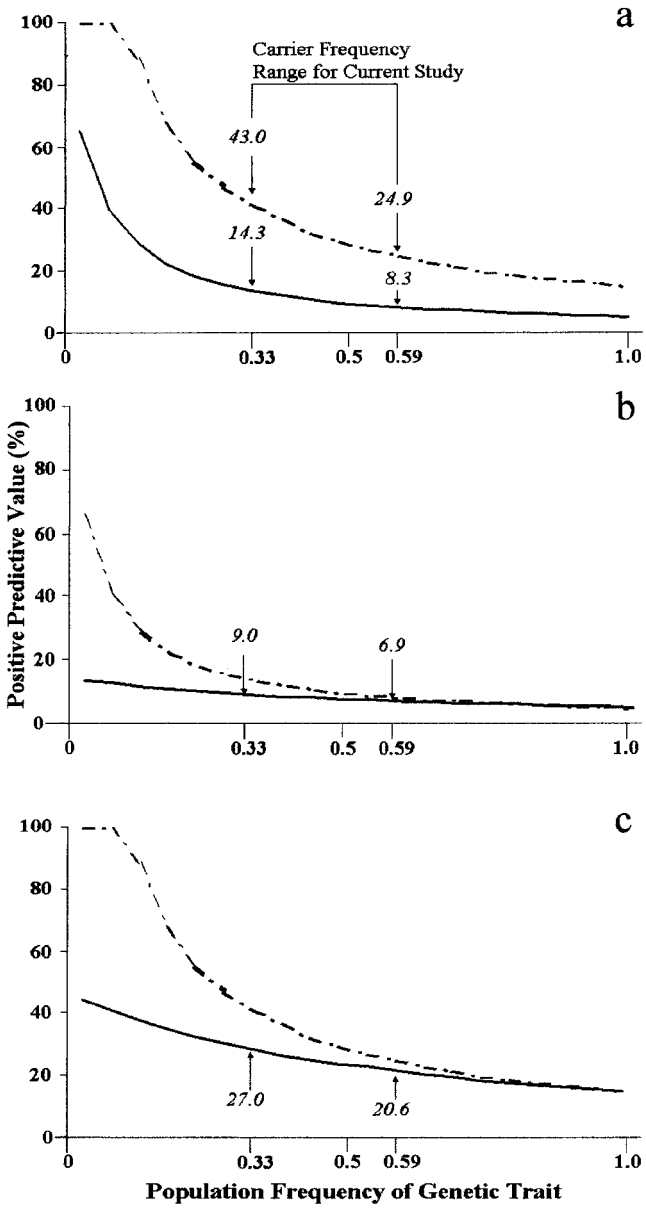


**FIGURE 3.** Graphical depictions of the positive predictive value of tests for *HLA-DPβ1*$^{E69}$. These curves are based on: (a) an odds ratio for carriers of 35 and prevalences of disease in beryllium-exposed workers of 5% (general, ———) and 15% (high-risk jobs, — – —); (b) generally exposed beryllium workers (disease prevalence 5%) and odds ratios for carriers of 3 (calculated from Richeldi et al., 1997, ———) and 35 (— – —); and (c) the sub-set of beryllium workers in high risk jobs (disease prevalence 15%) and odds ratios for carriers of 3 (calculated from Richeldi et al., 1997, ———) and 35 (— – —). The range of carrier frequencies encompasses those in the different racial populations studied here.

Richeldi et al., 1993, 1997; Wang et al., 1999; Saltini et al., 2001, respectively] may be low or otherwise inappropriate for evaluation of the positive predictive value, sensitivity, and specificity of the test. Moreover, some previous efforts to assess the positive predictive value of *HLA-DPβ1*$^{E69}$ testing appear to overlook the true prevalence of disease in the test group [Richeldi et al., 1993; Wang et al., 1999]. For example, in their previous study, Richeldi et al. [1993] compared allele frequencies in 33 cases (43%) and 44 controls; however, the cases and controls were geographically distinct. Similarly, Wang et al. [1999] studied 75 controls and 20 cases whose beryllium exposure occurred in Los Alamos, NM, the Rocky-Flats beryllium plant in the Denver, CO area, and other unspecified areas of the United States [Wang et al., 2001]. In contrast, the study by Richeldi et al. [1997] focused on workers from a beryllium-ceramics plant in the western United States of America. It was a population-based study and demonstrated a disease prevalence of 5%. Population-based predictive values are currently only deliverable from the study of workers at this beryllium plant [Richeldi et al., 1997].

The new evidence presented here shows that carrier frequencies for *HLA-DPβ1*$^{E69}$ are in general higher than previously reported, and vary between racial groups. With respect to the departure from Hardy–Weinberg equilibrium in Caucasians, we have compared these *HLA-DPβ1*$^{E69}$ data with genotype data for Caucasians ($n = 836$) enrolled in an ongoing molecular epidemiologic study of chronic beryllium disease. For the molecular epidemiologic study population, the allele frequency is almost exactly the same (0.22), but there is no departure from Hardy–Weinberg population laws ($\chi^2 = 1.88$, $P = 0.17$). Consequently, the carrier frequency among this latter Caucasian group is slightly higher (0.39 vs. 0.33 reported here).

Khoury et al. [1985] eloquently discussed and elegantly demonstrated how the interplay between disease and allele frequency affects the positive predictive value, sensitivity, and specificity of a genetic marker. Furthermore, Khoury et al. [1985] illustrated how these factors must be evaluated and considered prior to recommending use of a genetic marker as a screening tool. Specifically, as the marker frequency increases in a population, the positive predictive value will decrease even if the relative risk is high [Khoury et al., 1993]. Moreover, a reduced specificity accompanies a poor positive predictive value. As a consequence, a relatively large proportion of the population who does not have disease will be advised that the genetic test outcome is unfavorable with respect to their potential disease risk, or their employment potential in a particular industry. The *HLA-DPβ1*$^{E69}$ marker and chronic beryllium disease are an excellent example of these principles. Based on our results of the prevalence of *HLA-DPβ1*$^{E69}$ in the general population and published chronic beryllium disease rates among beryllium worker populations, the positive predictive value was low (8.3–14.3%), while the proportion of the population identified as positive for the marker, but without disease was high (at least 27–58%, derived from Table IV, i.e., 100–73 = 27 and 100–42 = 58 [Richeldi et al., 1993, 1997; Wang et al., 1999]). The concomitant specificity of this preemployment screening test is low (in the order of 60%).

These findings are consistent with those of Bartell et al. [2000] who concluded that the test had a general positive utility in CBD avoidance and cost savings, but that implementation of a genetic testing program needed to weigh the ethical, legal, and social issues as well. However, their findings were based solely on the results of one study that implicated the *HLA-DPβ1*0201* glutamic acid-containing variant [Richeldi et al., 1993]. More recently, certain of the non-**0201* glutamic acid containing variants have been proposed to be more strongly associated with CBD risk than the **0201* variants [Wang et al., 1999]. Even more recently, the specific importance of the supratypic *HLA-DPβ1*$^{E69}$ marker has been reasserted [Saltini et al., 2001], and this question remains unresolved. This dispute underscores the need for more carefully designed studies from both a molecular and epidemiologic perspective.

This simple example of potential misclassification based on *HLA-DPβ1*0201*/non-**0201* variants belies the true complexity of the *HLA-DPβ1*$^{E69}$ chronic beryllium disease association. Two other major issues require further consideration. First, allelic sub-sets typified by *HLA-DPβ1*$^{E69}$ may convey different levels of risk, e.g., *HLA-DPβ1*0201 <*1901<*1301<*0901 <*1001<*0601<*1701* [Wang et al., 1999]. These investigators subsequently demonstrated a similar trend in a group of 25 sensitized beryllium workers [Wang et al., 2001], although no evidence for an association between *HLA-DPβ1*$^{E69}$ and sensitization has been proffered by others [Saltini et al., 2001]. These issues may be difficult to completely resolve since large sample numbers will be required to test hypotheses about the many alleles of *HLA-DPβ1*$^{E69}$. Second, homozygosity could have profound significance for the positive predictive value since homozygosity is much less common than simple carrier status. As previously noted, *HLA-DPβ1*$^{E69}$ homozygotes may be at significantly higher risk than heterozygotes [Wang et al., 1999]. This important question may be more easily addressed than that of multiple *HLA-DPβ1*$^{E69}$ alleles.

Estimates of positive predictive value will increase as longitudinal follow-up allows for incident data to be determined [Newman, 1993]. Similarly, as specific sub-types or combinations of sub-types are identified, these too might lead to better clinical predictivity, and therefore, form the basis for a more accurate test. In the pilot pre-employment screening program, the beryllium materials manufacturer considered odds ratio estimates of 5–10 derived from research studies to be potentially significant to a person considering accepting employment with potential for exposure to beryllium particles. A third party contractor conducted pre- and post-test genetic counseling and testing, with provisions to ensure that the employer would not know which applicants chose to undergo the testing/counseling process and could not identify the genetic test result in any individual. Estimates of positive predictive value, presented here, show that odds ratios alone may convey a misleading perspective of risk. Justification for pursuing susceptibility testing should consider not only estimates of relative and absolute risks, but also positive and negative predictive value. Likewise, full risk counseling should include all of these parameters as well.

Chronic beryllium disease illustrates that the positive predictive value of tests for common genetic traits (with population frequency $> 30\%$) may be low in-spite of high odds of disease. Low positive predictive values minimize the effect of trait distribution among racial and ethnic groups, and consequently the utility of genetic testing based on common traits. If rare genetic traits are shown to have high positive predictive value for disease, legal, ethical, and social protection must evolve to ensure protection of groups characterized by increased risk of both disease and discrimination. The beneficial public health impact of genetic research that investigates common traits is most likely not through genetic testing, but rather the development of better protection strategies, followed by their implementation. Molecular epidemiologic association studies can identify genetic factors with which to develop transgenic animal models; in turn such models will be used to study disease mechanism, and investigate the efficacy of post-exposure interventions. Molecular epidemiologic investigations that include sound industrial hygiene and reliable exposure assessment matrices will help to elucidate gene–environment interactions. These latter studies are critical to meaningful risk assessment (genome-based risk assessment) that is needed for the development of exposure standards that protect all workers.

## ACKNOWLEDGMENTS

We thank Dr. Muin Khoury for helpful discussions. The thoughtful comments of Drs. Gregory Wagner and Cecil Burchfield are gratefully appreciated. Thanks also to Luke Thomas and Teresa Cummins for administrative support.

## REFERENCES

al-Daccak R, Wang FQ, Theophille D, Lethielleux P, Colombani J, Loiseau P. 1991. Gene polymorphism of HLA-DPB1 and DPA1 loci in caucasoid population: frequencies and DPB1–NDPA1 associations. Hum Immunol 31:277–285.

Bartell SM, Ponce RA, Takaro TK, Zerbe RO, Omenn GS, Faustman EM. 2000. Risk estimation and value of information analysis for three proposed genetic screening programs for chronic beryllium disease prevention. Risk Anal 20:87–99.

Cariappa A, Sands B, Forcione D, Finkelstein D, Podolsky DK, Pillai S. 1998. Analysis of MHC class II DP, DQ and DR alleles in Crohn's disease. Gut 43:210–215.

Fields S. 2001. Toxic beryllium: new solutions for a chronic problem. Environ Health Perspect 109:A74–A79.

Holden C. 1982. Looking at genes in the workplace. Science 217:336–337.

Khoury MJ, Newill CA, Chase GA. 1985. Epidemiologic evaluation of screening for risk factors: application to genetic screening. Am J Public Health 75:1204–1208.

Khoury MJ, Beaty TH, Cohen BH. 1993. Applications of genetic epidemiology in medicine and public health. In: Fundamentals of genetic epidemiology—Monographs in epidemiology and biostatistics, Vol 22. New York, Oxford: Oxford University Press. p 312–339.

Kreiss K, Mroz MM, Newman LS, Martyny J, Zhen B. 1996. Machining risk of beryllium disease and sensitization with median exposures below 2 $\mu g/m^3$. Am J Ind Med 30:16–25.

Lang L. 1994. Beryllium: a chronic problem. Environ Health Perspect 102:526–531.

Loudova M, Sramkova I, Cukrova V, Sieglova Z, Brdicka R. 1999. Frequencies of HLA-DRB1, -DQB1 and -DPB1 alleles in Czech population. Folia Biologica 45:27–30.

Lui Z, Lin J, Chen W, Jia Z, Pan D, Xu A. 2001. Sequence of complete exon 2 and partial intron 2 of HLA-DPB1*8001 allele. (GenBank accession # AF336231).

MacHulla HK, Schonermarck U, Schaaf A, Muller LP, Kloss C, Kruger J, Kunze G, Schonermarck G, Langner J. 2000. HLA-A, B, Cw and DRB1, DRB3/4/5, DQB1, DPB1 frequencies in German immunoglobulin A-deficient individuals. Scand J Immunol 52:207–211.

Magzoub MM, Stephens HA, Sachs JA, Biro PA, Cutbush S, Wu Z, Bottazzo GF. 1992. HLA-DP polymorphism in Sudanese controls and patients with insulin-dependent diabetes mellitus. Tissue Antigens 40: 64–68.

Marsh SG. 1998. Nomenclature for factors of the HLA system, update March 1998. WHO Nomenclature Committee for Factors of the HLA System. Tissue Antigens 51:681–683.

Marsh SG. 2000. Nomenclature for factors of the HLA system, update September 2000. WHO Nomenclature Committee for Factors of the HLA System. Tissue Antigens 56:565–566.

May J, Mockenhaupt FP, Loliger CC, Ademowo GO, Falusi AG, Jenisch S, Dippmann K, Schnittger L, Kremsner PG, Bienzle U, Meyer CG. 1998. HLA DPA1/DPB1 genotype and haplotype frequencies, and linkage disequilibria in Nigeria, Liberia, and Gabon. Tissue Antigens 52:199–207.

McTernan CL, Mijovic CH, Cockram CS, Barnett AH. 2000. The nucleotide sequence of two new DP alleles, DPA1*02015 and DPB1*8401, identified in a Chinese subject. Tissue Antigens 17: 661–666.

Moonsamy PV, Suraj VC, Bugawan TL, Saiki RK, Stoneking M, Roudier J, Magzoub MM, Hill AV, Begovich AB. 1992. Genetic diversity within the HLA class II region: ten new DPB1 alleles and their population distribution. Tissue Antigens 40:153–157.

Newman LS. 1993. To $Be^{2+}$ or not to $Be^{2+}$: immunogenetics and occupational exposure. Science 262:197–198.

Nishimaki K, Kawamura T, Inada H, Yagawa K, Nose Y, Nabeya N, Isshiki G, Tatsumi N, Niihira S. 2000. HLA DPB1*0201 gene confers disease susceptibility in Japanese with childhood onset type I diabetes, independent of HLA-DR and DQ genotypes. Diabetes Res Clin Pract 47:49–55.

Poulton KV, Kennedy LJ, Ross J, Thomson W, Mbanya JC, Ollier WE. 1998. A study of HLA-DPB1 phenotypes reveals DPB1*6301 in a rural population from Cameroon. Eur J Immunogenet 25:375–377.

Rani R, Sood A, Lazaro AM, Stastny P. 1999. Associations of MHC class II alleles with insulin-dependent diabetes mellitus (IDDM) in patients from North India. Hum Immunol 60:524–531.

Ravikumar M, Dheenadhayalan V, Rajaram K, Lakshmi SS, Kumaran PP, Paramasivan CN, Balakrishnan K, Pitchappan RM. 1999. Associations of HLA-DRB1, DQB1 and DPB1 alleles with pulmonary tuberculosis in south India. Tuber Lung Dis 79:309–317.

Richeldi L, Sorrentino R, Saltini C. 1993. HLA-DPB1 glutamate 69: a genetic marker of beryllium disease. Science 262:242–244.

Richeldi L, Kreiss K, Mroz MM, Zhen B, Tartoni P, Saltini C. 1997. Interaction of genetic and exposure factors in the prevalence of berylliosis. Am J Ind Med 32:337–340.

Rozmuller EH, van der Zwan AW, Voorter CE, Tilanus MG. 2000. DPB1*8501, a novel variant in the US Black population. Tissue Antigens 56:282–284.

Russell PJ. 1992. Population genetics. In: Davies G, Dolan K, editors. Genetics. 3rd edition. New York, NY: Harper-Collins. p 707–758.

Saltini C, Richeldi L, Losi M, Amicosante M, Voorter C, van den Berg-Loonen E, Dweik RA, Wiedemann HP, Deubner DC, Tinelli C. 2001. Major histocompatibility locus markers of beryllium sensitization and disease. Eur Respir J 18:677–684.

SAS Institute. 1999. User's Guide: Statistics, Version 7-1 Vol. 4. Cary, NC: SAS Institute, Inc.

Steiner LL, Wu J, Noreen HJ, Moehlenkamp C, Cavalli A, Davidson M, Johnson S, Winden T, Segall M, Begovich AB, Williams TM. 1999. Four new alleles identified in a study of 500 unrelated bone marrow donor-recipient pairs. Tissue Antigens 53:201–206.

Stokinger HE, Scheel LD. 1972. Hyper-susceptibility and genetic problems in occupational medicine—a consensus report. J Occup Med 15:564–573.

Varney MD, Tait BD. 2000. Identification of a novel DPB1* allele (DPB1*6901) and the occurrence of HLA haplotypes that extend to DPB1. Tissue Antigens 55:188–190.

Voorter CE, Tilanus MG, van den Berg-Loonen EM. 2000. Two new HLA DPB1 alleles identified by sequence-based typing: DPB1*8201 and DPB1*8301. Tissue Antigens 56:560–562.

Wang FQ, al-Daccak R, Ju LY, Lethielleux P, Charron D, Loiseau P, Colombani J. 1992. HLA-DP distribution in Shanghai Chinese—a study by polymerase chain reaction—restriction fragment length polymorphism. Hum Immunol 33:129–132.

Wang Z, White PS, Petrovic M, Tatum OL, Newman LS, Maier LA, Marrone BL. 1999. Differential susceptibilities to chronic beryllium disease contributed by different Glu69 HLA-DPB1 and -DPA1 alleles. J Immunol 163:1647–1653.

Wang Z, White PS, Petrovic M, Tatum OL, Newman LS, Maier LA, Marrone BL. 2001. Beryllium sensitivity is linked to HLA-DP genotype. Toxicology 165:27–38.

Weston A, Pan CF, Ksieski HB, Wallenstein S, Berkowitz GS, Tartter PI, Bleiweiss IJ, Brower ST, Senie RT, Wolff MS. 1997. p53 haplotype determination in breast cancer. Cancer Cancer Epidemiol Biomarkers Prev 6:105–112.

Williams RJ, Siegel FL. 1961. Propetology—a new branch of medical science? Am J Med 31:325–327.

Wolff MS, Berkowitz GS, Brower S, Senie R, Bleiweiss IJ, Tartter P, Pace B, Roy N, Wallenstein S, Weston A. 2000. Organochlorine exposures and breast cancer risk in New York City women. Environ Res Section A 84:151–161.

Wood AJJ. 2000. Racial differences in the response to drugs—pointers to genetic differences. N Engl J Med 344:1393–1396.

Zavon MR. 1962. Modern concepts of diagnosis and treatment in occupational medicine. Am Ind Hyg Assoc J 23:30–36.