Letter to the Editor

Am. J. Hum. Genet. 72:1071–1073, 2003

Resolving *ATM* Haplotypes in Whites

To the Editor:

In two recent studies, Bonnen et al. and Thorstenson et al. demonstrated extensive linkage disequilibrium distributed along *ATM* (GenBank accession number U82828) using SNPs (Bonnen et al. 2000; Thorstenson et al. 2001). In whites (Europeans), no recombination was observed along *ATM*. However, there are some discrepancies between the two articles, in that Thorstenson et al. found three haplotypes (H2, H3, and H4) in whites, whereas Bonnen et al. found five major haplotypes with frequencies >5% (2, 3, 15, 17, and 22) and two minor haplotypes in whites. Thorstenson et al. suggest that haplotype 2 and H4 may be equivalent, as may be 17 and H2 and 15 and H3. However, haplotype 22, which accounts for 35% of the population as determined by Bonnen et al., was not accounted for in the suggested equivalency. The two studies had only one overlapping SNP used to determine haplotypes, which may contribute to the discrepancy. Because we are interested in haplotyping *ATM* for association studies with breast cancer, we compared the haplotypes from the two studies in 159 individuals from 83 unrelated families with deleterious *BRCA1* mutations in order to determine which haplotypes are equivalent. Of the probands from the 83 families, all of whom carried *BRCA1* mutations, 72 were affected with cancer and 11 were not affected with cancer. In addition, we sought to determine the association of the three previously studied nonconservative coding region SNPs (S49C, D1853N, and P1054R) with each haplotype. As delineated by Thorstenson et al., there is a total of 12 nonconservative coding region SNPs in all populations, of which four appear in whites (S49C, F868L, D1853N, and P1054R). We were particularly interested in examining the association of the SNPs with the haplotypes, as Thorstenson et al. found that D1853N defined a single haplotype (H3), unlike Bonnen et al., who describe haplotype 15 (the suggested equivalent of H3) independently.

Bonnen typed 295 individuals from four ethnic groups (71 African Americans, 39 Asian Americans, 77 white European Americans, and 73 Hispanic Americans) for 14 SNPs that spanned 142 kb across *ATM*. Using the 14 SNPs, they predicted a total 22 of *ATM* haplotypes, using EMHAPFRE, with five predominant haplotypes having a frequency ⩾5%. The major haplotypes identified in white European Americans were 2 (29.2%), 3 (6.5%), 15 (17.5%), 17 (10%), and 22 (35.1%), as shown in table 1. In addition, they examined the association between three nonconservative coding region SNPs (S49C, D1853N, and P1054R) and the haplotypes they determined. Each nonconservative coding region SNP showed a significant association with a specific haplotype of *ATM*, as defined in their study (table 2). SNP1 (S49C) showed an association with haplotype 2, SNP2 (D1853N) with haplotype 15, and SNP3 (P1054R) with haplotype 17.

Thorstenson et al. typed 93 individuals from seven major human populations (18 from Africa, 9 from the Middle East, 12 from the Indian peninsula, 20 from Asia, 16 from Europe, 8 from Oceania, and 10 American Indians) for 17 SNPs (only one common to the 14 SNPs in the work of Bonnen et al.) spanning 146 kb across *ATM*. Ten of the 17 SNPs were found to be in complete linkage disequilibrium and were used to construct the *ATM* haplotypes. Seven haplotypes (H1–H7) were inferred using a maximum parsimony approach. In the European population, three major haplotypes were identified: H2 (40%), H3 (12.5%), and H4 (47%) (table 1). Thorstenson et al. also examined the association between these haplotypes

Table 1

Frequency of Haplotypes from Bonnen et al., Thorstenson et al., and Current Study Determined in Probands from Families and All Family Members

Study and Haplotypes	Frequency (%)		
		Current Study	
	Published	Probands	All Individuals
Bonnen et al.:			
2	29	33	39
3	6.5	4	3
15	17.5	16	14
17	10	12	14
22	35	35	30
Thorstenson et al.:			
H2	40	47	43
H3	12.5	17	14
H4	47	37	43

and the same amino acid variant SNPs as Bonnen et al. S49C showed an association with H4, P1054R showed an association with H2, and D1853N defined H3.

For our comparison, we constructed new *ATM* haplotypes by genotyping 159 individuals (318 alleles) from 83 families with deleterious *BRCA1* mutations; 150 were white non-Hispanic), and 9 were African American. These individuals are representative of our breast cancer study population. Our aim was to reconstruct *ATM* haplotypes as closely equivalent as possible to those of the other two studies, using the minimum number of SNPs from each paper that defined each haplotype. Thus, the following SNPs were genotyped: 10182, IVS46-257, IVS55+186, and IVS62-694 from the Bonnen study and IVS17-56 and D1853N from the Thorstenson study. The SNPs selected for this study allowed definition of all the major haplotypes in whites with haplotype frequencies >5%. Haplotypes 6 and 21, seen in table 2 of Bonnen et al., have haplotype frequencies <5% and, therefore, were not included in the study. Of the 159 samples typed using the SNPshot protocol on an ABI Prism 3100, all but two samples (1%) were consistent with the haplotype equivalencies shown in figure 1.

On the basis of our findings (shown in fig. 1), Bonnen's haplotype 22 and haplotype 17 are encompassed by Thorstenson's H2, and haplotypes 2 and 3 are encompassed by H4. Haplotype 15 is equivalent to H3. Our haplotype frequencies are consistent with those of Bonnen and Thorstenson in white individuals, as shown in table 1. For the two samples that did not fit into the equivalencies suggested in figure 1, one of the two samples contained haplotype 11, which was shown by Bonnen et al. to have a 1.3% frequency in the Asian population. It was seen in an individual homozygous for H4 and appears to be derived from haplotype 2. In the other sample, the haplotypes were not resolvable despite repeated genotyping.

As in the study by Bonnen et al, we report the percentage of the most frequent haplotype of the total number of alleles that are in the individuals with three nonconserved coding region SNPs (table 2). Similar to the

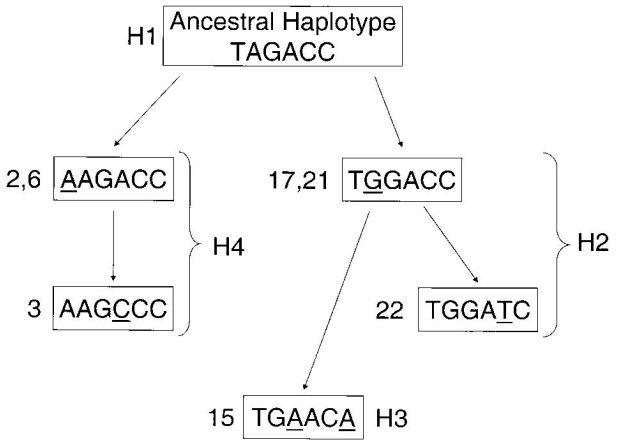

Figure 1 Phylogenetic relationship and equivalencies among major haplotypes in the work of Bonnen et al. and Thorstenson et al. Base pair changes defining the haplotypes are 10182T→A, IVS17-56G→A, 5557G→A (D1853N), IVS46-257A→C, IVS55+186C→T, and IVS62-694C→A.

study of Bonnen et al., haplotype 2 is the most frequently occurring allele in the individuals with S49C (cSNP1; 50%) but does not differ significantly from the percentage of haplotype 2 (of the total alleles) in the remaining individuals (38%; $P = .6$). Haplotype 17 is the most frequent allele in the individuals with P1054R (cSNP3; 46%), significantly more than in the alleles of the individuals without P1054R (9%; $P = .003$). Neither S49C nor P1054R is found exclusively on the haplotypes they are most frequently associated with, 2 and 17, respectively. For both S49C and P1054R, if the individual did not carry the most frequent haplotype (i.e., 2 and 17, respectively), he or she carried the haplotype derived from the most common haplotype (i.e., 3 and 22, respectively).

Unlike Bonnen et al., we did not find haplotype 15 in individuals without the D1853N SNP (cSNP2), and the frequency of haplotype 15 in the individuals with D1853N is entirely reflective of the rate of heterozygotes

Table 2

Most Frequent Haplotypes within Individuals Carrying Nonconservative Coding Region SNPs

	Bonnen et al.			Thorstenson et al.	
Nonconservative Coding Region SNPs	Associated Haplotype	Frequency of Associated Haplotype (%)		Associated Haplotype[b]	Frequency in Current Study of Associated Haplotype[b] (%)
		Current Study[a]	Published[a]		
S49C	2	50	64	H4	56
D1853N	15	66	57	H3	66
P1054R	17	46	52	H2	57

[a] Compares the percentage of the alleles with the most frequent haplotype in the group of individuals with the SNP, as published by Bonnen et al. and in the current study.

[b] Shows the percentage of total alleles contributed by the equivalent haplotypes (from Thorstenson et al.) in all the individuals with the SNP.

and homozygotes for D1853N. Our results are consistent with those of Thorstenson et al., who found the 1853N SNP defining a specific haplotype (H3). However, in general, our results are similar to those found by Bonnen et al. in the white population. In light of the interest in completing haplotype maps of the genome, this study illustrated two points that need to be taken into consideration in haplotype-association studies. First, haplotype association studies might miss functional SNPs similar to S49C, since haplotype 2 is no more frequent in carriers of S49C than noncarriers of S49C. Secondly, some nonconservative coding region SNPs, although associated with certain haplotypes, are not always seen in the context of the same haplotype, as seen with S49C and P1054R, whereas others are completely associated, as seen with D1853N. Our observation illustrates the importance of constructing phylogenetic trees to understand how haploypes might be grouped together for association studies. Thus, association studies using haplotype maps need to be constructed carefully with thought to the potential pitfalls demonstrated by this study.

Acknowledgments

We acknowledge the helpful comments of the two anonymous reviewers of this manuscript. This study was supported by the National Institutes of Health Grant K08 CA84030 (to K.L.N.).

RICHARD LETRERO,[1] BARBARA L. WEBER,[1,2] AND KATHERINE L. NATHANSON[1]
[1]*Department of Medicine and*
[2]*Abramson Family Cancer Research Institute*
University of Pennsylvania School of Medicine
Philadelphia

Electronic-Database Information

Accession number and URL for data presented herein are as follows:

GenBank, http://www.ncbi.nlm.nih.gov/Genbank/ (for genomic sequences of *ATM* [accession number U82828])

References

Bonnen PE, Story MD, Ashorn CL, Buchholz TA, Weil MM, Nelson DL (2000) Haplotypes at *ATM* identify coding-sequence variation and indicate a region of extensive linkage disequilibrium. Am J Hum Genet 67:1437–1451

Thorstenson YR, Shen P, Tusher VG, Wayne TL, Davis RW, Chu G, Oefner PJ (2001) Global analysis of *ATM* polymorphism reveals significant functional constraint. Am J Hum Genet 69:396–412

Address for correspondence and reprints: Dr. Katherine L. Nathanson, 513 BRB 2/3, 421 Curie Boulevard, University of Pennsylvania School of Medicine, Philadelphia, PA 19104. E-mail: knathans@mail.med.upenn.edu

0002-9297/2003/7204-0032$15.00