

## Localization of Psoriasis-Susceptibility Locus PSORS1 to a 60-kb Interval Telomeric to HLA-C

Rajan P. Nair,<sup>1</sup> Philip Stuart,<sup>1</sup> Tilo Henseler,<sup>4</sup> Stefan Jenisch,<sup>5</sup> Nicholas V. C. Chia,<sup>1</sup> Eckhard Westphal,<sup>5</sup> Nicholas J. Schork,<sup>6</sup> Jane Kim,<sup>7</sup> Henry W. Lim,<sup>7</sup> Enno Christophers,<sup>4</sup> John J. Voorhees,<sup>1</sup> and James T. Elder<sup>1,2,3</sup>

Departments of <sup>1</sup>Dermatology and <sup>2</sup>Radiation Oncology, University of Michigan, and <sup>3</sup>Ann Arbor Veterans Affairs Hospital, Ann Arbor; Departments of <sup>4</sup>Dermatology and <sup>5</sup>Immunology, University of Kiel, Kiel, Germany; <sup>6</sup>Department of Biostatistics, Case Western Reserve University, Cleveland; and <sup>7</sup>Department of Dermatology, Henry Ford Hospital, Detroit

Recent genome scans have established the presence of a major psoriasis-susceptibility locus in the human leukocyte antigen (HLA) complex on chromosome 6p21.3. To narrow the interval for candidate gene testing, we performed a linkage-disequilibrium analysis of 339 families, with the use of 62 physically mapped microsatellite markers spanning the major histocompatibility complex (MHC). As detected by use of the transmission/disequilibrium test (TDT), individual markers yielded significant linkage disequilibrium across most of the MHC. However, the strongest evidence for marker-trait disequilibrium was found in an ~300-kb region extending from the MICA gene to the corneodesmosin gene. Maximum-likelihood haplotypes were constructed across the entire MHC in the original sample and across a 1.2-Mb region of the central MHC in an expanded sample containing 139 additional families. Short (two- to five-marker) haplotypes were subjected to the TDT using a “moving-window” strategy that reduced the variability of TDT *P* values relative to the single-locus results. Furthermore, the expanded sample yielded a sharp peak of evidence for linkage disequilibrium that spanned ~170 kb and that was centered 100 kb telomeric to HLA-C. The 1.2-Mb interval was further dissected by means of recombinant ancestral haplotype analysis. This analysis identified risk haplotype 1 (RH1), which is a 60-kb fragment of ancestral haplotype 57.1, on all identifiable HLA risk haplotypes. One of these haplotypes exhibits significant linkage disequilibrium with psoriasis but does not carry Cw6, which is the HLA allele most strongly associated with the disease. These results demonstrate that RH1 is highly likely to carry the disease allele at PSORS1, and they exclude HLA-C and corneodesmosin with a high degree of confidence.

### Introduction

Psoriasis is a common, inflammatory, and hyperproliferative skin disease of suspected immunopathogenesis, with a strong genetic component (Elder et al. 1994; Christophers and Mrowietz 1999). Several confirmed and putative genetic loci for psoriasis have been described in the course of genome scanning (reviewed in Bhalerao and Bowcock 1998). The locus of strongest effect appears to reside within the major histocompatibility complex (MHC; Nair et al. 1997; Trembath et al. 1997; Burden et al. 1998; Leder et al. 1998). This locus has been designated as “PSORS1” (MIM 177900). The results of previous studies of linkage and/or association (Trembath et al. 1997; Jenisch et al. 1998; Balendran et

al. 1999; Oka et al. 1999) favor a location for PSORS1 in the class I or class III regions of the MHC, and strong allelic associations between psoriasis and human leukocyte antigen (HLA)–Cw6 have long been recognized in case-control studies (Tiilikainen et al. 1980). Associations between Cw6 and psoriasis are even more significant in juvenile-onset and familial psoriasis (Henseler and Christophers 1985; Enerback et al. 1997; Jenisch et al. 1998; Mallon et al. 1999). However, several lines of evidence suggest that HLA-Cw6 is not the PSORS1-susceptibility allele (Jenisch et al. 1998; Leder et al. 1998; Jenisch et al. 1999b). The most obvious evidence for this is the fact that many psoriatics do not carry HLA-Cw6. This could be explained by the existence of one or more psoriasis-susceptibility gene(s) unlinked to the MHC (locus heterogeneity), with Cw6 directly responsible for different proportions of the disease in different populations. However, locus heterogeneity cannot explain two other key observations. First, different psoriasis-associated HLA-Cw6-Bx haplotypes (Cw6-B57, Cw6-B13, and Cw6-B37) appear to carry different risks in populations that are widely separated in terms of geography and ethnicity (Ozawa et al. 1981; Roitberg-Tambur et al.

Received February 23, 2000; accepted for publication March 30, 2000; electronically published May 5, 2000.

Address for correspondence and reprints: Dr. James T. Elder, 3312 Cancer Center and Geriatric Center Building, Box 0932, 1500 East Medical Center Drive, University of Michigan, Ann Arbor, MI 48109-0932. E-mail: jelder@umich.edu

© 2000 by The American Society of Human Genetics. All rights reserved. 0002-9297/2000/6606-0012\$02.00

1994; Ikaheimo et al. 1996; Kim et al. 2000). Second, there is at least one HLA Cw6-negative HLA-B/C haplotype—Cw8-B65—that appears to segregate with psoriasis in some families (Jenisch et al. 1997).

An alternative explanation is that PSORS1 resides near HLA-C, with the disease allele in strong linkage disequilibrium with Cw6. Recently, strong associations have been observed between psoriasis and “allele 5” of corneodesmosin (CDSN [MIM 602593], originally known as the “S gene”), which resides ~160 kb telomeric to HLA-C (Zhou and Chaplin 1993; Allen et al. 1999; Tazi Ahnini et al. 1999). However, a more complete description of the allelic variation at CDSN revealed that allele 5 can be split into two alleles—known as “CD2” and “CD3”—that differ only at a single nucleotide (1215A/G). CD2 and CD3 were found only on known psoriasis-associated haplotypes that also carry HLA-Cw6: CD2 was found on Cw6-B57 and Cw6-B37, and CD3 was found on Cw6-B13 (Jenisch et al. 1999a). Thus, CDSN does not appear to account for risk of psoriasis any better than does HLA-Cw6.

The finding of linkage disequilibrium is not specific to psoriasis (or to other HLA-associated diseases), but it is also observed, in normal individuals, across ~3 Mb of the MHC, extending from HLA-DQ to HLA-A (reviewed in Todd, in press). This is a reflection of the fact that most modern-day HLA haplotypes can be accounted for by ~20 ancestral or extended haplotypes and their recombinants (Degli-Esposti et al. 1992c). These are designated as “EHxx.y,” where EH denotes an extended haplotype, where xx refers to the HLA-B allele carried by that haplotype, and where y refers to identified variants of that haplotype. Both Cw6 and the CD2 allele at CDSN are carried by the ancestral haplotype designated as “EH57.1” (Jenisch et al. 1999a). By making use of HLA types as markers, we (Jenisch et al. 1998) and others (Schmitt-Egenolf et al. 1996) have shown that familial psoriasis is associated with the class I end of the EH57.1 haplotype.

In the complete sequence of the MHC, 224 gene loci—128 of which are predicted to be expressed—have been identified (MHC Sequencing Consortium 1999). On the basis of the foregoing arguments, we suspected that PSORS1 might be a gene other than HLA-C or CDSN. We recognized that, although the HLA genes are extremely polymorphic, they are not sufficiently dense to map PSORS1 with high precision. To overcome this limitation, we screened a large number of potential polymorphisms identified from the emerging MHC sequence, and we used new and existing polymorphisms to perform high-density genotyping of a large collection of kindreds with familial psoriasis. In the present study, we report several analyses of these genotypes, all of which support localization of PSORS1 immediately telomeric to HLA-C. The most powerful analytical

method proved to be recombinant ancestral haplotype analysis. This method has previously been used to localize genes for insulin-dependent diabetes mellitus, myasthenia gravis, and Alzheimer disease as well as psoriasis (Todd et al. 1989; Degli-Esposti et al. 1992a; Degli-Esposti et al. 1992b; Schmitt-Egenolf et al. 1996; Jenisch et al. 1998). This analysis identified a 60-kb fragment of EH57.1 that is highly likely to carry the disease allele at PSORS1. This interval excludes both HLA-C and CDSN with a high degree of confidence.

## Subjects and Methods

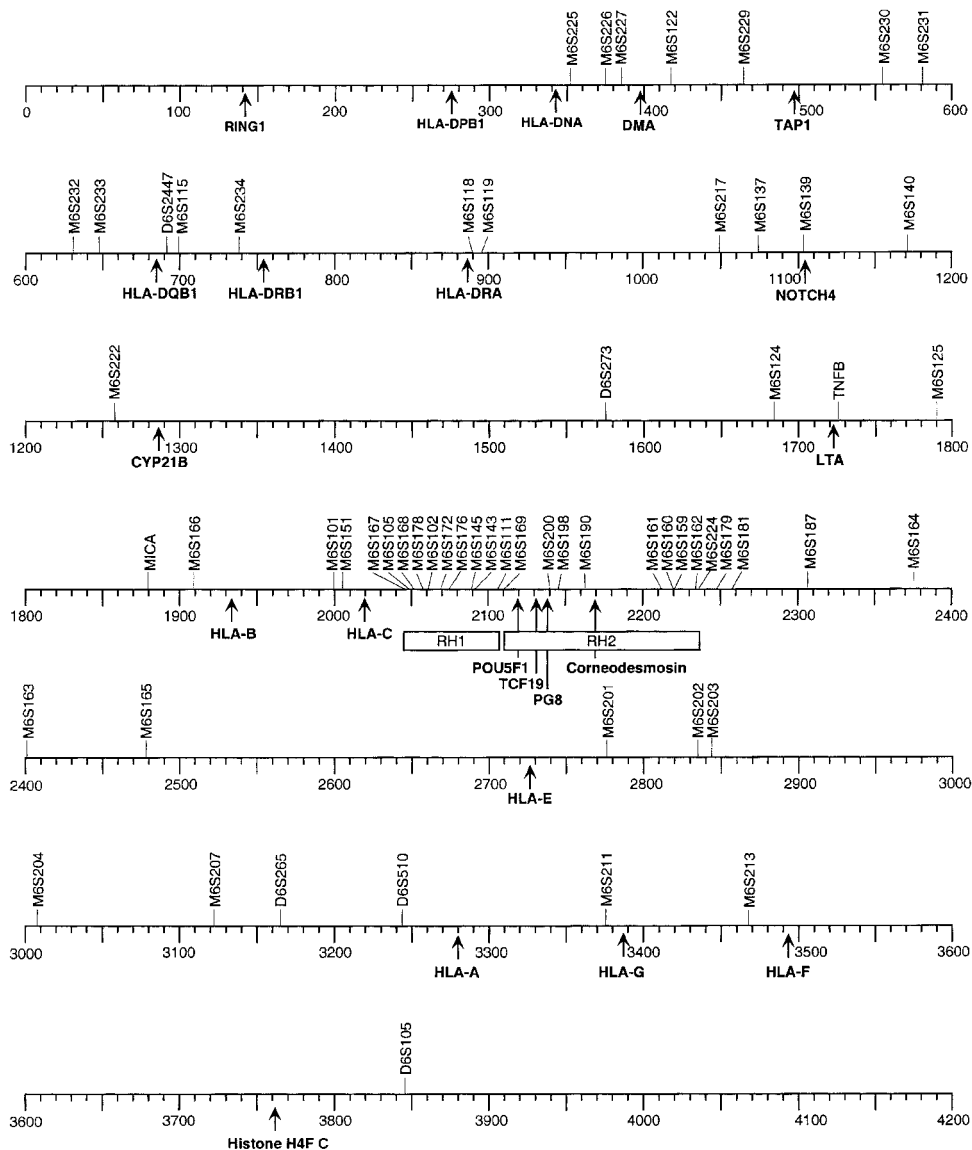
### Study Population

The study cohort consisted of families identified through the dermatology services of the University of Michigan Medical Center, the Ann Arbor Veterans Affairs Medical Center, the University of Kiel, and Henry Ford Hospital. Additional families were provided by the National Psoriasis Foundation Tissue Bank. On the basis of the established correlation between juvenile onset and a positive family history of psoriasis (Henseler and Christophers 1985), only those families in which the age at onset in the proband was <40 years were included. Individuals were considered to be affected if chronic plaque or guttate psoriasis lesions covered >1% of the total body-surface area (TBSA) or if at least two skin, scalp, nail, or joint lesions were clinically diagnostic of psoriasis (Nair et al. 1997). Informed consent was obtained from all subjects, under protocols approved by the institutional review boards of the participating institutions.

### Markers and Genotyping

A set of 62 microsatellite markers spanning 3.5 Mb was identified from published sources or by scanning of MHC-sequence submissions for short-nucleotide repeats (fig. 1). A complete description of these markers is available from the Web site of the Psoriasis Genetics Laboratory, Department of Dermatology, University of Michigan. As measured on the physical map of the MHC (available from the Sanger Centre Web site; see also GenBank [accession number NT\_001520]), the interval studied extended from 350 kb (~330 kb centromeric of HLA-DQB1) to 3,845 kb (~90 kb telomeric of histone H4) (fig. 1). The markers designated as “M6Sxxx” in figure 1 were developed in our laboratory at the University of Michigan, on the basis of the available sequence of the region. These markers have been deposited in the Genome Database. All other markers were identified from the Genome Database and are identified, in figure 1, by their assigned primary names in the Genome Database.

DNA was prepared from blood samples or from Ep-



**Figure 1** Physical map of the HLA complex, showing the microsatellite markers used in the present study. The map coordinates are based on the 3,838,986-bp consensus map available at the Sanger Centre Web site. The positions of the markers indicate the centromeric end of the amplimers, and the positions of the genes indicate the translation start codons. Positions of RH1 and RH2 are also indicated (see text for details).

stein-Barr virus-immortalized lymphoblasts and were genotyped as described elsewhere (Nair et al. 1995). Genotyping was performed on 478 families, each of which contained at least one affected individual with at least one typed parent. The primary sample, which consisted of 339 families, was typed for all 62 markers. An additional 139 families were typed for the middle 34 markers spanning 1.2 Mb, from 1,575 to 2,775 kb. The combined cohort of 478 families will be referred to as the “expanded sample” (table 1).

Allele sizes were determined by inspection of autoradiograms. When available, published allele sizes for

CEPH individual 134702 (available at the Center for Medical Genetics, Marshfield Medical Research Foundation Web site) were used as a reference. For the newly developed markers, allele sizes of the reference standard were assessed by use of the SeqMark DNA-sequencing ladder (Research Genetics). Genotyping problems causing Mendelian inheritance errors were discovered, and allele-frequency files were created with the use of PEDMANAGER and the LINKAGE package (Terwilliger and Ott 1994). Additional error checking was accomplished by means of visual inspection of the marker haplotypes generated by GENEHUNTER (see the Haplo-

**Table 1****Families Utilized in the Present Study**

FAMILY	SAMPLE	
	Primary	Expanded
Nuclear <sup>a</sup>	110	130
Extended <sup>b</sup>	81	97
Singleton <sup>c</sup>	148	251
Unrelated trios:		
With both parents	262	381
With one parent	77	97
Total, unrelated trios	339	478

<sup>a</sup> Two or more siblings with at least one parent available.

<sup>b</sup> Pedigrees with more than one nuclear family in two or more generations.

<sup>c</sup> An affected individual with one or both parents.

type Construction and Clustering section, below) for double recombinants.

Some individuals were also subjected to HLA and CDSN typing. HLA-C, HLA-DR, HLA-DQ, and CDSN typings were performed by means of DNA-based methods, and HLA-B types were determined serologically, as described elsewhere (Jenisch et al. 1998; Jenisch et al. 1999a).

#### Haplotype Construction and Clustering

Maximum-likelihood haplotypes for pedigrees were determined by use of GENEHUNTER, version 1.1, with the settings “max bits = 16” and “skip large = off” (Kruglyak et al. 1996). All uncertain phase assignments were detected and were converted to missing values.

Clusters of short haplotypes were formed by use of a “moving-window” strategy, with the window length varying from two to seven contiguous markers. Phase and allele number had to be known at all marker positions within the window, and no mismatches were allowed. A minimum of 10 founders were required to create a cluster in any given window. All founder haplotypes that fell into smaller groups or that had less than full characterization were lumped together into a single cluster for that window. The window was moved across the MHC one marker at a time, from centromere to telomere, and clustering was repeated for each window.

Clusters of long haplotypes were generated by use of an average-distance agglomerative hierarchical method with a percent difference metric, with the use of SYSTAT for Windows, release 8.0.3 (SPSS). Each long haplotype was comprised of 34 contiguous markers. To qualify for clustering, at least half of the alleles of a long haplotype were required to be typed and of known phase (>80% of the marker alleles were of known phase and were typed in 95% of the founder haplotypes). Untyped or phase-uncertain alleles were labeled as “missing,” rather than as “zero,” to avoid false matches and mismatches.

The criteria for assignment to a cluster were  $\geq 80\%$  homogeneity of the haplotypes within the cluster and a minimum of five founders. All smaller clusters were lumped together into a single cluster.

Clusters of short and long founder haplotypes were numbered and were assigned to all nonfounders in the pedigrees. Any recombination within the portion of the haplotype subjected to analysis caused it to be discarded for all individuals inheriting the recombinant haplotype.

HLA-B, HLA-C, HLA-DR, HLA-DQ, and CDSN alleles were not incorporated directly into either the long-haplotype determinations performed by GENEHUNTER or the haplotype-clustering algorithms. Alleles at these loci were assigned to haplotypes by inspection of the informative pedigrees.

#### Disequilibrium Testing

The transmission/disequilibrium test (TDT) was used for the determination of linkage disequilibrium between psoriasis and alleles at individual loci as well as between psoriasis and haplotypes generated by clustering. The multiallelic  $T_{mhet}$  statistic (Spielman and Ewens 1996) was used for the single-marker and short-haplotype scans. The biallelic TDT statistic (Spielman et al. 1993) was used for testing of individual haplotype clusters. In all analyses, only one trio per family was used. When more than one trio was present in a single family, we used the most fully typed trio. If a pedigree contained more than one equally typed trio, then one trio was randomly selected. Because results varied somewhat, depending on the seed used for the random-number generator, the analysis was repeated 15 times. For each marker, haplotype window, or haplotype, the run yielding the median  $P$  value is reported. Trios with only one typed parent were handled as recommended by Curtis and Sham (Curtis and Sham 1995). Inferred haplotypes and genotypes for untyped parents were not used for the TDT, because of the potential for bias (Curtis 1997). However, they were used when clustering founder haplotypes.

Pseudosupport intervals were determined with the use of a criterion of  $-3.0 \log_{10}$  units. This criterion is akin to a 95% confidence interval for 62 tests, since  $-\log_{10}(0.05/62) = 3.09$ . The interval was then conservatively extended by one marker outward from the last qualifying marker on either side. This criterion does not define a statistically valid support interval; to our knowledge, no such method exists for linkage-disequilibrium studies. However, it does provide a reasonable and objective method for comparison of the results of different analyses.

For recombinant-haplotype analysis of two contiguous haplotype fragments derived from one long ancestral haplotype, we extracted all founder haplotypes that were

fully characterized for all markers comprising the two fragments. These haplotypes were then grouped according to whether they perfectly matched either or both of the haplotype fragments; a mismatch was scored if at least one allele within the fragment differed from the parental haplotype. Each of these groups was then compared for risk, by use of the biallelic TDT.

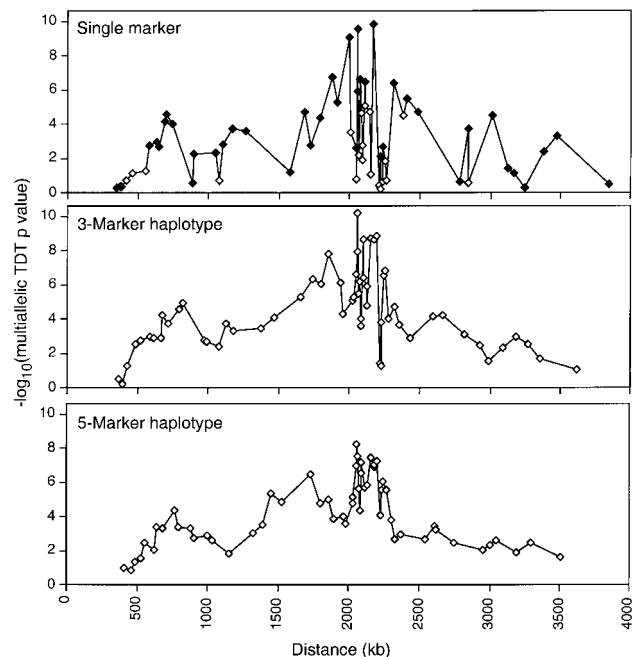
#### Comparison of Disease Phenotypes in Genetically Defined Subsets

Psoriatics that carried the risk haplotype (RH) identified by the aforementioned studies were compared with psoriatics that did not carry the RH, with respect to age at onset, percentage of TBSA involved, presence or absence of nail involvement, and presence or absence of arthritis, as determined by history or physical examination. One RH-positive and one RH-negative affected individual were randomly selected from each family. Only one individual was selected from those families in which all affected individuals were either RH-positive or RH-negative. The distributions for TBSA and age at onset were compared with the use of the Kolmogorov-Smirnov two-sample test; nail involvement and arthritis were analyzed by means of a  $\chi^2$  contingency test.

## Results

### Single-Marker TDT

As shown in the top panel of figure 2, the plot of  $T_{mhet}$   $P$  values as a function of physical map position in the MHC reveals linkage disequilibrium with psoriasis throughout much of the HLA region. For the primary sample of 339 families used in analysis of the entire MHC, the two most distant markers yielding multiallelic TDT  $P \leq .05$  were separated by a distance of 3.0 Mb, extending from M6S229 (map position 465 kb) to M6S213 (map position 3,468 kb). TDT  $P \leq .05$  was found for 43/57 markers spanning this region. After Bonferroni correction was done for 62 single-marker tests, the two most distant markers yielding  $P \leq .05$  were separated by a distance of 2.3 Mb, extending from just telomeric of HLA-DQB1 (D6S2447; 692 kb) to 270 kb centromeric of HLA-A (M6S204; 3,008 kb). Corrected TDT  $P \leq .05$  was found for 26/46 markers spanning this region. The range of the uncorrected multiallelic TDT  $P$  value was  $0.5-10^{-10}$ , with substantial variation found even for closely spaced markers. However, many of the markers yielding less significant  $P$  values were of low heterozygosity and, therefore, were less likely to be unique to the disease chromosome (fig. 2, top [unblackened diamonds]). Nevertheless, a peak was clearly evident in the data (fig. 2). The pseudosupport interval for this peak was 303 kb long and extended from 25 kb centromeric of HLA-B (M6S166; 1,909.4



**Figure 2** Multiallelic TDT analysis of the MHC region (primary sample). *Top*, TDT results for individual markers. Blackened diamonds indicate markers characterized by heterozygosity  $>0.7$ . The corresponding values for a three-marker haplotype window (*middle*) and the corresponding values for a five-marker haplotype window (*bottom*) are also shown. The midpoints of the three- and five-marker haplotypes are plotted on the X-axis.

kb) to 40 kb telomeric of CDSN (M6S161; 2,212.1 kb). This interval was subjected to additional analysis in the expanded sample of 478 families. In the expanded sample, the six markers yielding the most significant TDT  $P$  values were all contained within the 303-kb region identified by the primary sample (data not shown). Five of these markers yielded multiallelic TDT  $P < 5 \times 10^{-14}$ , which is nearly four orders of magnitude more significant than the  $P$  value of any marker lying outside the 303-kb interval.

### Short-Haplotype TDT

Multiallelic TDT  $P$  values were determined for two- to five-marker haplotypes, in both the primary and expanded samples; results for three- and five-marker haplotypes in the primary sample are shown (fig. 2, middle and bottom). The  $P$  value plots were much less “jagged” for the short haplotypes than for the single markers, and the disparity between low and high  $P$  values within the region of strongest disequilibrium steadily decreased as the haplotype window increased up to five markers. In the expanded sample, narrowing of the pseudosupport interval to  $\sim 170$  kb was observed as the length of the haplotype window increased from two to four markers.



the ordered dendrogram were grouped and were tested en masse by TDT. They, too, were found to carry risk (66.0% transmission;  $P = .011$ ). Inspection of the corresponding HLA types revealed that ~90% of the evaluable haplotypes in this group carried Cw6, together with various HLA-B alleles (alleles 13, 18, 37, 38, 44, 45, 57, 60, and 62). Most of the remaining evaluable haplotypes carried Cw8 and B65.

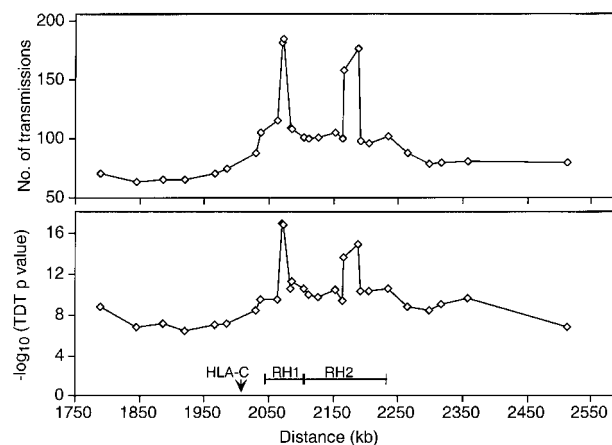
Inspection of the consensus-allele composition of the RHs revealed that they share all or part of a 185-kb region just telomeric of HLA-C. All RHs (including those that are too infrequent to fall within a testable cluster) share an identical allele sequence for an eight-marker region extending 42.5 kb from M6S105 to M6S143 (table 2). When extended to the first nonconserved flanking markers, this becomes a 59.4-kb segment extending from M6S167 to M6S111 (2,048.5–2,107.9 kb). We will refer to this haplotype as “RH1” (fig. 1). Of the RH clusters, all but one (cluster 17) share an identical allele sequence over a seven-marker region extending a distance of 93.5 kb, from M6S200 to M6S162; this becomes a 126.6-kb segment when it is extended to the first nonconserved flanking markers, M6S169 and M6S224 (2,109.9–2,236.5 kb). We will refer to this haplotype as “RH2.” All RHs—except for cluster 17—also share Cw6, although they carry a nonconserved marker (M6S167) between HLA-C and RH1.

#### Verification of RH1 and RH2 by Short Haplotypes That Perfectly Match EH57.1

Long-haplotype cluster 25 was determined to represent the alleles carried by EH57.1, because nearly all founder haplotypes in this cluster carry HLA-Cw6, B57, DR7, and DQ9, which are the expected EH57.1 alleles at these loci. Moreover, this cluster was very homogeneous: for 29/34 markers, the consensus allele was present in  $\geq 95\%$  of the founder haplotypes. Agreement with the consensus dropped toward the telomeric end of the cluster, since only 75%–89% of the haplotypes matched the consensus for the four most telomeric alleles. The remaining exception was M6S224, which appears to have a higher-than-normal mutation rate. It is heterogeneous in most clusters, and it produced nine apparent mutations in our sample, compared with two or fewer mutations seen for any of the other 33 markers within the 1.2-Mb region. Since long-haplotype cluster 25 is the most common RH in our sample and since it clearly derives from EH57.1, it is reasonable to assume that other risk clusters, which match EH57.1 only near HLA-C, derive their matching fragments from EH57.1 as well. On the basis of this assumption, a more powerful disequilibrium test is made possible by focusing only upon transmission of fragments of the EH57.1 haplotype, which is represented by long-haplotype cluster 25. For

this purpose, we used the 1.2-Mb region extending from D6S273 to M6S201 to analyze the expanded sample of 478 families. Overlapping seven-marker haplotypes were tested with the use of a biallelic TDT (i.e., EH57.1 vs. non-EH57.1) using the moving-window strategy described in the Subjects and Methods section. This window length was chosen because seven-marker haplotypes are long enough to provide reasonable assurance of IBD without being so long as to greatly reduce the number of evaluable haplotypes caused by missing marker genotypes or phase-uncertain alleles. This is also approximately the length of RH1 and RH2.

Graphs were produced for the number of transmissions to affected children and for the  $-\log_{10} P$  value for transmission of these haplotypes from heterozygous parents to affected offspring (fig. 3). We found a high level of linkage disequilibrium between psoriasis phenotype and EH57.1 genotype across the entire 1.2-Mb region (75%–80% transmission of seven-marker fragments of EH57.1 to offspring [data not shown]). However, the number of transmissions to affected individuals increased from a steady background of 60–80 to ~190 as the window passed through the two RH regions, with a distinct valley occurring between them (fig. 3, top). The corresponding significance levels increased by seven orders of magnitude across RH1 and by five orders of magnitude across RH2, when compared with the remainder of the 1.2-Mb region (fig. 3, bottom). These results indicate that many affected individuals in our sample inherited a large DNA segment found on the EH57.1 ancestral haplotype, and they explain why link-



**Figure 3** Short-haplotype biallelic TDT analysis (expanded sample). Both panels refer to data obtained for a seven-marker moving window. Symbols are plotted at the midpoint of each window. Perfect matches to cluster 25 constituted one allele, and all other haplotypes constituted the other allele. *Top*, Number of transmissions, to affected offspring, of the EH57.1 haplotype fragment within the moving window. *Bottom*,  $-\log_{10}$  of the TDT  $P$  value for the EH57.1 haplotype fragment contained within the moving window.

age disequilibrium with psoriasis is seen throughout much of the MHC region. However, sufficient ancestral recombinations have occurred to leave only RH1 and RH2 to be shared by the great majority of affected individuals inheriting any portion of the EH57.1 haplotype represented by cluster 25.

#### *Comparison of Disequilibrium for RH1 and RH2 on Ancestral Haplotype EH57.1*

As shown in table 2, only RH1 is found in all identifiable RHs, suggesting that RH1—not RH2—is likely to harbor PSORS1. However, the exclusion of RH2 (and, for that matter, HLA-C) as a candidate rests upon a single RH: cluster 17. Moreover, cluster 17 yields a 21:10 transmitted:nontransmitted (T:NT) ratio in the TDT (67.7% transmission), suggesting genuine risk, albeit with marginal significance (uncorrected  $P = .048$ ). To better compare the risk conferred by these two regions, we grouped haplotypes according to whether they perfectly matched either or both of the RH regions. This increased the power of the analysis, by including all haplotypes containing either of the two risk regions and not just those that happened to fall into a cluster containing at least five founders. Disequilibrium with psoriasis was then assessed for each haplotype group, by use of a biallelic TDT for independent trios. Haplotypes carrying both RH1 and RH2 showed strong preferential transmission to affected offspring (T:NT 150:41, transmission 78.5%;  $P = 3.1 \times 10^{-15}$ ). Haplotypes carrying RH1 but not RH2 also showed strong preferential transmission (T:NT 28:11, transmission 71.8%;  $P = .0065$ ), whereas haplotypes carrying RH2 but not RH1 did not (T:NT 6:9, transmission 40%;  $P = .44$ ). Moreover, the transmission ratios for the “RH1 only” and “RH2 only” groups were significantly different, despite small sample size ( $P = .030$ ). These results are in agreement with the long-haplotype analysis, indicating that RH1—not RH2—harbors the disease locus.

#### *Phenotypic Analysis of RH1-Positive versus RH1-Negative Psoriatics*

To determine whether genetic heterogeneity exists in the sample in the present study, we compared RH1-positive and RH1-negative psoriatics for their age at onset, TBSA involvement, nail involvement, and arthritis, as described in the Subjects and Methods section, above. Psoriatics carrying RH1 ( $n = 320$ ) had a significantly different age-at-onset distribution ( $P = .00125$ ), with a lower mean age (19.7 vs. 23.9 years), a lower SD (12.1 vs. 15.1 years), and a more skewed distribution (positive skew 1.03 vs. 0.94), compared with psoriatics lacking RH1 ( $n = 226$ ). No significant differences were seen for the other three variables.

## Discussion

Because several preexisting lines of evidence had suggested that PSORS1 might be a novel gene in the HLA class I region, we used the emerging human MHC sequence to develop a set of polymorphic markers that covered the MHC but were densest in the class I region. We then used these markers to type 339 families ascertained on the basis of juvenile-onset psoriasis in the proband. Analysis of these data, with use of the single-locus TDT, identified a peak of linkage disequilibrium that was centered just telomeric of HLA-C. A 1.2-Mb interval encompassing this peak was genotyped in an additional 139 families. Although TDT analysis of two- to four-marker haplotypes narrowed the width of the peak to 170 kb, five-marker haplotypes yielded no further localization, signaling a limit to the usefulness of haplotype-based TDT analysis. To further localize PSORS1, we turned to recombinant ancestral haplotype analysis. Use of a dense marker map to identify recombinants of an ancestral haplotype known to carry a disease gene is a powerful analytical tool, since comparison of recombinant haplotypes in unrelated individuals with psoriasis permits the study of a much larger number of informative recombinants than would typically be identified by a family study (Xiong and Guo 1997; Schork et al. 1998).

To identify long haplotypes shared across families, we applied cluster analysis, requiring  $\geq 80\%$  similarity and at least five founders per cluster. All RHs identified by biallelic TDT proved to be identical at the eight internal markers of RH1, which is a 59.4-kb segment encompassing these markers plus flanking markers M6S167 and M6S111. Moreover, all but one of the RHs were identical at the seven internal markers of RH2, which is a 126.6-kb segment encompassing these markers plus flanking markers M6S169 and M6S224 (table 2). When extended to their flanking nonconserved markers, as defined in the Subjects and Methods section, RH1 and RH2 are nearly contiguous, since they are separated only by the 2-kb interval between M6S111 and M6S169 (fig. 3). In aggregate, RH1 and RH2 corresponded closely to the 170-kb peak of linkage disequilibrium identified by haplotype-based TDT.

Cluster analysis is robust to missing data, marker mutation, and occasional genotyping errors. However, it carries the risk that similar haplotypes that are not IBD could be scored as IBD by the clustering algorithm. (This could be a problem even when all markers are fully characterized and when the cluster is 100% homogeneous, but this is unlikely.) Even though unintentional inclusion of a non-IBD haplotype would be predicted to weaken, rather than strengthen, the TDT results, it was still a concern. As a safeguard against false IBD assignment, we employed two methods of



short-haplotype analysis that required a perfect match to a predefined haplotype. This strategy also allowed for the inclusion of rare haplotypes excluded from the long-cluster analysis, thereby increasing power. The results of these analyses confirmed the existence of two RHs and demonstrated that RH1 is the minimal shared region common to all RHs. Moreover, all testable haplotypes that carry RH1 also carry risk for psoriasis. Taken together, these findings indicate that RH1 is highly likely to carry the disease allele at PSORS1.

RH1 extends from 29.2 to 88.6 kb telomeric of HLA-C and is found on at least one Cw6-negative haplotype cluster (Cw8-B65-CD1a; cluster 17). We have previously observed that the Cw8-B65 haplotype appears to segregate with psoriasis in some—but not in all—families (Jenisch et al. 1997). Although other plausible explanations exist (even cluster 25 does not demonstrate perfect segregation in the families under investigation), this observation may be explained by our identification of two distinct clusters carrying Cw8 and B65, only one of which carries RH1 (compare clusters 17 and 61 in table 2). Taken together, these studies definitively rule out HLA-C as PSORS1 for the first time. These studies also rule out several other nearby telomeric genes, including PG8, POU5F1, TCF19, and CDSN (fig. 1). Consistent with this finding, DNA sequencing of coding regions has failed to reveal any disease-specific sequence variations in the POU5F1 and TCF19 genes (R.P.N., P.S., N.V.C.C., and J.T.E., unpublished observations). Consistent with the fact that CDSN is contained within the boundaries of RH2 (fig. 1), cluster 25 (EH57.1) and cluster 17 differ at their CDSN alleles, with cluster 25-positive individuals carrying CD2 and with cluster 17-positive individuals carrying CD1a [Jenisch et al. 1999a; additional data not shown].

Only 61% of the affected individuals in our sample carry an identifiable copy of RH1. The remaining 39% of individuals may: (1) carry the same disease allele IBD at PSORS1, although the retained portion of RH1 may be too small to be identified by our analysis; (2) carry different disease allele(s) at PSORS1 that are represented too few times to be recognized; or (3) carry no disease allele at PSORS1. Consistent with prior associations between Cw6 and age at onset (Henseler and Christophers 1985), in the present study, stratification of the population for RH1 yielded a significant difference in the age-at-onset distribution. On the basis of these results, it is reasonable to conclude that heterogeneity exists in sample in the present study; however, we cannot distinguish between allelic heterogeneity, locus heterogeneity, and the existence of nongenetic causes. At the same time, these results also suggest that the disease allele carried on RH1 is a major determinant of psoriasis, since it is capable of exerting a detectable effect on the

phenotype. The importance of RH1 is underscored by its presence in 61% of affected individuals.

According to the recently published gene map of the HLA region (MHC Sequencing Consortium 1999), the interval defined by RH1 contains three transcription units: NOB4, HCGII-2, and HCGIX-3. Although these transcription units have been annotated as pseudogenes on the HLA map (MHC Sequencing Consortium 1999), they are, nevertheless, candidates for the PSORS1 gene and are under active investigation. However, because of the high degree of polymorphism manifested by the MHC, identification of the true disease-causing sequence variant will likely require genomic-sequence comparison of multiple nondisease chromosomes and the disease chromosome. If this proves to be the case, the relatively short length of RH1 (60 kb) should greatly simplify the identification of the PSORS1 gene.

The identification of RH1 has important implications for the population genetics of psoriasis. The existence of ancestral haplotypes has been suggested to result from suppression of recombination across certain segments of the MHC (Gaudieri et al. 1997) or to occur as a reflection of natural selection for particular allele combinations (Todd, in press). However, the results of recent studies (Collins et al. 1999) have suggested that the observed degree of linkage disequilibrium across the HLA complex may not be markedly atypical for the range of values exhibited by the human genome. Rather, the existence of ancestral haplotypes may simply reflect the origin of modern *Homo sapiens* from a limited founder population thought to have migrated out of Africa 100,000–200,000 years ago (Jorde et al. 1998), coupled with repeated population bottlenecks in the interim (Collins et al. 1999). This “bottleneck hypothesis” is attractive, since it can explain four curious features of the population genetics of psoriasis. First, psoriasis is associated with HLA-B13, HLA-B37, and HLA-B57 in many world populations (Tsuji et al. 1992), and all of these haplotypes carry Cw6 (reviewed in Elder et al. 1994). The data from the present study strongly suggest that the EH57.1 haplotype is the oldest chromosome to carry RH1 and that the Cw6-B13 and Cw6-B37 haplotypes are associated with disease because they often retain RH1. Second, there is a general correlation between the prevalence of psoriasis and that of Cw6, with the prevalence of both being highest in northern Europeans, intermediate in Asian populations, and nearly nonexistent in Amerindian populations (Farber and Nall 1974; Yip 1984; Tsuji et al. 1992; Mallon et al. 1999; Youn et al. 1999). These phenomena would be readily accounted for by genetic drift, with concomitant loss of Cw6 and RH1. Third, Cw6 is a common allele in Nigerians, Zimbabweans, and native black South Africans, among whom psoriasis is relatively uncommon. Conversely, Cw6 is rare in other African populations in

which psoriasis is common (Leder and Farber 1997). However, the Cw6-B57, Cw6-B37, and Cw6-B13 haplotypes are uncommon in Nigerians, Zimbabweans, and black South Africans (Tsuji et al. 1992). A parsimonious explanation would be that the RH1 disease determinant was present in premigration Africa but that it existed much closer to population equilibrium with HLA-C than it does in present-day populations. Fourth, although Cw6 carries a strongly increased relative risk of psoriasis in nearly all populations tested (Mallon et al. 1999), the proportion of Cw6-positive psoriatics and the associated relative risk varies substantially in different populations. While locus heterogeneity can explain variations in disease prevalence, it cannot explain why the association of psoriasis appears to vary substantially among the various HLA-Cw6-Bx haplotypes in geographically separated populations (Ozawa et al. 1981; Roitberg-Tambur et al. 1994; Ikaheimo et al. 1996; Kim et al. 2000). This observation could be explained by variable frequencies of recombination between Cw6 and RH1 in different population isolates. Detailed studies of marker haplotypes across RH1 and its flanking regions in various world populations will allow these theories to be tested.

## Acknowledgments

Our special thanks go to all the study subjects and referring physicians, for their participation in this study. We thank Drs. Christopher Bichakjian, Margaret Terhune, Michael Woodbury, Winfried Lenk, and Jaques Horstmann, for their assistance with clinical evaluation; Hilke Clasen, Claudia Gier, and Dorit Schuster, for HLA genotyping; and Dr. Wolfgang Müller-Ruchholtz, for advice and support. We also thank the National Psoriasis Foundation, for providing families from the National Psoriasis Foundation Tissue Bank and for its ongoing support of genetics research in psoriasis. This research was supported by U.S. Public Health Service (USPHS) awards P30 HG00209-03 and R01 AR4274-01 (to J.T.E., R.P.N., J.J.V., and N.J.S.), German Research Foundation award DFG-WE 905/1-1 (to T.H., S.J., and E.C.), the Ann Arbor Veterans Administration Hospital (support to J.T.E.), and the Babcock Memorial Trust. Portions of these studies were conducted at the General Clinical Research Center at the University of Michigan, which is funded by National Center for Research Resources, National Institutes of Health, USPHS grant M01EE00042.

## Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

GenBank, <http://www.ncbi.nlm.nih.gov/Genbank/index.html> (for MHC [accession number NT\_001520])  
Genome Database <http://www.gdb.org> (for markers designated as "M6Sxxx" in fig. 1)

Center for Medical Genetics, Marshfield Medical Research Foundation, <http://www.marshmed.org/markers/> (for reference allele sizes)  
Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for PSORS1 [MIM 177900] and CDSN [MIM 602593])  
PEDMANAGER, <ftp://ftp-genome.wi.mit.edu/distribution/software/pedmanager/>  
Psoriasis Genetics Laboratory, Department of Dermatology, University of Michigan, <http://www.psoriasis.umich.edu/hla2000/index.html> (for additional data not presented because of space constraints)  
Sanger Centre, <http://www.sanger.ac.uk/HGP/Chr6/MHC.shtml> (for the complete HLA sequence)

## References

- Allen MH, Veal C, Faassen A, Powis SH, Vaughan RW, Trembath RC, Barker JN (1999) A non-HLA gene within the MHC in psoriasis. *Lancet* 353:1589–1590
- Balendran N, Clough RL, Arguello JR, Barber R, Veal C, Jones AB, Rosbotham JL, et al (1999) Characterization of the major susceptibility region for psoriasis at chromosome 6p21.3. *J Invest Dermatol* 113:322–328
- Bhalerao J, Bowcock AM (1998) The genetics of psoriasis: a complex disorder of the skin and immune system. *Hum Mol Genet* 7:1537–1545
- Burden AD, Javed S, Bailey M, Hodgins M, Connor M, Tillman D (1998) Genetics of psoriasis: paternal inheritance and a locus on chromosome 6p. *J Invest Dermatol* 110:958–960
- Christophers E, Mrowietz U (1999) Psoriasis. In: Freedberg IM, Eisen AZ, Wolff K, Austen KF, Goldsmith LA, Katz SI, Fitzpatrick TB (eds) *Dermatology in general medicine*. Vol. 1. McGraw-Hill, New York, pp 495–521
- Collins A, Lonjou C, Morton NE (1999) Genetic epidemiology of single-nucleotide polymorphisms. *Proc Natl Acad Sci USA* 96:15173–15177
- Curtis D (1997) Use of siblings as controls in case-control association studies. *Ann Hum Genet* 61:319–333
- Curtis D, Sham PC (1995) A note on the application of the transmission disequilibrium test when a parent is missing. *Am J Hum Genet* 56:811–812
- Degli-Esposti MA, Abraham LJ, McCann V, Spies T, Christiansen FT, Dawkins RL (1992a) Ancestral haplotypes reveal the role of the central MHC in the immunogenetics of IDDM. *Immunogenetics* 36:345–356
- Degli-Esposti MA, Andreas A, Christiansen FT, Schalke B, Albert E, Dawkins RL (1992b) An approach to the localization of the susceptibility genes for generalized myasthenia gravis by mapping recombinant ancestral haplotypes. *Immunogenetics* 35:355–364
- Degli-Esposti MA, Leaver AL, Christiansen FT, Witt CS, Abraham LJ, Dawkins RL (1992c) Ancestral haplotypes: conserved population MHC haplotypes. *Hum Immunol* 34:242–252
- Elder JT, Nair RP, Guo SW, Henseler T, Christophers E, Voorhees JJ (1994) The genetics of psoriasis. *Arch Dermatol* 130:216–224
- Enerback C, Martinsson T, Inerot A, Wahlstrom J, Enlund F,

- Yhr M, Swanbeck G (1997) Evidence that HLA-Cw6 determines early onset of psoriasis, obtained using sequence-specific primers (PCR-SSP). *Acta Derm Venereol* 77:273-276
- Farber EM, Nall ML (1974) The natural history of psoriasis in 5,600 patients. *Dermatologica* 148:1-18
- Gaudieri S, Leelayuwat C, Tay GK, Townend DC, Dawkins RL (1997) The major histocompatibility complex (MHC) contains conserved polymorphic genomic sequences that are shuffled by recombination to form ethnic-specific haplotypes. *J Mol Evol* 45:17-23
- Henseler T, Christophers E (1985) Psoriasis of early and late onset: characterization of two types of psoriasis vulgaris. *J Am Acad Dermatol* 13:450-456
- Ikaheimo I, Silvennoinen-Kassinen S, Karvonen J, Jarvinen T, Tiilikainen A (1996) Immunogenetic profile of psoriasis vulgaris: association with haplotypes A2,B13,Cw6,DR7,DQA1\*0201 and A1,B17,Cw6,DR7,DQA1\*0201. *Arch Dermatol Res* 288:63-67
- Jenisch S, Henseler T, Nair RP, Guo S-W, Westphal E, Stuart P, Krönke M, et al (1998) Linkage analysis of human leukocyte antigen (HLA) markers in familial psoriasis: strong disequilibrium effects provide evidence for a major determinant in the HLA-B/C region. *Am J Hum Genet* 63:191-199
- Jenisch S, Koch S, Henseler T, Nair RP, Elder JT, Watts CE, Westphal E, et al (1999a) Corneodesmosin gene polymorphism demonstrates strong linkage disequilibrium with HLA and association with psoriasis vulgaris. *Tissue Antigens* 54:439-449
- Jenisch S, Nair RP, Henseler T, Elder JT, Marxen B, Krönke M, Westphal E (1997) Association of type I psoriasis with the Cw\*0802-DRB1\*0102-DQB1\*0501 haplotype in North-American multiplex families. In: Charron D (ed) Genetic diversity of HLA: functional and medical implications. Proceedings of the Twelfth International Histocompatibility Workshop. EDK Medical and Scientific, Sevres, France, pp 712-715
- Jenisch S, Westphal E, Nair RP, Stuart P, Voorhees JJ, Christophers E, Krönke M, et al (1999b) Linkage disequilibrium analysis of familial psoriasis: identification of multiple disease-associated MHC haplotypes. *Tissue Antigens* 53:135-146
- Jorde LB, Bamshad M, Rogers AR (1998) Using mitochondrial and nuclear DNA markers to reconstruct human evolution. *Bioessays* 20:126-136
- Kim TG, Lee HJ, Youn JI, Kim TY, Han H (2000) The association of psoriasis with human leukocyte antigens in Korean population and the influence of age of onset and sex. *J Invest Dermatol* 114:309-313
- Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 58:1347-1363
- Leder RO, Farber EM (1997) The variable incidence of psoriasis in sub-Saharan Africa. *Int J Dermatol* 36:911-919
- Leder RO, Mansbridge JN, Hallmayer J, Hodge SE (1998) Familial psoriasis and HLA-B: unambiguous support for linkage in 97 published families. *Hum Hered* 48:198-211
- Mallon E, Newson R, Bunker CB (1999) HLA-Cw6 and the genetic predisposition to psoriasis: a meta-analysis of published serologic studies. *J Invest Dermatol* 113:693-695
- MHC Sequencing Consortium (1999) Complete sequence and gene map of a human major histocompatibility complex. *Nature* 401:921-923
- Nair R, Guo S, Jenisch S, Henseler T, Lange E, Terhune M, Westphal E, et al (1995) Scanning chromosome 17 for psoriasis susceptibility: lack of evidence for a distal 17q locus. *Hum Hered* 45:219-230
- Nair RP, Henseler T, Jenisch S, Stuart P, Bichakjian CK, Lenk W, Westphal E, et al (1997) Evidence for two psoriasis susceptibility loci (HLA and 17q) and two novel candidate regions (16q and 20p) by genome-wide scan. *Hum Mol Genet* 6:1349-1356
- Oka A, Tamiya G, Tomizawa M, Ota M, Katsuyama Y, Makino S, Shiina T, et al (1999) Association analysis using refined microsatellite markers localizes a susceptibility locus for psoriasis vulgaris within a 111 kb segment telomeric to the HLA-C gene. *Hum Mol Genet* 8:2165-2170
- Ozawa A, Ohkido M, Tsuji K (1981) Some recent advances in HLA and skin diseases. *J Am Acad Dermatol* 4:205-230
- Roitberg-Tambur A, Friedmann A, Tzofoni EE, Battat S, Ben Hammo R, Safirman C, Tokunaga K, et al (1994) Do specific pockets of HLA-C molecules predispose Jewish patients to psoriasis vulgaris? *J Am Acad Dermatol* 31:964-968
- Schmitt-Egenolf M, Eiermann TH, Boehncke WH, Ständer M, Sterry W (1996) Familial juvenile onset psoriasis is associated with the human leukocyte antigen (HLA) class I side of the extended haplotype Cw6-B57-DRB1\*0701-DQA1\*0201-DQB1\*0303: a population- and family-based study. *J Invest Dermatol* 106:711-714
- Schork NJ, Thiel B, St Jean P (1998) Linkage analysis, kinship, and the short-term evolution of chromosomes. *J Exp Zool* 282:133-149
- Spielman RS, Ewens WJ (1996) The TDT and other family-based tests for linkage disequilibrium and association. *Am J Hum Genet* 59:983-989
- Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 52:506-516
- Tazi Ahnini R, Camp NJ, Cork MJ, Mee JB, Keohane SG, Duff GW, di Giovine FS (1999) Novel genetic association between the corneodesmosin (MHC S) gene and susceptibility to psoriasis. *Hum Mol Genet* 8:1135-1140
- Terwilliger JD, Ott J (1994) Handbook of human genetic linkage. Johns Hopkins University Press, Baltimore
- Tiilikainen A, Lassus A, Karvonen J, Vartiainen P, Julin M (1980) Psoriasis and HLA-Cw6. *Br J Dermatol* 102:179-184
- Todd JA. Multifactorial diseases: ancient gene polymorphism at quantitative trait loci and a legacy of survival during our evolution. In: Scriver CR, Beaudet A, Sly WS, Valle D, Vogelstein B, Childs B (eds) The metabolic and molecular bases of inherited disease, 8th ed. McGraw-Hill, New York (in press)
- Todd JA, Mijovic C, Fletcher J, Jenkins D, Bradwell AR, Barnett AH (1989) Identification of susceptibility loci for insulin-dependent diabetes mellitus by trans-racial gene mapping. *Nature* 338:587-589

- Trembath RC, Clough RL, Rosbotham JL, Jones AB, Camp RDR, Frodsham A, Browne J, et al (1997) Identification of a major susceptibility locus on chromosome 6p and evidence for further disease loci revealed by a two stage genome-wide search in psoriasis. *Hum Mol Genet* 6:813–820
- Tsuji K, Aizawa M, Sasazuki T (1992) HLA 1991: Proceedings of the Eleventh International Histocompatibility Workshop and Conference. Vol. 1. Oxford University Press, Oxford, pp 1091–1093
- Xiong M, Guo S-W (1997) Fine-scale genetic mapping based on linkage disequilibrium: theory and applications. *Am J Hum Genet* 60:1513–1531
- Yip SY (1984) The prevalence of psoriasis in the Mongoloid race. *J Am Acad Dermatol* 10:965–968
- Youn JI, Park BS, Park SB, Kim SD, Suh DH (1999) Characterization of early and late onset psoriasis in the Korean population. *J Dermatol* 26:647–652
- Zhou Y, Chaplin D (1993) Identification in the HLA class I region of a gene expressed late in keratinocyte differentiation. *Proc Natl Acad Sci USA* 90:9470–9474