S Gene (Corneodesmosin) Diversity and its Relationship to Psoriasis; High Content of cSNP in the HLA-Linked S Gene

Charlotta Enerbäck,*† Fredrik Enlund,* Annica Inerot,† Lena Samuelsson,* Jan Wahlström,* Gunnar Swanbeck,† and Tommy Martinsson*

*Department of Clinical Genetics and †Department of Dermatology, Sahlgrenska University Hospital/Östra, Göteborg, Sweden

Psoriasis is a heterogeneous disease in which several reports suggest the presence of a susceptibility gene in or in the proximity of the human leukocyte antigen complex in chromosome 6p. There is an association between HLA-Cw6 and young onset of the disease. The S gene (corneodesmosin), located 160 kb telomeric of HLA-C, is a strong candidate for psoriasis due to its reportedly exclusive expression in differentiating keratinocytes. We have studied this gene in a large Swedish psoriasis population and we report a strikingly high degree of polymorphism in the coding parts of the gene, 1 every 100 base pairs. We used a stratified approach to compare the polymorphic variants in patients and controls. A single nucleotide polymorphism in the coding region leading to an amino acid exchange (Ser→Phe) that differed significantly between patients and controls was identified (position 619). Owing to a high allele frequency in a larger control group, however, and an insignificant influence of the variant on the age at onset distribution curve based on a large psoriasis population, we could not confirm that this coding single nucleotide polymorphism was involved in disease etiology. We also examined the single nucleotide polymorphism in position 1243, recently proposed to have an influence on the pathogenesis of the disease. This polymorphism showed less association to the disease as compared with the single nucleotide polymorphism at positions 619 and 722. Such a high degree of variation present also in an HLA gene which is not involved in immune response indicates the difficulty involved in assessing the role of a specific allele in the pathogenesis of a complex disease in this region. A strong association effect due to linkage disequilibrium in an extended region in the HLA complex is also a complicating factor. ***Key words: corneodesmosin/HLA/psoriasis/S gene. J Invest Dermatol 114:1158–1163, 2000***

Psoriasis is a heterogeneous skin disease characterized by hyperproliferation of the epidermal keratinocytes and a dermal T cell infiltrate, predominately $CD4^+$ cells. We have previously presented a large population study suggesting a recessive mode of inheritance in our Swedish patient material (Swanbeck *et al*, 1994, 1995, 1997). The association of the disease to certain human leukocyte antigens (HLA), based on serology, has recently been confirmed by molecular based methods (Enerbäck *et al*, 1997; Mallon *et al*, 1997). The most consistently reported association has been to Cw6, an allele at the HLA-C locus (Tiilikainen *et al*, 1980). The frequency of Cw6 is higher among patients with young onset of the disease, especially in our population (Enerbäck *et al*, 1997). This association is not fully understood and it has not been made clear whether the association reflects the fact that Cw6 is the true disease-causing allele or is due to linkage disequilibrium with a nearby located gene. The fact that only 10% of Cw6-positive individuals are affected with psoriasis may favor the second alternative. The existence of a psoriasis gene in the HLA complex has been suggested by linkage studies in many populations, including ours (Nair *et al*, 1997; Trembath *et al*, 1997; Leder *et al*, 1998; Enlund *et al*, 1999). Further refinement of the region has been made possible by linkage disequilibrium effects, which have suggested the region containing HLA-B and -C, i.e., the β region of the HLA region (Leelayuwat *et al*, 1995), as the major candidate region (Jenisch *et al*, 1998). Our own data (Enlund *et al*, 1999) give the maximum LOD score value somewhat distal to the Genethon marker D6S276, which is considered to be located telomeric to the HLA region. A gene designated S, located about 160 kb telomeric of HLA-C in the class 1 region (Zhou and Chaplin, 1993), has recently been shown to encode the corneodesmosin protein, a protein that plays a major part in stratum corneum cohesion (Guerrin *et al*, 1998). Corneodesmosin is synthesized in the latest stages of keratinocyte differentiation and is located in the intercellular structures derived from desmosomes, i.e., the corneodesmosomes (Lundstrom and Egelrud, 1990; Serre *et al*, 1991). After further differentiation, the corneodesmosin is finally proteolyzed at the skin surface. This degradation is considered to be essential for normal desquamation to occur (Haftek *et al*, 1997). Before degradation, the protein has a molecular mass of approximately 55 kDa. The amino acid sequence contains a very high proportion of serine (27.5%), a feature shared by some other epidermal proteins, such as filaggrin, loricrin, and keratins 1 and 10 (Zhou and Chaplin, 1993). Corneodesmosin has previously been shown to be intensively expressed in psoriatic skin lesions, as well as other hyperproliferative skin disorders (Haftek *et al*, 1997). The

Manuscript received May 28, 1999; revised March 17, 2000; accepted for publication March 18, 2000.

Reprint requests to: Dr. C. Enerbäck, Department of Clinical Genetics, Sahlgrenska University Hospital/Östra, S-416 85, Göteborg, Sweden. Email: charlotta.enerback@sahlgrenska.se

Abbreviations: cSNP, coding single nucleotide polymorphism; TDT, transmission disequilibrium test.

gene is composed of two exons, separated by a 2.9 kb intron (Zhou and Chaplin, 1993). The corneodesmosin nucleotide sequence differs somewhat from the sequence initially reported for the S gene (Guerrin *et al*, 1998). An insertional G at position 1515 gives rise to an extension of the reading frame for another 27 amino acids. Moreover, the translational initiating site is now proposed to be located at nt15–17, in contrast to the earlier proposed nt63–65 in the S gene sequence (Zhou and Chaplin, 1993). This extends the sequence by 129 nucleotides.

Ishihara *et al* (1996) analyzed the S gene by direct sequencing in a Japanese population. The psoriasis patients were serologically typed as either Cw6 or Cw7 positive, although the association to the HLA region is much weaker in Japan (Ozawa *et al*, 1981; Nakagawa *et al*, 1991). Nine polymorphisms in the initial S gene sequence were found, with an equal distribution in patients and controls.

Recently, the involvement of the S gene in the pathogenesis of psoriasis has been suggested based on association analyses of position 1243. Allen *et al* (1999) showed linkage and association for an allele in the S gene using the transmission disequilibrium test (TDT) ($p = 0.000003$). Tazi Ahnini *et al* (1999) showed a higher odds ratio for homozygosity for the high risk allele in position 1243 (C/C) and coinheritance of Cw6 than for Cw6 alone. This was not shown for the young onset group, however, where the association to Cw6 is most pronounced. Furthermore, the same group recently showed a similar picture for guttate psoriasis, with an increased relative risk of the disease in the presence of a defined allele in the S gene. Cw6 conferred a higher relative risk but in the absence of Cw6 the S gene haplotype still showed a weak but significant association to the disease (Tazi Ahnini *et al*, 1999).

Interestingly, the major histocompatibility complex sequencing consortium recently presented the complete sequence and gene map of the HLA region (The MHC Sequencing Consortium, 1999). These authors now report that the candidate region for psoriasis has been narrowed down to a critical segment containing four genes, one of them the corneodesmosin (S) gene.

The aim of this study was to examine the potential pathogenic role of this interesting candidate gene in a Swedish psoriasis population, known to have a very strong association between psoriasis and the HLA-Cw6 allele (Enerbäck *et al*, 1997).

MATERIALS AND METHODS

Subjects In collaboration with the Swedish Psoriasis Association, blood samples were collected from 535 persons in 104 families, where at least two siblings were affected. The total number of affected siblings were 232 (153 affected sibpairs) with a median age of 41 y. The families consisted of 84 families with two, 16 families with three, and four families with four affected children. Only patients with psoriasis vulgaris were included; patients with arthritis or pustular forms were only included if they also showed signs of psoriasis vulgaris in the skin. Both parents were unaffected, in accordance with a recessive model. These families were examined thoroughly by one dermatologist (AI). In the TDT analysis, we used 232 parent-offspring trios from 104 independent psoriatic kindreds. Non-parametric linkage was based on the degree of allele sharing identical by descent. A maximum NPL value was reached at locus D6S276. Individuals in whom two were identical by descent could not be excluded at this locus (i.e., two siblings that have inherited two alleles from a common ancestor), were designated as HLA linked. These selected patients were chosen for the protein truncation test and direct sequencing. The patients were also Cw6 positive and those selected for sequencing also homozygous T for a polymorphism at position 722 in the S gene sequence.

Blood samples from 98 blood donors from the local hospital were also used in the study. They originate from the Swedish population. They were also analysed with respect to the polymorphism at position 722 in order to obtain a stratified control group that could be matched to the psoriasis patient sample with the same variant at this position.

General methods Genomic DNA was obtained from 10 ml blood samples using standard procedures (Ausubel, 1995). Primers were synthesized on an ABI 392 DNA/RNA Synthesizer. Polymerase chain reaction (PCR) was performed in a total volume of 20 µl with 125 ng genomic DNA, 20 pmol of each primer, 200 µM of each NTP, 1.5 mM $MgCl_2$, and 0.5 U Taq polymerase. For PCR conditions, see **Table I**.

Table I. Oligonucleotides used for PCR amplification of the S gene according to the position on the cDNA reference sequence[a]

Location	Primer	Position in gene sequence	Annealing temperature
Ex1	Ex1 f	nt 18–36	60°C
Intr1	Ex1 r	[b]	60°C
Ex2	S1:1	nt 100–119	57°C
Ex2	S3:2	nt 931–950	58°C
Ex2	S4:1	nt 891–910	56°C
Ex2	S5:2	nt 1583–1603	56°C
Ex2	S15	nt 1112–1131	56°C
Ex2	S16	nt 1302–1323	56°C
Ex2	619	nt 511–530[c]	56°C

[a]Genebank/EBI Data Bank, accession number AF030130.
[b]Intron sequence CGA CCA TCC AGT GAG GAG CA.
[c]Sequencing primer.

Analysis of the presence of HLA-Cw6 In the PCR-based typing of Cw6, we used sequence-specific primers, as described by Bunce *et al* (1995; Enerbäck *et al*, 1997). These primers amplify only a specific allele, due to exact matching to the nucleotide sequence specific for the serologic identity. The presence or absence of amplified product was determined after agarose-gel electrophoresis. This method does not distinguish heterozygosity from homozygosity for Cw6.

Genomic analysis The gene sequence and exon/intron boundaries were analyzed in the DNA clone UWGC: y24c203 (accession number: AC006163) obtained from Genemap 98 (http://www.ncbi.nlm.nih.gov/genemap98/). The sequence was tested for promoter properties by the NNPP promoter prediction program available at web site http://www-hgc.lbl.gov/projects/promoter.html

Protein truncation test Protein products were synthesized in a coupled *in vitro* transcription translation reaction using the TNT Coupled Reticulocyte Lysate System (Promega, Madison, WI), and analyzed by gel electrophoresis. Truncated proteins can be discriminated from full-size wild-type products and the size of the protein indicates the site of the mutation (Roest *et al*, 1993).

Restriction enzyme study

Position 722 Exon 2 was amplified by use of the primers S1:1 and S5:2, see **Table I**. The nonpolymorphic restriction sites at positions 370, 400, 1297, and 1511 gave rise to four cleavages; two at each side of the fragment. In this way, the polymorphism in position 722 was detected in the remaining 837 bp fragment (between restriction enzyme site 400 and 1297).

Position 1243 Primers flanking the polymorphism (S15 and S16) were used as previously described (Ishihara *et al*, 1996). Ten microliters of the PCR product was incubated with either *Msp1* (position 722) or *Hph1* (position 1243) and analyzed by agarose electrophoresis using standard procedures.

Sequencing analysis Exon 1 (nt18–99) was amplified with one primer pair (Ex1 f and Ex1 r) whereas exon 2 (nt100–1603) was amplified with two, overlapping, primer pairs (S1:1 and S3:2; S4:1 and S5:2), selected from the corneodesmosin nucleotide sequence (GenBank/EBI Data Bank, accession number AF030130) (**Table I**). This sequence is identical with that of the S gene, except for an elongation of the sequence of 129 nucleotides due to a probable earlier sequencing error in position 1515 leading to a shifted reading frame. Primers from the intron sequence were used for amplification of exon 1 whereas primers located in the exon were used for exon 2. This led to the total examination of exon 1 whereas 20 bp in both ends of exon 2 were not analyzed. The PCR products were purified with the QIA Quick PCR Purification Kit (Qiagen, Hilden, Germany) and sequenced by the direct sequencing method using the BigDye Terminator Cycle Sequencing Ready Reaction kit with AmpliTaq DNA polymerase, FS (PE Biosystems, Foster City, CA) and migrated on to an ABI PRISM 310 electrophoresis system (PE Biosystems). The PCR primers were also used as sequencing primers. Cycle conditions were chosen according to the manufacturer's instructions. The sequences were

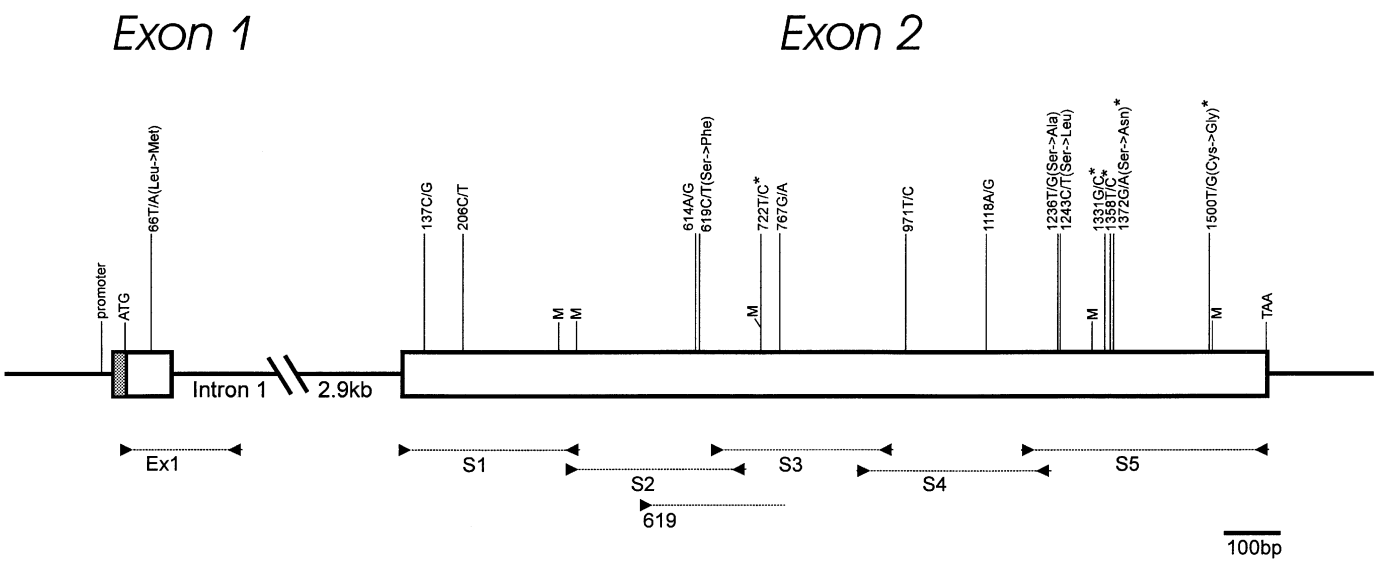

Figure 1. Genomic structure of the S gene. Exon 1 (nt18–99) and exon 2 (nt100–1603) are separated by a 2.9 kb intron. All polymorphic sites are given, the first letter indicates the consensus sequence. New polymorphic sites are indicated by an asterisk. Amino acid exchange is described. Restriction enzyme recognition site for *Msp1* (M) are not polymorphic, except in position 722. Below the primer pairs are given, as listed in **Table I**.

analyzed by the Sequence Navigator 1.0.1 and Factura 1.2.Or6 software programs, which identified heterozygous substitutions, marked by color, and homozygous substitutions, marked by an asterisk. All variants were visually inspected and they were always confirmed from the same or the opposite strand. The electrophoretogram from all sequences was screened visually to avoid undetected heterozygotes. The positions of all varying sites were indexed with regard to their location on the baseline reference sequence for the S gene obtained from GenBank.

The sequencing analyses of position 614 and 619 in 83 families and 73 controls were performed on an Perkin Elmer/ABI 377 automated Sequencer. The sequencing primer was located at a position 100 bp upstream from nt619.

Statistical analysis The TDT is a test for linkage in the presence of linkage disequilibrium (Spielman and Ewens, 1996). The TDT statistic tests for equal numbers of transmissions of marker alleles from heterozygous parents to affected offspring. Significantly different transmission provides evidence that the marker is linked to the disease locus. We used the Extended Transmission Disequilibrium Test computer program (Sham and Curtis, 1995).

Allele-wise and genotype-wise distribution among patients and controls was calculated with the chi-square test.

RESULTS

Genomic analysis We compared the sequence for the S gene (corneodesmosin GenBank/EBI Data Bank, accession number AF030130) with the corresponding gene sequence in the recently published human DNA clone UWGC: y24c203. New data on intron position, promoter, and flanking sequences were obtained. The 2.9 kb intron is located at nt99 instead of nt295 as previously reported (Ishihara *et al*, 1996). Given this new sequence information, coding sequence of the S gene was therefore not analyzed by the Japanese group. The first exon extends from nt18 to nt99 whereas the second exon ends in position 1603. Using the NNPP program, we identified a tentative promoter region located at 67–17 nt upstream to the first AUG in position 15–17. As pointed out by Zhou and Chaplin (1993), there are three alternative transcription-initiating sites (**Fig 1**).

Protein truncation test The protein truncation test was used for seven Cw6-positive psoriatics derived from families showing linkage to the HLA region and two healthy controls. After the coupled transcription–translation reaction, all samples showed an equal and expected size of ≈55 kDa corresponding to exon 2. No truncating mutations were found.

Restriction enzyme study The biallelic polymorphism (SNP) at the nucleotide position 722 in the second exon was identified by digestion of the entire exon with the restriction enzyme *Msp1*. This enzyme was used to identify other polymorphisms in the S gene sequence in a Japanese psoriasis population (Ishihara *et al*, 1996). The position of this newly identified SNP, which has not previously been reported, was verified by sequencing. This polymorphism creates the restriction enzyme recognition site for *Msp1*. The functional consequence is uncertain, as the variants do not lead to an amino acid change. One hundred and four families and 98 healthy blood donors were screened for this cSNP using *Msp1*. The allele frequency for T and C was equal in the control group whereas the noncleaved variant identical with the S gene sequence (T) was more common among psoriatics. To examine the transmission of this variant from heterozygous parents, the TDT test was used yielding a χ^2 of 16.2 ($p = 0.000057$) for the transmission. In the HLA-Cw6 negative families ($n = 22$) T in position 722 was transmitted 14 times and not transmitted 11 times, a difference that is not statistically significant. Although based on few families, we interpret this finding as showing that the allele Cw6 in the HLA-C locus is in linkage disequilibrium with T in position 722 in the S gene and that they could both be in linkage disequilibrium with the susceptibility gene. The SNP in position 1243 was screened with the restriction enzyme *Hph1* in 104 families and analyzed with TDT resulting in a χ^2 of 8.5 ($p = 0.0035$).

Sequencing analysis Blood samples from 20 patients homozygous for the noncleaved variant of the polymorphism at nt722 (i.e., T/T), HLA-Cw6 positive and where two siblings identical by descent could not be excluded, were analyzed with direct sequencing. Blood samples from 20 healthy controls that showed the same cleavage pattern for the polymorphism at nt722 were also sequenced, in addition to 10 healthy controls where we had no prior knowledge of the nucleotide sequence. All the polymorphic sites for the first 10 analyzed individuals in each group are listed in **Table II**. In all patients and controls, we found an insertion of G in position 1515, which extends the reading frame 27 amino acids. This is regarded as a sequencing error in the published S gene sequence (Zhou and Chaplin, 1993) and is therefore not shown in **Table II**. In all, 15 SNP were found in the coding sequence of the gene, which means one SNP every 100 bp in the coding region. Some of these substitutions result in an amino acid exchange whereas others do not (**Table II**). The SNP are

Table II. SNPs according to the position in the reference sequence of the S gene[a]

	66 **T** **T/A** **Leu/Met**	137 **C**	206 **C**	614 **A**	619 **C** **C/A** **Ser/Phe**	722* **T**	767 **G**	971 **T**	1118 **A**	1236 **T** **T/G** **Ser/Ala**	1243 **C** **C/T** **Ser/Leu**	1331* **G**	1358* **T**	1372* **G** **G/A** **Ser/Asn**	1500* **T** **T/G** **Cys/Gly**
Q1	T/A	C/G	..	G/G	T/T	C/C	..	C/C	..	..	T/T	G/C	T/C	..	..
Q2	A/A	C/G	..	A/G	C/T	T/C	..	T/C	A/G	..	C/T	G/C	T/C	..	..
Q3	A/A	C/G	..	A/G	C/T	T/C	..	T/C	A/G	..	C/T	G/C	T/C	..	..
Q4	T/A	C/G	..	A/G	T/T	T/C	..	T/C	A/G	..	C/T	G/C	T/C	..	..
Q5	nd	C/G	..	A/G	C/T	T/C	..	T/C	A/G	..	T/T	G/C	..	..	..
Q6	T/A	C/G	C/T	A/G	C/T	T/C	..	T/C	A/G	..	..	G/C	T/C	..	..
Q7	A/A	C/G	C/T	A/G	C/T	T/C	..	T/C	A/G	..	C/T	G/C	..	..	..
Q8	A/A	..	..	..	C/T	..	..	..	G/G	..	..	..	..	..	..
Q9	A/A	C/G	..	A/G	T/T	T/C	G/A	T/C	A/G	T/G	C/T	G/C	T/C	G/A	..
Q10	T/A	C/G	C/T	G/G	T/T	C/C	..	C/C	..	..	..	C/C	T/C	..	T/G
PSO1	A/A	..	..	..	T/T	..	G/A	T/C	G/G	T/G	..	..	..	..	..
PSO2	A/A	..	..	..	T/T	..	..	..	G/G	..	..	..	..	..	..
PSO3	A/A	..	..	..	T/T	..	..	..	G/G	..	..	..	..	..	..
PSO4	A/A	..	..	..	T/T	..	..	..	G/G	..	..	..	..	..	..
PSO5	A/A	..	..	..	T/T	..	G/A	T/C	G/G	T/G	..	..	..	..	..
PSO6	A/A	..	..	..	T/T	..	G/A	T/C	G/G	T/G	..	..	..	..	..
PSO7	A/A	..	..	..	T/T	..	G/A	T/C	G/G	T/G	..	..	..	..	..
PSO8	A/A	..	..	..	C/T	..	..	..	G/G	..	..	..	..	..	..
PSO9	A/A	..	..	..	T/T	..	G/A	T/C	G/G	T/G	..	..	..	..	..
PSO10	A/A	..	..	..	T/T	..	G/A	T/C	G/G	..	..	..	..	..	..
T/T Q1	A/A	..	..	..	T/T	..	A/A	C/C	G/G	G/G	..		..	..	..
T/T Q2	A/A	..	..	..	C/T	..	G/A	T/C	G/G	T/G	..	..	..	..	..
T/T Q3	A/A	..	..	..	C/T	..	G/A	T/C	G/G	T/G	..	..	..	..	..
T/T Q4	A/A	..	..	..	C/T	..	G/A	T/C	G/G	T/G	..	..	..	..	..
T/T Q5	A/A	..	..	..	C/T	..	G/A	T/C	G/G	T/G	..	..	..	..	..
T/T Q6	A/A	..	..	..	C/T	..	G/A	T/C	G/G	T/G	..	..	..	..	..
T/T Q7	A/A	..	..	..	C/T	..	..	..	G/G	..	..	..	..	..	..
T/T Q8	A/A	..	..	..	C/T	..	G/A	T/C	G/G	T/G	..	..	..	..	..
T/T Q9	A/A	..	..	..	C/T	..	G/A	T/C	G/G	T/G	..	..	..	..	..
T/T Q10	A/A	..	..	..	C/T	..	G/A	T/C	G/G	T/G	..	..	..	..	..

[a]Q1–Q10, Unselected controls. PSO1–10, Patients selected as described in the *Materials and Methods* and T/T in position 722. T/T Q1–10, Controls selected to be T/T in position 722. Letters under each nucleotide position and .. indicate consensus sequence. *New identified cSNP. Amino acid substitutions are presented. Marked area: Patients and controls stratified to be T/T in position 722 differ significantly in position 619.

equally distributed in both exons. Unexpectedly, there was a higher degree of S gene sequence homology between the reference S gene sequence and the sequence of the psoriasis patients than between the reference sequence and that of the unselected controls (**Table II**). As expected, the patients showed a similar pattern of polymorphisms to the controls stratified to have the same variant in nt722, as these are expected to belong to the same subpopulation. There is a strong tendency for linkage disequilibrium over this short stretch of DNA, which is a well-known feature of the HLA region. For example, homozygosity for C at position 1243 was in strong linkage disequilibrium with homozygosity for T at position 722. Interestingly, however, the patients and controls homozygous T/T at position 722 differed significantly at a position 103 bp upstream, at position 619 (boxed in **Table II**). Nucleotide T in this position results in an amino acid exchange, serine to phenylalanine. The controls (n = 20) had an equal distribution of C and T (22 *vs* 18) in position 619, whereas the pattern of the 20 psoriatics was different, 32 T *vs* 8 C, resulting in an allele-wise p-value of 0.007. In order to evaluate the significance of T in position 619 further, 83 unselected psoriasis families with at least two affected siblings and healthy parents were sequenced, in addition to 73 randomly selected controls. The three polymorphisms in positions 614, 619, and 722 were examined. The allele frequency of T in position 619 in the control material was higher than expected from the controls stratified to have T in position 722. Of 146 alleles, 106 were T and 40 C, which corresponds to an allele frequency in the normal population of 72.6% for T. In 83 unrelated psoriatics the allele frequency of T in position 619 was 83.1% (p = 0.025). In the TDT analysis, position 619 did not give a more significant p-value than that we obtained in position 722. The number of families informative for the analysis was smaller, but even if we extrapolated the data to have the same amount of alleles available for transmission, the p-value would not exceed that for position 722 (data not shown). The haplotype having an A in position 614 and T in positions 619 and 722 was the haplotype that conferred the highest significance.

To be able to estimate the functional consequence of a T in position 619, we related the presence of T/C and T/T in position 619 to the age at onset of psoriasis in these individuals. The age at onset distribution curves for these two patient groups were compared with the age at onset curve for 2547 psoriasis patients. These individuals had the same mean age as the probands. We found no influence of either 619 (T/T) (n = 92) or 619 (T/C) (n = 34) on the age at onset distribution curve. These results are shown in **Fig 2**.

We have previously shown a significant peak for age at onset of 21 or younger for Cw6 carriers (Enerbäck *et al*, 1997). The higher correlation of HLA-Cw6 with young age at onset than that of a T in position 619 argues against the importance of this variant in the pathogenesis of the disease in our population.

DISCUSSION

The presence of a gene in the HLA region involved in the pathogenesis of psoriasis vulgaris is now well established. Owing to the association to certain HLA antigens, above all Cw6, the first

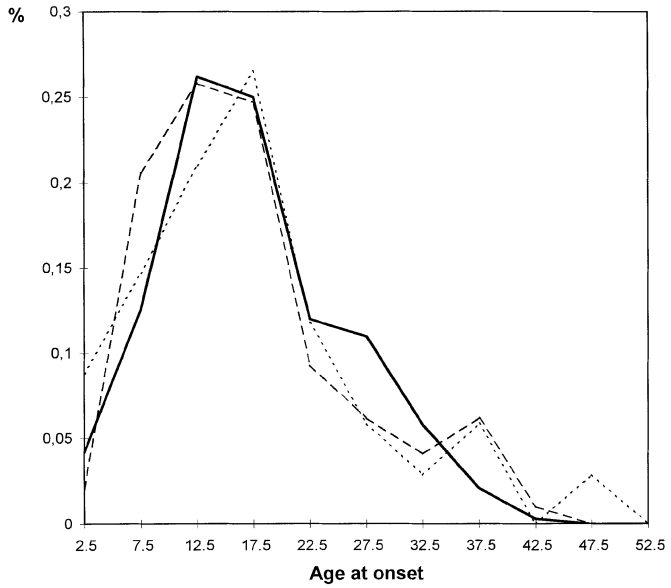

Figure 2. The presence of one or two T in position 619 does not seem to influence the age at onset when compared with a large patient material. Comparison of the age at onset distribution curves for two polymorphic variants in nucleotide position 619 of the S gene, T/T (n = 92) and T/C (n = 34), and the curve for 2547 Swedish psoriatics.

attempts to find linkage in population studies were made in this region. Several reports have been able to confirm that there are indications of a major susceptibility locus in the HLA region (Nair *et al*, 1997; Trembath *et al*, 1997; Leder *et al*, 1998; Enlund *et al*, 1999). This region, which spans around 3.5 cM, is estimated to contain over 80 coding sequences (Wells and Parnham, 1996). Many of these genes are involved in the immune response, such as the classical HLA genes HLA-A, HLA-B, and HLA-C. Genes with no obvious involvement in the immune system, however, are also present, such as the S gene (Zhou and Chaplin, 1993), coding for corneodesmosin (Guerrin *et al*, 1998).

One interesting feature of the classical HLA genes is their high degree of polymorphism, the highest reported in humans. The number of alleles of the HLA-C locus has recently reached 70, whereas for the HLA-B locus more than 240 alleles are known (HLA Informatics Group http.//www.anthonynolan.com/HIG/). The significance of this strikingly high allele frequency is not fully understood and has long been debated. The genetic mechanisms generating this variation are also under discussion. Gene conversion, intra-allelic recombination, and point mutation may all influence the diversity (Wells and Parnham, 1996).

The classical HLA genes encode proteins for receptors where the polymorphisms reflect differences in the specificity of the antigen-binding site. The degree of polymorphism in other human genes is at present under discussion. Some nonreceptor genes, such as several structural proteins, do not show any variations at all.

One recent publication described the SNP content in the lipoprotein lipase gene (LPL) (Nickerson *et al*, 1998). Of 88 variants detected in the LPL gene sequence, seven were biallelic polymorphisms in the coding sequence (SNP). The diversity, defined as the average number of nucleotide differences per base pair between any two chromosomes, was 0.0005 in coding DNA, which was reported to be 4-fold less than in noncoding DNA (0.0021). This could reflect a different effect of natural selection on coding and noncoding DNA.

In a material consisting of 10 healthy blood donors, we found 15 SNP in the 1585 bp coding region of the S gene, the product of which is corneodesmosin, an extracellular protein involved in keratinocyte adhesion. Six of these cSNP gave amino acids changes. This high degree of polymorphism (1 every 100 bp) corresponds to that found in the noncoding region of the LPL gene. The diversity in the coding region is 0.003; six times higher than in the LPL gene. It is surprising that such a variability is tolerated by a structural protein and it is tempting to speculate that the gene is under the influence of a different selection pressure in the HLA that moves towards variability. How such a sequence variation influences phenotypic variation is at this stage largely unresolved. The degree of variation would also have an impact on the possibility of detecting association to a disease allele.

We were interested in evaluating the possible importance of corneodesmosin in the pathogenesis of psoriasis: first, because it is a good candidate, considering its location closely telomeric of HLA-C, and secondly due to its specific expression in differentiating keratinocytes (Zhou and Chaplin, 1993). Ishihara *et al* (1996) reported nine SNP in a Japanese population, where the association to certain HLA types is not as prominent as in our population (Enerbäck *et al*, 1997). Another factor which warrants a new investigation is that the S gene sequence has recently been altered. We used HLA-Cw6-positive, HLA-linked psoriatics stratified for the presence of a strongly associated variation in the coding sequence of the gene itself and compared them with controls equally stratified, thereby providing a new way to assess the involvement of the gene in the susceptibility of the disease.

Although the patients and controls were stratified, they differed in a position close by, i.e., nt619. It is highly unlikely that this alteration is a recombination effect, due to the proximity to 722, where all 40 individuals were stratified to be homozygous T. These control patients did not show the same allele frequencies at position 619 as an unstratified control population. We interpret this to be an effect of linkage disequilibrium between 722 (T) and 619 (C) in the stratified control population and linkage disequilibrium between 722 (T) and 619 (T) in the stratified patient sample. The fact that 619 (T) is a common allele in both patients and controls obviously reduces the probability that this amino acid exchange is of functional importance in the development of the disease. In order to illustrate the influence of 619 (T), we compared the age at onset distribution curve for patients 619 (T/T) and 619 (T/C). A large psoriasis material was used for comparison. We found no significant difference between 619 (T/T) and 619 (T/C) and conclude that this position is not involved in psoriasis susceptibility; however, it is in strong linkage disequilibrium with the susceptibility gene.

The involvement of the S gene in the pathogenesis of psoriasis has recently been suggested, based on association studies of position 1243. This SNP leads to an amino acid exchange and is, according to our data (**Table II**), in linkage disequilibrium with the SNP in positions 619 and 722. We analyzed this polymorphism in 104 families using the restriction enzyme *Hph1*. In comparison with 722, the TDT test yielded a clearly higher p-value for 1243 (χ^2 16.2 for 722 and χ^2 8.5 for 1243), although the number of alleles available for transmission was not lower for 1243 than for 722. Also the frequency of the associated allele was equal in the general population (50%), which makes it possible to compare these two polymorphisms. For 1243, the risk allele was transmitted 165 times and not transmitted as much as 116 times, which makes this SNP less probable to be of causative significance for the disease.

To decide whether cSNP are involved in the etiology of a complex disease like psoriasis is not a straightforward task for many reasons. One is the known heterogeneity of the disease. Another problem in assessing the pathogenic effect of an associated variant of a gene is the lack of knowledge as to whether there is a single substitution or a combination of variants that influence the diseased phenotype. Certain variants in other genes may also contribute as well as an environmental component.

We conclude from this study that the S gene is not directly involved in the pathogenesis of psoriasis. There is a strikingly high degree of sequence variation in the coding regions studied and these cSNP are precisely localized dimorphisms in an interesting gene-rich region that could serve as mapping tools in other human diseases. We still have insufficient knowledge about the extent and nature of sequence variation in human genes. More research is

necessary to understand the pattern of diversity in different categories of human genes. A high degree of variation is a complicating factor in assessing the involvement of an allele in the pathogenesis of a complex disease. A strong association effect due to linkage disequilibrium in an extended region in the HLA complex is also a complicating factor; therefore, which site or combination of sites that may influence variation in risk of psoriasis could be difficult to predict.

We thank Staffan Nilsson for valuable help with the statistical analysis. This work was supported by the Swedish Psoriasis Association, the Welander Foundation, the Göteborg Medical Society, the Swedish Medical Research Council (B96–19X-11246–02B) and Glaxo-Wellcome Research Ltd.

REFERENCES

Allen MH, Veal C, Faassen A, Powis SH, Vaughan RW, Trembath RC, Barker JNWN: A non-HLA gene within the MHC in psoriasis. *Lancet* 353:1589–1590, 1999

Ausubel FM: *Current Protocols in Molecular Biology*. New York, Wiley Interscience, 1995

Bunce M, O'Neill CM, Barnardo M, Krausa P, Browning MJ, Morris PJ, Welsh KI: DNA typing for HLA-A, B, C, DRB and DQB1 by PCR-SSP. *Tissue Antigens* 46:355–367, 1995

Enerbäck C, Martinsson T, Inerot A, Wahlström J, Enlund F, Yhr M, Swanbeck G: Evidence that HLA-Cw6 determines early onset of psoriasis, obtained using sequence-specific-primers (PCR-SSP). *Acta Derm Venereol* 77:273–276, 1997

Enlund F, Samuelsson L, Enerbäck C, *et al*: Analysis of three suggested psoriasis susceptibility loci in a large Swedish material; confirmation of linkage to chromosome 6p (HLA-region), and to 17q, but not to 4q. *Hum Hered* 49:2–8, 1999

Guerrin M, Simon M, Montézin M, Haftek M, Vincent C, Serre G: Expression cloning of human corneodesmosin proves its identity with the S gene and allows improved characterisation of its processing during keratinocyte differentiation. *J Biol Chem* 273:22640–22647, 1998

Haftek M, Simon M, Kanitakis J, Marechal S, Claudy A, Serre G, Schmitt D: Expression of corneodesmosin in the granular layer and stratum corneum of normal and diseased epidermis. *Br J Dermatol* 137:864–863, 1997

Ishihara M, Yamagata N, Ohno T, *et al*: Genetic polymorphisms in the keratin-like S gene within the human major histocompatibility complex and association analysis on the susceptibility to psoriasis vulgaris. *Tissue Antigens* 48:182–186, 1996

Jenisch S, Henseler T, Nair RP, *et al*: 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, 1998

Leder RO, Mansbridge JN, Hallmayer J, Hodge S: Familial psoriasis and HLA-B: Unambiguous support for linkage in 97 published families. *Hum Hered* 48:198–211, 1998

Leelayuwat C, Pinelli M, Dawkins RL: Clustering of diverse replicated sequences in the MHC. evidence for en bloc duplication. *J Immunol* 155:692–698, 1995

Lundstrom A, Egelrud T: Evidence that cell shedding from plantar stratum corneum in vitro involves endogenous proteolysis of the desmosomal protein desmoglein I. *J Invest Dermatol* 94:216–220, 1990

Mallon E, Bunce M, Wojnarowska F, Welsh K: HLA-CW*0602 is a susceptibility factor in type I psoriasis, and evidence Ala-73 is increased in male type I psoriasis. *J Invest Dermatol* 109:183–186, 1997

Nair RP, Henseler T, Jenisch S, *et al*: Evidence for two psoriasis susceptibility loci (HLA and 17q) and two novel candidate regions (16q and 20p) by genome-wide scan. *Hum Mol Genet* 8:1349–1356, 1997

Nakagawa H, Akazaki S, Asahina A: Study of HLA class I, class II and complement genes in Japanese psoriatics. *Arch Dermatol Res* 283:281–284, 1991

Nickerson DA, Taylor SL, Weiss KM, *et al*: DNA sequence diversity in a 9.7-kb region of the human lipoprotein lipase gene. *Nature Genet* 19:233–240, 1998

Ozawa A, Ohkido M, Tsuji K: HLA haplotype in psoriasis (family study). *Jpn J Dermatol* 91:240, 1981

Roest PAM, Roberts RG, Van Sugino S, Ommen GJB, Den Dunnen JT: Protein truncation test (PTT) for rapid detection of translation-terminating mutations. *Hum Mol Genet* 2:1719–1721, 1993

Serre G, Mils V, Haftek M, *et al*: Identification of late differentiation antigens of human cornified epithelia, expressed in re-organised desmosomes and bound to cross-linked envelope. *J Invest Dermatol* 97:1061–1072, 1991

Sham PC, Curtis D: An extended transmission/ disequilibrium test (TDT) for multi-allele marker loci. *Ann Hum Genet* 59:323–326, 1995

Spielman RS, Ewens WJ: The TDT and other family-based tests for linkage disequilibrium and association. *J Hum Genet* 59:983–989, 1996

Swanbeck G, Inerot A, Martinsson T, Wahlström J: A population genetic study of psoriasis. *Br J Dermatol* 131:32–39, 1994

Swanbeck G, Inerot A, Martinsson T, Wahlström J, Enerback C, Enlund F, Yhr M: Age at onset and different types of psoriasis. *Br J Dermatol* 133:768–773, 1995

Swanbeck G, Inerot A, Martinsson T, *et al*: Genetic counselling in psoriasis: empirical data on psoriasis among first-degree relatives of 3095 psoriasis probands. *Br J Dermatol* 137:939–942, 1997

Tazi Ahnini R, Camp NJ, Cork MJ, Mee JB, Keohane SG, Duff GW, Giovine FS: Novel genetic association between the corneodesmosin (MHC S) gene and susceptibility to psoriasis. *Hum Mol Genet* 8:1135–1140, 1999

The MHC Sequencing consortium: Complete sequence and gene map of a human major histocompatibility complex. *Nature* 401:921–923, 1999

Tiilikainen A, Lassus A, Karvonen J, Vartiainen P, Julin M: Psoriasis and HLA-Cw6. *Br J Dermatol* 102:179–183, 1980

Trembath RC, Clough RL, Rosbotham JL, *et al*: 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, 1997

Wells RS, Parnham P: *HLA and MHC: Genes, Molecules and Function*. Oxford: BIOS Scientific, 1996, pp 77–94

Zhou Y, Chaplin DD: Identification in the HLA class I region of a gene expressed late in keratinocyte differentiation: *Proc Natl Acad Sci USA* 90:9470–9474, 1993