2008

Detailed genetic characterization of the interleukin-23 receptor in psoriasis

Steven J Schrodi

Detailed Genetic Characterization of the Interleukin-23 Receptor in Psoriasis

Veronica E. Garcia[1], Monica Chang[1], Rhonda Brandon[1], Yonghong Li[1], Nori Matsunami[2,3], Kristina P Callis-Duffin[4], Daniel Civello[1], Charles M Rowland[1], Nam Bui[1], Joseph J. Catanese[1], Gerald G. Krueger[3,4], Mark F. Leppert[2,4], Ann B. Begovich[1], Steven J Schrodi[1,*]

[1]Celera, 1401 Harbor Bay Parkway, Alameda, CA 94502, USA

[2]Department of Human Genetics, University of Utah, Salt Lake City, UT 84132, USA

[3]LineaGen Research Corporation, Salt Lake City, UT 84101, USA

[4]Department of Dermatology University of Utah, Salt Lake City, UT 84132, USA

* To whom correspondence should be addressed. Tel: 510-749-4270; Fax: 510-749-6250; E-mail: steven.schrodi@celera.com

Abstract

Using a multi-tiered, case-control association design, scanning 25,215 gene-centric SNPs, we previously identified two psoriasis susceptibility genes: *IL12B* and *IL23R*. These results have recently been confirmed. To better characterize the *IL23R* psoriasis-association, we used a fine mapping strategy to identify 59 additional *IL23R*-linked SNPs which were genotyped in our three independent, white North American sample sets (>2800 individuals *in toto*). A sliding window of haplotype association demonstrates co-localization of psoriasis susceptibility effects within the boundaries of *IL23R* across all sample sets, thereby decreasing the likelihood that neighboring genes, particularly *IL12RB2*, are driving the association at this region. Additional haplotype work identified two 5-SNP haplotypes with strong protective effects, consistent across our three sample sets (OR_{common} = 0.67; P_{comb} = 4.32E-07). Importantly, heterogeneity of effect was extremely low between sample sets for these haplotypes (P_{Het} = 0.961). Together, these protective haplotypes attain a frequency of 16% in controls, declining to 11% in cases. The characterization of association patterns within *IL23R* to specific predisposing/protective variants will play an important role in the elucidation of psoriasis etiology and other related phenotypes. Further, this work is essential to lay the foundation for the role of *IL23R* genetics in response to pharmaceutical therapy and dosage.

Introduction

Worldwide, psoriasis is the most common systemic autoinflammatory disorder affecting 0.5-3% of the general population. The disease is typified by overproliferation of keratinocytes and recruitment of immunocompetent cells, including CD4+ T lymphocytes, to the dermal tissues generating chronic inflammation.[1] Affecting joints and surrounding tissues in 10-30% of psoriatic patients, inflammatory arthritis markedly impacts mobility and can cause irreversible joint destruction. Several genetic and epidemiological studies have demonstrated that Crohn's disease, cardiovascular disease and possibly metabolic syndrome are co-morbid with psoriasis, particularly with more severe, earlier onset psoriasis.[2,3] Griffiths and Barker provide a current review of psoriasis etiology and clinical features.[4]

Although great strides have been made advancing the understanding of psoriasis etiology and pathogenesis, it is clear that much remains to be elucidated. It is well understood that psoriasis has a strong genetic component. Several family-based studies have shown marked increases in psoriasis risk as relatedness increases; most notably psoriasis risk rises from 20% in dizygotic twins to 70% in monozygotic twins.[5] For decades both association and linkage results have pointed to a strong predisposing effect from the MHC region, and recent association-based evidence has highlighted HLA-C*0602 as a likely primary candidate responsible for this effect.[6] Broadly speaking, the smaller susceptibility effects of non-MHC loci have taken longer to identify at high resolution. Following genome-wide linkage studies, evidence for psoriasis-associated variants segregating at 17q25 has been intriguing, albeit inconsistent in location and signal strength.[7-9] Recently, we have reported psoriasis-associated diplotypes at *IL12B*, confirming and extending the previous work by Tsunemi and colleagues[10], and predisposing/protective diplotypes at *IL23R* using a multi-staged, case-control design, scanning 25,215 gene-centric SNPs.[11] These results have been verified in four independent studies[12-15] and have paralleled association studies in related diseases. One of the *IL23R* missense SNPs implicated in psoriasis, R381Q, has also been strongly associated with several autoinflammatory phenotypes: Several studies have found

R381Q-mediated predisposition to adult inflammatory bowel disease (IBD), especially Crohn's disease[16-21] and pediatric Crohn's disease in individuals of European descent[22-24], two sizable studies reported association of multiple *IL23R* SNPs with ankylosing spondylitis (AS)[25,26] and a North American study demonstrated *IL23R* susceptibility for Graves' ophthalmopathy (GO).[27] *IL23R* polymorphisms have also been studied in celiac disease with mildly significant results[28] and multiple sclerosis with conflicting results.[25,28-31] In addition, Zhang and colleagues have recently reported a strong non-MHC psoriasis-association with haplotypes at *IL15* in Han Chinese[32] and we have published suggestive findings implicating *IL13* and/or other variants in the 5q31 cytokine-rich region.[33]

In this study we focused on the *IL23R* region and extended our initial genetic analysis by fine mapping the association signal over 336kb in our three case-control sample sets. Haplotype, association decay with linkage disequilibrium and conditional association analyses were performed to shed light on the psoriasis genetics in this region of chromosome 1p31.3.

Results

Genotyping for this study was performed on three independent sample sets consisting of white North American psoriasis cases and controls, totaling 1444 cases and 1382 controls. Basic demographic and clinical characteristics of these sample sets are described in a previous publication.[2] A type of genomic-control analysis was performed on pooled genotype data from the initial sample set which effectively ruled out large population stratification effects.[2]

Applying our fine mapping SNP selection algorithm (described in the **Materials and Methods**), we identified 59 additional SNPs for interrogation in our three sample sets, for a total of 61 SNPs covering 338kb (rs7530511, P310L; and rs11209026, R381Q were previously genotyped in all three sample sets). Thirty-one of these fine mapping SNPs were within the coding region or the 3'UTR of *IL23R*. **Figure 1a** shows a positional plot of the selected SNPs with the location of the genes in this region of

chromosome 1p31.3, with a more detailed view presented in **Figure 1b** focusing on individual SNPs genotyped within *IL23R*.

Of the 61 SNPs evaluated, eight had Mantel-Haenszel continuity-corrected P-values (MH P-values combine association evidence across the three samples sets, accounting for direction of effect) below 0.05. **Table S1** shows these results as well as allele frequencies, genotypic 2-df P-values and P-values for the exact test of Hardy-Weinberg equilibrium for each sample set. Mantel-Haenszel allelic odds ratio and 95% confidence intervals are also presented (these were calculated jointly across the three sample sets). Six of these eight significant SNPs reside within the *IL23R* coding region and the remaining two are located in the intergenic region between *IL23R* and *IL12RB2*. These eight SNPs include the original two missense SNPs previously reported, rs7530511 (P310L) and rs11209026 (R381Q).

Prior to tests of haplotype association, two types of graphical representations of linkage disequilibrium patterns were constructed: To characterize the pairwise correlation structure for the entire 338kb region, we generated a heatmap for cases and controls combined using the r^2 and D' statistics (**Figure 2a**). Individual pairwise LD values with corresponding SNPs are given in **Table S2**. All three sample sets were combined for this analysis. The r^2 heatmap shows an absence of solid block patterns. Rather, there are two very roughly-defined block structures in the region with slightly higher average r^2 values than the surrounding regions. These weak blocks are highly peppered with pairwise comparisons of low LD. More pronounced LD structure is displayed in the D' heatmap, with two or three blocks covering the region. Given the fine mapping SNP selection procedure employed where SNPs were genotyped if they exhibited high LD (as measured by r^2) with one of the originally associated missense SNPs (P310L or R381Q) and other SNPs were tagging SNPs reducing redundancy, we would not expect to observe strong block structures. From our r^2 data, the first weak block extends roughly from intron 3 of *C1orf141* into the 5' region of *IL12RB2* and the second weak block covers part of the first intron of *IL12RB2* through 28kb 5' of *SERBP1*.

As much of the association signal was driven by rs11209026 (R381Q) and other studies have identified this missense polymorphism as being strongly associated with the related phenotypes studied, we plotted the Mantel-Haenzsel P-value (combining the three

sample sets) as a function of r^2 with R381Q (**Figure 2b**). Under a model where R381Q is solely and causally responsible for the association patterns observed, we expect the approximate relationship: $\log P_M \approx r^2 \log P_D$; where P_M is the association P-value at a marker in linkage disequilibrium with the causative site, P_D is the association P-value at the causative site, in this case R381Q, r^2 is the pairwise LD measure between the two sites and equal numbers of genotypes are assayed at each site. This association decay analysis suggests that some SNPs in low LD with R381Q may independently contribute to disease status as they substantially depart from the expected relationship.

In our previous publication we constructed haplotypes for P310L and R381Q since haplotype association tests can be more informative than single marker test under many models where cis-effects play an important role.[34] Even though these SNPs were not in high LD, the number of double heterozygotes was small and hence linkage phase is unambiguous in the large majority of individuals. With these two missense polymorphisms, carrying the proline-arginine polypeptide-encoding gene shows susceptible effects whereas both the leucine-arginine and proline-glutamine polypeptides confer protective effects. For this study, we used the fine mapping data to scan this region positionally for haplotype effects using a sliding window approach. A window size of 3 adjacent SNPs was used. We present a positional plot of the global haplotype P-value for each window (**Figure 3**). The plot shows an analysis combined across sample sets using the Fisher's combined P-value method. The results indicate rather narrow peaks of association centering on *IL23R* intron 8 through intron 9 and including the R381Q polymorphism in exon 9 – a span of 12kb. Four SNPs generated this association signal: rs10789229, rs10889671, rs11209026 (R381Q), and rs10889674; with the first window (rs10789229-rs10889671-rs11209026) producing a combined global P-value (2-tailed in each sample set) of 1.28E-04 and the second window (rs10889671-rs11209026-rs10889674) yielding a combined global P-value (also 2-tailed in each sample set) of 6.42E-05. Although these haplotype association windows were less significant in comparison with rs11209026, further work (elaborated below) is required to determine if rs11209026 is the sole driver of this association peak due to differences in degrees of freedom and whether or not narrow hypotheses are tested. Through the association and LD analyses, it was apparent that although these four *IL23R* SNPs generated the peak

association signal, additional psoriasis-association effects may be possible through haplotypes derived from additional SNPs (*e.g.* P310L). Analysis of pairwise LD and association results indicates that within the eight SNPs having significant Mantel-Haenszel P-values, rs7530511 is highly correlated with rs10889671 (intron 8, *IL23R* SNP; r^2=0.943); and similarly, rs11209026 is highly correlated with rs11465804 (intron 8, *IL23R* SNP; r^2=0.852). We then proceeded to interrogate these data with several subsequent haplotype analyses.

We commenced another haplotype-based investigation by using the five SNPs that exhibited the strongest and most consistent single-SNP association signals: rs7530511, rs11465804, rs10889671, rs11209026 and rs1857292. These SNPs span 53kb from exon 7 in *IL23R* to the intergenic region between *IL23R* and *IL12RB2*. The five-SNP haplotypes were estimated and tested for association in the three sample sets. Five primary haplotypes were found (above 1% frequency), two of which conferred significant protection against psoriasis susceptibility (**Table 1a**). Together, these two (completely divergent) protective haplotypes, `TTAGT` and `CGGAA`, were present on 16% of control chromosomes, decreasing to 11% in cases. The effect of these protective haplotypes was consistent across sample sets (OR_{SS1} = 0.66, OR_{SS2} = 0.67, OR_{SS3} = 0.69), and the combined analysis was rather significant (P_{MH} = 4.32E-07). Importantly, the level of heterogeneity of effect was not significant across sample sets for the protective haplotype grouping of `TTAGT` and `CGGAA` versus all other haplotypes (P_{Het} = 0.961) as determined from the Mantel-Haenszel procedure to test odds ratios for homogeneity (see **Materials and Methods**).

We then systematically evaluated all possible combinations of these five SNPs in an exploratory analysis to see if we could eliminate one or more of these SNPs and retain or increase the significance of the association result. Eliminating rs11465804 and rs1857292 from the haplotypes yielded a simpler, slightly stronger association result for the protective haplotypes (`TAG` and `CGA` vs. Others; P_{MH} = 3.88E-08) (**Table 1b**), although it should be noted that the increase in significance may be due to stochastic effects given that many combinations of haplotypes were explored. Without these two SNPs, the frequencies of the resulting protective haplotypes increased to 20% in controls and 14% in cases.

Visual inspection of the LD, association and numerous haplotype results spawned a curiosity to better understand the physical extent of the protective haplotypes in this region. Hence, we expanded the haplotype analysis to include all contiguous markers such that the association signal was not substantially diminished by estimated historical recombination events. This region appeared to extend 55kb from P310L through the 3' region of *IL23R* to rs11209032 in the *IL23R-IL12RB2* intergenic region. Haplotype analysis was performed on all twenty-three markers in this region (**Table 2a**). Again, two common protective haplotypes were identified. Although the initial five- and three-marker protective haplotypes described above did not share alleles at any sites, the two protective haplotypes from the twenty-three marker analysis had twelve markers with the same alleles on both haplotypes. As no other common (>1%) haplotypes shared the alleles at these twelve markers, we reduced the analysis to those twelve markers and ran another haplotype analysis (**Table 2b**). Notably, the protective haplotype from this reduced set of SNPs, AGTTCCTCCCAG, carries substantial effects (freq in cases = 12%, freq in controls = 17%; OR_{MH} = 0.68; P_{MH} = 3.18E-07) and does not include P310L, R381Q or SNPs in high LD with these two missense SNPs (rs10889671, rs1857292, or rs11465804). In addition, this haplotype is remarkably similar in frequency and effect size across all three sample sets. This is not the case for the three marker haplotype (rs7530511-rs10889671-rs11209026) described previously with the TAG haplotype exhibiting the strongest protective effects in the Utah population-derived Sample Set 1 while the CGA haplotype was stronger in the remaining two sample sets (derived from the North American white population in general).

With this result, an immediate line of further inquiry is whether or not this 12-marker haplotype represents a variant contributing to psoriasis association independently of P310L and R381Q. To address this we dichotomized the 12-marker haplotypes into the protective haplotype described above and an aggregate of all other haplotypes, then we did the same for the rs7530511-rs11209026 haplotypes (TG and CA protective haplotypes combined together vs. CG and TA combined together). A diplotype-based, squared correlation coefficient r^2 statistic was then calculated between the two haplotype groupings. All individuals were used across all sample sets for this calculation. The resulting value, r^2=0.78, indicated a fairly high degree of correlation between the two

diplotype groupings, thereby suggesting that although they were not completely redundant, these were not independent effects.

The LD patterns and haplotype results appear to indicate that more than one polymorphism is contributing to the psoriasis association linked to *IL23R*. To formally investigate this, we performed a test of conditional association on SNPs having the most significant combined P-values ($P<0.005$) for the 2df genotype test and exhibiting significant Mantel-Haenzsel confidence intervals (95% CI excluding 1.0) for the allelic OR jointly calculated over the three sample sets. As some of these SNPs clustered into "LD groups" consisting of SNPs in very high LD and similar statistical significance for psoriasis association, a representative SNP was selected from each LD group when appropriate. Six SNPs met these criteria: the missense SNPs rs7520511 (P310L) and rs11209026 (R381Q), rs10889674 (putative transcription factor binding site, intron 9 *IL23R*), rs1857292 (3' of *IL23R*), rs11465804 (intron 8 of *IL23R*) and rs10889671 (intron 8 *IL23R*). Rs11465804 and rs10889671 were excluded from this analysis due to high LD with one of the missense SNPs (rs10889671-rs7530511 $r^2 = 0.943$; rs11465804-rs11209026 $r^2 = 0.852$) and since both missense SNPs had slightly elevated significance when compared to these LD counterparts. The conditional association permutation test revealed that the genotype association at R381Q remained significant after fixing the genotypes at P310L ($P_{comb}=0.00031$), or either of the remaining two SNPs ($P_{comb}=0.0183$, fixing genotypes at rs10889674; $P_{comb}=0.0027$, fixing genotypes at rs1857292). Conversely, the genotype association at P310L was also significant, albeit mildly so, following conditioning on R381Q genotypes ($P_{comb}=0.0299$); however the moderate LD between P310L and the other two SNPs removed the association at P310L. These results for the mutually conditionally independent association of the two missense SNPs were not unexpected given the very low amount of LD between these two SNPs. Hence, there is some evidence of at least two *IL23R*-linked polymorphisms independently contributing to psoriasis.

As other SNPs, not genotyped in this study, could possibly drive the association results observed here, we investigated the HapMap LD results for the CEU samples between genomic positions 67,225,114-67,725,113 on Build36.[35] Examining four key SNPs from this study, rs7530511, rs10889671, rs11465804, and rs11209026, we found

that seven SNPs were in substantial LD (r^2>0.50) with either rs7530511, rs10889671 or both: rs2863212 (*IL23R* intron 6), rs7528924 (*IL23R* intron 7), rs4655692 (*IL23R* intron 7), rs4655693 (*IL23R* intron 7), rs11804284 (*IL23R* intron 7), rs4655530 (*IL23R* intron 8), and rs2863209 (intragenic, within 8kb 3' of *IL23R*). The remaining two SNPs were only in substantial LD with each other ($r^2_{\text{rs11465804-rs11209026}} = 0.852$).

Discussion

We recently reported the finding that alleles at two missense SNPs segregating at *IL23R* were significantly associated with psoriasis. These results were replicated across our three sample sets and in external studies. Following up on our previous P310L-R381Q haplotype findings, fine mapping of the *IL23R*-linked region shows variants segregating at *IL23R* coding and flanking regions significantly associated with psoriasis. In particular, there are extended haplotypes in this region that protect against psoriasis susceptibility. Importantly, it also appears that at least two *IL23R* polymorphisms, P310L and R381Q, independently contribute to linkage disequilibrium with the psoriasis phenotype. In addition, both sliding window haplotype analyses and longer-range haplotype work pinpointed the association signal to the *IL23R* coding region. This is particularly important as the interleukin 12 receptor subunit-encoding gene, *IL12RB2*, is located 47kb from the 3' end of *IL23R* and some SNP pairs exhibit substantial and even perfect linkage disequilibrium between sites located in the coding regions of the two genes (as determined by genotyping in the CEU HapMap samples). Recent animal studies show that the *IL12RB2* knockout mouse develops an autoimmune/lymphoproliferative disorder with aberrant IL-12 signaling.[36] Hence, *IL12RB2* is a reasonable psoriasis candidate gene. However, the genetic results presented here seriously diminish the possibility that *IL12RB2* alleles are primarily responsible for the psoriasis predisposing effects that we have observed.

The HapMap project has general population genotype data on both missense SNPs, rs7530511 (P310L) and rs11209026 (R381Q).[35] At rs11209026, the A allele (minor allele) was found on CEU and YRI chromosomes (8 out of 120 CEU chromosomes and 2 out of 120 YRI chromosomes), but not on chromosomes from the

two East Asian samples (CHB and JPT). Hence, it is possible that rs11209026 may predispose some African and/or African-derived populations to autoimmunity and autoinflammatory traits, particularly if the effect size is larger in those subpopulations than European-derived samples. The rs7530511 SNP is polymorphic in all four HapMap sample sets, with varying frequencies: 15/120 CEU chromosomes, 2/90 CHB chromosomes, 3/88 JPT chromosomes, and 35/118 YRI chromosomes; thereby suggesting that the autoinflammatory effects ascribed to P310L for our North American white samples might translate to these other populations. For each of these SNPs, the minor allele in humans appears to be derived as many vertebrates including the chimpanzee, macaque, mouse, rat, cow, dog and chicken carry the major allele nucleotides at the orthologous sites.

Importantly, these genetic findings coupled with results from multiple areas of research ranging from molecular immunology to clinical biology implicate the IL-23/T_H-17 pathway as being central to chronic inflammatory conditions such as psoriasis and inflammatory bowel disease; the perturbation of which may disrupt the communication between the innate and adaptive immune responses. In sum, these studies demonstrate several key aspects of IL-23/T_H-17 pathobiology related to psoriasis: 1) Both IL-12p40 and IL-23p19 mRNA expression levels are significantly elevated in both non-lesional psoriatic skin versus normal skin as well as lesional psoriatic skin versus non-lesional psoriatic skin.[37,38] 2) IL-12 and IL-23 knockouts and IL23-deficient animal model experiments indicate that the systemic inflammatory effects, dermal inflammation and epidermal hyperplasia are often mediated through the IL-23/T_H-17 pathway[38,39], 3) T_H17 survival and expansion, key characteristics of epithelial inflammation and epidermal hyperplasia, occur in response to IL-23[40-42] 4) IL-23p19 antibodies inhibit proinflammatory cytokines in a mouse model of IBD[43], and 5) clinical studies have shown dramatic efficacy of IL-12p40 antibodies in reducing symptoms in a high percentage of psoriatic subjects[44,45] and those with active Crohn's disease.[46] These diverse studies have conspired to highlight the central function of the IL-23/T_H-17 axis in mediating chronic inflammatory disease pathogenesis, downplaying the role of IL-12. Hence, full genetic description of both *IL12B* and *IL23R*, genes encoding for critical proteins in the IL-23/T_H-17 response, is necessary to delineate specific variants

predisposing and protective of disease and to further our understanding of the molecular pathobiology of autoinflammatory phenotypes.

Along with psoriasis, *IL23R* appears to play an important role in predisposition to other autoinflammatory diseases including IBD (particularly adult and pediatric Crohn's disease)[16-24], AS[25,26], and GO.[27] It is possible that *IL23R* variants also underlie susceptibility to celiac disease[28], Graves' disease without ophthalmopathy[27] and multiple sclerosis[28,29], although the evidence is not currently strong for celiac disease and Graves' disease without opthalmopathy and is contradictory for multiple sclerosis.[25,30,31] Interestingly, multiple independent *IL23R* polymorphisms have been reported to be associated with AS, GO and Crohn's disease, suggesting a models of allelic heterogeneity within each disease where disruption of IL-23R function can occur from several distinct genetic insults. With AS, both R381Q and rs1343151 are replicated SNPs (only the former was associated with psoriasis in our study). Graves' disease offers a different view – unlike AS, Celiac disease, IBD and psoriasis, R381Q is not associated with Graves' disease or GO. Rather, two *IL23R* SNPs which are not psoriasis-associated, rs2201841 and rs10889677, are associated with GO; whereas P310L may be associated with Graves' disease. Although R381Q seems to be the major *IL23R* susceptibility polymorphism for Crohn's disease with the minor allele conferring protective effects as in psoriasis, other independent SNPs such as rs7517847 play a role in Crohn's but not psoriasis predisposition. Conversely, P310L appears to be significantly correlated with psoriasis, but not IBD. Understanding the complex, pleiotropic action of IL-23R and the specific variants that give rise to each of these autoimmune traits will undoubtedly increase progress in our understanding of the molecular pathophysiology underlying these chronic autoinflammatory conditions. Moreover, coupled with the recent advances in autoinflammatory biology, the *IL23R* variants described herein may lead to a deeper understanding of molecular action of targeted therapeutics such as the efficacious IL12/23 monoclonal antibodies[44-46] and prove useful in autoinflammatory pharmacogenetics.

Materials and Methods

Subjects

The subjects in all three sample sets were white North American individuals. Sample Set 1 consisted of 467 psoriasis cases and 500 controls, all residing in either Utah or southern Idaho. Sample Set 2 was obtained by the Genomics Collaborative Division of SeraCare Life Sciences (GCI) and included 498 cases and 498 controls. Lastly, BioCollections Worldwide and GCI provided Sample Set 3, composed of 481 cases and 424 controls. Details concerning these subjects were previously described.[11] All individuals included in this study were 18 years or older at time of enrollment. All protocols were approved by national and/or local institutional review boards, and informed written consent was obtained from all subjects.

Genotyping

Individual genotyping was performed using allele-specific kinetic PCR on 0.3 ng of DNA and the resulting data hand-curated prior to statistical analyses without knowledge of case/control status. Our genotyping accuracy is consistently better than 99%.[11]

SNP Selection

We undertook a multifaceted approach to identify SNPs to genotype individually in a fine-scale mapping effort in the *IL23R* region. A 336 kbp region was selected across a portion of *C1orf141* through *SERBP1*. This region was delineated on the basis of two criteria: 1) the decay of LD from the two *IL23R* SNPs, rs7530511 and rs11209026, originally identified to be associated in our sample sets[11], and 2) coverage of clear biological candidate genes nearby – in this case, *IL23R* and *IL12RB2*. Next, SNPs were selected in this 336kbp region to cover two genetic models: one of allelic heterogeneity where multiple variants segregating at the same gene or functional motif independently contribute to disease predisposition; and the second model where the association observed at the originally-identified SNPs, rs7530511 and rs11209026, was driven through LD with one or more untyped polymorphisms. To address these two possible models, we partitioned SNPs in the 336kbp region into those in moderate to high LD (r^2>0.20) with one of the original two associated SNPs, and those exhibiting weak LD with the original SNPs (r^2<0.20). The threshold value of r^2=0.20 was determined

analytically by solving for the r^2 value that would generate the observed results at these two original SNPs from an untyped marker having a reasonable disease model (relative risk below 2.25 with similar allele frequency). The r^2 values were calculated from the HapMap CEU data.[35] Next, we ran the tagging SNP program Redigo[47] on those SNPs in weak LD, selecting SNPs with the highest power to detect an arbitrary disease predisposing site in the region. Redigo uses a genotype-based approach that maximizes power to detect disease susceptibility SNPs. We then took all of the SNPs in the moderate-to-high LD group and reduced the set so that SNPs in extremely high LD (r^2>0.97) were represented by a single SNP. Lastly, any SNP with putative functional annotation was selected to be genotyped. In all, 61 SNPs, inclusive of the two SNPs fully genotyped in the original study, were identified and judged as being sufficient to cover both genetic models for the *IL23R* region.

Statistical Analysis

Several analyses were performed on individual SNPs. We used an in-house genetic analysis application to analyze much of the data. Hardy-Weinberg equilibrium testing was accomplished through the exact test of Weir.[48] A William's-corrected G-test was used to calculate P-values for genotypic association.[49] Approximate confidence intervals for the odds ratios were calculated using the typical estimate of the standard error of the log-odds ratio. The Mantel-Haenszel procedure to test odds ratios for homogeneity (test of heterogeneity of effect) was performed following Sokal and Rohlf (Chapter 17, Reference 49). P-values were combined across sample sets using either the continuity-corrected Mantel-Haenszel statistic (eqn 17.22, reference 49) or the Fisher's combined P-value (omnibus procedure).[50] Similarly, Mantel-Haenszel common odds ratios were calculated to combine data across sample sets.[51] A Monte Carlo simulation was written in XLISP-STAT to calculate 95% confidence intervals on the common odds ratios. Typically, 20,000 iterations of the Monte Carlo were performed unless results were not sufficiently converging, in which case 40,000 iterations were used.

Pairwise linkage disequilibrium was calculated using either the LDMax package where 2-SNP haplotypes were estimated through an EM algorithm and the standard r^2 statistic employed,[52,53] or, in some instances, we calculated an r^2 statistic using unphased

genotype or diplotype data. Given perfect phasing for the double heterozygotes and Hardy-Weinberg Equilibrium, these two methods yield identical values.

Sliding window haplotype association tests were performed by running Haplo.Stats[54] sequentially on adjacent sets of three SNPs. Plots of the global P-values from each window were plotted against the average position of the SNPs in the window. Additional haplotype and diplotype work was performed using the Pseudo-Gibbs sampling algorithm from the SNPAnalyzer program[55] to estimate phase, followed by a William's-corrected G test of homogeneity.

Similar to the haplotype method[56], a test of pairwise conditional independence (i.e. fixing the genotypes at one SNP and testing for the association at a second SNP) was performed through a permutation routine where case/control status is permuted against genotype data to generate a null distribution. For conditional independence hypotheses concerning only a small number of highly significant SNPs driving correlated SNPs to association solely through LD (such as is the case here), a permutation method has advantages over logistic regression models in that the P-values, given a sufficient number of iterations, will be appropriate regardless of LD levels, effect size and counts. Alternatively, we prefer logistic regression-based methods in situations that warrant inclusion of many SNPs and/or covariates with low to moderate LD/correlation levels and/or the hypothesis tested requires adjustment to be performed on more than one SNP. Typically, 2,000 iterations of the permutation were performed and P-values were calculated through a modeling procedure where a log-likelihood ratio test statistic is calculated for each of the permuted iterations. Next, the parameters of a gamma probability density are estimated from the permuted log-likelihood ratio test statistics and a P-value is calculated by integrating this null density from the observed log-likelihood ratio statistic. For a given number of permutation iterations, this modeling procedure gives more accurate P-values than simply taking the frequency of those permuted iterations that exceed the observed value.

Acknowledgements

We are grateful to all the individuals with psoriasis and control subjects for participation in this study. We thank J Sninsky and T White for their comments on the manuscript,

scientific advice and support. We also thank the members of the Celera High Throughput Lab and Computational Biology Group, S Prescott, M Paul, A Lindell, everyone at Lineagen, E Beasley, L McAllister and M Cargill for key help with this project. This study was supported in part by a Public Health Services research grant to the Huntsman General Clinical Research Center at the University of Utah, by National Center for Research Resources Grant M01-RR00064 and by generous gifts from the WM Keck Foundation and from the George S and Delores Dore Eccles Foundation. Finally, we note that as current employees of Celera, VE Garcia, M Chang, N Bui, JJ Catanese, Y Li, R Brandon, AB Begovich and SJ Schrodi may hold stock or stock options in Celera.

References

1. Chan JR, Blumenschein W, Murphy E, Diveu C, Wiekowski M, Abbondanzo S, *et al*. IL-23 stimulates epidermal hyperplasia via TNF and IL-20R2-dependent mechanisms with implications for psoriasis pathogenesis. *J Exp Med* 2006; 203:2577-2587.

2. Neimann AL, Porter SB, Gelfand JM. The epidemiology of psoriasis. *Expert Rev Dermatol* 2006; 1: 63-75.

3. Kremers HM, McEvoy MT, Dann FJ, Gabriel SE. Heart disease in psoriasis. *J Am Acad Dermatol* 2007; 57:347-354.

4. Griffiths CEM, Barker JN Pathogenesis and clinical features of psoriasis. *The Lancet* 2007; 370:263-271.

5. Farber EM, Nall ML, Watson W. *Arch Dermatol* 1974; 109:207-211.

6. Nair RP, Stuart PE, Nistor I, Hiremagalore R, Chia NV, Jenisch S *et al*. Sequence and haplotype analysis supports HLA-C as the psoriasis susceptibility 1 gene. *Am J Hum Genet* 2006; 78:827-851.

7. Helms C, Cao L, Krueger JG, Wijsman EM, Chamain F, Gordon D, *et al*. A putative RUNX1 binding site variant between SLC9A3R1 and NAT9 is associated with susceptibility to psoriasis. 2003; *Nat Genet* 35:349-356.

8. Capon F, Helms C, Veal CD, Tillman D, Burden AD, Barker JN, *et al*. Genetic analysis of PSORS2 markers in a UK dataset supports the association between RAPTOR SNPs and familial psoriasis. 2004; *J Med Genet* 41:459-460.

9. Stewart P, Nair RP, Abecasis GR, Nistor I, Hiremagalore R, Chia NV, *et al*. Analysis of RUNX1 binding site and RAPTOR polymorphisms in psoriasis: no evidence for association despite adequate power and evidence for linkage. 2006; *J Med Genet* 43:12-17.

10. Tsunemi Y, Saeki H, Nakamura K, Sekiya T, Hirai K, Fujita H, *et al*. Interleukin-12 p40 gene (IL12B) 3'-untranslated region polymorphism is associated with susceptibility to atopic dermatitis and psoriasis vulgaris. *J Dermatol Sci* 2002; 302:161-166.

11. Cargill M, Schrodi SJ, Chang M, Garcia VE, Brandon R, Callis KP, *et al*. A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. *Am J Hum Genet* 2007; 80:273-290.

12. Capon F, DiMeglio P, Szaub J, Prescott NJ, Dunster C, Baumber L, *et al*. Sequence variants in the genes for the interleukin-23 receptor (IL23R) and its ligand (IL12B) confer protection against psoriasis. *Hum Genet* 2007; 122:201-206.

13. Smith RL, Warren RB, Eyre S, Ho P, Ke X, Young HS, *et al*. Polymorphisms in the IL-12beta and IL-23R genes are associated with psoriasis of early onset in a UK cohort. *J Invest Dermatol* 2007; Nov 22 [Epub ahead of print]

14. Nair RP, Ruether A, Stuart PE, Jenisch S, Tejasvi T, Hiremagalore R, *et al*. Polymorphisms of the IL12B and IL23R genes are associated with psoriasis. *J Invest Dermatol* 2008; Jan 24 [Epub ahead of print].

15. Liu Y, Helms C, Liao W, Zaba LC, Duan S, Gardner J, *et al.* A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. *PLoS Genet.* 2008; 28:e1000041.

16. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, *et al*. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 2006; 314:1461-1463.

17. Libioulle C, Louis E, Hansoul S, Sandor C, Farnir F, Franchimont D, et al. Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. *PLoS Genet.* 2007; 3:e58.

18. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007; 447:661-678.

19. Parkes M, Barrett JC, Prescott NJ, Tremelling M, Anderson CA, Fisher SA, *et al.* Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet.* 2007; 39:830-832.

20. Raelson JV, Little RD, Ruether A, Fournier H, Paquin B, Van Eerdewegh, *et al.* Genome-wide association study for Crohn's disease in the Quebec founder population identifies multiple validated disease loci. *Proc Natl Acad Sci USA.* 2007; 104:14747-14752.

21. Tremelling M, Cummings F, Fisher SA, Mansfield J, Gwilliam R, Keniry A, et al. IL23R variation determines susceptibility but not disease phenotype in inflammatory bowel disease. *Gastroenterology* 2007; 132:1657-1664.

22. Van Limbergen J, Russell RK, Nimmo ER, Drummond HE, Smith L, Davies G, *et al*. IL23R Arg381Gln is associated with childhood onset inflammatory bowel disease in Scotland. *Gut* 2007; 56:1173-1174.

23. Dubinsky MC, Wang D, Picornell Y, Wrobel I, Katzir L, Quiros A, *et al*. IL-23 receptor (IL-23R) gene protects against pediatric Crohn's disease. *Inflamm Bowel Dis* 2007; 13:511-515.

24. Baldassano RN, Bradfield JP, Monos DS, Kim CE, Glessner JT, Casalunovo T, *et al*. Association of variants of the interleukin-23 receptor gene with susceptibility to pediatric Crohn's disease. *Clin Gastroenterol Hepatol* 2007; 5:972-976.

25. Wellcome Trust Case Control Consortium, Australo-Anglo-American Spondylitis Consortium (TASC), Burton PR, Clayton DG, Cardon LR, Craddock N, *et al*. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet.* 2007; 39:1329-1337.

26. Rueda B, Orozco G, Raya E, Fernandez-Sueiro JL, Mulero J, Blanco FJ, *et al*. The IL23R Arg381R non-synonymous polymorphism confers susceptibility to ankylosing spondylitis. Ann Rheum Dis. 2008; Jan 16 [Epub ahead of print].

27. Huber AK, Jacobson EM, Jazdzewski K, Concepcion ES, Tomer Y. IL-23R is a major susceptibility gene for Graves' ophthalmopathy: the IL-23/Th17 axis extends to thyroid autoimmunity. *J Clin Endocrinol Metab* 2007; 93:1077-1081.

28. Nunez C, Dema B, Cenit MC, Polanco I, Maluenda C, Arroyo R, et al. IL23R: a susceptibility locus for celiac disease and multiple sclerosis? *Genes Immun.* 2008; Mar 27 [Epub ahead of print].

29. Illes Z, Safrany E, Peterfalvi A, Magyari L, Farago B, Pozsonyi E, *et al*. 3'UTR C2370A allele of the IL-23 receptor gene is associated with relapsing-remitting multiple sclerosis. *Neurosci Lett.* 2008; 431:36-38.

30. Roos IM, Kockum I, Hillert J. The interleukin 23 receptor gene in multiple sclerosis: a case-control study. *J Neuroimmunol.* 2008; 194:173-180.

31. Begovich AB, Chang M, Caillier SJ, Lew D, Catanese JJ, Wang J, *et al.* The autoimmune disease-associated IL12B and IL23R polymorphisms in multiple sclerosis. *Hum Immunol.* 2007; 68:934-937.

32. Zhang X-J, Yan K-L, Wang Z-M, Yang S, Zhang G-L, Fan X, *et al*. Polymorphisms in interleukin-15 gene on chromosome 4q31.2 are associated with psoriasis vulgaris in Chinese population. *J Invest Derm* 2007; 127:2544-2551.

33. Chang M, Li Y, Yan C, Duffin KPC, Matsunami N, Garcia VE, et al. Variants in the 5q31 cytokine gene cluster are associated with psoriasis. *Genes Immun* 2007; 9:176-181.

34. Clark AG. The role of haplotypes in candidate gene studies. *Genet Epidemiol* 2004; 27:321-333.

35. International HapMap Consortium. A haplotype map of the human genome. *Nature* 2007; 437:1299-1320.

36. Airoldi I, Di Carlo E, Cocco C, Sorrentino C, Fais F, Cilli M, *et al*. Lack of il12rb2 signaling predisposes to spontaneous autoimmunity and malignancy. *Blood* 2005; 106:3846-3853.

37. Lee E, Trepicchio WL, Oestreicher JL, Pittman D, Wang F, Chamian F, et al. Increased expression of interleukin 23 p19 and p40 in lesional skin of patients with psoriasis vulgaris. *J Exp Med* 2004; 199:125-130.

38. Chan JR, Blumenschein W, Murphy E, Diveu C, Wiekowski M, Abbondanzo S, *et al*. IL-23 stimulates epidermal hyperplasia via TNF and IL-20R2-dependent mechanisms with implications for psoriasis pathogenesis. *J Exp Med* 2006; 203:2577-2587.

39. Ghilardi N, Kljavin N, Chen Q, Lucas S, Gurney AL, de Sauvage FJ. Compromised humoral and delayed-type hypersensitivity responses in IL-23-deficient mice. *J Immunol* 2004; 172:2827-2833.

40. Park H, Li Z, Yang XO, Chang SH, Nurieva R, Wang YH, *et al*. A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. *Nat Immunol* 2005; 6:1133-1141.

41. Harrington LE, Hatton RD, Mangan PR, Turner H, Murphy TL, Murphy KM, *et al*. Interleukin 17-producing CD4+ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. *Nat Immunol* 2005; 6:1123-1132.

42. Veldhoen M, Hocking RJ, Atkins CJ, Locksley RM, Stockinger B. TGFβ in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T-cells. *Immunity* 2006; 24:179-189.

43. Hue S, Ahern P. Buonocore S, Kullberg MC, Cua DJ, McKenzie BS, *et al*. Interleukin-23 drives innate and T cell-mediated intestinal inflammation. *J Exp Med* 2006; 203:2473-2483.

44. Kauffman CL, Aria N, Toichi E, McCormick TS, Cooper KD, Gottlieb AB, *et al*. A phase I study evaluating the safety, pharmacokinetics, and clinical response of a human IL-12 p40 antibody in subjects with plaque psoriasis. *J Invest Dermatol* 2004; 123:1037-1044.

45. Krueger GG, Langley RG, Leonardi C, Yeilding N, Guzzo C, Wang Y, *et al*. A human interleukin-12/23 monoclonal antibody for the treatment of psoriasis. *N Engl J Med* 2007; 356:580-592.

46. Mannon PJ, Fuss IJ, Mayer L, Elson CO, Sandborn WJ, Present D, et al. Anti-interleukin-12 antibody for active Crohn's disease. *N Engl J Med* 2004; 351:2069-2079.

47. Hu X, Schrodi SJ, Ross DA, Cargill M. Selecting tagging SNPs for association studies using power calculations from genotype data. *Hum Hered* 2004; 57:156-170.

48. Weir BS. *Genetic Data Analysis II: Methods for Discrete Population Genetic Data*. 1996; Sinauer, Sunderland, MA, USA.

49. Sokal RR, Rohlf FJ. *Biometry 3rd ed*. 1995; W.H. Freeman and Company, USA.

50. Fisher RA. *Statistical Methods for Research Workers, 12th ed.* 1954; Oliver & Boyd, Edinburgh.

51. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Nat Cancer Inst.* 1959; 22:719-748.

52. Excoffier L, Slatkin M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* 1995; 12:921-927.

53. Abecasis GR, Cookson WO. GOLD—graphical overview of linkage disequilibrium. *Bioinformatics* 2000; 16:182-183.

54. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet* 2002; 70:425-434.

55. Yoo J, Seo B, Kim Y. SNPAnalyzer: a web-based integrated workbench for single-nucleotide polymorphism analysis. Nucleic Acids Res 2005; 33(Web Server issue): W483-488.

56. Li H. A permutation procedure for the haplotype method for identification of disease-predisposing variants. *Ann Hum Genet* 2001; 65:189-196.

Figures

Figure 1a: Coverage of fine mapping SNPs on 1p31.3. With an average density of 1 SNP per 6kb, we genotyped a total of 61 SNPs across a 338kb region. SNPs were selected according to a fine mapping SNP selection algorithm where both traditional tagging SNPs were identified as well as SNPs in moderate to high LD with previously-associated P310L and R381Q polymorphisms. Gene boundaries are displayed. Positions and gene boundaries were determined by NCBI human genome assembly build 36.

Figure 1b: Individual IL23R SNPs genotyped. Across a 91kb region demarked by the *IL23R* gene boundaries, the 31 SNPs residing within *IL23R* are displayed in their respective positions.

Figure 2a: Pairwise linkage disequilibrium heatmap. Pairwise linkage disequilibrium between all 61 *IL23R*-region SNPs is presented in a heatmap using the D' and r^2 statistics. The upper triangle has the D' values and the lower triangle is composed of the r^2 values. Cases and controls from all three sample sets are combined. Both key missense SNPs, rs7530511 and rs11209026, are labeled.

Figure 2b: Decay of association with decreasing LD from R381Q. The combined P-value from the continuity-corrected Mantel-Haenzsel test for all individual SNPs is plotted as a function of linkage disequilibrium with R381Q, as measured by the r^2 statistic. Under the model where R381Q is the solely responsible for association patterns at *IL23R*, we expect an approximate logarithmic relationship as shown by the diagonal line. Those SNPs that may exhibit psoriasis effects independent of R381Q are circled.

Figure 3: Haplotype sliding window. A sliding window of haplotype association was performed. Estimated haplotypes using an EM algorithm and subsequent analyses were obtained from Haplo.Stats. A 3-SNP window was used. Global P-value for each window is plotted against the average position of the three SNPs in that window. Gene boundaries are displayed.

Table 1a: rs7530511- rs11465804-rs10889671-rs11209026-rs1857292 haplotypes

	Sample Set 1		Sample Set 2		Sample Set 3	
	No. (Frequency) in		No. (Frequency) in		No. (Frequency) in	
Haplotype[a,b]	Case	Control	Case	Control	Case	Control
CTGGA	754(0.818)	704(0.769)	791(0.802)	747(0.755)	795(0.828)	645(0.760)
TTAGT	69(0.075)	103(0.112)	68(0.069)	82(0.083)	74(0.077)	78(0.092)
CGGAA	36(0.039)	47(0.051)	45(0.046)	79(0.080)	33(0.034)	52(0.061)
CTGGT	31(0.034)	21(0.023)	31(0.031)	36(0.036)	22(0.023)	28(0.033)
TTAGA	23(0.025)	26(0.028)	20(0.020)	27(0.027)	28(0.029)	31(0.037)
Other	9(0.010)	15(0.016)	31(0.031)	19(0.019)	8(0.008)	14(0.017)

[a] Haplotype estimates were from the pseudo-Gibbs algorithm in the SNPAnalyzer program.

[b] These haplotypes consist of SNPs: rs7530511, rs11465804, rs10889671, rs11209026 and rs1857292 respectively.

	Sample Set 1			Sample Set 2			Sample Set 3			Combined
	No. (Frequency) in			No. (Frequency) in			No. (Frequency) in			
Haplotype	Case	Control	OR	Case	Control	OR	Case	Control	OR	P_{comb}[d]
Protective[c]	105(0.114)	150(0.164)		113(0.115)	161(0.163)		107(0.112)	130(0.153)		
All Other	817(0.886)	766(0.836)	0.66	873(0.885)	829(0.837)	0.67	853(0.889)	718(0.847)	0.69	4.32E-07

[c] TTAGT and CGGAA haplotypes combined.

[d] Continuity-corrected Mantel-Haenszel P-value.

Table 1b: rs7530511- rs10889671-rs11209026 haplotypes

	Sample Set 1		Sample Set 2		Sample Set 3		Combined
	No. (Frequency) in		No. (Frequency) in		No. (Frequency) in		
Haplotype[a,b]	Case	Control	Case	Control	Case	Control	P_{comb}[c]
CGG	783(0.852)	727(0.795)	830(0.844)	787(0.798)	818(0.855)	677(0.801)	3.88E-08
TAG	91(0.099)	128(0.140)	88(0.090)	108(0.110)	100(0.105)	107(0.127)	
CGA	39(0.042)	54(0.059)	51(0.052)	86(0.087)	33(0.035)	55(0.065)	
Other	6(0.007)	5(0.006)	14(0.014)	5(0.005)	6(0.006)	6(0.007)	

[a] Haplotype estimates were from the pseudo-Gibbs algorithm in the SNPAnalyzer program.

[b] These haplotypes consist of SNPs: rs7530511, rs10889671, and rs11209026, respectively.

[c] Continuity-corrected Mantel-Haenszel P-value for TAG + CGA

Table 2a: Twenty-three marker haplotypes

Table 2a

	Sample Set 1				Sample Set 2				Sample Set 3				Combined Analysis	
Haplotype[a,b]	Case	Control	OR	P[c]	Case	Control	OR	P[c]	Case	Control	OR	P[c]	OR[d]	P[e]
CTGGTGCTGGGTGCCCTACCAAA	248	217	1.18	0.133	266	252	1.07	0.506	280	220	1.18	0.127	1.139	0.034
CCGATAACGGTTAGCCTCCAACG	181	184	0.97	0.815	206	200	1.04	0.781	215	157	1.27	0.047	1.081	0.252
TTAATGACAGTTGCTCTCCCTAG	**65**	**100**	**0.62**	**0.0042**	**61**	**75**	**0.80**	**0.215**	**73**	**77**	**0.82**	**0.268**	**0.736**	**0.0022**
CCGATAATGGTTGCTATCTCAAG	85	76	1.12	0.510	60	72	0.82	0.281	65	69	0.82	0.281	0.921	0.440
CTAATGATGGGAGCCCTCCCAAG	79	62	1.28	0.162	76	70	1.09	0.667	75	52	1.30	0.168	1.216	0.064
CCGAGGCTGATTACCCTCCCAAG	**36**	**47**	**0.75**	**0.217**	**47**	**75**	**0.61**	**0.0089**	**31**	**51**	**0.52**	**0.0063**	**0.618**	**0.000113**
CTGGTGCTGGGTGCCCTACCAAG	35	38	0.91	0.176	37	36	1.03	1.000	36	34	0.93	0.808	0.955	0.792
CCGATAATGGTTGCTATCTAACG	23	32	0.70	0.220	25	26	0.96	0.888	37	26	1.27	0.372	0.961	0.860
CCGATAACGGTTAGCCTCCCAAG	21	23	0.90	0.762	34	30	1.14	0.703	20	17	1.04	1.000	1.035	0.906
Other	151	135	1.13	0.368	174	148	1.21	0.128	128	145	0.75	0.030	1.017	0.849

[a] Haplotypes built on: rs7530511, rs10489629, rs4655692, rs2201841, rs11465804, rs10489628, rs1343152, rs10789229, rs10889671, rs11209026, rs10889674, rs12085634, rs1343151, rs1008193, rs6693831, rs10889675, rs11465827, rs10889677, rs4655531, rs11209030, rs1857292, rs11209031, and rs11209032.

[b] Pseudo-Gibbs sampling algorithm in SNPAnalyzer used

[c] Fisher's Exact test

[d] Mantel-Haenszel common odds ratio

[e] Continuity-corrected Mantel-Haenszel P-value

Table 2b: Twelve marker reduced haplotypes

	Sample Set 1				Sample Set 2				Sample Set 3			
Haplotype[a,b]	Cases	Controls	OR	P-value[c]	Cases	Controls	OR	P-value[c]	Cases	Controls	OR	P-value[c]
GGGTCCTACCAA	256	217	1.374	0.055	270	255	1.125	0.476	280	225	1.238	0.227
AATTGCTCCACG	190	186	1.092	0.954	213	202	1.105	0.581	220	161	1.372	0.043
AGTTCCTCCCAG	**116**	**156**	**0.729**	**0.0070**	**124**	**174**	**0.668**	**0.0017**	**114**	**143**	**0.676**	**0.0029**
AATTCATCTCAG	90	75	1.291	0.255	71	77	0.932	0.609	76	72	0.963	0.668
AGGACCTCCCAG	85	66	1.394	0.127	78	74	1.082	0.800	77	48	1.530	0.077
GGGTCCTACCAG	36	39	0.961	0.724	38	38	1.019	1.000	38	35	0.994	0.905
AATTGCTCCCAG	34	42	0.836	0.350	56	57	1.000	0.923	32	38	0.760	0.223
AATTCATCTACG	24	34	0.727	0.183	28	26	1.100	0.890	37	27	1.270	0.524
Other	93	99	0.977	0.595	108	81	1.421	0.047	86	99	0.766	0.062

	Frequency			
Haplotype	Case	Control	OR[d]	P[e]
AGTTCCTCCCAG	0.123	0.172	**0.677**	**3.19E-07**
All Other	0.877	0.828		

[a] Haplotypes built on: rs2201841, rs10489628, 10889674, rs12085634, rs1008193, rs10889675, rs11465827, rs10889677, rs4655531, rs11209030, rs11209031, and rs11209032

[b] Haplotype estimates were from the pseudo-Gibbs algorithm in the SNPAnalyzer program.

[c] Fisher's Exact test

[d] Mantel-Haenszel Common OR

[e] Continuity-Corrected Mantel-Haenszel P-value