

## ARTICLE

## Cold sore susceptibility gene-1 genotypes affect the expression of herpes labialis in unrelated human subjects

John D Kriesel<sup>1</sup>, Amiteshwar Bhatia<sup>1</sup> and Alun Thomas<sup>2</sup>

Our group has recently described a gene on human chromosome 21, the Cold Sore Susceptibility Gene-1 (CSSG-1, also known as C21orf91), which may confer susceptibility to frequent cold sores in humans. We present here a genotype–phenotype analysis of CSSG-1 in a new, unrelated human population. Seven hundred fifty-eight human subjects were enrolled in a case/control Cold Sore Study. CSSG-1 genotyping, herpes simplex virus 1 (HSV1) serotyping, demographic and phenotypic data was available from 622 analyzed subjects. Six major alleles (H1–H6) were tested for associations with each of the self-reported phenotypes. The statistical analysis was adjusted for age, sex and ethnicity. Genotype–phenotype associations were analyzed from 388 HSV1-seropositive subjects. There were significant CSSG-1 haplotype effects on annual cold sore outbreaks ( $P=0.006$ ), lifetime cold sores ( $P=0.012$ ) and perceived cold sore severity ( $P=0.012$ ). There were relatively consistent trends toward protection from frequent and severe cold sores among those with the H3 or H5/6 haplotypes, whereas those with H1, H2, and H4 haplotypes tended to have more frequent and more severe episodes. Different alleles of the newly described gene CSSG-1 affect the expression of cold sore phenotypes in this new, unrelated human population, confirming the findings of the previous family-based study.

*Human Genome Variation* (2014) 1, 14024; doi:10.1038/hgv.2014.24; published online 20 November 2014

## INTRODUCTION

Herpes simplex labialis ('cold sores' or 'fever blisters') is a common and ubiquitous infection of the skin due to herpes simplex virus (HSV). The vast majority of cases are due to HSV type 1 (HSV1), although recurrent infections due to HSV type 2 have been reported. Roughly 20–40% of the population will experience labial or perioral outbreaks of vesicular herpetic lesions,<sup>1–4</sup> out of the ~60% of the population that are HSV1 seropositive.<sup>5</sup> The frequency of these outbreaks is variable, ranging, in some individuals, from rare episodes every 5–10 years, to monthly or more frequent outbreaks among a small proportion of subjects.<sup>6</sup> Herpes simplex labialis is usually mild, although uncomfortable and temporarily disfiguring for many persons. There can be a significant psychological impact of this prominent facial infection, particularly in young patients with frequent or severe recurrences.

Viral strain, environmental factors and genetics are believed to account for the observed differences in disease expression among those infected with HSV1. For instance, in humans fever, wind, sunburn and surgical manipulation of the ganglion are inducers of HSV reactivation.<sup>7–9</sup>

Our group previously identified a region of chromosome 21 that is significantly linked to Herpes simplex labialis disease. This region was defined using a genome-wide, family-based linkage study involving the Utah Centre d'Etude du Polymorphisme Humain families.<sup>10</sup> The region of nonrecombination believed to be driving the trait was refined by additional single-nucleotide polymorphisms (SNP) genotyping within this familial population. The association of individual SNP genotypes with Herpes simplex labialis frequency was evaluated using linkage analysis and transmission disequilibrium testing. These results indicated that human genotypes of C21orf91, which we call the 'Cold Sore

Susceptibility Gene-1' (CSSG-1), affects susceptibility to or protection from orolabial herpes outbreaks.<sup>11</sup>

The previous study implicating CSSG-1 in herpes disease expression was performed in a familial sample. It was of interest to determine whether the effects of CSSG-1 on herpes disease expression might also apply to the general population. Therefore, a case-control study in an entirely new, unrelated human population was performed to investigate the proposed effects of the gene CSSG-1 (c21orf91) on the susceptibility to cold sores.

## MATERIALS AND METHODS

## Instruments, standard techniques and procedures

A standardized, updated 'Cold Sore Questionnaire' was used as the instrument to collect phenotypic data from the participants (S11 Cold Sore Questionnaire). Six CSSG-1 SNP (Figure 1) were genotyped at the University of Utah Core Lab using standard TaqMan (Life Technologies, Carlsbad, CA, USA) genotyping methods. The PLINK whole-genome association analysis toolset computer application was used to group the SNPs into major haplotypes (groups of SNPs or alleles).<sup>12</sup> Blood was drawn and processed for recovery of genomic DNA. Serum was separated for type-specific HSV1 ELISA analysis (HerpeSelect, Focus Diagnostics, Cypress, CA, USA). The subjects were not asked about genital herpes lesions, nor were they serotyped for HSV2. Subjects were defined as HSV1 seropositive if the enzyme-linked immunosorbent assay (ELISA) index (EI) was >1.1 (manufacturer's recommendation) or EI >2.5 (more stringent definition, to avoid any false positives).

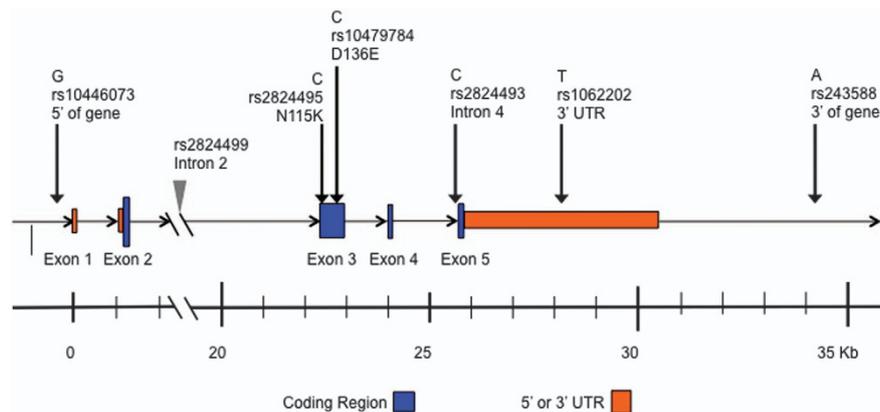
## Subjects and exclusions

Many of the subjects were recruited from the University of Utah Health Sciences Center. Other subjects were recruited largely from the University of Utah student body, staff and faculty. The subjects all provided written

<sup>1</sup>Department of Internal Medicine, Divisions of Infectious Diseases, University of Utah School of Medicine, Salt Lake City, UT, USA and <sup>2</sup>Department of Internal Medicine, Division of Genetic Epidemiology, University of Utah School of Medicine, Salt Lake City, UT, USA.

Correspondence: JD Kriesel (John.Kriesel@hsc.utah.edu)

Received 14 July 2014; revised 29 September 2014; accepted 29 September 2014



**Figure 1.** Cold sore susceptibility gene-1 (CSSG-1) map. The CSSG-1 gene spans 30,300 bp of chromosome 21. It is transcribed from the minus strand (for convenience shown left to right in this figure). Arrows indicate the locations of the six single-nucleotide polymorphisms (SNPs) within the gene. Each SNP is labeled with its wild-type nucleotide, rs (reference SNP) number and location within the gene. The two non-synonymous (amino-acid changing) SNPs are shown, D136E and N115K, both within exon 3. One SNP, rs2824499, used previously, is no longer considered here.<sup>11</sup>

informed consent, approved by the University of Utah Institutional Review Board, before participating. Prospective subjects with either a history of frequent cold sores ( $\geq 2$  per year, 'frequently affected') or no cold sores at all (0 lifetime cold sores, 'unaffected') were actively recruited into the study. Subjects with 0–2 cold sores per year ('unknown' or 'intermediate') were not recruited, but neither were they excluded from participation.

To ensure that the subjects were not closely related, relatives closer than 3rd cousins of the existing participants were excluded from the study. To avoid including persons with the same HSV1 strain in the analysis, subjects with identical addresses (e.g., marital partners) were identified and one of the pair was excluded from the analysis. Participants who were residents of a campus dormitory, all of whom had the same address, were all included in the analysis unless they shared surnames because it is unlikely that they were true cohabitants and, therefore, were unlikely to have similar HSV1 strains. Among pairs with the same address, HSV1 seropositives were preferentially retained over seronegatives and low seropositives. If both members of the pair had the same serostatus, the subject enrolled first was arbitrarily excluded. The exclusions were performed blinded to the subjects' phenotype and genotype. Additional criteria for exclusion from the analysis included insufficient sample, equivocal HSV1 serology or failure of the genotyping reaction. In these cases, the remaining partner was not excluded from the analysis. Thirty-three individuals with rare haplotypes (other than H1–H6) were excluded in order to avoid estimating haplotype effects from small numbers of observations.

### Phenotyping

The subjects were queried about their cold sore disease history using a printed questionnaire (S11 Cold Sore Questionnaire). They were asked to distinguish 'cold sores' on the lips or face from 'canker sores' inside the mouth. The subjects estimated their number of lifetime cold sores, their annual cold sore frequency and the perceived severity of their cold sores. The phenotypes were selected to capture as much of the complex nature of the disease as possible. Whether a person had frequent cold sores or none at all (i.e., 'affection status') was used in the previous study where the continuous phenotype 'annual cold sore frequency' also added significant useful information. Two other phenotypes, 'lifetime number of cold sores' and 'perceived cold sore severity', were proposed as methods to capture a broad range of cold sore phenotypes. For instance, persons who earlier in life experienced a large number of cold sores but who now have very few would be missed by the annual frequency phenotype, but recognized by the lifetime number phenotype. The phenotype definitions used for the study are included as Table S1.

### Statistical analysis

Least squares regression was used to evaluate the effects of CSSG-1 haplotypes on two binary responses and four ordinal responses in the presence of up to four covariates. The binary responses (discrete variables) were: HSV1 serostatus using the standard 1.1 EI cutoff and cold sore affection status ('frequent cold sores' or 'no cold sores'). The binary variables were analyzed using logistic regression. The ordinal responses

(continuous variables) were: EI, annual cold sore frequency, lifetime number of cold sores and perceived severity. The ordinal responses were analyzed using conventional linear regression.

For the 3 serostatus responses, we used all 622 subjects (HSV1 seropositives and seronegatives) and considered the subject's age, sex and ethnicity as covariates. For the remaining disease responses, we used only the 388 HSV1-seropositive individuals and considered age, sex and ethnicity as potential covariates. The covariates were included regardless of whether they were significant predictors of the response.

Haplotypes were used as the unit of analysis, with each haplotype being allocated the response and covariates of the individual carrying it. We note that this data departs from the assumptions implicit in regression analyses: the ordinal data generally is not Gaussian and contains some extreme values; both haplotypes in an individual's observed pair were allocated the same response and covariate values, and there may be correlations between covariates. Therefore, in addition to calculating the nominal *P* values for non-zero effects, we calculated bootstrap s.e. and *P* values by randomly resampling individual records. For each analysis we took 100,000 random bootstrap samples.

Effect sizes, conventional s.e. estimates, conventional *P* values, bootstrap s.e. estimates and bootstrap *P* values are reported to three decimal places. The *P* values are for two-sided tests in all cases. The sex effect is for males relative to the baseline for females, the ethnicity effect is for non-whites relative to the baseline for whites. Haplotype effects are relative to the baseline wild-type haplotype 1.

## RESULTS

### Demographics of the study population

Seven hundred fifty-eight human subjects were recruited into the Cold Sore Study. DNA and serum samples from 718 subjects were analyzed. After exclusions of cohabitants, subjects whose genotyping or serotyping failed and subjects with minor CSSG-1 haplotypes, a total of 622 subjects were examined for genotype–phenotype associations. Most of the analyzed population was Caucasian, as expected, given the demographics of the local Utah population. Seventy-eight of the total analyzed participants (13%) were African–American, Hispanic, Asian, Native American or other/mixed race. One hundred thirty-three participants (21%) did not report any race or ethnicity.

### Haplotype definition and frequency

Six major CSSG-1 haplotypes (i.e., alleles) were defined. The designations of the haplotypes and their frequencies within the study population are shown in Table 1. The locations of these SNPs within CSSG-1 are shown in Figure 1. A total of 1244 haplotypes were determined among the 622 analyzed subjects. (Each human subject has two copies of chromosome 21, and

**Table 1.** Determination of the CSSG-1 haplotypes and their frequencies

A.							
Haplotype	5' of gene rs10446073	Exon 3 (N115K) rs2824495	Exon 3 (D136E) rs10479784	Intron 4 rs2824493	3' UTR rs1062202	3''Tag rs243588	Haplotype frequency N (%)
H1	G	C (N)	C (D)	C	T	A	278 (22)
H2	G	C (N)	<b>G (E)</b>	<b>T</b>	T	<b>G</b>	427 (34)
H3	G	C (N)	C (D)	C	T	<b>G</b>	174 (14)
H4	G	<b>G (K)</b>	C (D)	C	<b>C</b>	A	265 (21)
H5	<b>T</b>	C (N)	C (D)	C	<b>C</b>	A	83 (7)
H6	G	C (N)	C (D)	C	<b>C</b>	A	17 (1)

B.				
Haplotype	Unaffected (N, %)	Affected (N, %)	Total (N, %)	Unknown (intermediate group)
H1	39 (25)	120 (75)	159 (21)	7
H2	77 (29)	185 (71)	262 (35)	12
H3	40 (38)	65 (62)	105 (14)	2
H4	45 (31)	101 (69)	146 (20)	10
H5	22 (39)	35 (61)	57 (8)	4
H6	5 (45)	6 (55)	11 (1)	1
Total	228 (31)	512 (69)	740 (100)	36

A. The CSSG-1 haplotypes were determined using the PLINK computer application. Haplotype assignments are equivalent to those derived from the familial population previously published.<sup>11</sup> Changes from the ancestral form of the gene (wild type) H1 are shown in bold. B. Haplotype frequencies among the 370 HSV1-seropositive participants, contributing 740 haplotypes, are shown. Thirty-six-seropositive individuals reported some cold sores but < 2 per year, making them of unknown or intermediate phenotype. Abbreviations: CSSG-1, cold sore susceptibility gene-1; UTR, untranslated region.

therefore two CSSG-1 genes with their respective haplotypes.) H2 was the most common haplotype (34%) followed by H1 (22%) and H4 (20%), H3 (14%), H5 (7%) and H6 (1%), as shown in Table 1. Haplotypes H5 and H6 were combined for the subsequent analysis because (a) the number of individuals with these haplotypes was low compared with the other four haplotypes, and (b) H5 and H6 are genetically similar, differing by only a single SNP in the 5' region of the gene.

#### HSV1 serotyping

Among the 622 analyzable participants, 234 (38%) were HSV1 seronegative (EI < 0.9) and 388 (62%) were HSV1 seropositive (EI > 1.1). The HSV-seropositive subjects were presumed to be HSV1 infected and, therefore, suitable for examining the effects of CSSG-1 haplotype on disease expression.

#### Serostatus (infection versus no infection with HSV1) and EI

Comparing the 388 seropositive and the 234 seronegative subjects, age had a highly significant effect on serostatus (Table 2). As expected, older subjects were more likely to be HSV1 seropositive: each year of age increased the log odds of being seropositive by 0.044 ( $P < 0.001$ ). Sex and ethnicity had no significant effects on serostatus. A higher EI value (EI > 2.5) was also used for this analysis owing to the difficulties with specificity of the assay.<sup>13,14</sup> The CSSG-1 haplotypes had no significant effect on HSV serostatus by either criteria (EI > 1.0 or > 2.5), except for the H5/H6 haplotype, which had a marginally significant effect on HSV1 serostatus where EI > 1.1 ( $P = 0.044$ ).

EI values ranged from 1.17 to 13.95 among the 388 HSV1-seropositive subjects. Age, but not sex or ethnicity, had a significant effect on the EI, with older individuals tending to have higher values. The CSSG-1 haplotypes did not significantly affect the EI values.

#### Affection status

Among the seropositives ( $N = 388$ ), logistic regression analysis revealed that sex and ethnicity, but not age, were significant predictors of affection status (Table 3). Women were more likely to be affected than men (74% vs. 60%,  $P = 0.01$ ), and whites were more likely to be affected than non-whites (71% vs. 52%,  $P = 0.02$ ). Each of the major CSSG-1 haplotypes was then examined individually for an effect on affection status, corrected for the covariates. Haplotypes 3 and 5/6 were marginally associated with the affection status ( $P = 0.050$ ,  $P = 0.056$ , respectively), with negative effect sizes demonstrating relative protection among these two haplotypes compared with the H1 wild type (Figure S13 Affection).

#### Annual cold sore frequency

The distribution of annual cold sore outbreaks among all the subjects, including seronegatives is shown in Figure 2. The 388 HSV1-seropositive subjects reported between 0 and 60 cold sores annually. Regression analysis revealed that sex ( $P = 0.001$ ) and ethnicity ( $P = 0.006$ ), but not age, were correlated with the number of annual outbreaks (Table 3). Again, whites and women had the most annual cold sore outbreaks. Haplotypes 3 ( $P = 0.006$ ) and 5/6 ( $P = 0.025$ ) were significantly associated with cold sore frequency, with negative effect sizes demonstrating relative protection among these two haplotypes compared with the H1 wild type (Figure S14 Annual Frequency).

#### Lifetime number of cold sores

The subjects reported 0 to 960 estimated lifetime cold sores. Again, 776 haplotypes from 388 HSV1-seropositive subjects were compared. Regression analysis revealed that age and sex, but not ethnicity, were significant covariates for this phenotype (Table 3). The median number of lifetime cold sores for each haplotype is shown in Figure S15 Lifetime. Haplotype 3 was significantly associated with lifetime cold sores ( $P = 0.012$ ), with a negative

**Table 2.** Summary of the regression analyses for the three cold sore serostatus phenotypes ( $N=622$  subjects)

Factor	Intercept	Sex	Ethnicity	Age	Hap 2	Hap 3	Hap 4	Hap 5/6
<i>Response variable = serostatus (EI &gt; 1.1)</i>								
Effect size	-1.102	0.226	0.093	0.044	0.142	0.003	-0.062	0.492
Conventional s.e.	0.222	0.128	0.183	0.006	0.164	0.205	0.181	0.264
Conventional <i>P</i> value	< <b>0.001</b>	0.078	0.611	< <b>0.001</b>	0.387	0.99	0.73	0.062
Bootstrap s.e.	0.304	0.183	0.271	0.008	0.164	0.213	0.182	0.256
Bootstrap <i>P</i> value	< <b>0.001</b>	0.207	0.723	< <b>0.001</b>	0.385	0.987	0.734	<b>0.044</b>
<i>Response variable = serostatus (EI &gt; 2.5)</i>								
Effect size	-1.253	0.225	0.216	0.044	0.065	-0.096	-0.16	0.317
Conventional s.e.	0.217	0.126	0.181	0.005	0.161	0.202	0.179	0.253
Conventional <i>P</i> value	< <b>0.001</b>	0.074	0.232	< <b>0.001</b>	0.689	0.634	0.371	0.21
Bootstrap s.e.	0.296	0.179	0.27	0.008	0.162	0.21	0.177	0.248
Bootstrap <i>P</i> value	< <b>0.001</b>	0.2	0.403	< <b>0.001</b>	0.69	0.644	0.363	0.191
<i>Response variable = EI</i>								
Effect size	1.933	0.387	0.11	0.083	0.196	0.122	-0.357	0.672
Conventional s.e.	0.432	0.256	0.37	0.01	0.33	0.415	0.368	0.5
Conventional <i>P</i> value	< <b>0.001</b>	0.131	0.766	< <b>0.001</b>	0.552	0.769	0.331	0.179
Bootstrap s.e.	0.567	0.359	0.512	0.013	0.335	0.44	0.363	0.475
Bootstrap <i>P</i> value	< <b>0.001</b>	0.281	0.834	< <b>0.001</b>	0.559	0.784	0.323	0.161

Subjects were defined as HSV1 seropositive if the EI was > 1.1 (manufacturer's recommendation) or EI > 2.5 (more stringent definition). The regression analysis was also performed on the continuous variable of EI. Significant values are in bold. Abbreviations: EI, enzyme-linked immunosorbent assay index; Hap, haplotype.

**Table 3.** Summary of the regression analyses for four herpes disease phenotypes ( $N=388$  HSV1-seropositive subjects)

Factor	Intercept	Sex	Ethnicity	Age	Hap 2	Hap 3	Hap 4	Hap 5/6
<i>Response variable = affection status</i>								
Effect size	0.955	-0.614	-0.786	0.013	-0.203	-0.531	-0.194	-0.616
Conventional s.e.	0.312	0.167	0.232	0.007	0.234	0.280	0.264	0.318
Conventional <i>P</i> value	<b>0.002</b>	< <b>0.001</b>	<b>0.001</b>	0.059	0.385	0.058	0.463	0.053
Bootstrap s.e.	0.413	0.242	0.340	0.010	0.228	0.271	0.243	0.319
Bootstrap <i>P</i> value	<b>0.019</b>	<b>0.010</b>	<b>0.021</b>	0.164	0.362	0.050	0.417	0.056
<i>Response variable = annual outbreaks</i>								
Effect size	5.277	-1.598	-1.827	-0.026	0.380	-1.110	0.758	-1.014
Conventional s.e.	0.708	0.401	0.591	0.015	0.525	0.664	0.598	0.753
Conventional <i>P</i> value	< <b>0.001</b>	< <b>0.001</b>	<b>0.002</b>	0.086	0.470	0.095	0.205	0.178
Bootstrap s.e.	0.979	0.490	0.626	0.022	0.501	0.408	0.692	0.449
Bootstrap <i>P</i> value	< <b>0.001</b>	<b>0.001</b>	<b>0.006</b>	0.200	0.451	<b>0.006</b>	0.263	<b>0.025</b>
<i>Response variable = lifetime outbreaks</i>								
Effect size	31.847	-37.739	-15.763	1.890	5.211	-28.389	0.768	-25.107
Conventional s.e.	15.299	8.663	12.780	0.330	11.343	14.354	12.913	16.267
Conventional <i>P</i> value	<b>0.038</b>	< <b>0.001</b>	0.218	< <b>0.001</b>	0.646	<b>0.048</b>	0.953	0.123
Bootstrap s.e.	19.811	10.654	23.411	0.496	11.410	11.353	13.137	12.765
Bootstrap <i>P</i> value	0.105	<b>0.001</b>	0.471	< <b>0.001</b>	0.654	<b>0.012</b>	0.957	0.050
<i>Response variable = perceived severity</i>								
Effect size	1.738	-0.465	-0.396	0.001	-0.208	-0.344	-0.110	-0.238
Conventional s.e.	0.143	0.081	0.119	0.003	0.106	0.134	0.121	0.152
Conventional <i>P</i> value	< <b>0.001</b>	< <b>0.001</b>	<b>0.001</b>	0.797	0.050	<b>0.010</b>	0.361	0.118
Bootstrap s.e.	0.191	0.113	0.172	0.004	0.103	0.130	0.112	0.163
Bootstrap <i>P</i> value	< <b>0.001</b>	< <b>0.001</b>	<b>0.025</b>	0.847	<b>0.047</b>	<b>0.008</b>	0.330	0.150

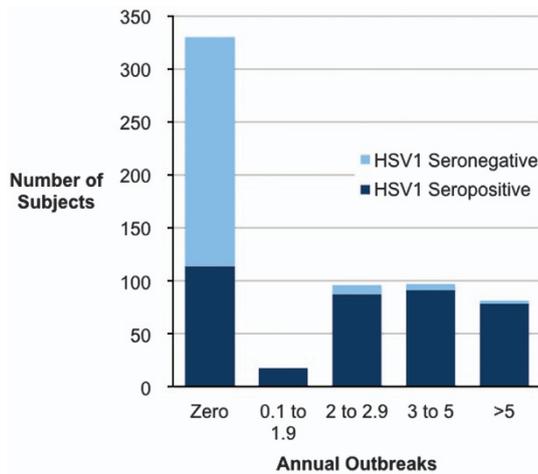
The regression analysis was also performed on the discrete variable 'affection status' and the continuous variables 'annual outbreaks', 'lifetime outbreaks', and 'perceived severity'. Significant values are in bold. Abbreviation: Hap, haplotype.

effect size demonstrating relative protection among this haplotype compared with the H1 wild type.

#### Perceived severity

The 388 HSV1-seropositive subjects reported cold sore severities between 0 (none) and 4 (severe). The median perceived cold

sore severity for each haplotype is shown in Figure S16 Severity. Regression analysis revealed that sex and ethnicity, but not age, were significant covariates. CSSG-1 haplotype H3 was significantly associated with perceived severity ( $P=0.008$ ), whereas the H2 haplotype association was marginally significant ( $P=0.047$ ).



**Figure 2.** Self-reported annual cold sore outbreaks among the 622 analyzed subjects. The bimodal distribution of HSV1-seropositive (EI > 1.1) subjects was due to recruitment and enrollment of only clearly 'affected' and 'unaffected' individuals.

## DISCUSSION

This study evaluated the effects of CSSG-1, a putative herpes susceptibility gene, on the expression of cold sores in a larger, unrelated human population. One or more CSSG-1 haplotypes significantly affected the reported number of annual cold sore outbreaks, the number of reported lifetime cold sores and perceived cold sore severity. CSSG-1 haplotype H3 had the most consistent effects on the measured cold sore phenotypes in this population. The CSSG-1 haplotypes generally did not affect HSV1 serostatus (i.e., infection with HSV1) or the EI (quantity and affinity of anti-HSV1 antibodies).

The age, sex and ethnicity-adjusted data here suggested significant protection from frequent and severe cold sores among those with the H3 or H5/6 haplotypes, whereas those with H1, H2 and H4 haplotypes tended to be subject to more frequent and more severe episodes. It is interesting that the H2 and H4 haplotypes associated with more frequent cold sores contain the amino-acid-changing SNPs, D136E and N115K, respectively. The D136E substitution is a conservative amino-acid change owing to the similarity between aspartic acid (D) and glutamic acid (E). The N115K substitution is predicted to have more effect on protein function than the D136E substitution due to the difference between asparagine (N, uncharged) and lysine (K, positively charged). In the previous familial study the H2 haplotype seemed more protective than the N115K containing H4 haplotype, a finding not confirmed in the present study. However, the H2 and H4 may provide for a lower lifetime number of cold sores compared with the wild-type H1 haplotype.

Although age had a highly significant effect on HSV1 serostatus and EI, the CSSG-1 haplotypes generally did not. As persons with more HSV1 outbreaks tend to have higher antibody titers,<sup>9</sup> it is possible that the effects of CSSG-1 haplotype were overshadowed by reactivation frequency. For instance, the protective H3 haplotype was shown to be associated with decreased reactivation frequency. It is possible that persons with the H3 haplotype also make more and/or higher affinity antibody against HSV1, but that this effect was not seen because these persons had fewer HSV1 outbreaks, which reduced their EI values. Therefore, the present study cannot exclude subtle effects of CSSG-1 genotypes on anti-HSV1 antibody production and affinity.

The relatively protective H3 and H5/H6 haplotypes carry SNPs in the 3' UTR (3' untranslated region or 3' Tag SNP) of the gene. As these SNPs do not occur in protein coding regions, we cannot say

whether these haplotypes exert their effect through alterations to the regulation or expression of CSSG-1 or through other genetic effects to nearby regions of the genome. Our data do not allow us to predict if these effects are associated with increased or decreased CSSG-1 protein production.

These results are different than the previous family-based study where some data suggested that the H4 and H5 haplotypes might be susceptible, whereas the H2 and H3 haplotypes could be protective.<sup>11</sup> However, in the previous study the H3 group had the lowest mean annual cold sore frequency, in agreement with the present study. The present study was designed to characterize the nature of the associations between the haplotypes and the measured phenotypes. The current study includes an analysis of the predictive values of genotype on phenotype, an analysis not performed in the previous study, which was focused on linkage.

This study showed only a marginally significant effect of CSSG-1 haplotypes H3 and H5/6 on the discrete variable of affection status ('frequently affected' versus 'unaffected'). The other phenotypes including annual cold sore frequency, lifetime cold sores and perceived severity are all continuous variables, which provide a more sensitive means to detect a genetic effect. The haplotype effects on these continuous variables were significant with H3 effects, for instance, all at significance levels of  $\leq 0.01$ .

Other genetic determinants of herpetic disease include the human genes UNC93B and Toll-like receptor 2.<sup>15,16</sup> Mutations in UNC93B, Toll-like receptor 3, STAT1 and NEMO genes were associated with an increased risk of HSV1 encephalitis in children.<sup>17</sup> Changes in Toll-like receptor 2 increased the number of genital outbreaks and the rate of genital shedding among 128 subjects with HSV2 infection.<sup>16</sup> Although interesting and potentially relevant, these studies involved investigating immunologic abnormalities in severe-HSV phenotypes such as encephalitis. A recent study in mice identified the calcitonin receptor as a possible susceptibility factor for herpes encephalitis.<sup>18</sup> Another interesting new genetic study mapped the severity of HSV1-induced disease to a region on mouse chromosome 16.<sup>19</sup> This important mouse herpes susceptibility region maps near the murine analog of CSSG-1 (c21orf91), suggesting that a single host locus may be the main genetic influence on HSV1 pathogenesis.

In conclusion, the present study confirms the activity of a previously unknown herpes susceptibility gene, CSSG-1, in a new population of unrelated humans. These new data are important because previous data in a large multigeneration linkage study suggests that CSSG-1 is the most important determinant of cold sore expression. The mechanism of CSSG-1 action is unknown.

## COMPETING INTERESTS

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

The authors wish to acknowledge the genotyping done at the University of Utah Genomics Core Facility. This investigation was supported by a grant from the Cold Sore Research Foundation, Salt Lake City, UT, USA; and by the United States Public Health Services research grants, R01-AI64349 and PO1-AI083215 (to PI Robert W Finberg).

## REFERENCES

- Ship II, Miller MF, Ram C. A retrospective study of recurrent herpes labialis (RHL) in a professional population, 1958-1971. *Oral Surg Oral Med Oral Pathol* 1977; **44**: 723-730.
- Embil JA, Stephens RG, Manuel FR. Prevalence of recurrent herpes labialis and aphthous ulcers among young adults on six continents. *Can Med Assoc J* 1975; **113**: 627-630.
- Grout P, Barber VE. Cold sores—an epidemiological survey. *J R Coll Gen Pract* 1976; **26**: 428-434.

- 4 Ship II, Morris AL, Durocher RT, Burket LW. Recurrent aphthous ulcerations and recurrent herpes labialis in a professional school student population I. Experience. *Oral Surg Oral Med Oral Pathol* 1960; **13**: 1191–1202.
- 5 Xu F, Sternberg MR, Kottiri BJ, McQuillan GM, Lee FK, Nahmias AJ *et al*. Trends in herpes simplex virus type 1 and type 2 seroprevalence in the United States. *JAMA* 2006; **296**: 964–973.
- 6 Spruance S, Kriesel J. Treatment of herpes simplex labialis. *Herpes* 2002; **9**: 64–69.
- 7 Pazin GJ, Ho M, Jannetta PJ. Reactivation of HSV after decompression of the trigeminal nerve root. *J Infect Dis* 1978; **138**: 405–409.
- 8 Shimizu F, Monma Y, Sekizawa T, Kamiyama K. Establishment of latent ganglionic infection with herpes simplex virus via maxillary gingiva and viral reactivation *in vivo* after trauma. *J Dent Res* 1989; **68**: 472–475.
- 9 Spruance SL, Kriesel JD, Evans TG, McKeough MB. Susceptibility to herpes labialis following multiple experimental exposures to ultraviolet radiation. *Antiviral Res* 1995; **28**: 57–67.
- 10 Hobbs MR, Jones BB, Otterud BE, Leppert M, Kriesel JD. Identification of a herpes simplex labialis susceptibility region on human chromosome 21. *J Infect Dis* 2008; **197**: 340–346.
- 11 Kriesel JD, Jones BB, Matsunami N, Patel MK, St Pierre CA, Kurt-Jones EA *et al*. C21orf91 genotypes correlate with herpes simplex labialis (cold sore) frequency: description of a cold sore susceptibility gene. *J Infect Dis* 2011; **204**: 1654–1662.
- 12 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
- 13 Nascimento MC, Ferreira S, Sabino E, Hamilton I, Parry J, Pannuti CS *et al*. Performance of the HerpeSelect (Focus) and Kalon enzyme-linked immunosorbent assays for detection of antibodies against herpes simplex virus type 2 by use of monoclonal antibody-blocking enzyme immunoassay and clinicovirological reference standards in Brazil. *J Clin Microbiol* 2007; **45**: 2309–2311.
- 14 Mark HD, Nanda JP, Roberts J, Rompalo A, Melendez JH, Zenilman J. Performance of focus ELISA tests for HSV-1 and HSV-2 antibodies among university students with no history of genital herpes. *Sex Trans Dis* 2007; **34**: 681–685.
- 15 Casrouge A, Zhang SY, Eidschchenk C, Jouanguy E, Puel A, Yang K *et al*. Herpes simplex virus encephalitis in human UNC-93B deficiency. *Science* 2006; **314**: 308–312.
- 16 Bochud PY, Magaret AS, Koelle DM, Aderem A, Wald A. Polymorphisms in TLR2 are associated with increased viral shedding and lesion rate in patients with genital herpes simplex virus Type 2 infection. *J Infect Dis* 2007; **196**: 505–509.
- 17 Sancho-Shimizu V, Zhang SY, Abel L, Tardieu M, Rozenberg F, Jouanguy E *et al*. Genetic susceptibility to HSV-1 encephalitis in mice and humans. *Curr Opin Allergy Clin Immunol* 2007; **7**: 495–505.
- 18 Abdelmagid N, Berezky-Veres B, Guerreiro-Cacais AO, Bergman P, Luhr KM, Bergstrom T *et al*. The calcitonin receptor gene is a candidate for regulation of susceptibility to herpes simplex type 1 neuronal infection leading to encephalitis in rat. *PLoS Pathog* 2012; **8**: e1002753.
- 19 Thompson RL, Williams RW, Kotb M, Sawtell NM. A forward phenotypically driven unbiased genetic analysis of host genes that moderate herpes simplex virus virulence and stromal keratitis in mice. *PLoS ONE* 2014; **9**: e92342.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>

Supplemental Information for this article can be found on the *Human Genome Variation* website (<http://www.nature.com/hgv>)