Published in final edited form as:

Genes Immun. 2019 February ; 20(2): 112–120. doi:10.1038/s41435-018-0013-4.

Genome-wide association study (GWAS) of human host factors influencing viral severity of herpes simplex virus type 2 (HSV-2)

Sarah E. Kleinstein[1,2], **Patrick R. Shea**[1], **Andrew S. Allen**[3], **David M. Koelle**[4,5,6,7,8], **Anna Wald**[4,5,9], and **David B. Goldstein**[1]

[1]Institute for Genomic Medicine, Columbia University, New York, NY 10032

[2]Department of Molecular Genetics and Microbiology, Duke University School of Medicine, Durham, NC 27708

[3]Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC 27708

[4]Department of Medicine, University of Washington, Seattle, WA 98195

[5]Vaccine and Infectious Diseases Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109

[6]Benaroya Research Institute, Seattle, WA 98101

[7]Department of Laboratory Medicine, University of Washington, Seattle, WA 98195

[8]Department of Global Health, University of Washington, Seattle, WA 98195

[9]Department of Epidemiology, University of Washington, Seattle, WA 98195

Abstract

Herpes simplex virus type 2 (HSV-2) is an incurable viral infection with severity ranging from asymptomatic to frequent recurrences. The viral shedding rate has been shown as a reproducible HSV-2 severity endpoint that correlates with lesion rates. We used a genome-wide association study (GWAS) to investigate the role of common human genetic variation in HSV-2 severity. We performed a GWAS on 223 HSV-2-positive participants of European ancestry. Severity was measured by viral shedding rate, as defined by the percent of days PCR+ for HSV-2 DNA over at least 30 days. Analyses were performed under linear regression models, adjusted for age, sex, and ancestry. There were no genome-wide significant (p<5E-08) associations with HSV-2 viral shedding rate. The top non-significant SNP (rs75932292, p=6.77E-08) associated with HSV-2 viral shedding was intergenic, with the nearest known biologically interesting gene (*ABCA1*) ~130Kbp downstream. Several other SNPs approaching significance were in or near genes with viral or neurological associations, including 4 SNPs in *KIF1B*. The current study is the first

Corresponding Author: David B. Goldstein, Ph.D., Institute for Genomic Medicine, Columbia University, 650 W. 168th Street, Room 1408, New York, NY, 10032, dg2875@cumc.columbia.edu, Phone: 212-305-0923, Fax: 212-305-3691.

Conflicts of Interest: SEK, PRS, and ASA have no conflicts.

Supplementary information is available at *Genes & Immunity's* website.

comprehensive genome-wide investigation of human genetic variation in virologic severity of established HSV-2 infection. However, no significant associations were observed with HSV-2 virologic severity, leaving the exact role of human variation in HSV-2 severity unclear.

Keywords

HSV-2; genital herpes; GWAS; human genetics

Introduction

Herpes simplex virus type 2 (HSV-2) is one of the most prevalent sexually transmitted infections worldwide, with a global prevalence estimated at 417 million and ~19.2 million new infections acquired per year (1). In the United States, the number of HSV-2 infected individuals has stabilized at ~16% of the population, indicating that transmission is a continuing public health problem (2). HSV-2 establishes lifelong latency upon infection and, to date, there are no vaccines, cures or even fully efficacious suppressive treatments (3).

Although 75-90% of HSV-2 infections are subclinical and asymptomatic (4), the severity varies widely. Symptomatic, clinical disease presents as painful, recurrent and often frequent outbreaks of genital lesions (5–7). Daily treatment with antiviral medications can reduce outbreak frequency, though it does not completely eliminate them (8). Due to its incurable status and adverse effects on quality of life, HSV-2 diagnosis may be associated with psychological distress (9,10). Further, HSV-2 infection during pregnancy, particularly primary and asymptomatic infections, can result in perinatal transmission to the infant, a rare but severe outcome (11). Children who acquire HSV-2 at birth experience significant morbidity and mortality, with survivors risking encephalitis or multi-organ disseminated disease, and over 50% developing central nervous system disease (11). Of additional public health concern, HSV-2 infection is associated with at least a three-fold increased risk of both acquiring and transmitting human immunodeficiency virus type 1 (HIV-1) (12,13).

There are currently no completely effective methods for interrupting HSV-2 transmission. While standard safe sex practices are recommended (5), condom usage risk reduction estimates have varied by gender and measurement, ranging from 30-96% (14,15), as active lesions and viral replication may occur on unprotected skin. Similarly, while antiviral treatment reduces HSV-2 transmission by ~50%, it does not completely ablate active viral replication as measured by viral shedding (8,11). Therefore, it is important to understand the host factors influencing HSV-2 severity in order to elucidate mechanisms to limit its impact on human health.

HSV-2 severity is a complex phenotype that can be measured by symptomatic lesion recurrences or viral shedding. Viral shedding from genital mucosa has previously been identified as a more objective representation of infection severity than the number of symptomatic recurrences (16,17). Days with active lesions show higher risk of viral shedding; thus, those with the most severe disease show both increased viral shedding and increased outbreaks (16). Further, though some asymptomatic individuals eventually recognize lesions, particularly following education on lesion identification, those who

remain asymptomatic have the lowest viral shedding (16). Thus, having quantitative information on viral shedding is valuable in determining virologic HSV-2 severity and limiting confounding from perception biases in lesion detection.

As a complex trait, infection severity is potentially influenced by a combination of viral (18), host (19–21), and environmental factors. The importance of host genetic variation in herpes pathogenesis has previously been demonstrated for certain rare phenotypes, such as herpes simplex virus type 1 encephalitis (HSE) (22,23). It is well-established that neuron-intrinsic deficiencies in toll-like receptor 3 (TLR3) pathway genes result in severe childhood HSE after primary HSV-1 infection (22,24). For HSV-2, candidate gene studies have implicated viral control and immune genes, including TLRs, in both susceptibility (25–28) and severity (19–21) during the chronic phase. Although these studies have detected variants potentially associated with herpes pathogenesis, they were limited to a small number of common allelic variants in candidate genes and none of these candidate gene associations have been replicated by other, independent studies. To date, only two genome-wide studies have investigated the role of human genetic factors in alphaherpes-related diseases: a genome-wide association study (GWAS) investigating herpes zoster susceptibility, which identified an association with *HCP5* in the HLA region and age of shingles onset (29), and a family-based linkage analysis that identified a 2.5MB region on chromosome 21 associated with HSV-1 susceptibility (30). Subsequent targeted sequencing pinpointed specific variants in the *C21orf91 (CSSG1)* gene that were responsible for the chromosome 21 association with HSV-1 susceptibility (31). Thus, while human genetics have been implicated in herpetic diseases, there remains a dearth of validated causal variants for HSV-2 susceptibility and severity.

It is important to utilize unbiased, genome-wide studies to definitively investigate the role of human genetics in disease etiology. The use of GWAS to investigate complex traits is a well-validated standard in human genetics. Though most complex disease traits previously studied relate to inherited diseases, host genetic factors influencing infectious diseases have been detected through genome-wide studies (32–35), including for other alphaherpesviruses (29,30). To date, no GWAS has been reported for HSV-2. In this study, we report the first genome-wide investigation of common human genetic variation influencing HSV-2 severity, as measured by the quantitative viral shedding rate.

Results

Demographic and clinical characteristics of participants included in the final analyses are described in Table 1. Briefly, most participants were symptomatic (83%), with more female participants than male (61% vs. 39%, respectively). Forty-four percent of participants were HSV-1 seropositive. The viral shedding rate ranged from 0-100%, with a median of 15% (see Figure 1). The full cohort had similar demographics to the genetically confirmed European subset used in this study (data not shown).

After multiple testing correction, there were no genome-wide significant associations ($p<5E-08$) with HSV-2 severity, as measured by the quantitative viral shedding rate and adjusted for age, gender, and ancestry (Figures 2-3). The 10 single-nucleotide

polymorphisms (SNPs) with the lowest p-values are listed in Table 2. The SNP that achieved the lowest p-value in our analysis was rs75932292, which was just below statistical significance (p=6.77E-08). rs75932292 is intergenic, with the nearest biologically relevant coding gene, ATP binding cassette subfamily A member 1 (*ABCA1*), located ~130Kbp downstream. To examine whether rs75932292 might tag functional variation in the *ABCA1* gene, we examined linkage disequilibrium (LD) within ±1MB of this SNP among 640 whole-genome sequenced (WGS) population controls of European ancestry. We identified several rare, functional variants in *ABCA1* that were in high LD with rs75932292 and could potentially account for any association with HSV-2 viral shedding, though their causality cannot be definitively determined (see Supplementary Table 1). As rs75932292 itself is relatively rare, with a minor allele frequency (MAF) of 0.07 in our cohort (MAF=0.05 for Europeans in the 1000 Genomes Project (36)), there were no individuals homozygous for the variant allele in our dataset. However, there was a slight trend toward increased viral shedding among carriers heterozygous for the variant genotype (GA) compared to homozygous wild-type (GG) individuals (data not shown).

Despite the lack of genome-wide significant associations, several potentially biologically interesting SNPs approached statistical significance. Four intronic SNPs in the kinesin family member 1B (*KIF1B*) gene were observed among the top results. These four SNPs (rs17034615, rs17034775, rs72865926, and rs72867415; all p=6.57E-06) were in perfect LD (r^2=1) among individuals of European ancestry in our cohort, with a MAF of 0.05 (identical to that expected for Europeans in the 1000 Genomes Project (36)), suggesting that they might all be tagging the same underlying variant. When LD was explored among the 640 WGS population controls of European ancestry, several rare, functional *KIF1B* variants were in high LD with the intronic variants (see Supplementary Table 2). There was also a slight trend toward increased viral shedding with the presence of any intronic SNP minor allele(s); however, this was primarily driven by heterozygous individuals, as there was only a single homozygous variant individual for three of the SNPs and none for rs72867415 (data not shown).

Three other SNPs also approached genome-wide significance and were in or near genes of plausible biological relevance to neurological and/or viral phenotypes, though none had previously been linked to HSV-2. These included two intergenic SNPs, rs56122323 (p=3.76E-07), located ~25KB downstream of mannosyl (alpha-1,6-)-glycoprotein beta-1,6-N-acetyl-glucosaminyltransferase, isozyme B (*MGAT5B*), and rs62377770 (p=4.29E-07), located ~58.5KB upstream of casein kinase 1 gamma 3 (*CSNK1G3*); as well as rs117944720 (p=5.43E-06), which is intronic in parkin RBR E3 ubiquitin protein ligase (*PARK2*).

Due to prior associations of HSV pathogenesis with immune-related genes, including a recent association with HLA-A*01 in a portion of this cohort (21), a targeted analysis limited to just the major histocompatibility complex (MHC) region of the genome was conducted. However, there were no suggestive associations with HSV-2 viral shedding when just the MHC region was considered (see Supplementary Figures 1-2). Further, while we did not directly test HLA haplotypes in this analysis, we genotyped 5 SNPs in nearly perfect LD (r^2~0.95) with HLA-A*0101. While these SNPs failed to reach genome-wide significance

(p=0.001) when tested using linear or Poisson regression models, we did observe ~10% higher frequency of these SNPs among the highest viral shedders (≥25% of days with viral shedding) compared to low/no shedders (<25% of days with viral shedding; Supplementary Table 3), suggesting that the HLA-A*01 haplotype may have a moderate influence on HSV2 shedding levels, but with an effect that requires a larger sample size to detect.

Additionally, a candidate gene analysis of ten non-MHC genes previously implicated in HSV-2 pathogenesis was also conducted. Though not statistically significant after Bonferroni correction, p-values were slightly lower than expected on the quantile-quantile (QQ) plot (no genomic inflation (lambda=1); see Supplementary Figure 3). The vast majority of the SNPs with the lowest p-values were on chromosome 10 in the mannose binding lectin 2 (*MBL2*) gene region. MBL2 is part of the innate immune system, where it activates the classical complement pathway and can detect viruses, including binding HSV-2 surface glycoproteins (27,37). The three SNPs with the lowest p-values (rs201381710, p=0.01; rs10824793, p=0.02; and rs4935047, p=0.02) shared nearly perfect LD (r^2>0.99) and were all intronic in *MBL2*. The SNP with the next lowest p-value in the non-MHC region candidate gene analysis was rs4696483 (p=0.02), which is intronic in *TLR2*. This SNP is not in LD with the previous *TLR2* candidate SNP rs1898830 in our cohort (r^2=0.03). The only other SNP with a p<0.05 was rs2147419 (p=0.04), which is intronic in *FAS*.

Discussion

This study represents the first GWAS of any measure of HSV-2 severity. Overall, the results failed to achieve statistical significance and there was no evidence for associations of common host genetic variation and HSV-2 viral shedding rate (as quantitatively measured by percent of days PCR+ for HSV-2 DNA at self-swabbed sites over a period of at least 30 days). Additionally, we were not able to replicate previously observed candidate gene associations with HSV-2 pathogenesis disclosed in targeted investigations, though we did see a slight, non-significant increase in frequency of HLA-A*0101 tagSNPs among high viral shedders relative to low viral shedders.

The top SNP (rs75932292, p=6.77E-08) was intergenic and just below the p<5E-08 threshold for genome-wide significance; it showed some evidence of linkage with potentially causal rare, functional variants in the downstream protein coding gene *ABCA1*. ABCA1 is a cholesterol/lipid regulator that is associated with several lipoprotein disorders (38–41). It has also been shown to interact with HIV-1 viral proteins (42–45), Newcastle disease virus (46), and hepatitis C (47). Rare variants in *ABCA1* might conceivably affect HSV-2 viral reactivation by altering lipids involved in membrane fusion or viral egress, hypotheses that are amenable to testing *in vitro*, as has been done with the viruses discussed above. There were several additional non-significant SNPs in the top results that are potentially of note, as they were present in genes previously associated with viral infections, including 4 intronic SNPs in *KIF1B*. KIF1B is a kinesin motor protein involved in anterograde transport of mitochondria and synaptic vesicle precursors. While other *KIF1B* mutations have been linked to Charcot-Marie-Tooth disease, neuroblastoma, and pheochromocytoma, *KIF1B* has also been identified in several studies related to hepatitis B virus-related hepatocellular carcinoma (48) and may act in early HIV-1 viral trafficking (49).

The role of *KIF1B* in anterograde synaptic transport could conceivably be related to HSV-2 reactivation because during viral reactivation from neurons HSV virion components are actively transported by this mechanism (50). While several rare, functional protein coding variants in *KIF1B* were in high LD with the identified intronic SNPs, it is unclear if these variants could affect viral trafficking without resulting in the severe phenotypes mentioned above and linked to known pathogenic *KIF1B* mutations.

In addition, several other SNPs approaching, but not reaching, genome-wide significance were noted for their location in or near genes with plausible biological linkage to viral replication or immunity, including: *MGAT5B*, an acetyl-glucosaminyltransferase isozyme that may be involved in processing HIV-1 protein glycosylation (51); *CSNK1G3*, a serine/threonine kinase that has been shown to bind the HIV-1 vpu protein *in vitro* (52); and finally *PARK2*, which has primarily been associated with Parkinson's Disease (53) but may also act in hepatitis C replication (54).

This study did not replicate any previous HSV-2 associations either in the full analysis or targeted candidate gene and MHC-region analyses. For primary HSV-1, it is well-established that deficiencies in the TLR3 pathway lead to the severe phenotype of HSE in children (22,23). It is not surprising that there were no associations with *TLR3* or genes, such as *UNC93B* or *TRIF*, that are upstream or downstream of *TLR3* and initiate type I interferon signalling. Mechanistic studies of these genes associated with pediatric primary HSV-1 have shown that they act intrinsically in neurons to reduce HSV-1 replication during the innate phase of the initial response (24), while recurrent shedding, the phenotype examined in the present study, relates to epithelial cell replication and immune cell function. Both HSV-2 viral shedding and lesion rates have previously been linked to two SNPs (rs4696480 and rs1898830) in another toll-like receptor, *TLR2* (20). *APOE* has also been linked to HSV severity, but was only associated with HSV-1 oral lesions, not HSV-2 genital viral shedding, the severity phenotype under consideration here (19). None of the top SNPs in the present study were in or near *APOE* or *TLR2*. Further, the *TLR2* SNP rs1898830 identified previously for HSV-2 was directly genotyped in our study and was not significant (p=0.57). Thus, associations with *APOE* or *TLR2*, the two genes previously implicated in candidate gene studies utilizing a portion of this cohort, were not able to be replicated at the genome-wide level in the current cohort.

In the non-MHC candidate gene analysis, though no SNPs reached significance after Bonferroni multiple testing correction, the top SNPs were primarily located in *MBL2*, a component of the innate immune system. A *MBL2* structural variant was previously identified as more common among participants with recurrent (symptomatic) HSV-2 than asymptomatic individuals or healthy controls in a small candidate gene study (27). The only additional SNPs with a $p<0.05$ were rs4696483 (p=0.02), which is intronic in *TLR2* and not in LD with the previously identified *TLR2* candidate SNP rs1898830, and rs2147419 (p=0.04), which is located in intron 2 of the *FAS* gene. *FAS* regulates activation-induced cell death and two polymorphisms, 1377G>A (rs2234767) and 670A>G (rs1800682), were previously implicated in HSV-2 susceptibility in a small candidate gene study of South African women that focused on three *FAS* and *FASLG* SNPs (28). Amongst these, neither the *FAS* SNPs (rs2234767, p=0.22; rs1800682, p=0.89) nor the candidate *FASLG* SNP

(rs763110, p=0.21) were significant in our analysis. The lack of replication of previously implicated HSV-2 SNPs underscores the need for rigorous replication of disease-associated genetic variants and the importance of determining biological mechanisms, if at all possible, particularly for variants identified through candidate gene studies.

While some non-significant SNPs were identified in potentially biologically plausible genes, we have been unable to demonstrate the presence of any common genetic variants robustly associated with HSV-2 viral shedding rate at the genome-wide level. Though the available sample size was modest for a GWAS, thus limiting our study power, viral shedding is a unique, robust and quantitative measure of HSV-2 virologic severity, making it important to investigate. *Post hoc* power calculations indicated that we had >99% power to detect a common SNP at MAF=5% that accounted for at least 25% of the variance of HSV-2 viral shedding, while we had 71.53% power to detect SNPs that accounted for 15% of the shedding variance, and only 27.29% power to detect SNPs that accounted for 10% of the shedding variance. Thus, while we had reasonable power to detect SNPs with large effect sizes, our ability to detect weaker effects was limited. However, the lack of statistical significance in this study suggests that there is no single common (MAF>5%) SNP that explains a large portion of HSV-2 viral shedding, adding valuable information about the genetic architecture of HSV-2 severity. It is possible that multiple common variants of smaller effect sizes act in HSV-2 severity or, as suggested by the *post hoc* linkage analyses, it may be that rare, rather than common, human genetic variation has a role in HSV-2 severity, as has been the case with many complex diseases (55–57) and which was not within the scope of our study design.

There remains active debate on the best surrogate measure for HSV-2 severity. Some previous studies have focused on active viral lesions or the dichotomy of asymptomatic or symptomatic diagnosis as a measure of HSV-2 severity, and measures of lesion recurrences are commonly used as an endpoint for drug or vaccine trials. While viral shedding rates correlate with lesion rates, the viral shedding rate has previously been shown to be a more accurate and consistent measure of HSV-2 viral reactivation than lesion rates, as viral shedding can occur at times lacking lesions, and some individuals who are initially asymptomatic may later recognize lesions, implying that the lesion rate is influenced by subject perception (16,58). Further, previously conducted candidate gene studies using a portion of this cohort implicated a role for host genetics in multiple measures of severity, including shedding rates (19–21). Though these associations were not replicated in the full GWAS of the current cohort, this is not uncommon, as most candidate gene associations are not replicated (59); indeed, we were unable to replicate any previous HSV-2 candidate gene associations in our GWAS analysis.

Despite the lack of evidence for a role of common host genetic variation of larger effect sizes in HSV-2 severity, as measured by the viral shedding rate in this study, it remains likely that host and viral factors interact with the environment to control HSV-2 reactivation, as with other herpesviruses (22,23,29,30). While these analyses were adjusted for age, gender, and ancestry, it is possible that additional factors might confound genetic associations with HSV-2 shedding severity, such as time since HSV-2 acquisition, HSV-2 inoculum size, or viral strain. HSV-2 severity as measured by both genital lesions and shedding rate can

decrease with time since acquisition (60). In order to interrogate the full range of HSV-2 virologic severity, including asymptomatic individuals, for whom information on time since acquisition is not available, we did not adjust for time since HSV-2 acquisition. However, inclusion of time since HSV-2 acquisition in the main analysis did not dramatically change the results (see Supplementary Table 4). While HSV-1 can cause genital herpes, oral HSV-1 seropositivity does not affect genital HSV-2 viral shedding (16). Thus, we included individuals co-infected with HSV-1, as all individuals in this cohort had Western blot confirmed genital HSV-2 and a majority of the global population has oral herpes.

Although we focused our study on a reasonably sized and well-characterized cohort of HSV-2 positive individuals of European ancestry, with unique quantification of HSV-2 viral shedding rate over at least a month, this represents a convenience cohort and may not be representative of the general population (16). Of note, approximately 80% of HSV-2 seropositive individuals are asymptomatic (4), while our cohort was 83% symptomatic, such that we may have under-represented persons with milder phenotypes. Larger studies of individuals across ethnicities and the HSV-2 severity spectrum will be needed to determine the role of human genetics, both for common and rare variation. In particular, we currently lack large cohorts with robust HSV-2 phenotyping, including information on the viral shedding rate, which will be important to gather for future studies in order to increase study power to detect genetic associations. HSV-2 reactivation remains a complicated and poorly understood process involving both host and viral factors, without a cure in sight. Though the role of human genetics in the rare and extremely severe HSV-1 caused HSE is undisputed, it remains unclear how strongly human genetic variation affects genital HSV-2 severity.

Materials and Methods

Participants

Western blot confirmed HSV-2 seropositive North American participants followed at the University of Washington were included in this study. All participants signed informed consent for genetics studies and institutional IRBs approved this study. This cohort has detailed phenotypic and quantitative information available, as has been described previously (16). Briefly, participants in this cohort were at least age 18, HIV-1 negative, not on antiviral treatments during the study period, and exhibited a wide range of disease severity by viral shedding and lesion rates. As part of the study protocol, quantitative data on daily viral shedding over a period of at least 30 days were collected prospectively, a window within which more than 77% of both asymptomatic and symptomatic individuals show viral shedding (17,58).

Viral shedding was determined using real-time PCR of self-sampled anogenital swabs with a cut-off of >150 copies of HSV-2 DNA per specimen, which has been validated to be an accurate measurement of HSV shedding (17). The viral shedding rate was calculated as the number of days PCR positive for HSV-2 DNA over the period of sampled days and has previously been shown to accurately represent viral behaviour (16). Any symptomatic HSV-2 episodes of lesions were recorded by self-reported diary entries and requested confirmatory clinical visits during the ≥30 day sampling period.

GWAS

Genotyping—A total of ~4.3 million SNPs in 307 participants were genotyped using the Illumina HumanOmni5Exome array platform (San Diego, CA). Of these, 191 participants were genotyped on the Omni5Exome4v1-1 array and 116 on the Omni5Exome4v1.0 array. The full cohort was combined for data analysis, including only non-monomorphic SNPs that were shared between the two arrays and where the DNA strand could be definitively determined (all symmetric (A/T or G/C) SNPs were excluded). The combined dataset included ~4 million SNPs.

Quality Control (QC)—A series of QC checks and all subsequent analyses were carried out to ensure sample integrity using PLINK (61). Gender was assessed for concordance between genetically-inferred and self-reported gender, with two discordant participants removed. Duplicate samples and cryptic relatedness (Identity by Descent>0.125) were identified based on genetic data; two pairs of duplicate samples were removed. Data quality was also evaluated at the marker level to remove low quality genotypes. Four participants with anomalously high or low heterozygosity (-0.07<F>0.07) were excluded and markers missing >1% of genotype calls were removed. In addition, rare variants with a MAF<5% were excluded. Following initial QC, there were 297 individuals of all ethnicities, 245 of whom self-reported as Caucasian. Quantitative estimates of genomic ancestry (principal components analysis (PCA) implemented with EIGENSTRAT software (62)) were then performed and compared with self-reported ethnicity, with outliers removed. Following QC and EIGENSTRAT, full phenotypic information and genotype data were available for 1 539 908 SNPs for 223 PCA-confirmed individuals of European ancestry.

Statistical Analysis—As 83% of the genotyped cohort self-reported as Caucasian, analyses were restricted to only the subset of participants PCA-confirmed as having European ancestry (N=223) in order to reduce population heterogeneity. We used the rate of HSV-2 viral shedding (quantified by the percent of days PCR+ for HSV-2 DNA over the sampling period) to measure HSV-2 virologic severity. The association of each SNP with the quantitative viral shedding was tested by linear regression under an additive model, including all 223 individuals of European ancestry, and adjusted for age, sex, and principal component (PC) axes that significantly contributed to the variance (PC1-3). Statistical models were examined to ensure they were robust and performed well under the assumptions of the tests. We used the standardized value of p<5E-08 (63–65) as the threshold for genome-wide statistical significance for our main regression analysis.

To investigate whether the top SNP and biological candidate *KIF1B* SNPs were tagging functional variation in nearby coding genes, LD within ±1MB was tested using D′ among 640 previously WGS population controls of European ancestry with IRB permission for use. Functional coding variation included stop gain/loss, frame-shift, start gain/loss, non-synonymous, splice site acceptor/donor, and indel variants.

Post hoc power analysis calculations of the quantitative viral shedding trait were conducted using the Genetic Power Calculator (66) for quantitative traits, under the following assumptions: sample size was N=223, marker allele frequency=5%, there was no dominance

(additive model), individuals were unrelated, there was perfect LD, the genome-wide statistical significance threshold was 5E-08, and either the power or detectable variance explained was varied.

Targeted Analyses of Candidate SNPs—Two targeted approaches were used to investigate SNPs with a higher prior probability of association with HSV-2. Given the central role of T-cells in several types of chronic viral infection and prior associations with other alphaherpesviradae, a targeted analysis of the MHC gene region (hg19/Ch37 chromosome 6: 29,570,005-33,377,699 (67)) was conducted. Other non-MHC genes previously implicated in HSV pathogenesis from candidate gene studies (*FASLG, TLR3, TLR2, MBL2, FAS, UNC93B1, TRAF3, TBX21, APOE*, and *C21orf91*) were tested for enrichment of association beyond that expected under the null hypothesis. In our cohort, there were 8791 SNPs in the MHC gene region and 131 SNPs among the non-MHC candidate genes genotyped with a MAF>5%. For targeted analyses of candidate SNPs, the threshold for statistical significance was determined using Bonferroni correction for the number of markers tested in each analysis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank Dr. Amalia Magaret for her statistical input. This work was supported by NIH grants P01 AI030731 (AW, DMK) and R01 AI094019 (DMK).

DMK is a consultant to Glaxo SmithKline and has research grants or contracts from Immune Design Corporation, Admedus Vaccines, Merck, and Sanofi Pasteur. DMK and AW are co-inventors on patents owned by the University of Washington. AW receives research funding from Genocea and Vical and is a consultant for Aicuris. DBG receives research funding from Janssen, AstraZeneca, Gilead, Biogen, and UCB, is a consultant for AstraZeneca, and holds a patent for *IL28B* findings.

References

1. Looker KJ, Magaret AS, Turner KME, Vickerman P, Gottlieb SL, Newman LM. Global Estimates of Prevalent and Incident Herpes Simplex Virus Type 2 Infections in 2012. PLoS ONE. 2015 Jan 21.cited 2016 Jul 1310(1)
2. Bradley H, Markowitz LE, Gibson T, McQuillan GM. Seroprevalence of Herpes Simplex Virus Types 1 and 2--United States, 1999-2010. J Infect Dis. 2013 Oct 16.
3. Johnston C, Gottlieb SL, Wald A. Status of vaccine research and development of vaccines for herpes simplex virus. Vaccine. 2016 Jun 3; 34(26):2948–52. [PubMed: 26973067]
4. Wald A, Zeh J, Selke S, Warren T, Ryncarz AJ, Ashley R, et al. Reactivation of genital herpes simplex virus type 2 infection in asymptomatic seropositive persons. N Engl J Med. 2000 Mar 23; 342(12):844–50. [PubMed: 10727588]
5. Kimberlin DW, Rouse DJ. Clinical practice. Genital herpes. N Engl J Med. 2004 May 6; 350(19): 1970–7. [PubMed: 15128897]
6. Sexually transmitted diseases treatment guidelines 2002. Centers for Disease Control and Prevention. MMWR Recomm Rep Morb Mortal Wkly Rep Recomm Rep Cent Dis Control. 2002 May 10; 51(RR-6):1–78.
7. [cited 2013 Oct 31] STD Facts - Genital Herpes [Internet]. Available from: http://www.cdc.gov/std/herpes/STDFact-herpes.htm

8. Corey L, Wald A, Patel R, Sacks SL, Tyring SK, Warren T, et al. Once-Daily Valacyclovir to Reduce the Risk of Transmission of Genital Herpes. N Engl J Med. 2004; 350(1):11–20. [PubMed: 14702423]

9. Ross K, Johnston C, Wald A. Herpes simplex virus type 2 serological testing and psychosocial harm: a systematic review. Sex Transm Infect. 2011 Dec 1; 87(7):594–600. [PubMed: 21903980]

10. Carney O, Ross E, Bunker C, Ikkos G, Mindel A. A prospective study of the psychological impact on patients with a first episode of genital herpes. Genitourin Med. 1994 Feb; 70(1):40–5. [PubMed: 8300099]

11. Sacks SL, Griffiths PD, Corey L, Cohen C, Cunningham A, Dusheiko GM, et al. HSV-2 transmission. Antiviral Res. 2004 Aug; 63(1):S27–35. [PubMed: 15450383]

12. Freeman EE, Weiss HA, Glynn JR, Cross PL, Whitworth JA, Hayes RJ. Herpes simplex virus 2 infection increases HIV acquisition in men and women: systematic review and meta-analysis of longitudinal studies. AIDS Lond Engl. 2006 Jan 2; 20(1):73–83.

13. Gray RH, Li X, Wawer MJ, Serwadda D, Sewankambo NK, Wabwire-Mangen F, et al. Determinants of HIV-1 load in subjects with early and later HIV infections, in a general-population cohort of Rakai, Uganda. J Infect Dis. 2004 Apr 1; 189(7):1209–15. [PubMed: 15031789]

14. Martin ET, Krantz E, Gottlieb SL, Magaret AS, Langenberg A, Stanberry L, et al. A Pooled Analysis of the Effect of Condoms in Preventing HSV-2 Acquisition. Arch Intern Med. 2009 Jul 13; 169(13):1233–40. [PubMed: 19597073]

15. Magaret AS, Mujugira A, Hughes JP, Lingappa J, Bukusi EA, DeBruyn G, et al. Effect of Condom Use on Per-act HSV-2 Transmission Risk in HIV-1, HSV-2-discordant Couples. Clin Infect Dis Off Publ Infect Dis Soc Am. 2016 Feb 15; 62(4):456–61.

16. Tronstein E, Johnston C, Huang ML, Selke S, Magaret A, Warren T, et al. Genital shedding of herpes simplex virus among symptomatic and asymptomatic persons with HSV-2 infection. JAMA J Am Med Assoc. 2011 Apr 13; 305(14):1441–9.

17. Magaret AS, Johnston C, Wald A. Use of the designation “shedder” in mucosal detection of herpes simplex virus DNA involving repeated sampling. Sex Transm Infect. 2009 Aug 1; 85(4):270–5. [PubMed: 19211593]

18. Szpara ML, Gatherer D, Ochoa A, Greenbaum B, Dolan A, Bowden RJ, et al. Evolution and diversity in human herpes simplex virus genomes. J Virol. 2013 Nov 13.

19. Koelle DM, Magaret A, Warren T, Schellenberg GD, Wald A. APOE genotype is associated with oral herpetic lesions but not genital or oral herpes simplex virus shedding. Sex Transm Infect. 2010 Jun; 86(3):202–6. [PubMed: 20410080]

20. Bochud PY, Magaret AS, Koelle DM, Aderem A, Wald A. Polymorphisms in TLR2 are associated with increased viral shedding and lesional rate in patients with genital herpes simplex virus Type 2 infection. J Infect Dis. 2007 Aug 15; 196(4):505–9. [PubMed: 17624834]

21. Magaret A, Dong L, John M, Mallal SA, James I, Warren T, et al. HLA Class I and II alleles, heterozygosity and HLA-KIR interactions are associated with rates of genital HSV shedding and lesions. Genes Immun. 2016 Dec; 17(7):412–8. [PubMed: 27853144]

22. Zhang SY, Jouanguy E, Ugolini S, Smahi A, Elain G, Romero P, et al. TLR3 Deficiency in Patients with Herpes Simplex Encephalitis. Science. 2007 Sep 14; 317(5844):1522–7. [PubMed: 17872438]

23. Zhang SY, Casanova JL. Inborn errors underlying herpes simplex encephalitis: From TLR3 to IRF3. J Exp Med. 2015 Aug 24; 212(9):1342–3. [PubMed: 26304982]

24. Lafaille FG, Pessach IM, Zhang SY, Ciancanelli MJ, Herman M, Abhyankar A, et al. Impaired intrinsic immunity to HSV-1 in human iPSC-derived TLR3-deficient CNS cells. Nature. 2012 Nov 29; 491(7426):769–73. [PubMed: 23103873]

25. Svensson A, Tunback P, Nordstrom I, Padyukov L, Eriksson K. Polymorphisms in Toll-like receptor 3 confer natural resistance to human herpes simplex virus type 2 infection. J Gen Virol. 2012 May 2; 93(Pt_8):1717–24. [PubMed: 22552940]

26. Svensson A, Bergin AMH, Löwhagen GB, Tunbäck P, Bellner L, Padyukov L, et al. A 3′-untranslated region polymorphism in the TBX21 gene encoding T-bet is a risk factor for genital

herpes simplex virus type 2 infection in humans. J Gen Virol. 2008 Sep 1; 89(9):2262–8. [PubMed: 18753235]

27. Seppänen M, Lokki ML, Lappalainen M, Hiltunen-Back E, Rovio AT, Kares S, et al. Mannose-binding lectin 2 gene polymorphism in recurrent herpes simplex virus 2 infection. Hum Immunol. 2009 Apr; 70(4):218–21. [PubMed: 19480845]
28. Chatterjee K, Dandara C, Gyllensten U, van der Merwe L, Galal U, Hoffman M, et al. A fas gene polymorphism influences herpes simplex virus type 2 infection in South African women. J Med Virol. 2010; 82(12):2082–2086. [PubMed: 20981796]
29. Crosslin DR, Carrell DS, Burt A, Kim DS, Underwood JG, Hanna DS, et al. Genetic variation in the HLA region is associated with susceptibility to herpes zoster. Genes Immun. 2015 Jan; 16(1): 1–7. [PubMed: 25297839]
30. 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 Feb 1; 197(3):340–6. [PubMed: 18199027]
31. 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 Dec 1; 204(11):1654–62. [PubMed: 22039568]
32. Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature. 2009 Sep 17; 461(7262):399–401. [PubMed: 19684573]
33. Fellay J, Ge D, Shianna KV, Colombo S, Ledergerber B, Cirulli ET, et al. Common genetic variation and the control of HIV-1 in humans. PLoS Genet. 2009 Dec.5(12):e1000791. [PubMed: 20041166]
34. Lingappa JR, Petrovski S, Kahle E, Fellay J, Shianna K, McElrath MJ, et al. Genomewide Association Study for Determinants of HIV-1 Acquisition and Viral Set Point in HIV-1 Serodiscordant Couples with Quantified Virus Exposure. PLoS ONE. 2011 Dec 12.cited 2013 Oct 296(12)
35. Petrovski S, Fellay J, Shianna KV, Carpenetti N, Kumwenda J, Kamanga G, et al. Common human genetic variants and HIV-1 susceptibility: a genome-wide survey in a homogeneous African population. AIDS Lond Engl. 2011 Feb 20; 25(4):513–8.
36. The 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature. 2015 Oct 1; 526(7571):68–74. [PubMed: 26432245]
37. Hook LM, Lubinski JM, Jiang M, Pangburn MK, Friedman HM. Herpes simplex virus type 1 and 2 glycoprotein C prevents complement-mediated neutralization induced by natural immunoglobulin M antibody. J Virol. 2006 Apr; 80(8):4038–46. [PubMed: 16571820]
38. Bodzioch M, Orsó E, Klucken J, Langmann T, Böttcher A, Diederich W, et al. The gene encoding ATP-binding cassette transporter 1 is mutated in Tangier disease. Nat Genet. 1999 Aug; 22(4): 347–51. [PubMed: 10431237]
39. Marcil M, Yu L, Krimbou L, Boucher B, Oram JF, Cohn JS, et al. Cellular cholesterol transport and efflux in fibroblasts are abnormal in subjects with familial HDL deficiency. Arterioscler Thromb Vasc Biol. 1999 Jan; 19(1):159–69. [PubMed: 9888879]
40. Brooks-Wilson A, Marcil M, Clee SM, Zhang LH, Roomp K, van Dam M, et al. Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency. Nat Genet. 1999 Aug; 22(4):336–45. [PubMed: 10431236]
41. Zwarts KY, Clee SM, Zwinderman AH, Engert JC, Singaraja R, Loubser O, et al. ABCA1 regulatory variants influence coronary artery disease independent of effects on plasma lipid levels. Clin Genet. 2002 Feb; 61(2):115–25. [PubMed: 11940086]
42. Mujawar Z, Tamehiro N, Grant A, Sviridov D, Bukrinsky M, Fitzgerald ML. Mutation of the ATP cassette binding transporter A1 (ABCA1) C-terminus disrupts HIV-1 Nef binding but does not block the Nef enhancement of ABCA1 protein degradation. Biochemistry (Mosc). 2010 Sep 28; 49(38):8338–49.
43. Cui HL, Grant A, Mukhamedova N, Pushkarsky T, Jennelle L, Dubrovsky L, et al. HIV-1 Nef mobilizes lipid rafts in macrophages through a pathway that competes with ABCA1-dependent cholesterol efflux. J Lipid Res. 2012 Apr; 53(4):696–708. [PubMed: 22262807]

44. Jacob D, Hunegnaw R, Sabyrzyanova TA, Pushkarsky T, Chekhov VO, Adzhubei AA, et al. The ABCA1 domain responsible for interaction with HIV-1 Nef is conformational and not linear. Biochem Biophys Res Commun. 2014 Jan 31; 444(1):19–23. [PubMed: 24406162]
45. Jennelle L, Hunegnaw R, Dubrovsky L, Pushkarsky T, Fitzgerald ML, Sviridov D, et al. HIV-1 protein Nef inhibits activity of ATP-binding cassette transporter A1 by targeting endoplasmic reticulum chaperone calnexin. J Biol Chem. 2014 Oct 17; 289(42):28870–84. [PubMed: 25170080]
46. Sheng XX, Sun YJ, Zhan Y, Qu YR, Wang HX, Luo M, et al. The LXR ligand GW3965 inhibits Newcastle disease virus infection by affecting cholesterol homeostasis. Arch Virol. 2016 Sep; 161(9):2491–501. [PubMed: 27357231]
47. Bocchetta S, Maillard P, Yamamoto M, Gondeau C, Douam F, Lebreton S, et al. Up-regulation of the ATP-binding cassette transporter A1 inhibits hepatitis C virus infection. PloS One. 2014; 9(3):e92140. [PubMed: 24646941]
48. Matsuura K, Isogawa M, Tanaka Y. Host genetic variants influencing the clinical course of Hepatitis B virus infection. J Med Virol. 2016 Mar; 88(3):371–9. [PubMed: 26255971]
49. Malikov V, da Silva ES, Jovasevic V, Bennett G, de Souza Aranha Vieira DA, Schulte B, et al. HIV-1 capsids bind and exploit the kinesin-1 adaptor FEZ1 for inward movement to the nucleus. Nat Commun. 2015 Mar 30.6:6660. [PubMed: 25818806]
50. Wisner TW, Sugimoto K, Howard PW, Kawaguchi Y, Johnson DC. Anterograde transport of herpes simplex virus capsids in neurons by both separate and married mechanisms. J Virol. 2011 Jun; 85(12):5919–28. [PubMed: 21450818]
51. Land A, Braakman I. Folding of the human immunodeficiency virus type 1 envelope glycoprotein in the endoplasmic reticulum. Biochimie. 2001 Aug; 83(8):783–90. [PubMed: 11530211]
52. Miyakawa K, Sawasaki T, Matsunaga S, Tokarev A, Quinn G, Kimura H, et al. Interferon-induced SCYL2 limits release of HIV-1 by triggering PP2A-mediated dephosphorylation of the viral protein Vpu. Sci Signal. 2012 Oct.95(245):ra73.
53. Ferreira M, Massano J. An updated review of Parkinson's disease genetics and clinicopathological correlations. Acta Neurol Scand. 2016 Jun 1.:n/a–n/a.
54. Hara Y, Yanatori I, Ikeda M, Kiyokage E, Nishina S, Tomiyama Y, et al. Hepatitis C Virus Core Protein Suppresses Mitophagy by Interacting with Parkin in the Context of Mitochondrial Depolarization. Am J Pathol. 2014 Nov; 184(11):3026–39. [PubMed: 25244949]
55. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat Rev Genet. 2010 Jun; 11(6):415–25. [PubMed: 20479773]
56. Need AC, Shashi V, Hitomi Y, Schoch K, Shianna KV, McDonald MT, et al. Clinical application of exome sequencing in undiagnosed genetic conditions. J Med Genet. 2012 Jun; 49(6):353–61. [PubMed: 22581936]
57. Cirulli ET, Lasseigne BN, Petrovski S, Sapp PC, Dion PA, Leblond CS, et al. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. Science. 2015 Mar 27; 347(6229):1436–41. [PubMed: 25700176]
58. Mujugira A, Huang ML, Selke S, Drolette L, Magaret AS, Wald A. High Rate of β-Globin DNA Detection Validates Self-Sampling in Herpes Simplex Virus Shedding Studies. Sex Transm Dis. 2015 Dec; 42(12):705–9. [PubMed: 26562701]
59. Siontis KCM, Patsopoulos NA, Ioannidis JPA. Replication of past candidate loci for common diseases and phenotypes in 100 genome-wide association studies. Eur J Hum Genet EJHG. 2010 Jul; 18(7):832–7. [PubMed: 20234392]
60. Phipps W, Saracino M, Magaret A, Selke S, Remington M, Huang ML, et al. Persistent genital herpes simplex virus-2 shedding years following the first clinical episode. J Infect Dis. 2011 Jan 15; 203(2):180–7. [PubMed: 21288817]
61. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007 Sep; 81(3):559–75. [PubMed: 17701901]
62. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006 Aug; 38(8):904–9. [PubMed: 16862161]

63. Sham PC, Purcell SM. Statistical power and significance testing in large-scale genetic studies. Nat Rev Genet. 2014 Apr 17; 15(5):335–46. [PubMed: 24739678]

64. Li MX, Yeung JMY, Cherny SS, Sham PC. Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. Hum Genet. 2012 May; 131(5):747–56. [PubMed: 22143225]

65. Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet Epidemiol. 2008 May; 32(4):381–5. [PubMed: 18348202]

66. Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinforma Oxf Engl. 2003 Jan; 19(1):149–50.

67. Shiina T, Hosomichi K, Inoko H, Kulski JK. The HLA genomic loci map: expression, interaction, diversity and disease. J Hum Genet. 2009 Jan 9; 54(1):15–39. [PubMed: 19158813]

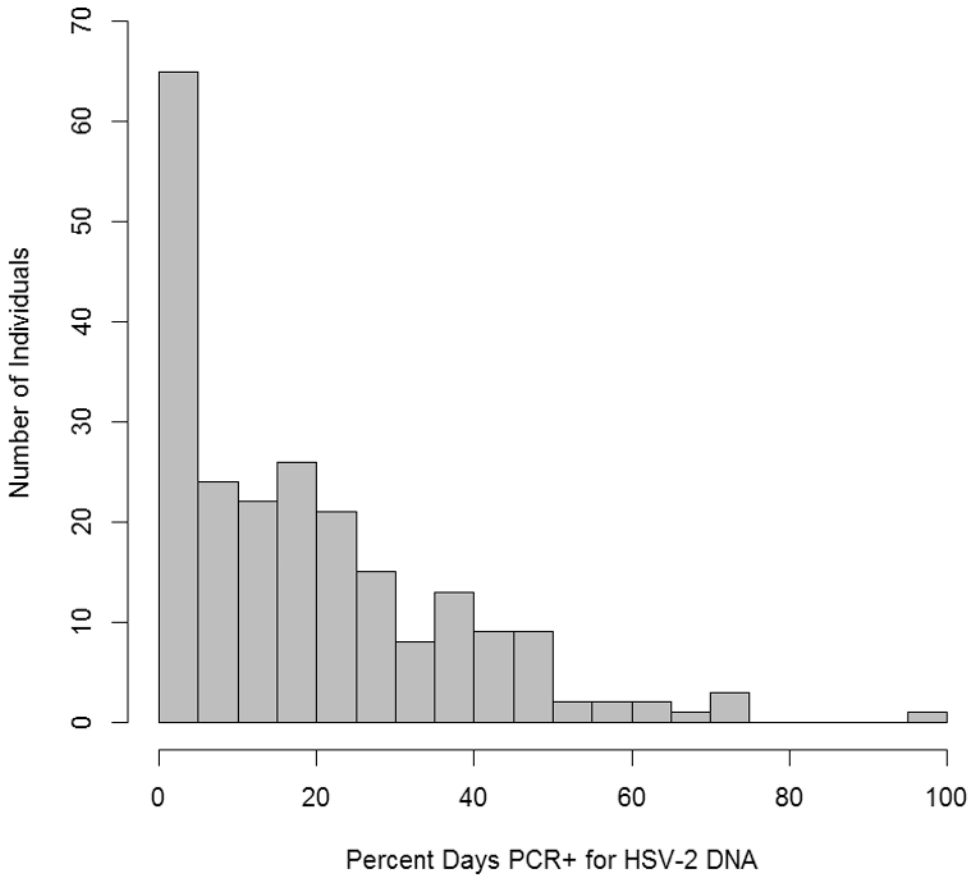

Figure 1.
Shedding rate (percent of days PCR+ for HSV-2 DNA) distribution among individuals of European ancestry (N=223).

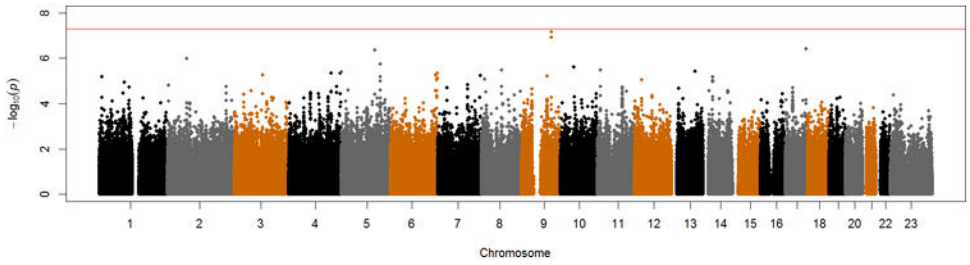

Figure 2.

Manhattan plot of the GWAS for HSV-2 viral severity among individuals of European ancestry (N=223). Linear regression model. The red line indicates the genome-wide statistical significance threshold (p<5E-08).

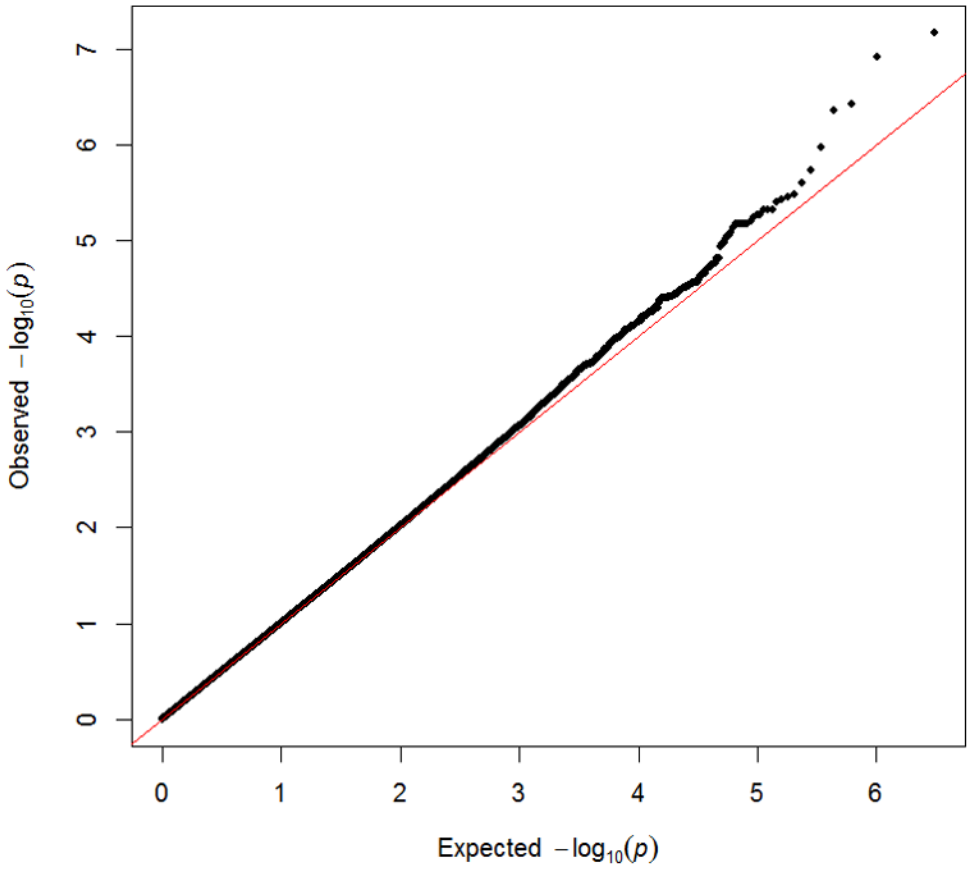

Figure 3.
QQ plot of the GWAS for HSV-2 viral severity among individuals of European ancestry (N=223). Linear regression model. No genomic inflation was observed (lambda=1.02).

Table 1

Participant demographics for the subset of the cohort included in the final analyses (N=223 with genetically confirmed European ancestry)

		N=223
Sex	Male, N (%)	86 (38.57%)
	Female, N (%)	137 (61.43%)
Age (years), median (range)		39.5 (22-76)
Diagnosis, N (%)	Symptomatic	186 (83.41%)
	Asymptomatic	37 (16.59%)
Days since diagnosis[1], median (range)		3407 (14-12649)
Median viral shedding rate, % (range)		15.3 (0-100%)
Median lesion rate[1], % (range)		6.3 (0-85.4%)
HSV-1 positive, N (%)		97 (43.5%)

[1]Data not available for all samples.

Table 2

The top 10 SNPs for HSV-2 viral severity among individuals of European ancestry (N=223). Linear regression analysis, adjusted for age, sex, and significant PC axes.

Rank	SNP	SNP Type	Nearest Gene	P-value
1	rs75932292	intergenic	-	6.77E-08
2	rs73664402	intergenic	-	1.21E-07
3	rs56122323	intergenic	*LOC105371899* and *MGAT5B*	3.76E-07
4	rs62377770	intergenic	*CEP120* and *CSNK1G3*	4.29E-07
5	rs55963884	intronic / upstream 2KB	*ARHGAP25*	1.06E-06
6	rs4912855	intergenic	*LOC101926941*	1.84E-06
7	rs11204209	intronic	*ZNF488*	2.51E-06
8	rs4910264	intronic	*LOC105376548* and *LOC105376550*	3.28E-06
9	rs59849217	intergenic	-	3.45E-06
10	rs75644638	intergenic	-	3.75E-06