The First Common Cold Sore Susceptibility Gene

Anthony L. Cunningham¹ ² and David Booth¹ ²

¹Centre for Virus Research, Institute of Immunology and Allergy Research, Westmead Millennium Institute, ²University of Sydney, New South Wales, Australia

(See the article by Kriesel et al, on pages 1654–62.)

The outcome of any interaction between a virus and host is determined by genetic and environmental factors and is manifest as either susceptibility to or chronicity of disease [1]. This interplay between host genetics and viral strain genomic sequence has been well defined in murine models of different viral infections, such as murine cytomegalovirus, retroviruses (Friend leukemia virus), and influenza virus [2–4]. In some human viral infections, eg, human immunodeficiency virus (HIV) infection, the effects of various host genes on viral acquisition, disease, and survival have been well defined. For example, homozygosity for a large deletion in the HIV coreceptor chemokine (C-C motif) receptor 5 (CCR5) is highly protective against viral acquisition [5]. KIR, CCR5, CCR2, and HLA-B57 variants delay HIV disease progression in white persons through immunologic mechanisms, whereas HLA-B35 accelerates it [6–8]. To date, however, only some of the variation in disease progression has been explained by genetic studies.

In human infections with herpes simplex virus types 1 (HSV-1) and 2 (HSV-2) there is added complexity in that both are shed asymptomatically in the saliva and genital secretions, respectively. Clinical lesions develop in only a portion of the shedding population. Worldwide, 50%–100% and 10%–90% of populations are infected with HSV-1 and HSV-2, respectively, but only a minority of infections present with disease, which varies remarkably in frequency and severity, eg, 20%–30% for herpes simplex labialis [8, 9]. In animal models, the viral strain, size of inoculum, and strain of animal help determine initial disease severity and latent viral burden in the spinal ganglia that, in turn, influences frequency of reactivation. Therefore, it is not surprising that a range of human genes might influence different phenotypes of herpes labialis, including initial disease severity, frequency and severity of recurrences, and/or frequency and extent of viral shedding (a determinant of spread or infectiousness). In mice and humans, several genes predisposing to severe initial disease and mortality have been identified. In mice, these include 3 loci on chromosome 6, 1 linked to the natural killer cell complex, and 2 near the tumor necrosis factor α gene locus, 1 of which is sex linked [10]; in humans, deficiency of Toll-like receptor (TLR) 3/UNC-93B has been linked to herpes simplex encephalitis [11]. In mice, such genetic resistance varies according to the route of inoculation and is eliminated by intracerebral inoculation. These findings are consistent with the observation that resistance to infection of the dorsal route ganglion, establishment of viral latency, and subsequent reactivation involves multiple genetic loci [11, 12]. Mice are poor models of recurrent herpetic infections, and influenza virus [2–4]. In some human viral infections, eg, human immunodeficiency virus (HIV) infection, the effects of various host genes on viral acquisition, disease, and survival have been well defined. For example, homozygosity for a large deletion in the HIV coreceptor chemokine (C-C motif) receptor 5 (CCR5) is highly protective against viral acquisition [5]. KIR, CCR5, CCR2, and HLA-B57 variants delay HIV disease progression in white persons through immunologic mechanisms, whereas HLA-B35 accelerates it [6–8]. To date, however, only some of the variation in disease progression has been explained by genetic studies.

In human infections with herpes simplex virus types 1 (HSV-1) and 2 (HSV-2) there is added complexity in that both are shed asymptomatically in the saliva and genital secretions, respectively. Clinical lesions develop in only a portion of the shedding population. Worldwide, 50%–100% and 10%–90% of populations are infected with HSV-1 and HSV-2, respectively, but only a minority of infections present with disease, which varies remarkably in frequency and severity, eg, 20%–30% for herpes simplex labialis [8, 9]. In animal models, the viral strain, size of inoculum, and strain of animal help determine initial disease severity and latent viral burden in the spinal ganglia that, in turn, influences frequency of reactivation. Therefore, it is not surprising that a range of human genes might influence different phenotypes of herpes labialis, including initial disease severity, frequency and severity of recurrences, and/or frequency and extent of viral shedding (a determinant of spread or infectiousness). In mice and humans, several genes predisposing to severe initial disease and mortality have been identified. In mice, these include 3 loci on chromosome 6, 1 linked to the natural killer cell complex, and 2 near the tumor necrosis factor α gene locus, 1 of which is sex linked [10]. In humans, deficiency of Toll-like receptor (TLR) 3/UNC-93B has been linked to herpes simplex encephalitis [11]. In mice, such genetic resistance varies according to the route of inoculation and is eliminated by intracerebral inoculation. These findings are consistent with the observation that resistance to infection of the dorsal route ganglion, establishment of viral latency, and subsequent reactivation involves multiple genetic loci [11, 12]. Mice are poor models of recurrent herpetic infections. However, in human recurrent herpes caused by HSV-2, older studies linked HLA-A1 and HLA-B27 deficiency to increased frequency of genital herpes. More recently, certain TLR2 polymorphisms were shown to predispose to more frequent genital herpes and genital HSV-2 shedding [13].

In the past, human studies have been restricted to single candidate genes suggested by pathogenetic studies or to a single chromosomal region (chromosome 6) encoding the HLA system, often not explaining the strong genetic predispositions suggested by familial or twin studies. Until recently, geneticists have been frustrated in their ability to localize familial traits to specific regions of the human genome and then to genes and precise genetic sequences. The sequencing of the human genome, the availability of dense genotyping primers across the genome, and powerful analytic methods have led to routine genome-wide analyses for a number of multigenic diseases, including infectious diseases. Perhaps unsurprisingly, the most common region affecting host response to viruses and other pathogens is still the HLA region. However, non–major histocompatibility complex regions have also been identified, most notably for spontaneous clearance and drug response to the hepatitis C virus [15–17]. Here a single genetic variant of interleukin (IL) 28B was found to predict viral clearance. This discovery has been translated directly into clinical practice, so that patient genotyping is already widely used by clinicians to predict who
will respond to current treatment. IL-28B genotyping is also used in clinical trials to stratify drug responses, and the protein it codes for, interferon λ, is in phase 2 clinical trials as a novel therapeutic agent. The rapid increase in new antiviral therapies for hepatitis C virus infection highlights the value of predicting individual responses to a specific regimen, enabled largely by genotyping of this gene.

In this issue of the Journal, Kriesel et al [18] follow up previous studies that identified a region of chromosome 21 linked to herpes simplex labialis, using genome-wide family-based linkage studies [19]. They used single-nucleotide polymorphism (SNP) genotyping to identify which of 6 possible candidate genes were associated with frequency of herpes labialis. Both 2-point linkage analysis and ParenTDT analysis were used. Linkage analysis measures the likelihood of a SNP being linked to frequently affected individuals more than unaffected individuals, compared with chance. The ParenTDT tests if SNP variants are transferred to affected offspring more often than expected by chance and combines this with a test to see if the variants are overrepresented in affected versus unaffected parents.

These 2 complementary techniques identified C21orf91 as a gene of interest, which was designated as the cold sore susceptibility gene 1 (CSSG-1). After transfection into 293T cells, this gene was expressed in the cytoplasm. It is known to be expressed in chick retinal precursor cells as well as anterior epithelial cells of the lens during development. Although retinal cells can be a site of HSV-1 infection, this provides little clue as to the function of CSSG-1. In addition, CSSG-1 messenger RNA has also been recorded at very low levels in CD4 Th1 cells and CD8 cells (Immunological Genome Database, www.immgen.org; IRIS Database, http://share.gene.com/clark.iris.2004/iris/iris.html), shown to be key infiltrating immune cells in recurrent herpes lesions. Alternatively CSSG-1 could also play a role in innate immunity, perhaps in dendritic cells, macrophages, γδT cells, or natural killer cells, all found in lesions [20–22] and some shown to be critical in HSV control in mice. Finally, it could be an intracellular restrictive factor in dorsal root ganglion neurons (cf retinal cells) or epithelial cells. The function of many restrictive intracellular proteins, like the antiviral interferon-stimulated genes, ISG15 and myxovirus resistance 1, or APOBEC3G for HIV, remain to be discovered [23–25]. Could CSSG-1 be an interferon-induced gene? On the interferome, it is not recognized as interferon responsive (www.interferome.org). However, because it is probably predominantly expressed in an untested tissue, this is not conclusive. It is noteworthy, if perplexing, that a SNP tagging the susceptibility CSSG-1 HSL susceptibility haplotype was recently identified as associated with high systolic blood pressure in a genome-wide association study (GWAS) of African Americans [26].

There is little doubt that many more host cell genes controlling HSV infection, initial and recurrent disease, and genital tract shedding, some perhaps homologous to the murine susceptibility genes, will be discovered as the techniques for discovery of rare gene variants improve. The current study used linkage to track down the genetic association in 43 pedigrees. New discoveries may come from GWASs, which use more accessible, case-control comparisons, rather than family studies. The cost of a GWAS is continuing to fall. New genotyping chips provide better genomic coverage, with rarer SNPs and SNPs that provide more dense coverage of haplotypes, including for the ethnic groups not previously tested, notably African Americans. The successes of GWASs in defining new genetic associations for common diseases have encouraged further efforts, now for diseases caused by HSV-1.

However, the failure to find regions of strong association with herpes simplex labialis in these 43 pedigrees points to small odds ratios for the genetic variants that do alter susceptibility. Furthermore, even a GWAS involving many thousands of cases for unequivocal phenotypes such as height or for the common autoimmune diseases fail to find most of the source of heritability. This could be because they only trawl for common variants, missing the rare variants of comparatively larger effect, or for several other possible reasons: the heritability is due to interactions between genetic variants, there is a long tail of association of genetic variants of ever diminishing effect, or the causative genes or causative SNPs occur in haplotype blocks of very high linkage disequilibrium. Strategies to combat this have been designed, including more comparisons to increase statistical power, use of more SNPs to capture all haplotypes, use of less common SNPs to refine haplotypes, and use of ethnic groups with smaller haplotype blocks.

The rapidly decreasing cost of next-generation sequencing opens up a host of new gene hunting strategies, especially for rare variants of large effect. Ultimately, biological studies are needed to establish which genes in a linkage disequilibrium block are the culprits and how they affect disease susceptibility. This is particularly so for the orphan genes, about which little is known. The spotlight now falls on C21orf91, first for replication of these results, then for functional studies, perhaps in the future for therapeutic applications, and, also, for studies to determine whether it plays a similar role in recurrent genital herpes.

Notes

Financial support. This work was supported by the National Health and Medical Research Council of Australia (grants 632638 to ALC and 1006759 to D. B.).

Potential conflicts of interest. A. L. C. has consulted for GlaxoSmithKline on herpes simplex vaccines and for GlaxoSmithKline, Novartis, and 3M on antiviral medications for oral and genital herpes. D. B. certifies no potential conflicts of interest.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Tuite A, Gros P. The impact of genomics on the analysis of host resistance to