Genetic Variations in the Receptor-Ligand Pair CCR5 and CCL3L1 Are Important Determinants of Susceptibility to Kawasaki Disease


Department of Pediatrics, University of California at San Diego School of Medicine, La Jolla, California; Veterans Administration Research Center for AIDS and HIV-1 Infection, South Texas Veterans Health Care System, Audie L. Murphy Division, San Antonio, and Department of Medicine, University of Texas Health Science Center at San Antonio; Departments of Pediatrics and Pediatric Cardiology, Harvard Medical School, Boston, Massachusetts

Kawasaki disease (KD) is an enigmatic, self-limited vasculitis of childhood that is complicated by development of coronary-artery aneurysms. The high incidence of KD in Asian versus European populations prompted a search for genetic polymorphisms that are differentially distributed among these populations and that influence KD susceptibility. Here, we demonstrate a striking, inverse relationship between the worldwide distribution of CCR5-Δ32 allele and the incidence of KD. In 164 KD patient-parent trios, 4 CCR5 haplotypes including the CCR5-Δ32 allele were differentially transmitted from heterozygous parents to affected children. However, the magnitude of the reduced risk of KD associated with the CCR5-Δ32 allele and certain CCR5 haplotypes was significantly greater in individuals who also possessed a high copy number of the gene encoding CCL3L1, the most potent CCR5 ligand. These findings, derived from the largest genetic study of any systemic vasculitis, suggest a central role of CCR5-CCL3L1 gene-gene interactions in KD susceptibility and the importance of gene modifiers in infectious diseases.

Recently, intense scrutiny has focused on determining CCR5-CCL3L1 gene-gene interactions and their role in susceptibility to infectious diseases [1]. This interest stems from CCR5 serving as (1) a high-affinity receptor for the potent chemokines CCL3 and CCL3L1, which are thought to play important roles in immunity and host defense [2], and (2) the major coreceptor with CD4 for cell entry of HIV-1 [3, 4]. Homozygosity for a 32-bp deletion in the coding region of CCR5 (CCR5-Δ32) results in the loss of CCR5 surface expression, and this genotype, in turn, affords significant protection against HIV-1 infection [3]. The distribution of the CCR5-Δ32 allele in human populations is thought to follow the migration patterns of individuals who descended from the Vikings, such that the frequency of the CCR5-Δ32 allele is highest in individuals of northern European descent. Thus, Icelanders and Swedes have among the highest CCR5-Δ32 allele frequencies (14.7% and 14.2%, respectively) [5], and the allele is virtually absent in Asians and Africans. It is speculated that the CCR5 signaling pathway is important for productive infection with pox viruses and that natural selection driven by epidemics of infectious diseases such as smallpox fixed the CCR5-Δ32 allele in a population of northern Europeans [6, 7].

Given the significant interest in elucidating the relationship between infectious disease susceptibility and the differential distribution of alleles among populations, we examined the relationship between the distribution of the CCR5-Δ32 allele and Kawasaki disease (KD), which is suspected to have an infectious trigger.
There were several compelling reasons to focus on KD. First, a striking racial-susceptibility pattern is evident, with overrepresentation of KD in children of Asian ancestry [8–10]. Second, the clinical, seasonal, and epidemiologic features of KD suggest an infectious etiology [11]. Third, studies support a genetic predisposition to KD [12, 13], including the growing recognition of KD pedigrees in Japan and the United States [14].

SUBJECTS, MATERIALS, AND METHODS

**Human subjects.** All patients with KD or a history of KD who (1) met 4 of the 5 standard clinical criteria [15] or (2) met 3 of the 5 standard clinical criteria and had coronary-artery abnormalities documented by echocardiography and for whom both biological parents agreed to donate DNA samples were entered into the study, after informed consent was obtained. This protocol was reviewed and approved by the Institutional Review Board at the University of California at San Diego and Boston Children’s Hospital.

**Clinical data.** Clinical data, including sex, ethnicity, race, age at disease onset, response to intravenous γ-globulin therapy, and coronary-artery status were recorded for all subjects. All echocardiograms during the first 2 months after disease onset were recorded as either normal (all 3 vessels within 2 SDs of the mean internal diameter for body surface area of the patient, according to the Newburger criteria) or abnormal [16]. For abnormal echocardiograms (z score, >2), the z score was recorded, and the patient was classified as “dilated” (z score, >2.0 but <4.0 and returned to <2.0 within 2 months of the follow-up period) or “aneurysm/ectasia” (z score, >4.0 with focal or persistent dilatation of the coronary-artery segment).

**DNA collection.** For children <6 years of age, 3 mL of blood was collected into tubes containing EDTA, and DNA was extracted using the Wizard Genomic DNA extraction kit (Promega), as described elsewhere [17]. This procedure routinely yielded 25–75 μg of polymerase chain reaction (PCR)–quality genomic DNA. For children with KD >6 years of age and their parents, 10 mL of Scope mouthwash was used to collect shed buccal cells for DNA extraction [18]. The yield was 10–200 μg of PCR-amplifiable genomic DNA.

**Genotyping.** The methods for genotyping CCR5 polymorphisms and for the generation of CCR5 haplotypes are described elsewhere [1, 19]. CCL3L1 gene copy number was estimated as described recently, and very extensive methods are described in the supplementary online material accompanying that article [1].

**Data analysis.** Polymorphism data were subjected to Mendelian checks by use of Pedcheck software (version 1; University of Pittsburgh) [20]. Where appropriate, haplotypes were inferred by use of Genehunter software (version 2) [21], and double crossovers within genes or families of genes on the same chromosome were flagged for examination. When available, prior information regarding linkage disequilibrium among polymorphisms in the same gene or family of genes in the same chromosomal region was used to identify potential genotyping errors. If the error could not be resolved by repeated genotyping, then that triad was deleted from the study, under the assumption that either an error occurred in the collection or labeling of 1 of the 3 samples in the triad or one of the parents is not the biological parent. When this protocol was followed, only 4 families were deleted from the data set.

**Statistical analysis.** The correlation between CCR5-Δ32 mutation frequency and KD incidence was compared using Spearman’s correlation coefficient. We used the transmission disequilibrium test (TDT) [22] to assess the transmission of each CCR5 haplotype in the patient-parent trios. We used the Stata software package (version 7.0; StataCorp) (command: symmetry) to conduct the TDT analyses.

A limitation of the TDT method is that it is suited to only those marker loci that are biallelic. We have shown previously that the CCR5 locus is essentially multiallelic (with 9 haplotypes). Thus, we used the extended TDT (E-TDT) for multiallelic loci [23]. It has been demonstrated that the E-TDT has sufficient power when linkage disequilibrium is strong. To conduct this analysis, we used the program ETDT (provided for public use by D. Curtis at http://www.mds.qmul.ac.uk/statgen/dcurts/software.html). We used the case/pseudocontrol analysis described by Cordell et al. [24]. This approach, while retaining several of the advantages of family-based designs, concurrently accounts for the effects of maternal genotype and parent of origin (imprinting). In this approach, the pseudo-controls are generated from the 3 untransmitted parental genotypes, and conditional logistic regression is used to test the association of the genotypes with disease. We conducted this analysis because it allows for a multivariate estimation of the phenotypic effects of the genotypes and could therefore be used in the context of the multiallelic CCR5 locus. For the assessment of the phenotypic effects of CCR5 on the genetic background conferred by CCL3L1, we considered 3 genetic backgrounds conferred by the CCL3L1 gene copy number: <2 copies, 2 copies, and >2 copies [1]. Within each of these categories, we conducted TDT, E-TDT, and case/pseudocontrol analyses to assess whether the genetic background conferred by CCL3L1 gene copy number influenced the phenotypic effects associated with the CCR5 haplotypes.

RESULTS

There was a striking inverse correlation between the incidence of KD [8, 10, 25–31] and the frequency of the CCR5-Δ32 allele [5, 32–37] in different geographic regions (figure 1). Thus, countries with the lowest incidences of KD had the highest frequencies of the CCR5-Δ32 allele. Conversely, in Japan, the country with the highest KD incidence in the world, the prevalence of the CCR5-Δ32 allele was virtually zero. Although many polymor-
Figure 1. World map depicting the frequency of the CCR5 Δ32 allele and the incidence of Kawasaki disease (KD) in different populations. Data are shown only for those countries for which both the KD incidence and CCR5 Δ32 allele frequency were available.

Phenotypes differ in frequency between Asians and other populations, this inverse relationship between KD and the CCR5 Δ32 allele prompted us to conduct a family-based study to test the hypothesis that the CCR5 Δ32 allele confers protection against developing KD. This cohort of 164 children with KD and their biological parents is, to the best of our knowledge, the largest number of affected subjects in any reported genetic study of KD. We employed 3 complementary statistical approaches: TDT [22], E-TDT [23], and case/pseudocontrol [24] analyses.

For a marker locus with 2 alleles, the TDT compares the transmission of either allele from heterozygous parents to the affected offspring; this test and the family-based design eliminate the risk of spurious associations due to population stratification, a common confounding factor in case-control association studies [38]. By use of the TDT, we found asymmetric transmission of the CCR5 Δ32 allele from 46 heterozygous parents to their affected children (transmitted–not transmitted ratio, 15:31) (table 1, model 1), supporting the hypothesis that the CCR5 Δ32 allele might protect against development of KD.

However, CCR5 is a multiallelic gene, and, in addition to the CCR5 Δ32 allele, we have previously described 8 CCR5 haplotypes that include polymorphisms in the promoter of CCR5 and coding regions of CCR5 and CCR2 [39]. These haplotypes are categorized into CCR5 human haplogroups A (HHA) to HHG*2, with the haplotype bearing CCR5 Δ32 designated as HHG*2. Because we and others have found that CCR5 haplotypes are associated with differential susceptibility to HIV-1 infection and disease progression [19, 40, 41], we next used the E-TDT to test the hypothesis that CCR5 haplotypes influence KD susceptibility. Complete CCR5 haplotypes were obtained from 164 KD patient-parent trios. The CCR5 locus in the study subjects was in Hardy-Weinberg equilibrium (multiallelic likelihood ratio test, χ² = 35.42; 28 df; P = .1579). The E-TDT analysis indicated that there was an overall significant association between asymmetric transmission of the CCR5 haplotypes from parents to their affected children (likelihood ratio test, χ² = 18.03; 7 df; P = .0119). Because of this overall association, we next determined which specific haplotype(s) accounted for the association. Three CCR5 haplotypes were associated with a significantly reduced risk of KD (table 1, model 2): (1) HHA, (2) HHIC, and (3) HHG*2, the CCR5 Δ32-containing haplotype, which was associated with a nearly 50% (95% confidence interval [CI], 0.25–0.93) reduction in the risk of KD. By contrast, the haplotype on which the CCR5 Δ32 mutation arose [39]—namely, CCR5-HHG*1—was associated with an increased risk of developing KD (table 1, model 2).

We next determined whether CCR5 haplotypes were associated with an altered risk of coronary-artery damage in children with KD. For this analysis, we compared 83 children with normal echocardiograms with 88 children who had either coronary-artery dilatation or frank aneurysms. There was a trend toward an association between possession of the CCR5-HHG*2 haplotype and a reduced risk of coronary-artery dilatation or aneurysm (odds ratio [OR], 0.36 [95% CI, 0.12–1.07]; P = .067), which is notable, considering the small sample size.

We recently made 2 observations related to HIV-1 susceptibility that could have direct bearing on the aforementioned findings [1]. First, we found that there is significant interin-
Table 1. Results of the transmission disequilibrium test (TDT) analysis in 198 patient-parent trios.

<table>
<thead>
<tr>
<th>Model, CCR5 genotype/haplotype, CCL3L1 gene dose</th>
<th>OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>CCR5Δ32</td>
<td>0.49 (0.25–0.93)</td>
</tr>
<tr>
<td>Model 2</td>
<td>HHA</td>
<td>0.57 (0.33–0.97)</td>
</tr>
<tr>
<td></td>
<td>HHC</td>
<td>0.67 (0.47–0.97)</td>
</tr>
<tr>
<td></td>
<td>HHG*1</td>
<td>2.25 (0.98–5.17)</td>
</tr>
<tr>
<td></td>
<td>HHG*2</td>
<td>0.49 (0.25–0.93)</td>
</tr>
<tr>
<td>Model 3</td>
<td>CCR5Δ32, &lt;2 CCL3L1 copies</td>
<td>0.50 (0.11–1.87)</td>
</tr>
<tr>
<td></td>
<td>CCR5Δ32, 2 CCL3L1 copies</td>
<td>0.21 (0.04–0.77)</td>
</tr>
<tr>
<td></td>
<td>CCR5Δ32, &gt;2 CCL3L1 copies</td>
<td>1.20 (0.31–4.97)</td>
</tr>
<tr>
<td>Model 4</td>
<td>HHC, &lt;2 CCL3L1 copies</td>
<td>0.44 (0.17–1.07)</td>
</tr>
<tr>
<td></td>
<td>HHG*2, 2 CCL3L1 copies</td>
<td>0.21 (0.04–0.77)</td>
</tr>
<tr>
<td></td>
<td>HHA, &gt;2 CCL3L1 copies</td>
<td>0.29 (0.09–0.73)</td>
</tr>
</tbody>
</table>

NOTE. Models 1 and 3, TDT analysis for association between possession of the CCR5Δ32 allele or haplotypes and Kawasaki disease susceptibility [1]. Second, we found that the underlying genetic background conferred by CCL3L1 copy number influenced the HIV-1 acquisition- and disease-influencing effects of CCR5 haplotypes [1]. Therefore, we tested whether the KD-influencing effects associated with CCR5 haplotypes were modified by CCL3L1 gene dosage. In this analysis, we stratified patients with KD into 3 groups: those who possessed CCL3L1 copy numbers that were <2, 2, or >2; 2 was the median copy number for the entire cohort. The TDT analyses for each CCL3L1 gene dose stratum showed that the KD-influencing effects of the CCR5Δ32 allele and CCR5 haplotypes were most evident in the context of certain CCL3L1 gene dose strata (table 1, models 3 and 4). Specifically, the effects of CCR5-HHG*2 and -HHA were more significant in individuals who also possessed median or high CCL3L1 copy numbers, respectively. Thus, individuals who possessed both CCR5-HHG*2 and 2 copies of CCL3L1 had a nearly 80% lower risk of developing KD (table 1, models 3 and 4). An association between KD susceptibility and possession of HHG*1 was not detected in this stratified analysis, suggesting that the effect of HHG*1 was distributed across the different CCL3L1 gene dose strata (table 1, models 3 and 4).

Recently, Cordell et al. proposed a unified framework for genetic association testing, based on conditional logistic regression analysis of cases and matched pseudocontrols derived from the genotypes of the patients and parents [24]. This method allows nuclear family data to be analyzed in a manner very similar to that used for case/control data, but by use of conditional logistic regression models. This approach complements the TDT and E-TDT analyses, and the findings further underscore the influence of CCR5 haplotypes on KD susceptibility, as well as the notion that these effects are most evident within the context of a specific genetic background that is dependent on CCL3L1 copy number (table 2).

DISCUSSION

Taken together, these TDT analyses demonstrate that genetic variation in CCR5 plays an influential role in KD susceptibility and may influence coronary-artery outcome in affected children. We noted striking parallels between CCR5 genotype and susceptibility and outcome in both KD and HIV-1/AIDS: (1) CCR5-HHA is the ancestral CCR5 haplotype [39] and is associated with resistance to KD as well as a reduced risk of progressing rapidly to AIDS in specific populations [19]; (2) CCR5-HHG*2, the CCR5Δ32–carrying allele, is associated with reduced KD susceptibility and protection from coronary-artery aneurysms, as well as a reduced risk of acquiring HIV-1 and progressing rapidly to AIDS; (3) the CCR5-HHG*1 haplotype is associated with increased KD susceptibility, as well as an increased rate of progression to AIDS [42]; and (4) the HIV-1/AIDS- and KD-influencing effects associated with CCR5 haplotypes are influenced by CCL3L1 gene dose [1].

In summary, we present, to our knowledge, both the largest and the first family-based association study of any systemic vasculitis. The inverse relationship between the distribution of the

Table 2. Results of the case/pseudocontrol analysis (n = 162 matched sets; n = 648 total).

<table>
<thead>
<tr>
<th>Model, CCR5 genotype/haplotype, CCL3L1 gene dose</th>
<th>OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>CCR5Δ32</td>
<td>0.69 (0.34–1.38)</td>
</tr>
<tr>
<td>Model 2</td>
<td>CCR5Δ32, &lt;2 CCL3L1 copies</td>
<td>0.67 (0.19–2.36)</td>
</tr>
<tr>
<td></td>
<td>CCR5Δ32, 2 CCL3L1 copies</td>
<td>0.27 (0.07–0.99)</td>
</tr>
<tr>
<td></td>
<td>CCR5Δ32, &gt;2 CCL3L1 copies</td>
<td>2.14 (0.55–8.36)</td>
</tr>
<tr>
<td>Model 3</td>
<td>HHA</td>
<td>0.46 (0.24–0.89)</td>
</tr>
<tr>
<td></td>
<td>HHG*1</td>
<td>3.31 (1.22–8.97)</td>
</tr>
<tr>
<td>Model 4</td>
<td>HHG*2, 2 CCL3L1 copies</td>
<td>0.27 (0.07–0.99)</td>
</tr>
<tr>
<td></td>
<td>HHA, &gt;2 CCL3L1 copies</td>
<td>0.24 (0.08–0.70)</td>
</tr>
</tbody>
</table>

NOTE. Models 1 and 3, conditional logistic regression analysis association between possession of the CCR5Δ32 allele or haplotypes and Kawasaki disease susceptibility; models 2 and 4, stepwise conditional logistic regression with an entry criterion of P = .1 for association of the CCR5Δ32 allele or haplotypes in the context of different CCL3L1 gene dose strata. For models 3 and 4, only significant associations are shown. All analyses were repeated after resolving the parent of origin of the alleles, and identical results were observed. CI, confidence interval; OR, odds ratio.
CCR5-Δ32 allele and KD incidence, the infectious disease hypothesis of KD, and the concordance in genetic determinants of HIV-1 and KD resistance collectively raise the speculative but intriguing possibility that CCR5 and its ligands, such as CCL3L1, play a critical role in some aspect of the host-pathogen interaction that triggers KD. These findings will need to be validated in independent cohorts. Of broader significance, an important implication of this study is that the search for genetic loci underlying infectious diseases is likely to be very complex and must take into account the influence of gene modifier effects, such as those observed for CCL3L1, on the phenotypic effects associated with CCR5 haplotypes.

Acknowledgments

The authors wish to thank our study nurses, Annette Baker and Jennifer Foley, as well as the following University of California at San Diego medical students, for their assistance in collecting DNA samples: David Bronstein, Jennie Buchanan, and Marina Dergun.

References


