A common dominant TLR5 stop codon polymorphism abolishes flagellin signaling and is associated with susceptibility to legionnaires' disease


Published in:
Journal of Experimental Medicine

DOI:
10.1084/jem.20031220

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
A Common Dominant TLR5 Stop Codon Polymorphism Abolishes Flagellin Signaling and Is Associated with Susceptibility to Legionnaires’ Disease

Thomas R. Hawn, Annelies Verbon, Kamilla D. Lettinga, Lue Ping Zhao, Shuying Sue Li, Richard J. Law, Shawn J. Skerrett, Bruce Beutler, Lea Schroeder, Alex Nachman, Adrian Ozinsky, Kelly D. Smith, and Alan Aderem

1 University of Washington Medical Center, Seattle, WA 98195
2 Institute for Systems Biology, Seattle, WA 98103
3 Department of Infectious Diseases, Academic Medical Center, 1105 AZ Amsterdam, Netherlands
4 Division of Medical Microbiology, Academic Hospital Maastricht, 6200 MD Maastricht, Netherlands
5 Fred Hutchinson Cancer Research Center, Seattle, WA 98109
6 Enodar BioLogic Corporation, Seattle, WA 98101
7 The Scripps Research Institute, La Jolla, CA 92037

Abstract

Although Toll-like receptors (TLRs) are critical mediators of the immune response to pathogens, the influence of polymorphisms in this gene family on human susceptibility to infection is poorly understood. We demonstrated recently that TLR5 recognizes flagellin, a potent inflammatory stimulus present in the flagellar structure of many bacteria. Here, we show that a common stop codon polymorphism in the ligand-binding domain of TLR5 (TLR5392STOP) is unable to mediate flagellin signaling, acts in a dominant fashion, and is associated with susceptibility to pneumonia caused by Legionella pneumophila, a flagellated bacterium. We also show that flagellin is a principal stimulant of proinflammatory cytokine production in lung epithelial cells. Together, these observations suggest that TLR5392STOP increases human susceptibility to infection through an unusual dominant mechanism that compromises TLR5’s essential role as a regulator of the lung epithelial innate immune response.

Key words: inflammation • immunity • genetic predisposition to disease • genetic markers • bacterial infections

Introduction

Pneumonia is the sixth leading cause of death in the United States and the most common cause of death due to infectious disease (1). Despite the seriousness of pneumonia, very little is known about genetic factors that predispose individuals to acquire lower respiratory tract infections. Recent advances in innate immunity and genetics are opening avenues to examine the immunologic and genetic factors that predispose individuals to acquire common infections such as pneumonia (2). The innate immune system enables the host to differentiate itself from invading microbes, discriminate among pathogens, and initiate a cascade of inflammatory molecules that influence formation of the acquired immune response as well as host survival. Toll-like receptors (TLRs) constitute a family of transmembrane proteins that differentially recognize pathogen-associated molecular patterns (PAMPs) through an extracellular domain and initiate inflammatory signaling pathways through an intracellular domain (3–6). Due to the central role of TLRs in the innate immune response, genetic variation in this gene family is predicted to alter susceptibility to infections in a PAMP-specific manner based on the ligand recognition of the individual TLR. Although initial papers have supported this hypothesis for TLR4, the
degree to which other TLRs affect susceptibility to infections is unknown (7–11). We demonstrated recently that TLR5 recognizes the flagellin protein, a potent inflammatory stimulus present in the flagellar structure of many bacteria (12, 13). Herein, we provide the first evidence that TLR5 influences human susceptibility to infection as well as describe an important role for TLR5 in the innate immune response of the lung epithelium.

Materials and Methods

Materials. RPMI 1640 medium, l-glutamine, penicillin, and streptomycin were obtained from Life Technologies, DMEM was obtained from JRH Biosciences, and Ham’s F12 was obtained from BioWhittaker. Ultrapure LPS was from Salmonella minnesota R595 (List Biological Labs, Inc.). Flagellin was purified from Salmonella typhimurium as described previously from strain TH4778, which is fisB/fisC (a gift from K. Hughes, University of Washington, Seattle, WA; references 12, 14).

Cells, Mice, and Bacteria. Chinese hamster ovary (CHO) cells (CHO-K1), A549 cells, HEK 293 cells, and Calu-3 cells were grown according to American Type Culture Collection recommendations. PBMCs were isolated from 50 ml of blood from individuals with a Ficoll gradient separation. MyD88−/− and TLR4−/− mice (a gift from S. Akira, Research Institute for Microbial Diseases, Osaka University, Japan) were derived on a 129Sv × C57Bl/6 background, backcrossed for six generations to C57Bl/6, and used for experiments along with littermate controls (15, 16). Bone marrow was harvested from these mice and grown in RPMI 1640 supplemented with 10% heat-inactivated FCS (Hyclone Laboratories), 20% L-cell-conditioned media, 100 U/ml penicillin, 100 μg/ml streptomycin, and 2 mM l-glutamine. Bone marrow–derived macrophages were used after 4–10 d grown according to the manufacturer’s Institutional Animal Care and Use Committees. Bone marrow–derived macrophages were used after 4–10 d grown according to American Type Culture Collection recommendations. Controls were drawn from the exhibitioners who worked at the flower show and were at high risk for exposure to Legionella pneumophila. 1,616 controls were contacted by letter to be in the analysis. The first 508 who completed the questionnaire had blood drawn for genetic analysis.

Statistics. Univariate analysis was performed for categorical variables with chi-square test, and for continuous variables using a Student’s t test. Single nucleotide polymorphism (SNP) data were analyzed using a recently developed methodology based on an estimating equation technique, the details of which have been described previously (22, 23). The basic idea is to estimate the distribution of haplotypes from SNP genotype data, to correlate all common haplotypes with LD phenotype with adjustment for demographic variables, and to assess interactions of haplotypes with exposure factors. The penetrance of paired haplotypes (H, h) and exposure (E) to LD outcome (D = 1) is quantified via a logistic regression model:

$$Pr(D = 1 | H,h,E) = \frac{1}{1 + \exp(-\alpha - \beta (H,h) - \gamma E - I(H,h,E,\delta))},$$

where unknown parameter $\beta$ in the function $G(H,h,\beta)$ quantifies disease associations with haplotypes, $\gamma$ quantifies the disease association with exposure, and $\delta$ in $I(H,h,\delta)$ quantifies the disease associations with the interaction of haplotypes and exposure. Estimating regression parameters ($\beta, \delta$) is of primary interest because their relevant statistics can be used to make scientific inferences. The key challenge to the estimation is that the haplotype structure for many individuals is not observed (i.e., phases of SNP alleles are not observed).

Molecular Biology. SNP discovery and genotyping was performed by PCR amplification of TLR5 from genomic DNA followed by sequencing. To generate templates for sequencing, the coding region of TLR5 was amplified from 25 ng of genomic DNA in two fragments with AmpliTaq DNA polymerase (Applied Biosystems). Primers for fragment A were as follows: T5-5, forward, 5′-GTATCTGCACACCGCACTCTGTAAGG-3′; and T5-20, reverse, 5′-GAGATCTTGAATGCTACACCCCG-3′. Primers for fragment B were as follows: T5-3, forward, 5′-GATACGGATATCTTCTTGAGTGGC-3′; and T5-21, reverse, 5′-TCGCAAACCTCTAAAGGAGAGCTG-3′. For genotyping by Ddel restriction digestion, a 276-bp PCR product was first generated with primers T5-11, forward, 5′-GTTAGCCAATGATTGGTGC-3′; and T5-22, reverse, 5′-CTGATTCAGTTTAGCTGGTGC-3′; and T5-23, digested with Ddel (New England Biolabs, Inc.), and separated on a 1.8% agarose gel. Functional studies, the coding region of TLR5 was amplified from genomic DNA by PCR using Pfu Turbo polymerase (Stratagene) and cloned into the pEF6/V5-His-TOPO vector (Invitrogen). PCR primers for the full length gene included T5-10, forward, 5′-ATGGGACGCCACCTGACCTCCT-3′; and T5-30, reverse, 5′-GGAGATGTGGTGCTCAGTTTGAACGG-3′. PCR primers for TLR5STOP were T5-10, forward, and T5-31, reverse, 5′-GGAGATGTCGCCACGTGTTTCCAGG-3′. To obtain the different polymorphic variants of TLR5, genomic DNA was selected from individuals with the desired genotype. The cloned variants of TLR5 were sequenced for verification by staining with Big Dye Terminator v3.0 and analyzed on a capillary sequencer (model ABI PRISM 3700; Applied Biosystems). Sequence was aligned and analyzed with the programs PHRED/PHRAP and CONSED (24).

Transfections. CHO cells were transfected as described previously (12). HEK293 cells were transfected as described previously in a 96-well plate format with calcium phosphate, using 5 × 10⁴ cells per well with 8 ng of TLR5, 8 ng of pRL-TK, and 80 ng of endothelial leukocyte adhesion molecule (ELAM)–luciferase per

1564 TLR5 and Legionella pneumophila
well (25). For some experiments, HEK293 cells were alternatively transfected with Polyfect (QIAGEN) according to the manufacturer’s instructions with 5 × 10^4 cells per well in a 96-well plate with 175 ng TLR5, 8 ng pRL-TK, and 80 ng of ELAM-luciferase per well. Stimulated cells were processed as described previously (12).

**Protein Analysis.** For the immunoblot, cells were lysed with 1% Triton X-100/150 mM NaCl/10 mM Hepes, pH 8.0, or passive lysis buffer (Promega) with a cocktail of protease inhibitors (Sigma-Aldrich). An immunoblot with proteins from cell lysates was probed with an anti-V5 epitope antibody (Serotec) followed by HRP-conjugated rabbit anti–mouse IgG (Zymed Laboratories) and developed with chemiluminescent reagents (Roche). Whole blood cytokine assays were prepared by diluting venous blood 1:5 with RPMI 1640, plating in a 96-well dish, stimulating for 18 h, and harvesting supernatants. For Calu-3 and A549 assays, cells were plated at a density of 2 × 10^4 per well in a 96-well dish, stimulated for 18 h, and supernatants were harvested. Cytokine levels were determined with a sandwich ELISA technique (Duoset; R&D Systems).

**Results**

**TLR5 SNP Discovery.** To discover novel SNPs with disease association, we PCR amplified and sequenced the TLR5 coding region in 40 healthy individuals. TLR5 is a type I transmembrane protein with a 642–amino acid leucine-rich extracellular domain, an 18-residue transmembrane domain, and a 198-residue cytoplasmic Toll homology signaling domain. Four SNPs were detected in TLR5. Most dramatic was a cytosine–thymidine transition at base pair 1174 that changed the arginine at amino acid 392 to a stop codon (Fig. 1A, TLR5^{392STOP}). This change is predicted to prematurely truncate TLR5 in the extracellular domain and cause the loss of the transmembrane domain and the entire signaling cytoplasmic tail. We verified this polymorphism with a DdeI restriction enzyme digestion, which selectively cut a 276-bp PCR product encompassing the 1174T variant (Fig. 1B). Two additional non-synonymous SNPs (A1775G [amino acid N592S] and T1846C [F616L]) alter residues in the ectodomain, whereas a fourth synonymous SNP (A2523G [K841K]) was in the cytoplasmic tail. The frequency of subjects who possessed the stop codon variant as heterozygotes (1174 CT) and homozygotes (1174 TT) was 7.5 and 0%, respectively. The serine variant at amino acid 392 and the leucine variant at 616 were present as heterozygote frequencies of 35.0 (1775 AG and 45.0% (1846TC), and homozygote frequencies of 5.0 (1775 GG) and 15.0% (1846 CC), respectively. The synonymous SNP at amino acid 841 (A2523G) was present as a heterozygote frequency of 20.0% (2523 AG) and a homoygote frequency of 0% (2523 GG). Except for minor deviations, the observed allelic frequencies of these four SNPs were consistent with expected allelic frequencies under the Hardy–Weinberg equilibrium.

cDNAs encoding the TLR5 polymorphisms were amplified by PCR and cloned into the pEF6-TOPO expression vector in frame with a COOH-terminal V5-epitope tag. The full-length constructs were labeled TLR5^{RNF}, TLR5^{RSF}, and TLR5^{RNL}, using the single letter designations of the amino acids at the three polymorphic amino acid positions (392, 592, and 616). When assessed by Western blot, TLR5^{RNF}, TLR5^{RNL}, and TLR5^{RSF} protein expression levels were similar in transfected CHO cells (Fig. 1C). TLR5^{392STOP} was expressed at higher levels in comparison to the full-length versions and was detected as a doublet band, which may be secondary to proteolysis or glycosylation. Next, we examined whether TLR5^{392STOP} was able to mediate flagellin signaling. We transfected CHO cells with TLR5 variants along with an NF-κB luciferase reporter plasmid (ELAM) and stimulated with purified *S. typhimurium* flagellin for 4 h (12). Flagellin induced NF-κB activity in cells expressing the full length TLR5 (Fig. 1D). In contrast, cells expressing TLR5^{392STOP} were unable to respond to flagellin. Identical results were ob-

![Figure 1](https://www.jem.org)
SNPs in TLR5 Are Associated with Susceptibility to LD. Next, we examined whether these TLR5 SNPs were associated with susceptibility to human infection in a PAMP-specific manner. L. pneumophila, first described in 1976 as the agent of LD, is a flagellated gram-negative bacteria that causes anywhere from 1 to 30% of cases of community-associated with susceptibility to human infection in a PAMP-specific manner.

Next, we determined the genotype and allele frequencies of individuals for the three nonsynonymous TLR5 SNPs (Table II and III). There was no significant departure from the Hardy–Weinberg equilibrium of the observed and expected frequencies of cases or controls for these SNPs. The 1174T (stop codon variant) and 1775G alleles were both associated with susceptibility to LD. The allele frequency of 1774T was 8.3% in the cases compared to 5.0% in the entire control group and 4.5% in the control group. In addition, we found that several other strains of L. pneumophila stimulated TLR5 activity, including the Bovenkarspel strain, which caused a recent epidemic of LD at a flower show in the Netherlands in 1999 (19). We also stimulated bone marrow–derived macrophages from wild-type, TLR4−/− mice with heat-killed L. pneumophila. Although TNF-α and IL-6 levels were abolished in MyD88−/− macrophages, there was no decrease in the TLR4−/− cells in comparison to wild-type controls (Fig. 2, B and C). In contrast, as a control, cytokine production was abolished in both TLR4−/− and MyD88−/− macrophages stimulated with S. typhimurium LPS. These results demonstrate that TLR4 is not involved in recognition of L. pneumophila by murine macrophages.

To determine whether TLR5 is associated with susceptibility to LD, we examined genotypes in a case-control cohort from the Bovenkarspel epidemic (19–21). Cases included individuals with radiologically confirmed pneumonia occurring during the epidemic time period around the West Fries Flower Show in the town of Bovenkarspel in the Netherlands. Controls were drawn from a pool of exhibitors who were likely to be exposed to the contaminated water product that caused the epidemic. All of the cases and controls were from the Netherlands and >95% of both groups were Dutch. To provide an additional control for population admixture, 89 of the 508 controls were matched to cases for their place of residence ± 25 km, as well as age and sex. The baseline characteristics of the case and control groups are shown in Table I. The mean age and the prevalence of diabetes were statistically different between cases and controls. In addition, smoking appeared to be marginally different between cases and controls. To formally assess the associations of age, sex, diabetes, and smoking status, our null hypothesis is that these variables do not associate with LD when comparing cases and controls (odds ratio [OR] = 1), in contrast to an alternative hypothesis of OR > 1 or OR < 1. A two-sided test was used for the purpose of evaluating statistical significance. Next, we adjusted the analysis by logistic regression with covariates for age (stratified into <40, 40–49, 50–60, and >60), gender, diabetes status, and smoking status. In the adjusted analysis, smoking status (OR = 2.84; 95% confidence interval [CI] = 1.55–5.21; P < 0.05) and age (OR in reference to 40–49 age group: OR = 0.071 (age, 0–39); CI = 0.01–0.57; P < 0.05; OR = 2.30 (age, 50–60); CI = 1.0–5.29; P = 0.05; and OR = 21.80 (age, >60); CI = 9.26–51.28; P < 0.05) had a significant association with LD. In contrast, gender (OR = 1.09; CI = 0.62–1.91) and diabetes status were not associated with LD (OR = 1.43; CI = 0.51–4.03).

Next, we determined the genotype and allele frequencies of individuals for the three nonsynonymous TLR5 SNPs (Tables II and III). There was no significant departure from the Hardy–Weinberg equilibrium of the observed and expected frequencies of cases or controls for these SNPs. The 1174T (stop codon variant) and 1775G alleles were both associated with susceptibility to LD. The allele frequency of 1174T was 8.3% in the cases compared with 5.0% in the entire control group and 4.5% in the...
matched control group. For the purpose of assessing genetic association, our formal null hypothesis is that individual alleles or their haplotypes do not associate with case-control status (i.e., OR = 1). The alternative hypothesis is that there is some association, with OR > 1 or OR < 1. Two-sided testing was used again for evaluating statistical significance. The OR for the association of allele 1174T with developing LD was 1.75 (95% CI = 1.00–3.05; P = 0.05) when compared with the entire control group. The association of 1174T with LD is even more striking because this stop codon variant was almost exclusively present as a heterozygote genotype, as there were only two homozygote individuals in the entire cohort (1174TT). The presence of both homozygote stop codon individuals in the

### Table I. Baseline Clinical Characteristics of Cases and Controls

<table>
<thead>
<tr>
<th></th>
<th>Cases n = 108</th>
<th>All controls n = 508</th>
<th>Matched controls n = 89</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>65 (60.2)</td>
<td>261 (52.4)</td>
<td>53 (60.2)</td>
</tr>
<tr>
<td>Smoking</td>
<td>47 (49.0)</td>
<td>172 (40.6)</td>
<td>32 (42.7)</td>
</tr>
<tr>
<td>History of</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COPD</td>
<td>8 (7.8)</td>
<td>39 (7.8)</td>
<td>4 (4.5)</td>
</tr>
<tr>
<td>DM</td>
<td>10 (10.1)</td>
<td>14 (2.8)</td>
<td>5 (5.6)</td>
</tr>
<tr>
<td>Malignancy</td>
<td>4 (4.0)</td>
<td>12 (2.5)</td>
<td>2 (2.3)</td>
</tr>
<tr>
<td>Rheum</td>
<td>2 (2.0)</td>
<td>13 (2.6)</td>
<td>4 (4.5)</td>
</tr>
<tr>
<td>Age</td>
<td>63.8 ± 10.8</td>
<td>44.9 ± 13.6</td>
<td>54.6 ± 11.2</td>
</tr>
</tbody>
</table>

Prevalence of baseline demographic and clinical information is indicated. Statistics are reported for each control population in reference to the cases. The odds ratios (OR) with 95% confidence intervals is indicated (CI). Chi-square test used to calculate two-sided significance values. Not all variables were known in all patients.

*For subgroups with small sample numbers, a Fisher’s exact test was used.

Mean ages compared with an independent sample t test.

%, number of patients tested for that variable; Rheum, rheumatologic disorder; COPD, chronic obstructive pulmonary disease; DM, diabetes mellitus.

### Table II. TLR5 SNP Genotype Frequencies in Cases and Controls

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Amino Acid</th>
<th>Cases n = 108</th>
<th>All controls n = 508</th>
<th>Matched controls n = 89</th>
<th>Total n = 616</th>
</tr>
</thead>
<tbody>
<tr>
<td>1174 CC</td>
<td>392RR</td>
<td>90 (0.833)</td>
<td>457 (0.905)</td>
<td>82 (0.921)</td>
<td>547 (0.892)</td>
</tr>
<tr>
<td>1174 CT</td>
<td>392R*</td>
<td>18 (0.167)</td>
<td>46 (0.091)</td>
<td>6 (0.067)</td>
<td>64 (0.104)</td>
</tr>
<tr>
<td>1174 TT</td>
<td>392*</td>
<td>0 (0)</td>
<td>2 (0.004)</td>
<td>1 (0.011)</td>
<td>2 (0.003)</td>
</tr>
<tr>
<td>1775 AA</td>
<td>592NN</td>
<td>72 (0.692)</td>
<td>383 (0.777)</td>
<td>71 (0.798)</td>
<td>455 (0.762)</td>
</tr>
<tr>
<td>1775 AG</td>
<td>592NS</td>
<td>28 (0.269)</td>
<td>106 (0.215)</td>
<td>18 (0.202)</td>
<td>134 (0.224)</td>
</tr>
<tr>
<td>1775 GG</td>
<td>592SS</td>
<td>4 (0.038)</td>
<td>4 (0.008)</td>
<td>0 (0)</td>
<td>8 (0.010)</td>
</tr>
<tr>
<td>1846 TT</td>
<td>616FF</td>
<td>28 (0.277)</td>
<td>126 (0.260)</td>
<td>21 (0.239)</td>
<td>154 (0.263)</td>
</tr>
<tr>
<td>1846 TC</td>
<td>616FL</td>
<td>49 (0.485)</td>
<td>252 (0.520)</td>
<td>53 (0.602)</td>
<td>301 (0.514)</td>
</tr>
<tr>
<td>1846 CC</td>
<td>616LL</td>
<td>24 (0.238)</td>
<td>107 (0.221)</td>
<td>14 (0.159)</td>
<td>131 (0.224)</td>
</tr>
</tbody>
</table>

TLR5 genotypes were determined in cases and controls. *, stop codon.
control group is likely due to the larger control sample size \((n = 508)\) as compared with the cases \((n = 108)\). In addition, although the controls were selected for an increased risk of contact with \(L. \text{ pneumophila}\), the actual exposure dose of each individual is not known. The OR for the comparison with the matched control group showed a similar trend, but was not significant due to the small sample size \((\text{OR } 1.93; 95\% \text{ CI } 0.74–5.03)\). The OR for the association of allele 1775G with LD was 1.60 \((95\% \text{ CI } 1.04–2.45; P = 0.03)\) compared with the entire control group and 1.83 \((95\% \text{ CI } 1.02–3.28; P = 0.04)\) for the matched cohort. There was no association of SNP T1846C with LD. To determine if coinheritance of 1174T and 1775G further increased the risk of pneumonia, we examined the association of TLR5 haplotypes with susceptibility to LD \((22, 23)\). However, no haplotypes containing both 1174T and 1775G were detected in either cases or controls (Table III, labeled 11). In the analysis by haplotypes, the association of 1174T-1775A (haplotype 10; \(\text{OR } 1.90; P = 0.03\)) and 1174C-1775G (haplotype 01; \(\text{OR } 1.68; P = 0.02\)) with susceptibility to LD was also present \((23)\). Next, we considered whether any variables might be confounders in the analysis, an unlikely possibility given that there is no overt biologic relationship of TLR5 to any of the variables in Table I. We repeated the analysis with an adjustment for sex and age (stratified at \(<40, 40–49, 50–60, \text{ and } >60\)) and found a similar trend in the association of TLR5 with LD (SNP 1174T; \(\text{OR } 1.56 \text{ [95\% CI } 0.82, 2.98\]); SNP 1775G, \(\text{OR } 1.71 \text{ [95\% CI } 1.07, 2.72\]}. In addition, there was no statistical association between any of the TLR5 SNPs and diabetes (unpublished data).

Because smoking has previously been shown to be significantly associated with LD, and this association was confirmed here, we investigated whether there was an interaction between smoking and TLR5 SNPs (Table IV; Table III. **TLR5 SNP Allele and Haplotype Frequencies in Cases and Controls**

<table>
<thead>
<tr>
<th>SNP</th>
<th>Cases (n = 108)</th>
<th>All controls (n = 508)</th>
<th>Matched controls (n = 89)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>BP</td>
<td>Allele freq</td>
<td>freq</td>
</tr>
<tr>
<td>1.</td>
<td>1174C</td>
<td>0</td>
<td>0.917</td>
</tr>
<tr>
<td></td>
<td>1174T</td>
<td>1</td>
<td>0.083</td>
</tr>
<tr>
<td>2.</td>
<td>1775A</td>
<td>0</td>
<td>0.827</td>
</tr>
<tr>
<td></td>
<td>1775G</td>
<td>1</td>
<td>0.173</td>
</tr>
<tr>
<td>3.</td>
<td>1846T</td>
<td>0</td>
<td>0.520</td>
</tr>
<tr>
<td></td>
<td>1846C</td>
<td>1</td>
<td>0.480</td>
</tr>
</tbody>
</table>

**Haplotype 1174–1775**

<table>
<thead>
<tr>
<th>No.</th>
<th>BP</th>
<th>Allele</th>
<th>freq</th>
<th>freq</th>
<th>OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>00</td>
<td>0.744</td>
<td>0.835</td>
<td>1</td>
<td>0.861</td>
<td>1</td>
</tr>
<tr>
<td>CG</td>
<td>01</td>
<td>0.173</td>
<td>0.115</td>
<td>1.68  (1.10, 2.59)</td>
<td>0.02</td>
<td>0.094</td>
</tr>
<tr>
<td>TA</td>
<td>10</td>
<td>0.083</td>
<td>0.049</td>
<td>1.90  (1.07, 3.35)</td>
<td>0.03</td>
<td>0.038</td>
</tr>
<tr>
<td>TG</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^0\) represents the common allele at each locus; 1 is the uncommon allele. Haplotype alleles for SNP1 (1174) and SNP2 (1775) are listed in order. Odds ratio (OR) and 95% confidence intervals represent a comparison of the cases with the two respective control groups with an unadjusted analysis.

**Table IV. Multivariate Analysis of Association of TLR5 SNPs with Legionnaires’ Disease in Smokers and Nonsmokers**

<table>
<thead>
<tr>
<th>SNP</th>
<th>Smokers</th>
<th>Nonsmokers</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>BP</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>1.</td>
<td>1174C</td>
<td>0 1</td>
</tr>
<tr>
<td></td>
<td>1174T</td>
<td>1 0.72 (0.18, 2.86)</td>
</tr>
<tr>
<td>2.</td>
<td>1775A</td>
<td>0 1</td>
</tr>
<tr>
<td></td>
<td>1775G</td>
<td>1 1.33 (0.54, 3.26)</td>
</tr>
</tbody>
</table>

**Haplotype 1174–1775**

<table>
<thead>
<tr>
<th>No.</th>
<th>BP</th>
<th>Allele</th>
<th>OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>00</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>CG</td>
<td>01</td>
<td>1.33 (0.55, 3.25)</td>
<td>0.53</td>
<td>1.99 (1.04, 3.80)</td>
</tr>
<tr>
<td>TA</td>
<td>10</td>
<td>0.76 (0.19, 3.05)</td>
<td>0.70</td>
<td>2.43 (1.00, 5.89)</td>
</tr>
</tbody>
</table>

Comparison is between cases \((n = 108)\) and the entire control group \((n = 508)\), excluding those with no information on smoking status. Odds ratio (OR) and 95% confidence intervals were calculated by a multivariate logistic regression analysis. All analyses are adjusted for gender and age (stratified at \(<40, 40–49, 50–60, \text{ and } >60\) ), and then ORs for smokers and nonsmokers were calculated separately through an interaction term.
The analysis revealed that the association of 1174T-1775A and 1174C-1775G haplotypes with LD was only found in the nonsmokers (ORs = 2.43 [CI = 1.00, 5.89; P = 0.05] and 1.99 [CI = 1.04, 3.80; P = 0.04], respectively), whereas the association became marginal among smokers. This result implies that smoking may be a dominant risk factor for LD and, thus, may mask genetic associations. Together, these results suggest that TLR5 SNPs are associated with LD in nonsmokers.

**TLR5 Is the Predominant Mediator of IL-8 Expression in Lung Epithelia.** These genetic findings raise the question of why a deficiency in TLR5 would be deleterious to the host because there are other PAMP receptors, including TLR4, which could presumably still recognize *L. pneumophila* and mediate an inflammatory response. To understand TLR5 regulation of signaling in pulmonary epithelia, we stimulated A549 lung cell lines and measured culture supernatant IL-8 levels. As seen in previous analyses, LPS was unable to stimulate IL-8 production, even at doses of 1,000 ng ml\(^{-1}\) (Fig. 3 A, not depicted) (31). In contrast, flagellin potently stimulated IL-8 in a dose-dependent fashion, starting at concentrations of 1 ng ml\(^{-1}\). Furthermore, although heat-killed wild-type *L. pneumophila* induced IL-8 production, a flagellin mutant strain (FlaA\(^{-}\)) was nonstimulatory (Fig. 3 B). Similar results were obtained for IL-6 and IL-8 production in Calu-3 lung epithelial cells (Fig. 3, C and D; and 3 B). In contrast to epithelial cells, both wild-type and FlaA\(^{-}\) strains were similarly able to induce TNF-\(\alpha\) production in a whole blood cytokine assay (Fig. 3 E). This control assay indicates that the FlaA\(^{-}\) mutant did not have a generalized loss of PAMPs that compromised its overall stimulatory capacity. Overall, these results suggest a dominant role for flagellin and TLR5 in regulating the pulmonary epithelial innate immune response to *L. pneumophila*.

**TLR5\(^{392STOP}\) Has a Dominant Effect.** Because individuals who are heterozygous for the TLR5 stop codon are more susceptible to LD, we examined whether TLR5\(^{392STOP}\) acts in a dominant fashion. We stimulated PBMCs from individuals with flagellin or LPS and measured IL-6 cytokine production by ELISA. Heterozygous stop codon individuals (392R\(^*\)) had significantly decreased IL-6 production in comparison to those with the wild-type arginine at amino acid 392 (Fig. 3 F, 392RR; P < 0.05). In comparison, there was no difference in LPS-induced IL-6 production (Fig. 3 G). This result indicates that TLR5\(^{392STOP}\) exerts a dominant effect and suggests that heterozygous TLR5 stop codon individuals are more susceptible to LD due to impaired production of proinflammatory cytokines. We reanalyzed the genetic association of TLR5 C1174T with susceptibility to LD with the assumption that TLR5\(^{392STOP}\) exhibits a dominant effect. The genotype frequency of 1174TC and TT (combined) was 16.7% in the cases compared with 9.5% in the entire control group and 7.9% in the matched control group. The OR for the association of genotype 1174T (C>T) with developing LD was 1.90 (95% CI = 1.06–3.42; P = 0.03) when compared with the entire group and 2.34 (95% CI = 0.93–5.90; P = 0.07) in comparison to the matched control group. This analysis confirmed and further strengthened the association of TLR5\(^{392STOP}\) with susceptibility to LD.

**Discussion**

In this paper, we show for the first time that a common TLR5 stop codon polymorphism acts in a dominant fashion and is associated with susceptibility to infection with a flagellated organism in humans. We further show that TLR5–flagellin interactions control IL-8 production in lung epithelial cell lines stimulated by *L. pneumophila*.

Although these results suggest an association of TLR5 with LD, we cannot exclude that these SNPs are in linkage
disease susceptibility to bacterial pneumonia has only been linked to a small number of genes, including mannose binding lectin and FcRIIa (33, 34). Our analytic approach highlights the importance of separating the effects of genetic and environmental variables on disease outcome because the genetic association of TLR5 with LD was exclusive to the nonsmoking population.

Our finding that flagellin is the major L. pneumophila stimulant of IL-8 production in lung epithelial cell lines has important implications for models of innate immunity. Previous investigators have demonstrated that the A549 lung epithelial cell line does not respond to LPS, even at doses of 1 μg ml⁻¹ (31, 35). Normal lung epithelia may down-regulate LPS recognition through TLR4 as a protective mechanism to avoid an overly exuberant inflammatory response (36). A previous paper found that Pseudomonas flagellin, but not LPS, stimulated IL-8 production in tracheal epithelial cells (37). Several recent works with intestinal and lung epithelial cells also demonstrate that flagellin is a major stimulus of proinflammatory cytokine production (38–44). Together, these results suggest that PAMP recognition in epithelial cells is more constrained than macrophages and may explain why a selective TLR5 deficiency can predispose an individual to an increased risk of pneumonia due to a flagellated organism.

If there is redundancy in the innate immune system, there may be no detriment to having a stop codon in TLR5 if an individual has a wild-type version of TLR4. First, lung epithelial cell lines recognize LPS poorly and suggest that there are cell and organ-specific differences in TLR regulation and function. Second, L. pneumophila LPS is ~1000-fold less stimulatory than S. typhimurium LPS (45–47). With such low stimulatory properties, L. pneumophila LPS may not be a major target of innate immune system recognition. Third, we have shown previously that TLR4-defective mice are not more susceptible to aerosolized infection with L. pneumophila (48). In addition, we found that TLR4 is not involved with recognition of heat-killed L. pneumophila by murine macrophages. These studies suggest that receptors other than TLR4 mediate recognition of L. pneumophila during an in vivo infection. Finally, even if TLR4 mediates L. pneumophila recognition, there may be qualitative or quantitative differences in the TLR4 induced gene expression profiles in comparison to TLR5 stimulation.

Although the importance of TLRs in regulating the innate immune response has been clearly demonstrated in animal models, the clinical significance of TLR polymorphisms in human disease susceptibility is only beginning to be elucidated. Recent analyses with TLR4 SNPs suggest a PAMP-specific influence of this gene on infectious disease susceptibility because TLR4 is associated with gram-negative bacterial infections (7, 8, 10, 11). The finding that TLR5 SNPs are associated with susceptibility to a flagellated bacterium supports the hypothesis that susceptibility profiles will be PAMP-dependent and vary between infections. The high frequency of the TLR5 stop codon mutation in apparently healthy individuals raises the question of whether there is an evolutionary advantage to having this genetic variant. Linkage papers have mapped a major susceptibility locus for systemic lupus erythematosus in humans and mice to chromosome 1q41 where TLR5 resides (49, 50). Given TLR5’s role in mediating inflammatory signaling pathways, it will be of interest to examine its possible association with an autoimmune disease such as systemic lupus erythematosus.

The dominant effect of the stop codon mutation implies that ~10% of individuals have severely impaired ability to recognize flagellated bacteria and may be more susceptible to infection. The presence of the stop codon in the extracellular ligand-binding domain, in contrast to the intracellular signaling domain, suggests that the dominant effect occurs through a unique mechanism. Because TLR5 acts as a homodimer, the stop codon variant may bind to the normal variant and inhibit its proper assembly or localization. Understanding this unusual mechanism may lead to novel insights for immunomodulatory treatment of inflammatory diseases.

We thank H.C. Boshuizen for establishing the database for controls, and A. Siegel and N. Khalid for statistical insights.

This work was supported by the National Institutes of Health grants to T.R. Hawn, K. Smith, L.P. Zhao, and A. Aderem. A Postdoctoral Physician Fellowship from the Howard Hughes Medical Institute supported T.R. Hawn.

Submitted: 22 July 2003
Accepted: 6 October 2003

References