INTRODUCTION

Background

Otitis media (OM) is the most common medical problem affecting infants and children of any age. Most children will experience one or more episodes of acute OM (AOM) in the first 2 years of life and 93% of all children will have at least one episode of AOM during childhood. Recurrent OM (ROM) occurs in 20% to 30% of the pediatric population. Chronic otitis media with effusion (COME) is defined as middle ear effusion lasting at least 3 months, which occurs in up to 10% of children, with a point prevalence of 20%. Of school-aged children who develop otitis media with effusion, 15% to 25% will develop chronic effusions lasting more than 3 months. OM can lead to serious complications, and the societal costs attributable to OM approach $5 billion annually. COME may lead to hearing loss and speech and language delays.

Despite well-known environmental risk factors, host genetic factors are increasingly implicated in susceptibility to OM. As one review of the subject stated, "There is incontrovertible evidence from epidemiological, anatomical, physiological, and immunological studies that susceptibility to recurrent acute otitis media and persistent otitis media with effusion is genetically determined." In another study, a positive family history was present in more than 50% of patients with ROM. In siblings and parents, as much as 60% to 70% of OM liability is due to genetic background. Some of the strongest evidence for a genetic predisposition comes from twin and triplet studies that demonstrated a heritability of 57% for acute and 72% for chronic ear infections. There is also a higher correlation for ROM in monozygotic twins as compared with dizygotic twins.

The specific genetic and immunologic factors underlying these findings are still unclear. Because the pathogenesis of OM is multifactorial, with contributions from immunity, inflammatory regulation, and eustachian tube function, the genetics of OM are likely...
to be complex. Thus, the identification of OM susceptibility genes could potentially allow for the development of molecular assays for stratification of OM risk in individual children. Such knowledge could impact clinical follow-up and decision making in treatment of children with OM.

The genetic factors associated with the otitis media-prone state have focused on twin and familial studies and genes identified via genome-wide association studies (GWAS). Studies of the following OM phenotypes have been published: recurrent acute otitis media (RAOM)/COME, RAOM, and COM. Gene Studies

A number of genes are implicated in OM pathogenesis. Analysis of the expression levels of single genes has been performed. For example, differences in mucin gene MUC5AC expression were noted in mucopurulent versus serous middle ear effusions. Candidate gene studies have been published on the following genes: TLR4, interleukins, tumor necrosis factor-α, F-Box Only protein 11, mucins, mannose-binding lectin 2, and surfactant protein A. Genome-wide association studies have identified the following chromosomes as putative susceptibility loci: 10q22.3, 3p25.3, 10q26.3, 17q12, and 12q13.43. Intricate immune response genes have been studied for their role in OM predisposition, yet conflicting results exist in the literature. Emonts et al. showed an association for polymorphisms in TLR4, among other innate immune response genes with ROM. Our previous genetic association analysis comparing DNA extracted from otitis-prone patients (RAOM and COME phenotype) undergoing tympanostomy tube placement to control patients undergoing nonotologic surgery did not reveal isolated single nucleotide polymorphisms (SNPs) in TLR4, TLR2, TLR9, and CD14 genes to be associated with COME. Heterogeneity in controls could have reduced the power to have detected an association.

TLR4 Animal Model

Animal models exist that help point to etiologic and genetic factors in OM, including the Jeff mouse (FBXO11 knockout) and the C3H/HeJ mouse (TLR4 deficient). The TLR4 mouse strain has a single amino acid substitution in the extracellular region of TLR4 that renders the receptor insensitive to endotoxin. Interestingly, 50% of C3H/HeJ mice develop chronic otitis media spontaneously, suggesting a key role for defective TLR signaling in OM pathogenesis. Because similar TLR4 mutations cause reduced responsiveness to endotoxin in humans, this C3H/HeJ mouse model of otitis media has significant translational potential for research concerned with the pathogenesis and treatment of COME. With this understanding of the TLR4–COM mechanism in the mouse model, we have focused on the COME phenotype to improve our chances of identifying SNPs of interest in the TLR4 gene and others.

Tag SNP Approach

Tag SNPs are representative SNPs in a region of a gene that are in high linkage disequilibrium (LD) with other SNPs, and therefore these SNPs “tag” other SNPs in the region. Using tag SNPs, it is possible to identify genetic variation without genotyping every SNP in a chromosomal region. This powerful approach allows association of SNPs that are tagged to be inferred. These tag SNPs are then analyzed for association to identify genes potentially increasing an individual’s susceptibility to disease. The correlation structure of SNPs within a gene allows for the identification of a subset of SNPs that will tag regions of the gene for association with a phenotype. These SNPs associated with the phenotype of interest are not necessarily causative, but point to a gene region of association with the phenotype of interest.

The eight genes used to select a panel of tag SNPs were chosen based on the functional evidence in the C3H/HeJ mouse for the TLR4 gene and from the literature review: TLR4, FBXO11, MUC2, MUC5AC/B, SCN1B, SMAD2, SMAD4, and SFTPD. Both TLR4 and SFTPD play important roles in the host defense against infectious microorganisms and in regulating the innate immune response to a variety of pathogen-associated molecular patterns. SFTPD encodes surfactant protein D. FBXO11, a member of the FBOX-only family, is a mediator of the transforming growth factor (TGF)-β pathway. The SMAD genes are also mediators of the TGF-β pathway and regulate cell proliferation, apoptosis, and cell differentiation. The mucin genes (MUC2, MUC5) are important in the production of mucin in OM and COME. Last, the SCN1B gene is a Na channel regulator and has been shown to have some association with the OM-prone state in children.

MATERIALS AND METHODS

Study Design

A case-control design was used for the study. Cases included children ages 18 months to 18 years undergoing tympanostomy tube placement for COME. For study purposes, COME was defined as a middle ear effusion persisting for at least 3 months. These criteria are in line with the practice guidelines of the American Academy of Otolaryngology, American Academy of Pediatrics, and American Academy of Family Practice. Diagnosis of COME was based on patient history, otomicroscopic examination, and tympanometry: A control group met the following inclusion criteria: children age 18 months to 18 years of age seen in the pediatric otolaryngology clinic for a nonotologic, nonadenotonsillar condition, without a history of any other chronic infection (chronic sinusitis) or other inflammatory indication (recurrent parotitis, neck abscess, chronic infectious state).

Exclusion criteria for both cases and controls included patients with Down syndrome, cleft lip and/or cleft palate, velocardiofacial syndrome, 22q deletion syndrome), genetic disorders with known otologic sequelae, prior otologic surgery, cholesteatoma or history of cholesteatoma, known immunodeficiency, and/or known inflammatory disease. Exclusion criteria for controls included current AOM or COME, history of COME, history of ROM, or tympanostomy tube insertion. An additional exclusion criterion was non-Caucasian ethnicity due to the varying incidence of SNPs in different ethnic populations.
DNA Collection and Extraction

A sample of the patient’s saliva was obtained in clinical using the Oragene DNA Self-Collection Kit (DNA Genotek Inc., Kanata, Ontario, Canada), in which the patient provides 1 to 2 mL of saliva into the kit. The kit is then closed, and the saliva is mixed with a preserving fluid to stabilize the DNA at room temperature. Genomic DNA was extracted from the saliva using the Oragene protocol (http://www.dnagenotek.com/ROW/pdf/PD-PR-006.pdf). The samples were then treated with RNase to digest contaminating RNA and the DNA precipitated with ethanol. DNA was assessed for purity and concentration using absorbance (optical density (OD)) via Nanodrop, aiming for an A260/A280 ratio of >1.8.

Table I. Phenotype Data

<table>
<thead>
<tr>
<th></th>
<th>Case, n = 100</th>
<th>Control, n = 79</th>
<th>Z score</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y (SD)</td>
<td>6.51 (3.47)</td>
<td>6.83 (4.20)</td>
<td>.585</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>57</td>
<td>42</td>
<td>.512</td>
<td>.6083</td>
</tr>
<tr>
<td>Female</td>
<td>43</td>
<td>37</td>
<td>.512</td>
<td>.6083</td>
</tr>
<tr>
<td>OM family history</td>
<td>65/93</td>
<td>33/79</td>
<td>3.712</td>
<td>.0002</td>
</tr>
<tr>
<td>Daycare</td>
<td>42/98</td>
<td>23/79</td>
<td>1.886</td>
<td>.0594</td>
</tr>
<tr>
<td>Smoking</td>
<td>16/99</td>
<td>11/79</td>
<td>0.413</td>
<td>.6793</td>
</tr>
<tr>
<td>Allergies</td>
<td>9/98</td>
<td>11/79</td>
<td>0.990</td>
<td>.3220</td>
</tr>
<tr>
<td>Asthma</td>
<td>9/99</td>
<td>12/79</td>
<td>1.253</td>
<td>.2101</td>
</tr>
</tbody>
</table>

*From 2-tailed t test. SD = standard deviation.

SNP Association Analysis

Exploratory data analysis was performed on the cohort data to identify outliers and differences in distributions. SNPs were assessed for genotyping quality and heterozygosity. Quality control across all SNPs and across all samples resulted in the removal of 22 SNPs: poorly formed genotyping clusters (six), homozgyous SNPs (seven), low call rate (78%) (one), minor allele frequency <1% (four), and out of Hardy Weinberg equilibrium in controls (four). This leaves 170 markers for the analysis (see Supporting Table 1 in the online version of this article). The statistical association of each individual SNP to case/control status was determined using a score test as implemented in R statistical genetics package GenABEL (R Foundation for Statistical Computing, Vienna, Austria, http://www.r-project.org/).

RESULTS

Demographic Data

Two hundred saliva samples were collected; 15 samples were lost in processing or did not have sufficient DNA in the collected saliva, leaving a remaining 185 subjects in the study (102 cases, 83 controls). Subject samples were removed for the following reasons: duplicate sample (one), call rate <97% (three), and related subjects as identified by Descent >95% (two). This leaves 179 subjects for analysis—100 cases and 79 controls.

Phenotype Data

Cases had an average duration of 7 months of middle ear effusions. Cases and controls were examined for differences in age or sex distributions; distribution of age was similar for cases and controls (P = .585) as well as evenly distributed for sex (P = .608) (Table I). There was no statistically significant difference in smoke
exposure in the home, environmental/food allergies, daycare attendance, or asthma between the two study populations. However, there was a significantly higher proportion of COME phenotype with a family history of otitis media (Table I). Diagnoses bringing control patients into the otolaryngology clinic included: hemangioma or vascular malformation (18), lymphadenopathy (four), dysphagia (six), branchial cleft anomaly or thyroglossal duct cyst (five), sensorineural hearing loss (four), neck mass (seven), vocal cord nodules (three), and other (32).

SNP Association Analysis

Twenty-two SNPs were removed from the final dataset (see Materials and Methods). Of the selected SNPs, a total of 13 were previously reported as associated with the otitis media population (dbSNP database). None of the SNPs suggestive for association in our study were from this group of SNPs (see Supporting Table I in the online version of this article). The mean heterozygosity for an SNP was $0.294 \pm 0.163$. The mean heterozygosity for a subject was $0.294 \pm 0.050$. Because of the number of SNPs tested, the P values were corrected for multiple testing through permutation testing (Table II). After multiple testing correction with permutation testing, the eight SNP markers failed to hold their P values ($P = 0.625–0.99$, Table II).

There were eight SNP markers located in four genes that had an unadjusted P value <.05 (Table II). These SNPs were found in the TLR4 gene (five), MUC5AC/B (one), SMAD2 (one), and SMAD4 (one) (Table II). Five of the eight SNPs were in the TLR4 gene and range in LD from 0.4 to 0.8 (TLR4 plot, Fig. 1). This level of LD indicates that these SNPs are most likely capturing one association signal for TLR4. Effect size for the associated SNPs ranged from 2.212 to 0.346. The TLR4 SNP plot (Fig. 1) shows that the SNP rs10116253 possibly lies in the promoter region of the TLR4 gene. The MUC5AC/B SNP (Fig. 2) lies within the gene region itself. The SMAD4 SNP (Fig. 3) lies within the gene region itself. The

### Table II. SNP Association Analysis

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene</th>
<th>Chromosome</th>
<th>MAF Caucasian</th>
<th>Location</th>
<th>P Value (Unadjusted)</th>
<th>P Value (Adjusted)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs11788318</td>
<td>TLR4</td>
<td>9</td>
<td>NA</td>
<td>INTERGENIC</td>
<td>0.008</td>
<td>0.655</td>
</tr>
<tr>
<td>rs4837494</td>
<td>TLR4</td>
<td>9</td>
<td>0.158</td>
<td>INTERGENIC</td>
<td>0.031</td>
<td>0.98</td>
</tr>
<tr>
<td>rs10116253</td>
<td>TLR4</td>
<td>9</td>
<td>0.25</td>
<td>INTERGENIC</td>
<td>0.007</td>
<td>0.625</td>
</tr>
<tr>
<td>rs1927914</td>
<td>TLR4</td>
<td>9</td>
<td>NA</td>
<td>INTERGENIC</td>
<td>0.023</td>
<td>0.94</td>
</tr>
<tr>
<td>rs1554973</td>
<td>TLR4</td>
<td>9</td>
<td>0.2</td>
<td>INTERGENIC</td>
<td>0.021</td>
<td>0.93</td>
</tr>
<tr>
<td>rs4963049</td>
<td>MUC5AC/B</td>
<td>11</td>
<td>0.133</td>
<td>INTERGENIC</td>
<td>0.033</td>
<td>0.99</td>
</tr>
<tr>
<td>rs8097137</td>
<td>SMAD2</td>
<td>18</td>
<td>0.217</td>
<td>INTERGENIC</td>
<td>0.032</td>
<td>0.985</td>
</tr>
<tr>
<td>rs17663887</td>
<td>SMAD4</td>
<td>18</td>
<td>NA</td>
<td>INTRON</td>
<td>0.011</td>
<td>0.756</td>
</tr>
</tbody>
</table>

*Adjusted P values corrected for multiple testing with permutation testing.

MAF = minor allele frequency from HapMap-CEU (CEPH:Utah Residents with Northern and Western European Ancestry).

Fig. 1. Single nucleotide polymorphism (SNP) locus plot for the TLR4 gene SNPs tested (27). The TLR4 gene resides on chromosome 9. Five of the 27 SNPs tested (labeled on the figure) had significant P values (see Table II). The SNP rs10116253 (marked by the purple diamond) may lie in a promoter region. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]
**SMAD2** SNP lies in between the **SMAD2** gene and the **ZBTB7C** gene (Fig. 4). The SNP appears to lie in an area of recombination. Whether the SNP truly belongs to **SMAD2** or **ZBTB7C** could depend on the annotation of the database used. In our case, the dbSNP database annotates this SNP to the **SMAD2** gene.

**DISCUSSION**

To increase the chances of encountering SNPs of interest, it was our intent to study a very narrow population of genetically and phenotypically similar OM patients. Previously published studies on the genetics of OM have either studied the ROM phenotype or the combined ROM/COME phenotype together. Otolaryngologists are well aware that the COME phenotype is distinct from ROM in terms of age of presentation and character of fluid present in the middle ear. Although there is substantial overlap with the AOM phenotype, we believe that the phenotypes are distinct. By careful screening, we were able to study the distinct COME phenotype and avoid the problems of a diluted or broad phenotype sample.

The two sample populations, case and control, were similar in terms of age, sex distribution, smoke exposure, allergy history, and asthma history. The exception to their similarity is that more subjects in the case group had a family history of OM. The increased family history in the case group makes intuitive sense if a predisposition to OM has a genetic contribution, and a familial association with the otitis-prone child has been shown in the past.
One limitation of the current study is the small number of genes and SNPs studied. However, the genes were chosen based on the experience with the C3H/HeJ TLR4-deficient mouse and its COM phenotype, as well as an extensive review of the literature for candidate genes. In addition, the sample size was relatively small for a tag-SNP study, but we limited the number of genes studied based on the power analysis performed. Last, when correction for multiple testing was performed, the SNPs with a significant association with COME did fail to retain their significance. Due to the multiple testing burden associated with the number of SNPs tested, it is possible that this study did not have the power to see an association. A replication study powered to test for the association of these eight SNPs with the COME phenotype is planned. Undoubtedly, the inheritance of OM is multifactorial and complex.

These eight SNPs associated with the COME phenotype are from genes governing immune response (TLR4), mucin production (MUC5B), and activation of TGF-β transcription/mediation of signaling pathways for cell proliferation/apoptosis and cell differentiation (SMAD2, SMAD4), all of which could play a role in the predisposition to OM. The confirmation of the phenotype in the mouse model (C3H/HeJ) lends weight to the gene’s possible role in COME.

Future studies are planned to confirm the eight SNPs in a novel population of COME children. Eventual GWAS will need to be performed on carefully screened COME patients to look for association with other genes of interest.

CONCLUSION

Eight candidate SNPs were found with unadjusted P values of <.05, five of them in the TLR4 gene. Due to the confirmation of COME in the TLR4-deficient mouse and the likely association of the other genes showing association with SNPs in this study with a predisposition to COME, these eight SNPs merit confirmation in a replication population. The innate immune response gene TLR4, among others, may confer susceptibility to COME.

BIBLIOGRAPHY