Determinants of host susceptibility to murine respiratory syncytial virus (RSV) disease identify a role for the innate immunity scavenger receptor MARCO gene in human infants

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Published Version
doi:10.1016/j.ebiom.2016.08.011

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:29408435

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Determinants of host susceptibility to murine respiratory syncytial virus (RSV) disease identify a role for the innate immunity scavenger receptor MARCO gene in human infants

Monica Higha,1, Hye-Youn Choa,1, Jacqui Marzeca,1, Tim Wiltshire b, Kirsten C. Verheina, Mauricio T. Caballero c, Patricio L. Acostac,d, Jonathan Ciencewickia, Zackary R. McCawa, Lester Kobzik e, Laura Miller-DeGraff a, Wes Gladwell a, David B. Peden f, M. Elina Serrac, Min Shig, Clarice Weinberg g, Oscar Suzukib, Xuting Wang h, Douglas A. Bell h, Fernando P. Polack c,i,⁎, Steven R. Kleebergera,⁎⁎

a Immunity, Inflammation, and Diseases Laboratory, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA
b Division of Pharmacotherapy and Experimental Therapeutics, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, USA
c Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina
d Department of Environmental Health, Harvard University School of Public Health, Boston, MA, USA
e Center for Environmental Medicine, Asthma and Lung Biology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
f Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA
g Genome Integrity & Structural Biology Laboratory, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA
h Department of Pediatrics, Vanderbilt University, Nashville, TN, USA

⁎ Correspondence to: F.P. Polack, Vanderbilt University, Department of Pediatrics, B-3307 MCN, 1161 21st Ave South, Nashville, TN 37232-2007, USA.
⁎⁎ Corresponding author at: Immunity, Inflammation, and Diseases Laboratory, National Institute of Environmental Health Sciences (NIEHS), 111 T.W. Alexander Drive, Building 101, MD D-201, Research Triangle Park, NC 27709, USA.
E-mail addresses: fernando.p.polack@vanderbilt.edu (F.P. Polack), kleeber1@niehs.nih.gov (S.R. Kleeberger).

1 These authors contributed equally to this work.

Abstract

Background: Respiratory syncytial virus (RSV) is the global leading cause of lower respiratory tract infection in infants. Nearly 30% of all infected infants develop severe disease including bronchiolitis, but susceptibility mechanisms remain unclear.

Methods: We infected a panel of 30 inbred strains of mice with RSV and measured changes in lung disease parameters 1 and 5 days post-infection and they were used in genome-wide association (GWA) studies to identify quantitative trait loci (QTL) and susceptibility gene candidates.

Findings: GWA identified QTLs for RSV disease phenotypes, and the innate immunity scavenger receptor MARCO was a candidate susceptibility gene; targeted deletion of MARCO worsened murine RSV disease. We characterized a human MARCO promoter SNP that caused loss of gene expression, increased in vitro cellular response to RSV infection, and associated with increased risk of disease severity in two independent populations of children infected with RSV.

Interpretation: Translational integration of a genetic animal model and in vitro human studies identified a role for MARCO in human RSV disease severity. Because no RSV vaccines are approved for clinical use, genetic studies have implications for diagnosing individuals who are at risk for severe RSV disease, and disease prevention strategies (e.g. RSV antibodies).

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

RSV is the primary cause for hospitalization during the first year of life in infants worldwide, and it is the leading cause of lower respiratory tract infection leading to bronchiolitis, pneumonia, and respiratory failure (Lozano et al., 2012). RSV is also a significant cause of respiratory illness in immunocompromised adults and the elderly (Falsey et al., 2005). While the majority of infected subjects present symptoms, the nature and severity of the symptoms vary among individuals (Ciencewicki et al., 2014; Caballero et al., 2015). For some, infection induces cold-like symptoms; others require hospitalization, and a small
percentage of cases require intensive care and may result in death. The wide variation in response to RSV infection suggests that susceptibility and disease are influenced by multiple host intrinsic factors (Stark et al., 2002; Miyairi and DeVincenzo, 2008; Gelfand, 2012; Feldman et al., 2015).

Previous in vivo mouse studies suggest that RSV infectivity is a multigenic trait, but did not identify the genes responsible for infection (Prince et al., 1979; Stark et al., 2002; Stark et al., 2010). The innate and acquired immune systems have been implicated in the pathogenesis of RSV infection, including surfactant association proteins, viral response receptors and proteins, and inflammatory cells and their products such as chemokines, cytokines, and soluble inflammatory mediators [e.g. (Mukherjee and Lukacs, 2013; Varga and Braciale, 2013)]. However, roles for each of these components in the pathogenesis of RSV disease are unclear.

RSV disease is complex and, while candidate gene single nucleotide polymorphisms (SNPs) have associated with RSV disease severity in infants [e.g. (Miyairi and DeVincenzo, 2008; Daley et al., 2012; Ciencwcki et al., 2014; Caballero et al., 2015)], the mechanisms of differential disease susceptibility remain unclear. To better understand the genetic basis of RSV disease, we performed a genome-wide association (GWA) study of disease following infection in inbred strains of mice to identify candidate quantitative trait loci (QTL) and genes that associate with specific disease phenotypes. We identified 7 significant and 8 suggestive QTLs and candidate susceptibility genes including macrophage receptor with collagenous structure (MARCO), an innate immunity scavenger receptor. Targeted deletion of mouse Marco worsened RSV disease phenotypes, consistent with a protective role in disease pathogenesis. We then identified and characterized in vitro a human MARCO promoter SNP that diminished mRNA expression basally and after RSV infection. The SNP is located in an antioxidant response element (ARE) of the MARCO promoter and is a binding site for the transcription factor nuclear factor erythroid-derived-2 like 2 (NFE2L2), also known as nuclear factor erythroid 2 (NF-E2)-related factor 2 (NRF2). Deletion of the ARE also reduced MARCO mRNA expression following RSV infection in airway epithelial cells. Moreover, compared to the wild-type C allele, we found that the MARCO promoter rs1318645 SNP G allele associated with increased disease severity in populations of RSV-infected infants. Our modeling of disease after RSV infection in mice thus identified a host susceptibility gene candidate that has implications for better understanding human disease.

2. Materials and Methods

2.1. Mice and Treatments

Mice from 30 inbred strains (male, 6–10 weeks) for the strain screen were purchased from Jackson Laboratory (Bar Harbor, ME) and housed singly with hardwood bedding (P) Murphy Forest Products Company, Montville, NJ) in pathogen-free facilities at NIEHS. A breeding colony of Marco deficient mice (Marco−/−) was developed and maintained at NIEHS (Arredouani et al., 2004; Cho et al., 2012). Age- and sex-matched C57BL/6 wild-type (WT) mice were purchased from Charles River Laboratories (Wilmington, MA) and used as controls for experiments with the Marco−/− mice. All mice were provided free access to water and pelleted open-formula rodent diet NIH-07 (Zeigler Brothers, Gardners, PA.). Constant air temperature (72 ± 3°F), relative humidity (50 ± 15%), and light/dark cycle (12 h on/12 h off) were maintained during the experiments. Mice of each strain were randomly assigned to either RSV infection or vehicle control groups. Mice were infected in the morning with 105 plaque-forming units (PFU) of human RSV-19 strain in 50 μl Hank’s balanced salt solution (HBSS) by intranasal instillation, and returned to their cages. HBSS containing HEP-2 cell lysates were intranasally instilled for vehicle controls. On post-instillation days 1 and 5, mice (6–12 per group) were killed by overdose with Nembutal (sodium pentobarbital; 104 mg/kg body weight). Phenotyping for response to RSV infection was performed in at least two separate experiments (n ≥ 3 per group), and power calculations based on previous studies were used to determine sample sizes. All animal use was approved by the NIEHS Animal Care and Use Committee. Animals were treated humanely and with regard to alleviation of suffering.

2.2. Lung Bronchoalveolar Lavage Fluid (BALF) and Cell Preparation

The right lung of each mouse was lavaged in situ consecutively four times with HBSS (35 ml/kg body weight, pH 7.2–7.4) and BALF was analyzed for cell differentials using standard morphological and cytological criteria as described previously (Saltini et al., 1984; Cho et al., 2001). Total protein content (a marker of lung permeability) was measured with the Bradford assay as described previously (Cho et al., 2001; Cho et al., 2007). Investigators were blinded to treatment for all lavage and histopathology samples (see below). Time points chosen for evaluation of inflammation and injury were based on a previous investigation of phenotype kinetics in lungs of a panel of 5 inbred strains of mice infected with RSV (data not shown). We found that the peaks of inflammation and injury phenotypes are time-dependent and usually occur either 1 or 5 days after infection as described previously (Cho et al., 2009).

2.3. Lung Histopathology

Left lung lobes were inflated intratracheally in situ and fixed with 10% neutral buffered zinc formalin for 24 h following procedures described previously (Cho et al., 2009). Fixed lungs were then placed in 70% ethanol until proximal and distal intrapulmonary axial tissues were excised and embedded in paraffin and sectioned (5 μm). Tissue sections were stained histochemically with hematoxylin and eosin (H&E) for histopathological analyses and alcin blue-periodic acid shift (AB/PAS) double staining to identify acidic and neutral mucosubstances (Cho et al., 2009).

2.4. Measurement of Stored Intraepithelial Mucus (Volume Density, VS)

Images of histological sections were captured using an Axioskop 40 microscope, camera, and Axiovision software. Black and white images were taken with a 10× objective lens and analyzed with Scion image analysis software. The mucus-containing area and the perimeter of basal lamina from large airways, visually judged to be third generation bronchus, were used for analysis. The volume of mucus per airway was calculated as described previously (Harkema et al., 1987).

VS values (nl/mm2 basal lamina) were calculated for each lung section and averaged for vehicle control and RSV infected groups. VS calculations were averaged for each treatment group and are reported as the mean ± SEM for each strain (n = 3/infection/strain).

2.5. Quantitative Real-time Polymerase Chain Reaction (qRT-PCR)

Total RNA was isolated from non-lavaged left lung pieces using RNeasy Mini Kits (Qiagen, Valencia, CA) following manufacturer procedures and as described previously (Cho et al., 2009). Quantification of gene expression was determined by ΔCt values obtained by subtracting fluorescence detected number of cycles to threshold (Ct) of 18s mRNA from those of target gene RNA in the same sample.

2.6. Heritability and Correlation of Phenotypes

Heritability estimates were calculated following methods described previously (Lighthart et al., 2004; Nichols et al., 2014). Heritability estimates > 35% were considered sufficient for GWA studies with 30 inbred strains of mice (Tsaih and Korstanje, 2009). Relatedness of RSV disease phenotypes was done using Pearson’s correlation coefficients (Gene Network, http://www.genenetwork.org). The maximum
response [1 or 5 days post-infection (PI)] for each phenotype was used for all analyses.

2.7. Genome-wide Association Mapping

GWA mapping was performed using the mean value for each phenotype for each strain. Data for each strain were distributed normally (no outliers removed), and no transformations were applied. We also generated a RSV disease index by performing a principal components (PC) analysis of each quantitated phenotype and then applied the first 3 PCs for each strain to haplotype association mapping. Two different haplotype-based approaches were used: SNPster (Genomics Institute, Novartis Research Foundation, San Diego, CA) and FastMap (University of North Carolina, Chapel Hill, NC). The algorithms (Pletcher et al., 2004; Gatti et al., 2009) and rationale for using these approaches have been described elsewhere (McClurg et al., 2006; Nichols et al., 2014). Experimental informative SNP genotypes for GWA mapping were selected as described previously (Benton et al., 2012). The SNPs used for the 30 strains of mice were fully genotyped (i.e. not imputed). $\log_{10} P$ values were considered significant if they exceeded 6.60 (genome-wide significance at $P < 0.05$) and suggestive if they exceeded 6.06 (genome-wide suggestive at $P \sim 0.20$). Only QTLs that were identified using both algorithms were considered further for candidate gene analyses. Candidate gene SNP and haplotype data were obtained from the Mouse Phenome Database (MPD; Jackson Laboratory: http://phenome.jax.org; dbSNP 138 annotation). Non-informative SNPs (i.e. no heterozygosity between strains or the minor allele was found in <10% of the strains) were removed from further haplotype and SNP analyses. Marco haplotype data were obtained from the MPD, and hierarchical clustering was performed as described previously (Cho et al., 2015).

2.8. Sequencing of MARCO Promoter and Functional Assessment

We used traditional walk-down sequencing to identify regions of MARCO that were polymorphic, and confirmed these sites by screening an ethnically diverse panel of individuals for allelic variation. Regions with SNPs at 5% or higher frequency were used in promoter deletion experiments to test for effects on transcriptional activity. Constructs were generated to encompass the full length of the MARCO promoter (−1388 to +98), and polymorphic regions were sequentially deleted to assess differential activity.

Initial promoter deletion assays identified a polymorphic region of the MARCO promoter with high activity (−300 and +98) that contained a putative NRF2-binding ARE. To determine whether the SNP region influenced transcript levels, the −300 to +98 region of MARCO containing the rs1318645, −156 C/G SNP (−157) was cloned upstream of the luciferase reporter gene and its expression analyzed by transient transfection using reporter assays. The −300 to −71 promoter region was further investigated by generating two constructs [−300 to +98 bp, (300-Luc) and −71 to +98 bp (71-Luc, lacking the SNP)] and comparing their promoter activity.

To examine effects of the promoter SNP on transcriptional activity, we amplified the 5′-flanking region of genomic MARCO promoter to generate polymorphic variants of the 300-Luc sequence [−156 C-Luc and −156 G-Luc]. All constructs were verified by sequencing, and transiently transfected using FuGENE HD (Promega, Carlsbad, CA) into BEAS-2B cells (The American Type Culture Collection, Manassas, VA), and analyzed for luciferase activity as described previously (Cho et al., 2015). For RSV infection we used the pGL4.20 vector, and Beas2B cells were transfected with MARCO promoter constructs [−156 C-Luc, −156 G-Luc or −71-Luc] overnight and were serum starved for 1 h during RSV infection (MOI of 4). Cells were harvested 3, 6, 24, 48 and 72 h post-infection. Luciferase activity was assessed in three independent samples and all experiments were repeated in triplicate.

2.9. Electrophoretic Mobility Shift Assay (EMSA)

EMSA allows visualization of transcription factor binding affinity to regulatory promoter sequences in the presence of variant alleles. DNA-protein binding specificity was further assessed using antibodies specific to the binding factor of interest. An aliquot (5 μg) of human lung nuclear protein (Active Motif, Carlsbad, CA) was incubated on ice with binding buffer for 15 min, followed by addition of $3 \times 10^6$ cpm [γ^32P] ATP (Amersham Biosciences, Piscataway, NJ) end-labeled wild type or variant probes (End-labeled probes: rs1318645 −156 C (wt) FORWARD 5′−CTG AGT GAG TGC ATT CAG AG−3′; REVERSE 5′−CTG AAT GCA CCC GCT ACT CCC AGG−3′; rs1318645 −156G (variant) FORWARD 5′−CTG AGT GAG TGC ATT CAG AG−3′; REVERSE 5′−CTG AAT GCA CCC GCT ACT CCC AGG−3′) and incubated for 30 min at room temperature. We used short (26 base pairs) fragment sequences for the EMSA experiments that overlap the consensus sequence for NRF2. We used polyclonal rabbit anti-NRF2 antisera (2 μl) or control IgG (2 μl) as controls and processed for EMSA as described previously (Marzec et al., 2007). Samples were electrophoresed and autoradiographed with an intensifying screen at −70 °C. Genotyping was done without knowledge of disease status (i.e. blinded to case or control). We tested for statistically significant differences in DNA-NRF2-antibody complex formation between wild-type, heterozygous, and variant genotypes using one-way ANOVA with SNK a posteriori comparisons of means.

2.10. Human RSV Disease Population

Two independent prospective case-control studies were conducted in Buenos Aires, Argentina from 2003 to 2006 (population 1) and 2010–2013 (population 2); these studies, with inclusion and exclusion criteria, were described previously (Caballero et al., 2015). Study participants were previously healthy full-term infants younger than 1 year of age and born after 15 September of the previous year (i.e. approximately 15 days after the end of RSV season in Buenos Aires). Participants had bronchiolitis symptoms for the first time in their lives, and clinical signs included wheezing with or without cough, rales, dyspnea, and increased respiratory rate and retractions of the respiratory muscles. Diagnosis was performed by trained pediatricians. Infants were tested for RSV, rhinoviruses, hPIV1–4, influenza virus A and B, and human metapneumovirus, and all participating subjects were infected only with RSV. Recruited cases were infants with bronchiolitis whose oxygen saturation upon enrollment was <93% in room air (the primary outcome). Oxygen saturation is a clinically relevant phenotype for discriminating between mild and severe RSV disease (Caballero et al., 2015). Infants with bronchiolitis were recruited as controls if oxygen saturation upon enrollment was ≥93% while breathing room air. Genotyping of infants was performed using whole blood samples (population 1) or nasal aspirates (population 2). Population sizes were based on calculations to provide sufficient statistical power (−80%) to find significant differences between mild and severe disease with a minor allele frequency of −0.2. The Institutional Review Boards (IRB) of the participating hospitals in Buenos Aires, as well as the Johns Hopkins University IRB (Baltimore, MD, USA), approved the protocol. We obtained informed consent from parents of each infant enrolled in the studies. The study conformed to standards indicated by the Declaration of Helsinki.

2.11. Allelic Discrimination Genotyping for MARCO rs1318645

Genomic DNA was isolated and characterized for purity and concentration as described previously (Caballero et al., 2015). The MARCO rs1318645 promoter variant was assessed by allelic discrimination with a predesigned SNP genotyping assay from Applied Biosystems. PCR was performed using 20 μl reactions with 5 ng of genomic DNA, 10 μl of TaqMan 2X PCR Master Mix, and 1.0 μl of 20 × pre-optimized
assay mix (C_9566534_20). The 20× mix consisted of 18 μM forward and reverse primers and 8 μM of each allele-specific fluorescently labeled (VIC or FAM) TaqMan MGB probe. Standard PCR cycling conditions were used as described previously (Caballero et al., 2015).

Allele-specific PCR products were detected and verified as described previously (Caballero et al., 2015). Five percent of the samples were repeated for quality control, and we found 100% concordance.

2.12. Statistical Analysis

Lavage and tissue parameters were analyzed by 3-way ANOVA (factors: exposure, time, strain) and Student-Newman-Keuls (SNK) comparisons of means procedures. Strain distribution patterns and Marco haplotype groups were analyzed by 1-way ANOVA and SNK comparisons of means. Statistical associations between candidate gene

![Image](image.png)

Fig. 1. Lung phenotypes in representative mouse strains with mild and severe disease after infection with RSV. Numbers of PMNs (a) and monocytes (b) found in BALF from C3H/HeJ (mild disease) and BALb/cByJ (severe disease) mice following vehicle or RSV infection (1, 5 days PI). Mean ± SEM (n = 6–12 mice/group). Groups were analyzed by 3-way ANOVA and Student-Newman-Keuls a posteriori tests. *P < 0.05 versus vehicle.+P < 0.05 versus C3H/HeJ. (c) Representative H&E stained lung sections from C3H/HeJ (mild disease) and A/J (severe disease) mice following vehicle and RSV infection (1, 5 days PI). Arrows = areas of increased airway inflammation and bronchial epithelial proliferation (hyperplasia). AV = Alveoli, BR = bronchus or bronchiole, BV = blood vessel. Bar = 100 μm. (d) Representative AB/PAS double stained lung sections from BTBR_T + _tf/J (mild disease) and SJL/J (severe disease) mice following vehicle and RSV infection (1, 5 days PI). (e) Circular node and edge representation of relatedness between RSV response phenotypes among all mouse strains. Pearson correlation coefficients (R) noted on edges (solid lines) indicate associations between phenotypes. R values ≥ 0.48, P < 0.05; R values ≥ 0.73, P < 0.01.
genotypes and phenotypes within mouse strains were calculated using the two-tailed Student’s t-test. Only the significance of non-synonymous coding SNPs was found within a gene multiple comparisons were performed. Because of potential bias effects of admixture in our human study populations, we followed methods described previously to test for evidence of bias due to population stratification (Epstein et al., 2007) using 33 ancestral informative markers (AIMs; Suppl. Table S1) that were evaluated in 300 infants with mild or severe RSV disease. Genotyping was performed on the Sequenom iPLEX platform by BioServe (Beltsville, MD, USA). In the two infant populations, we used $\chi^2$ analyses (dominant and recessive modeling; SigmaPlot 13.0) to test for associations between MARCO genotypes and RSV disease severity. We assessed potential confounding of demographic variables in the relationships between MARCO genotype and RSV disease severity using an additive model for multiple logistic regressions. Statistical significance in these and all other comparisons was accepted at $P < 0.05$.

3. Results

3.1. Phenotypic Characterization of RSV Disease in Inbred Mice

We developed a genetic model of RSV disease by infecting intranasally 30 strains of mice. Inflammation was assessed in BALF by measuring total protein concentration and counting inflammatory and epithelial cells 1 and 5 days PI. The time courses of inflammation and injury phenotypes indicated early (acute) and late phase responses in all of the strains tested. Relative to vehicle controls, mean numbers of total cells, polymorphonuclear leukocytes (PMNs), epithelial cells, and total protein in BALF of infected mice were significantly increased 1 day PI and declined 5 days PI in the majority (>50%) of strains (representative strains, Fig. 1a; all strains, Suppl. Table S2). Numbers of BALF macrophages, lymphocytes, and monocytes (representative strains, Fig. 1b; all strains, Suppl. Table S2) peaked at 5 days PI in the majority of strains. The differences in kinetics may reflect the sequential contributions of the innate and acquired immune systems in response to infection as others have found [e.g. (Kok et al., 2012; Stoppelenburg et al., 2013)]. Qualitative inspection of lung histopathology sections was consistent with BALF inflammation phenotypes (Fig. 1c). Mucus cell metaplasia is a clinical feature of RSV disease, and we found the mean amount of intraepithelial cell mucus (Fig. 1d; Suppl. Table S2) and lung RSV-N mRNA expression (Suppl. Table S2) were greatest 1 day PI in 59% and 70% of the strains, respectively.

We also found statistically significant inter-strain variation in the peak of each RSV disease phenotype. That is, significantly smaller numbers of inflammatory cells and other indicators of injury (mild disease) were found in a subset of strains (e.g. C3H/HeJ) compared to others with severe disease (e.g. BALB/cByJ) (Fig. 2a–d; Suppl. Table S2; Suppl. Fig. S1). Heritability estimates of RSV disease phenotypes ranged from 47.6% (BALF protein, 1 day PI) to 85.7% (intraepithelial mucus, 1 day PI) (Suppl. Table S2). To better understand mechanisms of response to RSV infection, we tested for correlation of disease phenotypes. Pearson’s correlation analyses of RSV disease phenotypes among 30 strains found that responses 1 day PI did not correlate with responses 5 days PI (data not shown). When maximum responses for each parameter were compared, few were significantly correlated (Fig. 1e; Fig. 2). Interestingly, RSV-N mRNA expression was not correlated with any response phenotype, a finding consistent with lack of correlation we found between RSV disease severity and RSV titers in infected children (Caballero et al., 2015). Large inter-strain variation and relatively high heritability of distinct disease phenotypes suggests the mouse model of response to RSV infection is multi-genic and mimics the variability of RSV disease severity found in human populations.

3.2. Identification of QTLs and Gene Candidates for RSV Disease Phenotypes

Because the phenotypic responses to RSV infection were largely not correlated, we reasoned that the phenotypes may have some QTLs in common, but likely also have unique QTLS. Therefore, we used GWA mapping to identify QTLs for each phenotype across the entire inbred strain data set. Using the maximum mean response phenotypes for each strain, mapping with SNPster identified significant (−$\log_{10} P > 6.6$) QTLS for RSV-induced changes in BAL monocytes (chromosomes 1 and 17; Fig. 3a, Table 1) and PMNs (chromosomes 3 and 8; Table 1), and the first PC (chromosomes 4, 11, and 13; Table 1). Suggestive QTLS (−$\log_{10} P > 6.0$) were also found for intraepithelial mucus content (chromosomes 7 and 15), BALF total protein concentration (chromosomes 9 and 19), and lung RSV-N mRNA expression (chromosomes 1 and 11) (Table 1). Importantly, all QTLS were also identified using FastMap (data not shown). GWA mapping of other RSV response phenotypes with SNPster did not identify significant or suggestive QTLS, or were not confirmed with FastMap.

To further investigate genetic contributions to RSV disease phenotypes we searched the MPD for informative SNPs in susceptibility gene candidates within significant and suggestive QTLS, and whether they

Fig. 2. Lung disease phenotypes in mice infected with RSV. Maximum numbers of BALF monocytes (a), lymphocytes (b), and PMNs (c), and RSV-N mRNA expression (d) in lung homogenates. Bars, means ± SEM ($n = 6–12/group$). Horizontal brackets (monocytes) indicate means are not significantly (P < 0.05) different from each other (1-way ANOVA and Student-Newman-Keuls comparisons of means).
associated with respective disease phenotypes. Genes were chosen for investigation if they had biological plausibility for a role in RSV disease and if the minor allele of candidate SNPs were present in ≥10% of the phenotyped inbred strains. Eleven gene candidates were identified in 9 QTLs that met these criteria and also contained SNPs that associated with differential susceptibility to an RSV disease phenotype (Table 2).
thus supporting a potential role for these genes in murine RSV disease severity.

We focused initially on the chromosome 1 QTL for monocyte RSV inflammation to validate gene candidates because this QTL (~1.299 Mb) had the greatest GWA = −Log10 P value (Table 2). The scavenger receptor gene Marco is located in the QTL (Fig. 3b) and is in linkage disequilibrium (LD) with the SNP that has the highest −Log P value in the QTL, and human MARCO is upregulated in children infected with RSV [i.e. biological plausibility (Fjaerl et al., 2006)]. Hierarchical clustering of RSV SNPs identified 3 major haplotypes in the strain panel (Fig. 3c; Suppl. Table S3). Mean numbers of RSV-induced BAL monocytes in strains from Marco haplotypes 2 and 3 were significantly greater than those in strains from wild-type haplotype 1 (Fig. 3d). Potentially functional, informative SNPs are located in Marco, including the non-synonymous coding SNP rs30741725 (A/T) that causes a threonine to serine substitution at the 476 amino acid residue in exon 17. Significantly increased numbers of BAL monocytes and lymphocytes 5 days PI were found in strains of mice with the minor T allele for this SNP compared with strains that have the major A allele (Table 2; Suppl. Fig. S2A and S2B). The majority of strains with the T allele have Marco haplotypes 2 or 3. The rs30741725 A/T SNP is in LD with other SNPs in the QTL with Marco (stated above).

3.3. Genetic Deletion of Marco Modifies RSV-induced Pulmonary Inflammatory Response

We further investigated the role of Marco on RSV disease severity by infecting Marco+/+ and Marco−/− mice with RSV. Relative to vehicle controls, RSV increased BAL monocytes in Marco÷/÷ and Marco÷−/− mice, and numbers of monocytes in Marco÷−/− mice were slightly lower compared to Marco÷/÷ mice 1 day PI, during the development of the monocyte response. However, numbers were significantly greater in Marco÷/−/− mice compared to those in Marco÷+/− mice at the peak of the response (5 days PI; Fig. 3e, left). Mean numbers of PMNs were also significantly increased in Marco÷/÷ and Marco÷−/− mice 1 day PI, but were significantly greater in Marco÷/−/− mice compared to Marco÷+/− mice, and remained significantly elevated at 5 days PI (Fig. 3e, right). BALF inflammation phenotypes were consistent with lung histopathology sections from Marco÷/÷ and Marco÷−/− mice (Fig. 3f). No significant genotype effects were found for other RSV BALF phenotypes (data not shown). Lack of resolution of the PMN response in Marco÷/−/− mice is consistent with the hypothesis that inflammatory cell clearance is impaired in these mice.

3.4. Functional Assessment of the Human MARCO Promoter

Based on our GWA investigation and identification of Marco as a gene candidate for differential RSV disease severity in the mouse, we queried dbSNP (build 129, 2008) to identify potentially relevant SNPs in human MARCO based on functional annotation. A missense SNP (rs61732282, Asp511Glu) in the scavenger receptor cysteine rich (SRCR) domain of exon 17 (similar to that found in mouse Marco), is predicted by SIFT (sift.jcvi.org) to have deleterious consequences, but the 1000 genomes project phase 1 (www.1000genomes.org) indicates the alternate allele frequency of 0.05 is too low to study in our populations of children with RSV-induced bronchiolitis. We then asked whether the MARCO promoter contains polymorphic regions that might alter transcriptional efficiency. We consecutively deleted sections of the promoter from −1388 to −71 and probed for transcriptional activity using luciferase reporter constructs. We identified a region close to the transcription start site (−300 to −71) that, when deleted, significantly reduced transcriptional activity (Fig. 4a). We then screened commercially available DNA samples (n = 40; Coriell Institute, Camden, NJ) for SNPs in the MARCO promoter region. Forty subjects (20 European, 20 African-American) were chosen to provide a detection rate of 98% for SNPs present in as little as 5% of the population. The screen identified a SNP (rs1318645, −156 C/G; Suppl. Fig. S3) that contained a putative antioxidant response element for NRF2 [C allele, position weight matrix (PWM) score = 10.9; G allele, PWM score = 7.5]. Mutation of the −156 site from C (wild type) to G completely reduced luciferase activity to that of the pGL3 basic vector (Fig. 4b).

Polymorphic and ARE deletion constructs cloned in BEAS-2B epithelial cells were also assessed for transcriptional activity following RSV infection. Robust MARCO activation was found 3, 6 and 24 h PI, and mutation from C to G at the −156 site reduced activity approximately 50% at 6 h PI, and constructs with ARE deletion caused similar reduction in activity (Fig. 4c). At 24 h PI, luciferase activity in constructs with ARE deletion was reduced by 50% compared to that of wild type −300C

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of QTL Type, Genomic Location (Mb), and Maximum Log10 P values for SNPs in 30 mouse strains infected with RSV.</td>
</tr>
<tr>
<td>Phenotype</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Monocytes</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>17</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>PMNs</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>Mucus</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>Protein</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>19</td>
</tr>
<tr>
<td>RSV N</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>PC 1</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>13</td>
</tr>
</tbody>
</table>

**P** = 0.05 versus Marco haplotype 1 (1-way ANOVA and Student-Newman-Keuls comparisons of means). (a) Overlaps with another QTL.
constructs (Fig. 4c). Transcription activity was close to basal levels 48 and 72 h PI (data not shown).

Gel shift analysis showed reduced NRF2 specific binding in the presence of the variant G allele in the potential ARE (Fig. 4d), suggesting binding efficiency is altered at this polymorphic site. Overall binding was increased with the G allele, suggesting additional factors may bind this regulatory locus.

3.5. Role of MARCO rs1318645 in RSV Disease Severity

Next we asked whether the MARCO SNP rs1318645 that caused reduced MARCO gene expression in vitro (see above) associated with increased risk of severe RSV disease in populations of infected infants from our prospective case-control study (Ciencewicki et al., 2014; Caballero et al., 2015). Demographic characteristics are described elsewhere (Caballero et al., 2015). Genotyped (Fig. 5a) infants with mild and severe infection had combined allelic frequencies (Suppl. Table S5) similar to those published previously (http://www.1000genomes.org/). We were unable to successfully genotype 39 infants in the first population and 25 in the second population, and there was no association of unsuccessful genotyping between cases and controls. Allelic frequencies satisfied Hardy-Weinberg equilibrium (P = 0.369) and, although admixture was present we found no evidence of bias due to population stratification (i.e. stratification was not predictive of differential RSV disease severity). \( \chi^2 \) analysis of RSV disease severity found the rs1318645 C/G polymorphism was associated with increased risk of severe RSV disease (Fig. 5b; \( \chi^2 = 5.15, P = 0.053 \)). The association was largely attributed to GG risk effect (recessive model) where infants homozygous for the G allele were at significantly increased risk of developing severe RSV disease (unadjusted OR: 1.615, 95% CI: 1.24, 2.35; adjusted OR: 1.676, 95% CI: 1.28, 2.45).

4. Discussion

The impact of RSV disease, particularly in vulnerable infants, remains a global public health concern (Lozano et al., 2012). Unfortunately, no RSV vaccines are approved for clinical use. Therefore, it is critical to identify host risk factors that lead to development of prevention and intervention strategies to reduce risk of RSV disease morbidity and mortality, as well as asthma incidence because RSV disease has been identified as a predisposing risk factor (Dakhamma et al., 2005; Edwards et al., 2012; Feldman et al., 2015). Our study used genome-wide association analyses of RSV disease parameters that model human disease to identify biologically plausible susceptibility gene candidates, including Marco. Moreover, a missense loss of function polymorphism in the human MARCO homolog also associated with severe RSV disease in two populations of infected infants, and illustrates the utility of a genetic mouse model for understanding human disease. Importantly, this study provides additional genetic insight that may lead to a means to identify individuals at risk for severe RSV disease.

Considerable evidence has suggested that genetic background is an important host determinant of RSV disease severity, but full characterization of susceptibility genes for this complex disease remains unclear. Previous genetic investigations of human RSV disease have questioned whether polymorphisms in gene candidates associate with innate and acquired immune responses to infection and have provided some understanding of susceptibility factors. Gene candidates in these studies were chosen because of prior knowledge about biological plausibility. We comprehensively characterized RSV disease phenotypes in a large panel of inbred strains representing genetic diversity that is present in the mouse genome. We found significant variation and robust heritability in each RSV disease phenotype between the strains. However, relatedness between maximum response phenotypes was minimal, suggesting multiple genetic pathways likely contribute to the responses.

Anh and Desmecht (2006) reported that pneumonia virus of mice (PMV, the murine form of RSV) infection in a panel of 6 strains caused strain-specific histopathologic inflammatory effects, decreased body weight, and changes in pulmonary function particularly 5–7 days PI. The studies are not directly comparable because mice were infected with different viruses and late phase PI times investigated, but the inter-strain variation in phenotypes is consistent with our findings. Others (Stark et al., 2002) compared RSV infection in 8 inbred strains and found significant inter-strain variation in infectivity at 4 days PI, but histopathologic changes were minimal and cellular effects were
discordant with the magnitude of infection likely due to the time point when acute virus effects on histopathology were no longer present.

We leveraged extensive genetic diversity among the 30 strains of mice to perform GWA of multiple RSV disease phenotypes and thus were not guided a priori by hypotheses about the role of specific genes. Significant and suggestive QTLs were identified for inflammatory cells, intraepithelial cell mucus, and RSV N mRNA expression. Several genes that may contribute to viral response, pulmonary inflammation, and cytokine trafficking were found within the identified QTLs [e.g., Pgs2 (Obata et al., 2013), Ramp1 (Li et al., 2014), Rhdfl2 (Isauree et al., 2013)], and putative functional SNPs in the genes associated with differential RSV disease phenotypes among the strains. SNPs in other QTL genes strongly associated with disease phenotypes and, while little is known about their function in infectious disease, they may be important targets for future investigation of their roles in RSV disease progression. We also found significant QTLs using the first PC as an RSV disease index. Two of these QTLs overlapped those found for lymphocytes, and suggests that further investigation of genes within the QTLs could provide greater understanding of disease susceptibility. Future GWA investigations using PC analyses may provide additional insight to disease pathogenesis beyond using single phenotypes. Interestingly, genome-wide linkage analyses of RSV infectivity in back-cross and F2 populations derived from C57BL/6J and AKR/J mice by Stark et al. (2010) identified a QTL on chromosome 6 that associated with increased susceptibility to infection, though disease phenotypes were not presented. While GWA in the present investigation identified suggestive QTLs on chromosomes 1 and 11 for RSV-N mRNA expression, none were detected on chromosome 6. Differences in study design (e.g. haplotype mapping with many strains versus linkage analysis with 2 strains, PI time differences) may have contributed to this discrepancy.

We focused on the gene candidate Marco, which was identified by GWA for monocytosis, and has been shown previously to have protective effects in the lung following inhalation of particulate matter, exposure to ozone, and bacterial infection (Arredouani et al., 2005; Dahl et al., 2007; Thakur et al., 2008; Thakur et al., 2009; Ghosh et al., 2011). MARCO gene expression was also upregulated in children infected with RSV (Jaaeri et al., 2006). Our hypothesis that Marco contributes to defense against RSV-induced pulmonary inflammation was supported by two lines of evidence. First, strains of mice with a Marco haplotype containing the T allele of missense coding SNP rs30741725 were at greater risk of developing more severe RSV disease compared to strains with the wild type A allele. Second, numbers of BALF monocytes and PMNs were significantly increased in Marco+/− mice compared to Marco+/- mice. The mechanism through which Marco modulates the inflammatory response to RSV is not completely understood. We initially hypothesized that Marco has a direct role in clearing RSV from the airways, perhaps binding the virus for internal degradation by macrophages. To test this hypothesis, we infected with RSV lung macrophages isolated from Marco+/- mice but found no differences in infectivity, with and without co-culturing with epithelial cells (data not shown). Marco also has an important role in clearance of apoptotic cells (Rogers et al., 2009; Getts et al., 2014), and recent studies have targeted Marco receptors with immune-modifying particles to resolve monocytic inflammation (Getts et al., 2014). It is plausible that

---

**Fig. 4.** Effect of a genetic polymorphism on in vitro MARCO promoter activity. (a) Deletion analysis of the MARCO promoter compared to pG3 basic vector. *P < 0.05 versus −300-Luc; group sizes = 4. (b) Mutagenesis analysis of the MARCO rs1318645 variant (−156G) compared to pG3 basic vector. Bars (means ± SEM) represent luciferase activity in basic vector, control −156 C allele, and mutant −156 G constructs. Flanking sequence for MARCO rs1318645 is shown at top with the NRF2 response element boxed. *P < 0.05 versus −156:C (P = 0.001); group sizes = 4. Statistical analyses, one-way ANOVA and Student-Newman-Keuls comparisons of means. (c) Luciferase activity of RSV infected MARCO reporter constructs. MARCO promoter constructs were transiently transfected and cells infected with RSV (MOI of 4). Bars (means ± SEM) represent luciferase responses after no RSV (control) or 3–72 h post-infection (PI). Cells were harvested 72 h post-infection and luciferase activity assessed. *P < 0.05 versus RSV-infected −300:C (WT); + P < 0.05 versus No RSV control; group sizes = 6. Statistical analyses, 2-way ANOVA and Student-Newman-Keuls comparisons of means. (d) Gel shift analysis of MARCO promoter variants. Gel shift analysis of MARCO −156 CC (wild type), GC and GG (variant) oligos show increased overall binding in the presence of the G allele. Incubation with NRF2 antibody shows reduced MARCO:NRF2 specific binding with the variant G allele. This suggests the MARCO G allele affects efficient binding to the NRF2 ARE site, and may alter transcriptional activity. A representative gel shift is shown; group sizes n = 3. Single arrowhead, overall binding; double arrowhead, MARCO:NRF2 specific binding; FP is free radiolabeled probe, B is bound. NRF2 is NRF2 antibody, Igs is control serum, and Cold is competition without antibody (40×).
loss of function mutations in Marco compromise clearance of monocytes and apoptotic cells which leads to accumulation of these cells in the airways, further contributing to RSV disease pathology. Impaired resolution of the inflammatory response to RSV infection and more severe disease in Marco−/− mice compared to Marco+/+ mice is consistent with this hypothesis.

It is important to note that Marco does not account for all of the monocyte inflammation caused by RSV infection. Targeted deletion of Marco significantly increased susceptibility to some RSV-induced phenotypes but did not account for all of the genetic variation identified by the strain screen. The background strain (C57BL/6 in this model, a moderate responder to RSV) likely contributes to part of this difference as the choice of background can influence the effect of targeted deletion in mice [e.g. (Baleato et al., 2008; Leontiev et al., 2012)]. Perhaps more importantly, other QTLs were also identified by the GWA and gene candidates within them also contribute to the genetic component of the phenotype. Further investigation of these QTLs should provide additional insight to the genetic contribution to RSV disease susceptibility.

Because Marco had an important role in modulating the inflammatory response elicited by RSV, we sought to determine whether human MARCO was also important in RSV disease. We initially queried whether the MARCO rs1318645 −156 C/G SNP caused loss of function because it was located in an ARE proximal to the transcription start site. Others had found that the transcription factor NRF2, which binds to AREs to increased gene transcription, had a role in Marco expression in mice (Harvey et al., 2011; Bonilla et al., 2013). We showed previously that murine Marco bears a potentially functional ARE for Nrf2 binding (−940 region, PWM score = 8.9) in a model of neonate acute lung injury (Cho et al., 2012). In the present study, we found that deletion of the ARE and mutating the −156 site in an ARE in human MARCO also significantly attenuated gene expression as well as MARCO:NRF2 specific binding. Moreover, deletion of the ARE also significantly reduced expression of MARCO in cells infected with RSV thus suggesting an important role for the ARE and the rs1318645 −156 C/G SNP in MARCO function, basally and in response to viral infection. This is supported by our previous study in which increased lung injury and inflammation were found in Nr2 deficient mice compared to wild type mice after RSV infection (Cho et al., 2009).

We further investigated the role of MARCO in human RSV disease by asking whether the missense MARCO rs1318645 G allele was more prevalent in infected children with severe disease. In two independent populations, infected children homozygous for the rs1318645 G allele were more likely to have severe RSV disease than children who were heterozygous or wild-type for the C allele. The significance of the risk effect was increased when the populations were combined. The MARCO promoter SNP increased the risk for severe RSV disease by approximately 62%. A previous investigation tested whether 9 non-synonymous coding variants or tag SNPs in MARCO associated with chronic obstructive pulmonary disease and found no association, although one of the SNPs (rs41279766) associated with the risk of sepsis (Thomsen et al., 2013). Another study found tag SNPs associated with increased risk of pulmonary tuberculosis, while another was associated with resistance (Bowdish et al., 2013), thus the role of MARCO in this disease is not clear.

A caveat for this investigation is that, although we tested for a number of potential co-factors in our studies, it is possible that other founders or unmeasured environmental factors may contribute to the differences in disease severity that associated with the MARCO rs1318645 G allele. For example, gender has been found to be important in susceptibility to some lung diseases in mouse models (Carey et al., 2007; Cabello et al., 2015). In the present study, we used only male mice. Although we are not aware of sexual dimorphism effects in RSV disease, male and female mice respond differently to influenza A infection (Larcombe et al., 2011). Another important RSV disease phenotype is airways hyperreactivity (Yoo et al., 2013). We chose to focus the present studies on the inflammation and injury responses to RSV infection, but investigations into the genetic basis of susceptibility to RSV-induced airway hyperreactivity may also identify important susceptibility genes. It is also feasible that different strain distribution patterns may be found with a different line of RSV. Lukacs et al. (2006) found that different strains of RSV of the same antigenic subgroup may cause different RSV-induced phenotypes in BALB/cj mice, and it is possible that other strains also respond differently to multiple RSV lines. Finally, it is important to emphasize that, while the rs1318645 SNP associated significantly with RSV disease in these populations, other gene polymorphisms also contribute to susceptibility, including myxovirus (influenza virus) resistance 1 (MX1) (Cienciewicki et al., 2014) and toll-like receptor 4 (TLR4) (Caballero et al., 2015), and interactions between these and other, unknown, genetic contributions remain to be clarified. It is likely that continued investigation of gene candidates identified with GWA for all of the phenotypes in the mouse model will add to our understanding of RSV disease susceptibility.

In conclusion, this translational investigation identified gene candidates, including Marco, for host susceptibility to multiple phenotypes of RSV disease in mice that closely mimic human disease, and a polymorphism in human MARCO associated with increased risk of RSV disease severity in infants. Continued investigation of the genetic basis of differential susceptibility to RSV disease may lead to a panel of informative SNPs to identify children who are at risk for disease severity. In the absence of an RSV vaccine, these individuals could be treated prophylactically with an RSV monoclonal antibody to prevent infection and sequelae associated with severe disease such as childhood asthma.
Funding Sources
Research related to the manuscript was supported by the Intramural Research Program of the National Institutes of Health, NIEHS (Z01 ES 100557).

Conflict of Interest Statement
The authors declare no conflicts of interest or competing financial interests.

Author Contributions

Acknowledgments
Dr. Michael Fessler and Donald Cook at the NIEHS provided excellent critical review of the manuscript.

Appendix A. Supplementary Data
Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ebio.2016.08.011.

References


