

Identification of *IFRD1* as a modifier gene for cystic fibrosis lung disease

YuanYuan Gu¹, Isaac T. W. Harley¹, Lindsay B. Henderson⁵, Bruce J. Aronow², Ilja Vietor⁷, Lukas A. Huber⁷, John B. Harley⁸, Jeffrey R. Kilpatrick⁸, Carl D. Langefeld⁹, Adrienne H. Williams⁹, Anil G. Jegga², Jing Chen², Marsha Wills-Karp³, S. Hasan Arshad¹⁰, Susan L. Ewart¹¹, Chloe L. Thio⁶, Leah M. Flick¹, Marie-Dominique Filippi⁴, H. Leighton Grimes³, Mitchell L. Drumm¹², Garry R. Cutting⁵, Michael R. Knowles¹³ & Christopher L. Karp¹

Lung disease is the major cause of morbidity and mortality in cystic fibrosis, an autosomal recessive disease caused by mutations in *CFTR*. In cystic fibrosis, chronic infection and dysregulated neutrophilic inflammation lead to progressive airway destruction. The severity of cystic fibrosis lung disease has considerable heritability, independent of *CFTR* genotype¹. To identify genetic modifiers, here we performed a genome-wide single nucleotide polymorphism scan in one cohort of cystic fibrosis patients, replicating top candidates in an independent cohort. This approach identified *IFRD1* as a modifier of cystic fibrosis lung disease severity. *IFRD1* is a histone-deacetylase-dependent transcriptional co-regulator expressed during terminal neutrophil differentiation. Neutrophils, but not macrophages, from *Ifrd1*-deficient mice showed blunted effector function, associated with decreased NF- κ B p65 transactivation. *In vivo*, *IFRD1* deficiency caused delayed bacterial clearance from the airway, but also less inflammation and disease—a phenotype primarily dependent on haematopoietic cell expression, or lack of expression, of *IFRD1*. In humans, *IFRD1* polymorphisms were significantly associated with variation in neutrophil effector function. These data indicate that *IFRD1* modulates the pathogenesis of cystic fibrosis lung disease through the regulation of neutrophil effector function.

Attention to the role of *CFTR* in regulating epithelial ion transport has failed to illuminate the path from gene to pathogenesis in cystic fibrosis (CF) lung disease. In CF, colonization and infection (paradigmatically with *Pseudomonas aeruginosa*) is associated with neutrophilic inflammation, the end result being progressive airway destruction². This inflammatory response is out of proportion to inciting infectious stimuli³, which may well be due to the compromise of lipid mediator pathways driving resolution of neutrophilic inflammation⁴.

To identify genetic modifiers of CF lung disease severity, we performed a genome-wide single nucleotide polymorphism (SNP) scan in the Genetic Modifier Study Group (GMSG) cohort, followed by validation of top candidates in the US CF Twin and Sib Study (CFTSS) cohort. The former enrolled Δ F508 *CFTR* homozygotes with extremes of lung function for age, for case-control association approaches to modifier gene identification⁵. The latter enrolled CF-affected twins and siblings with any *CFTR* genotype, and their parents, for transmission-based approaches¹. Severity status was quantified using lung function measures highly correlated with survival in CF⁶.

Genome-wide SNP analysis was performed using Affymetrix 100K microarrays in 320 patients from the GMSG cohort⁵: 160 with severe and 160 with mild lung disease, with DNA from 20 subjects pooled per microarray. To assess the robustness of the pooling approach, we first compared gene chip estimates of allele frequencies in pooled samples with individually genotyped frequencies in a subset of 93 SNPs. A high degree of correlation ($r^2 = 0.88$) was found. Second, the comparison of genome-wide allele frequencies in these CF patients with those from a similar pooled genome-wide scan in asthma patients and controls (from an Isle of Wight birth cohort study⁷) unambiguously identified *CFTR* as the disease-causing locus in CF. Of the top-ranked polymorphisms distinguishing the two cohorts, in terms of statistically significant differences in allele frequency, 34 out of 38 were clustered on chromosome 7, centred around *CFTR*, with a median uncorrected *P* value of 3×10^{-8} (Supplementary Fig. 1). Several of these SNPs would not pass Bonferroni correction for multiple testing, given the $>100K$ SNPs on the microarrays—despite the known biological significance of *CFTR*. Thus, in addition to generating false positive results, correction for multiple testing in genome-wide SNP scans can also generate false negative results.

To differentiate true from false association (or false lack of association), we prioritized follow-up efforts according to hierarchical criteria, focusing, first on regional clusters of SNPs exhibiting different allele frequencies (Supplementary Fig. 2), and second on regions containing genes with biological coherence: regulators of transcription, genes with known function in the immune system or in lung biology, and genes implicated in biological functions that are abnormal in the CF airway (for example, ion channels). The latter criterion was unlikely to facilitate the identification of genes with unknown function or of unexpected pathogenetic pathways. However, we aimed to identify true modifiers among false positives, not to identify all modifiers. This heuristic approach was used to select 6 regions/genes for follow-up study: *IFRD1*, *CEBPA/CEBPG*, *CHI3L2*, *C6*, *SLC4A3* and *ABCA1*.

The top-ranked locus in terms of clustering was *IFRD1*, 5 megabases (Mb) away from *CFTR*, the 3' region of which contained a cluster of nine SNPs (Supplementary Table 1 and Supplementary Fig. 2b) with significantly different allele frequencies between patient groups. Biology is addressed later. To confirm the genotyping in the pooled scan, as well as define which SNPs to pick for replication purposes, SNPs in *IFRD1* reaching significance in the pooling experiment, along with tagging

¹Division of Molecular Immunology, ²Division of Biomedical Informatics, ³Division of Immunobiology, and ⁴Division of Experimental Hematology & Cancer Biology, Cincinnati Children's Hospital Research Foundation and the University of Cincinnati College of Medicine, Cincinnati, Ohio 45229, USA. ⁵McKusick-Nathans Institute of Genetic Medicine, and ⁶Division of Infectious Diseases, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA. ⁷Biocenter, Division of Cell Biology, Innsbruck Medical University, Innsbruck A-6020, Austria. ⁸Arthritis & Immunology Program, Oklahoma Medical Research Foundation, and JK Autoimmunity Inc., Oklahoma City, Oklahoma 73104, USA. ⁹Department of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157, USA. ¹⁰The David Hide Asthma and Allergy Research Centre, Newport, Isle of Wight, PO30 5TG, UK. ¹¹Department of Large Animal Clinical Sciences, College of Veterinary Medicine, Michigan State University, East Lansing, Michigan 48824, USA. ¹²Departments of Pediatrics and Genetics, Case Western Reserve University, Cleveland, Ohio 44106, USA. ¹³Cystic Fibrosis–Pulmonary Research and Treatment Center, University of North Carolina, Chapel Hill, North Carolina 27599, USA.

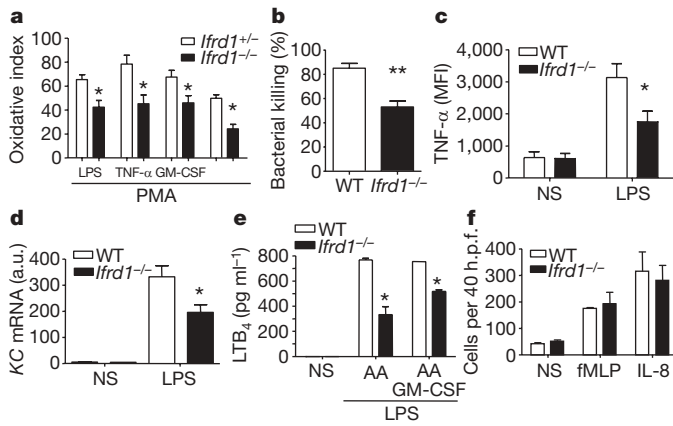


Figure 1 | IFRD1-deficient neutrophils exhibit blunted effector functions. **a**, Oxidative burst in peripheral blood neutrophils from *Ifrd1*^{-/-} mice and heterozygote littermate controls. **b**, Bacterial killing of *P. aeruginosa* by bone marrow neutrophils, calculated as: [(control c.f.u. – experimental c.f.u.)/control c.f.u.] × 100, in which c.f.u. denotes colony-forming units. Means and s.e.m. are shown, of six different mice per genotype, tested in duplicate. Similar results were seen with whole blood neutrophil killing assays in three separate experiments. **c**, Intracellular TNF- α expression by neutrophils, quantified by flow cytometry in LPS-stimulated whole blood isolated from *Ifrd1*^{-/-} mice and wild-type controls. **d**, KC mRNA expression, quantified by qRT-PCR in bone marrow neutrophils from wild-type or *Ifrd1*^{-/-} mice. Means and s.e.m. are shown, of six different mice per genotype, tested in duplicate. a.u., arbitrary units. **e**, LTB₄, quantified by ELISA in supernatants from bone marrow neutrophils from wild-type or *Ifrd1*^{-/-} mice incubated with arachidonic acid (AA), and stimulated with LPS in the presence or absence of GM-CSF. **f**, Quantification of chemotaxis by bone marrow neutrophils from wild type or *Ifrd1*^{-/-} mice in response to formyl-methionyl-leucyl-phenylalanine (fMLP) (1 μ M) or IL-8 (1 μ g ml⁻¹). h.p.f., high-power fields. Means and s.e.m. are shown, representative of three separate experiments for **a**, **c** and **e**. **P* < 0.05, ***P* < 0.01. NS, mock stimulation; PMA, phorbol-12-myristate-13-acetate.

SNPs throughout and flanking the region of the effect, were individually genotyped in the wider GMSG cohort. Although no association signal was observed for *IFRD1*-flanking markers, retention of association signal for a cluster of SNPs on the haplotype block containing the 3' *IFRD1* exons (Supplementary Tables 2, 3 and Supplementary Fig. 3) led us to pursue replication in a separate population.

Three *IFRD1* SNPs (rs7817, rs3807213 and rs6968084) that linkage disequilibrium analysis suggested captured the bulk of the variation observed at this locus were genotyped in patients in the CFTSS cohort (Supplementary Table 4). Notably, the family-based association test demonstrated significant association between the rs7817 polymorphism and both cross-sectional and longitudinal measures of lung function (Table 1a). The other two SNPs showed trends towards significance (rs3807213, *P* = 0.080; rs6968084, *P* = 0.082; for longitudinal and cross-sectional measures of lung function, respectively). A second, complementary method (quantitative transmission

disequilibrium test) verified the result derived for rs7817 (Table 1b). Notably, both methods showed that the heterozygote genotype ('CT') was associated with lower lung function than either homozygote ('CC' or 'TT'; data not shown). However, other genotype models (additive, recessive and dominant) could not be conclusively excluded. It was important to exclude linkage with *CFTR* alleles as the cause of the observed association between *IFRD1* SNPs and CF lung function. No correlation was detected between genotypes composed of *IFRD1* SNPs and the presence of 0, 1 or 2 copies of the common mutation Δ F508, or when *CFTR* mutations were grouped according to their association with exocrine pancreatic status. Furthermore, there was no evidence of linkage between the SNPs and the pulmonary phenotypes⁸, important for validating the association model used (a test of association in the absence, as opposed to the presence, of linkage). These data indicate that *IFRD1* polymorphisms contribute to lung function variation in CF independent of *CFTR*.

IFRD1 acts in a histone deacetylase (HDAC)-dependent manner to mediate transcriptional co-repression and co-activation⁹. Expression and genetic deletion studies have implicated *IFRD1* in cell differentiation and stress responses¹⁰. Available databases suggested highest expression in human blood cells¹¹. Flow cytometric analysis of such cells demonstrated greatest expression in neutrophils (Supplementary Fig. 4a). Similarly, quantitative PCR with reverse transcription (qRT-PCR) analysis of cells relevant to the CF airway indicated particular enrichment of expression in neutrophils (Supplementary Fig. 4b). Terminal differentiation of human and mouse neutrophils was associated with robust upregulation of *IFRD1* expression (Supplementary Fig. 5), something that was mirrored in expression databases¹². Neutrophilic differentiation of HL-60 cells also led to upregulation of *IFRD1* expression (Supplementary Fig. 6a), and short interfering RNA (siRNA)-mediated knockdown of *IFRD1* in such cells blunted oxidative burst capacity (Supplementary Fig. 6b, c) without altering visual morphology or the surface expression of CD11b (encoded by *ITGAM*) (data not shown).

No alteration in peripheral blood neutrophil count, morphology, or CD11b and Gr-1 (encoded by *Ly6g*) expression were observed in *Ifrd1*^{-/-} mice¹³ (data not shown). However, neutrophils from *Ifrd1*^{-/-} mice showed significant impairment of specific effector functions—including oxidative burst, bacterial killing, TNF- α , KC (CXCL1) and leukotriene B₄ (LTB₄) production, but not chemotaxis—compared with wild type and/or heterozygote littermate controls (Fig. 1). *In vivo* stimulation led to similar results: after intratracheal lipopolysaccharide (LPS) stimulation, *Ifrd1*^{-/-} mice exhibited significantly less TNF- α production on a per cell basis in neutrophils, but not macrophages, compared with wild-type controls (data not shown). No differences in airway neutrophil numbers or apoptosis were seen in these studies (data not shown). These functional effects showed specificity among myeloid cells; peritoneal macrophages from *Ifrd1*^{-/-} mice showed normal oxidative burst capacity and LPS-driven TNF- α production (Supplementary Fig. 7). Thus, *IFRD1* has an important role in the regulation of neutrophil effector function.

We subsequently analysed the role of *IFRD1* in modulating airway infection with *P. aeruginosa*. Genetic deficiency of *Ifrd1* was associated

Table 1 | Transmission analysis of *IFRD1* rs7817 in CF

a Transmission analysis using family-based association testing (PBAT module, Golden Helix)						
SNP	Alleles	Genetic model	<i>n</i>	<i>P</i> value	Clinical outcome measure	Effect
rs7817	C/T	Heterozygous distortion	248	0.004	Cross-sectional lung function*	–
rs7817	C/T	Heterozygous distortion	186	0.016	Longitudinal lung function†	–
b Transmission analysis using quantitative transmission disequilibrium testing (QTDT ²⁴)						
SNP	<i>F</i>	<i>N</i>	<i>P</i> value	Clinical outcome measure		
rs7817	4.10	467	0.0168‡	Cross-sectional lung function*		
rs7817	4.00	314	0.0187‡	Longitudinal lung function†		

* BayesFEV₁%pred@20yrs, estimated FEV₁ percentage-predicted values at age 20 years, as described¹.

† MaxFEV₁/CF%, maximum CF-specific percentile for FEV₁ in patient's most recent year of available data, as described¹.

‡ *P* < 0.05 after Bonferroni correction for two tests.

Minor allele frequencies: rs7817, 0.48 C; rs3807213, 0.40 C; rs6968084, 0.14 T

F, QTDT test statistic²⁴; *n*, number of informative families; *N*, number of informative individuals; Effect, phenotypic effect associated with the over-transmitted genotype (+, better function).

with significantly slower bacterial clearance (Fig. 2a). Notably, however, *Ifrd1*^{-/-} mice also had significantly ameliorated disease, with less weight loss, and less airway and systemic inflammation (Fig. 2b–f). To define whether this was due to haematopoietic cell IFRD1 expression, C57BL/6 (CD45.1) mice were lethally irradiated and reconstituted with bone marrow cells from CD45.2-expressing wild-type or *Ifrd1*^{-/-} mice. No differences in bone marrow reconstitution efficiency were observed (Supplementary Fig. 8a). Indeed, competitive reconstitution assays formally demonstrated the lack of a role for IFRD1 in early neutrophil development (Supplementary Fig. 8b). Wild-type mice reconstituted with IFRD1-deficient bone marrow cells mirrored the phenotype of *Ifrd1*^{-/-} mice during *P. aeruginosa* infection—with less efficient bacterial clearance, but less inflammation and disease (Supplementary Fig. 9). When reciprocal bone marrow transfers were performed (reconstituting lethally irradiated CD45.2-expressing wild type and *Ifrd1*^{-/-} mice with bone marrow cells from wild-type C57BL/6 (CD45.1) mice), no such differences in bacterial burden, inflammation or disease course were seen (Supplementary Fig. 10). Thus, IFRD1 modulation of the airway response to *P. aeruginosa* infection is largely dependent on haematopoietic cell expression of IFRD1.

Airway challenge with LPS was similarly associated with increased TNF- α and KC in bronchoalveolar lavage (BAL) fluid from wild-type, compared with *Ifrd1*^{-/-} mice (Supplementary Fig. 11a, b). *In vivo* HDAC inhibition blunted LPS-driven airway TNF- α and KC production, specifically in wild-type, but not in *Ifrd1*^{-/-} mice (Supplementary Fig. 11a, b). Furthermore, bone marrow transfer experiments showed that the effects of HDAC inhibition on LPS-driven airway TNF- α production—in BAL and by airway neutrophils (Supplementary Fig. 11c, d), but not by airway macrophages (data not shown)—was dependent on haematopoietic cell IFRD1 expression.

The effector functions blunted in the absence of IFRD1 are dependent on NF- κ B p65 (also known as RelA)^{14,15}. *Ifrd1*^{-/-} mice showed significantly decreased LPS-stimulated neutrophil NF- κ B p65 transactivation, compared with littermate controls (Supplementary Fig. 12a). As IFRD1, NF- κ B p65 and HDAC1 were co-immunoprecipitable in neutrophil nuclear extracts (Supplementary Fig. 12b), it seems probable that IFRD1 mediates its effects on neutrophils, at least in part, by direct interactions with NF- κ B. Although our data are compatible with IFRD1 being a co-activator of transcriptional activity or a co-repressor of an

inhibitor of transcriptional activity, the HDAC inhibition experiments suggest the latter. Co-immunoprecipitation analysis suggests the possibility of HDAC-mediated co-repression of an NF- κ B-driven transcriptional inhibitor of NF- κ B transactivation.

Although we may have not identified the causal variant(s)¹⁶, cogent hypotheses exist for how the identified SNPs may alter *IFRD1* expression and/or function (Supplementary Discussion). To test directly the association of *IFRD1* polymorphisms with neutrophil effector function, we studied neutrophils from healthy subjects. Notably, analysis of human peripheral blood neutrophils showed a significant association of *IFRD1* polymorphisms with quantitative measures of neutrophil effector function (Fig. 3). Taken together, these data suggest that IFRD1 modulates the course of CF airway disease through the regulation of neutrophil effector function.

Biology is rarely simple, however. The predictive value of mouse knockout models for the more subtle biological differences that probably result from human allelic polymorphisms, as well as the power of the functional data from human neutrophils to account for the effects of these polymorphisms on lung function over time, should not be overstated. It remains possible that neutrophils are not the only cells influenced by *IFRD1* polymorphisms in a fashion relevant to cystic fibrosis. There may, for example, be ways in which respiratory epithelia and neutrophils interact through *IFRD1* polymorphisms to modulate cystic fibrosis lung disease. On the whole, *Cftr* knockout and mutant mouse models have been disappointing; despite recapitulation of gut pathology, pulmonary phenotypes have been subtle. Whether this relates to different lung architecture in mice and humans, the influence of other genes in the mouse strains used, or stronger baseline immune counter-regulation in the mouse lung remains unclear. There is thus an essential problem with using

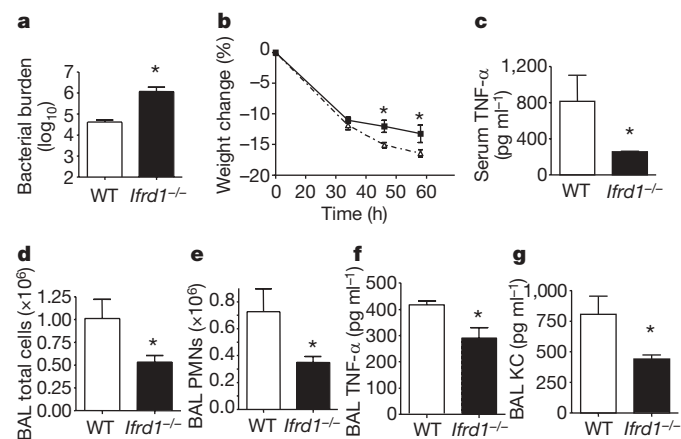


Figure 2 | Genetic deficiency of *Ifrd1* is associated with delayed bacterial clearance, but decreased neutrophilic inflammation and ameliorated disease, after airway challenge with mucoid *P. aeruginosa*. Wild type (WT) and *Ifrd1*^{-/-} mice were challenged intratracheally with *P. aeruginosa* (FRD1 strain), and analysed 48 h later. **a**, Lung bacterial burden is shown. **b**, Weight change is shown. Solid line, *Ifrd1*^{-/-}; dashed line, wild-type. **c**, Serum TNF- α levels. **d–g**, In BAL, the total number of cells (**d**), the number of neutrophils (**e**), TNF- α levels (**f**) and KC levels (**g**) are shown. No differences in parenchymal accumulation of neutrophils (quantified by the analysis of myeloperoxidase activity in blanching lungs) were observed. PMNs, polymorphonuclear leukocytes. Means and s.e.m. of six mice per group are shown; data are representative of three separate experiments. * $P < 0.05$.

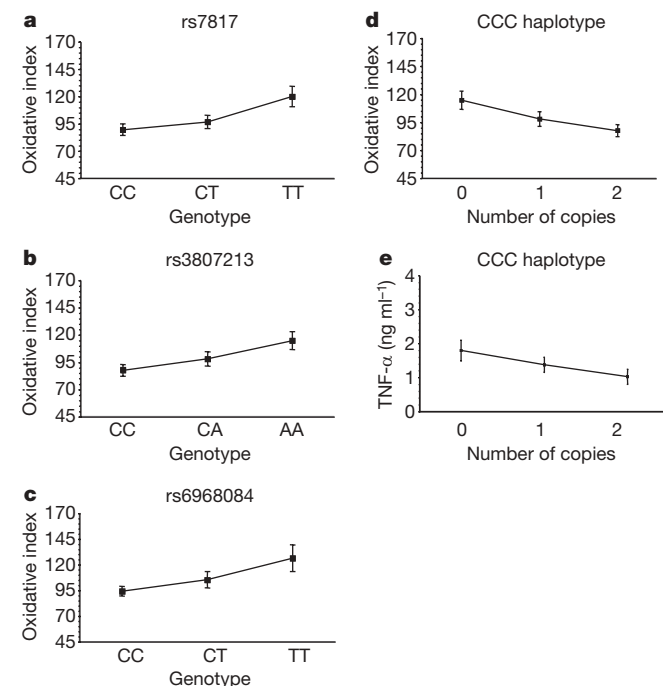


Figure 3 | *IFRD1* polymorphisms are associated with variation in human neutrophil effector function. Oxidative burst capacity following PMA stimulation, and LPS-driven TNF- α secretion, was quantified in neutrophils from healthy donors of self-reported European descent ($N = 36$).

a, Oxidative index for SNP rs7817. CC, $N = 13$; TC, $N = 13$; TT, $N = 10$. $P = 0.005$ (Wald test). **b**, Oxidative index for SNP rs3807213. CC, $N = 12$; AC, $N = 11$; AA, $N = 13$. $P = 0.007$. **c**, Oxidative index for SNP rs6968084. CC, $N = 19$; TC, $N = 15$; TT, $N = 2$. $P = 0.06$. **d**, Oxidative index for three-marker haplotype (rs7817, rs3807213 and rs6968084). 0 ($N = 13$), 1 ($N = 11$) or 2 ($N = 12$) copies of the CCC haplotype. $P = 0.007$. **e**, TNF- α levels, three-marker haplotype (rs7817, rs3807213 and rs6968084). CCC haplotype copy number as in **d**. $P = 0.03$. Data represent means \pm s.e.m.

such models to define whether the absence of *Ifrd1* (or the presence of mutant *Ifrd1* alleles) ameliorates CF lung disease, in the absence of a robust phenotype to ameliorate. In this light, the recent report of pigs with targeted disruption of *CFTR*¹⁷ may point the way to informative models.

Of the other top-ranked genes/regions from the initial scan, *C6*, *SLC4A3* and *ABCA1* did not survive genotyping in the wider GMSG cohort; *CHI3L2* failed replication in the CFTSS cohort (data not shown). However, after refinement of the association signal from the pooled scan in the locus containing *CEBPA* and *CEBPG* (Supplementary Table 5) via individual genotyping in the wider GMSG cohort (Supplementary Tables 6 and 7), replication was pursued in the CFTSS cohort. As shown in Supplementary Table 8, polymorphisms in the 40 kb intergenic region between *CEBPA* and *CEBPG* (genes oriented 5' to each other) were significantly associated with variation in CF lung function. Although there are clearly other possible ways in which these transcription factors may affect CF lung disease severity^{18,19}, *CEBP* α is essential for neutrophil development²⁰, a pathway that *CEBP* γ has also been implicated in²¹.

Despite considerable progress in CF therapy over recent decades, the norm is still an inexorable decline in pulmonary function. The identification of genes modifying CF lung disease, and delineation of the pathogenetic pathways that they influence, holds promise for the development of new therapeutic strategies. The current data suggest probable benefit for therapeutic targeting of neutrophils in this devastating disease. These data also suggest that the *IFRD1*/HDAC axis may provide a tractable therapeutic target in CF, and other diseases in which neutrophils have an important pathogenetic role.

METHODS SUMMARY

Genetic analysis. CF patients were from the GMSG⁵ and CFTSS¹ cohorts. Healthy controls were recruited at CCHMC. Studies were approved by the relevant Institutional Review Boards. Genome-wide analysis was performed using Affymetrix GeneChip 100K Human Mapping microarrays. Other genotyping was performed using Taqman PCR, AcycloPrime-FP SNP PCR, Illumina 610 Quad chips and Illumina SNP beadarray genotyping. Follow-up association analysis (GMSG cohort) was performed using SNP-GWA²². Association and transmission analysis (CFTSS cohort) was done using PEDSTATS v.0.6.6 (<http://www.sph.umich.edu/csg/abecasis/Pedstats>)²³, Interooled Stata 8, quantitative transmission disequilibrium testing (QTDT) v.2.5.0 (<http://www.sph.umich.edu/csg/abecasis/QTDT>)²⁴, and Golden Helix software (<http://www.goldenhelix.com>). Quantitative trait analysis (healthy donor cohort) was performed using PLINK version 1.03 (<http://pngu.mgh.harvard.edu/~purcell/plink/>).

Cellular assays. Surface and intracellular FACS staining was performed as described²⁵. Quantification of mRNA was performed by qRT-PCR, as described²⁵. Oxidative burst capacity was quantified by flow cytometry using the dihydrorhodamine 123 assay²⁶. TNF- α production was quantified by ELISA (BD Pharmingen) or by intracellular staining, as described²⁷. LTB₄ production was quantified by ELISA (Neogen). Bacterial killing was quantified as described²⁸. Neutrophil chemotaxis was quantified as described²⁹. Nuclear NF- κ B p65 DNA-binding activity was quantified using the EZ-Detect Transcription Factor ELISA (Pierce). Co-immunoprecipitation of nuclear proteins was performed using the Nuclear Complex Co-IP kit from Active Motif.

Mouse models. Mice were non-traumatically challenged intratracheally with *P. aeruginosa* (FRD1 strain). Forty-eight hours later, mice were euthanized. BAL cell analysis, lung bacterial burden and myeloperoxidase activity were quantified as described⁴. BAL and serum cytokines were quantified by ELISA (BD Pharmingen, R&D) or by the Cincinnati capture assay³⁰. Standard bone marrow cell transfer techniques were used. Suberoylanilide hydroxamic acid (SAHA), from Cayman, was administered intraperitoneally. Mice were non-traumatically challenged intratracheally with *P. aeruginosa* LPS (Sigma). Animal care was provided in accordance with National Institutes of Health guidelines. Studies were approved by the CCHMC Institutional Animal Care and Use Committee.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 3 December 2008; accepted 20 January 2009.
Published online 25 February 2009.

1. Vanscoy, L. L. *et al.* Heritability of lung disease severity in cystic fibrosis. *Am. J. Respir. Crit. Care Med.* **175**, 1036–1043 (2007).

- Welsh, M. J. *et al.* in *Metabolic and Molecular Basis of Inherited Disease* (eds Scriver, C. R., *et al.*) 5121–5188 (McGraw-Hill, 2001).
- Muhlebach, M. S. & Noah, T. L. Endotoxin activity and inflammatory markers in the airways of young patients with cystic fibrosis. *Am. J. Respir. Crit. Care Med.* **165**, 911–915 (2002).
- Karp, C. L. *et al.* Defective lipoxin-mediated anti-inflammatory activity in the cystic fibrosis airway. *Nature Immunol.* **5**, 388–392 (2004).
- Drumm, M. L. *et al.* Genetic modifiers of lung disease in cystic fibrosis. *N. Engl. J. Med.* **353**, 1443–1453 (2005).
- Schluchter, M. D. *et al.* Classifying severity of cystic fibrosis lung disease using longitudinal pulmonary function data. *Am. J. Respir. Crit. Care Med.* **174**, 780–786 (2006).
- Kurukulaaratchy, R. J. *et al.* Characterisation of atopic and non-atopic wheeze in 10 year old children. *Thorax* **59**, 563–568 (2004).
- Abecasis, G. R. *et al.* Merlin–rapid analysis of dense genetic maps using sparse gene flow trees. *Nature Genet.* **30**, 97–101 (2002).
- Vietor, I. *et al.* TIS7 interacts with the mammalian SIN3 histone deacetylase complex in epithelial cells. *EMBO J.* **21**, 4621–4631 (2002).
- Vietor, I. & Huber, L. A. Role of TIS7 family of transcriptional regulators in differentiation and regeneration. *Differentiation* **75**, 891–897 (2007).
- GNF SymAtlas. (<http://symatlas.gnf.org/SymAtlas>) (2008).
- Theilgaard-Monch, K. *et al.* The transcriptional program of terminal granulocytic differentiation. *Blood* **105**, 1785–1796 (2005).
- Vadivelu, S. K. *et al.* Muscle regeneration and myogenic differentiation defects in mice lacking TIS7. *Mol. Cell. Biol.* **24**, 3514–3525 (2004).
- Anrather, J. *et al.* NF- κ B regulates phagocytic NADPH oxidase by inducing the expression of gp91^{phox}. *J. Biol. Chem.* **281**, 5657–5667 (2006).
- Riedemann, N. C. *et al.* Regulation by C5a of neutrophil activation during sepsis. *Immunity* **19**, 193–202 (2003).
- Donnelly, P. Progress and challenges in genome-wide association studies in humans. *Nature* **456**, 728–731 (2008).
- Rogers, C. S. *et al.* Disruption of the *CFTR* gene produces a model of cystic fibrosis in newborn pigs. *Science* **321**, 1837–1841 (2008).
- Martis, P. C. *et al.* C/EBP α is required for lung maturation at birth. *Development* **133**, 1155–1164 (2006).
- Mullins, D. N. *et al.* CEBPG transcription factor correlates with antioxidant and DNA repair genes in normal bronchial epithelial cells but not in individuals with bronchogenic carcinoma. *BMC Cancer* **5**, 141 (2005).
- Zhang, D. E. *et al.* Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein α -deficient mice. *Proc. Natl Acad. Sci. USA* **94**, 569–574 (1997).
- Bjerregaard, M. D. *et al.* The *in vivo* profile of transcription factors during neutrophil differentiation in human bone marrow. *Blood* **101**, 4322–4332 (2003). (<http://www.phs.wfubmc.edu/public/bios/gene/downloads.cfm>).
- Wigginton, J. E. & Abecasis, G. R. PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics* **21**, 3445–3447 (2005).
- Abecasis, G. R. *et al.* A general test of association for quantitative traits in nuclear families. *Am. J. Hum. Genet.* **66**, 279–292 (2000).
- Divanovic, S. *et al.* Negative regulation of Toll-like receptor 4 signaling by the Toll-like receptor homolog RP105. *Nature Immunol.* **6**, 571–578 (2005).
- Richardson, M. P. *et al.* A simple flow cytometry assay using dihydrorhodamine for the measurement of the neutrophil respiratory burst in whole blood: comparison with the quantitative nitrobluetetrazolium test. *J. Immunol. Methods* **219**, 187–193 (1998).
- Atabani, S. F. *et al.* Natural measles causes prolonged suppression of interleukin-12 production. *J. Infect. Dis.* **184**, 1–9 (2001).
- Ellson, C. D. *et al.* Neutrophils from *p40^{phox}* mice exhibit severe defects in NADPH oxidase regulation and oxidant-dependent bacterial killing. *J. Exp. Med.* **203**, 1927–1937 (2006).
- Filippi, M. D. *et al.* Localization of Rac2 via the C terminus and aspartic acid 150 specifies superoxide generation, actin polarity and chemotaxis in neutrophils. *Nature Immunol.* **5**, 744–751 (2004).
- Finkelman, F. D. & Morris, S. C. Development of an assay to measure *in vivo* cytokine production in the mouse. *Int. Immunol.* **11**, 1811–1818 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was funded by grants from the National Cystic Fibrosis Foundation (C.L.K., M.L.D., G.R.C. and M.R.K.), the National Heart Lung and Blood Institute (G.R.C., M.R.K. and C.L.K.), the Wake Forest University Health Sciences Center for Public Health Genomics (C.D.L. and A.H.W.) and the Austrian Science Fund (I.V. and L.A.H.). We thank R. Pace and J. Yeatts for technical assistance, and D. Hassett for the *P. aeruginosa* FRD1 strain.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.L.K. (chris.karp@chmcc.org).

METHODS

Cohorts. The GMSG cohort consists of CF patients homozygous for $\Delta F508$ *CFTR* whose longitudinal FEV₁ measurements were in the highest or lowest quartile for age among $\Delta F508$ homozygotes. Enrolment criteria, data collection and genotyping have been described⁵. The study was approved by the institutional review boards (IRB) of all participating institutions. Patients and parents of minors provided written informed consent.

CF twins and siblings ($N = 1,118$) and their parents from 619 families were recruited by the CFTSS as previously described¹. Twenty-one dizygous and 49 monozygous twin pairs were included. Raw pulmonary function test data, *CFTR* genotypes, and height and weight measurements were obtained from medical records. In some cases in which genotypes were unavailable, *CFTR* exons were sequenced to identify mutations. Written informed consent or assent was obtained from all subjects. FEV₁ was used to derive cross-sectional (MaxFEV₁CF%) and longitudinal (AvgFEV₁CF% and EstFEV₁%pred) measures, as previously described¹. To include as many subjects as possible and to avoid randomly excluding one member of each pair, lung function measures were averaged for monozygous twin pairs and included in analyses only if the twins' values were within ten percentiles of each other (or ten per-cent-predicted), so as not to double-count genetically identical individuals. For monozygous twin pairs in which only one of the twins had pulmonary data, that twin's data was included.

Healthy controls (inclusion and exclusion criteria: standard for routine blood donation, plus exclusion for use of immunosuppressive medications or non-steroidal anti-inflammatory drugs in the 2 weeks before blood donation for functional assays) were recruited at CCHMC. Blood for neutrophil function studies was obtained from 45 participants, 36 of whom self-reported European ancestry. Owing to the small numbers of non-Europeans and the possibility of confounding due to stratification, analysis was restricted to these 36. Blood samples were blinded to haplotype and genotype status before functional analysis. All participants gave written informed consent; the study was approved by the CCHMC IRB.

Genotyping. Genome-wide analysis, using Affymetrix GeneChip 100K Human Mapping microarrays, was performed in 320 CF patients from the GMSG study⁵: 160 with severe lung disease (lowest quartile of FEV₁ for age); 160 with mild lung disease (highest quartile of FEV₁ for age); 308 who self-reported European ancestry. Each group of 160 was comprised of 80 males and 80 females. Each such group of 80 was divided into groups of 20 (from across the relevant quartile) for pooling purposes. Equimolar amounts of DNA, with an A_{260}/A_{280} ratio of 1.65–2 and an A_{260}/A_{230} ratio of 1.0–2.2, as quantified by NanoDrop spectrophotometry, were combined into pools containing 250 ng DNA. DNA pools were digested with XbaI or HindIII, adaptor-ligated, and PCR-amplified. Samples were separated on 4% agarose gels to ensure DNA fragmentation in the 100–300 bp range. PCR yields were compared between microarray chips to ensure uniformity ($>1,200$ ng μl^{-1} accepted), and PCR products were separated on 2% agarose gels to ensure the proper range of amplified product. GeneChip Genotyping software (v.4.0, Affymetrix, Inc.) was used for relative quality control assessment, detection rates, and allele distributions. Hybridization intensity comparisons of the case and control pools were used to identify significant allele frequency differences for each SNP. A set of 100,198 (out of 111,664) SNPs provided data of sufficient quality on all microarrays.

Taqman PCR genotyping, using assays from ABI, was performed (1) for assessment of the robustness of pooled estimates of allele frequency, by individual sample genotyping of the initial 320 patients in the GMSG cohort; (2) for individual genotyping of samples from 2,194 subjects in the CFTSS cohort (see later); and (3) for genotyping 91 normal healthy controls. AcycloPrime-FP SNP PCR assays (Perkin Elmer) were used to genotype 100 healthy controls (4). Autoclustering algorithms were used ([1] SDS Version 1, [2 and 3] SDS Version 2.3, both from ABI; and [4] FP Caller, from Johns Hopkins) to call SNPs. The call rates were (1) 99.9%, (2) 97.1%, (3) 98.6% and (4) 99.6%, respectively.

Individual sample genotyping of an expanded sample set of patients in the GMSG cohort, comprised of the 320 samples from the pooling experiment, plus a further 485 samples, for a total of 805 samples (261 severe, 541 mild), was performed. To minimize possible issues of population stratification, genetic analysis at this stage was confined to individuals self-reporting European ancestry (779 out of 805, including 241 and 538 patients with severe and mild lung disease, respectively). SNPs that reached significance in the pooling experiment, along with tagging SNPs³¹ throughout the region of the effect, were selected for follow-up. Tag SNPs were chosen on the basis of HapMap data using Tagger (<http://www.broad.mit.edu/mpg/tagger/>) (minor allele frequency threshold = 0.05; pairwise R^2 threshold = 0.8). Furthermore, a reported non-synonymous polymorphism in *IFRD1*, rs11542463, was assayed; it was monomorphic in the GMSG cohort. A final list of tag SNPs was chosen on the basis of predicted assay design scores for the SNP beadarray. Genotyping was done by custom Illumina

GoldenGate assays. An autoclustering algorithm was used on all SNPs. Clusters of SNPs were manually inspected when they had a low call rate (<98.5) or a low clustering score (<0.6). Three individuals (out of an initial 808) with DNA quality or gender reporting problems were excluded. The genotype success rate for each SNP was $>99.6\%$; the overall call rate was 99.95%. Eight samples assayed in duplicate as technical replicates had $>99.93\%$ concordance.

CFTSS subjects were genotyped by two methods, TaqMan (Applied Biosystems) and the Illumina 610 Quad chips. In total, approximately 2,200 individuals were typed for three *IFRD1* SNPs (rs6968084, rs3807213 and rs7817). Of these, 76% were typed by both methods, 20% by TaqMan only, and 4% by Illumina only. The discrepancy rates between the two methods were 0.42%, 0.06% and 0.49% for rs6968084, rs3807213 and rs7817, respectively. No Mendelian errors were detected in families typed using the Illumina platform, whereas five Mendelian errors were detected in families typed by TaqMan. Because the former method appeared to be more reliable, Illumina genotypes were used in cases where calls made by the two methods were different.

Genetic association data analysis. Allele frequencies for Affymetrix data were determined using adjustment factors for pooled samples³². Z^2 P values were used to rank all SNPs. A cluster analysis of Z^2 statistics was performed. Although a previous report has found evidence for minimal stratification in the GMSG cohort³³, the possibility of confounding owing to population substructure in the pooling step was investigated by applying the genomic control method to the Z^2 statistics for pooled DNA³⁴. Direct application of genomic control to pooled data assumes the variance due to pooling has properties delineated by Devlin *et al.*³⁴. Although the pooling experiment did not contain the technical replicates necessary to definitively satisfy these assumptions, genomic control was directly applied to the Z^2 statistics. Using this approach, the inflation factor $\lambda = 1.04$ when estimated using the mean, again suggesting that stratification is minimal in this population³⁵.

Follow-up association analysis in the GMSG cohort was performed using SNP-GWA²². Each SNP was tested for departures from Hardy–Weinberg equilibrium expectations. The additive genetic model test of association was the primary inference. Imputation analysis was performed using the gwas software, impute v0.4.2 and snptest v.1.1.5 (<http://www.stats.ox.ac.uk/~marchini/software/gwas/gwas.html>)³⁵.

For association and transmission analysis in the CFTSS cohort, genotype distributions were tested for Hardy–Weinberg equilibrium using the ‘unrelatedsOnly’ option in PEDSTATS v.0.6.6 (<http://www.sph.umich.edu/csg/abecasis/Pedstats>)²³, which performs an exact test in a subset of unrelated individuals, so as to avoid bias from correlated genotypes within families. Because correlation among sibling marker genotypes may invalidate the results of family-based tests of association in the presence of linkage, linkage between SNPs and pulmonary phenotypes was evaluated using Merlin software (MERLIN v. 1.1.2; <http://www.sph.umich.edu/csg/abecasis/Merlin/>).

Association between three *IFRD1* SNPs (rs6968084, rs3807213 and rs7817) and the three pulmonary phenotypes was analysed using the PBAT module implemented within Golden Helix software (Golden Helix, Inc., Golden Helix PBAT Software <http://www.goldenhelix.com>). Four genetic models were tested: additive, dominant, recessive and heterozygote distortion. The best-associated SNP and phenotypes resulting from PBAT analysis were tested for transmission disequilibrium by a second method, QTDT (<http://www.sph.umich.edu/csg/abecasis/QTDT>)²⁴, using the orthogonal model of association and assuming dominance. Complete and incomplete trios were used in both analyses. General statistics were performed in Intercooled Stata 8 (StataCorp).

Analysis of data on neutrophil oxidative index and TNF- α production in healthy donors was performed using PLINK version 1.03 (<http://pngu.mgh.harvard.edu/~purcell/plink/>) standard quantitative trait association options for genotypes (–assoc) and haplotypes (–hap-assoc) as indicated. Estimated Haplotypes were imputed using the Expectation-Maximization algorithm as implemented in PLINK using the (–hap-phase) option.

Cellular phenotypic and functional assays. Human neutrophils and mononuclear cells were isolated by Ficoll-Hypaque sedimentation, monocytes by leukapheresis and counter-current elutriation. CD34⁺ cells (CCHMC normal donor repository) were differentiated *in vitro* with rG-CSF (50 ng ml⁻¹) plus rSCF (50 ng ml⁻¹, both from Peprotech) for 8 days, followed by rG-CSF alone for 8 days. Primary tracheobronchial cells were collected from bronchial brushings from normal subjects (UCCOM bronchoscopy core). THP-1, BEAS-2B and HL-60 cells were from the American Type Culture Collection (ATCC). HL-60 cells were differentiated with 1.5% DMSO or with retinoic acid (1 $\mu\text{g ml}^{-1}$). Mouse peripheral blood leukocytes were isolated from whole blood after lysis of erythrocytes with ACK lysis buffer (Lonza). Neutrophils were isolated immunomagnetically from mouse bone marrow using Gr-1 beads (Miltenyi), a purification strategy yielding a highly purified population of mature neutrophils, as demonstrated by stained cytopins (data not shown). Mouse peritoneal exudate

macrophages were isolated after thioglycollate elicitation²⁵. Mouse haematopoietic progenitor (Lin⁻ c-kit⁺ sca-1⁺) cells were purified by flow-cytometric sorting²⁹, and differentiated *in vitro* for 11 days with rSCF (100 ng ml⁻¹), rMGDF (100 ng ml⁻¹) and rG-CSF (100 ng ml⁻¹; all from Amgen).

Surface and intracellular FACS staining was performed as described²⁵, using antibodies from Sigma (IFRD1), Molecular Probes (mouse IgG2a), eBioscience (CD11b, Gr-1, CD16, CD3, CD4, CD8 and CD19), an LSRII flow cytometer and FACSDiVa Software (BDPharmingen). Fc-receptor blockade was performed with human AB serum Gemini Bio; human cells) or blocking antibody to CD16 and CD32 (Fc block, eBioscience; mouse cells). Quantification of mRNA was performed by qRT-PCR²⁵, using a LightCycler (Roche) and the following primers: *IFRD1*, 5'-TGCAGCGTTAGCATCTGTTC; *IFRD1*, 3'-ACCAAAGCAAGTTGCAACAAG; *IFRD2*, 5'-TGTTTTTCAGCCGGTCTCTATGG; *IFRD2*, 3'-TGCCTGTCAAGGATGTGGC; ubiquitin, 5'-CACTTGGTCCTGCGCTTGA; ubiquitin, 3'-CAATTGGGAATGCAACAACCTTAT; *KC*, 5'-ACCAAACCGAAGTCAATAGC; *KC*, 3'-TCTCCGTTACTTGGGGACAC. Oxidative burst capacity was quantified by flow cytometry in mouse cells treated with PMA (Sigma), using the dihydrorhodamine 123 assay²⁶. HL-60 cells were mock transfected, or transfected by Nucleofection (Amaxa) with 90 pmol (45 nM) synthetic siRNA against *IFRD1*, or negative control siRNA, and incubated for 48 h, during differentiation to a neutrophil phenotype with DMSO. *IFRD1* siRNA sense/antisense: r(GGUGAGUUCUGAUUUUUAA)dTdT/r(UUAAUUAUCAGAACUCACC)dAdG; control (non-silencing) siRNA sense/antisense: r(UUCUCGAACGUGUCACGU)dTdT/r(ACGUGACACGUUCGGAGAA)dTdT. Oxidative burst was quantified by flow cytometry in human neutrophils by the dihydrorhodamine 123 assay²⁶. Fluorescence was quantified in neutrophils (CD11b⁺ CD15⁺ cells; antibodies from Biolegend; within the granulocyte gate set based on forward and side scatter characteristics) using an LSRII flow cytometer. TNF- α production by cells or in airways was quantified by ELISA (BD Pharmingen) or by intracellular staining²⁷ (anti-TNF- α from eBioscience). KC was quantified by ELISA (R&D) or by qRT-PCR. LTB₄ was quantified by ELISA (Neogen) after stimulation of neutrophils with *P. aeruginosa* LPS or GM-CSF, followed by incubation with arachidonic acid³⁶. Killing of *P. aeruginosa* (FRD1 strain) was quantified as described²⁸. Neutrophil chemotaxis was quantified as described²⁹. Nuclear NF- κ B p65 DNA-binding activity was quantified using the EZ-Detect Transcription Factor ELISA (Pierce). Co-immunoprecipitation of nuclear proteins was performed using the Nuclear Complex Co-IP kit from Active Motif.

Immunoprecipitating and immunoblotting antibodies were from Santa Cruz. Immunoreactive proteins were visualized by ECL (Amersham).

Mouse models. Six-to-eight-week-old *Ifrd1*^{-/-} mice¹³ on a C57BL/6 background (>10 generations) and wild-type controls, were challenged intratracheally (non-traumatically, as described⁴) with 6 \times 10⁶ CFU of *P. aeruginosa* (FRD1 strain). Forty-eight hours after challenge, mice were euthanized, BAL was performed, and serum and lungs were collected. Lung bacterial burden and myeloperoxidase activity were quantified by standard techniques⁴. BAL and serum cytokines were quantified by ELISAs (BD Pharmingen, TNF- α ; R&D, KC). CD45.1⁺ congenic C57BL/6 (B6.SJL-PtprcaPep3b/BoyJ) mice were lethally irradiated, and rescued with 2 \times 10⁶ bone marrow cells from wild-type or *Ifrd1*-deficient C57BL/6 (CD45.2⁺) mice. In these experiments, TNF- α was measured by the Cincinnati capture assay³⁰. Similarly, wild-type and *Ifrd1*^{-/-} C57BL/6 (CD45.2⁺) mice were lethally irradiated, and rescued with 2 \times 10⁶ bone marrow cells from wild-type (CD45.1⁺) mice. Reconstitution was monitored by flow cytometric analysis of peripheral blood cell populations, using monoclonal antibodies to CD45.2, CD11b, Gr-1, TCR, B220 and NK1.1 (eBioscience). Reconstituted mice were challenged with *P. aeruginosa* \geq 2 months after transplantation. To test formally their relative reconstitution ability, bone marrow cells from *Ifrd1*^{-/-} or wild-type mice (both CD45.2⁺) were transplanted into lethally irradiated wild type (CD45.1⁺) recipient mice, along with an equal number (1 \times 10⁶) of competitor bone marrow cells (CD45.1⁺). Mice were treated with SAHA (10 mg kg⁻¹; Cayman) intraperitoneally, followed 1 h later by intratracheal challenge with *P. aeruginosa* LPS (2 mg kg⁻¹; Sigma). Animal care was provided in accordance with National Institutes of Health guidelines. Studies were approved by the CCHMC Institutional Animal Care and Use Committee.

- de Bakker, P. I. *et al.* Transferability of tag SNPs in genetic association studies in multiple populations. *Nature Genet.* **38**, 1298–1303 (2006).
- Yang, H. C. *et al.* New adjustment factors and sample size calculation in a DNA-pooling experiment with preferential amplification. *Genetics* **169**, 399–410 (2005).
- Hillian, A. D. *et al.* Modulation of cystic fibrosis lung disease by variants in interleukin-8. *Genes Immun.* **9**, 501–508 (2008).
- Devlin, B. *et al.* Unbiased methods for population-based association studies. *Genet. Epidemiol.* **21**, 273–284 (2001).
- Marchini, J. *et al.* A new multipoint method for genome-wide association studies by imputation of genotypes. *Nature Genet.* **39**, 906–913 (2007).
- Gronert, K. *et al.* in *Methods in Molecular Biology* Vol. 120 (ed Lianos, E. A.) 119–144 (Humana Press, 1999).