Genetic Regulation of *Zfp30*, CXCL1, and Neutrophilic Inflammation in Murine Lung

Holly Rutledge,* David L. Aylor,† Danielle E. Carpenter,‡ Bailey C. Peck,‡ Peter Chines,‡
Lawrence E. Ostrowski,§ Elissa J. Chesler,** Gary A. Churchill,** Fernando Pardo-Manuel de Villena,*,††
and Samir N. P. Kelada*,§,1

*Department of Genetics, University of North Carolina, Chapel Hill, North Carolina 27599, †Department of Biology, North Carolina State University, Raleigh, North Carolina 27695, ‡National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, §Marsico Lung Institute, University of North Carolina, Chapel Hill, North Carolina 27599, **The Jackson Laboratory, Bar Harbor, Maine 04609, and ‡‡Lineberger Comprehensive Canter Center, University of North Carolina, Chapel Hill, North Carolina 27599

ABSTRACT Allergic asthma is a complex disease characterized in part by granulocytic inflammation of the airways. In addition to eosinophils, neutrophils (PMN) are also present, particularly in cases of severe asthma. We sought to identify the genetic determinants of neutrophilic inflammation in a mouse model of house dust mite (HDM)-induced asthma. We applied an HDM model of allergic asthma to the eight founder strains of the Collaborative Cross (CC) and 151 incipient lines of the CC (preCC). Lung lavage fluid was analyzed for PMN count and the concentration of CXCL1, a hallmark PMN chemokine. PMN and CXCL1 were strongly correlated in preCC mice. We used quantitative trait locus (QTL) mapping to identify three variants affecting PMN, one of which colocalized with a QTL for CXCL1 on chromosome (Chr) 7. We used lung eQTL data to implicate a variant in the gene *Zfp30* in the CXCL1/PMN response. This genetic variant regulates both CXCL1 and PMN by altering *Zfp30* expression, and we model the relationships between the QTL and these three endophenotypes. We show that *Zfp30* is expressed in airway epithelia in the normal mouse lung and that altering *Zfp30* expression *in vitro* affects CXCL1 responses to an immune stimulus. Our results provide strong evidence that *Zfp30* is a novel regulator of neutrophilic airway inflammation.

VARIATION in gene expression is thought to explain a substantial proportion of complex traits, including disease phenotypes (Nica and Dermitzakis 2008; Cookson *et al.* 2009; Nicolae *et al.* 2010). Indeed, one of the main conclusions from genome-wide associations (GWAS) studies of many diseases is that most of the causal variants appear to be regulatory in function, rather than affecting protein-coding gene regions (Hindorff *et al.* 2009; Nicolae *et al.* 2010; Maurano *et al.* 2012). In a simple scenario, a genetic variant causes a change in gene expression (*i.e.*, is an expression quantitative trait locus, eQTL), which in turn influences disease risk

(Emilsson *et al.* 2008). The same types of causal relationships between sequence variation, gene expression, and phenotypes have been documented in model organisms (Schadt *et al.* 2005; Harbison *et al.* 2009; Hageman *et al.* 2011).

We adopted this expression-centric view in an effort to identify new candidate genes for experimental asthma, which is composed of multiple phenotypes including granulocytic (eosinophilic and neutrophilic) inflammation of the airways, mucus metaplasia, and airway hyperresponsiveness. There is ample prior evidence that these traits have a genetic basis (Levitt and Mitzner 1988; Brewer *et al.* 1999; Whitehead *et al.* 2003; Kelada *et al.* 2011) and QTL have been identified for airway hyperresponsiveness on several chromosomes (De Sanctis *et al.* 1995, 1999; Karp *et al.* 2000; Ackerman *et al.* 2005; Camateros *et al.* 2010; Leme *et al.* 2010; Himes *et al.* 2013). However, a systematic evaluation of whether variation in gene expression underlies these traits has not been conducted.

To identify QTL and eQTL that may underlie the phenotypic QTL, we applied a house dust mite (HDM) mouse model

of allergic asthma to incipient lines of the Collaborative Cross (CC), a new mouse systems genetics resource (Collaborative Cross Consortium 2012). The CC is composed of a panel of recombinant inbred lines derived from eight-way crosses using five classical inbred strains (C57BL/6J, 129S1/SvImJ, A/J, NOD/ShiLtJ, NZO/HlLtJ) and three wild-derived inbred strains (WSB/EiJ, PWK/PhJ, and CAST/EiJ) (Collaborative Cross Consortium 2012). The genomes of the CC founder strains (Keane *et al.* 2011; Yalcin *et al.* 2011) and the features of the CC population have been described recently (Collaborative Cross Consortium 2012). Results obtained using incipient CC lines that were partially inbred (referred to as "preCC" mice) have been notable in terms of phenotypic diversity and identification of new QTL (Aylor *et al.* 2011; Bottomly *et al.* 2012; Kelada *et al.* 2012; Ferris *et al.* 2013; Kelada *et al.* 2014).

We measured response to HDM sensitization and challenge in the eight CC founder strains and 151 preCC mice, and the work here is focused primarily on one particular HDM-response phenotype, airway polymorphonuclear neutrophil (PMN) recruitment. PMNs are associated with severe asthma and decrements in lung function in humans (Little *et al.* 2002; Wenzel 2006). In addition to PMN recruitment, we measured an intermediate phenotype, the concentration of CXCL1 in lung lavage fluid, a hallmark PMN chemokine (Bozic *et al.* 1995; Frevert *et al.* 1995). We identified a locus on chromosome (Chr) 7 that controls both phenotypes and integrated lung eQTL data to identify a candidate gene, *Zfp30*, whose expression is controlled by a *cis*-regulatory variant in the same region. We propose that this genetic variant regulates both CXCL1 and PMN by altering *Zfp30* expression and we tested our model with structural equations and i*n vitro* experiments. Taken together, our results provide strong evidence that *Zfp30* is a novel regulator of neutrophilic airway inflammation.

Materials and Methods

Mice

We obtained 151 male preCC mice (ages 10–14 weeks) from Oak Ridge National Laboratory (Chesler *et al.* 2008; Kelada *et al.* 2012; Kelada *et al.* 2014). Each mouse was from an independent CC line that had undergone 5–14 generations of inbreeding. For each of the eight CC founder strains, we obtained five mice from The Jackson Laboratory. All mice were singly housed, with alpha-dri bedding, in the same facility at NIH under normal 12 hr light/dark cycles. All experiments conducted with mice herein were compliant with an IACUC protocol at an AALAC-approved facility.

Phenotyping protocol

Mice were sensitized with 10 µg of the immunodominant allergen from the *Dermatophagoides pteronyssinus* species of HDM, Der p 1, administered by intraperitoneal injection on days 0 and 7 of the protocol. On day 14, mice were challenged

with 50 µg of Der p 1, administered by orotracheal aspiration (Kelada *et al.* 2011; Kelada *et al.* 2014). Mice were phenotyped on day 17 by performing lavage followed by differential cell counts. CXCL1 was measured using a Millipore (Billerica, MA) luminex assay. According to recent sequencing data (Keane *et al.* 2011), no genetic variation in *Cxcl1* would affect detection of CXCL1 by the antibody used. Phenotype data are presented in Supporting Information, Table S1.

Heritability estimates

Analysis of variance was used to test whether phenotypes varied among CC founder strains. From these ANOVA models, we estimated broad-sense heritability (H^2) by calculating the interclass correlation (r_1) and the coefficient of genetic determination (g^2) (Festing 1979),

$$r_1 = (\text{MSB} - \text{MSW})/(\text{MSB} + (n-1)\text{MSW})$$

$$g^2 = (\text{MSB} - \text{MSW})/(\text{MSB} + (2n-1)\text{MSW}),$$

where MSB and MSW are the mean squares between and within, respectively, from the eight-way ANOVA model described above and n is the number of mice per strain ($n = 5/$ strain).

Narrow sense heritability (h^2) was estimated by regression of phenotypes on genotype for the peak locus for the *Zfp30* eQTL on chromosome 7. The allele effects for the eQTL indicated three functional alleles and the genotype data were coded accordingly. Genotypes and haplotype frequencies are provided in Table S2.

Genotyping and QTL mapping

Genotypes for 131of preCC mice used in these experiments have been previously reported (Kelada *et al.* 2012). Twenty additional mice are included in this study. Genotype data are provided in File S1. We genotyped each mouse at the University of North Carolina—Chapel Hill, using one of two Affymetrix SNP arrays (A or B) that were produced in development of the mouse diversity array (MDA) (Yang *et al.* 2009). After removing uninformative and poorly performing SNPs, these arrays contained 181,752 (A array) and 180,976 (B array) SNP assays, and the set of SNPs on each array did not overlap. Most mice (83%) were genotyped on the B array and the remaining were genotyped on the A array. These training arrays were annotated to NCBI build 36 of the mouse genome, but we mapped QTL boundaries to build 37 positions to integrate with other resources. We report NCBI build 37 positions in our results. We estimated the most probable ancestor for each SNP in each mouse using the GAIN algorithm (Liu *et al.* 2010) and reconstructed founder haplotypes on the basis of these results. We then merged the nonoverlapping SNP data sets from arrays A and B by imputing unobserved genotypes on the basis of inferred founder haplotype. For QTL mapping, we used HAPPY (Mott *et al.* 2000) to infer ancestry matrices for an additive genetic model. For efficiency, we then averaged the matrices across
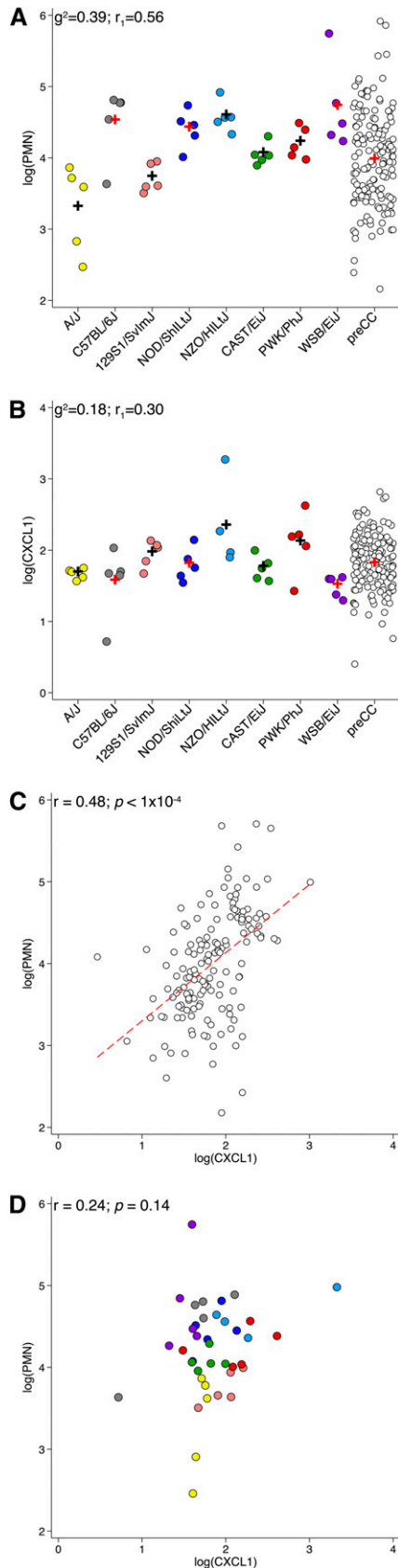
Figure 1 PMN and CXCL1 phenotype distributions and correlations in CC founder strains and preCC mice. (A) Log-transformed PMN counts in

SNPs between which GAIN inferred no recombination in the population, and this reduced the mapping data set to 27,059 intervals. We used BAGPIPE (Valdar *et al.* 2009) to fit a regression model and to report LOD scores. Significance thresholds were determined by permutation. We used the 1.5 LOD drop method to approximate confidence intervals for QTL (Dupuis and Siegmund 1999).

Gene expression analysis

We previously identified eQTL from whole lung RNA using data from a subset of the same mice described here ($n = 131$) (Kelada *et al.* 2014). We used these data to identify genes with eQTL in the QTL region on chromosome 7. We included genes whose eQTL peak resided in the confidence interval for the PMN and CXCL1 QTL, spanning 26–30 Mb, and limited the analysis to local eQTL, which we defined as within 10 Mb upstream or downstream of the gene's position (Kelada *et al.* 2014).

qRT–PCR and allele-specific PCR

To confirm that expression of *Zfp30* measured by microarray was accurate, we performed qRT–PCR. We designed a SYBR green-based qRT–PCR assays for *Zfp30* and two genes used for normalization *Rps29* and *Hprt*. We computed the geometric mean of the Ct of the two normalization genes, and this was used to calculate the delta Ct. All primer pairs given below were designed to avoid SNPs in Collaborative Cross founder strains:*Zfp30*-F 5′-TGTTGGAACAAGGGAAGG AG-3′, *Zfp30*-R 5′-GCCACGCCTTTGAATTCTT-3′; *Rps29*-F 5′-GGAGTCACCCACGGAAGT-3′, *Rps29*-R 5′-TCCATTCAG GTCGCTTAGTC-3′; *Hprt*-F 5′-GCCCTTGACTATAATGAGTA CTTCAGG-3′; and *Hprt*-R 5′-TTCAACTTGCGCTCATCTTA GG-3′.

To validate allele-specific expression of *Zfp30*, we employed pyrosequencing on lung cDNA samples from CC founder lines (C57BL/6J, WSB/EiJ, A/J, and CAST/EiJ and PWK/PhJ, $n = 3$/genotype) and F1 heterozygotes from CAST/EiJ × WSB/EiJ, PWK/PhJ × WSB/EiJ, and CAST/EiJ × 129S1/SvImJ crosses ($n = 6$/cross). We used PCR amplify to a region surrounding a SNP, rs45830063 (C/T), which is located in exon 5 of *Zfp30*, for which CAST/EiJ and PWK/PhJ have the alternate allele (T). PCR primers were 5′-AAGCCCTACGAATGTGGG-3′ and reverse 5′-TCGAAGGTCCGAGCTACAC-3′, and we used another primer, 5′-ATTCGTAGGGCTTCTCGCC-3′, in the sequencing reaction. Results are presented as percentage allele based on peak heights from the pyrograms.

---

lavage fluid. Color coding of strains is as follows: dark gray, C57BL/6J; light blue, NZO/HlLtJ; blue, NOD/ShiLtJ; green, CAST/EiJ; yellow, A/J; salmon, 129S1/SvImJ; red, PWK/PhJ; purple, WSB/EiJ. This color coding is used throughout the manuscript and related papers (Collaborative Cross Consortium 2012). g^2 and r_1 values denote estimates of broad-sense heritability. (B) Log-transformed CXCL1 levels (pg/ml) in lavage fluid. Plus (+) symbols represent the means. Pairwise correlation of PMN and CXCL1 in (C) preCC mice and (D) CC founders. Dashed line represents linear fit for preCC mice.
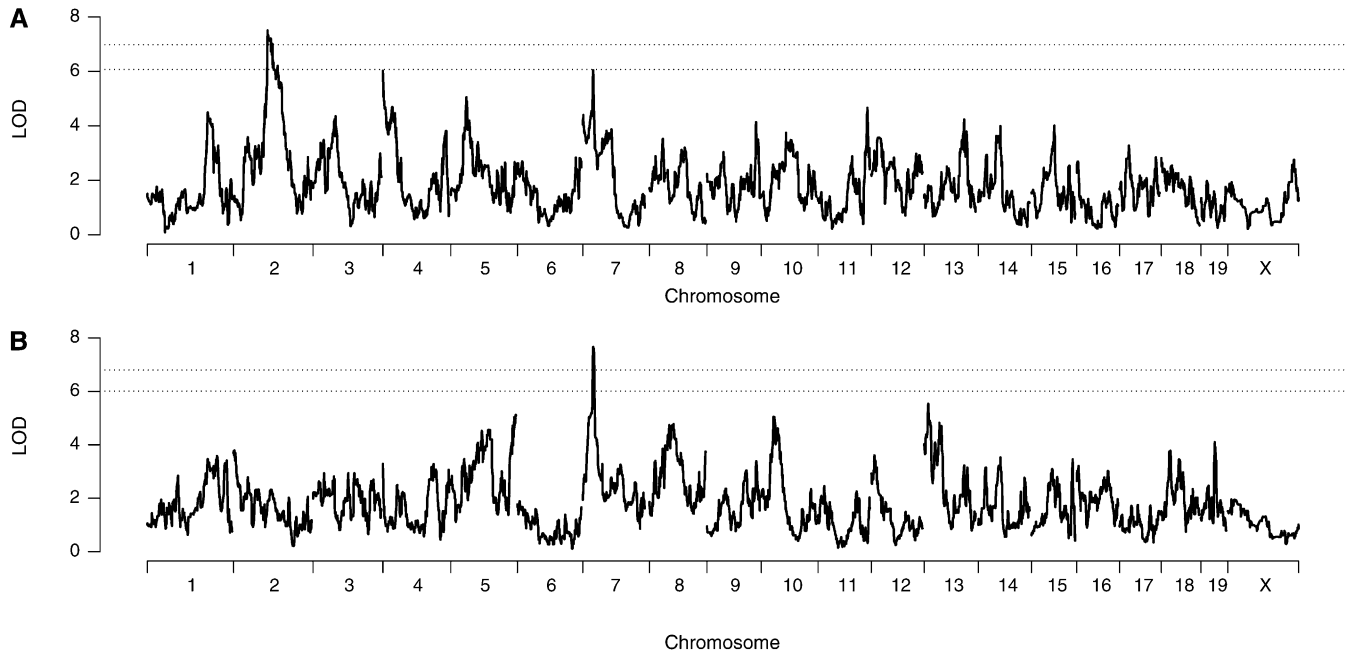
Figure 2 Genome scans of (A) PMN and (B) CXCL1. Dashed lines represent 95 and 80% significance thresholds. Note the colocalized QTL for both traits on chromosome 7.

Conditional correlation analysis

To identify eQTL that may underlie the QTL for PMN and CXCL1, we used conditional correlation. Specifically, we ran Chr 7 QTL scans for PMN, CXCL1, or gene expression, with covariates, either gene expression (GEX) or phenotype (*P*, *i.e.*, PMN or CXCL1), as follows:

Model 1(*P*) : $P = \text{QTL} + \varepsilon$
Model 2(*P*|GEX): $P = \text{QTL} + \text{GEX} + \varepsilon$
Model 3(GEX) : $\text{GEX} = \text{QTL} + \varepsilon$
Model 4(GEX|*P*) : $\text{GEX} = \text{QTL} + P + \varepsilon.$

A large drop in the LOD score (delta LOD) between models 1 and 2 score provides evidence to support a causal relationship (QTL → GEX → *P*). Conversely, a large LOD drop between models 3 and 4 is indicative of gene expression that is reactive to the phenotype, *i.e.*, QTL → *P* → GEX (Li *et al.* 2010).

Structural equation models

We used structural equation models (SEMs) (Shipley 2000) to test whether colocalization of the PMN and CXCL1 QTL on Chr 7 reflects a common cause, namely the *Zfp30* eQTL, or is the result of chance.

The causal model QTL → GEX → *P* is represented as two equations,

$$\text{GEX} = \text{QTL} + e$$

$$P = \text{GEX} + e,$$

and QTL genotype was coded as described in *Heritability estimates*. After fitting models for CXCL1 or PMN individually, we built a series of models that included both CXCL1

and PMN and identified the best-fit models using the Bayesian information criterion (BIC).

Detection of Zfp30 in the lung

In situ hybridization (ISH) was carried out on 18-μm-thick lung tissue sections from C57BL/6J mice using RNA probes generated from a plasmid containing *Zfp30* cDNA sequence corresponding to GenBank accession no. BC049120.
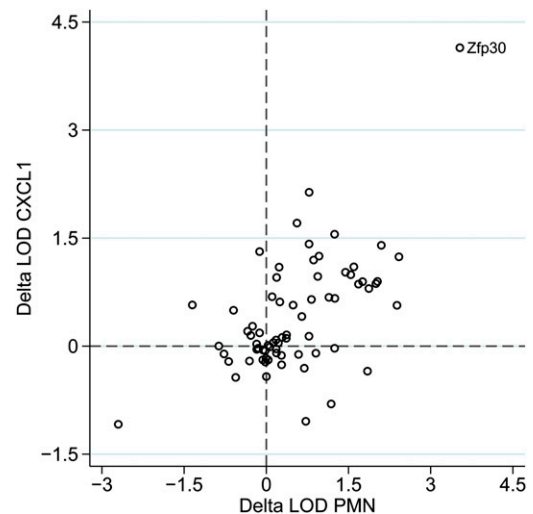

Figure 3 Conditioned QTL scans of CXCL1 and PMN with gene expression identify *Zfp30* as a strong candidate gene. *x*- and *y*-axes represent the change in the LOD score (delta LOD) at the peak locus on Chr 7 when expression of genes with local eQTL is included in the QTL model. *Zfp30* stands out because of the large delta LOD for both CXCL1 and PMN.
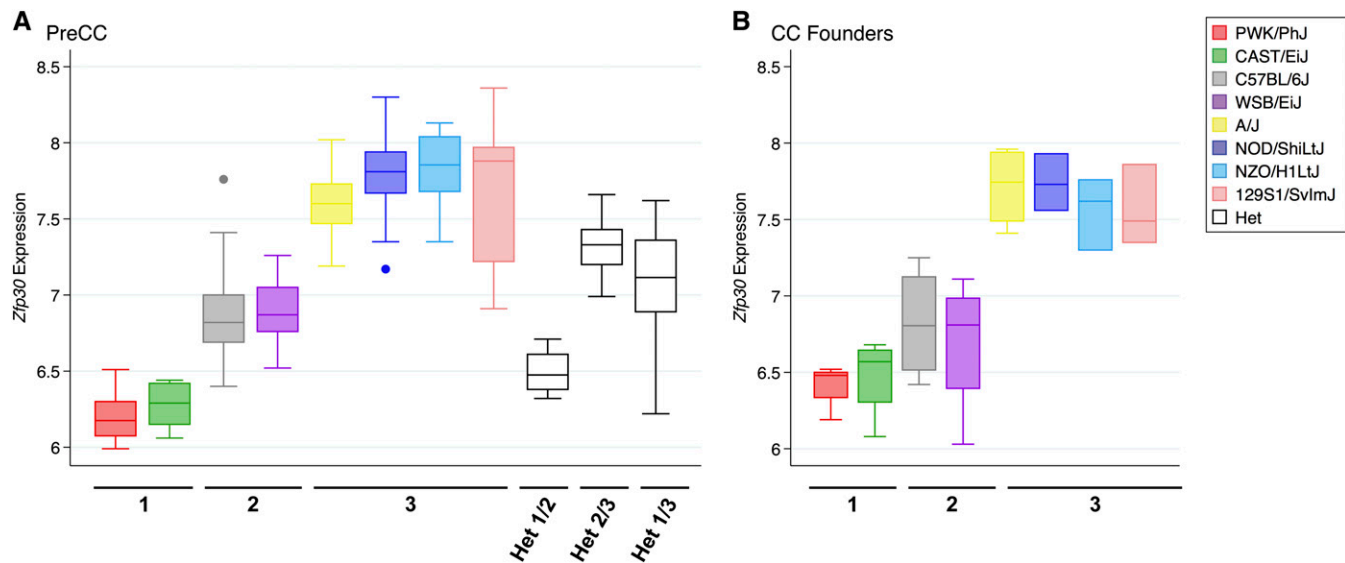
Figure 4 *Zfp30* gene expression among CC founder strains and preCC mice. Array-based measurements of *Zfp30* expression from 131 preCC mice partitioned by their founder haplotype at the peak locus for the eQTL on Chr 7 (A) and three to four mice per CC founder strain (B) are shown. Heterozygous preCC are shown at the far right in A.

In vitro assays with mouse airway epithelial cells

Primary mouse tracheal epithelial cells (MTECs) were grown under air–liquid interface conditions using a previously described protocol (You *et al.* 2002). Cells were harvested at selected time points (days 3, 5, 7, 11, 14, and 17) for RNA analysis. For knockdown experiments, we transfected an immortalized airway epithelial cell line, MLE12, with an siRNA for *Zfp30* designed by Dharmacon: 5′-GCAAA GAUCUUCACGUGUA-3′. Cells were plated and transfected with 10 pmol of siRNA and then treated with lipopolysaccharide (LPS) 48 hr later to stimulate CXCL1 production. Twenty-four hours after LPS treatment, supernatants were collected and CXCL1 was measured using an ELISA assay (R&D Systems). We verified that siRNA treatment reduced *Zfp30* levels using qRT–PCR since no antibody is available to detect ZFP30 protein.

Results

We measured HDM-induced airway PMN recruitment responses among the eight CC founder strains and 151 preCC mice by counting PMN in lung lavage fluid after HDM sensitization and challenge. We also measured concentrations of the well-known PMN chemokine CXCL1 in the same lung lavage fluid. For the sake of simplicity, we refer to these phenotypes as PMN and CXCL1, respectively. CC founder strains exhibited a range of PMN levels (Figure 1A) and the differences among strains provided evidence of heritability (Figure 1A, F-test $P < 1.0 \times 10^{-4}$). Estimates of broad sense heritability for PMN based on interclass correlations (r_1) and the coefficient of genetic determination (g^2) were 0.56 and 0.39, respectively. CXCL1 also differed significantly among CC founders (Figure 1B, F-test $P = 0.01$), although the

between- *vs.* within-strain variance was not as great as for PMN, leading lower estimates of broad sense heritability ($r_1 = 0.30$ and $g^2 = 0.18$).

Among preCC mice, there was substantial variation in both PMN and CXCL1 (Figure 1, A and B). The distributions of both phenotypes exceeded and interpolated the mean values of the eight founder strains, indicative of a polygenic basis of inheritance. Importantly, PMN and CXCL1 were strongly correlated among preCC mice (Figure 1C, Pearson $r = 0.48$, $P < 1.0 \times 10^{-4}$), whereas there was only weak evidence of correlation among CC founder strains (Figure 1D).

We then mapped QTL for PMN and CXCL1 (Figure 2) and identified one QTL for PMN on Chr 2 (Der p 1-induced pulmonary neutrophils, *Dpn1*; peak LOD = 7.54, $p_{adj} < 0.05$; 79.8–98.0 Mb) and two suggestive peaks on Chr 4 (LOD = 6.02, $p_{adj} < 0.2$; 3.3–10.0 Mb), and Chr 7 (LOD = 6.04, $p_{adj} < 0.2$; 24.6–29.8 Mb). For CXCL1, we identified a single QTL on Chr 7 (*Dpc1*; LOD = 7.66, $p_{adj} < 0.05$; 26.4–30.5 Mb), overlapping the PMN peak. Since *Cxcl1* is located on Chr 5, *Dpc1* is *trans*-regulator of CXCL1. The phenotypic correlation and colocalization of Chr 7 QTL for PMN and CXCL1 (*Dpc1*) suggested a variant linked to both phenotypes, and we focused our attention on this region for the rest of investigation.

We asked whether we could prioritize candidate genes for *Dpc1* using gene expression data. We queried a lung eQTL data set we recently generated (Kelada *et al.* 2014) and focused our search for candidate genes on genes with a local eQTL in *Dpc1* region on Chr 7 (24.6–30.5 Mb). Fifty-two genes corresponding to 70 array probes met this criterion (Table S3). We then used a conditional correlation approach to evaluate whether any of these 70 eQTL could explain the
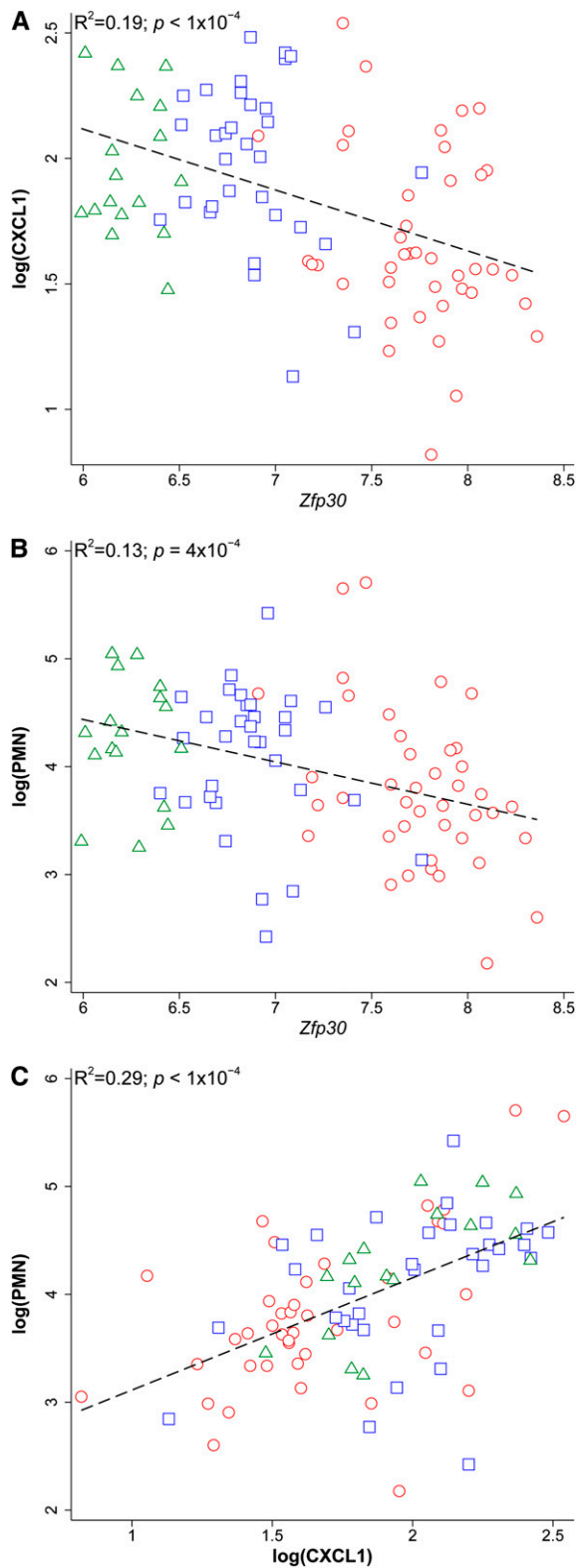
Figure 5 Pairwise correlations of *Zfp30* gene expression, CXCL1, and PMN as a function of Chr 7 locus CC founder genotype. Correlations between traits among homozygous preCC mice are shown, and mice are coded according to eQTL allele effect groups. Red circles represent group 3 homozygotes, blue squares represent group 2 homozygotes, and green triangles represent group 1 homozygotes. (A) CXCL1 *vs. Zfp30* gene expression. (B) PMN *vs. Zfp30* gene expression. (C) PMN *vs.* CXCL1. Dashed lines represent linear fit.

PMN and CXCL1 QTL. We evaluated whether inclusion of GEX into the QTL model for PMN or CXCL1 caused a substantial drop in the LOD score at the peak locus on Chr 7 (models 1 and 2 in *Materials and Methods*), which is indicative of a causal relationship (*i.e.*, QTL → GEX → *P*). We also tested if the gene expression trait is reactive to the phenotype (QTL → *P* → GEX) by comparing LOD scores obtained by from QTL models in which GEX is the phenotype and the covariate is either CXCL1 or PMN (models 3 and 4 in *Materials and Methods*). *Zfp30* was the only gene that had a local eQTL, showed a large drop in the LOD score for both PMN and CXCL1 (Figure 3) in the conditioned model, and did not show evidence of being reactive to PMN or CXCL1 (Table S3). Additionally, out of all 52 genes, the pairwise correlations between gene expression and CXCL1 and PMN were strongest for *Zfp30* ($r = -0.36$, $P < 1.0 \times 10^{-4}$ and $r = -0.39$, $P < 1.0 \times 10^{-4}$, respectively; Table S3).

To characterize the *Zfp30* eQTL, we examined the pattern of gene expression among CC founders and preCC mice (Figure 4). CC founders and preCC mice of the same haplotype had equivalent levels of *Zfp30* gene expression, indicative of a limited effect of genetic background, and were distributed in three groups composed of PWK/PhJ or CAST/EiJ alleles (group 1), C57BL/6J or WSB/EiJ alleles (group 2), and A/J, 129S1/SvImJ, NOD/ShiLtJ, or NZO/HlLtJ alleles (group 3). PreCC mice homozygous for group 3 alleles had 2.8 and 1.8 times higher *Zfp30* expression compared to preCC mice homozygous for group 1 and group 2 alleles, respectively ($P < 1 \times 10^{-4}$ for both comparisons). Likewise, preCC mice homozygous for C57BL/6J or WSB/EiJ alleles had 1.6 times higher expression than preCC mice homozygous for PWK/PhJ or CAST/EiJ alleles had ($P < 1.0 \times 10^{-4}$). Gene expression among preCC mice that were heterozygous at the eQTL peak ($n = 41$; far right side of Figure 4A) was consistent with a semidominant mode of inheritance.

We compared array-based measurements of *Zfp30* expression to qRT–PCR-based measurements among CC founder strains ($n = 3–4$ mice/strain) and found a strong correlation overall ($R^2 = 0.85$, $P = 0.001$; Figure S1). Further, we performed an analysis of allele-specific gene expression using pyrosequencing (Figure S2). We measured the amount of expressed reference allele (C) for rs45830063 among F1 mice ($n = 6–7$/cross) from crosses between CC founder lines from each of the three expression groups. As expected on the basis of array data, F1 heterozygotes of group 2 and 3 mice expressed more of the C allele than the T allele ($P < 0.05$ for both *t*-test comparisons), and we also observed higher C allele expression among group 1/3 heterozygotes compared with group 2/3 heterozygotes ($P < 0.05$).

Collectively, our results indicate that there are three functional alleles in CC founder strains and preCC mice and that *cis*-regulatory variation in *Zfp30* affects gene expression in a semidominant manner. The allele effects pattern is consistent with CC founder haplotypes spanning the majority of the *Zfp30* locus. We examined sequence of CC founder

strains (Keane *et al.* 2011) for *Zfp30*, including 3 kb of upstream sequence and 1.5 kb of downstream sequence (30,566,017–30,581,221 bp, containing 288 SNPs in total) and found that except for regions downstream of exon 4, the phylogeny of the CC founders supports three distinct alleles (Figure S3). More specifically, there are 140 SNPs in this delimited region and 99.3% of them ($n = 139$) are consistent with three alleles and therefore are potential causal variants (Table S4).

These three alleles account for 78% of variation in *Zfp30* expression (*i.e.*, $h^2 = 0.78$), which is consistent with the broad sense heritability we estimated ($H^2 = 0.58$–0.73) based on array data from the CC founder strains. Thus, the eQTL is the source of most variation in *Zfp30* expression. The *Zfp30* allele effects were consistent with those for CXCL1 ($h^2 = 0.14$) and PMN ($h^2 = 0.08$) (Figure 5). The narrow sense heritability estimate for CXCL1 approaches the lower bound of the broad sense heritability estimates (0.18–0.30), indicating that *Dpc1* is the major source of heritable variation of CXCL1 in the context of allergen challenge. PMN ($H^2 = 0.39$–0.56) has a more complex genetic architecture, with contributions from two other loci. This accumulated evidence led us to conclude that *Dpc1* is the *Zfp30* eQTL.

We used SEMs to test if the *Dpc1* effect on CXCL and PMN was independent of *Zfp30* expression. We evaluated the fit of causal (QTL → CXCL1 → PMN), reactive (QTL → PMN → CXCL1), and independent (CXCL1 ← QTL → PMN) models that included *Zfp30* gene expression (Figure 6). We found that two models (Figure 6, E and F) fit the data equally well, and both indicate that (1) the localization of the CXCL1 and PMN QTL are not independent, (2) *Zfp30* expression underlies both CXCL1 and PMN, and (3) *Zfp30* has an effect on PMN independent of CXCL1 (curved arrows). We cannot resolve the directionality of the causal relationship between CXCL1 and PMN. Nonetheless, we conclude that the association between *Dpc1* and both CXCL1 and PMN is secondary to *Dpc1*'s effect on *Zfp30* expression.

Given this preponderance of evidence for *Zfp30* as a novel factor in the airway inflammation, we sought to validate its role using *ex vivo* and *in vitro* approaches. We first confirmed its expression in mouse lung using ISH and found that *Zfp30* was expressed in airway epithelia in the lung and trachea (Figure 7A). Expression outside of the epithelia is also apparent, but expression appears to be most abundant in the epithelia. We examined *Zfp30* gene expression in airway epithelia in more detail using MTECs. *Zfp30* expression increased over the time course of MTEC growth and differentiation in air–liquid interface cultures, peaking at day 17, where expression was 24 times greater than in whole lung tissue (Figure 7B). These results demonstrate that *Zfp30* is expressed in well-differentiated airway epithelia, a cell type of the lung that is known to have an important role in innate and acquired immune responses. Finally, we tested whether altering *Zfp30* expression in immortalized mouse airway epithelial cells (MLE12) *in vitro* affected immune response to
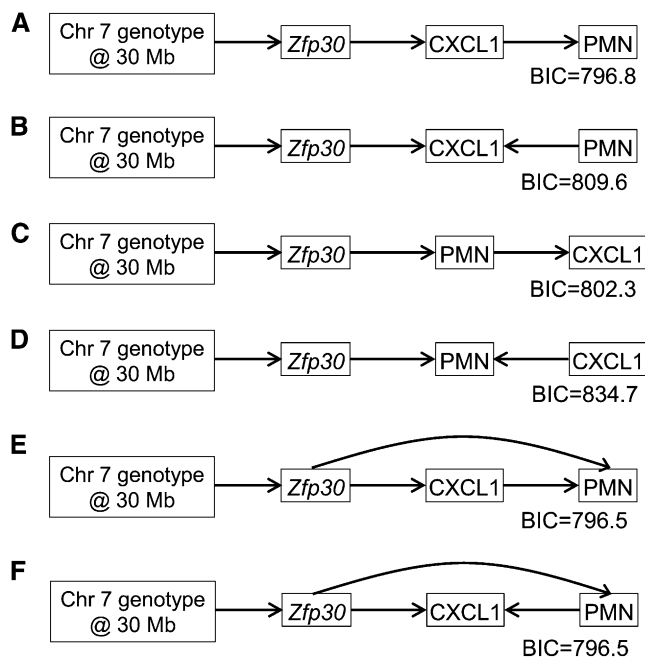

Figure 6 Structural equation models of Chr 7 QTL, *Zfp30* gene expression, CXCL1, and PMN. A series of linear models were built and the results are depicted graphically by nodes, which represent the phenotypes, and directed edges that indicate causal relationships. Models E and F are the best-fit models because they have the lowest Bayesian information criterion (BIC) values. Both models E and F place *Zfp30* upstream of CXCL1 and PMN and also indicate that *Zfp30* may have an effect on PMN independent of CXCL1. Model F also indicates that PMN may feedback on CXCL1.

LPS, a known inducer of CXCL1. As predicted on the basis of our *in vivo* preCC data showing a negative correlation between *Zfp30* expression and CXCL1, knockdown of *Zfp30* led to heightened CXCL1 in supernatants of cells exposed to LPS (Figure 7C, 43% change, $P < 0.05$). These results corroborate our findings from preCC mice and indicate that *Zfp30* is a novel regulator of immune responses in the lung.

Discussion

We used an inbred strain survey, QTL mapping, conditional correlation analysis, and causal inference to identify *Zfp30* as a candidate gene for neutrophilic response to HDM sensitization and challenge. The use of preCC mice in QTL mapping experiments afforded us some notable strengths. First, mapping resolution was on the order of 5 Mb for the phenotype QTL on Chr 7, a small size in comparison to typical F2-intercross studies. Additionally, by employing the hypothesis that functionally similar *Zfp30* alleles are similar in sequence and are descended from a common ancestor (Kelada *et al.* 2012), we identified a set of putative causal variants.

Our SEMs integrated four levels of information [genetic variation, *Zfp30* expression, CXCL1 (protein) expression, and PMN] and identified a novel regulator of a classical chemokine response. The best-fit models predict that decreasing
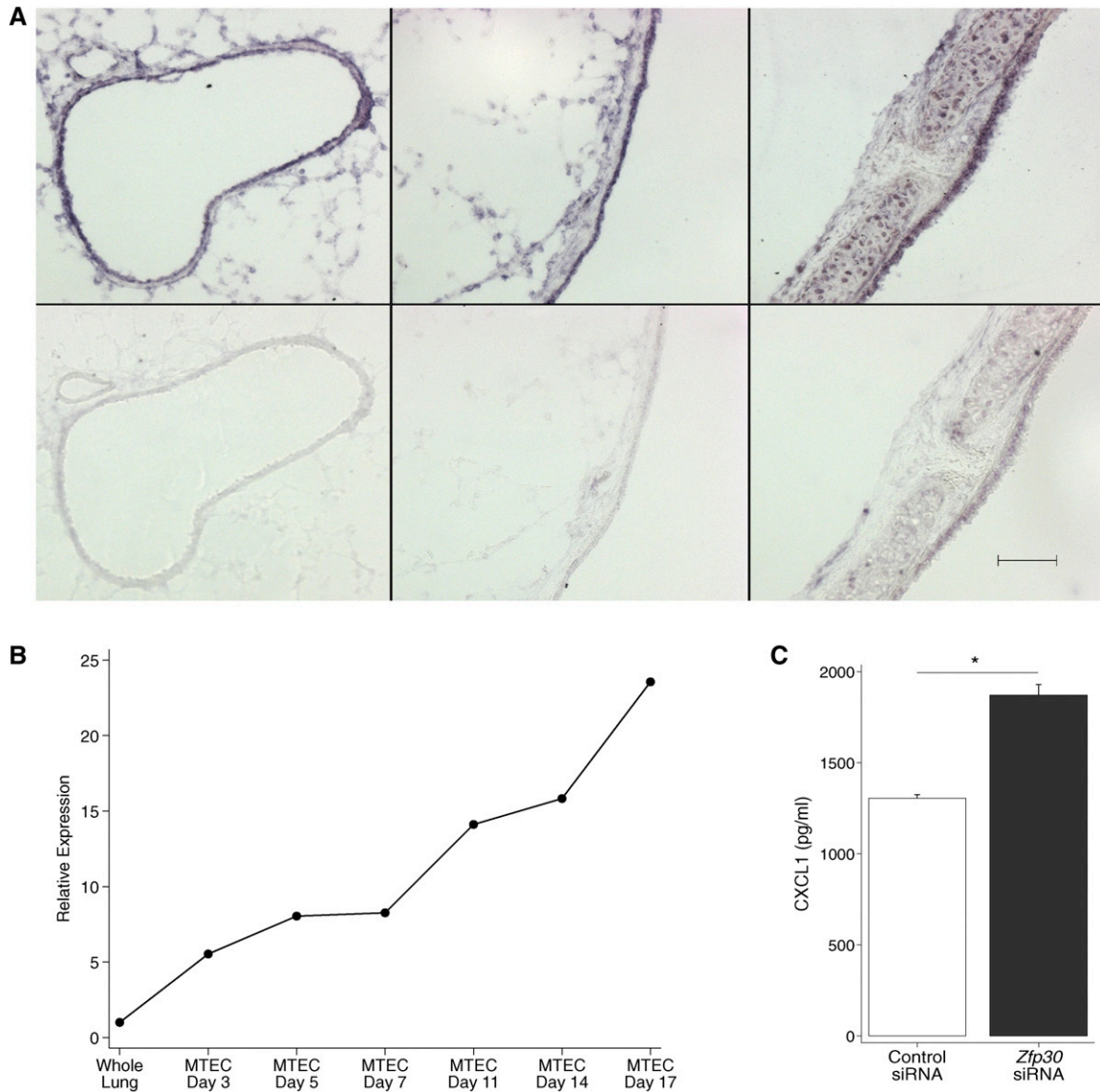
Figure 7 *Zfp30* expression in mouse airways and its effect on CXCL1. (A) *In situ* hybridization using a *Zfp30* RNA probe in the lung and trachea of a C57BL/6J mouse. Representative images are shown. Top: sections probed with anti-sense probe. Bottom: sections probed with the sense (control) probe. Left, medium airway; middle, large airway; right, trachea. Scale bar, 100 μm. (B) Time course of *Zfp30* gene expression in mouse tracheal epithelial cells (MTEC) grown in culture shows a correlation between MTEC differentiation and *Zfp30* expression. Cells are subjected to air–liquid interface on day 14. (C) siRNA-based knockdown of *Zfp30* enhances CXCL1 secretion in response to LPS treatment. MLE12 cells were stimulated with 100 ng/ml LPS after transfection with a control siRNA or siRNA designed for *Zfp30*. (*) $P < 0.05$. Data are representative of three experiments with similar results.

Zfp30 expression will increase CXCL1 and hence PMN. Subsequent *in vitro* assays validated these predictions. The SEMs also indicate that the relationship between CXCL1 and PMN may involve feedback. This is consistent with observations that PMN can secrete CXCL1 in a feed-forward loop (Scapini *et al.* 2000). These results provide a foundation for future experiments that will characterize the functional role of *Zfp30* in neutrophilic airway disease.

PMN are one of the first responders to innate immune stimuli. Their trafficking to the site of injury is controlled by the local gradient of CXC cytokines including CXCL1 and CXCL2 (MIP-2), as well as GM–CSF, which controls PMN re-

lease from the bone marrow. PMN contribute to the clearing of infection through a variety of mechanisms including the release of antibacterial proteins (Kolaczkowska and Kubes 2013). In humans, PMN are associated with asthma severity (Wenzel *et al.* 1997; Jatakanon *et al.* 1999) and some have suggested that the addition of PMN to eosinophilic inflammation may represent a distinct asthma subtype, one that is resistant to corticosteroids (Wenzel *et al.* 1999). Thus PMN are associated with both acute and chronic immune responses. At the present, it is not clear from our results whether the PMN response we observed was due to an acquired immune response to HDM *per se* or are the result of an innate immune

response, but the fact that *Zfp30* expression was primarily localized to airway epithelia (not an immune cell) suggests the latter. Innate immune responses are known to prime acquired immune responses through activation of pattern recognition receptors (Schenten and Medzhitov 2011); hence the role of *Zfp30* in both innate and acquired immune responses should be investigated further. It is worthwhile noting that two previous studies identified QTL for innate immune response phenotypes at approximately the same location on Chr 7. Using A/J × C57BL/6J (AxB) recombinant inbred strains, Matesic *et al.* (1999) found that variation at this locus modifies LPS-induced splenocyte proliferation responses. Likewise, Denny *et al.* (2003) found that this region controls survival and bacterial load after pulmonary infection with *Streptococcus pneumoniae* in an F2 intercross between BALB/cO1aHsd and CBA/CaO1aHsd strains. In both of these studies, the strains used had contrasting alleles at the *Zfp30* locus [according to Sanger data (Keane *et al.* 2011), BALB/c has the same allele as A/J and CBA has the same allele as C57BL/6J]. Hence these results are consistent with the hypothesis that the variation in *Zfp30* affects responses to a variety of immune stimuli.

We do not yet know how *Zfp30* controls CXCL1 and PMN influx. We examined gene expression data we previously collected and found that *Zfp30* expression does not change as a function of allergen challenge in whole lung or isolated airway epithelia (unpublished results). Thus baseline variation in *Zfp30* expression appears to be a determinant of the response to house dust mite allergen. It is somewhat surprising that a gene whose expression is not modulated by allergen challenge can play such an important role in neutrophil recruitment responses to allergen. How baseline variation in *Zfp30* expression modifies these responses remains to be determined.

Little is known about the function of ZFP30 (the protein product of *Zfp30)*, except from predictions based on amino acid sequence motifs. ZFP30 is a C2H2-type zinc finger that contains a Krupple-associated box (KRAB) domain, which are known to repress gene expression (Witzgall *et al.* 1994), perhaps through the induction of heterochromatin (Groner *et al.* 2010). *Zfp30* expression was negatively associated with CXCL1 protein (and PMN) but was not associated with lung *Cxcl1* RNA in preCC mice (not shown), suggesting that *Cxcl1* RNA is not a direct target of ZFP30 (in contrast to ZFP36, tristetraprolin, which is well-known regulator of *Cxcl1* RNA levels; Datta *et al.* 2008). The targets and sequence specificity of ZFP30 are clearly important questions to address in the future and will be necessary to understanding how variation in this gene affects immune response.

Allergic asthma is a complex trait that is influenced by multiple genetic loci and environmental exposures. Neutrophilic inflammation is one feature of the disease and is itself a complex trait. Our results suggest that the PMN response to HDM has contributions from at least three loci. We used a systems genetics approach to show that PMN is mediated by CXCL1 and *Zfp30* expression, and *Zfp30* expression is under local genetic control by *Dpc1*. This biological cascade

is undoubtedly a small piece of a larger puzzle. Nonetheless, we have tied a novel genetic variant to a complex biological response through its most proximal endophenotype, *Zfp30* expression.

Literature Cited

Ackerman, K. G., H. Huang, H. Grasemann, C. Puma, J. B. Singer *et al.*, 2005   Interacting genetic loci cause airway hyperresponsiveness. Physiol. Genomics 21: 105–111.

Aylor, D. L., W. Valdar, W. Foulds-Mathes, R. J. Buus, R. A. Verdugo *et al.*, 2011   Genetic analysis of complex traits in the emerging collaborative cross. Genome Res. 21: 1213–1222.

Bottomly, D., M. T. Ferris, L. D. Aicher, E. Rosenzweig, A. Whitmore *et al.*, 2012   Expression quantitative trait loci for extreme host response to influenza A in pre-collaborative cross mice. Genes Genomes Genet. 2: 213–221.

Bozic, C., L. Kolakowski, N. Gerard, and C. Garcia-Rodriguez, C. von Uexkull- Guldenband *et al.*, 1995   Expression and biologic characterization of the murine chemokine KC. J. Immunol. 154: 6048–6057.

Brewer, J. P., A. B. Kisselgof, and T. R. Martin, 1999   Genetic variability in pulmonary physiological, cellular, and antibody responses to antigen in mice. Am. J. Respir. Crit. Care Med. 160: 1150–1156.

Camateros, P., R. Marino, A. Fortin, J. Martin, E. Skamene *et al.*, 2010   Identification of novel chromosomal regions associated with airway hyperresponsiveness in recombinant congenic strains of mice. Mamm. Genome 21: 28–38.

Chesler, E., D. Miller, L. Branstetter, L. Galloway, B. Jackson *et al.*, 2008   The Collaborative Cross at Oak Ridge National Laboratory: developing a powerful resource for systems genetics. Mamm. Genome 19: 382–389.

Collaborative Cross Consortium, 2012   The genome architecture of the Collaborative Cross mouse genetic reference population. Genetics 190: 389–401.

Cookson, W., L. Liang, G. Abecasis, M. Moffatt, and M. Lathrop, 2009   Mapping complex disease traits with global gene expression. Nat. Rev. Genet. 10: 184–194.

Datta, S., R. Biswas, M. Novotny, P. G. Pavicic, T. Herjan *et al.*, 2008   Tristetraprolin regulates CXCL1 (KC) mRNA stability. J. Immunol. 180: 2545–2552.

De Sanctis, G. T., M. Merchant, D. R. Beier, R. D. Dredge, J. K. Grobholz *et al.*, 1995   Quantitative locus analysis of airway hyperresponsiveness in A/J and C57BL/6J mice. Nat. Genet. 11: 150–154.

De Sanctis, G. T., J. B. Singer, A. Jiao, C. N. Yandava, Y. H. Lee *et al.*, 1999   Quantitative trait locus mapping of airway responsiveness to chromosomes 6 and 7 in inbred mice. Am. J. Physiol. Lung Cell. Mol. Physiol. 277: L1118–L1123.

Denny, P., E. Hopes, N. Gingles, K. Broman, W. McPheat *et al.*, 2003   A major locus conferring susceptibility to infection by *Streptococcus pneumoniae* in mice. Mamm. Genome 14: 448–453.

Dupuis, J., and D. Siegmund, 1999   Statistical methods for mapping quantitative trait loci from a dense set of markers. Genetics 151: 373–386.

Emilsson, V., G. Thorleifsson, B. Zhang, A. S. Leonardson, F. Zink *et al.*, 2008   Genetics of gene expression and its effect on disease. Nature 452: 423–428.

Ferris, M. T., D. L. Aylor, D. Bottomly, A. C. Whitmore, L. D. Aicher *et al.*, 2013   Modeling host genetic regulation of influenza pathogenesis in the Collaborative Cross. PLoS Pathog. 9: e1003196.

Festing, M., 1979   *Inbred Strains in Biomedical Research*. Oxford University Press, New York.

Frevert, C. W., S. Huang, H. Danaee, J. D. Paulauskis, and L. Kobzik, 1995   Functional characterization of the rat chemokine KC and its importance in neutrophil recruitment in a rat model of pulmonary inflammation. J. Immunol. 154: 335–344.

Groner, A. C., S. Meylan, A. Ciuffi, N. Zangger, G. Ambrosini *et al.*, 2010   KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. PLoS Genet. 6: e1000869.

Hageman, R. S., M. S. Leduc, C. R. Caputo, S.-W. Tsaih, G. A. Churchill *et al.*, 2011   Uncovering genes and regulatory pathways related to urinary albumin excretion. J. Am. Soc. Nephrol. 22: 73–81.

Harbison, S. T., M. A. Carbone, J. F. Ayroles, E. A. Stone, R. F. Lyman *et al.*, 2009   Co-regulated transcriptional networks contribute to natural genetic variation in Drosophila sleep. Nat. Genet. 41: 371–375.

Himes, B. E., K. Sheppard, A. Berndt, A. S. Leme, R. A. Myers *et al.*, 2013   Integration of mouse and human genome-wide association data identifies *KCNIP4* as an asthma gene. PLoS ONE 8: e56179.

Hindorff, L. A., P. Sethupathy, H. A. Junkins, E. M. Ramos, J. P. Mehta *et al.*, 2009   Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc. Natl. Acad. Sci. USA 106: 9362–9367.

Jatakanon, A., C. Uasuf, W. Maziak, S. A. M. Lim, K. F. Chung *et al.*, 1999   Neutrophilic inflammation in severe persistent asthma. Am. J. Respir. Crit. Care Med. 160: 1532–1539.

Karp, C. L., A. Grupe, E. Schadt, S. L. Ewart, M. Keane-Moore *et al.*, 2000   Identification of complement factor 5 as a susceptibility locus for experimental allergic asthma. Nat. Immunol. 1: 221–226.

Keane, T. M., L. Goodstadt, P. Danecek, M. A. White, K. Wong *et al.*, 2011   Mouse genomic variation and its effect on phenotypes and gene regulation. Nature 477: 289–294.

Kelada, S. N. P., M. S. Wilson, U. Tavarez, K. Kubalanza, B. Borate *et al.*, 2011   Strain-dependent genomic factors affect allergen-induced airway hyperresponsiveness in mice. Am. J. Respir. Cell Mol. Biol. 45: 817–824.

Kelada, S. N. P., D. L. Aylor, B. C. E. Peck, J. F. Ryan, U. Tavarez *et al.*, 2012   Genetic analysis of hematological parameters in incipient lines of the Collaborative Cross. Genes Genomes Genet. 2: 157–165.

Kelada, S. N. P., D. E. Carpenter, D. L. Aylor, P. Chines, H. Rutledge *et al.*, 2014   Integrative genetic analysis of allergic inflammation in the murine lung. Am. J. Respir. Cell Mol. Biol. 51: 436–445.

Kolaczkowska, E., and P. Kubes, 2013   Neutrophil recruitment and function in health and inflammation. Nat. Rev. Immunol. 13: 159–175.

Leme, A., A. Berndt, L. Williams, S.-W. Tsaih, J. Szatkiewicz *et al.*, 2010   A survey of airway responsiveness in 36 inbred mouse strains facilitates gene mapping studies and identification of quantitative trait loci. Mol. Genet. Genomics 283: 317–326.

Levitt, R. C., and W. Mitzner, 1988   Expression of airway hyperreactivity to acetylcholine as a simple autosomal recessive trait in mice. FASEB J. 2: 2605–2608.

Li, Y., B. M. Tesson, G. A. Churchill, and R. C. Jansen, 2010   Critical reasoning on causal inference in genome-wide linkage and association studies. Trends Genet. 26: 493–498.

Little, S. A., K. J. MacLeod, G. W. Chalmers, J. G. Love, C. McSharry *et al.*, 2002   Association of forced expiratory volume with disease duration and sputum neutrophils in chronic asthma. Am. J. Med. 112: 446–452.

Liu, E. Y., Q. Zhang, L. McMillan, F. P.-M. de Villena, and W. Wang, 2010   Efficient genome ancestry inference in complex pedigrees with inbreeding. Bioinformatics 26: i199–i207.

Matesic, L. E., A. De Maio, and R. H. Reeves, 1999   Mapping lipopolysaccharide response loci in mice using recombinant inbred and congenic strains. Genomics 62: 34–41.

Maurano, M. T., R. Humbert, E. Rynes, R. E. Thurman, E. Haugen *et al.*, 2012   Systematic localization of common disease-associated variation in regulatory DNA. Science 337: 1190–1195.

Mott, R., C. J. Talbot, M. G. Turri, A. C. Collins, and J. Flint, 2000   A method for fine mapping quantitative trait loci in outbred animal stocks. Proc. Natl. Acad. Sci. USA 97: 12649–12654.

Nica, A. C., and E. T. Dermitzakis, 2008   Using gene expression to investigate the genetic basis of complex disorders. Hum. Mol. Genet. 285: 17: R129–R134.

Nicolae, D. L., E. Gamazon, W. Zhang, S. Duan, M. E. Dolan *et al.*, 2010   Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet. 6: e1000888.

Scapini, P., J. A. Lapinet-Vera, S. Gasperini, F. Calzetti, F. Bazzoni *et al.*, 2000   The neutrophil as a cellular source of chemokines. Immunol. Rev. 177: 195–203.

Schadt, E. E., J. Lamb, X. Yang, J. Zhu, S. Edwards *et al.*, 2005   An integrative genomics approach to infer causal associations between gene expression and disease. Nat. Genet. 37: 710–717.

Schenten, D., and R. Medzhitov, 2011   The control of adaptive immune responses by the innate immune system, pp. 87–124 in *Advances in Immunology*, Chap. 3, edited by W. A. Frederick. Academic Press, San Diego.

Shipley, B., 2000   *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations, and Causal Inference*. Cambridge Univ. Press, Cambridge, UK.

Valdar, W., C. C. Holmes, R. Mott, and J. Flint, 2009   Mapping in structured populations by resample model averaging. Genetics 182: 1263–1277.

Wenzel, S. E., 2006   Asthma: defining of the persistent adult phenotypes. Lancet 368: 804–813.

Wenzel, S. E., S. J. Szefler, D. Y. M. Leung, S. I. Sloan, M. D. Rex *et al.*, 1997   Bronchoscopic evaluation of severe asthma. Am. J. Respir. Crit. Care Med. 156: 737–743.

Wenzel, S. E., L. B. Schwartz, E. L. Langmack, J. L. Halliday, J. B. Trudeau *et al.*, 1999   Evidence that severe asthma can be divided pathologically into two inflammatory subtypes with distinct physiologic and clinical characteristics. Am. J. Respir. Crit. Care Med. 160: 1001–1008.

Whitehead, G. S., J. K. L. Walker, K. G. Berman, W. M. Foster, and D. A. Schwartz, 2003 Allergen-induced airway disease is mouse strain dependent. Am. J. Physiol. Lung Cell. Mol. Physiol. 285: L32–L42.

Witzgall, R., E. O'Leary, A. Leaf, D. Onaldi, and J. V. Bonventre, 1994 The Krüppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. Proc. Natl. Acad. Sci. USA 91: 4514–4518.

Yalcin, B., K. Wong, A. Agam, M. Goodson, T. M. Keane *et al.*, 2011 Sequence-based characterization of structural variation in the mouse genome. Nature 477: 326–329.

Yang, H., Y. Ding, L. N. Hutchins, J. Szatkiewicz, T. A. Bell *et al.*, 2009 A customized and versatile high-density genotyping array for the mouse. Nat. Methods 6: 663–666.

You, Y., E. J. Richer, T. Huang, and S. L. Brody, 2002 Growth and differentiation of mouse tracheal epithelial cells: selection of a proliferative population. Am. J. Physiol. Lung Cell. Mol. Physiol. 283: L1315–L1321.

Communicating editor: D. W. Threadgill

GENETICS

Genetic Regulation of *Zfp30*, CXCL1, and Neutrophilic Inflammation in Murine Lung

Holly Rutledge, David L. Aylor, Danielle E. Carpenter, Bailey C. Peck, Peter Chines, Lawrence E. Ostrowski, Elissa J. Chesler, Gary A. Churchill, Fernando Pardo-Manuel de Villena, and Samir N. P. Kelada
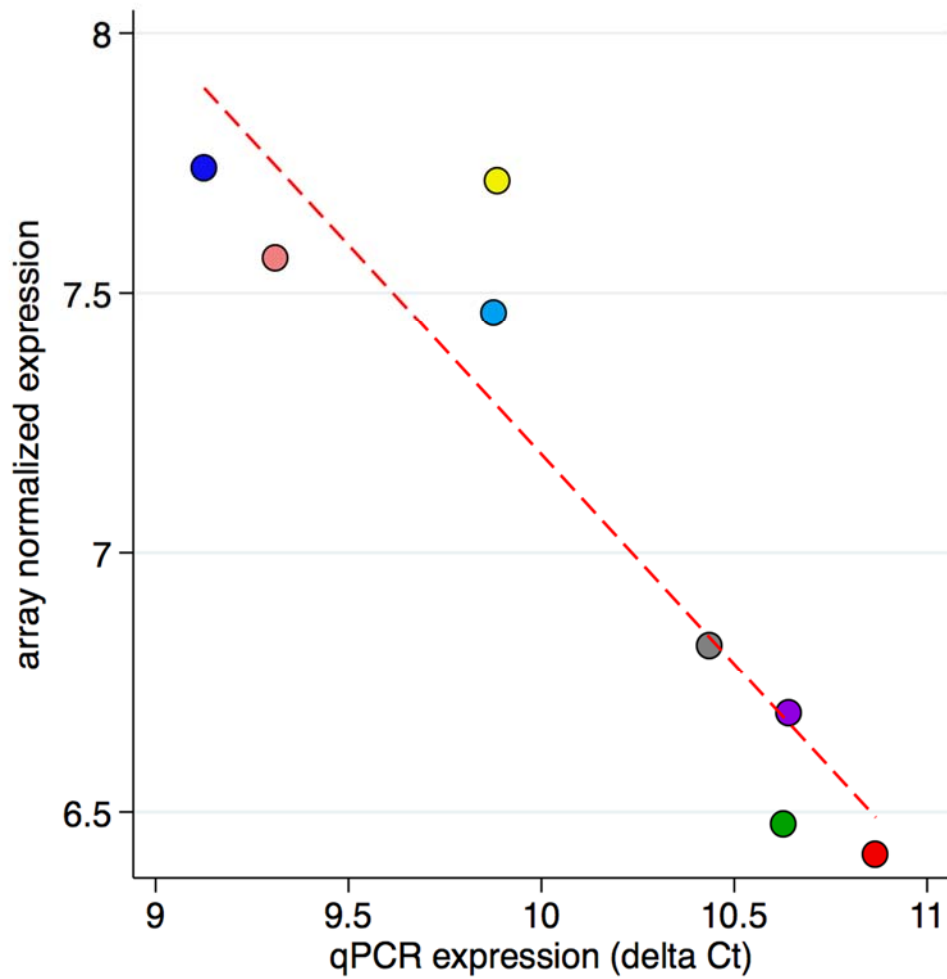
Figure S1 *Zfp30* **gene expression measured by microarray and qRT-PCR is strongly correlated.** Data points represent the mean of 2-4 mice per CC founder strain and are color coded as in Figure 1. Dashed red line represents linear fit of data. R^2 = 0.85 (p = 0.001).
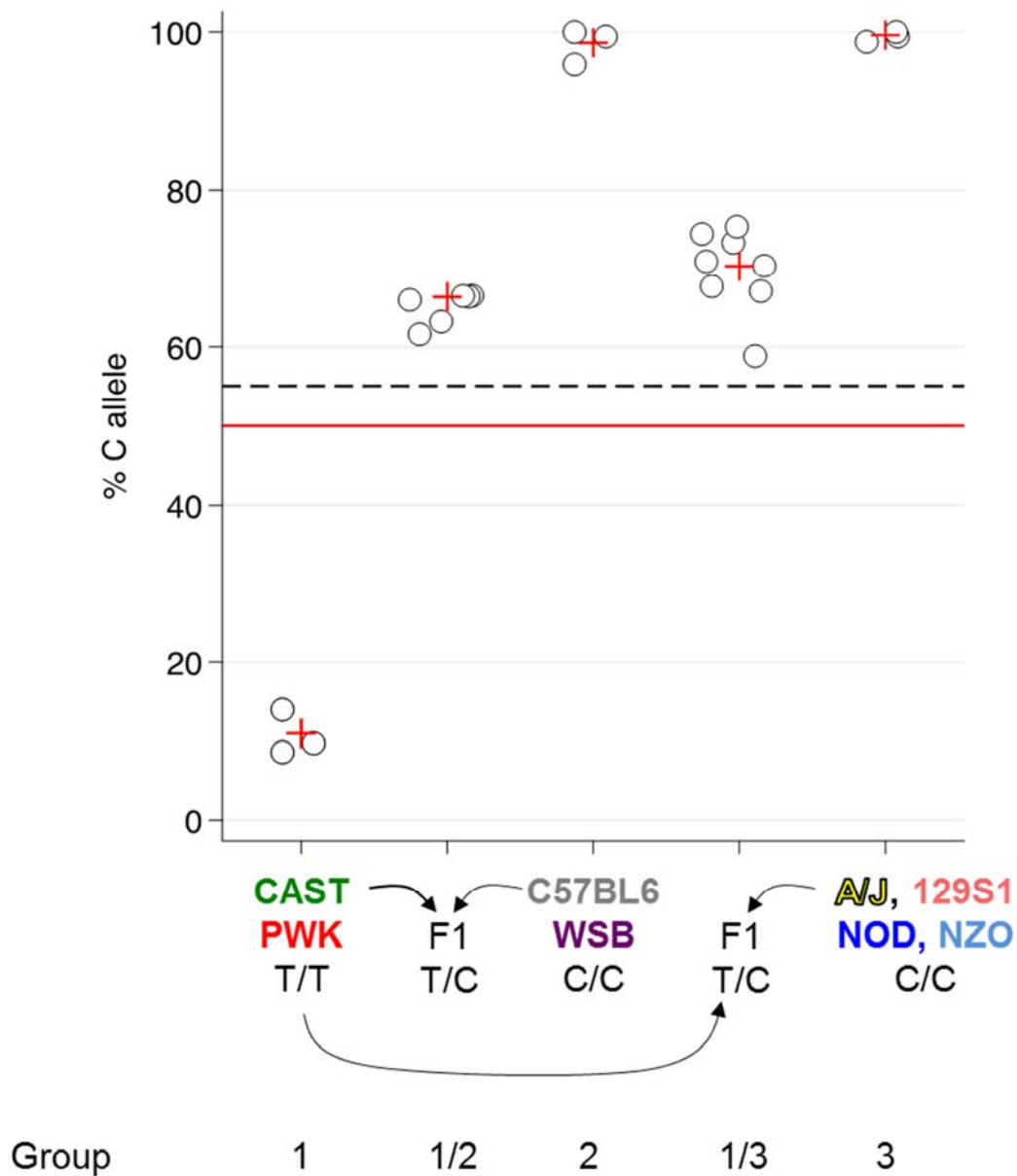
Figure S2 Allele-specific gene expression of *Zfp30*. Expression of alternate alleles of rs45830063 (C/T) was measured using pyro-sequencing on whole lung cDNA samples from CC founder lines (n=3/group) and F1 heterozygotes from crosses of CC founder lines from each *Zfp30* expression group (n=6-8/group). Strain groups on X-axis correspond to gene expression groups identified in Figure 4. Plus symbols denote mean. Dashed black line indicates mean of T/T and C/C homozygotes (55%). Both groups of F1 heterozygotes (1/2 and 1/3) express significantly higher levels of the C allele than T allele (>55%, one sided t-test *p*-value < 1 x 10[-4]) and Group 1/3 heterozygotes express higher levels of the C allele than Group 1/2 heterozygotes (70% vs. 66%, one sided t-test *p*-value = 0.04).

Concordant and Discordant SNPs in *Ztp30* (ENSMUST00000032803)

A

○ Concordant SNPs
● Discordant SNPs

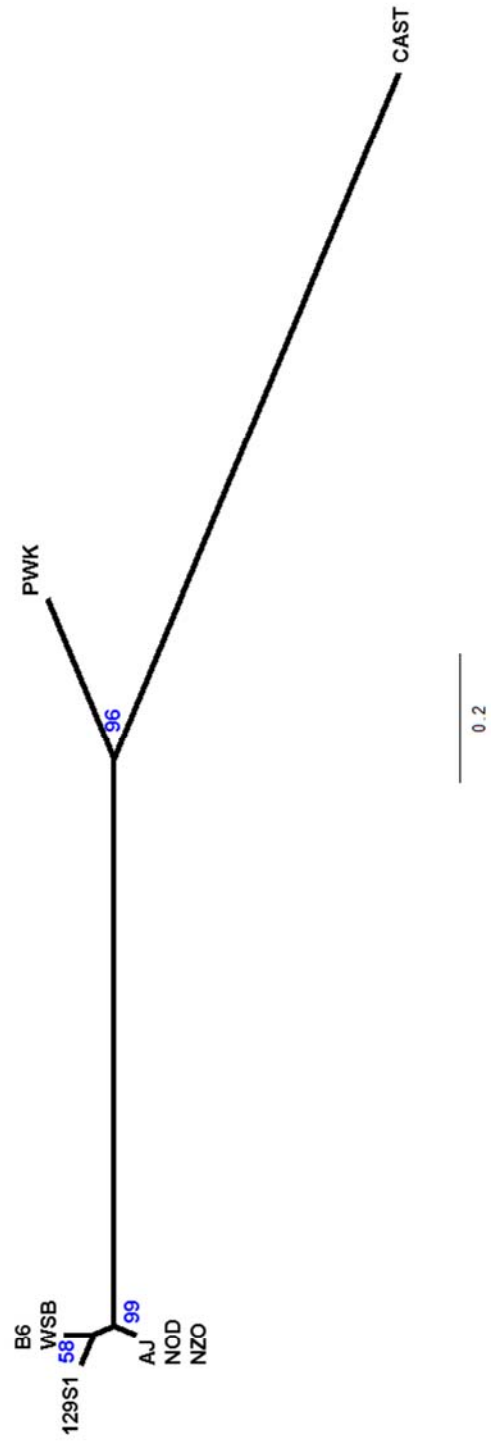Chr 7 position (Mb)

B

H. Rutledge *et al.*
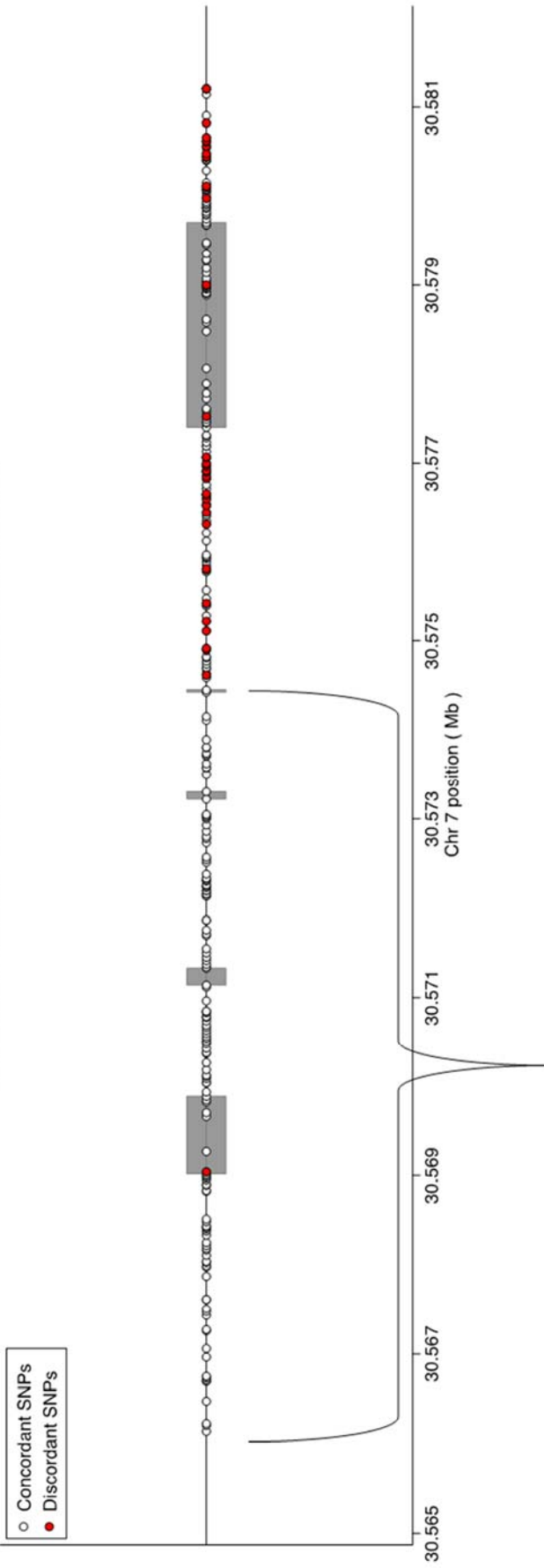
Figure S3 SNPs and haplotypes among CC founder strains at the *Zfp30* locus. A. SNPs (n = 242 SNPs) in *Zfp30* including 3kb of upstream and 1.5 kb of downstream sequence (30,566,017-30,581,221 bp) that are concordant (white circles) or discordant (red circles) with the pattern of gene expression shown in Figure 4. Gray boxes denote exons. **B.** Phylogeny of CC founder haplotypes for *Zfp30* sequence up to exon 4. Blue numbers indicate bootstrap support values.

File S1

Genotype Data. Genotype calls for preCC mice.

Available for download as a .gz file at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.168138/-/DC1

Tables S1-S4

Available for download at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.168138/-/DC1

Table S1   Phenotype data for CC founder strains and preCC mice

Table S2   Chr 7 Genotypes, *Zfp30* Expression Groups and Frequencies Used in Narrow Sense Heritability Regression Models and Structural Equation Models

Table S3   Genes with local eQTL in the PMN/CXCL1 QTL confidence interval on Chr 7.

Table S4   Potential Causal Variants for the *Zfp30* eQTL.