BASIC RESEARCH www.jasn.org

Uncovering Genes and Regulatory Pathways Related to Urinary Albumin Excretion

Rachael S. Hageman, Magalie S. Leduc, Christina R. Caputo, Shirng-Wern Tsaih, Gary A. Churchill, and Ron Korstanje

The Jackson Laboratory, Bar Harbor, Maine

ABSTRACT

Identifying the genes underlying quantitative trait loci (QTL) for disease is difficult, mainly because of the low resolution of the approach and the complex genetics involved. However, recent advances in bioinformatics and the availability of genetic resources now make it possible to narrow the genetic intervals, test candidate genes, and define pathways affected by these QTL. In this study, we mapped three significant QTL and one suggestive QTL for an increased albumin-to-creatinine ratio on chromosomes (Chrs) 1, 4, 15, and 17, respectively, in a cross between the inbred MRL/MpJ and SM/J strains of mice. By combining data from several sources and by utilizing gene expression data, we identified *Tlr12* as a likely candidate for the Chr 4 QTL. Through the mapping of 33,881 transcripts measured by microarray on kidney RNA from each of the 173 male F2 animals, we identified several downstream pathways associated with these QTL, including the glycan degradation, leukocyte migration, and antigen-presenting pathways. We demonstrate that by combining data from multiple sources, we can identify not only genes that are likely to be causal candidates for QTL but also the pathways through which these genes act to alter phenotypes. This combined approach provides valuable insights into the causes and consequences of renal disease.

J Am Soc Nephrol 22: 73–81, 2011. doi: 10.1681/ASN.2010050561

Chronic kidney disease is a growing medical problem caused by various environmental and genetic factors. Identifying the genes underlying common forms of kidney disease in humans has proven difficult, expensive, and time consuming. However, quantitative trait loci (QTL) for several complex traits, including renal phenotypes,[1] are concordant among mice, rats, and humans, suggesting that genetic findings from animal models are relevant to human disease. With respect to chronic kidney disease, QTL analysis using mice is likely to contribute new findings in the near future.

In addition to mapping the causative loci, it is of equal importance to identify the pathways regulated by the loci so that we gain a better understanding of the processes that drive renal damage. One approach to identify the genes that are driven by certain loci is a method known as genetical genomics, in which gene transcripts are treated as quantitative traits and mapped in a cross in the same way as any other phenotype.[2]

In this study we generated an F2 intercross between the kidney damage-susceptible SM/J (SM) and the nonsusceptible MRL/MpJ (MRL) mouse inbred strains. In addition to measuring the urinary albumin-to-creatinine ratio (ACR), we obtained a kidney expression profile from each mouse using Affymetrix arrays. This allowed us to identify the loci responsible for the difference in ACR between

Received May 28, 2010. Accepted August 18, 2010.

Published online ahead of print. Publication date available at www.jasn.org.

Correspondence: Dr. Ron Korstanje, The Jackson Laboratory, 600 Main Street, Bar Harbor, ME 04609. Phone: 207-288-6000; Fax: 207-288-6078; E-mail: ron.korstanje@jax.org

Table 1. Number and percentage of animals with albumin excretion in the parental, F1, and F2 males

	n	Animals with ACR > 0
MRL/MpJ	10	0
SM/J	14	5 (36%)
(MRL×SM)F1	32	9 (28%)
F2	172	97 (60%)

the two strains as well as the pathways that are driven by these loci in the kidney.

RESULTS

Differences in ACR between MRL and SM Mice Are Determined by Three Major Loci

Albumin and creatinine were measured in parental, F1, and 173 male F2 animals, and the ACRs were determined. In the parental animals only SM/J males show excretion of albumin. We observed a higher percentage of animals with albumin excretion in the F2 population compared with the F1 population, suggesting that some of the loci determining this phenotype are recessive (Table 1).

A three-stage QTL analysis was performed using the ACR phenotype. First, a main QTL scan was performed (Figure 1) in which significant QTL were found on chromosomes (Chrs) 1, 4, and 15 and one suggestive QTL was found on Chr 17 (data are summarized in Table 2). The allele-effect plots for these four loci suggest that the alleles are additive for the Chr 1 and Chr 4 loci and recessive for the Chr 15 locus, with the SM contributing to high ACR at all three loci. For the suggestive locus on Chr 17, it is the MRL allele that contributes to high ACR (Figure 2). Second, a pairscan was used to detect interaction among loci, but no significant interactions were found. Lastly, a multiple-regression model for ACR was fit with the main effect QTL (Table 3). The estimated proportion of phenotype variance explained by the four main-effect QTL is 24%. By dropping one term (QTL) at a time and comparing reduced models to the full model, we estimate that the Chr 4 locus explains the most variance (6.84%) and is the most significant term in the model.

Identification of *Tlr12* as a Candidate Gene That Determines Increased ACR in SM Mice

The Chr 4 locus at 53 centimorgans (cM) has been previously observed in an F2 intercross between C57BL/6J and A/J[3] (in which A/J contributed the high allele) and in a backcross between C57BL/6J and NZM/Aeg2410[4] (in which NZM contributed the high allele). Haplotype analysis can be applied under the assumption that the same gene underlies the QTL in all three crosses. In the analysis we assumed that the two strains contributing to the low allele were identical by descent, likewise for the three strains contributing to the high allele. In addition, we assumed that the high allele strains and low allele strains were different. We analyzed the 95% CI (from 66 to 130 Mb) using the Strain Comparison tool (cgd.jax.org/straincomparison). This resulted in 73 small regions containing only 42 genes, 41 of which had transcripts on the microarray (Table 4). For each of these genes, we tested for *cis*-linkage, performed a *t* test to identify significant differences between the MRL and SM alleles at the Chr 4 QTL loci, and used conditional genome scans as a graphical modeling technique to distinguish genes causal to ACR. Of these 42 genes, *Tlr12* is the most differentially expressed gene between F2 homozygous SM animals for the Chr 4 locus and homozygous MRL animals, with MRL having a two-fold higher expression in kidney ($P < 1.88 \times 10^{-27}$). This analysis revealed *Trl12*→ACR as the only causal relationship with significant change in the logarithm of odds (LOD) between ACR ($LOD(ACR|Trl12) = 1.55$) and the candidate genes in the Chr 4 QTL region (Table 4). We confirmed these observations by staining MRL and SM kidneys with an antibody against the TLR12 protein. Figure 3 shows staining in the distal convoluted tubuli in MRL, with no staining in SM. Zhang *et al.*[5] showed that TLR12 (in their paper called TLR11) is involved in the host defense system against bacteria in the kidney and that inactivation of the gene leads to a higher susceptibility for renal infection.[4]

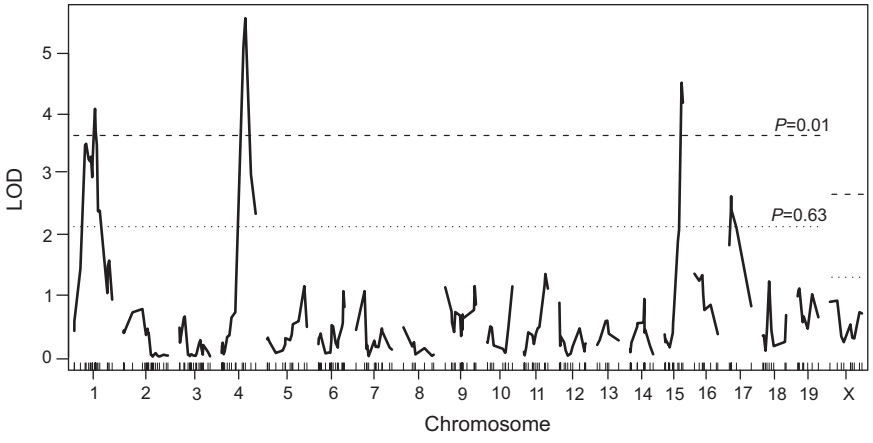

Figure 1. Genome-wide scan for ACR identifying three significant QTL ($P < 0.05$) on Chrs 1, 4, and 15, and one suggestive QTL ($P < 0.63$) on Chr 17.

Expression QTL in Kidneys Overlap the ACR Loci

Kidney samples from 173 F2 males were processed on Mouse Gene 1.0 ST Affymetrix GeneChips. Single trait QTL analysis was performed on 33,881 transcripts. Genetic loci and intervals that are most strongly linked to the variation in gene expression were identified for each individual transcript. Because the physical positions of all transcripts is known, we can visualize the results in a two-dimensional scatter plot in which the x-axis indicates the position of the controlling locus and the y-axis identifies the physical chromosomal position of transcripts and their

Table 2. Significant and suggestive main QTL for ACR

Locus	Chr	Peak (cM)	CI (cM)[a]	CI (Mb)	LOD	Marker[b]	High Allele	Mode[c]
Albq10	1	46	14 to 53	35.3 to 124.2	4.1	*rs13476079*	SM	A
Albq4	4	53	33 to 65	61.8 to 131.8	5.6	*rs3661158*	SM	A
Albq11	15	38	33 to 41	72.2 to 86.7	4.5	*rs13482713*	SM	R
	17	12	8 to 56	12.1 to 86.1	2.6	*rs3702604*	MRL	A

[a]CI = 95% confidence interval.
[b]Marker closest to the QTL peak.
[c]A, additive inheritance; R, recessive inheritance.

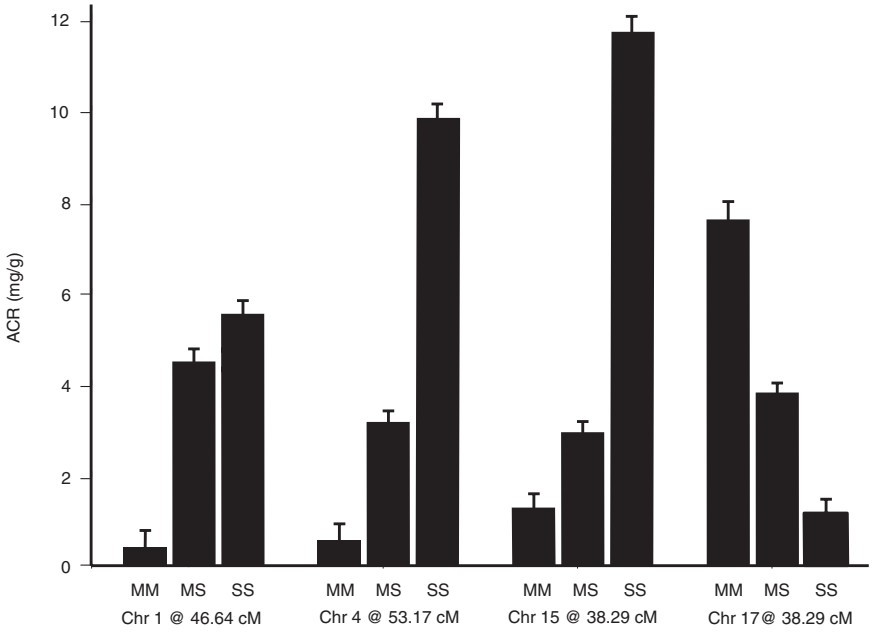

Figure 2. Allele-effect plots for UAE at the peak markers of the four identified QTL. All of the loci show an additive pattern. For the QTL on Chrs 1, 4, and 15 the high allele comes from SM, whereas for the Chr 17 QTL the high allele is contributed by MRL.

Table 3. Multiple-regression analysis of variance for ACR

Locus (cM)	LOD	Variance (%)	*F*	*P*
Chr 1 @ 46.2	3.15	3.34	3.59	0.03
Chr 4 @ 53.2	4.54	6.84	7.35	0.0008
Chr 15 @ 38.3	4.3	3.35	3.6	0.03
Chr 17 @ 12.4	1.24	3.04	3.27	0.04

Multiple regression was used to fit a QTL model for ACR. The full model explained 24% of the phenotype variance ($P = 2.4 \times 10^{-7}$). Each locus is dropped from the model, one at a time, and compared with the full model with all of the main effect terms included. LOD is the log-likelihood ratio from the comparison of the reduced model and the full model. Variance (%) is the estimated proportion of variance explained by each term (locus).

corresponding QTL peaks (Figure 4). A locus that controls a large number of transcripts is known as a transband. We have identified transbands on Chrs 1, 4, 15, and 17 that overlap the QTL found in our ACR analysis. Within these four transband regions, there are genes that are polymorphic between MRL and SM. These genes can potentially cause differences in both kidney gene expression and urinary albumin excretion.

Pathways and Genes in Transbands Relate to Kidney Function

Transbands are often thought to be the result of upstream allelic variants of genes in the locus region perturbing downstream genes and pathways. We set out to obtain a general characterization of the genes and pathways that arose in the transbands that overlapped the ACR QTL. Transcripts with suggestive QTL ($P < 0.63$) in broad transband regions (±20 cM) of QTL peaks for ACR on Chrs 1, 4, 15, and 17 were selected for further analyses. The lower threshold was chosen to ensure the inclusion of genes that may only be close to significant but may give rise to significant pathway signals. There were 4402, 4503, 2117, and 2579 transcripts in the Chr 1, 4, 15, and 17 transbands, respectively. Several genes that appeared in more than one transband (Figure 5) are associated with abnormal kidney phenotypes in mouse according to the Mammalian Phenotype Ontology[6] (Table 5 and Supplemental Table 1). Pathway enrichment analysis was performed on these transbands to detect over-represented KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways (Table 6). Genes in the Chr 4 transband are enriched for glycan structure degradation (10/25), likely a downstream consequence suggesting kidney damage. *Agt*, which encodes angiotensin and is part of the renin-angiotensin system, had significant QTL in the Chrs 1, 4, and 17 transbands; additionally, six of the 17 other genes in this pathway are regulated by the Chr 4 locus, enriching the transband ($P < 0.01$). The *Agt* transcript is positively correlated to ACR ($\rho = 0.4$ and $P = 1 \times 10^{-6}$) and is regulated by many of the same loci. Genes in the Chr 1 region are enriched for the *N*-glycan degradation pathway (7/15) and leukocyte transendothelial migration (24/106). The Chr 17 region is enriched for antigen processing and presentation (21/79).

DISCUSSION

Mapping of QTL has proven to be successful in identifying genomic regions that are associated with multigenic diseases. A fundamental problem in analysis of this type is the subsequent identification of the causative gene, because often the large chromosomal intervals still contain hundreds of genes. Recent developments of bioinformatic tools use the growing genomic resources and the realization that, as loci are concordant between species, data from different

Table 4. Genes in the Chr 4 QTL interval after narrowing the region by haplotype comparison

Gene	Start (bp)	Chr 4 LOD Score				*cis*-QTL	High Allele	*P*
		M0: ACR	M0:ACR|GEX	N0:GEX	N0:GEX|ACR			
Tlr4	66488618	5.6	5.41	1.23	1.26	N	SM	0.632
Ptprd	75587143	5.6	5.28	1.71	1.52	N	MRL	0.478
Mpdz	80924409	5.6	5.46	3.61	3.87	Y	MRL	0.059
Nfib	81936077	5.6	4.68	1.08	0.71	N	MRL	0.389
Zdhhc21	82444642	5.6	5.18	16.69	16.75	Y	SM	4.59×10^{-05}
Cer1	82527655	5.6	5.31	1.70	1.58	N	Het	0.609
Frem1	82543831	5.6	6.60	2.56	3.49	N	SM	0.002
Cachd1	100449280	5.6	5.37	0.91	0.82	N	MRL	0.675
Sgip1	102413902	5.6	5.40	2.13	2.42	N	MRL	0.020
Wdr78	102690477							
Insl5	102710670	5.6	5.35	0.50	0.51	N	MRL	0.839
Mier1	102786995	5.6	4.72	1.79	1.42	N	MRL	0.010
Slc35d1	102843254	5.6	4.19	7.09	5.73	Y	SM	2.06×10^{-05}
Dab1	103291964	5.6	5.32	2.92	2.65	N	Het	0.567
Gm1027	106353022	5.6	2.40	6.96	4.62	Y	SM	0.000
Tmem59	106851004	5.6	5.42	1.76	1.74	N	SM	0.789
Zyg11a	107854539	5.6	5.51	1.82	2.21	N	Het	0.314
Zcchc11	108132031	5.6	4.59	3.04	1.93	N	MRL	0.002
Hyi	118032595	5.6	3.65	16.74	17.03	Y	SM	7.24×10^{-11}
Szt2	118035348	5.6	5.27	1.60	1.59	N	Het	0.573
Med8	118081942	5.6	5.32	2.08	2.04	N	MRL	0.207
Elovl1	118100698	5.6	5.03	2.02	1.91	N	SM	0.103
Tmem125	118213546	5.6	5.37	0.54	0.55	N	SM	0.268
Ppt1	122513485	5.6	4.34	7.90	7.23	Y	SM	0.000
Cap1	122536290	5.6	5.27	0.90	0.93	N	SM	0.546
Mfsd2a	122624093	5.6	4.64	5.53	4.77	Y	SM	6.56×10^{-06}
Mycl1	122672895	5.6	5.37	0.68	0.65	N	Het	0.998
Oxct2b	122793507	5.6	5.31	0.71	0.70	N	MRL	0.477
Ppie	122804358	5.6	5.09	1.53	1.23	N	MRL	0.156
Akirin1	123411802	5.6	5.37	1.95	2.06	N	SM	0.933
Rhbdl2	123465117	5.6	4.02	14.27	12.47	Y	MRL	4.74×10^{-10}
Psmb2	126354874	5.6	5.36	1.61	1.62	N	MRL	0.374
Tcfap2e	126393250	5.6	5.38	1.02	0.89	N	Het	0.936
Zmym4	126539183	5.6	4.44	3.06	2.88	N	MRL	0.003
Zmym1	126724338	5.6	5.23	1.33	1.36	N	MRL	0.279
Zmym6	126754627	5.6	5.95	1.42	1.49	N	MRL	0.233
Csmd2	127665101	5.6	5.37	1.14	0.91	N	Het	0.789
Hmgb4	127937456	5.6	5.28	1.02	0.80	N	MRL	0.512
Zscan20	128260783	5.6	5.42	1.84	2.01	N	MRL	0.012
Tlr12	128292690	5.6	1.33	40.18	38.86	Y	MRL	1.88×10^{-27}
Phc2	128331982	5.6	5.29	1.66	1.59	N	Het	0.127
A3galt2	128432608	5.6	5.51	1.55	1.65	N	Het	0.675

The Chr 4 LOD scores for the ACR model (M0: $ACR = Q + \varepsilon$) and conditional on the gene expression trait (M1: $ACR = Q + GEX + \varepsilon$), likewise, the GEX model (N0: $GEX = Q + \varepsilon$) and the conditional model (N1: $GEX = ACR + Q + \varepsilon$). Transcripts that are *cis*-QTL at the suggestive level ($P < 0.05$) are indicated, along with the high allele and the *P* value for the ANOVA contrast MRL *versus* SM. *Trl12* is the only candidate that upon conditioning reduces the ACR LOD peak below the suggestive level ($P < 0.63$). Conditioning *Trl12* on ACR has little effect on the LOD score. Taken together, this suggests that *Trl12* is the most likely upstream candidate gene.

species can be combined to narrow QTL regions for the identification of candidate genes. We applied these techniques along with graphical modeling methods to gene expression data to identify *Trl12* as the most likely candidate underlying the Chr 4 QTL determining the difference in ACR between MRL and SM. In addition to mapping the QTL that are responsible for the difference in urinary albumin excretion between MRL and SM, we also mapped gene transcripts from the kidneys of the same F2 animals using microarrays. We found that the same loci driving the ACR phenotype also drive the expression of many downstream genes and pathways that directly relate to kidney function.

By using the ACR as a convenient measure for urinary albumin excretion in an intercross between MRL (which does not excrete any detectable albumin in the urine) and SM (which excretes albumin), we identified three signifi-

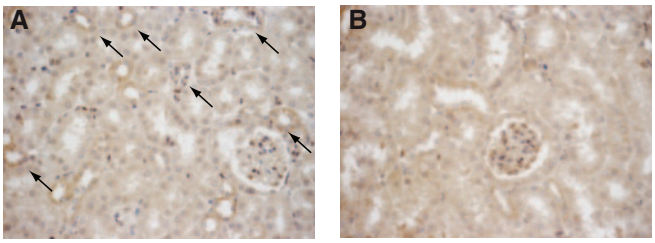

Figure 3. A TLR12 antibody stains the distal convoluted tubules of MRL kidneys (A, indicated by the arrows), but not of SM kidneys (B).

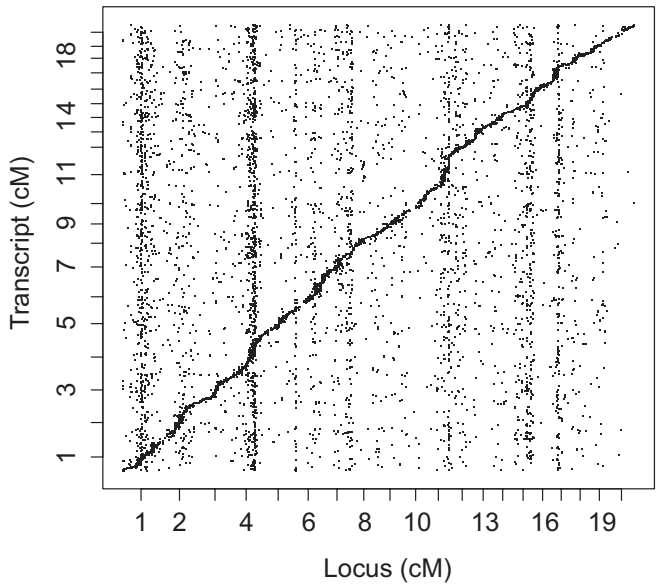

Figure 4. Expression-QTL map for transcripts with significant ($P < 0.05$, LOD >3.85) QTL. The x-axis indicates the position of the QTL locus, and the y-axis identifies the physical position of the transcript. For each transcript, the locus corresponding to the highest LOD score was selected for each chromosome. If LOD scores exceeded the threshold, they are indicated (black dots). *cis*-Acting QTL can be seen on the diagonal. Transbands of *trans*-acting QTL can be seen as vertical lines. Clear transbands are present on Chrs 1 and 4.

cant QTL on Chrs 1, 4, and 15 and one suggestive QTL on Chr 17. Comparing these results with previous studies, recently summarized by Garrett *et al.*,[1] we see that the homologous region of the Chr 1 locus has been mapped in a rat cross[7] and a human linkage study,[8] the Chr 4 locus has been mapped in two mouse crosses, B6 × A[3] and B6 × NZM,[4] and the homologous region of the Chr 15 locus has been mapped in a rat cross.[9] This concordance between loci can be used to narrow the interval and exclude genes in the region when we assume that it is the same gene underlying these loci. This method is best demonstrated for the Chr 4 locus, in which we used haplotype comparison to narrow the number of candidate genes to 44 genes. Graphical modeling techniques utilizing the microarray data revealed *Tlr12*→ACR as the most significant causal relationship. In addition, *Tlr12* is known to be involved in kidney infection;

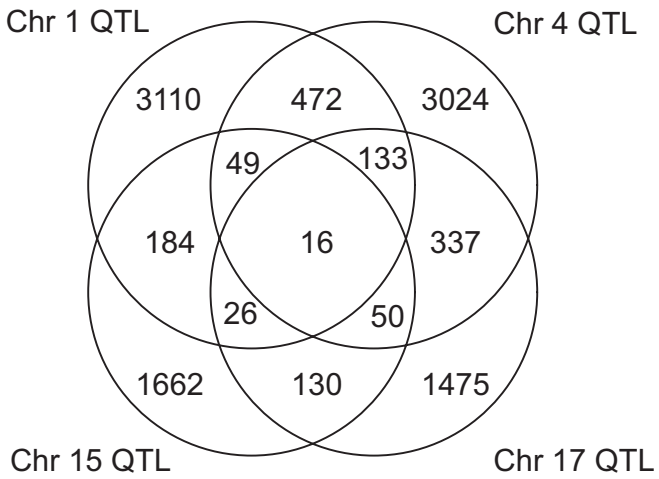

Figure 5. A significant number of transcripts have QTL in more than one of the ACR QTL regions. For example, 3110 transcripts have suggestive QTL in the same Chr 1 region as the ACR QTL, and 472 transcripts have QTL in both the Chr 1 and Chr 4 regions of ACR QTL. Not shown: 412 genes are common to both Chr 1 and Chr 17, and 422 genes are common to Chrs 4 and 15.

we therefore considered this to be the most likely candidate gene of the 44. Because there are no nonsynonymous single-nucleotide polymorphisms (SNPs) between the strains that were used in the crosses, according to the SNP databases of the Mouse Phenome Database (www.jax.org/phenome) and the Mouse Genomes Project at the Sanger Institute (www.sanger.ac.uk/resources/mouse/genomes/), we assumed that a difference in urinary albumin excretion would be caused by a difference in *Tlr12* expression. Our microarray data showed a two-fold higher expression of the gene in MRL compared with SM, and staining in the kidney using an antibody against TLR12 supported this. The antibody stained the distal convoluted tubuli in MRL, but we did not observe any detectable staining in SM.

In addition to finding the causative genes for ACR, it is also of major importance to identify the pathways that are associated with this phenotype. An effective way of doing this is by looking at the genes that are regulated by the ACR QTL. We observed that the regulation of many renal transcripts is controlled by factors that are located within the same intervals as for the ACR. This suggests that the same genes that are responsible for the differences in urinary albumin excretion between MRL and SM may also be responsible for differences in pathway regulation. To identify the pathways that differ between the two strains, we performed a pathway-enrichment analysis. Some of the most interesting pathways we identified are *N*-glycan degradation and leukocyte transendothelial migration controlled by the Chr 1 locus, glycan degradation and the renin-angiotensin system controlled by the Chr 4 locus, and the antigen-presenting pathway controlled by the Chr 17 locus. The involvement of the *N*-glycan pathway in the differences between MRL and SM makes sense in light of the differences in susceptibility to urinary tract infections. Often the first step in

Table 5. Genes annotated to kidney function in mouse that are present in more than one transband associated with ACR QTL

Gene	Chr	Start (bp)	Transband			
			Chr 1	Chr 4	Chr 15	Chr 17
Abca1	4	53043661	×	×		
Agt	8	127080676	×		×	×
Agtr1a	13	30428326	×			×
Aqp2	15	99409487	×	×		
Atp12a	14	56983865	×	×		
Bcam	7	20341487	×	×		×
Cdkn2c	4	109333481	×	×		×
Cybb	X	9013154	×		×	
Fgfr2	7	137305967	×	×		
Gdnf	15	7761011	×			×
Golm1	13	59736357	×	×		
Igh-6	12	114657161			×	×
Lamc2	1	154969886	×		×	
Leprot	4	101320353		×		×
Limk2	11	3243300			×	×
Mpv17	5	31443344	×			×
Mt1	8	96703141	×	×		
Mut	17	41071656	×			×
Nfat5	8	109817370	×	×		
Nfe2l2	2	75513571		×		×
Pdgfb	15	79826330	×		×	×
Pou3f3	1	42753991	×		×	
Ppt1	4	122513489	×		×	×
Prkdc	16	15637959			×	×
Ptprf	4	117880818	×		×	×
Rfc1	5	65653097	×			×
Rhcg	7	86738255	×	×	×	
Slc14a1	18	78296833			×	×
Slc19a1	10	76496004	×	×		
Slc4a1	11	102210141	×	×		
Trp53	11	69393861	×	×		
Tsc22d1	14	76814768			×	×
Twsg1	17	66272429		×		×
Vegfa	17	46153942			×	×
Wt1	2	104966686		×		×
Yes1	5	32913606	×	×		

a bacterial infection is the recognition of host glycans by bacterial lectins and vice versa. A common urinary tract infection by *Escherichia coli* is mediated by the mannose-specific adhesin, FimH. Another example of bacterial lectins involved in infection includes the adhesin on uropathogenic P-fimbriated *E. coli* that recognizes galabiose. Infection caused by a decrease in TLR12 in the kidney, where bacteria will bind and degrade host glycans, will most likely lead to changes in expression of genes involved in glycan and antigen-presenting pathways.

In conclusion, in this study we demonstrated that we can identify QTL for urinary albumin excretion, and by combining previously published data and gene expression analysis, we can select testable candidate genes. Furthermore, gene expression data shed light on the pathways and genes that play important roles in the phenotype. This combination of data will provide valuable insight in the understanding of the causes and consequences of renal diseases.

CONCISE METHODS

Mice and Phenotyping

Female MRL/MpJ (MRL) mice were crossed with male SM/J (SM) mice; their progeny were intercrossed to produce 371 F2 animals. Because of the low penetrance of the phenotype in females, we used only the 173 males from the cross for this study. The mice were housed in a climate-controlled facility with a 14-h:10-h light-dark cycle with free access to food and water throughout the experiment. After weaning, the mice were maintained on a chow diet (5K52; LabDiet®, St. Louis, Missouri). Spot urine samples were taken between 8 and 10 weeks, and albumin and creatinine concentrations were measured on a Beckman Synchron CX5 Chemistry Analyzer. Actual mouse albumin concentrations were calculated by linear regression from a standard curve generated with mouse albumin standards (Kamiya Biomedical Company, Seattle, Washington).[10] All of the experiments were approved by the Jackson Laboratory's Animal Care and Use Committee.

Genotyping

F2 mice were tail-tipped at 2 weeks of age. DNA was extracted using the Gentra kit that utilizes a phenol/chloroform extraction method. DNA was genotyped by the High Throughput Sequenom and Illumina Genotyping facility (http://www.hpcgg.org/), using a 760-SNP array; 258 of these SNPs were polymorphic between MRL and SM for an average distance between markers of 5.5 cM, ranging from 0 to 25.1 cM. The genetic map was derived using Mouse Map converter, and SNP identifications were converted to the cM position.[11]

Gene Expression Profiling

Total RNA was isolated using the TRIzol Plus method according to the manufacturer's protocols including "on the column" DNase digestion. RNA quality was assessed using an Agilent 2100 Bioanalyzer instrument and RNA 6000 Nano LabChip assay (Agilent Technologies, Palo Alto, California). Total RNA was reverse transcribed followed by second strand cDNA synthesis. RNA was labeled and hybridized to the mouse gene 1.0 ST microarray (1M) following the manufacturer's protocols (Affymetrix, Santa Clara, California).

Data were imported into R (http://www.r-project.org/) and processed using the *affy* package from Bioconductor (http://www.bioconductor.org/). The data were normalized using robust multi-array average expression with no background correction.[12] Image plots, sample-based clustering, and histograms were used to detect sample outliers and technical artifacts. A custom cdf file mogene10stvmmenstcdfV11.0.1 was obtained from Brainarray (http://brainarray.mbni.med.umich.edu/brainarray/default.asp) and used to define 33,881 probe sets identified by Ensembl transcript identifications.

Quantitative Trait Locus Analysis

The gene expression data were transformed using a van der Waerden normal score.[13] The distribution of the ACR phenotype is skewed. We

Table 6. Over-represented KEGG pathways ($P < 0.01$) in transbands on Chrs 1, 4, 15, and 17

Pathway	*P*	No. of Genes	Genes
Chr 1			
aminoacyl-tRNA biosynthesis	1.79×10^{-4}	14 (37)	*Aars2, Dars, Dars2, Eprs, Hars2, Iars, Iars2, Lars2, Sars, Tars, Tars2, Wars, Wars2, Yars*
cell cycle	0.001	27 (174)	*Anapc1, Anapc4, Atm, Ccna2, Ccnd1, Cdc16, Cdc23, Cdk2, Cdk4, Cdkn1, Cdkn2c, Check2, Cul1, Mad2l1, Mad2l2, Mcm2, Mcm3, Mcm6, Mdm2, Orcsl, Orc3l, Orc4l, Orc6l, Pik1, Smc1a, Trp53*
ECM-receptor interaction	0.001	21 (79)	*Agm, Chad, Col2a1, Col3a1, Col4a1, Col4a2, Col4a4, Fndc3a, Gp5, Hmmr, Itga1, Itga3, Itga6, Itgav, Itgb, Lamc3, Npnt, Sdc3, Spp1, Sv2a, Sv2c*
N-glycan degradation	0.002	7 (15)	*Ag, Hexa, Lct, Manba, Neu1, Neu2, Neu3*
alanine and aspartate metabolism	0.003	11 (33)	*ars2, Acy3, Adsl, Adss, Asl, Ass1, Dars, Dars2, Gad1, Pdha1, Pet112l*
leukocyte transendothelial migration	0.007	24 (106)	*Actn2, Actn3, Cdh5, Cldn14, Cldn6, Cldn9, Ctnnb1, Ctnndl, Cybb, Ezr, Gnai2, Itgb1, Mmp2, Ncf2, Ncf4, Ocln, Pecam1, Pik3r3, Plcg2, Rhoa, Rock1, Thy1, Vav2, Vcam1*
DNA replication	0.009	12 (33)	*Fen1, Mcm2, Mcm3, Mcm7, Pola1, Pola2, Pold3, Prim1, Prim2, Rcf1, Rcf3, Rpa1*
Chr 4			
glutathione metabolism	2.45×10^{-4}	14 (35)	*Ggt7, Gpx3, Gsr, Gsta3, Gsta4, Gstk1, Gstm1, Gsto1, Gsto2, Gstt1, Gstt2, Idh1, Mgst3, Txndc12*
glycine, serine, and threonine metabolism	0.001	15 (42)	*Abp1, Akr1c13, Akr1c6, Akr1e1, Chdh, Gars, Gatm, Gldc, Maoa, Maob, Phgdh, Pisd, Sars, Shmt2, Tarsl2*
bisphenol A degradation	0.001	5 (7)	*Akr1c13, Akr1c6, Akr1e1, Pon2, Pon3*
fatty acid metabolism	0.002	14 (41)	*Acaa2, Acadm, Acads, Acat1, Acox3, Acsl1, Acsl5, Acsl6, Aldh3a2, Cpta1a, Cyp4a10, Cyp4a12b, Dci, Hsd17b10*
glycan structures, degradation	0.002	10 (25)	*Arsb, Fuca1, Fuca2, Galns, Gib1, Gusb, Hgsnat, Hyal1, Man2b1*
arginine and proline metabolism	0.003	11 (30)	*Aldh4a1, Arg2, Asi, Ass1, Ckmt1, Gatm, Glud1, Lcmt1, Oat, P4 ha2, Prodh*
urea cycle and metabolism of amino groups	0.003	10 (26)	*Abp1, Acy1, Aldh3a2, Arg2, Asl, Ass1, Gatm, Maoa, Maob, Srm*
glycerophospholipid metabolism	0.007	16 (57)	*Ard1a, Cris1, Dgkb, Dgkq, Dgkz, Etnk1, Gnpat, Gpd1, Gpd2, Lypla2, Pcyt1a, Pisd, Pla2g6, Pld2, Ppap2c, Ptdss1*
glycosaminoglycan degradation	0.007	6 (13)	*Arsb, Galns, Gib1, Gusb, Hgsnat, Hyal1*
renin-angiotensin system	0.008	8 (17)	*Ace, Agt, Agtra1, Anpep, Lnpep, Mas1, Mme, Ren1*
Chr 15			
nitrogen metabolism	2.34×10^{-5}	9 (20)	*Amt, Asna, Car12, Car14, Car5a, Car5b, Car6, Gls, Gls2*
fatty acid metabolism	3.40×10^{-5}	13 (41)	*Acaa1b, Acadvl, Acox1, Adh1, Adh5, Aldh2, Aldh7a1, Cpt1b, Cyp4a10, Dci, Gcdh, Hsd*
drug metabolism, cytochrome P450	2.40×10^{-4}	16 (68)	*Adh1, Adh5, Cyp2c39, Cyp2c44, Cyp2d22, Cyp2d9, Fmo5, Gsta2, Gstk1, Gstm2, Gstm7, Gsto2, Gstp1, Gstl2, Mgst3, Ugt2a3*
metabolism of xenobiotics by P450	0.001	14 (62)	*Adh1, Adh5, Cyp2c39, Cyp2c44, Dhdh, Gsta2, Gstk1, Gstm2, Gstm7, Gsto2, Gstp1, Gstl2, Mgst3, Ugt2a3*
glutathione metabolism	0.003	9 (35)	*Gtg1, Gsta2, Gstk1, Gstm2, Gstm7, Gsto2, Gstp1, Gstl2, Mgst3*
biosynthesis of unsaterated fatty acids	0.003	8 (29)	*Acaa1b, Acot2, Acot5, Acot8, Acot9, Acox1, Elov16, Fads2*
butanoate metabolism	0.004	9 (36)	*Abat, Akr1c13, Aldh2, Aldh7a1, Hmgcl, Hmgcs2, Hsd17b10, Hsd17b4, Pdha1*
alanine and aspartate metabolism	0.007	8 (33)	*Aars, Abat, Adssl1, Agxt2, Asns, Aspa, Ddo, Pdha1*
Chr 17			
antigen processing and presentation	3.10×10^{-6}	21 (79)	*B2m, H2-Ab1, H2-DMa, H2-DMb1, H2-DMb2, H2-Ea, H2-Eb1, H2-K1, H2-L, H2-M1, H2-M3, H2-Q1, H2-Q1, H2-Q8, H2-Q9, H2-T10, H2-T22, H2-T23, Hspa1l, Nfya, Tap1*
cell adhesion molecules	2.59×10^{-5}	28 (139)	*Cdh3, Cdh4, Cldn11, Cldn23, Cldn7, H2-Ab1, H2-DMa, H2-DMb1, H2-DMb2, H2-Ea, H2-Eb1, H2-K1, H2-L, H2-M1, H2-M3, H2-Q1, H2-Q1, H2-Q8, H2-Q9, H2-T10, H2-T22, H2-T23, Jam2, Negr1, Nrxn3, Ptprf, Sele, Vcam1*

The number of genes present in the transband is given, and the number of genes in the pathway represented on the microarray is indicated in parentheses. The MGI symbols (http://www. informatics.jax.org/) for the genes involved in each pathway on the transband are given.

transformed ACR to the scale log10(ACR + 1) (Supplemental Figure 1). The data were still non-normal because of the large number of unaffected individuals. To protect against spurious linkage that might arise from distributional assumptions, we applied a permutation test to derive the LOD thresholds. Conditional genome probabilities were calculated with a 2-cM step size and an assumed genotype error rate of 0.001. Genome scans were performed with Haley-Knott regression for Chrs 1 to 19 and the X chromosome. 10,000 permutation tests were performed for the autosomes, and 167,131 permutations were performed for the X chromosome for ACR and the gene expression traits.[14] The LOD score thresholds were set at the suggestive level ($P < 0.63$) as 2.23 for the autosomes and 1.44 for the X chromosome and at the significant level ($P < 0.05$) as 3.65 for the autosomes and 2.73 for the X chromosome.[15] Statistical analysis was performed in R/qtl.[16]

A genome-wide level significant measure was calculated for each gene expression trait, and multiple test adjustments were made for the autosomes and X chromosome. The P values for traits (T_i) were calculated from the max LOD score for the autosomes ($LOD_{max}T_i^A$) and for the X chromosome ($LOD_{max}T_i^X$) and the corresponding permutations (Supplemental Figure 2). The Q value package was used to estimate the false discovery rate and the proportion of null hypotheses (π_0).[17] A genome-wide threshold of Q value <0.1 was applied, yielding 8248 significant gene expression traits.

Conditional genome scans are used as a graphical modeling technique for estimating statistical independence between variables. In this approach, a QTL regression model for ACR (M0: $ACR = Q + \varepsilon$) is compared with a model with the candidate gene expression trait (GEX) included as a covariate (M1: $ACR = Q + GEX + \varepsilon$) to determine the extent to which the candidate gene expression acts as a mediator of the QTL effect. Likewise, a QTL model for GEX is constructed (N0: $GEX = Q + \varepsilon$) and compared with a QTL model that includes ACR as a covariate (N1: $GEX = Q + ACR + \varepsilon$). This approach has been used in graphical modeling strategies for QTL and expression-QTL data.[8,19] We applied this strategy to infer causal relationships between candidate genes within the Chr 4 QTL region. In our application, GEX is causal to ACR if the following are satisfied: (*1*) after conditioning ACR on GEX, the LOD score in the Chr 4 region is reduced below the suggestive level ($P < 0.63$, LOD = 2.23) and (*2*) after conditioning GEX on ACR the LOD score in the Chr 4 region is not reduced below the suggestive level. This is a stringent criterion that will reveal the most causal candidates.

Pathway Analysis

The Gene Ontology Consortium has established a controlled vocabulary, gene ontology, to describe gene and gene product relationships.[20] KEGG pathways have been developed to describe relationships over metabolic, signaling, and disease pathways.[21] Hyper-geometric tests were performed with the *GOstats* package with annotation from the package *org.Mm.eg.db* from bioconductor (http://www.bioconductor.org/) to detect over-represented KEGG and gene ontology categories in the QTL regions.[22] The gene universe (background) was defined as the 12,656 unique Entrez Gene Identifiers present in the data set. Gene sets were tested for significance at the threshold level $P < 0.01$ in KEGG and gene ontology categories. This type of pathway analysis reveals over-represented pathways within the regions that overlap with QTL for the ACR phenotype.

Histology

After cervical dislocation, the left kidney from each of ten 13-week-old MRL and SM males was removed. The kidney was halved with a sterile razor blade and fixed overnight in 4% paraformaldehyde. The kidneys were paraffin embedded, sectioned, and stained with hematoxylin and eosin and Periodic Acid Schiff to determine the general histology of the animals. A section from each kidney was also stained with the primary antibody against TLR12 (ab47097; Abcam), using heat-induced antigen retrieval with Tris-EDTA on a Ventana Medical Systems Discovery XT Automated Immunostainer.

ACKNOWLEDGMENTS

This work was funded with grants from DK069381 from the NIDDK (R.K.), the NHLBI NSRA fellowship 1F32 HL095240-02 (RSH), the NIGMS GM076468 (G.A.C.), an American Heart Association fellowship (M.S.L.), AG-NS-0421-07 (S.-W.T.) from the Ellison Medical Foundation, a Glenn Award for Research in Biologic Mechanisms of Aging (S.-W.T.) from the Glenn Foundation for Medical Research, and a National Cancer Institute Cancer Core grant (CA034196) to the Jackson Laboratory. We thank Dana Godfrey, Sue Grindle, and Ken Walsh for excellent technical assistance; Joanne Currer for writing assistance; and Jesse Hammer for preparation of the figures.

DISCLOSURES

None.

REFERENCES

1. Garrett MR, Pezzolesi MG, Korstanje R: Integrating human and rodent data to identify the genetic factors involved in chronic kidney disease. *J Am Soc Nephrol* 21: 398–405, 2010
2. Jansen RC, Nap JP: Genetical genomics: The added value from segregation. *Trends Genet* 17: 388–391, 2001
3. Doorenbos C, Tsaih SW, Sheehan S, Ishimori N, Navis G, Churchill G, Dipetrillo K, Korstanje R: Quantitative trait loci for urinary albumin in crosses between C57BL/6J and A/J inbred mice in the presence and absence of Apoe. *Genetics* 179: 693–699, 2008
4. Morel L, Rudofsky UH, Longmate JA, Schiffenbauer J, Wakeland EK: Polygenic control of susceptibility to murine systemic lupus erythematosus. *Immunity* 1: 219–229, 1994
5. Zhang D, Zhang G, Hayden MS, Greenblatt MB, Bussey C, Flavell RA, Ghosh S: A toll-like receptor that prevents infection by uropathogenic bacteria. *Science* 303: 1522–1526, 2004
6. Smith C, Goldsmith C, Eppig J: The Mammalian Phenotype Ontology as a tool for annotating, analyzing and comparing phenotypic information. *Genome Biol* 6: R7, 2005
7. Murayama S, Yagyu S, Higo K, Ye C, Mizuno T, Oyabu A, Ito M, Morita H, Maeda K, Serikawa T, Matsuyama M: A genetic locus susceptible to the overt proteinuria in BUF/Mna rat. *Mamm Genome* 9: 886–888, 1998
8. Rogus JJ, Poznik GD, Pezzolesi MG, Smiles AM, Dunn J, Walker W, Wanic K, Moczulski D, Canani L, Araki S, Makita Y, Warram JH,

Krolewski AS: High-density single nucleotide polymorphism genome-wide linkage scan for susceptibility genes for diabetic nephropathy in type 1 diabetes: Discordant sibpair approach. *Diabetes* 57: 2519–2526, 2008
9. Schulz A, Standke D, Kovacevic L, Mostler M, Kossmehl P, Stoll M, Kreutz R: A major gene locus links early onset albuminuria with renal interstitial fibrosis in the MWF rat with polygenetic albuminuria. *J Am Soc Nephrol* 14: 3081–3089, 2003
10. Grindle S, Garganta C, Sheehan S, Gile J, Lapierre A, Whitmore H, Paigen B, DiPetrillo K: Validation of high-throughput methods for measuring blood urea nitrogen and urinary albumin concentrations in mice. *Comp Med* 56: 482–486, 2006
11. Cox A, Ackert-Bicknell CL, Dumont BL, Ding Y, Bell JT, Brockmann GA, Wergedal JE, Bult C, Paigen B, Flint J, Tsaih SW, Churchill GA, Broman KW: A new standard genetic map for the laboratory mouse. *Genetics* 182: 1335–1344, 2009
12. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP: Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 31: e15, 2003
13. Lehmann, E: *Nonparametrics: Statistical Methods Based on Ranks*, San Francisco, Holden-Day, 1975
14. Broman K, Sen S, Owens S, Manichaikul A, Southard-Smith E, Churchill G: The X chromosome in quantitative trait locus mapping. *Genetics* 174: 2151–2158, 2006
15. Lander E, Kruglyak L: Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. *Nat Genet* 11: 241–247, 1995
16. Broman K, Wu H, Sen S, Churchill G: R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19: 889–890, 2003
17. Storey J: A direct approach to false discovery rates. *J R Stat Soc Series B* 64: 479–498, 2002
18. Li R, Tsaih SW, Shockley K, Stylianou IM, Wergedal J, Paigen B, Churchill GA: Structural model analysis of multiple quantitative traits. *PLoS Genet* 2: e114, 2006
19. Chaibub Neto E, Keller MP, Attie AD, Yandell BS: Causal graphical models in systems genetics: A unified framework for joint interference of causal network and genetic architecture for correlated phenotypes. *Ann Appl Stat* 4: 320–339, 2010
20. Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C, Richter J, Rubin GM, Blake JA, Bult C, Dolan M, Drabkin H, Eppig JT, Hill DP, Ni L, Ringwald M, Balakrishnan R, Cherry JM, Christie KR, Costanzo MC, Dwight SS, Engel S, Fisk DG, Hirschman JE, Hong EL, Nash RS, Sethuraman A, Theesfeld CL, Botstein D, Dolinski K, Feierbach B, Berardini T, Mundodi S, Rhee SY, Apweiler R, Barrell D, Camon E, Dimmer E, Lee V, Chisholm R, Gaudet P, Kibbe W, Kishore R, Schwarz EM, Sternberg P, Gwinn M, Hannick L, Wortman J, Berriman M, Wood V, de la Cruz N, Tonellato P, Jaiswal P, Seigfried T, White R: The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res* 32: D258–D261, 2004
21. Kanehisa M, Goto S: KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28: 27–30, 2000
22. Falcon S, Gentleman R: Using GOstats to test gene lists for GO term association. *Bioinformatics* 23: 257–258, 2007

Supplemental information for this article is available online at http://www.jasn.org/.