RESEARCH ARTICLE

Genomic dissection and prioritizing of candidate genes of QTL for regulating spontaneous arthritis on chromosome 1 in mice deficient for interleukin-1 receptor antagonist

YANHONG CAO[1,2], JIFEI ZHANG[5], YAN JIAO[2], JIAN YAN[3], FENG JIAO[2], XIAOYUN LIU[2], ROBERT W. WILLIAMS[4], KAREN A. HASTY[2], JOHN M. STUART[3] and WEIKUAN GU[2]*

[1]*Institute of Kaschin-Beck Disease, Center for Endemic Disease Control, Centers for Disease Control and Prevention, Harbin Medical University, Harbin, 150081, People's Republic of China*
[2]*Department of Orthopaedic Surgery, Campbell Clinic and Pathology,* [3]*Department of Medicine and* [4]*Department of Anatomy and Neurobiology, University of Tennessee Health Science Center, Memphis, TN 38163, USA*
[5]*Mudanjiang Medical College, Mudanjiang, 157001, People's Republic of China*

Abstract

Rheumatoid arthritis is a heterogeneous disease with clinical and biological polymorphisms. IL-1RN is a protein that binds to interleukin-1 (IL-1) receptors and inhibits the binding of IL-1-alpha and IL-1-beta. IL-1RN levels are elevated in the blood of patients with a variety of infectious, immune, and traumatic conditions. Balb/c mice deficient in IL-1ra (mouse gene of IL-1RN) develop spontaneous autoimmune arthritis while DBA/1 mice deficient in IL-1ra do not. Previously, we identified a major QTL that regulates the susceptibility to arthritis in Balb/c mice with IL-1ra deficiency. In this study, we found that the QTL may contain two peaks that are regulated by two sets of candidate genes. By haplotype analysis, the total genomic regions of candidate genes were reduced from about 19 Mbp to approximately 9 Mbp. The total number of candidate genes was reduced from 208 to 21.

[Cao Y., Zhang J., Jiao Y., Yan J., Jiao F., Liu X., Williams R. W., Hasty K. A., Stuart J. M. and Gu W. 2012 Genomic dissection and prioritizing of candidate genes of QTL for regulating spontaneous arthritis on chromosome 1 in mice deficient for interleukin-1 receptor antagonist. *J. Genet.* **91**, 119–128]

Introduction

Rheumatoid arthritis (RA), like other common polygenic autoimmune diseases, is characterized by genetic risk factor make-up and phenotypic heterogeneity. Previously, by whole-genome mapping, we obtained a major quantitative trait locus (QTL) on chromosome 1 that regulates spontaneous arthritis in interleukin-1 receptor antagonist (IL-1ra)-deficient mice (Jiao *et al.* 2011). The QTL is flanked by microsatellite markers *D1mit265* and *D1mit540* with peaks near *D1Mit55* and *D1Mit425* at 82.6 cM. The QTL accounts for as much as 12% of the phenotypic variation in susceptibility to spontaneous arthritis (Jiao *et al.* 2011). According to data in the Ensembl genome database, the QTL region contains 208 known transcripts. Using an integrative approach, we obtained a list of candidate genes through bioinformatics, gene expression profiles, and molecular pathway analysis. According to known gene functions reported in PubMed and Online Mendelian Inheritance in Man (ONIM), *Mr1*, *Pla2g4a*, and *Fasl* are likely candidate genes. According to differential expression levels from microarray analysis and gene function, 11 genes were selected as favourable candidates. Three of the 11 genes, *Prg4*, *Ptgs2*, and *Mr1*, correlated with the IL-1ra pathway (Jiao *et al.* 2011).

However, the majority of candidate genes from the three approaches appear different from each other. Only *Mr1* and *Ptgs2* appear in both bioinformatics and pathway analyses. Further analysis of single nucleotide polymorphisms (SNP) of genes and gene expression profiles at the site of disease may improve candidate gene selection. One limitation of our previous analysis was that data were not from spleen, the tissue most relevant to spontaneous arthritis (Jiao *et al.* 2011). At the time of our previous study, gene expression profiles were available from only livers and kidney from 46 B×D RI mouse strains in the GeneNetwork (http://www.genenetwork.org/). The gene network analysis therefore was limited to genes at a detectable level in both kidney and liver of those mouse strains.

*For correspondence. E-mail: wgu@utusc.edu.

Keywords. arthritis; interleukin-1 receptor antagonist; microarray; mouse; QTL; SNP.

Using gene expression profiles and gene network data from spleen and SNP data, we report here a more extensive analysis of candidate genes within the QTL region. Our SNP analysis indicated that there are two subloci within the QTL region. We successfully narrowed the QTL region and identified a most favourable candidate gene.

Materials and methods

Mice

BALB/c-based IL-1ra−/− mice (KO) were kindly provided by Dr Yoichiro Iwakura (Center for Experimental Medicine, Institute of Medical Science, University of Tokyo, Tokyo, Japan). They were bred to DBA/1 mice to generate DBA/1-based IL-1ra −/− mice that are arthritis resistant. F_2 mice were generated using two parents, BALB/c and DBA/1; both are IL-1ra deficient (Jiao *et al.* 2011).

RNA extraction

RNA from IL-1ra-deficient mice was extracted from different parts of brains by using a Trizol reagent (Invitrogen, Carlsbad, USA). Total RNA was purified using the RNeasy MinElute Cleanup kit (Qiagen, Valencia, USA). RNA quality and integrity were checked by Agilent Bioanalyzer (Agilent, Santa Clara, USA).

Microarray experiment

High-quality total RNA (200 ng) with an RNA integrity score of more than seven was used to generate cDNA and cRNA by using the Illumina® TotalPrep™ RNA Amplification kit (Ambion, Foster City, USA). For each of six individual samples,1.5 μg of cRNA sample were hybridized overnight to the Mouse-6 v1B BeadChip in a multiple-step procedure according to the manufacturer's instructions; the chips were washed, dried, and scanned on the BeadArray Reader (Illumina, San Diego, USA), and raw data were generated using BeadStudio 2.3.41 (Illumina, San Diego, USA).

Raw data were normalized with the quantile method by using BeadStudio software. Upon detection ($P < 0.05$) of probe signals and fold change (≤ 2) of expression level filtering, statistical analysis was performed using EDGE software (Longmont, USA) to identify differentially expressed genes, which were subject to hierarchical and functional clustering using Cluster/TreeView (Frederick, USA) and bioinformatics DAVID tools (http://david.abcc.ncifcrf.gov), respectively. To identify differentially expressed genes between different groups, a cutoff value of 2-fold changes was chosen based on published studies (Jeffrey *et al.* 2002; Townsend and Hartl 2002; Leek *et al*. 2006; Yan *et al.* 2007; Sharov *et al.* 2008).

Bioinformatics evaluation of candidates

A gene network analysis was performed using the bioinformatics approach. To explore the potential interaction between candidate genes and IL-1ra, we analysed the interaction between potential candidate genes and IL-1, IL-1r, and IL-1ra by using gene expression profiles of 108 strains, most of the RI strains were from C57BL/6J (B) and DBA/2J (D) mice. The 108 samples included 81 B×D strains; both parental strains and both reciprocal F_1 hybrids (B6D2F1 and D2B6F1), and 26 other common, inbred strains were quantified. In most cases, two arrays were processed per strain (http://www.genenetwork.org/dbdoc/UTHSC_SPL_RMA_1210.html).

The gene expression profiles of spleen from 108 mice were uploaded to the GeneNetwork database (http://www.genenetwork.org/) and interaction among genes were determined using gene expression profiles from mouse spleen at GeneNetwork. The gene expressions of spleen of the mice were profiled using the GeneChip Mouse Gene 1.0 ST (Affymetrix, Santa Clara, USA) array that contains approximately 34,728 probe sets that target approximately 29,000 well-defined transcripts (RefSeq mRNA isoforms) and essentially all known protein coding genes in mice. Matrix correlation analyses among genes in each set were conducted following instructions of GeneNetwork.

SNP searching

Known SNP genomic regions within the QTL region were examined using available SNP databases such as the Roche SNP database (http://mousesnp.roche.com/), MGI mouse SNP database (http://www.informatics.jax.org/javawi2/servlet/WIFetch?page=snpQF), and the NCBI SNP database (http://www.ncbi.nlm.nih.gov/snp/).

Linkage analysis

To confirm whether the QTL for spontaneous arthritis on chromosome 1 is regulated by two sets of genes, QTX (http://www.mapmanager.org/mmQTX.html) was used to analyse the QTL region on chromosome 1 that regulates spontaneous arthritis. Data collection for arthritis phenotype and the procedure for genotypes have been reported (Zhou *et al.* 2005; Jiao *et al.* 2011). Data, organization and patterns requested by the software QTX and the procedures followed those previously reported (Jiao *et al.* 2007).

Results

SNP distribution pattern suggests two sets of genetic factors regulating the QTL

In the region between *D1mit265* and *D1mit540*, there are 66 SNPs (see table 1 in electronic supplementary material at http://www.ias.ac.in/jgenet). However, only half of them

were polymorphic between Balb/c and DBA/1 mice. Further, polymorphic markers are not evenly distributed along this genome region (figure 1). All 33 polymorphic markers are concentrated in the two genomic regions. While 18 markers are located between 150,927,537 bp and 157,073,549 bp, the other 15 are located between 165,318,402 bp and 169,640,436 bp (figure 1). The distribution patterns of SNP raise a question whether one or two genes or sets of genes regulate spontaneous arthritis under QTL. Looking at our previous mapping figure, we realized that there are two peaks on this QTL (figure 1; Jiao *et al.* 2011).

As indicated by the sizes of those two genomic regions, assuming the candidate gene regions are correct, the total genomic region for QTL is reduced from the original 19 Mbp to approximately 9 Mbp.

Linkage remapping with a different program confirms the two peaks of the QTL

In our previous report, R/qtlbim software was used for Bayesian QTL mapping of ordinal traits to locate QTL on chromosome 1 (Jiao *et al.* 2011). Although it appeared that the QTL has two peaks, they were not very obvious (figure 1; Jiao *et al.* 2011). We therefore conducted another linkage mapping on mouse chromosome 1 by using another QTL mapping program, QTX (http://www.mapmanager.org/mmQTX.html), with a total of 561 F_2 progeny. While we used the same marker, *D1mit265*, at one end of the QTL, we used mostly different markers on the other sites of the QTL region. As shown in figure 2, two obvious peaks appear for this QTL. *D1mit55* is located between 157,013,984 and 157,014,134 bp, in the genomic region of the first set of polymorphic markers between Balb/c and DBA/1 mice. *D1mit110* is located between 169,688,648 and 169,688,784 bp, along the edge of the second set of polymorphic markers. On the distal end of the QTL, *D1mit426* is located between 182,341,840 and 182,341,923 bp outside the previous mapping marker, *D1mit540*, which is located between 170,685,914 and 170,686,059 bp. With the new mapping information and previous data, we are convinced that two sets of genetic elements contribute to the phenotype of QTL.

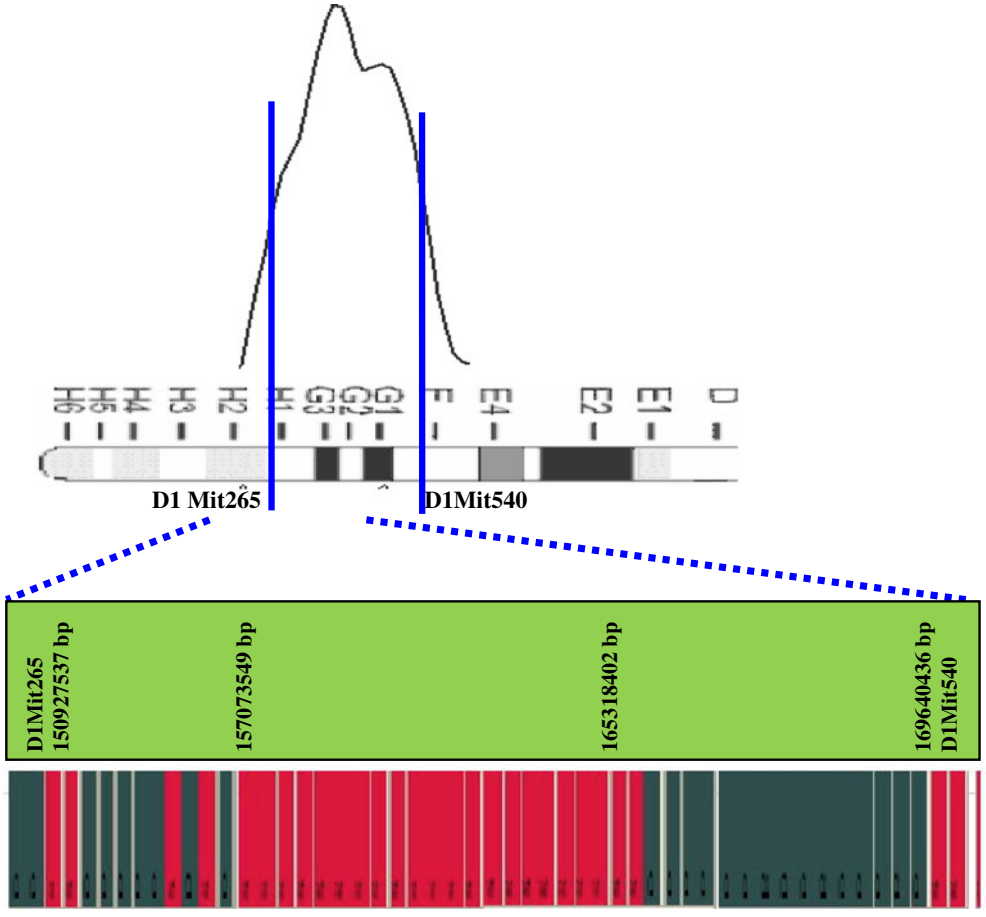


Figure 1. SNP distribution along mouse chromosome 1 between Balb/c and DBA/1. All 33 polymorphic markers are concentrated in the two genomic regions. While 18 markers are located between 150,927,537 bp and 157,073,549 bp, the other 15 are located between 165,318,402 bp and 169,640,436 bp.

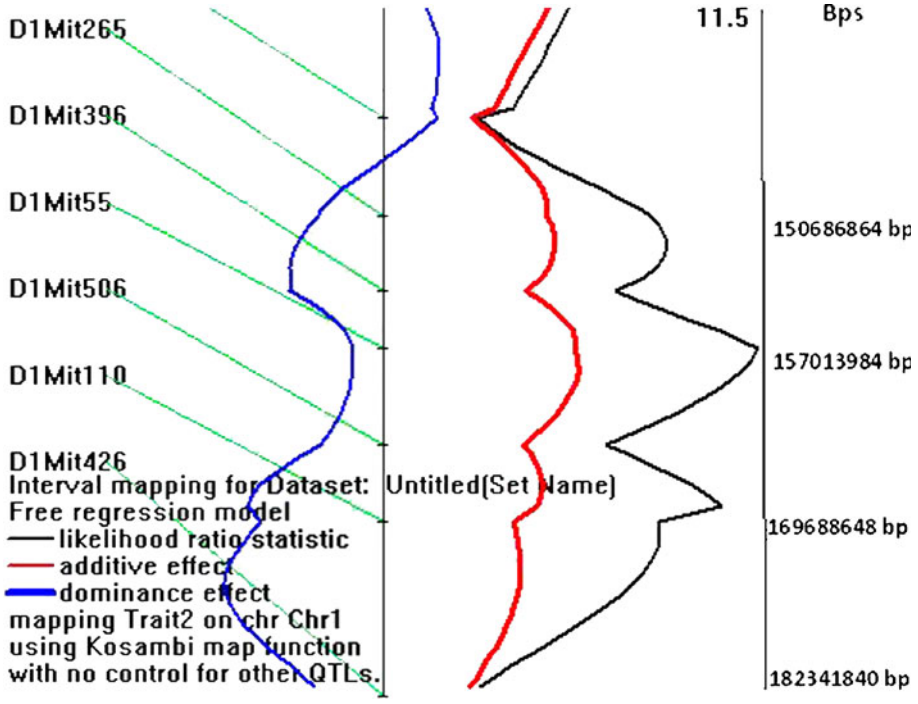


Figure 2. QTL mapping indicates potentially two genes or sets of genes regulate spontaneous arthritis under the QTL on mouse chromosome 1. *D1mit55* is located between 157,013,984 and 157,014,134 bp. *D1mit110* is located between 169,688,648 and 169,688,784 bp. *D1mit426* is located between 182,341,840 and 182,341,923 bp outside the previous mapping marker, *D1mit540*, which is located between 170,685,914 and 170,686,059 bp.

Candidate genes in two potential candidate regions

The total number of candidate genes with SNP in two genomic regions is 21 (tables 1 and 2). We then examined the potential candidate genes that contain SNP in the two genomic regions. In the first region, between 150,927,537 bp and 157,073,549 bp, we found 18 SNP located in 10 genes (table 1). Within those genes, cytosolic phospholipase

Table 1. Candidate genes and SNP in the first candidate region of the QTL on chromosome 1.

SNP ID (dbSNP build 128)	Map position (NCBI build 37) bp	Rs orient	Gene : dbSNP function class	DBA/1J	BALB/cByJ	Differential expression
rs13476172	Chr. 1: 150927537	f		G	A	–
rs13476173	Chr. 1: 151137796	f		C	T	–
rs4222678	Chr. 1: 151719657	r	Pla2g4a : intron	A	G	None
rs13476174	Chr. 1: 151748341	f	Pla2g4a : coding-synonymous	A	G	None
rs13476175	Chr. 1: 151907070	f		C	G	
rs13476176	Chr. 1: 152240071	f	BC003331 : mRNA-UTR, Tpr : mRNA-UTR	T	C	None
rs13476177	Chr. 1: 152491007	f	Hmcn1 : within coordinates of	C	T	–
rs13476179	Chr. 1: 152919406	f		A	G	–
rs6398113	Chr. 1: 153539523	f	Fam129a : intron	C	T	–
rs6196928	Chr. 1: 154430872	f	Rgl1 : intron	G	A	None
rs13476183	Chr. 1: 154621425	f	Arpc5 : coding-synonymous	C	A	None
rs13476184	Chr. 1: 155032005	f	Lamc2 : intron	G	G	None
rs13476185	Chr. 1: 155557262	f		T	C	
rs13476186	Chr. 1: 155739113	f	5830403L16Rik : within coordinates of, Teddm1 : coding-synonymous	T	C	None
rs13476187	Chr. 1: 156052564	f	Zfp648 : within coordinates of	T	C	–
rs13476188	Chr. 1: 156575913	f		C	G	–
rs13476189	Chr. 1: 156774308	f		C	T	–
rs13476190	Chr. 1: 157073549	f	BC034090 : coding-nonsynonymous	G	A	–

Table 2. Candidate genes and SNP in the second candidate region of the QTL on chromosome 1.

SNP ID (dbSNP build 128)	Map position (NCBI build 37)	Rs orient	Gene : dbSNP function class	DBA/1J	BALB/cByJ	Differential expression
rs6323784	Chr. 1: 165318402	f	Gorab : intron	T	C	None
rs13476212	Chr. 1: 165540566	f		C	T	–
rs13476214	Chr. 1: 166007852	f	Sell : intron	C	T	None
rs13476215	Chr. 1: 166341970	f	Nme7 : intron	A	G	–
rs6154568	Chr. 1: 166650341	f		C	T	
rs13459054	Chr. 1: 167145883	r	Tiprl : coding-synonymous	C	G	None
rs13476217	Chr. 1: 167324177	f	Dcaf6 : intron	A	G	
rs6234151	Chr. 1: 167706427	f		T	C	
rs6363233	Chr. 1: 167838560	f	Pou2f1 : intron	T	C	None
rs13476219	Chr. 1: 168178248	f	Ildr2 : locus-region (upstream)	T	C	None
rs13476220	Chr. 1: 168426134	f	Gm4846 : coding-synonymous	A	G	–
rs13476221	Chr. 1: 168913139	f	Gm16701 : within coordinates of	G	A	–
rs13476222	Chr. 1: 169310123	f	Mgst3 : intron	T	C	None
rs6272538	Chr. 1: 169483394	f		G	C	–
rs13476223	Chr. 1: 169640436	f	Lmx1a : intron	C	G	None

A2-alpha (*Pla2g4a*) has been considered as one of the favourite candidate genes (Jiao *et al.* 2011). *Pla2g4a* contains two polymorphisms between DBA/1 and Balb/c mice. One is located in the intron at position 151,719,657 bp, while the other is located in the coding region at 151,748,341 bp. However, the one in the coding region is a synonymous polymorphism. Three previously favourite candidate genes have no polymorphisms between Balb/c and DBA/1 mice: prostaglandin-endoperoxide synthase 2 (*Ptgs2* or *Cox2*) between 151,951,015 bp and 151,951,160 bp; the major histocompatibility complex, class I-related (*Mr1*) between 156,975,076 bp and 156,993,934 bp; and proteoglycan 4 (*Prg4*) between 152,296,946 bp and 152,313,295 bp.

The second set of candidate genes is in a polymorphic region without disruption between DBA/1 and Balb/c mice. Thus, this region is a haploid type that contains 11 genes, each with at least one polymorphism (table 2).

According to our gene expression analysis using spleens from mice at one month to four months (see table 2 in

Table 3. Dysregulated genes specific in IL-1RN deficient BALB/c mice.

Category	Probe ID	Gene symbol	Gene title	Fold change
Protein folding				
	1444500_at	*Ahsa1*	Activator of heat shock 90-kDa protein ATPase homolog 1	−2.5
	1458729_at	*Fkbp4*	FK506 binding protein 4	−2.8
	1416288_at	*Dnaja1*	DnaJ (Hsp40) homolog, subfamily A, member 1	−42.8
	1416755_at	*Dnajb1*	DnaJ (Hsp40) homolog, subfamily B, member 1	−2.5
	1437497_a_at	*Hspca*	Heat shock protein 1, alpha	−2.2
	1423566_a_at	*Hsp105*	Heat shock protein 105	−5.4
	1452388_at	*Hspa1a*	Heat shock protein 1A	−20.9
	1427126_at	*Hspa1b*	Heat shock protein 1B	−19.8
	1431182_at	*Hspa8*	Heat shock protein 8	−4.5
Cell cycle				
	1417656_at	*Mybl2*	Myeloblastosis oncogene-like 2	2.4
	1452047_at	*Cacybp*	Calcyclin binding protein	−2.3
	1452040_a_at	*Cdca3*	Cell division cycle associated 3	2.3
	1442569_at	*Cct4*	Chaperonin subunit 4 (delta)	−2.9
	1434611_at	*Rnf123*	Ring finger protein 123	2.0
Proteolysis and peptidolysis	1451019_at	*Ctsf*	Cathepsin F	2.3
	1424936_a_at	*Dnahc8*	Dynein, axonemal, heavy chain 8	−2.1
	1426785_s_at	*Mgll*	Monoglyceride lipase	2.1
	1448107_x_at	*Klk6*	Kallikrein 6	−5.2
Transcription regulation	1456350_at	*Zfp451*	Zinc finger protein 451	−2.2
	1455210_at	*Zhx2*		4.4
	1437187_at	*E2f7*	E2F transcription factor 7	2.4

Table 3 (*contd*).

Category	Probe ID	Gene symbol	Gene title	Fold change
Transport				
	1453255_at	*Slc43a1*	Solute carrier family 43, member 1	2.2
	1416464_at	*Slc4a1*	Solute carrier family 4 (anion exchanger), member 1	2.1
	1417636_at	*Slc6a9*	Solute carrier family 6 (neurotransmitter transporter member 9)	2.3
	1424455_at	*Gprasp1*	G protein-coupled receptor associated sorting protein 1	−2.4
Signal transduction				
	1446518_at	*Rock1*	Rho-associated coiled-coil forming kinase 1	−2.3
	1442263_at	*Rgs13*	regulator of G-protein signaling 13	−3.4
	1425456_a_at	*Map2k3*	Mitogen-activated protein kinase kinase 3	2.2
	1436739_at	*Agtr1*	Angiotensin receptor 1	2.0
	1419127_at	*Npy*	Neuropeptide Y	2.7
	1428391_at	*Rab3il1*	RAB3A interacting protein (rabin3)-like 1	2.1
Immune response				
	1418809_at	*Pira1*	Paired-Ig-like receptor A1	6.3
	1425247_a_at	*Igh-4*	Immunoglobulin heavy chain 4 (serum IgG1)	5.0
	1422924_at	*Tnfsf9*	Tumour necrosis factor superfamily member 9	−3.4
	1424775_at	*Oas1a*	2'−5' Oligoadenylate synthetase 1A	−16.3
	1421264_at	*Butr1*	Butyrophilin related 1	2.3
	1419569_a_at	*Isg20*	Interferon-stimulated protein	2.1
	1451962_at	*Igk-V8*	Similar to monoclonal antibody 17-1A, light chain	2.0
Others				
	1452132_at	*0610007A15Rik*	RIKEN cDNA 0610007A15 gene	2.2
	1433110_at	*5830474E16Rik*	RIKEN cDNA 5830474E16 gene	−2.0
	1435660_at	*5830484A20Rik*	RIKEN cDNA 5830484A20 gene	2.9
	1452848_at	*5930418K15Rik*	RIKEN cDNA 5930418K15 gene	−2.2
	1436434_at	*9230110J10*	hypothetical protein 9230110J10	2.2
	1439450_x_at	*A230046K03Rik*	RIKEN cDNA A230046K03 gene	−2.5
	1440927_x_at	*A330102K04Rik*	RIKEN cDNA A330102K04 gene	3.4
	1427503_at	*AI324046*	Expressed sequence AI324046	5.1
	1433765_at	*B230113M03Rik*	RIKEN cDNA B230113M03 gene	2.4
	1423678_at	*BC017643*	cDNA sequence BC017643	−2.1
	1452218_at	*BC018601*	cDNA sequence BC018601	−2.4
	1460399_at	*BC018601*	cDNA sequence BC018601	−2.1
	1425704_at	*BC022224*	cDNA sequence BC022224	2.3
	1454630_at	*BC034054*	cDNA sequence BC034054	2.2
	1440227_at	*BF642829*	Expressed sequence BF642829	−4.7
	1442845_at	*C130075A20Rik*	RIKEN cDNA C130075A20 gene	−2.1
	1460645_at	*Chordc1*	Cysteine, histidine-rich domain containing zinc-binding protein 1	−2.5
	1448111_at	*Ctps2*	Cytidine 5'-triphosphate synthase 2	−6.2
	1423693_at	*Ela1*	Elastase 1, pancreatic	−4.4
	1422924_at	*EST*	EST	−3.4
	1435579_at	*EST*	Transcribed locus	2.1
	1436320_at	*EST*	CDNA clone IMAGE:1328649	2.2
	1439348_at	*EST*	EST	−2.4
	1441475_at	*EST*	EST	−2.5
	1447167_at	*EST*	EST	−2.3
	1447706_at	*EST*	EST	−2.2
	1448057_at	*EST*	Transcribed locus	2.4
	1455210_at	*EST*	EST	4.4
	1457769_at	*EST*	3 Days neonate thymus cDNA,	−2.1
	1439290_at	*Fxr1h*	Fragile X mental retardation gene 1,	−2.1
	1420901_a_at	*Hk1*	Hexokinase 1	2.1
	1448325_at	*Myd116*	Myeloid differentiation primary response gene 116	2.3
	1422890_at	*Pcdh18*	Protocadherin 18	−2.3
	1455391_at	*Rad23a*	RAD23a homolog (S. cerevisiae)	2.1
	1425288_at	*Samd11*	Sterile alpha motif domain containing 11	3.0
	1416688_at	*Snap91*	Synaptosomal-associated protein 91	−2.0
	1450556_at	*Spnb1*	Spectrin beta 1	2.3
	1452666_a_at	*Tmcc2*	Transmembrane and coiled-coil domains 2	2.4
	1419063_at	*Ugt8*	UDP-Glucuronosyltransferase 8	2.5

electronic supplementary material), none of the genes is differentially expressed between IL-1ra and wildtype Balb/c mice.

One previous favorite candidate gene, Fas ligand (TNF superfamily, member 6) (*Fasl*) (163,710,820–163,718, 626 bp) is located outside of the two candidate regions in the middle of a nonpolymorphic region.

Expression level of Il-1ra relevant genes in spleen of KO and wildtype mice

Gene expression levels of spleen from mice at 1 to 4 months of age were analysed. A total of 216 genes showed differential expression between KO and wildtype Balb/c mice at both 1 and 4 months of age (see table 2 in electronic supplementary material). To characterize the biological meaning of the phase-specific changes, functional profiling was performed using DAVID tools (http://david.abcc.ncifcrf.gov). Table 3 shows major functional clusters associated with the early and late phases of disease. In the early phase of disease, melanin biosynthesis and mast cell-derived signalling activities were suppressed, whereas response to external stimulus/stress and signalling associated with acute-phase response were stimulated. In the late phase of disease, protein folding, cellular protein metabolism, and protein binding activities were inhibited, but haematopoiesis, alternative splicing, phosphorylation, and regulatory action of the cell cycle were enhanced.

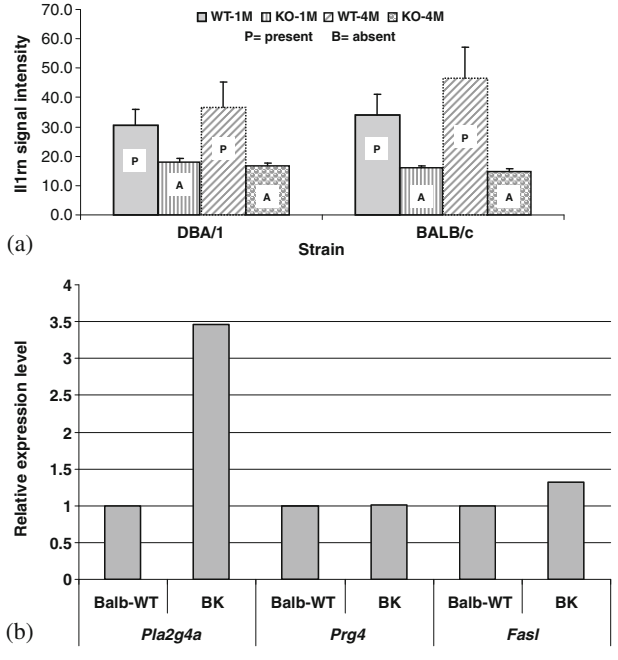


Figure 3. The expression level of IL-1RN and three candidate genes. Data for expression level of IL-1RN is from microarray (3A); while the expression of *Pla2g4a*, *Prg4*, and *Fas1* are from real time PCR (3B).

In particular we evaluated the validity of gene engineering in diseased mice by checking the expression pattern of IL-1RN in the study mice. As expected, expression was significantly suppressed in KO mice of BALB/c background (figure 3a). The expression level of IL-1RN in DBA/1 KO mice was also examined against that in DBA/1 background (figure 3a). Based on the MAS5 detection algorithm, all individual KO mice showed 'absent' calls for IL-1RN expression, suggesting that this transcript was undetectable. Meanwhile, no significant difference was seen in the expression of *IL-1a* and *IL-1b* between wildtype and KO mice (data not shown).

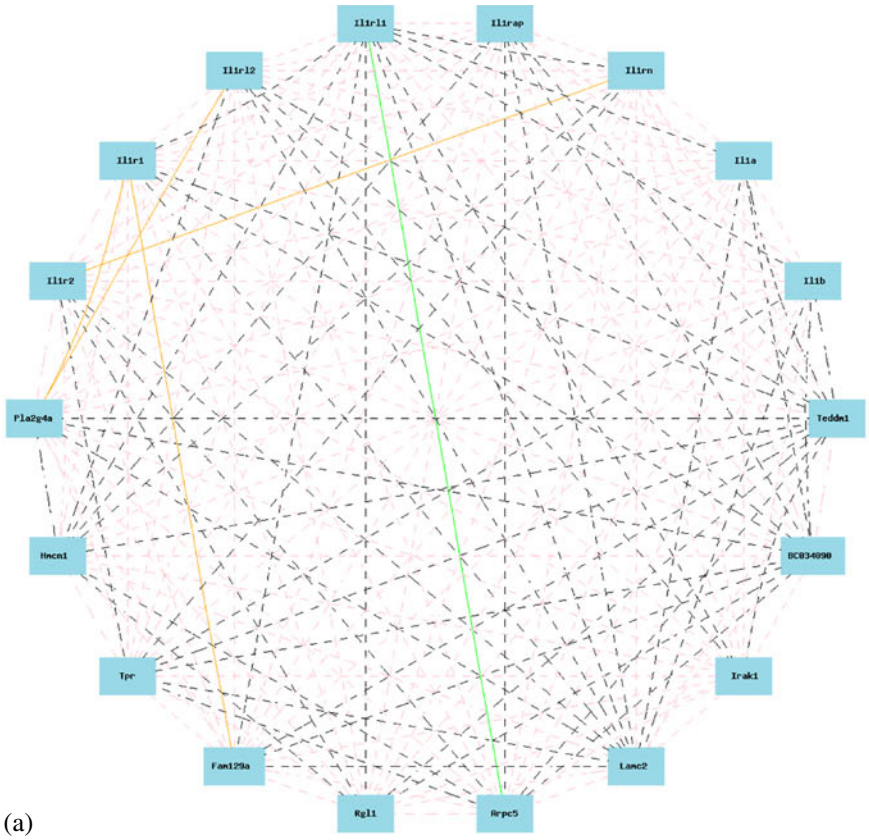
The expression level of three candidate genes, *Pla2g4a*, *Prg4*, and *Ptgs2* was also evaluated by real time PCR using BioMark HD system (BioMark, San Francisco, USA). Preliminary data indicated that the expression level of *Pla2g4a* in Balb/c knockout mice is higher than that in wildtype Babl/c mice; while the expression levels of other two genes showed no significant difference between Balb/c knockout and wildtype (figure 3b).

Pathway analysis

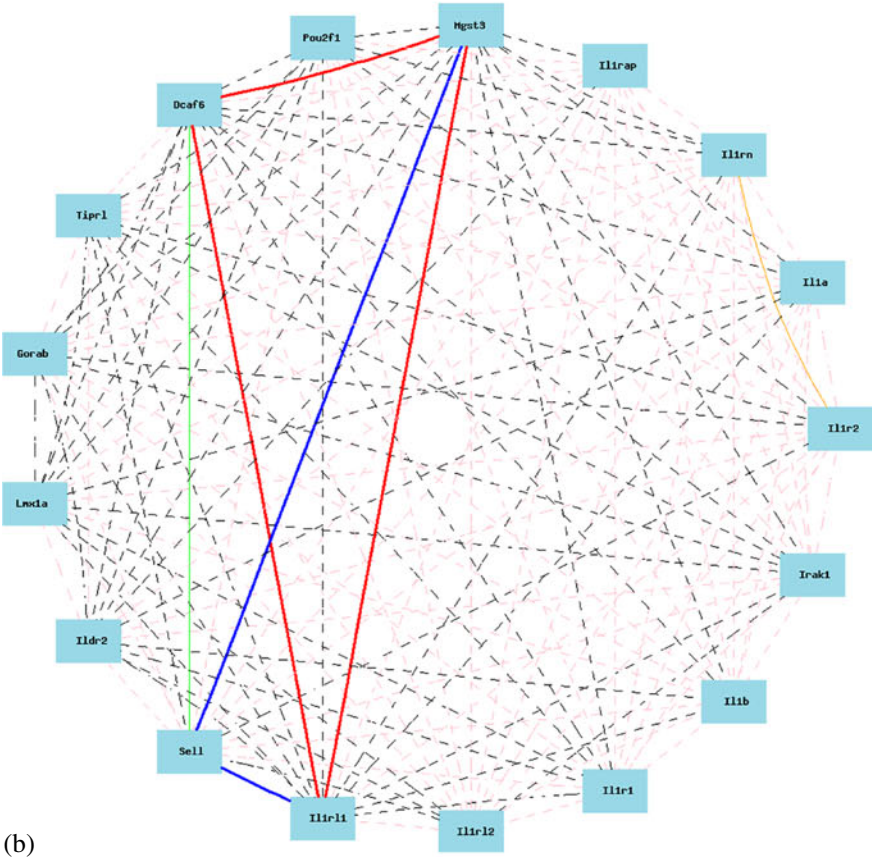
Our investigation of gene networks focussed on the polymorphic genes from two sets of candidate regions. The interactions of the genes listed in tables 1 and 2 with the relevant genes of IL-1RN were examined. The relevant IL-1RN genes included IL-1 alpha (*IL-1a*), IL-1 beta (*IL-1b*), IL-1 receptor type I (*IL-1r1*), IL-1 receptor type II (*IL-1r2*), IL-1 receptor-associated kinase 1 (*Irak1*), IL-1 receptor-like 1 (*IL-1rl1*), IL-1 receptor-like 2 (*IL-1rl2*), IL-1 receptor accessory protein (*IL-1rap*), and IL-1 receptor antagonist (*IL-1ra*).

In the first region of candidate genes, eight genes were analysed against nine relevant IL-1RN genes (see table 3 in electronic supplementary material). There are five significant interactions among there genes, including four positive and one negative interaction (figure 4a). The four positive interactions were between *Fam129a* and *IL-1r1*, *IL-1ra* and *IL-1r2*, *Pla2g4a* and *IL-1r1*, and *Pla2g4a* and *IL-1r2*. One negative interaction occurred between *Arpc5* and *IL-1r1*. *Pla2g4a* interacted with both *IL-1r1* and *IL-1r2* and stood out as the most favourite candidate; it has also been linked to arthritis (Hegen *et al.* 2003; Sharov *et al.* 2008; Raichel *et al.* 2008), and has a polymorphism between Balb/c and DBA/1 mice (table 1).

In the second region of candidate genes, also eight genes were analysed against nine relevant IL-1RN genes (see table 4 in electronic supplementary material). There were eight significant interactions among there genes: five positive and three negative (figure 4b). The five positive interactions were between *Decaf6* and *IL-1rl1*, *Decaf6* and *Mgst3*, *IL-1ra* and *IL-1r2*, *Mgst3* and *IL-1rL1*, and *Sell* and *Pou2fl*. Three negative interactions occurred between *Sell* and *Decaf6*, *Sell* and *IL-1rl1*, and *Sell* and *Mgst3*. There was no obvious candidate gene, although *Sell* had the most interactions with other genes.



(a)



(b)

Figure 4. The potential gene–gene interactions among candidate genes and their pathways.

Table 4. Regulatory genomic regions of candidate genes in first QTL region.

Dataset	Trait ID	Symbol	Description	Location	Mean	N cases	Max LRS	Max LRS location Chr. and Mb
UTHSC_SPL_RMA_1210	10358982	*Mr1*	Major histocompatibility complex, class I-related	Chr.1: 156.975008	9.357	109	21.5	Chr.1: 156.456242
UTHSC_SPL_RMA_1210	10358434	*Pla2g4a*	Phospholipase A2, group IVA (cytosolic, calcium-dependent)	Chr.1: 151.676748	9.674	109	14.7	Chr.1: 149.096503
UTHSC_SPL_RMA_1210	10358476	*Prg4*	Proteoglycan 4 (megakaryocyte stimulating factor, articular superficial zone protein)	Chr.1: 152.296999	6.881	109	11.2	Chr.11: 13.574420
UTHSC_SPL_RMA_1210	10350516	*Ptgs2*	Prostaglandin-endoperoxide synthase 2	Chr.1: 151.947230	6.987	109	13.1	Chr.11: 35.986496

Potential interactions between candidate genes in QTL region and rest of the genome

We assume that there are complicated interactions between genes within the QTL regions and the rest of the genome. To explore this aspect, we further examined the transcriptome for regulation of candidate genes within the first QTL region using gene expression profiles generated using spleen mRNA from B×D population. In particular, we focussed on several favourite candidates, *Pla2g4a*, *Ptgs2*, *Mr1*, and *Prg4*. As shown in table 4, *Pla2g4a* and *Mr1* are regulated by genes within or nearby the QTL region on chromosome 1, while *Ptgs2* and *Prg4* are mainly regulated by genomic regions on chromosome 11. Although *Fasl* was not in any of two potential QTL regions, we analysed its interaction with other genomic region, just in case it interacts with genes within those two genomic regions. It turns out that *Fasl* are regulated by several genomic regions on chromosome 1, 2, 4, 6, 8, 11 and X, none of them contributes a significant amount (data not shown).

Discussion

Our data suggest that two sets of candidate genes may regulate the QTL for RA susceptibility on chromosome 1 in IL-1RN deficient Balb/c mice, and the data reduced the size of candidate genomic regions from about 19 Mbp to approximately 9 Mbp. One region is 5 Mbp, located between 150,927,537 bp and 157,073,549 bp, and the other is 4 Mbp, located between 165,318,402 bp and 169,640,436 bp. More importantly, establishing two sub-QTL regions clarified future directions for identifying and confirming causal genes for the susceptibility of spontaneous arthritis in this chromosomal region.

Our study also greatly reduced the number of candidate genes. The total number of candidate genes in our previous analysis (Jiao *et al.* 2011) was 208; through SNP and haplotype analyses, we reduced the candidate genes to 21 (tables 1 and 2). While the remaining genes cannot be completely ruled out as candidates, we at least have prioritized 21 genetic elements. Surprisingly, four of our favourite candidate genes either do not have SNP or are not in the haploid type region. Only one gene, *Pla2g4a*, stands out as the most favoured candidate gene in the first genomic region. Our SNP and gene network information on the candidate genes in the two genomic regions have established the foundation for further investigation of the candidate genes.

The rapid development of functional genomics and molecular biology resources has dramatically enhanced the gene discovery and functional analysis. Identifying causal genes for QTL has been the bottleneck in the study of molecular mechanisms of QTL. Our research indicates that, with those resources and advanced biotechnologies, identifying causal genes from QTL will be much easier and quicker than was possible decades ago. Nevertheless, information from some resources may not be as accurate as we want and should

improve. Further studies are needed to confirm the candidacy of the genes and phenotypes of the two sets of genomic regions.

Acknowledgements

This work was supported by the Center of Genomics and Bioinformatics and Center of Connective Tissue Research at the University of Tennessee Health Science Center; the Veterans Administration Medical Center in Memphis; and the National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health (R01 AR51190 to WG; R01 AR50785 to JS). The authors thank Dr Yoichiro Iwakura, Center for Experimental Medicine, Institute of Medical Science, University of Tokyo, Tokyo, Japan, for providing IL-1RA deficient mice for this study.

References

Hegen M., Sun L., Uozumi N., Kume K., Goad M. E., Nickerson-Nutter C. L. *et al.* 2003 Cytosolic phospholipase A2-alpha-deficient mice are resistant to collagen-induced arthritis. *J. Exp. Med.* **197**,1297–1302.

Jeffrey S. S., Fero M. J., Børresen-Dale A. L. and Botstein D. 2002 Expression array technology in the diagnosis and treatment of breast cancer. *Mol. Interv*. **2**, 101–109.

Jiao Y., Chiu H., Fan Z., Jiao F., Eckstein E. C., Beamer W. G. and Gu W. 2007 Quantitative trait loci that determine mouse tibial nanoindentation properties in an F2 population derived from C57BL/6J x C3H/HeJ. *Calcif. Tissue Int.* **80**, 383–390.

Jiao Y., Jiao F., Yan J., Xiong Q., Shriner D., Hasty K. *et al.* 2011 Identifying a major locus that regulates spontaneous arthritis in IL-1ra-deficient mice and analysis of potential candidates. *Genet. Res. (Camb)*. **18**, 1–9.

Leek J. T., Monsen E., Dabney A. R. and Storey J. D. 2006 EDGE: extraction and analysis of differential gene expression. *Bioinformatics* **22**, 507–508.

Raichel L., Berger S., Hadad N., Kachko L., Karter M., Szaingurten-Solodkin I. *et al.* 2008 Reduction of cPLA2alpha overexpression: an efficient anti-inflammatory therapy for collagen-induced arthritis. *Eur. J. Immunol.* **38**, 2905–2915.

Sharov A. A., Falco G., Piao Y., Poosala S., Becker K. G., Zonderman A. B. *et al.* 2008 Effects of aging and calorie restriction on the global gene expression profiles of mouse testis and ovary. *BMC Biol.* **3**, 6–24.

Townsend J. P. and Hartl D. L. 2002 Bayesian analysis of gene expression levels: statistical quantification of relative mRNA level across multiple strains or treatments. *Genome Biol.* **3**, RESEARCH0071.

Yan J., Jiao Y., Jiao F., Stuart J., Donahue L. R., Beamer W. G. *et al.* 2007 Effects of carbonic anhydrase VIII deficiency on cerebellar gene expression profiles in the wdl mouse. *Neurosci. Lett.* **413**, 196–201.

Zhou F., He X., Iwakura Y., Horai R. and Stuart J. M. 2005 Arthritis in mice that are deficient in interleukin-1 receptor antagonist is dependent on genetic background. *Arthritis Rheum.* **52**, 3731–3738.

Received 19 October 2011, in revised form 11 January 2012; accepted 14 February 2012
Published on the Web: 12 June 2012