

Supporting Information

Goto et al. 10.1073/pnas.0906776106

SI Materials and Methods

Mapping BR2 and BR4 Crossover Breakpoints. Segments (3–5 kb) of the 27-kb crossover breakpoint region were amplified by PCR in a RoboCycler (Stratagene) with the Roche Expand Long Template PCR System (in Buffer 3) using an initial 2-min heating step at 95 °C followed by 30 or 35 cycles that included: 95 °C for 20 s, 58–62 °C for 30 s (temperature varied depending on the primer length), and 68 °C for 3–5 min (interval defined by the expected product size). Reactions were completed with a final incubation at 68 °C for 7 min. Amplified segments were end-sequenced. SNP matches with one parent or the other defined the origin of each BR2 and BR4 segment and guided further sequencing until the two breakpoints were resolved. PCR amplification and sequencing primers are listed in [Table S3 A and C](#).

Haplotype Sequencing. Full haplotype sequencing was carried out using previously described primers and procedures (1). Sequencing of BR2, BR4, and their parent haplotypes was extended nearly 3 kb beyond the BGI locus to fully delineate the crossover breakpoints using primers listed in [Table S3 B and C](#).

Typing of Birds in Viral Challenge Trials. Blood samples (0.2 mL) were collected from all birds at 1 week of age and shipped to the City of Hope Beckman Research Institute for typing. DNA was extracted, as previously described (2). The segment of MHC-B DNA identified early in this study as different between BR2 and BR4 was PCR amplified using primers #359 and #342 ([Table S3C](#)). Amplifications were made with the Expand Long Template PCR System (Roche) using 2.5 μ L of 10 \times Buffer 3, 0.5 μ L primer #359 (400 nM), 0.5 μ L primer #342 (400 nM), 1.0 μ L dNTP (10 mM stock), 1 μ L template DNA (50–500 ng), 0.25 μ L enzyme mix (1.25 U), and sterile distilled water to bring the volume to 25 μ L. PCR amplifications were begun by heating for 2 min at 95 °C, followed by 35 cycles of 30 s at 94 °C, 30 s at 60 °C and 2 min at 72 °C, and finished by a final incubation at 72 °C for 7 min. The PCR amplification products (5 μ L) were electrophoresed in 0.8% TBE gels at 150 V for 35 min. The gels were then stained with ethidium bromide and images recorded with a Gel Documentation System (Bio-Rad). The BR2 and BR4 haplotypes were distinguished by the size of the amplified product, approximately 1,800 bp for BR2 and 2,000 bp for BR4.

Genome-Wide SNP Typing. As part of a large, genome-wide SNP typing project, eight DNA samples from homozygous BR2 birds and eight from homozygous BR4 birds within the GaHV-2 challenge trial, along with an UCD003 line samples were typed (3). Five adjacent SNPs on chromosome 2 were examined in 34 additional BR2 samples using primer pairs listed in [Table S3D](#).

3'-RACE Assays To Determine BGI Transcript 3'-UTR Length. 3'-RACE was carried out with a SMART12 RACE cDNA Amplification Kit (Clontech). The 5'-primer was designed to be BGI mRNA-specific (BG1SP) by targeting the splice product of the last two exons (5'-ACGAAGCAAAGAGGCGATGG-3'). The 3'-primer was 5'-AAGCAGTGGTATCAACGCAGAGTAC(T)₃₀AG-3'. First-strand cDNA syntheses and 3'-RACE PCR reactions were performed according to the kit directions. The PCR reactions consisted of one cycle of 94 °C for 5 min; five cycles of 94 °C for 30 s, 62 °C for 30 s, 72 °C for 2 min; 28 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 2 min; and finally 72 °C for 7 min. Products were analyzed in ethidium bromide stained

1.2% agarose gels and images captured with a Bio-Rad Gel Documentation System.

Reverse-Transcriptase Real-Time Quantitative PCR Assay for Gene Expression. To examine the relative expression of BGI*BR2 and BGI*BR4, line 003.R2 and 003.R4 eggs were obtained from breeding stock at the University of New Hampshire, shipped to City of Hope, hatched, and raised in the Animal Resource Center at City of Hope. Birds were euthanized with CO₂, tissues quickly collected into RNAlater (Ambion) and stored at –20 °C. Total RNA was purified using RNeasy Protect Midi Kits following the manufacturer's protocol (Qiagen). Residual DNA was digested with Qiagen RNase-Free DNase Sets. Then cDNA was synthesized from 2 μ g total RNA using reverse transcriptase and random hexamer primers from Invitrogen. A 1/10 dilution of the reverse transcription product (6 μ L) was used for each reaction. BGI-specific primer sequences were designed to cross exon boundaries to prevent production of qPCR products from genomic DNA. The sense (5'-ACGAAGCAAAGAGGCGATGG-3') and anti-sense (5'-AGATTGGAGTCCCATGTGTG-3') primers were designed to amplify a fragment of 102 bp. RNA was diluted to 1 in 10⁻⁵ and the 18S rRNA reference gene control amplified using sense (5'-cgaacgagactctggcatgct-3') and anti-sense (5'-catcacagactgttattgctc-3') primers designed to yield a product of 118 bp. Batches of the specific sense and anti-sense primers at 400 nM concentration were prepared and used in amplification reactions in the presence of iQ SYBR Green Supermix (Bio-Rad). Initially, the samples were heated in the iQ5 Multicolor Real-time PCR Detection System (Bio-Rad) cycler for 3 min at 95 °C, then 40 cycles of amplification were carried out at 95 °C for 15 s and 60 °C for 25 s. Subsequently, data from an 85 °C cycle melt curve were collected to check primers. No template and no reverse transcriptase controls were included to confirm that reactions were free from artifacts of primer dimerization and genomic DNA contamination. All reactions were run in triplicate. The amplified transcripts were quantified using the comparative C_T method. Briefly, the threshold cycle (CT) from each test sample was subtracted from the CT for 18S, resulting in Δ C_T.

Firefly/Renilla Dual Luciferase Reporter Assays. To investigate the influence of the 3'-UTRs on protein expression, BGI*BR2 and BGI*BR4 (short and long) full-length 3'-UTR sequences (482, 578, and 702 bp, respectively) were PCR amplified from the 3'-RACE reaction cDNA and subcloned into the *Photinus* firefly luciferase pGL3-control (SV40 promoter) vector (Promega) at the *Xba*I site located at the 3'-end of the luciferase gene. The 3'-UTR cDNA was obtained using primers incorporating the *Xba*I restriction site (underlined in the primer sequences following). The primer pair for BGI*BR2 was 5'-AGTTCTAGAT-ACCCAAACCAAAGAGG-3' (upstream) and 5'-AGTTCTAGATTGATTACCACCTGC-3' (downstream). The same upstream primer, 5'-AGTTCTAGATACCCAAACCAAAGAGG-3', was coupled with two different downstream primers to clone the short and long forms of the BGI*BR4 3-UTR, 5'-AGTTCTAGAAAGGTGATGACAAAGC-3', and 5'-AGTTCTAGATTATACACCTGCAGG-3', respectively. Chicken LMH (Line M Hepatoma) cells (4) were seeded in 24-well dishes (5 \times 10⁴ cells/mL). The luciferase reporter gene constructs (400 ng) and the control *Renilla* luciferase reporter (TK promoter) vector (10 ng) for normalization of transfection efficiency were co-transfected into LMH cells in triplicate wells using Transfec-

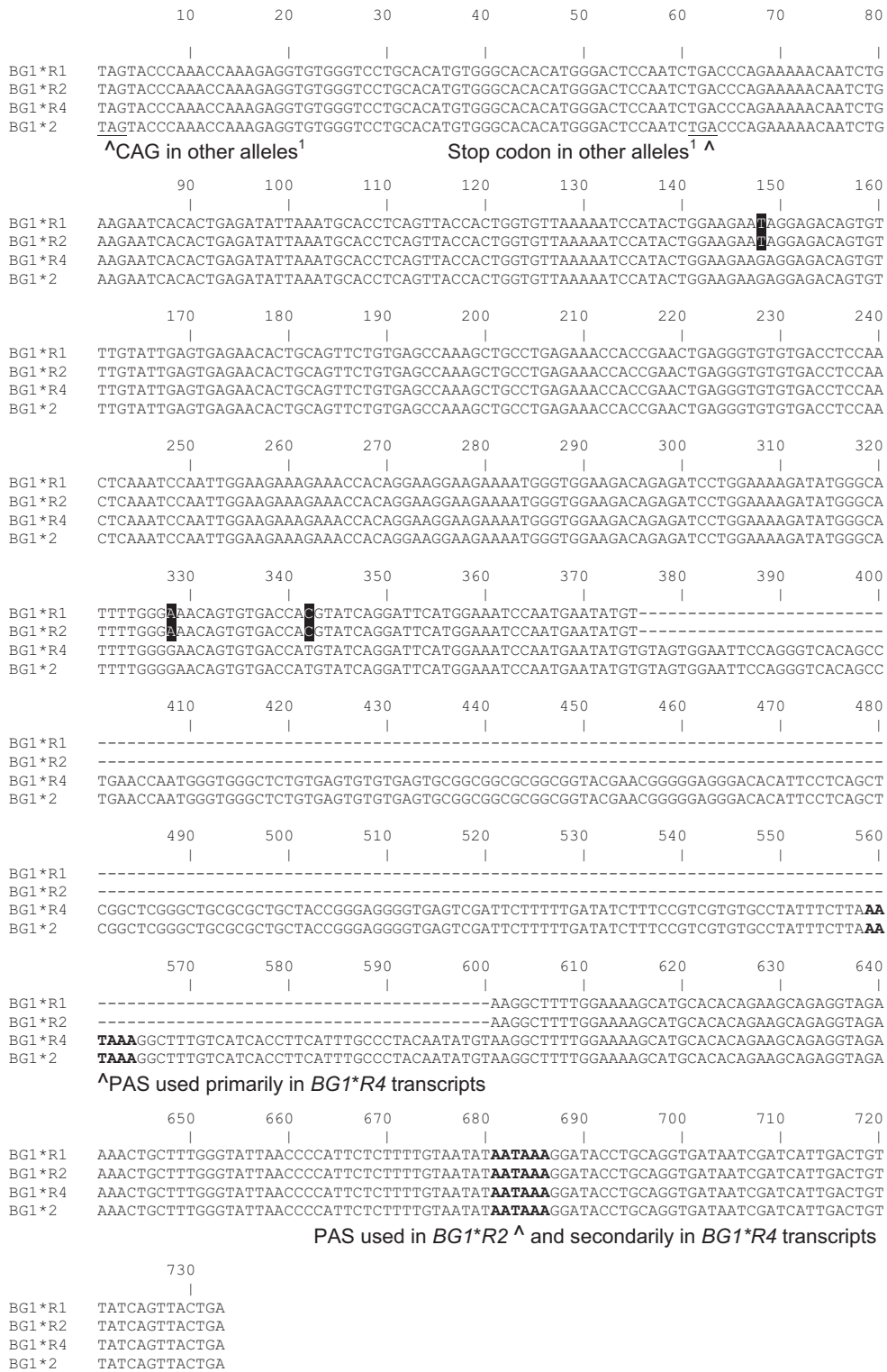
tin TM Lipid Reagent 6 (Bio-Rad) at a ratio of 1 μg plasmid to 2.5 μL reagent. Twenty-four hours after co-transfection, the cells were incubated in medium containing 100 ng/mL phorbol myristate acetate at 37 °C for 6 h to stimulate SV40 promoter-driven expression. Cells were lysed in passive lysis buffer following the manufacturer's recommended procedure. Luminescence was measured in a luminometer (20/20n Turner Biosystems). Relative luciferase activity was determined from the ratio between the activity levels of firefly and *Renilla* luciferase. Data shown (mean \pm SD) are from triplicate samples within a single experiment. Similar results were obtained in three independent experiments.

Expression of FLAG-Epitope Tagged BG1, Tyrosine Phosphorylation, and Phosphatase Association. *BG1**R2 cDNA made from *BR2* liver RNA was cloned into Bluescript. The FLAG epitope was inserted between exon 1 and exon 2. This construct was then directionally cloned into pMX-PIE (gift from Drs. G. Ehrhardt and M.D. Cooper, University of Alabama, Birmingham, AL) using the *EcoR*I and *Xho*I sites in the vector. Restriction sites were added to the *BG1* construct by splice overlap PCR reactions using forward primer: 5'-TTAGAATTCATGCACTTTCTAATTGGGC-3' and reverse primer: 5'-TTACTCGAGCTAAGGATAATCAGACTT-3'. The pMX-PIE vector expresses the gene of interest upstream of an internal ribosomal entry site element and the enhanced green fluorescent gene. The construct was transfected into LMH cells, puromycin selection applied and cells monitored for GFP expression. The GFP-positive cells grown out were found to be surface positive for FLAG when prepared for flow cytometry with M2 and secondary anti-mouse

IgG-APC (allophycocyanin) and DAPI staining. Samples were analyzed in a CyAn flow cytometer.

For tyrosine phosphorylation studies and phosphatase association studies, LMH cells ($2 \times 10^5/\text{mL}$) stably expressing FLAG-BG1 were seeded 24 h before treatment. Pervanadate (0.2 mM Na_3VO_4) was added to the culture media for the times indicated. Cells were collected, washed, pelleted, and lysed in 1% Nonidet P-40, 50 mM Tris·HCl (pH 7.5), 5 mM EDTA, and 150 mM NaCl, to which was added immediately before use 40 $\mu\text{g}/\text{mL}$ PMSF, 0.2 mM Na_3VO_4 , 1 mM Na_2MoO_4 , 5 mM β -glycerophosphate, and 10 $\mu\text{L}/100 \mu\text{L}$ final volume protease inhibitor mixture (Sigma P8840). The lysate was cleared by centrifugation at $12,000 \times g$ for 15 min. Aliquots of the lysate were incubated with M2 antibody overnight at 4 °C. Then a 50% slurry of Protein G Sepharose 4 Fast Flow beads (30 μL) (GE Healthcare Bio-Sciences Corp.) was added and the mixture incubated at 4 °C for 2 h. The beads were washed five times each with lysis buffer (1 mL). Beads were boiled for 5 min in 100 μL loading buffer and spun down. Proteins in the supernatant were separated by 9% or 12% SDS/PAGE and then transferred to PVDF membranes (Millipore Corporation). The membranes were subsequently incubated at room temperature in SuperBlock Dry Blend Blocking Buffer in TBS (Pierce) supplemented with 0.05% Tween-20, and then with M2 (1:500 dilution), anti-phosphotyrosine 4G10 antibody (1:200) (Upstate Biotechnology) or anti-SHP-2 SH-PTP2 (C-18) (1:200) (Santa Cruz Biotechnology) in blocking buffer for 1 h. After four washes, the membranes were incubated with secondary antibody at 1:7,500 dilutions (LI-COR Odyssey goat anti-mouse IRDye 800CW, 926–32210 or goat anti-rabbit IRDye 800CW, 827–08365) for 1 h. After four washes in TBS with 0.05% Tween-20, the membranes were visualized with the LI-COR Biosciences Odyssey Imager.

- Hosomichi K, et al. (2008) Contribution of mutation, recombination, and gene conversion to chicken MHC-B haplotype diversity. *J Immunol* 181:3393–3399.
- Goto R, et al. (1988) Isolation of a cDNA clone from the B-G subregion of the chicken histocompatibility (B) complex. *Immunogenetics* 27:102–109.
- Muir WM, et al. (2008) Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. *Proc Natl Acad Sci USA* 105:17312–17317.
- Kawaguchi T, Nomura K, Hirayama Y, Kitagawa T (1987) Establishment and characterization of a chicken hepatocellular carcinoma cell line, LMH. *Cancer Res* 47:4460–4464.
- Muir WM, et al. (2008) Review of the initial validation and characterization of a 3K chicken SNP array. *World's Poultry Science Journal* 64:219–225.



¹In sequences from SI ref. 1

Fig. S1. Alignment of the 3'-untranslated regions (3'-UTRs) of *BG1*R2* and *BG1*R4* with the 3'-UTRs of the *BG1* alleles in the parent haplotypes. The *BG1*R2* 3'-UTR is identical to that in the *BG1*R1* parent allele as demonstrated by three SNPs (highlighted in black) and the absence of the 225-bp insert. *BG1*R4* and the *BG1*02* both contain a 225-bp insert that provides an alternative polyadenylation signal (PAS) in addition to the PAS shared among all four alleles. Note that all four alleles possess a TAG stop codon upstream of the TGA stop found in other *BG1* alleles.

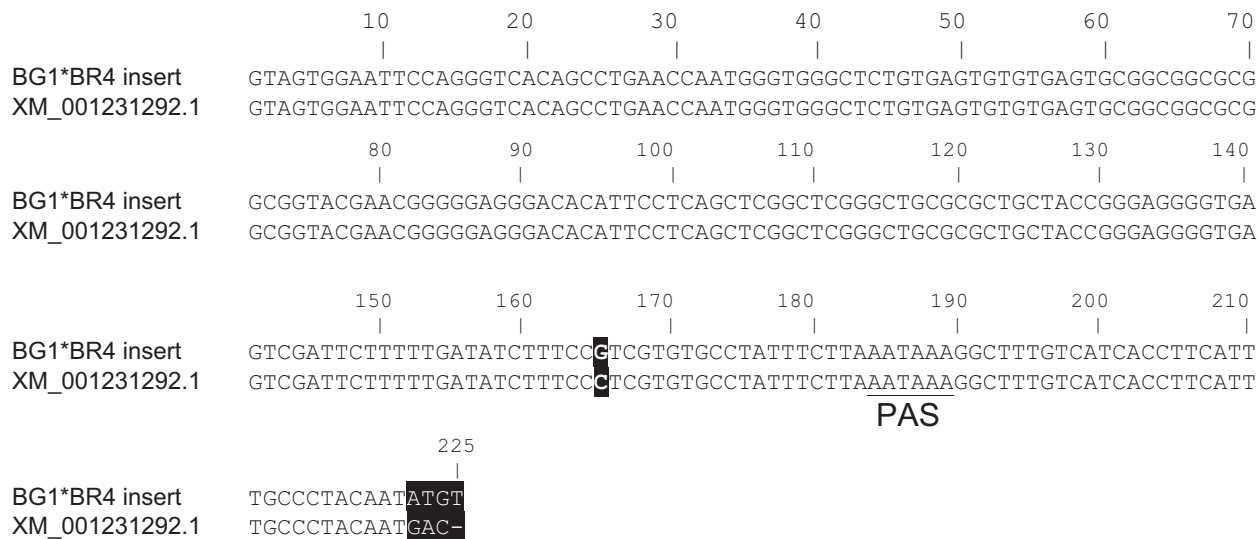


Fig. S2. Alignment of 225-bp insert sequence in *BG1*BR4* with a portion of XM.001231292.1 from *Gallus gallus*. The 225-bp insert in *BG1*BR4* is essentially identical to a portion of XM.001231292.1 that includes a PAS sequence. XM.001231292.1 is predicted to encode mRNA for a Pro-Pol-dUTPase polyprotein containing RNaseH, dUTPase, integrase, a protease and a reverse transcriptase. Many more highly similar sequences are present in the *G. gallus* genome.

Signal Peptide

BG1*R2 MHFLLGCNHPSFTLPWRTLLPYLVALHLLQPGSA
BG1*R4 MHFLLGCNHPSFTLPWRTLLPYLVALHLLQPGSA
BG1*12 MHFLLGCNHPSFTLPWRTLLPYLVALHLLQPGSA

IgV-like ectodomain

BG1*R2 QLRVVAPSLHVTANVGQDVVLRCLSPCKDAWSSDIRWIQHRTSGFVHHYQNGEDLEQMEEYKGRTELLRDGLSDGNLDRITAVSTSDSGSYSCAVQDGDGYADAVVDLEVSD
BG1*R4 QLRVVAPSLHVTANVGQDVVLRCLSPCKDAWSSDIRWIQHRTSGFVHHYQNGEDLEQMEEYKGRTELLRDGLSDGNLDRITAVSTSDSGSYSCAVQDGDGYADAVVDLEVSD
BG1*12 QLRVVAPSLRVTAIVGQDVVLRCLSPCKDAWSSDIRWIQHRTSGFVHHYQNGEDLEQMEEYKGRTELLRKLSDGNLDRITAVSTSDSGSYSCAVLDGDGYADAVVDLEVSD

Transmembrane domain

BG1*R2 PFSQIIHPWKVALAVIVTILVGSFVITVFLYRKK
BG1*R4 PFSQIIHPWKVALAVIVTILVGSFVITVFLYRKK
BG1*12 PFSQITHPWKVALAVIVTILVGSFVITVFLYRKK

Region of Coiled Coil (presented with breaks to that correspond to individual exons)

BG1*R2 AAETTKQ KGKDAEL ERMDAKL GTLAAEL ERRDAKL ETLVESL ERRNAEF -EKLASDL ERRNAQL DKLASDL VQQTKAV
BG1*R4 AAETTKQ KGKDAEL ERMDAKL GTLAAEL ERRDAKL ETLVESL ERRNAEF -EKLASDL ERRNAQL DKLASDL VQQTKAV
BG1*12 -AETTKQ KGKDAEL EGMDAKL GTLAAEL ERRDAKL ETLVENL ERRNTEF AKKLAEL ERRNAQL DKLASDL VQQTKAV

Terminal Domain (presented with break reflecting two exons)

BG1*R2 EKLNSQWSKLQSLKLTKSDTIQNNFIGYEKSPQAVNYSPLSNPEKHHEAK RRWYIKSDYP
BG1*R4 EKLNSQWSKLQSLKLTKSDTIQNNFIGYEKSPQAVNYSPLSNPEKHHEAK RRWYIKSDYP
BG1*12 EKLNSQWSKLQSLKLTKSDTIQNNFIGYEKSPQAVNYSPLSNPEKHHEAK RRWYIKSDYPQYPNQRGVGAHVGHGTII

ITIM

Fig. S3. Alignment of the predicted amino acid sequences for *BG1*R2* and *BG1*R4*. Alignment illustrates the truncation of the R2 and R4 *BG1* isoforms as compared to other alleles, represented here by the amino acid sequence of *BG1*B12*. Dashes (-) are inserted to optimize alignment. Spaces inserted into the sequence denote exon boundaries.

Table S1. SNP test for genetic uniformity of the 003.R2 and 003.R4 congenic lines using SNPs evenly spaced throughout the genome

	SNPs	%
Scored across all 16 samples	2,702	—
Identical among eight 003.R2 and eight 003.R4 samples	2,167*	80.2
SNPs segregating in both eight 003.R2 and eight 003.R4 samples	186	6.9
SNPs segregating only among eight 003.R2 samples	262	9.7
SNPs segregating only among eight 003.R4 samples	82	3
Opposite SNPs fixed in two sample sets but segregating in lines	5	0.2

The samples were analyzed as part of a larger study (5). The sixteen samples were identical at 2167 SNP alleles (80.2%), suggesting that the 003.R2 and 003.R4 lines are indeed highly inbred and similar, but have not become fully homozygous across the genome (See SI Materials and Methods for explanation of the residual diversity). Among the 535 SNPs that remain segregating, 186 were segregating in both sample sets. The remaining alleles, except for five, were variously identical in the sample set of one line, but segregating in the other. Only five adjacent SNP alleles on Chromosome 2 were found to be entirely opposite within this initial comparison of sixteen samples. At these five loci, the eight 003.R4 samples were all found to have the same allele as UCD003. In contrast, the eight 003.R2 samples all carried the opposite allele. Typing of 34 additional line 003.R2 samples showed that the SNP alleles of all five loci are still segregating in line 003.R2, indicating that these loci are not fixed with opposite SNP alleles nor are the allelic frequencies at these loci highly skewed in different directions between the 003.R2 and 003.R4 lines. * This includes six SNP loci located on Chromosome16.

Table S2. Reverse transcriptase quantitative PCR assays. Expression of *BG1* in young birds from line 003.R2 (*BG1R2) and line 003.R4 (*BG1**R4) was compared using qPCR**

	ΔC_T		ΔC_T	
	003.R2	003.R4	003.R2	003.R4
	6-day old		21-day old	
Spleen	13.87*	13.63	12.5	12.6
Liver	5.07*	4.61	8.37	7.92
Thymus	- (n.d.)	-	11.7	12.7
Bursa	-	-	12.34	12.92
Brain	-	-	15.25*	14.16*
Proventriculus	-	-	11.78	10.87
Gizzard	-	-	14.79	14.99
Small Intestine	-	-	6.08	6.81
Pancreas	-	-	9.79*	8.07
Lung	-	-	12.29*	11.81
Kidney	-	-	6.83*	6.94*
Heart	-	-	15.25*	14.16*
Skin	-	-	15.71	15.89

*Indicates samples from single individuals; n.d. = not determined.

Table S3. Primers for SNP mapping, sequencing and SNP typing. A. Amplification primers for SNP mapping crossover breakpoints

Region	Gene	Nucleotide position from AB268588	Primer Name	Primer Direction	Primer Sequence (5' to 3')	Length (bp) from AB268588	SNP detection primers
1	<i>BTN-1</i>	119039	355	F	TATGGCACAAAAGGTGACGGC	2524	355, 356
		121563	356	R	TTCGTAGTCCAGAGAGATGCG		
2	<i>BTN-1</i>	120918	369	F	GGAGAGGCGAAGAAAAGTTGG	1324	369, 353, 356, 370
		122242	370	R	TTTGCTTTCTACAGAAATAGC		
3	<i>BTN-1</i>	121441	353	F	AAGGAGCCGTCAGCTTAACC	2570	353, 354
		124011	354	R	GGACAGCCAACAAAACACTCC		
4	<i>BTN-2</i>	123819	351	F	AACGTCACCCTGAGTCTTCC	2662	351, 352
		126481	352	R	TTTGGGATCAGCGATGTCTCC		
5	<i>BTN-2</i>	126366	349	F	ATACTGCGTCAGAGCATCTGG		360, 361, 366, 367, 377
		128203	367	R	GACTCTCTCCACAACGACC		
6	<i>Leu-tRNA/BG1</i>	128545	342	F	GGATGAAGACTTACCCTGTTCC	3018	342, 348, 358, 359, 362, 363, 368
		131563	348	R	AGTGCTGGGCTTTGTGATGG		
7	<i>Blec4</i>	136423	336	F	CCTGCAATAGGAAATCCATCTTG	979	335, 336
		137402	335	R	CTTGCAAGGATTTGGTGCC		

Table S3D. Primers for SNP typing region on Chromosome 2

No.	Illumina SNP identifier	Primer Sequence (5' to 3')
520	snp-51-46-74632-S-2	CCAAGTATCTTGTGTTAGC
521		TACCCACACTTAACTGGC
522	snp-51-11-29905-S-3	CACATGCACATCGTGAAC
523		TACAGACGTCCGGAATTC
550	snp-51-17-21450-S-3	AAGGCAGCAGGCAGAGC
525		TCACATACAGGATGATGAC
526	snp-51-27-153706-S-3	TGTGTTTCCTACCAGCGC
527		GCTGTCCATCCTCTCTTG
570	SNP-51-41-17623-S-2	TTGCAGAAGCGAGAGAATC
571		TGGAACATCCCAGAACAGC