Molecular Ecology (2006) **15**, 1263–1273 doi: 10.1111/j.1365-294X.2005.02753.x

A cDNA macroarray approach to parasite-induced gene expression changes in a songbird host: genetic response of house finches to experimental infection by *Mycoplasma gallisepticum*

ZHENSHAN WANG,*‡ KRISTY FARMER,† GEOFFREY E. HILL† and SCOTT V. EDWARDS*§
*Department of Biology, University of Washington, Seattle, WA 98195, †Department of Biological Sciences, Auburn University, Auburn, AL 36849, USA §Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

Abstract

In 1994, the bacterial parasite *Mycoplasma gallisepticum* expanded its host range and swept through populations of a novel host — eastern US populations of the house finch (*Carpodacus mexicanus*). This epizootic caused a dramatic decline in finch population numbers, has been shown to have caused strong selection on house finch morphology, and presumably caused evolutionary change at the molecular level as finches evolved enhanced resistance. As a first step toward identifying finch genes that respond to infection by *Mycoplasma* and which may have experienced natural selection by this parasite, we used suppression subtractive hybridization (SSH) and cDNA macroarray approaches to identify differentially expressed genes regulated by the *Mycoplasma* parasite. Two subtractive cDNA libraries consisting of 16 512 clones were developed from spleen using an experimentally uninfected bird as the 'tester' and an infected bird as 'driver', and vice versa. Two hundred and twenty cDNA clones corresponding 34 genes with known vertebrate homologues and a large number of novel transcripts were found to be qualitatively up- or down-regulated genes by high-density filter hybridization. These gene expression changes were further confirmed by a high throughout reverse Northern blot approach and in specific cases by targeted Northern analysis. BLAST searches show that heat shock protein (HSP) 90, MHC II-associated invariant chain (CD74), T-cell immunoglobulin mucin 1 (TIM1), as well as numerous novel expressed genes not found in the databases were up- or down-regulated by the host in response to this parasite. Our results and macroarray resources provide a foundation for molecular co-evolutionary studies of the *Mycoplasma* parasite and its recently colonized avian host.

Keywords: alternative splicing, cDNA macroarray, granzyme A, house finch, HSP90, *Mycoplasma gallisepticum*, suppression subtractive hybridization, TIM1

Received 14 June 2005; revision accepted 24 August 2005

Introduction

Host–parasite systems are intrinsically interesting because of the potential for rapid co-evolutionary dynamics and for evolutionary and behavioural modification of host populations. Host populations that have recently colonized a new geographic area offer opportunities not only for examining host–parasite interactions in a novel host environment, but also for comparisons with ancestral populations from which the recent colonist arose. Using a variety of behavioural, demographic and genetic approaches, we have been studying a decade-old host–parasite interaction in an avian population that has itself been established in a novel environment for about 50 years. In so doing we hope to provide a multidimensional view of a

Correspondence: Scott Edwards, Present address: Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA. Fax: 617-495-5667; E-mail: sedwards@fas.harvard.edu.
‡Present address: Department of Pharmacology, University of Washington, Seattle, WA 98195, USA.

recently evolved and historically well-documented host–parasite interaction and thereby provide a detailed view of the potential for parasites to modify vertebrate hosts and to undergo evolutionary changes themselves.

House finches (*Carpodacus mexicanus*) are passerine birds native to western North America from southern British Columbia to Oaxaca, Mexico, and were introduced to Long Island, New York, in 1940 from populations originating in coastal California (Hill 1993, 2002). A newly recognized disease caused by a bacterial parasite *Mycoplasma gallisepticum* (MG) emerged in eastern US house finch populations in the mid-1990s (Fischer *et al*. 1997). The bacterium is a common pathogen in domestic turkeys and chickens, which may have served as the source of the epizootic. Both captive and wild infected birds have red, swollen, watery, or crusty eyes; in extreme cases the eyes become swollen shut or crusted over, and the birds become essentially blind (Roberts *et al*. 2001a, b; Farmer *et al*. 2002). Conjunctivitis was first noticed in house finches during the winter of 1993–1994 in Virginia and Maryland. The disease later spread to states along the Atlantic coast of the United States, and has now been reported throughout most of eastern North America, as far north as Quebec, Canada, as far south as Florida and as far west as Montana; the dynamics of the epizootic was one of the best documented for any avian population (Dhondt *et al*. 1998). The disease is most common in the eastern population of house finches. However, a few reports of the disease have been confirmed in American goldfinches (*Carduelis tristis*), purple finches (*Carpodacus purpureus*), evening grosbeaks (*Coccothraustes verspertinus*), and pine grosbeaks (*Pinicola enucleator*), all members of the family Fringillidae. In contrast to the several studies of house finch pathology caused by MG or variability among MG strains (Farmer *et al*. 2002; Pillai *et al*. 2003), at present there is almost no information about the molecular response of house finches to *Mycoplasma gallisepticum* in the literature. Indeed, despite the growing base of information on gene expression changes induced by pathogens in domestic chickens, and the cloning of a variety of immune response genes in chickens and other avian species, virtually nothing is known of genetic responses to infection in any wild songbird. Identification of genes that exhibit differential expression between infected and healthy birds is the first step toward understanding the molecular mechanisms of immune response in natural populations, as well as identifying genes that may show intriguing evolutionary dynamics in response to naturally occurring epizootics.

Genetic responses to parasite infection provide a natural link between genetics and a variety of mechanisms for sexual selection, mate choice and plumage evolution in birds. Disease and parasites are well known as important modulators of ornamental traits including particularly colourful plumage (Hill 2006). The immune response is thought to be both an important cost to birds of both sexes and a benefit, in terms of enhanced clearance of parasites and ability to devote greater resources to costly signals (Kilpimaa *et al*. 2004). Immune responses can thereby influence individual fitness through primary mate choice as well as extra-pair matings. Although house finches are among the best-studied models for the myriad links between plumage brightness, parasites and disease (summarized in Hill 2002), the genetic basis of resistance to parasites is virtually unknown in any bird species. Major histocompatibility complex (MHC) genes are one candidate locus for involvement in immune response and sexual selection in wild birds, but are only one of several possible groups of genes whose expression and evolution might be influenced by parasites (Hess *et al*. 2000; Hess & Edwards 2002).

Several approaches are available to identify differentially expressed genes among cell populations in different physiological states, such as traditional subtracted hybridization (Sargent & Dawid 1983), differential display reverse transcriptase (DDRT) polymerase chain reaction (PCR) (Liang & Pardee 1992), cDNA amplified fragment length polymorphisms (AFLP; Bachem *et al.* 1996), representational difference analysis (RDA; Hubank & Schatz 1994), serial analysis of gene expression (SAGE; Velculescu *et al.* 1995), and differential analysis of library expression (DAZLE; Li *et al*. 2004). Although all these methods have been successfully applied to identify differentially expressed genes in various contexts, one limitation common to all of them is their difficulty of identifying rare expressed transcripts, or need to physically separate single-stranded and double-stranded cDNAs during laboratory analysis, which can be challenging. Microarray analysis offers a fundamentally new approach to identification of differentially expressed genes that is now routine and commonly applied to model species, many of which have large expressed sequence tag (EST) databases or whole-genome sequences (Quackenbush 2001; Yang & Speed 2002). Microarrays were relatively new at the inception of this study. Although this approach is being applied increasingly to non-model species (Oleksiak *et al*. 2002; Renn *et al*. 2004; Cossins & Crawford 2005), without modification (Hegarty *et al*. 2005) it can be prohibitively expensive and is still technically beyond the capabilities of many evolutionary ecologists.

For these reasons, to study gene expression changes in house finches induced by *Mycoplasma*, we elected to use a new PCR-based cDNA subtraction method, termed suppression subtractive hybridization (SSH), which is technically straightforward, does not require any genomic knowledge of the study organism, and overcomes some of the problems of transcript rarity and differences in mRNA abundance. This last advantage comes about by incorporating a hybridization step that normalizes (equalizes) sequence abundance during the course of subtraction by

standard hybridization kinetics (Diatchenko *et al*. 1996). It eliminates any intermediate steps for physical separation of single-stranded and double-stranded cDNAs, unlike previous SSH approaches requires only one round of subtractive hybridization, and can achieve greater than 1000-enrichment for differentially expressed cDNAs. SSH is a highly effective method for generating subtracted cDNA libraries, dramatically increases the probability of obtaining low-abundance differentially expressed cDNAs and simplifies the analysis of the subtracted cDNA libraries.

In a fashion similar to microarrays, nylon membranes containing robotically spotted, differentially expressed cDNAs, or macroarrays, have been used to compare patterns of gene expression in thousands of genes in a single hybridization (Rast *et al*. 2000). The emerging technology of cDNA macroarray hybridization offers the possibility for providing a rapid, high throughout method to screen an SSH cDNA library. However, there are several disadvantages to SSH-macroarray analysis of gene expression compared to microarrays. Despite some controversy about results obtained in microarrays made by different manufacturers, it is generally agreed that microarrays can be highly sensitive even to changes in abundance of rare transcripts. Once the microarrays for a given species are produced, the actual hybridization protocols are technically straightforward, probably more so than the SSH approach that we have employed. Most importantly, whereas microarrays have the potential to determine the relative level of expression of transcripts relative to controls, SSH only allows determination of up- or down-regulation, without any information on the relative expression level in the native cell populations. Nonetheless, like microarray analysis, SSH can identify rare transcript differences and in principle could be applied to study gene expression changes in closely related species (Renn *et al*. 2004). Thus, to our knowledge for the first time in an avian population, we have employed an SSH-macroarray protocol to identifying gene expression changes in house finches experimentally infected with MG. Our hope was to provide a molecular portrait of genetic response to MG in house finches to provide a foundation for further studies examining gene expression and molecular evolution in a rapidly evolving host–parasite system. By focusing specifically on genes expressed in the house finch spleen, the site of antibody production and a variety of other immunogenetic loci, such as cytokines, we hoped to enrich our survey for genes of importance to the house finch immune system (John 1994a, b).

Materials and methods

Bird materials and inoculation

Hatching-year house finches that had fledged from their nests approximately 6 weeks earlier were captured in traps at feeding stations in Auburn, Alabama, during July and August 2000–2002. At the time of capture all birds were examined for symptoms of MG infection and symptomatic birds were released. Blood was collected from asymptomatic birds and serum from each blood sample was screened for MG-specific antibodies using the serum plate agglutination assay (SPA) described by Luttrell *et al*. (1996, 1998; InterVet, Inc.). SPA-negative birds were retained for this study and held in outdoor aviaries with food and water *ad libitum* for 2 weeks, and then retested by SPA. Thus, all birds began the study with no prior exposure to MG.

Birds in the exposure group were inoculated with a culture of house finch MG via a bilateral ocular route (Farmer *et al*. 2002). Mycoplasmal infection was confirmed in inoculated birds and lack of MG infection was confirmed in noninoculated birds using SPA and PCR approaches 2 weeks after inoculation (Roberts *et al*. 2001b). The spleens of infected and uninfected birds were collected and stored in RNAlater (Ambion Inc.) before isolation of total RNA.

Isolation of total RNA and cDNA synthesis

Total RNA was prepared by using TRIzol regent (Invitrogen Life Technologies). The quantity and quality of RNA samples were determined spectrophotometrically by the absorbance at 260 nm and the absorbance ratio at 260/280 nm, and checked by electrophoresis on a 1.2% agarose/formaldehyde gel. One microgram total RNA each from the spleen of the uninfected and the infected house finches was used to prepare double-stranded cDNA using the SMART PCR cDNA Synthesis Kit (Clontech Laboratories, Inc.) according to the manufacturer's protocol.

Generation of subtracted libraries by SSH

To target MG-regulated genes in house finches, we first isolated total RNA from each of two infected and two uninfected female house finches. In our experimental protocol (Fig. 1), 'tester' RNAs served as the RNA pool which was hybridized to RNA from a 'driver', which served as a control for genes expressed in both pools. However, RNA from infected birds could serve as either 'tester' or 'driver' depending on whether one wishes to isolate transcripts that are up- or down-regulated in infected birds, respectively. SSH procedures were performed using the Clontech PCR-Select™ cDNA Subtraction Kit with minor modifications (Clontech). Briefly, tester and driver double-stranded cDNAs were synthesized from uninfected and infected birds. Then the tester and driver cDNAs were digested with the restriction enzyme *Rsa*I to yield blunt-ended cDNA populations. The tester cDNAs were then subdivided into two subsamples, and a different adaptor primer was ligated to each subsample. Two hybridizations are then performed. In

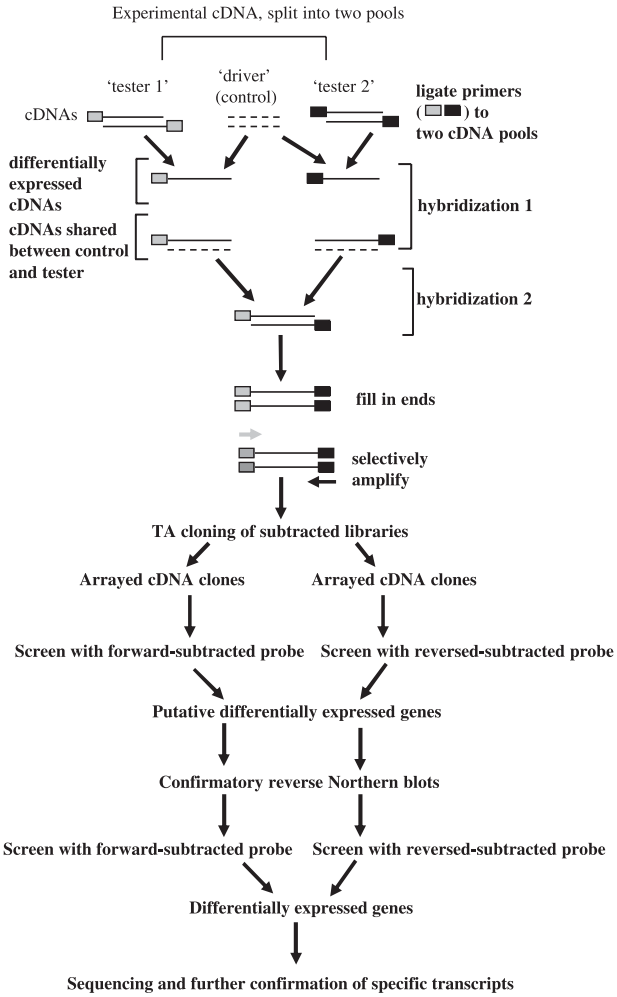

Fig. 1 A schematic diagram of the suppression subtractive hybridization procedure and screening strategy used in this study. Modified from literature from the Clontech kit used for the procedure.

the first hybridization, after heating to produce single-stranded cDNAs, an excess of driver cDNA is added to each of the two subsamples of tester cDNA in order to equalize and enrich the differentially expressed cDNAs. In the second hybridization, single-stranded cDNA from the two primary hybridization samples are allowed to anneal to generate double-stranded tester molecules with different adaptor primers. After filling in the overhanging ends with nucleotides and DNA polymerase, the blunt-ended molecules generated are then subjected to PCR with complements to the adaptor primers in order to amplify the desired differentially expressed sequences. Secondary PCR amplification is performed using nested primers provided by the kit to further reduce any background PCR products and to enrich for differentially expressed sequences. These nested PCR products were purified using QIAquick™ PCR Purification Kit (QIAGEN), cloned into TOPO cloning vector and transformed into chemically competent TOP 10 cells (Invitrogen Inc.). White clones were picked and arrayed into 384-well microtitre plates with a Q-bot clone-picking robot (Genetix, Inc.) and spotted on high density grid filters (22×22 cm) in a 4×4 spotting duplicate pattern using a TAS-BioGrid robot (BioRobotics).

Differential hybridization screening with subtracted probes

In the following steps and in Fig. 1, 'forward' subtracted probes refer to cDNA pools made from RNA in which the infected bird was tester, and 'reverse' subtracted probes refer to cDNA pools made from RNA from infected birds as driver (and uninfected birds as tester; Fig. 1). Although in principle all of the cDNAs isolated by the protocol should be differentially expressed, in practice hybridization steps are required to differentiate between false positives and truly differentially expressed transcripts. Hybridizing the macroarray filters with either total unsubtracted cDNA from infected or uninfected birds is a logical procedure at this point, but in practice this approach can miss many rarely transcribed sequences. Instead, after restriction digestion with *Rsa*I so as to remove the adaptor primers, the forward and reverse subtracted and amplified cDNA pools were labelled with ^{32}P used and as probes to screen duplicate cDNA macroarrays as described in the Clontech kit. Differentially expressed genes are confirmed by contrasting hybridization signals using the two probes: clones hybridizing with the forward-subtracted probe but not the reverse-subtracted probe and vice versa are inferred to be differentially expressed. Genes are inferred to be up-regulated by *Mycoplasma* when positive by this screen and derived from experiments in which the cDNA from an infected bird served as the 'tester'. By contrast, genes are inferred to be down-regulated when positive for this screen and derived from experiments in which cDNA from an uninfected bird is the 'tester' and that from an infected bird is the 'driver'.

Reverse Northern blots and screening

Candidate clones identified from macroarray screening were further confirmed by a reverse Northern blot hybridization procedure in which the inserts of differentially expressed cDNA clones were amplified by colony PCR using priming sites in the cloning vector, resolved electrophoretically and transferred in parallel onto nylon membranes (von Stein *et al*. 1997). Filters were then hybridized with equivalent amounts of ^{32}P-labelled forward- and reverse-subtracted PCR mixtures as in the macroarray hybridization; again, clones differentially hybridizing with the two probe sets were inferred to be differentially expressed.

Sequencing and database searches

Single-pass DNA sequences of selected cDNA clones (GenBank Accession nos DR782681–DR782901) were searched against GenBank using BLASTX version 2.2.10, a version containing sequences from the chicken genome project (International Chicken Genome Sequencing Consortution 2004). Each clone was categorized as known from orthologues in other species, similarity to a database EST, or a novel sequence not found in the database. Similarities that showed *e*-values less than e^{-08} with more than 100 nucleotides were considered significant.

Northern blot hybridizations

To further confirm differential expression in a panel of experimentally infected finches unrelated to those from which the cDNA libraries were made, total RNA (10 μg) from five infected and five uninfected birds was fractionated by agarose-formaldehyde gel electrophoresis (1% agarose/formaldehyde). The RNA was transferred to a nylon membrane (Amersham-Pharmacia) by capillary action and cross-linked by UV irradiation (Stratalinker; Stratagene). Candidate cDNA clones identified as differentially expressed in the previous experiments and as commonly up- or down-regulated in the database searches were used as probes on a panel of RNAs from infected and uninfected birds. The hybridizations were performed with the ECL nonradioactive labelling kit according to the manufacturer's instructions (Amersham Biosciences).

Results

Two cDNA libraries totaling 16 512 clones were constructed. Seven thousand two hundred ninety-six (7296) clones were derived from SSH experiments in which an uninfected bird was 'tester' and an infected bird was 'driver'; 9216 clones were derived from the reverse experiment. Clones of the libraries were picked and arrayed into 43 384-well microtitre plates and spotted on high density nylon grid filters (22×22 cm) in a 4×4 pattern in which each clone was spotted in duplicate for use as a positive control during hybridization. Two identical macroarray filters were screened with the restriction digested forward- and reverse-subtracted cDNA probes from SSH; these are probes in which the cDNA populations have been enriched and normalized. Clones that differ strongly in their hybridization intensity between the two identical macroarray filters are candidates for differential expression. An example hybridization is shown in Fig. 2. From these experiments, approximately 400 clones were preliminarily identified as differentially expressed finch genes that were up- or down-regulated as a result of infection with the *Mycoplasma* parasite.

Since the amount of DNA of each spotted clone on the macroarray filters was only about 50 ng (Granjeaud *et al*. 1999), even weak hybridization signals that could represent false positives should not immediately be excluded. In order to reduce the number of such false positives and to identify a manageable number of clones, a second screening involving reverse Northern blot was carried (von Stein *et al*. 1997). An example of reverse Northern blot hybridization is shown in Fig. 3. By comparing the results from reverse Northern blot hybridizations with the forward- and reverse-subtracted probes, 220 clones from the ~400 preliminarily identified assumed positive clones were further confirmed as strongly differentially expressed. Of the 220 clones, 97 clones were up-regulated (Table 1), and 123 clones (Table 2) were down-regulated in birds experimentally infected with *Mycoplasma*.

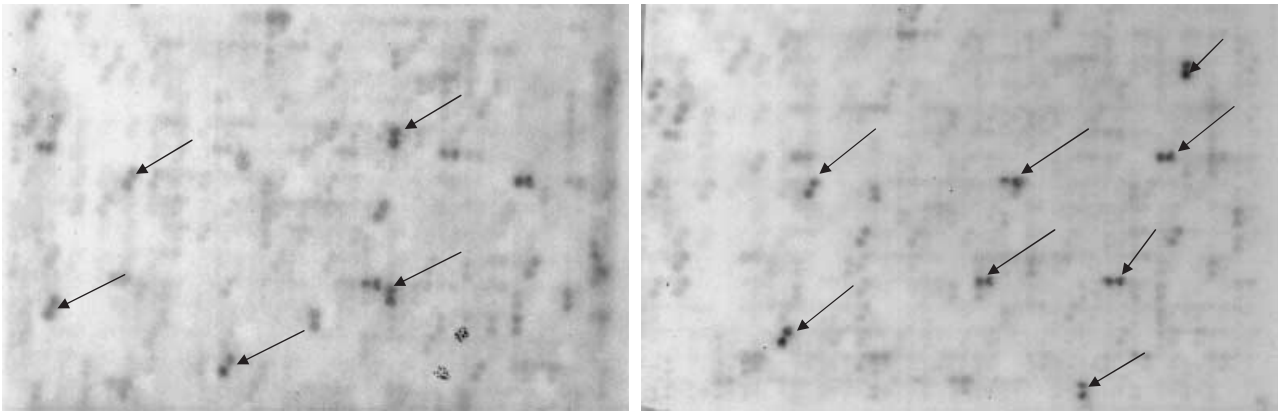

Fig. 2 Example of comparative hybridization profiles for cDNA macroarrays consisting of subtracted cDNA from an uninfected bird. The left panel is probed with subtracted uninfected cDNAs. The right panel is probed with subtracted infected cDNAs. The arrows indicate genes that are putatively differentially expressed between uninfected and infected individuals.

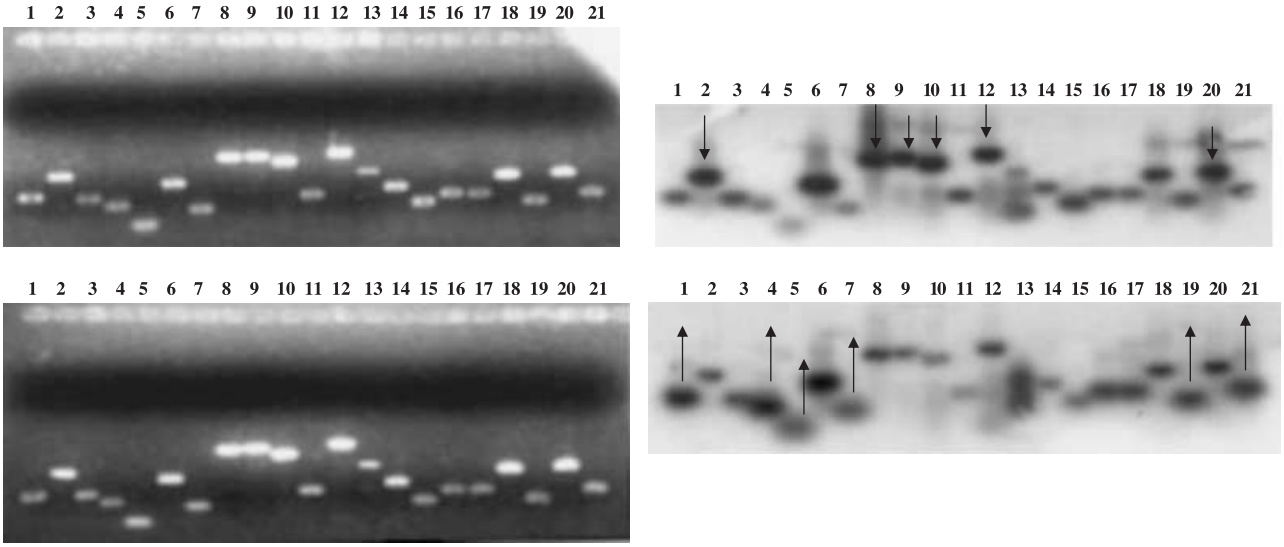

Fig. 3 An example of a reverse Northern blot hybridization to confirm putatively differentially expressed genes. Right panel: equal amounts of PCR-amplified putative differentially expressed cDNA clone inserts were resolved on 1.5% agarose gels in parallel to make two identical filters. Left panel: duplicate filters were hybridized with subtracted uninfected (upper) and infected (down) complex cDNA probes as in the macroarray hybridization. Arrows indicate differentially expressed genes between uninfected and infected individuals.

Table 1 Summary of up-regulated SSH cDNA clones in house finch

Homologue	Accession no.	*e*-value	Identity	Species	Redundancy
Heat shock protein 90	CAA30251.1	6.00E-67	101/103 (98%)	*Gallus gallus*	29
Predicted: similar to Splicing factor, arginine/serine-rich 12	XP_424756.1	5.00E-48	95/95 (100%)	*Gallus gallus*	16
Predicted: similar to TIM1	XP_414570.1	3.00E-19	40/60 (66%)	*Gallus gallus*	5
MGC69002 protein	AAH56044.1	6.00E-27	55/94 (58%)	*Xenopus laevis*	3
TPA: putative ISG12(2) protein	CAE00390.1	2.00E-11	35/61 (57%)	*Gallus gallus*	2
granzyme A	NP_989788.1	1.00E-26	59/96 (61%)	*Gallus gallus*	2
MRP protein	AAT01286.1	4.00E-17	44/59 (74%)	*Coturnix japonica*	1
Hspca protein	AAH49124.1	3.00E-11	28/32 (87%)	*Mus musculus*	1
K123 protein	NP_990350.1	2.00E-20	42/55 (76%)	*Gallus gallus*	1
Predicted: similar to KIAA1424 protein	XP_418602.1	1.00E-93	166/166 (100%)	*Gallus gallus*	1
Predicted: *Gallus gallus* hypothetical gene supported by CR352378	XM_429565	*2.00E-75	*182/193 (94%)	*Gallus gallus*	1
Novel					35
Total					97

*Nucleotide BLAST results.

All 220 clones identified by macroarray and reversed Northern blot hybridization were minipreped and subjected to DNA sequence determination to further characterize them. Of these, the sequences of 216 clones were successfully obtained by single-pass sequencing. A summary of BLASTX GenBank search results is presented in Fig. 4. Detailed information for each unique sequence is described in Table 1 and Table 2. Among the 97 clones inferred to be up-regulated, 62 (64%) show similarity to genes in other species, and 9 (9.2%) show highest identity with genes in chicken or other birds. Similarly, among the 123 clones inferred to be down-regulated in infected birds, 41 (33%) clones found significant matches with genes from other species. There was a high percentage of novel clones which had no significant matches in the database.

Several of the differentially expressed finch transcripts with identifiable homologues [HSP90, argenine-serine rich splicing factor, translation initiation factor, mitochondrial cytochrome oxidase subunit I (COI), elongation factor (EF) 1α, T-cell immunoglobulin mucin (TIM) 1, hemoglobin alpha-A chain, high mobility group (HMG)-14 protein, MHC class II-associated invariant chain (CD74)] are

Table 2 Summary of down-regulated SSH cDNA clones in house finch

Homologue	Accession no.	*e*-value	Identity	Species	Redundancy
Cytochrome oxidase subunit I	AAD32527.1	3.00E-70	134/148 (90%)	*Vidua chalybeata*	10
MHC class II-associated invariant chain Ii	NP_001001613.1	4.00E-48	87/131 (66%)	*Gallus gallus*	3
Invariant chain	AAT36345.1	8.00E-11	33/68 (48%)	*Gallus gallus*	3
ep37-L2	BAA21834.1	6.00E-13	31/56 (55%)	*Cynops pyrrhogaster*	3
Translation elongation factor eEF-1 alpha	I50226	9.00E-49	92/93 (98%)	*Gallus gallus*	2
NADH dehydrogenase subunit 4	CAA77204.1	7.00E-36	78/132 (59%)	*Corvus frugilegus*	2
Predicted: similar to nonhistone chromosomal protein HMG-14A	XP_420260.1	5.00E-11	32/34 (94%)	*Gallus gallus*	2
Cytochrome oxidase subunit III	AAQ05165.1	5.00E-12	33/37 (89%)	*Passerina amoena*	1
unnamed protein product	BAB71316.1	6.00E-29	65/70 (92%)	*Homo sapiens*	1
Hemoglobin alpha-A chain	P07407	1.00E-44	89/100 (89%)	*Eurasian tree sparrow*	1
Lectin, galactoside-binding, soluble	NP_006489.1	2.00E-24	54/103 (52%)	*Homo sapiens*	1
Hypothetical protein	NP_701705.1	5.00E-21	49/78 (62%)	*Plasmodium falciparum 3D7*	1
Rps24 protein	AAH58140.1	9.00E-15	38/47 (80%)	*Rattus norvegicus*	1
Unknown (protein for IMAGE: 4134193)	AAH22412.1	1.00E-27	58/62 (93%)	*Homo sapiens*	1
Predicted: similar to L-plastin (lymphocyte cytosolic protein 1)	XP_417047.1	1.00E-42	89/91 (97%)	*Gallus gallus*	1
Predicted: similar to expressed sequence AI132321	XP_421399.1	1.00E-17	53/82 (64%)	*Gallus gallus*	1
DOCK10 protein	AAH15018.1	2.00E-21	50/52 (96%)	*Homo sapiens*	1
Ubiquitin C	NP_062613.2	8.00E-45	92/96 (95%)	*Mus musculus*	1
Predicted: similar to hypothetical protein MGC14126	XP_422669.1	6.00E-41	83/117 (70%)	*Gallus gallus*	1
Predicted: similar to partitioning-defective protein 3 homologue	XP_418571.1	3.00E-34	51/73 (69%)	*Gallus gallus*	1
Predicted: similar to hypothetical protein	XP_426168.1	2.00E-17	44/44 (100%)	*Gallus gallus*	1
Homo sapiens mRNA for squalene epoxidase	D78130	*3.00E-13	*40/40 (100%)	*Homo sapiens*	1
Homo sapiens adult retina protein (LOC153222)	NM_153607	*7.00E-29	*86/92 (93%)	*Homo sapiens*	1
Novel					82
Total					123

*Nucleotide BLAST results.

known to have important roles in vertebrate immunity. Among the up-regulated clones, 29 (30%) corresponded to HSP90, an important gene in general cell proliferation and immunological defence against disease (see review in Wallin *et al*. 2002). This prompted us to confirm its differential expression by Northern blot analysis with RNAs from uninfected and infected individual birds (Fig. 5). The Northern blot analysis indicates HSP90 is up-regulated in all the infected birds. This further confirms the reverse Northern blot and macroarray results.

Discussion

Identifying genes in natural populations of hosts whose expression is modified by pathogens is a central goal of molecular ecology. In the present study, using a combination of suppression subtractive hybridization (SSH) and cDNA macroarray approaches, we have successfully identified about 100 differently expressed genes in house finch whose expression was influenced by infection with MG. Although our survey is almost certainly cursory and incomplete, it nonetheless provides a molecular sketch of an immunogenetic response to MG in a wild bird. Although we do not know whether any of the differentially expressed genes we have identified are responding specifically to MG or are more general in their activity against pathogens, ours is an important foundation for further analysis of host–parasite co-evolution in the house finch–*Mycoplasma* system. SSH is a powerful technique for identifying differently expressed transcripts that are candidates for further evolutionary analysis (Diatchenko *et al*. 1996). Although SSH has been previously used in identifying differently expressed genes in many organisms, such as bacteria (Dwyer *et al*. 2004), plants (Bassani *et al*. 2004), rodents (Kirsch *et al*. 2001), as well humans (Liang *et al*. 2004), to our knowledge, the data presented here is the first large scale functional genomic study by using this technique in an avian species. The approach is particularly appropriate for application to species whose genomes and expressed sequences have been very little characterized.

Among the genes identified as differentially expressed in our study, almost half were homologous to sequences

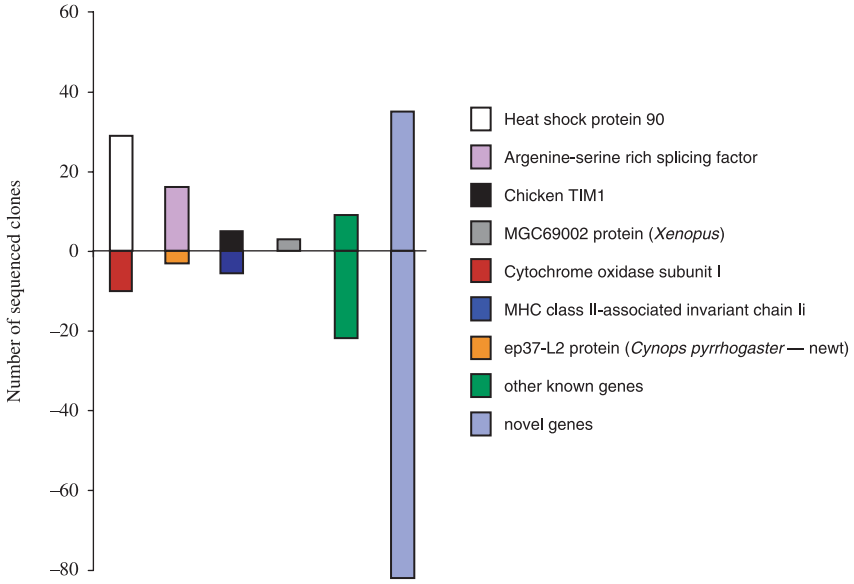

Fig. 4 Categories of differentially expressed house finch genes showing *Mycoplasma* induced gene expression changes. DNA sequences were searched against GenBank using BLASTX 2.2.10 version. Similarities that showed *e*-values less than 1e^{-08} with more than 100 nucleotides were considered significant.

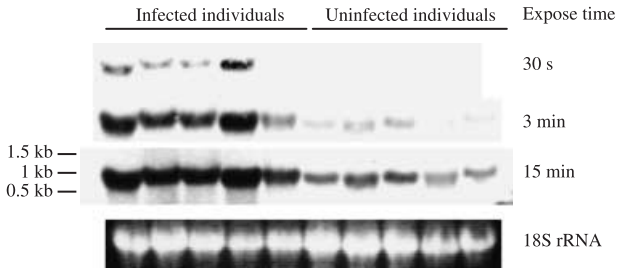

Fig. 5 A Northern blot hybridization of RNA from infected and uninfected birds using a finch HSP90 transcript as probe. Each lane was loaded 10 µg total RNA. The 18S rRNA is a control for equal amounts of RNA in each lane.

known from other species. We briefly summarize the functions of transcripts that appear noteworthy with regard to the immunogenetics of house finches. Among the up-regulated genes identified (Table 1), the most abundant were heat shock protein (HSP) 90, arginie/serine-rich plicing factor 12 (SFRS12), and TIM1.

- HSP90 has important roles in many immune and defence systems in vertebrates and invertebrates, and is evolutionarily quite conserved (Wallin *et al*. 2002).
- SFRS12 is a member of serine-arginine-rich (SR) family and regulates alternative splicing of RNAs by both positively and negatively modulating the activity of other SR proteins; the unique EK domain of this protein functions in inhibiting both constitutive and alternative splicing (Li *et al*. 2003). As such, SFRS12 could have diverse roles in modulating expression responses to environmental stimuli.
- TIM1 is a member of the recently described TIM (T-cell immunoglobulin mucin) gene family (Kuchroo *et al*. 2003). TIM1 has a demonstrated role as the cellular receptor for hepatitis A virus in humans, and is important for developmental differentiation of T-helper 1 and 2 cells (Kuchroo *et al*. 2003; Umetsu & Dekruyff 2004).
- Interferon-stimulated gene (ISG) 12 encodes a protein of ~M(r) 12 000 and belongs to a family of genes induced by interferon α (Martensen *et al*. 2001). Interferons (IFNs) are in turn a family of cytokines with growth inhibitory, and antiviral functions (Smidt *et al*. 2003).
- Granzyme A (GzmA) is an abundant protease in granules of cytotoxic T lymphocytes (CTLs) and NK cells. It plays an important role in generating reactive oxygen species (ROS) as part of the pathway causing cell death and mitochondrial damage leading to a state very similar to classical apoptosis (Martinvalet *et al*. 2005).

In the down-regulated known genes identified in our study (Table 2), the most frequently encountered were cytochrome oxidase subunit I (COI), MHC class II-associated invariant chain Ii, ep37-L2, EF1-a, NADH dehydrogenase subunit 4, and high mobility group (HMG) protein 14A. Of these, only the MHC class II-associated invariant chain appears to be strongly linked to specific immunological functions. It has been demonstrated that MHC class II-associated invariant chain is crucial to bind MHC class II proteins during MHC II biogenesis and function in antigen presentation (Stumptner-Cuvelette & Benaroch 2002). Cytochrome oxidase subunit I is known to play an important role in mitochondrial associated diseases, but is otherwise known as a key component of the oxidative respiratory pathway common to all eukaryotes (Lenaz *et al*. 2004). It is tempting to speculate that the downregulation of several mitochondrial genes that we observed could be linked to our finding that granzyme A is upregu-

lated in infected finches (Table 1), because expression of granzyme A is known to be a direct causes of mitochondrial degredation and DNA damage in pathogen-infected cells (Beresford *et al*. 1999; Martinvalet *et al*. 2005). EF-1α is an ancient and ubiquitous protein that plays a crucial role in translation of proteins by forming a ternary complex with GTP and aminoacyl-tRNA, thereby mediating the binding of aminoacyl-tRNA to the ribosome (Chang & Traugh 1998). It can be interpreted as a transcript functioning in increased cellular production of proteins, and is unlikely to be a genetic response specific to MG infection. HMG14a gene is a unique member of the high mobility group gene family with a role as a nonhistone chromosomal protein (Browne & Dodgson 1993), and is likely associated here with increased chromosomal modification as a prelude to increased expression of diverse genes. It is intriguing that several of these genes would have been predicted to be up-regulated rather than down-regulated in response to MG infection. We have shown elsewhere that MHC class II transcripts in house finches also appear down-regulated by infection with MG (Hess *et al*. submitted). It may be that MG exerts a suppressive role on some aspects of host response, including genes of immunological import. Several pathogens, such as cytomegaloviruses, are known to result in reduced expression of immunologically important genes in vertebrate hosts, presumably as a mechanism of immune evasion (Hengel *et al*. 1999). Alternatively, down-regulation of these genes could be adaptive if it alleviates the costs of an immune response and avoids immunopathology, particularly when individuals are stressed. Hanssen *et al*. (2004) have suggested shown that immune responses in wild birds entail costs to individuals that are otherwise experiencing stress, such as during the nesting season. Adaptive avoidance of immune response (Raberg *et al*. 1998) should be considered in future investigations of house finch regulation of immunogenetic responses.

In the present study, almost half of the differently expressed clones identified have no known matches in the existing databases. These genes are called 'orphan' genes (Schmid & Aquadro 2001; Domazet-Loso & Tautz 2003), and the biological roles of such orphan genes here remain to be evaluated. As our study was being conducted, the chicken genome was sequenced; despite this, it may be that many expressed sequences in chicken remain to be discovered and characterized. As the database grows, it is likely that many of the genes currently designated as 'novel' in house finches will find matches. Orphan genes may also be the result of rapid evolution such that sequence similarity is lost even within relatively short evolutionary time spans (Schmid & Aquadro 2001; Domazet-Loso & Tautz 2003). The neutral theory hypothesizes a correlation between the functional importance of a gene and its evolutionary conservation (Kimura 1983). However, ESTs involved in immunity or in genomic conflicts, such as sexually antagonistic genes, often exhibit rapid evolution (Endo *et al*. 1996; Swanson *et al*. 2001; Tiffin & Hahn 2002; Barrier *et al*. 2003; Swanson *et al*. 2004). A more mundane explanation, however, is that our sequences failed to cover critically important residues and domains that might aid in identification. It will be important to further investigate whether these orphan genes comprise a distinct group of house finch transcripts that are genuinely distinct from chicken genes and perhaps are derived at a higher level within passerines or other avian clades.

What are the next steps in the study of the immunogenetic response of house finches to *Mycoplasma*? A logical question for evolutionary biologists is whether any of the genes identified in our screen may have participated in the adaptation of house finch populations to selection imposed by the parasite in the eastern United States. Our study has provided a lengthy list of possible candidate genes whose evolutionary trajectory in pre- and postepizootic populations can now be explored. Adaptive evolution at the molecular level is more likely to be discovered in the regulatory networks underlying the expression of house finch genes, rather than in the coding regions of these genes; indeed, several of the genes whose expression was altered in our study are highly evolutionarily conserved in vertebrates (e.g. HSP90). Thus one possible approach would be to look for adaptive mutations in the upstream regulatory elements of genes such as HSP90 in evolving populations of house finches, as has been carried out for heat shock regions in the *Drosophila* genome (Mckechnie *et al*. 1998; Feder 2000; Lerman *et al*. 2003). HSP90 is itself highly pleiotropic and an agent of canalization in *Drosophila*, and may be linked to morphological change in house finch populations during the epizootic (Rutherford & Lindquist 1998; Rutherford 2003). At the other extreme, there may be no necessary link between the presumably plastic regulatory responses observed in our study and adaptation at the population level, or differential survival of susceptible and resistant birds to infection in the wild. At the very least, our study has produced a preliminary description of an immunogenetic response in a wild bird which we hope can help identify genes in this and other species that may be important in modulating infection by pathogens on both ecological and evolutionary timescales.

In summary, by combining using SSH, inverse Northern blot, and cDNA macroarray approaches, we have identified about 100 differently expressed genes in house finch which are regulated by MG parasite. Among them, many known genes identified play important roles in disease defence systems. The novel orphan genes identified in the present study might have unique roles to understand the dynamics of this disease system. Our results and macroarray resources provide a foundation toward to understanding the ecological genetics of infectious disease and

molecular interactions between the *Mycoplasma* parasite and its recently colonized avian host.

Acknowledgements

We are grateful to Dr Shizhen Qin of the Institute for Systems Biology for allowing use of the robotic clone picker, and to Dr Chris T. Amemiya for assistance in arraying on membranes. We wish to thank Edwards laboratory members at University of Washington for substantial assistance during the research and Suzanne Rutherford, Craig Moritz and Martin Feder for advice and discussion. Two anonymous reviewers provided helpful comments on the manuscript. This work was supported by National Science Foundation grants (DEB-IRCEB 0077804) to Geoffrey Hill, Scott Edwards, and Sharon Roberts.

References

Bachem CW, Hoeven RSvd, Bruijn SMd *et al.* (1996) Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. *Plant Journal*, **9**, 745–753.

Barrier M, Bustamante C, Yu J, Purugganan M (2003) Selection on rapidly evolving proteins in the *Arabidopsis* genome. *Genetics*, **163**, 723–733.

Bassani M, Neumann P, Gepstein S (2004) Differential expression profiles of growth-related genes in the elongation zone of maize primary roots. *Plant Molecular Biology*, **56**, 367–380.

Beresford PJ, Xia Z, Greenberg AH, Lieberman J (1999) Granzyme A loading induces rapid cytolysis and a novel form of DNA damage independently of caspase activation. *Immunity*, **10**, 585–594.

Browne D, Dodgson J (1993) The gene encoding chicken chromosomal protein HMG-14a is transcribed into multiple mRNAs. *Gene*, **124**, 199–206.

Chang Y, Traugh J (1998) Insulin stimulation of phosphorylation of elongation factor 1 (eEF-1) enhances elongation activity. *European Journal of Biochemistry*, **251**, 201–207.

Cossins AR, Crawford DL (2005) Fish as models for environmental genomics. *Nature Reviews Genetics*, **6**, 324–333.

Dhondt AA, Tessglia DL, Slothower RL (1998) Epidemic mycoplasmal conjunctivitis in house finches from eastern North America. *Journal of Wildlife Diseases*, **34**, 265–280.

Diatchenko L, Lau YF, Campbell AP *et al.* (1996) Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proceedings of the National Academy of Sciences, USA*, **93**, 6025–6030.

Domazet-Loso T, Tautz D (2003) An evolutionary analysis of orphan genes in *Drosophila*. *Genome Research*, **13**, 2213–2219.

Dwyer K, Lamonica J, Schumacher J *et al.* (2004) Identification of *Bacillus anthracis* specific chromosomal sequences by suppressive subtractive hybridization. *BMC Genomics*, **5**, 15.

Endo T, Ikeo K, Gojobori T (1996) Large-scale search for genes on which positive selection may operate. *Molecular Biology and Evolution*, **13**, 685–690.

Farmer K, Hill G, Roberts SR (2002) Susceptibility of a naive population of house finches to *Mycoplasma gallisepticum*. *Journal of Wildlife Diseases*, **38**, 282–286.

Fischer JR, Stallknecht DE, Luttrell P, Dhondt AA, Converse KA (1997) Mycoplasmal conjunctivitis in wild songbirds: the spread of a new contagious disease in a mobile host population. *Emerging Infectious Diseases*, **3**, 69–72.

Granjeaud S, Bertucci F, Jordan B (1999) Expression profiling: DNA arrays in many guises. *BioEssays*, **21**, 781–790.

Hanssen SA, Hasselquist D, Folstad I, Erikstad KE (2004) Costs of immunity: immune responsiveness reduces survival in a vertebrate. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **271**, 925–930.

Hegarty MJ, Jones JM, Wilson ID *et al.* (2005) Development of anonymous cDNA microarrays to study changes to the *Senecio* floral transcriptome during hybrid speciation. *Molecular Ecology*, **14**, 2493–2510.

Hengel H, Reusc U, Gutermann A *et al.* (1999) Cytomegaloviral control of MHC class I function in the mouse. *Immunological Reviews*, **168**, 167–176.

Hess CM, Edwards SV (2002) The evolution of the major histocompatibility complex in birds. *Bioscience*, **52**, 423–431.

Hess CM, Gasper J, Hoekstra H, Hill C, Edwards SV (2000) MHC class II pseudogene and genomic signature of a 32-kb cosmid in the House Finch (*Carpodacus mexicanus*). *Genome Research*, **10**, 613–623.

Hill GE (1993) House finch. In: *Birds of North America* (Eds A. Poole, P. Stettenheim, F. Gill). Philadelphia: American Ornithologists' Union.

Hill GE (2002) *A Red Bird in a Brown Bag: The Function and Evolution of Colorful Plumage in the House Finch*. Oxford University Press, New York.

Hill GE (2006) Environmental regulation of ornamental coloration. In: *Bird Coloration*, Vol. 1. *Mechanisms and Measurements* (eds Hill GE, McGraw KJ), pp. 507–560. Harvard University Press, Cambridge, Massachusetts.

Hubank M, Schatz DG (1994) Identifying differences in mRNA expression by representational difference analysis of cDNA. *Nucleic Acids Research*, **22**, 5640–5648.

International Chicken Genome Sequencing Consortium (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, **432**, 695–716.

John JL (1994a) The avian spleen: a neglected organ. *Quarterly Review of Biology*, **69**, 327–351.

John JL (1994b) Nematodes and the spleen: an immunological relationship. *Experientia*, **50**, 15–22.

Kilpimaa J, Alatalo RV, Siitari H (2004) Trade-offs between sexual advertisement and immune function in the pied flycatcher (*Ficedula hypoleuca*). *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **271**, 245–250.

Kimura M (1983) *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge, UK.

Kirsch T, Wellner M, Luft F, Haller H, Lippoldt A (2001) Altered gene expression in cerebral capillaries of stroke-prone spontaneously hypertensive rats. *Brain Research*, **910**, 106–115.

Kuchroo Vijay K, Umetsu Dale T, Dekruyff Rosemarie H, Freeman Gordon J (2003) The TIM gene family: emerging roles in immunity and disease. *Nature Immunology*, **3**, 454–462.

Lenaz G, Baracca A, Carelli V *et al.* (2004) Bioenergetics of mitochondrial diseases associated with mtDNA mutations. *Biochemica et Biophysica Acta*, **1658**, 89–94.

Lerman DN, Michalak P, Helin AB, Bettencourt BR, Feder ME (2003) Modification of heat-shock gene expression in *Drosophila melanogaster* populations via transposable elements. *Molecular Biology and Evolution*, **20**, 135–144.

Li H, Gu X, Dawson VL, Dawson TM (2004) Identification of calcium- and nitric oxide-regulated genes by differential analysis

of library expression (DAzLE). *Proceedings of the National Academy of Sciences, USA*, **101**, 647–652.

Li J, Hawkins I, Harvey C *et al.* (2003) Regulation of alternative splicing by SRrp86 and its interacting proteins. *Molecular and Cellula Cell Biology*, **23**, 7437–7447.

Liang P, Pardee AB (1992) Differential display of eukaryotic messager RNA by means of the polymerase chain reaction. *Science*, **257**, 967–971.

Liang G, Zhang X, Wang L *et al.* (2004) Identification of differentially expressed genes of primary spermatocyte against round spermatid isolated from human testis using the laser capture microdissection technique. *Cell Research*, **14**, 507–512.

Luttrell MP, Fischer JR, Stallknecht DE, Kleven SH (1996) Field investigations of *Mycoplasma gallisepticum* infection in house finches (*Carpodacus mexicanus*) from Maryland and Georgia. *Avian Diseases*, **40**, 335–341.

Luttrell MP, Stallknecht DE, Fischer JR, Sewell CT, Kleven SH (1998) Natural *Mycoplasma galliseptium* infection in a captive flock of house finches. *Journal of Wildlife Diseases*, **34**, 289–296.

Martensen PM, Sogaard TM, Gjermandsen IM *et al.* (2001) The interferon alpha induced protein ISG12 is localized to the nuclear membrane. *European Journal of Biochemistry*, **268**, 5947–5954.

Martinvalet D, Zhu P, Lieberman J (2005) Granzyme A induces caspase-independent mitochondrial damage, a required first step for apoptosis. *Immunity*, **22**, 355–370.

Mckechnie SW, Halford MM, McColl G, Hoffmann AA (1998) Both allelic variation and expression of nuclear and cytoplasmic transcripts of *Hsr-omega* are closely associated with thermal phenotype in *Drosophila*. *Proceedings of the National Academy of Sciences, USA*, **95**, 2423–2428.

Oleksiak MF, Churchill GA, Crawford DL (2002) Variation in gene expression within and among natural populations. *Nature Genetics*, **32**, 261–266.

Pillai S, Mays HL, Jr., Ley D *et al.* (2003) Molecular variability of house finch Mycoplasma gallisepticum isolates as revealed by sequencing and restriction fragment length polymorphism analysis of the pvpA gene. *Avian Diseases*, **47**, 640–648.

Quackenbush J (2001) Computational analysis of microarray data. *Nature Reviews in Genetics*, **2**, 418–427.

Raberg L, Grahn M, Hasselquist D, Svensson E (1998) On the adaptive significance of stress-induced immunosuppression. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **265**, 1637–1641.

Rast JP, Amore G, Calestani C *et al.* (2000) Recovery of developmentally defined gene sets from high-density cDNA macroarrays. *Developmental Biology*, **228**, 270–286.

Renn S, Aubin-Horth N, Hofmann H (2004) Biologically meaningful expression profiling across species using heterologous hybridization to a cDNA microarray. *BMC Genomics*, **5**, 42.

Roberts SR, Nolan PM, Hill GE (2001a) Characterization of *Mycoplasma gallisepticum* infection in captive house finches (*Carpodacus mexicanus*) in 1998. *Avian Diseases*, **45**, 70–75.

Roberts SR, Nolan PM, Lauerman LH, Li LQ, Hill GE (2001b) Characterization of the mycoplasmal conjunctivitis epizootic in a house finch population in the southeastern USA. *Journal of Wildlife Diseases*, **37**, 82–88.

Rutherford SL (2003) Between genotype and phenotype: protein chaperones and evolvability. *Nature Reviews Genetics*, **4**, 263–274.

Rutherford SL, Lindquist S (1998) Hsp90 as a capacitor for morphological evolution. *Nature*, **396**, 336–342.

Sargent TD, Dawid IB (1983) Differential gene expression in the gastrula of *Xenopus laevis*. *Science*, **222**, 135–139.

Schmid K, Aquadro C (2001) The evolutionary analysis of 'orphans' from the *Drosophila* genome identifies rapidly diverging and incorrectly annotated genes. *Genetics*, **159**, 589–598.

Smidt K, Hansen L, Sogaard T *et al.* (2003) A nine-nucleotide deletion and splice variation in the coding region of the interferon induced ISG12 gene. *Biochemica et Biophysica Acta*, **1638**, 227–234.

von Stein OD, Thies WG, Hofmann M (1997) A high throughput screening for rarely transcribed differentially expressed genes. *Nucleic Acids Research*, **25**, 2598–2602.

Stumptner-Cuvelette P, Benaroch P (2002) Multiple roles of the invariant chain in MHC class II function. *Biochemica et Biophysica Acta*, **1542**, 1–13.

Swanson W, Clark A, Waldrip-Dail H, Wolfner M, Aquadro C (2001) Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. *Proceedings of the National Academy of Sciences, USA*, **98**, 7375–7379.

Swanson W, Wong A, Wolfner M, Aquadro C (2004) Evolutionary expressed sequence tag analysis of *Drosophila* female reproductive tracts identifies genes subjected to positive selection. *Genetics*, **168**, 1457–1465.

Tiffin P, Hahn M (2002) Coding sequence divergence between two closely related plant species: *Arabidopsis thaliana* and *Brassica rapa* ssp. *pekinensis*. *Journal of Molecular Evolution*, **54**, 746–753.

Umetsu D, Dekruyff R (2004) Regulation of tolerance in the respiratory tract: TIM-1, hygiene, and the environment. *Annals of the New York Academy of Science*, **1029**, 88–93.

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW (1995) Serial analysis of gene expression. *Science*, **270**, 484–487.

Wallin R, Lundqvist A, More S *et al.* (2002) Heat-shock proteins as activators of the innate immune system. *Trends in Immunology*, **23**, 130–135.

Yang YH, Speed T (2002) Design issues for cDNA microarray experiments. *Nature Reviews Genetics*, **3**, 579–588.

Dr. Zhenshan Wang's work has focused on the ecological genomics of plants and animals. Dr. Geoffrey Hill's interests focus on the function and evolution of ornamental traits in birds. With his colleagues he has studied the effects of pathogens on plumage coloration and is very interested in whether ornamental traits might signal good genes. Kristy Farmer is completing her graduate work on avian diseases and host/pathogen co-evolution. Dr. Scott Edwards' interests are in avian molecular evolution, population genetics and ecological genomics.
