Changes in transcript abundance relating to colony collapse disorder in honey bees (*Apis mellifera*)

Reed M. Johnson[a,1], Jay D. Evans[b], Gene E. Robinson[a], and May R. Berenbaum[a,2]

[a]Department of Entomology, University of Illinois at Urbana-Champaign, Urbana, IL 61801; and [b]Bee Research Laboratory, Agricultural Research Service, United States Department of Agriculture, Beltsville, MD 20705

Contributed by May R. Berenbaum, July 14, 2009 (sent for review February 18, 2009)

Colony collapse disorder (CCD) is a mysterious disappearance of honey bees that has beset beekeepers in the United States since late 2006. Pathogens and other environmental stresses, including pesticides, have been linked to CCD, but a causal relationship has not yet been demonstrated. Because the gut acts as a primary interface between the honey bee and its environment as a site of entry for pathogens and toxins, we used whole-genome microarrays to compare gene expression between guts of bees from CCD colonies originating on both the east and west coasts of the United States and guts of bees from healthy colonies sampled before the emergence of CCD. Considerable variation in gene expression was associated with the geographical origin of bees, but a consensus list of 65 transcripts was identified as potential markers for CCD status. Overall, elevated expression of pesticide response genes was not observed. Genes involved in immune response showed no clear trend in expression pattern despite the increased prevalence of viruses and other pathogens in CCD colonies. Microarray analysis revealed unusual ribosomal RNA fragments that were conspicuously more abundant in the guts of CCD bees. The presence of these fragments may be a possible consequence of picorna-like viral infection, including deformed wing virus and Israeli acute paralysis virus, and may be related to arrested translation. Ribosomal fragment abundance and presence of multiple viruses may prove to be useful diagnostic markers for colonies afflicted with CCD.

microarray | picorna-like virus | ribosomal RNA

As the premier managed pollinator in the United States, the western honey bee, *Apis mellifera*, contributes more than $14 billion to agriculture annually (1). Beginning in fall 2006, the American apiculture industry experienced catastrophic losses of unknown origin. The phenomenon, called colony collapse disorder (CCD), was identified by a set of distinctive characteristics, including the absence of dead bees in or near the colony and the presence of abundant brood, honey, and pollen despite vastly reduced numbers of adult workers (2). Losses were estimated at 23% over the winter of 2006–2007 (3) and at 36% over the winter of 2007–2008 (4).

Using metagenomics, Cox-Foster et al. (3) compared the microbial flora of honey bees in hives diagnosed with CCD and in ostensibly healthy hives, evaluating pathogens with respect to their association with diagnosed CCD. The most predictive pathogen was Israeli acute paralysis virus (IAPV), a picorna-like virus (Dicistroviridae) hitherto unreported in the United States. However, a later study found IAPV in U.S. bees before the appearance of CCD (5), discounting an exclusive causal relationship. The metagenomic analysis identified other pathogens associated with CCD bees, including the picorna-like viruses Kashmir bee virus (KBV) and deformed wing virus (DWV), and 2 species in the microsporidian genus *Nosema* (3), one of which, *N. ceranae*, was subsequently linked to collapses of colonies in Spain (6), although this may be unrelated to CCD in the United States. A high prevalence of multiple pathogens in CCD bees suggests that a compromised immune response may be integral to CCD.

Pesticides also have been suspected to play a role in CCD. Sublethal exposures leading to behavioral disruptions consistent with the failure of foragers to return to their hives have been associated with the neonicotinoid pesticides (7). The long-term use of combinations of in-hive pesticides for control of honey bee parasites also may have contributed to otherwise unexplained bee mortality (8).

To differentiate among possible explanations for CCD, we used whole-genome microarray analysis, comparing gut gene expression in adult worker bees from healthy and CCD colonies. We assayed the gut because it is the principal site of pesticide detoxification and an integral component in the immune defense against pathogens in *A. mellifera*. Samples from colonies varying in CCD severity were collected on the U.S. east and west coasts in the winter of 2006–2007 and compared with healthy ("historical") controls collected in 2004 and 2005.

Results

Gene Expression Differences. The microarray contained oligonucleotide probes representing 9,867 different genes (with duplicate spots), based on gene predictions and annotation from the honey bee genome sequencing project; in addition, 2,729 probes specific for ESTs matching no existing annotation were represented on the array (9–11), including 21 probes derived from expressed sequence tags (ESTs) specific for rRNA [supporting information (SI) Table S1].

Of the 6,777 probes with expression above background level, 1,305 probes showed significant differences in expression in at least 1 of the 5 contrasts [false discovery rate (FDR), $P < .01$; and fold-change >2) (Fig. 1]. The west coast versus historical comparison alone generated a list of 948 differentially expressed probes; however, 668 of these were not differentially expressed in any of the other comparisons.

There was little concordance in the lists of differentially expressed genes between the east and west coast comparison of guts taken from "severe" and "mild" colonies. Just 54 probes shared differential expression between the east and west coast severe versus mild comparisons, with 191 probes showing unique differential expression. That so many genes are differentially transcribed in the severe versus mild comparisons suggests these classifications may not be equivalent across geographical samplings.

Reflecting the strong differences in gene expression, hierarchical cluster analysis, using expression values for all probes,

Author contributions: R.M.J., J.D.E., G.E.R., and M.R.B. designed research; R.M.J. and J.D.E. performed research; J.D.E. and G.E.R. contributed new reagents/analytic tools; R.M.J., J.D.E., G.E.R., and M.R.B. analyzed data; and R.M.J., J.D.E., G.E.R., and M.R.B. wrote the paper.

The authors declare no conflict of interest.

Data deposition: The data reported in this paper have been deposited in the ArrayExpress database, www.ebi.ac.uk/arrayexpress (accession no. E-MEXP-2292).

[1]Present address: Department of Entomology, University of Nebraska-Lincoln, Lincoln, NE 68583.

[2]To whom correspondence should be addressed. E-mail: maybe@illinois.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0906970106/DCSupplemental.

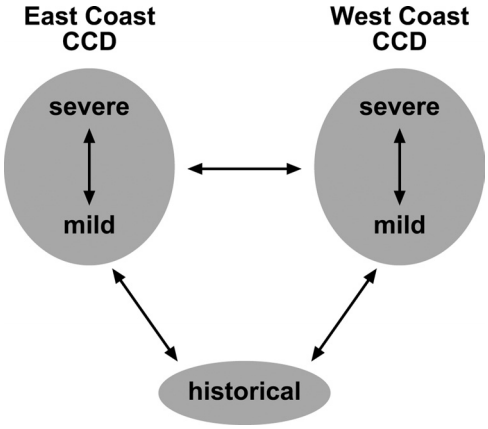

Fig. 1. Experimental design for microarray comparisons of CCD and healthy bees. East coast CCD includes bees collected in Florida and Pennsylvania, and west coast CCD includes bees from California. Guts from bees remaining in colonies that were classified as mild or severe were compared with each other and together, through a reference gut RNA sample, with guts of healthy historical bees collected in 2004 and 2005. A total of 22 microarrays were hybridized.

grouped the samples broadly by health status and geography (Fig. 2). All healthy bees clustered together, as did all CCD bees, which were further clustered according to geography, with west coast California bees forming one group and east coast Pennsylvania and Florida bees clustering together in another group. Differences in relative colony health at the time of collection, either mild or severe CCD, were not reflected in the cluster analysis.

Although geographic variation among the east and west coast CCD samples is evident, if CCD is a distinct phenomenon, then it should be possible to identify a list of CCD-related genes that show differential expression in both east and west coast CCD bees relative to healthy bees. But because these comparisons rely on the historical samples, which may carry their own geographic biases, we further winnowed the gene list by using only genes differentially expressed in the east coast severe versus mild comparison (Fig. 3). This comparison was informative; expression of 65 transcripts was consistently up or down when comparing relatively sick bees (CCD or severe) and relatively healthy bees (historically healthy or mild) (Fig. 2); 29 probes on this list correspond to annotated genes, 35 probes correspond to ESTs, and 1 probe corresponds to a pathogen, DWV.

Gene Ontology Functional Analyses of Gene Expression Differences. Genes differentially expressed between groups were categorized according to Gene Ontology (GO) (12). FlyBase orthologs of bee genes that were differentially expressed were used for GO enrichment analyses (Table 1). In the cellular component ontology, genes associated with lipid particle (GO:0005811), cytosolic small ribosomal subunit (GO:0022627), and mitochondrial respiratory chain (GO:0005746) generally demonstrated reduced expression in west coast bee guts relative to historical samples.

In the molecular function ontology, genes related to transcription factor activity (GO:0003700) were more highly expressed in both east coast and west coast CCD bees and in severe CCD bees. West coast bees also showed reduced expression of both heme binding (GO:0020037) and carboxylesterase activity (GO:0004091) categorizations, which include cytochrome P450 monooxygenases (P450s) and carboxylesterases (COEs), the principal xenobiotic metabolizing enzymes. One P450, CYP4G11, was expressed at a lower level in all 3 contrasts tested.

In the biological process ontology, genes related to development (GO:0016348 and GO:0007509) and the peptidoglycan catabolic process (GO:0009253) were overrepresented but did not demonstrate consistent directional differences. The peptidoglycan recognition proteins (PGRPs) contribute to immunity by detecting bacteria (13); PGRP-S2 was up-regulated and PGRP-S3 was down-regulated in CCD bees. Transcripts encoding sallimus, a large complex protein that mediates muscle elasticity (14), were more abundant in west coast bees. This protein also may contribute to immunity, because it contains Ig repeats and its transcripts are overexpressed in *Anopheles gambiae* mosquitoes after bacterial infection (15).

The gene *mblk-1* was more highly expressed in CCD bees; the *D. melanogaster* ortholog of this gene plays a role in hormone-triggered cell death (16) and larval midgut histolysis (17). Thus, overexpression of *mblk-1* may indicate apoptosis in the guts of CCD bees. However, of the 39 apoptosis-related (GO:0006915) genes on the microarray, only *mblk-1* and *GB14659*, which is similar to *apoptotic peptidase activating factor 1*, were more highly expressed in the guts of CCD bees.

Pathogens in CCD Bees. Twenty-two probes specific for 8 bee pathogens were present on the microarray [KBV, DWV, black queen cell virus (BQCV), acute bee paralysis virus, sacbrood virus, *Ascophaera apis*, *Nosema apis*, and *Paenibacillus larvae*] (Fig. S2), providing an opportunity to survey the samples for pathogens. In some cases, the number of colonies sampled was small with respect to standard pathogen surveys, so the results are best interpreted in a qualitative rather than quantitative manner. Three pathogens were detectable at different levels among the sampled bees. Chalkbrood (*A. apis*) RNA was more abundant in both east coast CCD bees and in severe CCD bees. BQCV RNA was more abundant in CCD bees, as was DWV, but DWV also was more abundant in east coast severe CCD bees. The DWV probe on the microarray also may have detected the closely related Kakugo virus or *Varroa destructor* virus 1 (18), with which it shares 91% identity. A probe specific for IAPV (19) was not included on the array and likely is not detected by the probe for KBV, with which it shares just 60% identity. The presence of *Nosema*, IAPV, and other viruses was surveyed with quantitative PCR (qPCR), indicating a generally higher viral load in CCD colonies (Table 3).

Validation of Microarray Results with qPCR. Expression changes of 8 transcripts were verified by qPCR. Five of 8 transcripts had similar expression profiles when measured using the same RNA samples analyzed by both methods (Table S4). Biological validation of microarray results was performed using qPCR on gut samples from healthy and CCD bees collected in California in 2007, a year after the samples used for microarray analysis were obtained. Differential expression of 3 of 8 transcripts was confirmed: CCD-associated up-regulation of the ESTs QW33 and jdeC15, corresponding to 28S and 5.8S rRNA subunits. DWV also was more abundant in CCD gut samples (Table 2).

Further qPCR validation of markers was performed with abdomens of bees from 147 colonies in healthy and CCD apiaries in 6 states. Only an rRNA-coding EST, QW33, was differentially expressed between CCD and healthy abdomens (Table 2). These CCD bees also were more likely to harbor both a larger number of viruses (as detected by these methods) and nonviral pathogens, as in the original samples (Table 3).

Discussion

No simple explanations for the cause of CCD emerge from the microarray analysis. Expression of detoxification and immune gene transcripts, which would indicate toxins or disease as the cause of CCD, was largely unchanged. However, considerable geographic variation existed among CCD bees, with west coast bees seemingly more severely affected. Transcripts of genes related to basic cellular processes involving ribosomal and

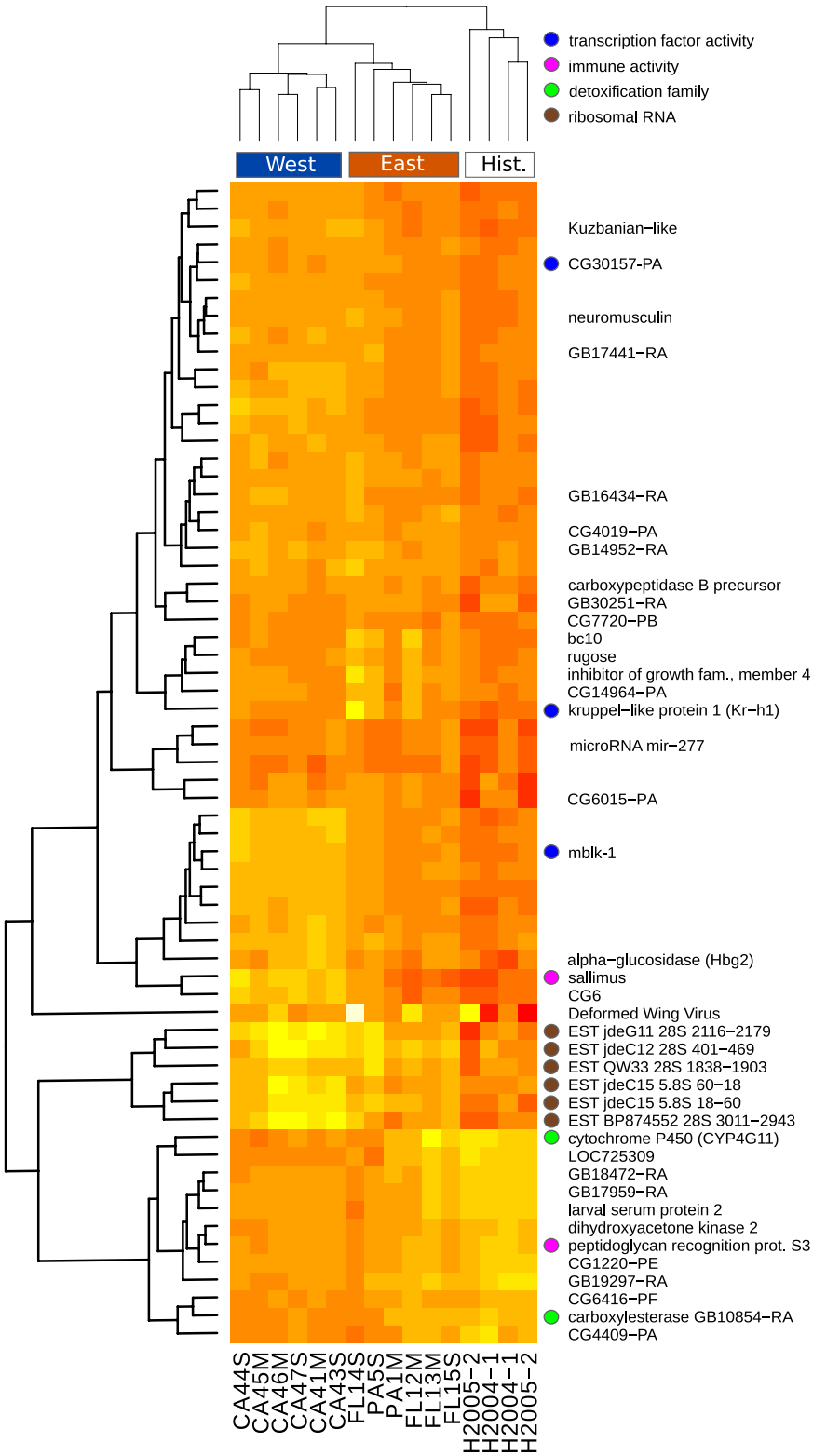

Fig. 2. Heatmap showing expression values for 65 probes demonstrating differential expression in 3 contrasts: east coast CCD versus historical, west coast CCD versus historical, and east coast severe versus east coast mild [$P < .01$ (FDR); fold-change >2]. Yellow indicates increased transcript abundance and red indicates decreased transcript abundance relative to the mean for all colonies. Probes lacking a description correspond to EST sequences for which no matching gene has been found. More information about the probes is provided in Table S5. CA, California; PA, Pennsylvania; FL, Florida. Numbers correspond to colony number. S, severe CCD; M, mild CCD. Colonywise clustering was performed using expression values for all 6,777 probes. Euclidean distances were calculated and clustered using the ''complete'' method.

mitochondrial function were generally less abundant in west coast bees.

P450s and COEs play multiple metabolic roles and are particularly important in the detoxification of natural and synthetic toxins in bees and other insects (20). However, the few genes in these superfamilies differentially transcribed with CCD likely serve functions other than detoxification. Expression of CYP4G11 was elevated in CCD bees, but the function of this P450 remains uncharacterized. Its ortholog in *D. melanogaster*, CYP4G15, is expressed only in the central nervous system, where it may be associated with ecdysteroid metabolism (21). The single COE overexpressed in CCD bees, GB10854, differs from other insect COEs in terms of its catalytic site, such that it may be incapable of detoxificative carboxylester hydrolysis (22).

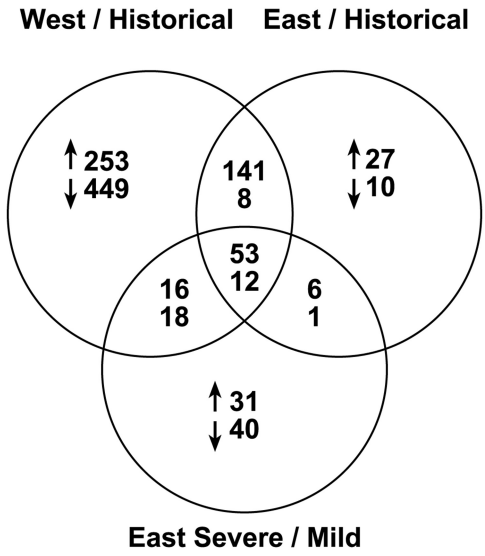

Fig. 3. Venn diagram showing the number of transcripts up-regulated and down-regulated in 3 comparisons: west coast (California) CCD versus historical, east coast (Florida and Pennsylvania) CCD versus historical, and east coast severe CCD versus mild CCD.

Although 88 known immune-related genes were included on the microarray, these also were mostly unchanged in expression in CCD bees. Given the association between pathogens and CCD (3), the finding that so few immune genes were differentially expressed in CCD bees was unexpected. Insect guts can produce antimicrobial peptides, including the apidaecins in honey bees (23), yet CCD-related changes in transcription of these genes were not detected. This lack of response does not rule out viral infection, however, because antimicrobial proteins are ineffective against viruses. In *D. melanogaster*, fungal or bacterial infection induces antimicrobial peptides, but flies infected with *Drosophila* C virus (DCV), another picorna-like virus, did not produce more antimicrobial peptides (24) and demonstrated elevated transcription of only a small number of immune genes relative to fungal and bacterial infection (25).

Many of the transcripts differentially expressed in CCD bees were detected by probes based on ESTs matching no existing annotation (10, 11). Of the 2,729 geneless ESTs on the array, 28 were differentially expressed in CCD bees. Of all of the ESTs on which probes were based, 13 were discovered to contain rRNA sequence, and 5 of these were more highly expressed in CCD bee guts. These rRNA-coding ESTs were not identified as rRNA previously, because the *A. mellifera* rRNA sequence is not available in online repositories, and identification of these probes was accomplished only by manually transcribing the published rRNA sequence from printed text (26). Had these ESTs been previously identified as rRNA, the associated probes likely would have been excluded from the microarray.

The presence of rRNA on these microarrays is surprising, because both the EST projects on which the probes were based and the microarray hybridization rely on a polyadenylated 3′ tail for the initial priming step in the reverse-transcription reaction. Ribosomal RNA is transcribed by RNA polymerase I, and the resulting transcript is not polyadenylated; as such, oligo(dT)-based reverse-transcription reactions should not reverse-transcribe rRNA. Yet transcripts of rRNA have long appeared in EST projects, and their reverse transcription has been attributed to poly(A)-rich internal sequences hybridizing to the oligo(dT) or to genomic DNA contamination (27). This differential expression of polyadenylated rRNA transcripts in insects currently defies technical explanation; differential expression of these transcripts associated with age and caste in a termite (*Reticulotermes flavipes*) EST project (28) suggests an underlying biological explanation.

One possible explanation for the presence of poly(A) rRNA sequences in bees is that they are degradation intermediates. Polyadenylation is a well-known marker of RNA degradation in bacteria and organelles (29), and poly(A) addition is known to contribute to eukaryotic rRNA degradation (30). Polyadenylated 25S rRNA comprises ≈0.02% of all of the 25S rRNA transcripts in fission yeast *Schizosaccharomyces pombe* (31). Another explanation for these sequences is the "ribosomal filter hypothesis," in which conventional mRNA transcripts include tens to hundreds of nucleotides similar to rRNA sequence in either the sense or antisense direction (32). Transcripts of genes containing rRNA sequence can bind to ribosomal subunits and alter translation efficiency (33). Short sections of rRNA function as an internal ribosome entry site (IRES) to interact with the ribosome and alter translation. Indeed, functions have been assigned to some mRNAs containing large sections of rRNA sequence (34), although none of the predicted genes annotated from the honey bee genome contains sufficient rRNA sequence to hybridize to any of the rRNA probes on the array (Table S1). A related hypothesis posits that long sequences of rRNA-like transcript may aid in protein folding, taking on part of the enzymatic activity of rRNA (35).

The picorna-like virus DWV was more abundant in CCD bees analyzed in the microarray experiment and a broader sampling of bee colonies revealed that CCD bees carry a larger number of different picorna-like viruses (Table 3). Picorna-like viruses, including IAPV and KBV, that have been associated with CCD (3) initiate transcription of viral proteins by the ribosome through an IRES rather than through the 5′-methylation cap that initiates translation of most mRNA (36). The IRES sequences from both DWV and *Varroa destructor* virus 1 effectively enhance translation of a reporter gene (36). Picorna-like viruses in mammals also halt translation of host mRNA through cleavage of the translation initiation factor eIF4G and the poly(A)-binding protein (37), leaving the ribosomes incapable of binding host mRNA and giving the viral RNA little competition for its IRES-mediated binding to host ribosomes (38).

While there is no direct evidence that elevated poly(A) rRNA is a result of picorna-like viral infection, one consequence of the viral disabling of ribosomal function may well be increased ribosomal degradation. Picornavirus infection in mammals both reduces protein production and causes strings of translating polyribosomes to break down, and these idle ribosomal subunits may be more susceptible to degradation (39). Ribosomal degradation also could occur in CCD bees independent of viral infection; for example, starving bacteria demonstrate elevated ribosomal degradation (39). Regardless of the cause, the abundance of poly(A) rRNA may be a useful diagnostic marker for determining CCD status, although the presence of poly(A) rRNA does not in itself indicate CCD, because healthy bees, including those used to generate EST libraries, contain these fragments.

Although gene transcript analysis did not clearly identify a specific cause for CCD, our study documents several patterns suggestive of a causal mechanism. The reduced protein synthetic capabilities that would accompany ribosomal hijacking by multiple picorna-like viruses would leave bees unable to respond to additional stresses from pesticides, nutrition, or pathogens. Although any interpretation of the presence of these rRNA fragments is speculative, the reported interaction between bee picorna-like viruses and rRNA is suggestive of a possible root cause of CCD.

To establish a causal relationship, the quantitative association between multiple picorna-like virus infections and polyadenyl-

Table 1. GO analyses for differentially expressed genes in 3 contrasts

Comparison	GO ID	Term	Annotated	Significant	Expected	*P* value	Change in expression
Cellular component							
East vs. historical	GO:0035102	PRC1 complex	3	2	0.09	.003	+2/−0
	GO:0005576	Extracellular region	124	11	3.6	.008	+6/−5
West vs. historical	GO:0005811	Lipid particle	205	46	26.72	<.001	+6/−40
	GO:0022627	Cytosolic small ribosomal subunit	30	10	3.91	.003	+0/−10
	GO:0005746	Mitochondrial respiratory chain	53	14	6.91	.006	+4/−10
Molecular function							
East vs. historical	GO:0003700	Transcription factor activity	144	15	3.91	<.001	+14/−1
	GO:0008599	Protein phosphatase type 1 reg. activity	2	2	0.05	.001	+2/−0
	GO:0043565	Sequence-specific DNA binding	75	8	2.04	.001	+8/−0
	GO:0008745	N-acetylmuramoyl-L-alanine amidase	3	2	0.08	.002	+1/−1
	GO:0005102	Receptor binding	41	6	1.11	.003	+3/−3
	GO:0005001	Transmembrane receptor protein tyrosine phosphatase	4	2	0.11	.004	+2/−0
	GO:0042834	Peptidoglycan binding	4	2	0.11	.004	+1/−1
	GO:0005372	Water transporter activity	5	2	0.14	.007	+1/−1
West vs. historical	GO:0043565	Sequence-specific DNA binding	75	18	9.2	.003	+12/−6
	GO:0003700	Transcription factor activity	144	29	17.66	.004	+21/−8
	GO:0004091	Carboxylesterase activity	26	9	3.19	.008	+3/−6
	GO:0020037	Heme binding	49	12	6.01	.013	+1/−11
East mild vs. severe	GO:0003700	Transcription factor activity	144	10	2.92	.001	+4/−6
	GO:0004558	Alpha-glucosidase activity	7	2	0.14	.008	+2/−0
Biological process							
East vs. historical	GO:0016348	Imaginal disc-derived leg joint morphogenesis	2	2	0.05	.001	+2/−0
	GO:0045449	Regulation of transcription	274	17	7.5	.001	+16/−1
	GO:0007415	Defasciculation of motor neuron axon	3	2	0.08	.002	+2/−0
	GO:0007509	Mesoderm migration	4	2	0.11	.004	+2/−0
	GO:0009253	Peptidoglycan catabolic process	4	2	0.11	.004	+1/−1
	GO:0035286	Leg segmentation	7	4	0.19	.007	+4/−0
West vs. historical	None						
East mild vs. severe	GO:0006541	Glutamine metabolic process	4	2	0.09	.003	+2/−0

GO terms significantly enriched ($P < 0.1$, Fisher's exact test) in the 3 gene lists are presented. The number of differentially expressed genes and the direction of change in expression relative to the more distressed second group in each comparison are listed for each GO term.

ated rRNA fragment abundance merits further exploration. In addition, the consequences of viral infection and CCD on the function of ribosomes should be explored through assays of translational efficiency. Because of the potential for translational interference, studies on immune suppression should focus on bioassays or protein abundance rather than on immune gene transcripts.

Our results also indicate an unappreciated variation in gene expression patterns and pathogen loads with geography. This variation provides insight into the different stresses facing bees and clearly demonstrates that diagnostic surveys must sample extensively across numerous bee populations. Colony surveillance via assay of rRNA-like transcript abundance may provide

Table 2. Comparison of fold-change expression differences as measured by microarray and qPCR

		Fold-change microarray			qPCR	
Transcript name	Entrez ID	W vs. H	E vs. H	S vs. M	Gut	Abdomen
Deformed wing virus	AJ489744.2	4.05	5.34	7.33	1.36*	3.00
Bee brain EST	BI515001	1.21	0.95	0.68	—	−1.21
mblk-1	NM_001011629.1	3.97	1.84	2.69	−1.51	—
Cytochrome P450 (CYP4G11)	NM_001040233.1	−4.01	−1.82	−7.44	−0.3	4.45
Peptidoglycan recognition prot. S3	XM_001123180.1	−1.94	−1.15	−2.47	−0.48	2.47
28S rRNA 401−469 (EST jdeC12)		4.67	3.58	3.94	0.72	—
5.8S rRNA 18−60 (EST jdeC15)		5.29	3.15	2.34	0.61*	—
28S rRNA 1838−1903 (EST QW33)	BG101565	2.92	2.80	3.99	1.92*	2.47*

Microarray: W vs. H, California 2006 CCD vs. historical; E vs. H, Florida and Pennsylvania CCD vs. historical; S vs. M, severe CCD vs mild CCD. QPCR: gut, California 2007 CCD vs. California 2007 healthy ($n = 60$); abdomen, pooled abdomen samples collected from healthy and CCD bees in Pennsylvania, Florida, and California and healthy bees collected in Massachusetts and Illinois during 2006–2008 ($n = 147$) (see Table S3). Transcripts showing similar trends in abundance in microarrays and qPCR samples ($P < .05$, one-tailed Mann-Whitney test) are indicated by * and using pooled abdomen samples collected from healthy and CCD bees in Pennsylvania, Florida, and California and healthy bees collected in Massachusetts and Illinois during 2006–2008 ($n = 147$).
*Significance ($P < .05$).

Table 3. Proportion of colonies with detectable levels of viral and microsporidial pathogens, as measured by qPCR

Year	Status	State	Colonies	ABPV	KBV	IAPV	DWV	SBV	Number of viruses	*N. apis*	*N. ceranae*	Number of Pathogens
2006	Healthy	MA/PA	14	0%	14%	7%	64%	36%	1.21 ± 0.80	0%	50%	1.71 ± 0.99
	CCD	FL	24	38%	38%	25%	46%	8%	1.54 ± 1.67	8%	42%	2.04 ± 2.03
	CCD	CA	57	51%	21%	21%	58%	30%	1.81 ± 1.41	40%	60%	2.81 ± 1.74
2007	Healthy	CA	14	86%	7%	14%	29%	7%	1.43 ± 1.02	7%	86%	2.36 ± 1.00
	CCD	CA	16	69%	44%	25%	44%	6%	1.88 ± 0.72	0%	94%	2.81 ± 0.83

Total number of different viruses and pathogens detected are summarized (± SD). More pathogens were detected in bees from colonies identified as suffering from CCD in both 2006 and 2007 ($P < .05$; one-way Mann-Whitney test), and CCD colonies overall contained more viruses ($P < .05$; Fisher's combined probability test).

an earlier indication of CCD status than has hitherto been available and allow beekeepers to take actions to reduce losses.

Materials and Methods

Biological Material and RNA Extraction. Remaining adult bees in colonies diagnosed with CCD at apiaries in Florida, California, and Pennsylvania were collected during winter 2006–2007. CCD colony health was scored at the time of collection as either severe or mild, depending on the apparent strength of the colony. Historical bees, collected before the appearance of CCD and hence ostensibly healthy, were collected in 2004 and 2005 from colonies set up on new equipment with no miticide treatments in apiaries of Pennsylvania State University near State College, Pennsylvania (Table S3). Bees for the reference (ostensibly healthy) sample were collected from 5 hives near Urbana, Illinois in July 2007. To date, there have been no confirmed cases of CCD in central Illinois. All bees were immediately frozen on dry ice and kept at −80 °C until dissection.

Dissection, RNA extraction, microarray, statistical and qPCR validation analyses are described in SI Text. Pooled samples of RNA from 6 guts collected from each CCD colony were hybridized to 2 arrays, incorporating a dye swap (Fig. S1). Two separate pools of RNA from 6 guts each were created for both historical samplings, and each was hybridized against the reference on a single microarray. Experiments were designed to meet Minimum Information About a Microarray Experiment standards, and all microarray data obtained in these studies were deposited at ArrayExpress [www.ebi.ac.uk/arrayexpress (accession no. E-MEXP-2292)].

ACKNOWLEDGMENTS. We thank D. vanEngelsdorp, J. Pettis, and D. Cox-Foster for sample collection and colony health scores, and D. Lopez and T. Newman for laboratory assistance. We also thank I. Lipkin, J. Willis, and J. Gillespie for an insightful and constructive review. This work was supported by U.S. Department of Agriculture Grants USDAG 2007–37610-18499 (to S. Ratcliffe, G.E.R., and M.B.), USDAG 2008–3532-18831 (to M.R.B.), and USDA-NRI 2004–35604-14277 (to G.E.R.).

1. National Academy of Sciences (2007) *Status of Pollinators in North America* (National Academies Press, Washington, DC).
2. vanEngelsdorp D, Underwood R, Caron D, Hayes J (2007) An estimate of managed colony losses in the winter of 2006–2007: A report commissioned by the apiary inspectors of America. *Am Bee J* 147:599–603.
3. Cox-Foster DL, et al. (2007) A metagenomic survey of microbes in honey bee colony collapse disorder. *Science* 318:283–287.
4. vanEngelsdorp D, Hayes J, Jr, Underwood RM, Pettis J (2008) A survey of honey bee colony losses in the U.S., fall 2007 to spring 2008. PLoS One 3:e4071.
5. The Honeybee Genome Sequencing Consortium (2006) Insights into social insects from the genome of the honeybee, *Apis mellifera*. *Nature* 443:931–949.
6. Chen YP, Evans JD (2007) Historical presence of Israeli acute paralysis virus in the United States. *Am Bee J* 147:1027–1028.
7. Higes M, et al. (2009) Honeybee colony collapse due to *Nosema ceranae* in professional apiaries. *Environ Microbiol Rep* 1:110–113.
8. Decourtye A, Devillers J, Cluzeau S, Charreton M, Pham-Delègue M (2004) Effects of imidacloprid and deltamethrin on associative learning in honeybees under semi-field and laboratory conditions. *Ecotoxicol Environ Saf* 57:410–419.
9. Johnson RM, Pollock HS, Berenbaum MR (2009) Synergistic interactions between in-hive miticides in *Apis mellifera*. *J Econ Entomol* 102:474–479.
10. Evans JD, Wheeler DE (2000) Expression profiles during honeybee caste determination. *Genome Biol* 2:research0001.1–0001.6.
11. Whitfield CW, et al. (2002) Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee. *Genome Res* 12:555–566.
12. Ashburner M, et al. (2000) Gene Ontology: Tool for the unification of biology. *Nat Genet* 25:25–29.
13. Werner T, Borge-Renberg K, Mellroth P, Steiner H, Hultmark D (2003) Functional diversity of the *Drosophila PGRP-LC* gene cluster in the response to lipopolysaccharide and peptidoglycan. *J Biol Chem* 278:26319–26322.
14. Burkart C, et al. (2007) Modular proteins from the *Drosophila sallimus* (*sls*) gene and their expression in muscles with different extensibility. *J Mol Biol* 367:953–969.
15. Garver LS, Xi Z, Dimopoulos G (2008) Immunoglobulin superfamily members play an important role in the mosquito immune system. *Dev Comp Immunol* 32:519–531.
16. Lee T, Marticke S, Sung C, Robinow S, Luo L (2000) Cell-autonomous requirement of the USP/EcR-B ecdysone receptor for mushroom body neuronal remodeling in *Drosophila*. *Neuron* 28:807–818.
17. Buszczak M, Segraves WA (2000) Insect metamorphosis: Out with the old, in with the new. *Curr Biol* 10:830–833.
18. Lanzi G, et al. (2006) Molecular and biological characterization of deformed wing virus of honeybees (*Apis mellifera* L.). *J Virol* 80:4998–5009.
19. Maori E, et al. (2007) Isolation and characterization of Israeli acute paralysis virus, a dicistrovirus affecting honeybees in Israel: Evidence for diversity due to intra- and inter-species recombination. *J Gen Virol* 88:3428–3438.
20. Feyereisen R (2005) Insect cytochrome P450. In *Comprehensive Molecular Insect Science*, eds Iatrou K, Gilbert L, Gill S (Elsevier Pergamon), Vol 4, pp 1–77.
21. Maïbèche-Coisne M, Monti-Dedieu L, Aragon S, Dauphin-Villemant C (2000) A new cytochrome P450 from *Drosophila melanogaster*, CYP4G15, expressed in the nervous system. *Biochem Biophys Res Comm* 273:1132–1137.
22. Claudianos C, et al. (2006) A deficit of detoxification enzymes: Pesticide sensitivity and environmental response in the honeybee. *Insect Mol Biol* 15:615–636.
23. Li W, Ma G, Zhou X (2006) Apidaecin-type peptides: Biodiversity, structure–function relationships and mode of action. *Peptides* 27:2350–2359.
24. Sabatier L, et al. (2003) Pherokine-2 and −3. *Eur J Biochem* 270:3398–3407.
25. Roxström-Lindquist K, Terenius O, Faye I (2004) Parasite-specific immune response in adult *Drosophila melanogaster*: A genomic study. *EMBO Rep* 5:207–212.
26. Gillespie JJ, Johnston JS, Cannone JJ, Gutell RR (2006) Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: Hymenoptera): Structure, organization, and retrotransposable elements. *Insect Mol Biol* 15:657–686.
27. Gonzalez IL, Sylvester JE (1997) Incognito rRNA and rDNA in databases and libraries. *Genome Res* 7:65–70.
28. Scharf ME, Wu-Scharf D, Zhou X, Pittendrigh BR, Bennett GW (2005) Gene expression profiles among immature and adult reproductive castes of the termite *Reticulitermes flavipes*. *Insect Mol Biol* 14:31–44.
29. Slomovic S, Portnoy V, Liveanu V, Schuster G (2006) RNA polyadenylation in prokaryotes and organelles: Different tails tell different tales. *Crit Rev Plant Sci* 25:65–77.
30. Slomovic S, Laufer D, Geiger D, Schuster G (2006) Polyadenylation of ribosomal RNA in human cells. *Nucleic Acids Res* 34:2966–2975.
31. Win TZ, et al. (2006) Requirement of fission yeast Cid14 in polyadenylation of rRNAs. *Mol Cell Biol* 26:1710–1721.
32. Mauro VP, Edelman GM (2002) The ribosome filter hypothesis. *Proc Natl Acad Sci USA* 99:12031–12036.
33. Tranque P, Hu MC, Edelman GM, Mauro VP (1998) rRNA complementarity within mRNAs: A possible basis for mRNA–ribosome interactions and translational control. *Proc Natl Acad Sci USA* 95:12238–12243.
34. Chang Y, Stockinger MP, Tashiro H, Lin CG (2008) A novel noncoding RNA rescues mutant SOD1-mediated cell death. *FASEB J* 22:691–702.
35. Kong Q, Stockinger MP, Chang Y, Tashiro H, Lin CG (2008) The presence of rRNA sequences in polyadenylated RNA and its potential functions. *Biotech J* 3:1041–1046.
36. Ongus JR, Roode EC, Pleij CWA, Vlak JM, van Oers MM (2006) The 5′ non-translated region of *Varroa destructor* virus 1 (genus *Iflavirus*): Structure prediction and IRES activity in *Lymantria dispar* cells. *J Gen Virol* 87:3397–3407.
37. Joachims M, Van Breugel PC, Lloyd RE (1999) Cleavage of poly(A)-binding protein by enterovirus proteases concurrent with inhibition of translation in vitro. *J Virol* 73:718–727.
38. Hellen CT, Sarnow P (2001) Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev* 15:1593–1612.
39. Zundel MA, Basturea GN, Deutscher MP (2009) Initiation of ribosome degradation during starvation in *Escherichia coli*. *RNA* 15:977–983.

Supporting Information

Johnson et al. 10.1073/pnas.0906970106

SI Materials and Methods

Biological Material and RNA Extraction. Remaining adult bees in colonies diagnosed with CCD at apiaries in Florida, California, and Pennsylvania were collected during the winter of 2006–2007. CCD colony health was scored at the time of collection as either "severe" or "mild," depending on the apparent strength of the colony. "Historical" bees, collected before the appearance of CCD and hence ostensibly healthy, were collected in 2004 and 2005 from colonies set up on new equipment with no miticide treatments in apiaries of Pennsylvania State University near State College, PA (Table S3). All bees were immediately frozen on dry ice and kept at −80 °C until dissection.

To dissect guts, each abdomen was placed in 500 μL of cold RNAlater ICE (Ambion) and kept at −20 °C for 36–48 h. Guts were rapidly dissected on a chilled glass plate and returned to −80 °C until RNA extraction. Gut RNA was isolated by first grinding in TRIzol (Invitrogen) and then purifying over an RNeasy minicolumn (Ambion). RNA isolated from individual guts was equally pooled among 5–7 guts and then used for reverse transcription and hybridization of microarrays.

Bees for the "reference" (ostensibly healthy) sample were collected from 5 hives near Urbana, IL in July 2007. To date, there have been no confirmed cases of CCD in central Illinois. Guts were removed from 15 cold-anesthetized live bees from each colony by pulling the last abdominal segment with forceps and then freezing the gut on dry ice. Guts were extracted using TRIzol and RNeasy Midi columns (Ambion) and then pooled to make the reference sample.

For the comparison of abdomen RNA using qPCR, abdomens were detached from frozen bees. Eight abdomens from the same colony were ground in liquid nitrogen, and RNA was isolated using TRIzol.

Microarray Design. Oligonucleotide-based microarrays (UIUC Honey Bee oligo 13K v1) were fabricated at the University of Illinois Keck Center for Comparative and Functional Genomics and represent 9,867 gene predictions for the honey bee genome (1) as well as 2,729 probes specific for transcripts discovered through EST projects on bee brain (2) and larvae (3) that match no existing genes. In addition, 22 probes specific for honey bee pathogens were included. Seventy-mer oligonucleotides representing each sequence were designed using an established algorithm (4), synthesized by Illumina, Inc, and double-spotted on glass microscope slides along with appropriate positive and negative control spots. A total of 13,440 probes specific for RNA transcripts were present on the array.

Experimental Design. Experiments were designed to meet Minimum Information About a Microarray Experiment standards, and all microarray data obtained in these studies were deposited at ArrayExpress [www.ebi.ac.uk/arrayexpress (accession no. E-MEXP-2292)]. A combination looped, common-reference microarray design was used to compare historical and CCD samples. A blend of RNA isolated from healthy colonies (collected near Urbana, IL in July 2007) served as a reference (5). The microarray experiment compared the guts of bees from mildly and severely afflicted colonies in apiaries experiencing CCD on the east coast (Florida and Pennsylvania) and west coast (California) with a common reference (Fig. S1). Pooled samples of RNA from 6 guts collected from each CCD colony were hybridized to 2 arrays, incorporating a dye swap. Two separate pools of RNA from 6 guts each were made for both historical samplings, and each was hybridized against the reference on a single microarray.

Microarray Hybridizations. Microarrays were hybridized, washed, and scanned following established protocols (6). RNA (5 μg) was hybridized to 6 μg of oligo(dT)$_{18}$ primer. Then single-stranded cDNA was synthesized with 400 U ArrayScript (Ambion) combined with chilled 10× first-strand buffer, 20 U Rnase OUT (Invitrogen), and 0.5 mM amino-allyl dNTP. After incubation overnight at 42 °C, the reaction was stopped by incubating with 15 μL of 0.1 N NaOH at 70 °C and then neutralized with 15 μL of 0.1 N HCl. The Qiaquick PCR purification kit (Qiagen), with modified Tris-free buffers, was used to purify cDNA, which was then dried down in a SpeedVac, thoroughly resuspended in 4.5 μL of coupling buffer, and then dye-coupled with Cy dye (40 nmol, RPN5661; GE Healthcare) in 4.5 μL of DMSO during a 1-h incubation. Dye-coupled cDNA was purified separately with a PCR purification kit; samples were then combined with 1× hybridization buffer (25% formamide, 5× SSC, 0.1% SDS), applied to arrays, and hybridized for 24 h at 42 °C. A series of 10-min agitated washes removed excess probe: 1× SSC and 0.2% SDS at 42 °C, 0.1× SSC at room temperature, and 2 washes with 0.1× SSC and 0.2% SDS at room temperature. Arrays were spin-dried and stored under argon until being scanned with an Axon 4000B scanner. Images were analyzed with GenePix Pro (Molecular Devices).

Statistical Analyses. The LIMMA/Bioconductor/R statistical package was used for statistical analysis of the intensity data from the arrays (7–9). NORMEXP was used for background correction, followed by PRINTTIPLOESS correction within arrays and SCALE correction between arrays. DUPCOR from the LIMMA package was used to estimate the correlation between duplicate spots on the arrays. After normalization, data from negative and positive control probes were removed from further analysis, as were spots with an expression level below the level of negative control spots on at least 5 of the 22 arrays. Corrected intensity values for each colony were fit to a linear model and ranked in order of evidence for differential expression using EBAYES (9) on 3 contrast matrices: east coast versus historical, west coast versus historical and east coast severe versus mild, and west coast severe versus mild and west coast versus east coast (Fig. 3). A P value $<.01$ after FDR correction and a fold-change >2 was established as the cutoff for genes to be considered differentially expressed. A relatively permissive P value combined with a fold-change cutoff has been shown to produce more reliably reproducible gene lists in microarray experiments (10).

Hierarchical Cluster Analysis and Heatmaps. An experimentwide hierarchical cluster analysis was performed using expression data for 6,777 probes that met the filtering criteria. Euclidean distance calculation was followed by complete clustering, and was supported with 1,000 bootstrap replicates (11). This experimentwide clustering of colonies was used for all dendrograms, and these same clustering methods were used for genewise clustering and the production of heatmaps.

GO Analysis. Bee genes with putative *D. melanogaster* orthologs (12) were annotated based on *D. melanogaster* GO classifications (version 08/27/2008) from the GO website (http://www.geneontology.org/). GO classifications specific to honey bee genes from

Uniprot (http://www.uniprot.org/) were included as well. When more than one probe on the microarray was specific to a particular gene, the expression data corresponding to the most 3′ probe were chosen. A BioPerl script (13) was used to automate bioinformatic tasks. A total of 3,512 probes were associated with GO annotation and served as the background for GO analyses. Inside the R environment, TOPGO (14) was used to test for GO term enrichment in the list of most differentially expressed genes using Fisher's exact test.

qPCR Methods. qPCR procedures followed established protocols (15). Synthesis of cDNA began with hybridization of oligo(dT)$_{12–18}$ and oligo(dT)$_{20}$ to 1 μg of RNA, with the addition of DNase. Shorter oligo(dT) primers were included to increase the sensitivity to pathogen transcripts. Reverse transcripts were synthesized with SuperScript III (Invitrogen). Primer pairs were designed to produce 120- to 180-bp products and were designed using Primer 3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) (Table S2). PCR reactions were quantified using SYBR-Green I dye (Applied Biosystems) and 40 cycles of a thermal profile consisting of 94 °C for 20 s, 60 °C for 30 s, 72 °C for 1 min, and 78 °C for 20 s, using a Bio-Rad I cycler. Fluorescence was measured multiple times during the 78 °C step, and expression was normalized against the geometric mean CT value for 3 honey bee housekeeping genes (*RPS5*, *MGsT*, and *actin*).

Ribosomal RNA BLAST Methods. Nuclear and mitochondrial rRNA sequences (16) were divided into overlapping 70-nt segments, with a 5-nt difference between each. BLASTN (17), with default settings, was used to compare these rRNA segments with all predicted honey bee genes (5) and all microarray probes. Probes with at least 40 nt matching rRNA were labeled in analyses as rRNA (Table S1).

1. The Honeybee Genome Sequencing Consortium (2006) Insights into social insects from the genome of the honeybee, *Apis mellifera*. *Nature* 443:931–949.
2. Evans JD, Wheeler DE (2000) Expression profiles during honeybee caste determination. *Genome Biol* 2:research0001.1–0001.6.
3. Whitfield CW, et al. (2002) Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee. *Genome Res* 12:555–566.
4. Bozdech Z, et al. (2003) Expression profiling of the schizont and trophozoite stages of *Plasmodium falciparum* with a long-oligonucleotide microarray. *Genome Biol* 4:R9.
5. Kerr KF, Serikawa KA, Wei C, Peters MA, Bumgarner RE (2007) What is the best reference RNA? And other questions regarding the design and analysis of two-color microarray experiments. *OMICS* 11:152–165.
6. Alaux C, et al. (2009) Regulation of brain gene expression in honey bees by brood pheromone. *Genes Brain Behav* 8:309–319.
7. R Development Core Team (2008) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria).
8. Gentleman R, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5:R80.
9. Smyth GK (2005) Limma: Linear models for microarray data. In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor R*. eds Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W (Springer, New York), pp 397–420.
10. Shi L, et al. (2008) The balance of reproducibility, sensitivity, and specificity of lists of differentially expressed genes in microarray studies. *BMC Bioinform* 9:S10.
11. Suzuki R, Shimodaira H (2006) Pvclust: An R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* 22:1540–1542.
12. Elsik CG, et al. (2007) Creating a honey bee consensus gene set. *Genome Biol* 8:R13.
13. Stajich JE, et al. (2002) The BioPerl toolkit: Perl modules for the life sciences. *Genome Res* 12:1611–1618.
14. Alexa A, Rahnenführer J, Lengauer T (2006) Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* 22:1600–1607.
15. Evans JD (2006) Beepath: An ordered quantitative-PCR array for exploring honey bee immunity and disease. *J Invert Pathol* 93:135–139.
16. Gillespie JJ, Johnston JS, Cannone JJ, Gutell RR (2006) Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: Hymenoptera): Structure, organization, and retrotransposable elements. *Insect Mol Biol* 15:657–686.
17. Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402.

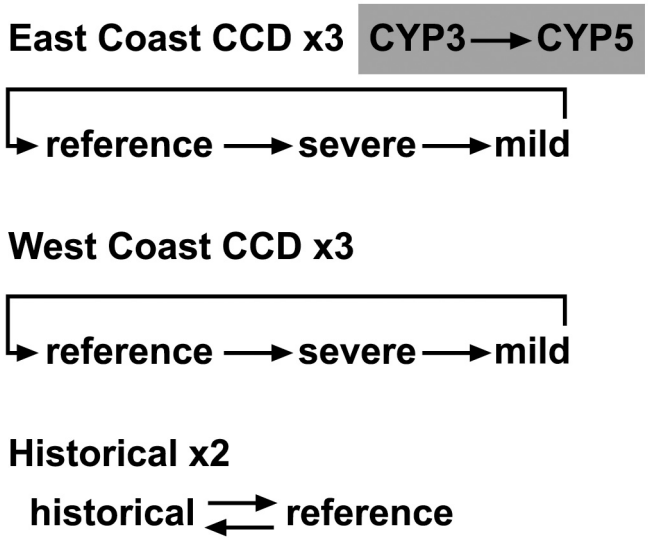

Fig. S1. Summarized experimental design for microarray comparisons of mild CCD, severe CCD and historical gut RNA samples. The east coast loop was repeated 3 times with pooled gut RNA isolated from guts of 5–7 bees collected at 2 mild and 2 severe CCD colonies in Florida and 1 mild and 1 severe colony in Pennsylvania in late 2006. The west coast loop was repeated 3 times, with bees collected at 3 mild and 3 severe CCD colonies in California in late 2006. The historical comparison was repeated twice using bees collected from healthy Pennsylvania colonies in fall 2004 and 2005. All indirect comparisons were made using a common reference gut RNA sample isolated from healthy Illinois bees collected in summer 2007.

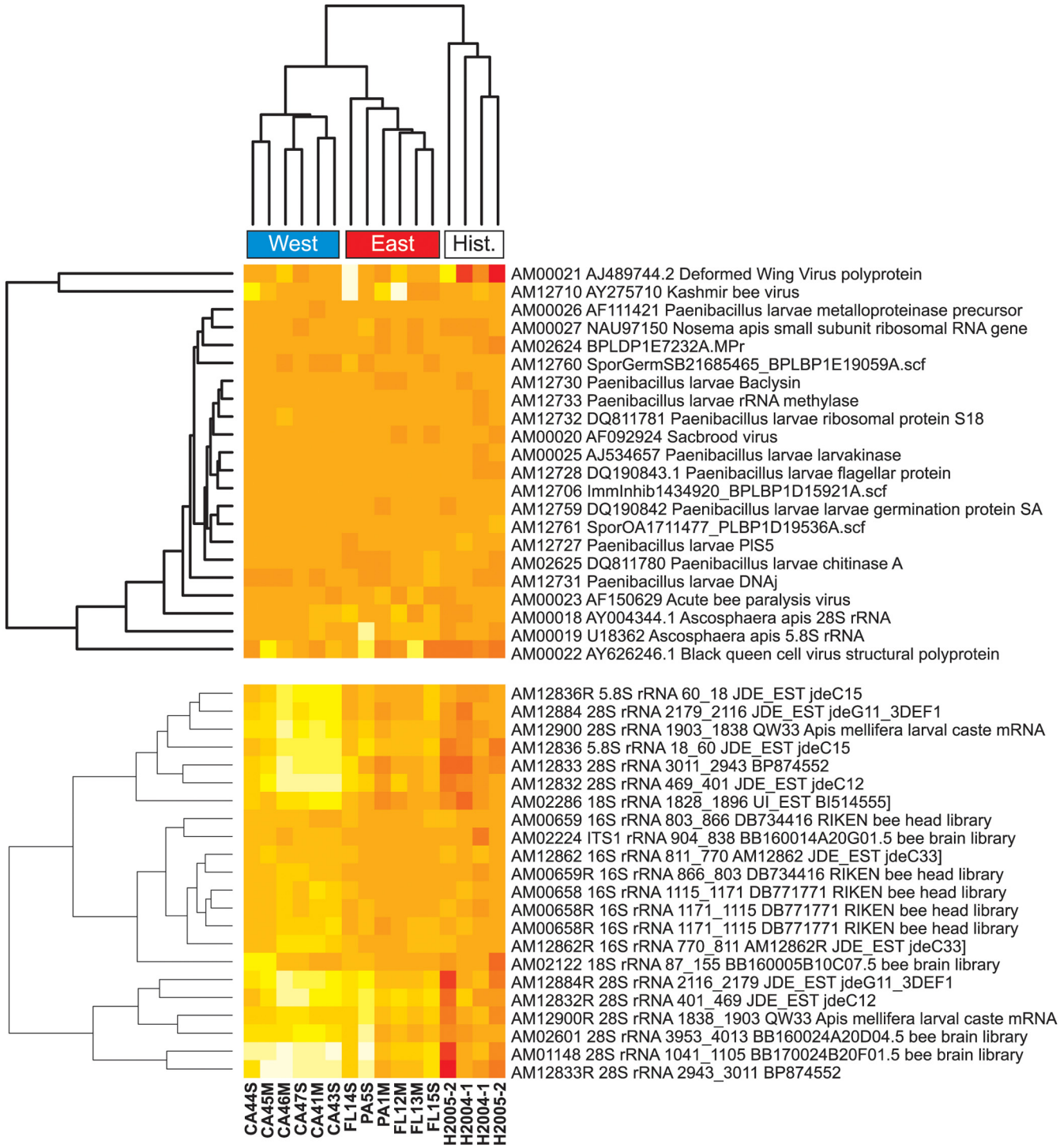

Fig. S2. Heatmap comparing microarray expression values for all forward probes specific for pathogens and all probes specific for rRNA. ''AM'' designations correspond to microarray probe number. Yellow indicates increased transcript abundance, and red indicates decreased transcript abundance relative to the mean for all colonies. CA, California; PA, Pennsylvania; FL, Florida. Numbers correspond to colony number. S, severe CCD; M, mild CCD. Colonywise clustering was performed using expression values for all 6,777 probes.

Table S1. Top BLASTN hits between rRNA (16) and annotated genes from the honey bee genome (A) and microarray probes (B)

(A)

Annotated gene	Start	End	Orientation	rRNA match	Match length	E-value	Identity
GB16311-RA	300	324	+	18S rRNA 291–360	25	5.10E-002	92%
GB14230-RA	11441	11469	+	18S rRNA 1011–1080	29	5.10E-002	90%
GB17239-RA	429	457	+	ITS1 rRNA 81–150	29	5.10E-002	90%
GB12491-RA	929	955	+	ITS2 rRNA 1276–1345	27	7.90E-001	89%
GB13774-RA	242	266	−	12S rRNA 1–70	25	5.10E-002	92%
GB18978-RA	1202	1228	−	12S rRNA 571–640	26	7.90E-001	96%
GB30293-RA	1706	1730	−	16S rRNA 591–660	25	5.10E-002	92%
GB17261-RA	3993	4019	+	16S rRNA 916–985	27	7.90E-001	89%
GB12764-RB	1897	1923	+	16S rRNA 1011–1080	26	7.90E-001	96%
GB10944-RA	769	797	+	16S rRNA 1341–1371	29	1.60E-002	90%
GB16604-RA	186	214	−	16S rRNA 1341–1371	29	1.60E-002	90%
GB12296-RA	473	501	+	16S rRNA 1341–1371	30	4.00E-003	93%

(B)

Probe	Start	End	Orientation	rRNA match	Match length	E-value	Identity	EST
AM02286	3	70	+	18S rRNA 1826–1895	68	8.00E-034	100%	UI_EST BI514555
AM02122	2	70	+	18S rRNA 86–155	69	1.00E-029	97%	BB160005B10C07.5 bee brain library
AM02224	3	69	−	ITS1 rRNA 836–905	67	3.00E-033	100%	BB160014A20G01.5 bee brain library
AM12836	3	45	+	5.8S rRNA 16–85	43	7.00E-019	100%	JDE_EST jdeC15
AM12836R	13	55	−	5.8S rRNA 6–75	43	7.00E-019	100%	JDE_EST jdeC15
AM12832	1	69	−	28S rRNA 401–470	69	2.00E-034	100%	JDE_EST jdeC12
AM12832R	1	69	+	28S rRNA 401–470	69	2.00E-034	100%	JDE_EST jdeC12
AM01148	1	65	+	28S rRNA 1041–1110	65	1.00E-029	98%	BB170024B20F01.5 bee brain library
AM12900	3	68	−	28S rRNA 1836–1905	66	1.00E-032	100%	QW33 *A. mellifera* larval caste mRNA
AM12900R	3	68	+	28S rRNA 1836–1905	66	1.00E-032	100%	QW33 *A. mellifera* larval caste mRNA
AM12884	1	64	−	28S rRNA 2116–2185	64	1.00E-014	91%	JDE_EST jdeG11_3DEF1
AM12884R	6	69	+	28S rRNA 2111–2180	64	1.00E-014	91%	JDE_EST jdeG11_3DEF1
AM12833R	3	70	+	28S rRNA 2941–3010	68	8.00E-034	100%	BP874552
AM12833	3	70	−	28S rRNA 2941–3010	68	8.00E-034	100%	BP874552
AM02601	8	68	+	28S rRNA 3946–4015	61	1.00E-029	100%	BB160024A20D04.5 bee brain library
AM12862	20	61	−	16S rRNA 751–820	42	6.00E-016	98%	JDE_EST jdeC33
AM12862R	25	66	+	16S rRNA 746–815	42	6.00E-016	98%	JDE_EST jdeC33
AM00659	3	70	+	16S rRNA 801–870	68	8.00E-034	100%	DB734416 RIKEN bee head library
AM00659R	3	70	−	16S rRNA 801–870	68	8.00E-034	100%	DB734416 RIKEN bee head library
AM00658	3	70	+	16S rRNA 1101–1170	68	8.00E-034	100%	DB771771 RIKEN bee head library
AM00658R	3	70	−	16S rRNA 1101–1170	68	8.00E-034	100%	DB771771 RIKEN bee head library

Probe EST sources include JDE_EST and QW33 (2), BB and UI_IST (3), and DB (1).

Table S2. Primer sequences used in qPCR experiments: (A) annotated genes, (B) ESTs, and (C) pathogens

Microarray			qPCR primer	
Probe	Description	Entrez ID	Forward	Reverse
(A)				
AM02500	Transcription factor mblk-1 (Mblk-1)	NM_001011629	ACGTTCGTGCTTTTCTTCGT	TGCCACTTCCTTCGTTTCTC
AM04590	Cytochrome P450 monooxygenase (Cyp4 g11)	NM_001040233	CAAAATGGTGTTCTCCTTACCG	ATGGCAACCCATCACTGC
AM05407		GB12792	GGCCGAACTCCTGACCAC	TAATAACAATCTGATTTGCGTGTG
AM07887	Similar to CG6619-PA	XM_624420	GAACCAAAATTTTTACGCGAAT	GCCTCTGCTTGTTTTGCAGT
AM10450	Peptidoglycan recognition protein (PGRP-S3)	XM_001123180	TTTGAAAATTTCCTATGAAAGCA	TTTTTAATTGGTGGAGATGGAAA
(B)				
AM01773	UI_EST BI510059	NW_001253565	TCCAACGATGATTTCTGACG	TCCTGCAGATTTTTCGGGTA
AM02286	UI_EST BI514555, 18S subunit rRNA		CGTCCCTGCCCTTTGTACAC	CCAGTCCGAAGACCTCACTAAATC
AM02446	BB160017A10B06.5 bee brain EST	BI515001	CGCGCGAGTAAGAGAAAGAG	TCGAACAAGGGAAACGAAAC
AM12832R	JDE_EST jdeC12, 28S subunit rRNA		AGGAATGAGCCCAGCACTGA	ACTTGGACGCAGTTCGACATC
AM12836	JDE_EST jdeC15, 5.8S subunit rRNA		GCGCGTCAACGTGTGAACT	CCGTGGACCGCAATGTG
AM12900R	QW33 larval caste EST, 28S subunit rRNA	BG101565	TTAAGCAACCAACGCCTTTC	GGATCATGAAGCCACGAGAT
(C)				
AM00021	Deformed wing virus polyprotein gene	AJ489744	GAGATTGAAGCGCATGAACA	TGAATTCAGTGTCGCCCATA
AM00023	Acute bee paralysis virus	AF150629	ACCGACAAAGGGTATGATGC	CTTGAGTTTGCGGTGTTCCT
AM12710	Kashmir bee virus	AY275710	TGAACGTCGACCTATTGAAAAA	TCGATTTTCCATCAAATGAGC
AM00020	Sacbrood virus	AF092924	GGGTCGAGTGGTACTGGAAA	ACACAACACTCGTGGGTGAC
	IAPV-F1a		GCGGAGAATATAAGGCTCAG	CTTGCAAGATAAGAAAGGGGG
	IAPV_B4SO427_L17		CGAACTTGGTGACTTGAAGG	GCATCAGTCGTCTTCCAGGT
	IAPVpwF16		ACCCCCAACTGCTTTCAACAG	CTGGATATAGTACATTAATGTCCTGC
	Nosema ceranae	NCU26533	CAATATTTTATTATTTTGAGAGA	TATATTTATTGTATTGCGCGTGCA
AM00027	Nosema apis	NAU97150	CAATATTTTATTGTTCTGCGAGG	TATATTTATTGTATTGCGCGTGCT

Table S3. Collection data for bees included in microarray and qPCR comparisons

Experiment	Status	State	Year	Apiaries	Colonies	Individuals/colony
Microarray gut	Healthy	Pennsylvania	2004 & 2005	1	2	12
	CCD	Florida	2006	1	4	6
	CCD	Pennsylvania	2006	1	2	6
	CCD	California	2006	1	6	6
qPCR gut	Healthy	California	2007	1	16	2
	CCD	California	2007	1	14	2
qPCR abdomen	Healthy	Massachusetts	2006	1	11	8
	Healthy	Pennsylvania	2004 & 2005	1	2	8
	Healthy	Illinois	2007	4	12	8
	Healthy	California	2007	1	14	8
	CCD	California	2006	9	69	8
	CCD	Florida	2006	3	32	8
	CCD	California	2007	1	16	8

Gut collections for microarray analysis from 2006 included equal numbers of colonies categorized as severe or mild CCD: 2 severe and 2 mild in Florida, 1 severe and 1 mild in Pennsyvania, and 3 severe and 3 mild in California.

Table S4. Technical validation of microarray results with qPCR

ID	Reporter name	Entrez	Rho	*P* value
AM01773	UI_EST BI510059	NW_001253565.1	0.79	<0.01*
AM02446	BB160017A10B06.5 bee brain EST	BI515001	0.19	0.27, NS
AM02500	Transcription factor mblk-1	NM_001011629.1	0.38	0.1, NS
AM04590	Cytochrome P450 (CYP4G11)	NM_001040233.1	0.65	0.01*
AM05407		GB12792-RA	0.76	<0.01*
AM07887	Similar to CG6619-PA	XM_624420.2	−0.6	0.98, NS
AM10450	Peptidoglycan recognition protein S3	XM_001123180.1	0.69	0.01*
AM12900R	QW33 larval caste EST, 28S subunit rRNA	BG101565	0.69	0.01*

Technical validation of transcripts found differentially expressed on microarray using same mRNA samples for array and qPCR. NS, not significant.
*Significant correlation ($P < .05$; Spearman's rank correlation test; $n = 13$).

Table S5. Supplemental data for probes appearing in Fig. 2.

Heatmap label	Probe no.	Detailed description
	AM02591	BB160024A10D03.5 Bee Brain Library
	AM02425	BB160019B20A08.5 Bee Brain Library
Kuzbanian-like	AM11733	GB19179-RA similar to Kuzbanian-like CG1964-PA
	AM01939	BB160007A10F09.5 Bee Brain Library
CG30157-PA	AM06238	GB13640-RA similar to CG30157-PA
	AM02330	BB160016B20F04.5 Bee Brain Library
	AM02241	BB160014B10C02.5 Bee Brain Library
Neuromusculin	AM06930	GB14330-RA similar to neuromusculin CG8779-PA
	AM02567	NM_001011629.1 Mblk-1 UI_EST BI516759
GB17441-RA	AM10013	GB17441-RA
	AM02574	BB160003B10F04.5 Bee Brain Library
	AM02324	BB160016B10E04.5 Bee Brain Library
	AM02046	BB160009B20F08.5 Bee Brain Library
	AM01778	BB170022A10A06.5 Bee Brain Library
	AM01773	NW_001253565.1 UI_EST BI510059
	AM12776	NW_001254336.1
	AM02577	NW_001253018.1 UI_EST BI516848
GB16434-RA	AM09013	GB16434-RA
	AM02412	EST082 female antennae library
CG4019-PA	AM07977	GB15389-RA similar to CG4019-PA, isoform A
GB14952-RA	AM07543	GB14952-RA
	AM00430	DB780023 RIKEN honey bee head library
Carboxypeptidase B precursor	AM06256	GB13658-RA similar to Carboxypeptidase B precursor
GB30251-RA	AM00225	GB30251-RA
CG7720-PB	AM04230	GB11610-RA similar to CG7720-PB, isoform B
bc10	AM10352	GB17779-RA similar to bc10 CG4867-PA
Rugose	AM08593	GB16012-RA similar to rugose CG6775-PA, isoform A
Inhibitor of growth fam., mem. 4	AM09163	GB16584-RA similar to inhibitor of growth family, member 4
CG14964-PA	AM11696	GB19142-RA similar to CG14964-PA
Kruppel-like protein 1	AM07461	GB14867-RA kruppel-like protein 1 (Kr-h1)
	AM01978	BB160008A10C02.5 Bee Brain Library
microRNA mir-277	AM00065	NW_001253371.1 microRNA mir-277
	AM02281	BB160015B10D03.5 Bee Brain Library
	AM01686	NW_001253388.1 UI_EST BI509259
CG6015-PA	AM03671	GB11047-RA similar to CG6015-PA
	AM02500	NM_001011629.1 Mblk-1 UI_EST BI516137
	AM02017	BB160008B20E05.5 Bee Brain Library
mblk-1	AM01826	GB17328-RA Mblk-1
	AM00552	DB762728 RIKEN honey bee head library
	AM02463	BB160020B20E04.5 Bee Brain EST Library
	AM02099	NM_001011629.1 Mblk-1 UI_EST BI513090
	AM01315	BB170022A10A10.5 Bee Brain Library
	AM01284	BB170017B10G02.5 Bee Brain Library
Alpha-glucosidase (Hbg2)	AM12618	GB20070-RA alpha-glucosidase (Hbg2)
Sallimus	AM01016	XM_001121572.1 similar to sallimus CG1915-PC, isoform C
CG6619-PA	AM07887	GB15297-RA similar to CG6619-PA
Deformed Wing Virus	AM00021	AJ489744.2 Deformed Wing Virus gene for polyprotein
EST 28S 2116–2179	AM12884R	28S rRNA 2116–2179 JDE_EST jdeG11_3DEF1
EST 28S 401–469	AM12832R	28S rRNA 401–469 JDE_EST jdeC12
EST 28S 1838–1903	AM12900R	28S rRNA 1838–1903 QW33 larval caste EST
EST 5.8S 60–18	AM12836R	5.8S rRNA 60–18 JDE_EST jdeC15
EST 5.8S 18–60	AM12836	5.8S rRNA 18–60 JDE_EST jdeC15
EST 28S 3011–2943	AM12833	28S rRNA 3011–2943 BP874552 adult brain library
Cytochrome P450 (CYP4G11)	AM04590	GB11973-RA cytochrome P450 monooxygenase (CYP4G11)
LOC725309	AM02502	XM_001120351.1 hypothetical protein LOC725309
GB18472-RA	AM11034	GB18472-RA
GB17959-RA	AM10532	GB17959-RA
Larval serum protein 2	AM12781	GB30362-RA similar to Larval serum protein 2 CG6806-PA
Dihydroxyacetone kinase 2	AM05667	GB13058-RA similar to dihydroxyacetone kinase 2
Peptidoglycan recognition prot. S3	AM10450	GB17879-RA Peptidoglycan recognition protein S3
CG1220-PE	AM06214	GB13615-RA similar to CG1220-PE, isoform E
GB19297-RA	AM11850	GB19297-RA hypothetical protein LOC725074
CG6416-PF	AM02897	GB30104-RA similar to CG6416-PF, isoform F
Carboxylesterase GB10854-RA	AM03479	GB10854-RA
CG4409-PA	AM05331	GB12716-RA similar to CG4409-PA