Comparative Brain Transcriptomic Analyses of Scouting Across Distinct Behavioural and Ecological Contexts in Honey Bees

(Supplementary material)

This supplementary material includes:

Additional methods with references

Figure S1- S5

Table S1, S3, S4, S5 *

*Table S2 is an Excel spreadsheets (Table S2_GO enrichment data.xlsx) uploaded as a separately file.

Additional Methods

(a) Brain dissection and RNA extraction

Bee heads were freeze-dried at -80°C, and then whole brains were dissected out in 100% ethanol on dry ice and stored at -80°C [1]. RNA was isolated from whole brains using TRIzol® Reagent (Invitrogen, CA) and RNeasy® Mini Kit (Qiagen Sciences, MD). RNA extraction was carried out as per manufacturer's instructions for total RNA and for on-column DNase I treatment (Qiagen). RNA yields and purity were determined with a NanoDrop® ND-1000 UV-Vis Spectrophotometer (Thermo Fisher Scientific, DE).

(b) Transcriptomic analysis

Brain dissection and RNA extraction was performed according to standard methods (detailed description in the ESM). We used a second-generation honey bee whole genome microarray, which was primarily based on the Honey Bee Genome Sequencing Consortium Official Gene Set 2.0 [2] and was previously characterized [3]. The microarray has 13,614 probes, henceforth referred to as genes. A total of 120 bees were analyzed in the study. For food scouting, 30 scouts and 30 recruits were used from three colonies, with 10 bees per behavioural group per colony. For nest scouting, 30 scouts and 30 recruits were used from two colonies, with 15 bees per behavioural group per colony. The bees were compared in a loop design [4, 5] with a total of 138 microarrays (Fig. S2). RNA was extracted from bee brains, and 500 ng per brain were amplified using amino-allyl MessageAmpII kits (Ambion/Applied Biosystems, TX). Of this amplified aRNA, 15 ug were conjugated with Cy3 or Cy5 mono NHS ester (GE Healthcare). From this conjugated aRNA, 3ug of Cy3 and 3ug of Cy5 labeled aRNA were combined and fragmented. The labeled aRNA was loaded on custom-printed microarray slides and hybridized overnight using Agilent coverslips (Agilent, CA) in a rotating hybridization oven at 42°C. Hybridized arrays were scanned on an Axon 4000B scanner and the images analyzed using GENEPIX 6.1 software (Agilent, CA). Images were manually edited for aberrant spots.

Fluorescence intensity spots were filtered if flagged by the scanning software or if the intensity did not surpass the median intensity of the negative control spots for each channel. A Loess transformation was used to adjust the log2 transformed gene expression intensities for dye effects within each microarray, and duplicated spots within the microarray elements were averaged. Gene intensities within each microarray were centralized to allow analysis across microarrays, and microarray elements with more than one missing observation were removed from the analysis.

References:

1 Schulz, D. J. & Robinson, G. E. 1999 Biogenic amines and division of labor in honey bee colonies: behaviorally related changes in the antennal lobes and age-related changes in the mushroom bodies. *J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol.* **184**, 481–488.

2 Honey bee Genome Sequencing Consortium 2006 Insights into social insects from the genome of the honey bee *Apis mellifera*. *Nature* **443**, 931–949. (doi: 10.1038/nature05260)

3 Alaux, C., Le Conte, Y., Adams, H. A., Rodriguez-Zas, S., Grozinger, C. M., Sinha, S. & Robinson, G. E. 2009 Regulation of brain gene expression in honey bees by brood pheromone. *Genes Brain Behav* **8**, 309–319. (doi:10.1111/j.1601-183X.2009.00480.x)

4 Kerr M.K., Churchill G.A. 2001 Experimental design for gene expression microarrays. *Biostatistics* **2**, 183–201. (doi: 10.1093/biostatistics/2.2.183.)

5 Whitfield C.W., Cziko A.M., Robinson G.E. 2003 Gene expression profiles in the brain predict behavior in individual honey bees. *Science* **302**, 296-299. (doi: 10.1126/science.1086807)

Figure S1. Experimental design, with effect of role analyzed by comparing scouts with recruits, and effect of context analysed by comparing foragers (food) and nest hunters (nest).

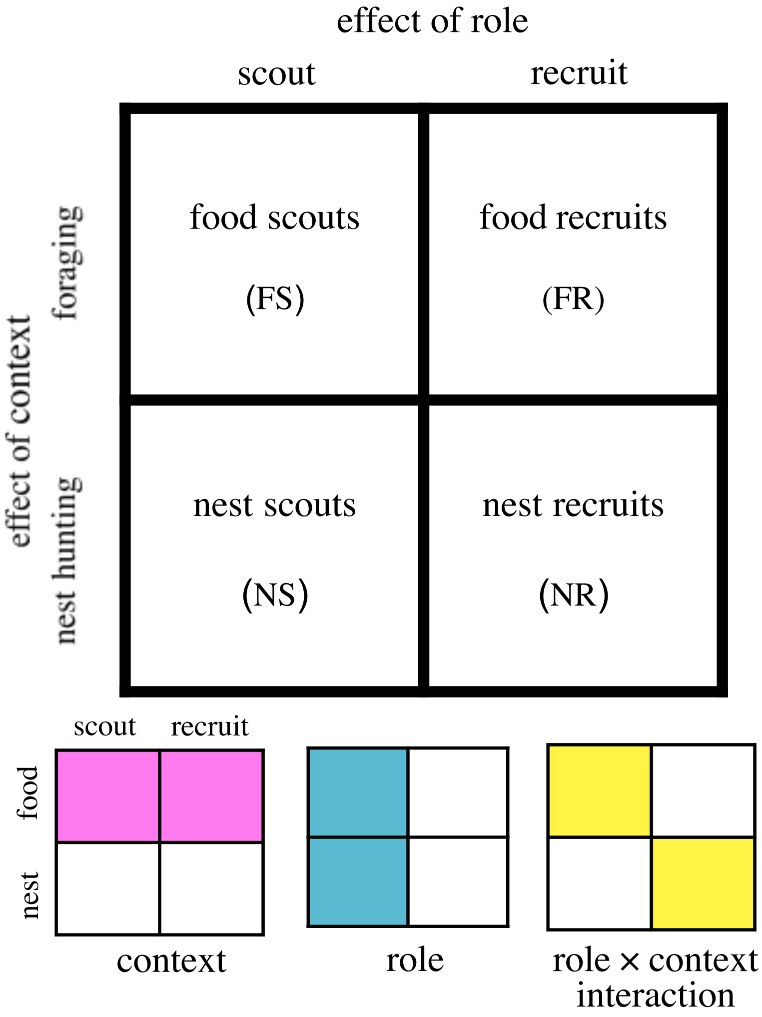

Figure S2 Microarray loop design for the comparative brain gene expression between two scouting behaviours. FS: food scouts, FR: food recruits, NS: nest scouts, NR: nest recruits. B1-12, B1-13 and B1-14 were three colonies used to collected food scouts and recruits. B1-12 and B2-17 were two colonies used to collected nest scouts and recruits. A total of 138 arrays were used. Detailed description of sample use was in the Material and Methods section of the main text.

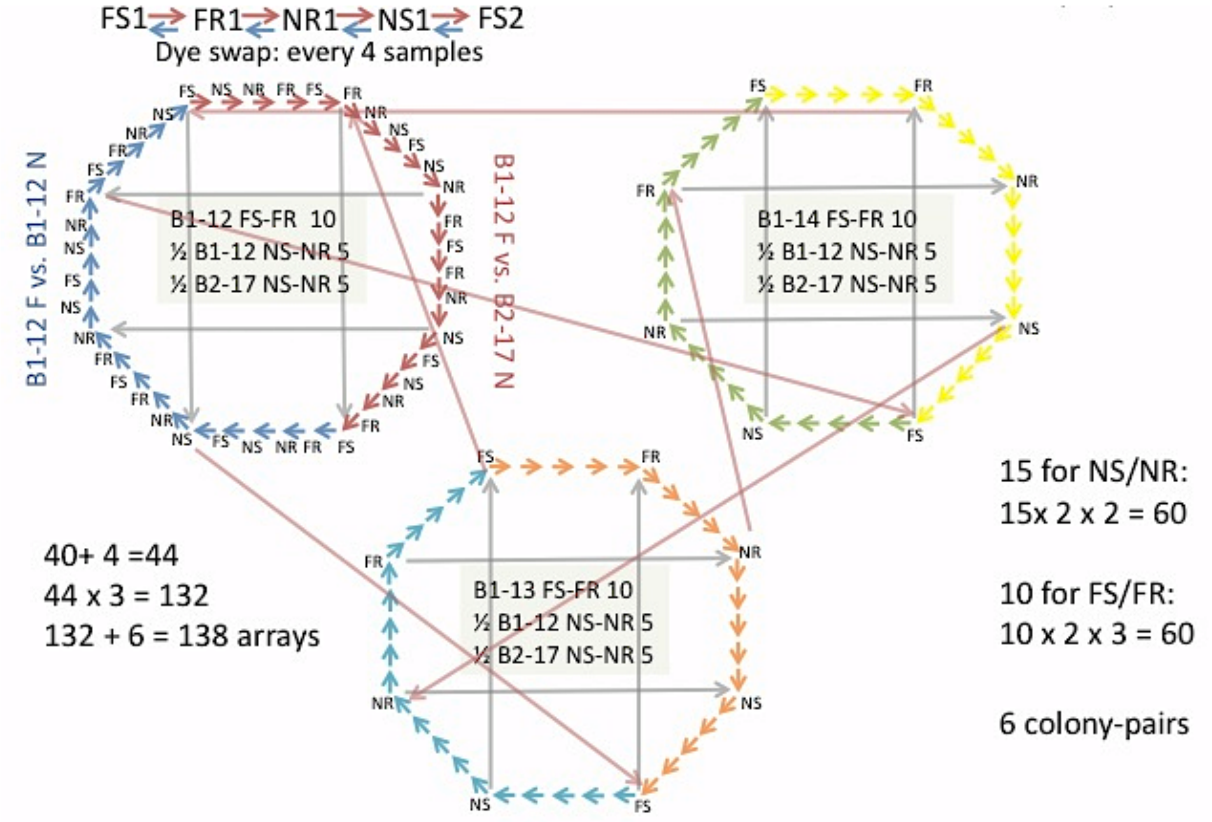

Figure S3. Linear Discriminant Analysis (LDA) failed to separate the four behavioural groups either by context or by role when using 5 neural signaling genes in the serotonin and tyramine systems. (a) Food scouts and food recruits. (b) Nest scouts and nest recruits. FS: food scouts (red), FR: food recruits (blue), NS: nest scouts, NR: nest recruits, with different colonies denoted as subscript 1, 2 and 3. More detailed gene information is listed in Table S1.

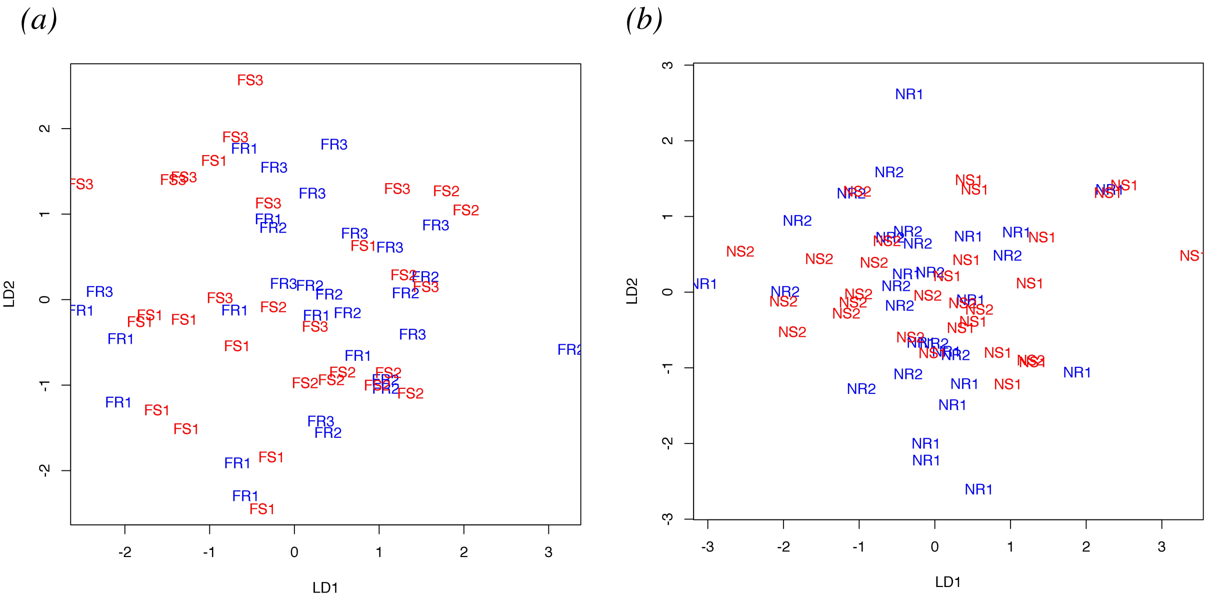

Figure S4. Log 2 brain expression values of *odor binding protein 4* (*obp4*) gene **(A)** among the four behavioural groups during scouting, and **(B)** across different behaviours. In (a), the *obp4* expression patterns shows an up-regulation pattern and the highest fold change between nest scouts and recruits and a moderate down-regulation in food scouts in comparison to food recruits (both experimented in the field environment, FDR<0.05, contrast $p<0.05$). In (b), the *obp4* gene was up-regulated in food scouts compared to non-scouts (experiments in semi-natural enclosure), and in foragers compared to nurses. This gene was down-regulated in active foragers compared to inactive foragers during the morning (AM) and afternoon (PM), all at FDR<0.05, contrast $p<0.005$ level. *Obp4* was also found to be significantly down-regulated (at FDR<0.05 without contrast test) in high pollen-hoarding foragers, but the fold change was very marginal.

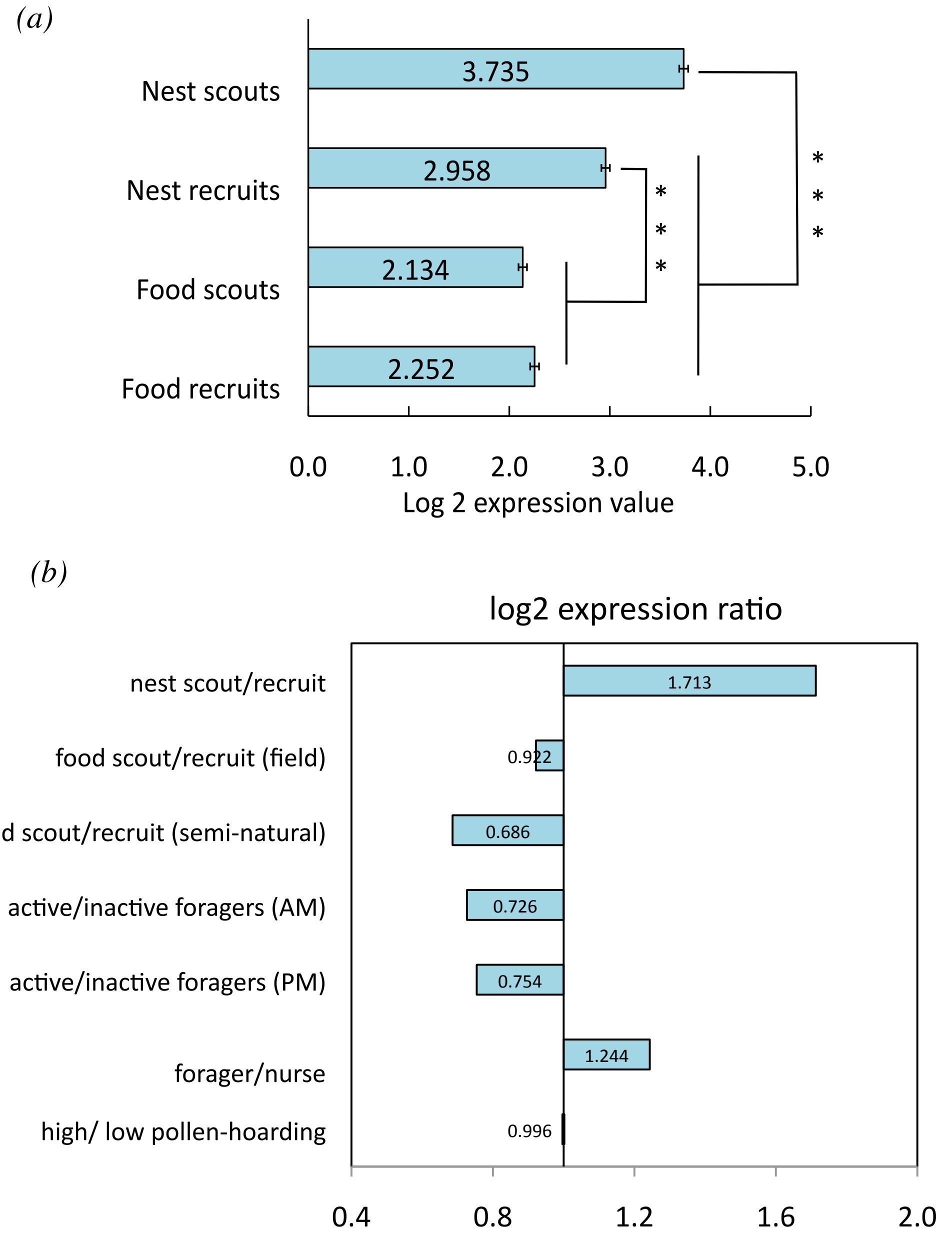

Figure S5. Hierarchical Clustering result was shown in heatmap for the 89 best predictor genes across the expression values of four behavioural groups (group mean). FS: food scouts, FR: food recruits, NS: nest scouts, NR: nest recruits (gene list: Table S3).

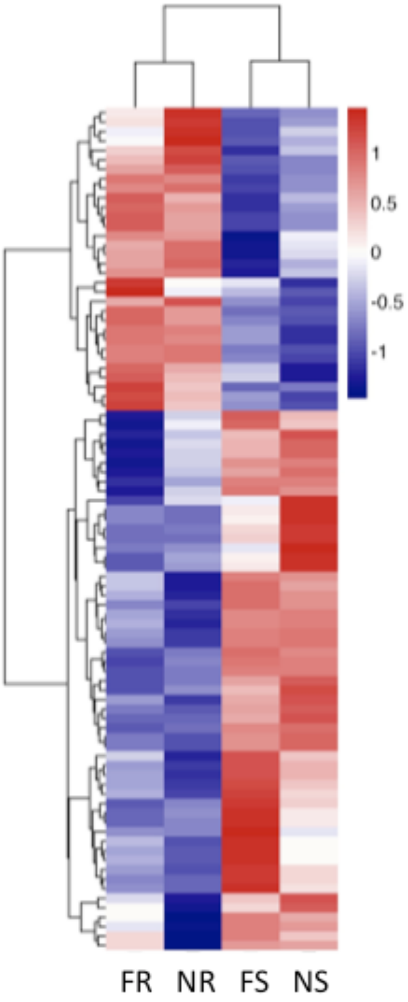

Table S1. Gene information for the 16 neural signaling genes studied as a part of shared molecular signatures of both types of scouting. "Functional description" based on full names of the genes (NCBI, updated 15-Aug-2013).

Array ID	Gene name	GB_OGS 2.0	GB_OGS 3.2	Functional description
AM06883	nAChRa5	GB17254	GB50159	nicotinic acetylcholine receptor alpha5 subunit
AM10791	CheT-1-like	GB13547	GB41372	high-affinity choline transporter 1-like
AM03745	CheT 1	GB11121	GB43293	high-affinity choline transporter 1
AM01620 &AM12780	DopR1	GB 30031_3	GB50192	dopamine receptor 1
AM10492	DopR2	GB17921	GB50155	dopamine receptor 2
AM09330	Gat-1b	GB16752	GB51198	sodium- and chloride-dependent gaba transporter 1b
AM11925	Gat-a	GB19372	GB54918	sodium- and chloride-dependent gaba transporter 1
AM05627	mGAR-B-R3	GB13016	GB53009	metabotropic gamma-aminobutyric acid type b receptor subtype 3
AM12526	Gad1	GB19979	GB40118	glutamate decarboxylase-like
AM11182	mGluR1	GB18621	GB44968	metabotropic glutamate receptor 1
AM08175	mGluR2	GB15591	GB44216	metabotropic glutamate receptor 2
AM04064	Glu-RI	GB11443	GB40973	ionotropic glutamate receptor 1
AM08958	Eaat-2	GB16377	GB49785	excitatory amino acid transporter 2
AM04250	Eaat-4	GB11630	GB53002	excitatory amino acid transporter 4
AM03889	OctR1	GB11266	GB52910	octopamine receptor
AM11426	Oct-b-2R	GB18869	GB49696	octopamine receptor beta-2R

Table S2. Functional analysis results for differentially expressed genes in nest scouts and food scouts, respectively, based on GO and KEGG pathway enrichment analyses, with up-regulated and down-regulated genes are tested separately. (Tab label: GOenrich_Nest_UP, Nest_down, GOenrich_Food-UP, Food_down). The spreadsheets also includes GO analysis results for differentially expressed genes in two main effects: role and contexts, and their interaction, respectively (Table label: GOenrich_Role_UP, Role_down, GOenrich_Context_UP, Context_down, GOenrich_interaction). Enrichment was determined in two ways: more stringent test is a hypogeometric test with Benjamini-FDR<0.05, which results heighted in blue; or a EASE score with p-value <0.05, which analysis showed in Fig. 1b were based. CC: cellular component; MF: molecular function; BP: biological process. (Detailed description of this GO analysis is in the Materials and Methods).

An Excel spreadsheet named "Table S2_GO enrichment data.xlsx" is provided as a separate file and a part of the electronic supplementary material.

Table S3. Hormone signaling gene expression in (a) food scouting behaviour, (b) nest scouting behaviour. * Significant for "task" and "interaction", but not for "role", despite both significant in food and nest scouting. (c) Hormone signaling genes that were significant for effect of "role" between scouts and recruits (samples were pooled from both food and nest contexts, so here the expression patterns were an overall pattern across Table S3 A and B) ** Significant for the interaction between "role" and "context".

(a) foraging:

Array probe	GB number (OGS 2.0)	Gene name	Expression pattern
08021	30298	Ecdysone receptor (EcR), transcription variation 1	Scout down
08226	17814	Hormone receptor-like 38, CG1864 * (probable nuclear hormone receptor hr38)	Scout down
09450	16873	Ftz transcription factor 1, CG4059, loc726450	Scout up

(b) nest-hunting:

Array probe	GB number (OGS 2.0)	Gene name	Expression pattern
09226	16648	Ultraspiracle Protein (USP), CG4380	Scout up
03985	11364	Ecdysone-induced protein 75, CG8127	Scout up
03986	11364	Ecdysone-induced protein 75, CG8127	Scout up
03987	11364	Ecdysone-induced protein 75, CG8127	Scout up
10132	16648	DopEcR, isoform A. loc413040	Scout up
03384	17814	Ecdysone-regulated gene E74	Scout up
08226	17814	Hormone receptor-like 38, CG1864 * (probable nuclear hormone receptor hr38)	Scout up
09450	16873	Ftz transcription factor 1, CG4059, loc726450	Scout up
06669	30150	Broad-complex (Br-c), CG11491	Scout up

(c) role:

Array probe	GB number (OGS 2.0)	Gene name	Expression pattern
09226	16648	Ultraspiracle Protein (USP), CG4380 **	Scout up
03144	17330	Juvenile hormone acid methyltransferase, CG17330	Scout down
03985	11364	Ecdysone-induced protein 75, CG8127 **	Scout up
03986	11364	Ecdysone-induced protein 75, CG8127**	Scout up
10132	16648	DopEcR, isoform A. loc413040 **	Scout up
09450	16873	Ftz transcription factor 1, CG4059, loc726450	Scout up
06669	30150	Broad-complex (Br-c), CG11491 **	Scout up

Table S4. Class prediction results using uncorrelated shrunken centroid (USC) method. *(a)* Best predictor genes selected from 557 “role-only” genes. Delta and rho are user-selected parameters. Gene lists from delta=0.5 are subsets of delta=0, and within each delta value, the selected gene number increases while smaller gene lists are the subsets of the larger gene lists. *(b)* Cross-context class prediction results, testing how well the identifier genes best for predicting food scouts and recruits predict nest scouts and recruits and *vice versa*. USC: delta=0, rho=0 was used for both tests in the table. *P* value with an asterisk denotes significance (*p*<0.005).

(a)

USC: delta=0	**rho=0.5**	**rho=0.6**	**rho=0.7**	**rho=0.8**	**rho=1**	**no selection**
no. of predictor genes	89	167	302	428	509	557
no. of correct prediction	111	110	110	109	108	94
% of correct prediction	92.5 %	91.7%	91.7%	90.8%	90.0%	78.3%
USC: delta=0.5	**rho=0.5**	**rho=0.6**	**rho=0.7**	**rho=0.8**	**rho=1**	
no. of predictor genes	81	145	247	325	371	
no. of correct prediction	110	111	111	108	106	
% of correct prediction	91.7%	92.5%	92.5%	90.0%	88.3%	

(b)

significant gene set	**training individual profiles**	**no. of best identifier genes** (% of correct prediction)	**testing individual profiles**	**no. of correct prediction**	**binomial test *p* value**
food scouting (1003 genes)	30 food scouts & 30 recruits	95 genes (95%)	30 nest scouts & 30 recruits	31/60 bees (58%)	0.449
nest scouting (1032 genes)	30 nest scouts & 30 recruits	108 genes (95%)	30 food scouts & 30 recruits	39/60 bees (65%)	0.0134 *

Table S5. Gene ID and annotation of 89 best predictor genes in OGS 3.2. Dmel_r5.42 are the up-to-date gene annotation in *Drosophila* genome (by May 2012). #N/A: not available in gene annotations or *Drosophila* orthologs. (Bold font: 16 transcripts currently have no gene information, thus have all #N/A except Probe IDs).

Probe ID	GB Number OGS3.2	Gene Names	Dmel_r5.42_Gene_ID	Dmel_r5.42_Name
AM06501R	GB54890	kynurenine 3-monooxygenase	FBgn0000337	cn-PA
AM11801	GB46713	elongation factor 2	FBgn0000559	Ef2b-PA
AM12505	GB42011	follicle cell protein 3c-1	FBgn0000644	Fcp3C-PA
AM07018	GB48999	transcription factor ap-4	FBgn0001994	crp-PA
AM02905	GB44133	tubulin beta-1 chain	FBgn0003889	betaTub85D-PA
AM12856R	GB54085	#N/A	FBgn0004828	His3.3B-PA
AM04761	GB49031	ribosomal protein s18	FBgn0010411	RpS18-PB
AM08533	GB42312	mitochondrial isoform 1	FBgn0024891	ferrochelatase-PA
AM03767	GB40778	udp-n-acetylglucosamine transporter	FBgn0024994	Csat-PB
AM10675	GB52717	#N/A	FBgn0025709	CG8083-PA
AM10502	GB43265	equilibrative nucleoside transporter 1-like isoform 1	FBgn0026585	Ent2-PA
AM04357	GB43719	alpha-catulin-like isoform 3	FBgn0029105	alpha-catenin-related-PA
AM10585	GB43117	nucleoside diphosphate-linked moiety x motif mitochondrial-like	FBgn0030528	CG11095-PA
AM07719	GB44686	adp-ribosylation factor-like protein 16-like	FBgn0031254	CG13692-PA
AM05627	GB53009	gamma-aminobutyric acid type b receptor subunit 2	FBgn0031275	GABA-B-R3-PA
AM04152	GB52600	uncharacterized protein loc100871763	FBgn0031632	CG15628-PA
AM06376	GB54929	#N/A	FBgn0032005	Snx6-PB
AM05883	GB43205	uncharacterized protein loc100867964	FBgn0032598	ChLD3-PA
AM05890	GB43379	protease m50 membrane-bound transcription factor site 2 protease	FBgn0033656	S2P-PA
AM12313	GB41602	sugar transporter	FBgn0034045	CG8249-PA
AM02800	GB48112	abhydrolase domain-containing protein 16a-like	FBgn0035519	CG1309-PA
AM02667	GB46008	PREDICTED: adrenodoxin, mitochondrial-like [Apis mellifera]	FBgn0035529	CG1319-PB
AM07547	GB55515	inositol oxygenase-like	FBgn0036262	CG6910-PB

Table S5 (cont.)

AM10298	GB40580	integrator complex subunit 9	FBgn0036570	IntS9-PA
AM06329	GB49994	ribosomal protein l26e	FBgn0036825	RpL26-PA
AM10697	GB44936	histone-arginine methyltransferase carmer-like isoform 1	FBgn0037770	Art4-PA
AM09435	GB45806	unc93-like protein mfsd11-like	FBgn0038053	CG18549-PA
AM11809	GB50572	osiris 14	FBgn0040279	Osi14-PA
AM04616	GB42671	glypican 6	FBgn0041604	dlp-PB
AM02383	GB46479	uncharacterized protein loc100872175	FBgn0044324	Chro-PB
AM02866	GB42844	guanine nucleotide exchange factor dbs	FBgn0050440	CG30440-PA
AM08404	GB52980	probable g-protein coupled receptor 158	FBgn0051195	CG31195-PB
AM10876	GB47354	neurotrimin- partial	FBgn0051646	CG31646-PA
AM00120	GB51398	l-aminoadipate-semialdehyde dehydrogenase-phosphopantetheinyl transferase	FBgn0052099	eap-PA
AM06335	GB44917	low density lipoprotein receptor-related protein	FBgn0052432	CG32432-PA
AM12190	GB51578	mitochondrial import inner membrane translocase subunit tim50-c-like	FBgn0250874	ttm50-PA
AM01698	GB52172	sh2 domain-containing adapter protein f	FBgn0259109	CG42251-PC
AM09976	GB40089	voltage-dependent l-type calcium channel subunit beta-2-like	FBgn0259822	Ca-beta-PI
AM02994	GB42487	calpain-c	FBgn0260450	CalpC-PA
AM05789	GB43054	#N/A	FBgn0261259	Fhos-PC
AM07970	GB47052	gene model 996	FBgn0261802	CG42748-PG
AM10360	GB40150	forkhead box protein P	FBgn0262477	FoxP-PC
AM03831	GB53419	pyrazinamidase nicotinamidase-like	#N/A	#N/A
AM11223	GB45906	protein lethal essential for life-like	#N/A	#N/A
AM11755	GB55300	#N/A	#N/A	#N/A
AM06799	GB43761	uncharacterized protein loc100866401	#N/A	#N/A
AM08113	GB44455	poor imd response upon knock-in	#N/A	#N/A
AM02267	GB51435	#N/A	#N/A	#N/A
AM01635	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM09485	GB54184	bridging integrator 3 homolog	#N/A	#N/A
AM00574	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM01504	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM02307	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM09723	GB48133	l-lactate dehydrogenase a-like 6a-like	#N/A	#N/A
AM01423	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM00665	GB50486	#N/A	#N/A	#N/A
AM03859	GB41579	neuronal calcium sensor 2	#N/A	#N/A
AM02126	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM05033	GB49828	zinc finger cchc-type and rna-binding motif-containing protein 1-like	#N/A	#N/A
AM06506	GB50674	uncharacterized protein loc100879076	#N/A	#N/A
AM03212	GB43788	enhancer of split mgamma	#N/A	#N/A
AM07265	GB53750	upf0454 protein c12orf49 homolog	#N/A	#N/A
AM04631	GB40086	sodium-dependent multivitamin transporter	#N/A	#N/A
AM02964	GB50124	protein bv8	#N/A	#N/A
AM05101	GB52845	sh3 and cysteine-rich domain-containing protein 2	#N/A	#N/A
AM12118	GB41499	tetratricopeptide repeat protein 5	#N/A	#N/A

Table S5 (cont.)

AM11465	GB42877	polyamine-modulated factor 1	#N/A	#N/A
AM08986	GB44965	selenophosphate synthetase	#N/A	#N/A
AM02331	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM08728	GB45121	#N/A	#N/A	#N/A
AM00622	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM00499	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM12175	GB55541	ubiquitin-conjugating enzyme e2c-binding	#N/A	#N/A
AM12902R	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM01771	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM04023	GB46086	#N/A	#N/A	#N/A
AM08966	GB50806	#N/A	#N/A	#N/A
AM01678	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM02799	GB55582	#N/A	#N/A	#N/A
AM00664	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM00693	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM07754	GB52339	microtubule-associated proteins 1a 1b light chain 3a-like	#N/A	#N/A
AM07930	GB50946	uncharacterized protein loc100869298	#N/A	#N/A
AM10769	GB46912	tudor domain-containing protein 12	#N/A	#N/A
AM04830	GB50880	uncharacterized protein loc100866375	#N/A	#N/A
AM00646	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM08161	GB50625	#N/A	#N/A	#N/A
AM02489	**#N/A**	**#N/A**	**#N/A**	**#N/A**
AM03598	GB54137	transcription factor with ap2 domain	#N/A	#N/A