Frequent Recent Origination of Brain Genes Shaped the Evolution of Foraging Behavior in Drosophila

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Chen, Sidi, Maria Spletter, Xiaochun Ni, Kevin P. White, Liqun Luo, and Manyuan Long. “Frequent Recent Origination of Brain Genes Shaped the Evolution of Foraging Behavior in Drosophila.” Cell Reports 1, no. 2 (February 2012): 118–132.</th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1016/j.celrep.2011.12.010">http://dx.doi.org/10.1016/j.celrep.2011.12.010</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Elsevier</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Sun Apr 24 12:01:15 EDT 2016</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/101698">http://hdl.handle.net/1721.1/101698</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Creative Commons Attribution</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td><a href="http://creativecommons.org/licenses/by-nc-nd/3.0/">http://creativecommons.org/licenses/by-nc-nd/3.0/</a></td>
</tr>
</tbody>
</table>
Frequent Recent Origination of Brain Genes Shaped the Evolution of Foraging Behavior in *Drosophila*

Sidi Chen,1,5 Maria Spletter,2,4 Xiaochun Ni,1,3 Kevin P. White,1,3 Liquan Luo,2 and Manyuan Long1,*

1Department of Ecology and Evolution, The University of Chicago, Chicago, IL 60637, USA
2Howard Hughes Medical Institute and Department of Biology, Stanford University, Stanford, CA 94305, USA
3Institute for Genomics and Systems Biology, The University of Chicago and Argonne National Laboratory, Chicago, IL 60637, USA
4Present address: Laboratory of Muscle Dynamics, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany
5Present address: Department of Biology and Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
*Correspondence: mlong@uchicago.edu
DOI 10.1016/j.celrep.2011.12.010

SUMMARY

The evolution of the brain and behavior are coupled puzzles. The genetic bases for brain evolution are widely debated, yet whether newly evolved genes impact the evolution of the brain and behavior is vaguely understood. Here, we show that during recent evolution in *Drosophila*, new genes have frequently acquired neuronal expression, particularly in the mushroom bodies. Evolutionary signatures combined with expression profiling showed that natural selection influenced the evolution of young genes expressed in the brain, notably in mushroom bodies. Case analyses showed that two young retrogenes are expressed in the olfactory circuit and facilitate foraging behavior. Comparative behavioral analysis revealed divergence in foraging behavior between species. Our data suggest that during adaptive evolution, new genes gain expression in specific brain structures and evolve new functions in neural circuits, which might contribute to the phenotypic evolution of animal behavior.

INTRODUCTION

Throughout the animal kingdom, the nervous system plays a fundamental role in processing sensory information and forming proper behavioral responses (Beatty, 1995). Originating from a simple protobrain of a single common ancestor, brains have evolved to be the structural and functional center of the nervous system in most metazoan taxa (Shepherd, 1994). There is a large diversity of central nervous systems in both structural organization and functional specialization (Denes et al., 2007). One of the primary functions of the brain, regardless of its level of complexity, is to control the behavior of an animal (Carew, 2000). Behaviors evolve constantly (Brown, 1975; Lowe et al., 2003), and the evolution of behavior is associated with evolution of the brain (Oro, 2004).

Recently, extensive efforts have focused on understanding the genetic basis and molecular mechanisms for brain evolution. Evidence shows that evolutionary changes in the size, shape, structure, and function of the brain are highly correlated with various genetic changes, such as adaptive evolution in protein sequences (Zhang, 2003), regulatory elements (Haygood et al., 2007), and noncoding RNAs (Pollard et al., 2006); changes in gene expression (Cáceres et al., 2003); and occasionally the birth and death of genes (Burki and Kaessmann, 2004; Popesco et al., 2006). However, we know little about how recently evolved genes shape the brain and animal behavior.

New genes are the novel genetic loci, physically distinct and derived segments of DNA that originate in an evolutionary time with newly encoded functional transcripts (Long and Langley, 1993; Long et al., 2003). It is well known that new genes frequently originate during evolution and can acquire functions in different biological processes in various organisms (Kaessmann et al., 2009; Long et al., 2003). However, little is known about young genes that are expressed and potentially function in the brain. We first identified newly evolved genes with neuronal expression in *Drosophila*. Further characterization of their expression patterns revealed significantly enriched expression of new genes in the mushroom bodies (MB). We then detected the signatures of natural selection for the evolution of these genes. Furthermore, we examined the roles of these genes in behavior. We found that animals deficient in two recently originated MB-expressed retrogenes, Xcpb1 and Desr, showed foraging behavior phenotypes.

RESULTS

Identification of Young Brain Genes

To identify recently evolved brain genes in *Drosophila*, we curated genes that originated in the *Drosophila* genus (Table S1) from the literature (Chen et al., 2010; Zhang et al., 2010). Among *D. melanogaster* genes that are younger than 25 million years (Myr), arising after the divergence from *D. pseudoobscura*, we found that 48.8% (161/330) were expressed in the brain, as detected by RT-PCR (Figures 1A and S1, Data Set S1, Extended Experimental Procedures). We refer to those young genes with...
distinct brain expression as “young brain genes” and other genes as “young nonbrain genes.”

We examined the brain gene origination events. Young brain genes can be generated by DNA/RNA-based gene duplication or de novo (Figure 1B). For duplicated genes, 65.2% (105/161) moved to new genomic locations distant from those of their parental genes (Figures S1A and S1B). On the Drosophila phylogeny, we observed a burst of brain gene formation between 3 and 6 Myr ago (Figure 1A), a significant enrichment of young brain over young nonbrain genes (Fisher’s exact test to compare 28/132/155, two tailed, p < 0.01). Young brain genes are enriched on the X chromosome with a 44% excess over random expectation ($\chi^2$ test, degrees of freedom = 5, p = 1.5 × 10^{-3}), whereas young nonbrain genes are not (Figure 1C), implying that the sex chromosome has gained more brain genes than autosomes recently. Young brain genes encode various protein domains, which are enriched in several biological processes for protein level regulation, such as regulation of kinase activity and phosphorylation, whereas young nonbrain genes are enriched in a unique term proteolysis (Data Set S2).

**Young MB Structures Recruited an Excess of Young Brain Gene Expression**

We next determined the expression pattern of young brain genes at cellular resolution in the adult brain by using enhancer trap lines, as they often mimic the expression pattern of the genes adjacent to the insertion site of the P-element (Brand and Perrimon, 1993). We obtained 97 Gal4 enhancer trap lines identified from GETDB (Hayashi et al., 2002) and CBD (Bourbon et al., 2002), representing 35 newly evolved genes. We identified 30 lines that drive clear UAS-mCD8GFP expression patterns in substructures of the brain, representing 17 genes younger than 25 Myr (Tables S2 and S3, Extended Experimental Procedures). The proportion of genes expressed in the brain identified by enhancer trap (48.6%, or 17/35) agreed with that by RT-PCR (48.8%, or 161/330). Additionally, expression patterns from the few genes with available mRNA in situ hybridization data were consistent with those from the enhancer trap lines (Bourbon et al., 2002; Bousum et al., 2008; Hong and Ganetzky, 1996; Tomancak et al., 2007).

Collectively, young brain genes were expressed in neurons projecting to most major neuropils in the brain of D. melanogaster (Figure 2, Table S2). Different genes showed distinct expression patterns in one or more structures. For example, hog (CG32595), a ~6- to 11-Myr-old X-linked Forkhead-associated transcription factor known to be involved in neuronal cell migration and differentiation (Bousum et al., 2008), was expressed in all major brain structures we scored (Figure 2, Table S2). In contrast, CG71825, a gene that encodes...
a putative hypoxia-induced protein, was expressed specifically in a small subset of neurons in the protocerebra and SOG (Figure 2, Table S2). These data demonstrated that recently originated genes acquired stereotypic expression patterns in substructures of the brain, which implies that they acquired neuronal regulation of gene expression.
We next examined how often brain-expressed genes are expressed in the MB. From our set of 35 young genes, of the 17 brain-positive genes, 82% (14/17) are expressed in MBs (Figure S2, Tables S2 and S3). By contrast, from 1934 randomly chosen genes, of the 1231 genes that are expressed in the brain, only 34% (429/1231) are expressed in MBs (E.C. Marin and L.L., unpublished data). An independent enhancer-trap-based study estimated a similar rate of 23% (65/281) for randombrain-expressed genes with MB expression (Kelso et al., 2004). While the basal probability of brain expression is similar between young and random genes in the genome, young genes are significantly enriched in the MB (Fisher’s exact test, p = 0.0018 and p = 0.022, respectively). Given that enhancer trap collections represent a relatively random sampling of genomic loci with respect to brain expression (Brand and Perrimon, 1993; Hayashi et al., 2002), these data suggest that the MB is a favored tissue for new genes when they acquire expression in the brain.

The MB consists of three distinct types of neurons, including the α/β, α'/β', and γ neurons (Crittenden et al., 1998; Lee et al., 1999; Tanaka et al., 2008). Interestingly, all of the MB-positive young brain genes are expressed in the α/β neurons, while only four show expression in α'/β' or γ or α'/β' neurons (Figure S2, Tables S2 and S3). Previous work has shown that the γ lobe is the most ancestral, while the α/β lobes are derived and the most heterogeneous (Strausfeld and Li, 1999a, Strausfeld and Li, 1999b). The preferential expression of young brain genes in the α/β lobes suggests that the derived substructures may have frequently recruited new genes during recent evolution.

**Expression Profiling of the MB Transcriptome**

To examine the expression profile of MBs at the genomic level, we profiled the transcriptomes of dissected MBs in parallel with dissected whole brains by RNA sequencing and confirmed the quality of the data set (Figures 3A and SSA–SSD). In brain and MB RNA-seq samples, young brain genes are expressed at higher levels than young nonbrain genes, and young MB genes identified by enhancer trap are expressed significantly higher than young non-MB genes (Figures 3C and 3D, Wilcoxon test, p < 0.001 for all comparisons). The correlation between young and parental gene expression is weak (Figure S3E), suggesting that the expression pattern and level of most young genes are not directly inherited from their parental genes (Extended Experimental Procedures). These data provide a genome-wide data set for testing differential gene expression in the brain and MBs.

To estimate the differential expression for each gene between MB and whole brain, we used a generalized linear model framework as previously described (Marioni et al., 2008). A total of 2,940 genes were identified as differentially expressed (multiple testing corrected p value < 0.001), including 58 young genes (< 60 Myr) and 2,272 old genes (≥ 60 Myr) (Data Set S4). In the old gene group, only half (48.9%, or 1,112/2,272) of the differentially expressed genes are MB-enriched, while in the young gene group, a significantly higher proportion of differentially expressed genes (91.4%, or 53/58) are MB-enriched (Fisher’s exact test, two-tailed, p < 0.0001, Figures 3B and S3B). This observation was recapitulated with a more stringent young gene data set (Table S1), revealing a stronger pattern: 97.0% (32/33) of young genes are MB-enriched (Figures 3B and S3B, young versus old, p < 1 × 10⁻⁵). Similar patterns were not observed in either MB versus testis or brain versus testis differential expression analyses (Data Sets S5 and S6). These data revealed an excess of MB-enriched genes and a paucity of MB-depleted genes in the young gene group, suggesting that differentially expressed young genes tend to be enriched in the MBs.

**Natural Selection on Young Brain and Young MB Genes**

We examined natural selection on young brain and/or MB genes. Using Drosophila polymorphism data (Begun et al., 2007) (DPGP) (Data Set S3), we estimated θ (Experimental Procedures); i.e., the proportion of amino acids under positive selection (Eyre-Walker and Keightley, 2009). A higher θ is indicative of stronger positive selection. We found that the θ of young brain genes (+0.596) is significantly higher than that of young nonbrain genes (0.525), random genes (0.462), or old brain genes (0.464) (Figures 3E and S3F) (Wilcoxon test, p = 0.00001 for all comparisons).

Young MB genes identified by enhancer trap have strong positive selection (θ = +0.634), significantly higher than young non-MB genes (0.344) (Figure 3F, Wilcoxon test, p = 0.00002). At the genomic level, MB-enriched genes (θ = +0.534) showed a stronger signature than MB-depleted genes (0.494) (Figure 3F, Wilcoxon test, p = 0.00002). These data suggest that stronger positive Darwinian selection has shaped the evolution of young brain genes and possibly influenced their expression in MBs.

**Evolution and Expression Pattern of a Young Brain Gene, Xcbp1**

Young brain/MB genes may have offered fitness advantages for selection. A young MB gene, Xcbp1, showed a strong signal of positive selection (Dn = 195, Ds = 80, Pn = 27, Ps = 23, McDonald and Kreitman test, p = 0.02), a starting point for phenotypic examination to test this hypothesis.

Xcbp1 originated roughly 5 Myr ago, after the D. melanogaster-D. simulans clade diverged from D. yakuba but before the clade split (Figure 4A, Table S2). Its parental gene, Cnx99a, encodes a 1 Calnexin that binds to calcium ions and generally acts as a chaperone to facilitate folding of glycoproteins such as rhodopsins (Rosenbaum et al., 2006; Figure 6A). During its initial retroposition event, Xcbp1 retained only the CDS portion from its parental gene Cnx99a (Figure 4B) (Bai et al., 2007) and integrated into an X-linked cluster of five previously identified neuronal genes (Figure 4B) by nesting into the first intron of paralytic, which encodes the major voltage-sensitive sodium channel in Drosophila (Hong and Ganetzky, 1996). This genomic region contains multiple binding sites of well-known neuronal transcription factors such as Dichaete, Disconnected, Jumeau, and Senseless (Nègre et al., 2011). Xcbp1 likely hitchhiked the regulatory elements from this neuronal gene environment and acquired a distinct neuronal expression pattern, as shown previously by in situ hybridization (Hong and Ganetzky, 1996).

To identify neurons that express Xcbp1, we utilized the Gal4 enhancer trap line pG156 (Bourbon et al., 2002), hereafter called Xcbp1-Gal4, which is inserted 1 kb 5' to the Xcbp1 transcription start site on the sense strand of Xcbp1 (Figure 4C). Xcbp1-Gal4 drives UAS-mCD8GFP expression in the peripheral nervous system.
Figure 3. Adaptive Evolution of Young Brain and Young MB Genes

(A) Schematic representation of RNA-seq showing the workflow of experimental (dissection, RNA-sequencing) and analytic procedures, with a heat map of young gene expression in MB, brain, and testis at the bottom.

(B) Statistical summary of MB-enriched and MB-depleted genes in differentially expressed gene sets; 60 Myr, genes younger than 60 million years (n = 447, DE = 58); 35 Myr, genes younger than 35 million years (n = 279, DE = 33); Old, genes older than 60 Myr (n = 9275, DE = 2272); Random, randomly sampled genes from the genome (n = 11820, DE = 2940); DE, number of differentially expressed genes.
system, especially in olfactory receptor neurons (ORNs) in the antennae and maxillary palps (Figure 4D), which are the primary olfactory sensory organs (Stocker, 1994). Furthermore, Xcbp1-Gal4 is also expressed in α/β MB neurons (Figures 4E and S2), and the MB is a higher brain center known to be involved in olfactory learning and memory (Keene and Waddell, 2007). Using eyFlp and UAS > Stop > mCD8GFP to limit expression to only ORNs, we verified that Xcbp1-Gal4-positive ORNs projected into the antennal lobes, especially to glomeruli VC3m, VC3i, DC4, V, DL5, and DM1 (Figures 4E S4A, and S4B), which are innervated by coeloconic and basiconic ORNs that are likely receptive to food odors (Benton et al., 2009; Hallem et al., 2006). In particular, the DM1 glomerulus has a major role in attraction to food odors (Semmelhack and Wang, 2009), suggesting that Xcbp1+ neurons might participate in this process.

Xcbp1 Facilitates Foraging

Xcbp1-Gal4 is primarily expressed in brain centers and sensory organs that are involved in food source detection and odor processing and therefore may have a role in foraging behavior. We designed an assay to measure the foraging ability of Drosophila (Figure 5A), and we first tested two wild-type (WT) D. melanogaster lines (Canton-S and Oregon-R). We observed that starved flies efficiently found and immediately started consuming food (Figure 5B, Extended Experimental Procedures, Movie S1). More than half of the WT flies found food within 3 hr, and we observed no significant differences between the WT strains (two-sample Komogorov-Smirnov (KS) test, not significant (NS), Figure 5C). We also quantified foraging performance strains (two-sample Komogorov-Smirnov (KS) test, not significant (NS), Figure 5C). We also quantified foraging performance strains (two-sample Komogorov-Smirnov (KS) test, not significant (NS), Figure 5C). We also quantified foraging performance strains (two-sample Komogorov-Smirnov (KS) test, not significant (NS), Figure 5C). We also quantified foraging performance strains (two-sample Komogorov-Smirnov (KS) test, not significant (NS), Figure 5C). We also quantified foraging performance strains (two-sample Komogorov-Smirnov (KS) test, not significant (NS), Figure 5C). We also quantified foraging performance strains (two-sample Komogorov-Smirnov (KS) test, not significant (NS), Figure 5C). We also quantified foraging performance strains (two-sample Komogorov-Smirnov (KS) test, not significant (NS), Figure 5C).

We tested whether these Xcbp1-positive (Xcbp1+) neurons are responsible for foraging behavior. Using a temperature-sensitive dominant-negative Dynamin allele, ShibireT1 (Kittamoto, 2001), we specifically inactivated synaptic transmission in Xcbp1-Gal4+ neurons using Xcbp1-Gal4. When Xcbp1-Gal4 / UAS-ShibireT1 flies were shifted to restrictive temperature (31°C) before testing to inhibit synaptic transmission in Xcbp1+ neurons, there was a sharp reduction in foraging performance as compared to WT flies (Figures S5A and S5B). Both Xcbp1-Gal4 and UAS-ShibireT1 control flies behaved like WT flies at 24°C and 31°C (Figure S5). These data showed that Xcbp1+ neurons are necessary for foraging.

We then measured the foraging behavior of D. melanogaster when Xcbp1 was knocked down by RNA interference (RNAi). Constitutive tubulin::Gal4-driven Xcbp1 RNAi (UAS::ds-Xcbp1) (Dietz et al., 2007) reduced Xcbp1 transcript level sharply compared to controls, with no off-target effect on Cnx99a (Figure 5B). Compared to controls, Xcbp1 RNAi flies (tubulin::Gal4 > > UAS::ds-Xcbp1) showed a significant reduction in foraging ability (KS test, p = 2 × 10^-5 Figure 5E). This was recapitulated with both Act5C::Gal4 and Xcbp1-Gal4 (KS test, p = 0.0001, respectively, Figures 5E and SF). These data indicated that a decrease in Xcbp1 expression reduces the foraging ability of D. melanogaster.

Multiple modalities are important for foraging. We investigated the locomotion and flight abilities of Xcbp1 RNAi flies with previous methods (Stockinger et al., 2005). Both locomotion and flight were normal for Xcbp1 RNAI and controls (Figures S5A and S5H). This suggested that neither a locomotion defect nor a flight defect was responsible for the foraging phenotype. Given the prominent expression of Xcbp1 in the olfactory system, it is possible that defects in olfactory neurons contribute to the foraging deficiency in Xcbp1-Gal4-driven Xcbp1 knockdown (Figure S5E).

Since Xcbp1-Gal4 is expressed in MBs, which are essential for the olfactory response, we used the pan-MB Gal4 driver OK107 to silence Xcbp1 in the MBs. OK107-driven Xcbp1-RNAi also led to a reduction in foraging ability equivalent to constitutive or Xcbp1-Gal4 driven RNAi (Figure 5I), suggesting that Xcbp1 expression in the MBs is critical for efficient foraging.

Olfaction is essential for foraging (Osborne et al., 2001). The phenotype of Xcbp1 is consistent with its expression pattern in the olfactory systems for food cue sensing (Carlson, 1996) and olfactory signal processing (Laurent, 2002). Xcbp1 expression in MB is required for foraging, suggesting that MB processing is critical, though multiple levels might be involved. Together, these data suggest that Xcbp1 evolved to participate in the neuronal circuit regulating foraging behavior.

Evolution of Xcbp1 Function

Cnx99a is ancestral and conserved, whereas Xcbp1 evolved rapidly under positive selection. At origination, Xcbp1 inherited a calnexin-like structure from its parental gene Cnx99a (Figures 6A, S6A, and S6B). Subsequently, Xcbp1 protein has rapidly evolved 36 amino acid (aa) substitutions, a large de novo 56 aa lysine-glutamate rich insertion in the putative cytosolic

(See also Figure S3B)
(C) Relative gene expression abundance (mean RPKM values from biological replicates) of young brain genes (y.brain) and young nonbrain genes (y.nonbrain) detected by RT-PCR, in brain and MB RNA-seq data sets. Error bars represent mean +/- SEM.
(D) Relative gene expression abundance of young MB genes (ET.MB) and young genes not expressed in the MB (ET.nonMB) detected by enhancer trap, in brain and MB RNA-seq data sets. Error bars represent mean +/- SEM.
(E) Estimation of alpha, the proportion of nonsynonymous substitutions subjected to positive selection, revealed positive selection on young brain genes. Young brain genes (young br.), have significantly higher alpha compared to young nonbrain genes (young nb.), random young genes (random young), or random genes (random), representing the genomic background from random sampling. The “random” error bar is too small to display. Error bars represent the 95% confidence interval.
(F) Estimation of alpha revealed positive selection on MB genes. MB-enriched genes have significantly higher alpha compared to MB-depleted genes; young MB genes (ET.MB) have significantly higher alpha compared to young non-MB genes (ET.nonMB). The “MB-depleted” error bar is too to display. Stars denote significance in statistical comparison, if applicable. Error bars represent the 95% confidence interval.
(See also: Figure S3)
Figure 4. Origin and Expression Pattern of Xcbp1
(A) Drosophila phylogeny showing the origination of Xcbp1.
(B) Scheme illustrating how the CDS of Cnx99a (green) retroposed (orange) and inserted into the first intron of paralytic (para), generating Xcbp1 (not drawn to scale).
(C) Illustration of the genomic location and orientation of Xcbp1-Gal4 (not drawn to scale).
(D) Expression of Xcbp1-Gal4 in D. melanogaster antenna and maxillary palp (GFP, green; scale bars represent 50 μm).
(E) Expression of Xcbp1-Gal4 in D. melanogaster brain (GFP, green; neuropil, red). Top panel: Xcbp1-Gal4 labeling of MB and AL in a half-brain view, full stack (E1). Magnified anterior (E2) and posterior (E3) views of the MB showing labeling of α and β lobes of the MB and the calyx. Bottom panel: ORN-specific labeling by the intersection of eyFlp and Xcbp1-Gal4 expression. Antennal nerve-specific expression is visible (arrows) in whole-brain view, full stacks (E4). Glomerular labeling of DC4, VC3, V, and DM1 is clearly visible in single planes of the middle (E5) and posterior (E6) layers of the AL. Scale bars represent 50 μm.
(See also: Figure S4)
Figure 5. Xcbp1 Influences Foraging Behavior

(A) Scheme of the foraging assay system design.

(B) Semiquantitative RT-PCR showing the specific and efficient RNAi knockdown of Xcbp1. Sample genotypes are shown above and assayed genes on the right.
C-terminal domain, and two small deletions in the putative luminal N-terminal domain (Figures 6A, S6A, and S6B).

Cnx99a has been reported to be important for rhodopsin maturation and photoreceptor survival (Rosenbaum et al., 2006). Consistently, we found that Cnx99a-RNAi using eyespecific GMR-Gal4 caused a disruption in ommatidial structure (Figure 6C). RNAi knockdown of Cnx99a by Act5C-Gal4 caused both shortening and bluntness in notum bristles (Figure 6D), suggesting that Cnx99a might also be involved in bristle development. RNAi knockdown of Xcbp1 in the eye also resulted in a disruption of ommatidial structure (Figure 6C), but knockdown with Act5C-Gal4 did not result in a bristle defect (Figure 6D). Neuronal knockdown (by Elav-Gal4) of Xcbp1 reduced foraging efficiency, whereas Cnx99a knockdown with the same driver did not (Figure 6B). Both Xcbp1 and Cna99a are expressed in the brain; however, whereas Xcbp1 is enriched in the MB over the brain (LRT, p = 0.007), Cnx99a is not (LRT, p = 0.24) (Data Sets S1–S6). The expression of Cnx99a in the brain is consistent with its function in rhodopsin maturation, while the MB-enriched expression of Xcbp1 might be related to foraging behavior. In the last five Myr, Xcbp1 inherited and maintained an ancestral role in eye development and might have acquired a novel role in foraging behavior.

Divergence of Foraging Efficiency and MB Gene Origination

We surveyed the foraging behaviors of several Drosophila species with or without Xcbp1. Four species within the D. melanogaster clade—D. melanogaster, D. simulans, D. sechellia, and D. mauritiana—possess Xcbp1 orthologs (Figure 4A). These species showed high FSIs that were not significantly different (Figure 7A). Three outgroup species—D. yakuba, D. erecta, and D. pseudoobscura—lack an Xcbp1 ortholog and showed significantly slower foraging speed (test for equality in slopes, p < 1 × 10⁻⁵ (Figure 7A)). These data revealed diversity in the foraging speed of closely related species with the same food cue, possibly because different species have adapted to different foraging cues (Harry et al., 1998).

We used the OK107-Gal4 line to drive RNAi of other young MB genes in MB neurons and assayed foraging. Out of eight tested, an 11- to 25-Myr-old MB gene, Drosophila Elm2-Sant retrogene (Desr, CG31875), also showed a significantly reduced FSI when knocked down in the MB (Figures 7B and S7B). Desr originated by retroposition from the short splicing isoform of the parental gene CoREST (Figures 7C and 7D), which is a transcriptional corepressor essential for nervous system development (Dallman et al., 2004). By retroposition, Desr moved into a genomic location near a neuronal gene, Bib (Figure 7C). Desr also underwent rapid protein evolution compared to its paralogs (Figures 7D). Desr recruited novel 5' and 3' exons and possibly adjacent enhancers (Figure 7C), acquiring expression in brain structures including ALs and MBs (Figure 7E). These data reveal a second, more ancient event of neuronal gene origination that influenced foraging behavior in D. melanogaster (Figures 7A and 7F). The influences of Xcbp1 and Desr on foraging in other species awaits further study, as different species may have adapted to different foraging cues in unique niches and outgroup species may have a greater foraging ability when tested with food cues most appropriate to their ecology.

DISCUSSION

The evolution of the brain and complex behaviors is an intriguing process. Although most of the basic components of synaptic vesicles and postsynaptic protein complexes are evolutionarily conserved (Jiménez and Davletov, 2007), the neuronal roles of clade-specific genes, especially those of recent origin, have rarely been studied. Our data in D. melanogaster reveal that new genes with neuronal expression have originated frequently during Drosophila evolution. During origination, many young genes either nested in or jumped close to other neuronal genes (Table S2) and might have recruited neuronal enhancers from local genomic environments (Nègre et al., 2011). Indeed, RT-PCR, RNA-seq expression profiling, and enhancer-trap expression verified neuronal expression for many genes in our data set. This finding strengthens enhancer hitchhiking by new genes (Kaessmann et al., 2009). Intriguingly, there is an enrichment of young brain genes on chromosome X compared to either young nonbrain genes or to autosomes, implying a coevolution of the sex chromosome and the brain in terms of new gene origination. Evolutionary analysis suggests that natural selection is actively engaged in the evolution of young brain genes.

We observed a relationship in the evolution of new genes and MB expression at the genomic level. These observations support the idea that young genes rapidly evolved brain expression, and MBs are hotbeds for novel gene expression in the brain. Interestingly, many young MB genes are expressed in the evolutionarily young α/β subtype of MB neurons. The

(C) Foraging curve of WT D. melanogaster strains showing the time-dependent increase in the percentage of flies that found food. Flies do not crawl to compartment 2 if the food source is not present (negative control).
(D) FSI measurements of WT D. melanogaster strains.
(E) Foraging curve for Xcbp1 RNAi in D. melanogaster with constitutive Gal4s and Xcbp1-Gal4 (pG156) or in control (WT, Gal4 only; UAS only) animals. Genotypes are shown in the inset legend.
(F) FSI measurements of control or Xcbp1 RNAi knockdown animals. Loss of Xcbp1 results in a significant decrease in foraging ability. Genotypes are shown at the bottom of each column.
(G–H) Xcbp1 RNAi knockdown animals show no significant impairment of either flight (G) or locomotion (H) when compared to controls. Genotypes are shown at the bottom of each column.
(I) FSI measurements of control or Xcbp1 MB RNAi knockdown animals. Loss of Xcbp1 in MB results in a significant decrease in foraging ability. Genotypes are shown at the bottom of each column.

In barplots, data are represented as mean ± SEM, where error bars denote SEM values; asterisks denote statistical significance (p < 0.01 unless otherwise noted) in all pairwise comparisons; n.s., not significant (p > 0.01 unless otherwise noted) in all pairwise comparisons.

(See also: Figure S5)
MBs are major control centers for a variety of neuronal functions, such as olfactory information processing, learning, memory, and foraging behavior (Davis, 2001; Krashes et al., 2007; Osborne et al., 2001). Drosophila MBs are relatively more complex than those of paleopteran and thysanuran insects (Strausfeld et al., 1998). MBs are thought to have high genetic and synaptic plasticity (de Belle and Heisenberg, 1994; Marin et al., 2002). Such plasticity might enable new genes to be expressed in the

Figure 6. Conservation and Divergence of the Xcbp1-Cnx99a Gene Pair
(A) Scheme of the XCBP1 and CNX99A proteins. N-terminal domains (NTDs), brown; transmembrane domain (TM), green; C-terminal domains (CTDs), cyan. Amino acid substitutions, deletions, and expansions are noted on the Xcbp1 diagram.
(B) FSI measurements of control, Xcbp1 RNAi knockdown, and Cnx99a RNAi knockdown animals. Genotypes are shown at the bottom of each column. Data are represented as mean ± SEM. The asterisk denotes statistical significance (p < 0.01); n.s., not significant (p > 0.01).
(C) SEM images of eye morphology of control animals (C1) or animals with Cnx99a (C2) or Xcbp1 (C3) RNAi knockdown. Yellow arrows point to normal ommatidial structures in control animals and defective or overproliferated ommatidial structures in both Cnx99a and Xcbp1 RNAi animals. Scale bars represent 100 μm.
(D) SEM images of notum morphology of control animals (D1) or animals with Cnx99a (D2) or Xcbp1 (D3) RNAi knockdown. Red arrowheads denote shorter and blunter bristles in Cnx99a RNAI that are unaffected in control or Xcbp1 RNAI flies. Scale bars represent 100 μm.

(See also: Figure S6)
Figure 7. Evolution of Foraging Behavior Driven by New Brain Gene Origination in Drosophila

(A) Phylogenetic comparison of FSIs in closely related Drosophila species that are color coded according to the phylogeny in (D).

(B) MB-specific knockdown of Desr by the OK107 driver shows a significant decrease in FSI, correlating with the increase in foraging ability observed between nodes E and D (see also D and Figure S4B). Data are represented as mean ± SEM. The asterisk denotes statistical significance (ANOVA, p < 0.01); n.s., not significant (ANOVA, p > 0.01).

(C) Scheme illustrating how the 11- to 25-Myr-old MB gene Desr (orange) retroposed from the X-linked parent CoREST (green) into a genomic location on Chr. 2L adjacent to bib (gray), and recruited new 5' and 3' exons (deep orange). Gene models and distances are not drawn to scale.

(D) Protein sequence alignment of DESR showed homology to the CoREST N-terminal protein, resulting from retroposition of the CoREST short isoform. Regional alignment of DESR to its homologs showed evidence for rapid amino acid divergence; DESR arose after the D. pseudoobscura – D. melanogaster split and was lost in D. erecta. Amino acids color coded by conservation level. Only partial representative sequences are shown for simplicity.
MB and gradually acquire beneficial functions and become integrated as new genetic components.

Changes in the brain have been reported to be associated with rapid protein evolution or cis-regulatory mutations. The evolution of language ability is associated with the FOXP2 transcription factor (Konopka et al., 2009). Increased brain complexity in human is thought to be associated with gene expansion (Pope et al., 2006) and the emergence of noncoding genes (Pollard et al., 2006). The expansions of the soluble ligand gene family in the neural crest are vertebrate innovations (Martinez-Morales et al., 2007). In Drosophila, a newborn RNA gene was related to male courtship behavior (Dai et al., 2008). We showed that new genes frequently arose, acquired expression in distinct subsets of neurons, and regulate foraging behavior in D. melanogaster. These brain genes encode many different protein domains (Table S2 and Data Set S2), suggesting that the genetic network in the brain may evolve with the origination of new genes that recombine existing protein function with novel neuronal expression.

Animal behaviors are constantly evolving (Brown, 1975; Evans, 1962), and change in existing behavioral genes (Chang et al., 2011; Fitzpatrick et al., 2007; Wheeler et al., 1991) has been suggested as an underlying mechanism. The frequent origin of new genes in the nervous system provides an alternative. The recent origin of both Xcbp1 and Desr contributed to foraging success for D. melanogaster. Foraging is critical for animal survival (de Bono and Maricq, 2005; Sokolowski, 1980) because in the natural environment failure to locate and consume food can lead to starvation and death. Foraging success is highly correlated with sensitivity to sensory cues coming from food sources (Asahina et al., 2008); mutations that confer advantages in food cue sensing and/or olfactory processing that might lead to higher foraging success could be favored and quickly get fixed. Foraging behavior is polymorphic (Fitzpatrick et al., 2007) and constantly evolving (Osborne et al., 1997; Sokolowski et al., 1997). Xcbp1 and Desr are such macromutations, although how they interact with previously identified foraging behavior genes and circuits are still open questions (Osborne et al., 2001; Root et al., 2011). We provide two examples of newly evolved brain genes that have facilitated foraging behavior in D. melanogaster. Additional components of the foraging circuit and the mechanisms underlying foraging behaviors in other Drosophila species await future investigation.

EXPERIMENTAL PROCEDURES

Identification and Expression of Young Brain Genes

Newly evolved genes in the D. melanogaster genome were curated from the literature, and their ages were verified with the use of multiple-species genomic sequence alignments. We obtained a larger data set of 886 young genes (≤ 60 Myr), and a more stringent dataset of 566 young genes (≤ 35 Myr) (Table S1). We designed primers and performed RT-PCR for a set of 330 genes that are younger than 25 Myr (Extended Experimental Procedures, Data Set S1).

Expression Profiling of Brains and MBs

Whole brains and GFP+ MB tissues were dissected from 1- to 7-day-old adult flies from the line OK107-Gal4 > > UAS-mCD8GFP with a fluorescence dissecting scope. Total RNA samples were prepared via phenol-chloroform extraction followed by QIAGEN MinElute Kit purification. Single-end RNA sequencing Library preparations and sequencing were performed with the use of the standard Illumina protocols on Solexa (Illumina). Read mapping and gene expression analysis are described in Extended Experimental Procedures.

Evolutionary Analysis

Evolutionary analyses of new-gene origination were performed as described in Vibransovski et al. (2009) and Zhang et al. (2010). Primary polymorphism data were from DPGP, and we filtered nucleotides with a Phred score < 30 as “N.” Polymorphic frequency spectra were analyzed with Polymorphorama (Haddrill et al., 2008). Estimation of α was carried out with DoFE (Eyre-Walker and Keightley, 2009).

Fly Stocks and Crosses

Enhancer trap lines used in this study were ordered from the Gal4 Enhancer Trap Insertion Database (GETDB) or the Centre de Biologie du Développement (CBD). RNAi lines for Xcbp1 (GD5597) and Cnx99a (GD1335, GD42397) were ordered from the VDRC. UAS:ShibireT1 was provided by Dr. Kitamoto. Other fly stocks used in this paper include UAS-mCD8GFP, TubP-Gal4, Elav-Gal4, OK107-Gal4, Ey/Ftp; UAS-FRT-Stop-FRT-UAS-mCD8GFP (Lee and Luo, 1999), Act5C-Gal4 (Flybase), and GMR-Gal4 (Rebay lab). WT stocks of D. melanogaster (Can-S and Oregon-R), D. simulans (MD 197), D. mauritiana (148 g), D. sechelia (Robertson), D. yakuba (Tai 6), D. erecta (151.4), and D. pseudoobscura (MV-25) were obtained from the University of California at San Diego Stock Center.

Immunohistochemistry, Confocal Imaging, and Brain Structure Scoring

Brains were dissected and stained via previously established protocols (Wu and Luo, 2006). Antibodies used include rat-anti-mCD8 (Developmental Studies Hybridoma Bank, DSHB), mouse-anti-nic82 (DSHB), goat-anti-rat-Alexa488 (Invitrogen), and goat-anti-mouse-Cy3 (Invitrogen). Samples were imaged on a Leica SP5 or Zeiss LSM510 confocal microscope. Anatomical structures were identified via methods outlined in Chou et al., 2010; Jefferis et al., 2002; Lassieu et al., 1999; Marn et al., 2002.

Foraging Behavior Assay

We measure foraging behavior under standard experimental conditions (see Extended Experimental Procedures) by using a behavioral assay system (Figure 5A) comprised of two transparent compartments linked by a 20 × 2 mm tube: the first 7 × 2.5 × 2.5 cm compartment contains no food, while the second 25 × 1.5 × 1.5 cm compartment contains a yogurt-based food mixed with fruit odors such as strawberry and banana (Halleen and Carlson, 2004a, b, 2006). A cotton plug was inserted into the end of the second compartment to allow air exchange.

Cuticle Preparation and Scanning Electron Microscope Imaging

WT and RNAi F1 adult cuticle samples were air-desiccated, mounted on metal specimen holders with double-stick carbon disks, and coated with 8.0 nm of platinum/palladium alloy. Samples were imaged with a NanoSEM scanning electron microscope (FEI) with the standard lens at 5 kV.

(E) Expression pattern of Desr-Gal4 in the brain, particularly in MB axons (top panel) and calyces (ca, middle panel), the CB (top panel), and the AL (bottom panel). (F) Xcbp1 and Desr (red stars) are shown on the Drosophila phylogeny. The node for the last common ancestor of the whole group (note E) is colored blue, while common ancestors younger than Desr but older than Xcbp1 (nodes C and D) are labeled yellow, and the common ancestors younger than Xcbp1 (nodes A and B) are labeled in green. Groups of animals with slow foraging behavior are highlighted in blue, fast foraging in green, and intermediate foraging in yellow. (See also: Figure S7)
ACCESSION NUMBERS

Sequencing data were deposited in the GEO database under accession number GSE33783.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, three tables, six data sets, and one movie and can be found with this article online at doi:10.1016/j.cell.2011.12.010.

LICENSING INFORMATION

This is an open-access article distributed under the terms of the Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 Unported License (CC-BY-NC-ND; http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode).

ACKNOWLEDGMENTS

We thank Dr. Bourbon, Dr. Tomaru, and Dr. Kitamoto for fly stocks; GETDB and CBD for enhancer trap lines; TRIP and VDRC for RNAi lines; and Dr. Charles Langley and DPGP for population genomics data. We thank Dr. Vytas Bindokas for confocal imaging, Dr. Phillipe Laisse and Dr. Leslie Vosshall for neuroanatomy, Dr. Shujuan Lu for molecular biology, Dr. Wei Du and Xiaoxi Zhuang for fly genetics and assays, and Grace Lee and Bin He for sequence analysis. We thank Dr. Urs Schmidt-Ott, Dr. Margarida Moreira, and Benjamin Kinsky for critically reading the manuscript. We thank all Long and Luo lab members for support and discussions. S.C. was supported by an NSF Doctoral Dissertation Improvement Grant (DEB-1110607); M.S. was supported by an NIH-NRSA fellowship (F31DC008277); and an NSF CAREER award (MCB0238168), and an NSF grant (MCB 1051826). K.P.W. was supported by an NIH grant (P50GM081892). L.L. was supported by an NIH grant (R01GM078070-01A1), an NSF CAREER award, and Benjamin Krinsky for critically reading the manuscript. We thank all Long and Luo lab members for support and discussions. S.C. was supported by an NSF Doctoral Dissertation Improvement Grant (DEB-1110607); M.S. was supported by an NIH-NRSA fellowship (F31DC008277); and an NSF CAREER award (MCB0238168), and an NSF grant (MCB 1051826). K.P.W. was supported by an NIH grant (P50GM081892). L.L. was supported by an NIH grant (R01DC005982) and HHMI.

Received: March 7, 2011
Revised: November 7, 2011
Accepted: December 21, 2011
Published online: February 16, 2012

REFERENCES


