Feeding-Related Traits Are Affected by Dosage of the foraging Gene in Drosophila melanogaster

Aaron M. Allen,†,* Ina Anreiter,† Ina M. Neville,‡ and Marla B. Sokolowski*†,‡

*Department of Cell and Systems Biology, University of Toronto, Ontario M5S 3G5, Canada; †Department of Ecology and Evolutionary Biology, University of Toronto, Ontario M5S 3B2, Canada; and ‡Child and Brain Development Program, Canadian Institute for Advanced Research (CIFAR), Toronto, Ontario MSG 1Z8, Canada.

ABSTRACT Nutrient acquisition and energy storage are critical parts of achieving metabolic homeostasis. The foraging gene in Drosophila melanogaster has previously been implicated in multiple feeding-related and metabolic traits. Before foraging’s functions can be further dissected, we need a precise genetic null mutant to definitively map its amorphic phenotypes. We used homologous recombination to precisely delete foraging, generating the for0 null allele, and used recombineering to reintegrate a full copy of the gene, generating the forDAG rescue allele. We show that a total loss of foraging expression in larvae results in reduced larval path length and food intake behavior, while conversely showing an increase in triglyceride levels. Furthermore, varying foraging gene dosage demonstrates a linear dose-response on these phenotypes in relation to foraging gene expression levels. These experiments have unequivocally proven a causal, dose-dependent relationship between the foraging gene and its pleiotropic influence on these feeding-related traits. Our analysis of foraging’s transcription start sites, termination sites, and splicing patterns using rapid amplification of cDNA ends (RACE) and full-length cDNA sequencing, revealed four independent promoters, pr1–4, that produce 21 transcripts with nine distinct open reading frames (ORFs). The use of alternative promoters and alternative splicing at the foraging locus creates diversity and flexibility in the regulation of gene expression, and ultimately function. Future studies will exploit these genetic tools to precisely dissect the isoform- and tissue-specific requirements of foraging’s functions and shed light on the genetic control of feeding-related traits involved in energy homeostasis.

KEYWORDS foraging gene; behavior; fat; larva; null mutant

FEEDING is critical to the development and survival of all organisms. During development, Drosophila larvae eat almost continuously. They experience rapid growth and deposit large amounts of triglycerides as energy stores. Once sufficient energy stores are reached, larvae stop feeding and alter their behavior to find pupariation sites (Edgar 2006). The size that a larva reaches and the level of stored nutrients has profound effects on survivorship, adult body size, and reproductive success (Bakker 1962). Consequently, the coordination of larval movement, feeding, and energy storage is critical to the health of both the larva and adult. This coordination requires the communication of multiple tissue systems in Drosophila such as the brain, endocrine tissues, and fat body (Leopold and Perrimon 2007). As such, perturbations in a variety of tissue systems are sufficient to alter feeding behavior.

The Drosophila melanogaster foraging gene has become a classic model for the genetic influences on feeding-related behaviors (Sokolowski 2001). The differences in locomotion on a nutritive medium (path length) between the rover and sitter strains was mapped primarily to the foraging (for) gene, also known as dg2, which encodes a cGMP-dependent protein kinase (PKG) (de Belle et al. 1989; Kalderon and Rubin, 1989; Osborne et al. 1997). foraging is highly conserved at both the sequence and phenotypic levels (Manning et al. 2002; Fitzpatrick and Sokolowski 2004; Sokolowski 2010). The rover and sitter strains were subsequently shown...
to differ in a suite of other behavioral and physiological traits, such as food intake and metabolism (Kaun et al. 2007; Kent et al. 2009), and as such foraging appears to be pleiotropic.

Much of the work conducted to date on the foraging gene relied on the rover and sitter foraging allelic variants (de Belle et al. 1989, 1993; Kaun et al. 2007, 2008). In rover and sitter strains, the entirety of their second chromosomes, where foraging resides, differ (Bauer and Sokolowski 1984; Sokolowski 1980). There is thus the potential for other loci to contribute to the phenotypic differences between these strains. To understand the contributions of the allelic variants of the foraging gene in these two strains, we need a precise understanding of the gene structure, its products, and its amorphic phenotypes. Since a genetic null allele of foraging would facilitate an examination of its potential pleiotropic effects on feeding-related behaviors in the larvae, we generated a precise deletion of the foraging gene by homologous recombination (HR; Gong and Golic 2003). To further examine the connection between genotype and phenotype we manipulated gene dosage using recombination-mediated genetic engineering (recombineering; Warming et al. 2005; Venken et al. 2006). Using these engineered allelic combinations, our analyses revealed an unequivocal role for foraging in larval movement, food intake, and energy storage. These experiments provide the solid foundation required for future research into determining tissue- and isoform-specific functions of foraging, as well as determining the causal differences between the rover and sitter allelic variants.

Materials and Methods

**Fly strains and rearing**

The rover and sitter strains have been in the lab for over 30 years and share a common first and third chromosome. They have been isogenized multiple times over the years, most recently between 2010 and 2012. The strains differ in their second pair of chromosomes where foraging resides. The path lengths of these rover and sitter strains are comparable to those collected from the field (Sokolowski 1980; Sokolowski et al. 1997) and the relative path length differences remain regardless of how the design of the path length assay has changed over the years (Anreiter et al. 2016). Fly strains for HR included y1w6 b, hsFLP, hsl-SceI/Y, hs-his and y1w6 b; eyFLP5 and y1w6 b, eyFLP2; P{Crey}1b; snasa30CyO strain (Siegal and Hartl 1996; BDSC #766) was used to resolve theloxP sequences to remove w+mc, and delete foraging. The following foraging deletion strains were obtained from Bloomington Drosophila Stock Center: w118; Df(2L)Exel7018/CyO (Parks et al. 2004; BDSC #7789), y1 w6 b; Df(2L)drm-P2, P{lacW}ND-PDSWk10101/SM6b (Green et al. 2002; BDSC #6507), and w118; Df(2L)ED243, P(3'RS+3.3')ED243/SM6a (Ryder et al. 2004; Belay et al. 2007; BDSC #24122). The foraging for4, for dup, and for Rac alleles were generated in this paper (see below); the for0 mutants were maintained over a CyO, act-GFP balancer chromosome. Strains were reared in 40-ml vials with 10 ml of food and 170-ml bottles with 40 ml of food at 25 ± 1°C with a 12L:12D photoperiod with lights on at 0800 hr. The fly yeast–cornmeal–molasses–agar food recipe contained 1.5% sucrose, 1.4% agar, 3% glucose, 1.5% cornmeal, 1% wheat germ, 1% soy flour, 3% molasses, 3.5% yeast, 0.5% propionic acid, 0.2% Tegosept, and 1% ethanol in water. Mid third instar larvae were developmentally synchronized as described in Anreiter et al. (2016). Briefly, 5- to 7-day-old adults were allowed to oviposit on grape juice and agar media (45% grape juice, 2.5% ethanol, 2.5% acetic acid, 2% agar in water) for 20 hr. The following day, any hatched larvae were cleared from the media, followed by a 4-hr incubation; newly hatched larvae were seeded into a 100-mm diameter Petri dish containing 30 ml of food. The plates were incubated at 25 ± 1°C in a 12L:12D photoperiod with lights on at 0800 hr until they reached mid third instar (72 ± 2 hr post hatch).

**Gene model characterization**

All primer design, sequence analysis, chromatograph editing, contig assembly, in-silico cloning, and digestion confirmation was performed using the Geneious 8.1.7 software package (Kearse et al. 2012). All primer sequences are listed in Supplemental Material, Table S1. Transcription start sites (TSSs) were identified with 5’-rapid amplification of cDNA ends (RACE) experiments that used homopolymeric tailing (Michelson and Orkin 1982; Sambrook and Russell 2001) and RNA ligase using GeneRacer (Thermo Fisher Scientific, cat# L150202). Total RNA was extracted from pooled mid third instar larvae and adult flies from our rover and sitter strains with TriZOL Reagent (Thermo Fisher Scientific, cat# 15596018). RNA was reverse transcribed with Superscript III (Thermo Fisher Scientific, cat# 18080044) and primed with random hexamers and oligo dT primers. Terminal Transferase (New England Biolabs, Beverly, MA, cat# M0315S) was used to add a poly-guanoisine tail to the 5’ end of the isolated cDNA. The 5’ ends of the transcripts were amplified with an oligo dC primer and gene-specific primer (comRT-R.3) targeting the coding sequence for the catalytic domain of the foraging gene. Transcription end site (TESs) were identified by 3’-RACE using the GeneRacer Kit (ThermoFisher Scientific, cat# L150202) following the manufacturer’s instructions, using a gene-specific primer (43S) targeting the catalytic domain. Splice variants were identified by RT-PCR using forward primers targeting the exons near the TSSs (TSS1-F, TSS2-F, TSS3-F, TSS4-F) and a reverse primer targeting the terminal exon (comORF3’-R). The resulting amplions were cloned into the pGEM-Teasy vector (Promega, Madison, WI, cat# A1360). 144 RACE clones and 240 splice variant clones were then sequenced by Sanger sequencing on an ABI 3130xl Capillary Sequencer using BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific, cat# 4337454).
Ends-out gene targeting

A 5-kb homology arm 5’ to the foraging gene was amplified with PCR using the primers HA1-F and HA1-R, with BstWI and Ascl sites, respectively. The 5’-homology arm was cloned into the BstWI and Ascl sites in the pW25 ends-out gene targeting vector (Gong and Golic 2004). An attP sequence (Thorpe and Smith 1998) was cloned into the KpnI and NotI sites. The attP sequence was amplified with PCR (attP-F, attP-R) from the pcaryP vector (Groth et al. 2004) obtained from Addgene. A KpnI was added to the forward primer and NotI and SbfI sites were added to the reverse primer. A 5.2-kb homology arm 3’ to the foraging gene was amplified with PCR using primers HA2-F and HA2-R, and adding SbfI and NotI sites to the forward and reverse primers, respectively. The 3’-homology arm was cloned into the SbfI and NotI sites in the pW25-HA1-attP vector resulting in the pW25EO-for-attP vector.

The pW25EO-for-attP vector was injected into a w^{118} strain using P element transgenesis (performed by Genetic Services). This yielded an X-chromosome transformant. A series of crosses were conducted to mobilize the targeting construct and to allow HR (as in Demir and Dickson 2005). Briefly, ~1000 females containing the element were crossed into an hs-Flp, hs-I-SceI-containing background, and embryos and L1 larvae were heat shocked at 38°C for 2 hr on two consecutive days. Approximately 8000–10,000 progeny were then crossed into an ey-Flp background and were screened for red eye color. These crosses yielded eight integrations at the foraging locus; however, none of these had a deletion of the gene. The recombinants had integrated at either the 5’- or 3’-end of the locus, corresponding to the two homology arms. There are several reasons why this might have occurred, for one, foraging is a large gene and the size difference relative to the considerably smaller cloned construct might have impeded proper alignment. Alternatively, after sequencing the alleles where our construct had integrated (forHR1 and forHR2), we found that the I-SceI restriction sites remained uncut causing the cloned targeting construct to not linearize, which would have prevented spanning of the endogenous locus. This could have been due to inadequate heat shock, or a less than functional hs-I-SceI transgene. Fortunately, it was still possible to generate a complete deletion of the foraging gene using the laxP sites in pW25, which were intended for removal of the w^{+mc} gene. To accomplish this, we crossed the 5’ and 3’ recombinants to make a trans-heterozygote in a hs-Cre background (BDSC #766). Embryos and L1 larvae were heat shocked at 38°C for 2 hr on two consecutive days. Single male flies were then isolated and balanced. Twenty-four iso-male populations were established and screened for a deletion of the foraging locus. Four populations were identified as deletions, for\textsuperscript{a}, and eight as duplications, for\textsuperscript{dup}. The remainder showed no recombination.

Recombineering

The P[acman] vector was redesigned to be paired with the pW25-attP vector to have a minimal footprint upon reintegration of the engineered locus of interest. The BAC backbone was isolated by digesting the P[acman] cut with SalI and SphI. The w^{+mc} was isolated by digesting the P[acman] vector with EcoRI and SphI. The attB sequence was amplified with PCR using the following primers, attB-F and attB-R. The att-B-F primer included SalI, loxP, Ascl, NotI, PacI, and AsSI sequences. The attB PCR product was cloned into the pSCA-ampl/kan (Agilent’s Stratadone) and digested with SalI and EcoRI. The purified P[acman], w^{+mc}, and the attB fragment were ligated together. This resulted in a vector with loxP, MCS, attB, w^{+mc} synteny. We called the resulting vector P[attlox]. The P[acman] clone CH321-64J02 BAC from Children’s Hospital Oakland Research Institute (http://www.chori.org) was used as the source of the foraging gene sequence. A gap repair protocol (like Venken et al. 2006) was used to trim the larger BAC down to a 39.3-kb segment containing foraging. Left and right homology arms were amplified with PCR using the following primers: LA-F with an Ascl, LA-R with a NotI, RA-F with a NotI, and RA-R with an AsSI. The left and right arms were cloned into the Ascl and AsSI sites of the P[attlox] vector creating P[GAP]for. P[GAP]for was linearized with NotI and transformed into induced SW102 cells (Warming et al. 2005) containing the CH321-64J02 BAC. Recombination between the P[GAP]for and the CH321-64J02 BAC yielded a foraging-specific BAC in the reengineered P[acman] vector, P[attlox] for. As the BAC from the P[acman] library used here was generated using DNA isolated from the y\textsuperscript{1}cn\textsuperscript{1}, bw\textsuperscript{1}, sp\textsuperscript{1} strain, containing a naturally occurring copia transposable element in the foraging gene, we used galK selection (Warming et al. 2005) to remove the copia transposable element. A 579-bp left homology arm and a 580-bp right homology arm flanking the copia element were cloned into the XhoI/EcoRI and BamHI/XbaI sites of the pGalK vector. The left arm was amplified with L-copia-F, and L-copia-R with an EcoRI site. The PCR product contained an internal XhoI site. The right arm was amplified with R-copia-F with a BamHI site, and R-copia-R. The PCR product contained an internal XbaI site. The XhoI/XbaI LA-galK-RA fragment was transformed into SW102 cells containing the P[attlox] for vector. The for-copia-galK BAC-containing cells were then transformed with a linear oligo, for-copia (Sigma Genesys). The resulting “wild-type” BAC was transformed into the EPI300 cells (Wild et al. 2002, acquired from Epicentre) and incorporated into the fly’s genome using φC31 integration into the attP2 landing site on the third-chromosomes (Groth et al. 2004). Transgenesis was done by Genetic Services.

Western blot analysis

Twenty mid third instar larvae (72 ± 2 hr post hatch) were homogenized on ice in 400 μl of lysis buffer (50 mM Tris-HCl pH 7.5, 10% glycerol, 150 mM NaCl, 1% Triton-X 100, 5 mM EDTA, 1× Halt Protease Inhibitor Cocktail (cat# 1862209)). Samples were centrifuged at 16,000 RCF at 4°C and the supernatant was transferred to a new tube and placed on ice. Protein quantification was performed with Pierce BCA Protein Assay Kit (cat# 23227). Twenty micrograms of protein were denatured for 5 min at 100°C. The samples were run on a 4%
10 wells (0.5 mm depth, 9.5 cm diameter) arranged in a plates (37 cm width, 60 cm length, 0.5 cm height) with was measured using custom black rectangular Plexiglas M mean had robust stability values (mean coef 2 reference genes were run (values (Eff(Target) variance (0.0291) and was used to calculate relative expression in other studies (Ling and Salvaterra 2011; Ponton –Tub84B had the lowest coef –Tub84B and Act5c) and both were initial selected based on stability values found in other studies (Ling and Salvaterra 2011; Ponton et al. 2011). Two reference genes were run (α-Tub84B and Act5c) and both had robust stability values (mean coefficient variance = 0.0296, mean M value = 0.0854). α-Tub84B had the lowest coefficient variance (0.0291) and was used to calculate relative expression values (Eff(Ref)/Eff(Target)) to determine differences between genotypes. Sample collection and data processing followed abbreviated MIQE recommendations (Taylor et al. 2010).

Path length
A detailed description of the foraging path length protocol is given in Anreiter et al. (2016). Briefly, foraging path length was measured using custom black rectangular Plexiglas plates (37 cm width, 60 cm length, 0.5 cm height) with 10 wells (0.5 mm depth, 9.5 cm diameter) arranged in a 2-by-5 pattern (Sokolowski et al. 1997). In the present study, mid third instar larvae (72 ± 2 hr post hatch) were randomly selected from the food plates and washed in a few drops of water. A homogenous yeast suspension (2:1 w/w) was spread across the wells creating a thin even layer in each well. Individual larvae were placed in the center of each well and covered with the lid of a 10-cm Petri dish. Larvae could move for 5 min after which the path length of each larva was traced onto the Petri dish lid. Path lengths were digitized using Fiji (Schindelin et al. 2012).

Food intake
We measured larval food intake using a new assay described here. Becton, Dickinson Falcon cell strainers (cat# 352350) were placed in 35-mm diameter Petri dishes and used to isolate and selectively feed groups of larvae. Fluorescent liquid food was prepared with 0.5% fluorescein (Sigma), 5% sucrose, and 5% yeast extraction New sentence. Third instar larvae (72 ± 2 hr post hatch) were removed from food plates, washed, and transferred in groups of 10 to cell strieners with 800 μl of liquid food. The above-mentioned Petri dish size and food volume resulted in the food being able to drain into the space below the mesh of the cell strainer, but remaining in contact with the mesh, flowing up by capillary action when it was eaten by the larvae. Once placed in the strainer, larvae were left to feed for 10 min, after which the strieners were lifted out of the food and rinsed with water, followed by three more washes with water. Washed individual larvae were placed into 0.2-ml wells of 96-well PCR plates and frozen at −20°. Frozen larvae were homogenized in 150 μl of 1× PBS with a 5/32” stainless steel ball bearing (OPS Diagnostics) and agitation using a Qiagen Tissue Lyser. Samples were centrifuged at 3500 RCF for 15 min. Twenty microliters of the supernatant was then mixed with 180 μl of 1× PBS in a fluorimeter plate (Corning, cat# 3631) and excited at 488 nm. Emission was measured at 562 nm. Homogenates from larvae fed with food containing no fluorescein were used as controls.

Triglyceride analysis
Groups of 20 larvae (72 ± 2 hr post hatch) were homogenized in 1 ml of 0.1% Tween 20 in 1× PBS. Samples were incubated at 70° for 5 min, chilled on ice for 2 min, and used for triglyceride and total protein quantification. To measure triglyceride content, 20 μl of the above supernatant was mixed with 200 μl Infinity TAG Reagent (Thermo Scientific, cat# TR22421) in a 96-well spectrophotometer plate (Corning, cat# 351172). Samples were incubated at 37° for 10 min, and absorbance was measured at 540 nm. To quantify total protein levels for standardization of the triglyceride measures, 20 μl of the above supernatant was quantified for protein content using the Pierce BCA Protein Assay Kit (cat# 23227), following the manufacturer’s instructions. BCA samples were incubated at 37° for 30 min, and absorbance was measured at 562 nm. Standard curves for triglycerides and albumin protein were used to calculate triglyceride and
protein concentration from the obtained absorbance values. Lipid levels are displayed as microgram glycerol per milligram protein.

**Statistical analysis**

Statistical analysis was performed in R (R Core Team 2013). The effects of genotype as well as blocking factors such as the date of the experiment were modeled with general linear models using the lm function, and post hoc t-tests were run to compute pairwise statistical significances. Multiple comparisons were corrected using the Holm method (Holm 1979). α of P = 0.05 was considered significant. Results of statistical analyses are provided in the figure captions. Graphs were plotted in R and edited in Inkscape.

**Data availability**

Strains are available upon request. Information in supplemental material: Table S1 contains primer names and sequences used in this study. Figure S1 describes the forIR alleles, and the generation of the forO null and fordup alleles. Figure S2 shows verification of the forO allele. Figure S3 shows construction and verification of the forBAC construct. Figure S4 shows path length, food intake and fat levels of hemizygous rover and sitter larvae. Figure S5 shows foraging gene dosage and allelic contributions to foraging gene expression in whole larvae. Figure S6 provides foraging transcript sequences. File S1 provides the figure legends for Figures S1-S6.

**Results**

**Characterizing the transcriptional complexity of the foraging gene**

Previous studies have differed in the number and variety of foraging transcripts reported (Kalderon and Rubin 1989; Osborne et al. 1997; Stapleton et al. 2002). To produce an accurate model of the foraging gene, we first exhaustively identified the transcription start sites (TSSs) and transcription end sites (TESs) represented in mRNA transcripts produced by sequencing 144 RACE clones. We identified four separate TSSs and one TES, supporting a gene model for foraging that contains four independent minimal promoters (pr1, pr2, pr3, pr4, Figure 1) whose RNA products all splice into a common 3'‐exon. A model containing multiple TSSs and one shared TES has also been shown for foraging's orthologs in other organisms (Ørstavik et al. 1997; Stansberry et al. 2001). These promoters fall into two categories: pr1, 2 and 4 are peaked, whereas pr3 is broad. Peaked TSSs are typically found in regulated promoters, whereas broad TSSs are found in constitutive promoters (Hoskins et al. 2011). All four promoters contain clear consensus core promoter Initiator (Inr) and Downstream Promoter Elements.

We next surveyed the alternative splicing complexity found in foraging mRNA. We isolated and sequenced 240 full-length cDNA clones which represented 21 distinct mRNA transcripts, encoding nine distinct open reading frames (ORFs; Figure 1). Due to the shared 3'‐end of the transcripts, these variants differ in their 5'‐UTR and the corresponding N-terminal coding sequences of their ORFs. The coding sequences for the cGMP‐binding domains as well as the kinase domain of the protein products were all located in the shared exons. Ten of the 21 transcripts are currently annotated in Flybase (Atrill et al. 2016) making 11 transcripts novel. Six of these 11 transcripts exhibit a novel instance of exon skipping in the foraging gene (RL, RM, RN, RO, RP, RS), all of which reduce N‐terminal coding sequence. Although these results suggest a more complex gene structure compared to previous reports, all the new variants identified here include the same originally described 12 exons (Kalderon and Rubin 1989).

**Generation of a novel and precise foraging null allele (forO)**

To generate a null mutation of the foraging gene we used the ends‐out gene targeting system (Gong and Golic 2004). Regions of homology on either side of the foraging gene (Figure 2A) were cloned in the pW25 vector. The intended recombination event would delete all of foraging, leaving a loxP and attP site (Figure 2B). Mobilization of this targeting construct yielded eight second chromosome integrations, none of which deleted the foraging locus as expected. However, we did find separate integration events at both the 5'‐ and 3'‐end of the locus (Figure S1). These recombination events allowed us to generate the intended null mutation of foraging (Figure 2B) by using the loxP site‐specific recombination sequences in the pW25 vector, which were intended for removal of the white+mC gene used to identify integration events (Figure S1). To accomplish this, we induced a recombination event resulting in the deletion of the entire locus by crossing the 5'‐loxP element (integrand forIR) to the 3'‐loxP element (integrand forIR), generating a trans‐heterozygous mutant in the presence of a Cre recombinase. This approach successfully generated a null allele of foraging, which we called forO, and a duplication allele, which we called fordup. PCR, Southern blotting, sequencing, and restriction enzyme analysis (Figure S2 and data not shown) confirmed that the Cre recombinants were indeed deletions of the foraging locus. Critically, RT‐PCR and western blot experiments of homozygous forO larvae did not detect any RNA or protein (Figure 2, D and E). Thus, our newly generated forO mutant is a true null allele of the foraging gene.

**Generation and genomic integration of a foraging BAC**

To manipulate foraging gene dosage, we used recombiner‐generating to generate a novel transgenic copy of the entire foraging locus that we named {forBAC}. The P[acman] library was used as a source of our {forBAC} allele (Venken et al. 2006). This BAC library was generated from the y1;cn1;bw1;sp1 strain that contained a copia transposable element within the foraging gene. Extensive DNA sequencing has shown that this copia element does not occur in any of our laboratory strains, nor have we detected it in any allele of the foraging gene other
than the reference genome line (data not shown). We successfully removed this copia element using the galK-selection method (Warming et al. 2005; Figure S3). Furthermore, the gap repair method (Venken et al. 2006) we used here yielded a BAC spanning the 35-kb foraging locus, removing the rest of the large genomic region included in BAC CHORI-64J02 (Figure S3). The resulting P[attP2] allele (Figure S3) was successfully incorporated into the fly’s genome using φC31 integrase into the attP2 landing site (Groth et al. 2004), generating the {forBAC} allele (Figure 2C). RT-PCR and western blot analyses of the {forBAC} allele showed similar expression patterns to the wild-type sitter strain (Figure 2, D and E).

**foraging BAC rescue of for0 pupal lethality**

The generation of a for0 allele allowed us to determine whether foraging is an essential gene. We found that homozygous for0 larvae died in the late pupal (pharate adult) stage. The recessive pupal lethality of homozygous for0 flies was fully rescued by a single copy of the {forBAC} allele, indicating that the 35-kb genomic fragment of the foraging gene in the BAC contained sufficient regulatory elements to fully restore the lethality of foraging null to wild-type; all adults were viable and fertile (data not shown). The survivorship of the for0 null mutants from first larval instar into the pupal stage was comparable to that of wild-type animals. The fact that late pharate homozygous for0 animals dissected from their pupal cases did not survive (data not shown) suggested that the lethality did not result from an inability to eclose from their pupal cases. Several previously published genomic deletions spanning all or part of the foraging gene were unable to complement the pupal lethality of the for0 mutant. We tested the Df(2L)Exel7018 (Parks et al. 2004), Df(2L)drmP2 (Green et al. 2002), and Df(2L)ED243 (Ryder et al. 2004) deletions, all of which failed to complement the for0 allele (data not shown). Like the for0 homozygous null, all three heterozygous combinations of these deletions with foro were pharate adult lethal.

**foraging gene dosage and allelic contributions significantly affect larval foraging path length**

Metabolic homeostasis requires a balance of energy expenditure, energy acquisition, and energy storage. Food search and navigation of a nutritional environment is vital to this process. We first examined the larval path length phenotypes of the newly generated strains that differed in the dosage of the foraging gene. As mentioned previously, the recombination event that generated the foro null allele also generated an allele with a duplication of the foraging gene locus, fordup (Figure S1). Larvae homozygous for the fordup allele showed longer path lengths than larvae with homozygous foro null alleles (Figure 3A). We next examined the larval path length for the null (foro), rescue (foro;{forBAC}), and over expresser (foro;{forBAC}). The insertion of a {forBAC} into the foro genetic background, which we call our rescue strain (foro;{forBAC}), significantly increased mean larval path lengths on yeast (Figure 3A). Furthermore, insertion of a {forBAC} into the foro genetic background (foro;{forBAC}) increased mean larval path lengths significantly more, relative to both foro;{forBAC} and foro. These results showed that increasing foraging gene expression by increasing gene dosage resulted in an increase in larval path length on yeast. We then investigated the allelic relationships between the wild-type rover (for8) and sitter (for4) alleles on larval path lengths, and asked if our gene dosage experiments can inform on the nature of rovers and sitters. As previously shown (Sokolowski 1980; de Belle et al. 1989; Osborne et al. 1997; Kaun et al. 2007), rover larvae had longer mean larval path lengths on yeast than sitters, and rover/sitter heterozygotes were indistinguishable from rover homozygotes (Figure 3C) confirming the dominance of the rover allele over the sitter allele. The rover/sitter difference in path
length was maintained when the alleles were hemizygous with the for0 allele (Figure S4A).

**foraging gene dosage and allelic contributions significantly affect larval food intake**

Once a larva has found a food source, it needs to ingest the food to acquire the nutrients needed for metamorphosis. Feeding rate can affect body size, which in turn can affect survival. We moved on to measure food intake, the next step in energy acquisition. We developed a novel fluorescence-based food intake assay. By using a liquid food medium and cell strainer dishes, larvae can be rapidly assayed and cleaned and frozen prior to quantification with a fluorometer. Foraging null larvae (for0) had lower food intake than duplication allele (fordup) larvae (Figure 4A). The null allele also had lower mean food intake compared to the rescue (for0;{forBAC}) and over expresser (fors;{forBAC}) strains (Figure 4B), suggesting that increased foraging gene dosage decreased triglyceride levels. We then assayed rovers and sitters to see if they differed in triglyceride levels, and to see what their phenotype implies about their relative expression considering the gene dosage. We found that rovers had lower mean triglyceride levels than sitters and that when the alleles were hemizygous with the for0 allele (Figure S4B). This suggested that the pattern of dominance for food intake differed from that for larval path length, with the sitter allele being dominant to the rover allele for food intake.

**Increased foraging gene dosage significantly decreases triglyceride levels**

Fat stores are critical for survival during metamorphosis. After feeding, the acquired nutrients need to be digested and absorbed. Absorbed nutrients are primarily stored as triglycerides in the fat body of the larva. With our results showing that foraging plays a role in both energy expenditure and intake measures, we then moved on to measure fat levels, a component of energy storage. We used our foraging dosage-specific allelic series to analyze lipid storage. Whole larval homogenates were incubated with the Infiinity TAG Reagent (Thermo Scientific), which contains a lipase, dissociating fatty acids from the glycerol backbone of triglycerides (as well as monoglycerides and diglycerides). Free glycerol is then colorimetrically quantified with a spectrophotometer. The for0 null larvae had higher triglyceride levels than the fordup duplication larvae (Figure 5A). Similarly, the for0 null mutants had significantly higher mean triglycerides than the rescue (for0;{forBAC}) and the over expresser (fors;{forBAC}) strains (Figure 5B), suggesting that increased foraging gene dosage decreased triglyceride levels. We then assayed rovers and sitters to see if they differed in triglyceride levels, and to see what their phenotype implies about their relative expression considering the gene dosage. We found that rovers had lower mean triglyceride levels than sitters and that
rover/sitter heterozygotes had an intermediate level of mean triglycerides (Figure 5C). This rover/sitter difference was maintained when the alleles were hemizygous with the for0 allele (Figure S4C).

**foraging gene dosage and allelic contributions to foraging gene expression**

We next set out to verify that our allelic combinations of *foraging* used to manipulate gene dosage do indeed confer differences in gene expression. We used reverse transcription quantitative PCR (RT-qPCR) to characterize the expression levels of the *foraging* gene in whole larval homogenates. We probed the common coding region of the gene to capture all known transcripts, as well as the four promoter regions (Figure 1 and Figure 6A). We identified clear differences in expression between our homozygous *foraging* null (for0), our homozygous genomic rescue (for0;forBAC), and homozygous over expresser (for1;forBAC) larvae for all regions assayed (Figure 6, B–F). As expected for0 showed no expression, for0;forBAC showed significantly higher expression than the null, and for1;forBAC showed significantly higher expression than both the null and the rescue. We saw no difference between homozygous rovers, sitters, and rover/sitter heterozygotes, but hemizygous rovers and sitters (for0/for0 and for0/for0) had half the expression of *foraging* (Figure S5). Osborne et al. (1997) showed using northern analyses that the *foraging* mRNA levels in whole adult rovers were higher than in sitters. In contrast, in the present study using RT-qPCR we assayed *foraging* mRNA levels from 72-hr post hatch whole larvae and found no differences in expression between rovers and sitters. The effects of the different allelic combinations were similar across the different promoters and were consistent with the pattern seen for the common coding region. The only exception to this pattern was the expression of pr1, which had lower expression in the (forBAC) allele relative to the for1 allele, suggesting that the expression pattern of our rescue line differs in pr1 expression compared to wild type (Figure 6B). There were, however, large differences in the level of expression between the different promoters. Notably, pr1 and pr3 made up most of the *foraging* expression (Figure 6, B and D, respectively). Conversely, pr2 and pr4 only represent a small fraction of the total *foraging* expression (Figure 6, C and D, respectively). The signal seen with the pr3 primer set also captured some of the isoforms produced from pr1 and pr2 since the transcript RB has no unique element and is nested within RA, RH, and RI. These data would suggest that in the whole larva, pr1 and pr3 contribute roughly equally to the overall expression levels seen from *foraging* and they represent most of the expression, with pr2 and pr4 having overall low expression levels, or are expressed in only a small subset of *foraging*-expressing cells.

**Modeling the effect of foraging gene dosage on path length, food intake, and triglyceride phenotypes**

We then modeled the continuous effect of gene copy number on the path length, food intake, fat levels, and gene expression phenotypes. We performed linear regressions on the data (Figure 3B, Figure 4B, Figure 5B, and Figure 6F), but with the genotype variable exchanged for gene dosage number (null = 0, rescue = 2, over = 4). The phenotypes were normalized to rescue strain and are plotted as a percent. From
these regression lines (solid lines) and their 95% confidence intervals (faded boxes), we can clearly see the linear gene dosage effect of the foraging gene on the phenotypes of path length, food intake, and triglyceride levels (Figure 7A–C). The effect of gene dosage on foraging gene expression is added for comparison (Figure 7D). The regressions were highly significant, and had $R^2$ values of between 0.3 and 0.4 for path length, food intake, and fat, but naturally much higher for gene expression, $R^2$ = 0.98 (Figure 7A–D). Aside from gene expression, food intake had the steepest slope and fat levels had the shallowest (Figure 7E). Therefore, increasing foraging gene expression had a larger relative effect on food intake than it did on path length (food intake vs. path length, $t_{(230)} = 3.21$, $P = 1.46e−3$) or fat levels (food intake vs. fat, $t_{(280)} = 11.8$, $P = 1.99e−26$).

Discussion

Meeting energy requirements and maintaining energy homeostasis is critical for every organism. The coordination of food searching, food acquisition, and nutrient storage is necessary for meeting the challenges of changing environments. Selecting for one of these traits has been shown to have knock-on effects on the others (Sewell et al. 1974; Burnet et al. 1977; Joshi and Mueller 1988; Sokolowski et al. 1997). The extensive behavioral literature on the foraging gene serves as an important example of how allelic variants can be used to infer the functions of a gene. Much of the work conducted to date on the foraging gene relied on the rover and sitter allelic variants (de Belle et al. 1989, 1993; Kaun et al. 2007) and studies using cDNA constructs driving expression beyond wild-type levels (Osborne et al. 1997; Mery et al. 2007; Kaun et al. 2008; Burns et al. 2012). However, it is important to note that this approach may only uncover a subset of foraging’s functions and does not allow for a genetic dissection of the relationship between them.

As a first step toward a comprehensive characterization of the foraging gene, we carried out an exhaustive molecular analysis, defining the transcriptional complexity of foraging expression in larvae. We have defined four independent promoters, producing 21 mRNA transcripts between them, 11 of which are novel. This complexity lends support to the idea that foraging’s gene products might be differentially regulated to produce its multiple phenotypes. All foraging transcripts encode ORFs with a common cGMP-binding and kinase domain; however, isoforms notably differ in their N-termini. Furthermore, among the novel transcripts, there is a class that exhibits exon skipping, relative to the other transcripts produced from the same promoter. This exon skipping results in a roughly 300–500 (depending on the transcript) amino acid truncation at the N-termini, which includes both the autoinhibitory domain and dimerization domain. Interestingly, the N-termini of PKGs are critical in determining PKG-substrate interaction (Pearce et al. 2010). So, foraging’s function not only depends on which cell a promoter drives expression, but also what substrate foraging’s proteins can bind.

A precise genetic null allele allows us to conclusively associate foraging’s phenotypes, and lay the foundation for future genetic dissection of foraging’s isoform- and tissue-specific functions. In the present study, we precisely deleted the entire 35-kb foraging gene using HR to generate a for0 null mutant, and we generated a transgenic (forRBC) allele using recombineering. We could study the resulting loss-of-function phenotypes in D. melanogaster larvae despite the pupal lethality of the homozygous for4 null mutant. We have shown that foraging plays a critical role in each of the feeding-related phenotypes studied here. Increasing foraging gene copy number, both by insertion of the (forRBC) allele in a wild-type background, as well as the duplication at the locus itself, increased both larval path length on yeast and food intake while reducing triglyceride levels.

We have provided compelling evidence for a causal relationship between the foraging gene and its three larval phenotypes, using gene dosage. The null mutant generated here also provides a basis for conditional rescue using transcriptspecific cDNA lines, and tissue-specific GAL4 drivers. These tools will allow the rescue of the amorphic phenotype with spatial- and temporal-specific expression using the various foraging isoforms. Mapping the promoters and their transcripts involved in foraging-related phenotypes will allow us to narrow down DNA sequences in the locus that contain putative regulatory elements. This will in turn make it possible to select candidate SNPs between rovers and sitters that might be driving differences seen in expression and behavior. The fact that our (forRBC) construct rescues the null phenotypes, suggests that the 35-kb locus of the foraging gene contains the required regulatory elements to drive the expression
necessary for viability, path length, food intake, and triglyceride levels. We have sequenced the foraging gene of rovers and sitters and found that there are 300 SNPs that differ between these lines. The vast majority of these differences lie in the noncoding region of the gene, suggesting key allelic differences are likely associated with the regulation of gene expression rather than the modification of the function of the kinase itself. Nevertheless, this number of SNPs is far too extensive to address each SNP as a candidate causal SNP. Thus, narrowing down putative regulatory regions will be the first step to address the underlying DNA differences that drive rover and sitter behavior.

We quantified the expression of the individual promoters of foraging in whole larvae and saw striking differences in expression levels. These different expression levels show that the isoforms of foraging are differentially regulated. This differential regulation may result in different isoforms being expressed in different tissues at different times, and at different levels. Microarray data from Fly Atlas and RNA-Seq data from modENCODE support this possibility. foraging is expressed in many tissues (Chintapalli et al. 2007) throughout development (Graveley et al. 2011). Furthermore, isoform-specific exons are differentially expressed in different tissues (Brown et al. 2014). foraging's function may be tissue-specific or distributed across many tissue systems for any given phenotype. foraging's tissue-specific requirements for the larval phenotypes studied here are currently not known.

Whether foraging's role in these phenotypes is due to acute or developmental functions of its gene products is also unknown. Formation of the proper circuits for feeding behavior are developmentally regulated (Friedrich et al. 2016), and foraging mutants have been shown to alter nervous system development in the embryo (Peng et al. 2016). It is possible that a portion of foraging's function in the larva may be due to such a role in embryo development. Alternatively, foraging may be functioning acutely during the behavior to elicit its phenotypic effects. hugin neurons, or their targets, are candidate cells for foraging function since acute manipulation of these cells can alter larval locomotion and feeding (Schoofs et al. 2014). A lack of foraging in D. melanogaster could also affect larval path length or feeding due to altered muscle function, making the muscles less sensitive to incoming neuronal stimuli. foraging's ortholog in mammals has been characterized for its role in muscle function (Pfeifer et al. 1998;
As for the triglyceride level phenotype, foraging is highly expressed in the fat body of the larva (Chintipalli et al. 2007; Brown et al. 2014), and may function locally in the fat body to alter triglyceride levels. Foraging's orthologs have previously been shown to affect fat levels (Raizen et al. 2006; Miyashita et al. 2009). Future studies will unravel foraging's tissue-specific requirements for larval path length, food intake, and triglyceride levels.

The use of a kinase such as foraging that can regulate the activity of many targets might allow the coregulation of physiologically related phenotypes such as path length, food intake, and fat in response to external inputs or internal physiological state. This would be important since behavioral responses are complex, originating from the interaction of different cellular processes in different tissues (e.g., changes in movement on food might be brought on by an interaction of fat metabolism, nervous system, and muscle activation), and need to be coordinated. Without knowing the precise tissues or cells where foraging expression is required to alter its associated behaviors, it is difficult to speculate on the nature of the gene networks in which it resides. That being said, previous genetic evidence has suggested possible intergenic interactions with insulin-signaling and foxo-signaling (Kent et al. 2009; Kanao et al. 2012).

Foraging has been implicated in a variety of adult phenotypes (Sokolowski 2010). Given the pupal lethality seen in the foraging null mutant, it is possible that some of the previously associated adult phenotypes may be developmental in origin, possibly resulting from differential neurogenesis. In the present study, the pupal lethality of the foraging null restricted its use to larval rather than adult phenotypes. We have recently, however, overcome this limitation by using temperature-sensitive foraging RNAi conditional-knockdown, specifically during adulthood (data not shown). These results along with the results of the present study show that foraging is not required for viability specifically in the larval and adult stages, but is required for viability during pupal development.

In the present study, we described and characterized the structure of the foraging gene. We generated a foraging null mutant, for0, using HR. This mutant was pupal lethal, showing that foraging gene products are not required for viability during larval development. This allowed us to evaluate the feeding-related behavior and metabolic state of the for0 null larvae. We have shown that foraging influences how larvae navigate their food environment, how they ingest nutrients, and further deposits those nutrients into fat. Our deletion study combined with the genomic rescue experiments have unequivocally proven a causal, dose-dependent relationship between the foraging gene and its pleiotropic influence on these feeding-related traits. Future studies will shed further light onto foraging's regulation and the nature of the rover and sitter alleles.

Acknowledgments

We thank Scott Douglas, Jean-Christophe Billeter, and Joel Levine for guidance and advice; the Sokolowski Lab for helpful discussions; Jeff Dason, Stephen Goodwin, and the two anonymous reviewers for comments on the manuscript; and the Bloomington Drosophila Stock Center for fly lines. This research was supported by grants from the Natural Sciences and Engineering Council of Canada (NSERC) and the Canadian Institutes for Health Research (CIHR) to M.B.S.

Literature Cited


Fitzpatrick, M. J., and M. B. Sokolowski, 2004 In search of food: exploring the evolutionary link between cGMP-dependent protein kinase (PKG) and behaviour. Integr. Comp. Biol. 44: 28–36.


Communicating editor: M. F. Wolfner