

Duplication and Sub/Neofunctionalization of *Malvolio*, an Insect Homolog of *Nramp*, in the Subsocial Beetle *Nicrophorus vespilloides*

Elijah C. Mehlferber,^{*,1} Kyle M. Benowitz,^{†,2} Eileen M. Roy-Zokan,^{†,3} Elizabeth C. McKinney,^{*,†} Christopher B. Cunningham,^{†,4} and Allen J. Moore^{*,†,5}

^{*}Department of Entomology, and [†]Department of Genetics, University of Georgia, Athens, Georgia 30602

ORCID IDs: 0000-0002-9958-3247 (E.C.M.); 0000-0002-7685-1827 (K.M.B.); 0000-0002-1657-1073 (E.M.R.-Z.); 0000-0003-2426-5567 (E.C.M.); 0000-0003-0178-1833 (C.B.C.); 0000-0002-1498-3322 (A.J.M.)

ABSTRACT With growing numbers of sequenced genomes, increasing numbers of duplicate genes are being uncovered. Here we examine *Malvolio*, a gene in the natural resistance-associated macrophage protein (*Nramp*) family, that has been duplicated in the subsocial beetle, *Nicrophorus vespilloides*, which exhibits advanced parental behavior. There is only one copy of *Mvl* in honey bees and *Drosophila*, whereas in vertebrates there are two copies that are subfunctionalized. We first compared amino acid sequences for *Drosophila*, beetles, mice, and humans. We found a high level of conservation between the different species, although there was greater variation in the C-terminal regions. A phylogenetic analysis across multiple insect orders suggested that *Mvl* has undergone several independent duplications. To examine the potential for different functions where it has been duplicated, we quantified expression levels of *Mvl1* and *Mvl2* in eight tissues in *N. vespilloides*. We found that while *Mvl1* was expressed ubiquitously, albeit at varying levels, expression of *Mvl2* was limited to brain and midgut. Because *Mvl* has been implicated in behavior, we examined expression during different behavioral states that reflected differences in opportunity for social interactions and expression of parental care behaviors. We found differing expression patterns for the two copies, with *Mvl1* increasing in expression during resource preparation and feeding offspring, and *Mvl2* decreasing in these same states. Given these patterns of expression, along with the protein analysis, we suggest that *Mvl* in *N. vespilloides* has experienced sub/neofunctionalization following its duplication, and may be evolving differing and tissue-specific roles in behavior and physiology.

KEYWORDS

behavior
burying beetle
gene expression
parenting
phylogenetics

The process of gene duplication is one of the primary mechanisms hypothesized to play a role in the evolution of novel phenotypes (Ohno 1970; Ditmar and Liberles 2010; Innan and Kondrashov 2010; Wagner 2011). When duplicate genes are maintained, one copy often becomes free to mutate and acquire new functions, as it is no longer constrained by the selective pressure to perform its previous role (Ohno 1970; Maere and Van de Peer 2010). This process can take two nonexclusive paths: subfunctionalization or neofunctionalization (Nadeau and Sankoff 1997; Nowak *et al.* 1997; Wagner 1998; Force *et al.* 1999; Maere and Van de Peer 2010). In the former process, both genes lose a portion of their function, so that the two duplicated genes together recapitulate the function of the ancestral gene (Force *et al.* 1999; Maere and Van de Peer 2010). In the latter process, one duplicate evolves a novel function absent from the ancestral gene (Ohno 1970; Maere and Van de Peer 2010). Neofunctionalization may also arise following subfunctionalization (Maere and Van de Peer 2010). Given the proposed role that gene

duplication has in the production of new phenotypes, it follows that more derived organisms with novel traits will provide good systems for investigating the divergence of duplicated genes. Genetic influences on behavior may require neofunctionalization of gene duplications to overcome constraints that would otherwise arise through pleiotropy. For example, G-protein coupled receptors are cell surface receptors where gene duplications have evolved to become diversified and specialized for different behaviors (Katz and Lillvis 2014). Gene duplication and neofunctionalization have been implicated in the evolution of insect behaviors as diverse as *vitellogenin*'s influences on ant queen and worker social behavior and tasks (Corona *et al.* 2013), olfactory receptors related to shifts to herbivory (Goldman-Huertas *et al.* 2015), and opsin genes related to color vision and foraging preferences (Feuda *et al.* 2016).

With the advent of improved bioinformatics and sequencing, we are acquiring information on genomes of nonmodel organisms at an

accelerated pace, many of which are studied primarily for their derived novel traits and not for their genetic accessibility. Such genomes often reveal duplicated genes that may have previously been unknown. We recently sequenced, assembled, and annotated the genome of such an organism, the burying beetle *Nicrophorus vespilloides* (Cunningham *et al.* 2015b). Burying beetles (*Nicrophorus* spp.) are unusual among beetles for their parenting behavior. Burying beetles breed on vertebrate carrion, which they shape into a ball, prepare with antimicrobial secretions, and bury. The larvae then crawl onto the carcass and one or both parents care for the offspring. Parenting in burying beetles is more than provisioning food for developing offspring, as in many insects; it involves direct and extensive prolonged social interactions. Burying beetles not only prepare and maintain a carcass for food, they feed their offspring by regurgitating predigested carrion directly into their mouths. Parenting in this taxon is thus both complex and extensive, and strongly selected as it influences the fitness of offspring (Eggert *et al.* 1998; Lock *et al.* 2004). This behavior is highly derived, and therefore we predicted that sub/neofunctionalization could be important in the evolution of complex parenting. Therefore, we examined this genome for evidence of duplications.

A candidate duplicate gene we found in the genome of *N. vespilloides* is *Malvolio*, a transporter of divalent cations and homolog of the *Nramp* (natural resistance-associated macrophage proteins) family in vertebrates (Folwell *et al.* 2006). *Malvolio* has been ascribed a role in behavior as well as cation transport (Evans *et al.* 2001). Moreover, there is evidence that *Nramp1* and *Nramp2* have subfunctionalized in mammals and fish (Techau *et al.* 2007; Neves *et al.* 2011). In fish, for example, *Nramp1* has been lost but a duplication of *Nramp2* and subsequent subfunctionalization recovers the primary roles of iron transport and defense against pathogens (Neves *et al.* 2011). These roles are achieved by one form with ubiquitous expression and the other copy with localized expression, primarily in the immune cells and neuronal cells (Searle *et al.* 1998; Evans *et al.* 2001; Techau *et al.* 2007; Neves *et al.* 2011). However, in the best-studied insects, honey bees and *Drosophila*, there is only one copy of *Malvolio*. This gene is ubiquitously expressed and believed to function in many of the same roles as the human homolog of *Nramp2*, but is also known to effect behavior (Rodrigues *et al.* 1995). *Mvl* influences the transition between nurse and forager roles in honey bees (Ben-Shahar *et al.* 2004; Søvik *et al.* 2015) and food choice in *Drosophila* (Orgad *et al.* 1998; Søvik *et al.* 2017). Given the role this gene plays in behavior, and its duplication in *N. vespilloides*, we

hypothesize that it has undergone sub/neofunctionalization and plays a role in the unique social behavior of *N. vespilloides*.

To begin our investigation into whether the duplication of *Malvolio* in beetles facilitated sub/neofunctionalization in relation to its effects on behavior, we first examined *Malvolio* protein sequences across mammals and insects. We then built a gene phylogeny to determine the evolutionary history of duplication in this gene in insects. Next, to ask whether this gene displays behavior consistent with sub/neofunctionalization, we measured gene expression of *Mvl1* and *Mvl2* across eight tissue types in *N. vespilloides*. Finally, we examined expression of the two *Malvolio* copies in head tissue collected from beetles before, during, and after parenting, and found changes in expression during parenting in opposite directions for the two copies. We further compared this expression to RNA-seq data available for *Drosophila melanogaster*. We suggest that the *Malvolio* duplicates in *N. vespilloides* are in the process of evolutionary divergence, with neofunctionalization as a possible endpoint.

MATERIALS AND METHODS

Identification and comparison of protein and gene sequences

To verify the putative duplication that we found in the genome, we searched for putative *N. vespilloides* *Malvolio* homologs using BLASTp [v2.3.0+; default search settings; Camacho *et al.* (2009)] with *D. melanogaster* (NP_524425.2) and *Tribolium castaneum* (XP_967521.1) *Mvl* sequences. We obtained sequences from the National Center for Biotechnology Information (NCBI) or UniProt databases. We used BLAST on these *Mvl* sequences against the proteome produced from the annotated *N. vespilloides* genome (Cunningham *et al.* 2015b). Further identification of putative *Mvl1* and *Mvl2* of *N. vespilloides* was done using BLASTp with default settings into NCBI's nonredundant insect protein database, and by BLASTing (BLASTp, default settings) them into *D. melanogaster* and *T. castaneum* proteomes alone to establish if all of sequences were reciprocal best BLAST (RBB) hits for each other.

To visualize protein conservation across both *Mvl* copies, we aligned protein sequences from *N. vespilloides*, *T. castaneum*, *D. melanogaster*, *Homo sapiens*, and *Mus musculus* using ClustalW and produced box shade plots with the MobyE@Pasteur web portal (<http://mobile.pasteur.fr>). We used NCBI protein BLAST with default settings to determine the percent similarity between *N. vespilloides* *Mvl1* and *Mvl2*, and *D. melanogaster* *Mvl*. The NCBI BLAST used both the sequences for *Mvl1* (XP_967521.1) and *Mvl2* (XP_973779.1) from *T. castaneum*. Protein sequences were then aligned using Clustal Omega (McWilliam *et al.* 2013), and a model test was performed in MrBayes v3.2 (Ronquist *et al.* 2012) to determine the most appropriate model of protein evolution, which was WAG (Whelan and Goldman 2001).

For the phylogenetic analysis, we included all insect *Mvl* sequences and *Mvl* proteins we could identify from NCBI, with the exception that we did not include every *Drosophila* spp. Due to the large number of published *Drosophila* genomes, and to avoid redundancy, we only included *D. melanogaster*. To provide a representative sample of insect species we searched for *Mvl* in Lepidoptera, but there are currently no assembled and annotated genomes of the order that contain a copy of *Mvl*. Thus, our analysis includes all available insect orders. A Bayesian phylogenetic analysis was conducted in MrBayes for 5,000,000 generations with a sample frequency of every 100 generations. The consensus tree was compiled after discarding the first 25% of trees sampled, and the resultant tree was rooted with human and mouse *Nramp1* and *Nramp2*, and a *Crassostrea gigas* *Mvl* outgroup. We collected all unique *Mvl* isoform sequences from NCBI for *N. vespilloides*, *T. castaneum*,

Copyright © 2017 Mehlferber *et al.*

doi: <https://doi.org/10.1534/g3.117.300183>

Manuscript received July 17, 2017; accepted for publication August 16, 2017; published Early Online August 22, 2017.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material is available online at www.g3journal.org/lookup/suppl/doi:10.1534/g3.117.300183/-/DC1.

¹Present address: Integrative Biology, 3040 Valley Life Sciences Bldg. #3140, University of California, Berkeley, CA 94720-3140.

²Present address: Department of Entomology, University of Arizona, Tucson, AZ 85721.

³Present address: Plant Genome Mapping Laboratory, University of Georgia, Athens, GA 30602.

⁴Present address: Department of Biosciences, Swansea University, SA2 8PP Swansea, UK

⁵Corresponding author: College of Agricultural and Environmental Sciences University of Georgia 109 Conner Hall 147 Cedar Street Athens, GA 30602-7503. E-mail: ajmoore@uga.edu

and *D. melanogaster* and used NCBI protein BLAST at default settings to determine variation between isoform amino acid sequences.

Comparison of gene expression

We maintained *N. vespilloides* as an actively outbred colony at the University of Georgia. We founded the colony with beetles collected from the wild near the University of Exeter, Cornwall, UK, and new wild individuals were introduced to the colony yearly to maintain genetic variation. We isolated individuals as larvae and housed them individually in 4 × 7 cm biodegradable circular deli containers (Eco products, Boulder, CO) filled with 2.5 cm of moist soil (FoxFarm, Samoa, CA). Individuals were kept in an incubator (Percival Scientific, Perry, IA) set at 22 ± 0.1°, under a 15:9 light:dark cycle. Upon reaching adulthood they were fed two decapitated mealworms (*Tenebrio*) once a week.

For the comparison of expression across different tissues, we collected eight tissues—brain, fat bodies, hindgut, midgut, thoracic musculature, Malpighian tubules, testes, and ovaries—from five virgin female beetles at 26–30 d post adult eclosion (testes came from five males of the same age and rearing conditions). These same tissues types were previously examined for octopamine expression in Cunningham *et al.* (2015a), except for testes, but on separate tissue collections. We dissected beetles in ice cold PBS, starting with the brain and then moving on to the internal organs. We cleaned fat and connective tissue from each organ and placed them in separate 1.5 ml vials with 300 µl of RNAlater (Applied Biosystems, Foster City, CA) on ice. Dissection times for brains were 10 min or less, and the total time of dissections was <30 min. After dissection, we stored organs overnight at 4° and then moved them to –20° until RNA extraction. See Cunningham *et al.* (2015a) for further details.

We collected whole heads from 10 individuals in each of five behavioral states to examine changes associated with changes in behavior: virgins, individuals mated but not provided with the resources necessary to breed, mated individuals provided with a mouse carcass to prepare and that stimulates egg laying, individuals actively caring for and provisioning food to begging offspring, and individuals that had completed parental care and had dispersed away from the carcass and larvae. See Roy-Zokan *et al.* (2015) and Cunningham *et al.* (2016) for further details.

RNA was extracted using a Qiagen RNeasy micro kit (Qiagen, Venlo, The Netherlands) for the brain tissue and larval hemolymph and a Qiagen RNeasy lipid kit for all other tissue. The extractions were performed with 350 µl QIAzol (Qiagen) as the lysis buffer and 150 µl chloroform (J.T. Baker, Center Valley, PA). DNA was removed using DNase I (Qiagen) according to manufacturer's instructions. After the final RNA product was obtained, it was quantified with the Qubit 2.0 fluorometer (Qubit Systems, Kingston, ON, Canada) according to manufacturer's instructions. The RNA was then stored until the time of cDNA production in a freezer set to –80°. cDNA was created using 500 ng total RNA and the Quanta Biosciences qScript reverse transcriptase master mix (QuantaBio, Beverly, MA) following the manufacturer's instructions. The RNA template was then eliminated using RNase H (New England BioLabs, Ipswich, MA) and the single-stranded cDNA was quantified using the Qubit 2.0 fluorometer according to manufacturer's instructions. The resulting cDNA was then stored at –20°.

Using the two *Malvolio* gene sequences, eight primer pairs (four primer pairs per gene) were produced by utilizing Integrated DNA Technology (IDT, Coralville, IA) and Primer 3 v.4.0 (Koressaar and Remm 2007; Untergasser *et al.* 2012). These primer pairs were then validated by estimating PCR efficiency and observing the number of amplicons generated by each pair. The primer efficiency was determined by running a qRT-PCR reaction with stock cDNA (produced

using same methods as experimental cDNA from whole-body samples) diluted to 1:4, 1:16, 1:64, 1:256, and 1:1024 concentrations, while amplicons were observed in the Melt Curve Analysis. These primer pairs had efficiency levels of 1.805 (*Mvl1*) and 1.7852 (*Mvl2*).

The quantification of gene expression was accomplished using a qRT-PCR reaction with the Roche LightCycler 480 using Roche LightCycler 480 SYBR I Green Master Mix (Roche Applied Science, Indianapolis, IN). Each biological replicate ($N=5$) was run with three technical replicates, using 10 µl reactions containing 5 µl SYBR mix, 2 µl of 1.5 ng/µl cDNA, and 3 µl of an equal mixture of forward and reverse primers at 1.33 µmol/liter each. The LightCycler was run according to manufacturer's instructions for the enzyme activation step, followed by 45 cycles of amplification at 60° and a disassociation curve step to measure the number of amplicons produced in the reaction. Each reaction included three primers: *Mvl1*, *Mvl2*, and TATA-binding protein as the reference gene (*Mvl1*- forward: CGACGATGACGGGAA CTTATG reverse: TTGCGATGGATCTGGTGAAG *Mvl2*- forward: GGTATCGTGGGAGCAGTTATC reverse: GCTGCTCTCGATGAG GTAATAG *tbp*- forward: CACCATGACTCCAGCAGAT reverse: ACGTGCATGCAGAGCTATCTT).

Statistical analysis of gene expression

We analyzed the log of the relative expression differences in *Mvl1* and *Mvl2*, where relative expression was quantified as $2^{-\Delta\Delta CT}$ and was relative to expression in ovaries, where expression was negligible for both genes. We made comparisons among the tissues types in JMP Pro 13 (SAS Institute, Cary, NC) using ANOVA on log-transformed relative expression, which was more normally distributed than relative expression, with specific pairwise comparisons made using Fisher's least significant difference (LSD) test.

Comparisons of expression of *Mvl1* and *Mvl2* across different behavioral states were made as described in Roy-Zokan *et al.* (2015) and Cunningham *et al.* (2016), using $2^{-\Delta\Delta CT}$ with relative expression standardized to virgins. We used virgins as the comparison as this is the behavioral/physiological state of individuals used in the tissue comparison, and that we have used in previous studies as a nonsocial state. ANOVA on log-transformed relative expression was used to determine statistically significant changes in expression.

To provide comparative context for our results on expression in different behavioral states and tissues, we collected *D. melanogaster* RNA-seq tissue data from FlyBase (FBref0221009) (Gelbart and Emmert 2013). We then averaged normalized expression across each exon to produce one value for each tissue and behavioral state. These data were then visualized in JMP Pro 13, allowing descriptive comparisons with our expression patterns in *N. vespilloides*.

Data availability

All data and reagents are available on request. Data were deposited in Dryad (doi: 10.5061/dryad.110qd). All accession numbers for sequences used in the phylogenetic comparison are available in Supplemental Material, File S1.

RESULTS

Phylogenetic analysis of Malvolio across insects

Box shade plots illustrating sequence homology between *N. vespilloides*, *T. castaneum*, *D. melanogaster*, *H. sapiens*, and *M. musculus* indicate a high level of conservation between the different species, especially in the transmembrane regions (Figure 1). The first external loop and the consensus transport motif are highly conserved as well. The primary differences between the proteins are observed in the N-terminal and

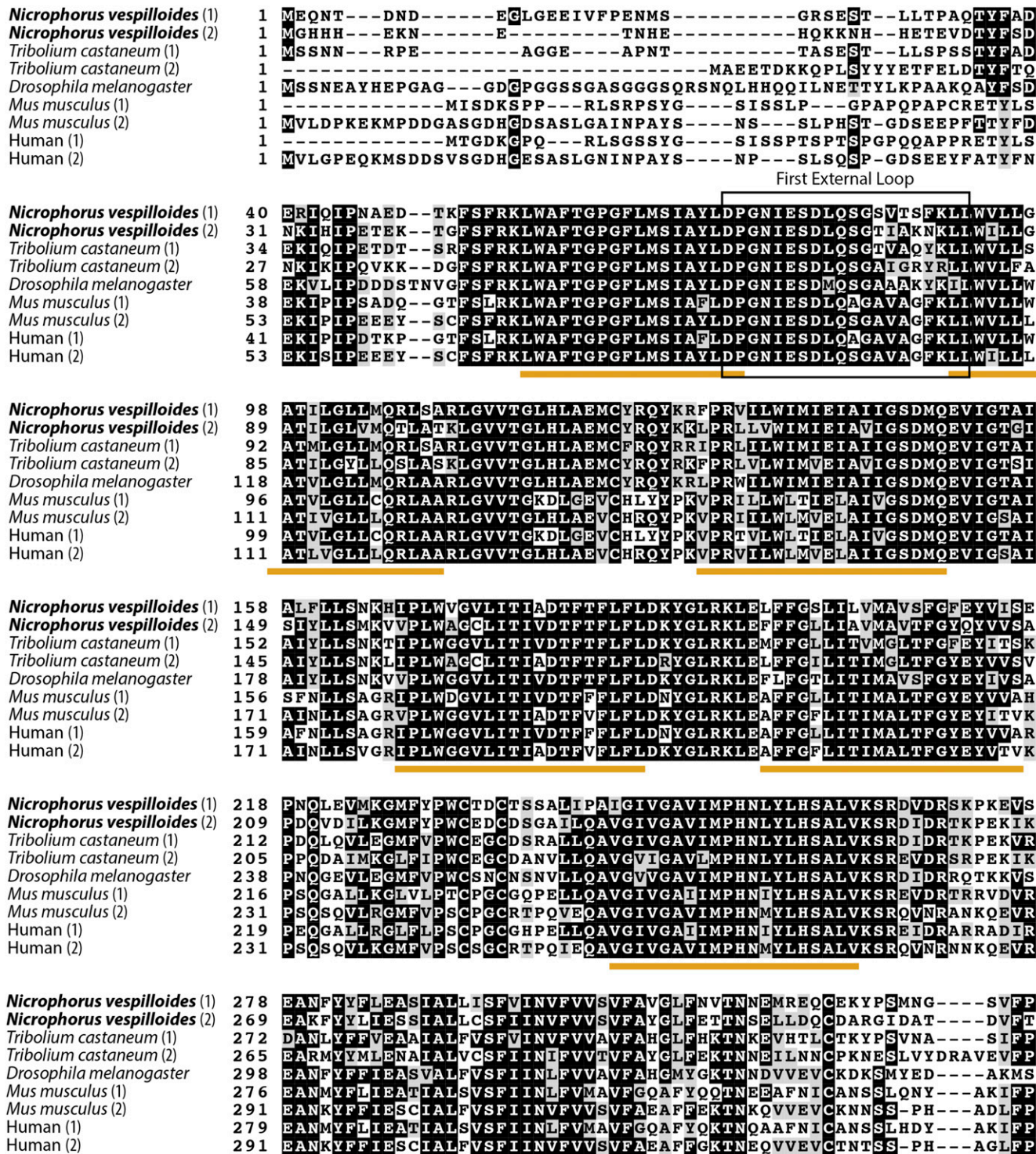


Figure 1 Amino acid alignment of *N. vespilloides* Mvl1 and Mvl2, *T. castaneum* Mvl1 and Mvl2, *D. melanogaster* Mvl, and the homologous *M. musculus* and human Nramp1 and Nramp2. Shaded regions represent a >50% similarity among sequences. Underlined sequences indicate putative transmembrane domains, while other regions of interest are indicated by labeled boxes. Specific amino acids of interest are highlighted and labeled with a letter. A, B, and C represent amino acids that indicate symporter or antiporter activity, with the following patterns; A:S indicates symporter, A indicates antiporter, B:N or D indicates symporter, K or R indicates antiporter, C: LY indicates symporter, and YF indicates antiporter. D and E are conserved amino acid positions that are part of the intracellular localization motif, which determines the intracellular localization of the protein in humans (Tabuchi et al. 2002).

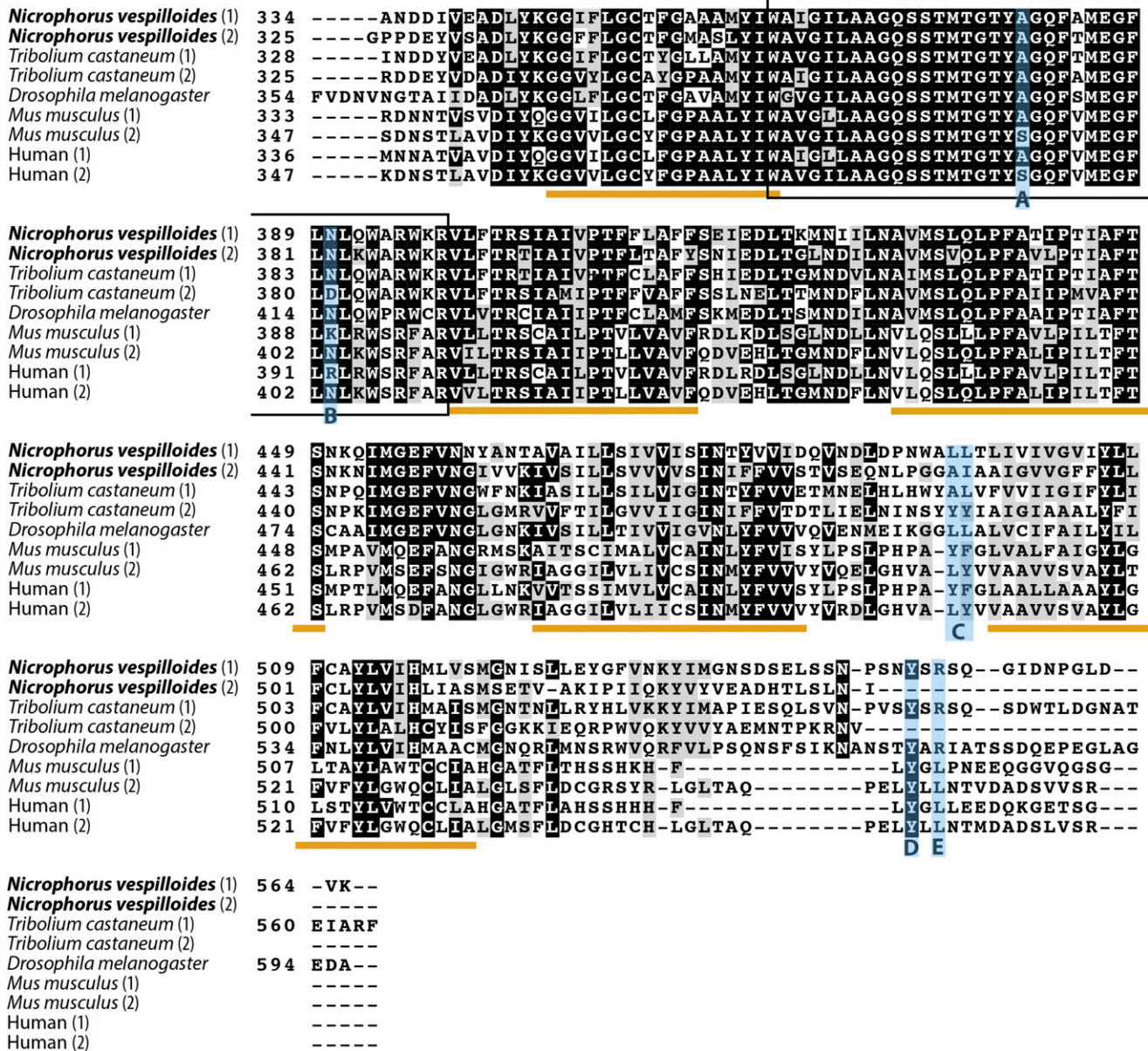


Figure 1 Continued.

C-terminal ends. Overall, *N. vespilloides* Mvl1 shared 66% similarity and Mvl2 shared 64% similarity with *D. melanogaster* Mvl.

N. vespilloides Mvl1 has two isoforms differing in only the last several amino acids at the C-terminal end. The intracellular localization motif, which in humans is composed of the amino acids located at positions 555 (D on Figure 1) and 557 (E on Figure 1) and has been implicated in the intracellular localization of Nramp2, marks the beginning of the alternative splice site. The amino acid at the insect equivalent of human position 555 is always a Y; meanwhile, the amino acid at the equivalent of position 557 differs between the two isoforms, with isoform X1 having an R and isoform X2 having an S. This pattern is also seen in the isoforms of *T. castaneum* Mvl1 and *D. melanogaster* Mvl, with each isoform having a Y at the 555 position, and either an R or an S at the 557 position. Neither *N. vespilloides* nor *T. castaneum* Mvl2 has different isoforms.

Our phylogenetic analysis (Figure 2) shows that *Mvl* has undergone several independent gene duplications that have been maintained both in insects and other animals. Among insects, *Mvl* appears to have duplicated separately in hemipterans (true bugs), Coleoptera (beetles), and wasps. Other than wasps, among the Hymenoptera, bees and ants have only one copy of *Mvl*. The coleopteran duplication appears to have preceded the split of beetles and Hymenoptera, which have lost *Mvl2*. Therefore, *Mvl1* shows greater homology to all hymenopteran *Mvl* genes.

Tissue-specific expression

Expression of *Mvl1* in brain, fat bodies, Malpighian tubules, midgut, ovaries, testes, and thoracic musculature varied across the different tissue types ($F_{7,32} = 44.361$, $P < 0.0001$). Expression in fat bodies was statistically significantly higher than in other tissues (Figure 3a). Hindgut, midgut, and thoracic musculature had moderate levels of

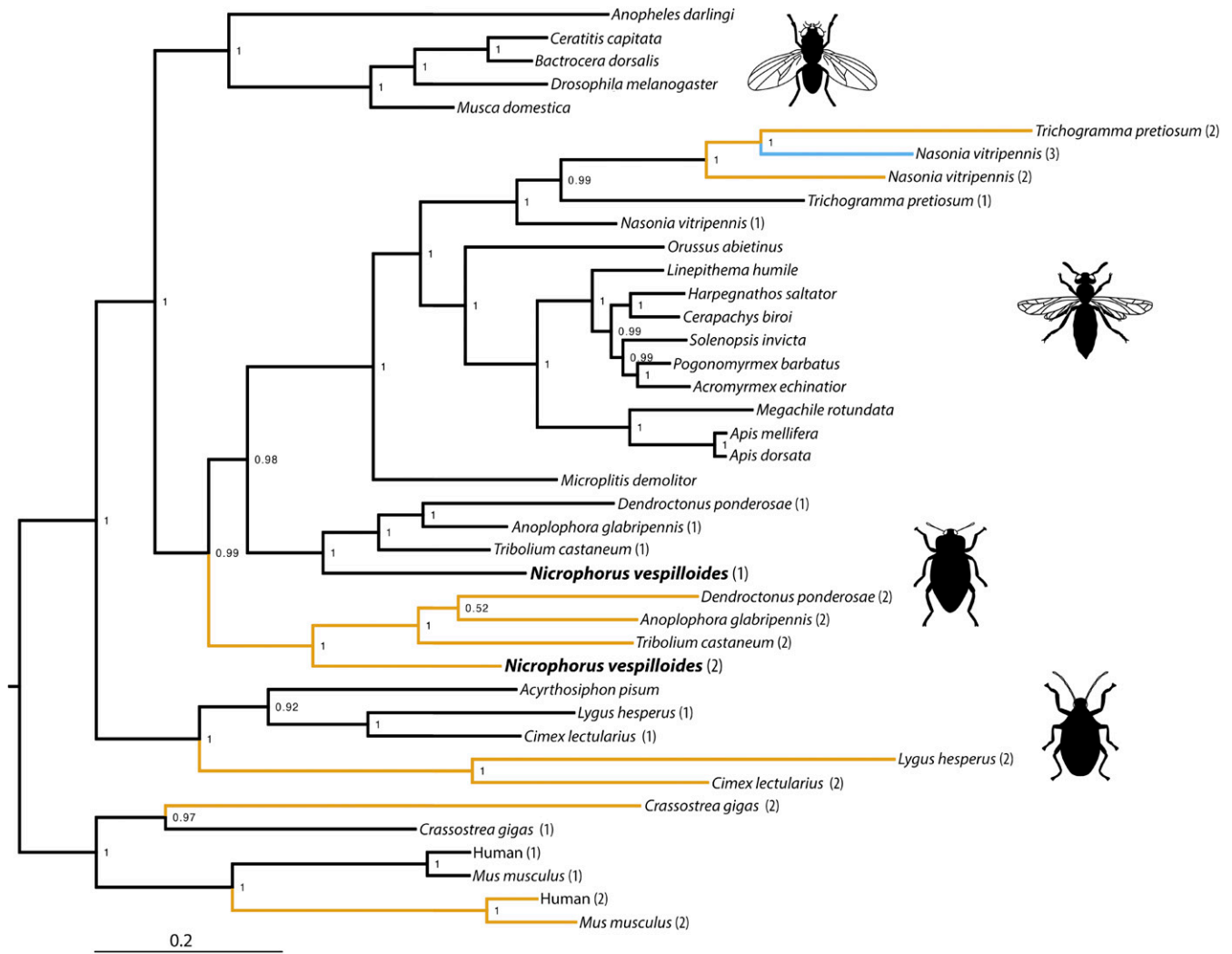


Figure 2 Phylogenetic relationships of *Malvolio*. Duplications are indicated with yellow and major insect groups are indicated pictorially. Included in this tree are mammals [human (*Homo sapiens*) and mouse (*Mus musculus*)] and oyster (*Crassostrea gigas*) as outgroups, and all insect orders where we could find homologs of *Malvolio* including Hemiptera [bed bug (*Cimex lectularius*), pea aphid (*Acyrtosiphon pisum*), and western tarnished plant bug (*Lygus hesperus*)], Hymenoptera (bees: *Apis dorsata*, *Apis mellifera*, *Metamicroptera rotundata*; wasps: *Orussus abietinus*, *Microplitis demolitor*, *Nasonia vitripennis*, *Trichogramma pretiosum*); ants: *Acromymex echinatio*, *Pogonomymex barbatus*, *Solenopsis invicta*, *Ooceraea biroi*, *Harpegnathos saltator*, *Linepithema humile*), Coleoptera (beetles: *Tribolium castaneum*, *Nicrophorus vespilloides*, *Anoplophora glabripennis*, *Dendroctonus ponderosae*), and Diptera (flies: *Ceratitis capitata*, *Bactrocera dorsalis*, *Drosophila melanogaster*, *Musca domestica*, *Anopheles darlingi*).

expression, while expression was relatively low in testes, Malpighian tubules, brains, and ovaries.

Expression patterns across tissues of *Mvl2* differed from those of *Mvl1* (Figure 3b). Overall, there was statistically significantly different expression across the different tissue types ($F_{7,32} = 37.420$, $P < 0.0001$) although in all tissues expression was much lower than that of *Mvl1*. Expression was highest in midgut and brain, with low expression in fat bodies, hindgut, and testes. Expression was negligible, and sometimes undetectable, in Malpighian tubules, ovaries, and thoracic muscle (Figure 3b).

Expression across different behavioral states

Expression patterns across behavioral states differed for *Mvl1* compared with *Mvl2*, with the patterns of expression being opposite for *Mvl1* and *Mvl2* across these states. Overall, there were statistically significant changes in expression across the behavioral states in *Mvl1* ($F_{4,45} =$

3.4087, $P = 0.0162$) (Figure 4a), with a significant increase in expression in resource preparation ($P = 0.0044$) and caring for offspring ($P = 0.0032$). There was no overall statistically significant difference in expression across behavioral states for *Mvl2* ($F_{4,43} = 2.2682$, $P = 0.077$) (Figure 4b) although expression decreased during social interactions, and resource preparation showed significantly lower expression than either virgin ($P = 0.019$) or post care ($P = 0.0285$).

RNA-seq data from *D. melanogaster* shows a pattern of relatively stable *Mvl* expression over time in the heads of virgin females, while the heads of mated females have a steady decrease in *Mvl* expression. Thus, the single copy of *Drosophila Mvl* has a similar expression to *N. vespilloides Mvl2* in the brain during mating. By contrast, *Drosophila Mvl* expression increases several-fold in the ovaries of mated females compared with virgin females. *Drosophila Mvl* expression in the digestive system and in testes is similar to *Mvl1* expression in *N. vespilloides* (Figure S1).

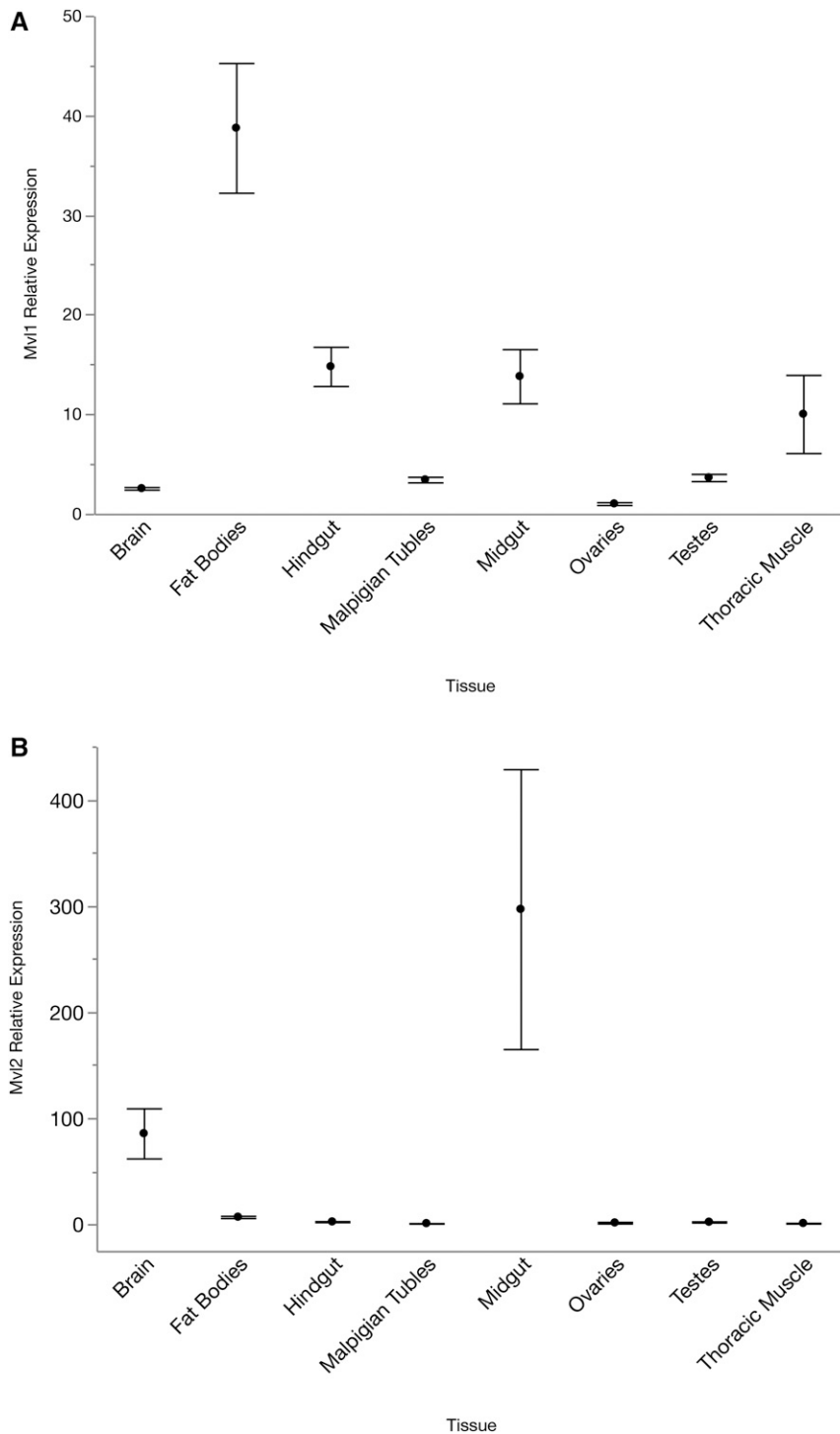


Figure 3 Relative expression of *Mvl1* (A) and *Mvl2* (B) in different tissue types. Two extreme values of *Mvl2* relative expression in the midgut were removed; removal of these two does not alter any statistical interpretations. Means \pm 1 SE are presented.

DISCUSSION

Gene duplication is a major factor in evolution (Ohno 1970; Innan and Kondrashov 2010; Wagner 2011), particularly where there is neofunctionalization, as the duplicated gene can permit access to variation that may have otherwise been constrained. Here we examined *Malvolio*, a gene that typically functions as a transporter of divalent cations. Examining the genome of the subsocial beetle, *N. vespilloides*, we found that *Malvolio* was duplicated in this insect. Two other factors suggested

that it would be informative to examine this duplication further: first, *Malvolio* is the homolog of *Nramp* in vertebrates, a gene that is duplicated and subfunctionalized (Techau *et al.* 2007; Neves *et al.* 2011) and second, *Malvolio* is known to play a role in social behavior in bees (although they have only one copy), suggesting that a duplication could therefore be sub/neofunctionalized for behavior. Our study had two components, the first analyzing and comparing *Mvl* protein sequence and the second examining expression of the duplicated genes. These

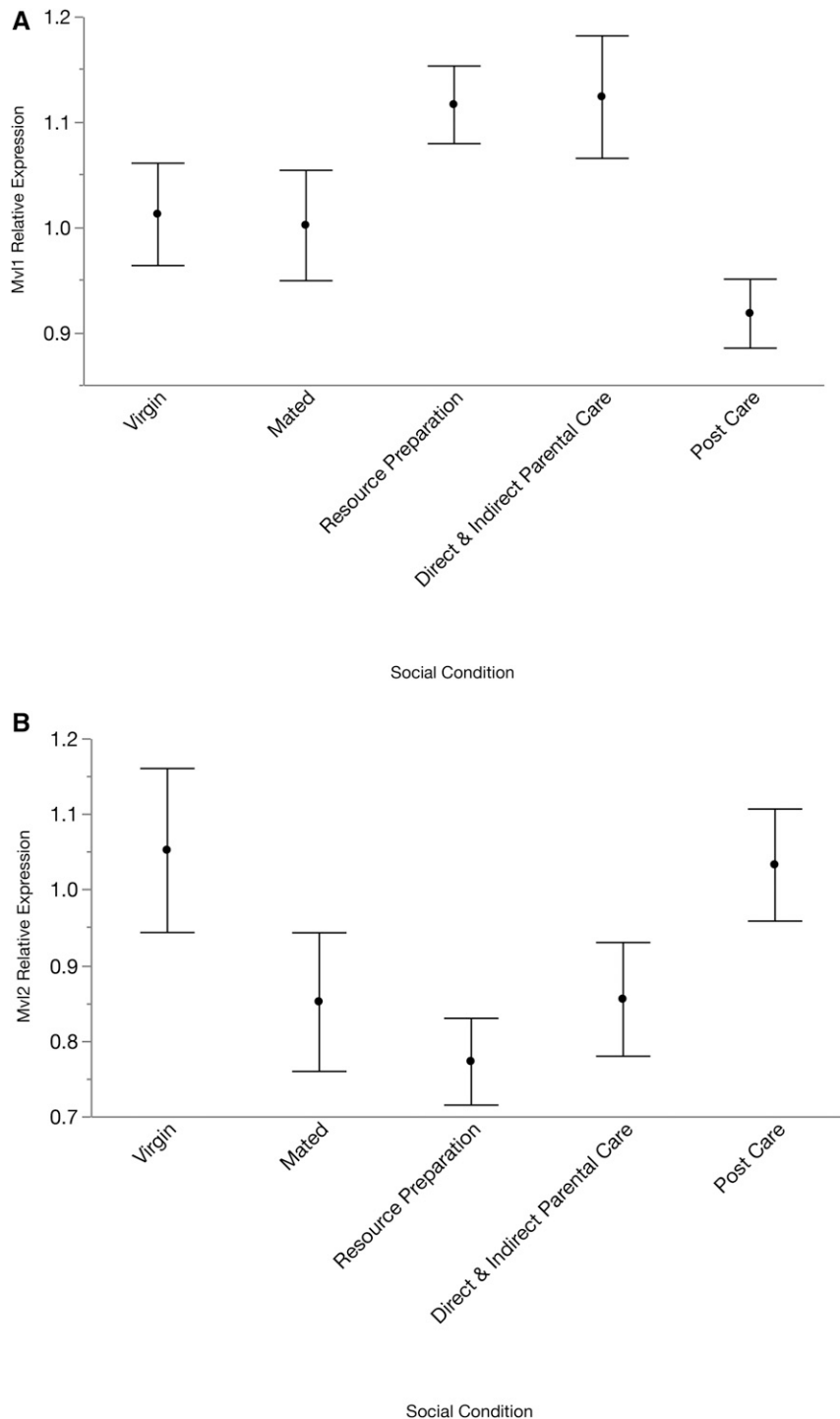


Figure 4 Relative expression of *Mvl1* (A) and *Mvl2* (B) in female heads across different physiological/behavioral states. Virgin females had no social experience or exposure to a carcass resource required for mating and oogenesis; mated females were placed with a male for 48 hr but not provided with a carcass; resource preparation were mated females provided with a carcass for 48 hr to prepare for reproduction and provisioning of offspring; direct and indirect care were females sampled during the most active period of direct feeding of offspring and in the act of regurgitating food to the offspring; post care females had dispersed from the carcass and had no further interactions with larvae for at least 24 hr. All individuals were 21 d of age when sampled. Means \pm 1 SE are presented.

studies provide several lines of evidence supporting the neofunctionalization of *Mvl* in *N. vespilloides* specifically, suggesting that *Mvl1* has retained a more ancestral function while *Mvl2* has diverged.

Our phylogenetic analysis showed that many species of insects have duplicate copies and, furthermore, that these duplications appear to have occurred in multiple lineage-specific events. *Malvolio* duplicates have arisen and persisted in at least three different insect lineages (and in wasps, a second duplication event appears to have led to at least one species having three copies of *Malvolio*). Given the tendency of dupli-

cated genes to remain redundant and eventually be removed from the genome, this suggests that *Malvolio* may possess qualities that have been found to encourage persistence after a duplication event (Kondrashov *et al.* 2002; Papp *et al.* 2003; Davis and Petrov 2004; Jordan *et al.* 2004; Marland *et al.* 2004).

The first evidence supporting sub/neofunctionalization of the *Mvl* genes specifically in *N. vespilloides* comes from analysis of the comparative structure of both copies. Specifically, we examined four regions of the gene: the first external loop, which is responsible for metal ion

binding specificity (Cohen *et al.* 2003; Southon *et al.* 2008); the consensus transport motif, which is important for the formation of voltage-gated potassium channels (Cellier *et al.* 2007); a region indicating whether the protein will function as a symporter or transporter (Techau *et al.* 2007; Southon *et al.* 2008); and the C-terminal region, which determines intracellular localization (Tabuchi *et al.* 2002). The first three of these regions are highly similar between *Mvl1* and *Mvl2* in *N. vespilloides*, suggesting that there may be overlap in transport specificity and mechanism. However, there does appear to be significant variation in the C-terminal region. Specifically, our alignment of *Mvl* and *Nramp* proteins suggests that amino acids known to affect localization of human *Nramp2* (Tabuchi *et al.* 2002) are conserved in *N. vespilloides* *Mvl1* but completely absent in *Mvl2*. Furthermore, *Mvl1* undergoes alternative splicing at this site, which suggests that, as in humans (Tabuchi *et al.* 2002), different splice forms of *Mvl1* may locate to different parts of the cell. These localizations are likely different from those of *Mvl2*, where these amino acids are absent. This predicted variation in cellular localization would support a divergence in tissue- or organ-level function between *Mvl1* and *Mvl2* despite the overall similarity between these gene copies.

We next examined tissue-specific expression of both genes to further investigate whether sub/neofunctionalization may be responsible for the maintenance of both *Malvolio* duplicates in *N. vespilloides*. Our data show that *Mvl1* is expressed in all eight measured tissues, with relatively low variance in gene expression within a tissue. In contrast to *Mvl1*, expression of *Mvl2* was limited to only two tissues, the brain and the midgut. This pattern is roughly consistent with tissue- and stage-specific data from *Tribolium*, which also shows high and ubiquitous *Mvl1* expression as opposed to low and inconsistent, but detectable, *Mvl2* expression (Dippel *et al.* 2014). This suggests that *Mvl1* may have maintained a conserved homeostatic role throughout the coleopteran lineage, consistent with the necessity of manganese transport on the cellular level (Culotta *et al.* 2005). Differences in expression between specific tissues may be related to other well-established functions of *Mvl* and its homologs, such as intracellular immunity (Evans *et al.* 2001; Cellier *et al.* 2007). *Mvl2*, on the other hand, appears not to be required for basic tissue function, and thus may be subject to weaker pleiotropic constraints.

Finally, we examined the expression patterns of both genes in the head, which includes both brain and fat body, in relation to reproductive and parental care behavior to further examine the possibility that the function of *Mvl2* has diverged from that of *Mvl1* in *N. vespilloides*. This species of beetle is not a genetic model organism, although we have a sequenced genome (Cunningham *et al.* 2015b). Instead, *N. vespilloides* is biologically interesting for its unusually elaborate parenting and social interactions (Parker *et al.* 2015). Previous research has shown that genes differentially expressed during parenting are detected in the specific social conditions we sampled (Parker *et al.* 2015; Roy-Zokan *et al.* 2015; Cunningham *et al.* 2016, 2017). Our hypothesis that there may be a behavioral function was based on the studies showing *Malvolio* is involved in caste differentiation in honey bees (Ben-Shahar *et al.* 2004) as well as feeding behavior in *Drosophila* (Søvik *et al.* 2017). We have hypothesized that feeding pathways are coopted to influence parental provisioning behavior (Cunningham *et al.* 2016, 2017), which fits the known roles of *Malvolio*, making this gene a strong candidate for influencing parenting. We found that the two copies do show differences in expression in head tissue associated with changes in behavior and social interactions. Whereas *Mvl1* expression increases during parenting, *Mvl2* appears to decrease during the same behavioral stages. These opposing expression patterns suggest that even though both gene copies have retained roles in social behavior, these roles have diverged.

Drosophila Mvl expression in the head shows a similar trend to *Mvl2* in *N. vespilloides*, where mated individuals have decreased expression. Interestingly, this is the opposite of *Mvl1* in *N. vespilloides*, where the expression increased after mating. The relatively high expression of *Mvl2* in the midgut of *N. vespilloides* likely indicates a role in digestive iron intake, though further experiments would be required to determine if this function is independent or complementary to *Mvl1* in the midgut.

Given the tissue- and stage-specific expression patterns of *Mvl1* and *Mvl2*, it appears likely that these genes have undergone neofunctionalization in burying beetles. Definitive evidence will require functional studies. It is intriguing that expression is opposite for *Mvl1* and *Mvl2* during parent-offspring social interactions. Data from honey bees, in which *Malvolio* is not duplicated, show that a single copy can account for both behavioral and other gene functions (Ben-Shahar *et al.* 2004; Søvik *et al.* 2015), suggesting that divergence in gene function between copies could be obtained by subfunctionalization alone. However, if this were the case, we would predict that one copy would have completely lost its association with behavior in *N. vespilloides*. Instead, we observe the evolution of opposing gene expression patterns between copies, meaning the expression patterns of at least one gene copy must be derived. Given the highly divergent and elaborate social interactions during parenting, and extensive parenting, in this species, this suggests that *Malvolio* may be coopted for further behavioral evolution. Furthermore, divergence in tissue-specific expression patterns, as we observed here, is often associated with neofunctionalization (Huminiacki and Wolfe 2004, Li *et al.* 2005). Therefore, our data are consistent with neofunctionalization. It may be that the expression patterns of *Mvl2* in the brain and midgut are still evolving, and understanding whether expression is being gained or lost in these tissues along with explicitly functional studies would help resolve this question.

In conclusion, *N. vespilloides* produces two copies of *Malvolio*, a gene which is commonly duplicated and maintained in insects and vertebrates. Our sequence and expression data suggest that, in *N. vespilloides*, *Malvolio* has experienced neofunctionalization following its duplication, potentially with an enhanced role in behavior. Although further functional studies are needed to eliminate subfunctionalization as an explanation for the patterns we see, our data do suggest that the two copies are not equivalent. Finally, we suggest that the predilection for duplicates of this gene to be maintained may reflect a tendency for sub- or neofunctionalization of *Mvl* in other systems as well.

ACKNOWLEDGMENTS

We thank Trish Moore and Michelle Ziadie for helpful discussions. Two anonymous reviewers provided excellent suggestions that expanded and improved our study. This research was supported by a National Science Foundation grant IOS-1354358 to A.J.M., with E.C. Mehlferber supported by a Research Experiences for Undergraduates supplement.

Author contributions: E.C. Mehlferber and A.J.M. conceived and designed the study, E.C. Mehlferber carried out all experiments along with C.B.C. (genome comparisons), E.M.R.-Z. (phylogenetic comparison), and E.C. McKinney (gene expression). E.C. Mehlferber, K.M.B., and A.J.M. performed bioinformatics and statistical analyses, and wrote the paper with input from all authors.

LITERATURE CITED

Ben-Shahar, Y., N. L. Dudek, and G. E. Robinson, 2004 Phenotypic deconstruction reveals involvement of manganese transporter *malvolio* in honey bee division of labor. *J. Exp. Biol.* 207: 3281–3288.

- Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos *et al.*, 2009 BLAST plus: architecture and applications. *BMC Bioinformatics* 10: 1.
- Cellier, M. F., P. Courville, and C. Campion, 2007 Nramp1 phagocyte intracellular metal withdrawal defense. *Microbes Infect.* 9: 1662–1670.
- Cohen, A., Y. Nevo, and N. Nelson, 2003 The first external loop of the metal ion transporter DCT1 is involved in metal ion binding and specificity. *Proc. Natl. Acad. Sci. USA* 100: 10694–10699.
- Corona, M., R. Libbrecht, Y. Wurm, O. Riba-Grognuz, R. A. Studer *et al.*, 2013 Vitellogenin underwent subfunctionalization to acquire caste and behavioral specific expression in the harvester ant *Pogonomyrmex barbatus*. *PLoS Genet.* 9: e1003730.
- Culotta, V. C., M. Yang, and M. D. Hall, 2005 Manganese transport and trafficking: lessons learned from *Saccharomyces cerevisiae*. *Eukaryot. Cell* 4: 1159–1165.
- Cunningham, C. B., M. K. Douthit, and A. J. Moore, 2015a Expression of octopaminergic receptor genes in 4 nonneural tissues in female *Nicrophorus vespilloides* beetles. *Instr. Sci.* 22: 495–502.
- Cunningham, C. B., L. Ji, R. A. W. Wiberg, J. Shelton, E. C. McKinney *et al.*, 2015b The genome and methylome of a beetle with complex social behavior, *Nicrophorus vespilloides* (Coleoptera: Silphidae). *Genome Biol. Evol.* 7: 3383–3396.
- Cunningham, C. B., K. VanDenHeuvel, D. B. Khana, E. C. McKinney, and A. J. Moore, 2016 The role of neuropeptide F in a transition to parental care. *Biol. Lett.* 12: 20160158.
- Cunningham, C. B., M. Badgett, R. B. Meagher, R. Orlando, and A. J. Moore, 2017 Ethological principles predict the neuropeptides co-opted to influence parenting. *Nat. Commun.* 8: 14225.
- Davis, J. C., and D. A. Petrov, 2004 Preferential duplication of conserved proteins in eukaryotic genomes. *PLoS Biol.* 2: e55.
- Dippel, S., G. Oberhofer, J. Kahnt, L. Gerischer, L. Opitz *et al.*, 2014 Tissue-specific transcriptomics, chromosomal localization, and phylogeny of chemosensory and odorant binding proteins from the red flour beetle *Tribolium castaneum* reveal subgroup specificities for olfaction or more general functions. *BMC Genomics* 15: 1141.
- Dittmar, K., and D. Liberles (Editors), 2010 *Evolution After Gene Duplication*. Wiley-Blackwell, Hoboken, NJ.
- Eggert, A.-K., M. Reinking, and J. K. Müller, 1998 Parental care improves offspring survival and growth in burying beetles. *Anim. Behav.* 55: 97–107.
- Evans, C. A., M. S. Harbuz, T. Ostefeld, A. Norrish, and J. M. Blackwell, 2001 Nramp1 is expressed in neurons and is associated with behavioural and immune responses to stress. *Neurogenetics* 3: 69–78.
- Feuda, R., F. Marlitz, M. A. Bentley, and P. W. H. Holland, 2016 Conservation, duplication, and divergence of five opsin genes in insect evolution. *Genome Biol. Evol.* 8: 579–587.
- Folwell, J. L., C. H. Barton, and D. Shepherd, 2006 Immunolocalisation of the *D. melanogaster* Nramp1 homologue Malvolio to gut and Malpighian tubules provides evidence that Malvolio and Nramp2 are orthologous. *J. Exp. Biol.* 209: 1988–1995.
- Force, A., M. Lynch, F. B. Pickett, A. Amores, Y. L. Yan *et al.*, 1999 Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151: 1531–1545.
- Gelbart, W. M., and D. B. Emmert, 2013 FlyBase high throughput expression pattern data. ModENCODE. Available at: <http://flybase.org/reports/FBBrf0212041.html>. Accessed: October 29, 2013.
- Goldman-Huertas, B., R. F. Mitchell, R. T. Lapoint, C. P. Faucher, J. G. Hildebrand *et al.*, 2015 Evolution of herbivory in Drosophilidae linked to loss of behaviors, antennal responses, odorant receptors, and ancestral diet. *Proc. Natl. Acad. Sci. USA* 112: 3026–3031.
- Huminiacki, L., and K. H. Wolfe, 2004 Divergence of spatial gene expression profiles following species-specific gene duplications in human and mouse. *Genome Res.* 14: 1870–1879.
- Innan, H., and F. Kondrashov, 2010 The evolution of gene duplications: classifying and distinguishing between models. *Nat. Rev. Genet.* 11: 97–108.
- Jordan, I. K., Y. I. Wolf, and E. V. Koonin, 2004 Duplicated genes evolve slower than singletons despite the initial rate increase. *BMC Evol. Biol.* 4: 22–32.
- Katz, P. S., and J. L. Lillis, 2014 Reconciling the deep homology of neuromodulation with the evolution of behavior. *Curr. Opin. Neurobiol.* 29: 39–47.
- Kondrashov, F. A., I. B. Rogozin, Y. I. Wolf, and E. V. Koonin, 2002 Selection in the evolution of gene duplications. *Genome Biol.* 3: RESEARCH0008.
- Koressaar, T., and M. Remm, 2007 Enhancements and modifications of primer design program Primer3. *Bioinformatics* 23: 1289–1291.
- Li, W. H., J. Yang, and X. Gu, 2005 Expression divergence between duplicate genes. *Trends Genet.* 21: 602–607.
- Lock, J. E., P. T. Smiseth, and A. J. Moore, 2004 Selection, inheritance, and the evolution of parent-offspring interactions. *Am. Nat.* 164: 13–24.
- Maere, S., and Y. Van de Peer, 2010 Duplicate retention after small- and large-scale duplications, pp. 31–56 in *Evolution After Gene Duplication*, edited by Dittmar, K., and D. Liberles. Wiley-Blackwell, Hoboken, NJ.
- Marland, E., A. Prachumwat, N. Maltsev, Z. Gu, and W. H. Li, 2004 Higher gene duplicabilities for metabolic proteins than for nonmetabolic proteins in yeast and *E. coli*. *J. Mol. Evol.* 59: 806–814.
- McWilliam, H., W. Li, M. Uludag, S. Squizzato, Y. M. Park *et al.*, 2013 Analysis tool web services from the EMBL-EBI. *Nuc. Ac. Res.* 41: W597–W600.
- Nadeau, J. H., and D. Sankoff, 1997 Comparable rates of gene loss and functional divergence after genome duplications early in vertebrate evolution. *Genetics* 147: 1259–1266.
- Neves, J. V., J. M. Wilson, H. Kuhl, R. Reinhardt, L. F. Castro *et al.*, 2011 Natural history of SLC11 genes in vertebrates: tales from the fish world. *BMC Evol. Biol.* 11: 106.
- Nowak, M. A., M. C. Boerlijst, J. Cooke, and J. Maynard Smith, 1997 Evolution of genetic redundancy. *Nature* 388: 167–171.
- Ohno, S., 1970 *Evolution by Gene Duplication*. Springer, New York.
- Orgad, S., H. Nelson, D. Segal, and N. Nelson, 1998 Metal ions suppress the abnormal taste behavior of the *Drosophila* mutant malvolio. *J. Exp. Biol.* 201: 115–120.
- Papp, B., C. Pál, and L. D. Hurst, 2003 Dosage sensitivity and the evolution of gene families in yeast. *Nature* 424: 194–197.
- Parker, D. J., C. B. Cunningham, C. A. Walling, C. E. Stamper, M. L. Head *et al.*, 2015 Transcriptomes of parents identify parenting strategies and sexual conflict in a subsocial beetle. *Nat. Commun.* 6: 8449.
- Rodrigues, V., P. Y. Cheah, K. Ray, and W. Chia, 1995 malvolio, the *Drosophila* homologue of mouse NRAMP-1 (Bcg), is expressed in macrophages and in the nervous system and is required for normal taste behaviour. *EMBO* 14: 3007–3020.
- Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling *et al.*, 2012 MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61: 539–542.
- Roy-Zokan, E. M., C. B. Cunningham, L. E. Hebb, E. C. McKinney, and A. J. Moore, 2015 Vitellogenin and vitellogenin receptor gene expression is associated with male and female parenting in a subsocial insect. *Proc. R. Soc. Lond. B Biol. Sci.* 282: 20150787.
- Searle, S., N. A. Bright, T. I. Roach, P. G. Atkinson, C. H. Barton *et al.*, 1998 Localisation of Nramp1 in macrophages: modulation with activation and infection. *J. Cell Sci.* 111: 2855–2866.
- Southon, A., A. Farlow, M. Norgate, R. Burke, and J. Camakaris, 2008 Malvolio is a copper transporter in *Drosophila melanogaster*. *J. Exp. Biol.* 211: 709–716.
- Søvik, E., C. J. Perry, A. LaMora, A. B. Barron, and Y. Ben-Shahar, 2015 Negative impact of manganese on honeybee foraging. *Biol. Lett.* 11: 20140989.
- Søvik, E., A. LaMora, G. Seehra, A. B. Barron, J. G. Duncan *et al.*, 2017 *Drosophila* divalent metal ion transporter Malvolio is required in dopaminergic neurons for feeding decisions. *Genes Brain Behav.* 16: 506–514.
- Tabuchi, M., N. Tanaka, J. Nishida-Kitayama, H. Ohno, and F. Kishi, 2002 Alternative splicing regulates the subcellular localization of divalent metal transporter 1 isoforms. *Mol. Biol. Cell* 13: 4371–4387.

- Techau, M. E., J. Valdez-Taubas, J. F. Popoff, R. Francis, M. Seaman *et al.*, 2007 Evolution of differences in transport function in Slc11a family members. *J. Biol. Chem.* 282: 35646–35656.
- Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth *et al.*, 2012 Primer3-new capabilities and interfaces. *Nuc. Ac. Res.* 40: e115.
- Wagner, A., 1998 The fate of duplicated genes: loss or new function? *Bio-Essays* 20: 785–788.
- Wagner, A., 2011 The molecular origins of evolutionary innovations. *Trends Genet.* 27: 397–410.
- Whelan, S., and N. Goldman, 2001 A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* 18: 691–699.

Communicating editor: J. Comeron