An Eye on Trafficking Genes: Identification of Four Eye Color Mutations in *Drosophila*

Paaqua Grant,*^{,1} Tara Maga,*^{,†,2} Anna Loshakov,*^{,3} Rishi Singhal,*^{,4} Aminah Wali,^{†,5}

Jennifer Nwankwo,^{†,6} Kaitlin Baron,^{*,7} and Diana Johnson^{*,8}

*Department of Biological Sciences and [†]Undergraduate Summer Research Program, Department of Biological Sciences, The George Washington University, Washington, DC 20052

ABSTRACT Genes that code for proteins involved in organelle biogenesis and intracellular trafficking produce products that are critical in normal cell function. Conserved orthologs of these are present in most or all eukaryotes, including *Drosophila melanogaster*. Some of these genes were originally identified as eye color mutants with decreases in both types of pigments found in the fly eye. These criteria were used for identification of such genes, four eye color mutations that are not annotated in the genome sequence: *chocolate, maroon, mahogany*, and *red Malpighian tubules* were molecularly mapped and their genome sequences have been evaluated. Mapping was performed using deletion analysis and complementation tests. *chocolate* is an allele of the *VhaAC39*-1 gene, which is an ortholog of the Vacuolar H⁺ ATPase AC39 subunit 1. *maroon* corresponds to the *Vps16A* gene and its product is part of the HOPS complex, which participates in transport and organelle fusion. *red Malpighian tubule* is the *CG12207* gene, which encodes a protein of unknown function that includes a LysM domain. *mahogany* is the *CG13646* gene, which is predicted to be an amino acid transporter. The strategy of identifying eye color genes based on perturbations in quantities of both types of eye color pigments has proven useful in identifying proteins involved in trafficking and biogenesis of lysosome-related organelles. Mutants of these genes can form the basis of valuable *in vivo* models to understand these processes.

Supplemental material is available online at www.g3journal.org/lookup/suppl/doi:10.1534/g3.116.032508/-/DC1.

Drosophila melanogaster is an extremely useful model organism for both classical genetics and molecular biology. It is also one of the earliest organisms to have its DNA sequence determined and extensively annotated. However, many of the genes identified and used in classical genetics have not been matched with their corresponding genes in the molecular sequence. This study identifies four genes coding for eye color mutations.

Certain eye color genes regulate vesicular transport in cells. Enzymes and other substances needed for pigment synthesis are transported to the pigment granule, a lysosome-related organelle (Dell'Angelica *et al.* 2000; Reaume *et al.* 1991; Summers *et al.* 1982). *D. melanogaster* has a redbrown eye color caused by the presence of two classes of pigments, pteridines (red) and ommochromes (brown). There are independent pathways for the synthesis of each (Reaume *et al.* 1991; Summers *et al.* 1982). Some eye color mutations are caused by the lack of individual enzymes in these pathways. Examples include bright red *vermillion* (tryptophan 2,3-dioxygenase), which is incapable of synthesis of any ommochromes, and dark brown *sepia* (GSTO4), which lacks an enzyme used to synthesize a subset of pteridines (Baillie and Chovnick 1971; Kim *et al.* 2006; Searles and Voelker 1986; Searles *et al.* 1990; Walters *et al.* 2009; Wiederrecht and Brown 1984). The first mutant described in *Drosophila, white*, lacks both types of pigments in the eye

KEYWORDS

genetic analysis vesicular transporters LysM domain in eukaryotes

Copyright © 2016 Grant et al.

doi: 10.1534/g3.116.032508

Manuscript received June 17, 2016; accepted for publication August 2, 2016; published Early Online August 23, 2016.

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.

¹Present address: Integrated Resources, Inc (IRI), 6180 Willow Lane, Farmington, NY 14425.

²Present address: Familial Breast Cancer Program (Genetic Program), University of Illinois – Chicago, Chicago, IL 60612.

³Present address: The Commonwealth Medical College, Scranton, PA 18509.
⁴Present address: The George Washington University School of Medicine, Washington, DC 20052.

⁵Present address: Curriculum in Genetics and Molecular Biology, Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599.

⁶Present address: Pharmacology & Experimental Therapeutics, Sackler Graduate School of Biomedical Sciences, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111.

⁷Present address: GW Medical Faculty Associates, 2150 Pennsylvania Avenue, NW Suite 8-416, Washington, DC 20037.

⁸Corresponding author: Department of Biological Sciences, The George Washington University, 800 22nd Street, NW Suite 6000, Washington, DC 20052. E-mail: dejohnsn@gwu.edu

Table 1 Stocks used for locating eye color mutants

cho Mapping	All Autosomal Mapping	mahogany Mapping, cont'd.
P{XP}XPG-L Exelixis Stock Collection	P{ry ^{+t7.2} =hsFLP}1, y ¹ w ¹¹¹⁸ ; Dr ^{Mio} /TM3 ry [*] Sb ¹	w ¹¹¹⁸ ;Df(3R)P{XP}CG31121 ^{d06890} to PBac{WHJ ^{f01730} /TM6B, Tb Constructed by authors
w ¹¹¹⁸ /Binsinscy	w ¹¹¹⁸ ; wg ^{Sp-1} /CvO; sens ^{Ly-1} /TM6B, Tb ¹	,
w ¹¹¹⁸ ; MKRS, P{ry ^{+t7.2} = hsFLP}86E/TM6B, Tb ¹		red Mapping
w ¹¹¹⁸ /FM7c	maroon Mapping	w ¹¹¹⁸ ; Df(3R)Exel7321/TM6B, Tb ¹
y ² cho ² flw ¹	ma ¹ fl ¹	w ¹¹¹⁸ ; Df(3R)Exel6267, P{w ^{+mC} = XP-U}Exel6267/TM6B, Tb ¹
Df(1)10-70d, cho ¹ sn ³ /FM6	w ¹¹¹⁸ ; Df(3R)BSC507/TM6C, Sb ¹ cu ¹	w ¹¹¹⁸ ; PBac{ w ^{+mC} = RB}su(Hw)e ⁰⁴⁰⁶¹ /TM6B_Tb ¹
Df(1)ED6716 w ¹¹¹⁸ /FM7h	w ¹¹¹⁸ ; PBac{RB}Aats-trp ^{e00999} /TM6B Exelixis Stock Collection	w ¹¹¹⁸ ; P{XP}trx ^{d08983} /TM6B, Tb ¹
w ¹¹¹⁸ P+PBac{XP3.WH3}BSC877/ FM7h/Dp(2;Y)G P{hs-hid}Y	w ¹¹¹⁸ ; P{XP} ^{d00816} /TM6B Exelixis Stock Collection	w ¹¹¹⁸ ; Df(3R)P{XP}trx ^{d08983} to PBac{RB}su(Hw) ^{e04061} /TM6B, Tb ¹ Constructed by authors
Df(1)BSC834 w ¹¹¹⁸ /Binsinscy	w ¹¹¹⁸ ; Df(3R)Exel9036, PBac{WH}Exel9036/TM6B, Tb ¹	
w ¹¹¹⁸ ; <i>P{XP}^{d03180}</i> Exelixis Stock Collection	w ¹¹¹⁸ ;Df(3R)ED5339, P{3'.RS5+3.3'}ED5339/TM6C, cu ¹ Sb ¹	
w ¹¹¹⁸ ; PBac{RB}VhaAC39-1 ^{e04316} Exelixis Stock Collection	w ¹¹¹⁸ ;Df(3R) PBac{RB}Aats-TrpRS ^{e00999} to P{XP} ^{d00816} /TM6B, Tb ¹ Constructed by authors	Additional Stocks for Rescue Crosses
w^{1118} ; $P\{w^{+mC} = XP\}yin^{d02176}$		w*; P{w ^{+mc} GAL4-ninaE.GMR}12
<i>P{XP}ec^{d00965}</i> Exelixis Stock Collection	mahogany Mapping	w*;Cy/P{w ^{+mc} GAL4-ninaE.GMR}12 Constructed by authors
PBac{WH} ^{f06086} Exelixis Stock Collection	mah ¹	w*; Kr ^{lf-1} /CyO
w ¹¹¹⁸ ;Df(1) P{XP}ec ^{d00965} Exelixis	w ¹¹¹⁸ ; PBac{WH} ^{f01730} /TM6B Exelixis Stock	w*; Kr ^{lf-1} /CyO; Df(3L)Ly,
Stock Collection		sens ^{Ly-1} /TM6C, Sb ¹ Tb ¹
PBac{WHJ ^{rucuse} Exelixis Stock Collection w ¹¹¹⁸ ;Df(1) P{XP}ec ^{d00965} to PBac{WHJ ^{rucuse} /FM7h Constructed	w ¹¹¹⁸ ; P{w ^{+mc} = XP}CG31121 ^{auos90} /TM6B w ¹¹¹⁸ ; Df(3R)BSC494/TM6C, Sb ¹ cu ¹	w ¹¹¹⁸ ; Kr ^{I-1} /CyO; TM3, Sb ¹ /D ¹ w [*] ; Kr ^{IF-1} /CyO; CxD/TM6C, Sb ¹ Tb ¹
by authors w ¹¹¹⁸ PBac{RB}VhaAC39-1 ^{e04316} to P{XP}yin ^{d02176} /FM7h Constructed	w ¹¹¹⁸ ; Df(3R)Exel6200, P{XP-U}Exel6200/TM6B, Tb ¹	w*; Kr ^{lf-1} /CyO; TM3, Ser ¹ /D ¹
by authors w ¹¹¹⁸ ; Df(1)P{XP} ^{d03180} to PBac{RB}VhaAC39-1 ^{e04316} /FM7h Constructed by authors	w ¹¹¹⁸ ; Df(3R)BSC318/TM6C, Sb ¹ cu ¹	w ¹¹¹⁸ ; TM3, Sb ¹ /CxD

Unless otherwise noted, the stocks were obtained from the Bloomington Drosophila Stock Center.

(Morgan 1910). The *white* gene codes for a subunit of an ABC transporter located in the membrane of the pigment granule, and is required for transport of substrates into pigment granules (Goldberg *et al.* 1982; Mackenzie *et al.* 1999; Summers *et al.* 1982). Lloyd *et al.* (1998) developed a hypothesis that single mutations causing perturbations in the amount of both types of pigments in the eyes could result from defects in genes associated with intracellular transport. They named these genes "the granule group." Their concept has been supported by both identification of existing eye color genes for which the genomic coding sequences were originally unknown, and by studies in which induced mutations or reduced expression of orthologs of known transporter genes resulted in defective eye color phenotypes.

Vesicular trafficking is critical. Each cell in the body is highly organized with a number of compartments. These must contain specific proteins and have distinct characters, such as pH, as well as specific molecules attached to them. Vesicular trafficking is accomplished by a large number of protein complexes and individual effector proteins, and the list of these is still growing. One mode of delivery involves endocytic trafficking resulting in the sorting of proteins, lipids, and other materials and directing them to specific vesicles/organelles (Li *et al.* 2013). Many of the relevant proteins were first identified in yeast and have been highly conserved (Bonifacino 2014). D. melanogaster eye color genes provide in vivo metazoan models of these proteins and their interactions in trafficking. For example, the adaptor protein 3 (AP3) complex consists of four proteins: the β 3 (RUBY), δ 3 (GARNET), μ 3 (CAR-MINE), and σ 3 (ORANGE) subunits. It is involved in cargo selection for vesicles, transport of soluble N-ethylmaleimide-sensitive factor attachment protein receptor complexes, biogenesis of lysosomes, and lysosome-derived organelles, including formation of synaptic vesicles (Besteiro et al. 2008; Cowles et al. 1997; Faundez et al. 1998; Kent et al. 2012; Mullins et al. 1999, 2000; Ooi et al. 1997). The homotypic vacuole fusion and protein sorting (HOPS) complex in metazoans has four core proteins coded for by the *Vps16A*, *Vps11*, *Vps18* (dor), and *Vps33* (car) genes. Two additional proteins coded for by Vps39 and Vps41 (lt) can also be part of the complex. The HOPS complex participates in endocytic transport, endosome maturation, and fusion with lysosomes (Nickerson et al. 2009; Solanger and Spang 2013). Rab GTPases are molecular switches, active when coupled with GTP. The guanine exchange factor (GEF) catalyzes the conversion of GDP to GTP, activating its associated Rab GTPase. Rabs contribute to cargo selection, vesicle movement via microfilaments and actin, and fusion of membranes. Two eye color mutants, lightoid (Rab32) and claret (its putative

Table 2 Existing deletions or transposable elements used to produce deletions

Gene Deletion or Transposable Elements	Complements Mutant
chocolate	
P{XP} ^{d03180} and PBac{RB}VhaAC39-1 ^{e04316}	Yes
PBac{RB}VhaAC39-1 ^{e04316} and P{XP}yin ^{d02176}	Yes
P{XP}ec ^{d00965} and PBac{WH}f06086	No
Df(1)BSC834	Yes
Df(1)BSC877	Yes
Df(1)ED6716	No
maroon	
Df(3R)ED5339	No
Df(3R)Exel9036	Yes
Df(3R)BSC507	No
mahogany	
Df(3R)BSC318	No
Df(3R)Exel6200	Yes
Df(3R)BSC494	Yes
P{XP}CG31121 ^{d06890} and PBac{WH}f01730	Yes
red Malpighian tubules	
Df(3R)Exel6267	No
Df(3R)Exel7321	Yes
P{XP}trx ^{d08983} and PBac{RB}su(Hw) ^{e04061}	No

Data from complementation experiments are given for each deletion.

GEF), have been shown to affect pigment granule morphology and autophagy; lightoid's transcript has also been shown to be enriched in neurons. Human Rab32 participates in the transport of enzymes involved in melanin production to the melanosome, another lysosomerelated organelle (Chan et al. 2011; Hutagalung and Novick 2011; Ma et al. 2004; Wang et al. 2012a; Zhen and Stenmark 2015). The human gene Lysosomal Trafficking Regulator Protein (LYST) is required for normal size and number of lysosomes and biogenesis of cytotoxic granules. Mutations in the gene are associated with Chediak-Higashi syndrome, which causes defects in immunity, prolonged bleeding, and oculocutaneous albinism. The fly ortholog, mauve, recapitulates some of these characteristics. It has overlarge pigment granules, abnormal eye color, and is susceptible to bacterial infection. It also lacks the ability to produce mature autophagosomes (Huizing et al. 2008; Rahman et al. 2012; Sepulveda et al. 2015). At least three sets of Biogenesis of Lysosome-related Organelle Complexes, BLOC1, BLOC2, and BLOC3, are required for normal development of these organelles (Dell'Angelica. 2004). The classic eye mutant, *pink*, is a member of BLOC2, an ortholog of Hermansky-Pudlak Syndrome 5 (HPS5) and shows genetic interactions with AP3 genes, garnet and orange, and the HOPS gene carnation (Di Pietro et al. 2004; Falcon-Perez et al. 2007; Syrzycka et al. 2007). Finally, Drosophila orthologs of genes originally identified as BLOC1 genes in other organisms have produced an abnormal eye color phenotype when their expression was inhibited by RNAi. These include four of the BLOC1 genes, Blos1, Pallidin, Dysbindin, and Blos4 (Bonifacino 2004; Cheli et al. 2010; Dell'Angelica et al. 2000).

Using the perturbation in amounts of both ommochromes and pteridines as a criterion, we selected and identified the genomic sequence coding for each of four eye color genes in *D. melanogaster: chocolate (cho), maroon (ma), mahogany (mah),* and *red Malpighian tubules (red)*. Two of the genes, *cho* and *ma*, are VATPase and *VPS16a* sub-units, respectively. The roles of these annotated genes in vesicular transport have been previously characterized. The *mah* mutation also codes for a gene involved in transfer of amino acids to granules. Finally, the

■ Table 3 Nonsynonymous differences, deletions, and insertions between the Drosophila melanogaster genome sequence and mutant alleles of maroon, chocolate, mahogany, and red Malpighian tubules

Position	DNA Change	Nonsynonymous Changes
Х	VhaA	AC39-1-chocolate
X:3,882,405	G > T	W330L
3R	V	os16A-maroon
9267133	G > A	M4I
9267260	G > A	A24T
9268607-9268615	Deletion	Deletion 422-I M R-424
9269662	A > T	E712D
3R	CG	13646-mahogany
24949138	roo insert	
24949334	T > C	1460Tª
3R	(CG12207 -red
		red ¹
14300442	A > C	N67H ^b
		red ^{K1}
14298973	G > A	G51S ^b

^dThis part of the exon may not be translated when the *roo* LTR is present. The amino acid positions listed are for isoforms PA, PD, and PE. In isoforms PB, PF, and PG, the positions are N90H for red^1 and G74S for red^{K1} .

last gene, *red*, codes for a predicted protein with a LysM domain, but the protein's function is unknown.

MATERIALS AND METHODS

Nucleic acid isolation, RT-PCR, and production of transgenic flies

Fly DNA was isolated from single flies using the method of Gloor et al. (1993). The Parks et al. (2004) "5 Fly Extraction" was used to isolate purer DNA from groups of five to 10 flies at a time. The second technique was modified by using single microcentrifuge tubes and a tissue grinder for homogenizing. RNA was isolated using TRIzol (Life Technologies). cDNA was prepared using Superscript II (Life Technologies) and an oligo dT primer. Genomic DNA for wild-type versions of cho, ma, and mah and cDNA from the CG12207 transcript A were amplified by PCR and TOPO-TA cloned (Life Technologies). Each gene's DNA/cDNA was subcloned into the pUAST vector (Drosophila Genomic Resource Center) in order to use the GAL4/UAS technique (Brand and Perrimon 1993; Duffy 2002). Clones were sequenced to verify that they had the wild-type sequence. The constructs were isolated with Plasmid Midi Prep kits (Qiagen). Injections of the pUAST clones to produce transgenic flies were performed by Rainbow Transgenic Flies.

Fly stocks and crosses

Fly stocks were maintained at 25° on Instant *Drosophila* Medium (Carolina Biological). Jim Kennison provided the EMS-induced red^{K1} stock and its OreR progenitor. Other stocks were obtained from the Bloomington and Exelixis Stock Collections (Table 1). Deletions were made using the Flp-FRT methods described by Parks (Parks *et al.* 2004). The chromosome sequence coordinates of specific existing deletions and deletions that were made are shown in Supplemental Material, Table S1. Recombinants were chosen based on their eye color. PCR was used to verify the deletions.

Deletion mapping crosses were made between each deletion stock (Table 2) and its corresponding homozygous mutant stock. Rescue crosses required that eye color phenotypes be evaluated in w^+ flies. The GAL4/UAS system was used (Brand and Perrimon 1993; Duffy



Figure 1 The VhaAC39-1 gene complements the chocolate gene. (A) Wild-type genotype and phenotype. (B) cho/Y with transgene showing the wild-type phenotype. The transgenic male is hemizygous for the mutant allele and carries the Gal4 driver from w^{*}; P{GAL4-ninaE.GMR}12 and one copy of the VhaAC39-1 transgene. (C) cho/Y showing the mutant phenotype.

2002). The driver employed for eye specific expression was Gal4-ninaE. GMR12. Stocks with balancers were used to produce w^+ ; $Cy/P\{GAL4-ninaE.GMR\}12$ homozygous mutant stocks for *ma*, *mah*, and *red*. These flies were crossed at 27° with the Cy/pUAST transgene stocks that were also homozygous for the appropriate third chromosome mutant allele and the phenotype was assayed in the F₁s. For the X-linked *cho* gene, w^+cho/w^+cho ; $Cy/P\{GAL4-ninaE.GMR\}12$ females were crossed with wY; Cy/pUAST-cho transgene stock flies and the phenotype was evaluated in the male F₁. Rescue was assayed for at least three independent transgene lines for each gene.

Photomicroscopy

Flies were photographed using a Leica DFC425 digital camera mounted on a Leica M205A stereomicroscope. A series of 15–70 images were taken at different focal planes with the software package Leica Application Suite version 3.0 (Leica Microsystems, Switzerland) and assembled using the software Helicon Focus 6 (Helicon Soft, Ukraine).

Sequence analysis

Sequencing was performed by the DNA Analysis Facility on Science Hill at Yale University, the Biological Sciences Sequencing Facility at the George Washington University, and Macrogen USA. Primers for cloning and sequencing are listed in Table S2. Sequences were compiled using Sequencher (Gene Codes). Predicted protein alignments were performed using T-Coffee or PSI-Coffee (Di Tommaso *et al.* 2011;

		0320	0	346	9
CHOCOLATE	HEVKLDVYAF	LQQFHFGVFY	AYLKLKEQEC	RNIVLIAECV	AQKHRAKIDN
D.melanogaster	H E V K L D <mark>V Y</mark> A F	LQQFHFGVFY	AYLKLKEQEC	RNIVWIAECV	AQKHRAKIDN
Zebrafish	HEVKLN <mark>KL</mark> AF	L N Q F H F S V F Y	A Y V K L K E Q E C	RNIVWIAECI	A Q R H R A K I D N
Platyfish	HEVKLN <mark>KL</mark> AF	L N Q F H F S V F Y	A Y V K L K E Q E C	RNIVWIAECI	A Q <mark>R</mark> H R A K I D N
X.Laevis	HEVKLN <mark>KL</mark> AF	L N Q F H F G V F Y	A F V K L K E Q E C	RNVVWIAECI	AQRHRAKIDN
Takifugu	HEVKLN <mark>KL</mark> AF	L N Q F H F S V F Y	A Y V K L K E Q E C	RNIVWIAECI	A Q <mark>R</mark> H R A K I D N
Alligator	HEVKLN <mark>KL</mark> AF	L N Q F H F G V F Y	A F V K L K E Q E C	RNIVWIAECI	AQRHRTKIDN
C.elegans	H E V K L N <mark>V H</mark> S Y	L H Q F H F G V F Y	AFIKLKEQEM	RNIIWIAECI	S Q <mark>R</mark> H R T K I D N
D.virilis	HEVKLN <mark>VF</mark> AF	L Q Q F H F G V F Y	A Y L K L K E Q E C	RNIVWIAECV	A Q <mark>K H R A</mark> K I D N
Ascaris	Y E V K L N <mark>V M</mark> S Y	L H Q F H F G V F Y	AFIKLKEQEM	RNIIWIAECI	S Q <mark>R</mark> H R T K I D N
Mosquito	HEVKLN <mark>VY</mark> AF	M Q Q F H F G V F Y	S Y L K L K E Q E C	RNIVWIAECV	AQKHRAKIDN
Giant Honeybee	HEVRLN <mark>VH</mark> AF	L Q Q F H F G V F Y	S Y L K L K E Q E C	RNIVWIAECV	A Q <mark>K H R A </mark> K I D N
Tribolium	HEVRLN <mark>VH</mark> AF	L Q Q F H F G V F Y	S Y L K L K E Q E C	RNIVWIAECV	A Q <mark>K H R A</mark> K I D N
Turkey	HEVKLN <mark>KL</mark> AF	L N Q F H F G V F Y	A F V K L K E Q E C	RNIVWIAECI	AQRHRTKIDN
M.musculus	HEVKLN <mark>KL</mark> AF	L N Q F H F G V F Y	A F V K L K E Q E C	RNIVWIAECI	AQRHRAKIDN
Homo Sapiens	HEVKLN <mark>KL</mark> AF	L N Q F H F G V F Y	A F V K L K E Q E C	RNIVWIAECI	AQRHRAKIDN
Felis catus	HEVKLN <mark>KL</mark> AF	L N Q F H F G V F Y	AFVKLKEQEC	RNIVWIAECI	AQRHRAKIDN
Saccharomyces	L E M E L C <mark>R D</mark> A F	TQQFAISTVW	AWMKSKEQEV	RNITWIAECI	AQNQRERINN
Consistency	8 * 9 8 * 8 5 4 9 9	86**997999	877*9****7	**99*****9	9 * 7 9 * 7 9 * 9 *

The color assignments are:

Unconserved 0 1 2 3 4 5 6 7 8 9 10 Conserved

Notredame *et al.* 2000) or PRALINE. PSI-Coffee alignment figures were produced using Boxshade (Source Forge) and PRALINE (Simossis and Heringa 2003; Simossis *et al.* 2005).

Data availability

Sequences of mutant alleles have been deposited in Genbank (accession nos. KU665627 – *cho¹*, KU682283 – *ma¹*, KU682282 – *mah¹*, *red* KU711835 - *red¹* and *red* KU711836 - *red^{K1}*). Other gene sequences are available by request. Aligned nucleotide sequences for the *red* mutants and OreR are in Figure S1. Table S2 lists the primers used in cloning and sequencing. Nucleotide changes in *cho*, *ma*, and *mah* are in Table S3. The identifiers for all proteins used in alignments are given in Table S4. The distribution of nucleotide substitutions in the two *red* stocks and OreR are shown in Table S5.

RESULTS

cho codes for a subunit of vesicular ATPase

cho is an X-linked recessive eye color mutation that is also associated with brown pigmentation in the Malpighian tubules. It was originally described by Sturtevant (Lindsley and Zimm 1968; Sturtevant 1955). Both the amounts of ommochromes and pteridines present in *cho* eyes are decreased compared to wild type (Ferre *et al.* 1986; Reaume *et al.* 1991). The cytological position of *cho* is 3F1–3F4 and Sturtevant placed it close to *echinus* (*ec*). Three overlapping deficiencies, Df(1)BSC834,

Figure 2 Alignment of predicted partial protein sequences for CHOCOLATE orthologs in insects, fish, reptiles, birds, mammals, and yeast. Multiple sequence alignment, conservation scoring, and coloring were performed by PRALINE. 0 is the least conserved alignment position, increasing to 10 for the most conserved alignment position. Asterisks in the consistency sequence indicate identity in all sequences. The CHO sequence is not included in the consistency rating. The predicted CHO sequence for the region surrounding the missense mutation (shaded) is in the first line. The orthologous sequences were obtained for some vertebrates and yeast. The whole protein is highly conserved and the tryptophan at position 335 is constant except for the CHO sequence which has leucine. Species, gene, and protein identifiers are in Table S4.



Figure 3 The Vps16A wild-type allele complements the *maroon* mutation. (A) Wild-type genotype and phenotype. (B) *ma/ma* mutant genotype plus the Vps16A transgene and P{GAL4-ninaE.GMR}12 driver produces a wild-type phenotype. (C) *ma/ma* mutant genotype and phenotype.

Df(1)ED6716, and Df(1)BSC877, which together removed regions around and including *ec*, were used for the first deletion mapping (Table 2). Only heterozygotes for *cho* and Df(1)ED6716 showed the cho phenotype, indicating that *cho* was in the region removed exclusively by Df(1)ED 6716 (Table 2). Genomic deficiencies were made using the Flp-FRT method (Parks *et al.* 2004). Deletion mapping with these identified a 47.8-Kbp region that contained the *cho* gene. Within it were three candidate genes: *VhaAC39-1*, a subunit of vacuolar ATPase; *CG42541*, a member of the Ras GTPase family; and *CG15239*, which has an unknown function (Table 2).

Each of these genes was amplified from *cho* DNA and sequenced. The coding regions of *CG15239* and *CG42541* did not contain any nonsynonymous sequence changes in the *cho* mutant flies. The coding region of *VhaAC39-1* had a change from guanine to thymine at position X:3,882,405 that resulted in a nonsynonymous change, W330L, in the deduced protein sequence (Table 3 and Table S3). The UAS-VhaAC39-1 transgene (carried on chromosome 2) and the Gal4-ninaE.GMR12 driver were used to fully rescue *cho/cho* stocks (Figure 1). The residue at position 330 is highly conserved. Alignments of predicted orthologs from insects, other invertebrates, including yeast and vertebrates all showed a tryptophan at their corresponding positions (Figure 2). The conservation and phenotype change caused by *cho* mutant alleles both show the importance of this tryptophan in protein function.

ma codes for a component of the HOPS complex, *Vps16A*

The *ma* allele is recessive and homozygotes have darker eyes than normal and yellow Malpighian tubules (Bridges 1918). Homozygotes also have decreased amounts of both pteridines and ommochromes (Nolte 1955). Deletion mapping results (Table 2) indicated that the region in which *ma* resides contains *Vps16A*, a gene involved in endosomal transport and eye pigment (Pulipparacharuvil *et al.* 2005); *Aats-trp*, a tRNA synthase (Seshaiah and Andrew 1999); *CG8861*, a predicted Lymphocyte Antigen super family 6 gene (Hijazi *et al.* 2009); *HP1E*, a member of the heterochromatin protein 1 family (Vermaak *et al.* 2005); and *CG45050*, a gene coding for a protein with Zinc finger domains (dos Santos *et al.* 2015).

The Vps16A gene was sequenced in ma flies because of its role in the HOPS complex. The DNA sequence has 10 nucleotide substitutions and a 9-bp deletion (Table S3). The protein has three missense substitutions and a deletion of three contiguous amino acids (Table 3). Homozygous mutant ma^1 flies had a near wild-type phenotype when rescued by the UAS-Vps16A transgene driven by the Gal4-ninaE. GMR12 driver (Figure 3).

The cause of impaired protein function is unclear. The first two substitutions (M4I and A24T) are not conservative and are found in variable regions in the protein (Figure 4A). The aspartic acid of the E712D substitution is found in the corresponding site in the majority of

Drosophila species and in many other species (Figure 4B). The deletion sequence is in a region with moderately conservative changes between amino acids and it does lie between two regions that are conserved from fly to man (Figure 4C). The spacing between these conserved regions is

Α	10		B	B 790800
MAROON	MPIIYNT	SEW <mark>FKVRPD</mark> - YYRK		R S D F K V P D R R F
Derec	MPIMYNT	GEW <mark>FKVRPD</mark> - YYRK	VELATP	R S DFKVPERRF
Dyaku	MPIMYNT	SEW <mark>FKVRPD</mark> - YYRK	V <mark>EL</mark> ATP	R S DFKVPERRF
Dmela	MPIMYNT	GEW <mark>FKVRPD</mark> - YYRK	V <mark>EL</mark> ATP	R S EFKVPDRRF
Dsech	MPIMYNT	SEW <mark>FKVRPD</mark> - YYRK	V <mark>EL</mark> ATP	R S DFKVPDRRF
Dsimu	MPIMYNT	GEW <mark>FKVRPD</mark> - YYRK	V <mark>EL</mark> ATP	R S DFKVPDRRF
Dpseu	MPIMYNT	SEW <mark>FKVRPD</mark> - YYRK	V <mark>EL</mark> ATP	R N DYKVPDRRF
Dpers	MPIMYNT	GEW <mark>FKVRPD</mark> - YYRK	V <mark>EL</mark> ATP	R N DYKVPDRRF
Danan	MPIMYNT	SEW <mark>FKVRPD</mark> - YYRK	V <mark>EL</mark> PTP	R N GFKVPERRF
Dgrim	MPIMENT	GEW <mark>FKVRPD</mark> - YYRK	VELETP	K S DYKVPDRRY
Dmoja	MPIMNNT	GEW <mark>FKVRPD</mark> - YYRK	VELATP	K S DYKVPDRRF
Dviri	MPIMYNT	GEW <mark>FNVQPD</mark> - YYRK	VELATP	K T EFNVPDRRF
Dwill	MPIMYNT	GEW <mark>FKVRPD</mark> - YYRK	V <mark>EL</mark> DTP	K N DYKVPDRRF
Musca	MPIMYNT	GDW <mark>FVLRPT</mark> - TYRK	F <mark>DL</mark> YPP	K N DFKVPDRRF
Aedes	MSLLYNT	GDWFS <mark>LGHGN</mark> SYRK	IDLYTM	K S EFKVPERRY
Tribolium	MSAAMLT	ADWFL <mark>lgrdl</mark> yfrki	FEIYTM	R N DYKIPDKRF
Apis dorsata	M-LLMLT	ADCFP <mark>lgrdv</mark> yfrki	F <mark>el</mark> ypl	K S EYRISDRRY
Poecilia	<mark>MAFIT</mark>	ANW <mark>NPLG-DA</mark> FYRK	TELYEM	Y R DFRVPDKRY
Danio rerio	<mark>MAFV</mark> T	ANWNPLGEA - FYRK	I <mark>el</mark> yem	Y K DFRVPDKRF
Xenopus	<mark>MDFY</mark> T	ANWYP <mark>MGSV</mark> - YYRKI	FDCYRM	Y K EFKIPDKRY
Takifugu	<mark>MTFIT</mark>	ANW <mark>NPLGD</mark> A - FYRK	TELYDM	Y R DFKVPEKRY
Anolis	<mark>SNMDCYT</mark>	ANWDPLGEAA YYRK		Y R DEKIPDKRE
Gallus gallus	<mark>MDCY</mark> T	ANWNPLGEEA FYRKI	FELYTM	Y R DEKIPDKRY
Macaca	<mark>MDCY</mark> T	ANWNPLGDSA FYRK	YELYSM	A R DFRIPDKRL
Alligator	<mark>MAC</mark> -S	NNWNPLGEQT FYRR	S <mark>EL</mark> YSM	Y R EFKIPDKRF
Mus musculus	<mark>MDCY</mark> T	ANWNP <mark>lgdsa</mark> fyrk	YELYSM	A R DFRIPDKRL
Felis catus	<mark>MDCY</mark> T	ANWNPLGDSA FYRK	YELYSM	A R DFRIPDKRL
Bos taurus	<mark>MDCY</mark> T	ANWNPLGDSA FYRK	YELYSM	A R DFRIPDKRL
Homo sapiens	<mark>MDCY</mark> T	ANWNPLGDSA FYRK	YELYSM	A R DFRIPDKRL
Rattus	<mark>MDCY</mark> T	ANWNPLGDSA FYRK	YELYSM	A R DERIPOKRL
Caenorhabditi	GELCL RP	SS <mark>SVFLG</mark> - <mark>DQ</mark> QLYF	TQEYLR	K Q QARLTDKQV
Ascaris suum	CLMGSEL	QDW <mark>VYLD</mark> - <mark>TL KLRK</mark> I	D <mark>VL</mark> FKR	K K NFKLNDKLC
Saccharomyce	<mark>MKNP</mark> S	FDW <mark>ERLK</mark> - <mark>DV</mark> FYRSI	RAIGEL	V K KFKISEKKL
Consistency	0002274448	668 <mark>4485351</mark> 6898	588555	4 <mark>5 78898889</mark> 6

Figure 4 A PRALINE alignment of regions of the predicted MAROON protein which have changed compared to the D. melanogaster sequence. Alignment, conservation scoring and coloring was performed by PRALINE. The *ma* sequence is not included in the consistency rating. 0 is the least conserved alignment position, increasing to 10 for the most conserved alignment position. Asterisks in the consistency line indicate identity for all sequences. Species, gene and protein identifiers are in Table S4. MA amino acid changes are shaded black in the first line for each alignment. (A) The first two amino acid changes are in regions with low conservation. Note, however, that the residue at position 7 is not found in any other species. The M at that site is conserved in all Drosophila. The A to T change at position 28 is also not found in other species. The position is variable in insects, but vertebrates usually have a Y. (B) The E to D change at position 797 is in a moderately conserved area. D is found in many organisms at this site. (C) The deletion of three amino acids in the MA protein shown in line one at positions 461-463 lies between two sets of conserved sequences, 1, 2, 3, and 4 (underlined). The distance between the 2 and 3 regions is conserved in all species shown except C. elegans and S. cerevisiae.

MANDON EVKSDEVLSM CRE KTOLA VSETEARSY EFCPETQKS LET YVEGG INN HNDE Y NRLENV NTLENERTA MPTTFKGFSH Draku YKADEVLSM CRE KTOLA VSECTEARSY EFCPETQKS LET YVEGG INN HNDE YKADEVLSM CRE KTOLA VSECTEARSY EFCPETQKS LET YVEGG INN HNDE YKADEVLSM CRE KTOLA VSECTEARSY EFCPETQKS LET YVEGG INN HNDE YMTINELENV NTLENERTA MPTTFKGFSH Damb YKSDEVLSM CRE KTOLA VSECTEARSY EFCPETQKS LET YVEGG INN HNDE YMTINELENV NTLENERTA MPTTFKGFSH Damb YKSDEVLSM CRE KTOLA VSECTEARSY EFCPETQKSL LET YVEGG INN HNDE YMTINELENV NTLENERTA MPTTFKGFSH Dama YKSDEVLSM CRE KTOLA VSECTEARSY EFCPETQKSL LET YVEGG INN HNDE YMTINELENV NTLNHERTA MPTTFKGFSH Dama YKSDEVLSM CRE KTOLA VSECTEARSY EFCPETQKSL LET YVEGG INN HNDE YMTINELENV NTLNHERTA MPTTFKGFSH Dama YKSDEVLSM CRE KTOLA VSECTEARSY EFCPETQKSL LET YVEGG INN HNDE YMTINELENV NTLNHERTA MPTTFKGFSH	С	41	0	420	430	440	0 456		460		. 470	48	0	490
Derec SYKSDEVISM CREKIDLA VAECIEAASY EFCPETKSS LIKTWYEKGG IRRMNDDE YRINRILRV LNTLHHERIA MPITFKQFSH Deah SYKSDEVISM CREKIDLA VSECIEAASY EFCPETKSS LIKTWYEKGG IRNMNDDE YRINRILRV LNTLHHERIA MPITFKQFSH Deah SYKSDEVISM CREKIDLA VSECIEAASY EFCPETKSS LIKTWYEKGG IRNMNDDE YRINRILRV LNTLHHERIA MPITFKQFSH Deah SYKSDEVISM CREKIDLA VSECIEAASY EFCPETKSS LIKTWYEKGG IRNMNDDE YRINRILRV LNTLHHERIA MPITFKQFSH Deam SYKSDEVISM CREKIDLA VSECIEAASY EFCPETKSS LIKTWYEKGG IRNMNDDE YRINRILRV LNTLHHERIA MPITFKQFSH Deam SYKSDEVISM CRDRIDLA YTECIEAASF EFCTETKSS LIKTWYEKGG IRNMNDDE YRINRILRV LNTLHHERIA MPITFKQFSH Deam SYKSDEVISM CRDRIDLA YTECIEAASF EFCTETKSS LIKTWYEKGG ILNMNDE YRINRILRV LNTLHHERIA MPITFKQFSH Deam SYKSDEVISM CRDRIDLA YTECIEAASF EFCTETKSS LIKTWYEKGG ILNMNDE YRINRILRV LNTLHHERIA MPITFKQFSH Deam SYKSDEVISM CRDRIDLA YTECIEAASF EFCTETKSS LIKTWYEKGG ILNMNDE YRINRILRV LNTLHHERIA MPITFKQFSH Deam SYKSDEVISM CRTNIELA WECIEAASY EFCPETKSS LIKTWYEKGG ILNMNDE YRINRILRV LNTLHHERIA MPITFKQFSH Dama SYKSDEVISM CRTNIELA WECIEAAAY EFCPETKSS LIKTWYEKGG ILNMNDE YRINRILRV LNTLHHERIA MPITFKQFSH Dama SYKSDEVISM CRTNIELA WECIEAAAY EFCPETKSS LIKTWYEKGG ILNMNDE YRINRILRV LNTLHHERIA MPITFKQFSH Musa MQADEVISS FRORNNW DECIEAAAY EFCPETKSS LIKTWYEKGG ILNMNDE YRINRILRV LNTLHHERIA MPITFKQFSH Musa MQADEVISS FRORNNW DECIEAAAY EFCPETKSS LIKTWYEKGG ILNMNDE YRINRILRV LNTLHHERIA MPITFKQFSH Musa MAADEVISS FRORNNW DECIEAAAY EFCPETKSS LIKTWYEKGG INGMNDE YRINRILRV LNTLHHERIA MPITFKQFSH Musa MAADEVISS FRORNNW DECIEAAAY EFCPETKSS LIKTWYEKGG INGMNDE YRINRILRV LNTLHHERIA MPITFKQFSH Musa MAADEVISS FRORNNW DECIEAAAY EFCPETKSS LIKTWYEKGG INGMNDE YRINRILRV LNTLHHERIA MPITFKQFSH Musa MAADEVISS KAADEVISS RAALAY EFCPETKSS LIKTWYEKGG ING NAPEE YNNDE YERKUR LNTLHKKIA MPITFKQFSH Musa MAADEVISS KAADEVISS RAALAY EFCPETKSS LIKAWEKSS LAAASEKGF DAN-ANDEY YRINRILRV LAALMPHTA YDYTKG MI MAADAYADY NG MAAAYA FCPETKKS LIKAWEKSS LAAASEKKC DAN-EPSPE YNNOCOLN LAALMPHTA YDYTKGG INTLHKYEK MAADEYSKYK YKK NGA YNNY YCHYKKAA YNNY YNY K	MAROON	S Y K S D E Y L <mark>S M</mark>	C R E K I	DLA VSECI	EAA <mark>S</mark> Y EF	<mark>C</mark> P E T <mark>Q K S L</mark>	L R T <mark>A </mark> Y F G K <mark>G</mark> F	I R N	H N <mark>P D E N</mark>	Y MR I L	RV LNT	LRHERIA	MPITFK	Q F S H
Dowa SYKADEVISM CREKIDLA VSECIEAASY EFCPETKKS. LETKYVEKGE Dowa SYKSDEVISM CREKIDLA VSECIEAASY EFCPETKKS. LETKYVEKGE TRRMNDE VRINRTLRV LITLHAERIA MPITKKES Dowa SYKSDEVISM CREKIDLA VSECIEAASY EFCPETKS. LETKYFEKGE TRRMNDE VRINRTLRV LITLHAERIA MPITKKES Dowa SYKSDEVISM CREKIDLA VSECIEAASY EFCPETKS. LETKYFEKGE TRRMNDE VRINRTLRV LITLHAERIA MPITKKES Dowa SYKSDEVISM CRDRIDLA TELIEAASF EFCTETKS. LETKYFEKGE TRRMNDE VRINRTLRV LITLHAERIA MPITKKES Dowa SYKSDEVISM CRDKIDLA VELIEAASF EFCTETKS. LETKYFEKGE TRRMNDE VRINRTLRV LITLHAERIA MPITKKES Dowa SYKSDEVISM CRSNIELA VMELIEAASF EFCTETKS. LETKYFEKGE TRRMNDE VRINRTLRV LITLHAERIA MPITKKES Dowa SYKSDEVISM CRSNIELA VMELIEAASF EFCTETKS. LETKYFEKGE TRRMNDE VRINRTLRV LITLHAERIA MPITKKES Dowa SYKSDEVISM CRSNIELA VMELIEAASF EFCTETKS. LETKYFEKGE TRRMNDE VRINRTLRV LITLHAERIA MPITKKES Dowa SYKSDEVISM CRSNIELA VMELIEAASY EFCPTKS. LETKYFEKGE TRRMNDE VRINTLRV LITLHAERIA MPITKKES Dowa SYKSDEVISM CRSNIELA VMELIEAASY EFCPTKS. LETKYFEKGE TRRMNDE VRINTLRV LITLHAERIA MPITKKES Dowa SYKSDEVISM CRSNIELA VMELIEAASY EFCPTKS. LETKYFEKGE Dowa SYKSDEVISM CRSNIELA VMELIEAASY EFCPTKS. LETKYFEKGE Dowa SYKSDEVISM CRSKILA VMELIEAASY EFCPTKS. LETKYFEKGE TRXMNDE VTRILRV LAVKVEKTA MPITKKES DOWA SYKSDEVISM CRSKILA VMELIEAASY EFCPTKS. LEASFEKCE DOWA SYKSDEVISM CRSKILA VMELIEAASY EFCFTKS. LEASFEKCE DOWA SYKSDEVISM CRSKILA VMELIEAASY EFCFTKS. LEASFEKCE DOWA SYKSDEVISM CRSKILA VMELIEAASY EFCFTKS. LEASFEKCE DOWA SYKSDEVISM LIKAKEKEKY LICKYFEKS TRAMNDE YKRETTRVK LIKA	Derec	S <mark>Y</mark> KSDEYL <mark>SM</mark>	CREKI	DLA V <mark>A</mark> ECI	EAA <mark>s</mark> y ef	C P E T Q K S L	L R T <mark>A Y</mark> Y G K <mark>G</mark> F	I R N	HN <mark>PDE N</mark>	YI <mark>RIM</mark> RI <mark>L</mark>	RV LNT	LRHERIA	MPITF	Q F S H
Deeda SYKSDEVLSM CREKIDLA VSECIEAGY EFCPETOKS. LETMYEGKGE IRMHNDE VMRIMRILRV LMILHERIA MPITEKOFSM Dimu SYKSDEVLSM CREKIDLA VSECIEAGY EFCPETOKS. LETMYEGKGE IRMHNDDE VMRIMRILRV LMILHERIA MPITEKOFSM Dimu SYKSDEVLSM CREKIDLA VSECIEAGY EFCPETOKS. LETMYEGKGE IRMHNDDE VMRIMRILRV LMILHERIA MPITEKOFSM Dimu SYKSDEVLSM CREKIDLA VSECIEAGY EFCPETOKS. LETMYEGKGE IRMHNDDE VMRIMRILRV LMILHERIA MPITEKOFSM Dimu SYKSDEVLSM CREKIDLA VSECIEAGY EFCPETOKS. LETMYEGKGE ILMHNDDE VMRIMRILRV LMILHERIA MPITEKOFSM Dimu SYKSDEVLSM CREKISLA VDECIEAGY EFCPETOKS. LETMYEGKGE ILMHNDDE VMRIMRILRV LMILHERIA MPITEKOFSM Dimi SYKSDEVLSM CREKILA VNECIEAGY EFCPETOKS. LETMYEGKGE ILMHNDE VMRIMRILRV LMILHERIA MPITEKOFSM Dimi SYKSDEVLSM CREKILA VNECIEAGY EFCPETOKS. LRTMYEGKGE ILMHNDE VMRIMALLV LMILHERIA MPITEKOFSM Dimi SYKSDEVLSM CREKILA VNECIEAGY EFCPETOKS. MRTMYEGKGE ILMHNDE VMRIMALLV LMILHERIA MPITEKOFSM Dimi SYKSDEVLSM CREKILA VNECIEAGY EFCPETOKS. MRTMYEGKGE ILMHNDE VMRIMALLV LMILHERIA MPITEKOFSM Dimi SYKSDEVLSM CREKILA VNECIEAGY EFCPETOKS. MRTMYEGKGE IDHNPEE VMRILRILRV LMILHERIA MPITEKOFSM Dimi SYKSDEVLSM CREKILA VNECIEAGY EFCPITOKS. MRTMYEGKGE IDNNPEE VMRILRILRV LMILHERIA MPITEKOFSM Dimi SYKSDEVLSK CREKILA VNECIEAGY EFCPITOKS. MRTMYEGKGE IDNNPEE VMRILRILRV LMILHERIA MPITEKOFSM Dimi SYKSDEVLSK CREKILA VNECIEAGY EFCPITOKS. MRTMYEGKGE IDNNPEE VMRILRIZUL NILHERIA MPITEKOFSM Dimi SYKSDEVLSK CREKILA VNECIEAGY EFCPITOKS. MRTMYEGKGE IDNNPEE VMRILRIZUL NILHERIA MPITEKOFSM Dimi SYKSDEVLSK CREKILA VNECIEAGY EFCPITOKS. MRTMYEGKGE IDNNPEE VMRILRIZUL NILHERIA MPITEKOFSM Dimi SYKSDEVLSK CREKILA VNECIEAGY EFCPITOKS. MRTMYEGKGE IDNNPEE VMRILRIZUL NILHERIA MPITEKOFSM Dimi SYKSDEVLSK CREKILA VNECIEAGY EFCPITOKS. MRTMYEGKGE IDNNPEE VMRILRIZUL NILHKEKIA MPITEKOFSM Dimi SYSSECHTER KOSTALSKA VGEVAGAM EYOPDTOKAL MRAKEKGE IDNNPEE VMRILRIZUL NILHKEKIA MPITEKOFSM Dimi SYSSECHTER VKOLEMA VGEVAGAM EYOPDTOKAL MRAKEKGE IDANNPEE VMRICALURV NARKEKGE IDNNPEE VMRICALURV NARKEK	Dyaku	S <mark>y</mark> k a d e y l <mark>s m</mark>	CREKI	DL <mark>A V</mark> SECI	EAA <mark>s</mark> y ef	C P E T Q K S L	L R T <mark>A Y</mark> Y G K <mark>G</mark> F	I R N I	HN <mark>PDK </mark>	YI <mark>RIM</mark> RI <mark>L</mark>	RV LNT	LRHERIA	MPITFK	QFSH
Dach SYKSDEVLSM CREKIDLA VSETIEAGY EFCPETOKS. LETHVEGGE TRMHNDE VMRTMRLEV. LATLRHERIA MPITEKOFSM Demu SYKSDEVLSM CREKIDLA VSETIEAGY EFCPETOKS. LETHVEGGE TRMHNDE VMRTMRLEV. LATLRHERIA MPITEKOFSM Demu SYKSDEVLSM CREKIDLA VTETEAGSF EFCPETOKS. LETMVEGGE TRMHNDE VMRTMRLEV. LATLRHERIA MPITEKOFSM Demo SYKSDEVLSM CREKIDLA VTETEAGSF EFCPETOKS. LETMVEGGE TRMHNDE VMRTMRLEV. LATLRHERIA MPITEKOFSM Demo SYKSDEVLSM CREKIDLA VTETEAGSF EFCPETOKS. LETMVEGGE TRMHNDE VMRTMRLEV. LATLRHERIA MPITEKOFSM Demo SYKSDEVLSM CREKIDLA VETEAGSF EFCPETOKS. LETMVEGGE TLFHNDE VMRTMRLEV. LATLRHERIA MPITEKOFSM Demo SYKSDEVLSM CREKIGLA VNETEAGY EFCPETOKS. LATMVEGGE TLFHNDE VMRTMRLEV. LATLRHERIA MPITEKOFSM Demo SYKSDEVLSM CREKIGLA VNETEAGY EFCPETOKS. LATMVEGGE TLFHNDE VMRTMRLEV. LATLRHERIA MPITYKOFSM DMM SYKSDEVLSM CREKIGLA VNETEAGY EFCPETOKS. LATMVEGGE TLFHNDE VMRTMRLEV. LATLRHERIA MPITYKOFSM DMM SYKSDEVLSM CREKIGLA VNETEAGY EFCPETOKS. LATMVEGGE TLFHNDE VMRTMRLEV. LATLRHEKA MPITYKOFSM DMM SYKSDEVLSM CREKIGLA VNETEAGY EFCPETOKS. LATMVEGGE TLFHNDE VMRTMRLEV. LATLRHEKA MPITYKOFSM DMM SYKSDEVLSM CREKIGLA VNETEAGY EFCPETOKS. LATMVEGGE TLFHNDE VMRTMLTLEV. LATLRHEKA MPITYKOFSM DMM SYKSDEVLSM CREKIGLA VNETEAGY EFCPETOKS. LATMVEGGE TLFHNDE VMRTMLTLEV. LATLRHEKA MPITYKOFSM DMM SYKSDEVLSM CREKIGLA VNETEAGY EFCPETOKS. LATMVEGKE TLFHNDE VMRTMLTLEV. LATLRHEKA MPITYKOFSM DMM SYKSDEVLSM CREKIGLA VNETEAGY EFCPETOKS. LATMVEGKE TLFNNDE VMRTMLTLEV. LATLRHEKA MPITYLEV PACHD SYKSDEVLSM CREKIGLA VNETEAGY EFCPETOKS. LATMVEGKE TLFNNDE VMRTMLTLEV. LATLRHEKA MPITYLEV PACHD DMM SYKSDEVLSM CREKIGLA VOQUAAAN VETTEGAAY EFCPETOKS. LATMVEGKE TLFCNNDE VMRTMLTRUK LATRNETOKO MBM SS SKADEVLKE TKEGKUGA VQUAAAH VYDDTTAAL MKANSFKC TLFCNNDE VTATLRHEKA NARVFKG SKO PLITTREV PACHD SKADEVLKE TKEGKUGA VQUAAAH VYDDTTAAL MKANSFKC LDFVSDA VYNNLTTRLEV. NARVFKGS SKO PLITTREV PACHD SKADEVLKE TKEGKUGA VQUAAAH VYDDTTAAL MRASFKC LDFSADE VYNCEALKV. NARVFKGS SKO PLITTREV KANDYGU SKADEVLKE TKEGKUGA VQUEAAAH HEPTYTT LARASFKC LDFSAD	Dmela	S <mark>Y</mark> K S D E Y L <mark>S M</mark>	C R E K I	DL <mark>A V</mark> SECI	EAA <mark>s</mark> y ef	C P E T Q K S L	L R T A <mark>Y</mark> F G K <mark>G</mark> F	I R N I	HN <mark>PDE N</mark>	YM <mark>RIM</mark> RI	RV LNT	LRHERIA	MPITFK	QFSH
Domu SYKSDEVLGN CREKIDLA VSECTEAASY EFCPETOKS. LETENYEKKEF INNHHPDE YMRINRILRV LNTLRHERA PPITFKÖFSH Dens SYKSDEVLSN CRDRIDLA VTECIEAASY EFCPETOKS. LETENYEKKEF INHHPDE YMRINRILRV LNTLRHERA PPITFKÖFSH Dana SYKSDEVLSN CRDKNSLA VDECIEAASY EFCPETOKS. LETENYEKKEF INHHPDE YMRINRILRV LNTLRHERA PPITFKÖFSH Dana SYKSDEVLSN CRTNELA VNECIEAASY EFCPETOKS. LETENYEKKEF INHHPDE YMRINRILRV LNTLRHERA PPITFKÖFSH Domi SYKSDEVLSN CRTNELA VNECIEAASY EFCPETOKS. LATENYEKKEF INHHPDE YMRINRILRV LNTLRHERA PPITFKÖFSH Domi SYKSDEVLSN CRTNELA VNECIEAASY EFCPETOKS. LATENYEKKEF INHHPDE YMRINRILRV LNTLRHKRA AQUTYKÖFSH Domi SYKSDEVLSN CRTNELA VNECIEAASY EFCPETOKS. LATENYEKKEF INHHPDE YMRINRILRV LNTLRHKRA AQUTYKÖFSH Domi SYKSDEVLSN CRTNIELA VNECIEAASY EFCPETOKS. LATENYEKKEF INHHPEE YMRINRILRV LNTLRHKKA AQUTYKÖFSH Domi SYKSDEVLSN CRTNIELA VNECIEAASY EFCPETOKS. LATENYEKKEF INHNPEE YMRINRILRV LNTLRHKKA AQUTYKÖFSH Domi SYKSDEVLSN CRTNIELA VNECIEAASY EFCTETOKS. LATENYEKKEF INHNPEE YMRINRILRV LNTLRHKKA AQUTYKÖFSH Domi SYKSDEVLSN CRTNIELA VNECIEAASY EFCTETOKS. LATENYEKKEF INHNPEE YMRINRILRV LNTLRHKKA AQUTYKÖFSH Domi SYKSDEVLSN CRTNIELA VNECIEAASY EFCTETOKS. LATENYEKKEF INNNPEE YMRINRILRV LNTLRHKKA AQUTYKÖFSH Domi SYKSDEVLSN CRTNIELA VNECIEAASY EFCTETOKS. LATENYEKKEF INNNPEE YMRINRILRV LNALNNPRA PPITKRÖFSH MAKO SHQADEVLSI FRDRNNV VDECIEAASY EFCTETOKS. LATENYEKKEF INNNPEE YMRINRILRV LNALNNPRA PPITKRÖFSH MAKO SHQADEVLSI REV VOLONA VQUEAAAH YDPITKRÍF Domi SKADEVLSI VKDKLADA IKACIDASH EFOFETIKLI KAANFEKKEF INNNPE YWNKTRILLV LNALNNPRA BPITFYDEY Poedia SQKADEVLRE IKEQNILGAA VQUEAAAH YDPITKAL HAANFEKKEF INNNPE YWNKTRILLV LNALNNPRA BPITFYDEY Poedia SQKADEVLRE IKEQNILGAA VQUEAAAH YDPITKAL HAANFEKKEF INNNPE YWNKTRILLV LNALNNPRA BPITFYDEY Poedia SQKADEVLRE IKEQNILGAA VQUEAAAH YDPITKAL HAASEKKEF INNPE PPIFFYDE YVNKTRILLNY LNALNNPYG BPITFYDEY Poedia SQKADEVLRE IKEQNILGAA VQUEAAAH YDPITKAL HAASEKKEF INNPE PPIFFYDE YVNKTRILLV LNALNNPYG BPITYDEYN MASSO SQKADEVLRE IKEQNILGAA VQUEAAAH	Dsech	S <mark>Y</mark> KSDEYL <mark>SM</mark>	CREKI	DL <mark>A V</mark> SECI	EAA <mark>s</mark> y ef	C P E T Q K S L	L R T A <mark>Y</mark> F G K <mark>G</mark> F	IRN	HN <mark>PDE N</mark>	YM <mark>RIMRI</mark>	RV LNT	LRHERIA	MPITFK	QFSH
Dpeu SYKSDEVLSM CRDRIDLA VIECTEAASF EFCTETOKSL LATANFGKAF IFGHNPDE YMRIMRILRV LNTLRHERIA MPLIFKQFSH Dana SYKSDEVLSM CRDRIDLA VIECTEAASF EFCTETOKSL LATANFGKAF IFGHNPDE YMRIMRILRV LNTLRHERIA MPLIFKQFSH Dana SYKSDEVLSM CRDKISLA VECTEAASY EFCPTOKSL LATANFGKAF IGHNPDE YMRIMRILRV LNTLRHERIA MPLIFKQFSH Dana SYKSDEVLSM CRSNIELA VHECTEAASY EFCPTOKSL LATANFGKAF ILLHNPEE YMRIMRILRV LNTLRHERIA MPLIFKQFSH David SYKSDEVLSM CRSNIELA VHECTEAASY EFCPTOKSL KATAVFGKGF ILLHNPEE YMRIMRILRV LNTLRHKRIA MQLTYKQFSH Dvid SYKSDEVLSM CRSNIELA VHECTEAASY EFCPTOKSL KATAVFGKGF ILGHNPDE YMRIMRILRV LNTLRHKRIA MQLTYKQFSH Dvid SYKSDEVLSM CRSNIELA VHECTEAASY EFCPETOKSL KATAVFGKGF IGAHNPDE YMRIMRILRV LNTLRHKRIA MPLTFKQFSH Dvid SYKSDEVLSM CRSNIELA VHECTEAAAY FFCPETOKSL KATAVFGKGF IGAHNPDE YMRIMRILRV LNTLRHKRIA MPLTFKQFSH Dvid SYKSDEVLSM CRSNIELA VHECTEAAAY FFCPETOKSL KATAVFGKGF IGAHNPDE YMRIMRILRV LNTLRHKRIA MPLTFKQFSH Dvid SYKSDEVLSM CRSNIELA VHECTEAAAY FFCPETOKSL LATANFGKGF IGAHNPDE YMRIMRILRV LNTLRHKRIA MPLTFKQFSH Dvid SYKSDEVLSM CRSRIGLA VNECTEAAAY FFCPETOKSL LATANFGKGF IGAHNPDE YMRIMRILRV LNALRHKRIA PPLTFKQFSH Dvid SYKSDEVLSM CRSRIGLA VNECTEAAAY FFCPETOKSL LATANFGKGF IGAHNPDE YMRIMRILRV LNALRHKRIA PPLTFKQFSH Dvid SYKSDEVLSM CRSRIGLA VNECTEAAAY FFCPETOKSL LATANFGKGF IGAHNPDE YMRIMRILRV LNALRHKRA PPLTFKQFSH Dvid SYKSDEVLSM CRSKIGAA VQCTEAAAY FFCPETOKSL LATANFGKGF SKTHNPDE YMRIMRILRV LNALRHFKA PPLTFKQFSH Dvid SYKSDEVLSM LYCKIDAA IKACIDGASH FFOFETOKIL MKARYFKKFF LGGHNPDE YMRIMRILRV LNALRFPIG PPLTKRFT AGGS SKADEVLRE IKEQKLESA VQCTEAAH FYDPTOKIL MKARYFKKFF LGGHNPDE YMRIMRILRV LNALRFFIG PPLTKRFT Dvid SYKSDEVLSM LYCKIDAA IKACIDGASH FOFETOKIL MKARYFKKFF LGGHNPDE YWNIRCHTRV LNALRFFTG DVIT ADAGS SKADEVLRE IKEQKLESA VQCTEAAH HYDDYDYSSI LRAASFGKCF LGNFNPD FYTTCRELRV LNALRFFTG PPLTKTILRFFT ANGIN SYKSDEVLRE IKEQKLESA VQCTEAAH HYDDYDYSSI LRAASFGKCF LGNFPPS FVHTQQURV LNALRFFTG PPLTKTILFTY ANGINS SKADEVLRE IKEQKLESA VQCTEAAGY HHPPTYSSI LRAASFGKCF LGNFPPS FVHTQQURV LNALRFFTG IPLTYSYKG ADAGSY	Dsimu	S <mark>Y</mark> K S D E Y L <mark>S M</mark>	CREKI	DL <mark>A V</mark> SECI	EAA <mark>s</mark> y ef	C P E T Q K S L	L R T A <mark>Y</mark> F G K <mark>G</mark> F	IRN	HN <mark>PDE N</mark>	YM <mark>RIMRI</mark>	RV LNT	LRHERIA	MPITFK	QFSH
Ders SYKSDEVLSK CRDKIDLA VTELIEAASY EFCTETIKSL LATAHFEKAF IEG-HNDDE YMRINRILKV LATLAHERIG IPLIFKQFTH Dana SYKSDEVLSK CRDKIDLA VDECIEAASY EFCTETIKSL LATAHFEKAF IEGHNDDE YMRINRILKV LATLAHERIG IPLIFKQFTH Dana SYKSDEVLSK CRTNIELA VNECIEAASY EFCTETIKSL LATAHFEKAF ILNHPDE YMRINRILKV LATLAHERIG IPLIFKQFTH Dana SYKSDEVLSK CRTNIELA VNECIEAASY EFCTETIKSL LATAHFEKAF ILNHPDE YMRILRILKV LATLAHERIG IPLIFKQFTH Dana SYKSDEVLSK CRTNIELA VNECIEAASY EFCTETIKSL LATAHFEKAF ILNHPDE YMRILRILKV LATLAHERIG ADLIFKQFTH Dani SYKSDEVLSK CRTKIGLA VNECIEAASY EFCTETIKSL LATAHFEKAF IFDHNPDE YMRILRILKV LATLAHERIG ADLIFKQFSH Dani SYKSDEVLSK CRTKIGLA VNECIEAASY EFCTETIKSL LATAHFEKAF INDNHPDE YMRILRILKV LATLAHERIG ADLIFKQFSH Dani SYKSDEVLSK CRTKIGLA VNQCIEAASY EFCTETIKSL LATAHFEKAF IFDHNPDE YMRILRILKV LATLAHERIG ADLIFKQFSH Dani SYKSDEVLSK CRTKIGLA VNQCIEAASY EFCTETIKSL LATAHFEKAF INDNHPDE YMRINTILKV LATLAHERIG ADLIFKQFSH Musca SHQADEVLSL FRDRMWA DECIEAAGY EFCTATUKSL KARAYFEKAF INDYMNSDK YWIIRLKV LATAHPRIA VPITIRFYF Tribuim SHRANEYICL VQDLAKA DQCINAAGY EFDPEVYKMI IRAAFFEKAF INDYMNSDK YWIIRLKV LAAVHPAIG IPLIFYQFFH Dani SKADGYDL VKDKLQAA IXACIDGASH EFDFETOKIL MKAANFEKAF SKTIDPEY YWNGRTLRV LAAVHPAIG IPLIFYQFFY Poecilia SKADGYDL VKDKLQAA IXACIDGASH EFDFETOKIL MKAANFEKAF ISTIDPEY YWNGRTLRV LAAVHPAIG IPLIFYQFY Poecilia SKADGYLRE IKEQULDAA VRQUVEAAAH EYDPTQTAI WKAASFEKCF ISNFPPEQ FVSNCRDLRV LAAVHPAIG IPLIFYQFYY Poecilia SKADGYLRE IKEQULDAA VRQUVEAAAH EHPPETOKIL LAASFEKCF IDFFSADP FVECORLRV LAAVNOVNG IPLIFYQYYQ Tahigua SKADGYLRE IKQULDAA VQQUEAAGH EHPETOKIL LAASFEKCF IDKFPPES FVKRCQDLRV LAAVNOVNG IPLIFYQYYQ Tahigua SKADGYLRE IKQULDAA VQQUEAAGH EHPETOKIL LAASFEKCF IDKFPPES FVKRCQDLRV LAAVNOVNG IPLIFYQYYQ Tahigua SKADGYLRE IKQULDAA VQQUEAAGH EHPETOKIL LAASFEKCF IDKFFPES FVKRCQDLRV LAAVNOVNG IPLIFYQYYQ Tahigua SKADGYLRE IQELGQUTQA VQQUEAAGH EHPETOKIL LAASFEKCF IDKFFPES FVKRCQDLRV LAAVNOVNG IPLIFYQYYQ Tahigua SKADGYLRE IQELGQUTQA VQQUEAAGH EHPETOKIL LAASFEKCF IDKFFPES FVKRCQDLRV LAAVNOVNG IPLIFYQY	Dpseu	S <mark>y</mark> k s d e y l <mark>s m</mark>	CRD RI	DLA V <mark>TECI</mark>	EAA <mark>s</mark> f ef	C T E T Q K S L	L R T A <mark>N</mark> F G K <mark>A</mark> F	IFG	HN <mark>PDE N</mark>	YM <mark>RIMRI</mark>	RV LNT	LRHERIA	MPLTFK	QFSH
DANAN SYKSDEVLSK CRDKNSLA VDECIEAASY EFCPETOKSL LRTHMFEKGF ILNHNDDE YMRINRILKV LKTLKMERIG IPLIFKGFTH Dorin SYKSDEVLSK CRTNIELA VNECIEAASY EFCPETOKSL MRTNYFKGF ILEHNDDE YMRINRILKV LKTLKMERIA MQLTYKGFHH Dorin SYKSDEVLSK CRSDIELA VNECIEAASY EFCPETOKSL MRTNYFKGF ILEHNDDE YMRINRILKV LKTLKMERIA MQLTYKGFHH DWII SYKSDEVLSK CREKIGLA VNECIEAASY EFCPETOKSL KKANYFGKAF IPSHNPDE YMRILRILKV LKILKMERIA MPLTYKGFSH Musca SHQADEVLSK CREKIGLA VNECIEAASY EFCPETOKSL KKANYFGKAF IPSHNPDE YMRILRILKV LKILKMERIA MPLTYKGFSH Musca SHQADEVLSK CREKIGLA VNECIEAAGY EFCPETOKSL KRANYFGKAF IPSHNPDE YMRILRILKV LKILKMERIA MPLTYKGFSH Musca SHQADEVLSK CREKIGLA VNECIEAAGY EFCFITOKSL KRANYFGKAF IPGHNPDE YMRINRILKV LKILKMERIA MPLTYKGFSH Musca SHKSDEFLCL IQORLASA VGEVDAAGQ EFDSHTOKSL IRAAMFGKGF IAYMNSDE YIMCKVIKV LNALKPERIA MPLTYKGFSH Musca SHKSDEFLCL IQORLASA VGEVDAAGQ EFDSHTOKSL IRAAMFGKGF IXY-MNSDK YVNIIRLKV LKALKPRIA VPLTIKGFTN Tribolum SHKADGYMDI VKDKLDAA IXACIDGASH EFDFETOKLL MKANKFGKGF ISKT-IDPEY YVNGCRILKV LKAVKPRAG IPLTYGFTN Poecilis SKADGYMDI VKDKLDAA IXACIDGASH EFDFETOKLL MKANKFGKGF ISKT-IDPEY YVNGCRILKV LKAVKPFAG IPLTYGFTY Poecilis SKADGYNDI VKDKLDAA IXACIDGASH EFDFETOKLL MKANKFGKGF VEKYSPAD FVTTCKELKV LKAVKPSSG MPITHTGFKQ Donioren SQKADDYLRE IKEQUVLGAA VREVEAAGH EHOPETOKSL LRAASFGKCF VEKYSPAD FVTTCKELKV LKAVKVSSG UPITHTGFKQ Senopus SQKADDYLRE IKEQULEAA VQEVEAAGH EHOPETOKSL LRAASFGKCF VEKYSPAD FVTTCKELKV LKAVKVSSG LPITYDGFKK Anolis SQKADEYLRE IKEQULEAA VQEVGAARY EVDPQTQKSL LRAASFGKCF IDKFFPDS FVHKCQOLKV LKAVKDYHG IPITYSQYKQ Boliusgalus SQKADEYLRE IQELGQUTQA VQQIEAAGH EHOPETOKSL LRAASFGKCF IDRFFPDS FVHKCQOLKV LKAVKDYHG IPITYSQYKQ Bosturus SQKADEYLRE IQELGQUTQA VQQIEAAGH EHOPETOKSL LRAASFGKCF IDRFFPDS FVHKCQOLKV LKAVKDYHG IPITYSQYKQ Bosturus SQKADEYLRE IQELGQUTQA VQQIEAAGH EHOPETOKSL LRAASFGKCF IDRFFPDS FVHKCQOLKV LKAVKDYHG IPITYSQYKQ Bosturus SQKADEYLRE IQELGQUTQA VQQIEAAGH EHOPETOKSL LRAASFGKCF IDRFFPDS FVHKCQOLKV LKAVKDYHG IPITYSQYKQ Bosturus SQKADEYLRE IQELGQUTQA VQQIEAAGH EHOPETOKSL L	Dpers	S <mark>Y</mark> KSDEYL <mark>SM</mark>	C R D R I	DLA V <mark>TECI</mark>	EAA <mark>s</mark> f ef	C T E T Q K S L	L R T A <mark>N</mark> F G K <mark>A</mark> F	I F G	HN <mark>PDE N</mark>	YM <mark>RIM</mark> RI	RV LNT	LRHERIA	MPLTF	QFSH
Dgrim ŠVKADEVISM CRTNIELA VNECIEAAGY EFCPTQKL NETAVGRGF ILENNPDE VLRIRILRV INTLENKRIA MQLTYKGFMH Dmaja SVKADEVISM CRTNIELA VNECIEAAGY EFCPTQKSL NETAVGRGF IPLNNPEE VNRILRILRV INTLENKRIA MQLTYKGFMH DWI SVKSDQVISI CRSDIELA VNECIEAAGY EFCPTQKSL NEAMYFGKGF IPLNNPEE VNRILRILRV INTLENKIA MPLTYKGFM Musco SHQADEVISI CRSDIELA VNECIEAAGY EFCTTQKSL NEAMYFGKGF INGNNPDE VNRINTILRV INTLENKIA MPLTYKGFM Musco SHQADEVISI CRDRIASA VGEVDAAGQ EFDSNTQKSL IRANFGKGF INGNNPDE VNRINTILRV INTLENKIA MPLTYKGFM Musco SHKSDEFLCI 1QDRIASA VGEVDAAGQ EFDSNTQKSL IRANFGKGF INGNNPDE VNRINTILRV INTLENKIA MPLTYKGFM Adeds SHKSDEFLCI VKODIAKA VDCINAKGY EFDFTQKI IRANFGKGF INGNNPDE VNRINTILRV INALENPRIA VPITIKGFM Tibbium SHRANEVICI VKODIAKA VDCINAKGY EFDFTQKI IRANFGKGF INGNNPDE VNNRTRI VI NANRHAKIG IPLTFTQLFF Adeds SHKSDEFLCI VKODIAKA VDCINAKGY EFDFTQKI IRANFGKGF INGNNDGK VVNIIRLLRV INANRHAKIG IPLTFTQLFF Poecila SQKADEVIRE IKEGNVGDA VRQEVEAAAH EYDPDTQKAL NKANKFGKGF ITA'NSDK VVNIIRLLRV INANRHAKIG IPLTFTQLFF Poecila SQKADEVIRE IKEGNUGDA VRQEVEAAAH EYDPDTQKAL NKANKFGKGF INGFPPEQ FVSMCRDLRV INANRHAKIG IPLTFTQLFK Panoja SQKADEVIRE IKONNIGDA VKOLVEAAAH EYDPDTQKAL NKANKFGKGF INGFPPEQ FVSMCRDLRV INANRHAKIG IPLTFTQLFK Andis SQKADEVIRE IKONNIGDA VKOLVEAAAH EYDPDTQKAL IRAASFGKCF INGFPPEQ FVSMCRDLRV INANRHAKIG IPLTFTQFKK Masoca SQKADEVIRE IKONNIGDA VKOLVEAAAH EYDPDTQKSI IRAASFGKCF IDKFPPES FVETCRDLRV INANRHAKIG IPLTYSGYKG Masousous SQKADEVIRE IQEGQUTQA VQQIEAAGH EHEPTQKSI IRAASFGKCF IDKFPPES FVETCRDLRV INANRHAKIG IPLTYSGYKG Masousous SQKADEVIRE IQEGQUTQA VQQIEAAGH EHQPDMQKSI IRAASFGKCF IDKFPPES FVETCRDLRV INANRHAKIG IPLTYSGYKG Masousous SQKADEVIRE IQEGQUTQA VQQIEAAGH EHQPDMQKSI IRAASFGKCF IDKFPPS FVETCRDLRV INANRDVHIG IPLTYSGYKG Masousous SQKADEVIRE IQEGQUTQA VQQIEAAGH EHQPDMQKSI IRAASFGKCF IDKFPPS FVHCQDLRV INANRDVHIG IPLTYSGYKG Masousous SQKADEVIRE IQEGQUTQA VQQIEAAGH EHQPDMQKSI IRAASFGKCF IDKFPPS FVHCQDLRV INANRDVHIG IPLTYSGYKG Homosopiens SQKADEVIRE IQEGQUTQA VQQIEAAGH EHQPDMQKSI IRAASFGKCF IDKFPPS FVHCQ	Danan	S <mark>Y</mark> KSDEYL <mark>SM</mark>	C R D K M	SLA V <mark>DECI</mark>	EAA <mark>s</mark> y ef	C P E T Q K S L	L R T A <mark>H</mark> F G K <mark>G</mark> F	ILN	HN <mark>PDE N</mark>	YM <mark>RIMRI</mark>	RV LNT	LRHERIG	IPLTFK	QFTH
Dmoja SYKSDEYLSM CRSNIELA VNECTEARAY EFCPETOKSL MARAYEGKGF IPLMAPEE YMAILRILRV LNTLRHEKIA MPLTYKÖFSH DMI SYKSDEYLSM CREKIGLA NNECTEARAY EFCPETOKSL MARAYEGKGF IPLMAPEE YMAILRILRV LNTLRHEKIA MPLTYKÖFSH Musca SKADEYLSL FRDRMWA VDECTEARAY EFCTETOKSL MARAYEGKGF IPAMAPDE YMAILRILRV LNTLRHEKIA MPLTYKÖFSH Musca SKADEYLSL FRDRMWA VDECTEARAY EFCTETOKSL MARAYEGKGF IPAMAPDE YMAILRILRV LNTLRHEKIA MPLTYKÖFSH Musca SKADEYLSL FRDRMWA VDECTEARAY EFCTETOKSL MARAYEGKGF IPAMAPDE YMAILRILRV LNALRPRIA VPLTIKÖFSH Musca SKADEYLSL KEDKLOAA VOCTAAGO EFCSTIVKSL MARAYEGKGF IPAMAPDE YIMALRILRV LNALRPRIA VPLTIKÖFSH Aedes SHKSDEFLCL IQORLASA VGECVDARGO EFCSTIVKSL MARAYEGKGF IPAMAPDE YIMALRILRV LNALRPRIA VPLTIKÖFN Thöbium SKANEYTCL VKQDLAKA VOCTAAGO EFDFEVKLL MAAAFGKGF SKTDDFF YVWACRTLRV LNALRPRIG IPLTYTÖFT Opecilia SQKADEYLRE IKEQULGAA VROCVERAAH EYDPETOKLL MAAAFGKGF SKTDDFF YVWACRTLRV LNAVRESSG MPLIHTÖFKQ Xenopus QKADEYLRE IKEQSULGSA VQCVERAGH HEPETOKSL LRAASFGKCF LSNFPPEQ FYSMCRDLRV LNAVRESSG MPLIHTÖFKQ Yenopus QKADEYLRE IKEQSULEEA VKQCVGARY EVDEPTOKSL LRAASFGKCF VKKSSPAD FVYTKCORLRV LNAVRSSG LPLITYTÖFKM Anolis GKADEYLRE IKDQULLPEA VKQCTGARY EVDEPTOKSL LRAASFGKCF LDNFSADP FVETCRDLRV LNAVRVSSG LPLITYD FKM Anolis QKADEYLRE IKDQULLPEA VKQCTGARY EVDEPTOKSL LRAASFGKCF LDNFSADP FVETCRDLRV LNAVRVSSG LPLITYD FKM Macaca SQKADEYLRE IKDQULLPEA VQCIERAGY HEPETOKSL LRAASFGKCF LDRFSPES FVMRCQDLRV LNAINDVIG IPLITYDYKQ Gallusdafus QKADEYLRE IKDQULGEA VQCIERAAGY HEPETOKSL LRAASFGKCF LDRFSPES FVMRCQDLRV LNAINDVHG IPLITYSYKQ Bostaurus SQKADEYLRE IQELGQLTQA VQCIERAAGH HEPETOKSL LRAASFGKCF LDRFSPES FVMRCQDLRV LNAINDVHG IPLITYSYKQ GostaDEYLRE IQELGQLTQA VQCIERAAGH HEPETOKSL LRAASFGKCF LDRFSPES FVMRCQDLRV LNAINDVHG IPLTYSYKQ Bostaurus SQKADEYLRE IQELGQLTQA VQCIERAAGH HEPETOKSL LRAASFGKCF LDRFSPES FVMRCQDLRV LNAINDVHG IPLTYSYKQ GostaDEYLRE IQELGQLTQA VQCIERAAGH HEPETOKSL LRAASFGKCF LDRFPPDS FVMRCQDLRV LNAINDVHG IPLTYSYKQ GostaDEYLRE IQELGQLTQA VQCIERAAGH HEPETOKSL LRAASFGKCF LDRFPPDS FVMRCQDLRV LNAINDVHG IPL	Dgrim	S <mark>y</mark> kadeyl <mark>sm</mark>	C R T N I	ELA VNECI	EAA <mark>s</mark> y ef	C P V T Q K K L	M R T A <mark>Y</mark> F G K <mark>G</mark> F	ILE	HN <mark>PDE N</mark>	Y L <mark>R I L R I L</mark>	RV LNT	LRNKRIA	MQLTYK	QFMH
DVMI ŠYKSDQYLSL CRS DIELA VNETIEAAAY FCCTETOKSL MKAAYGKAF TPS HNPEE VMRILRILRV LNTLRHOK A MPLTYDFSH MUSCA SMADDYLSL FRD RIASA VGETAAAY FCCTETOKSL LNTNNFGKGF ING HNPDE VMRIMRILRV LNTLRHEKTA MPLTFKGFSH MUSCA SMADDYLCL TQC RLASA VGETVDAAGQ FFD STQKSL LNTNNFGKGF ING HNPDE VMRIMRILRV LNTLRHEKTA MPLTFKGFSH Addes SHKSDEFLCL IQD RLASA VGETVDAAGQ FFD STQKSL INAN FGKGF LPG YNPDG YIRMLRTLRV LNALREPIYG MPLTLRGFNH Tibolim SHRANKYICL VKC DLAKA VDQCINAVGY FED FEVTKIL IKAAFGKGF LPG YNPDG YIRMLRTLRV LNALREPIYG MPLTLRGFNH Tibolim SKRANKYICL VKC DLAKA VQCURAAH EYD PDTQKAL MRARSFGKGF LTE FSPDQ FVTTCRELRV LNAVNDFKIG IPLTFTQLFFA Agis dosta SKADEYLRE IKEQNYLGDA VRQCVEAAAH EYD PDTQKAL MRARSFGKGF LTE FSPDQ FVTTCRELRV LNAVNDFKIG IPLTHTQFKQ Tanjua SQKADEYLRE IKEQSHLEEA VKQCVGAARY EYD PDTQKAL MRARSFGKGF LSN FPPEQ FVSKCRDLRV LNAVNDYTIG IPLTHTQFKQ Takjuu SQKADEYLRE IKOONLLPEA VRQCVEAAGH EHEP FTYKIL LRAASFGKGF LAD FSADP FVETCRELRV LNAVNSSVG LPLTHTQFKQ Gallus gallus SQKADEYLRE IKEQSHLEEA VKQCVGAARY EYD PDTQKSL LRAASFGKGF LAD FSADP FVETCRELRV LNAVNSSVG LPLTHTQFKQ Macca SQKADEYLRE IKDONLLPEA VRQCIEAAGY EHEP FTYKIL LRAASFGKGF LDR FFPES FVETCRLRV LNAINDYGG IPLSFDYNQ Gallus gallus SQKADEYLRE IKDOKLLPEA VQCIEAAGY EHEP FTYKSL LRAASFGKGF LDR FFPES FVETCRLRV LNAINDYGG IPLSFDYNQ Gallus gallus SQKADEYLRE IKDQKLLPEA VQCIEAAGY EHEP FTYKSL LRAASFGKGF LDR FFPES FVETCRLRV LNAINDYGG IPLTFTQYKR Macca SQKADEYLRE IQEGQUTQA VQCIEAAGH EHEP PMXSL LRAASFGKGF LDR FFPES FVETCRLRV LNAINDYGG IPLTFTQYKR Macca SQKADEYLRE IQEGQUTQA VQCIEAAGH EHEP PMXSL LRAASFGKGF LDR FFPES FVETCRLRV LNAINDYNG IPLTFYSYKQ Gallus gallus SQKADEYLRE IQEGQUTQA VQCIEAAGH EHEP PMXSL LRAASFGKGF LDR FFPES FVETCRLRV LNAINDYNG IPLTFYSYKR Macca SQKADEYLRE IQEGQUTQA VQCIEAAGH EHEP PMXSL LRAASFGKGF LDR FFPES FVETCRLRV LNAINDYNG IPLTFYSYKR Macca SQKADEYLRE IQEGQUTQA VQCIEAAGH EHEP PMXSL LRAASFGKGF LDR FFPES FVETCRLRV LNAINDYNG IPLTFYSYKR Macca SQKADEYLRE IQEGQUTQA VQCIEAAGH EHEP PMXSL LRAASFGKGF LDR FFPES FVETCRLRV LNAINDYNG IPLTFY	Dmoja	S <mark>Y</mark> K S D E Y L <mark>S M</mark>	C R S N I	ELA V <mark>N</mark> ECI	EAA <mark>ay e</mark> f	C P E T Q K S L	M R T A <mark>Y</mark> F G K <mark>G</mark> F	IPL	H N <mark>P E E P</mark>	YM <mark>RILRI</mark> L	RV LNT	LR <mark>HEK</mark> IA	MPLTYK	Q F S H
DWIII SYTSDEVLSM CREKIGLA VNQCIEAASY EFCTATOKSL MRTANFGKGF IHGHNPDE VMRINRILEV LNILMHKKA MPITFKOFSH Musca SHQADEYLSL FRDRMNVA VDECIEAAGY EFCTATOKSL MRTANFGKGF IHGHNPDE VMRINRILEV LNILMHKKA MPITFKOFSH Musca SHQADEYLSL FRDRMNVA VDECIEAAGY EFCTATOKSL MRTANFGKGF IPAHNPDE VMRURTLRV LNALRHPRIA VPLIKGFHY Ades SHXSDEFLCL IQORMNVA VDECIEAAGY EFDEVQMNL IKAANGGKGF IAAHNPDE VIRMLRILEV LNALRHPRIA VPLIKGFHY Tibolium SHANEYICL VKQDLAKA VDQCINAGY EFDEVQMNL IKAANGGKGF SKTIDPEY VNMCRILV LNALRHPRIA INTITGEV Apis dorsate SHKADGYMDL VKDKLDAA IKACIDGASH EFDFETQKLL MKANFGKGF SKTIDPEY VNMCRILV LNAVRDPKIG IPLTYTOFTV Poecila SQKADEYLRE IKEQSLISEA VQCVEAAGH EHEPETQKTL LRAASFGKCF LSNFPPEQ FVSMCRDLRV LNAVRDYTG IPLTYTOFTV Poecila SQKADEYLRE IKEQSLISEA VQCVEAAGH EHEPETQKIL LRAASFGKCF LSNFPPEQ FVSMCRDLRV LNAVRDYTG IPLTYTOFTV Poecila SQKADEYLRE IKDQNLLSDA VENCIEAAGY EHEPETQKIL LRAASFGKCF VKKYSPAD FVTHCQDLRI LNAINDYQIG IPLTMTYPKK Anois SQKADEYLRE IKDQNLLPEA VRQCIEAAGY EYDPTYNSL LRAASFGKCF VDKFFPES FVETCRDLRV LNAINDYQIG IPLTMTYPKK Maccac SQKADEYLRE IKDQNLLPEA VRQCIEAAGY EHEPETQKSL LRAASFGKCF LDRFPPES FVRCQDLRV LNAINDYQIG IPLTFYQYKR Maccac SQKADEYLRE IKDQQLIQ VQCIEAAGH EHQPDMVSL LRAASFGKCF LDRFPPES FVRCQDLRV LNAINDYQIG IPLTFYQYKR Maccac SQKADEYLRE IKDQQLIQ VQCIEAAGY EHEPETQKSL LRAASFGKCF LDRFPPES FVRCQDLRV LNAINDYNG IPLTYSQYKQ Aligator SQKADEYLRE IQEGQUTQA VQCIEAAGY EHEPETQKSL LRAASFGKCF LDRFPPES FVRCQDLRV LNAINDYNG IPLTYSQYKQ Muscusuus SQKADEYLRE IQELGUTQA VQCIEAAGH EHQPMMVSL LRAASFGKCF LDRFPPES FVRCQDLRV LNAINDYNG IPLTYSQYKQ Bostaurus SQKADEYLRE IQELGUTQA VQCIEAAGH EHQPMVKSL LRAASFGKCF LDRFPPES FVRCQDLRV LNAINDYNG IPLTYSQYKQ Adacas SQKADEYLRE IQELGUTQA VQCIEAAGH EHQPMKSL LRAASFGKCF LDRFPPES FVRCQDLRV LNAINDYNG IPLTYSQYKQ Bostaurus SQKADEYLRE IQELGUTQA VQCIEAAGH EHQPMKSL LRAASFGKCF LDRFPPES FVRCQDLRV LNAINDYNG IPLTYSQYKQ Bostaurus SQKADEYLRE IQELGUTQA VQCIEAAGH EHQPMKSL LRAASFGKCF LDRFPPES FVRCQDLRV LNAINDYNG IPLTYSQYKQ Bostaurus SQKADEYLRE IQELGUTQA VQCIEAAGH EHQPMKSL LRAASFGK	Dviri	S <mark>Y</mark> K S D Q Y L <mark>S L</mark>	C R S D I	ELA V <mark>N</mark> ECI	EAA <mark>ay e</mark> f	C P E T Q K S L	M K A A <mark>Y</mark> F G K <mark>A</mark> F	I P S I	HN <mark>PEE </mark>	YM <mark>RILRI</mark> L	RV LNT	LR <mark>HDK</mark> IA	MPLTY	QFSH
MUSCI SHQADEYLSL FRD - RNAVA VDECIEAAGY EFCTATORSL MRTAYEGKGE TPA - HNPDE YIRMLRTLRV LNALRHPETA VPLTLQFLV Aces SHKSDEFLCL IQD - RLASA VGECUDAAGQ EFDENTKKSL TRAAMFGKGE LPG - YNPDG YIRMCRVLRV LNALRHPETA VPLTLQFHV Fibolum SHRANEYLCL VKO - DLAKA VOOCINAAGY EFDEFVKLL TRAAMFGKGE LPG - YNPDG YIRMCRVLRV LNALRHPTA VPLTLQFHV Apis dorsata SHKADGYMDL VKD - KLDAA IKACIDGASH EFDEFTQKLL MKARKFGKGE SKT - IDPEY YVNGCRTLRV LNAVRDPKIG IPLTYTQFTV Poelia SQKADEYLRE IKEQNVLGDA VRQCVEAAGH EYDPDTQKAL MKARSFGKGE LSN - FPPEQ FVSMCRDRV LNAVRESYG MPLTHTQFKQ Annor SQKADEYLRE IKEQNULGDA VRQCVEAAGH EYDPDTQKAL MKARSFGKGE LSN - FPPEQ FVSMCRDRV LNAVRESYG MPLTHTQFKQ Annor SQKADEYLRE IKEQNLEEA VRQCVEAAGH EYDPDTQKSL LRAASFGKGE LSN - FPPEQ FVSMCRDRV LNAVROYTIG IPLTYTQFTV Gedius gdius SQKADEYLRE IKDQNLLPEA VRQCIEAASY EYDPDTXSL LRAASFGKGE LAD - FSADP FVETCRDRV LNAVRVSYG LPLTYDQFKH Anols SQKADEYLRE IKDQNLLPEA VRQCIEAASY EYDPDTXSL LRAASFGKGE LAD - FFPES FVETCRDRV LNAIRDYQIG IPLTFQVKQ Galus gdius SQKADEYLRE IKDQNLLPEA VRQCIEAAGY EHEPETQKSL LRAASFGKGE LAA - FFPES FVETCRDRV LNAIRDYQIG IPLTFQVKQ Mus musculs SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKGE LDR - FPPDS FVHMCQDLRV LNAIRDYQIG IPLTFQVKQ Mus musculs SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKGE LDR - FPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQVKQ Aligator SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKGE LDR - FPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQVKQ Felis catus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKGE LDR - FPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQVKQ Felis catus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKGE LDR - FPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQVKQ Felis catus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKGE LDR - FPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQVKQ Felis catus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKGE LDR - FPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQVKQ Felis catus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPMQKSL LRAASFGKGE LDR - FPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQVKQ Felis catus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPMQKSL LRAASFGKGE LDR - FPPDS FVHMCQDLRV	Dwill	S <mark>Y</mark> KSDEYL <mark>SM</mark>	CREKI	GLA V <mark>N</mark> QCI	EAA <mark>s</mark> y ef	C T E T Q K S L	L R T A <mark>N</mark> F G K <mark>G</mark> F	IHG	HN <mark>PDE N</mark>	YM <mark>RIMRI</mark>	RV LNT	LR <mark>HEK</mark> IA	MPLTFK	QFSH
AedesSHKSDEFLCLIQD RLASAVGECVDAAGQEFDSHTQKSLIRAAHFGKSFLPG YNPDGYIEMCRVLRVLNALREFIVGMPLTLRQFMHApisdorstaSHKADGYDLVKO KLDAAIKALDGGASHEFDFETQKLLIRAAGFGKGFITAY NNSDKYVNIIRLLRVLNAVRPKIGIPLIFTQLQFApisdorstaSKADGYLREIKEQNVLGDAVRQCVEAAGHEFDFETQKLLNKAASFGKGFITA FDPEQYVNGRTLRVLNAVRPSSVGMPLHTQFKQDanio rerioSQKADDYLREIKEQSLISEAYQQCVEAAGHEHPETQKILIRAASFGKCFISN FDPEQFVSMCRDLRVLNAVRPSVGIPLIHTQFKQYandiyuSQKADDYLREIKEQSLISEAYQQCVEAAGHEHPETQKILIRAASFGKCFISN FPPEQFVSMCRDLRVLNAVRPSVGIPLIHTQFKQYandiyuSQKADDYLREIKOQNLLSAVRQCIEAAGYEHPETQKILIRAASFGKCFVZK SPADFVTMCQDLRVINAIRDYQIGIPLIFTQYKRAnolisSQKADEYLREIKOQNLLPEAVRQCIEAAGYEHPETQKSLIRAASFGKCFVDK FTPESFVETCRDLRVINAIRDYQIGIPLIFTQYKRMacacaSQKADEYLREIKOQULPAVRQCIEAAGYEHPETQKSLIRAASFGKCFIDK FPPDSFVMCQDLRVINAIRDYQIGIPLIFTQYKRMusamusculusSQKADEYLREIKOQULPAVRQCIEAAGHEHPETQKSLIRAASFGKCFIDK FPPDSFVMCQDLRVINAIRDYHIGIPLITYQYKQMusamusculusSQKADEYLREIQELGQLTQAVQQCIEAAGHEHQPDMQKSLIRAASFGKCFIDR FPPDSFVMCQDLRVINAIRDYHIGIPLITYQYKQMusamusculusSQKADEYLREIQELGQLTQA<	Musca	S <mark>H Q A D E Y L <mark>S L</mark></mark>	FRD RM	NVA V <mark>d</mark> eci	EAA <mark>gy e</mark> f	C T A T Q K S L	M R T A <mark>Y</mark> F G K <mark>G</mark> F	IPA	H N <mark>P D E N</mark>	Y I <mark>R M L R T</mark> L	RV LN <mark>A</mark>	LRHPRIA	VPLTIK	
Tribolium Apis dorste SHKAADGYMOL VKQDLAKA VDQTINAJGY EF0PEVQKML IRAQFGKCF IAYMNSDK YVMIIRLLRV LNAVROPKIG IPLTFTQLQF Apis dorste SHKADGYMOL VKQKLDAA IKALIDGASH EF0FETQKLL HKAAKFGKGF SKTLDPEY YVNMCRTLRV LNAVROPKIG IPLTFTQLQF Donio rein SQKADEYLRE IKEQSLLSEA VQQCVEAAGH EHEPETQKTL LRAASFGKCF LTEFSPDQ FVTCRELRV LNAVROYTIG IPLTHTQFKQ Xenopus SQKADEYLRE IKEQSLLSEA VQQCVEAAGH EHEPETQKTL LRAASFGKCF LSNFPPEQ FVSMCROLRV LNAVROYTIG IPLTHTQFKQ Anolis SQKADEYLRE IKONLIPEA VKQCVGAARY EY0PQTQKSL LRAASFGKCF LSNFSPEQ FVTCROLRV LNAVROYTIG IPLTHTQFKQ Gallus gallus SQKADEYLRE IKONLIPEA VKQCVGAARY EY0PQTQKSL LRAASFGKCF LDKFSADP FVETCRELRV LNAVROYTIG IPLTHTQFKQ Gallus gallus SQKADEYLRE IKONLIPEA VSQCIEAAGY EHEPETQKSL LRAASFGKCF LDKFSPES FVETCRDLRV LNAVROYTIG IPLTFQYKR Macca SQKADEYLRE IKDQKLLPEA VSQCIEAAGY EHEPETQKSL LRAASFGKCF LDKFSPES FVETCRDLRV LNAVROYTIG IPLTFQYKR Maususus SQKADEYLRE IQELGQLTQA VQQTIEAAGH EHQDMQKSL LRAASFGKCF LDRFSPES FVETCRDLRV LNAVROYTIG IPLTFQYKR Mausussus SQKADEYLRE IQELGQLTQA VQQTIEAAGH EHQDMQKSL LRAASFGKCF LDRFPPES FVMCQDLRV LNAIROYHIG IPLTYSQYKQ Musumsusus SQKADEYLRE IQELGQLTQA VQQTIEAAGH EHQDMQKSL LRAASFGKCF LDRFPPDS FVMCQDLRV LNAIROYHIG IPLTYSQYKQ Homs speins SQKADEYLRE IQELGQLTQA VQQTIEAAGH EHQDMQKSL LRAASFGKCF LDRFPPDS FVMCQDLRV LNAIROYHIG IPLTYSQYKQ Rattus SQKADEYLRE IQELGQLTQA VQQTIEAAGH EHQDMQKSL LRAASFGKCF LDRFPPDS FVMCQDLRV LNAIROYHIG IPLTYSQYKQ Ceenorhabdits GHNSFAAST VIQDIVKA IDDISTACD TWQPEGKLL LKAASFGKCF LDRFPPDS FVMCQDLRV LNAIROYHIG IPLTYSQYKQ Ceenorhabdits GHNSFAAST VIQDIVKA IDDUSTAL LRAASFGKCF LDRFPPDS FVMCQDLRV LNAIROYHIG IPLTYSQYKQ Ceenorhabdits GHNSFAAST VIQDIVKA IDDUSTALD HQPEGKLL LKAASFGKCF LDRFPPDS FVMCQDLRV LNAIROYHIG IPLTYSQYKQ Ceenorhabdits GHNSFAAST VIQDIVKA IDDUSTACD TWQPEGKL LKAASFGKCF LDRFPPDS FVMCQDLRV LNAIROYHIG IPLTYSQYKQ Ceenorhabdits GHNSFAAST VIQDIVKA IDDUSTACD TWQPEGKL LKAASFGKCF LDRFPPDS FVMCQDLRV LNAIROYHIG IPLTYSQYKQ Ceenorhabdits GHNSFAAST VIQDIVKA IDDUSTACD TWQPEGKL LKAASFGKCF LDRFPPDS FVMCQDLRV LNAIROYHIG IPLTY	Aedes	S H K S D E F L <mark>C L</mark>	IQD RL	AS <mark>A V</mark> GECVI	DAA <mark>G</mark> Q EF	<mark>D S H T Q K S L</mark>	I R A <mark>A H</mark> F G K <mark>S</mark> F	L P G '	Y N <mark>P D G </mark>	Y I E M C R V L	RV LN <mark>A</mark>	LREPIVG	MPLTL	QFNH
APIS dorsata Poedia SQKADEYLRE IKEQUVLGDA VRQCVEAAAH EVDPTOKAL MKAAKFGKGF Danio rerio SQKADEYLRE IKEQUVLGDA VRQCVEAAAH EVDPTOKAL MKAAKFGKGF Conorio SQKADEYLRE IKEQULGDA VRQCVEAAAH EVDPTOKAL MKAAKFGKGF Takingu SQKADEYLRE IKEQULGDA VRQCVEAAAH EVDPTOKAL MKAAKFGKGF Takingu SQKADEYLRE IKEQSHLEEA VKQCVEAAAH EVDPTOKAL MKAAKFGKGF Takingu SQKADEYLRE IKEQSHLEEA VKQCVEAAAH EVDPTOKSL LRAASFGKCF Anois SQKADEYLRE IKDQNLLPEA VRQCIEAAGY EHDPETOKSL LRAASFGKCF Gallsgallus SQKADEYLRE IKDQNLLPEA VRQCIEAAGY EHDPETOKSL LRAASFGKCF Gallsgallus SQKADEYLRE IKDQNLLPEA VRQCIEAAGY EHDPETOKSL LRAASFGKCF Anois SQKADEYLRE IKDQNLLPEA VRQCIEAAGY EHDPETOKSL LRAASFGKCF Gallsgallus SQKADEYLRE IKDQNLLPEA VRQCIEAAGY EHDPETOKSL LRAASFGKCF Aligotor SQKADEYLRE IKDXEQLSEA VQQCIEAAGH EHDPOMOKSL LRAASFGKCF Hussensselus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHDPOMOKSL	Tribolium	S H R A N E Y I <mark>C L</mark>	<mark>V К Q</mark> <mark>D L</mark>	AKA V <mark>D</mark> QCI	NAV <mark>g</mark> y ef	D P E V Q K M L	I R A <mark>A </mark> F G K <mark>C</mark> F	IAY	MN <mark>SD</mark> K	YV <mark>NIIRL</mark>	RV LN <mark>A</mark>	V R D P K I G	IPLTFT	QLQF
Poediia SQKADEVLRE IKEQNVLGBA VRQCVEAAAH EVDPDTQKAL MRAASFGKCF LTEFSPDQ FVTTCRELRV LNAVRESSNG MPLTHTQFKQ Daniario SQKADEVLRE IKEQNLGBA VQQCVEAAAH EHEPETQKL LRAASFGKCF LSNFPPEQ FVTKCRELRV LNAVRQVTIG IPLTHTQFKQ Xenopus SQKADEVLRE IKDQNHLSDA VENIEAAGY EHEPETQKSL LRAASFGKCF LSNFSPEQ FVTKCRELRV LNAVRVVSVG LPLTYPQFKH Anolis SQKADEVLRE IKDQNLLPEA VRQIEAAGY EHEPETQKSL LRAASFGKCF LADFSADP FVETCRELRV LNAVRVVSVG LPLTYPQFKH Anolis SQKADEVLRE IKDQNLLPEA VRQIEAAGY EHEPETQKSL LRAASFGKCF LDDFSADP FVETCRELRV LNAVRVVSVG IPLTFTQYKR Macca SQKADEVLRE IKDQKLLPEA VSQCIEAAGY EHEPETQKSL LRAASFGKCF LDRFPPES FVETCRDLRV LNAIRDYQIG IPLTFTQYKR Alligator SQKADEVLRE IKDKEQLSEA VQQIEAAGH EHEPETQKSL LRAASFGKCF LDRFSPES FVMCQDLRV LNSIRDYQIG IPLTFTQYKR Bastaurus SQKADEVLRE IQELGQLTQA VQQIEAAGH EHEPETQKSL LRAASFGKCF LDRFSPES FVMCQDLRV LNSIRDYQIG IPLTFAQYKQ Humosapiens SQKADEVLRE IQELGQLTQA VQQIEAAGH EHEPETQKSL LRAASFGKCF LDRFSPES FVMCQDLRV LNAIRDYHIG IPLTYQYKQ Bostaurus SQKADEVLRE IQELGQLTQA VQQIEAAGH EHEPETQKSL LRAASFGKCF LDRFSPES FVMCQDLRV LNAIRDYHIG IPLTYQYKQ Humosapiens SQKADEVLRE IQELGQLTQA VQQIEAAGH EHEPETQKSL LRAASFGKCF LDRFSPES FVMCQDLRV LNAIRDYHIG IPLTYQYKQ Katus SQKADEVLRE IQELGQLTQA VQQIEAAGH EHEPPEQKSL LRAASFGKCF LDRFSPES FVMCQDLRV LNAIRDYHIG IPLTYQYKQ Homosapiens SQKADEVLRE IQELGQLTQA VQQIEAAGH EHEPPEQKSL LRAASFGKCF LDRFPPDS FVMCQDLRV LNAIRDYHIG IPLTYQYKQ Katus SQKADEVLRE IQELGQLTQA VQQIEAAGH EHEPPEQKSL LRAASFGKCF LDRFPPDS FVMCQDLRV LNAIRDYHIG IPLTYSQYKQ Ratus SQKADEVLRE IQELGQLTQA VQQIEAAGH EHQPDMXKL LRAASFGKCF LDRFPPDS FVMCQDLRV LNAIRDYHIG IPLTYSQYKQ Cenorhabditis GHNSFAAST VIQDLYKA IDDISTACO TWQPEGKLL LKAASFGKCF LDRFPPDS FVMCQDLRV LNAVRDYHIG IPLTYSQYKQ Cacharomyces AFKAIEILKN FVLEKG VLDCIAAAH GFDHSMKKL LKAASFGKCF LDRFPPDS FVMCQDLRV LNAVRDYHIG IPLTYQYKQ Cacharomyces AFKAIEILKN FVLEKG VLDCIAAAH GFDHSMKKL LKAASFGKCF LDRFPPDS FVMCQDLRV LNAVRDYHIG IPLTYQYKQ Consistenv 9587888855 5660048449 947*979956 963776*7* 88759595757 7348655875 8756574999 *67954497 7899569755	Apis dorsata	S <mark>h</mark> k a d g y m <mark>d l</mark>	VKD KL	DA <mark>A I</mark> KACII	DGA <mark>S</mark> H EF	D F E T Q K L L	M K A A <mark>K</mark> F G K <mark>G</mark> F	<mark>S К T</mark>	ID <mark>PEY</mark>	Y V <mark>N M C R T</mark> L	RV LN <mark>A</mark>	V R H P A I G	IPLTY	QFTV
Danio rerio SQKADEYLRE IKEQSLLSEA VQQCVEAAGH EHEPETQKTL LRAASFGKCF Takiguu SQKADEYLRE IKEQSMLEEA VKQCVGAAGY EYDPQQKSL LRAASFGKCF Takiguu SQKADEYLRE IKEQSMLEEA VKQCVGAAGY EYDPQQKSL LRAASFGKCF Gallusgallus SQKADEYLRE IKEQSMLEEA VKQCVGAAGY EYDPQQKSL LRAASFGKCF Gallusgallus SQKADEYLRE IKEQSMLEEA VQQCIEAAGY EHEPETQKSL LRAASFGKCF Gallusgallus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDR - FFPDS FVMCQDLRV LNAIRDYDIG IPLTFQYKR Musmusulus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDR - FFPDS FVMCQDLRV LNAIRDYHIG IPLTYSQYKQ Homo sopiens SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDR - FFPDS FVMCQDLRV LNAIRDYHIG IPLTYSQYKQ Actus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDR - FFPDS FVMCQDLRV LNAIRDYHIG IPLTYSQYKQ Musmusulus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDR - FFPDS FVMCQDLRV LNAIRDYHIG IPLTYSQYKQ Homo sopiens SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDR - FFPDS FVMCQDLRV LNAIRDYHIG IPLTYSQYKQ Rattus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDR - FFPDS FVMCQDLRV LNAIRDYHIG IPLTYSQYKQ MSGXADEYLRE IQELGQLTQA VQCIEAAGH EHQPDMQKSL LRAASFGKCF LDR - FFPDS FVMCQDLRV LNAIRDYHIG IPLTYSQYKQ Caenorhobidis TGNNSFAAST VIC - DLYKA IDDCISTACD TWQPEEQKLL LKAASFGKCF LDR - FPPDS FVMCQDLRV LNAIRDYHIG IPLTYSQYKQ Coenorhobidis TGNNSFAAST VIC - DLYKA VQCLFAAAH QFDHSMQKL LKAASFGKCF LDR - FPPDS FVMCQDLRV LNAIRDYHIG IPLTYSQYKQ Coenorhobidis TGNNSFAAST VIC - DLYKA VQCLFAAAH QFDHSMQKL LKAASFGKCF LDR - FPPDS FVMCQDLRV LNAIRDYHIG IPLTYSQYKQ Coenorhobidis TGNNSFAAST VIC - DLYKA VQCLFAAAH QFDHSMQKL LKAASFGKCF DOS SOKADEYLRE IQEAQHAAAFGYCF DAS SOKAFFT	Poecilia	S <mark>Q</mark> K A D E Y L <mark>R E</mark>	IKEQNVL	GDA V <mark>RQCV</mark> I	ЕАА <mark>АН Е</mark> Ү	<mark>D P D T Q K A L</mark>	M R A <mark>A </mark> F G K <mark>C</mark> F	LTE	F S <mark>P D Q I</mark>	F V T T C R E L	RV LN <mark>A</mark>	V R <mark>E S S V</mark> G	MPLTH	Q F K Q
XenopusSQKADDYLREIKDQNHLSDAVENCIEAAGYEHAPEMQKSLLRAASFGKCFVEK - YSPADFVTMCQDLRILNAIHDYQIGIPLTMTQYKQTakfuguSQKADEYLRELKEQSMLEEAVKQCVGAARYEYDPQTQKSLLRAASFGKCFLAD - FSADPFVETCRELRVLNAIRDYQIGIPLTMTQYKQAnoisSQKADEYLREIKDQNLLPEAVRQCIEAAGYEQEPETQKSLLRAASFGKCFLAD - FSADPFVETCRELRVLNAIRDYQIGIPLTFAQYKQGallus gallusSQKADEYLREIQELGQLTQAVQQCIEAAGYEHEPETQKSLLRAASFGKCFIDK - FFPESFVMCQDLRVLNAIRDYQIGIPLTFAQYKQMaccaSQKADEYLREIQELGQLTQAVQQCIEAAGYEHEPETQKSLLRAASFGKCFLDR - FSPESFVMCQDLRVLNAIRDYQIGIPLTFAQYKQMusmusculusSQKADEYLREIQELGQLTQAVQQCIEAAGHEHQPDMQKSLLRAASFGKCFLDR - FSPESFVMCQDLRVLNAIRDYHIGIPLTYSQYKQMusmusculusSQKADEYLREIQELGQLTQAVQQCIEAAGHEHQPDMQKSLLRAASFGKCFLDR - FSPESFVMCQDLRVLNAIRDYHIGIPLTYSQYKQMusmusculusSQKADEYLREIQELGQLTQAVQQCIEAAGHEHQPDMQKSLLRAASFGKCFLDR - FSPEDSFVMCQDLRVLNAIRDYHIGIPLTYSQYKQHomo sapiensSQKADEYLREIQELGQLTQAVQQCIEAAGHEHQPDMQKSLLRAASFGKCFLDR - FSPEDSFVMCQDLRVLNAIRDYHIGIPLTYSQYKQGaeonhabditis TGHNSFAASTVIQ - DLYKAIDDCISTACDTWQPEEQKLLLKAASFGKCFLDR - FSPEDSFVMCQDLRVLNAIRDYHIGIPLTYSQYKQSaccrissuumSGKADEYLREIQELGQL	Danio rerio	S <mark>Q</mark> K A D E Y L <mark>R E</mark>	IKEQSLL	SEA VQQCV	EAA <mark>G</mark> H EH	E P E T Q K T L	L R A <mark>A </mark> F G K <mark>C</mark> F	L S N	FP <mark>PEQ</mark> I	F V <mark>S M C R D</mark> L	RV LNA	V R D Y T I G	IPLTH	Q F K Q
TakifuguSQKADEYLRELKEQSMLEEAVKQCVGAARYEYDPQTQKSLLRAASFGKCFLAD FSADPFVETCRELRVLNAVRVSSVGLPLTYPQFKHAnoisSQKADEYLREIKDQNLLPEAVRQCIEAASYEQEPEIQKSLLRAASFGKCFVDK FTPESFVETCRELRVLNAVRVSSVGLPLTYPQFKHMacacaSQKADEYLREIKDQKLLPEAVSQCIEAAGYEHEPETQKSLLRAASFGKCFIDK FFPESFVEMCQDLRVLNSIRDYQIGIPLTFYQYKQMacacaSQKADEYLREIKDQKLLPEAVQCIEAAGHEHEPETQKSLLRAASFGKCFIDK FFPESFVEMCQDLRVLNAVRDYHGIPLTYSQYKQAlligatorSQKADEYLREIQELGQLIQAVQCIEAAGHEHEPETQKSLLRAASFGKCFLDR FFPESFVEMCQDLRVLNAIRDYHGIPLTYSQYKQMus musculusSQKADEYLREIQELGQLIQAVQQCIEAAGHEHPPDMQKSLLRAASFGKCFLDR FFPDSFVEMCQDLRVLNAIRDYHGIPLTYSQYKQBos tarursSQKADEYLREIQELGQLIQAVQQCIEAAGHEHPPDMQKSLLRAASFGKCFLDR FFPDSFVEMCQDLRVLNAIRDYHGIPLTYSQYKQHomo sapiensSQKADEYLREIQELGQLIQAVQQCIEAAGHEHQPDMQKSLLRAASFGKCFLDR FFPDSFVEMCQDLRVLNAIRDYHGIPLTYSQYKQCaeorhabditis TGHNSFAASTVIQ DLYKAIDDCISTACDTWQPEQKLLLKAASFGKCFLDR FPPDSFVHMCQDLRVLNAIRDYHGIPLTYSQYKQSqKADEYLREIQELGQLIQAVQQCIEAAGHEHQPDMQKSLLRAASFGKCFLDR FPPDSFVMCQDLRVLNAIRDYHGIPLTYSQYKQHomo sapiensSQKADEYLREIQELGQLIQAVQQCIE	Xenopus	S <mark>Q</mark> K A D D Y L <mark>R E</mark>	IKDQNHL	SDA VENCI	EAA <mark>G</mark> Y EH	A P E M Q K S L	L R A <mark>A </mark> S F G K <mark>C</mark> F	V E K Y	Y S <mark>P A D </mark>	F V T M C Q D L	RI LN <mark>A</mark>	IH <mark>DYQ</mark> IG	IPLTMT	т <mark>оч</mark> ко
Anolis Gallus gulus Galusgalus by KADEYLRE IKDQKLLPEAIKDQNLLPEA VRQCIEAAGY VQCIEAAGY EQEPEIQKSL EMEPETQKSL EMEPETQKSL LRAASFGKCFVDKFTPES FVERCQLRV LDRFTPES FVERCQLRV LDRFTPES FVERCQLRV LDRFTPES FVERCQLRV LDR-FTPES FVERCQ	Takifugu	S <mark>Q</mark> K A D E Y L <mark>R E</mark>	L K E Q S M L	ΕΕΑ VKQCV	GAA <mark>r</mark> y ey	D P Q T Q K S L	L R A <mark>A </mark> F G K <mark>C</mark> F	L A D	FS <mark>ADP I</mark>	F V E T C R E L	RV LNA	V R <mark>V S S</mark> V G		Q F K H
Gallus gallus Gallus gallus SQKADEYLRE INDEXTREMENT INDEXTR	Anolis	S <mark>Q</mark> K A D E Y L <mark>R E</mark>	IKDQNLL	PEA VRQCI	EAA <mark>S</mark> Y EQ	E P E I <mark>Q K S L</mark>	L R A <mark>A </mark> S F G K <mark>C</mark> F	VDK	FT <mark>PES</mark> I	F V E T C R D L	RV LN <mark>A</mark>	IR <mark>DYQ</mark> IG	IPLSFC	Q Y R Q
Macaca SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDRFPPDS FVHMCQDLRV LNAVRDYHIG IPLTYSQYKQ Alligator SQKADEYLRE IKDKEQLSEA VQQCIEAAGY EHEPETQKSL LRAASFGKCF LDRFPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQYKQ Musmusculus SQKADEYLRE IQELGQLIQA VQQCIEAAGH EHPPDMQKSL LRAASFGKCF LDRFPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQYKQ Felis catus SQKADEYLRE IQELGQLIQA VQQCIEAAGH EHPPDMQKSL LRAASFGKCF LDRFPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQYKQ Bos taurus SQKADEYLRE IQELGQLIQA VQQCIEAAGH EHPDPDMQKSL LRAASFGKCF LDRFPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQYKQ Bos taurus SQKADEYLRE IQELGQLIQA VQQCIEAAGH EHPDPDMQKSL LRAASFGKCF LDRFPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQYKQ Rattus SQKADEYLRE IQELGQLIQA VQQCIEAAGH EHPDPMQKSL LRAASFGKCF LDRFPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQYKQ Rattus SQKADEYLRE IQELGQLIQA	Gallus gallus	S <mark>Q</mark> K A D E Y L <mark>R E</mark>	I K D Q K L L	PEA V <mark>SQ</mark> CI	EAA <mark>gy e</mark> h	E P E T Q K S L	L R A <mark>A </mark> F G K <mark>C</mark> F	IDK	F P <mark>P E S I</mark>	F V <mark>R M C Q D</mark> L	RV LNS	IR <mark>DYQ</mark> IG	IPLTFT	I <mark>QY</mark> KR
Alligator SQKADEYLRE IKDKEQLSEA VQQCIEAAGY EHEPETQKSL LRAASFGKCF LDRFSPES FVRMCQDLRV LNAIRDYHIG IPLTFAQYKQ Musmusculus SQKADEYLRE IQELGQLIQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDRFPPDS FVRMCQDLRV LNAIRDYHIG IPLTYTQYKQ Bostaurus SQKADEYLRE IQELGQLQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDRFPPDS FVRMCQDLRV LNAIRDYHIG IPLTYSQYKQ Homo sopiens SQKADEYLRE IQELGQLIQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDRFPPDS FVRMCQDLRV LNAIRDYHIG IPLTYSQYKQ Caenorhobditis TGHNSFAAST VIQDLYKA IDDCISTACD TWQPEEQKLL LKAASFGKCF LDRFPPDS FVRMCQDLRV LNAIRDYHIG IPLTYSQYKQ Caenorhobditis TGHNSFAAST VIQDLYKA IDDCISTACD TWQPEEQKLL LKAASFGKCF LDRFPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQYKQ Saccharomyces APKAIEILKN FVLEKG VLDCIAAAID EFEPKLQKML LNAASYGKAS Consistency 9587888855 5660048849 947 979956 963776*7 887 59 959 7340055875 8756574999 *67954497 7899659755 1 2 2 3 4	Macaca	S <mark>Q</mark> K A D E Y L <mark>R E</mark>	IQELGQL	TQ <mark>A VQ</mark> QCI	EAA <mark>G</mark> H EH	<mark>Q</mark> P D M <mark>Q K S L</mark>	L R A <mark>A S</mark> F G K <mark>C</mark> F	L D R	FP <mark>PDS</mark> I	F V H M C Q D L	RV LN <mark>A</mark>	V R D Y H I G	IPLTYS	<mark>Q Y K Q</mark>
Mus musculus SQKADEYLRE IQELGQLIQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDR FPPDS FVHMCQDLRV LNAIRDYHIG IPLTYTQYKQ Pelis catus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHRPDMQKSL LRAASFGKCF LDR FPPDS FVHMCQDLRV LNAIRDYHIG IPLTYTQYKQ Bos taurus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHRPDMQKSL LRAASFGKCF LDR FPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQYKQ Homo sapiens SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDR FPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQYKQ Katus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDR FPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQYKQ Caeorhabditis TGHNSFAAST VIQ DLYKA IDDCISTACD TWQPEEQKLL LKAASFGKCF LDR FPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQYKQ Sacaris suum SHGVYSYFKM IEQ QFDKA VQCLFAAAH QFDHSMQKRL LKAASLGNAL LQR HDPSQ FVMCQDLRV LNAVRQYPIG MPLSFLQSND Saccharomyces APKAIEILKN FVLEKG	Alligator	S <mark>Q</mark> K A D E Y L <mark>R E</mark>	I K D K E Q L	SEA VQQCI	EAA <mark>G</mark> Y EH	E P E T <mark>Q K S L</mark>	L R A <mark>A </mark> S F G K <mark>C</mark> F	L D R	FS <mark>PES</mark> I	F V R M C Q D L	RV LN <mark>A</mark>	I R <mark>D Y Q</mark> I G	IPLTF	I <mark>QY</mark> KQ
Feliscatus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHRPDMQKSL LRAASFGKCF LDR FPPDS FVRMCQDLRV LNAIRDYHIG IPLTYSQYKQ Bostaurus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHRPDMQKSL LRAASFGKCF LDR FPPDS FVRMCQDLRV LNAIRDYHIG IPLTYSQYKQ Homosopiens SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHRPDMQKSL LRAASFGKCF LDR FPPDS FVRMCQDLRV LNAIRDYHIG IPLTYSQYKQ Katus SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHPPDMQKSL LRAASFGKCF LDR FPPDS FVRMCQDLRV LNAIRDYHIG IPLTYSQYKQ Caenorhobditis TGHNSFAAST VIQ DLYKA IDDCISTACD TWQPEEQKLL LKAASFGKCF IDR FPPDS FVMCQDRV LNAIRDYHIG IPLTYSQYKQ Sacaris suum SHGVYSYFKM IEQ QFDKA VQCLFAAAH QFDHSMQKRL LKAASIGNAL LQR HDPSQ FVMCQNCHVMRV LNAVRQPYIG MPLSFLQSND Saccharomyces APKAIEILKN FVLEKS VLDCIAAAID EFEPKLQKML LNAASYGKAS LQYKSFDASI FVMACNTIKL LNAVRQPYIG MPLSFLQSND Saccharomyces 9587888855 </td <td>Mus musculus</td> <td>S <mark>Q</mark> K A D E Y L <mark>R E</mark></td> <td>IQELGQL</td> <td>IQA VQQCI</td> <td>EAA<mark>G</mark>H EH</td> <td><mark>Q</mark> P D M <mark>Q K S L</mark></td> <td>L R A <mark>A </mark>S F G K <mark>C</mark> F</td> <td>L D R </td> <td>FP<mark>PDS</mark>I</td> <td>F V H M C Q D L</td> <td>RV LN<mark>A</mark></td> <td>IR<mark>DYH</mark>IG</td> <td>IPLTYT</td> <td>Т<mark>Q Y K Q</mark></td>	Mus musculus	S <mark>Q</mark> K A D E Y L <mark>R E</mark>	IQELGQL	IQA VQQCI	EAA <mark>G</mark> H EH	<mark>Q</mark> P D M <mark>Q K S L</mark>	L R A <mark>A </mark> S F G K <mark>C</mark> F	L D R	FP <mark>PDS</mark> I	F V H M C Q D L	RV LN <mark>A</mark>	IR <mark>DYH</mark> IG	IPLTYT	Т <mark>Q Y K Q</mark>
Bos taurus SQKADEYLRE IQELGQLPQA VQQCIEAAGH EHWPDMQKSL LRAASFGKCF LDRFPPDS FVRMCQDLRV LNAIRDYHIG IPLTYSQYKQ Homo sapiens SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDRFPPDS FVHMCQDLRV LNAVRDYHIG IPLTYSQYKQ Rattus SQKADEYLRE IQELGQLIQA VQQCIEAAGH EHQPDMQKRL LRAASFGKCF LDRFPPDS FVHMCQDLRV LNAIRDYHIG IPLTYSQYKQ Caenorhabditis TGHNSFAAST VIQDLYKA IDDCISTACD TWQPEQKLL LKAAFGMAY Ascaris suum SHGVYSYFKM IEQQFDKA VRQCLFAAAH QFDHSMQKRL LKAASLGNAL Saccharomyces APKAIEILKN FVLEKG VLDCIAAAID EFEPKLQKML LNAASYGKAS Consistency 9587888855 5666048449 947*97956 963776*7* 887*59*959 7348055875 8756574999 **67954497 7899659755 1 2 2 3 4	Felis catus	S <mark>Q</mark> K A D E Y L <mark>R E</mark>	IQELGQL	TQA VQQCI	EAA <mark>G</mark> H EH	R P D M Q K S L	L R A <mark>A </mark> F G K <mark>C</mark> F	L D R	FP <mark>PDS</mark> I	F V <mark>R M C Q D</mark> L	RV LN <mark>A</mark>	IR <mark>DYH</mark> IG	IPLTYS	<mark>Q Y K Q</mark>
Homo sapiens SQKADEYLRE IQELGQLTQA VQQCIEAAGH EHQPDMQKSL LRAASFGKCF LDRFPPDS FVHMCQDLRV LNAVRDYHIG IPLTYSQYKQ Rattus SQKADEYLRE IQELGQLIQA VQQCIEAAGH EHQPDMQKRL LRAASFGKCF LDRFPPDS FVHMCQDLRV LNAVRDYHIG IPLTYSQYKQ Caenorhobditis TGHNSFAAST VIQDLYKA IDDCISTACD TWQPEEQKLL LKAARFGMAY TNT-TPDTTK LMRAIKEIRV LNELMVYTG IPLTHRQFRA Saccris suum SHGVYSYFKM IEQQFDKA VRQCLFAAAH QFDHSNQKRL LKAASIGNAL QRHDPSQ FVMCQDLRV LNAVRQYIG MPLSFLQSND Saccris suum SHGVYSYFKM IEQQFDKA VRQCLFAAAH QFDHSNQKRL LKAASIGNAL QRHDPSQ FVMCQNVRV LNAVRQYIG MPLSFLQSND Saccris suum SHGVYSYFKM IEQQFDKA VRQCLFAAAH QFDHSNQKRL LNAASYGKAS GVXSFDASI FVNACNTIKL LNCFRSFG IFLTYEQYKQ Saccharomyces APKAIEILKN FVLEKG VLDCIAAAID EFEPKLQKML LNAASYGKAS B756574999 *67954497 7899659755 Consistency 9587888855 5660048449 947*979956 963776*7* </td <td>Bos taurus</td> <td>S <mark>Q</mark> K A D E Y L <mark>R E</mark></td> <td>IQELGQL</td> <td>PQ<mark>A VQ</mark>QCI</td> <td>EAA<mark>GH E</mark>H</td> <td>W P D M <mark>Q K S L</mark></td> <td>L R A <mark>A </mark>F G K <mark>C</mark> F</td> <td>L D R </td> <td>F P <mark>P D S I</mark></td> <td>F V R M C Q D L</td> <td>RV LN<mark>A</mark></td> <td>I R <mark>D Y H</mark> I G</td> <td>IPLTYS</td> <td><mark>Q Y K Q</mark></td>	Bos taurus	S <mark>Q</mark> K A D E Y L <mark>R E</mark>	IQELGQL	PQ <mark>A VQ</mark> QCI	EAA <mark>GH E</mark> H	W P D M <mark>Q K S L</mark>	L R A <mark>A </mark> F G K <mark>C</mark> F	L D R	F P <mark>P D S I</mark>	F V R M C Q D L	RV LN <mark>A</mark>	I R <mark>D Y H</mark> I G	IPLTYS	<mark>Q Y K Q</mark>
Rattus SQKADEYLRE IQELGQLIQA VQQCIEAAGH EHQPDMQKRL LRAASFGKCF LDRFPPDS FVHMCQDLRV LNAIRDYHIG IPLTYTQYKQ Caenorhabditis IGENSFAAST VIQDLYKA IDDCISTACD TWQPEEQKLL LKAAFGMAY TNT-TPDTTK LMRAIKEIRV LNELRMVRTG IPLTYTQYKQ Ascaris suum SHGVYSYFKM IEQQFDKA VRQCLFAAAH QFDHSMQKRL LKAASLGNAL LQRHDPSQ FVMCQDLRV LNAVRQPYIG MPLSFLQSND Saccharomyces APKAIEILKN FVLEKG VLDCIAAAID EFEPKLQKML LNAASYGKAS LQYKSFDASI FVNACNTIKL LNCFRSFG IFLTVEYRC Consistency 9587888855 5660048449 947*979956 963776*77 887*59*959 8756574999 **67954497 7899659755 1 2 3 4	Homo sapiens	S <mark>Q</mark> K A D E Y L <mark>R E</mark>	IQELGQL	TQA VQQCI	EAA <mark>G</mark> H EH	Q P D M Q K S L	L R A <mark>A </mark> F G K <mark>C</mark> F	L D R	FP <mark>PDS</mark> I	F V H M C Q D L	RV LN <mark>A</mark>	V R <mark>D Y H</mark> I G	IPLTYS	<mark>Q Y K Q</mark>
CaenorhobditisTGHNSFAAST VIQDLYKA IDDCISTACD TWQPEEQKLL LKAARFGMAY TNT-TPDTTK LMRAIKEIRV LNELRMVRTG IPLTHRQFRA Ascaris suum SHGVYSYFKM IEQQFDKA VRQCLFAAAH QFDHSMQKRL LKAARSLGNAL LQRHDPSQ FVDMCHVMRV LNAVRQPYIG MPLSFLQSND Saccharomyces APKAIEILKN FVLEKG VLDCIAAAID EFEPKLQKML LNAASYGKAS FVNACNTIKL LNCFRSFG IFLTVEEYRC Consistency 9587888855 5660048849 947*979956 963776*7* 887*59*959 340055875 8756574999 **67954497 7899659755 1 2 3 4	Rattus	S <mark>Q</mark> K A D E Y L <mark>R E</mark>	IQELGQL	IQ <mark>A VQ</mark> QCI	EAA <mark>GH E</mark> H	Q P D M <mark>Q K R L</mark>	L R A <mark>A </mark> S F G K <mark>C</mark> F	L D R	F P <mark>P D S I</mark>	F V H M C Q D L	RV LN <mark>A</mark>	IR <mark>DYH</mark> IG	IPLTYT	Т <mark>Q Y</mark> K Q
Ascaris suum SHGVYSYFKM IEQQFDKA VRQCLFAAAH QFDHSMQKRL LKAASLGNAL LQRHDPSQ FVDMCHVMRV LNAVRQPYIG MPLSFLQSND Saccharomyces APKAIEILKN FVLEKG VLDCIAAAID EFEPKLQKML LNAASYGKAS LQVKSFDASI FVNACNTIKL LNCFRSFG IFLTVEEYRC Consistency 9587888855 5660048449 947 979956 963776*77 887 59 959 7340055875 8756574999 **67954497 7899659755 1 2 2 3 4	Caenorhabditis	T <mark>G</mark> H N S F A A <mark>S T</mark>	VIQDL	YKA I <mark>d</mark> dci:	STA <mark>C</mark> D TW	Q P E E <mark>Q K L L</mark>	L K A A <mark>R</mark> F G M <mark>A</mark> Y	T N T - T I	PD <mark>TTK I</mark>	L M R A I K E I	RV LNE	L R <mark>M V R</mark> T G	IPLTH	L <mark>QFRA</mark>
Saccharomyces APKAIEILKN FVLEKG VLDCIAAAID EFEPKLQKML LNAASYGKAS LQYKSFDASI FVNACNTIKL LNCFRSFG IFLTVEEYRC Consistency 9587888855 5660048449 947*979956 963776**7* 887*59*959 7340055875 8756574999 **67954497 7899659755 1 2 3 4 4	Ascaris suum	S <mark>H</mark> G V Y S Y F <mark>K M</mark>	IEQQF	DKA VRQCLI	FAA <mark>AH QF</mark>	<mark>D</mark> H S M <mark>Q K R L</mark>	L K A A <mark>S</mark> L G N <mark>A</mark> L	L Q R I	hd <mark>psq</mark> I	F V D M C H V M	RV LN <mark>A</mark>	V R <mark>Q P Y</mark> I G	MPLSFL	Q S N D
Consistency 9587888855 5660048449 947 979956 963776 7 887 59 959 7340055875 8756574999 767954497 7899659755 	Saccharomyces	A <mark>P</mark> KAIEIL <mark>KN</mark>	F V L	EK <mark>G VLDCI</mark>	AAA <mark>I</mark> D EF	E <mark>pkl<mark>qk</mark>ml</mark>	L N A <mark>S</mark> Y G K <mark>A</mark> S	LQYKS	FD <mark>ASI I</mark>	F V <mark>N A C N T</mark> I	KL LNC	FR <mark>S</mark> FG	IFLTVE	EYRC
<u>1</u> 2 <u>3</u> 4	Consistency	9 <mark>5</mark> 8 7 8 8 8 8 <mark>5 5</mark>	5 <mark>6600</mark> 4 <mark>8</mark>	44 <mark>9 9</mark> 47*9	7 9 9 <mark>5</mark> 6 9 6	<mark>3</mark> 7 7 6 * * 7 *	8 8 7 <mark>* 5</mark> 9 * 9 <mark>5</mark> 9	<mark>7</mark> 3400	55 <mark>875</mark>	8 7 <mark>5 6 5</mark> 7 4 <mark>9</mark>	99 ** <mark>6</mark>	7 9 <mark>5 4 4</mark> 9 7	7 8 9 9 <mark>6</mark> 5	9755
<u> 1 2 3 4</u>										_				
1 2 3 4														
				1			2				3		4	

The color and number assignments are:

Unconserved 0 1 2 3 4 5 6 7 8 9 10 Conserved

Figure 4 Continued.

invariant among the investigated species, with the exception of *Caeno-rhabditis elegans* and *Saccharomyces cerevisiae*.

mah contains domains found in amino acid transporters and permeases

mah is a recessive eye color mutant on the right arm of the third chromosome with a recombination map position of 88. It was discovered by Beadle (Lindsley and Zimm 1968). Homozygotes have decreases in ommochrome pigments and six of 10 pteridine pigments (Ferre *et al.* 1986).

Deficiency mapping showed that the possible *mah* genes were *Nmnat* and *CG13646* in the 96B cytological region (Table 2). *Nmnat* is Nicotinamide mononucleotide adenylyltransferase and is involved in

NAD synthesis and photoreceptor cell maintenance (Zhai *et al.* 2006). The *CG13646* gene contains transmembrane domains and an amino acid transporter domain (dos Santos *et al.* 2015; Romero-Calderon *et al.* 2007). Along with functions that could relate to pigmentation, both genes had expression patterns consistent with an eye color gene (dos Santos *et al.* 2015) *Nmnat* and *CG13646* were sequenced using DNA from the *mah* strain.

The *Nmnat* allele sequence matched the Flybase sequence in its coding region and introns. The *CG13646* gene had 10 single base pair substitutions and an insertion of 433 bp, including a duplication of a 5-bp insert sequence in its fifth exon (Table S3). The additional DNA is a solo insert of a *roo* LTR (Meyerowitz and Hogness 1982; Scherer *et al.* 1982). For *CG13646*, only one substitution



Figure 5 The *CG13646* gene rescues the *mahogany* gene. (A) Wild-type genotype and phenotype. (B) *mah/mah* mutant genotype and *CG13646* transgene under the control of the Gal4 driver *P*{*GAL4-ninaE.GMR*}12 shows the wild-type phenotype. (C) *mah/mah* genotype shows the mutant phenotype.

										460			•							470
MAH	v	D	L	Е	M	E	P	L	L	v	т	s	P	P	т	т	P	R	G	С
Dmel	v	D	L	E	M	E I	P	L	L	v	I	s	P	P	т	т	P	R	G	С
Dper	v	E	L	E	M	E	P	L	L	v	A	A	ĸ	P	N	s	P	R	G	С
Dpse	v	E	L	E	M	E I	Р	L	L	v	A	A	ĸ	P	N	s	P	R	G	С
Dere	v	D	L	E	M	E I	Р	L	L	v	т	s	P	P	т	т	P	R	G	С
Dsec	v	E	L	E	M	E I	Р	L	L	v	т	s	P	P	т	т	P	R	G	С
Dsim	v	E	L	E	M	E I	P	L	L	v	т	s	P	P	т	т	P	R	G	С
Dyak	v	D	L	E	M	E I	P	L	L	A	т	т	P	P	т	т	P	H	G	С
Dana	I	E	L	E	M	E I	P	L	L	v	s	т	P	s	т	т	P	R	G	С
Dvir	I	E	L	E	M	E I	P	L	L	v	v	A	ĸ	P	н	т	т	Q	G	С
Dmoj	v	E	L	E	M	E I	P	L	L	v	A	ĸ	ĸ	P	н	s	s	P	G	С
Dgri	I	D	L	E	M	E I	P	L	L	v	E	т	ĸ	s	H	т	P	R	G	L
Dwil	v	E	L	E	M	E I	P	L	L	V	A	s	Q	P	н	s	P	R	G	С
Consistency	9	8	*	*	*	*	*	*	*	9	5	6	5	8	5	8	8	7	*	8

Figure 6 PRALINE alignment and consistency scores of predicted residues 450–470 with the MA-HOGANY protein sequence in *Drosophila* species. 0 is the least conserved alignment position, increasing to 10 for the most conserved alignment position. Asterisks in the consistency rating indicate identity for all. The MAH sequence is not included in the consistency rating. The change in sequence in the MAH protein is shaded black in the first line. A number of other species, *D. erecta*, *D. simulans*, *D. yakuba*, and *D. sechellia*, have a T at the same position as mahogany. Species, gene, and protein identifiers are in Table S4.

The color assignments are:

Unconserved 0 1 2 3 4 5 6 7 8 9 10 Conserved

mutation was nonsynonymous, a T > C at 3R:24949334 causing an I460T change if that position were transcribed and translated (Table 3). The UAS-CG13646 transgene under the control of the eye-specific GAL4-ninaE.GMR12 driver rescued *mah/mah* flies (Figure 5).

The insertion would be expected to cause production of a truncated inactive protein product. If the translation of the last exon proceeds through the insert, there will be an early termination signal adding three amino acids coded for by *roo* and deleting 130 amino acids. Comparing the *CG13646* orthologous proteins in 12 *Drosophila* species shows the threonine substitution is at the corresponding position in *D. melanogaster's* close relatives, *D. simulans*, *D. erecta*, and *D. sechellia*, implying that this substitution would be functional if the *mah* transcript were properly spliced and translated (Figure 6).

Unlike the genes discussed above, *mah*'s protein product shows very high sequence similarity to predicted orthologous proteins in other *Drosophila* species but markedly decreased similarity to orthologs in other insects (Figure S1A). It shows relatively low similarity to vertebrate proteins. While Psi Blast searches with the *CG13646* protein found sequence similarity with human and mouse GABA vesicular transporters, another *D. melanogaster* predicted protein, VGAT, shows greater sequence similarity and is the presumptive ortholog of the mammalian protein (Figure S1B).

red is coded for by a gene with a LysM domain and an unknown function

The *red* gene's phenotype is caused by a recessive allele producing flies with dark red-brown eyes and rusty red-colored Malpighian tubules. This gene is located on the third chromosome at cytogenetic position 88B1-B2 and was discovered by Muller (Lindsley and Zimm, 1968). *red* flies show decreases in both ommochromes and some pteridines (dos Santos *et al.* 2015; Ferre *et al.* 1986).

Deletion mapping revealed *red* was coded for by one of three genes, *CG12207*, CG3259, or su(Hw) (Table 2). The su(Hw) gene codes for a DNA-binding protein and mutants have well described phenotypes that do not involve the eyes. The CG3259 gene is involved in micro-tubule binding and the *CG12207* gene codes for a product of unknown function. The exons of *CG12207* and the complete CG3259 gene were

sequenced in three stocks, red^1 , red^{K_1} , and the OreR stock that is the red^{K_1} progenitor.

Compared to the Genbank reference sequence for *CG12207* and *CG3259*, the three stocks had substitutions at 34 sites for *CG12207*. The most telling comparison is red^{K_1} vs. its progenitor. The red^{K_1} stock contained four changes that were absent from OreR (Table 4 and Table S5). One was a missense change producing G51S in the LysM domain of protein isoform A (Table 3). The others were two synonymous changes and one substitution in an untranslated region. The red^1 stock had 26 substitutions compared to the Genbank reference sequence and 12 compared to OreR (Table S5). It carried only one missense mutation, A to C, that produced an N67H change in the LysM domain of the protein isoform A (Table 3 and Table 4).

The *CG3259* sequences from the three stocks had 25 substitutions compared to the Genbank reference sequence (Table 4). The red^{K_1} stock had two changes in sequence that were missing in OreR, one synonymous and one in an untranslated region. These two stocks shared the same predicted protein sequence with four amino acids that varied from the Genbank reference protein sequence. The red^1 flies shared the replacements coding for the four amino acid substitutions and had four more missense changes, resulting in a substitution of eight amino acids compared with the Genbank reference sequence (Table 4 and Table S5).

Tab	ole 4 Ty	pes a	and number	's of s	substitutio	ons in <i>red</i> ¹ , i	red ^{ĸ1} ,	and
OreR	stocks	for	CG12207	and	CG3259	compared	with	the
Genbank reference sequences for CG12207 and CG3259								

			Stock	
Gene	Туре	red ¹	red^{K1}	OreR
CG12207	Missense	1	1	0
CG12207	Synonymous	4	2	0
CG12207	UTR	21	15	14
Total		26	18	14
CG3259	Missense	9	5	5
CG3259	Synonymous	9	6	4
CG3259	UTR	2	3	3
Total		20	14	12



Figure 7 CG 12207 partially complements the red/red genotype. (A) Wild-type genotype and phenotype. (B) red/red genotype and partially wild-type phenotype with a *CG12207* transgene and the *P* {*GAL4-ninaE.GMR*}12 driver present. (C) Red mutant phenotype in red/red fly.

The nucleotide sequences for both genes from OreR and the two *red* stocks were compared to the *Drosophila* Genomic Reference Panel, a database showing nucleotide polymorphisms in a panel of 200 inbred lines of *D. melanogaster* (Mackay *et al.* 2012). All of the *CG12207* replacements were found in the panel, except for each missense change in *red¹* and *red^{K1}* (Figure S2). All of the CG3259 changes also were found in the panel, except one shared by OreR and *red^{K1}*. That substitution resided in an intron. When the CG3259 predicted protein sequences were found in other species and were at variable sites in the protein sequence (data not shown).

Given the presence of unique amino acid changes in *red CG12207* sequences, the UAS-*CG12207* cDNA-A construct was chosen for a rescue experiment. It was able to partially or fully complement *red* flies when expressed with the *GAL4-ninaE.GMR12* driver, demonstrating that *CG12207* is the *red* gene (Figure 7). These data are consistent with the Johnson Laboratory analysis comparing the reference gene sequence to an earlier restriction map that included the *red* gene (Breen and Harte 1991) and the predictions of Cook and Cook found in Flybase (FBrf0225865).

The single domain identified in CG12207 is the LysM domain and it is the site of both red^1 and red^{K1} missense mutations. It is present in all six transcripts of the gene, but the function(s) of the predicted proteins are unknown. The domain in CG12207 is highly conserved in predicted *Drosophila* orthologs (Figure 8). Further, the LysM domain is present in Eubacteria, plants, animals, and fungi.

Examples of LysM domains similar in sequence to that found in *red* were identified b querying NCBI Blast using the *CG12207* LysM domain. These were aligned using T-Coffee (Notredame *et al.* 2000) and the G and N sites mutated in *red* alleles were perfectly conserved from rice to man (Figure 9). In addition, the N site appears conserved in

other LysM motifs (Laroche et al. 2013; Zhang et al. 2009). Glycine is
often found in tight turns in proteins and even a seemingly neutral
change to serine may influence protein conformation (Betts and Russell
2003). While changes of asparagines to histidines are found in proteins,
the exclusive presence of asparagines in the LysM domains examined
supports the idea that the amino acid is important for normal protein

DISCUSSION

function.

The cho gene is essential and functions to acidify cellular compartments

The *cho* gene, *VhaAC39-1*, codes for Vacuolar H⁺ ATPase AC39 subunit *d*. The V-ATPase is found in all eukaryotes and is present in lysosomes, endosomes, and clathrin-coated vesicles. The two sectors of V-ATPases are V_1 and V_0 , each of which contains multiple proteins. V_1 is found outside the organelle/vesicle and interacts with ATP, ADP, and phosphate. V_0 is found in the plasma membrane and transports H⁺ into the compartment (Beyenbach and Wieczorek 2006; Marshansky *et al.* 2014). Subunit *d* of the V_0 complex may function in connecting the V_1 complex to the V_0 complex through interactions with V_1 protein subunits (Allan *et al.* 2005). Intracellular V-ATPases play important roles in trafficking, such as separation of ligands and receptors in endosomes, and degradation in lysosomes (Wang *et al.* 2012b). Acidification regulates and mediates the trafficking of cellular receptors and ligands, such as Notch and Wnt.

The mutation in the *cho* gene is a tryptophan-to-leucine mutation. Tryptophan is the rarest amino acid in eukaryotic proteins, while leucine is the most common (Gaur 2014). Owegi *et al.* (2006) investigated the effects of a change of the corresponding tryptophan to alanine (W325A) in yeast. Loss of the tryptophan decreased the assembly of

CG12207 CG12207	RED ¹ RED ^{K1}	IRHIVEKTDTLQGIALKYGCTTEQIRRA <mark>H</mark> RLFASDSLFLRQFLLV IRHIVEKTDTLQ <mark>S</mark> IALKYGCTTEQIRRANRLFASDSLFLRQFLLV
Dmel		IRHIVEKTDTLQGIALKYGCTTEQIRRANRLFASDSLFLRQFLLV
Dwil		IRHMVDKSDTLQGIALKYGCTTEQLRRANRLFASDSLFLRQFLLV
Dgri		IRH <mark>T</mark> VEKTDTLQGI <mark>S</mark> LKYG <mark>A</mark> TTEQIRRANRLFASDSLFLRQFLLV
Dvir		IRH <mark>T</mark> VEKTDTLQGI <mark>S</mark> LKYG <mark>A</mark> TTEQIRRANRLFASDSLFLRQFLLV
Dmoj		IRH <mark>T</mark> VEKTDTLQGI <mark>S</mark> LKYG <mark>A</mark> TTEQIRR <mark>V</mark> NRLYASDSLFLRQFLLV
Dpse		IRHIVEKTDTLQGIALKYGCTTEQIRRANRLFASDSLFLRQFLLV
Dper		IRHIVEKTDTLQGIALKYGCTTEQIRRANRLFASDSLFLRQFLLV
Dana		IQHIVEKTDTLQGIALKYGCTTEQIRRANRLFASDSLFLRQFLLV
Dere		IRHIVEKTDTLQGIALKYGCTTEQIRRANRLFASDSLFLRQFLLV
Dyak		IRHIVEKTDTLQGIALKYGCTTEQIRRANRLFASDSLFLRQFLLV
Dsec		IRHIVEKTDTLQGIALKYGCTTEQIRRANRLFASDSLFLRQFLLV
Consensus		* * * * * *****

Figure 8 Alignment of LysM Domains in CG12207 orthologs in 11 Drosophila species and in RED¹ and RED^{K1}. The multiple sequence alignment was performed by PSI-Coffee and the illustration made using BoxShade. RED¹ and RED^{K1} proteins were not included in the consensus calculation. Residues that match the consensus sequence are shaded black. Gray regions represent changes to similar amino acids. White regions indicate substitutions to less similar amino acids. Asterisks in the consensus denote identity in all sequences. The red¹ allele (top line) codes for a substitution of G to S and red^{K1} (second line) produces an N to H substitution (both marked in yellow). In wild-type D. melanogaster and the other 10 species, these sites are conserved. Species, gene, and protein identifiers are in Table S4.

RED ¹	IRHIVEKTDT LQGIALKYGC TTEQIRRARR LFASDSLFLR QFLLV-	
RED ^{K1}	IRHIVEKTDT LQSIALKYGC TTEQIRRANR LFASDSLFLR QFLLV-	
Alligator	VEH <mark>RLRAG</mark> DT L <mark>QGIALKY</mark> GV TMEQIKRAN <mark>K LFTN</mark> DCIFLR K <mark>T</mark> L	
Collared Flycatcher	VEH <mark>RLSAG</mark> DT L <mark>Q</mark> GIALKYGV TMEQIKRAN <mark>K LFTN</mark> DCIFLR K <mark>T</mark> L	
Chicken	VEH <mark>RLSAG</mark> DT L <mark>Q</mark> GIALKYGV TMEQIKRANK LFTNDCIFLR K <mark>T</mark> L	
Takifugo	VEH <mark>RVTDS</mark> DT L <mark>Q</mark> GIALKYGV TMEQIKRANK LFSNDCIFLR N <mark>S</mark> L	
Rainbow smelt	MEH <mark>RVIDS</mark> DI LQGIALKYGV IMEQIKRANK L <mark>FSNDC</mark> IFLR N <mark>S</mark> L	
Slime mold	IE <mark>HVLQSQ</mark> DT L <mark>QGLALKY</mark> SS TVGDIKRVN <mark>K IWKD</mark> DTLFLK K <mark>SL</mark> FI-	
House cat	LE <mark>HQLAPG</mark> DT L <mark>AGLALKY</mark> GV TMEQIKRAN <mark>R LYTNDS</mark> IFLK K <mark>T</mark> L <mark>YI</mark> -	
D.virilis	IRH <mark>TVEKT</mark> DT L <mark>Q</mark> GISLKYGA TTEQIRRANR LFASDSLFLR QFL <mark>LV</mark> -	
Mosquito	IRH <mark>DVDKT</mark> DT L <mark>Q</mark> GIALKYGC SMEQIRRINR LLPTDTIFLR PFLMI-	
Ciona	- <mark>kh</mark> slsksdt l <mark>q</mark> gialkygt tteelrrink lyssdsmfir sylmv-	
Flour beetle	IKH <mark>YVSNT</mark> DT L <mark>Q</mark> GIALKYDV TIEQIRRVNR LWASDSLFLK EY <mark>L</mark> LI-	
Homo sapiens	LEH <mark>QLEPG</mark> DT LAGLALKYGV TMEQIKRANR L <mark>YTN</mark> DSIFLK K <mark>L YI</mark> -	
Opossum	VEH <mark>PLEPG</mark> DT L <mark>aglalky</mark> gv tmeqikranr lytndsiflk k <mark>tlyi</mark> -	
House fly	VRH <mark>VVEKT</mark> DT L <mark>QGIALKY</mark> GC TTEQIRRANR L <mark>FAS</mark> DSLFLR Q <mark>FL</mark> MI-	
D. melanogaster	IRH <mark>IVEKT</mark> DT L <mark>Q</mark> GIALKYGC TTEQIRRANR LFASDSLFLR QFL <mark>LV</mark> -	
Mouse	VEH <mark>RVRAG</mark> DT L <mark>Q</mark> GIALKYGV TMEQIKRANK LFTNDCIFLK K <mark>TL</mark> SI-	
Sea urchin	IQH <mark>EIQPG</mark> ET L <mark>Q</mark> GISIKYAV PVEQIKRAN <mark>k LFN-ND</mark> IFMR K <mark>YL</mark> SI-	
Arabidopsis	H <mark>riskf</mark> dt l <mark>agvaiky</mark> gv evadvkkm <mark>nn lvtd</mark> lomfal k <mark>sl</mark> oip	
Malo sina	LEH <mark>RVSRM</mark> DT L <mark>aglaiky</mark> gv eisdikrans L <mark>MTD</mark> SQMFAH K <mark>il</mark> li-	
Corn	LEH <mark>EVTRM</mark> DT L <mark>agiaiky</mark> gv eisdikrans l <mark>vtd</mark> sqmfah k <mark>sl</mark> li-	
Xenopus	IEH <mark>CLSPS</mark> DT L <mark>Q</mark> GIALKYGV TMEQIKRANK L <mark>FST</mark> DCIFLR K <mark>S</mark> L	
Bee	LKH <mark>TVLTT</mark> DT L <mark>Q</mark> GIALKYGV TTEQIRRVNR L <mark>WASDS</mark> LFLR E <mark>H</mark> L <mark>FI</mark> -	
Consistency	66 <mark>*3</mark> 8 <mark>544</mark> 99 * <mark>6</mark> *899**77 76779897*6 9 <mark>565</mark> 758987 6 <mark>4</mark> *250	

Figure 9 PRALINE alignment of predicted LysM domains from a variety of animals and plants. Asterisks in the consistency line indicate identity for all sequences. The RED¹ and RED^{K1} protein sequences are not included in the consistency rating. Sequences were chosen by their similarity to the CG12207 LysM domain sequence. The sites that were mutated in RED¹ and RED^{K1} (shaded black) are normally completely conserved from rice and corn to man. All species, protein, and gene identifiers are in Table S4.

The color assignments are:



the V_0V_1 active enzymes, ATPase activity, and proton transport to about 10% of the normal rate.

The *cho/VhaAC39-1* gene has several other characteristics that are consistent with its role as an important component of trafficking. RNAi experiments demonstrated that its knockdown in *D. melanogaster* resulted in complete lethality (Mummery-Widmer *et al.* 2009). Its ortholog in mice is also essential (Miura *et al.* 2003). The viability of *cho* flies indicates a malfunction of VHAAC39-1 that results in a decrease in activity *vs.* a loss of protein, as seen in knockdowns. Further, *cho* is ubiquitously expressed (Chintapalli *et al.* 2007; Hammonds *et al.* 2013; Tomancak *et al.* 2002, 2007). Studies of *cho*'s effects in all parts of the body could reveal gene interactions between VHAAC39-1 and proteins of other mutant genes participating in the same pathways.

The ma gene is essential and participates in the HOPS complex used in trafficking, endosome maturation, and fusion with lysosomes

The *ma*/*Vps16A* gene product is one of four core proteins in the HOPS complex in metazoans. This role has been confirmed in flies. Pulipparacharuvil *et al.* (2005) have shown that *VPS16A* complexes with DOR and CAR proteins. Takáts *et al.* (2014) demonstrated that the HOPS complex is required for fusion of autophagosomes with the lysosomes. Disruption of the HOPS complex also resulted in increased metastasis and growth of tumors in *Drosophila* (Chi *et al.* 2010). Recent work in humans has shown that VPS 16 is required to recruit VPS 33A to the HOPS complex. The lack of either of these proteins prevented the fusion of lysosomes with endosomes or autophagosomes (Wartosch *et al.* 2015).

In *D. melanogaster*, a *VPS16A* knockdown in the eye resulted in a change in eye color and retinal degeneration due to defects in lysosomal delivery and the formation of pigment granules. An organism-wide knockdown of *VPS16A* caused death (Pulipparacharuvil *et al.* 2005). Like *cho*, *ma* is an essential gene and the *ma* allele produces a partially functional protein. In addition, *ma* RNA is maternally deposited and the gene is expressed widely in larvae and adults (Chintapalli *et al.* 2007; Hammonds *et al.* 2013; Tomancak *et al.* 2002, 2007). The study of *ma* mutant flies for other possible interacting genes would produce an *in vivo* system capable of revealing new gene networks and their sites of action.

The red gene's product has an unknown function and a LysM domain, which is part of a superfamily found in bacteria, plants, and animals

The role of the *red/CG12207* gene product is unknown. The fact that the two *red* mutant alleles had missense substitutions in their LysM domains indicates that the domain is important for the protein's function. In animals, the LysM domain is found either by itself or in combination with TLDC motifs found in putative membrane-bound proteins (Zhang *et al.* 2009). The RED protein has a single LysM domain near the N terminal end of the protein. It lacks the transmembrane region found in most LysM proteins of bacteria, fungi, and plants, and, presumably, remains inside cells *vs.* on their surfaces.

In plants, some LysM proteins function in immune responses. Evidence in animals is mixed. Shi *et al.* (2013) reported that a red swamp crayfish gene carrying the LysM domain, PcLysM, shows increases in its mRNA accumulation when the crayfish are challenged with bacteria. Knockdown of PcLysM mRNA in the animals is

accompanied by a decrease in the antimicrobial response. Laroche *et al.* (2013) reviewed two microarray studies that measured Zebrafish mRNA levels in response to challenges with bacteria. The studies failed to detect a change in the quantities of LysM domain containing mRNAs (Laroche *et al.* 2013). Four *D. melanogaster* genes contain a LysM domain, *CG15471*, *CG17985*, and *mustard* (*mtd*). Only *mtd* also contains a TLDC domain. It is the only *Drosophila* LysM-containing gene that has been investigated with respect to innate immunity. *mtd* has a mutant allele that increases fly tolerance to *Vibrio* infection and decreases the transcription of at least one antimicrobial peptide involved in innate immunity. However, the *mtd* transcript that is most influential in changing sensitivity lacks the LysM domain and carries a TLDC domain (Wang *et al.* 2012b).

The CG12207 protein has two different N terminal sequences. Neither of these appears to be a signal sequence (Petersen et al. 2011). Like the cho and ma genes, CG12207 is maternally deposited and is ubiquitously expressed. Its highest expression is observed in Malpighian tubules (Chintapalli et al. 2007; Hammonds et al. 2013; Tomancak et al. 2002, 2007). Two of the five gene interactions listed in the Flybase Interactions Browser for CG12207 are with proteins related to trafficking: CG16817 influences Golgi organization and ZnT63C transports Zn (Guruharsha et al. 2011). The Golgi apparatus contributes cargo to the endosomes and lysosomes that may be involved in vesicular trafficking. ZnT63C moves Zn out of the cytoplasm either into intracellular compartments or outside the cell membrane (Kondylis et al. 2011; Wang et al. 2009). Another Zn transporter, Catsup, has been shown to disrupt trafficking of NOTCH, Epidermal Growth Factor Receptor, and Drosophila Amyloid Precursor-Like proteins (Groth et al. 2013).

The MAH protein is predicted to be an amino acid transporter that is not essential

InterPro analysis of the MAH predicted protein identified 11 transmembrane helices and a conserved domain found in amino acid transporters (Mitchell *et al.* 2015). The NCBI Conserved Domain Database identified the domain in MAH as the Solute Carrier (SLC) families 5 and 6-like; solute binding domain. Two studies (Romero-Calderon *et al.* 2007; Thimgan *et al.* 2006) fail to place *CG13646* into the *D. melanogaster* SLC6 protein group and Romero-Calderon *et al.* (2007) suggest that the protein may be an amino acid permease.

Unlike *cho* and *ma*, *mah* is not essential because the predicted protein in the *mah/mah* fly is a truncated, and probably inactive enzyme, yet these flies are viable. The expression of *mah* is limited, being ranked as present in the larval central nervous system, adult eye, Malpighian tubule, and testes (Chintapalli *et al.* 2007). Its time of highest expression is the white prepupal stage when eye pigmentation begins (Graveley *et al.* 2011).

Disruption of both types of pigments is a good criterion to identify genes involved in Drosophila vesicular trafficking

Four candidate genes were identified as possible granule group genes based on one characteristic, decreases in the amounts of both ommochromes and pteridines found in fly eyes. Of these, *ma* is the best example of such a gene since it is a member of the HOPS complex, like some previously identified granule group genes. The vesicular ATPase, *cho*, is also very important in vesicle maturation and function. The *mah* gene may well be important in transport to pigment granules given that it appears to be a membrane protein similar to other amino acid carriers. The *red* gene's function and significance in transport are unknown. *cho* and *ma* can be used to study the involvement of their products in a variety of processes when their proteins are expressed *in vivo*. These two genes are essential, but the mutants are visible and viable without partial RNAi knockdowns. Further, since the genes are ubiquitously expressed they could be tested in screens detecting changes in tissues and organs other than the eye. They can also be used in screens in which the tested genes are knocked down by RNAi. The use of mutants in whole organisms to understand the effects of genes and genetic interactions is very powerful.

Trafficking is a complex phenomenon in which many genes participate. A simple criterion allows the discovery of more *Drosophila* trafficking genes and more alleles of identified genes. There is a set of unmapped, eye color mutants that have had their relative pteridine and ommochrome levels tested that could be identified using the techniques of deletion mapping and sequencing. More genes/ alleles could be discovered whose use would contribute to understanding vesicular trafficking.

ACKNOWLEDGMENTS

The authors thank Courtney Smith for many helpful discussions and advice. Ligia Rosario Benavides, Thiago da Silva Moreira, and Robert J. Kallal of the Gustavo Hormiga laboratory assisted us with photomicroscopy. Jim Kennison kindly provided the red^{K1} strain and the Oregon R strain from which it was derived. T.M., A.W., and J.N. were supported by the Undergraduate Research Program for the Biological Sciences Department. The work was supported by intramural funding through The George Washington University Facilitating Fund grants, Columbian College Facilitating Funds and Dilthey grants. The authors have no conflicts of interest.

LITERATURE CITED

- Allan, A. K., J. Du, S. A. Davies, and J. A. Dow, 2005 Genome-wide survey of V-ATPase genes in *Drosophila* reveals a conserved renal phenotype for lethal alleles. Physiol. Genomics 22: 128–138.
- Baillie, D. L., and A. Chovnick, 1971 Studies on the genetic control of tryptophan pyrrolase in *Drosophila melanogaster*. Mol. Gen. Genet. 112: 341–353.
- Besteiro, S., D. Tonn, L. Tetley, G. H. Coombs, and J. C. Mottram, 2008 The AP3 adaptor is involved in the transport of membrane proteins to acidocalcisomes of Leishmania. J. Cell Sci. 121: 561–570.
- Betts, M. J., and R. B. Russell, 2003 Amino acid properties and consequences of subsitutions, in *Bioinformatics for Geneticists*, edited by Barnes, M. R. M. R., and I. C. Gray. John Wiley & Sons, Chichester.
- Beyenbach, K. W., and H. Wieczorek, 2006 The V-type H+ ATPase: molecular structure and function, physiological roles and regulation. J. Exp. Biol. 209: 577–589.
- Bonifacino, J. S., 2004 Insights into the biogenesis of lysosome-related organelles from the study of the Hermansky-Pudlak syndrome. Ann. N. Y. Acad. Sci. 1038: 103–114.
- Bonifacino, J. S., 2014 Vesicular transport earns a nobel. Trends Cell Biol. 24: 3–5.
- Brand, A. H., and N. Perrimon, 1993 Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118: 401–415.
- Breen, T. R., and P. J. Harte, 1991 Molecular characterization of the *tri-thorax* gene, a positive regulator of homeotic gene expression in *Drosophila*. Mech. Dev. 35: 113–127.
- Bridges, C. J., 1918 Maroon-a recurrent mutation in Drosophila. Proc. Natl. Acad. Sci. USA 4: 316-318.
- Chan, C. C., S. Scoggin, D. Wang, S. Cherry, T. Dembo *et al.*, 2011 Systematic discovery of rab GTPases with synaptic functions in *Drosophila*. Curr. Biol. 21: 1704–1715.

- Cheli, V. T., R. W. Daniels, R. Godoy, D. J. Hoyle, V. Kandachar et al., 2010 Genetic modifiers of abnormal organelle biogenesis in a Drosophila model of BLOC-1 deficiency. Hum. Mol. Genet. 19: 861–878.
- Chi, C., H. Zhu, M. Han, Y. Zhuang, X. Wu *et al.*, 2010 Disruption of lysosome function promotes tumor growth and metastasis in *Drosophila*. J. Biol. Chem. 285: 21817–21823.
- Chintapalli, V. R., J. Wang, and J. A. Dow, 2007 Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. Nat. Genet. 39: 715–720.
- Cowles, C. R., G. Odorizzi, G. S. Payne, and S. D. Emr, 1997 The AP-3 adaptor complex is essential for cargo-selective transport to the yeast vacuole. Cell 91: 109–118.
- Dell'Angelica, E. C., 2004 The building BLOC(k)s of lysosomes and related organelles. Curr. Opin. Cell Biol. 16: 458–464.
- Dell'Angelica, E. C., C. Mullins, S. Caplan, and J. S. Bonifacino, 2000 Lysosomerelated organelles. FASEB J. 14: 1265–1278.
- Di Pietro, S. M., J. M. Falcon-Perez, and E. C. Dell'Angelica, 2004 Characterization of BLOC-2, a complex containing the Hermansky-Pudlak syndrome proteins HPS3, HPS5 and HPS6. Traffic 5: 276–283.
- Di Tommaso, P., S. Moretti, I. Xenarios, M. Orobitg, A. Montanyola *et al.*, 2011 T-coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. Nucleic Acids Res. 39: W13–W17.
- dos Santos, G., A. J. Schroeder, J. L. Goodman, V. B. Strelets, M. A. Crosby et al., 2015 FlyBase: introduction of the *Drosophila melanogaster release* 6 reference genome assembly and large-scale migration of genome annotations. Nucleic Acids Res. 43: D690–D697.
- Duffy, J. B., 2002 GAL4 system in *Drosophila*: a fly geneticist's swiss army knife. Genesis 34: 1–15.
- Falcon-Perez, J. M., R. Romero-Calderon, E. S. Brooks, D. E. Krantz, and E. C. Dell'Angelica, 2007 The *Drosophila* pigmentation gene *pink* (*p*) encodes a homologue of human Hermansky-Pudlak syndrome 5 (HPS5). Traffic 8: 154–168.
- Faundez, V., J. T. Horng, and R. B. Kelly, 1998 A function for the AP3 coat complex in synaptic vesicle formation from endosomes. Cell 93: 423–432.
- Ferre, J., F. J. Silva, M. D. Real, and J. L. Mensua, 1986 Pigment patterns in mutants affecting the biosynthesis of pteridines and xanthommatin in *Drosophila melanogaster*. Biochem. Genet. 24: 545–569.
- Gaur, R. K., 2014 Amino acid frequency distribution among eukaryotic proteins. IIOABJ 5: 6.
- Gloor, G. B., C. R. Preston, D. M. Johnson-Schlitz, N. A. Nassif, R. W. Phillis et al., 1993 Type I repressors of P element mobility. Genetics 135: 81–95.
- Goldberg, M. L., R. Paro, and W. J. Gehring, 1982 Molecular cloning of the white locus region of Drosophila melanogaster using a large transposable element. EMBO J. 1(1): 93–98.
- Graveley, B. R., A. N. Brooks, J. W. Carlson, M. O. Duff, J. M. Landolin et al., 2011 The developmental transcriptome of *Drosophila melanogaster*. Nature 471: 473–479.
- Groth, C., T. Sasamura, M. R. Khanna, M. Whitley, and M. E. Fortini, 2013 Protein trafficking abnormalities in *Drosophila* tissues with impaired activity of the ZIP7 zinc transporter *catsup*. Development 140: 3018–3027.
- Guruharsha, K. G., J. F. Rual, B. Zhai, J. Mintseris, P. Vaidya et al., 2011 A protein complex network of Drosophila melanogaster. Cell 147: 690–703.
- Hammonds, A. S., C. A. Bristow, W. W. Fisher, R. Weiszmann, S. Wu et al., 2013 Spatial expression of transcription factors in Drosophila embryonic organ development. Genome Biol. 14: R140–2013–14–12-r140.
- Hijazi, A., W. Masson, B. Auge, L. Waltzer, M. Haenlin *et al.*, 2009 Boudin is required for septate junction organisation in *Drosophila* and codes for a diffusible protein of the Ly6 superfamily. Development 136: 2199–2209.
- Huizing, M., A. Helip-Wooley, W. Westbroek, M. Gunay-Aygun, and W. A. Gahl, 2008 Disorders of lysosome-related organelle biogenesis: clinical and molecular genetics. Annu. Rev. Genomics Hum. Genet. 9: 359–386.
- Hutagalung, A. H., and P. J. Novick, 2011 Role of rab GTPases in membrane traffic and cell physiology. Physiol. Rev. 91: 119–149.

- Kent, H. M., P. R. Evans, I. B. Schafer, S. R. Gray, C. M. Sanderson *et al.*, 2012 Structural basis of the intracellular sorting of the SNARE VAMP7 by the AP3 adaptor complex. Dev. Cell 22: 979–988.
- Kim, J., H. Suh, S. Kim, K. Kim, C. Ahn et al., 2006 Identification and characteristics of the structural gene for the Drosophila eye colour mutant sepia, encoding PDA synthase, a member of the omega class glutathione S-transferases. Biochem. J. 398: 451–460.
- Kondylis, V., Y. Tang, F. Fuchs, M. Boutros, and C. Rabouille, 2011 Identification of ER proteins involved in the functional organisation of the early secretory pathway in *Drosophila* cells by a targeted RNAi screen. PLoS One 6: e17173.
- Laroche, F. J., C. Tulotta, G. E. Lamers, A. H. Meijer, P. Yang *et al.*, 2013 The embryonic expression patterns of zebrafish genes encoding LysM-domains. Gene Expr. Patterns 13: 212–224.
- Li, X., A. G. Garrity, and H. Xu, 2013 Regulation of membrane trafficking by signalling on endosomal and lysosomal membranes. J. Physiol. 591: 4389–4401.
- Lindsley, D. L., and G. G. Zimm, 1968 The Genome of Drosophila Melanogaster. Carnegie Institute, Washington, DC.
- Lloyd, V., M. Ramaswami, and H. Kramer, 1998 Not just pretty eyes: Drosophila eye-colour mutations and lysosomal delivery. Trends Cell Biol. 8: 257–259.
- Ma, J., H. Plesken, J. E. Treisman, I. Edelman-Novemsky, and M. Ren, 2004 Lightoid and claret: a rab GTPase and its putative guanine nucleotide exchange factor in biogenesis of Drosophila eye pigment granules. Proc. Natl. Acad. Sci. USA 101: 11652–11657.
- Mackay, T. F., S. Richards, E. A. Stone, A. Barbadilla, J. F. Ayroles *et al.*, 2012 The *Drosophila melanogaster* genetic reference panel. Nature 482: 173–178.
- Mackenzie, S. M., M. R. Brooker, T. R. Gill, G. B. Cox, A. J. Howells et al., 1999 Mutations in the *white* gene of *Drosophila melanogaster* affecting ABC transporters that determine eye colouration. Biochim. Biophys. Acta 1419: 173–185.
- Marshansky, V., J. L. Rubinstein, and G. Gruber, 2014 Eukaryotic V-ATPase: novel structural findings and functional insights. Biochim. Biophys. Acta 1837: 857–879.
- Meyerowitz, E. M., and D. S. Hogness, 1982 Molecular organization of a *Drosophila* puff site that responds to ecdysone. Cell 28: 165–176.
- Mitchell, A., H. Y. Chang, L. Daugherty, M. Fraser, S. Hunter *et al*, 2015 The InterPro protein families database: the classification resource after 15 years. Nucleic Acids Res. 43: D213–D221.
- Miura, G. I., G. J. Froelick, D. J. Marsh, K. L. Stark, and R. D. Palmiter, 2003 The d subunit of the vacuolar ATPase (Atp6d) is essential for embryonic development. Transgenic Res. 12: 131–133.
- Morgan, T. H., 1910 Sex limited inheritance in *Drosophila*. Science 32: 120–122.
- Mullins, C., L. M. Hartnell, D. A. Wassarman, and J. S. Bonifacino, 1999 Defective expression of the mu3 subunit of the AP-3 adaptor complex in the *Drosophila* pigmentation mutant *carmine*. Mol. Gen. Genet. 262: 401– 412.
- Mullins, C., L. M. Hartnell, and J. S. Bonifacino, 2000 Distinct requirements for the AP-3 adaptor complex in pigment granule and synaptic vesicle biogenesis in *Drosophila melanogaster*. Mol. Gen. Genet. 263: 1003–1014.
- Mummery-Widmer, J. L., M. Yamazaki, T. Stoeger, M. Novatchkova, S. Bhalerao *et al.*, 2009 Genome-wide analysis of *Notch* signalling in *Drosophila* by transgenic RNAi. Nature 458: 987–992.
- Nickerson, D. P., C. L. Brett, and A. J. Merz, 2009 Vps-C complexes: gatekeepers of endolysosomal traffic. Curr. Opin. Cell Biol. 21: 543–551.
- Nolte, D. J., 1955 The eye pigmentary system of *Drosophila* VI. The pigments of the ruby and red groups of genes. J. Genet. 53: 1-10.
- Notredame, C., D. G. Higgins, and J. Heringa, 2000 T-coffee: a novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302: 205–217.
- Ooi, C. E., J. E. Moreira, E. C. Dell'Angelica, G. Poy, D. A. Wassarman et al., 1997 Altered expression of a novel adaptin leads to defective pigment granule biogenesis in the *Drosophila* eye color mutant garnet. EMBO J. 16: 4508–4518.

Owegi, M. A., D. L. Pappas, M. W. Finch, Jr, S. A. Bilbo, C. A. Resendiz *et al.*, 2006 Identification of a domain in the V0 subunit d that is critical for coupling of the yeast vacuolar proton-translocating ATPase. J. Biol. Chem. 281: 30001–30014.

Parks, A. L., K. R. Cook, M. Belvin, N. A. Dompe, R. Fawcett et al., 2004 Systematic generation of high-resolution deletion coverage of the Drosophila melanogaster genome. Nat. Genet. 36: 288–292.

Petersen, T. N., S. Brunak, G. von Heijne, and H. Nielsen, 2011 SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8: 785–786.

Pulipparacharuvil, S., M. A. Akbar, S. Ray, E. A. Sevrioukov, A. S. Haberman et al., 2005 Drosophila Vps16A is required for trafficking to lysosomes and biogenesis of pigment granules. J. Cell Sci. 118: 3663–3673.

- Rahman, M., A. Haberman, C. Tracy, S. Ray, and H. Kramer, 2012 Drosophila mauve mutants reveal a role of LYST homologs late in the maturation of phagosomes and autophagosomes. Traffic 13: 1680–1692.
- Reaume, A. G., D. A. Knecht, and A. Chovnick, 1991 The rosy locus in Drosophila melanogaster: xanthine dehydrogenase and eye pigments. Genetics 129: 1099–1109.

Romero-Calderon, R., R. M. Shome, A. F. Simon, R. W. Daniels, A. DiAntonio *et al.*, 2007 A screen for neurotransmitter transporters expressed in the visual system of *Drosophila melanogaster* identifies three novel genes. Dev. Neurobiol. 67: 550–569.

Scherer, G., C. Tschudi, J. Perera, H. Delius, and V. Pirrotta, 1982 B104, a new dispersed repeated gene family in *Drosophila melanogaster* and its analogies with retroviruses. J. Mol. Biol. 157: 435–451.

Searles, L. L., and R. A. Voelker, 1986 Molecular characterization of the Drosophila vermilion locus and its suppressible alleles. Proc. Natl. Acad. Sci. USA 83: 404–408.

Searles, L. L., R. S. Ruth, A. M. Pret, R. A. Fridell, and A. J. Ali, 1990 Structure and transcription of the *Drosophila melanogaster vermilion* gene and several mutant alleles. Mol. Cell. Biol. 10: 1423–1431.

Sepulveda, F. E., A. Burgess, X. Heiligenstein, N. Goudin, M. M. Menager et al., 2015 LYST controls the biogenesis of the endosomal compartment required for secretory lysosome function. Traffic 16: 191–203.

Seshaiah, P., and D. J. Andrew, 1999 WRS-85D: a tryptophanyl-tRNA synthetase expressed to high levels in the developing *Drosophila* salivary gland. Mol. Biol. Cell 10: 1595–1608.

Shi, X. Z., J. Zhou, J. F. Lan, Y. P. Jia, X. F. Zhao *et al.*, 2013 A lysin motif (LysM)-containing protein functions in antibacterial responses of red swamp crayfish, *Procambarus clarkii*. Dev. Comp. Immunol. 40: 311–319.

Simossis, V. A., and J. Heringa, 2003 The PRALINE online server: optimising progressive multiple alignment on the web. Comput. Biol. Chem. 27: 511–519.

Simossis, V. A., J. Kleinjung, and J. Heringa, 2005 Homology-extended sequence alignment. Nucleic Acids Res. 33: 816–824.

Solanger, J., and A. Spang, 2013 Tethering complexes in the endocytic pathway: CORVET and HOPS. FEBS J. 280: 2743–2757.

Sturtevant, A. H., 1955 New mutants report. D. I. S. 29: 75.

Summers, K., A. Howells, and N. Pyliotis, 1982 Biology of eye pigmentation in insects. Adv. Insect Physiol. 16: 119–166.

- Syrzycka, M., L. A. McEachern, J. Kinneard, K. Prabhu, K. Fitzpatrick et al., 2007 The pink gene encodes the Drosophila orthologue of the human Hermansky-Pudlak syndrome 5 (HPS5) gene. Genome 50: 548–556.
- Takáts, S., K. Pircs, P. Nagy, A. Varga, M. Karpati *et al.*, 2014 Interaction of the HOPS complex with syntaxin 17 mediates autophagosome clearance in *Drosophila*. Mol. Biol. Cell 25: 1338–1354.
- Thimgan, M. S., J. S. Berg, and A. E. Stuart, 2006 Comparative sequence analysis and tissue localization of members of the SLC6 family of transporters in adult *Drosophila melanogaster*. J. Exp. Biol. 209: 3383–3404.
- Tomancak, P., A. Beaton, R. Weiszmann, E. Kwan, S. Shu et al., 2002 Systematic determination of patterns of gene expression during Drosophila embryogenesis. Genome Biol. 3: RESEARCH0088.

Tomancak, P., B. P. Berman, A. Beaton, R. Weiszmann, E. Kwan et al., 2007 Global analysis of patterns of gene expression during *Drosophila* embryogenesis. Genome Biol. 8: R145.

Vermaak, D., S. Henikoff, and H. S. Malik, 2005 Positive selection drives the evolution of *rhino*, a member of the heterochromatin protein 1 family in *Drosophila*. PLoS Genet. 1: 96–108.

Walters, K. B., P. Grant, and D. L. Johnson, 2009 Evolution of the GST omega gene family in 12 Drosophila species. J. Hered. 100: 742–753.

- Wang, C., Z. Liu, and X. Huang, 2012a Rab32 is important for autophagy and lipid storage in *Drosophila*. PLoS One 7: e32086.
- Wang, X., Y. Wu, and B. Zhou, 2009 Dietary zinc absorption is mediated by ZnT1 in *Drosophila melanogaster*. FASEB J. 23: 2650–2661.

Wang, Z., C. D. Berkey, and P. I. Watnick, 2012b The *Drosophila* protein MUSTARD tailors the innate immune response activated by the immune deficiency pathway. J. Immunol. 188: 3993–4000.

Wartosch, L., U. Gunesdogan, S. C. Graham, and J. P. Luzio, 2015 Recruitment of VPS33A to HOPS by VPS16 is required for lysosome fusion with endosomes and autophagosomes. Traffic 16: 727–742.

Wiederrecht, G. J., and G. M. Brown, 1984 Purification and properties of the enzymes from *Drosophila melanogaster* that catalyze the conversion of dihydroneopterin triphosphate to the pyrimidodiazepine precursor of the drosopterins. J. Biol. Chem. 259: 14121–14127.

Zhai, R. G., Y. Cao, P. R. Hiesinger, Y. Zhou, S. Q. Mehta *et al.*, 2006 Drosophila NMNAT maintains neural integrity independent of its NAD synthesis activity. PLoS Biol. 4: e416.

- Zhang, X. C., S. B. Cannon, and G. Stacey, 2009 Evolutionary genomics of LysM genes in land plants. BMC Evol. Biol. 9: 183.
- Zhen, Y., and H. Stenmark, 2015 Cellular functions of rab GTPases at a glance. J. Cell Sci. 128: 3171–3176.

Communicating editor: C. Gonzalez