Insect Biochemistry and Molecular Biology 62 (2015) 168–173

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Molecular evolution and expression of the CRAL_TRIO protein family in insects

Gilbert Smith*, Adriana D. Briscoe

Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA

ARTICLE INFO

Article history:
Received 1 December 2014
Received in revised form
29 January 2015
Accepted 3 February 2015
Available online 13 February 2015

Keywords:
Alpha-tocopherol transfer protein
Retinoid-binding protein
CRAL_TRIO
SEC14
Gene duplication

ABSTRACT

CRAL_TRIO domain proteins are known to bind small lipophilic molecules such as retinal, inositol and Vitamin E and include such gene family members as PINTA, α-tocopherol transfer (ATT) proteins, retinoid binding proteins, and clavesins. In insects, very little is known about either the molecular evolution of this family of proteins or their ligand specificity. Here we characterize insect CRAL_TRIO domain proteins and present the first insect CRAL_TRIO protein phylogeny constructed by performing reciprocal BLAST searches of the reference genomes of *Drosophila melanogaster*, *Anopheles gambiae*, *Apis mellifera*, *Tribolium castaneum*, *Bombyx mori*, *Manduca sexta* and *Danaus plexippus*. We find several highly conserved amino acid residues in the CRAL_TRIO domain-containing genes across insects and a gene expansion resulting in more than twice as many gene family members in lepidopterans than in other surveyed insect species, but no lepidopteran homolog of the PINTA gene in *Drosophila*. In addition, we examined the expression pattern of CRAL_TRIO domain genes in *Manduca sexta* heads using RNA-Seq data. Of the 42 gene family members found in the *M. sexta* reference genome, we found 30 expressed in the head tissue with similar expression profiles between males and females. Our results suggest this gene family underwent a large expansion in lepidopteran, making the lepidopteran CRAL_TRIO domain family distinct from other holometabolous insect lineages.

1. Introduction

The CRAL_TRIO domain is an N-terminal ligand binding region within a larger protein domain (SEC14) that takes its name from common sequence elements in cellular retinaldehyde binding protein (CRALBP) and the signaling domain of the TRIO guanine exchange factor (Panagabko et al., 2003). The domain was first characterized three-dimensionally from structures of yeast phosphatidylinositol transfer protein SEC14p (Sha et al., 1998) and human supernatant protein factor (SPF; Stocker et al., 2002) but is present in most organisms and is likely to be of ancient evolutionary origin (Saito et al., 2007). The domain structure is comprised of alternating α-helices and β-strands (Fig. 1) that bind small lipophilic molecules such as retinal, inositol, squalene, and Vitamin E, the latter of which is composed of α-, β-, γ-, and δ-tocopherols (Panagabko et al., 2003). CRAL_TRIO domains are found alongside many other domains in proteins, including the Golgi dynamics (GOLD), protein tyrosine phosphatase (PTP), RasGAP, RhoGAP, RhoGEF, pleckstrin homology (PH), macro, spectrin repeats (SPEC), Src homology (SH3) and protein kinase domains (Gupta et al., 2012; Saito et al., 2007).

Relatively few members of the large CRAL_TRIO gene family have been studied for ligand specificity. Those that have been studied include human α-tocopherol transfer protein (α-TTP), SPF (which binds squalene), cellular retinaldehyde binding protein (CRALBP, which binds retinal) and *Saccharomyces cerevisiae* phosphatidylinositol transfer protein (SEC14p, which binds phosphatidylinositol and phosphatidylcholine). Ligand specificity of these related proteins has been compared in one study that sought to explain why mammals are able to selectively reuptake α-tocopherol regardless of the mixture of dietary tocopherols ingested (Panagabko et al., 2003). In humans, mutations in α-TTP can result in a neurodegenerative disease called ataxia with isolated vitamin E deficiency (AVED) that causes a variety of neurological symptoms (Ouahchi et al., 1995). While, of the four proteins compared, α-TTP exhibited the highest binding affinity for α-tocopherol, promiscuity in ligand binding was observed in the other three proteins, suggesting that caution should be exercised when attempting to infer

* Corresponding author.
E-mail address: gsmith1@uci.edu (G. Smith).

http://dx.doi.org/10.1016/j.ibmb.2015.02.003
0965-1748/

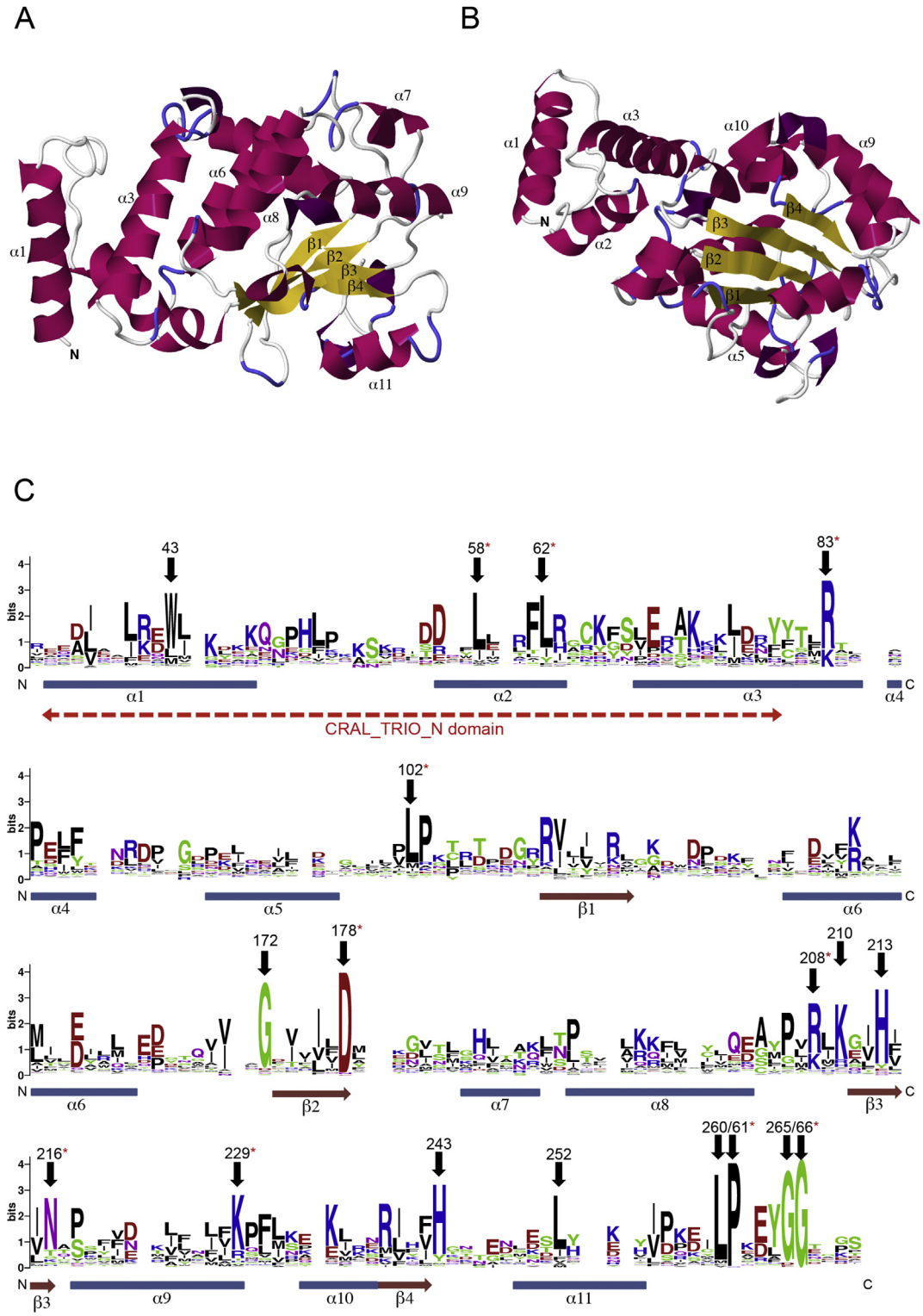

Fig. 1. Protein structure and conservation of the SEC14 domain in insects. The three-dimensional protein structure was predicted using I-TASSER and visualized in Jmol from two different viewpoints (a & b) using the consensus sequence of 159 insect sequences (α-helices shown in pink, β-strands in yellow and 3_{10} helices purple). The sequence logo (c) was created in WebLogo from 159 insect sequences and includes the approximate positions of α-helices (blue bars) and β-strands (red arrows) numbered from the N- to C-terminus, the majority of which are labeled in a & b. The height of each stack is proportional to the sequence conservation, measured in bits, and the height of each letter is proportional to the frequency of residues at that position. The position of the CRAL_TRIO_N domain is shown (dashed red line). Black arrows denote highly conserved amino acid positions with a bit score of >2.5 and frequency of >70% at that site. Numbers above arrows show the position of that amino acid within the yeast SEC14p protein (NCBI accession: NP_013796). Red stars indicate conserved sites between the insect consensus and yeast SEC14p sequences.

function of these proteins from single ligand studies. The SEC14 domain contains a large flexible lipid binding hydrophobic pocket that is closed off by the N-terminal domain, and docking studies suggest different lipids may bind SEC14 in several different ways (Saito et al., 2007). Few studies have identified key residues that interact with particular ligands. However, structural and mutational analysis of yeast SEC14 suggests that residues E207 and K239 are critical for phosphatidylinositol head group binding (Sha et al., 1998), and AVED causing mutations in human α-TTP are located in the binding pocket (Bromley et al., 2013; Min et al., 2003).

Besides α-tocopherol binding and transport by α-TTP, and phosphatidylinositol/phosphatidylcholine exchange and transfer observed in SEC14p, it has been proposed that other gene family members are involved in lipid-mediated regulatory functions in organelles and in intracellular traffic (Saito et al., 2007). For example, members of the clavesin (clathrin vesicle-associated SEC14 proteins; Katoh et al., 2009) gene family are found abundantly in mammalian brain tissue on clathrin-coated vesicles that originate from the *trans*-Golgi network and form the major protein transport pathway from the secretory system to endosomes/lysosomes (Katoh et al., 2009). Experimental work by Katoh et al. (2009) indicates that clavesins specifically function in neurons in a transport pathway between early endosomes and mature lysosomes and that one of the molecules they bind is phosphatidylinositol 3,5-bisphosphate (PtdIns(3,5)P_2). Mutations in mice that deplete PtdIns(3,5)P_2 produce neurodegeneration phenotypes with little effect on other tissues (Zhang et al., 2007).

In insects, the function of only one member of the CRAL_TRIO domain protein, PINTA, has been determined. PINTA or prolonged depolarization afterpotential (PDA) is not apparent was discovered in a screen for *Drosophila* eye mutants that are deficient in rhodopsin biogenesis (Wang and Montell, 2005). Visual pigments or rhodopsins are composed of an opsin protein covalently linked to derivatives of Vitamin A (all-*trans*-retinol), such as 11-*cis*-retinal. Although the cells that compose the vertebrate and invertebrate eye differ significantly, the function of PINTA is somewhat reminiscent of human CRALBP, which is involved in the transport of the visual pigment chromophore, 11-*cis*-retinal, in the photoreceptors of the human eye (Crabb et al., 1988). Humans that carry mutations in the CRALBP gene may develop retinitis pigmentosa or retinitis punctata albescence (Fishman et al., 2004). In *Drosophila*, PINTA is localized to the retinal pigment cells, where all-*trans* 3-hydroxyretinol is converted to 11-*cis* 3-hydroxyretinal (Pak et al., 2012; Wang and Montell, 2005). Ligand-binding assays indicate that PINTA preferentially binds all-*trans*-retinol (Wang and Montell, 2005). It has been proposed that PINTA either sequesters the 3-hydroxy version of all-*trans*-retinol in retinal pigment cells, generating a concentration gradient that promotes uptake of Vitamin A in these cells, and/or that PINTA facilitates the presentation of all-*trans*-retinol to proteins that participate in the oxidation of all-*trans*-retinol to all-*trans*-retinal (Pak et al., 2012; Wang and Montell, 2005; but also see Voolstra et al., 2010).

In the course of annotating vision-related proteins of the hawkmoth (*Manduca sexta)* genome, we discovered that lepidopteran genomes appear to contain a large number of CRAL_TRIO domain proteins. While many of these gene family members are given names in GenBank such as 'α-tocopherol transfer (ATT) protein', 'clavesin' or 'retinaldehyde binding protein', in fact nothing is known experimentally about the function of any of these proteins in insects, except for PINTA. To help begin to bridge this gap in our knowledge, here we characterize for the first time the CRAL_TRIO domain in insects and its molecular phylogeny. Specifically, we determine phylogenetic relationships between gene family members from insects with reference genomes, including the newly sequenced genome of *M. sexta*. We find a number of conserved sites within the SEC14 domain. Further, we see large variation in the size of the gene family, with lepidopterans (butterflies and moths) having the highest number of gene family members among the species examined. We also make use of RNA-Seq data to examine the expression of gene family members in *M. sexta*.

2. Materials and methods

We searched the predicted proteins of the reference genomes of *M. sexta* and *Danaus plexippus* for all homologs of the CRAL_TRIO domain family using tblastn and query sequences from *Drosophila melanogaster, Anopheles gambiae, Apis mellifera, Tribolium castaneum* and *Bombyx mori* homologs, as well as reciprocal tblastn searches of the nr/nt or RefSeq RNA databases (top hits with e-values of $<1 \times 10^{-10}$). We then performed tblastn searches of a *de novo* assembly of the *M. sexta* transcriptome (ArrayExpress accession number E-MTAB-2066; Smith et al., 2014) to find mRNA contigs that were identical to the predicted peptides. Comparison of predicted peptide sequences from the reference genome against the predicted peptide mRNA sequences permitted us to identify and correct errors in the predicted protein sequences. Isoforms found in GenBank were excluded from further analysis, with the first annotated transcript selected for analysis. The identity of the CRAL_TRIO domain-containing proteins was confirmed through a comparison of their SEC14 domains to the Pfam database of HMM-based gene family models using HMMER (Finn et al., 2011). Amino acid sequences were aligned using MUSCLE with default settings. Neighbor-joining analysis of 455 amino acid sites was performed using the Poisson model and pairwise deletion in MEGA (Tamura et al., 2011). Robustness of the phylogeny was tested using 500 bootstrap replicates.

2.1. 3D modeling and sequence logo of the insect CRAL_TRIO/SEC14 domain

The three-dimensional protein structure of the SEC14 domain (containing the CRAL_TRIO domain) was predicted using I-TASSER (Roy et al., 2010), which uses both multiple threading alignments and *ab initio* modeling with further refinement to obtain the most likely 3D structure. Several shorter sequences were omitted from this analysis in order to minimize gaps in the alignment. Thus the 3D structure was predicted from the consensus sequence of 159 insect SEC14 domains, producing a top model with a confidence score (C-score) of 0.87. The three dimensional structure was visualized in Jmol (Jmol Team, 2002) and secondary structures were predicted using Jalview (Waterhouse et al., 2009) from the alignment. A sequence logo was created in WebLogo (Crooks et al., 2004) to examine conservation of the SEC14 domain across insects. Highly conserved amino acid positions were determined as those with a bit score of >2.5 with a frequency of >70% at a particular position.

2.2. RNA-Seq analysis

Bar plots were made using the R package ggplot2 (Wickham, 2009) to visualize the expression levels of members of the CRAL_TRIO domain gene family in *M. sexta*. This data was obtained from Smith et al. (2014) and consisted of mRNA extracted from the heads of four male and four female *M. sexta* adults sequenced on the Illumina platform (50 bp single-end reads), followed by *de novo* transcriptome assembly using Trinity (Grabherr et al., 2011) and RSEM (Li and Dewey, 2011) to assemble contigs that represent mRNA transcripts and obtain expression levels. Raw count levels were normalized by transcript length using the FPKM (Fragments per kilobase of exon per million fragments mapped). FPKM was further normalized between libraries in the R package NOIseq

(Tarazona et al., 2011), using the trimmed mean of M-values (TMM) normalization method. Differences in gene expression between males and females were tested using ANOVA, and p-values were corrected for multiple tests using a False Discovery Rate (FDR) calculated in the R package qvalue (Dabney and Storey, 2014). Genes were defined as significantly differentially expressed when the FDR was <0.05.

3. Results and discussion

Our GenBank searches yielded 43 predicted proteins in the CRAL_TRIO domain family from *B. mori*, 12 from *D. melanogaster*, 14 from *A. gambiae* and 18 from *T. castaneum*. We also found 6 from *A. mellifera*, although this number may have been underestimated due to the poor initial annotation of this genome (Elsik et al., 2014). Our searches of whole-genome sequences yielded 42 predicted CRAL_TRIO domain proteins in the *Manduca sexta* genome (Table S1) and 38 in the *D. plexippus* genome. A variety of isoforms were also evident in each species. A comparison of the blast-identified CRAL_TRIO proteins to existing Pfam HMM protein family models showed the top hit for each to be the CRAL_TRIO domain family (Pfam accession PF00650). All but four proteins had top hit e-values of $<1 \times 10^{-10}$, the remaining four having e-values of $<1 \times 10^{-5}$. *M. sexta* e-values of matches to the CRAL_TRIO domain family are shown in Table S1.

The insect SEC14 consensus protein structure prediction included four β-strands (Fig. 1A and B), compared to six in the yeast phosphatidylinositol transfer protein (Sha et al., 1998) and five in the human supernatant protein factor (Stocker et al., 2002). The overall insect SEC14 structure resembled that of human and yeast SEC14 domain structures, which includes a deep hydrophobic pocket with the β-strands making up the floor of the pocket. Sequence logo analysis of the SEC14 domain, in which the CRAL_TRIO_N domain is embedded (Fig. 1C), indicates the presence of 18 highly conserved amino acid residues 12 of which are conserved compared to yeast SEC14p and 6 of which appear to be conserved across insects (Fig. 1C). The majority of these conserved residues (12 out of 18) were associated with secondary structures

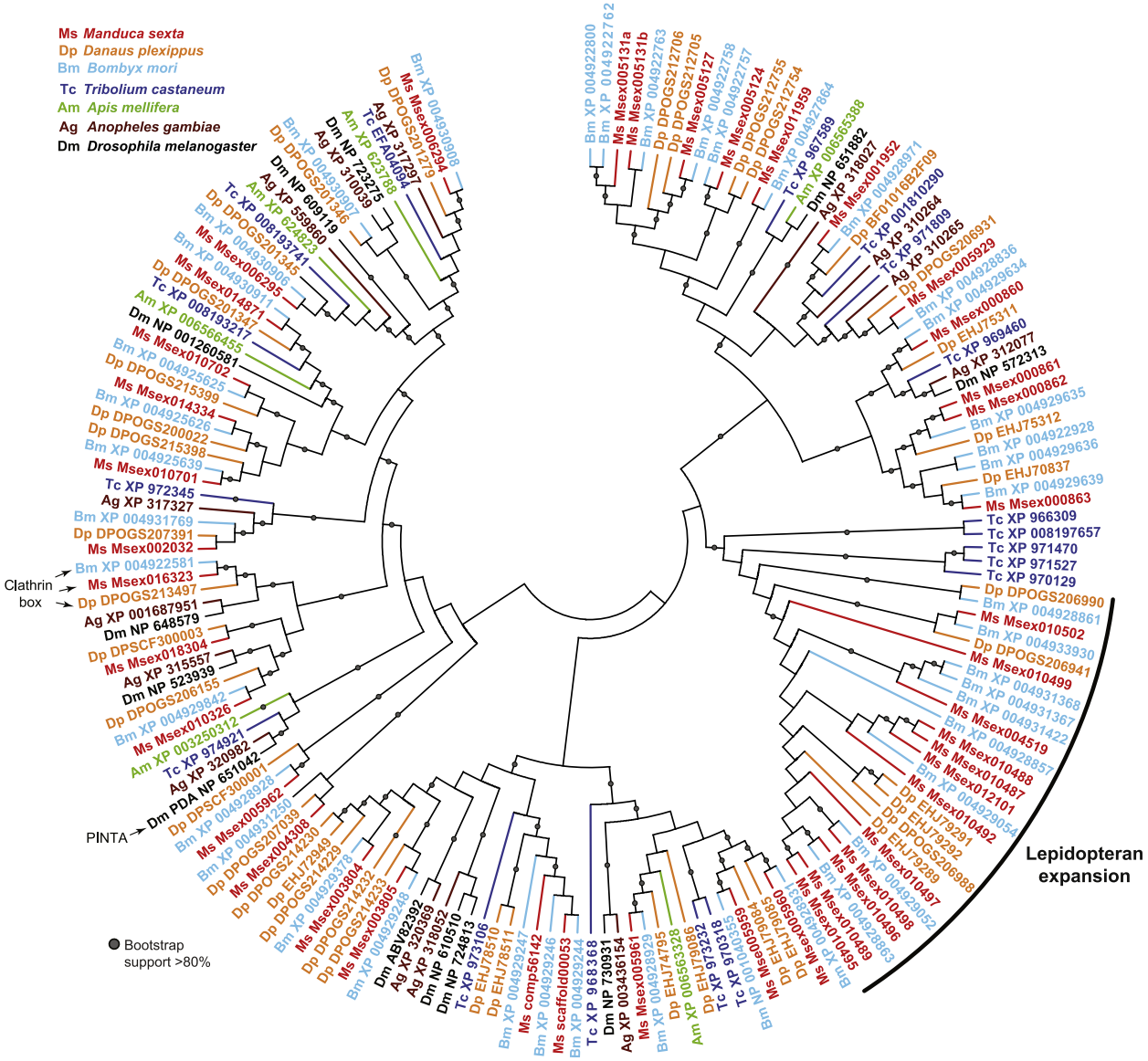

Fig. 2. Phylogeny of the insect CRAL_TRIO domain proteins. The neighbor-joining tree was constructed in MEGA with 500 bootstrap replicates. Bootstrap support values of >80% are shown as dots on branches.

(α-helices or β-strands). Gupta et al. (2012) surveyed CRAL_TRIO domains across a broad range of species, including yeast, plants, fish and humans, yet only one residue (position D178 in yeast SEC14p) is conserved in both their study and here, providing further evidence that the six conserved residues found here could be insect specific.

Phylogenetic analysis of the gene family (excluding isoforms) suggests numerous lineage-specific gene duplications in insects but especially in lepidopterans (Fig. 2). Many gene duplicates are in adjacent positions on the same genomic scaffold suggesting they are tandem arrays of gene duplicates (e.g., *M. sexta* genes Msex005959-Msex005962). In particular, nine genes from the lepidopteran expansion are tandemly arrayed on a single scaffold in *M. sexta* (Msex010487-Msex010499), of which four were expressed in the head tissues (Table S1). Similar patterns were seen for *B. mori* and *D. plexippus* genes within the lepidopteran expansion, and duplication patterns suggest independent expansion events in each species (Fig. 2). This suggests an important role for gene duplication during the evolution of lepidopteran CRAL_TRIO proteins. Orthologs of the PINTA gene in *Drosophila* were not found in any of the three lepidopteran genomes searched. Additionally, we found evidence of mRNA expression in *M. sexta* heads for 30 CRAL_TRIO domain genes (Fig. 3). The majority of these genes had appreciable levels of expression and one in particular, comp17579_c0 (Msex0051227; Table S1), was expressed at a very high level. Few differences in the expression of CRAL_TRIO domain genes were seen between males and females. Only a single gene (comp20575_c0/Msex014871) showed significant differential expression after FDR correction, having a higher expression level in males (Fig. 3).

Clavesins, which are predominantly associated with clathrin-coated vesicles in mammalian brain tissue, contain a clathrin box motif, LLALD, which is putatively involved in clathrin binding (Katoh et al., 2009). We found this motif, which is the result of a four amino acid insertion in the C-terminal domain, to be conserved in three orthologous lepidopteran proteins: *M. sexta* Msex016323, *B. mori* XP_004922581 and *D. plexippus* DPOGS213497 (Fig. 2). Msex016323 (comp35578_c0) is also one of the 30 CRAL_TRIO domain genes expressed in heads, though at relatively low levels compared to other CRAL_TRIO genes (Fig. 3). We did not find this insertion and precise motif in any of the lepidopteran proteins found in GenBank named clavesin-1-like or clavesin-2-like except for *B. mori* XP_004929842 and Msex010326 (comp20338_c0), which contained a related motif, VLALD/N, in a similar position in the protein. Unlike comp35578_c0 (Msex016323), comp20338_c0 *is* expressed at appreciable levels in *Manduca sexa* heads.

It is notable that we did not find an ortholog of the PINTA gene in lepidopterans (Fig. 2). Currently, the identities of the retinoid binding proteins (RBPs) required for transformation of vitamin A to the chromophore remain unclear in most insects. In *Drosophila*, all-*trans* 3-hydroxyretinal can be converted into 11-*cis* 3-hydroxyretinal by light but it is also the case that 11-*cis* 3-hydroxyretinal can be made from dietary Vitamin A. The fly

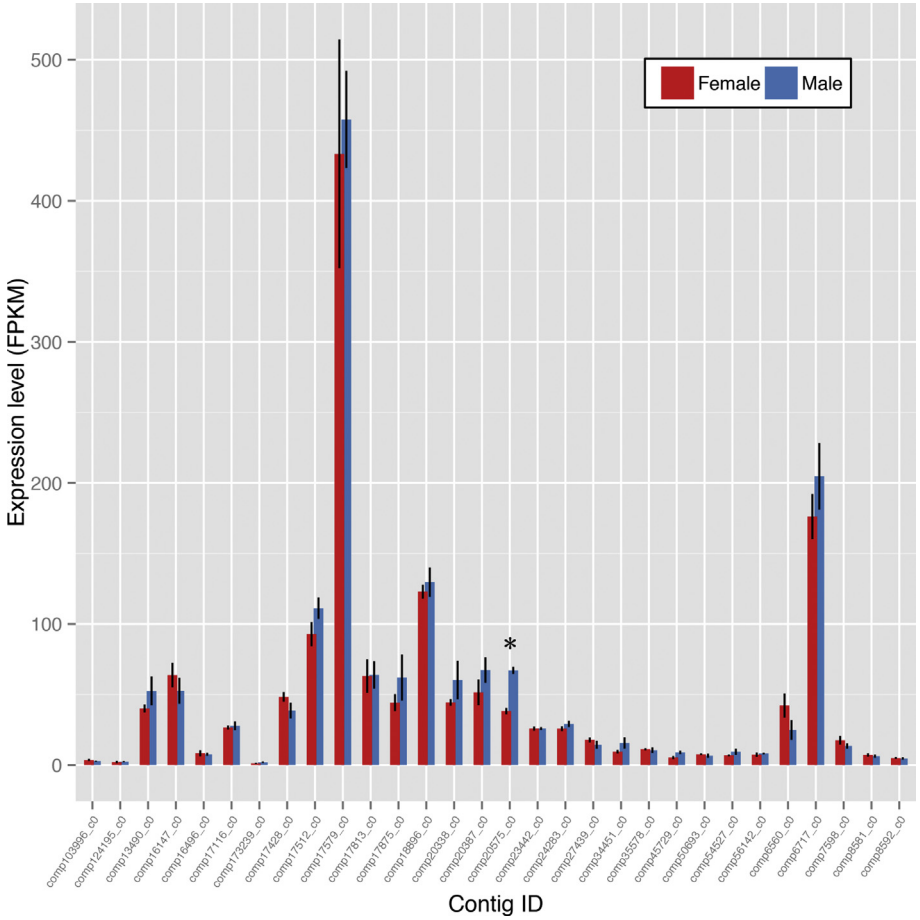

Fig. 3. Expression levels of members of the CRAL_TRIO domain gene family in the head tissue of *Manduca sexta*. Expression level is the mean normalized FPKM for female (red) and male (blue) RNA sequencing libraries. Error bars are the standard errors of mean female and male expression across four biological replicates of each sex. Contig IDs are from the Trinity assembly. Significant sex differences are denoted by asterisks, with significance defined as an FDR of <0.05.

chromophore does not rapidly dissociate from the opsin after light excitation, whereas in butterflies it can (Bernard, 1983a, 1983b; Stavenga and Hardie, 2011). This suggests that in lepidopteran species, production of the chromophore may be even more dependent on enzymatic processes than in *Drosophila*. It is therefore possible that some of the newly-identified CRAL_TRIO domain proteins identified here have taken on that role as retinoid-binding proteins involved in the visual cycle.

We surveyed insect proteins containing the lipid binding CRAL_TRIO domain, including the newly sequenced *M. sexta* genome. We found that the CRAL_TRIO domain gene family has undergone a considerable expansion in lepidopteran species and that a number of amino acid residues within the wider SEC14 domain appear to demonstrate insect-specific conservation. Further, the majority of these genes are expressed in the head tissue in *M. sexta*. Our results suggest that the evolution of CRAL_TRIO domain-containing proteins might play an important role in insect evolution and in particular in the evolution of lepidopteran species, which is potentially linked to the evolution of visual systems.

Acknowledgments

We would like to thank two anonymous reviewers for their constructive comments. This project was funded by a National Science Foundation grant: IOS-1257627.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ibmb.2015.02.003.

References

Bernard, G.D., 1983a. Bleaching of rhabdoms in eyes of intact butterflies. Science 219, 69–71.
Bernard, G.D., 1983b. Dark-processes following photoconversion of butterfly rhodopsins. Biophys. Struct. Mech. 9, 277–286.
Bromley, D., Anderson, P.C., Daggett, V., 2013. Structural consequences of mutations to the α-tocopherol transfer protein associated with the neurodegenerative disease ataxia with vitamin E deficiency. Biochemistry 52, 4264–4273.
Crabb, J.W., Goldflam, S., Harris, S.E., Saari, J.C., 1988. Cloning of the cDNAs encoding the cellular retinaldehyde-binding protein from bovine and human retina and comparison of the protein structures. J. Biol. Chem. 263, 18688–18692.
Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190.
Dabney, A., Storey, J.D., 2014. Qvalue: Q-value Estimation for False Discovery Rate Control. R package version 1.36.0.
Elsik, C.G., Worley, K.C., Bennett, A.K., Beye, M., Camara, F., Childers, C.P., de Graaf, D.C., Debyser, G., Deng, J.X., Devreese, B., Elhaik, E., Evans, J.D., Foster, L.J., Graur, D., Guigo, R., Hoff, K.J., Holder, M.E., Hudson, M.E., Hunt, G.J., Jiang, H.Y., Joshi, V., Khetani, R.S., Kosarev, P., Kovar, C.L., Ma, J., Maleszka, R., Moritz, R.F.A., Munoz-Torres, M.C., Murphy, T.D., Muzny, D.M., Newsham, I.F., Reese, J.T., Robertson, H.M., Robinson, G.E., Rueppell, O., Solovyev, V., Stanke, M., Stolle, E., Tsuruda, J.M., Van Vaerenbergh, M., Waterhouse, R.M., Weaver, D.B., Whitfield, C.W., Wu, Y.Q., Zdobnov, E.M., Zhang, L., Zhu, D.H., Gibbs, R.A., HGSC Production Teams, Honey Bee Genome Sequencing Consortium, 2014. Finding the missing honey bee genes: lessons learned from a genome upgrade. BMC Genomics 15, 86.
Finn, R.D., Clements, J., Eddy, S.R., 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 39 (Web Server issue), W29–W37.
Fishman, G.A., Roberts, M.F., Derlacki, D.J., Gimsby, J.L., Yamamoto, H., Sharon, D., Nishiguchi, K.M., Dryja, T.P., 2004. Novel mutations in the cellular retinaldehyde-binding protein gene (*RLBP1*) associated with retinitis punctata albescens: evidence of interfamilial genetic heterogeneity and fundus changes in heterozygotes. Arch. Ophthalmol. 122, 70–75.
Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.
Gupta, A.B., Wee, L.E., Zhou, Y.T., Hortsch, M., Low, B.C., 2012. Cross-species analyses identify the BNIP-2 and Cdc42GAP homology (BCH) domain as a distinct functional subclass of the CRAL_TRIO/Sec14 superfamily. PloS one 7, e33863.
Jmol Team, 2002. Jmol: an Open-source Java Viewer for Chemical Structures in 3D. http://www.jmol.org/.
Katoh, Y., Ritter, B., Gaffry, T., Blondeau, F., Hönig, S., McPherson, P.S., 2009. The clavesin family, neuron-specific lipid- and clathrin-binding Sec14 proteins regulating lysosomal morphology. J. Biol. Chem. 284, 27646–27654.
Li, B., Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinforma. 12, 323.
Min, K.C., Kovall, R.A., Hendrickson, W.A., 2003. Crystal structure, of human alpha-tocopherol transfer protein bound to its ligand: implications for ataxia with vitamin E deficiency. P. Natl. Acad. Sci. U. S. A. 100, 14713–14718.
Ouahchi, K., Arita, M., Kayden, H., Hentati, F., Ben Hamida, M., Sokol, R., Arai, H., Inoue, K., Mandel, J.L., Koenig, M., 1995. Ataxia with isolated vitamin E deficiency is caused by mutations in the alpha-tocopherol transfer protein. Nat. Genet. 9, 141–145.
Pak, W.L., Shino, S., Leung, H.T., 2012. PDA (prolonged depolarization afterpotential)-defective mutants: the story of *nina's* and *ina's*—*pinta* and *santa maria*, too. J. Neurogenet. 26, 216–237.
Panagabko, C., Morley, S., Hernandez, M., Cassolato, P., Gordon, H., Parsons, R., Manor, D., Atkinson, J., 2003. Ligand specificity in the CRAL-TRIO protein family. Biochemistry 42, 6467–6474.
Roy, A., Kucukural, A., Zhang, Y., 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. 5, 725–738.
Saito, K., Tautz, L., Mustelin, T., 2007. The lipid-binding SEC14 domain. Biochim. Biophys. Acta 1171, 719–726.
Sha, B., Phillips, S.E., Bankaitis, V.A., Luo, M., 1998. Crystal structure of the *Saccharomyces cervisiae* phosphatidylinositol-transfer protein. Nature 391, 506–510.
Smith, G., Chen, Y.R., Blissard, G.W., Briscoe, A.D., 2014. Complete dosage compensation and sex-biased gene expression in the moth *Manduca sexta*. Genome Biol. Evol. 6, 526–537.
Stavenga, D.G., Hardie, R.C., 2011. Metarhodopsin control by arrestin, light-filtering screening pigments, and visual pigment turnover in invertebrate microvillar photoreceptors. J. Comp. Physiol. A 197, 227–241.
Stocker, A., Tomizaki, T., Schulze-Briese, C., Baumann, U., 2002. Crystal structure of the human supernatant protein factor. Structure 10, 1533–1540.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
Tarazona, S., Garcia-Alcalde, F., Dopazo, J., Ferrer, A., Conesa, A., 2011. Differential expression in RNA-seq: a matter of depth. Genome Res. 21, 2213–2223.
Voolstra, O., Oberhauser, V., Sumser, E., Meyer, N.E., Maguire, M.E., Huber, A., von Lintig, J., 2010. *NinaB* is essential for *Drosophila* vision but induces retinal degeneration in opsin-deficient photoreceptors. J. Biol. Chem. 285, 2130–2139.
Wang, T., Montell, C., 2005. Rhodopsin formation in *Drosophila* is dependent on the PINTA retinoid-binding protein. J. Neurosci. 25, 5187–5194.
Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., Barton, G.J., 2009. Jalview version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189–1191.
Wickham, H., 2009. ggplot2: Elegant Graphics for Data Analysis. Springer, New York.
Zhang, Y., Zolov, S.N., Chow, C.Y., Slutsky, S.G., Richardson, S.C., Piper, R.C., Yang, B., Nau, J.J., Westrick, R.J., Morrison, S.J., Meisler, M.H., Weisman, L.S., 2007. Loss of Vac14, a regulator of the signaling lipid phosphatidylinositol 3,5-bisphosphate, results in neurodegeneration in mice. Proc. Natl. Acad. Sci. U. S. A. 104, 17518–17523.