J Mol Evol (2002) 55:470–475
DOI: 10.1007/s00239-002-2342-0

JOURNAL OF MOLECULAR EVOLUTION

The Major Yolk Proteins of Higher Diptera Are Homologs of a Class of Minor Yolk Proteins in Lepidoptera

Thomas W. Sappington

United States Department of Agriculture–Agricultural Research Service, IFNRRU, 2413 East Highway 83, Weslaco, TX 78596, USA

Received: 16 January 2002 / Accepted: 16 April 2002

Abstract. In most oviparous animals, including insects, vitellogenin (Vg) is the major yolk protein precursor. However, in the higher Diptera (cyclorrhaphan flies), a class of proteins homologous to lipoprotein lipases called yolk polypeptides (YP) are accumulated by oocytes instead of Vg, which is not produced at all. Lepidopterans (moths) produce Vg as the major yolk protein precursor, but also manufacture a class of minor yolk proteins referred to as egg-specific proteins (ESP) or YP2s. Although the lepidopteran ESP/YP2s are related to lipoprotein lipases, previous attempts to directly demonstrate their homology with higher-dipteran YPs were unsuccessful. In this paper, a multiple alignment of amino acid sequences was constructed using a shared lipid binding motif as an anchor, to demonstrate that lepidopteran ESP/YP2s, higher-dipteran YPs, and lipoprotein lipases are indeed homologous. Phylogenetic analyses of the aligned sequences were performed using both distance-based and parsimony strategies. It is apparent that the higher dipterans did not requisition a lipoprotein lipase to replace Vg as a yolk protein precursor, but instead utilize a class of proteins with an evolutionary history of use as minor constituents of yolk in other insects.

Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

email: tsapping@weslaco.ars.usda.gov

Key words: Yolk protein — Vitellogenesis — Vitellogenin — Egg specific protein — Lipase — Insects — Lipoprotein

Introduction

The majority of oviparous animals from annelids, nematodes, and insects to vertebrates provision their eggs with vitellogenin (Vg), a product of the vitellogenin gene family (Spieth et al. 1991; Hafer et al. 1992; Chen et al. 1997; Sappington and Raikhel 1998). In insects, Vg is synthesized in the fat body, an organ functionally analogous to the vertebrate liver. It is secreted into the hemolymph, transported to the surface of the oocyte, internalized, and routed to a developing yolk body where it crystallizes as vitellin, the major yolk protein (Raikhel and Dhadialla 1992; Snigirevskaya et al. 1997). Vitellogenins are members of a greater superfamily of large lipid transfer proteins (Babin et al. 1999) that includes the vertebrate serum proteins apolipoprotein B-100 (apoB) and von Willebrand factor (vWf) (Byrne et al. 1989; Chen et al. 1997), among others.

In lepidopterans (moths), several types of smaller yolk protein precursors unrelated to vitellogenins are synthesized in the ovary by the follicular epithelium. Some have been molecularly characterized, including the egg-specific protein (ESP) of the silk moth *Bombyx mori* (Inagaki and Yamashita 1989), and the follicular epithelium yolk protein (YP2) of the pyralid moths *Plodia interpunctella* and *Galleria mellonella*, (Shirk and Perera 1998). Though not as abundant as Vg, they nevertheless constitute 25–40% of the yolk protein (Shirk and Perera 1998). Similarly, the follicular epithelium as well as the fat body of higher (cyclorrhaphan) dipterans (flies) such as *Drosophila* produce several homologous yolk protein precursors referred to by convention as YPs (for yolk polypeptides), which are products of *Yp* genes (Bownes 1992). YP sequences from several species have been reported, and, although they lack enzymatic activity, they are

A

```
             +                    *  **   +             +          +*
PiYP2     GDVADWRGIKYAVGPNSKVLKK.DIESIFYEAQKSMKHPSEETKE.KYHEMYESVSAENI  237
BmESP     GDSVVWKNIEMASGPNSPVQTEQDIEDIFGDSLKTWDHFTDDAKKNTFHDAISETQRENN  185
GmYP2     GDAISSR.ASKSLGPHSQIKKKEQIEKLFADAQMSMKHIDEDSQK.TFHKMYELAVQKED  135
DmYP1     IPALDDFTIERLENMNLERGAE.LLQQVYHLSQIHHNVEPNYVP..SGIQVYVPKPNGDK  103
DmYP2     LPNLKEVTLKKLQEMSMEEGAT.LLDKLYHLSQFNHVFKPDYTPEPSQIRGYIVGERGQK  110
DmYP3     VPSLNDITWERLENQPLEQGAK.VIEKIYHVGQIKHDLTPSFVPSPSNVPVWIIKSNGQK  100
MdYP1     TPSVNELTCEELENMPLEKGDT.LMCKLYHLSQIDYSVSPNFCPSPTNVPVHSFNNKGEK   99
MdYP2     LPSLKEIKLKQLEEMSAFEGAD.LINRLYHLAQATQALEPTYAPRASEIPAFLVTPDNQK  112
MdYP3     TPSLDELTFEELEKMPLEKGAK.LMRKIYHLAQIENSVSPNFVPSPSNVPVYIFNGKGEK   99
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
HsGL                                                      HGLFGKLHPGS.   28
RnLL                                                      HGLFGKLGPGN.   27
DmL1                                                      ETPEDRLQRKNI   54
DmL3

              **             +*       *   +   +              *
PiYP2     EDARLNAT.QLLNKYQYPVEEHVVRTDDGYFLTLFRISKQTEKDTTDEVVQKPVVLLMHS  296
BmESP     EDFHLNAT.ELLKKHQYPVEEHTVRTDDGYHLTVLRIPPTHQTRDD...KKKPVALLMHG  241
GmYP2     EDVYFNAT.QLLKKYGYPVEEHTIQTGDGYYLTVFRIMKYTARRTP.SVASKGVVLLMHG  193
DmYP1     TVAPLNEMIQRLKQKQNFGEDEVTIIVTGLPQTSETVKKAT.RKLVQAYMQRYNLQQQRQ  162
DmYP2     IEFNLNTLVEKVKRQQKFGDDEVTIFIQGLPETNTQVQKAT.RKLVQAYQQRYNLQPYET  169
DmYP3     VECKLNNYVETAKAQPGFGEDEVTIVLTGLPKTSPAQQKAM.RRLIQAYVQKYNLQQLQK  159
MdYP1     ETSNLNKHSTLPEEKPKFDEQEVTVFITGLPQSLEDVKTPN.TKLIQSYIQRYTKKPE..  156
MdYP2     .SIQLNELPRVAREQSHCGKQEVTVFITGLPSKLESVKEAT.RSLTQAYMQRYNDESS..  168
MdYP3     ETCNLNNYVDIAKNKPKFGEQEVTVFITGLPQSLDDVKKAN.TRLIQAYIQRYSQKPT..  156
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
HsGL      PEVTMNIS.QMITYWGYPNEEYEVVTEDGYILEVNRIPYG..KKNSGNTGQRPVVFLQHG   85
RnLL      PEANMNIS.QMITYWGYPQQEYEVVTEDGYILGVYRIPHG..KNNSENIGKRPVVYLQHG   84
DmL1      KQDSTLSVDKLIAKYGYESEVHHVTTEDGYILTMHRIRKQ......GA....PPFLLQHG  104
DmL3      ..RPISDCGERIEDDGYPMERHEVVTSDNYILTMHRIPYS...PKTGESSNRPVAFLMHG   77

                                                   *++     +         *
PiYP2     MLGSADDWLLMGPGQSLAYLLADQGYD.........VWLGNARGNRYTRHHVN.HHAAKA  346
BmESP     LLGSADDWLLMGPSKSLAYMLSDAGYD.........VWLGNVRGNKYSRSHVS.KHPALN  291
GmYP2     LYGSAR.LAPHGPRILLAYLLADEGYE.........VWLGNVRGNKYGRQHVS.KHPAQK  242
DmYP1     HGKNGNQDYQDQSNEQRKNQRTSSEEDYSEEVKNAKTQSGDIIVIDLGSKLNTYERYAML  222
DmYP2     TDYS.N...EEQSQRSSSEEQQTQRRK..QNGEQDDTKTGDLIVIQLGNAIEDFEQYATL  223
DmYP3     NAQE.....QQQQLKSSDYDYTSSEEA.ADQWKSAKAASGDLIIIDLGSTLTNFKRYAML  213
MdYP1     ........APQGEDQS..............KWENEKPVGGHLVVIDLGHAITNVERYATL  194
MdYP2     YYQNSATSSSSSHYHLNKKQQRSDSDE...DNSNQKP.SGCLVVVKFGDTISDFEEHATV  224
MdYP3     ........PPRDDDKS..............KWENEQPVGGHLVVIDLGHTITDMERYASL  194
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
HsGL      LLASATNWISNLPNNSLAFILADAGYD.........VWLGNSRGNTWARRNL....YYSP  132
RnLL      LIASATNWIANLPNNSLAFMLADAGYD.........VWLGNSRGNTWSRKNV....YYSP  131
DmL1      LVDSSAGFVVMGPNVSLAYLLADHNYD.........VWLGNARGNRYSRNHT....TLDP  151
DmL3      MLSSSSDWVLMGPERSLAYMLADAGYD.........VWMGNARGNTYSKAHK....YWPT  124
```

Fig. 1. Multiple alignment of lepidopteran ESP/YP2s with those of dipteran YP, vertebrate lipase, and *Drosophila* lipase amino acid sequences. Positions with at least two functionally similar residues among the lepidopteran ESP/YP2s and at least four functionally similar amino acid residues among the dipteran YPs are boxed, as are any similar residues among the lipases. Residues completely conserved at a position among yolk proteins are indicated by bold white letters on a black background. Positions in which all nine yolk protein sequences contain a functionally similar residue are marked with *, and those with eight of nine functionally similar residues are marked with +. The conserved lipid binding site (Sato and Yamashita 1991; Bownes 1992; Shirk and Perera 1998) which was used to anchor the multiple alignment, is marked with ### over the alignment. Overscore indicates positions used for phylogenetic analyses. Dashed line separates YP and ESP sequences from lipase sequences. Numbers at the right indicate amino acid residue positions from the initial methionine. PiYP2, *Plodia interpunctella* (Lepidoptera) yolk protein 2 (Shirk and Perera 1998); BmESP, *Bombyx mori* (Lepidoptera), egg-specific protein (Inagaki and Yamashita 1989; Sato and Yamashita 1991); GmYP2, *Galleria mellonella* (Lepidoptera) yolk protein 2 (Accession No. U69881); DmYP1-DmYP3, *Drosophila melanogaster* (Diptera) yolk proteins 1–3, respectively (Yan et al. 1987); MdYP1–MdYP3, *Musca domestica* (Diptera) yolk proteins 1–3, respectively (White and Bownes 1997); HsGL, *Homo sapiens* (human) gastric lipase (Bodmer et al. 1987); RnLL, *Rattus norvegicus* (rat) lingual lipase (Docherty et al. 1985); DmL1 and DmL3, *D. melanogaster* lipases 1 and 3, respectively (Pistillo et al. 1998).

B

```
        *    +        *      *+                *  **  *      +
                                               ##########
PiYP2   DFWRYS.......NDDIALHDLPAMIDYALKTTGQRKLFYVGYDLGTTAFFALASTRPEY   399
BmESP   DFWKFS.......NDEIALHDLPAIIDHVLDISGQERLHYIGHSQGATTFFALMSEQPSY   344
GmYP2   DFWQFR.......VDEIARVDLPSLIDYVLQITGQKKLYYVGYDQGTTAFLPMASTMPEY   295
DmYP1   DIEKTGAKIGKWIVQMVNELDMPFDT...........IHLIGQNVGAHVAGAAAQEFTRL   271
DmYP2   NIERLGEIIGNRLVELTNTVNVPQEI...........IHLIGSGPAAHVAGVAGRQFTRQ   272
DmYP3   DVLNTGAMIGQTLIDLTNK.GVPQEI...........IHLIGQGISAHVAGAAGNKYTAQ   261
MdYP1   NVKETGKMIGKTLAELEKESNVDLED...........LHVIGQGIGANVAGAAGKAFKDV   243
MdYP2   DTEKVGKEVGKFWLQLLEKTNCNRDN...........VHLIGSNLGANIAGAAGRQYTKV   273
MdYP3   DVKETGKMIGKTFAELMDECDVDVED...........MHVVAQGIATNVGGSAGKDFKDI   243
HsGL    DSVEFWAFS....FDEMAKYDLPATIDFIVKKTGQKQLHYVGHSQGTTIGFIAFSTNPSL   188
RnLL    DSVEFWAFS....FDEMAKYDLPATINFIVQKTGQEKLHYVGHSQGTTIGFIAFSTNPTL   187
DmL1    DESKFWDFS....WHEIGMYDLPAMIDHVLKVTGFPKLHYAGHSQGCTSFFVMCSMRPAF   207
DmL3    YWQIFWNFS....WNEIGMYDVPAMIDYVLAKTGQQQVQYVGHSQGTTVYLVMVSERPEY   180

          *+*   *  +* *     +        *    *         +         *    *     *
PiYP2   .NNKVAMMYALSPMAYMSHVRS.PLVKMIAPDSPFYNNLKQYLKDGEFKPSKELVYTMGG   457
BmESP   .NEKIVSMHALSPIVYMNYVRS.PLFRMIAPTSKFYQYIHDQVGHGAFEPGKHLIETFGG   402
GmYP2   .GDKIIKMYAMAPMVYMSH....SMVRMIAPPSDLHEQLSPYFMDGEFHPSKELLKTLGG   349
DmYP1   TGHKLRRITALDPTKIYGKPEE.RLT.GLARGDAEFVDAIHTSVYGMGTPIRSGDVDFYP   329
DmYP2   TGHKLRRITALDPTKIYGKPEE.RLT.GLARGDADFVDAIHTSAYGMGTSQRLANVDFFP   330
DmYP3   TGHKLRRITGLDPAKVLSKRPQ.ILG.GLSRGDADFVDAIHTSTFAMGTPIRCGDVDFYP   319
MdYP1   TTHKLGRITVLDPARQVGKDPK.VLT.GLSRCSVKFVDVIHTSALGMGTTRRVGDVDFFS   301
MdYP2   TNHQLRRITGLDPVKCFAKDPE.TLT.GLARGDAEFVDAIHTTANSMGTSARAADVDFYP   331
MdYP3   TTHKLDRITALDPARQVAKNPK.VLS.GLARGSANFVDAIHTSALGMGTTRRVGDVDFFP   301
HsGL    .AKRIKTFYALAPVATVKYTKS.LINKLRFVPQSLFKFI...FGDKIFYPHNFFDQFLAT   243
RnLL    .AKKIKTFYALAPVATVKYTQS.PLKKISFIPTFLFKLM...FGKKMFLPHTYFDDFLGT   242
DmL1    .NDKVVSMQALAPAVYAKETEDHPYIRAISLYFNSLVGS...SIREMF....NGEFRF..   256
DmL3    .NDKIKSAHLLGPAAYMGNMKS.PLTRAFAPILGQPNAIVEVCGSMEFMPSNKFKQDLGI   238

        +  +     *         *+                 +   *  +     +*++        ++
PiYP2   EMLENEIGCKKIASNVNFVMSGMNVDNMDVKSI.RVIMGHLPAGGSTRQVRQYGQAVATH   516
BmESP   AACREKLGCRHVCNNLNYVISGINVYNQDADIV.PVVMAHLPAGTSARVMKQYGQNVASH   461
GmYP2   EMCEKEIGCRKICSNLNFVMSGVNLEHMEPAQI.PMITGHLRWR.FDQTNKAVQPIVASR   407
DmYP1   N..GPAAGVP.GASNVVEAAMRATRY..FAESVRPGNERSFPAV.PANSLQQYKQNDG.F   382
DmYP2   N..GPSTGVP.GADNVVEATMRATRY..FAESVRPGNERNFPSV.AASSYQEYKQNKG.Y   383
DmYP3   N..GPSTGVP.GSENVIEAVARATRY..FAESVRPGSERNFPAV.PANSLKQYKEQDG.F   372
MdYP1   N..GTCQGVP.GSRNAIDAQARATRL..FGETVRPGNSRNFPAV.EASSLIQYRNNDG.Y   354
MdYP2   E..GPNEAVP.GADNIVESPMRAVRY..FAESVVPGNERNFPAE.SAHSLNEYKNSNT.S   385
MdYP3   H..GPCEGVP.GTRNVIEAQARATRF..YAESVRPGNSRNFPAL.EASSLKQYRNKDS.Y   354
HsGL    EVCSREMLNL.LCSNALFIICGFDSKN.FNTSRLDVYLSHNPAGTSVQNMFHWTQAVK.S   300
RnLL    EVCSREVLDL.LCSNTLFIFCGFDKKN.LNVSRFDVYLGHNPAGTSVQDFLHWAQLVR.S   299
DmL1    .LCRMTEETERLCIEAVFGIVGRNWNE.FNRKMFPVILGHYPAGVAAKQVKHFIQIIK.S   314
DmL3    EMCQATSPYADMCANEIFLIGGYDTEQ.LDYELLEHIKATSPAGASVNQNLHFCQEYN.S   296
```

Fig. 1. Continued.

clearly related evolutionarily to lipoprotein lipases (Bownes et al. 1988; Baker 1988; Terpstra and Ab 1988; Bownes 1992). Unlike the case in Lepidoptera, however, higher-dipteran YPs are the major component of the yolk and Vg is not synthesized at all. A retinoid- and fatty acid-binding glycoprotein in *D. melanogaster* is homologous to Vg (Kutty et al. 1996), but it appears to be a member of the lipophorin family and thus a paralog.

Shirk and Perera (1998) demonstrated that the YP2s of *P. interpunctella* and *G. mellonella* are homologous to the *B. mori* egg-specific protein (ESP) (Inagaki and Yamashita 1989). Sato and Yamashita (1991) recognized that ESP is related to vertebrate lipases, and Shirk and Perera (1998) demonstrated that the pyralid YP2s are as well. If extensive regions of lepidopteran ESP/YP2s and dipteran YPs are both related to the same vertebrate lipases, the principle of transitive homology (Pearson 1996) dictates they must be related to one another. Shirk and Perera (1998) recognized that a 10-amino acid region containing the lipid binding sites of vertebrate lipases is conserved in the insect sequences, but concluded from a larger multiple alignment that it was located about 45 amino acids further downstream in the dipteran YPs than in the lepidopteran ESP/YP2s. Their alignment shows no evidence for homology between these two types of yolk protein precursors.

C

```
          * *  *  *      + *+     +   * *    *          *
PiYP2     EFR.MYDYGSEVNQEVYGD.RVPPVYDVTKIR..TPVALYFSEHDWLAHPKDVLRLKEQL   574
BmESP     DFR.KYNYGAETNMKVYGA.SEPPSYDLSKVS..APVNLYHSHDAWLAHPKDVEKLQENL   519
GmYP2     EFR.MYDHGAKINKKMYGS.VQPPVYDVSKIQ..TPVVLYYSEEDWLSHPKDVERLHREL   465
DmYP1     GKR.AY.MGIDTAHDLEGD.YILQVNPKSPFGRNAPAQKQSSYHG.VHQAWNTNQDSKDY   438
DmYP2     GKR.GY.MGIATDFDLQGD.YILQVNSKSPFGRSTPAQKQTGYHQ.VHQPWRQSSSNQGS   439
DmYP3     GKR.AY.MGLQIDYDLRGD.YILEVNAKSPFGQRSPAHKQAAYHG.MHHAQN           420
MdYP1     GKR.TY.MGIATHRDISGD.YMLEVNAESPYGKRTPARKQKSYHG.FHQTSYA.KSNENY   409
MdYP2     GRR.IY.MGIITTFKVEGD.YMLQVNTKSPFGRSTPVQKQQNVHG.VHKSWKMSSSRDEE   440
MdYP3     GKR.AY.MGIATRRDTTGD.YILEVNEQTPFGKRSAPQ.QRSVQS.FN.......SENY   401
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
HsGL      GKFQAYDWGSPVQNRMHYDQSQPPYYNVTAMN..VPIAVWNGGKDLLADPQDVGLLLPKL   358
RnLL      GKFQAFNWGSPSQNMLHYNQKTPPEYDVSAMT..VPVAVWNGGNDILADPQDVAMLLPKL   357
DmL1      GRF.APYSYSSNKNMQLYRDHLPPRYNLSLVT..VPTFVYYSTNDLLCHPKDVESMCDDL   371
DmL3      GKF.RKFDYTALRNPYEYGSYFPPDYKLKNAK..APVLLYYGANDWMCDVSDVRKLRDEL   353


PiYP2     PNVTEYYQVP.EEYFSHMDFLYSQKAPVVVYKNLINSINNNIHKN                  616
BmESP     PNVKQSFEVPEQQHFTDLDFQFSKKAPDTVYQKLMENMQNNS                     559
GmYP2     PNVTEYYKVP.EGYFAHMDYQHYKKAPEMVYTRLIKSMNSSS                     504
DmYP1     Q                                                              439
DmYP2     RRQ                                                            442
DmYP3
MdYP1
MdYP2
MdYP3
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
HsGL      PNLIYHKEIP...FYNHLDFIWAMDAPQEVYNDIV.SMISEDKK                   398
RnLL      SNLLFHKEIL...AYNHLDFIWAMDAPQEVYNEMI.SMMAED                     395
DmL1      GNVTGKYLVP.QKEFNHMDFLWAIDVRKMLYRRMLEVLGKVPEGS(+18)             433
DmL3      PNMALDYLVP.FEKWAHLDFIWGTEARKYVYDEVLKQMQSYE                     394
```

Fig. 1. Continued.

Distant homologies can be detected with multiple sequence comparisons (Livingstone and Barton 1996; Henikoff et al. 1997), despite low overall conservation and low pairwise similarities, based on transitive homology (Pearson 1996), positional patterns of similar residues within the sequence (Patthy 1996), and by the types of residues conserved (Jornvall et al. 1987). In addition to the three *D. melanogaster* YPs, the three lepidopteran ESP/YP2s, and the two vertebrate lipases analyzed by Shirk and Perera (1998), two *D. melanogaster* lipases and three other dipteran YPs from the house fly (*Musca domestica*) were included in a multiple alignment constructed completely by hand (Fig. 1). When the sequences were aligned using the lipid-binding domain as an anchor, it became clear that dipteran YPs and lepidopteran ESP/YP2s are indeed homologous. 11.6% of the positions in the alignment (up to the last residue of *D. melanogaster* YP3, the shortest sequence) contain functionally conserved residues in all nine insect yolk protein sequences, and 20.2% of the positions contain similar residues in at least eight of nine sequences (Fig. 1). As is characteristic of diverged but homologous proteins, a disproportionate number (five of 10) of completely conserved residues are glycine or proline (Fig. 1), residues especially important in secondary structure (Jornvall et al. 1987). Disproportionate conservation of glycine, proline, and cysteine residues confirmed homology among insect, nematode, and vertebrate Vgs despite low overall conservation (Chen et al. 1997).

The sequences in the multiple alignment were used to reconstruct the phylogeny of the yolk proteins using both neighbor-joining and maximum-parsimony algorithms in PAUP* (Swofford 1998) (Fig. 2). In both cases, 1000 bootstrap replications were run and the 50% majority-rule consensus tree is presented. The vertebrate and *Drosophila* lipases were assigned as outgroups to root the trees. The topology of both trees is similar in that the dipteran YPs are located within a single clade, as are the lepidopteran ESP/YP2s. In addition, the two yolk protein clades clustered into a single larger clade with strong bootstrap support (83–88%) (Fig. 2), confirming that the two groups share a common ancestor not shared by lipases, a class of proteins to which both types of yolk proteins have long been known to be homologous (Bownes et al. 1988; Sato and Yamashita 1991; Bownes 1992; Shirk and Perera 1998).

Thus, it is apparent that the higher Diptera did not requisition lipases for use as their major yolk protein precursors, but instead employ a class of proteins with an evolutionary history of use as minor constituents of yolk in other insects. Experiments involving transplant of ovaries into male *B. mori* showed that viable eggs can be produced in the absence of Vg, using only ESP as the major yolk protein (Yamashita and Irie 1980). The lepidopteran ESP/YP2s are pro-

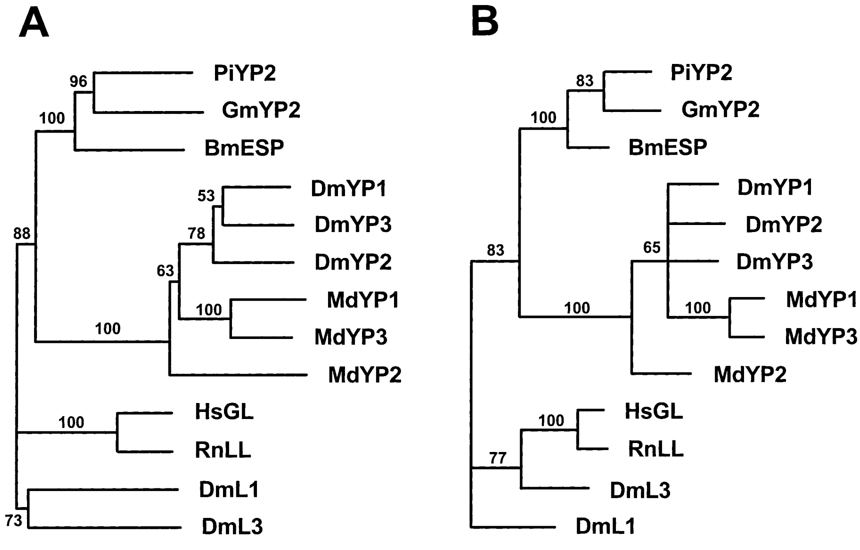

Fig. 2. Phylogenetic reconstructions of lepidopteran ESP/YP2s and dipteran Yps based on amino acid sequences. **(A)** Phylogram generated by the distance-based neighbor-joining method. Branch lengths are proportional to distance. (**B**) Phylogram generated by the maximum parsimony method. Branch lengths are proportional to the number of changes assigned to each branch. Both trees represent the 50% majority-rule consensus of 1000 bootstrap replications and were constructed using PAUP* (Swofford 1998). The positions and alignment used in the analyses are indicated in Fig. 1. Abbreviations and references are in the caption to Fig. 1.

duced exclusively in the ovary by the follicular epithelium (Shirk et al. 1984; Sato and Yamashita 1991; Shirk and Perera 1998), as are a few YPs of the higher Diptera (Houseman and Morrison 1986; Handler 1997). However, most higher dipteran YPs are produced by both the follicular epithelium and the fat body (Isaac and Bownes 1982; White and Bownes 1997). It is not clear which is the ancestral state.

It seems likely that proteins related to *Drosophila* and lepidopteran YPs eventually will be identified as minor constituents in the yolk of mosquitoes (or other lower Dipterans) where Vg is the major component. Why and how Vg was lost by the higher Diptera remains unknown. But now the evolutionary sequence of events seems less improbable, because only the abandonment of Vg need be explained, not both its abandonment and the simultaneous *de novo* recruitment of a lipase to replace it.

Acknowledgments. I thank Kate Aronstein, Patrick Moran, Erik Mirkov, Paul Shirk, and an anonymous colleague for helpful comments in their critical reviews of the manuscript.

References

Babin PJ, Bogerd J, Kooiman FP, Van Marrewijk WJA, Van der Horst DJ (1999) Apolipophorin II/I, apolipoprotein B, vitellogenin and microsomal triglyceride transfer protein genes are derived from a common ancestor. J Mol Evol 49:150–160

Baker ME (1988) Is vitellogenin an ancestor of apolipoprotein B-100 of human LDL and human lipoprotein lipase? Biochem J 255:1057–1060

Bodmer MW, Angal S, Yarranton G, Harris TJR, Lyons A, King DJ, Pieroni G, Riviere C, Verger R, Lowe PA (1987) Molecular cloning of a human gastric lipase and expression of the enzyme in yeast. Biochim Biophys Acta 909:237–244

Bownes M (1992) Why is there sequence similarity between insect yolk proteins and vertebrate lipases? J Lipid Res 33:777–790

Bownes M, Shirras A, Blair M, Collins J, Coulson A (1988) Evidence that insect embryogenesis is regulated by ecdysteroids released from yolk proteins. Proc Natl Acad Sci USA 85:1554–1557

Byrne BM, Gruber M, Ab G (1989) The evolution of egg yolk proteins. Prog Biophys Mol Biol 53:33–69

Chen J-S, Sappington TW, Raikhel AS (1997) Extensive sequence conservation among insect, nematode, and vertebrate vitellogenins reveals ancient common ancestry. J Mol Evol 44:440–451

Docherty AJ, Bodmer MW, Angal S, Verger R, Riviere C, Lowe PA, Lyons A, Emtage JS, Harris TJ (1985) Molecular cloning and nucleotide sequence of rat lingual lipase cDNA. Nucleic Acids Res 13:1891–1903

Hafer J, Fischer A, Ferenz H-J (1992) Identification of the yolk receptor protein in oocytes of *Nereis virens* (Annelida, Polychaeta) and comparison with the locust vitellogenin receptor. J Comp Physiol 162B:148–152

Handler AM (1997) Developmental regulation of *yolk protein* gene expression in *Anastrepha suspensa*. Arch Insect Biochem Physiol 36:25–35

Henikoff S, Greene EA, Pietokovski S, Bork P, Attwood TK, Hood L (1997) Gene families: the taxonomy of protein paralogs and chimeras. Science 278:609–614

Houseman JG, Morrison PE (1986) Absence of female-specific protein in the hemolymph of stable fly *Stomoxys calcitrans* (L.) (Diptera: Muscidae). Arch Insect Biochem Physiol 3:205–213

Inagaki S, Yamashita O (1989) Complete amino acid sequence of *Bombyx* egg-specific protein deduced from cDNA clone. Arch Insect Biochem Physiol 10:131–139

Isaac PG, Bownes M (1982) Ovarian and fat body vitellogenin synthesis in *Drosophila melanogaster*. Eur J Biochem 123:527–534

Jornvall H, Hempel J, von Bahr-Lindstrom H, Hoog JO, Vallee BL (1987) Alcohol and aldehyde dehydrogenases: structures of the human liver enzymes, functional properties and evolutionary aspects. Alcohol Suppl 1:13–23

Kutty RK, Kutty G, Kambadur R, Ducan T, Koonin EV, Rodriguez IR, Odenwald WF, Wiggert B (1996) Molecular characterization and developmental expression of a retinoid- and fatty acid-binding glycoprotein from *Drosophila*. J Biol Chem 271:20641–20649

Livingstone CD, Barton GJ (1996) Identification of functional residues and secondary structure from protein multiple sequence alignment. Meth Enzymol 266:497–512

Patthy L (1996) Consensus approaches in detection of distant homologies. Meth Enzymol 266:184–198

Pearson WR (1996) Effective protein sequence comparison. Meth Enzymol 266:227–258

Pistillo D, Manzi A, Tino A, Boyl PP, Graziani F, Malva C (1998) The *Drosophila melanogaster* lipase homologs: A gene family with tissue and developmental specific expression. J Mol Biol 276:877–885

Raikhel AS, Dhadialla TS (1992) Accumulation of yolk proteins in insect oocytes. Annu Rev Entomol 37:217–251

Sappington TW, Raikhel AS (1998) Molecular characteristics of insect vitellogenins and vitellogenin receptors. Insect Biochem Mol Biol 28:277–300

Sato Y, Yamashita O (1991) Structure and expression of a gene coding for egg-specific protein in the silkworm, *Bombyx mori*. Insect Biochem 21:495–505

Shirk PD, Perera OP (1998) 5′ coding region of the follicular epithelium yolk polypeptide 2 cDNA in the moth, *Plodia interpunctella*, contains an extended coding region. Arch Insect Biochem Physiol 39:98–108

Shirk PD, Bean D, Millemann AM, Brookes VJ (1984) Identification, synthesis, and characterization of the yolk polypeptides of *Plodia interpunctella*. J Exp Zool 232:87–98

Snigirevskaya ES, Hays AR, Raikhel AS (1997) Secretory and internalization pathways of mosquito yolk protein precursors. Cell Tiss Res 290:129–142

Spieth J, Nettleton M, Zucker-Aprison E, Lea K, Blumenthal T (1991) Vitellogenin motifs conserved in nematodes and vertebrates. J Mol Evol 32:429–438

Swofford DL (1998) PAUP*: Phylogenetic analysis using parsimony (*and other methods) version 4. Sinauer Associates, Sunderland, MA

Terpstra P, Ab G (1988) Homology of *Drosophila* yolk proteins and the triacylglycerol lipase family. J Mol Biol 202:663–665

White NM, Bownes M (1997) Cloning and characterization of three *Musca domestica* yolk protein genes. Insect Mol Biol 6:329–341

Yamashita O, Irie K (1980) Larval hatching from vitellogenin-deficient eggs developed in male hosts of the silkworm. Nature 283:385–386

Yan YL, Kunert CJ, Postlethwait JH (1987) Sequence homologies among the three yolk polypeptide (*Yp*) genes in *Drosophila melanogaster*. Nucleic Acids Res 15:67–85