ELSEVIER

Biochimica et Biophysica Acta 1397 (1998) 131–136

BIOCHIMICA ET BIOPHYSICA ACTA
BBA

Short sequence-paper

A novel DEAD-box RNA helicase exhibits high sequence conservation from yeast to humans

Arri Eisen, Martin Sattah, Tal Gazitt, Karama Neal, Paul Szauter, John Lucchesi *

Department of Biology, Emory University, Atlanta, GA 30322, USA

Received 8 September 1997; revised 7 November 1997; accepted 19 November 1997

Abstract

We have identified a novel *Drosophila* protein, DBP80, that exhibits significant similarity to mouse mDEAD5, yeast TIF1/2, and mammalian eIF-4A. DBP80 is a member of a subclass of DEAD-box proteins that contains a distinct domain, PX(I/R)ILLKR(E/D)EETLEGIKQ(F/Y)(F/Y), in addition to the seven canonical helicase domains. © 1998 Elsevier Science B.V.

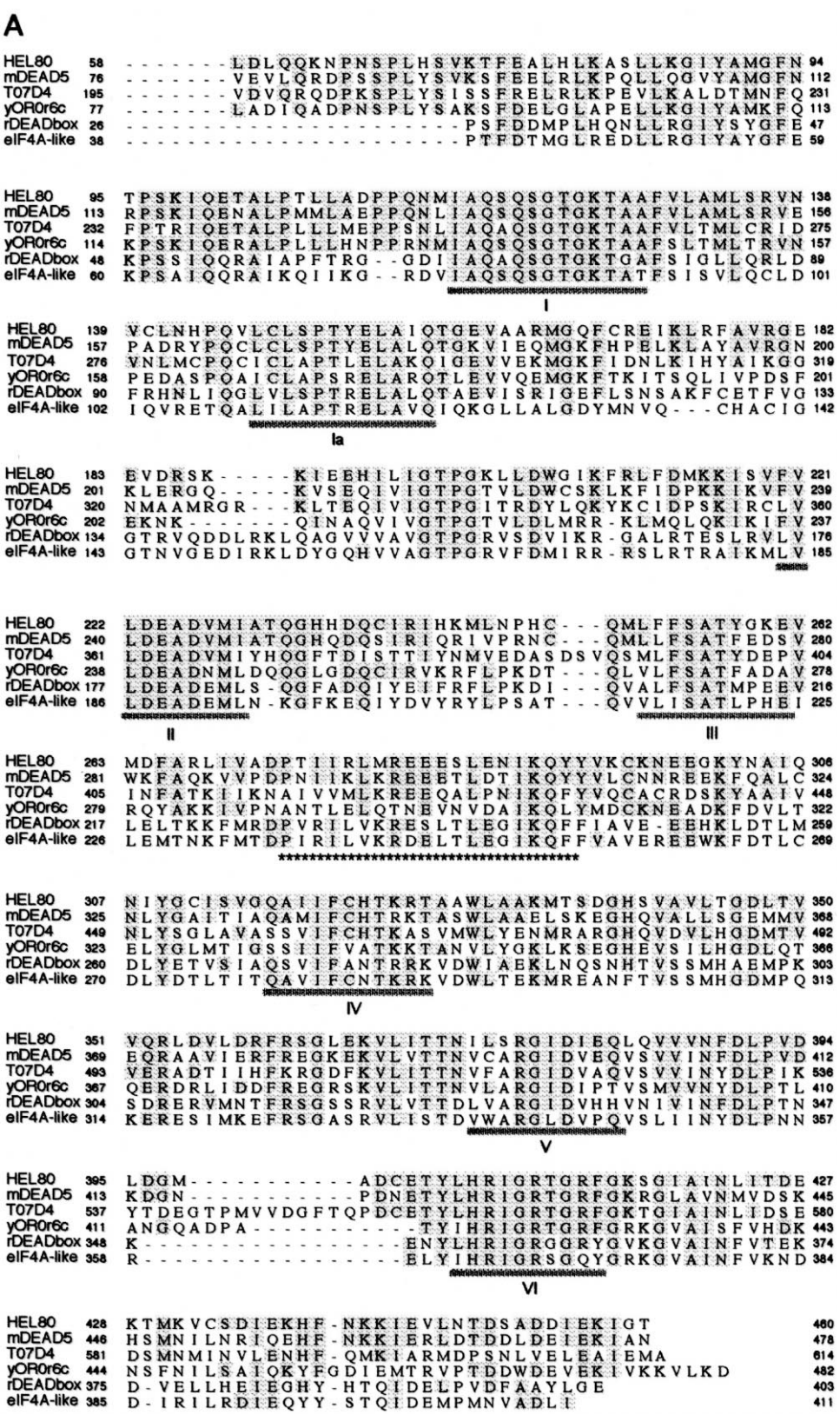
Keywords: Helicase; Translation initiation; Gene evolution; (*Drosophila*)

Helicases, which may be encoded by as much as 1% of the total eukaryotic genome, are defined by the presence of seven protein motifs (underlined in Fig. 1) that have been conserved throughout evolution [1]. The function of these motifs has been deduced primarily from mutational analyses of eIF-4A, the canonical DEAD-box helicase, so named because of the specific amino acid sequence of the third motif. Motif I (AXXGXGKT) is box A of the ATP-binding domain [2,3]; motif II (VLDEADRMLD), the ATP-binding B-box, is required for ATP hydrolysis and couples that activity to RNA unwinding; motif III (SAT) is necessary for helicase activity, and mutations in motif VI (HRIGRXXR) can affect RNA binding, RNA helicase activity, and ATP binding and hydrolysis [4]. DEAD-box helicases from the entire evolutionary spectrum have been implicated in nearly all stages of RNA metabolism, including pre-mRNA splicing [5], ribosome assembly [6,7], mRNA stability [8] and degradation [9], and translation initiation [10]. We report here the molecular characterization and cytological mapping of the *Drosophila* gene *Dbp80*. The protein that it encodes is localized in the cytoplasm and exhibits a striking level of similarity to a subclass of eukaryotic DEAD-box helicases some of which are known to be involved in translation initiation.

The *Drosophila* dosage compensation protein, MLE, a member of the DEAD-box RNA helicase family involved in dosage compensation [11], was used as a 'bait' in the yeast two-hybrid system [12,13] to screen for interactors from a *Drosophila*

* Corresponding author. Department of Biology, O.W. Rollins Research Center, 1510 Clifton Road, Atlanta, GA 30322, USA. Fax: +1-404-727-2880; E-mail: lucchesi@biology.emory.edu

PII S0167-4781(97)00213-3


A
HEL80 58 - - - - - - - LDLQQKNPNSPLHSVKTFEALHLKASLLKGIYAMGFN 94
mDEAD5 76 - - - - - - - VEVLQRDPSSPLYSVKSFEELRLKPQLLQGVYAMGFN 112
T07D4 195 - - - - - - - VDVQRQDPKSPLYSISSFRELRLKPEVLKALDTMNFQ 231
yOR0r6c 77 - - - - - - - LADIQADPNSPLYSAKSFDELGLAPELLKGIYAMKFQ 113
rDEADbox 26 - PSFDDMPLHQNLLRGIYSYGFE 47
eIF4A-like 38 - PTFDTMGLREDLLRGIYAYGFE 59
HEL80 95 TPSKIQETALPTLLADPPQNMIAQSQSGTGKTAAFVLAMLSRVN 138
mDEAD5 113 RPSKIQENALPMMLAEPPQNLIAQSQSGTGKTAAFVLAMLSRVE 156
T07D4 232 FPTRIQETALPLLLMEPPSNLIAQAQSGTGKTAAFVLTMLCRID 275
yOR0r6c 114 KPSKIQERALPLLLHNPPRNMIAQSQSGTGKTAAFSLTMLTRVN 157
rDEADbox 48 KPSSIQQRAIAPFTRG - - GDIIAQAQSGTGKTGAFSIGLLQRLD 89
eIF4A-like 60 KPSAIQQRAIKQIIKG - - RDVIAQSQSGTGKTATFSISVLQCLD 101
I
HEL80 139 VCLNHPQVLCLSPTYELAIQTGEVAARMGQFCREIKLRFAVRGE 182
mDEAD5 157 PADRYPQCLCLSPTYELALQTGKVIEQMGKFHPELKLAYAVRGN 200
T07D4 276 VNLMCPQCICLAPTLELAKQIGEVVEKMGKFIDNLKIHYAIKGG 319
yOR0r6c 158 PEDASPQAICLAPSRELARQTLEVVQEMGKFTKITSQLIVPDSF 201
rDEADbox 90 FRHNLIQGLVLSPTRELALQTAEVISRIGEFLSNSAKFCETFVG 133
eIF4A-like 102 IQVRETQALILAPTRELAVQIQKGLLALGDYMNVQ - - - CHACIG 142
Ia
HEL80 183 EVDRSK - - - - - KIEEHILIGTPGKLLDWGIKFRLFDMKKISVFV 221
mDEAD5 201 KLERGQ - - - - - KVSEQIVIGTPGTVLDWCSKLKFIDPKKIKVFV 239
T07D4 320 NMAAMRGR - - - KLTEQIVIGTPGITRDYLQKYKCIDPSKIRCLV 360
yOR0r6c 202 EKNK - - - - - - - QINAQVIVGTPGTVLDLMRR - KLMQLQKIKIFV 237
rDEADbox 134 GTRVQDDLRKLQAGVVVAVGTPGRVSDVIKR - GALRTESLRVLV 176
eIF4A-like 143 GTNVGEDIRKLDYGQHVVAGTPGRVFDMIRR - RSLRTRAIKMLV 185
HEL80 222 LDEADVMIATQGHHDQCIRIHKMLNPHC - - - QMLFFSATYGKEV 262
mDEAD5 240 LDEADVMIATQGHQDQSIRIQRIVPRNC - - - QMLLFSATFEDSV 280
T07D4 361 LDEADVMIYHQGFTDISTTIYNMVEDASDSVQSMLFSATYDEPV 404
yOR0r6c 238 LDEADNMLDQQGLGDQCIRVKRFLPKDT - - - QLVLFSATFADAV 278
rDEADbox 177 LDEADEMLS - QGFADQIYEIFRFLPKDI - - - QVALFSATMPEEV 216
eIF4A-like 186 LDEADEMLN - KGFKEQIYDVYRYLPSAT - - - QVVLISATLPHEI 225
II
III
HEL80 263 MDFARLIVADPTIIRLMREEESLENIKQYYVKCKNEEGKYNAIQ 306
mDEAD5 281 WKFAQKVVPDPNIIKLKREEETLDTIKQYYVLCNNRREKFQALC 324
T07D4 405 INFATKIIKNAIVVMLKREEQALPNIKQFYVQCACRDSKYAAIV 448
yOR0r6c 279 RQYAKKIVPNANTLELQTNEVNVDAIKQLYMDCKNEADKFDVLT 322
rDEADbox 217 LELTKKFMRDPVRILVKRESLTLEGIKQFFIAVE - EEHKLDTLM 259
eIF4A-like 226 LEMTNKFMTDPIRILVKRDELTLEGIKQFFVAVEREEWKFDTLC 269
****
HEL80 307 NIYGCISVGQAIIFCHTKRTAAWLAAKMTSDGHSVAVLTGDLTV 350
mDEAD5 325 NLYGAITIAQAMIFCHTRKTASWLAAELSKEGHQVALLSGEMMV 368
T07D4 449 NLYSGLAVASSVIFCHTKASVMWLYENMRARGHQVDVLHGDMTV 492
yOR0r6c 323 ELYGLMTIGSSIIFVATKKTANVLYGKLKSEGHEVSILHGDLQT 366
rDEADbox 260 DLYETVSIAQSVIFANTRRKVDWIAEKLNQSNHTVSSMHAEMPK 303
eIF4A-like 270 DLYDTLTITQAVIFCNTKRKVDWLTEKMREANFTVSSMHGDMPQ 313
IV
HEL80 351 VQRLDVLDRFRSGLEKVLITTNILSRGIDIEQLQVVVNFDLPVD 394
mDEAD5 369 EQRAAVIERFREGKEKVLVTTNVCARGIDVEQVSVVINFDLPVD 412
T07D4 493 VERADTIIHFKRGDFKVLITTNVFARGIDVAQVSVVINYDLPIK 536
yOR0r6c 367 QERDRLIDDFREGRSKVLITTNVLARGIDIPTVSMVVNYDLPTL 410
rDEADbox 304 SDRERVMNTFRSGSSRVLVTTDLVARGIDVHHVNIVINFDLPTN 347
eIF4A-like 314 KERESIMKEFRSGASRVLISTDVWARGLDVPQVSLIINYDLPNN 357
V
HEL80 395 LDGM - - - - - - - - - - ADCETYLHRIGRTGRFGKSGIAINLITDE 427
mDEAD5 413 KDGN - - - - - - - - - - PDNETYLHRIGRTGRFGKRGLAVNMVDSK 445
T07D4 537 YTDEGTPMVVDGFTQPDCETYLHRIGRTGRFGKTGIAINLIDSE 580
yOR0r6c 411 ANGQADPA - - - - - - - - - - TYIHRIGRTGRFGRKGVAISFVHDK 443
rDEADbox 348 K - - - - - - - - - - - - - - - - ENYLHRIGRGGRYGVKGVAINFVTEK 374
eIF4A-like 358 R - - - - - - - - - - - - - - - - ELYIHRIGRSGQYGRKGVAINFVKND 384
VI
HEL80 428 KTMKVCSDIEKHF - NKKIEVLNTDSADDIEKIGT 460
mDEAD5 446 HSMNILNRIQEHF - NKKIERLDTDDLDEIEKIAN 478
T07D4 581 DSMNMINVLENHF - QMKIARMDPSNLVELEAIEMA 614
yOR0r6c 444 NSFNILSAIQKYFGDIEMTRVPTDDWDEVEKIVKKVLKD 482
rDEADbox 375 D - VELLHEIEGHY - HTQIDELPVDFAAYLGE 403
eIF4A-like 385 D - IRILRDIEQYY - STQIDEMPMNVADLI 411

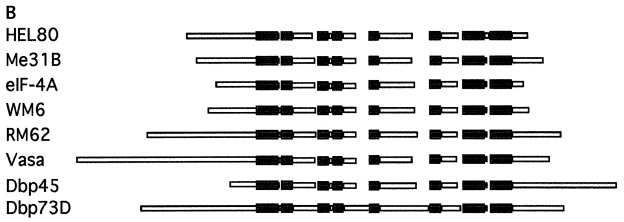


Fig. 1. Sequence comparison of DBP80 to related proteins. (A) Comparison of DBP80 to its most similar proteins across species. Identical residues are indicated with shading, and the seven conserved helicase motifs are underlined; the sequence specific to the subclass of DEAD-box helicases is indicated by asterisks: mouse mDEAD5 [5]; *C. elegans* T07D4 (GB accession no. 899474); *S. cerevisiae* yOR0r6c [20]; *L. braziliensis* rDEADbox [28]; human eIF4A-like (GB accession no. P38919). (B) Schematic representation of the known *Drosophila* DEAD-box helicases with conserved domains indicated by black boxes: Me31B [26]; eIF-4A [29]; WM6 [30]; RM62 [31]; Vasa [32,33]; Dbp45A [34]; Dbp73D [35].

imaginal disc cDNA library. A specific interactor, encoded by a 1.5 kb cDNA, was identified. Sequence analysis of this cDNA revealed an open reading frame of 1386 bp, beginning with the first ATG in the sequence. The deduced protein, designated DBP80, is 462 aa in length with a calculated M_r of 51441.

The full-length cDNA clone was used as a probe for in situ hybridization to larval salivary gland polytene chromosomes. *Dbp80* is located in the extreme proximal region of the left arm of chromosome 3, i.e., in centromeric heterochromatin (Fig. 2). Although several lethal mutations have been mapped to this region [14], the relative density of known genes is several orders of magnitude smaller in this region and in other regions of pericentric heterochromatin than in euchromatic regions of comparable length [15].

Northern analysis was performed with poly(A)$^+$ RNA derived, independently, from male and female adults; using the full-length *Dbp80* cDNA as a probe, a single 2.0 kb message was detected (Fig. 3).

The *Dbp80* gene was subcloned into pMT-HA, a vector specifically modified from its precursor pMK33/pMtHy [16] that expresses proteins under the control of the *Drosophila* metallothionein promoter and allows fusion of the protein to an HA-tag at the N- or C-terminus (A. Pannuti and W. Gu, in preparation). *Drosophila* S2 cells were transfected using Ca_2PO_4 [17] and transformants were selected for hygromycin resistance. Transformed cells were treated with Cu_2SO_4 and, 60 h later, were fixed, stained with Hoechst's to identify the nuclei and with anti-HA antibody [17], to identify the fusion protein. The results indicate that DBP80/HA is localized predominantly in the cytoplasm (not shown), strongly suggesting that it is not involved in dosage compensation, a transcriptional regulatory mechanism. The interaction of DBP80 with MLE in the yeast two-hybrid system could have occurred through a spurious 'three-hybrid' complex interaction involving RNA [18].

The deduced DBP80 amino acid sequence was derived from a cDNA clone (GenBank accession number AF005239) and was used in a BLAST search of the GenBank data base; sequences were aligned using the MACAW program. DBP80 exhibits significant similarity to proteins in the DEAD-box family of RNA helicases (Fig. 1). It is most similar to a group of DEAD-box helicases that includes those required for the ATP-dependent binding of mRNA to ribosomes to initiate translation [5,19]. This group of helicases also includes mDEAD5, a mouse putative helicase isolated using PCR with primers directed to conserved sequences of DEAD-box helicases [5]. DBP80 and mDEAD5 are comparable in length (460 aa and 472 aa, respectively) and are approximately 55% identical and 66% similar, when conservative substitutions are allowed. The only significant difference between the two proteins occurs at the N-terminus. Based on sequence comparisons, the proteins to which DBP80 and mDEAD5 are most similar have been subdivided into two classes [5]. While the sequence and spacing of the seven helicase motifs is highly conserved among proteins in both of these classes, the class to which DBP80 belongs includes proteins that tend to be smaller and that contain a conserved sequence, PX(I/R)ILLKR(E/D)EETLEGIKQ(F/Y)(F/Y), downstream from the SAT domain (Fig. 1).

DBP80 is the eighth member of the DEAD-box protein family to be discovered in *Drosophila* (Fig. 1). Numerous DEAD-box proteins also occur in other eukaryotes: at least five in mice [5], more than twenty in yeast [20], and at least eight in tobacco [21]. Several members of the DEAD-box protein family

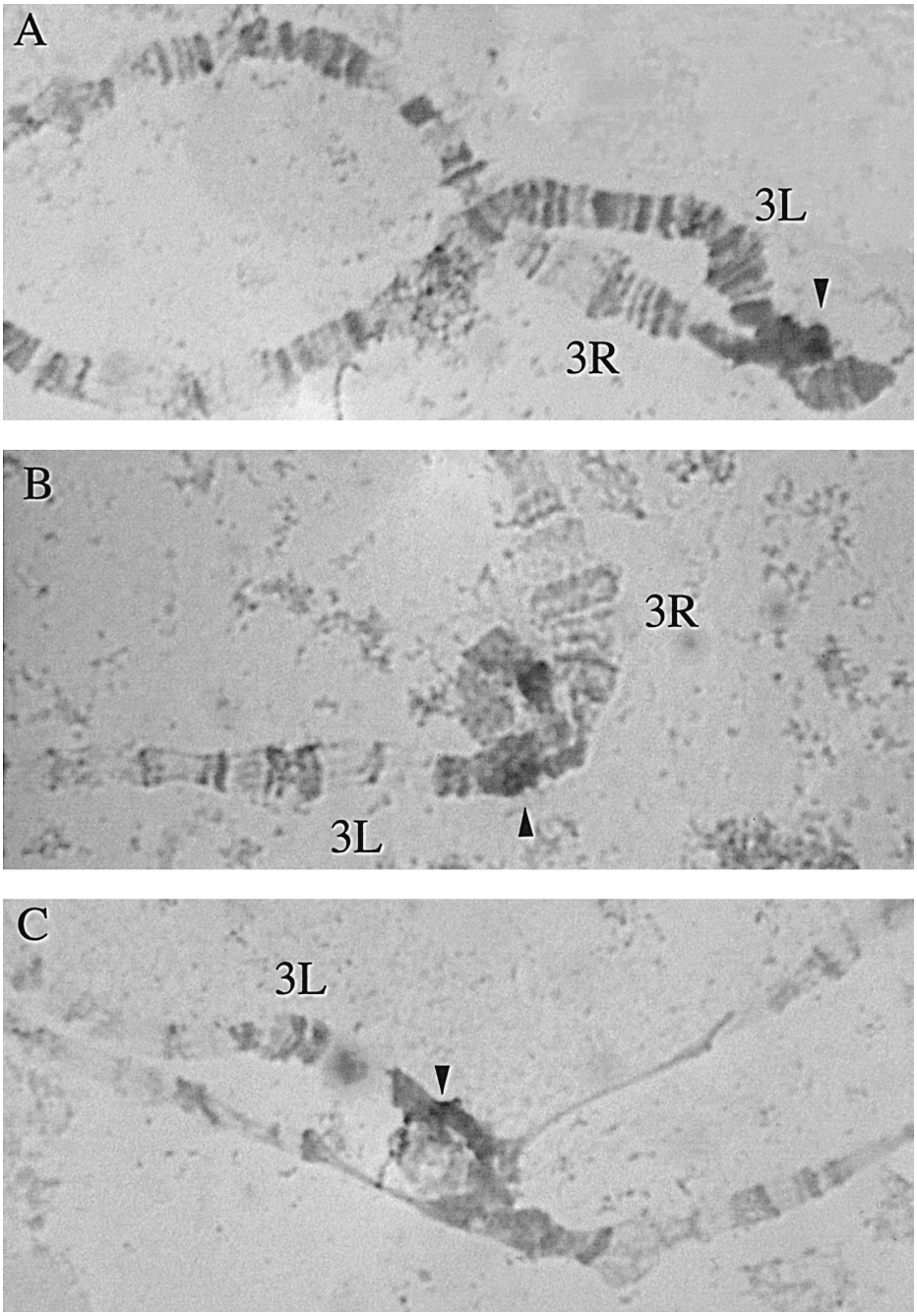


Fig. 2. Localization of *Dbp80* on polytene chromosomes. Polytene chromosomes were isolated from a wild-type Samarkand strain and prepared for in situ hybridization by the method of Pardue [36]. A *Dbp80* fragment was labeled with digoxygenin [37] and used in the hybridization procedure. The arrows indicate the region of hybridization.

have been shown to be functional translation initiation factors [10,22]. These considerations suggest that helicases may be used for the developmental or tissue-specific regulation of gene function at the level of translation. In fact, several eIF-4A-like proteins have a restricted distribution that may reflect such

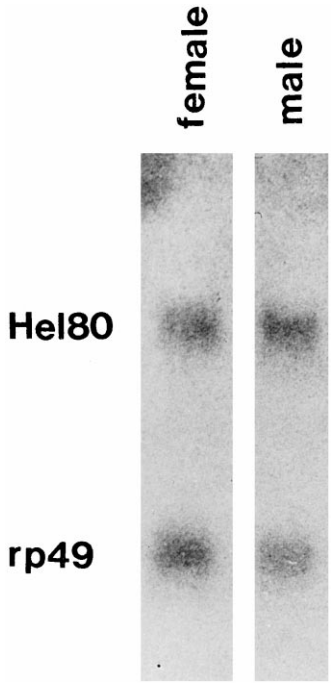


Fig. 3. Northern analysis of poly(A) RNA from adult Samarkand males and females using a *Dbp80* cDNA probe. The upper 2.0 kb band corresponds to the *Dbp80* transcript, and the lower band to the *ribosomal protein-49* transcript, which was used to normalize the amount of RNA loaded. Full-length probes were labeled with ^{32}P using random priming.

specificity: NeIF-4A8 is found only in the male gametophytes of *Nicotiana tabacum* [23]; An3 is localized to the animal pole in the developing *Xenopus* oocyte [24]; *ste13* is required only for nitrogen starvation-induced G1 arrest and initiation of sexual development in *S. pombe* [25]; the *Drosophila* homologue of the *ste13* gene, *Me31B*, is preferentially expressed in the female germline and may be required for oogenesis [26], while *Drosophila Vasa* is localized first to the nurse cells and then to the pole plasm during oogenesis [27].

References

[1] A.E. Gorbalenya, E.V. Koonin, Russian Academy of Medical Sciences and National Institutes of Health, 1993, pp. 419–429.
[2] A. Pause, N. Sonenberg, EMBO J. 11 (1992) 2643–2654.
[3] J.E. Walker, M. Saraste, M.J. Runswick, N.J. Gay, EMBO J. 1 (1982) 945–951.
[4] A. Pause, N. Methot, N. Sonenberg, Mol. Cell. Biol. 13 (1993) 6789–6798.
[5] S.L. Gee, J.G. Conboy, Gene 140 (1994) 171–177.
[6] T.L. Ripmaster, G.P. Vaughn, J.L. Woolford, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 11131–11135.
[7] A.B. Sachs, R.W. Davis, Science 247 (1990) 1077–1079.
[8] I. Iost, M. Dreyfus, Nature 372 (1994) 193–196.
[9] B. Py, C.F. Higgins, H.M. Krisch, A.J. Carpousis, Nature 381 (1996) 169–172.
[10] S.R. Schmid, P. Linder, Mol. Microbiol. 6 (1992) 283–292.
[11] L. Richter, J.R. Bone, M.I. Kuroda, Genes Cells 1 (1996) 325–336.
[12] T. Sato, M. Hanada, S. Bodrug, S. Irie, N. Iwama, L.H. Boise, C.B. Thompson, E. Golemis, L. Fong, H.G. Wang, J.C. Reed, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 9238–9242.
[13] J. Gyuris, E. Golemis, H. Chertkov, R. Brent, Cell 75 (1993) 791–803.
[14] G.E. Marchant, D.G. Holm, Genetics 120 (1988) 519–532.
[15] K.S. Weiler, B.T. Wakimoto, Annu. Rev. Genet. 29 (1995) 577–695.
[16] M.R. Koelle, W.S. Talbot, W.A. Segraves, M.T. Bender, P. Cherbas, D.S. Hogness, Cell 67 (1991) 59–77.
[17] F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl, Current Protocols in Molecular Biology, 1994–1997.
[18] D.J. SenGupta, B. Zhang, B. Kraemer, P. Pochart, S. Fields, M. Wickens, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 8496–8501.
[19] S. Blum, S.R. Schmid, A. Pause, P. Buser, P. Linder, N. Sonenberg, H. Trachsel, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 7664–7668.
[20] Chang Tien-Hsien, J. Arenas, H. Abelson, Proc. Natl. Acad. Sci. U.S.A. 87 (1990) 1571–1575.
[21] K.A. Brander, T. Mandel, G.W. Owttrim, C. Kuhlemeier, Biochim. Biophys. Acta 1261 (1995) 442–444.
[22] R.Y. Chuang, P.L. Weaver, Z. Liu, T.H. Chang, Science 275 (1997) 1468–1471.
[23] K.A. Brander, C. Kuhlemeier, Plant Mol. Biol. 4 (1995) 637–649.
[24] R. Gururajan, D.L. Weeks, Biochim. Biophys. Acta 1350 (1997) 169–182.
[25] H. Maekawa, T. Nakagawa, Y. Uno, K. Kitamura, C. Shimoda, Mol. Gen. Genet. 244 (1994) 456–464.
[26] T.D. Valoir, M.A. Tucker, E.J. Belikoff, L.A. Camp, C. Bolduc, K. Beckingham, Proc. Natl. Acad. Sci. U.S.A. 88 (1991) 2113–2117.
[27] L. Liang, W. Diehl-Jones, P. Lasko, Development 120 (1994) 1201–1211.
[28] Y.A. Skeiky, J.A. Guderian, D.R. Benson, O. Bacelar, E.M. Carvalho, M.M. Kubin, R. Badaro, G. Trinchieri, S.G. Reed, J. Exp. Med. 181 (1995) 1527–1537.
[29] R. Dorn, H. Morawietz, G. Reuter, H. Saumweber, Mol. Gen. Genet. 237 (1993) 233–240.

[30] E. Warbrick, D. Glover, Mol. Gen. Genet. 245 (1994) 654–657.
[31] D.R. Dorer, A.C. Christensen, D.H. Johnson, Nucleic Acids Res. 18 (1990) 5489–5494.
[32] P.F. Lasko, M. Ashburner, Nature 335 (1988) 611–617.
[33] B. Hay, L.Y. Jan, Y.N. Jan, Cell 55 (1988) 577–587.
[34] C.A. Lavoie, M. Harvey, P.F. Lasko, Biochim. Biophys. Acta 1216 (1993) 140–144.
[35] L.F. Patterson, M. Harvey, P.F. Lasko, Nucleic Acids Res. 20 (1992) 3063–3067.
[36] M. Pardue, *Drosophila melanogaster*: Practical Uses in Cell and Molecular Biology, Academic Press, 1994.
[37] E.R. Schmidt, J. Gutenberg, A simplified and efficient protocol for nonradioactive in situ hybridization to polytene chromosomes with a DIG-labeled probe 1-97-99, Boehringer Mannheim, 1996.