The embryonic RNA helicase gene (*ERH*): a new member of the DEAD box family of RNA helicases

Jane SOWDEN,* Wendy PUTT,* Katie MORRISON,* Rosa BEDDINGTON† and Yvonne EDWARDS*‡

*MRC Human Biochemical Genetics Unit, The Galton Laboratory (UCL), Wolfson House, 4 Stephenson Way, London, NW1 2HE, and †National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K.

DEAD box proteins share several highly conserved motifs including the characteristic Asp-Glu-Ala-Asp (D-E-A-D in the amino acid single-letter code) motif and have established or putative ATP-dependent RNA helicase activity. These proteins are implicated in a range of cellular processes that involve regulation of RNA function, including translation initiation, RNA splicing and ribosome assembly. Here we describe the isolation and characterization of an embryonic RNA helicase gene, *ERH*, which maps to mouse chromosome 1 and encodes a new member of the DEAD box family of proteins. The predicted ERH protein shows high sequence similarity to the testes-specific

mouse PL10 and to the maternally acting *Xenopus* An3 helicase proteins. The ERH expression profile is similar to that of An3, which localizes to the animal hemisphere of oocytes and is abundantly expressed in the embryo. ERH is expressed in oocytes and is a ubiquitous mRNA in the 9 days-post-conception embryo, and at later stages of development shows a more restricted pattern of expression in brain and kidney. The similarities in sequence and in expression profile suggest that *ERH* is the murine equivalent of the *Xenopus* An3 gene, and we propose that ERH plays a role in translational activation of mRNA in the oocyte and early embryo.

INTRODUCTION

RNA helicases are ATP-binding proteins which unwind helical secondary structures in RNA molecules [1]. The DEAD box family of RNA helicases are defined by their sequence similarity to eIF-4A, a eukaryotic translation initiation factor which shows single-stranded RNA-dependent ATPase activity *in vitro*, and in combination with a second initiation factor, eIF-4B, has RNA helicase activity [2–4]. eIF-4A functions as part of a protein complex with other translation initiation factors, including eIF-4B. It is thought that eIF-4A is directed to the cap structure at the 5' end of mRNA by the cap-binding factor (eIF-4E), where it unwinds mRNA 5' secondary structure to generate a single-stranded RNA region that facilitates ribosome binding [5].

All members of the DEAD box family of proteins have nine highly conserved amino acid regions in common with eIF-4A [2,6] and, in particular, they share a D-E-A-D box (Asp-Glu-Ala-Asp) which represents a special version of the B motif of ATP-binding proteins [7]. The family is increasing in number and has members in species ranging from *Escherichia coli* to humans [1,8]. Proteins with related D-E-A-H and D-E-X-H motifs have also been identified, and they include putative DNA and RNA helicases. Members of all three families form the helicase superfamily II (70 members) which share seven conserved motifs [9].

In addition to eIF-4A, three other DEAD box helicases, vasa [10], p68 [11] and An3 [12], have demonstrated ATP-hydrolysis-dependent RNA helicase activity *in vitro*, and it is notable that their helicase activity does not require the co-operation of other proteins. These helicases are larger than eIF-4A and have additional N-terminal and C-terminal sequences, which may provide additional specificity to the RNA helicase or even additional protein function. DEAD box RNA helicases are

implicated in a range of cellular processes in addition to translation initiation; these include RNA splicing, ribosome assembly, oogenesis, spermatogenesis and cell growth and division [1,8]. Recently the role of two *E. coli* DEAD box helicases, deadD and SrmB, in stabilizing mRNA has been demonstrated [13].

In the present paper we describe the cloning and characterization of a novel mouse DEAD box RNA helicase which we have designated the embryonic RNA helicase (ERH). The predicted ERH protein is most closely related to the PL10-like subgroup of helicases [1], which include mouse PL10 (gene symbol *DIPas1* [14]), *Xenopus* An3 [12], yeast DED1 and DBP1 [15,16], *Drosophila* vasa [17] and *C. elegans* glh-1 [18]. Several of these helicases show distinctive temporal and spatial patterns of expression, implying they have specific roles in development and in tissue-specific regulation. Here we describe the pattern of expression of ERH during development and its chromosome assignment, and we compare the ERH sequence with other DEAD box proteins. These findings suggest that *ERH* is the mouse equivalent of the *Xenopus* An3 gene, and we discuss its possible role in translational activation.

MATERIALS AND METHODS

Characterization of cDNAs: library screening and DNA sequencing

Embryos for all procedures were obtained from matings of C57BL/6 × DBA mice. cDNA probes for differential screening of an early mouse embryo notochord cDNA library (J. Sowden, K. Morrison, W. Putt, R. Beddington and Y. Edwards, unpublished work) in λExcell (Pharmacia) were prepared as follows. mRNA (4 µg) from adult mouse liver and from newborn-mouse

Abbreviations used: *ERH*, embryonic RNA helicase gene; dpc, days post conception; RT-, reverse transcription; RFLV, restriction-fragment-length variation; CPE, cytoplasmic polyadenylation element.

‡ To whom correspondence should be addressed.

The complete sequence of the longest cDNA fragment, 3185 bp, of *ERH* has been deposited with the EMBL Nucleotide Sequence Database under the accession number Z38117.

```

TGAgcctgctctgtagtaggtcacccctgccaaacaagctaatatggaaaccacatgtaac 2121
ttagccagactataccttgtgtagcttcaagaactcgcagtagcattaccagctgtgattc 2181
tccactgaaatTTTTTTTaaaggagctcaaggtcacaaagtaagatgaaaggaacaatc 2241
agcagccctgtctgaaggtggtttgaagacttcattgctgtagtttggattaacgccctc 2301
cccgcctacccttatccaaactgcatttataattttgtgactgaggatcatttgtttgt 2361
taatgtactgtgcctttaactttagacaactTTTTTatTTTgatgtcctgttggctcagta 2421
atgctcaagattccaattgtttttgacaaaataaatttactgaacttgggctaaaatcaa 2481
accttggcacacaggtgtgatacaacttaacaggaatcatcgattcatccataaataata 2541
taaggaaaaactcaagtggtagcctgtcttaaggcttttgatacttgagattggggaaa 2601
aacaacaacaacacgtcttgaagcatattaatggaattagtttctaagtggcaaacgtg 2661
attaagttaaaagttctgatttgcctactctatcctgtataggtatttagaacctgataa 2721
tagtctttaaacaagccattccagtcagtgatgaggtgatgtatggatacatgcatacatt 2781
caaagcactgttctcaaagttaatgcaagtaaatcacagcaattcctctttcaatgtttag 2841
gcagatcggttaactatgagctagccaaatgtgggcatgttattacagggaaagttaaag 2901
gtctgataacttgaaataggttttaggagaattcatctacttagactTTTTTaaatgcctgc 2961
cataaaaaattgaaatggtagaatggctgaccacagcaatgaccagccctcacctagggc 3021
tctggatgatttttgggtctaataacgcagtgctagtgttgatgttttttgggtcaagatggg 3081
aatgaacaggaagaattatgcagcaggcttataTTTTTaaatgcccattcacattactctgt 3141
tcaagctgcgttgagatgttaaactggcttactatagactttgt (a)n 3185

```

Figure 1 3' untranslated sequence of ERH mRNA

The TGA stop codon is shown in upper-case letters. The polyadenylation signal gttaaa is boldly underlined. The CPE-like sequences are underlined. Positions of forward and reverse primers used in PCR are indicated by arrows below the sequence. *EcoRI* and *SacI* sites delineate the fragment used as a hybridization probe in RNA *in situ* hybridization experiments. The sequence is numbered from the TGA stop codon starting at position 2062.

brain and limb was incubated in 50 μ l of first-strand cDNA synthesis buffer (Amersham), together with 1 mM dGTP, 1 mM dTTP, 25 μ M dCTP, 25 μ M dATP, 100 μ Ci of [32 P]dCTP (3000 Ci/mmol), 100 mCi of [32 P]dATP (3000 Ci/mmol), oligo(dT) (2.5 μ l supplied with the Amersham kit) and 50 units of reverse transcriptase, for 75 min at 42 °C. Duplicate filters of 2×10^8 recombinant phages were plated and screened according to standard procedures [19]. Filters were finally washed at high stringency in $0.1 \times$ SSC ($1 \times$ SSC is 0.15 M NaCl/0.015 M sodium citrate)/0.1% SDS for 30 min at 65 °C. A total of 50 phage clones which did not hybridize to the adult cDNA were excised *in vivo* as pExcell plasmid and cultured in microtitre plates. Nucleotide sequences were determined by the dideoxy chain-termination method [20] using single-stranded DNA as template. Nucleotide- and amino- acid-sequence analyses were performed

using programs from the University of Wisconsin Genetics Computer Group, including pileup and bestfit, and the BASIC local alignment search tool (BLAST) to search Genbank, EMBL and SwissProt databases [21].

PCR and Northern-blot analysis

Reverse transcription (RT-) PCR of oocyte and 9.5- days-post-coitus (dpc) embryo RNA was carried out using sequence-specific primers for *ERH*, the Brachyury gene *T* and the phosphoglucosyltransferase gene *Pgm-2*. Total RNA prepared from mouse oocytes and from 9.5 dpc mouse embryos using the acid guanidinium thiocyanate/phenol/chloroform technique [22] was reverse-transcribed in 35 μ l of a reaction mix containing 1 nmol of random hexamers, 400 units murine-Moloney-leukaemia-

Figure 3 ERH is expressed ubiquitously in the early 9 dpc mouse embryo and shows a more restricted pattern of expression later in development

Dark-field micrographs of sections through mouse embryos hybridized with antisense (panels A, D, E and F) and sense (panels B and C) ERH riboprobe. Panels A and B, transverse sections through a 9 dpc mouse embryo at the level of the heart. Panels C and D, transverse sections through a 14 dpc embryo in the region of the developing metanephros. Panel E, a frontal section through a 13.5 dpc embryo; and panel F, a sagittal section through a 15.5 dpc embryo. Abbreviations: nt, neural tube; k, kidney; br, brain; t, tongue; li, liver; tp, tooth primordia; lu, lung; and o, olfactory bulb. In panels A–E the scale bar represents 250 μ m; in panel F the bar represents 1 mm.

Figure 4 ERH is expressed in the neurons of the adult brain

Dark-field micrograph of a parasagittal section of an adult brain hybridized with the ERH riboprobe (panel A, antisense; panel B, sense). Abbreviations: o, olfactory bulb; e, ependymal cells; cp, choroid plexi; d, dentate gyrus; ce, cerebellum; h, hippocampus; th, thalamus; co, colliculi; p, pontine nucleus. The scale bar represents 1 mm.

site in the coding sequence, nt922, varied (C/T) among the cDNAs isolated; T was present in one cDNA and C in two. This T-to-C mutation changes the amino acid at position 283 from Tyr to His and is likely to represent a genetic polymorphism. The 5' ATG at nucleotide 76 which conforms to the Kozak [24] consensus is the most likely translation initiation start site and gives a deduced protein sequence of 662 amino acids and a 5' leader sequence of 75 bases. The coding sequence is followed by a 1124-nucleotide 3' untranslated sequence with a poly(A)₁₉ tail following a GUUAAA polyadenylation signal [25] at position

3159. The 3' untranslated sequence, which was used to generate ERH-specific probes and primers, is given in Figure 1.

For ease of reference we have designated the novel helicase-like gene, coding for the embryonic RNA helicase, *ERH*, on the basis of its expression pattern (see below). Translation of the *ERH* open reading frame reveals that the nine peptide motifs (I–IX) which are conserved among the DEAD box helicases are all present in *ERH* (Figure 2): the ATPase A motif A---GKT (starting at position 224); PTRELA (position 274); GG (position 302); TPGR (position 323); DEAD (position 347); SAT (position

382), F--T (position 447); RG-D (position 503) and HRIGR--R (position 527). Database searches found that the ERH protein has the highest number of amino acid identities with the PL10-like subgroup of helicases [1]; these include mouse PL10 (95%), *Xenopus* An3 (80%), yeast DED1 (53%) and DBP1 (53%), *Drosophila* vasa (42%) and *Caenorhabditis elegans* glh-1 (38%), which are more related to each other than other members of the family of DEAD box proteins. Other less related DEAD box proteins include eIF-4A, the eukaryotic translation initiation factor, p68, a mammalian nuclear protein found only in dividing cells [26], and *E. coli* SrmB, involved in 50 S ribosome assembly and mRNA stability [13,27].

ERH is most closely related to mouse PL10 (87% nucleotide identity, 95% amino acid identity) and *Xenopus* An3 (80% nucleotide identity, 88% amino acid identity) and these sequences, together with that of eIF-4A [28], are compared in Figure 2. The complete eIF-4A protein sequence is equivalent to the core region (amino acids 151–575) of ERH, which contains the motifs (I–IX). In this region, 32% identity is found between eIF-4A and ERH, mainly due to conservation of the core motifs. Conservation across the entire core sequence is more marked amongst the PL10 like proteins; there is 95% identity between An3 and ERH and 99% between PL10 and ERH.

In the PL10-like proteins the core region is flanked by conserved N-terminal (150 amino acids, ERH; 149 amino acids, PL10; 192 amino acids, An3) and C-terminal (87 amino acids, ERH; 85 amino acids, PL10; 80 amino acids, An3) extensions which are not found in eIF-4A. The N-terminal extension is conserved with 59% identity between ERH, PL10 and An3 after realignment for a 44-amino-acid insertion in An3. Similarly, at the C-terminus, 77% of residues are conserved among PL10, An3 and ERH. Detailed sequence analyses show that the overall amino acid conservation between ERH and An3 (88%) is higher than the conservation between PL10 and An3 (85%). This similarity is most marked in the N-terminal extension, where ERH and An3 show 68% identity in amino acid sequence, whereas PL10 and An3 show 59% identity.

Chromosome localization

The chromosome location of ERH was determined using the European Collaborative Interspecific Backcross [28a]. The symbol for the *ERH* gene in the mouse genetic databases will be *Ddx35*. We identified an *EcoRV* restriction-fragment-length variation (RFLV) between the two parental mice, *Mus spretus* and C57BL/6, using a 2.2 kb *EcoRI*–*SacI* fragment, including untranslated and coding sequence of *ERH* as the hybridization probe. Both species showed a non-variable band of 6 kb and a variable band of 2.4 kb in the *Mus spretus* and 3.3 kb in the C57BL/6 parental DNA samples. Inheritance of the RFLV was examined in 44 random-backcross progeny and assigned the *ERH* gene (*Hlr5*) to chromosome 1, 11.36 cM $\pm$ 4.78 cM proximal to the marker *Col3a-1*. Its position in relation to flanking markers on chromosome 1 is as follows: cen–8.64 cM–*Ddx3*–11.36 cM–*Col3a-1*–30 cM–*D1Mit12*. *Col3a-1* is the upper anchor marker on the MBx mouse map.

Expression of *ERH*

We have analysed the spatial and temporal localization of the *ERH* mRNA during development using *in situ* hybridization to mouse embryo sections and an 35 S-labelled riboprobe covering the 3' untranslated region of the *ERH* gene (Figure 1). There appear to be two phases in the expression of ERH in the developing mouse embryo. Early in development, 8.5–9.5 dpc, ERH is ubiquitously expressed at moderately high levels (Figures

Figure 5 ERH Northern-blot analyses

(a) Approx. 4 μ g of mouse polyadenylated enriched RNA from 11.5 dpc mouse dorsal tissue ['dorsal t.' (neural tube, somites and notochord)], newborn-mouse skeletal muscle ('sk. mus'), liver, brain, heart, lung and kidney, and adult skeletal muscle, liver and brain, was hybridized with a 272 bp *SacI*–*EcoRI* fragment from the 3' untranslated region of *ERH*. (b) Ethidium bromide staining of same gel shows RNA loading levels.

3, panels A and B). As development proceeds, ERH becomes more restricted in its expression. At 11.5 dpc, ERH mRNA is most abundant in the neural tube, but continues, during subsequent stages of development, to be expressed at moderate levels in a number of other sites, for example in the mesenchyme of limbs and the face and in liver (Figure 3, panel E). At 14 dpc, ERH mRNA is high in the developing brain and caudal neural tube, but is absent from the marginal layer of the neural tube which largely consists of tracts of nerve fibres (Figure 3, panel D). ERH is increased in abundance in the developing metanephros (Figure 3, panels C and D) and lung. Figure 6 (below) shows the pattern of ERH expression in the 15 dpc embryo. In addition to the abundant signal in the brain and spinal cord, ERH mRNA has accumulated in teeth primordia, the lung and in the developing limb. In adult life ERH mRNA is relatively less abundant, but shows a complex pattern of neuronal expression in the brain (Figure 4). ERH mRNA is a feature of the granule cells of the cerebellum and dentate gyrus, the pyramidal cells of the hippocampus, the ependymal cells lining the ventricles, choroid plexi and olfactory bulb. ERH mRNA also accumulates in the thalamic nuclei, the dorsal region of the colliculi and the pontine nucleus.

ERH was not abundant in neonatal (1–12 h after birth) or adult tissues, as judged by detection of the 3.2 kb mRNA after Northern-blot analysis, although low levels could be detected in neonatal brain and adult skeletal muscle. A typical autoradiograph is shown in Figure 5, which includes a sample from 11.5 dpc embryo dorsal tissue (which includes somites, neural tube and notochord) for comparison.

Our sequence analyses identified a close structural relationship between *ERH* and the *Xenopus* gene *An3*. Since in earliest

Figure 6 ERH expression in oocytes

PCR amplifications of identical aliquots of randomly primed first-strand cDNA synthesized from mouse oocyte RNA (O+) and 9.5 dpc whole mouse embryo RNA (9.5 dpc) using primers specific for ERH, *T* (T/Brachyury) and phosphoglucosyltransferase1 (PGM). No amplification product was observed when RT-PCR was carried out without reverse transcriptase (O-). C is a control PCR without cDNA. M, marker.

development *An3* is a maternally contributed mRNA and is abundant in oocytes [12], we were prompted to investigate whether ERH was similarly expressed in oocytes. RNA was prepared from mouse oocytes and a cDNA pool synthesized using random oligonucleotide primers. PCR amplifications of the oocyte cDNA were carried out using *ERH*-specific primers derived from the 3' untranslated region of *ERH* (Figure 1) alongside control amplifications using primers specific for the mRNA encoding the ubiquitously expressed phosphoglucosyltransferase1 (*Pgm-2*; [29]) and for the developmentally important *T* gene, which is first expressed in 7 dpc early-gastrulation-stage mouse embryos [30,31]. ERH mRNA was readily detected in oocytes, whereas *T* was absent and *Pgm-2* was only weakly detected (Figure 6). These data indicate that ERH is a maternal mRNA, and in this context it was of interest to find four cytoplasmic polyadenylation element (CPE)-like sequences in the 3' untranslated region of *ERH* (Figure 1). CPEs with the consensus UUUUUUAU [32,33] mediate the translational activation of stored/dormant maternal mRNAs during oogenesis and embryogenesis. Two of the CPE-like sequences in *ERH* form an 11 bp repeat 166 bp apart, but within 216 bp of the nuclear polyadenylation signal, and are identical with the UUUUAAA sequence in the hypoxanthine:guanine phosphoribosyltransferase gene which is functionally active as a CPE [34].

DISCUSSION

We have identified a putative RNA helicase which we have designated the embryonic RNA helicase (ERH) on the basis of its expression pattern. This protein is a new member of the DEAD box family of RNA helicases. The RNA unwinding and ATPase activities of DEAD box proteins eIF-4A and vasa have been ascribed to particular conserved motifs: A---GKT for ATP binding, DEAD and SAT for ATP hydrolysis and RNA unwinding, and HRIGR--R for ATP hydrolysis and RNA binding

[6,10,35]. ERH shows conservation of these functionally important motifs and is likely to have a similar helicase activity.

ERH shows highest similarity to the PL10-like subgroup of helicases which include examples from yeast, DED1 and DPB-1 [15,16], *Drosophila*, vasa [17], *Xenopus* An3 [12] and mouse, PL10, mDEAD2 and mDEAD3 [14,36]. mDEAD2 and mDEAD3 were cloned recently from mouse erythroleukaemia cells [36] and are unusual among the PL10-like helicases in showing a broad abundant pattern of expression in most adult somatic tissue, apart from lung and spleen. Other PL-10 like helicases are differentially expressed during development and appear to be particularly important in the male and female germ lines. For example, the *Drosophila* protein vasa is expressed primarily in the female germ line and is involved in oogenesis and specification of the anterior-posterior axis of the embryo [17], whereas mouse PL10 is specifically expressed at high levels in male germ cells from postnatal day 16 and is thought to play a role in the spermatogenic process [14].

Although ERH shows highest similarity both at the nucleotide (87%) and amino acid (95%) level to mouse PL10, their patterns of tissue expression are distinct. Our data show that ERH is expressed throughout the post-gastrula mouse embryo and, in later developmental stages, shows a more restricted pattern of expression in the brain, lung and kidney. ERH parallels the embryonic expression pattern of *Xenopus* An3, which begins after gastrulation when An3 mRNA is distributed nearly uniformly throughout the embryo [12]. In adult life, An3 mRNA, like that of ERH, is present in several tissues, including kidney and lung. Moreover, ERH and An3 show high sequence identity (80% at the nucleotide level and 88% at the amino acid level), despite the large evolutionary distance between mouse and *Xenopus*. An3 mRNA was originally identified as a maternally contributed message and is one of only a few mRNAs known to localize to the animal hemisphere of *Xenopus* oocytes and early embryos [12,37,38]. An3 protein is present in oocyte cytoplasm and in the nucleus, where it co-localizes with nucleoli, which are the sites for transcription and processing of rRNA and the initial steps of ribosome assembly. It has been proposed that An3 in the nucleus might be involved in the processing of large maternal pools of rRNA during oogenesis [12]. The presence of ERH mRNA in the mouse oocyte and the occurrence of four cytoplasmic polyadenylation elements in the *ERH* 3' untranslated region support the view that *ERH* is the mouse homologue of *Xenopus* An3 and is maternally transcribed.

Maternal mRNA is composed of many diverse sequences, some of which are translated during oocyte maturation, fertilization or early embryogenesis. Maternal control of development is exerted by differential translational activation of these dormant mRNAs. The processes which regulate translational activation of target mRNAs have not been clearly identified; while cytoplasmic extension of their polyadenylated tails has been implicated [32,33], it is not yet understood how the lengthened tail is recognized by the translational apparatus. One attractive possibility is that DEAD box RNA helicases, such as ERH, in the cytoplasm play a direct role in this translational activation. Local unwinding of maternal mRNAs by the oocyte-specific RNA helicase ERH could render them accessible for cytoplasmic polyadenylation or could modulate mRNA binding to ribosomes and influence translation initiation in a similar way to the ubiquitous helicase eIF-4A [6].

The role of ERH during development needs to be clearly established. Of particular relevance to this question is the possibility that particular DEAD box helicases act on either a single RNA substrate or on specific subsets of mRNA substrates and in this way regulate different cellular processes. The recent

discovery that the *E. coli* DEAD box protein DbpA can hydrolyse ATP only in the presence of bacterial 23 S ribosomal RNA [39] is the first *in vitro* demonstration of a specific RNA sequence requirement by a DEAD box protein for ATPase and helicase activity. The conserved regions of DEAD box helicases do not include any of the major RNA binding motifs characterized to date (reviewed in [40]), although recent work has shown that the three conserved arginine residues in the HRIGR-R motif of eIF-4A are essential for RNA binding [6]. The N- and C-terminal extensions which are characteristic of the PL10-like proteins may provide RNA binding specificity. Indeed individual proteins appear to employ different RNA-binding motifs, and this may be the key to their different functions. For example, the recently identified *C. elegans* DEAD box protein Glh-1, which is expressed in germ cells, has four retroviral-type zinc-finger motifs in its N-terminus which may be involved in RNA binding [18], whereas the *Drosophila* vasa protein has an RGG box RNA-binding motif in its N-terminal region consisting of five closely spaced Arg-Gly-Gly repeats [17]. It is of considerable interest that, despite their evolutionary distance, ERH and An3 show high levels of conservation in the N- and C-terminal extensions, and this is particularly marked at the N-terminus. Analyses of the extension sequences show that neither ERH nor An3 have the zinc-finger motif or the RGG box, but it seems likely that other, as yet uncharacterized, motifs may be involved in the RNA binding of these two proteins.

We thank John Gearhart (The Johns Hopkins University) for the 8.5 dpc mouse embryo cDNA library and Orla McGuinness (University of Cambridge) for mouse oocytes.

REFERENCES

- Schmid, S. and Linder, P. (1992) *Mol. Microbiol.* **6**, 283–292
- Linder, P., Lasko, P., Ashburner, M. et al. (1989) *Nature (London)* **337**, 121–122
- Rozen, F., Ederly, I., Meerovitch, K., Dever, E., Merrick, W. and Sonnenberg, N. (1990) *Mol. Cell. Biol.* **10**, 1134–1144
- Ray, B., Lawson, T., Kramer, J. et al. (1985) *J. Biol. Chem.* **260**, 7651–7658
- Thach, R., (1992) *Cell* **68**, 177–180
- Pause, A., Methot, N. and Sonnenberg, N. (1993) *Mol. Cell. Biol.* **13**, 6789–6798
- Walker, J., Saraste, M., Runswick, M. and Gay, N. (1982) *EMBO J.* **1**, 945–951
- Fuller-Pace, F. (1994) *Trends Cell Biol.* **4**, 271–274
- Gorbalenya, A., Koonin, V., Dochenko, A., and Blinov, V. (1989) *Nucleic Acids Res.* **17**, 4713–4729
- Liang, L., Diehl-Jones, W. and Lasko, P. (1994) *Development* **120**, 1201–1211
- Hirling, H., Scheffner, M., Restlt, T. and Stahl, H. (1989) *Nature (London)* **339**, 562–564
- Gururajan, R., Perry-O'Keefe, H., Melton, D. and Weeks, D. (1991) *Nature (London)* **349**, 717–719
- Iost, I. and Dreyfus, M. (1994) *Nature (London)* **372**, 193–196
- Leroy, P., Azari, P., Sassoon, D., Wolgemuth, D. and Fellous, M. (1989) *Cell* **57**, 549–559
- Jamieson, D. and Beggs, J. (1991) *Mol. Microbiol.* **5**, 805–812
- Jamieson, D., Rahe, B., Pringle, J. and Beggs, J. (1991) *Nature (London)* **349**, 715–717
- Lasko, P. and Ashburner, M. (1988) *Nature (London)* **335**, 611–617
- Roussell, D. and Bennett, K. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 9300–9304
- Maniatis, T., Fritsch, E. and Sambrook, J. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor laboratory Press, Cold Spring Harbor, NY
- Sanger, F., Nicklen, S. and Coulson, A. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Altschul, S., Gish, W., Miller, W., Myers, E., and Lipman, D. (1990) *J. Mol. Biol.* **215**, 403–410
- Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.* **162**, 156–159
- Lyons, G., Buckingham, M., Tweedie, S. and Edwards, Y. (1991) *Development* **111**, 233–244
- Kozak, M. (1986) *Cell* **44**, 283–292
- Proudfoot, N. and Brownlee, G. (1976) *Nature (London)* **263**, 211–214
- Ford, M., Anton, I. and Lane, D. (1988) *Nature (London)* **332**, 736–738
- Nishi, K., Morel-Deville, F., Hershey, J., Leighton, T. and Schnier, J. (1988) *Nature (London)* **336**, 496–498
- Nielsen, P. and Trachsel, H. (1988) *EMBO J.* **7**, 2097–2105
- Breen, M., Deakin, L., Macdonald, B. et al. (European Backcross Collaborative Group) (1994) *Hum. Mol. Genet.* **3**, 621–627
- Whitehouse, D., Putt, W., Lovegrove, J. et al. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 411–415
- Herrmann, B., Labeit, S., Poustka, A., King, T. and Lehrach, H. (1990) *Nature (London)* **343**, 617–622
- Wilkinson, D., Bhatt, S. and Herrmann, B. (1990) *Nature (London)* **343**, 657–6594
- Richter, J. (1993) *Dev. Genet.* **14**, 407–411
- Salles, F., Darrow, A., O'Connell, M. and Strickland, S. (1992) *Genes Dev.* **6**, 1202–1212
- McGrew, L., Dworkin-Rastl, E., Dworkin, M. and Richter, J. (1989) *Genes Dev.* **3**, 803–815
- Pause, A. and Sonnenberg, N. (1992) *EMBO J.* **11**, 2643–2654
- Gee, S. and Conboy, J. (1994) *Gene* **140**, 171–177
- Rebagliati, M., Weeks, D., Harvey, R. and Melton, D. (1985) *Cell* **42**, 769–777
- Weeks, D. and Melton, D. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2798–2802
- Fuller-Pace, F., Nicol, S., Reid, A. and Lane, D. (1993) *EMBO J.* **12**, 3619–3626
- Burd, C. and Dreyfuss, G. (1994) *Science* **265**, 615–621