The EMBO Journal vol.6 no.3 pp.761–766, 1987

EGF homologous sequences encoded in the genome of *Drosophila melanogaster*, and their relation to neurogenic genes

Elisabeth Knust, Ursula Dietrich, Ulrich Tepass, Kirsten A.Bremer, Detlef Weigel[1], Harald Vässin and Jose A.Campos-Ortega

Institut für Entwicklungsphysiologie der Universität zu Köln, Gyrhofstrasse 17, 5000 Köln 41, FRG

[1]Present address: Max-Planck-Institut für Entwicklungsbiologie, Spemannstrasse 37, 7400 Tübingen 1, FRG

Communicated by J.A.Campos-Ortega

The function of the neurogenic genes of *Drosophila melanogaster* is required for a normal pattern of commitment of neural and epidermal progenitor cells. In the course of searching for a molecular basis for the functional interrelationships that exist between the neurogenic genes, fragments of cloned DNA from the genes master mind (*mam*), Delta (*Dl*), Enhancer of split [*E(spl)*] and Notch (*N*) were hybridized to each other. Strong cross-hybridization was observed between a fragment of the *Dl* gene and a fragment of the *N* gene encoding a peptide with homology to several proteins of mammals, including the epidermal growth factor (EGF). Sequencing of this *Dl* fragment revealed an open reading frame encoding four EGF-like repeats with homology to the repeats found in the *N* gene. Screening genomic and cDNA libraries under conditions of reduced stringency with *Dl* and *N* probes that encode EGF-like repeats uncovered several cross-hybridizing clones, suggesting that other *Drosophila* genes may also encode such peptides. Part of a cross-hybridizing cDNA clone, derived from a gene located at position 95F on the third chromosome, was sequenced and found to encode five repeats with homology to those encoded by *N* and *Dl*. Preliminary evidence on the spatial pattern of transcription indicates that the gene at position 95F is regulated in its expression, as it is transcribed in all ectodermal derivatives, with the exception of the central nervous system. Indirect evidence suggests that this clone may derive from the crumbs (*crb*) gene, which is likely to be an hitherto unknown neurogenic gene. The chromosomal location of the other cross-hybridizing fragments does not correspond to any other neurogenic gene known to us.

Key words: EGF-like repeats/Notch and Delta/neurogenesis/*Drosophila melanogaster*/middle repetitive sequences

Introduction

The separation of neuroblasts from the ectoderm into the inner part of the embryo is one of the first steps of central nervous system (CNS) development in insects (Poulson, 1950; Bate, 1982; Hartenstein and Campos-Ortega, 1984; Doe and Goodman, 1985). In *Drosophila melanogaster*, this process is under the control of the neurogenic genes (ref. to Campos-Ortega, 1985). Loss of function mutations in any of these genes result in the development of all cells of the neurogenic ectoderm as neuroblasts (Poulson, 1937; Lehmann *et al.*, 1983).

Previous genetic analyses have revealed the existence of a complex pattern of functional relationships between the neurogenic genes (Campos-Ortega *et al.*, 1984; Dietrich and Campos-Ortega, 1984). In particular the results of gene dosage studies are consistent with the notion that the neurogenic genes are links of a chain of epistatic relationships (Vässin *et al.*, 1985; de la Concha *et al.*, in preparation). The DNA of four neurogenic genes, i.e. Notch (*N*, Artavanis-Tsakonas *et al.*, 1983; Kidd *et al.*, 1983), master mind (*mam*, Weigel *et al.*, 1987), Delta (*Dl*, H.Vässin *et al.*, in preparation) and Enhancer of split [*E(spl)*, E.Knust *et al.*, in preparation] has been cloned, which allowed us to search for a molecular basis for these functional interrelationships.

The sequence of the neurogenic gene *N*, as determined recently by Wharton *et al.* (1985b), uncovered a particularly interesting structure. *N* encodes a putative transmembrane protein, the extracellular domain of which contains 36 repeated units with homology to a group of mammalian proteins that includes the epidermal growth factor (EGF) (Wharton *et al.*, 1985b; Kidd *et al.*, 1986). We have looked for sequence homologies between the *mam*, *Dl* and *E(spl)* locus and fragments of the *N* gene. A small fragment of the *Dl* gene turned out to be homologous to the stretch of the *N* gene that encodes the EGF-like repeats. We used these fragments of the *N* and *Dl* genes to screen genomic and cDNA libraries under conditions of low stringency, and detected a few other cross-hybridizing clones. One of the clones was partially sequenced and was found to share this homology. Indirect evidence suggests that this gene may correspond to the crumbs (*crb*) locus, which is likely to be another neurogenic gene.

Results

EGF-like repeats in the Dl locus

DNA clones of the neurogenic genes *mam*, *Dl* and *E(spl)* were hybridized to each other and to cosmid *132D4*, which comprises ~25 kb DNA of the *N* region (Artavanis-Tsakonas *et al.*, 1983). Extensive homology was found to the repeated element opa, which encodes a polyglutamine stretch (Wharton *et al.*, 1985a). Regions of cross-hybridization to the opa sequence were found once in the DNA of *Dl* and *E(spl)* (unpublished) respectively, and several times in the DNA of *mam* (Weigel *et al.*, 1987).

In addition, under conditions of low stringency, cross-hybridization was also detected between a fragment of cosmid *132D4* encoding EGF-like repeats and part of the *Dl* locus. The region of cross-hybridization in *Dl* could be assigned to a 600-bp *Hin*cII fragment within 25 kb of genomic DNA to which several *Dl* mutations have been mapped by Southern blotting (Vässin, 1986). Moreover, the same *Hin*cII fragment was found to be part of a cDNA isolated from a library of 3- to 12-h embryos (K.Bremer, H.Vässin and E.Knust, unpublished). Apart from the 600-bp fragment, the cDNA contains a second cross-hybridizing fragment adjacent to it. To verify the structural homology to the *N* sequence, the nucleotide sequence of the cross-hybridizing *Hin*cII *Dl* fragment and of the adjacent region of the

```
    1  TGCAGTCCGAATCCCTGCATAAACGGTGGAAGCTGTCAGCCGAGCGAAACGTGTATTTGC   60
       ---------+---------+---------+---------+---------+---------+
       C  S  P  N  P  C  I  N  G  G  S  C  Q  P  S  E  T  C  I  C

   61  CCCAGCGGATTTTCGGGAACGAGGTGCGAGACCAACATTGACGATTGCCTTGCGCACCAG  120
       ---------+---------+---------+---------+---------+---------+
       P  S  G  F  S  G  T  R  C  E  T  N  I  D  D  C  L  A  H  Q

  121  TGCGAGAACGGAGGCACCTGCATAGATATGGTCAACCAATATCGCTGCCAATGCGTTCCC  180
       ---------+---------+---------+---------+---------+---------+
       C  E  N  G  G  T  C  I  D  M  V  N  Q  Y  R  C  Q  C  V  P

  181  GGCTTTCATGGCACCCACTGTAGTACGAAAGTTGACTTGTGCCTCATCAGACCGTGTGCC  240
       ---------+---------+---------+---------+---------+---------+
       G  F  H  G  T  H  C  S  T  K  V  D  L  C  L  I  R  P  C  A

  241  AATGGAGGAACCTGCTTGAATCTCAACAACGATTACCAGTGCACCTGTCGTGCGGGATTT  300
       ---------+---------+---------+---------+---------+---------+
       N  G  G  T  C  L  N  L  N  N  D  Y  Q  C  T  C  R  A  G  F

  301  ACTGGCAAGGATTGCTCCGTGGACATCGATGAGTGCAGCAGTGGACCCTGTCATAACGGC  360
       ---------+---------+---------+---------+---------+---------+
       T  G  K  D  C  S  V  D  I  D  E  C  S  S  G  P  C  H  N  G

  361  GGCACTTGCATGAACCGCGTCAATTCGTTCGAATGCGTGTGTGCCAATGGTTTCAGGGGC  420
       ---------+---------+---------+---------+---------+---------+
       G  T  C  M  N  R  V  N  S  F  E  C  V  C  A  N  G  F  R  G

  421  AAGCAGTGCGATGAGGAGTCCTACGATTCGGTGACCTTCGATGCCCACCAATATGGAGCG  480
       ---------+---------+---------+---------+---------+---------+
       K  Q  C  D  E  E  S  Y  D  S  V  T  F  D  A  H  Q  Y  G  A

  481  ACCACACAAGCGAGAGCCGATGGTTTAGCCAATGCCCAGGTAGTCCTAATTGCTGTTTTC  540
       ---------+---------+---------+---------+---------+---------+
       T  T  Q  A  R  A  D  G  L  A  N  A  Q  V  V  L  I  A  V  F

  541  TCCGTTGCGATGCCTTTGGTGGCGGTTATTGCGGCGTGCGTGGTCTTCTGCATGAAGCGC  600
       ---------+---------+---------+---------+---------+---------+
       S  V  A  M  P  L  V  A  V  I  A  A  C  V  V  F  C  M  K  R
```

Fig. 1. Nucleotide sequence of part of the embryonic cDNA of the *Dl* locus, and the predicted amino acid sequence. The four repeated units are underlined. Single-letter amino acid designations are A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr.

Dl gene was determined. The nucleotide sequence, as well as the deduced amino acid sequence, are shown in Figure 1. Four repeated units, comparable to the sequences encoding the EGF-like repeats described for the *N* locus, are found in this part of the *Dl* cDNA (see below). The homology between the cDNA of *Dl* and the region of the cDNA of *N* encoding the EGF-like repeats ranges from 57 to 66%, depending on the part of the *N* sequence considered. For the sake of comparison, the homology between the nucleotide sequences of the individual EGF-like repeats encoded by *N* was found to range from 59 to 95%. At the amino acid level, the homology between the *Dl* and the *N* repeats was found to be 45%.

Using the same conditions of hybridization at low stringency, no cross-hybridization was detected between DNA fragments from *N* and *Dl* that encode EGF-like repeats and DNA clones from the *E(spl)* region. The *E(spl)* region has been defined by mapping several breakpoints and other mutations of the *E(spl)* gene at the molecular level (E.Knust *et al.*, in preparation). A weak cross-hybridization was observed to a fragment in the neighbourhood of the *mam* gene; sequencing of this fragment, however, showed that this hybridization was not due to structural homology (data not shown).

Other genes with homology to the EGF-like sequence

One strong and several weaker fragments were detected in genomic Southern blots probed under conditions of low stringency with fragments encoding the EGF-like repeats from either the *N* or the *Dl* locus (Figure 2). The bands giving the strong hybridization signals correspond in size to the unique fragments from *N* and *Dl*, respectively, which were used as probes. The presence of several additional cross-hybridizing genomic fragments suggested to us the existence of additional genes encoding EGF-like repeats in the genome of *D. melanogaster*. To determine whether these additional fragments derive from other neurogenic genes, a genomic library (approximately eight genome equivalents) and two cDNA libraries (from 3- to 12- and 12- to

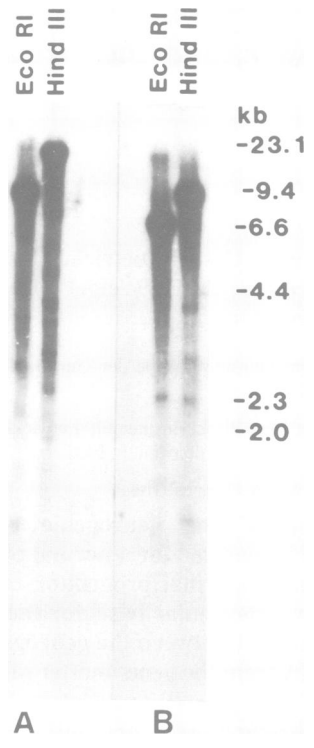

Fig. 2. Detection of additional sequences in the genome of *Drosophila* with homology to EGF-like sequences. Genomic DNA digested with either *Eco*RI or *Hin*dIII was hybridized under low stringency conditions (see Materials and methods) to fragments of *N* (**A**) or *Dl* (**B**) encoding EGF-like repeats. The autoradiogram was overexposed in order to demonstrate weakly hybridizing fragments.

Table I. Chromosomal location of clones that cross-hybridize to DNA probes encoding EGF-like repeats from the *Dl* and *N* genes

Probe	Location	Library
Delta	3C7 (*N*)	genomic
	17E/F	cDNA
	21E	genomic
	25D	cDNA
	69A/B	genomic
	92A2 (*Dl*)	cDNA
	95F	genomic
	98D3-7	cDNA
	98E	cDNA
Notch	2D/E	cDNA
	3C7 (*N*)	genomic
	24E	cDNA
	28E/F	cDNA
	52E	cDNA
	76D/E	cDNA
	92A2 (*Dl*)	cDNA
	95F	genomic and cDNA

24-h embryos, respectively), were screened under conditions of low stringency with the same *N* and *Dl* fragments encoding the EGF-like repeats. Several clones were isolated, and their chromosomal origin was determined by *in situ* hybridization of ^{3}H-labelled DNA to polytene chromosomes from larval salivary glands. The localization of these clones in the genome is given in Table I. Several of the clones were found to derive from either *Dl* or *N*; of the remaining ones only the clone at 2D-E hybridiz-

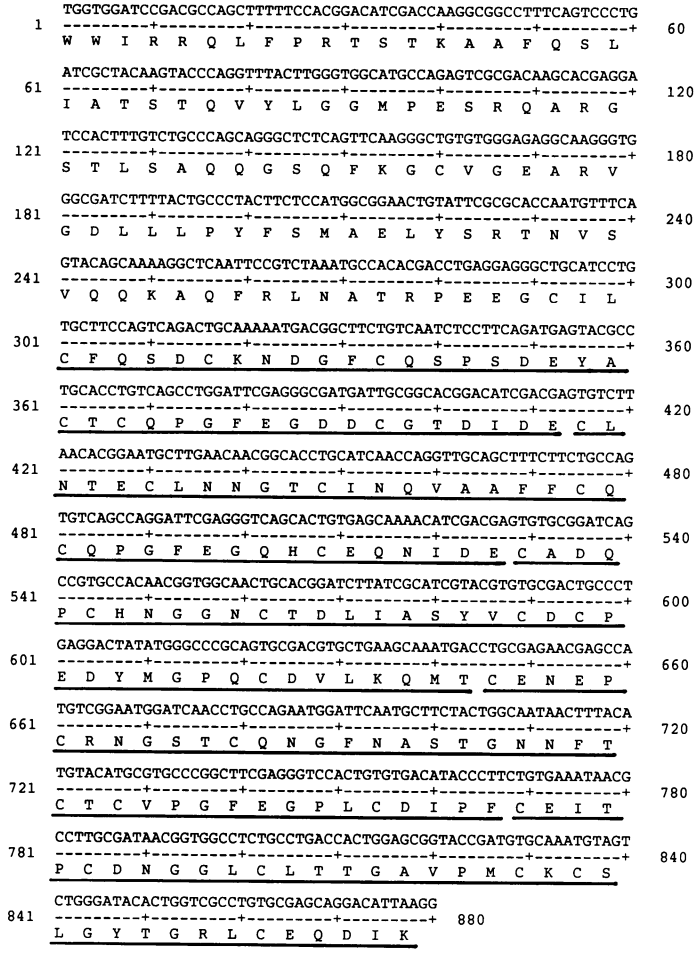

Fig. 3. Nucleotide sequence of part of the embryonic cDNA of the 95F region, and the predicted amino acid sequence. The five repeat units are underlined. Single-letter amino acid designations are the same as in Figure 1.

ed to the neighbourhood of another known neurogenic gene, pecanex (*pcx*) (LaBonne and Mahowald, 1985). In subsequent experiments, however, this clone did not hybridize to the cosmid *cos9*, which includes the *pcx* gene (Haenlin *et al.*, 1985) (data not shown).

One of the clones giving a relatively strong hybridization signal derives from the chromosomal location 95F on the right arm of the third chromosome (Table I). Two findings make this clone particularly interesting. First, it was the only clone detected with the probes of both *Dl* and *N*. Second, both a genomic and a cDNA clone were isolated from this region. Both the genomic and cDNA clone from 95F were partially sequenced to ascertain a possible structural homology to the sequences encoding EGF-like repeats. The nucleotide sequence of the cDNA clone and the amino acid sequence deduced from it are shown in Figure 3. The degree of sequence homology between the EGF-like repeats encoded by the 95F and the *N* cDNA was found to vary between 57 and 61%. At the amino acid level, the homology between the peptide encoded by the gene at 95F and the peptides encoded by *N* and *Dl* was found to be ~40 and 33%, respectively.

The optimal alignment of parts of the predicted amino acid sequences of both the *Dl* and the 95F gene fragments revealed an arrangement similar to the consensus sequence of all 36 EGF-like repeats encoded in the *N* gene (Wharton *et al.*, 1985b) and thus similar to the mammalian EGF and other proteins (Figure 4). The homology is chiefly based on the spacing of six cysteine residues within each repeated unit, as well as on the conservation of other amino acid residues (cf. Wharton *et al.*, 1985b). The *Dl* sequence encodes four complete repeat units, while the clone from 95F encodes a minimum of five such repeats (Figure 4).

The full extent of the open reading frames including the EGF-like repeats of *Dl* and 95F cDNAs will be published elsewhere.

The spatial pattern of transcription of the gene at 95F

With the aim of obtaining indications about the function of the gene from 95F, we performed *in situ* hybridization to tissue sections of embryos at different stages of development. ^{35}S-labelled single-stranded RNA probes were synthesized from a 4-kb cDNA clone that includes the sequences encoding the EGF-like repeats. Figure 5 shows some examples of the hybridization pattern. Transcripts are clearly detectable at the blastoderm stage, from the beginning of cell formation (early stage 5, according to Campos-Ortega and Hartenstein, 1985). Transcription is particularly intense in the neurogenic region of the ectoderm, as well as within the anlagen of the dorsal epidermis and the amnioserosa (Figure 5A). A weak signal is detected in the endodermal anlagen, and no signal is apparent in the mesodermal region (Figures 5B and C). Strikingly, hybridization is localized in the apical, immediately supranuclear sector of the blastoderm cells. At the beginning of germ band extension (stage 8), no more transcripts are detectable in endodermal derivatives (midgut primordia; not shown). From stage 9 onwards, in embryos with extended germ band, transcripts are limited to the epidermis and to parts of the foregut and to the entire hindgut. They become particularly abundant in the tracheal pits (Figure 5E), while the primordium of the nervous system is devoid of transcripts (Figure 5D and E). This pattern remains constant until the completion of embryogenesis.

Discussion

Several regions were identified in the genome of *D. melanogaster* which cross-hybridize to DNA fragments that encode repeated peptide units with homology to EGF and other mammalian proteins. The *N* gene has been found to encode 36 EGF-like repeats that are tandemly arranged (Wharton *et al.*, 1985b; Kidd *et al.*, 1986). Here, we describe the partial sequence of two additional genes, *Dl* and a gene at 95F, both of which also contain stretches encoding putative EGF-like peptides. *N* and *Dl* are neurogenic genes. The function of the gene at 95F is still subject to investigation; two arguments provide indirect evidence, which suggests that it might correspond to *crb*, a gene identified by Jürgens *et al.* (1984). On the one hand, both genes have a similar chromosomal location (*crb* has been mapped to the interval 95E–96A by Jürgens *et al.*, 1984). On the other hand, the spatial pattern of transcription of the gene at 95F shows a striking congruence with the parts affected in mutant *crb* embryos. Transcription of the gene at 95F is restricted to ectodermal derivatives, i.e. epidermis, fore- and hindgut; at early stages transcripts are particularly abundant within the neurogenic region of the ectoderm. Except for the blastoderm stage, when a low level of transcription is also observed in the endodermal anlagen, expression is neither detectable in mesodermal primordia from early stages onwards, nor in endodermal primordia during postblastoderm stages of embryogenesis. The phenotypic abnormalities exhibited by embryos homozygous for crb^{-} mutations are primarily restricted to ectodermal derivatives. The embryos display considerable neural hyperplasia, along with epidermal hypoplasia and morphogenetic defects of both fore- and hindgut

```
Dl  (cDNA)     1    CSPNPCINGGSC...QPSET.....CICPSGFSGTRCETNIDD    34
Dl    "       35    CLAHGCENGGTCIDYGNQYR.....CQCVPGFHGTHCSTKVDL    72
Dl    "       73    CLIAPCANGGTCLNLNNDYQ.....CTCRAGFTGKDCSVDIDE   110
Dl    "      111    CSSGPCHNGGTCMNRVNSFE.....CVCANGFRGKQCDEESYD   148

95F(cDNA)    101    CFQSDCKNDGFCQSPSDEYA.....CTCQPGFEGDDCGTDIDE   138
95F   "      139    CLNTECLNNGTCINQVAAFF.....CQCQPGFEGQHCEQNIDE   176
95F   "      177    CADQPCHNGGNCTDLIASYV.....CDCPEDYMGPQCDVLKQMT  215
95F   "      216    CENEPCRNGSTCQNGFNASTGNNFTCTCVPGFEGPLCDIPF..   256
95F   "      257    CEITPCDNGGLCLTTGAVPN.....CKCSLGYTGRLCEQDIK.   293
                                          F           F
Notch consensus     CXSXPCXNGGTCXDXXXX X.....CXCXXG XGXXCEXXXDX
                                          Y           Y
```

Fig. 4. Optimal alignment of the EGF-like repeats encoded in the *Dl* and 95F cDNA clones with the consensus sequence of all 36 EGF-like repeats encoded in the *N* gene (Wharton *et al.*, 1985b). Notice the correct spacing of the cysteine residues and, in addition, the conservation of some other amino acids.

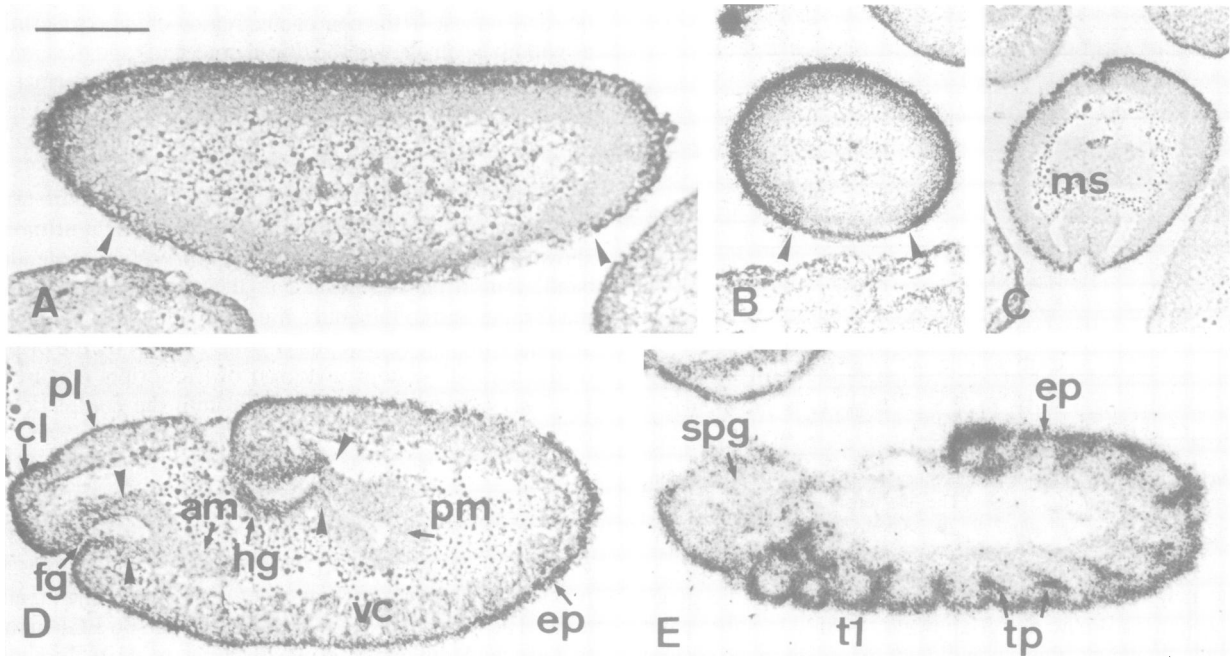

Fig. 5. Hybridization of a probe from 95F to sections of embryos at different stages. Bright field photographs, same magnification for all pictures [scale bar in (**A**) 100 μm]. (**A**) shows a sagittal and (**B**) a transversal section of embryos at the blastoderm stage. The region between the arrowheads in both pictures is devoid of silver grains; it corresponds to the anlage of the mesoderm where no transcripts are detectable. Contrarily, transcripts are abundant in the remaining cells, particularly in the neurogenic and non-neurogenic regions of the ectoderm (compare with the fate map of Hartenstein *et al.*, 1985). Notice the high concentration of transcripts present in the apical portions of the blastoderm cells. (**C**) shows a section through a gastrulating embryo. No hybridization can be seen within the mesodermal primordium (ms); the remaining cells show abundant transcripts. (**D**) and (**E**): in the embryo at stage 11 (extended germ band) transcripts continue to be present in the entire epidermal primordium (ep), and in parts of both foregut (fg) and hindgut (hg), being particularly abundant in the epidermis surrounding the clypeolabrum (cl in **D**) and the tracheal pits (tp in **E**). Notice the absence of transcripts over the epidermis covering the procephalic lobe (pl in **D**) and the supraoesophageal ganglion (spg in **E**). Transcription ends abruptly within the foregut, at the presumptive boundary between the primordia of pharynx and oesophagus (arrowheads), and at the boundary between hindgut and posterior midgut (pm). All other regions of the embryo are devoid of transcription. This pattern of expression is maintained throughout embryogenesis. Other abbreviations: am: anterior midgut: tl: prothorax; vc: ventral cord.

(unpublished observations). Thus, the phenotype of *crb*$^-$ mutations is perfectly compatible with the pattern of expression of the gene at 95F described above. Neural hyperplasia and epidermal hypoplasia are severe in *crb*$^-$ mutants and reminiscent of the phenotypic defects displayed by loss-of-function mutations of neurogenic genes in general (Poulson, 1937; Lehmann *et al.*, 1983), suggesting that *crb* is a neurogenic gene. However, further studies on the embryonic development of *crb*$^-$ mutants are necessary to decide whether the neural hyperplasia shown by these mutants is due to a neurogenic misrouting of ectodermal cells, as in the neurogenic mutants, or due to a different mechanism.

We do not yet know whether any of the remaining cross-hybridizing genes identified in this work do actually exhibit sequence homology to the repeated DNA motif present in *N*, *Dl* and 95F. Nor do we know whether any (or all) of those genes will eventually turn out to be neurogenic genes. In any case, we would like to emphasize that EGF-coding sequences are probably not a characteristic feature of neurogenic genes in general, in spite of their presence in both *N* and *Dl*: neither the DNA of

mam nor of *E(spl)* cross-hybridized to DNA encoding EGF-like repeats under the conditions we used. Therefore, the neurogenic genes *mam* and *E(spl)* probably do not encode EGF-like polypeptides, although they may contain a more distantly related class of repeats, which we are unable to detect under the conditions used. Also, none of the cross-hybridizing clones described here derived from the location of any known neurogenic gene (e.g. *amx*, *bib*, *neu*, etc; see Results concerning *pcx*).

Recently, part of the *lin-12* gene of the nematode *Caenorhabditis elegans* has been cloned and sequenced (Greenwald, 1985). This DNA encodes a peptide with 11 repeated units homologous to the EGF. The *lin-12* locus controls certain binary decisions during development (Greenwald *et al.*, 1983; Sternberg and Horvitz, 1984). The function of *lin-12* is to some extent similar to that of the neurogenic genes of *Drosophila*, since the latter also participate in a binary decision during development, i.e. the segregation between the neural and the epidermal lineages. Moreover, both in *C. elegans* and in *Drosophila*, cell−cell interactions seem to play an important role in the commitment of cell fate. This conclusion can be drawn from results of laser ablation experiments in *C. elegans* (Sulston and White, 1980; Sulston *et al.*, 1983). and from cell transplantation experiments in *Drosophila* (Technau and Campos-Ortega, 1986). It is conceivable that proteins containing EGF-like peptides take part in the process of cell communication that leads to cell commitment during development.

However, the biological role of the EGF-like peptides in the proteins encoded by these three (and presumably more) genes of *Drosophila* is indeed not understood. The mammalian EGF itself, a 53-amino acid polypeptide, has been shown to induce cell division and cell differentiation in cultured cells of ectodermal and mesodermal origin (for review see Carpenter and Cohen, 1979). Apparently, the EGF is released from a large precursor protein that exhibits the structural features of a transmembrane protein (Gray *et al.*, 1983; Scott *et al.*, 1983). The precursor contains nine related cysteine-rich peptide units — the EGF repeats. Several mammalian proteins are known to contain one or several EGF-like repeats. These proteins have in common only that they are either membrane-bound, e.g. the bovine and human LDL receptor (Russel *et al.*, 1984; Südhof *et al.*, 1985) or secreted, e.g. EGF (Gray *et al.*, 1983; Scott *et al.*, 1983), maturing transforming growth factor α (TGF-α, Marquardt *et al.*, 1984), blood clotting factors IX, X and protein C (Anson *et al.*, 1984; Doolittle *et al.*, 1984; Foster and Davie, 1984).

Our current information on the structure of the gene products of the *N* and *Dl* genes does not allow any firm conclusion to be drawn concerning their mode of action. The putative *N* gene product is a transmembrane protein, which contains in its extracellular domain 36 EGF-like repeats (Wharton *et al.*, 1985b; Kidd *et al.*, 1986). It is not known whether one or several of these repeated units are cleaved from a precursor protein and function as peptide hormones, in a way reminiscent of the EGF itself. Alternatively, the entire membrane-bound protein may function as a receptor, or even interact directly with the neighbouring cells. Genetic mosaics show that the gene products of *N* and *Dl* are unable to diffuse over long distances (Dietrich and Campos-Ortega, 1984; Hoppe and Greenspan, 1986); these data argue against a hypothetical function of these gene products as peptide hormones. The available experimental evidence suggests rather the participation of *N* and *Dl* in cell interactions requiring immediate cell−cell contact. When cells of the neurogenic region of N^- and Dl^- embryos are transplanted individually into the neurogenic region of wild-type embryos, the transplanted cells behave like wild-type cells and give rise to both neural and epidermal progenies (G.M.Technau and J.A.Campos-Ortega, unpublished), although the transplanted mutant cells would have invariably developed as neuroblasts in the mutant embryo. This result indicates that the fate of the mutant cells is determined by the neighbouring wild-type cells and thus that neither *N* nor *Dl* are cell autonomous in their mode of expression.

Thus, concerning the role of neurogenic genes in early neurogenesis, the current evidence points to their participation in cell communication, mediated by membrane-bound processes. There is no indication as to the role played by the EGF-like repeats in these processes. It is conceivable that their only function is to give a particular shape to the proteins containing them. The positions of six cysteine residues, which are evolutionarily conserved, would allow the formation of a specific three-dimensional structure via disulfide bridges. The importance of a three-dimensional structure for correct biological activity has been shown for the EGF (Savage *et al.*, 1973) and has been proposed for the TGF-α (Marquardt *et al.*, 1984). Considered from this point of view, the structural homology between these proteins would indicate evolutionary relationships, as proposed by Doolittle *et al.* (1984). The presence of homologous repeated peptides in different proteins may be the result of a divergent evolution from a hypothetical common ancester, distributed to additional genes by duplication and translocation (see Doolittle *et al.*, 1984; Patthy, 1985, for further discussion).

Materials and methods

Preparation of genomic DNA and Southern blot hybridization

Genomic DNA from adult flies was prepared as described by Weigel *et al.* (1987). Blotting the genomic DNA onto GeneScreenPlus membrane (NEN/Dupont) was carried out according to Reed and Mann (1985), transfer to nitrocellulose filters according to Maniatis *et al.* (1982). Hybridization was performed as described by Weigel *et al.* (1987).

Hybridizations under conditions of low stringency were carried out overnight at 60°C in 6 × SSPE, 5 × Denhardt, 50 mM phosphate buffer pH 7.0, 10 mM pyrophosphate, 0.1% SDS, 1 mg/ml calf-thymus DNA and $1-4 \times 10^7$ c.p.m. of ^{32}P-labelled nick-translated probe. The filters were washed three times for 20 min in 4 × SSPE, 0.1% SDS at 50°C (protocol according to H.Jäckle, personal communication; see Schuh *et al.*, 1986).

In situ hybridization to polytene chromosomes

Preparation of polytene chromosomes and hybridization conditions were essentially as described by Pardue (1985), with the modifications described by Weigel *et al.* (1987).

Sequencing

Sequencing was carried out essentially as described by Maxam and Gilbert (1980) and Sanger *et al.* (1980). For the dideoxy sequencing, restriction fragments were cloned into M13mp18 or M13mp9 vectors into the corresponding restriction sites. Long inserts were shortened by *Bal*31 digestion, eluted and ligated into the appropriate M13 vectors (Poncz *et al.*, 1982). The nucleotide sequences were determined for both strands. The nucleotide sequence and the predicted amino acid sequence were analyzed on a VAX/VMS computer version V4.4, using the programs of the University of Wisconsin (Devereux *et al.*, 1984).

In situ hybridization to embryonic tissue sections

In situ hybridization to embryonic tissue sections were performed essentially as described by Ingham *et al.* (1985) with laboratory modifications. DNA used as a probe for hybridization was subcloned into pGem-2 (Promega) and ^{35}S-labelled RNA was synthesized using either the SP6 or the T7 RNA polymerase.

Other procedures

Whole restriction-digested phage DNA or gel-purified fragments were subcloned into pGem-2 (Promega). Small-scale preparations were performed as described by Willimzig (1985). Large-scale preparations were carried out according to Holmes and Quigley (1981).

Phage DNA preparations, nick-translation and screening of recombinant phage libraries were as described in Maniatis *et al.* (1982). For a description of the solutions used see Maniatis *et al.* (1982).

Embryonic cDNA libraries in λgt10 were obtained from L.Kauvar (Poole *et al.*, 1985).

Acknowledgements

We thank Ferdi Grawe and Sigrid Baars for technical assistance during different parts of this work, L.Kauvar for the gift of cDNA libraries, S.Artavanis-Tsakonas for cosmid *132D4*, Eliane Mohier for cosmid *cos9*, C.Nüsslein-Volhard for *crb* mutations and Michael Brand for critical reading of the manuscript. Ursula Dietrich was a fellow of the Boehringer-Ingelheim Fond and Detlef Weigel was supported by the Studienstiftung des Deutschen Volkes. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, SFB 74).

References

Anson,D.S., Choo,K.H., Rees,D.J.G., Giannelli,F., Gould,K., Huddleston,J.A. and Brownlee,G.G. (1984) *EMBO J.*, **3**, 1053–1060.

Artavanis-Tsakonas,S., Muskavitch,M.A.T. and Yedvobnick,B. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 1977–1981.

Bate,C.M. (1982) *Neurosci. Res. Prog. Bull.*, **20**, 803–806.

Campos-Ortega,J.A. (1985) *Trends Neurosci.*, **8**, 245–250.

Campos-Ortega,J.A. and Hartenstein,V. (1985) *The Embryonic Development of Drosophila melanogaster*. Springer Verlag, Berlin, Vol. VIII.

Campos-Ortega,J.A., Lehmann,R., Jimenez,F. and Dietrich,U. (1984) In Sharma,S.C. (ed.), *Organizing Principles of Neural Development*. Plenum Press, New York, pp. 129–144.

Carpenter,C. and Cohen,S. (1979) *Annu. Rev. Biochem.*, **48**, 193–216.

Devereux,J., Haekerli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387–395.

Dietrich,U. and Campos-Ortega,J.A. (1984) *J. Neurogenet.*, **1**, 315–332.

Doe,C.Q. and Goodman,C.S. (1985) *Dev. Biol.*, **111**, 193–205.

Doolittle,R.F., Feng,D.F. and Johnson,M.S. (1984) *Nature*, **307**, 558–560.

Foster,D. and Davie,E.W. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 4766–4770.

Gray,A., Dull,T.J. and Ullrich,A. (1983) *Nature*, **303**, 722–725.

Greenwald,I. (1985) *Cell*, **43**, 583–590.

Greenwald,I.S., Sternberg,P.W. and Horvitz,H.R. (1983) *Cell*, **34**, 435–444.

Haenlin,M., Steller,H., Pirrotta,V. and Mohier,E. (1985) *Cell*, **40**, 827–837.

Hartenstein,V. and Campos-Ortega,J.A. (1984) *Wilhelm Roux Arch. Dev. Biol.*, **193**, 308–325.

Hartenstein,V., Technau,G.M. and Campos-Ortega,J.A. (1985) *Wilhelm Roux Arch. Dev. Biol.*, **194**, 213–216.

Holmes,D.S. and Quigley,M. (1981) *Anal. Biochem.*, **114**, 193–197.

Hoppe,P.E. and Greenspan,R.J. (1986) *Cell*, **46**, 773–783.

Ingham,P.W., Howard,K.R. and Ish-Horowicz,D. (1985) *Nature*, **318**, 439–445.

Jürgens,G., Wieschaus,E., Nüsslein-Volhard,C. and Kluding,H. (1984) *Wilhelm Roux's Arch. Dev. Biol.*, **193**, 283–295.

Kidd,S., Lockett,T.J. and Young,M.W. (1983) *Cell*, **34**, 421–433.

Kidd,S., Kelley,M.R. and Young,M.W. (1986) *Mol. Cell. Biol.*, **6**, 3094–3108.

LaBonne,S.G. and Mahowald,A.P. (1985) *Dev. Biol.*, **110**, 264–267.

Lehmann,R., Jimenez,F., Dietrich,U. and Campos-Ortega,J.A. (1983) *Wilhelm Roux's Arch. Dev. Biol.*, **192**, 62–74.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Marquardt,H., Hunkapiller,M.W., Hood,L.E. and Todaro,G.J. (1984) *Science*, **223**, 1079–1082.

Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–560.

Pardue,M.L. (1985) In Hames,B.D. and Higgins,S.J. (eds), *Nucleic Acid Hybridization*. IRL Press, Oxford, pp. 179–202.

Patthy,L. (1985) *Cell*, **41**, 657–663.

Poncz,M., Solowiejczyk,D., Ballantine,M., Schwartz,E. and Surrey,S. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4298–4302.

Poole,S.J., Kauvar,L.M., Drees,B. and Kornberg,T. (1985) *Cell*, **40**, 37–43.

Poulson,D.F. (1937) *Proc. Natl. Acad. Sci. USA*, **23**, 133–137.

Poulson,D.F. (1950) In Demerec,M. (ed.), *Biology of Drosophila*. Wiley, New York, pp. 168–274.

Reed,K.C. and Mann,D.A. (1985) *Nucleic Acids Res.*, **13**, 7207–7221.

Russel,D.W., Schneider,W.J., Yamamoto,T., Luskey,K.L., Brown,M.S. and Goldstein,J.L. (1984) *Cell*, **37**, 577–585.

Savage,C.R.,Jr, Hash,J.H. and Cohen,S. (1973) *J. Biol. Chem.*, **248**, 7669–7672.

Sanger,F., Coulson,A.R., Barrell,B.G., Smith,A.J.H. and Roe,B.A. (1980) *J. Mol. Biol.*, **143**, 161–178.

Schuh,R., Aicher,W., Gaul,U., Cote,S., Preiss,A., Maier,D., Seifert,E., Nauber,U., Schröder,C., Kemler,R. and Jäckle,H. (1986) *Cell*, **47**, 1025–1032.

Scott,J., Urdea,M., Quiroga,M., Sanchez-Pescador,R., Fong,N., Selby,M., Rutter,W.J. and Bell,G.I. (1983) *Science*, **221**, 236–240.

Sternberg,P.W. and Horvitz,H.R. (1984) *Annu. Rev. Genet.*, **18**, 489–524.

Südhof,T.C., Goldstein,J.L., Brown,M.S. and Russell,D.W. (1985) *Science*, **228**, 815–822.

Sulston,J.E. and White,J.G. (1980) *Dev. Biol.*, **78**, 577–597.

Sulston,J.E., Schierenberg,E., White,J.G. and Thomas,J.N. (1983) *Dev. Biol.*, **100**, 64–119.

Technau,G.M. and Campos-Ortega,J.A. (1986) *Wilhelm Roux's Arch. Dev. Biol.*, **195**, 445–454.

Vässin,H. (1986) *Genetische Charakterisierung und molekulare Klonierung des Delta-Locus von Drosophila melanogaster*. PhD Thesis, University of Cologne.

Vässin,H., Vielmetter,J. and Campos-Ortega,J.A. (1985) *J. Neurogenet.*, **2**, 291–308.

Weigel,D., Knust,E. and Campos-Ortega,J.A. (1987) *Mol. Gen. Genet.*, in press.

Wharton,K.A., Yedvobnick,B., Finnerty,V.G. and Artavanis-Tsakonas,S. (1985a) *Cell*, **40**, 55–62.

Wharton,K.A., Johansen,K.M., Xu,T. and Artavanis-Tsakonas,S. (1985b) *Cell*, **43**, 567–581.

Willimzig,M. (1985) *Trends Genet.*, **1**, 158.

Received on November 11, 1986; revised on January 9, 1987