The EMBO Journal vol.14 no.14 pp.3487–3495, 1995

vnd, a gene required for early neurogenesis of *Drosophila*, encodes a homeodomain protein

Fernando Jiménez[1], Linda E.Martin-Morris[2,3], Leonardo Velasco, Hsin Chu[4,5], Javier Sierra, Daniel R.Rosen[2,6] and Kalpana White[2,4]

Centro de Biología Molecular 'Severo Ochoa', Universidad Autónoma/CSIC, 28049 Madrid, Spain, [2]Biology Department, Brandeis University, Waltham, MA 02254, USA, [4]Biochemistry Department, Brandeis University, Waltham, MA 02254, USA and [5]Volen National Center for Complex Systems, Brandeis University, Waltham, MA 02254, USA
[3]Present address: Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, WA 98104 , USA
[6]Present address: Wadsworth Center for Laboratories and Research, Empire State Plaza, Albany, NY 12201-0509, USA

[1]Corresponding author

The development of the central nervous system in *Drosophila* is initiated by the segregation of neuroblasts, the neural progenitors, from the embryonic neuroectoderm. This process is guided by at least two classes of genes: the *achaete–scute* complex (*AS*-C) proneural genes and the neurogenic genes. It has been known for some time that loss-of-function mutations in the *AS*-C result in neural hypoplasia and the first observed defect is failure of segregation of a fraction of neuroblasts. Loss-of-function mutations at the *ventral nervous system defective* (*vnd*) locus are known to lead to similar phenotypic defects in early neurogenesis. More recently, the *vnd* locus has been implicated in the regulation of the proneural *AS*-C genes and the neurogenic genes of the *Enhancer of split* complex. In this paper we report the identification of a transcript associated with the *vnd* locus, the transcript distribution in embryogenesis, which is compatible with the nervous system mutant phenotypes described for this gene, and that the protein product is a member of the NK-2 homeodomain family. We discuss these findings within the framework of early *Drosophila* neurogenesis and the known phenotypes associated with the *vnd* locus.
Key words: *Drosophila* neurogenesis/neuroblasts/NK-2 homeodomain/prepattern genes/proneural genes

Introduction

During *Drosophila* embryogenesis, neuroblasts, the progenitors of central nervous system (CNS) neurons, delaminate individually from the neural ectoderm. This process occurs in five temporal waves, S1–S5, over a 3 h period, and results in an invariant spatio-temporal pattern of neural stem cells (Hartenstein and Campos-Ortega, 1984; Doe, 1992). It is now recognized that each individual neuroblast is singled out from an equivalent group of cells termed 'proneural cluster' within the neuroectoderm (reviewed in Campos-Ortega, 1993). A large body of work has elucidated a pivotal role for the proneural genes of the *achaete–scute* complex (*AS*-C), which encode basic-helix–loop–helix proteins, in the formation of proneural clusters and in the subsequent delamination of the neuroblasts (reviewed in Campuzano and Modolell, 1992; Campos-Ortega, 1993). Studies on the early waves of neuroblast segregation show that all cells of the S1 and S2 proneural clusters express *AS*-C genes. As development proceeds, a single cell within each cluster, the presumptive neuroblast, delaminates and continues to express the proneural genes (Cabrera *et al.*, 1987; Romani *et al.*, 1987; Martín-Bermudo *et al.*, 1991; Skeath and Carroll, 1992; Ruiz-Gómez and Ghysen, 1993). The remaining cells of the cluster cease to express proneural genes and begin expressing genes of the *Enhancer of split* gene complex [*E(spl)*-C] which also encodes several basic-helix–loop–helix proteins (Knust *et al.*, 1992; Jennings *et al.*, 1994).

The first indication that additional proneural genes exist came from the studies conducted by Jiménez and Campos-Ortega (1987, 1990). They observed that the entire deletion of *AS*-C or mutations at the vital locus *ventral nervous system defective* (*vnd*) (White, 1980; White *et al.*, 1983) result in elimination of 25% of the embryonic neuroblast complement and that the loss of neuroblasts in each case preferentially occurs in different spatial domains. Further, in embryos that are deleted for *AS*-C as well as the *vnd* locus, elimination of up to 50% of the neuroblasts was observed. These observations led to the proposal that *vnd* encodes a function that is important for the commitment of a neuroblast population that is largely distinct from that dependent on the *AS*-C. More recently the role of *vnd* during proneural cluster formation has been further probed. Prior to the segregation of S1 neuroblasts, each hemisegment of the neuroectoderm can be described, using the convention of Skeath *et al.* (1994), as a three-by-four matrix of proneural clusters consisting of medial, intermediate and lateral columns and A, B, C, D transverse rows. During studies on genetic regulatory mechanisms that govern *AS*-C expression in S1 and S2 neuroblasts, Skeath *et al.* (1994) discovered that *vnd* was essential for the *AS*-C expression in every alternate proneural cluster of the medial column of S1 neuroblasts. Furthermore, they proposed that *vnd* is required at two distinct steps. First, in the formation of the medial cluster; and second, for the maintenance of proneural gene expression in the presumptive neuroblast. *vnd*-effected regulation in the medial neuroectoderm was also demonstrated for two *E(spl)*-C-*lacZ* reporter genes (Kramatschek and Campos-Ortega, 1994). The *E(spl)*-C-*lacZ* reporter gene expression is severely reduced over the entire neuroectoderm in embryos carrying a deletion of *AS*-C [*Df(1)sc*^{19}]; whereas, only the expression in the medial neuroectodermal clusters is missing in *vnd* mutant embryos. Moreover, in

Df(1)RT184 embryos that are deleted for *AS*-C and *vnd* (in addition to the known genes *svr* and *elav*), neuroectodermal expression of *E(spl)*-C-*lacZ* is totally abolished, suggesting a synergistic effect in the deletion that includes both *AS*-C and *vnd*. Together these studies make a strong case that *vnd* directly or indirectly regulates gene expression, at least of *AS*-C and *E(spl)*-C genes, during segregation of S1 neuroblasts (Kramatschek and Campos-Ortega, 1994; Skeath *et al.*, 1994).

Identification of the protein product of the *vnd* gene is an essential prerequisite to illuminating how *vnd* acts in neurogenesis, how it activates *AS*-C and *E(spl)*-C genes in specific domains, and whether it regulates *AS*-C or *E(spl)*-C directly or indirectly. We report in this paper the delimitation of the *vnd* genomic region, sequence analysis of a cDNA derived from this region which most likely encodes the VND protein, and *in situ* localization of the transcript in wild type and *vnd* mutant embryos. The deduced protein product of the cDNA is a homeodomain protein. A partial sequence of this protein was previously deduced by Kim and Niremberg (1989) from a genomic sequence isolated in a search for new homeodomain proteins in *Drosophila*, and the predicted novel homeodomain for this sequence was named NK-2. We provide evidence that strongly supports the notion that this homeodomain protein is encoded by the *vnd* gene.

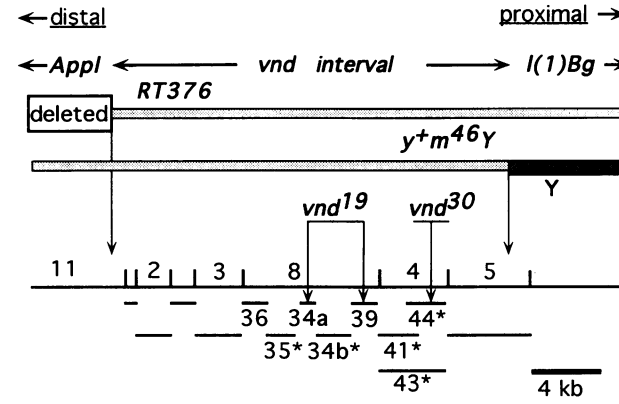

Fig. 1. Schematic representation of the genomic *vnd* region. The solid line represents the DNA. The *Eco*RI sites and the sizes of the fragments are indicated just above the DNA line. Just below it, the subclones of the genomic DNA used in this study are indicated. *Appl* and *l(1)Bg* represent the nearest known flanking loci; distal and proximal indicate the direction of telomere and centromere respectively. The top line indicates the extent of the terminal deletion *Df(1)RT376* which is $Appl^{-}vnd^{+}$ and the line below it indicates the extent of chromosome duplicated in $y^{+}m^{46}Y$ which is vnd^{+} $l(1)Bg^{-}$. Also indicated are breakpoints associated with vnd^{19} and vnd^{30}, the two alleles associated with chromosomal rearrangements. The subclones 35, 34b, 43, 41 and 44 that are recognized by the cDNA are marked with an asterisk.

Results

Genomic interval that includes vnd

The *vnd* locus is flanked by the *Amyloid precursor protein-like* (*Appl*) locus distally and the *l(1)Bg* locus proximally within the salivary segment 1B9-11, at the tip of the X-chromosome (Rosen *et al.*, 1989; Martin-Morris and White, 1990; Lindsley and Zimm, 1992). We have refined the previously published *vnd* genomic interval (Rosen *et al.*, 1989) by analyzing additional chromosomal rearrangements that flank the *vnd* gene but still retain the *vnd* function. In particular two important breakpoints were defined: a terminal deletion, *Df(1)RT376*, which is $Appl^{-}$ but vnd^{+}, and an X chromosome translocation, $y^{+}m^{46}Y$, which is vnd^{+} but $l(1)Bg^{-}$ (Figure 1) (Martin-Morris, 1991). These two breakpoints border ~23 kb of genomic DNA within which the *vnd* locus must reside. In this interval a breakpoint of a hypomorphic mutation, vnd^{30}, which is associated with a chromosomal rearrangement, was mapped within the 4 kb *Eco*RI fragment just distal to the proximal 5 kb *Eco*RI fragment (Figure 1) (Rosen, 1987; Martin-Morris, 1991). Within the 23 kb interval we also mapped vnd^{19}, a strong embryonic lethal mutation which is also associated with a chromosomal rearrangement. In this case the lesion at the molecular level is more complex and two breakpoints have been mapped, one in each of the two fragments 34a and 39 (Figure 1).

Analysis of a transcriptional unit from the vnd genomic interval

To determine the transcriptional activity of this region, we performed *in situ* hybridization to wild type embryos using individual DNA subclones that covered ~17 kb out of the proximal ~19 kb of the genomic interval that defines *vnd* (Figure 1). A reliable hybridization signal localized to the ventral-most neuroectoderm was found with subclones 35, 43 and the two subclones of 43 (41 and 44). Subclones 41 and 44 share an overlap of ~0.5 kb.

The subclone 43 was used to probe an embryonic cDNA library and a single clone containing an ~3 kb cDNA insert was isolated. *In situ* hybridization with an antisense riboprobe generated from the 3 kb cDNA reproduces entirely the pattern observed previously with genomic fragments. Comparision of the cDNA sequence (see below) with the genomic sequence of subclone 41 shows that the 1417 bp of 3′ end of the cDNA are entirely contained within this subclone. The *Pst*I site which marks the distal end of subclone 44 is at position 2594 in the cDNA, thus, the very 3′ 335 bp of the cDNA are in the ~0.5 kb region shared between the subclones 41 and 44 (Figure 1). The cDNA also recognizes subclones 35 and 34b by Southern blotting and restriction digest mapping of subclone 35 and comparision with the restriction sites found in the cDNA sequence suggests that the 5′ end of the cDNA is in this subclone. These results indicate that the corresponding transcript is encoded within the *vnd* region and that the direction of the transcription is distal to proximal.

Developmental expression

Since *vnd* mutant embryos show defects in the first waves of neuroblast segregation we studied the transcriptional pattern of this gene during embryogenesis by *in situ* hybridization, using a riboprobe generated from the cDNA (Figures 2 and 3). Expression is first observed in the blastoderm as two longitudinal stripes that flank the prospective mesoderm (Figures 2A and 3A). A transient pair-rule modulation is evidenced in the staining during stage 6 (Figure 3B), but soon the continuity of the stripes is restored when they join at the midline, following mesoderm invagination (Figure 3D). Simultaneously, the anterior-most portion of each stripe begins to resolve in the head region as a distinct cell cluster whose fate in

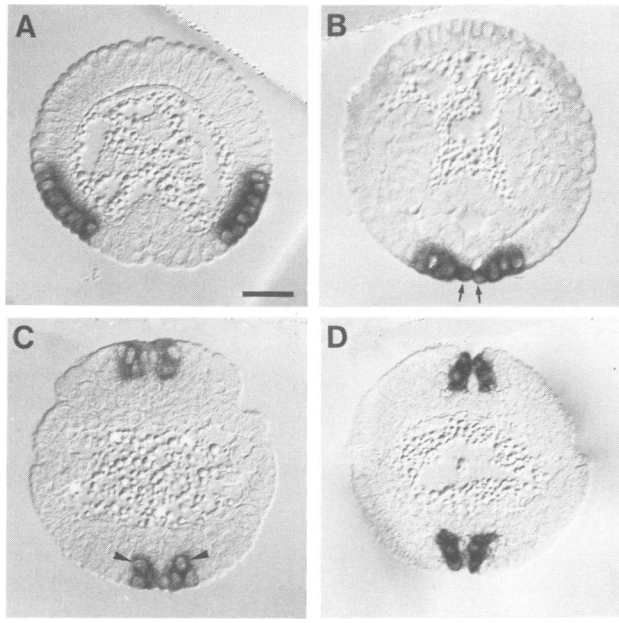

Fig. 2. Expression of the *vnd* transcript visualized in transverse sections of wild type embryos. (**A**) At the beginning of gastrulation two anterio-posterior stripes are observed. Each stripe is seven to eight cells wide along the dorso-ventral axis. (**B**) During stage 8, the two stripes fuse at the ventral midline. Each stripe now consists, along the anterior-posterior axis, of one mesectodermal cell (arrow) and three to four neighboring neuroectodermal cells from which S1 neuroblasts of the medial column delaminate at the end of this stage. (**C**) In stage 9, transcription is detected in medial neuroblasts (arrowheads) while the width of the stripe in the neuroectoderm is restricted to two to three cells. (**D**) In early stage 11, transcription at the midline in mesectodermal cells, which have already delaminated into the neuroblast layer, is barely detected and in the neuroectodem has been further restricted to the ventral-most cell row. Within the CNS, high levels of expression are found in medial neuroblasts and in ganglion mother cells and neurons overlying them. Scale bar: 20 μm.

subsequent stages has not been further followed. Due to cell rearrangements associated with germ-band extension (Irvine and Wieschaus, 1994), the width of each stripe is reduced by half, from about eight cells at the blastoderm (Figures 2A and 3A) to about four cells at stage 8 (Figures 2B and 3D). One of these cells forms part of the mesectoderm and the remaining cells constitute the ventral-most region of the neuroectoderm (Figure 2B), where S1 medial proneural clusters form during embryonic stages 7 and 8 (Martín-Bermudo *et al.*, 1995). In stage 9, after segregation of S1 and S2 neuroblasts, signal is seen only in those of the medial column and the width of the neuroectodermal stripes begins to diminish (Figure 2C). During stage 10, when S3 neuroblasts segregate, the width of the stripes is further reduced and concomitantly gaps of non-expressing cells first appear (Figure 2F). In early stage 11, expression in the CNS is mostly confined to medial neuroblasts and their presumptive progeny (Figure 2D), although a lateral expansion of the domain of transcription becomes evident (Figure 3G). The pattern evolves during subsequent stages towards a series of broad transverse rows (Figure 3I) and, by stage 16, transcription begins to decline and is practically restricted to a single row of neural cells per segment (Figure 3J). In addition to the nervous system, expression is also observed beginning at germ-band retraction in anterior and posterior midgut (Figure 3I and J).

The spatial and temporal aspects of expression of this cDNA are consistent with the known phenotypes associated with the *vnd* mutations. The transcript is expressed early enough to account for the defects associated with S1 and S2 neuroblast segregation and perturbations associated with the expression of *AS*-C and *E(spl)*-C genes (Jiménez and Campos-Ortega, 1990; Kramatschek and Campos-Ortega, 1994; Skeath *et al.*, 1994). Moreover, the spatial domains of expression largely correspond with the associated defects, except with those which affect the intermediate column of neuroblasts (see Discussion).

Expression in vnd mutants

We studied transcription patterns in a number of *vnd* mutant alleles. For most alleles studied, vnd^{2}, vnd^{3}, vnd^{6}, vnd^{24}, vnd^{27} and vnd^{28}, all embryonic lethals, roughly the same modulation of the pattern of wild type signal was observed as exemplified for vnd^{6} in Figure 3. Activation of the gene at the blastoderm stage seems normal, but soon a substantial loss of signal in the longitudinal stripes is readily observed at gastrulation (Figure 3C). The signal in the neuroectoderm and CNS in subsequent stages of development is also greatly diminished (Figure 3E and H). Expression in vnd^{19} embryos is undetectable (data not shown), as would be expected from the complex rearrangement observed in this mutant. In the case of vnd^{30} no significant differences were observed between mutant and wild type embryos.

Analysis of the open reading frame

The nucleotide sequence of the 3 kb cDNA was determined. Sequence analysis revealed a single large open reading frame (ORF) with a first methionine at position 342 and a stop codon at position 2508 (Figure 4A). The sequence preceding the first and second methionine codons at positions 342 and 747 is CGATATG and CACCATG respectively. The second methionine codon shows a closer match to the *Drosophila* translational initiation sequence c/aAAa/cAT (Cavener, 1987). Assuming that the first methionine codon is used for the translational start, the conceptual translation of this sequence yields a 722 amino acid polypeptide.

A search for protein homology revealed that part of this protein sequence was identical to a known partial protein sequence from *Drosophila* that predicted a new homeodomain: NK-2 (Kim and Nirenberg, 1989). The corresponding genomic fragment was mapped to salivary band region 1C1-5 by chromosomal *in situ* hybridization (Kim and Nirenberg, 1989), close to the cytogenetic map position of *vnd* in 1B9 (Lindsley and Zimm, 1992). The NK-2 homeodomain protein family now comprises proteins from mouse, *Xenopus*, *Caenorhabditis elegans*, leech and planaria (Figure 4B) (García-Fernández *et al.*, 1991; Price *et al.*, 1992; Lints *et al.*, 1993; Nardelli-Haefliger and Shankland, 1993; Saha *et al.*, 1993; Okkema and Fire, 1994). In addition to the homeobox and the acidic domain originally pointed out for NK-2 (Kim and Nirenberg, 1989), many members of the family also have a second conserved domain of 24 amino acids called the NK-2 box (Saha *et al.*, 1993) (Figure 4A). Several stretches rich in alanines are observed (Figure 4A).

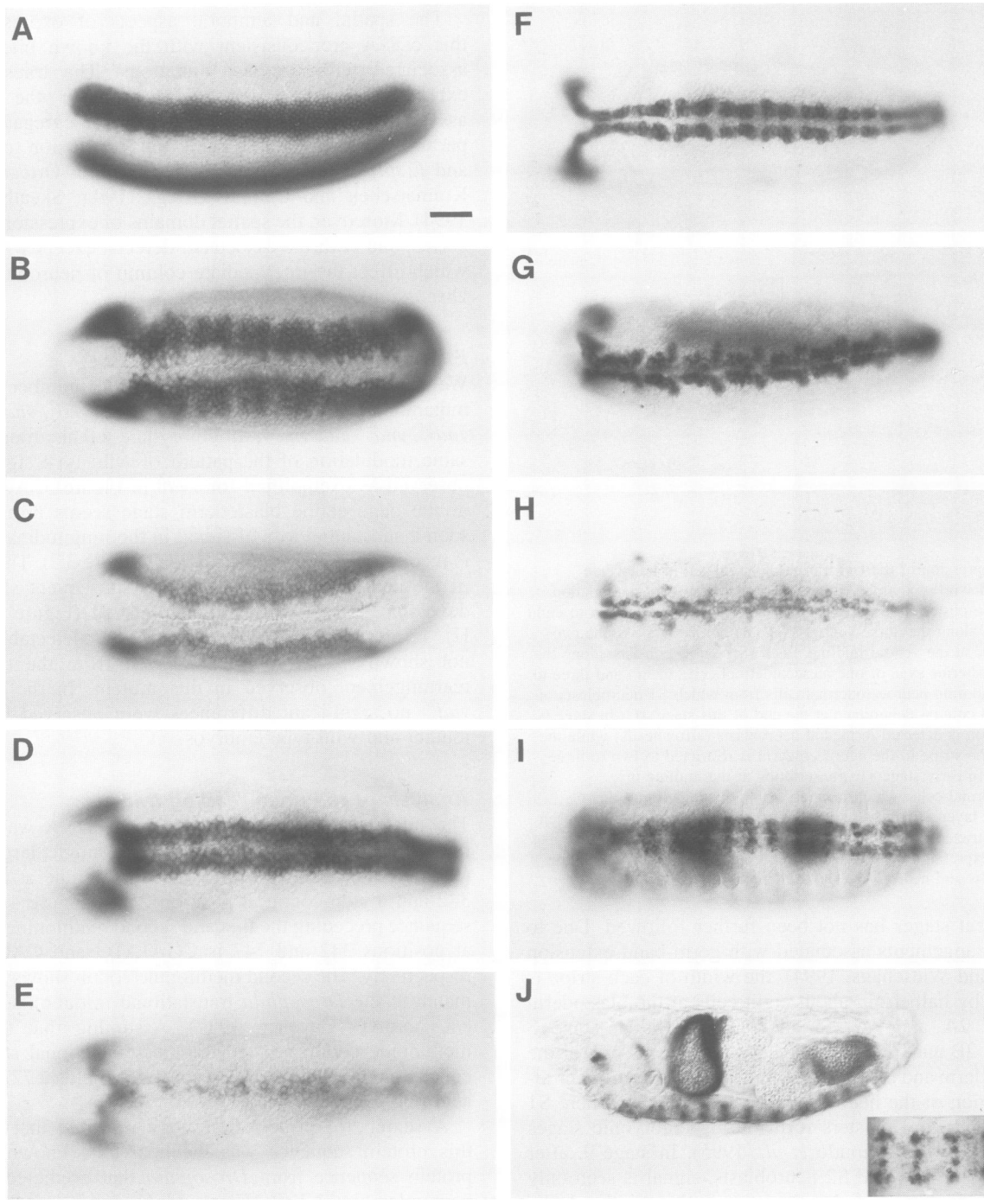

Fig. 3. Embryonic expression of the *vnd* transcript in whole mounts of staged wild type and *vnd* mutant embryos. Embryo progeny of the $vnd^{6}/FM7$ stock were hybridized to a riboprobe, and mutant embryos and their siblings with vnd^{+} allele were analyzed. The siblings carrying vnd^{+} alleles show patterns identical to those observed in wild type embryos. Panels A, B, D, F, G, I and J depict wild type (sibling carrying a vnd^{+} allele) and C, E and H depict mutant (vnd^{6}) embryos. All embryos are orientated with anterior to the left, and ventral views are shown, except when indicated. (**A**) Expression at the blastoderm stage (ventro-lateral view) is restricted to two ventro-lateral stripes that extend along most of the length of the embryo. (**B**) During stage 6 a rough pair-rule modulation is observed in the wild type. (**C**) The mutant embryos are readily identified at this stage by the lower level of transcription. (**D**) During stage 7/early stage 8, the two stripes fuse at the ventral midline forming the mesectoderm and the ventral-most part of the neuroectoderm. (**E**) In contrast to what is seen in D for wild type embryos, the expression is barely detectable in a comparable stage mutant embryo. (**F**) During stage 10, the width of the stripes in the neuroectoderm decreases and gaps of two or three non-expressing cells appear at a segmental periodicity. (**G**) In an early stage 11 wild type embryo, focused at the level of the developing CNS, expression is mainly detected in medial neuroblasts and their progeny, and, with a segmental periodicity, more lateral expression is seen in a small cluster of neural cells. (**H**) Notice the low levels of transcription in the stage 11 mutant embryo as compared with the wild type. (**I**) During stage 13, the transcript in the CNS is becoming restricted to stripes of cells perpendicular to the anterior-posterior axis. Out of focus, the transcript is also detected in the anterior and posterior areas of the midgut. (**J**) In stage 16 (lateral view), expression is lower in the CNS and further restricted to discrete cell clusters (see inset, ventral view), but persists in the midgut. Scale bar: 50 μm.

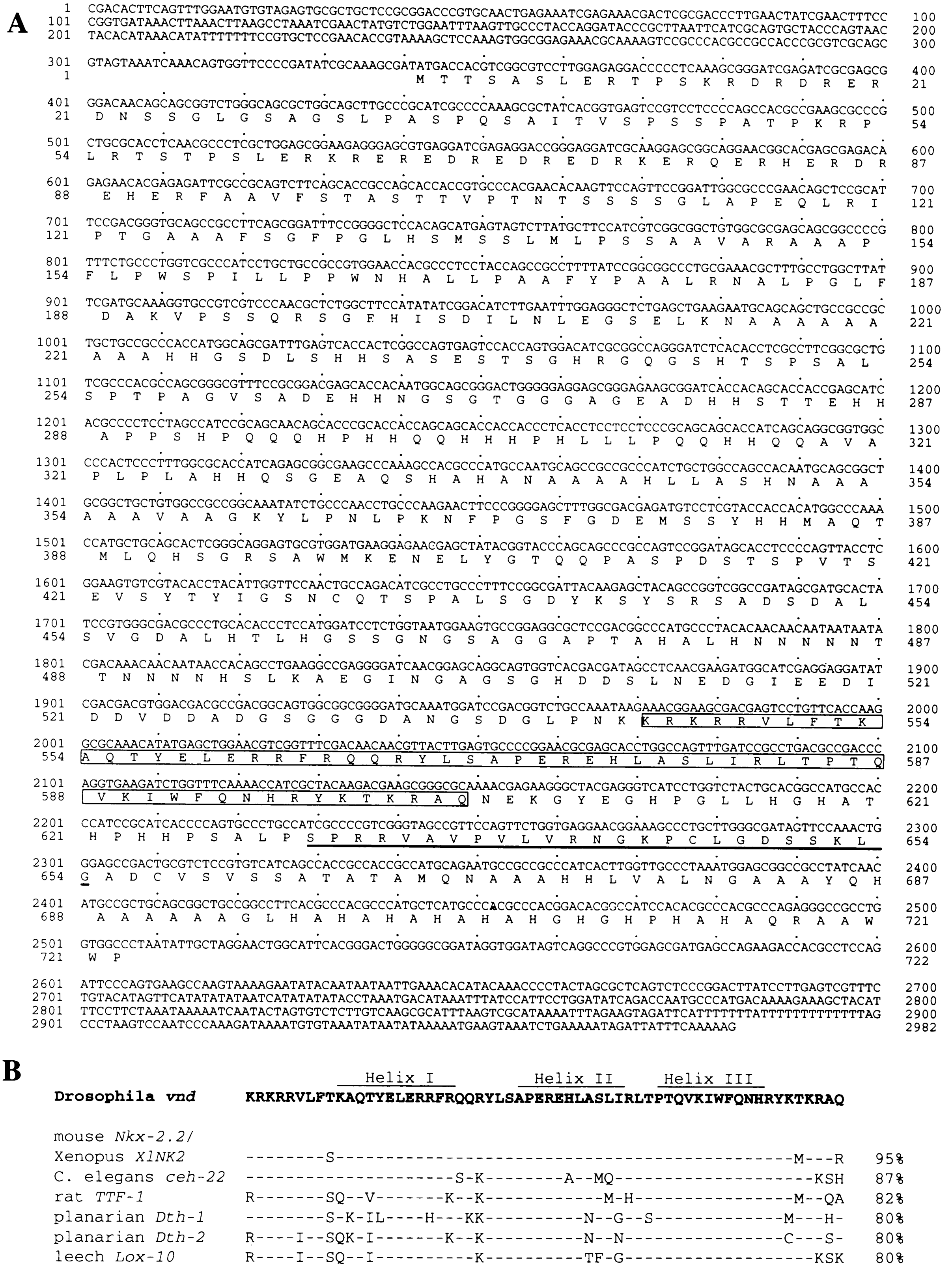

Fig. 4. Nucleotide and predicted amino acid sequence of the ORF of the *vnd* cDNA. (**A**) DNA sequence and predicted protein sequence of the only large open reading frame are depicted. The numbering for the protein sequence begins at ATG_{342} and ends before the stop codon at TAA_{2509}. The sequence between amino acids 476 and 631 is identical to a partial sequence previously published by Kim and Nirenberg (1989). As indicated by these authors, it contains a homeodomain (boxed sequence) and an acidic region between residues 508 and 527. In addition, it has a NK-2 box (underlined) conserved between several NK-2 homeodomain proteins. The protein also has three alanine-rich regions between residues 215 and 223 (9/9), 336 and 359 (15/24) and 670 and 719 (22/49). (**B**) Comparison between homeodomains of NK-2 domain proteins from different species. Identical amino acids are indicated by '–' Percentage of amino acid identity with the VND homeodomain as indicated.

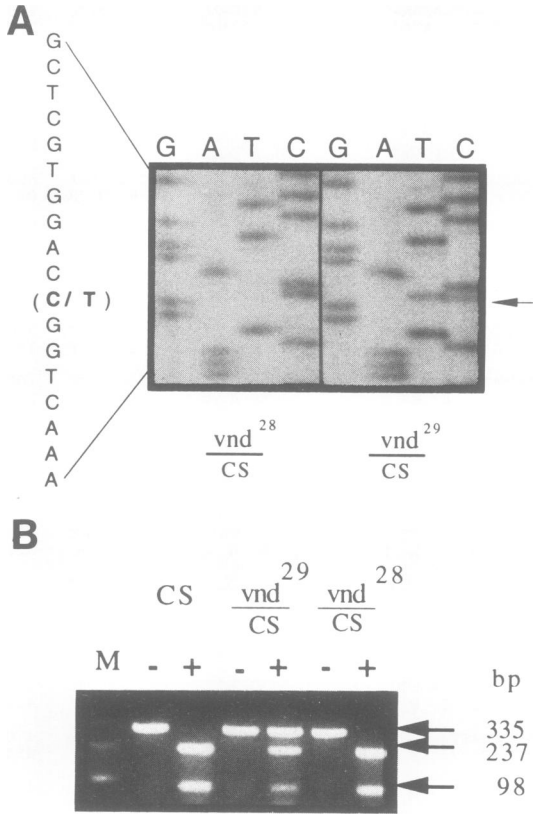

Fig. 5. *vnd*^{29} has a base pair substitution within the homeodomain. **(A)** The left four lanes show the DNA sequence of the antisense strand in *vnd*^{28}*/CS* DNA and the right four lanes show corresponding sequence in *vnd*^{29}*/CS* DNA. Notice the band at the same position in both C and T lanes in the *vnd*^{29}*/CS* DNA (indicated by an arrow). **(B)** The 335 bp band was PCR amplified from genomic DNAs from *CS*, *vnd*^{29}*/CS* and *vnd*^{28}*/CS* and cleaved with *Hae*III. '–' denotes the lanes with the uncleaved PCR product, '+' denotes the lanes in which the PCR product was digested with *Hae*III. All three DNAs yield the 237 and 98 bp fragments, but in *vnd*^{29}*/CS* the 335 bp uncleaved band from the mutant DNA is evident.

A vnd mutation has an amino acid substitution within the homeodomain

We examined whether some of the *vnd* mutations mapped within the sequence that spans the homeodomain and the NK-2 box (nucleotides 1837–2478). To detect mutations, DNA from heterozygous mutant flies of genotypes *vnd*^{mutant}/+ was amplified by PCR, melted and reannealed. In all, six *vnd* mutants, *vnd*^{2}, *vnd*^{3}, *vnd*^{5}, *vnd*^{6}, *vnd*^{28} and *vnd*^{29}, were analyzed. Reannealed DNA was subjected to heteroduplex analysis as described in Materials and methods. DNA from *vnd*^{29}*/CS* flies revealed additional bands (data not shown). The region was further narrowed down using similar methodology to a 335 bp fragment (nuceotides 1837–2172). Further sequencing of *vnd*^{29}*/CS* DNA and the control *vnd*^{28}*/CS* DNA revealed a G to A change at nucleotide position 2073 (Figure 5A) in *vnd*^{29} but not in *vnd*^{28}, an allele which was induced in the same parental chromosome, using nitroso urea as a mutagen which favors single nucleotide substitutions. This change in *vnd*^{29} altered an *Hae*III site, which was further confirmed by restriction digests of PCR amplified DNA from *vnd*^{29}*/CS*, *vnd*^{28}*/CS* and *CS* (Figure 5B).

The mutation in *vnd*^{29} leads to an A_{578}–T_{578} amino acid substitution in the second helix of the homeodomain. This position is conserved in mouse *NKx-2.2*, rat *TTF-1*, *Xenopus XINc2* and *C.elegans ceh-22*, but not in planaria *Dth1* and *Dth2* (N) and the leech *Lox-10* (T) (Figure 4B). Although a T is observed in the leech *Lox-10* homeodomain, we suspect that in the context of the entire homeobox this substitution is likely to be relevant, and impair the function of the VND protein. We propose that it is this change that is likely to be responsible for the mutant phenotype.

Discussion

In this paper we have presented a molecular characterization of the *Drosophila vnd* locus. We delimited the genomic region in which *vnd* must reside, identified and sequenced a transcript from the region, determined its pattern of expression in the wild type and mutant embryos and molecularly characterized three mutant *vnd* alleles. Our results, when viewed in light of existing information, make a strong case that the transcript we have identified is encoded by the *vnd* locus for the following reasons:

(i) The cDNA is derived from the genomic region in which *vnd* must reside, which is defined by the breakpoint of *Df(1)RT376* and the distal breakpoint of the duplication $y^{+}m^{46}Y$. Moreover, we have previously mapped seven other X chromosome terminal deletions that are genetically *vnd*$^{-}$ *l(1)Bg*$^{+}$; three map to the 8 kb *Eco*RI genomic fragment (Rosen *et al.*, 1989) and four others map proximal to the 8 kb *Eco*RI fragment (Rosen, 1987).

(ii) Analysis of a lethal mutation, *vnd*^{19}, shows that the transcription unit is disrupted by the genomic lesion in this mutation. Additionally *vnd*^{30}, a hypomorphic mutation which partially complements some of the *vnd* alleles (Rosen, 1987), maps just 3′ to the cDNA. The genomic mapping of the 1B breakpoint of *vnd*^{30} suggests that the entire ORF is retained in the *vnd*^{30} chromosome and thus it is in accord with the hypomorphic nature of this allele.

(iii) The expression pattern of the transcript in early embryogenesis is compatible with the known early phenotypes observed in *vnd* mutations. The first reported defect in *vnd* embryos is loss at stage 10 of a subpopulation of neuroblasts restricted to the medial and intermediate columns (Jiménez and Campos-Ortega, 1990). Later, it was shown that, after segregation of S1 NBs at late stage 8, medial column neuroblasts of rows B and D are lost reliably and a few neuroblasts of row A and C are lost in medial and intermediate columns (Skeath *et al.*, 1994). The transcript expression is restricted to the medial column where the earliest and most consistent effects of *vnd* mutation are observed.

(iv) Recent data have implicated *vnd* as a regulator of expression of *AS*-C and *E(spl)*-C genes in specific domains (Kramatschek and Campos-Ortega, 1994; Skeath *et al.*, 1994). Furthermore, the product encoded by *vnd*, a homeodomain protein, is befitting for a gene implicated in the regulation of other genes. This protein was initially identified by Kim and Nirenberg (1989) as a *Drosophila* NK-2 homeodomain protein in a search for genes encoding novel homeodomain proteins. We have identified an amino acid substitution, A_{578} to T_{578}, in the second helix of the homeodomain in a lethal mutant allele which could compromise the function of the mutant protein.

(v) *In situ* hybridization studies in the mutants attest to the transcript being derived from the *vnd* gene. A transcriptional defect is seen in embryos for all lethal alleles tested: transcript level is low in the mutant embryo from stage 6 and continues to be consistently lower in the ensuing stages. Importantly, in vnd^{19}, no hybridization signal was observed, suggesting that this rearrangement, which disrupts the transcriptional unit, is a transcript null. In contrast, the signal in the vnd^{30} embryos is indistinguishable from the wild type as may be expected from a hypomorphic allele. That all tested lethal alleles show a very similar defect in mutant embryos in *in situ* hybridization suggests that the *vnd* gene autoregulates or that a *vnd*-regulated gene may in turn control *vnd* expression. In support of the autoregulation postulate, Nirenberg and colleagues have determined a consensus binding site for the NK-2 homeodomain T(T/C)AAGT(G/A)G (Tsao *et al.*, 1994) and observed 25 high affinity binding sites within the 2.2 kb 5′ region to the NK-2 transcription start site (M.Nirenberg, personal communication).

What functions can be assigned to the VND protein? We suggest that one function of *vnd* is that of a prepattern gene, namely, one component of the genetic system that establishes coordinates for the activation of the proneural clusters (Ghysen and Dambly-Chaudière, 1989; Skeath *et al.*, 1992), a possibility initially raised by Skeath *et al.* (1994). *vnd* would act as an intermediate gene that translates, to specifically activate proneural genes in the ventral-most neuroectoderm, the positional coordinates previously established by other dorso-ventral patterning genes in the early embryo. These latter genes would be primary and *vnd* a secondary prepattern gene. A recent demonstration by Mellerick and Nirenberg (1995) that the expression of NK-2 homeobox gene (which this paper identifies as the *vnd* locus) is regulated by dorso-ventral patterning genes like *dorsal*, *snail*, *twist*, and *decapentaplegic*, supports this latter suggestion. In its role as a prepattern gene, *vnd* would act in concert with other genes for, although *vnd* is expressed along the entire medial neuroectoderm during stages 7 and 8, it is essential for only every other S1 medial neuroblast. Other genes, probably the pair-rule (Skeath *et al.*, 1992), would suffice to set up *AS*-C expression in the medial proneural clusters that do not require *vnd*.

Two important questions remain unanswered: are there other genes that serve similar functions in the medial and lateral regions of the neuroectoderm, and how does *vnd* affect the neuroblasts in the medial column where it is not expressed? The observation by Jiménez and Campos-Ortega (1990) that S1 and S2 neuroblasts of the intermediate column are frequently lacking in *vnd* mutants, although the gene is not expressed in the corresponding neuroectodermal domain during the relevant period, underscores the likely complexity of the *vnd* function in the neuroectoderm and suggests that another component of the function could be regulation of an inductive signal that would explain the non-autonomy of action. This kind of phenomenon has a precedent in *wingless*, which is required for either the formation or correct determination of neuroblasts derived from neuroectodermal anterior-posterior stripes adjacent to the *wingless*-expressing stripe (Chu-LaGraff and Doe, 1993).

It is possible that *vnd* may be involved in the maintenance of certain neurons, as enhanced cell degeneration within the neural primordium has been observed in the *vnd* mutants (Jiménez and Campos-Ortega, 1990). Two factors may contribute to the enhanced cell degeneration: inductive degeneration caused due to absence of certain neurons and/or failure of cells to realize or maintain specific cell fates. Further studies are needed to clarify the significance of *vnd* expression after neuroblast segregation.

Materials and methods

Fly stocks

The following *vnd* alleles were used: vnd^{2}, vnd^{3}, vnd^{5} and vnd^{6} are described in White *et al.* (1983) and vnd^{24}, vnd^{27}, vnd^{28} and vnd^{29} are described in Lindsley and Zimm (1992). In addition, two chromosomal rearrangements associated with *vnd* mutations, the lethal allele vnd^{19} *[T(1;3)1B6;88A]* (Lindsley and Zimm, 1992) and the hypomorphic allele vnd^{30} *[T(1;2L; Inv1B-4F;25E)]*, both from the collection of Dr Lefevre, were used. vnd^{30} was genetically characterized by Rosen (1987). The wild type stock for the PCR analysis was Canton Special (*CS*).

The terminal *Df(1)RT* deletions used in these experiments have previously been described (Mason *et al.*, 1986). The duplication $y^{+}m^{46}Y$ chromosome was identified by screening a series of deletions, generated in the duplication $y^{+}scY$ generously donated by R.Voelkar, for breakpoints just distal to *vnd* by complementation tests with lethal mutations in *vnd* and *l(1)Bg*.

Breakpoint mapping

Genomic DNA from adult flies was extracted, DNA blots were prepared and hybridized with radioactive probes, prepared from subclones from the *vnd* genomic region, according to standard methods previously described (Campos *et al.*, 1987). Appropriate control DNAs for the mapping of *Df(1)RT* chromosomes and of duplication $y^{+}m^{46}Y$ consisted of parental chromosomes in which these chromosomes were induced (Campos *et al.*, 1987). To define the breakpoints in vnd^{19}, DNA from genotypes vnd^{19}/*FM7*, *FM7*, nrg^{14}/*FM6* and *FM6* was analyzed. nrg^{14}, used as a control, is a *neuroglian* (but not *vnd)* mutant allele generated on the same parental chromosome as vnd^{19}. To define vnd^{30}, DNAs from $vnd^{30}/y^{+}scY$, *Df(1)y74k24.1*/$y^{+}scY$ and Amherst M56i were used; Amherst M56i is the parental chromosome for vnd^{30} and *Df(1)y74k24.1* is a terminal deletion that deletes upto *l(1)Bi*.

In situ hybridization

In situ hybridization followed essentially the protocol of Tautz and Pfeifle (1989). The antisense riboprobe was generated from the 1.5 kb fragment at the 3′ end of the cDNA. DNA and RNA probes were labelled with digoxigenin with the DIG DNA and RNA labelling kits from Boehringer. After the staining reaction, embryos were embedded in Epon as indicated previously (Martín-Bermudo *et al.*, 1995). Occasionally, embryos were sectioned at 5 μm. Staging was according to Campos-Ortega and Hartenstein (1985).

Isolation of cDNA and sequencing of DNA

A cDNA library made in λ-ZAP with poly(A)$^{+}$ RNA from 0–12 h old embryos (a gift from Sergio González-Crespo) was screened with DIG-labelled probes made from subclone 43. To sequence the cDNA, the cDNA insert was subcloned into pBlueScript KS(+) (Stratagene) and both strands were sequenced by the dideoxy chain termination method (Sanger *et al.*, 1977). Sequenase Version 2.0 Sequencing kit (Amersham) or the fmol DNA Sequencing system (Promega) was used according to manufacturer's instructions. Custom-made oligonucleotides were used. The sequence was analyzed using the University of Wisconsin GCG software package (Devereux *et al.*, 1984).

PCR amplification and mutation detection

Genomic DNA isolated from flies heterozygous for *vnd* alleles was used as template for PCR amplification of a 335 bp genomic fragment containing the homeodomain. The two 19mer primers used were 5′-GGATCAACGGAGCAGGCAG-3′ and 5′-ACCCTCGTAGCCCTTC-TCG-3′. For PCR amplification, ~200 ng of genomic DNA were used as template, and the amplification conditions were 95°C for 2 min, followed by 95°C for 30 s, 57.5°C for 30 s and 72°C for 30 s for 30 cycles, in GeneAmp PCR system 9600 (Perkin Elmer Cetus). Amplifica-

tion buffer contained 1.2 mM $MgCl_2$ and 1 U of *Taq* DNA polymerase (Promega), in a total volume of 50 µl. Aliquots of 1.0 µl were analyzed by 3% agarose gel electrophoresis and 10 µl of the sample were prepared for heteroduplex analyses according to the manufacturer's recommendations (MDE; AT Biochem). Heteroduplexes were visualized by staining with ethidium bromide. The double-stranded PCR fragment with altered band mobility detected in the heteroduplex analysis was subjected to direct dideoxy sequencing. The detected mutation was verified by restriction enzyme digestion. The PCR fragments from different *vnd* alleles were cleaved by *Hae*III and electrophoresed on a 3% agarose gel.

Accession number

The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X87141.

Acknowledgements

We are grateful to Drs R.Voelkar, J.Lim, J.Mason and N.Perrimon for stocks and Dr M.Nirenberg for sharing information prior to publication. Special thanks to Chong-Hyun Kim for preliminary *in situ* hybridization studies. K.W. and F.J. thank Dr J.Campos-Ortega for his continued interest in the *vnd* gene. We thank colleagues of our laboratories and Drs S.Birren, S.Campuzano and J.Modolell for helpful comments on the manuscript. This work was supported by DGICYT Grant PB93-0099 to F.J., an institutional grant of Fundación Ramón Areces to the Centro de Biología Molecular, and National Institutes of Health Grant GM-33205 to K.W. L.M.-M. was supported by NIH Predoctoral Trainee Grant NS-07292 and NIMH predoctoral fellowship MH-09824.

References

Cabrera,C.V., Martínez-Arias,A. and Bate,M. (1987) The expression of three members of the *achaete-scute* gene complex correlates with neuroblast segregation in *Drosophila*. *Cell*, **50**, 425–433.

Campos,A.R., Rosen,D.R., Robinow,S.N. and White,K. (1987) Molecular analysis of the locus *elav* in *Drosophila melanogaster*: a gene whose embryonic expression is neural specific. *EMBO J.*, **6**, 425–431.

Campos-Ortega,J.A. (1993) Early neurogenesis in *Drosophila melanogaster*. In Bate,M. and Martínez Arias,A. (eds), *The Development of Drosophila melanogaster*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 1091–1129.

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

Campuzano,S. and Modolell,J. (1992) Patterning of the *Drosophila* nervous system: the *achaete-scute* gene complex. *Trends Genet.*, **8**, 202–207.

Cavener,D.R. (1987) Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. *Nucleic Acids Res.*, **15**, 1353–1361.

Chu-LaGraff,Q. and Doe,C.Q. (1993) Neuroblast specification and formation regulated by *wingless* in the *Drosophila* CNS. *Science*, **261**, 1594–1597.

Devereux,J., Haeverli,P. and Smithies,O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.*, **12**, 387–395.

Doe,C.Q. (1992) Molecular markers for identified neuroblasts and ganglion mother cells in the *Drosophila* central nervous system. *Development*, **116**, 855–863.

García-Fernández,J., Baguñá,J. and Saló,E. (1991) Planarian homeobox genes: cloning, sequence analysis and expression. *Proc. Natl Acad. Sci. USA*, **88**, 7338–7342.

Ghysen,A. and Dambly-Chaudière,C. (1989) Genesis of the *Drosophila* peripheral nervous system. *Trends Genet.*, **5**, 251–255.

Hartenstein,V. and Campos-Ortega,J.A. (1984) Early neurogenesis in wildtype *Drosophila melanogaster*. *Roux's Arch. Dev. Biol.*, **193**, 308–325.

Irvine,K.D. and Wieschaus,E. (1994) Cell intercalation during *Drosophila* germband extension and its regulation by pair-rule segmentation genes. *Development*, **120**, 827–841.

Jennings,B., Preiss,A., Delidakis,C. and Bray,S. (1994) The Notch signalling pathway is required for *Enhancer of split* bHLH protein expression during neurogenesis in the *Drosophila* embryo. *Development*, **120**, 3537–3548.

Jiménez,F. and Campos-Ortega,J.A. (1987) Genes of the subdivision 1B of the genome of *Drosophila melanogaster* and their participation in neural development. *J. Neurogenet.*, **4**, 179–200.

Jiménez,F. and Campos-Ortega,J.A. (1990) Defective neuroblast commitment in mutants of the *achaete-scute* complex and adjacent genes of *D. melanogaster*. *Neuron*, **5**, 81–89.

Kim,Y. and Nirenberg,M. (1989) *Drosophila* NK-homeobox genes. *Proc. Natl Acad. Sci. USA*, **86**, 7716–7720.

Knust,E., Schrons,H., Grawe,F. and Campos-Ortega,J.A. (1992) Seven genes of the *Enhancer of split* complex of *Drosophila melanogaster* encode helix–loop–helix proteins. *Genetics*, **132**, 505–518.

Kramatschek,B. and Campos-Ortega,J.A. (1994) Neuroectodermal transcription of the *Drosophila* neurogenic genes *E(spl)* and *HLH-m5* is regulated by proneural genes. *Development*, **120**, 815–826.

Lindsley,D.L. and Zimm,G.G. (1992) *The Genome of Drosophila melanogaster*. Academic Press, San Diego.

Lints,T.J., Parsons,L.M., Hartley,L., Lyons,I. and Harvey,R.P. (1993) *Nkx-2.5*: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants. *Development*, **119**, 419–431.

Martín-Bermudo,M.D., Martínez,C., Rodríguez,A. and Jiménez,F. (1991) Distribution and function of the *lethal of scute gene* product during early neurogenesis in *Drosophila*. *Development*, **113**, 445–454.

Martín-Bermudo,M.D., Carmena,A. and Jiménez,F. (1995) Neurogenic genes control gene expression at the transcriptional level in early neurogenesis and in mesectoderm specification. *Development*, **121**, 219–224.

Martin-Morris,L.E. (1991) Molecular analysis of two neighboring genes in *Drosophila*: *ventral nervous system condensation defective* and *Amyloid protein precursor-like*. Ph.D Thesis, Brandeis University, Waltham, MA.

Martin-Morris,L.E. and White,K. (1990) The *Drosophila* transcript encoded by the β-amyloid protein precursor-like gene is restricted to the nervous system. *Development*, **110**, 185–195.

Mason,J.M., Voelker,R.A., Rosen,D., Campos,A.R., White,K. and Lim,J.K. (1986) Localization of terminal deficiency breakpoints on the X chromosome. *Drosophila Information Service*, **63**, 164–165.

Mellerick,D.M. and Nirenberg,M. (1995) Dorsal–ventral patterning genes restrict *NK-2* homeobox gene expression to the ventral half of the central nervous system of *Drosophila* embryos. *Dev. Biol.*, in press.

Nardelli-Haefliger,D. and Shankland,M. (1993). *Lox10*, a member of the *NK-2* homeobox gene class, is expressed in a segmental pattern in the endoderm and in the cephalic nervous system of the leech *Helobdella*. *Development*, **118**, 877–892.

Okkema,P.G. and Fire,A. (1994) The *Caenorhabditis elegans* NK-2 class homeoprotein CEH-22 is involved in combinatorial activation of gene expression in pharyngeal muscle. *Development*, **120**, 2175–2186.

Price,M., Lazzaro,D., Pohl,T., Mattei,M.-G., Rüther,U., Olivo,J.-C., Duboule,D. and Di Lauro,R. (1992) Regional expression of the homeobox gene *Nkx-2.2* in the developing mammalian forebrain. *Neuron*, **8**, 241–255.

Romani,S., Campuzano,S. and Modolell,J. (1987) The *achaete-scute* complex is expressed in neurogenic regions of *Drosophila* embryos. *EMBO J.*, **6**, 2085–2092.

Rosen,D.R. (1987) Molecular cloning of the *ventral nervous system condensation defective* (*vnd*) locus of *Drosophila melanogaster*, a neurogenesis-associated gene. Ph.D Thesis, Brandeis University, Waltham, MA.

Rosen,D., Martin-Morris,L., Luo,L. and White,K. (1989) A *Drosophila* gene encoding a protein resembling the human β-amyloid protein precursor. *Proc. Natl Acad. Sci. USA*, **86**, 2478–2482.

Ruiz-Gómez,M. and Ghysen,A. (1993) The expression and role of a proneural gene, *achaete*, in the development of the larval *Drosophila* nervous system. *EMBO J.*, **12**, 1121–1130.

Saha,S.M., Michel,R.B., Gulding,K.M. and Grainger,R.M. (1993) A *Xenopus* homeobox gene defines dorsal–ventral domains in the developing brain. *Development*, **118**, 193–202.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.

Skeath,J.B. and Carroll,S.B. (1992) Regulation of proneural gene expression and cell fate during neuroblast segregation in the *Drosophila* embryo. *Development*, **114**, 939–946.

Skeath,J.B., Panganiban,G., Selegue,J. and Carroll,S.B. (1992) Gene regulation in two dimensions: the proneural *achaete* and *scute* genes are controlled by combinations of axis-patterning genes through a common intergenic control region. *Genes Dev.*, **6**, 2606–2619.

Skeath,J.B., Panganiban,G.F. and Carroll,S.B. (1994) The *ventral nervous*

system defective gene controls proneural gene expression at two distinct steps during neuroblast formation in *Drosophila*. *Development*, **120**, 1517–1524.

Tautz,D. and Pfeifle,C. (1989) A non-radioactive *in situ* hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma*, **98**, 81–85.

Tsao,D.H.H., Gruschus,J.M., Wang,L.-H., Nirenber,M. and Ferretti,J.A. (1994) Elongation of helix III of the NK-2 homeodomain upon binding to DNA: A secondary structure study by NMR. *Biochemistry*, **33**, 15053–15060.

White,K. (1980) Defective neural development in *Drosophila* embryos deficient for the tip of the X chromosome. *Dev. Biol.*, **80**, 332–344.

White,K., DeCelles,N.L. and Enlow,T.C. (1983) Genetic and developmental analysis of the locus *vnd* in *Drosophila melanogaster*. *Genetics*, **104**, 433–448.

Received on April 4, 1995; revised on May 10, 1995