

**TITLE**

Lineage-unrelated neurons generated in different temporal windows and expressing different combinatorial codes can converge in the activation of the same terminal differentiation gene

**AUTHORS**

María Losada-Pérez<sup>1</sup>, Hugo Gabilondo<sup>1</sup>, Delia del Saz<sup>1</sup>, Magnus Baumgardt<sup>2</sup>, Isabel Molina<sup>1</sup>, Yolanda León<sup>1</sup>, Ignacio Monedero<sup>1</sup>, Fernando Díaz-Benjumea<sup>3</sup>, Laura Torroja<sup>1</sup> and Jonathan Benito-Sipos<sup>1,\*</sup>

1) Departamento de Biología, Universidad Autónoma de Madrid, Cantoblanco, E 28049 Madrid, SPAIN.

2) Department of Clinical and Experimental Medicine, Linköping University, SE-581 83 Linköping, SWEDEN.

3) Centro de Biología Molecular-Severo Ochoa, Universidad Autónoma de Madrid- C.S.I.C., Cantoblanco, E 28049 Madrid, SPAIN.

(\*) Author for correspondence (e-mail: [jonathan.benito@uam.es](mailto:jonathan.benito@uam.es) )

Total character count: 8460

Figures: 6

**RUNNING TITLE**

Neuropeptidergic cell fate specification

Correspondence (J. B-S.): C/ Darwin, 2. Edificio de Biología.Unidad Fisiología Animal. Universidad Autónoma de Madrid. Campus de Cantoblanco. 28049 Madrid (Spain).

Phone: +34 914972818; Fax: +34 914978344; e-mail: [jonathan.benito@uam.es](mailto:jonathan.benito@uam.es)

## SUMMARY

It is becoming increasingly clear that the activation of specific terminal differentiation genes during neural development is critically dependent upon the establishment of unique combinatorial transcription factor codes within distinct neural cell subtypes. However, it is still unclear to which extent these codes are shared by lineage-unrelated neurons expressing the same terminal differentiation genes. Additionally, it is not known if the activation of a specific terminal differentiation gene is restricted to cells born at a particular developmental time point. Here, we utilize the terminal differentiation gene *FMRFa* which is expressed by the Ap4 and SE2 neurons in the *Drosophila* ventral nerve cord, to explore these issues in depth. We find that the Ap4 and SE2 neurons are generated by different neural progenitors and use different combinatorial codes to activate FMRFa expression. Additionally, we find that the Ap4 and SE2 neurons are generated in different temporal gene expression windows. Extending the investigation to include a second *Drosophila* terminal differentiation gene, *Leukokinin*, we find similar results, suggesting that neurons generated by different progenitors might commonly use different transcription factor codes to activate the same terminal differentiation gene. Furthermore, these results imply that the activation of a particular terminal differentiation gene is temporally unrestricted.

## KEY WORDS

*Drosophila*, terminal differentiation, combinatorial code, neuropeptidergic cell identity, temporal genes, FMRFa

## 1. INTRODUCTION

Cell diversity underpins the extensive adaptive radiation of multicellular organisms. Therefore, a large amount of work in the last two decades has aimed to understanding how such cellular complexity is generated during development. In animals, this cellular diversity is especially high in the central nervous system (CNS). Frequently we use the expression of a neurotransmitter (Nt) or a neuropeptide (Np) to define the terminal differentiation phenotype of a neuron. Hence, the common denominator of a given neuron type is the expression of a battery of genes that code for a set of proteins involved in the Nt/Np synthesis and transport. In this manner, the same neuronal fate may arise from different lineage origins. However, although numerous studies have

described components of the transcriptional regulatory networks that activate terminal differentiation genes in various neuron types (review in (Ma, 2006), how lineage-unrelated neurons acquire the same Nt/Np fate is a fundamental question that still remains under debate.

Several different models have been proposed for explaining how the genes that characterize a given neural cell type become activated within a differentiating cell. Based on detailed analyses mostly performed in *Caenorhabditis elegans* (*C.elegans*), one model suggests that terminal differentiation is controlled by a single, or a combination of a few, “terminal selector genes” (reviewed in (Hobert, 2008). Terminal selector genes are transcription factors which assign individual neuronal identities by directly and coordinately controlling the expression of the battery of genes that define a neuronal phenotype. This mechanism of terminal fate specification seems to operate in different types of neurons in *C. elegans* (reviewed in (Hobert, 2008). Moreover, elegant studies in the *C. elegans* dopaminergic system have shown that lineally unrelated neurons acquire their shared phenotype by the action of a common selector gene, the ETS transcription factor *ast-1*, that activates the expression of several genes of the dopamine (DA) pathway which share a common cis-regulatory AST-1 binding motif (DA motif)(Flames and Hobert, 2009). The vertebrate ETS transcription factor Etv1 utilizes an equivalent mechanism to control DA identity in the dopaminergic cells of the mouse olfactory bulb (Flames and Hobert, 2009); however, in this case, data suggest that other ETS transcription factors must be necessary to specify the DA phenotype in other areas of the vertebrate CNS.

The terminal selector gene model implies the existence of at least one committed transcription factor for each neuron type. However, as CNS complexity increases, the increase in neuronal diversity largely overcomes the rise in the number of transcription factors encoded in the genome. This suggests that additional regulatory mechanisms must exist to account for the escalation of neuronal types. Accordingly, a second model has been proposed based on the combinatorial use of transcriptional regulators: a unique combination of regulators, instead of a single regulator, activates the expression of a distinct battery of terminal differentiation genes within a particular neuronal cell type (review in (Ma, 2006; Molyneaux et al., 2007). Examples of combinatorial codes that specify the identity of particular neuronal types have been described in *Drosophila* and

in vertebrates (Benito-Sipos J., 2010; Lee et al., 2008; Miguel-Aliaga et al., 2008; Shirasaki and Pfaff, 2002; Skeath and Thor, 2003; Song et al., 2009). Perhaps the best documented case relates to the differentiation of the *Drosophila* Ap1 (Tvb) and Ap4 (Tv) neurosecretory cells, which express the neuropeptide genes *Nplp1* and *FMRFa*, respectively (Baumgardt et al., 2009; Baumgardt et al., 2007).

Ap1 and Ap4 neurons are both generated by the neuroblast (NB) NB 5-6T lineage during a Castor/Grainyhead temporal window (Baumgardt et al., 2009; Baumgardt et al., 2007), and use a complex and partially overlapping combination of regulators for activating the *FMRFa* and *Nplp1* genes in each postmitotic neuron (Allan et al., 2005; Allan et al., 2003; Baumgardt et al., 2007). Strikingly, miss-expression of the basic code *ap/dimm/dac* leads almost exclusively to strong ectopic *FMRFa* activation, but simply replacing one element of the code (*dac* with *col*) changes ectopic expression to almost exclusively the *Nplp1* gene. This fact might suggest that a unique simple transcriptional code could be used to specify a common terminal identity in lineage-unrelated neurons. However, in other developmental contexts, different codes can activate transcription of the same gene in a cell-specific manner. One of the most exhaustive analyses entails the striped expression pattern of the pair-rule gene *even skipped* (*eve*), which is established by five stripe-specific *cis*-regulatory modules, each of which responds in a distinctive way to gradients of positional information through their unique set of *trans*-acting factor binding sites (Andrioli et al., 2002; Arnosti et al., 1996; Frasch et al., 1987; Fujioka et al., 1999; Goto et al., 1989; Harding et al., 1989; Macdonald et al., 1986; Small et al., 1996). In fact, a few studies in *Drosophila* aimed to address the implication of single players of the code, rather than the entire code, have shown that individual genes are required for neuropeptide activation in some cell subtypes, but not in others (Benveniste et al., 1998; Herrero et al., 2007; Herrero et al., 2003). Still, whether these neuronal subtypes originate from different or equivalent precursors has not been established. Indeed, there are very few examples of neuropeptidergic identities in *Drosophila* in which the lineage is known; these include the Ap1 and Ap4 neurons (Baumgardt et al., 2009; Baumgardt et al., 2007), the EW3 corazoninergic neuron (Isshiki et al., 2001; Karcavich and Doe, 2005; Novotny et al., 2002), and the ABLK leucokinergergic neuron (Benito-Sipos J., 2010). Interestingly, in all these cases, the neuropeptidergic fate arises from the last embryonic temporal window of the lineage (Baumgardt et al., 2009; Benito-Sipos J., 2010; Isshiki et al., 2001). Whether this

feature is common to all neuropeptidergic identities, and whether it contributes to the activation of the specific terminal differentiation genes, remains unknown.

In this study we aimed to establish if neurons originated from different lineages use similar mechanisms to acquire a common terminal phenotype. Specifically, we asked whether activation of a neuropeptidergic gene in lineage-unrelated neurons requires the same combinatorial code, and if it depends on the temporal window in which the neurons are born. To tackle these questions, we have taken advantage of the well-documented *Drosophila* terminal differentiation gene *FMRFa*. The *FMRFa* neuropeptide gene is expressed in about 17 different cell subtypes in the CNS. Two of them, Ap4 and SE2, are the unique cell subtypes identifiable in the embryonic Ventral Nerve Cord (VNC) (Fig. 1A). We show that these two *FMRFa*midergic cells are generated by different NBs and that they are born under distant temporal gene windows. Moreover, we find that none of the factors required for *FMRFa* activation in the Ap4 neuron is necessary in the SE2 cell, except for *dimm*, which is known to act as a master gene of the neuropeptidergic fate. The use of different combinatorial codes to specify the same terminal phenotype in lineage-unrelated neurons is not exclusive of the *FMRFa* fate, because we have found similar results for the *Leucokinin (LK)* neuropeptide gene. Together, our data support the proposition that the use of different genetic combinatorial codes to specify the same terminal fate, in neurons generated by different neuroblasts during distinct temporal gene expression windows, could be a common strategy in developmental neurobiology.

## [Figure.1]

## 2. RESULTS

The expression of the *Drosophila* neuropeptide gene *FMRFa* commences at late embryonic stage 17. In the VNC, *FMRFa* is specifically expressed in two subsets of cells: the six lateral thoracic Ap4 neurons (one in each thoracic hemisegment) and the two medial SE2 cells (in the second suboesophageal segment) (Fig. 1A). The Ap4 cell belongs to a lateral cluster of four neurons, the Apterous (Ap) cluster, which also includes the Ap1/Nplp1 neuron and two neurons with unknown destiny, herein denoted Ap2 and Ap3. The Ap cluster neurons are the last four cells to be born in the large NB

5-6T lineage of 20 cells (Baumgardt et al., 2007). They are born directly, without a Ganglion Mother Cell (GMC) intermediate, from the Neuroblast (NB), with the birth order of: Ap1/Nplp1, Ap2, Ap3, and Ap4/FMRFa (Baumgardt et al., 2009). A complex combinatorial code necessary to activate the *FMRFa* gene in the Ap4 neurons has recently been established (Fig. 1C)(Allan et al., 2005; Allan et al., 2003; Baumgardt et al., 2009). The Ap4 neurons are generated during a Castor/Grainyhead temporal window. Cas plays a critical role during Ap4 neuron determination. Cas activates the transcription co-factor Dachshund (Dac) and the COE class transcription factor Collier/Knot (Col). Col acts to activate the LIM-HD transcription factor Ap and the transcription factor Eyes absent (Eya) (Fig. 1B and C). Ap and Eya activate the bHLH protein Dimmed (Dimm) and participate in the proper axon pathfinding. In this manner, Ap4 neuron projects its axon out of the VNC at the dorsal midline and innervates the Dorsal Neurohemal Organ (DNH; Fig. 1A). When the axon reaches the DNH, a target-derived TGF- $\beta$ /BMP retrograde signal is activated, which results in the expression in Ap4 of the phosphorylated Smad protein Mothers against dpp (pMad) (Fig. 1B and 1C). Castor also activates expression of the Krüppel type zinc finger gene *squeeze* (*sqz*), which in turn activates the transcriptional cofactor *nab*. The co-expression of Sqz/Nab is crucial for the late downregulation of Col in T1 segment, allowing the proper specification of Ap4. These complex processes lead to the expression of the *FMRFa* gene in Ap4 cells (Fig. 1B and 1C) (Allan et al., 2005; Allan et al., 2003; Baumgardt et al., 2009).

To summarize, the Ap4 neuron is generated by the NB 5-6T lineage during a Castor/Grainyhead temporal window. *ap*, *dac* and *sqz* are key components of the combinatorial code specifying Ap4. Eya and Ap are generic markers for the Ap cluster. FMRFa and pMad are specific markers for Ap4 cells, and Nplp1 and Col are specific markers to recognize Ap1 in the Ap cluster.

## [Figure.2]

### 2.1. SE2 and Ap4 arise from different Neuroblast

Previous studies have shown that the Ap4 cell is produced by the thoracic NB 5-6 (Baumgardt et al., 2007). Even though the disparate locations of SE2 and Ap4 neurons suggest that they are born from different NB, differences in cell migration might alternatively explain how two neurons arising from equivalent NBs in different

segments could display different final localizations (reviewed in (Edenfeld et al., 2005). To discard this later possibility, we tested if the SE2 cell is generated by the NB5-6 by analyzing colocalization between the NB 5-6-specific transgenic marker *ladybird early(K)-Gal4 (lbe(K)-Gal4)* (Baumgardt et al., 2009) and anti-FMRFa antibody. In line with previous reports, we found colocalization between *lbe(K)-Gal4* and Ap4 (Fig. 2A; (Baumgardt et al., 2009), but we did not find *lbe(K)-Gal4>UAS-GFP* expression in SE2 cells (Fig. 2B). This result demonstrates that the SE2 and Ap4 cells are generated by different cell lineages.

In order to identify the progenitor NB of the SE2 cell, we co-labeled late embryonic CNSs with the FMRFa antibody and the available collection of markers commonly used to identify VNC NBs. Given that the FMRFa expression is first detected during late embryonic development (from stage 17 onward), and that we do not have other earlier markers to identify those neurons, the present co-expression analysis was carried out at stage 18h After Egg Laying (AEL). Performing lineage tracing at this late developmental stage poses a problem, since it is not well known how the expression of the lineage markers used to identify NBs at stage 11 changes late in embryogenesis. An exception to this is *gooseberry (gsb)*, which has been shown to be lineage specific throughout all embryonic stages (Buenzow and Holmgren, 1995).

We did not find co-expression of FMRFa and *engrailed-Gal4*, *gooseberry-lacZ*, *mirror-lacZ*, *huckebein-lacZ*, *eagle-Gal4* or Seven-up antibody in the SE2 cells (Fig. 2C-H), which points to that the SE2 cell is generated by NB 3-1, 3-5 or 4-1 (Broadus et al., 1995; Chu-LaGraff and Doe, 1993; Doe, 1992). However, NB 4-1 is an unlikely progenitor of the SE2 cell since it always expresses Seven-up, and hence does probably not undergo a *hb* temporal window (Doe, 1992).

There are not any available markers to distinguish between the NB 3-1 and 3-5. However, since the SE2 neurons show a very medial position, the NB 3-1 is the more likely progenitor of the SE2 cell.

In summary, we show that the SE2 and Ap4 cells are generated by different cell lineages. In addition, our results show that the NB 3-1 is a likely candidate for being the progenitor of the SE2 cell.

### [Figure.3]

## 2.2. The SE2 and Ap4 cells are born in different temporal windows

Neural progenitors generate distinct cell types at different stages. In the *Drosophila* embryonic CNS, a serial cascade of transcription factors has been identified and found to act in most, if not all NBs, to change progenitor competence over time. This cascade is called the ‘temporal gene cascade’, and it consists of the *hunchback-kruppel-pdm-castor-grainyhead* genes (Cleary and Doe, 2006; Grosskortenhaus et al., 2005; Isshiki et al., 2001; Kambadur et al., 1998; Mettler et al., 2006; Novotny et al., 2002; Pearson and Doe, 2003; Tran and Doe, 2008; Tsuji et al., 2008). The Ap4 cell is born in the last temporal gene window of the NB 5-6, the Castor/Grainyhead window (Baumgardt et al., 2009). To determine whether the SE2 cell is born in the same temporal gene window, we examined the SE2/FMRFa expression in mutants for each temporal gene (genotypes in legend of the Fig. 3 and Experimental Procedures).

We found that the SE2/FMRFa expression was wild type in all these mutants (Fig. 3A-E) but *kr* mutants, where we occasionally saw additional or fewer SE2 cells than wild type, although the average phenotype was not significantly different from wild type (Fig 3B and 3K). Unfortunately we were not able to study *hb* loss-of-function, since the death of *hb* mutants during early stages precluded the identification of FMRFa cells at stage 17 (in spite of using *hb<sup>P1</sup> hb<sup>FB</sup>* mutants to remove only *hb* CNS expression (Isshiki et al., 2001). However, the *kr* phenotype suggested that the SE2 cell was born in the *hb* temporal window, since the lack of *kr* could extend the *hb* window and thus produce more Hb<sup>+</sup> cells with the SE2/FMRFa fate. To test this hypothesis we studied the expression of Hb and Kr in the SE2 cells. Indeed, we found Hb, but not Kr, expression in the SE2 cells (Fig. 3G and 3H). Moreover, Hb protein was also detected in ectopic SE2 cells in *kr* mutants (Fig. 3I), indicating that they are born due to an extended Hb window.

Our results supported the hypothesis whereby the SE2 is generated in the *hb* temporal window. To reinforce this hypothesis, we analyzed the SE2 cells both in a *hb* misexpression (*elavGal4>UAS-hb*) and *seven up* mutant background, since it has been reported that *seven up* is necessary to switch off *hb* expression in NBs (Kanai et al., 2005; Mettler et al., 2006). Both genetic backgrounds lead to robust ectopic

SE2/FMRFa neurons (Fig. 3F and L), further supporting the birth of SE2 cells in the Hb temporal window.

To summarize, our results indicate that Ap4 and SE2 neurons are born in different embryonic temporal gene windows: the SE2 cell is born in the first temporal window (*hunchback*), while the Ap4 cell is born in the last temporal window (*castor/grainyhead*).

### [Figure.4]

#### 2.3. Ap4 and SE2 neurons express different markers

Given that the markers to identify the postmitotic Ap4 neurons are very well defined, we wished to assess whether these markers are present in SE2 cells. For this purpose, we performed co-localization experiments of each of these markers with the anti-FMRFa antibody in the SE2 cells. We began studying the generic markers for the Ap cluster: Eya and Ap (Fig. 1B). We found that whereas Eya was present in Ap4 cells (inset in Fig. 4A), it was entirely absent from SE2 neurons (arrowheads in Fig. 4A). Similarly, Ap was also absent from the SE2 cell (data not shown; (Benveniste et al., 1998). We next examined the expression of the Zinc Finger transcription factor Squeeze (*Sqz*) and the transcription co-factor Dachshund (*Dac*) in SE2 cells, which are both expressed in all but one (Ap1) Ap cluster cells (Fig. 1B; (Allan et al., 2003; Miguel-Aliaga et al., 2004). Using *sqzGal4>UAS-GFP*, we found that *sqz* was not expressed by the SE2 neurons (arrowheads in Fig. 4B; (Allan et al., 2003). Likewise, we did not detect anti-Dac immunoreactivity in SE2 cells (arrowheads in Fig. 4C). However, both *sqz* and *Dac* were present in Ap4 (inset in Fig. 4B and 4C; Fig. 1B and 4H; (Allan et al., 2003; Miguel-Aliaga et al., 2004).

We next examined the neuropeptidergical marker Dimm (Gauthier and Hewes, 2006; Hewes et al., 2003) in the SE2 cells. Dimm is a bHLH transcription factor that promotes a complete program for neurosecretory cell differentiation (Hamanaka et al., 2010; Park et al., 2008a). Expression of Dimm in most, but not all, peptidergic neurons suggested that Dimm confers neurosecretory identity to cells (Park et al., 2008b). To determine whether Dimm is expressed by SE2 cells, we co-labeled for Dimm and FMRFa antibodies. We saw that Dimm protein is present in SE2 neurons (arrowheads in Fig. 4D).

We next examined a specific marker for Ap4 in the Ap cluster: pMad, a phosphorylated form of the Smad transcription factor Mad, which indicates that the TGF/BMP route is active. We observed that anti-pMad immunoreactivity was not detected in SE2 neurons (arrowheads in Fig. 4E), whereas it was found in Ap4 (inset in Fig. 4E; (Allan et al., 2003)).

Given that the SE2 cells express different markers than Ap4 neurons, we addressed the possibility that they express markers specific of the other neuropeptidergic neuron in the Ap cluster, the Ap1/Nplp1 cell. We analyzed Nplp1 expression in SE2 cells, and, as expected, found no co-localization between Nplp1 and FMRFa in either the SE2 cells (arrowheads in Fig. 4F) nor in the Ap4 cells (arrow in Fig. 4F; (Baumgardt et al., 2007)). Finally, we examined the expression in SE2 cells of one of the most crucial players involved in specifying Ap1 identity, the COE class transcription factor Col (Baumgardt et al., 2007). We did not find anti-Col immunoreactivity in the SE2 neurons (arrowheads in Fig. 4G).

To summarize, the SE2 cells do not express any of the markers that characterize the Ap4 neuron: Eya, Ap, Sqz, Dac or pMad. Neither do they express Col, a specific marker expressed by the other peptidergic cell in the Ap cluster (Ap1). However, the SE2 neurons express Dimm protein. Therefore, the Dimm transcription factor is the only identified marker shared by the Ap4 and the SE2 cells.

### [Figure.5]

#### **2.4. The combinatorial code involved in Ap4 specification is not required to specify the SE2/FMRFa terminal identity**

The fact that the typical Ap4 neuron markers are not expressed by the SE2 cells at 18 h After Egg Laying (AEL) stage do not rule out the possibility that those genes are expressed in the SE2 neurons during previous stages or in the SE2 neuron precursors. In this manner, they could still be involved in SE2 specification. Unfortunately, it is not possible to identify the SE2 cells before 18 h AEL stage, since FMRFa is the only marker for SE2 identification, and it is not expressed until 18 h AEL stage. Therefore, to discard an earlier effect of the Ap4 regulators in SE2 specification, we studied loss-

of-function mutations for these and other important genes involved in the Ap4/FMRFa specification.

The basic combinatorial code for Ap4 specification consists of *ap*, *dac*, *sqz* and *dimm* genes. Pan-neuronal misexpression of this code leads to strong ectopic FMRFa expression (Allan et al., 2003; Baumgardt et al., 2007). To test the role of these genes in specification of the SE2 FMRFa phenotype, we examined anti-FMRFa immunoreactivity in loss-of-function mutant backgrounds for each gene. In addition we also studied SE2 cells in *eya* mutant background (*eya<sup>CliII-D</sup>*), since *eya* is a crucial player for Ap4 specification (Miguel-Aliaga et al., 2004). Consistent with previous reports (Allan et al., 2003; Benveniste et al., 1998; Miguel-Aliaga et al., 2004), we did not find differences in SE2/FMRFa expression in *ap*, *dac*, *sqz*, or *eya* mutants with respect to wild type (Fig. 5A-E). In contrast, intensity of FMRFa in SE2 cells was moderately weak in *dimm* mutants (Fig. 5F). The difference in SE2/FMRFa signal intensity between wild type and *dimm* mutants was significant (Fig. 5F, 5K). These data are consistent with the expression of Dimm in SE2 cells, and with previous studies on *dimm* function and expression (Fig. 4)(Hewes et al., 2003; Park et al., 2008b).

In addition, we studied the possible role of *nab* in SE2 specification. *nab* encodes a transcriptional cofactor without a DNA binding domain, and it has been previously reported that Nab acts as coactivator of Sqz for specifying Ap4 cell identity (Terriente Felix et al., 2007). Nevertheless, Nab is known to have other partners in different contexts (Benito-Sipos J., 2010; Terriente Felix et al., 2007), and thus it could be involved in SE2 fate specification without Sqz. We studied the SE2/FMRFa expression in *nab* mutant background (*nab<sup>R52</sup>* and *nab<sup>SH143</sup>*), and in that genetic context the SE2 cells are unaffected (Fig. 5G). In addition, since *col* is a crucial player in the specification of the Ap cluster (Baumgardt et al., 2007), we studied SE2/FMRFa expression in *col* mutants. According to previous results, SE2 neurons are properly specified in that genetic background (Fig 5I; (Baumgardt et al., 2007).

FMRFa expression in the Ap4 neuron is crucially dependent on a target-derived BMP signal mediated by the BMP ligand Gbb, which is accessed by Ap4 axons at the DNH (Allan et al., 2003; Marques et al., 2002). Although pMad is not expressed by the SE2 cell at 18 h AEL stage (this work; Fig. 4E), we wanted to discard the possibility that

BMP signaling is involved in SE2 specification at earlier stages. We addressed this question by analyzing *wit*, *gbb*, and *baboon* mutants (*wit* and *gbb*: type-II BMP receptor; *baboon*: type-I BMP receptor; (Zheng et al., 2003), and we did not find differences with respect to wild type in SE2/FMRFa expression (data not shown). This indicates that BMP signaling is not necessary for SE2/FMRFa specification. To rule out other retrograde signal involved in SE2 specification, we interfered with retrograde axonal transport by expressing a dominant-negative version of the P150/Glued dynactin motor complex component (*UAS-Glued<sup>DN</sup>*) with the *elav-Gal4* pan-neural driver (Allen et al., 1999). In *elav-Gal4>UAS-Glued<sup>DN</sup>* 18 h AEL embryos we observed a wild type SE2/FMRFa expression (Fig. 5H). Hence, retrograde signaling is not necessary for proper SE2 specification.

In summary, the combinatorial code specifying Ap4 fate (*ap*, *dac*, *sqz*) is not required to specify the SE2/FMRFa fate. In *dimm* mutants, SE2 cells are properly specified, but FMRFa signal intensity is lower than wild type. In addition, retrograde axonal transport is not necessary in SE2 specification.

### [Figure.6]

#### **2.5. The use of different combinatorial codes to specify the same terminal differentiation phenotype in lineage-unrelated cells is not exclusive of the *FMRFa* gene**

We have shown how, at least two, different combinatorial codes converge in specifying the same fate. Is this strategy exclusive of the FMRFa system, or is it extended in animal neurodevelopment? To gain insights into this question, we studied another specific terminal differentiation gene, the *Leucokinin* (*LK*) gene, which defines the LK identity. The LK system is particularly interesting since LK expression is highly restricted in the VNC; it is expressed in 18 neurons, divided in two subsets of cells: the SELK cells comprise four neurons in the suboesophageal ganglion, and the ABLK cells consist of 14 abdominal cells, one in each abdominal hemineuromere, from A1 to A7 (Fig.6A)(Cantera and Nassel, 1992; Herrero et al., 2003). Moreover, previous reports have shown that *ap* and *sqz* are differentially required for LK activation in different leucokineric cells, suggesting that *Lk* activation can be accomplished by distinct mechanisms (Herrero et al., 2007). In addition, a new set of genes involved in the

specification of ABLK neurons have recently been identified (Benito-Sipos J., 2010). However, whether SELK and ABLK cells are lineally related is unknown, and the effect of the genes that control ABLK identity (Benito-Sipos J., 2010) has not been analyzed in SELK cells. Therefore, we studied anti-LK immunoreactivity in those mutant backgrounds in which ABLK specifications is known to be disturbed (Benito-Sipos J., 2010; Gauthier and Hewes, 2006; Hewes et al., 2003).

ABLK neurons have been shown to originate from NB 5-5 during the Castor/Grainyhead temporal window (Benito-Sipos J., 2010). Even though the disparate locations of SELK and ABLK neurons suggest that they are born from different NB, we tested if the SELK cell is generated by the NB5-5 by analyzing colocalization between *Gooseberry-lacZ* (*gsb-lacZ*) (a neuroblast marker for row 5, 6 and NB7-1; (Buenzow and Holmgren, 1995; Duman-Scheel et al., 1997; Gutjahr et al., 1993; Skeath et al., 1995) and anti-LK antibody. We did not find *Gsb-lacZ* expression in SELK cells (Fig. 6B). This result suggests that the SELK and ABLK cells are generated by different cell lineages. We next examined whether SELK cells are also born in the Castor/Grainyhead temporal window. To address this question we studied LK expression in temporal gene mutant backgrounds (except *hb*, since *hb* mutant did not live until stage18, when LK expression begins). We found that SELK cells are missing only in *cas* mutants, but not in *grh* mutants (Fig. 6E, 6G and 6P). These results suggest that SELK cells are born in a *cas* temporal window, without *grh* involvement. In order to reinforce this evidence we tested pan-neural *cas* miss-expression, and, as we expected, we found significant ectopic SELK neurons (Fig. 6F and 6P). Taken together, these results show that SELK neurons are born in a different temporal window than ABLK cells.

Some of the genetic requirements for SELK and ABLK specification have been recently identified. In the VNC, SELK cells require *Sqz*, but not *Ap*, for LK expression, while in ABLK neurons both *sqz* and *ap* are dispensable (Herrero et al., 2007). In the ABLK cell (Benito-Sipos J., 2010), Notch signaling is active (Notch ON), whereas its Notch OFF sibling undergoes apoptosis. In addition, *nab*, *klumpfuss* (*klu*) and *jumeaux* (*jumu*) genes are crucial for ABLK specification. *klu* encodes a Zn-finger protein and *jumu* encodes a new member of the winged-helix family of transcription factors. Also, *Dimm* is required in ABLK cells to maintain high levels of LK expression (Hewes et al., 2003).

First, we studied if specification of the SELK fate requires Notch ON condition (Benito-Sipos J., 2010). However, unlike the ABLK cells, SELK cells seemed to be Notch OFF, since in *sanpodo* mutants SELK/LK expression was normal (Fig. 6H and 6P), whereas SELK cells were missing when Notch signaling was activated in SELK cells (*elavG4>UAS-Notch intra*) (Fig. 6I and 6P). Moreover, the SELK cell does not seem to have a sibling LK-expressing cell undergoing apoptosis, given that in a scenario where apoptosis is inhibited (*elav-Gal4>UAS-p35*), the number of SELK cells is wild type (Fig. 6J and 6P).

We next investigated the role of key players in ABLK specification in the SELK scenario. We have found that the intensity of LK expression is lower in SELK than in ABLK neurons (Fig. 6A-B). This aspect lead us to study Dimm expression in the SELK cells, because Dimm expression has been reported to correlate with Np levels (Park et al., 2008b). Confirming previous results (Park et al., 2008b), we saw that SELK neurons did not express Dimm (Fig. 6C-C''). Next, we studied LK expression in *nab* mutant background, and we found that SELK cells are completely missing (Fig. 6K and 6P). However, whereas *sqz* is not involved in ABLK specification (Benito-Sipos J., 2010; Herrero et al., 2007), it is crucial for expression of LK in SELK cells, since we did not find SELK cells in *sqz* mutant (Fig. 6L and 6P; (Herrero et al., 2007)). This result suggests that Sqz is the partner of Nab in this context. To further investigate this idea, we studied the simultaneous misexpression of *nab* and *sqz* (*elav-Gal4>UAS-Nab, UAS-Sqz*), and we found that it leads to robust ectopic SELK/LK expression (Fig. 6M and 6P). Given that Nab does not have a DNA binding site, our results support that Nab is working with a different partner in SELK and ABLK neurons. Finally, we examined the SELK/LK expression in *jumu* and *klu* mutant background. Contrary to the effect on ABLK cells (inset in Fig. 6N; 6O), we did not find absence of SELK cells in those mutants; in fact we found ectopic SELK cells in *klu* mutant (Fig. 6O and 6P).

In summary, SELK and ABLK cells are born in different temporal windows from distinct NBs. SELK cells are Notch OFF instead of Notch ON, and they do not have a LK-expressing sibling cell undergoing apoptosis. Nab is involved in SELK specification, but in contrast to the ABLK scenario, Sqz seems to be the partner of Nab in these cells. Besides, neither *jumu* nor *klu* are required for SELK fate. Therefore, the LK fate is achieved in SELK and ABLK by different combinatorial codes.

### 3. DISCUSSION

An important question in development is how lineage unrelated neurons converge to express the same terminal differentiation gene. In the *C. elegans* dopaminergic system, a single transcription factor, Ast-1, acts in all cell types to coordinately activate expression of the genes required for Dopamine synthesis and transport (Flames and Hobert, 2009). In this study, we show that in *Drosophila*, a different mechanism seems to operate to specify particular neuropeptidergic fates. Our results demonstrate that two different cell lineages generate FMRFa expressing cells during distant temporal windows, and that activation of *FMRFa* in these cells use distinct combinatorial codes. In addition, we have extended our findings to another specific differential fate, the LK system, and we have found similar results. Collectively, our data suggest that neurons of diverse origins in different temporal windows can achieve the same terminal fate by the use of different genetic combinatorial codes.

#### 3.1. The same specific neuropeptidergic fate can originate from distinct NBs and at different temporal gene windows.

Generation of cell diversity occurs at different levels during the developmental process, from the early precursor to the final postmitotic cell. In *Drosophila*, each NB in a segment is provided with an individual identity, which will determine the unique properties of its descendants. Therefore, specific terminal phenotypes might be determined at the NB level, so that two neuron types that share a particular trait –such as expression of a Np or Nt- might arise from equivalent NBs, and differ in other phenotypic attributes –such as location or axonal projection- by the action of segment specific genes -such as homeobox genes (Skeath and Thor, 2003). However, our study provides the first demonstration in *Drosophila* for the existence of lineage-unrelated neurons that express the same specific neuropeptidergic gene. We have shown that the FMRFamidergic neurons Ap4 and SE2 originate from different NBs. In addition, our data strongly suggest that SELK and ABLK neurons also derive from distinct NBs, since SELK cells do not express *GsbZ* -a specific marker for NB from rows 5 and 6, and NB7-1, whereas ABLK cells, which generate from NB5-5, express it (Benito-Sipos J., 2010). However, whether *GsbZ* expression is maintained until stage 18 has not been

definitely established, and the lack of earlier markers for SELK cells prevented us from analyzing *GsbZ* at earlier stages.

In *Drosophila*, identity is also achieved at the level of the GMC, through the action of temporal genes that are inherited from the NB. In those few cases in which the lineages of specific neuropeptidergic cells are known, the neuropeptidergic neurons are always born during the last temporal window of the lineage. (Baumgardt et al., 2009; Benito-Sipos J., 2010; Isshiki et al., 2001). This intriguing observation prompted us to examine if this characteristic is shared by all neuropeptidergic fates. We demonstrate that the same neuropeptidergic identity can be generated at different temporal gene windows: SE2 cells are born in a *hb* window, while Ap4 neurons originate in a *cas/grh* window, and ABLK cells are born in a *cas/grh* window whereas SELK cells are generated in a *cas* (but not *grh*) window. Moreover, the fact that *kr* mutants occasionally show ectopic SE2/FMRFa cells, which express Hb, suggests that *hb* is not the last temporal gene of that lineage, but that there is at least one additional posterior *kr* window. Likewise, the loss of most mid Dimm cells in *kr* mutants further support the birth of neuropeptidergic fates from temporal windows other than the last one of the lineage.

### **3.2. Different combinatorial codes converge to activate the same specific terminal differentiation gene.**

Since neither the NB nor the temporal gene window is determinant for the activation of a specific terminal differentiation gene, the question rises: do lineage-unrelated cells that express the same terminal differentiation gene share code? Or is there a common gene that could act as a terminal selector gene in these cells? Previous accounts have shown how different genes are involved in the specification of one cell type, but not another, within a neuronal cell fate. For instance, the *Drosophila* LIM-homeodomain transcription factor *islet* controls the neurotransmitter expression and pathfinding behavior in some serotonergic and dopaminergic neurons, but its function is cell-type dependent (Thor and Thomas, 1997). It has also been shown that *ap* is necessary for the Ap4 fate but not for SE2 specification (Benveniste et al., 1998), and that *ap* and *sqz* are crucial for the LHLK fate, whereas only *sqz*, but not *ap*, is required in SELK cells, and both are dispensable for ABLK specification (Herrero et al., 2007; Herrero et al., 2003). Examples of genes controlling the terminal differentiation of single neuronal subtypes and that show cell-type-specificity have also been described in *C. elegans* (Altun-

Gultekin et al., 2001; Wightman et al., 2005; Zheng et al., 2005) and vertebrates (reviewed in (Goridis and Rohrer, 2002; Howard, 2005). Our work is the first systematic analysis in which the complete genetic combinatorial code known to determine a specific neuropeptidergic fate (FMRFa) has been studied in neurons of different origins (Ap4 and SE2). The code consists of *ap*, *dac*, *sqz* and *dimm*. This code is necessary and sufficient to specify the Ap4 neuron fate (Allan et al., 2005; Baumgardt et al., 2007). However, we have shown that, except for *dimm*, no single gene of the code is shared between Ap4 and SE2. In addition, other players involved in Ap4 specification, such as *eya*, *nab*, *col*, or BMP signaling, have no effect whatsoever on SE2/FMRFa. Similarly, the two leucokinergic cell types within the *Drosophila* VNC, SELK and ABLK, only coincide in their requirement for Nab to activate *Lk* expression, but Nab seems to act with a different partner in each cell type. Even more, expression of LK is altered in *klu* mutants in both SELK and ABLK cells, but the opposite phenotypes observed in both neuron types indicate a completely different mechanism of regulation. All together, our data demonstrate that at least two different combinatorial codes can converge to specify the same terminal phenotype. Further analysis will be needed to find the combinatorial code involved in SE2 and SELK specification, but this will involve identifying early markers for these two cell types.

### **3.3. Are there terminal selector genes for Neuropeptidergic fates?**

The term terminal selector gene has been applied in neurobiology to define transcription factors that assign individual identities to neurons by directly controlling post-mitotic transcription of the core battery of genes that define that identity (Hobert, 2008). For example, *ast-1* has been characterized as the terminal selector gene for the *C. elegans* dopaminergic fate, because it directly and coordinately activates the expression of all dopamine pathway genes in all dopaminergic cell types, and its loss prevents terminal differentiation of all dopaminergic neuronal subtypes. Putative selector genes for neurotransmitter phenotypes have been described also in vertebrates (Cheng et al., 2004). In *Drosophila*, however, detailed analyses on neuropeptidergic fates suggest the existence of complex specific post-mitotic combinations of transcription factors for determining terminal cell identity. Moreover, our work shows that, at least for *FMRFa* and *Lk*, different neuronal subtypes use different codes to activate Np gene expression. Overall, these data indicate that, thus far, no individual gene has been identified that is both necessary and sufficient for the FMRFa or LK fate.

So, does this imply that there are no terminal selector genes for the neuropeptidergic fates? Regarding expression of a terminal differentiation phenotype, there is a fundamental difference between a neurotransmitter (Nt) and a neuropeptide (Np) fate. Expression of classical Nt requires transcriptional orchestration of the genes encoding the specific enzymes and transporters involved in its synthesis and packaging, an outcome that can be easily achieved by the use of terminal selector genes. On the other hand, expression of a Np requires transcriptional activation of the Np gene itself, together with enzymes and proteins for peptide processing and packaging; but unlike with classical Nt, these proteins show very low specificity, are expressed broadly, and are used to process and transport very different types of Np (Hamanaka et al., 2010; Park et al., 2008a). Indeed, recent studies in *Drosophila* indicate that expression of the genes encoding these Np processing proteins is coordinately activated by the transcription factor Dimm (Hamanaka et al., 2010; Park et al., 2008a), which acts as a master regulator for neurosecretory identity by promoting a complete program of neurosecretory cell differentiation (Hamanaka et al., 2010). In such a way, Dimm is responsible for setting up the intracellular conditions that permit accumulation of functional Nps in neurosecretory cells (Hamanaka et al., 2010; Park et al., 2008b). Consistent with this function, we find reduced, but not complete absence of FMRFa in *dimm* mutants in Ap4 and SE2 cells, and low levels of LK peptide in SELK cells, which do not normally express Dimm (Park et al., 2008b). Therefore, Dimm must be part of any code specifying neurosecretory identity, but it determines a generic neurosecretory state, rather than expression of a specific neuropeptide. This is in agreement with our results, which show that Dimm is the only gene required by both Ap4 and SE2 cells for proper FMRFa expression. However, Dimm has also been reported to control transcription of Np genes (Gauthier and Hewes, 2006; Hewes et al., 2003), probably through the formation of dimers with other transcription factors, like Ap, Chip, or Sqz (Allan et al., 2005). In this manner, *dimm* could also act as a terminal selector gene (Hobert, 2008): on one hand, it would confer the cell with the necessary conditions for the Np functional environment (Hamanaka et al., 2010); on the other hand, it would combine with cell-specific factors to help activate a specific Np gene within subsets of neurosecretory cells (Allan et al., 2005; Baumgardt et al., 2007).

A unique combination of Dimm with cell-specific transcription factors could constitute the code for specifically activating a particular neuropeptidergic fate in all neurosecretory cell subtypes. However, we have not found any transcription factor other than Dimm common to all neurosecretory FMRFa cells. Moreover, promoter analyses have shown that different *FMRFa* enhancers direct expression in Ap4 and SE2 cells (Benveniste et al., 1998; Schneider et al., 1993). SELK and ABLK cells also use distinct combinations of transcription factors for *Lk* expression. Therefore, why do cells use different combinatorial codes to activate the same terminal differentiation phenotype? We propose two alternative hypotheses to explain our findings. First, different cell subtypes often processed the same propeptide differently, producing unique combinations of functional peptides (Sossin et al., 1989). In addition, physiological stimuli can have diverse outcomes on Np processing and expression in different cell types. Therefore, distinct cell types could require activation of the Np gene together with specific sets of Np processing genes, and thus make use of different combinatorial codes. Alternatively, different combinatorial codes could act as upstream regulators that coordinately activate distinct aspects of the neuronal phenotype. For example, the combinatorial code of the Ap4 cell is necessary not only for *FMRFa* expression, but also for proper axon pathfinding (Allan et al., 2005; Allan et al., 2003). In such case, there may be terminal selector genes for each trait, which have not been yet identified. These two hypotheses are not exclusive, and further work is needed to deepen our understanding of this fascinating problem.

## 4. EXPERIMENTAL PROCEDURES

### 4.1. Fly Stocks

Fly stocks were raised and crosses were performed at 25°C on standard medium. The following fly mutant alleles were used: *sqz<sup>ie</sup>* (Allan et al., 2003); *ap<sup>P44</sup>*; *ap<sup>md544</sup>* (referred to as *ap-Gal4*; (O'Keefe et al., 1998); *cas<sup>Δ1</sup>* (Mellerick et al., 1992)(provided by S. Thor); *grh<sup>IM</sup>* (Nusslein-Volhard et al., 1985); *dim<sup>P1</sup>* (Hewes et al., 2003), (provided by S. Thor); *hb<sup>P1</sup> hb<sup>FB</sup>*, a genetic combination that removes Hb CNS expression (Hulskamp et al., 1994; Isshiki et al., 2001) (provided by C.Doe); *jumu<sup>11683</sup>* and *klu<sup>212IR51C</sup>* (Cheah et al., 2000; Yang et al., 1997) (provided by W. Chia); *col<sup>1</sup>/col<sup>3</sup>* (provided by Stefan Thor) (Baumgardt et al., 2007); *kr<sup>1</sup> kr<sup>CD</sup>* to remove Kr CNS expression (Isshiki et al., 2001; Romani et al., 1996) (provided by C.Doe); *nab<sup>SH143</sup>*, *nab<sup>R52</sup>* (Terriente Felix et al.,

2007); *spdo*<sup>G104</sup> (Skeath and Doe, 1998); *dac*<sup>4</sup> (Mardon et al., 1994); *eya*<sup>Cli-III</sup> (Pignoni et al., 1997); *svp*<sup>1</sup> (Kanai et al., 2005); *gsb*<sup>01155</sup> (referred to as *gsb-lacZ*) (Duman-Scheel et al., 1997) (provided by Stefan Thor); *mirror-lacZ* (a gift of Sonsoles Campuzano); *hkb-lacZ*; *en-Gal4*. Transgenic stocks were: *elav-Gal4* (DiAntonio et al., 2001); *UAS-cas* and *UAS-hb* (provided by S. Thor) (Kambadur et al., 1998); *UAS-p35* (provided by Manuel Calleja); *UAS-Glued*<sup>DN</sup> (Allen et al., 1999); *UAS-N<sup>intra</sup>* (provided by A. Baonza); *lbe(K)-Gal4*; *UAS-nmEGFP*. (provided by S. Thor). Mutants were kept over *CyO*, *Act-GFP*; *CyO*, *Dfd-EYFP*; *TM3*, *Ser*, *Act-GFP*; *CyO*, *twi-Gal4*, *UAS-GFP*; *TM3*, *Sb*, *Ser*, *twi-Gal4*, *UAS-GFP*; or *TM6*, *Sb*, *Tb*, *Dfd-EYFP* balancer chromosomes. As wild type, *orizo 2* was often used. Unless otherwise stated, flies were obtained from the Bloomington Drosophila Stock Center.

#### 4.2. Immunohistochemistry

The antibodies used were: Guinea pig  $\alpha$ -Col (1:1,000), guinea pig  $\alpha$ -Dimm (1:1,000), chicken  $\alpha$ -proNplp1 (1:1,000), rabbit  $\alpha$ -Krüppel (1:500) and rabbit  $\alpha$ -proFMRFa (1:1,000) (all of them provided by S. Thor) (Baumgardt et al., 2007); mAb  $\alpha$ -Seven up (1:50) (a gift of Y. Hiromi); rabbit  $\alpha$ -pMad (1:200) (provided by M. Calleja); rabbit anti-Leucokinin (1:50) (provided by D. Nässel); rabbit  $\alpha$ -Hunchback (1:1,000) (provided by R. Pflanz); mAb  $\alpha$ -Eya10H6 (1:250) and mAb  $\alpha$ -Dac (1:25) (from Developmental Studies Hybridoma Bank, Iowa City, IA, US). All polyclonal sera were pre-absorbed against pools of early embryos. Secondary antibodies were conjugated with FITC, Rhodamine-RedX or Cy5, and used at 1:500 (Jackson ImmunoResearch, PA, US). Embryos were dissected in PBS, fixed for 25 minutes in 4% PFA, blocked and processed with antibodies in PBS with 0.2% Triton-X100 and 4% donkey serum. Slides were mounted with Vectashield (Vector, Burlingame, CA, US). In all cases wild-type and mutant embryos were stained and analyzed on the same slide.

#### 4.3. Confocal Imaging, Data Acquisition and Staining quantification

A Zeiss META 510 Confocal microscope was used to collect data for all fluorescent images; confocal stacks were merged using LSM software or Adobe Photoshop CS4. Where appropriate, images were false colored to facilitate for color-blind readers or to facilitate the understanding of the paper.

For staining quantification we use the Mean pixel luminosity data. Mean pixel luminosity for the cell area was measured for each neuron using Adobe Photoshop.

#### 4.4. Statistical Methods

Statistical analysis was performed using Microsoft Excel. Quantifications of observed phenotypes were performed using Student's two-tailed t test, assuming equal variance.

#### ACKNOWLEDGMENTS

We are grateful to S.Thor, M. Calleja, S. Campuzano, Y.Hiromi, R. Pflanz, A. Baonza, C.Doe, D. Nässel, W. Chia and the Developmental Studies Hybridoma Bank and the Bloomington Stock Center for sharing antibodies and fly lines. We thank I.Canal and S. Thor for critically reading the manuscript. We also want to thank C.Sánchez, V.Labrador and A. Muñoz for their technical assistance in confocal microscopy. This work was supported by a grant from the Spanish Ministerio de Ciencia e Innovación (BFU-2008-04683-C02-02 to L.T.).

#### REFERENCES

- Altun-Gultekin, Z., Andachi, Y., Tsalik, E.L., Pilgrim, D., Kohara, Y. and Hobert, O. (2001) A regulatory cascade of three homeobox genes, *ceh-10*, *ttx-3* and *ceh-23*, controls cell fate specification of a defined interneuron class in *C. elegans*. *Development* 128, 1951-69.
- Allan, D.W., Park, D., St Pierre, S.E., Taghert, P.H. and Thor, S. (2005) Regulators acting in combinatorial codes also act independently in single differentiating neurons. *Neuron* 45, 689-700.
- Allan, D.W., St Pierre, S.E., Miguel-Aliaga, I. and Thor, S. (2003) Specification of neuropeptide cell identity by the integration of retrograde BMP signaling and a combinatorial transcription factor code. *Cell* 113, 73-86.
- Allen, M.J., Shan, X., Caruccio, P., Froggett, S.J., Moffat, K.G. and Murphey, R.K. (1999) Targeted expression of truncated glued disrupts giant fiber synapse formation in *Drosophila*. *J Neurosci* 19, 9374-84.
- Andrioli, L.P., Vasisht, V., Theodosopoulou, E., Oberstein, A. and Small, S. (2002) Anterior repression of a *Drosophila* stripe enhancer requires three position-specific mechanisms. *Development* 129, 4931-40.
- Arnosti, D.N., Barolo, S., Levine, M. and Small, S. (1996) The *eve* stripe 2 enhancer employs multiple modes of transcriptional synergy. *Development* 122, 205-14.
- Baumgardt, M., Karlsson, D., Terriente, J., Diaz-Benjumea, F.J. and Thor, S. (2009) Neuronal subtype specification within a lineage by opposing temporal feed-forward loops. *Cell* 139, 969-82.
- Baumgardt, M., Miguel-Aliaga, I., Karlsson, D., Ekman, H. and Thor, S. (2007) Specification of neuronal identities by feedforward combinatorial coding. *PLoS Biol* 5, e37.
- Benito-Sipos J., E.-G.A., Moris-Sanz M., Baumgardt M., Thor S. and Díaz-Benjumea F.J. (2010) A genetic cascade involving *klumpfuss*, *nab* and *castor* specifies the abdominal leucokinergetic neurons in the *Drosophila* CNS. *Development* In press.

- Benveniste, R.J., Thor, S., Thomas, J.B. and Taghert, P.H. (1998) Cell type-specific regulation of the *Drosophila* FMRF-NH2 neuropeptide gene by Apterous, a LIM homeodomain transcription factor. *Development* 125, 4757-65.
- Broadus, J., Skeath, J.B., Spana, E.P., Bossing, T., Technau, G. and Doe, C.Q. (1995) New neuroblast markers and the origin of the aCC/pCC neurons in the *Drosophila* central nervous system. *Mech Dev* 53, 393-402.
- Buenzow, D.E. and Holmgren, R. (1995) Expression of the *Drosophila* gooseberry locus defines a subset of neuroblast lineages in the central nervous system. *Dev Biol* 170, 338-49.
- Cantera, R. and Nassel, D.R. (1992) Segmental peptidergic innervation of abdominal targets in larval and adult dipteran insects revealed with an antiserum against leucokinin I. *Cell Tissue Res* 269, 459-71.
- Cleary, M.D. and Doe, C.Q. (2006) Regulation of neuroblast competence: multiple temporal identity factors specify distinct neuronal fates within a single early competence window. *Genes Dev* 20, 429-34.
- Cheah, P.Y., Chia, W. and Yang, X. (2000) Jumeaux, a novel *Drosophila* winged-helix family protein, is required for generating asymmetric sibling neuronal cell fates. *Development* 127, 3325-35.
- Cheng, L., Arata, A., Mizuguchi, R., Qian, Y., Karunaratne, A., Gray, P.A., Arata, S., Shirasawa, S., Bouchard, M., Luo, P., Chen, C.L., Busslinger, M., Goulding, M., Onimaru, H. and Ma, Q. (2004) *Tlx3* and *Tlx1* are post-mitotic selector genes determining glutamatergic over GABAergic cell fates. *Nat Neurosci* 7, 510-7.
- Chu-LaGraff, Q. and Doe, C.Q. (1993) Neuroblast specification and formation regulated by wingless in the *Drosophila* CNS. *Science* 261, 1594-7.
- DiAntonio, A., Haghighi, A.P., Portman, S.L., Lee, J.D., Amaranto, A.M. and Goodman, C.S. (2001) Ubiquitination-dependent mechanisms regulate synaptic growth and function. *Nature* 412, 449-52.
- Doe, C.Q. (1992) Molecular markers for identified neuroblasts and ganglion mother cells in the *Drosophila* central nervous system. *Development* 116, 855-63.
- Duman-Scheel, M., Li, X., Orlov, I., Noll, M. and Patel, N.H. (1997) Genetic separation of the neural and cuticular patterning functions of gooseberry. *Development* 124, 2855-65.
- Edenfeld, G., Stork, T. and Klambt, C. (2005) Neuron-glia interaction in the insect nervous system. *Curr Opin Neurobiol* 15, 34-9.
- Flames, N. and Hobert, O. (2009) Gene regulatory logic of dopamine neuron differentiation. *Nature* 458, 885-9.
- Frasch, M., Hoey, T., Rushlow, C., Doyle, H. and Levine, M. (1987) Characterization and localization of the even-skipped protein of *Drosophila*. *EMBO J* 6, 749-59.
- Fujioka, M., Emi-Sarker, Y., Yusibova, G.L., Goto, T. and Jaynes, J.B. (1999) Analysis of an even-skipped rescue transgene reveals both composite and discrete neuronal and early blastoderm enhancers, and multi-stripe positioning by gap gene repressor gradients. *Development* 126, 2527-38.
- Gauthier, S.A. and Hewes, R.S. (2006) Transcriptional regulation of neuropeptide and peptide hormone expression by the *Drosophila* dimmed and cryptocephal genes. *J Exp Biol* 209, 1803-15.
- Goridis, C. and Rohrer, H. (2002) Specification of catecholaminergic and serotonergic neurons. *Nat Rev Neurosci* 3, 531-41.
- Goto, T., Macdonald, P. and Maniatis, T. (1989) Early and late periodic patterns of even-skipped expression are controlled by distinct regulatory elements that respond to different spatial cues. *Cell* 57, 413-22.
- Grosskortenhaus, R., Pearson, B.J., Marusich, A. and Doe, C.Q. (2005) Regulation of temporal identity transitions in *Drosophila* neuroblasts. *Dev Cell* 8, 193-202.

- Gutjahr, T., Patel, N.H., Li, X., Goodman, C.S. and Noll, M. (1993) Analysis of the gooseberry locus in *Drosophila* embryos: gooseberry determines the cuticular pattern and activates gooseberry neuro. *Development* 118, 21-31.
- Hamanaka, Y., Park, D., Yin, P., Annangudi, S.P., Edwards, T.N., Sweedler, J., Meinertzhagen, I.A. and Taghert, P.H. (2010) Transcriptional orchestration of the regulated secretory pathway in neurons by the bHLH protein DIMM. *Curr Biol* 20, 9-18.
- Harding, K., Hoey, T., Warrior, R. and Levine, M. (1989) Autoregulatory and gap gene response elements of the even-skipped promoter of *Drosophila*. *EMBO J* 8, 1205-12.
- Herrero, P., Magarinos, M., Molina, I., Benito, J., Dorado, B., Turiegano, E., Canal, I. and Torroja, L. (2007) Squeeze involvement in the specification of *Drosophila* leucokineric neurons: Different regulatory mechanisms endow the same neuropeptide selection. *Mech Dev* 124, 427-40.
- Herrero, P., Magarinos, M., Torroja, L. and Canal, I. (2003) Neurosecretory identity conferred by the apterous gene: lateral horn leucokinin neurons in *Drosophila*. *J Comp Neurol* 457, 123-32.
- Hewes, R.S., Park, D., Gauthier, S.A., Schaefer, A.M. and Taghert, P.H. (2003) The bHLH protein Dimmed controls neuroendocrine cell differentiation in *Drosophila*. *Development* 130, 1771-81.
- Hobert, O. (2008) Regulatory logic of neuronal diversity: terminal selector genes and selector motifs. *Proc Natl Acad Sci U S A* 105, 20067-71.
- Howard, M.J. (2005) Mechanisms and perspectives on differentiation of autonomic neurons. *Dev Biol* 277, 271-86.
- Hulskamp, M., Lukowitz, W., Beermann, A., Glaser, G. and Tautz, D. (1994) Differential regulation of target genes by different alleles of the segmentation gene hunchback in *Drosophila*. *Genetics* 138, 125-34.
- Isshiki, T., Pearson, B., Holbrook, S. and Doe, C.Q. (2001) *Drosophila* neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny. *Cell* 106, 511-21.
- Kambadur, R., Koizumi, K., Stivers, C., Nagle, J., Poole, S.J. and Odenwald, W.F. (1998) Regulation of POU genes by castor and hunchback establishes layered compartments in the *Drosophila* CNS. *Genes Dev* 12, 246-60.
- Kanai, M.I., Okabe, M. and Hiromi, Y. (2005) seven-up Controls switching of transcription factors that specify temporal identities of *Drosophila* neuroblasts. *Dev Cell* 8, 203-13.
- Karcavich, R. and Doe, C.Q. (2005) *Drosophila* neuroblast 7-3 cell lineage: a model system for studying programmed cell death, Notch/Numb signaling, and sequential specification of ganglion mother cell identity. *J Comp Neurol* 481, 240-51.
- Lee, S., Lee, B., Joshi, K., Pfaff, S.L., Lee, J.W. and Lee, S.K. (2008) A regulatory network to segregate the identity of neuronal subtypes. *Dev Cell* 14, 877-89.
- Ma, Q. (2006) Transcriptional regulation of neuronal phenotype in mammals. *J Physiol* 575, 379-87.
- Macdonald, P.M., Ingham, P. and Struhl, G. (1986) Isolation, structure, and expression of even-skipped: a second pair-rule gene of *Drosophila* containing a homeo box. *Cell* 47, 721-34.
- Mardon, G., Solomon, N.M. and Rubin, G.M. (1994) dachshund encodes a nuclear protein required for normal eye and leg development in *Drosophila*. *Development* 120, 3473-86.
- Marques, G., Bao, H., Haerry, T.E., Shimell, M.J., Duchek, P., Zhang, B. and O'Connor, M.B. (2002) The *Drosophila* BMP type II receptor Wishful Thinking regulates neuromuscular synapse morphology and function. *Neuron* 33, 529-43.
- Mellerick, D.M., Kassis, J.A., Zhang, S.D. and Odenwald, W.F. (1992) castor encodes a novel zinc finger protein required for the development of a subset of CNS neurons in *Drosophila*. *Neuron* 9, 789-803.

- Mettler, U., Vogler, G. and Urban, J. (2006) Timing of identity: spatiotemporal regulation of hunchback in neuroblast lineages of *Drosophila* by Seven-up and Prospero. *Development* 133, 429-37.
- Miguel-Aliaga, I., Allan, D.W. and Thor, S. (2004) Independent roles of the dachshund and eyes absent genes in BMP signaling, axon pathfinding and neuronal specification. *Development* 131, 5837-48.
- Miguel-Aliaga, I., Thor, S. and Gould, A.P. (2008) Postmitotic specification of *Drosophila* insulinergic neurons from pioneer neurons. *PLoS Biol* 6, e58.
- Molyneaux, B.J., Arlotta, P., Menezes, J.R. and Macklis, J.D. (2007) Neuronal subtype specification in the cerebral cortex. *Nat Rev Neurosci* 8, 427-37.
- Novotny, T., Eiselt, R. and Urban, J. (2002) Hunchback is required for the specification of the early sublineage of neuroblast 7-3 in the *Drosophila* central nervous system. *Development* 129, 1027-36.
- Nusslein-Volhard, C., Kluding, H. and Jurgens, G. (1985) Genes affecting the segmental subdivision of the *Drosophila* embryo. *Cold Spring Harb Symp Quant Biol* 50, 145-54.
- O'Keefe, D.D., Thor, S. and Thomas, J.B. (1998) Function and specificity of LIM domains in *Drosophila* nervous system and wing development. *Development* 125, 3915-23.
- Park, D., Shafer, O.T., Shepherd, S.P., Suh, H., Trigg, J.S. and Taghert, P.H. (2008a) The *Drosophila* basic helix-loop-helix protein DIMMED directly activates PHM, a gene encoding a neuropeptide-amidating enzyme. *Mol Cell Biol* 28, 410-21.
- Park, D., Veenstra, J.A., Park, J.H. and Taghert, P.H. (2008b) Mapping peptidergic cells in *Drosophila*: where DIMM fits in. *PLoS One* 3, e1896.
- Pearson, B.J. and Doe, C.Q. (2003) Regulation of neuroblast competence in *Drosophila*. *Nature* 425, 624-8.
- Pignoni, F., Hu, B., Zavitz, K.H., Xiao, J., Garrity, P.A. and Zipursky, S.L. (1997) The eye-specification proteins So and Eya form a complex and regulate multiple steps in *Drosophila* eye development. *Cell* 91, 881-91.
- Romani, S., Jimenez, F., Hoch, M., Patel, N.H., Taubert, H. and Jackle, H. (1996) Kruppel, a *Drosophila* segmentation gene, participates in the specification of neurons and glial cells. *Mech Dev* 60, 95-107.
- Schneider, L.E., Roberts, M.S. and Taghert, P.H. (1993) Cell type-specific transcriptional regulation of the *Drosophila* FMRFamide neuropeptide gene. *Neuron* 10, 279-91.
- Shirasaki, R. and Pfaff, S.L. (2002) Transcriptional codes and the control of neuronal identity. *Annu Rev Neurosci* 25, 251-81.
- Skeath, J.B. and Doe, C.Q. (1998) Sanpodo and Notch act in opposition to Numb to distinguish sibling neuron fates in the *Drosophila* CNS. *Development* 125, 1857-65.
- Skeath, J.B. and Thor, S. (2003) Genetic control of *Drosophila* nerve cord development. *Curr Opin Neurobiol* 13, 8-15.
- Skeath, J.B., Zhang, Y., Holmgren, R., Carroll, S.B. and Doe, C.Q. (1995) Specification of neuroblast identity in the *Drosophila* embryonic central nervous system by gooseberry-distal. *Nature* 376, 427-30.
- Small, S., Blair, A. and Levine, M. (1996) Regulation of two pair-rule stripes by a single enhancer in the *Drosophila* embryo. *Dev Biol* 175, 314-24.
- Song, M.R., Sun, Y., Bryson, A., Gill, G.N., Evans, S.M. and Pfaff, S.L. (2009) Islet-to-LMO stoichiometries control the function of transcription complexes that specify motor neuron and V2a interneuron identity. *Development* 136, 2923-32.
- Sossin, W.S., Fisher, J.M. and Scheller, R.H. (1989) Cellular and molecular biology of neuropeptide processing and packaging. *Neuron* 2, 1407-17.
- Terriente Felix, J., Magarinos, M. and Diaz-Benjumea, F.J. (2007) Nab controls the activity of the zinc-finger transcription factors Squeeze and Rotund in *Drosophila* development. *Development* 134, 1845-52.

- Thor, S. and Thomas, J.B. (1997) The *Drosophila* islet gene governs axon pathfinding and neurotransmitter identity. *Neuron* 18, 397-409.
- Tran, K.D. and Doe, C.Q. (2008) Pdm and Castor close successive temporal identity windows in the NB3-1 lineage. *Development* 135, 3491-9.
- Tsuji, T., Hasegawa, E. and Isshiki, T. (2008) Neuroblast entry into quiescence is regulated intrinsically by the combined action of spatial Hox proteins and temporal identity factors. *Development* 135, 3859-69.
- Wightman, B., Ebert, B., Carmean, N., Weber, K. and Clever, S. (2005) The *C. elegans* nuclear receptor gene *fax-1* and homeobox gene *unc-42* coordinate interneuron identity by regulating the expression of glutamate receptor subunits and other neuron-specific genes. *Dev Biol* 287, 74-85.
- Yang, X., Bahri, S., Klein, T. and Chia, W. (1997) Klumpfuss, a putative *Drosophila* zinc finger transcription factor, acts to differentiate between the identities of two secondary precursor cells within one neuroblast lineage. *Genes Dev* 11, 1396-408.
- Zheng, X., Chung, S., Tanabe, T. and Sze, J.Y. (2005) Cell-type specific regulation of serotonergic identity by the *C. elegans* LIM-homeodomain factor LIM-4. *Dev Biol* 286, 618-28.
- Zheng, X., Wang, J., Haerry, T.E., Wu, A.Y., Martin, J., O'Connor, M.B., Lee, C.H. and Lee, T. (2003) TGF-beta signaling activates steroid hormone receptor expression during neuronal remodeling in the *Drosophila* brain. *Cell* 112, 303-15.

## FIGURE LEGENDS

### Figure 1.

#### **Expression of FMRFa in 18-h AEL VNC and genetic pathways specifying the Ap4 neuron.**

- (A) Staining for FMRFa in VNC. The six Ap4 neurons reside laterally and project axons to the midline and exit the VNC into the DNH. The SE2 neurons reside anteriorly and project axons posteriorly along the dorsal midline. Anterior is up.
- (B) Previous studies have identified several markers specifically expressed in subsets of Ap neurons, acting to specify their identities (see text for references).
- (C) Part of the genetic cascade acting to specify the Ap4 neuron. *cas* triggers *col*, *sqz*, *grh* and *dac* expression late in the lineage. *col* plays a critical early role in establishing a “generic” Ap neuron fate in all four Ap neurons, by activating *ap* and *eya*. *ap*, *eya*, *grh* and *dac* specify the terminal Ap4 fate. See text for details.

### Figure 2.

#### **SE2 and Ap4 are generated by different Neuroblasts.**

Expression of *lbe(K)-Gal4* reveals the NB 5-6 lineage. Overlap of FMRFa (green) and the *lbe(K)-Gal4* (magenta).

(A-A'') As previous reports indicate, the Ap4/FMRFa neuron (green; arrowhead) expresses *lbe(K)-Gal4* (magenta) and therefore is generated by NB5-6. See results for references.

(B-B'') SE2/FMRFa cells (green; arrowheads) do not express *lbe(K)-Gal4* (magenta), indicating that SE2 neurons do not arise from NB5-6.

(C-H) Overlap of SE2/FMRFa (green) and different NB markers (magenta): *engrailed-Gal4* (C), *gooseberry-lacZ* (D), *mirror-lacZ* (E), *huckebein-lacZ* (F), *eagle-Gal4* (G) and Seven-up antibody (H). SE2/FMRFa cells do not express any of them.

Genotype: (A-B) *lbe(K)-Gal4>UAS- nmEGFP*; (C) *engrailed-Gal4>UAS- nmEGFP*; (D) *gooseberry-lacZ*; (E) *mirror-lacZ*; (F) *huckebein-lacZ*; (G) *eagle-Gal4>UAS- nmEGFP*.

Anterior is up in all images.

### Figure 3.

#### SE2 neurons are not specified in Cas/Grh window.

SE2/FMRFa expression in control (A), *kr<sup>1</sup> kr<sup>CD</sup>* (B and I; the images show two examples of ectopic SE2 cells), *pdm<sup>(Df(2L)ED773)</sup>* (C), *cas<sup>A1</sup>/cas<sup>A1</sup>* (D), *grh<sup>IM</sup>/grh<sup>IM</sup>* (E), and *elav-Gal4>UAS-Hb* (F).

(G) Overlap of SE2/FMRFa (green) and Hunchback (magenta). Arrowheads indicate SE2. Hunchback is expressed by SE2 cells (G'').

(H) Overlap of SE2/FMRFa (green) and Krüppel (magenta). Arrowheads indicate SE2. Krüppel is not expressed by SE2 cells (H'').

(I) Overlap of SE2/FMRFa (green) and Hunchback (magenta) in *kr* mutants. Arrowheads indicate the bonafide and ectopic SE2 cells.

(J) Cartoon summarizing our findings. See text for details and references.

(K) Quantification of observed phenotypes in *kr* mutants (n=54 VNCs). The number of SE2 neurons is not significantly affected in *kr* mutants (Student t-test,  $p < 0.01$ ).

(L) Quantification of observed phenotypes in *elav-Gal4>UAS-Hb* and *svp<sup>1</sup>/svp<sup>1</sup>* mutants (n=10 and 16 VNCs respectively). The number of SE2 neurons is significantly affected in both genetic backgrounds mutant (Student t-test,  $p < 0.01$ ).

Anterior is up in all images.

### Figure 4.

#### Dimm is the only Ap4 marker expressed by SE2 neurons

(A-E) Overlap of FMRFa (green) and different Ap4 markers (magenta): Eya (A), *sqz* (B), Dac (C), Dimm (D), pMad (E). Arrowheads always indicate SE2 neurons and Ap4 is shown in the inset. To avoid overlapping merged images, in (B, C and E) the confocal stack is focusing at the SE2 level. Only Dimm is expressed by SE2 neurons.

(F-G) Overlap of FMRFa (green) and two Ap1 markers (magenta): Nplp1 (F) and Col (G). Arrowheads indicate SE2 neurons and Ap1 is shown in the inset. Neither Nplp1 nor Col are expressed by SE2 cell.

(H) Summary of the expression of Eya, Sqz, Dac, Dimm, pMad, Nplp1 and Col within the Ap cluster and SE2 neurons of the stage-18 embryonic *Drosophila* VNC. Note that only Dimm is expressed in all FMRFa neurons.

Anterior is up in all images.

Genotypes: (A,C-G) *orizo2*. (B) *sqz-Gal4>UAS-GFP*.

## Figure 5.

### Ap4 and SE2 neurons use different genetic combinatorial code.

Expression of FMRFa in SE2 neurons in mutants for genes involved in Ap4 specification:

(A) Control; (B) *ap<sup>P44</sup>/ap<sup>P44</sup>*; (C) *dac<sup>4</sup>/dac<sup>4</sup>*; (D) *sqz<sup>ie</sup>/sqz<sup>ie</sup>*; (E) *eya<sup>Cli-III</sup>/eya<sup>10</sup>*; (F) *dimm<sup>P1</sup>/dimm<sup>P1</sup>*; (G) *nab<sup>SH143</sup>/nab<sup>R52</sup>*, (H) *elav-Gal4>UAS-Glued<sup>DN</sup>* and (I) *col<sup>1</sup>/col<sup>3</sup>*. In

all genotypes the number of SE2/FMRFa cells is normal. Ap4 neuron is shown in the inset. Anterior is up in all images.

(J) Quantification of observed phenotypes (n≥10 VNCs in all genotypes). The SE2 neuron number is not significantly affected in any mutant (Student t-test, p< 0.01). See genotypes above.

(K) Quantification of staining intensity of FMRFa in control and *dimm* mutants (n=11 VNCs). The asterisk (\*) denotes that staining intensity of FMRFa is significantly affected in *dimm* mutants (Student t-test, p ≤0.01). Wild type and mutant VNCs were stained and analyzed on the same slide.

## Figure 6.

**The Leucokinin system also uses different genetic combinatorial codes to produce distinct neuron subtypes.**

(A) Staining for Leucokinin in 18h AEL VNC. The fourteen ABLKs neurons reside laterally in abdominal segments (A1-A7) whereas SELKs neurons occupy middle positions in the first suboesophageal segment.

(B and B') Overlap of *Gsb-lacZ* (green) and Leucokinin (magenta). (B) SELK cells (arrowhead) do not express *Gsb-lacZ*. (B') ABLK (arrowhead) expresses *Gsb-lacZ*.

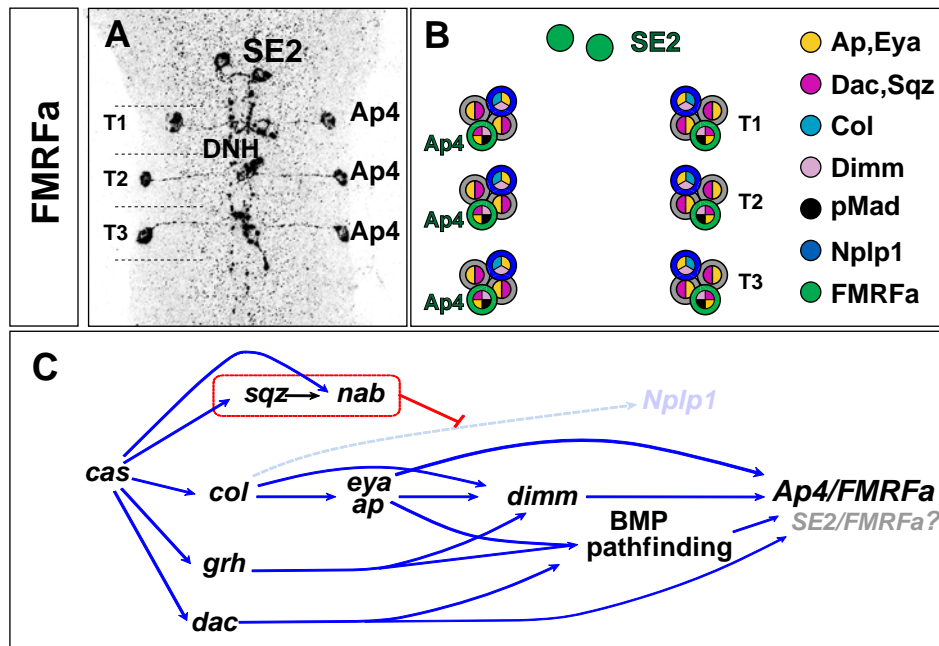
(C- C'') Overlap of Dimm (green) and Leucokinin (magenta). SELK neurons (encircled with dashed line) are Dimm-negative.

(D-O) Leucokinin expression in Control (D), *cas<sup>Δ1</sup>/cas<sup>Δ1</sup>* (E), *elav-Gal4>UAS-Cas* (F), *grh<sup>IM</sup>/grh<sup>IM</sup>* (G), *spdo<sup>G104</sup>* (H), *elav-Gal4>UAS-Notch<sup>intra</sup>* (I), *elav-Gal4>UAS-p35* (J), *nab<sup>SH143</sup>/nab<sup>R52</sup>* (K), *sqz<sup>ie</sup>/sqz<sup>ie</sup>* (L), *elav-Gal4>UAS-Nab, UAS-Sqz* (M), *jumu<sup>11.683</sup>* (N), and *klu<sup>212IR51C</sup>* (O). ABLK neuron is shown in the inset.

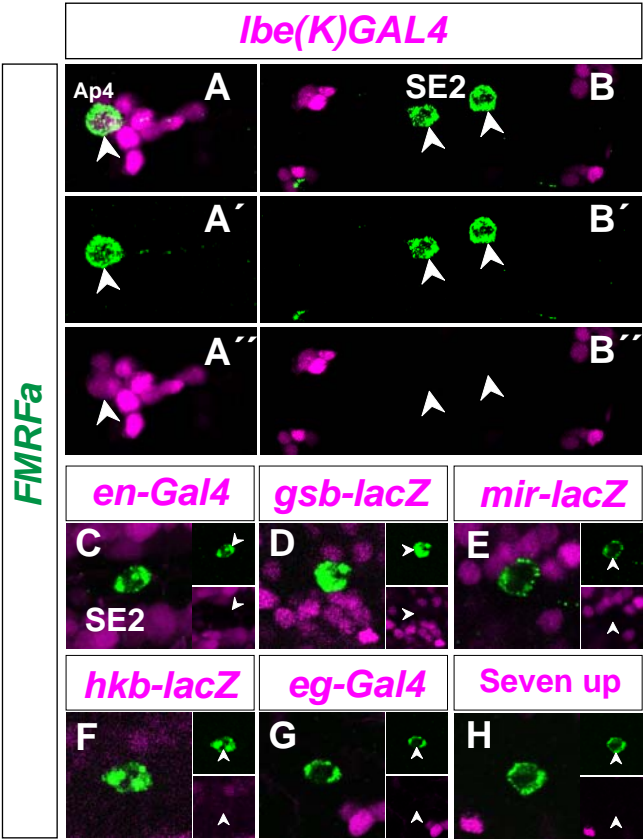
(P) Quantification of observed phenotypes (n≥20 hemineuromeres in all genotypes).

Asterisks (\*) denote significant difference compared to control (Student t-test, p< 0.01).

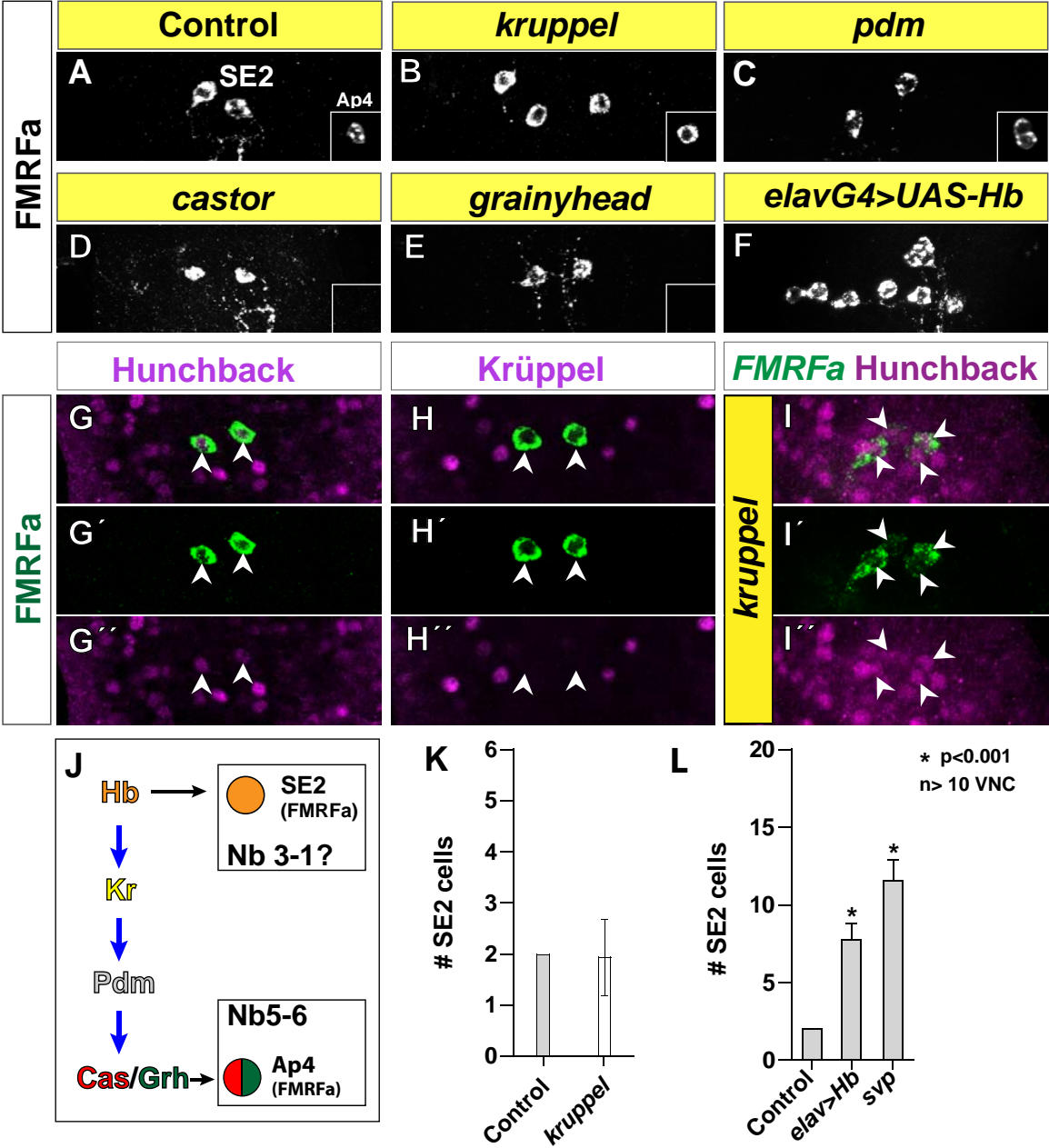
See genotypes above. Anterior is up in all images.



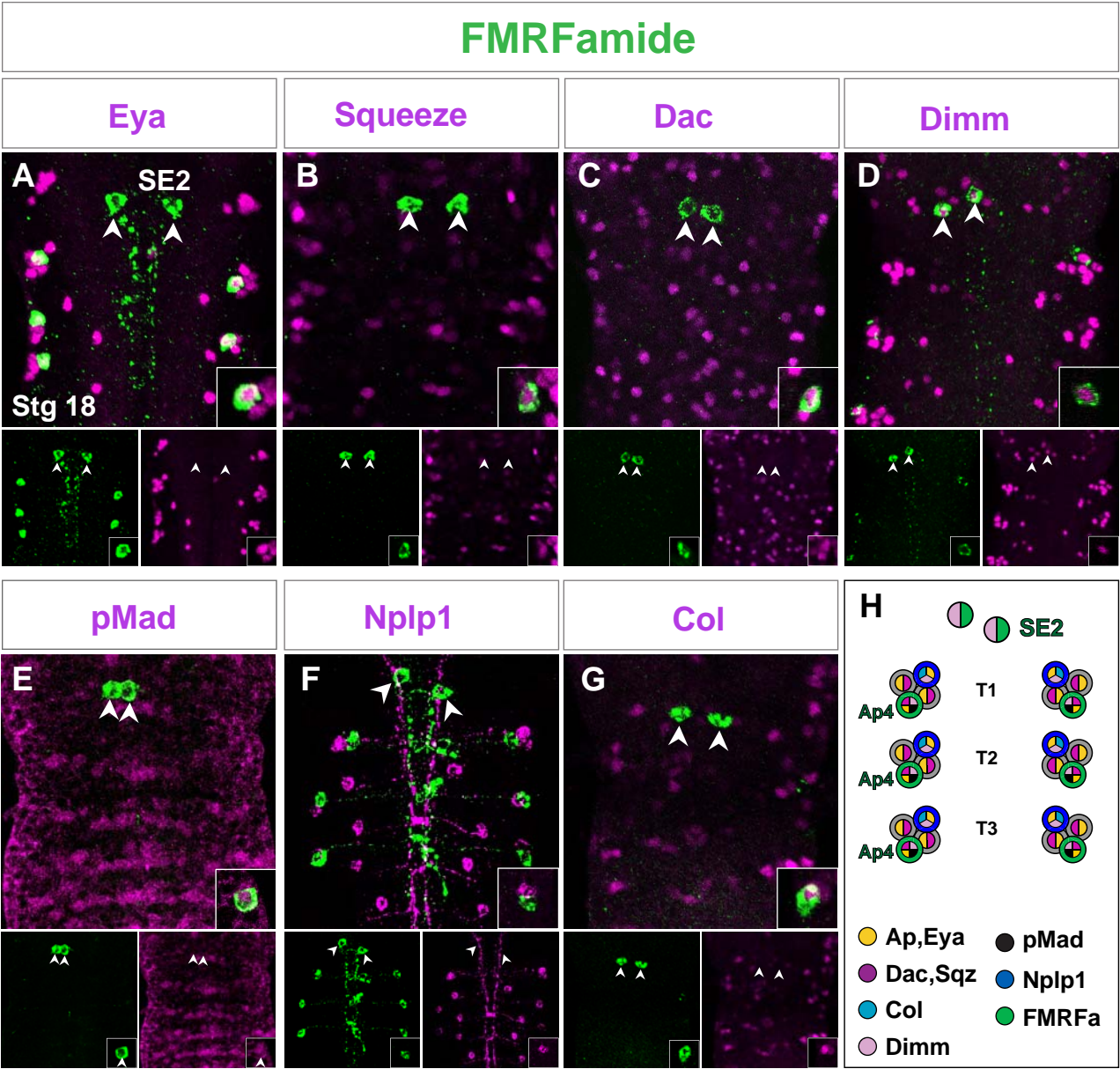
# Figure 2



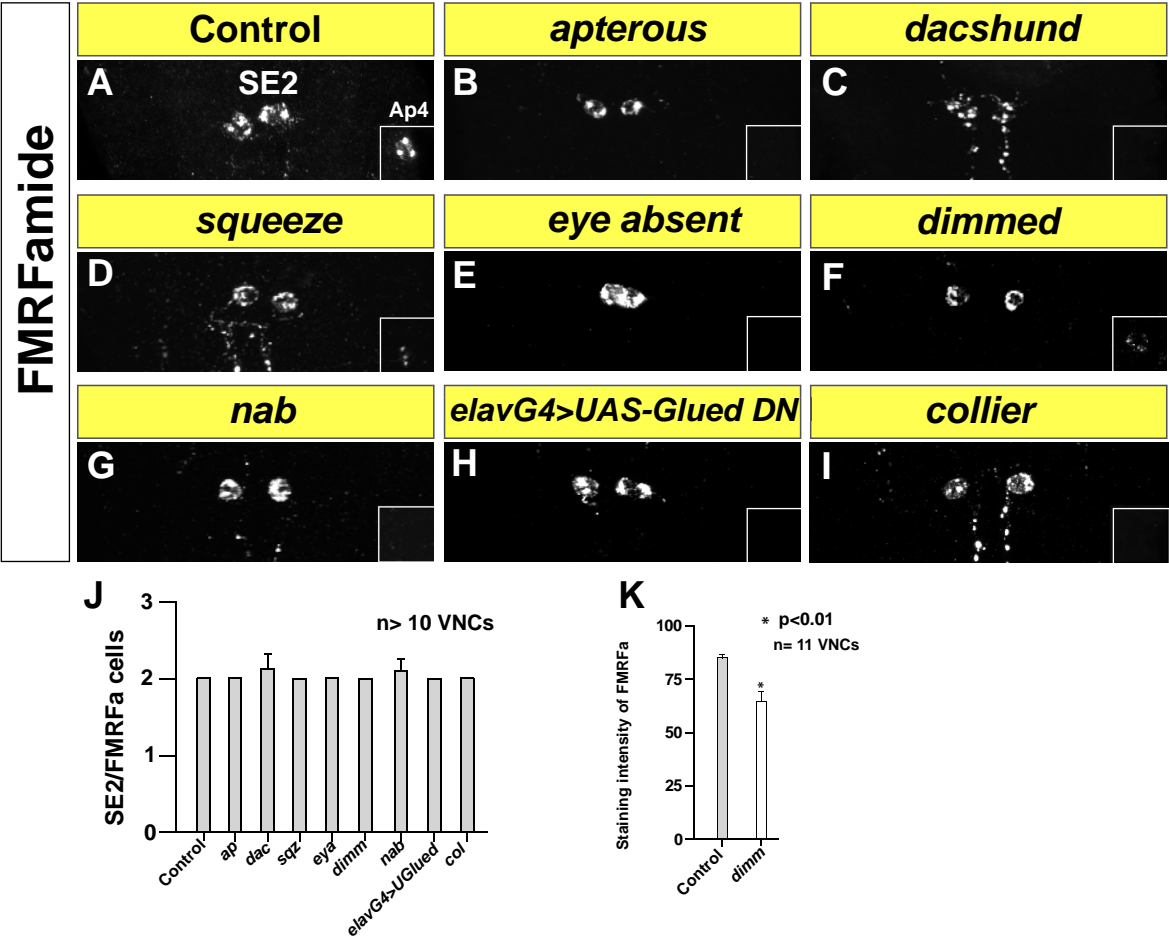
# Figure 3



# Figure 4



# Figure 5



# Figure 6

