Post-transcriptional regulation of transcription factor codes in immature neurons drives neuronal diversity

Wenyue Guan, Stéphanie Bellemin, Mathilde Bouchet, Lalanti Venkatasubramanian, Camille Guillermín, Anne Laurençon, Chérif Kabir, Aurélien Darnas, Christophe Godin, Séverine Urudy, et al.

To cite this version:
Wenyue Guan, Stéphanie Bellemin, Mathilde Bouchet, Lalanti Venkatasubramanian, Camille Guillermín, et al.. Post-transcriptional regulation of transcription factor codes in immature neurons drives neuronal diversity. Cell Reports, 2022, 39 (13), pp.110992. 10.1016/j.celrep.2022.110992. hal-03880028

HAL Id: hal-03880028
https://hal.science/hal-03880028
Submitted on 6 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License
Post-transcriptional regulation of transcription factor codes in immature neurons drives neuronal diversity

Highlights

- 29 leg MNs from the same neuroblast express 19 combinations of TFs (TF codes)
- The TF codes are expressed as a function of MN birth order
- Imp and Syp post-transcriptionally regulate at least five TFs
- Changing Imp and Syp level modifies nine TF codes and the axonal targeting accordingly

Authors

Wenyue Guan, Stéphanie Bellemin, Mathilde Bouchet, ..., Séverine Urdy, Richard S. Mann, Jonathan Enriquez

Correspondence

jonathan.enriquez@ens-lyon.fr

In brief

Guan et al. discovered that the individual morphological fates of motoneurons are determined by the transcription factor mRNAs they express together with gradients of RNA-binding proteins that determine which TF mRNAs are translated, ultimately producing a set of TF codes that are unique for each neuron.
Post-transcriptional regulation of transcription factor codes in immature neurons drives neuronal diversity

Wenyue Guan,1 Stéphanie Bellemin,1,4 Mathilde Bouchet,1,4 Lalanti Venkatasubramanian,2,4 Camille Guillermin,1 Anne Laurençon,1 Chérif Kabir,1 Aurélien Darnas,1 Christophe Godin,3 Séverine Urdy,1 Richard S. Mann,2 and Jonathan Enriquez1,5,*

1Institut de Génomique Fonctionnelle de Lyon, UMR5242, Ecole Normale Supérieure de Lyon, Centre National de la Recherche Scientifique, Université Claude Bernard-Lyon 1, 46 Allée d’Italie, 69364 Lyon Cedex 07, France
2Departments of Biochemistry and Molecular Biophysics, and Neuroscience, Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY 10027, USA
3Laboratoire Reproduction et Développement des Plantes, Univ Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRA, Lyon, France
4These authors contributed equally
5Lead contact
*Correspondence: jonathan.enriquez@ens-lyon.fr
https://doi.org/10.1016/j.celrep.2022.110992

SUMMARY

How the vast array of neuronal diversity is generated remains an unsolved problem. Here, we investigate how 29 morphologically distinct leg motoneurons are generated from a single stem cell in Drosophila. We identify 19 transcription factor (TF) codes expressed in immature motoneurons just before their morphological differentiation. Using genetic manipulations and a computational tool, we demonstrate that the TF codes are progressively established in immature motoneurons according to their birth order. Comparing RNA and protein expression patterns of multiple TFs reveals that post-transcriptional regulation plays an essential role in shaping these TF codes. Two RNA-binding proteins, Imp and Syp, expressed in opposing gradients in immature motoneurons, control the translation of multiple TFs. The varying sensitivity of TF mRNAs to the opposing gradients of Imp and Syp in immature motoneurons decrypts these gradients into distinct TF codes, establishing the connectome between motoneuron axons and their target muscles.

INTRODUCTION

Locomotion is a stereotyped behavior used by animals to find food, mates, or escape from predators. The rhythmic and precise pattern of locomotion is linked to the coordinated contraction of muscles that are innervated by a complex wiring of motoneuron (MN) axons controlling the timing and the intensity of muscle activity. One central challenge is to decipher how stem cells generate a huge diversity of MN morphologies. Here, we used a combination of genetic and computational approaches to understand how, in Drosophila, a single stem cell gives rise to 29 morphologically unique MNs.

Transcription factors (TFs) are central regulators of the morphological specification of neurons, including MNs. In vertebrates, during development, Hox6 and Hox10 paralogs at the brachial and lumbar levels of the spinal cord distinguish MNs that target leg muscles from those that target body wall muscles. Subsequently, limb-targeting MNs are further refined into pools, where all MNs in a single pool target the same muscle. Each pool is molecularly defined by the expression of pool-specific TFs (Dasen and Jessell, 2009; Philippidou and Dasen, 2013). In Drosophila embryos, subclasses of MNs innervating the larva body-wall muscles are also morphologically specified by unique combinations of TFs controlling ventral versus dorsal targeting (Fujikawa et al., 2003; Garces and Thor, 2006; Landgraf et al., 1999; Broihier and Skeath, 2002; Broihier et al., 2004; Certel and Thor, 2004; Oyallon et al., 2012; Thor and Thomas, 1997; Thor et al., 1999). The morphological specification of Drosophila leg MNs seems to be controlled at the single-cell level by a combination of morphology-specifying TFs (mTFs). Seven MNs innervating limb appendages differentially express unique combinatorial codes of five mTFs that determine most of their morphological fates (Enriquez et al., 2015).

Immature neurons expressing the TF codes are generated by dedicated stem cells. These stem cells are regulated in space and time to generate an enormous amount of morphological diversity at the right time and in the right place (Kohwi and Doe, 2013; Sagner and Briscoe, 2019). Drosophila neurons are generated by neuroblasts (NBs), specialized stem cells dedicated to the generation of neurons and glia (Doe and Skeath, 1996; Prökop and Technau, 1991; Truman and Bate, 1988). As they divide, NBs express a temporal sequence of TFs (tTFs) that contribute to the generation of neuronal diversity. In the embryonic ventral nerve cord (VNC) (analogue of the human spinal cord), most
Figure 1. Lin A, a model to study how the MN-muscle connectome is built
(A) Drawing of a fly showing the CNS and Lin A/15 MNs.
(B) VNC with six Lin A/15 labeled with myr:GFP (green) and immunostained with anti-BRP (blue). Arrowheads indicate MN cell bodies.
(C) T1 leg with Lin A/15 axons labeled with myr:GFP (green).
(D) T1 leg with a Lin A/15 MARCM clone expressing mCD8:GFP (green) and all muscles were labeled with Mhc-RFP (red). The number of MNs innervating the corresponding leg segment are indicated in (C). Leg muscles innervated by Lin A/15 are indicated in (D2).
(E) Schematic showing the link between birth order and muscle targeting, modified from Baek and Mann (2009) (see link between birth order and muscle targets in STAR Methods). Top: schematic of the cell body of LinA/15 iMNs is shown. The numbers inside indicate their birth order, and the abbreviations below indicate the name of the MNs based on the nomenclature from Baek and Mann (2009). Bottom: schematic of a T1 leg innervated by Lin A/15 is shown; the muscles innervated (legend continued on next page)
NBs express a sequence of five tTFs (Isshiki et al., 2001; Li et al., 2013a), whereas in medulla NBs of the visual system and intermediate neural progenitors of the Drosophila larval brain, a different series of tTFs have been described (Bayraktar and Doe, 2013; Konstantinides et al., 2022; Li et al., 2013b). In vertebrates, neural stem cells, e.g., in the cerebral cortex, retina, and spinal cord, use analogous strategies, suggesting that the regulatory logic of tTFs is evolutionarily conserved (Alsió et al., 2013; Delile et al., 2019; Elliott et al., 2008; Jacob et al., 2008; Mattar et al., 2015; Okano and Temple, 2009).

In Drosophila, neuronal diversity can be generated by NBs as they age via a second mechanism involving two RNA-binding proteins (RBPs), insulin growth factor (IGF)-II mRNA-binding protein (Imp) and Syncrip (Syp). These RBPs are expressed in opposing temporal gradients in brain NBs from high to low and low to high, respectively (Liu et al., 2015; Ren et al., 2017; Syed et al., 2017). In the NBs of the mushroom body, a structure in the brain processing olfactory inputs, these two RBPs regulate the translation of the TF Chinmo, which in turn controls the temporal identity of the mushroom-body NB progeny in a concentration-dependent manner. This gradient strategy allows the generation of many neurons of the same type, within a particular time window, as observed in mammalian progenitors.

Here, we used a lineage called Lin A/15 that produces 29 of the ~50 MNs with unique morphologies (Baek and Mann, 2009; Enriquez et al., 2018) as a model and discovered a mechanism controlling the generation of a large amount of neuronal diversity in a short time window by a single stem cell. We report the discovery of 16 TFs expressed in combination and determining 19 different TF codes in immature MNs (iMNs) just before their morphological differentiation. These TF codes are not established at the birth of the iMNs but are gradually shaped during development. By comparing the expression profile of six mRNAs with their corresponding proteins, we revealed that post-transcriptional regulation of TFs plays a key role in shaping the TF codes. We then examined the function of two known RBPs, Imp and Syp, in shaping the TF code in iMNs. We first demonstrated that, in earlier born iMNs, Syp protein was undetectable and the level of Imp was high, whereas in the later born iMNs, the level of Syp protein was high and Imp was low. We discovered that this opposite expression pattern in iMNs of both RBPs is essential to shape the axon-muscle connectome by regulating the translation of at least five TFs examined. Finally, we revealed that Imp controls the fate of the Lin A/15 progenies directly in iMNs. With these results, we propose a model where RBPs act as temporal factors to specify the fate of iMNs by post-transcriptionally sculpting a complex set of TF codes.

RESULTS

Lin A/15, a model to study how the MN-muscle connectome is built during development

Previous studies have characterized the morphologies of individual Lin A/15 MNs by MN-driven GFP labeling, whereby the presumptive muscle target is identified through the localization of the terminal branches in the leg (Baek and Mann, 2009; Brierley et al., 2009, 2012).

Here, we extend these earlier studies by driving the expression of RFP in all leg muscles (Mhc-RFP), in addition to driving the expression of GFP in 28 Lin A/15 MNs with the MARCM technique (Lee and Luo, 1999; Figure 1D). Lin A/15 MNs innervate 9 out of the 14 leg muscles: all five muscles in the tibia, three muscles in the femur, and one muscle in the trochanter. The first-born MN from Lin A/15, which cannot be visualized with the “one-spot” MARCM technique (see STAR Methods), innervates a body-wall muscle and can be genetically labeled with GFP by using a Lin A/15 tracing system (Figures 1A–1C; see STAR Methods). Based on this and previous studies, the tight correlation between the birth order of Lin A/15 MNs and their muscle targets can be schematically summarized (Figure 1E; see STAR Methods).

Correlation between birth order and the spatial organization of iMNs

The late third instar (L3) is a key developmental stage between the end of MN production and the establishment of axon-muscle connections (Baek et al., 2013; Venkatasubramanian et al., 2019). We thus chose this stage as an entry point to understand how MN diversity is generated.

We first determined how the spatial organization of iMNs is related to birth order (Figures 1F–1K). In late L3, the NB has a ventral location, whereas iMNs are located more dorsally and anteriorly (Figures 1G and 1H). When larvae are fed with the DNA-label EdU at early time points, the last-born MNs (EdU+) are located ventrally and posteriorly near to the NB, whereas older MNs (EdU−) are located dorsally and anteriorly distant from the NB. Moreover, EdU+ iMNs and EdU− iMNs are barely intermingled (Figures 1I–1K). This demonstrates that, up to the end of L3, iMNs maintain distinct spatial positions that are dictated by their birth dates. Importantly for our study, the most ventral iMNs, which are the last-born iMNs, undergo

by Lin A/15 are color coded based on their innervation. The line between the cell body and leg muscles indicates the relationships between MN birth order and muscle innervation. Muscle nomenclature is based on Soler et al. (2004): d, depressor; f, femur; l, levator; m, muscle; r, reductor; t, tendon; ta, tarsus; tI, tibia; tr, trochanter.

(F) Drawing of the anterior region of an L3 larva showing the CNS and Lin A/15 iMNs. (G and H) 3D reconstruction of six (G) and one (H) right thoracic hemisegment 2 (T2R) Lin A/15 in an L3 larva labeled with myr:GFP (green). Ventral (G) and lateral (H) views are shown; axes: A, anterior; L, lateral; V, ventral. (H2) Plot of the relative position of each Lin A/15 cell in (H1) from a lateral perspective is shown. Lin A/15 proliferative glia (PG) are in white, iMNs are in blue, GMCs are in white, and the NB is in cyan. (H3–H6) Confocal sections of the Lin A/15 in (H1) and (H2) immunostained with anti-Elav (blue) and Dpn (cyan).

(I–K) Plots of the relative position of each Lin A/15 cell from a lateral perspective in L3 larvae fed with EdU at indicated time points: EdU+ cells (red) and EdU− cells (blue). The horizontal axis indicates the time point of EdU feeding. On the right of each graph, confocal sections of Lin A/15 are shown.

(L) Schematic of Lin A/15 cell bodies in an L3 larva and of an adult leg that shows the correlation between the position of the iMNs and the muscle targeting of adult MNs.
apoptosis during the pupal stages (Guán et al., 2021), whereas the first-born 29 iMNs morphologically differentiate during metamorphosis (Figure 1L). This clear relationship between birth order and the spatial organization of iMNs was then used to understand how iMNs are morphologically specified by TFs.

A complex combination of TFs in iMNs prefigures morphological diversity of LinA/15 MNs

To identify the TFs expressed in Lin A/15 iMNs at late L3 that could control the stereotypic wiring of muscle innervation, we performed an immunostaining screen of GFP-expressing iMNs with a collection of around 220 antibodies directed against different TFs, covering ~35% of all TFs coded by the Drosophila genome. In total, 45 TFs were identified in Lin A/15, with 16 TFs expressed in different subpopulation of iMNs (Figures 2A–2G illustrate the example of Jim and Figure S1 other TFs). We then focused our study on these 16 TFs, since their differential expressions could be at the origin of the large diversity of MN morphologies generated by Lin A/15 NB.

We therefore developed a method based on the correlation between the relative position of iMNs from Lin A/15 NB and their birth order to assign the correct combination of TFs to each iMN (see STAR Methods). We named this method the positive cell cluster detection (PCCD) method. In this method, x, y, and z coordinates and on and off expression of a given TF were assigned to each Lin A/15 MN for at least 15 Lin A/15 samples per immunostaining (Figure 2H illustrates two Lin A/15 samples with anti-Jim immunostaining). With these coordinates, each Lin A/15 MN is ranked along an x’ axis according to its relative distance from the NB, such that the lowest rank (1) corresponds to the greatest distance from the NB (Figure 2I illustrates the MN ranking of two samples of anti-Jim immunostaining). Then, the frequency expression is calculated as a function of x’ (Figure 2J illustrated frequency of Jim expression in 15 samples analyzed). The distributions of frequencies are smoothed by application of a Savitzky–Golay filter. The peak of each distribution is identified, and the positive cell cluster of iMNs associated with each peak is assigned (Figure 2K illustrates the identification of iMN cluster expressing Jim; Figure 2L shows similar cluster predictions for all other TFs). We schematized our results for the 29 first-born Lin A/15 iMNs (Figure 2M). One TF (Eyeless) was only expressed in the supernumerary iMNs eliminated by apoptosis during pupal stages (Guán et al., 2021). To validate and refine this analysis, 12 co-immunostainings for different TFs were performed (Figure 2N illustrates Chinmo and Jim co-staining; Figure S2 shows all co-stainings). The result of the co-stainings reveals that PCCD is very accurate, because in most cases, none or only one cell correction was required (STAR Methods; compare Figure 2M versus Figure 2O). In summary, we could assign 19 different TF codes for 29 iMNs: 10 iMNs shared unique TF codes, 16 iMNs share a code with other iMNs born sequentially, and 3 iMNs born in the same time window have similar codes (Figure 2O).

TF-code diversity is progressively established in iMNs

We then determined how the expression pattern of the 15 differentially expressed TFs is established by analyzing their expression from the end of L2, when the NB begins to proliferate, until late L3.

We classified Lin A/15 TFs into three categories based on their spatial and temporal expression dynamics. The first category encompasses 10 TFs whose expression starts in the ganglion mother cell (GMC) or postmitotic neurons (Figures 3A–3H and S5; note that Pros and FoxP were placed in this category, even though they were also detected in the NB cytoplasm). The expression dynamics of the TFs in this category were further subdivided into two sub-categories based on their temporal expression dynamics. Category 1.1 is composed of seven TFs expressed and maintained by subpopulations of iMNs since their birth (Lov, RunxA, Zfh2, Kr, and Zfh1) or whose expression turns
on after a delay (Jim and FoxP) (Figures 3A–3D and S3). Three TFs from the other sub-category (category 1.2) are expressed in all new-born iMNs (Oli, Cas, and Pros). However, their expression is only maintained in a subpopulation of iMNs (Figures 3E–3H and S3). The second category (category 2) is composed of Br, Run, and Nvy, which are TFs expressed continuously in the Lin A/15 NB from late L2 until late L3. All new-born iMNs expressed these TFs, but their expression is maintained in different groups of iMNs (Figures 3I–3L and S3). Last, the third category (category 3) is composed by Chinmo and Mamo, which are the only TFs having a temporal expression in the NB (Figures 3M, 3O, and S3). As described for other lineages (Dillard et al., 2018; Maurange et al., 2008; Syed et al., 2017), Chinmo is expressed in the NB in early stages and only maintained in the first-born MNs. However, the expression of Mamo in the NB is completely uncoupled from its expression in iMNs, suggesting that Mamo function in NB versus MNs might not be linked (Figure S3).

These results showed that the TF codes are gradually established in iMNs by a de novo expression of TFs in GMCs and iMNs (category 1) and by the selective maintenance and/or repression of TFs in iMNs (categories 1, 2, and 3). These expression dynamics suggest the existence of upstream regulators shaping the TF codes by induction and maintenance and/or repressive mechanisms.

Post-transcriptional regulation governs the establishment of TF codes

In view of the complex dynamics of TF protein expression, we wondered to which extent it could be shaped by post-transcriptional regulation. To address this question, we analyzed the RNA expression pattern of six TFs selected as representatives of the protein-expression profiles categorized above: jim and oli (category 1), br and nvy (category 2), and mamo and chinmo (category 3).

We performed single-molecule fluorescence in situ hybridization (smFISH) and developed a pipeline to quantify the number of RNA spots per Lin A/15 iMN in L3 VNC (Figures 4A–4I and S4). This pipeline combined 3D segmentation of Lin A/15 cells (Machado et al., 2019) and a computational quantification of the number of RNA spots in each Lin A/15 cell (Raj et al., 2008; see STAR Methods). The number (N) of RNA dots detected in each iMN as a function of the iMN’s relative distance to the NB was calculated (Figures 4E–4I and S4). The number of RNA were also calculated for clusters of contiguous iMNs (N<sup>x’</sup>–<sup>x”</sup>), where x’ and x” are the iMN furthest and closest to the NB in a selected cluster. For example, jim RNA was expressed at higher levels in ventral (young) iMNs proximal to the NB (N<sup>x’</sup> = 52 and N<sup>x”</sup> = 54) than the dorsal (old) iMNs (N<sup>x’’</sup> = 27) and was barely detectable in the NB (N<sup>NB</sup> = 9). This ubiquitous expression of jim RNA from high to low in young and old iMNs, respectively, contrasts with the expression of the Jim protein detected in only eight iMNs (Figure 2) and not detected at all in young iMNs, including in the supernumerary MNs removed during pupal stages, where jim RNA is the most expressed (Figure 2 versus Figure 4). The RNA of chinmo, mamo, br, nvy, and oli were also present ubiquitously in iMNs, while the proteins were detected in subpopulation of iMNs (Figure 2 versus Figures 4F–4I and S4 and Table S1).

The comparison of the RNA expression pattern of six TFs with their respective proteins suggest that post-transcriptional regulation plays a key role in shaping TF codes in iMNs.

Opposite spatial gradients of Imp and Syp control the specificity of axonal targeting

To investigate further post-transcriptional regulation, we investigated the expression dynamics and function on muscle innervation of two known RBPs, Imp and Syp, that are major players in the temporal specification of VNC and central brain lineages in Drosophila (Liu et al., 2015; Ren et al., 2017; Syed et al., 2017). In Lin A/15 NB, both Imp and Syp proteins follow opposing temporal expression gradients in accordance with previously described studies of other NBs in the CNS (Guan et al., 2021).

In iMNs, Imp and Syp proteins have opposing gradients according to their birth date (Figure 4K). By analyzing the co-expression of Imp and Syp in iMNs with Jim (expressed in iMNs 16–23; Figures 2O and 4J), we concluded that Imp is highly expressed in iMNs 1–23 and weakly expressed in iMNs 24–29, whereas Syp is highly expressed in iMNs 24–29 in gradient manner and not detected in older iMNs (Figure 4L). Moreover, the detection of primary transcripts of imp and syp by intronic probes reveals that both RBPs are actively transcribed in iMNs and not only passively inherited from the NB (Figure S5). These results suggest that opposite gradients of Imp and Syp in iMNs result from their inheritance from the NB and their de novo transcription.

We next examined the function of Imp and Syp in shaping the correct connectome between Lin A/15 axons and leg muscles (Figure 5). Distal region of the tibia, which in wild-type (WT) Lin A/15 is innervated by last-born MNs expressing a low level of Imp during development, was not affected in imp<sup>–/–</sup> MARCM clones. By contrast, all the other regions were less innervated (Figures 5A, 5F, 5K, and 6; Table S2). In syp<sup>–/–</sup> MARCM clones, an opposite phenotype was observed. The distal region...
of the tibia, innervated by last-born MNs expressing a high level of Syp during development, was no longer innervated by Lin A/15 axons, whereas other leg regions were correctly innervated. Interestingly, we observed more axonal branches than normal in the distal region of the femur in Syp−/−, suggesting that abolishing Syp function is sufficient to re-direct axonal targeting from distal tibia to distal femur (Figures 5B, 5G, and 5K; Table S2). In Syp−/− clones, more neurons are produced (Figure S6), probably due to the survival of the supernumerary and a late decommissioning of the NB as observed in Syp RNAi LinA/15 (Guan et al., 2021). The ectopic expression of Imp in all Lin A/15 cells, including the Syp+ cells (which normally express a low level of Imp), induced a similar innervation phenotype and an increase of the MN produced to that observed in Syp−/− MARCM clones (Figures 5C, 5H, and 5K; Table S2). The result highlights the capacity of Imp to inhibit iMNs to divide into distal tibia and to promote the targeting of the distal femur. Only a minor innervation phenotype (the talm is not innervated) resulted from the overexpression of Syp in Lin A/15 (Figure S6), suggesting a dorsal prevalence of Imp over Syp. Although, Imp and Syp negatively cross-regulate each other in the brain and VNC NBs (Liu et al., 2015; Ren et al., 2017; Syed et al., 2017), Imp and Syp mutual inhibition was not observed in our genetic manipulations in late L3 VNC, indicating that both RBPs act independently on specifying axonal targeting (Figure S7).

In summary, our results show that Syp inhibits iMN targeting of the distal femur and promotes iMN targeting of distal tibia, whereas Imp has the opposite effect, with Imp required in early-born MNs for the correct targeting of the femur and proximal tibia. Finally, the strong increase in axonal branching of the distal femur in Syp−/− and Imp-overexpressing Lin A/15 suggests that the surviving MNs are maybe integrated into the axon-muscle connectome.

**Imp and Syp shape the LinA/15 TF codes**

The function of Imp and Syp in shaping the axon-muscle connectome raised the question of whether they act through the post-transcriptional regulation of the TF codes. To address this question, we focused on those six TFs whose expression has been characterized at the level of the RNA and protein—Mamo, Chinmo, Br, Nvy, Jim, and Oli (Figures 2 and 4)—and analyzed changes in their expression patterns when the levels of Imp and Syp were modified.

We tested the epistatic relationships between Imp and Syp and the TF codes in Syp−/− and Imp-overexpressing Lin A/15 clones because both genetic backgrounds induced similar innervation phenotypes and were therefore anticipated to induce similar variations in TF expression. Br was expressed in fewer iMNs in the absence of Syp (Figures 6A, 6D, 6J, and 6Q), whereas it was unaffected by Imp overexpression (Figures 6A, 6G, 6M, and 6Q). Mamo expression was unaffected by the absence of Syp, but it was extended to more ventral iMNs with Imp overexpression (Figures 6C, 6F, 6L, 6O, and 6Q). The expression of Chinmo and Jim was extended to ventral iMNs and Nvy was expressed in fewer iMNs with both the absence of Syp and the overexpression of Imp (Figures 6A, 6B, 6D, 6E, 6G, 6H, 6J, 6K, 6M, 6N, and 6Q). Interestingly, only the ventral cluster of Nvy+ iMNs (Br+) were affected by Imp overexpression, but not the dorsal cluster of Nvy+ iMNs (Br−) (Figures 6Q and S7). Oli expression was unaffected by both the absence of Syp and the overexpression of Imp (Figures 6Q and S7).

Finally, by using VGlut-Gal4 to label iMNs with GFP, we could observe some of the iMNs that are eliminated by apoptosis during later stages (VGlut > GFP did not label all of them because the driver is expressed with a delay in iMNs). However, the weak expression of the GFP in late-born iMNs and their ventral localization allowed us to recognize the iMNs that are eliminated by apoptosis in a later stage and to reveal that they express a code similar to MNs that targets the distal femur (Jim+, Nvy− or low, Br low) in Syp−/− and Imp-overexpressing Lin A/15 (Figures 6D1–6D4, 6E1–6E4, 6G1–6G4, and 6H1–6H4). These observations combined with the excessive amount of extra targeting into distal femur in Syp−/− and Imp-overexpressing Lin A/15 (Figure 5) suggest that the surviving MNs are integrated in the axon-muscle connectome because their
fate is modified. In summary, Imp and Syp have opposite effects on the expression of several TFs: Imp promoted the expression of Chinmo, Mamo, and Jim proteins and inhibited the expression of Nvy, whereas Syp inhibited the expression of Chinmo and Jim and promoted the expression of Br and Nvy (Figures 6P and 6Q).

**TF code is functional in building the axon-muscle connectome**

The opposite effects of Imp and Syp on TF expression could explain why Syp⁻/⁻ and Imp-overexpressing Lin A/15 induced similar innervation phenotypes. To investigate such a possibility, we next analyzed the function of Chinmo, Br, Nvy, Jim, and

---

**Figure 5. Function of Imp, Syp, and TF codes in building axonal-muscle connectome**

(A–J) Axonal targeting phenotypes of WT tub > GFP⁺ P35 (A and E), WT VGlu > GFP (B–D), Imp⁻/⁻ tub > P35 (F), Syp⁻/⁻ (G), tub > Imp (H), nvy⁻/⁻ (I), and jim KD (J) Lin A/15. Axons are mCD8:GFP⁺ (green); muscles were labeled with Mhc-RFP (red) in (B), (D), (E), (G), (I), and (J). Insets show leg regions most affected. Arrowheads and arrows point to normal targeting (white), the absence or reduction of targeting (orange), and extra targeting (blue). Asterisks indicate an absence of trochanter targeting in Imp⁻/⁻ tub > P35 (F) that was not considered as part of the phenotypes because the WT tub-P35 showed similar defects (A). In WT tub > GFP⁺ P35 LinA/15, the supernumerary neurons targeted a new region in the coxa (n = 8/9). Note: for each indicated phenotype, a minimum of six legs are analyzed; see Table S2 for details.

(K) Schematic of the axonal targeting phenotypes. See Figure S6 for Imp⁻/⁻, br⁻/⁻, chinmo⁻/⁻, and Oli⁻/⁻ phenotypes and the number of Lin A/15 cells in each genetic background. See Figure 7 for UAS-Jim phenotype. See Table S2 for the penetrance.
Cell Reports
Article

WT VGlut>GFP

Syp/- VGlut>GFP

Br Chinmo
Nvy Jim
Mamo

Br Chinmo
Nvy Jim
Mamo

tub>Imp VGlut>GFP

Syp/- VGlut>GFP vs control

Br Chinmo
Nvy Jim
Mamo

tub>Imp VGlut>GFP vs control

Control
Syp/

Imp OE

Control
Syp/

Imp OE

Control
Syp/

Imp OE

Control
Syp/

Imp OE

Syp/

Imp OE

(legend on next page)
Oli in shaping the axon-muscle connectome (the attempts to analyze mamo were unsuccessful; Figure 5 and Figure S6).

Nvy is expressed in two different MN clusters. One of the clusters corresponds to the last-born MNs that mostly innervate the distal region of the tibia (Figure 2O). In nvy−/− clones, the distal region of the tibia is not or less innervated (Figures 5D, 5I, and 5K; Table S2). Br is expressed in MNs targeting the distal femur and distal tibia (Figure 2O). In br−/− clones, the distal femur is less innervated, a phenotype associated with a mistargeting of the tibia (Figure 5E; Table S2). In jim knockdown clones, the distal femur is less innervated by the four Jim-expressing IMNs that normally target this region (Figures 5E, 5J, and 5K; Table S2). In chinmo−/− clones, both the proximal and distal regions of the femur were less innervated (Figure S6; Table S2). This axonal targeting defect at the distal femur, a region normally innervated by MNs never expressing Chinmo, suggests a non-autonomous function of Chinmo in these MNs (see discussion). The MNs expressing Oli are all affected with variable penetrance in Oli−/− Lin A/15 MARCM clones (Figure S6; Table S2).

These phenotypes demonstrate that the TF code is functional in building the axon-muscle connectome.

**Imp and Syp specify MN axonal targeting through reprogramming TF codes in IMNs**

The phenotypic effects of Imp−/−, Syp−/−, or Imp-overexpressing (UAS-Imp) MARCM clones have more consistent effects on axon-muscle connectivity than removing a single TF (Figure 5 and Figure S6). The effects of Imp and Syp on axon-muscle connectivity are linked to the post-transcriptional regulation of multiple Lin A/15 TFs, thus explaining why removing one single TF cannot recapitulate the innervations phenotypes observed when levels of Imp and Syp are modified. However, some genetic manipulations modifying the level of TFs induced similar phenotypes to the ones generated in Syp−/− or Imp-overexpressing Lin A/15. For example, Nvy and Jim are TFs necessary for specifying axonal targeting into distal tibia and distal femur respectively (Figures 5I-5K). In Syp−/− or Imp-overexpressing Lin A/15, Nvy is repressed and Jim is ectopically translated in the IMNs that target distal tibia (Figures 6E, 6H, and 6Q), thus accordingly, the innervation in distal tibia is reduced and the axonal targeting at distal femur is increased (Figures 5G, 5H, and 5K).

In summary, the epistatic relationships of Imp and Syp with these TFs strongly imply that these RBPs shape the architecture of the axon-muscle connectome by shaping TF codes in IMNs.

**Post-transcriptional regulation of TFs in IMNs by Imp shape the axon-muscle connectome**

To establish the correct TF codes in IMNs, Imp and Syp could function in either the NB and/or IMNs. To discriminate between these two possibilities, we induced Imp expression in IMNs without affecting its temporal expression in the NB or GMC, using the VGlut-Gal4 driver (Figure 7).

In this genetic background, muscle innervation phenotypes were similar to those generated by Imp overexpression in all Lin A/15 cells, including the NB (Figure 7B versus Figure 5H). In particular, the distal region of the tibia was not or was less innervated than normal. Moreover, similar epistatic relationships were observed between Imp and the TF codes to those observed by Imp overexpression in all Lin A/15 cells (Figures 7D-7J, VGlut > Imp versus Figures 6G-6I, 6M-6O, tub > Imp). In VGlut > Imp Lin A/15, the number of IMNs expressing Chinmo, Mamo, Jim, and Br was higher, although the number of IMNs expressing Nvy was not statistically lower in VGlut > Imp MARCM clones (Figure 7J), while it was in tub > Imp MARCM clones (Figure 6N). Interestingly, the overexpression of Jim only in IMNs, including last-born MNs expressing a low level of Imp in WT, induces a similar phenotype as the one in VGlut > Imp Lin A/15 and characterized by decreased innervation of the distal tibia (Figure 7C; Table S2).

These results show that the level of Imp shapes the axon muscle connectome through the regulation of TFs in IMNs. Together with above results, we propose a model where Imp and Syp shape the axon muscle connectome through the post-transcriptional regulation of TFs translation in IMNs (Figure 7K; see discussion).

**DISCUSSION**

Lin A/15 NB divides every 70 min (Baek and Mann, 2009), and at each division, it produces a GMC or intermediate mother cell (IMC) that generates an IMN with a unique developmental fate. Based on our results, we postulate that the Lin A/15 NB divides too fast to implement neuronal identity using only transcriptional mechanisms to generate TF codes. The codes need time to be established in IMNs and thus depend on two layers of
regulation: the transcriptional expression of TF RNA and their post-transcriptional refinement. The transcriptional upstream regulators of the mRNAs coding for the TFs remain to be identified. In our system, Imp and Syp, which are known to control the temporal identity of NBs, are maintained as opposing gradients in iMNs according to their birth order. These opposite expression patterns are progressively decypted into sharp mTF codes in iMNs, allowing for the generation of a large amount of neuronal diversity. Our study is a proof of concept that the temporal identity of the stem cell defined by two opposite gradients of RBPs can be deciphered by iMNs into TF codes. If we take into consideration the level of expression of the TF analyzed, our results reveal that both RBPs shape nine different TF codes. Below, we propose a model on how iMNs decode this temporal information.

**From stem cells to immature neurons**

This developmental logic regulating MNs targeting contrasts with mushroom body neurons. Lin A/15 NB produces few neurons with a high amount of morphological diversity, whereas the MB NBs generate a higher number of neurons with less morphological diversity. In Lin A/15, Imp and Syp are two key components of the post-transcriptional machinery shaping the TF codes, demonstrating that the post-transcriptional regulation of TFs is also compatible with lineages generating a large amount of neuronal diversity within a short period of time. Our results revealed that the differential sensitivities of TFs RNA to the relative concentrations of Imp and Syp explain how a gradient of two RBPs is transformed into unique TF codes. The molecular basis of these differentially sensitivities could be found in the physical interaction between RNAs and RBPs since jim RNA is expressed from low to high in iMN 1–iMN 15, and pros are known to immunoprecipitate with Imp and/or Syp (McDermott et al., 2014; Yang et al., 2017a). Investigating further into the mechanisms of Imp- and Syp-TF interaction by, for example, switching UTRs between TF RNAs, which are the regions interacting with RBPs, could help determine whether the range of proteins’ expression is changed accordingly. Moreover, the expression gradient of some TF RNAs could also explain how two gradients of RBPs are decoded into sharp TF code. For example, jim RNA is expressed from low to high in IMN 1–IMN 29 and Jim protein might not be present in iMNs 1–15 because the threshold of Jim RNA needed to have detectable protein is not reached. MNs 24–29 do not express Jim protein, even if Jim RNA is highly expressed because of the inhibition of Jim translation by Syp. As a consequence, only the iMNs 16–23, expressing high level of Jim RNA and Imp protein, expressed Jim protein.

**Imp and Syp are multi-tasked proteins determining several facets of neuronal identity**

Here, we have shown that Imp and Syp are important to determine one aspect of the MN identity, its morphology. However, both proteins might also determine other aspects of MN fate. In our parallel study, we have demonstrated that both RBPs act in iMNs to determine the survival of Lin A/15 MNs (Guan et al., 2021). Syp prevents the survival of iMNs, whereas Imp promotes it. Interestingly, both proteins might also control the biological identity of MNs. We were able to generate Lin A/15 Imp−+/− MARCM clones with a pan-cellular tub-Gal4 driver, but not with VGlut-Gal4, an enhancer trap transgene of the gene coding for the vesicular glutamate transporter that is expressed by all Drosophila MNs (Mahr and Aberle, 2006). The TFs controlling these molecular features are called terminal selectors (Allan et al., 2005; Eade et al., 2012; Hobert, 2011, 2016). Good candidates for terminal selectors could be genes expressed in all iMNs, such as Lim 3 and Nkx6 found in our screen (Figure S1).

**Muscle development is linked to innervation by MNs**

The development of the locomotor system needs communication between different tissues, such as muscles and MNs. Muscle development was affected in several genetic backgrounds in addition to muscle innervation defects. The effect on muscle development is due to a non-autonomous effect of MNs on muscles because the genetic backgrounds only affected MNs. For example, in nvy−/−, Jim-overexpressing, and Imp-overexpressing Lin A/15 clones, the tarm1/2 is atrophic or absent when it was not innervated by LinA/15 axons (Figures 5 and 7). This dependence of muscle development on innervation by MNs might result from a lack of electrical activity or an absence of the neurotrophic factors secreted by MN axons. Another possibility could be that myogenesis is directly induced by innervation. Our work would stimulate further investigations on the dialogue between MN and muscles during development.

**Limitations of the study**

Even though our work emphasizes the role of Imp and Syp in iMNs, we could not exclude the possibility that both RBPs might also be important for the temporal identity of the NB, presumably through
governing the expression of temporal TFs. Chinmo, a TF expressed in a limited time window in Lin A/15 NB (Figure 3), serves as a putative downstream effector of Imp and Syp in specifying the temporal identity of Lin A/15 NB. This could also explain the decrease in tirm innervation in chinmo−/− Lin A/15 clones that we interpret as a function of Chinmo in the NB (Figure S6).

Transcriptional upstream regulators inducing the expression of the TF RNA remain to be identified. We propose that spatial selectors expressed in a gradient manner could be the upstream regulators of TF mRNA expression, which are expressed ubiquitously as a gradient, such as jim or mamo. Candidates for spatial selectors could be the TFs found in our expression screen, such as RunxB, Islet, or Jumu, that have a gradient expression pattern in iMNs.

Finally, we cannot exclude that TFs could function in parallel with RBPs to regulate mTF transcription, as has been shown in larval MNs (Seroka et al., 2020). Moreover, our results also show that not all mTFs, such as Oli, are sensitive to Imp and Syp concentration, suggesting that other post-transcriptional machinery remains to be identified.

**STAR METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHOD DETAILS**
  - Genetic Crosses for each figure
  - Lineage tracing and MARCM Clonal Analysis
  - Leg imaging
  - Immunostaining of larval CNS
  - Immunostaining of adult VNC
  - Primary and secondary antibodies
  - Image acquisition of immunostained VNC/CNS and adult legs
  - 5-′Ethynyl-2′deoxyuridine (EdU) labelling
  - Probe design and preparation for smiFISH
  - smiFISH, sample preparation, and hybridization
  - Image acquisition of larval VNC after smiFISH
  - smiFISH analysis
  - Schematics
  - Cloning
  - MIMIC swap to generate Mi[Trojan-VGlut-LexA::-GAD.2]
  - Positive Cell Cluster Detection (PCCD) method (Figure 2, see also PCCD note book in key resources table)
  - Test of the PCCD method by co-staining
  - Lin A/15 MARCM clones vs Lin A/15 memory system
  - Link between birth order and muscle targets: Corrections of (Baek and Mann, 2009)
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

**SUPPLEMENTAL INFORMATION**

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.110992.

**ACKNOWLEDGMENTS**

We thank Alain Vincent, Filipe Pinto-Teixeira, and Steve Garvis for comments on the manuscript. This work was funded by the Atp—which is also supported by: NIH (no. 219999), and FRM (no. AJE2107537445) to J.E.; the NIH: R01 NS070644 to R.S.M.; and the GWIS-Nell Mondy Trust and Elizabeth Weisburger Fellowships to L.V. We acknowledge the contribution of SFR Biosciences (Arthro-tool facility, UAR344/CNRS, USB/INSERM, ENS de Lyon, UCBL) and of the IGFL facility platform.

**AUTHOR CONTRIBUTIONS**


**DECLARATION OF INTERESTS**

The authors declare no competing interests.

Received: February 23, 2022

Revised: April 15, 2022

Accepted: June 1, 2022

Published: June 28, 2022

**REFERENCES**


### KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibodies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rabbit anti-Castor</td>
<td>Gift from Claude Desplan</td>
<td>N/A</td>
</tr>
<tr>
<td>mouse anti Br-core</td>
<td>DSHB</td>
<td>Cat#2E9.D7; RRID:AB_528104</td>
</tr>
<tr>
<td>guinea-pig anti-Jim</td>
<td>Gift from Claude Desplan</td>
<td>N/A</td>
</tr>
<tr>
<td>mouse anti-Propero</td>
<td>DSHB</td>
<td>Cat#MR1A; RRID:AB_528440</td>
</tr>
<tr>
<td>guinea-pig anti-Runt</td>
<td>Gift from Kuniaki SAITO</td>
<td>N/A</td>
</tr>
<tr>
<td>rabbit anti-Nvy</td>
<td>Gift from Richard S. Mann</td>
<td>N/A</td>
</tr>
<tr>
<td>rabbit anti-RunXA</td>
<td>Gift from Claude Desplan</td>
<td>N/A</td>
</tr>
<tr>
<td>guinea-pig anti-FoxP</td>
<td>Gift from Chris Q Doe</td>
<td>N/A</td>
</tr>
<tr>
<td>guinea-pig anti-Lov</td>
<td>Gift from Claude Desplan</td>
<td>N/A</td>
</tr>
<tr>
<td>rabbit anti-Zfh1</td>
<td>Gift from Jim Skeath</td>
<td>N/A</td>
</tr>
<tr>
<td>guinea-pig anti-Kr</td>
<td>Gift from Claude Desplan</td>
<td>N/A</td>
</tr>
<tr>
<td>rat anti-Zfh2</td>
<td>Gift from Chris Q Doe</td>
<td>N/A</td>
</tr>
<tr>
<td>rabbit anti-Oli</td>
<td>Gift from Claude Desplan</td>
<td>N/A</td>
</tr>
<tr>
<td>rabbit anti-Mamo</td>
<td>Gift from Oren Schuldiner</td>
<td>RRID:AB_2666566</td>
</tr>
<tr>
<td>guinea-pig anti-Chinmo</td>
<td>Gift from Claude Desplan</td>
<td>N/A</td>
</tr>
<tr>
<td>rabbit anti-Eyeless</td>
<td>Gift from Claude Desplan</td>
<td>N/A</td>
</tr>
<tr>
<td>mouse anti-Elav</td>
<td>DSHB</td>
<td>Cat#9F8A9; RRID:AB_2314364</td>
</tr>
<tr>
<td>rat anti-Elav</td>
<td>DSHB</td>
<td>Cat#7E8A10; RRID:AB_528218</td>
</tr>
<tr>
<td>guinea-pig anti-Dpn</td>
<td>Gift from Jim Skeath (Skeath et al., 2017)</td>
<td>RRID:AB_2314299</td>
</tr>
<tr>
<td>rat anti-Imp</td>
<td>Gift from Claude Desplan</td>
<td>N/A</td>
</tr>
<tr>
<td>rabbit anti-Syp</td>
<td>Gift from Claude Desplan</td>
<td>N/A</td>
</tr>
<tr>
<td>goat anti-mouse Alexa 647</td>
<td>Invitrogen</td>
<td>Cat#A32728; RRID:AB_2633277</td>
</tr>
<tr>
<td>donkey anti-rat Alexa 647</td>
<td>Jackson</td>
<td>Cat#712-605-153; RRID:AB_2340694</td>
</tr>
<tr>
<td>goat anti-mouse Alexa 555</td>
<td>Invitrogen</td>
<td>Cat#A32727; RRID:AB_2633276</td>
</tr>
<tr>
<td>goat anti-rabbit Alexa 555</td>
<td>Invitrogen</td>
<td>Cat#A32732; RRID:AB_2633281</td>
</tr>
<tr>
<td>goat anti-Rat Alexa 555</td>
<td>Abcam</td>
<td>Cat#fab150166</td>
</tr>
<tr>
<td>donkey anti-guinea-pig DyLight405</td>
<td>Jackson</td>
<td>Cat#706-475-148; RRID:AB_2340470</td>
</tr>
<tr>
<td><strong>Chemicals, peptides, and recombinant proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formaldehyde</td>
<td>Thermo Scientific</td>
<td>Cat#28908</td>
</tr>
<tr>
<td>PBS</td>
<td>Dutscher</td>
<td>Cat#X0515-500</td>
</tr>
<tr>
<td>Triton</td>
<td>Sigma</td>
<td>Cat#T8787-100mL</td>
</tr>
<tr>
<td>BSA</td>
<td>Sigma</td>
<td>Cat#A7906-500 g</td>
</tr>
<tr>
<td>Vectashield mounting medium</td>
<td>Vector Laboratories</td>
<td>Cat#H1000</td>
</tr>
<tr>
<td><strong>Critical commercial assays</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clicl-iT EdU imaging kit</td>
<td>Invitrogen</td>
<td>Cat#C10340</td>
</tr>
<tr>
<td><strong>Experimental models: Organisms/strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UAD-KD (attP2[68A4])</td>
<td>(Awasaki et al., 2014)</td>
<td>N/A</td>
</tr>
<tr>
<td>dpn::KDRT-stop-KDRT::CRE (su[Hw]attP8 [8E10])</td>
<td>(Awasaki et al., 2014)</td>
<td>N/A</td>
</tr>
<tr>
<td>act&gt;loxP- stop-loxP&gt;LexA::P65 (attP40 [25C7])</td>
<td>(Lacin and Truman, 2016)</td>
<td>Tzumin Lee</td>
</tr>
<tr>
<td>lexAop-myr::GFP (su[Hw]attP5[50F1])</td>
<td>(Awasaki et al., 2014)</td>
<td>N/A</td>
</tr>
<tr>
<td>R10c12::GAL4 (3rd chromosome, attp2)</td>
<td>(Lacin and Truman, 2016)</td>
<td>N/A</td>
</tr>
</tbody>
</table>

(Continued on next page)
**RESOURCE AVAILABILITY**

**Lead contact**
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jonathan Enriquez (jonathan.enriquez@ens-lyon.fr).

**Materials availability**
All fly stocks made in this study, as well as the nucleotide sequences of the plasmids used, are available from the lead contact without restriction.

**Data and code availability**
Microscopy data reported in this paper will be shared by the lead contact upon request. PCCD note book and smFISH analysis source code have been deposited at Zenodo and is publicly available as of the date of publication. DOIs are listed in the key resources table. Any other additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

**EXPERIMENTAL MODEL AND SUBJECT DETAILS**

Experimental model for this study was the vinegar fly *Drosophila melanogaster*. A full list of strains used in the paper is included in the Key resources table and Genetic Crosses for each Figure. Unless otherwise described, flies were maintained on a standard cornmeal medium and kept at 25°C in a 12:12 h light:dark cycle. Males and females were chosen at random.
# METHOD DETAILS

## Genetic Crosses for each figure

<table>
<thead>
<tr>
<th>Figures</th>
<th>Summary</th>
<th>Genetic crosses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figures 1B and 1C</td>
<td>WT (LinA/15 tracing labeling)</td>
<td>( R10c12\text{-GAL4} ) TO ( dpn&gt;\text{KDRT-stop-KDRT&gt;CRE;} ) act&gt;( \text{loxP-stop-loxP&gt;LexA::P65, lexAop-myr::GFP; UAS-KD} )</td>
</tr>
<tr>
<td>Figures 1G–1K</td>
<td></td>
<td>( R10c12\text{-GAL4} ) TO ( dpn&gt;\text{KDRT-stop-KDRT&gt;CRE;} ) act&gt;( \text{loxP-stop-loxP&gt;LexA::P65, lexAop-myr::GFP; UAS-KD} )</td>
</tr>
<tr>
<td>Figure 1D</td>
<td>WT (LinA/15 MARCM Clone)</td>
<td>( y,w, \text{hs-Flp}^{1,22}); VGlut-Ga4, UAS-mCD8::GFP, Mhc-RFP, FRT42D/CyO; TM6B/MKRS TO ( y,w, \text{hs-Flp}^{1,22}); VGlut-Ga4, UAS-mCD8::GFP, Mhc-RFP, FRT42D, tubP-Gal80/CyO; MKRS/TM6B</td>
</tr>
<tr>
<td>Figures 2B–2G</td>
<td>WT (LinA/15 tracing labeling)</td>
<td>( R10c12\text{-GAL4} ) TO ( dpn&gt;\text{KDRT-stop-KDRT&gt;CRE;} ) act&gt;( \text{loxP-stop-loxP&gt;LexA::P65, lexAop-myr::GFP; UAS-KD} )</td>
</tr>
<tr>
<td>Figure 2N</td>
<td></td>
<td>( y,w, \text{hs-Flp}^{1,22}); VGlut-Ga4, UAS-mCD8::GFP, Mhc-RFP, FRT42D/CyO; TM6B/MKRS TO ( y,w, \text{hs-Flp}^{1,22}); VGlut-Ga4, UAS-mCD8::GFP, Mhc-RFP, FRT42D, tubP-Gal80/CyO; MKRS/TM6B</td>
</tr>
<tr>
<td>Figure 3</td>
<td></td>
<td>( y,w, \text{hs-Flp}^{1,22}); VGlut-Ga4, UAS-mCD8::GFP, Mhc-RFP, FRT42D/CyO; TM6B/MKRS TO ( y,w, \text{hs-Flp}^{1,22}); VGlut-Ga4, UAS-mCD8::GFP, Mhc-RFP, FRT42D, tubP-Gal80/CyO; MKRS/TM6B</td>
</tr>
<tr>
<td>Figures 4A–4C</td>
<td>WT and Syp-/-</td>
<td>( y,w, \text{hs-Flp}^{1,22}); VGlut-Ga4, UAS-mCD8::GFP, Mhc-RFP, FRT42B, tubP-Gal80/TM6B TO ( y,w, \text{hs-Flp}^{1,22}); VGlut-Ga4, UAS-mCD8::GFP, Mhc-RFP/CyO; FRT82B, Sypf03755/TM6B for Figure 5B and 6A–6C</td>
</tr>
<tr>
<td>Figures 4J and 4K</td>
<td></td>
<td>( y,w, \text{hs-Flp}^{1,22}); VGlut-Ga4, UAS-mCD8::GFP, Mhc-RFP, FRT42B, tubP-Gal80/TM6B TO ( y,w, \text{hs-Flp}^{1,22}); VGlut-Ga4, UAS-mCD8::GFP, Mhc-RFP/CyO; FRT82B, Sypf03755/TM6B for Figures 5G and 6D–6F</td>
</tr>
<tr>
<td>Figure 5B</td>
<td>WT and tub&gt;imp</td>
<td>( y,w, \text{hs-Flp}^{1,22}); VGlut-LexA::GAD,13LexAop-mCD8GFP, FRT42D/CyO; tubP-GalGal4/TM6B TO ( y,w, \text{hs-Flp}^{1,22}); FRT42D,tubP-Gal80/CyO; MKRS/TM6B for Figure 5C</td>
</tr>
<tr>
<td>Figure 5G</td>
<td></td>
<td>( y,w, \text{hs-Flp}^{1,22}); VGlut-LexA::GAD,13LexAop-mCD8GFP, FRT42D/CyO; tubP-GalGal4/TM6B TO ( y,w, \text{hs-Flp}^{1,22}); FRT42D,tubP-Gal80/CyO; MKRS/TM6B for Figure 5C</td>
</tr>
<tr>
<td>Figures 6A–6C</td>
<td>WT and Syp-/-</td>
<td>( y,w, \text{hs-Flp}^{1,22}); VGlut-LexA::GAD,13LexAop-mCD8GFP, FRT42D/CyO; tubP-GalGal4/TM6B TO ( y,w, \text{hs-Flp}^{1,22}); FRT42D,tubP-Gal80/CyO; MKRS/TM6B for Figure 5C</td>
</tr>
<tr>
<td>Figures 6D–6F</td>
<td></td>
<td>( y,w, \text{hs-Flp}^{1,22}); VGlut-LexA::GAD,13LexAop-mCD8GFP, FRT42D/CyO; tubP-GalGal4/TM6B TO ( y,w, \text{hs-Flp}^{1,22}); FRT42D,tubP-Gal80/CyO; MKRS/TM6B for Figure 5C</td>
</tr>
<tr>
<td>Figure 5C</td>
<td>WT and tub&gt;imp</td>
<td>( y,w, \text{hs-Flp}^{1,22}); VGlut-Gal4, UAS-P35/CyO; tubP-GalGal4/QUAS-mCD8::GFP/TM2 TO ( y,w, \text{hs-Flp}^{1,22}); FRT19A/AFM7; VGlut-Gal4, UAS-mCD8::GFP, Mhc-RFP, FRT42D/CyO; tubP-GalGal4/QUAS-mCD8::GFP/TM2</td>
</tr>
<tr>
<td>Figure 5A</td>
<td>WT and tub&gt;imp</td>
<td>( y,w, \text{hs-Flp}^{1,22}); VGlut-Gal4, UAS-P35/CyO; tubP-GalGal4/QUAS-mCD8::GFP/TM2 TO ( y,w, \text{hs-Flp}^{1,22}); FRT19A/AFM7; VGlut-Gal4, UAS-mCD8::GFP, Mhc-RFP, FRT42D/CyO; tubP-GalGal4/QUAS-mCD8::GFP/TM2</td>
</tr>
<tr>
<td>Figure 5F</td>
<td>WT and tub&gt;imp</td>
<td>( y,w, \text{hs-Flp}^{1,22}); VGlut-Gal4, UAS-P35/CyO; tubP-GalGal4/QUAS-mCD8::GFP/TM2 TO ( y,w, \text{hs-Flp}^{1,22}); FRT19A/AFM7; VGlut-Gal4, UAS-mCD8::GFP, Mhc-RFP, FRT42D/CyO; tubP-GalGal4/QUAS-mCD8::GFP/TM2</td>
</tr>
<tr>
<td>Figure 5D</td>
<td>WT and nvy-/-</td>
<td>( y,w, \text{hs-Flp}^{1,22}); VGlut-Gal4, UAS-mCD8::GFP, Mhc-RFP, FRT42D,tubP-Gal80/CyO; UAS-mCD8::GFP/MKRS TO ( y,w, \text{hs-Flp}^{1,22}); UAS-mCD8::GFP, VGlut-Gal4, FRT42D/CyO; UAS-mCD8::GFP/MKRS for Figure 5D</td>
</tr>
<tr>
<td>Figure 5I</td>
<td>WT and nvy-/-</td>
<td>( y,w, \text{hs-Flp}^{1,22}); VGlut-Gal4, UAS-mCD8::GFP, Mhc-RFP, FRT42D,tubP-Gal80/CyO; UAS-mCD8::GFP/MKRS TO ( y,w, \text{hs-Flp}^{1,22}); UAS-mCD8::GFP, VGlut-Gal4, FRT42D,CyO; UAS-mCD8::GFP/MKRS for Figure 5I</td>
</tr>
</tbody>
</table>

(Continued on next page)
**Lineage tracing and MARCM Clonal Analysis**

LinA/15 tracing labeling is achieved by immortalizing Gal4 expression in Lin A/15 neuroblasts and its descendants (Awasaki et al., 2014; Lacin and Truman, 2016). Fly strains used to specifically label LinA/15 are listed in Genetic crosses for each figure.

All fly strains used to introduce MARCM (Lee and Luo, 1999) clones labeling LinA/15 are listed in Genetic Crosses for each Figure. To introduce MARCM clones, embryos were collected for 12hs in vials and incubated for 24hs at 25°C. First-instar larvae (0-24h ALH) were heat shocked at 37°C for 20 min to induce mosaic clones in L3 larvae and at 35°C for 15min to induce mosaic clones in adults. All flies were then raised at 25°C degrees at the exception of Imp-/- tub>GFP+P35 flies being raised at 29°C after heat shock to boost the efficiency of P35 (Figure 5F). P35 is added in the genetic background only when less MNs are produced in a given genetic perturbation, it is a way to exclude the possibility that the axonal targeting defect is the result of less MNs produced. In terms of the GAL4 driver (tub vs VGlut), the reason for choosing tub instead of VGlut are as follows: Figures 5 F and S6: Imp-/- tub>GFP: VGlut>GFP failed to label LinA MNs: our hypothesis is that Imp is necessary for VGlut expression; Figure 5J: jim KD tub>GFP: a better knocking down efficiency is observed for RNAi-jim using tub driver compared to VGlut driver, probably due to the delayed onset of VGlut in MNs compared to tub. Figure S6: chinmo-/-, tub>GFP: chinmo-/- MARCM clones was impossible to generate with VGlut because of the proximity of genomic localization between chinmo (Chr2L, 22A) and VGlut (Chr2L, 22E) that make recombination events very rare, and thus the chromosome carrying both elements nearly impossible to recover.

**Leg imaging**

Legs were dissected, fixed, and imaged as described in (Guan et al., 2018).

**Immunostaining of larval CNS**

Inverted L3 larvae were fixed in 4% paraformaldehyde in PBS for 20 min at room temperature and blocked in the blocking buffer for one hour. L3 larval or pupal CNS were carefully dissected in PBS and then incubated with primary antibodies overnight (≥12h) and secondary antibodies in dark for one day (≥12h) at 4°C. Fresh PBST-BSA (PBS with 0.3% Triton X-100, 1% BSA) was used for the blocking, incubation, and washing steps: five times for 20 min at room temperature after fixation and after primary/secondary antibodies. Larval/pupal CNS were mounted onto glass slides using Vectashield anti-fade mounting medium (Vector Labs). Slides were either imaged immediately or stored at 4°C.

**Immunostaining of adult VNC**

After removing the abdominal and head segments, the thorax of the flies was opened and fixed in 4% paraformaldehyde in PBS for 25 min at room temperature and blocked in the blocking buffer for 1 h. After dissection, adult VNC was incubated with primary antibodies for 1 day and secondary antibodies in dark for 1 day at 4°C. Fresh PBST-BSA (PBS with 0.3% Triton X-100, 1% BSA) was used for the blocking, incubation, and washing steps: five times for 20 min at room temperature after fixation and after primary/secondary antibodies.
primary/secondary antibodies. VNC was mounted onto glass slides using Vectashield anti-fade mounting medium (Vector Labs). Slides were either imaged immediately or stored at 4°C.

**Primary and secondary antibodies**

We screened around 220 antibodies against transcription factors (at least 10 CNSs/TF) from various sources including: a collection of antibodies generated by the modENCODE provided to us by C. Desplan (Li et al., 2013a); Developmental Studies Hybridoma Bank and various gifts from the Drosophila community. Antibodies used for the 16 TFs (expressing in different clusters of Lin A/15 MNs) are listed in key resources table. Other TfNs are screened following: Grh, Insv, Vvl, E93, Lin3, Nlx6, RunXB, Antp, Islet, Jumu, Gt, Psg, Seq, Ind; Adf1, E2F1, E2F2, Zelda, Sd, Snr1, Inv, ham, Hth, Tll, Grm, Ato, CG8108, CG12391, Dalao, B-R21, Br-Z3, vestigial, twist, gooseberry, gooseberry neuro, POU domain protein 2, sal, rotund, mad, labial, deformed, midline, virlle, even skipped, achaete-c, knot, ladybird early, eagle, sloppy paired 1, Dichaete, apontic, LIM homeobox 1, extra-extra, Escargot, dachshund, period, Grain II, SVP, Hb9, Toy, Eg, En, Gm, Hkb, Msh, Nub, Eya, common split domain, klumpfluss, pointed, cryptochrome, Sox15, knirps, sloppy paired 2, Huckebein, Lim1, pdm1, acj6, abrupt, anterior open, pterid, bicone, brain-specific homeobox, caulif, cut, defplected proventriculus, empty spiracles, engrailed, extradendric, eyegone, eyes absent, fushi tarazu, pebbled, intermediate neuroblasts defective, odd paired, odd skipped, Optix, paired, Ptx1, Retinal Homeobox, senseless, Sex combs reduced, so, snail, spineless, teashirt, Vx1, Vx2, Zelda, BarH1, CG3065, CG3281, CG7963, CG8478, cropped, CTCF, D1, Dorsochriss2, Drop, Dsp1, elB, Fed3, GATAe, Hnf4, Hr39, Jra, J(1) sc, pho, pho, Smox, Sox21a, Sox21b, SoxN, srp, sug, Taf12, tin, zeste, abdominal A, cubitus interruptus, crocodile, disco, deformed wings, fusel, Meics, ocellis, p53, pangolin, pita, stat92E, stc, Suppressor of Hairless, Trithorax-like, wdn, won, tramtrack, Ubx, Mef2, Bro, CG1832, CG3075, E(bx), E(z), knirps-like, Polycomblike, Rel, dwg, ftz-f1, hunchback.

**Image acquisition of immunostained VNC/CNS and adult legs**

Multiple 0.5-μm-thick and 1-μm-thick sections in the z axis for larval/pupal CNS or adult VNC and adult legs respectively were imaged with a Leica TCS SP8 or a Zeiss LSM 780 confocal microscope. Objective used: 40×1.3 oil objective for larval/pupal CNS or adult VNC; 20× glycerol objective for adult leg samples. Binary images for z stack images were generated using NIH ImageJ.

**5-Ethynyl-2′-deoxyuridine (EdU) labelling**

To mark MNs born at different time points (Figures 1I–1K), L2 or L3 larvae (48, 72 and 96 h after egg laying [AEL]) were transferred from standard fly food to fly food containing 250 mM EdU. Larvae were dissected at 120 h AEL and dissected CNS were fixed in 4% paraformaldehyde in PBS for 20 min at room temperature, followed by a quick wash with PBST (PBS with 0.3% Tween-20). EdU labelling was then detected using Click-iT EdU imaging kit (Invitrogen) according to manufacturer’s instructions. An immunostaining was then performed as described in the Immunostaining section.

**Probe design and preparation for smiFISH**

Primary probes against mRNA sequences of genes of interest (common sequence of all isoforms of genes of interest; up to 48 probes per gene) were designed using the Biosearch Technologies stellaris RNA FISH probe designer tool (free with registration, https://biosearchtech.com). The probe sequences for each gene used in this study have been deposited at Zenodo (Zenodo: https://doi.org/10.5281/zenodo.6592760) and is publicly available as of the date of publication. The following sequence was added to the 5′ end of each 20 nucleotide (nt) probe: CCTCCTAAGTTTCGAGCTGGACTCAGTG, which is the reverse complement of the X flap sequence. This sequence was added to the 5′ end-labeled with Quasar 570, is synthesized by Biosearch Technologies. Fluorophore-labeled probe sets are prepared as detailed in (Tsanov et al., 2016), by hybridizing the fluorophore-labeled X Flap sequence and 5′-extended primary probe sets.

**smiFISH, sample preparation, and hybridization**

Dissected larval CNS from third instar larvae were fixed in 4% paraformaldehyde (in PBS with 0.3% Triton X-100) for 20 minutes at room temperature. Samples were washed for 3 times of 15 min with PBST (PBS with 0.3% Tween-20) before a pre-hybridization wash in smiFISH wash buffer (10% deionised formamide (stored at -80°C C) and 2X SSC in DEPC water) at 37°C for 30 min. Samples were then incubated with hybridized fluorophore-labeled probe sets diluted in smiFISH hybridization buffer (10% deionized formamide, 2X SSC and 5% dextran sulphate in DEPC water) at 37°C for 12 to 14 h. The working concentration of each probe sets are: jim (400 nM), chinmo (800 nM), br (1 μM), nvy (2 μM), mamo (2 μM) and oil (400 nM). Samples were washed for 40 min at 37°C followed by three times of 15 min in smiFISH wash buffer at room temperature, and 10 min of washing in PBST (PBS with 0.3% Tween-20) before sample mounting. Protocol adapted from (Yang et al., 2017b).
**Image acquisition of larval VNC after smFISH**

Images were acquired on a Leica TCS SP8 microscope using a 40x/1.3 oil objective. Image stacks were taken with the following settings: format 4400x4400 (223.56 μm * 223.56 μm (XY) and 44.95 μm (Z)), speed 700 Hz, unidirectional, sequential line scanning, line averaging 8, pinhole 1 airy unit. Image stacks were acquired with a 350 nm z interval. Parameters for the three channels used in this study are: DAPI excitation 405 nm, laser 2%, detection 412–496nm; GFP excitation 488nm, laser 4.5%, detection 494–530nm; Quasar 570 excitation 548nm, laser power 5%, detection 555–605nm.

**smFISH analysis**

See also smFISH analysis source code in the Key resources table. Each Lin A/15 cell was segmented in 3D in ImageJ/Fiji (Rueden et al., 2017; Schindelin et al., 2012; Schneider et al., 2012), using the LimeSeg plugin (Machado et al., 2019), on the GFP channel, with the following parameters: D_0 = 6–6, F pressure = 0.01, Z_scale = 6.8, Range in d0 units = 4–6, Number of integration step = 4–6, real XY pixel size = 50. For subsequent analysis, each segmented cell was exported into a separate ply file which was then imported in MATLAB as a point cloud (The Math Works, Inc.). The original stacks were imported in MATLAB using the Bio-Formats toolbox (https://www.openmicroscopy.org/bio-formats/downloads/). These stacks were then cropped around each cell using the point clouds generated by individual cell segmentation with LimeSeg.

The mRNA spots were detected in 3D, in the mRNA channel of these cropped stacks, using the method described by (Raj et al., 2008). In short, the spots were identified computationally by running a Matlab image processing script that runs the raw data through a filter (Laplacian of a Gaussian) designed to enhance spots of the correct size and shape while removing the slowly varying background. The filtered stacks are then thresholded to disregard remaining background noise. In order to choose an optimal threshold, all possible thresholds are computed. The thresholds were always chosen manually and close to the plateau. A ‘check File stack’ for each cell was generated in order to visualize the accuracy of the spot detection for a given threshold. In most of our samples, common thresholds were chosen for all the cells of a given Lin A/15. However, specific thresholds were occasionally chosen for some cells. The parameters that gave the best visual detection of mRNA spots in our datasets—check files were generated were as follow: width of the Gaussian filter = 5; variance of the Gaussian filter = 0.25; threshold = 0.11 for jim, 0.08 for chinmo, 0.15 for mamo, 0.09 for br, 0.29 for nvy and 0.15 for oli. The detected spots in each cell had a normal distribution of diameters with a mean of 0.25 +/- 0.05 micrometers and their maximum intensities displayed a unimodal distribution, arguing that the detected spots are mostly individual molecules. A custom Matlab routine transformed the point clouds corresponding to segmented cell volumes obtained from the LimeSeg plugin into an alpha shape (critical radius = 100) and then returned the number of mRNA contained in this cell volume using the Matlab function inShape.

**Schematics**

All schematics were done with Microsoft PowerPoint.

**Cloning**

Pattb-20XUAS-hsp70-Jim-sv40. The full coding sequence of jim (2469bp) where the Xhol site in exon 1 was mutated to CTCCAG was ordered from Genewiz and cloned (Xhol/Xbal fragment) in the plasmid pJFRC165-20XUAS-IVS-R::PEST (https://www.addgene.org/32142/). The resulting Pattb-20XUAS-hsp70-jim-sv40 construct was inserted in position 86f on chromosome III by injection into embryos carrying attP-86f landing site.

**MIMIC swap to generate Mi[Trojan-VGlut-LexA::GAD.2]**

The plasmid pBS-KS-attB2-SA2-T2A-LexA::GADfluw-Hsp70 was ordered from Addgene (https://www.addgene.org/78307/) and inserted in the MIMIC Mi[MIC]VGlut[MIO4979] (BDSC_38078) line inserted into the relevant intron in Phase 2 of the VGlut gene.

**Positive Cell Cluster Detection (PCCD) method** (Figure 2, see also PCCD note book in key resources table)

The Positive Cell Cluster Detection (PCCD) method aims to link the expression of a given TF to the birth-order of an immature MN (iMN) by using the correlation between the birth-order of iMNs and their spatial organization. In our Lin A/15 model, the EdU experiments (Figure 1) reveal a good correlation between the birth order of iMNs and their spatial distance from the NB in 3rd instar larvae: young born iMNs are farther away from the NB compared to older iMNs. The final goal of this method is to predict the TF code expression pattern in each iMNs in a third instar larva.

The method followed a series of steps

1. Step 1: From the imaging, assign spatial x, y, z coordinates and the expression (on/off) of a given TF to each Lin A/15 cell (N > 15, Number of Lin A/15 immunostained for a given TF).

2. Step 2: Calculate the Euclidean distance between the NB and the x, y, z, coordinates of each iMN (relative distance).

3. Step 3: Order iMNs in each Lin A/15 according to their distance to NB. This presents each Lin A/15 as an ordered sequence of iMNs (this defines the x axis position where cell #1 is defined as the furthest from the NB, i.e. the oldest iMNs on average). Then calculate the frequency of expression of all TFs as a function of their rank in each ordered Lin A sequence.

4. Step 4: Apply a filter (Savitzky-Golay) to smooth each distribution.
Step 5: Define the position in the sequence of the positive cell cluster(s) by using a peak detection method. Determine its length (average number of cells expressing a given TF in all Lin A/15 samples analyzed). Then find the position of the positive cell cluster with this average length compatible with the smoothed TF distribution. The position and its length are represented by a horizontal line.

Step 6: Assemble all positive cell clusters for each TF on the same graph to reveal combinatorial TF code for each iMN. Convert the x’ axis to a birth order axis (1–29) since the distance between iMN and the NB is tightly linked to their birth order. Define the coverage index at the border of all cell clusters.

More details about the method include

Savitzky-Golay filter (Step 4): The Savitzky-Golay algorithm (polynomial filter) was set with a window of size 11 and a polynomial order of 3 (see scipy.signal.savgol_filter function from the scipy python library).

Peak detection method (Step 5): Peaks were detected as local maxima in the normalized TF distributions. Local maxima were determined according to local conditions. They had a minimal height (h_min = 0.02), a minimal distance from other peaks (d_min = 15), and a minimal prominence (p_min = 0.1). The prominence of a peak measures how much a peak is emerging clearly locally in the signal. It is defined as the vertical distance between the peak and the altitude of the largest region it dominates. These values were found to yield best peak interpretations over the whole set of TFs. We used the function scipy.signal.find_peaks of the scipy library.

Positive cell clusters (Step 5): For each TF, the average number “p” of positive cells was computed in each Lin A/15 iMN observed Cluster. The procedure varied according to whether only one peak was detected or more than one (i.e. 2 in our data).

Case of a single detected peak (e.g. Jim): The altitude of the peak at which the signal width below the peak is exactly “p”. The cluster of positive cells was assumed to correspond to all the cells expressing the TF.

Case of two detected peaks (e.g. oli).

The sequence was split into the regions defined by each peak. Then the average number of positive cells “p1” and “p2” are computed for each of the two regions. Then the method proceeds within each region and its average number of positive cells as in the case of a single detected peak. This determines both the estimated length and the position of the two positive cell clusters.

The method leaves some uncertainty as to whether a cell located at the boundary of the cluster coverage (Step 6) should or not be considered as positive. If we consider that a cell at integer position x extends from x-0.5 to x+0.5, we can define a coverage index at the border (number from 0 to 1), revealing how much the positive cell cluster falls into the region around the cell (at −0.5 to +0.5).

For example, the length of the Jim positive cell cluster is 8.2 and its position is [14.8–23.1], which makes the two cells, cell 15 (14.8 belongs to [15–0.5, 15 + 0.5] and cell 23 (23.1 belongs to [23–0.5,23 + 0.5], situated at the boundary uncertain of Jim expression. The coverage index is thus calculated for both cells, 0.7 (i.e: 15.5–14.8) for cell 15 and 0.6 (i.e: 23.1–22.5) for cell 23 (Figure 2). Since the average number of cells expressing Jim is 8.3, either cell, other than both, could be considered as positive and the method cannot distinguish between them. The coverage index at the border is indicated for cell cluster when inferior to 1.

Test of the PCCD method by co-staining

The cell cluster detection method predicts the combinations of TFs expressed in each immature MN. To validate the positive-cell cluster detection method and to determine its accuracy, we tested its predictions by performing co-stainings. Among 256 possible co-stainings, we chose N = 12 co-staining combinations, serving as a proof-of-concept. We chose those combinations to test two parameters of the positive-cell cluster detection method: the accuracy of the position of the positive-cell cluster and the accuracy of the number of cells per cluster when 2 cell clusters were detected (length of the positive cell cluster).

The accuracy of the position of the positive cell cluster:

The number of cells expressing a given TF (length of the positive-cell cluster) is not a prediction of the model when only one cluster is detected because this parameter is based on real data: average number of positive cells expressing a given TF from all our experiments. However, we tested the accuracy of the position on the x’ axis of the positive-cell cluster by performing co-staining between TFs that have partially overlapping cluster coverage.

Kr and Zfh2 co-staining

Prediction of the PCCD: The model predicts 3 to 5 cells expressing both TFs with a higher probability of only 3 cells because there is a higher chance for cell #5 rather than cell #1 to express Kr (coverage index at cell #5 for Kr is 0.9 and 0.4 at cell #1) and low chance for cell #5 to express Zfh2 (coverage index at cell #5 for Zfh2 is 0.3). Result of the co-staining: Number of cells expressing Kr and Zfh2 = 3. We concluded, as predicted by the model, that cell #5 expresses Kr whereas cell #1 did not, and cell #5 is Zfh2 negative.

Lov and Zfh2 co-staining

Prediction of the PCCD: The model predicts 1 to 3 cells expressing both TFs with a higher probability of 2 cells because there is higher chance for cell #3 rather than cell #7 to express Lov (coverage index at cell #3 for Lov is 0.6 and 0.4 at cell #7) and low chance for cell #5 to express Zfh2 (coverage index at cell #5 for Zfh2 is 0.3). Result of the co-staining: Number of cells expressing Lov and Zfh2 = 2. Lov antibody seems to be sensitive to fixation conditions and we found one more cell expressing low level of Lov after optimizing the fixation conditions (Figure S2). Consequently, we corrected the length of the positive cell cluster of Lov from 4 to 5 and included both cell #3 and cell #7. We included cell #3 as the cell expressing low level of Lov because the cell expressing a low level of Lov is Zfh2+.

Chimmo and Jim co-staining

Prediction of the PCCD: The model predicts to have 3 to 5 cells expressing both TFs. Result of the co-staining: Number of cells expressing Chimmo and Jim = 4. Here, we have two possibilities: cell #15 is Jim+ and cell #19 is Chimmo- or cell #15 is Jim- and cell #19
is Chinmo+. In the PPCD method the length of the positive cell cluster of Chinmo is 18.5. After including the Chinmo+ cells from co-staining experiments the average number of cell expressing Chinmo is closer to 19. This variation between samples is due to the weak expression of Chinmo in one Jim+ cell localized ventrally, which is sometimes not detected (Figure S2). We thus concluded that cell #23 expresses Jim while cell #15 does not.

Conclusion: the combinatorial expression of TFs predicted by the method seems to be accurate since the co-staining validated the predictions.

The number of cells per cluster when 2 cell clusters are detected (length of positive cell cluster) and accuracy of the position of the positive cell cluster.

We tested the length of the positive dorsal cluster (left cluster) when two clusters were detected, by performing co-staining between TFs that were predicted to be completely overlapping in the dorsal cluster and not the ventral one. We then tested the accuracy of the position of the positive cell cluster as described in the previous paragraph. We tested the length, as well as the position, of cell clusters expressing RunxA, Zfh1, Oli and Nvy.

**RunxA and chinmo co-staining (testing the length of the dorsal RunxA + cluster)**

Prediction of the PPCD: The model predicts a length of the dorsal RunxA + cluster = 5.1 and a number of cells expressing RunxA and Chinmo = 5. Result of the co-staining: Number of cells expressing RunxA and Chinmo = 5. We concluded that the number of cells expressing RunxA in the dorsal is cluster is correct.

**RunxA and jim co-staining (testing the position of the dorsal RunxA + cluster)**

Prediction of the PPCD: The model predict 0 to 1 cell expressing both. Result of the co-staining: Number of cells expressing RunxA and Jim = 0. Because we demonstrated in the previous paragraph that cell #15 is Jim+, we could not conclude if the cell #10 or the cell #15 is RunxA+.

**Zfh1 and chinmo co-staining (testing the length of the dorsal Zfh1+ cluster)**

Prediction of the PPCD: The model predicts a length of the dorsal Zfh1+ cluster = 10.8 and a number of cells expressing Zfh1 and Chinmo = 10 or 11. Result of the co-staining: Number of cells expressing Zfh1 and Chinmo = 12. Based on these results, we made a correction of the length of the dorsal Zfh1+ cluster from 10.8 to 12.

**Zfh1 and Zfh2 co-staining (testing the position of the dorsal Zfh1+ cluster)**

Prediction of the PPCD: The model predicted 2 to 4 cells expressing both TFs. Result of the co-staining: Number of cells expressing Zfh1 and Zfh2 = 2. We concluded that cell #2 does not express Zfh1 and cell #5 does not express Zfh2.

**Zfh1 and Kr co-staining (testing the position of the dorsal Zfh1+ cluster)**

Prediction of the PPCD: The model predicts 2 to 4 cells expressing both TFs. Result of the co-staining: Number of cells expressing Zfh1 and Kr = 3. We had concluded (from Kr and Zfh2 co-staining) that cell #5 is Kr+, confirming that cell #2 does not express Zfh1.

Based on these three co-stainings including Zfh1, we concluded that cell #13 expresses Zfh1 while cell #2 does not. Moreover, we made a correction by adding cell #14 into the Zfh1+ cluster to respect a length of the Zfh1+ cluster = 12.

**Oli and Zfh2 co-staining (testing the length of the dorsal Oli + cluster)**

Prediction of the PPCD: The model predicts a length of the dorsal Oli + cluster = 2.8 and a number of cells expressing Oli and Zfh2 = 2 or 3. Result of the co-staining: Number of cells expressing Oli and Zfh2 = 2. Based on these results we made a correction of the length of the Oli+ cell cluster for the first Oli cluster. The length of the Oli+ cell cluster = 2 instead of 2.8.

**Oli and Kr co-staining (testing the position of the dorsal Oli + cluster)**

Prediction of the PPCD: The model predicts to have 1 to 3 cells expressing both. Result of the co-staining: Number of cells expressing Oli and Kr = 1. We concluded that cell #1 expresses Oli whereas cell #3 does not.

**Nvy and Chinmo (testing the length of the dorsal Nvy + cluster)**

Prediction of the PPCD: The model predicts a length of the dorsal Nvy + cluster = 3.6 and a number of cells expressing Nvy and Chinmo = 3 or 4. Result of the co-staining: Number of cells expressing Nvy and Chinmo = 4. Notably, we notice that the dorsal Nvy+ cluster express higher level of nvy compare to the ventral cluster.

**Nvy and Jim (testing the position of the dorsal Nvy + cluster)**

Prediction of the PPCD: The model predicts 0 to 2 cells expressing both TFs. Result of the co-staining: Number of cells expressing Nvy and Jim = 2. However, these two cells express a very low level of Nvy and are located ventrally in the 14 samples analyzed. We concluded that cell #12 instead of #16 is Nvy+. See below for comments on the ventral clusters and why the PCCD method did not predict that 2 cells of the ventral Nvy+ cluster express Jim.

Conclusion: The length of the cluster coverage when two clusters are detected, is accurate. No corrections or sometimes corrections of one cell to cluster length were made.

Note for the ventral clusters: The PCCP method probably underestimated the number of cells that should be included on the left region of clusters located near the NB. Indeed, the frequency of TF expression in a cell close to the NB was low, not because of the variation in the cell positioning between samples, but due to the fact that the cell may not have been born in some samples. This would artifactually reduce the frequency of TF expression at positions close to the NB. This bias could be overcome in a future version of the PCCD method. However, the Nvy and Jim co-staining revealed that the errors were small. Finally, all ventral clusters had a gradient expression from high (ventral) to low (dorsal). Consequently, the boundary of the left region of these clusters was not sharp due to weak TF expression that was sometimes difficult to detect.
Lin A/15 MARCM clones vs Lin A/15 memory system

The one-spot MARCM technique makes it possible to visualize one of the daughter cells derived from a common progenitor (Lee and Luo, 1999). Consequently, the generation of an NB MARCM Lin A/15 clone at the larval stage, when the NB is quiescent, makes it possible to visualize the first MN born or the next 28 MNs, but not all together. However, the Lin A/15 memory system (Awasaki et al., 2014; Lacin and Truman, 2016) allows the visualization of all 29 MNs.

Link between birth order and muscle targets: Corrections of (Baek and Mann, 2009)

A previous study revealed a link between the Lin An MN birth order and the muscles they target (Baek and Mann, 2009). In our study (Figure 1), we made two corrections by adding two MNs (Lin A Tr1 and B1), not described in (Baek and Mann, 2009). The relative birth orders of the 28 Lin A/15 MNs have been characterized by inducing GMCs MARCM clones at different time points during larval stages. The generation of GMC MARCM clones has allowed the determination of the relative birth order of 27 MNs out the 28 MNs. In our study, when we generated an NB MARCM clone to label the 28 Lin A/15 MNs, we always marked an MN targeting the trochanter segment, an MN which was not considered to be part of Lin A/15 (Baek and Mann, 2009). We named this MN, Lin A/15 Tr1. Based on the shape of the axonal terminal branches, we propose that this MN has been misidentified (Baek and Mann, 2009) as Lin G, a lineage supposedly producing a single MN. Here, we propose that Lin G is not a lineage by itself (discovered by generating an NB MARCM clone) but a GMC MARCM clone. Moreover, we conclude that Lin A/15 Tr1 MN is the first born MN of the 28 MNs by comparing in the Baek and Mann article (2009), the last time point where the Tr1 GMC MARCM clone can be induced with the other GMC MARCM clones. Finally, the NB one-spot MARCM technique, cannot label the first Lin A/15 MN. The twin spot QMARCM/MARCM has revealed that the first-born MN from Lin A/15 targets a body wall muscle (Enriquez et al., 2018). Here, we name this MN: B1.

QUANTIFICATION AND STATISTICAL ANALYSIS

Graphs of the relative position of each LinA/15 cell were generated with Microsoft Excel. The spatial coordinates were assigned to each cell using the cell counter plug-in of NIH ImageJ software. The coordinates of each cell were normalized with Microsoft Excel in order to have the position of the Lin A/15 NB at the origin of the plot graph. For samples where the NB was not labeled, the coordinates of each cell were then normalized to a cell located most anteriorly.

The plots of the number of TF-expressing cells (TFs are as indicated in each graph) were generated with Prism (GraphPad Software). All error bar represents standard deviation of a minimum of 7 Lin A/15 samples, each dot represents a single Lin A/15 sample analyzed. Student’s t test was performed to compare the difference in between indicated groups. Differences of p < 0.05 were considered significant. *0.01 < p < 0.05; **0.001 < p < 0.01; ***0.0001 < p < 0.001; ****p < 0.0001.