PNAS

Supporting Information for

C. elegans Sine oculis/SIX-type homeobox genes act as homeotic switches to define neuronal subtype identities

Cyril Cros and Oliver Hobert

Oliver Hobert
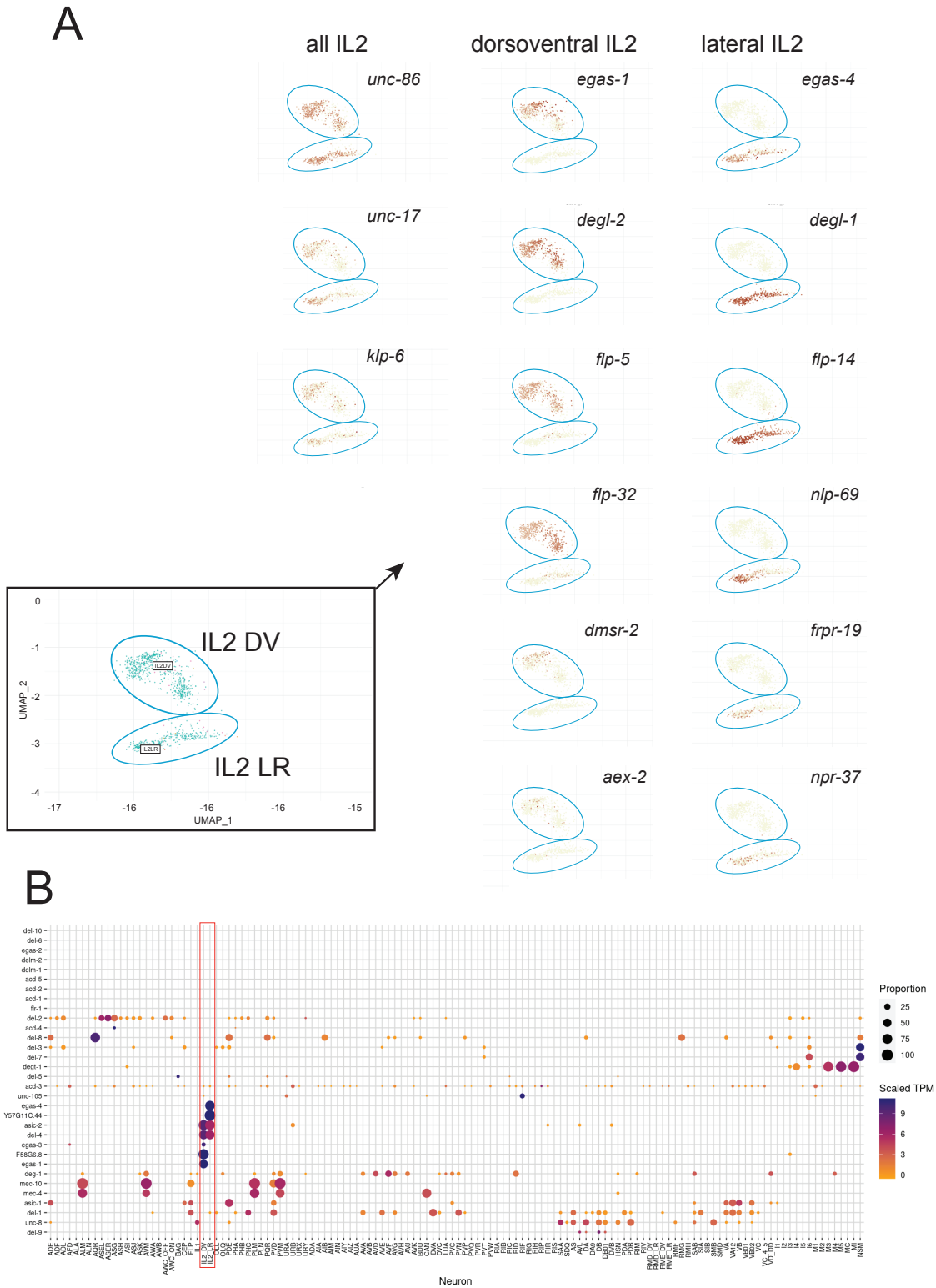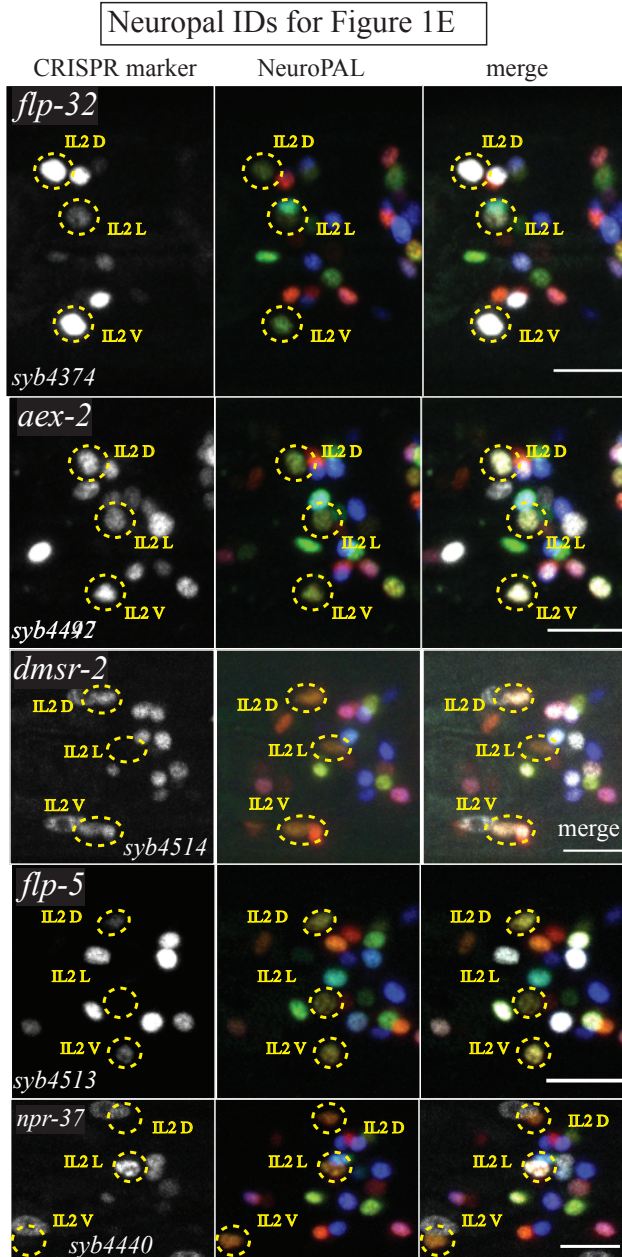Email: or38@columbia.edu

This PDF file includes:

Fig.S1: IL2 subtype markers revealed by scRNA analysis
Single cell transcriptome data, displayed with the Cengen App (Taylor et al., 2021).
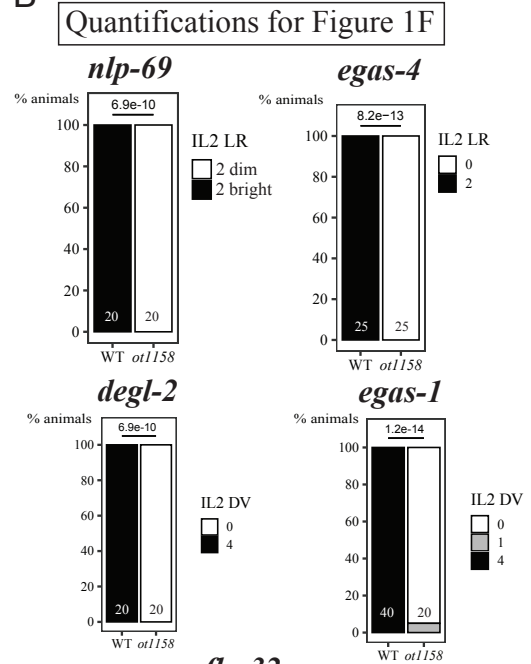**A:** Single cell UMAP plots for each individual IL2 markers.
B: Expression of all members of the DEG/ENaC/ASIC family of ion channels throughout the *C. elegans* nervous system.
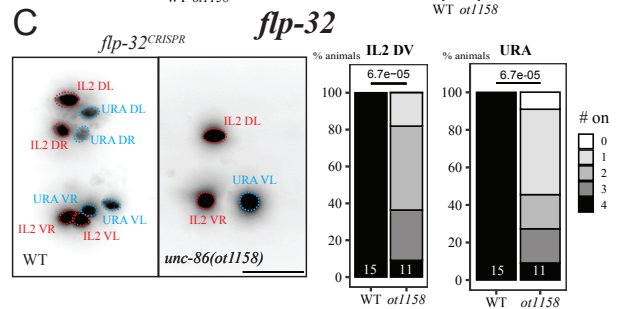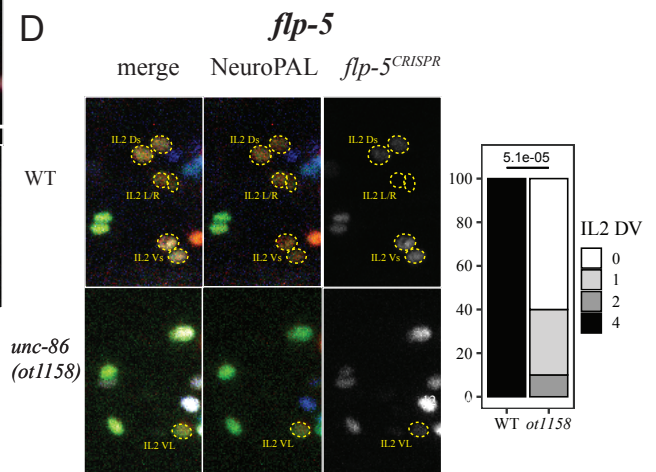
Fig.S2: NeuroPAL IDs and scoring for *unc-86* dependency of IL2 subtype markers

A: NeuroPAL pictures for Fig 1F CRISPR reporters *flp-32(syb4374), aex-2(syb4447), dmsr-2(syb4514), flp-5(syb4513), npr-37(syb4440)*

B: Quantification of Fig 1 G *unc-86* dependent IL2 lateral markers *nlp-69(syb4512) egas-4 / otIs833* and IL2 dorsoventral markers *degl-2(syb5229) egas-1 / otIs846*, p-values by Fisher two-sided test.

C: *flp-32* IL2 DV and URA expression is dependent on *unc-86,* using the *flp-32(syb4374)* reporter in the *unc-86(ot1158)* null allele. *flp-32* is expressed at a much weaker level in IL2 L/R and was not scored here

D: *flp-5* IL2 DV expression is *unc-86* dependent, using the *flp-5(syb4513)* reporter in the *unc-86(ot1158)* null allele.
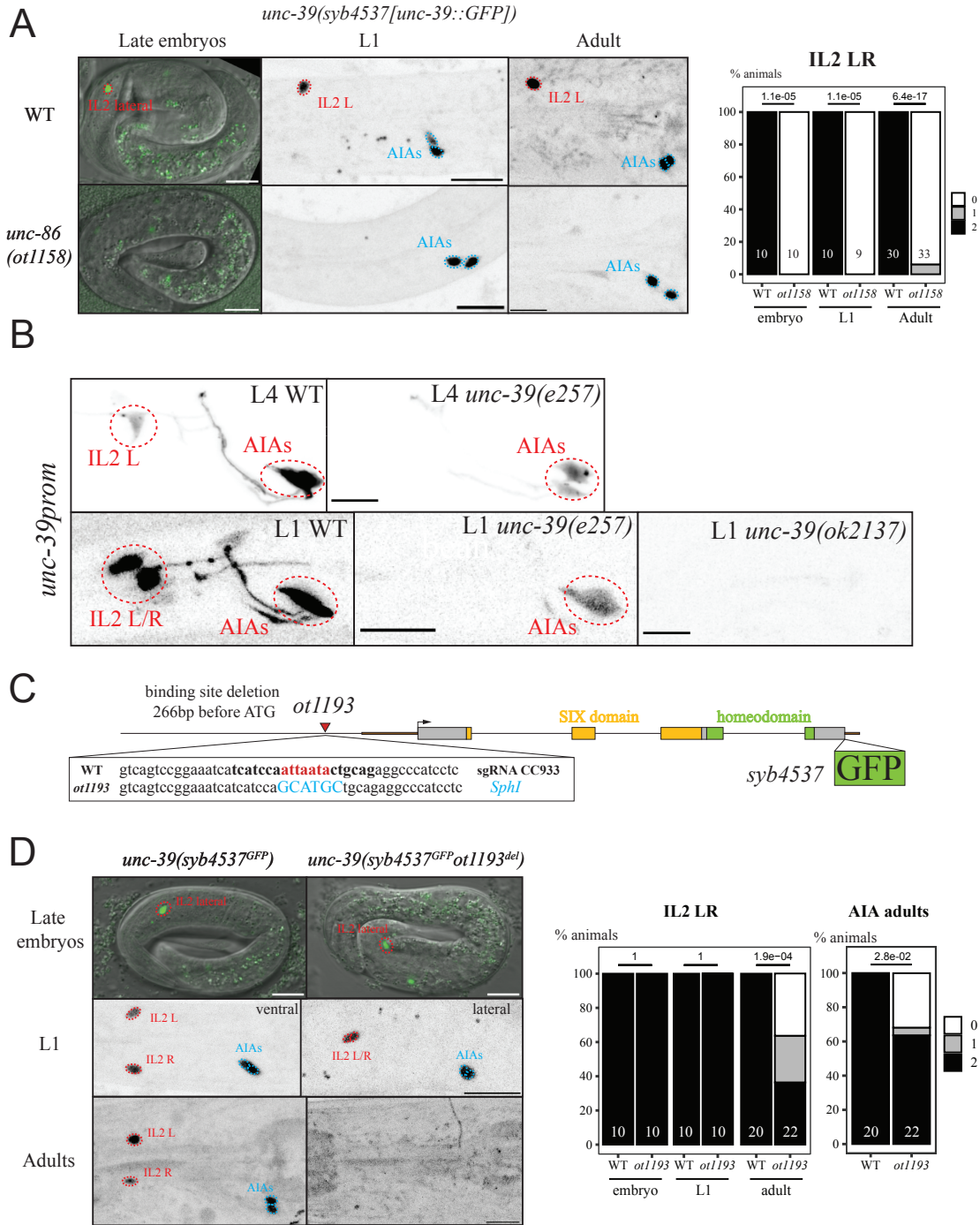
3

Fig.S3: *unc-39* regulation

A: Initiation and maintenance of *unc-39* expression in the lateral IL2 neurons (measured with the CRISPR/Cas9-engineered reporter allele *syb4537*) is *unc-86* dependent. AIA neurons do not express *unc-86* and *unc-39* expression is unaffected in these neurons.

B: A 2.3kb *unc-39* promoter fusion (chromosomal integrant *otIs854*) displays diminished expression in both lateral IL2 and AIA in the *unc-39* mutants, with stronger effects in IL2 vs AIA, and in the *ok2137* null animals (which can only be scored as arrested L1) vs the *e257* hypomorph (which can be scored at later stages).

C: Molecular details of the *unc-39(syb4537^{GFP}o1193^{bs_del})* putative binding site deletion.

D: The homeodomain binding site mutation engineered with CRISPR/Cas9 into the *unc-39* reporter allele (*unc-39(syb4537^{GFP}ot1193^{bs_del})*) does not affect embryonic initiation but maintenance of *unc-39* expression in adult lateral IL2 and AIA neurons.

5

Fig.S4: *unc-39* **gene model and functional analysis.**

A: *unc-39* gene model with SIX and homeodomain protein domains, and mutant alleles (null *ok2137* and *unc-39*[R203Q] hypomorphs *e257, ot1171, ot1172* and *ot1173*).

B: Effect of *unc-39* on various IL2 fate markers *ins-1(syb5452) dmsr-2(syb4514) flp-32(syb4374)* (from **Fig.2E/F**), with NeuroPAL images included for cell identification purposes. Note that a pair of pharynx neurons (blue circle, with no NeuroPAL coloring) express *ins-1*.

6

C: Quantifications for **Fig.2E/F** marker losses with markers *nlp-69(syb4512[nlp-69::SL2::gfp::H2B]), ins-1(syb5452[ins-1::SL2::gfp::H2B]), flp-14(syb3323[flp-14::SL2::gfp::H2B]), degl-1 (otIs825)* and *egas-4 (otIs833)* for the lateral IL2; *dmsr-2(syb4514[dmsr-2::SL2::gfp::H2B]), flp-32(syb4374[flp-32::SL2::gfp::H2B]), degl-2(syb5229[degl-2::SL2::gfp::H2B])* and *egas-1(otIs846)* for the dorsoventral IL2. Dorsoventral marker gain was counted as all or nothing, reporting extra cell counts (0,1 or 2 cells past the usual 4 cells). For *flp-32* the very dim lateral IL2 wild type expression was not counted as proper expression, in the mutant IL2 are always brighter than this. Fisher test p-values are shown on top, number of worms on the bottom.

D: Pan-IL2 identity is not affected in the *unc-39* null or the *unc-39*[R203Q] hypomorphic allele, as assessed by unaffected expression of the *unc-17prom9* fragment (*otEx6986*), the *cho-1* fosmid reporter (*otIs354*), the *klp-6* reporter *myIs13*. the IL1 *flp-3* reporter *otIs703*, the *unc-17* reporter allele *syb4491* or the NeuroPAL color code (which also derives from *klp-6).*

E: Quantifications for **Fig.2E/F** *unc-39* misexpression experiments. Lateral markers, *egas-4 (otIs833)* and *nlp-69(syb4512[nlp-69::SL2::gfp::H2B]),* are fully off in *e257* control*,* but were counted as on even if only weakly expressed. Dorsoventral markers, *egas-1(otIs846) and degl-2(syb5229[degl-2::SL2::gfp::H2B])*, are counted as off if visibly weaker than in the *e257* control. Each dot is a worm, on the y-axis is the number of neurons expressing the marker. There are sometimes more neurons than just the six IL2 for *egas-4* and *nlp-69*. Wilcoxon rank sum tests were done for each line vs the hypomorph with p-values on top, number of worms on the bottom.

F: Effect of ectopic *unc-39* on *flp-32(syb4374[flp-32::SL2::gfp::H2B]),* expression (from **Fig.2E**), with NeuroPAL images included, as well as quantifications below for IL2 on the left, URA on the right. The *unc-39(ot1073)* animals are shown in a ventral view.
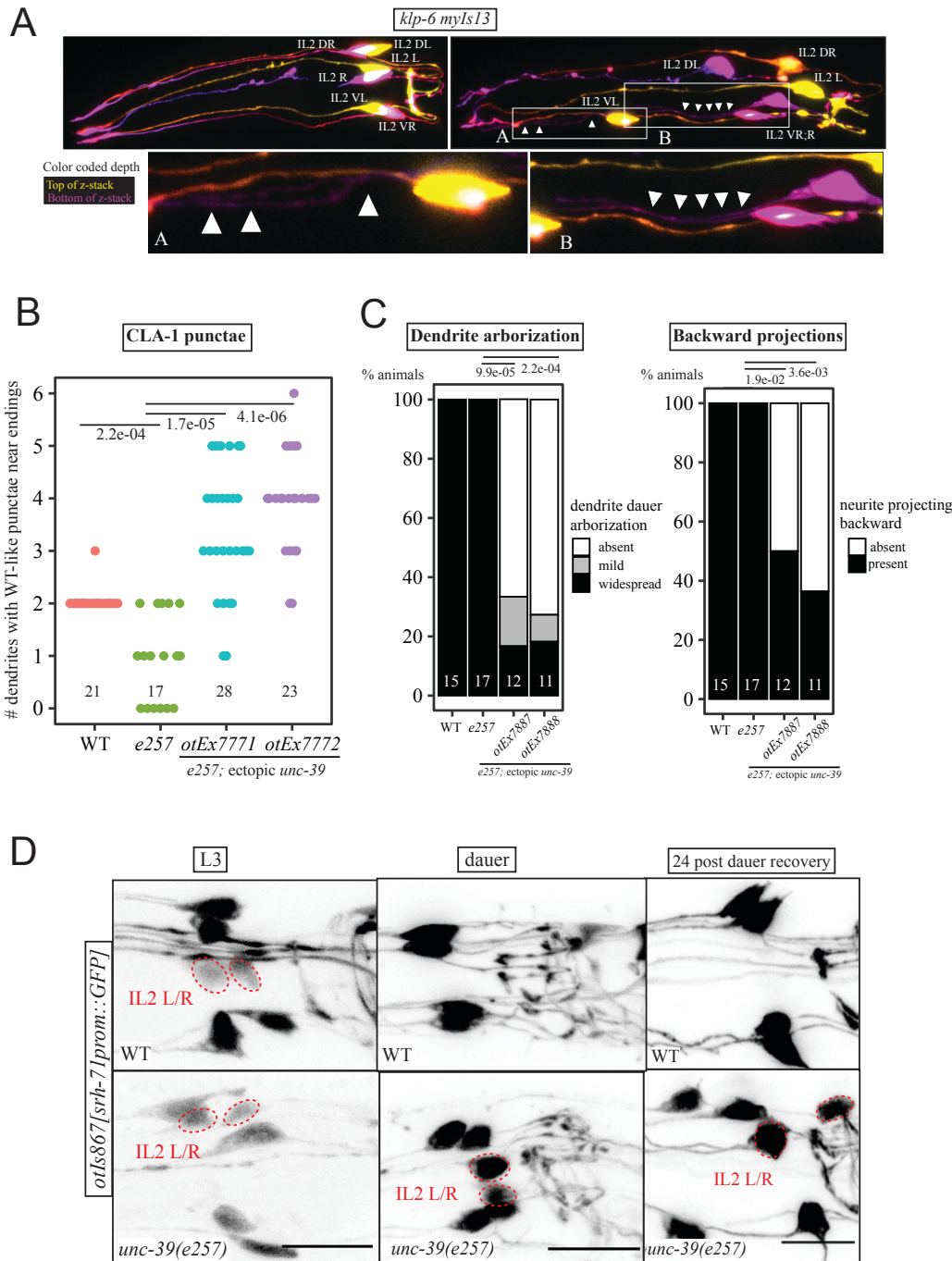
Fig.S5: *unc-39* affects IL2 morphology
A: Top: same images as Fig.3A of a *klp-6 (myIs13)* cytoplasmic reporter. To better show fasciculation it is now depth-coded (orange is top of the stack, purple on bottom). On the left a wildtype image of the left/right pairs show up as yellow/purple, but on the right *unc-39* mutant panel, the right lateral IL2 joins the right ventral IL2 cell body and their dendrite fasciculate. **Bottom:** blow-up images.
B: GFP::CLA-1 punctae quantification, by taking the number of dendrites per worm that have multiple bright punctae before the cilia, like in the wild type lateral IL2. P-values are computed by Wilcoxon rank sum test.
C: Dauer arborization phenotype quantification. We scored if branching was similar to wild type and *e257* mutant animals, reduced or fully gone. We then scored whether we observed neurites projecting backward.
D: The same *srh-71* reporter *(otIs867)* is shown left to right at L3, in dauer and 24 hours after replating dauers on fresh food. Wildtype animals are on top, *unc-39* mutants at the bottom. Notice that after recovery the lateral IL2 are gone in the wild type, but are equally bright as the dorsoventral IL2.
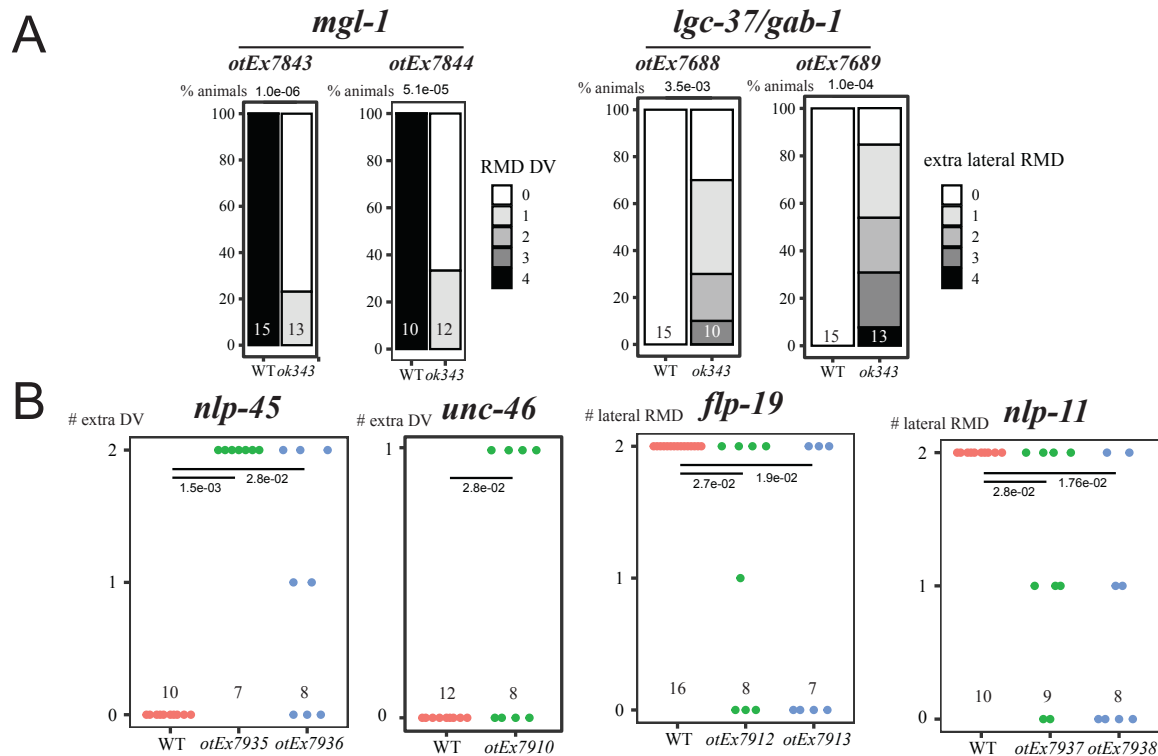
Fig.S6: Quantification of *ceh-32* loss- and gain-of-function experiments
A: Quantifications for the loss of *mgl-1* in the dorsoventral RMD across the two lines (*otEx7843*, *otEx7844*) and the gain of *lgc-37* in the lateral RMD (counting only extra cells beyond the usual two lateral RMD), in lines *otEx7688* and *otEx7689*. P-values were obtained by two-sided Fisher test.
B: Quantification for the misexpression experiments. Dorsoventral markers *nlp-45(ot1032)* and *unc-46 otIs568* are gained (only one lateral RMD at most for *unc-46).* The lateral markers *flp-19(syb3278)* and *nlp-11(syb4759)* are lost. Note that only one line is shown for *unc-46,* with only one lateral RMD being reprogrammed at most*;* another line *otEx7911* was considered but had no effect on the lateral RMD. P-values were obtained by Wilcoxon rank sum test.
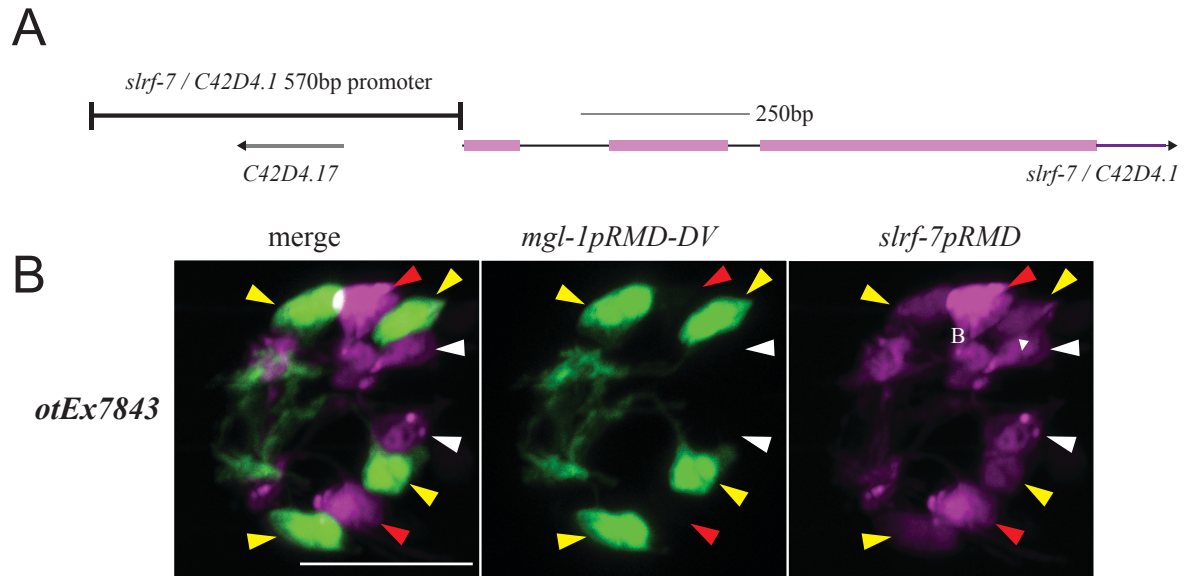
A

slrf-7 / C42D4.1 570bp promoter

250bp

C42D4.17

slrf-7 / C42D4.1

B

merge	mgl-1pRMD-DV	slrf-7pRMD

otEx7843

Fig.S7: Characterization of the *C42D4.1/slrf-7* pan-RMD driver
A: Genome structure of our *C42D4.1/slrf-7* promoter we use as a pan-RMD driver
B: Annotated close up of the *otEx7843* panel from **Fig.4C**, demonstrating that *slrf-7* is expressed in all RMDs. The animal is in a ventral position. The yellow triangles point to cells expressing both *slrf-7* and a *mgl-1* fragment known to be specific of the dorsoventral RMD subtype. The red triangles point to the lateral RMD, between the RMD of the previous subtype, with a brighter *slrf-7* expression in this subtype. White triangles point to the dorsal SAAs, based on both the CeNGEN scRNA atlas and the fact that they make projections anteriorly in the sublaterals. The ventral SAA rarely express *slrf-7*, depending on transgenic lines.
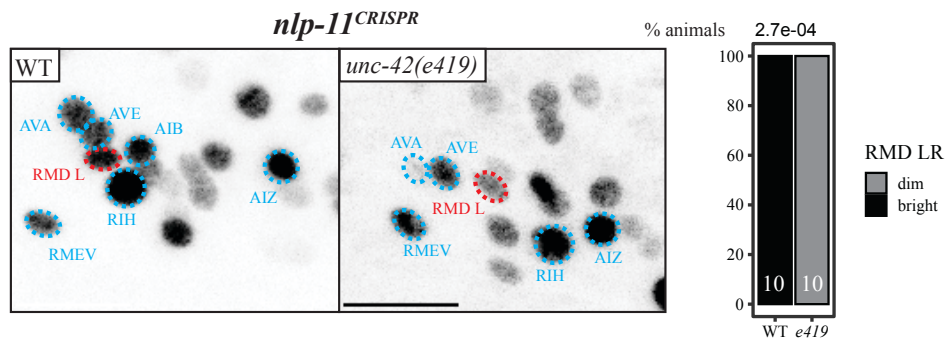
Fig.S8: Effect of *unc-42* on *nlp-11* expression
The lateral RMD subtype marker *nlp-11(syb4759)* becomes dim in the *unc-42(e419)* mutant. The lateral RMD are circled in red, neighbouring neurons are circled in blue. We used *unc-42* independent neurons like RMEV or RIH to assess brightness changes. Significance was calculated by Fisher two-sided test.

Table S1: Strain list

Strain Name	Mutant or knock-in	Array name	DNA on array	Reference
OH17019		otIs825	degl-1prom::GFP unc-122prom::GFP	This study
OH17072		otIs833	egas-4prom::TagRFP unc-122prom::GFP	This study
OH17217		otIs846	egas-1promGFP unc-122prom::GFP	This study
PHX5452	ins-1(syb5452[ins-1::SL2::gfp::H2B])			This study
PHX3323	flp-14(syb3323 [flp-14::T2A::3XNLS::GFP])			This study
PHX4512	nlp-69(syb4512[nlp-69::SL2::gfp::H2B])			This study
PHX4440	npr-37(syb4440[npr-37::SL2::gfp::H2B])			This study
PHX5229	degl-2(syb5229[degl-2::SL2::gfp::H2B]),			This study
PHX4374	flp-32(syb4374[flp-32::SL2::gfp::H2B])			This study
PHX4447	aex-2(syb4447[aex-2::SL2::gfp::H2B])			This study
PHX4513	flp-5(syb4513[flp-5::SL2::gfp::H2B]),			This study
PHX4514	dmsr-2(syb4514 [dmsr-2::SL2::gfp::H2B])			This study
OH17241	unc-86(ot1158)			This study
PHX4537	unc-39(syb4537[unc-39::gfp])			This study
OH17368		otIs854	unc-39prom::TagRFP	This study
OH17657	unc-39(syb4537ot1193[unc-39binding_site_mutant::unc-39::gfp])			This study
CB257	unc-39(e257)			(Brenner 1974)
OH16543	unc-39(ok2137)	otEx7581	unc-39 fosmid WRM0636cG07 NotI cut @10 + inx-6p18::GFP @6	This study
OH17378	nlp-69(syb4512[nlp-69::SL2::gfp::H2B]) unc-39(ot1172[unc-39^{R203Q}])			This study
OH15363	him-5(e1490)	otIs669	Bright NeuroPAL on V	(Yemini et al. 2021)
OH17491	unc-39(ot1173[unc-39^{R203Q}]) him-5(e1490)	otIs669		This study
OH17709	nlp-69(syb4512) unc-39(ot1172)	otEx7870	unc-39 mix 1*	(Flames and Hobert 2009)
OH17710	nlp-69(syb4512) unc-39(ot1172)	otEx7871	unc-39 mix 1*	(Flames and Hobert 2009)
OH17713	unc-39(e257)	otIs833 otEx7874	unc-39 mix 1*	(Flames and Hobert 2009)
OH17714	unc-39(e257)	otIs833 otEx7875	unc-39 mix 1*	(Flames and Hobert 2009)

OH17709	degl-2(syb5229) unc-39(e257)	otEx7868	unc-39 mix 1*	(Flames and Hobert 2009)
OH17710	degl-2(syb5229) unc-39(e257)	otEx7869	unc-39 mix 1*	(Flames and Hobert 2009)
OH17711	unc-39(e257)	otIs846 otEx7872	unc-39 mix 1*	(Flames and Hobert 2009)
OH17712	unc-39(e257)	otIs846 otEx7873	unc-39 mix 1*	(Flames and Hobert 2009)
OH17709	nlp-69(syb4512) unc-39(ot1172)	otEx7870	unc-39 mix 1*	(Flames and Hobert 2009)
OH15056	him-5(e1490) pha-1(e2123)	otIs354 otEx6986	otIs354[cho-1::YFP fosmid] otEx6986[unc-17prom9::NLS:TagRFP pha-1 rescue]	(Flames and Hobert 2009; Serrano-Saiz et al. 2020)
OH15601		myIs13; otIs703	myIs13[klp-6prom::GFP] otIs703[flp-3prom::mCherry]	myIs13: (Schroeder et al. 2013)
OH17741	unc-39(e257)	myIs13 otIs703 otEx7887	unc-39 mix 1*	This study
OH17742	unc-39(e257)	myIs13 otIs703 otEx7888	unc-39 mix 1*	This study
OH16970		otIs815	klp-6p::mCherry + klp-6p::GFP::cla-1(S)	This study
OH17357		otIs815 otEx7771	unc-39 mix 2*	This study
OH17358		otIs815 otEx7772	unc-39 mix 2*	This study
PHX4693	che-7(syb4693[che-7::SL2::gfp::H2B]),			This study
OH17377	che-7(syb4693) unc-39(ot1171[unc-39R203Q])			This study
OH17502		otIs867	srh-71prom::GFP 50ng/ul; pha-1 rescue 50ng/ul	This study
OH16477	ceh-32(ot1040[ceh-32::gfp(SEC)])			(Reilly et al. 2020)
OH15338	ceh-32(ok343)	otEx7146	ceh-32fosmid WRM0637dA10 ;myo-2::RFP	This study
OH17709	nlp-69(syb4512) unc-39(ot1172)	otEx7870	unc-39 mix 1*	(Flames and Hobert 2009)
OH17954	ceh-32(ok343)	otEx7843	mgl-1promRMDD/V (pFZ50 @30) slrf-7promRMD::TagRFP (pCC265 @30)	(Flames and Hobert 2009)
OH17955	ceh-32(ok343)	otEx7844	mgl-1promRMDD/V (pFZ50 @30) slrf-7promRMD::TagRFP (pCC265 @30)	(Flames and Hobert 2009)
OH16909	ceh-32(ok343) pha-1(e2190)	otEx7688	mgl-1promRMDL/R (pMG249 @30) slrf-7promRMD::TagRFP	(Flames and Hobert 2009)

			(pCC265 @20) pha-1 rescue @50	
OH16911	*ceh-32(ok343) pha-1(e2190)*	*otEx7689*	*mgl-1promRMDL/R (pMG249 @30) slrf-7promRMD::TagRFP (pCC265 @20) pha-1 rescue @50*	(Flames and Hobert 2009)
OH16380	*nlp-45(ot1032[nlp-45::T2A::GFP:H2B])*			(Sun and Hobert 2021)
OH13026		*otIs568*	*unc-46fosmid::SL2::H2B::m CHOPTI; pha-1 rescue*	(Gendrel et al. 2016)
PHX3278	*flp-19(syb3278 [flp-19::T2A::3×NLS::GFP])*			This study
PHX4759	*nlp-11(syb4759[nlp-11::SL2::GFP::H2B])*			This study
OH17794	*nlp-11(syb4759[nlp-11::SL2::GFP::H2B]) unc-42(e419) him-5(e1490)*	*otIs669*		This study
OH17322	*him-8(e1489) ceh-32(ot1040[ceh-32::gfp(SEC)]) unc-42(ot1164)*	*otIs696* (Yemini et al. 2021)	Dim NeuroPAL mix	This study
OH17900	*nlp-45(ot1032[nlp-45::T2A::GFP:H2B])*	*otEx7935*	*ceh-32* mix 1*	This study
OH17901	*nlp-45(ot1032[nlp-45::T2A::GFP:H2B])*	*otEx7936*	*ceh-32* mix 1*	This study
OH17806		*otIs568 otEx7909*	*ceh-32* mix 2*	This study
OH17906	*nlp-11(syb4759[nlp-11::SL2::GFP::H2B])*	*otEx7937*	*ceh-32* mix 1*	This study
OH17907	*nlp-11(syb4759[nlp-11::SL2::GFP::H2B])*	*otEx7938*	*ceh-32* mix 1*	This study
OH17809	*flp-19(syb3278 [flp-19::T2A::3×NLS::GFP])*		*ceh-32* mix 2*	This study
OH17810	*flp-19(syb3278 [flp-19::T2A::3×NLS::GFP])*		*ceh-32* mix 2*	This study

The *otIs703[flp-3prom::mCherry]* transgenes and the plamids used in the IL2 *CLA-1* strain were kindly provided by Maryam Majeed, the *ceh-32* rescue strain OH15338 was created by Molly Reilly, and *unc-86(ot1158)* by Tessa Tekieli.
pFZ50(*mgl-1promRMDD/V::GFP*): (Zhang et al. 2014)
pMG249(*lgc-37prom5.4kbRMDL/R::GFP*): (Gendrel et al. 2016)

* Ectopic expression mixes for *unc-39* and *ceh-32* misexpression

- ***unc-39* mix 1**:
  Ex[unc-17p9::unc-39::unc-39(pCC298@25ng/ul) inx18pAVG::TagRFP(pCC220@25ng/ul)] used for most *unc-39* misexpression
- ***unc-39* mix 2:**
  Ex[unc-17p9::unc-39::unc-39(pCC298@15ng/ul) unc-122prom::mCherry @30ng/ul)] used for the *unc-39* misexpression in the *otIs815* IL2 *CLA*-1 strain

- ***ceh-32* mix 1:**
  Ex[slrf-7p::ceh-32 (pCC327@40ng/ul); slrf-7P::tagRFP (pCC265@30ng/ul)] *ceh-32* misexpression and labelling in all RMDs.
- ***ceh-32* mix 2:**
  Ex[slrf-7p::ceh-32 (pCC327@40ng/ul); inx-6prom(2TAAT-deletion)::GFP (pAB1@30ng/ul)] *ceh-32* misexpression with a procorpus coinjection marker, no RMD labelling

SI References

Brenner S. 1974. The genetics of Caenorhabditis elegans. *Genetics* **77**: 71-94.
Flames N, Hobert O. 2009. Gene regulatory logic of dopamine neuron differentiation. *Nature* **458**: 885-889.
Gendrel M, Atlas EG, Hobert O. 2016. A cellular and regulatory map of the GABAergic nervous system of C. elegans. *eLife* **5**.
Reilly MB, Cros C, Varol E, Yemini E, Hobert O. 2020. Unique homeobox codes delineate all the neuron classes of C. elegans. *Nature* **584**: 595-601.
Schroeder NE, Androwski RJ, Rashid A, Lee H, Lee J, Barr MM. 2013. Dauer-specific dendrite arborization in C. elegans is regulated by KPC-1/Furin. *Curr Biol* **23**: 1527-1535.
Serrano-Saiz E, Gulez B, Pereira L, Gendrel M, Kerk SY, Vidal B, Feng W, Wang C, Kratsios P, Rand JB et al. 2020. Modular Organization of Cis-regulatory Control Information of Neurotransmitter Pathway Genes in Caenorhabditis elegans. *Genetics* **215**: 665-681.
Sun H, Hobert O. 2021. Temporal transitions in the post-mitotic nervous system of Caenorhabditis elegans. *Nature* **600**: 93-99.
Yemini E, Lin A, Nejatbakhsh A, Varol E, Sun R, Mena GE, Samuel ADT, Paninski L, Venkatachalam V, Hobert O. 2021. NeuroPAL: A Multicolor Atlas for Whole-Brain Neuronal Identification in C. elegans. *Cell* **184**: 272-288 e211.
Zhang F, Bhattacharya A, Nelson JC, Abe N, Gordon P, Lloret-Fernandez C, Maicas M, Flames N, Mann RS, Colon-Ramos DA et al. 2014. The LIM and POU homeobox genes ttx-3 and unc-86 act as terminal selectors in distinct cholinergic and serotonergic neuron types. *Development* **141**: 422-435.