Homo- and heterodimerization of bHLH transcription factors balance stemness and bipotential differentiation

Aleix Puig-Barbé[1,2], Vinicius Dias Nirello[3], Sina Azami[4], Bruce A Edgar[5,6], Patrick Varga-Weisz[3,7], Jerome Korzelius[5,8] and Joaquín de Navascués[1,7,✉]

[1] European Cancer Stem Cell Research Institute, School of Biosciences, Cardiff University, Hadyn Ellis Building, Maindy Road, Cardiff CF24 4HQ, UK.

[2] Current address: EMBL-European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK.

[3] International Laboratory for Microbiome Host Epigenetics, Department of Genetics, Evolution, Microbiology, and Immunology, Institute of Biology, University of Campinas, Campinas, SP 13083-862, Brazil

[4] Max Planck Institute for Biology of Ageing, Joseph-Stelzmann-Straße 9B, 50931 Köln, Germany

[5] DKFZ/ZMBH Alliance, University of Heidelberg, Heidelberg, Germany.

[6] Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112, USA.

[7] School of Life Sciences, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK.

[8] School of Biosciences, University of Kent, Canterbury, Kent, CT2 7NZ, UK.

✉ j.denavascues@essex.ac.uk

Running title: Da, Emc and Sc balance stemness and differentiation in the fly intestine

Key words: bHLH / stemness / differentiation / bipotent stem cell / intestinal stem cell

ABSTRACT

Multipotent adult stem cells must balance self-renewal with differentiation into various mature cell types. How this activity is molecularly regulated is poorly understood. By using genetic and molecular analyses in vivo, we show that a small network of basic Helix-Loop-Helix (bHLH) transcription factors controls both stemness and bi-potential differentiation in the *Drosophila* adult intestine. We find that homodimers of Daughterless (Da, homolog to mammalian E proteins) maintain the self-renewal of intestinal stem cells and antagonise the activity of heterodimers of Da and Scute (Sc, homolog to ASCL and known to promote intestinal secretory differentiation). We find a novel role for the HLH factor Extramacrochaetae (Emc, homolog to Id proteins), titrating Da and Sc to promote absorptive differentiation. We further show that Emc prevents committed absorptive progenitors from de-differentiating, underscoring the plasticity of these cells. Switching physical interaction partners in this way enables the active maintenance of stemness, but primes stem cells for differentiation along two alternative fates. Such regulatory logic could be recapitulated in other bipotent stem cell systems.

INTRODUCTION

The regulation of stem cell fate decisions hinges on the control of transcription. Central to this control are gene regulatory networks, whose activity and dynamics steer cells along particular differentiation pathways (Graf and Enver, 2009; Levine and Davidson, 2005; Moris et al., 2016). These pathways can often be construed as a succession of binary steps that are regulated by the cross antagonism of transcription factor pairs (Graf and Enver, 2009; Simon et al., 2018). However, for multipotent stem cells, several fate options may be simultaneously available through higher-dimensional switches (Cinquin and Demongeot, 2002; Cinquin and Demongeot, 2005). To understand the choice between several alternative fates, knowledge of the interactions between transcriptional regulators is essential.

Intestinal stem cells (ISCs) are a clear example of multipotency in adult tissues. In *Drosophila* and mammals, ISCs face a triple choice between self-renewal and differentiation into one of two mature lineages: secretory or absorptive (reviewed in Crosnier et al., 2006; Jiang and Edgar, 2012; Philpott and Winton, 2014). The secretory lineage comprises several cell types in vertebrates, but only enteroendocrine (EE) cells in *Drosophila*. Absorptive cells are enterocytes (ECs). In *Drosophila*, ISCs produce lineage-specific precursors through distinct molecular triggers. High Notch signalling induces formation of enteroblasts (EBs), which will give rise to ECs. Expression of the bHLH (basic Helix-Loop-Helix) transcription factors Scute (Sc) or Asense (Ase), members of the *achaete-scute Complex* (*AS-C*, homologs of *ASCL* mammalian genes) induces the formation of EE precursor cells (pre-EEs), which quickly turn into mature EE cells (Bardin et al., 2010; Chen et al., 2018; Micchelli and Perrimon, 2006; Ohlstein and Spradling, 2006; Zeng and Hou, 2015). Whether ISCs face a multiple choice, or a sequence of binary decisions, is not clear (Perdigoto *et al*, 2011; Gervais & Bardin, 2017). Moreover, very little is known about potential molecular mechanisms that could regulate, simultaneously, self-renewal and bi-potential differentiation.

Transcription factors of the bHLH family control fate in multiple developmental contexts, from sex determination and mesoderm specification, to neurogenesis and immune cell type specification (reviewed in Bertrand et al., 2002; Murre, 2005; Murre et al., 1994). Their HLH motif mediates dimerization, while the preceding region, rich in basic amino acids, allows DNA binding (Massari and Murre, 2000); other domains

provide transcriptional regulation. The class I of bHLH factors comprises factors such as E2A, E2.2 and HEB (E proteins, encoded by *TCF3/4/12* in mammals, and a single gene, *daughterless –da–*, in *Drosophila*), which are usually expressed in broad spatial domains. By contrast, class II bHLH factors (A proteins) such as MYOD, TWIST and ASCL (MyoD, Twi, and proteins encoded by genes of the *AS-C*, in *Drosophila*) are usually expressed in more restricted patterns. Class II factors generally heterodimerise with a class I member to bind DNA, while class I factors can either homo- or heterodimerise (Murre et al., 1994). The class V comprises HLH factors lacking the basic domain, known as *Inhibitors of DNA binding* (Id proteins; a single gene, *extra macrochaetae –emc–*, in *Drosophila*), as heterodimerisation with them renders dimers unable to bind DNA (Murre et al., 1994). Combinatorial, antagonistic or cooperative interactions confer unique properties to regulatory networks based on these factors (Amoutzias et al., 2008).

Class II bHLH factors are known to regulate differentiation in the metazoan intestine (reviewed in Hartenstein et al., 2017; Philpott and Winton, 2014). In *Drosophila*, Sc and Ase can initiate EE differentiation (Amcheslavsky et al., 2014; Bardin et al., 2010; Chen et al., 2018; Guo and Ohlstein, 2015), while other bHLH genes maintain EE function (*dimmed*, homolog of *NeuroD*; Beebe et al., 2015), or promote their functional diversity (*tap*, homolog of *Neurogenins*; Hartenstein et al., 2017). On the other hand, Da is required for ISC maintenance, as ISCs mutant for *da* differentiate (Bardin et al., 2010; Lan et al., 2018). However, which interaction partner binds to Da to maintain stemness is not known, and how different bHLH factors dimerise to allow differentiation has not been explored. Here we identify the Da homodimer as the critical bHLH complex maintaining ISC self-renewal and elucidate the role of the HLH factor Emc in titrating Da and Sc to promote absorptive differentiation. We further find that Da:Da and Da:Sc dimers functionally cooperate in the promotion of ISC, but act antagonically with respect the EE fate. Our results suggest that a network of bHLH factors acts as a three-way switch to regulate self-renewal and bipotential differentiation in the adult fly gut.

RESULTS

emc is expressed preferentially in the absorptive lineage

bHLH factors are essential for correct differentiation during adult intestinal homeostasis across metazoans (Hartenstein et al., 2017; Philpott and Winton, 2014). This suggests that *emc*, the only *Drosophila* Id factor (Campuzano, 2001; Garrell and Modolell, 1990), may have a regulatory function in the adult fly intestine, where it is strongly transcribed (Brown et al., 2014; Chintapalli et al., 2007; Dutta et al., 2015). We examined its expression pattern using the homozygous viable protein trap line $emc^{CPTI\text{-}002740}$ (Lowe et al., 2014) and detected Emc in all EBs and most ECs in the posterior midgut (Fig 1A-B). Some EEs and ISCs showed very low levels of expression (Fig 1C-D), placing *emc* expression as a marker for the absorptive lineage. This pattern of protein expression agreed with our analysis of two published scRNAseq experiments (Hung *et al*, 2020; Li *et al*, 2022). This integrated dataset, comprising ~9800 cells (Fig S1A) showed increased *emc* expression in EBs and ECs (Fig 1E), with similar pattern of enrichment respect to EEs and ISCs to bona fide markers of the absorptive lineage, like *myoIA* (Fig 1F) or *nub*, Trypsin-like serine proteases β and ι, lysosomal α-mannosidases V and VI, *klumpfuss*, and members of the *E(spl)* gene complex *mα*, *mβ* and *m3* (Fig S1B-I). Moreover, expression of *emc* along the transcriptional trajectory from ISC/EBs to ECs of the posterior midgut (pECs) (Fig 1G) increases progressively, whereas it decreases in the trajectory leading to the EE fate (Fig 1H). Taking expression pattern together with the described role of several bHLH proteins in intestinal homeostasis suggested a function of *emc* in promoting EC differentiation.

Emc is required for ISC and EB differentiation

We examined the effect of *emc* loss in the adult posterior midgut by generating *emc* mutant MARCM clones (Lee and Luo, 1999). Clones of the null allele emc^{AP6} (Ellis, 1994) Fig S2A) were depleted in differentiated cells and enriched in ISCs, as revealed by expression of the Notch ligand, Delta (Dl) (Fig 2A-B, Fig S2E and Table S1) (Ohlstein and Spradling, 2007). Clones for the strong hypomorphic allele emc^{1} and the predicted transcriptional null $emc^{LL02590}$ (Fig S2A) showed similar enrichment in Dl^{+} ISCs respect to wild-type clones (Fig 1C-E and Fig S2E). We then knocked-down *emc* in ISCs and EBs, and traced their offspring using the *escargot* flip-out (FO) approach ("esg^{TS}-*FO*"; Jiang et al., 2009) (Fig 2F). Loss of *emc* blocked the differentiation of ISCs and EBs, and sometimes the typically small nests of 1-3 esg^{+} cells expanded into larger

clusters (Fig 2H) (de Navascués et al., 2012; Ohlstein and Spradling, 2006). This was suggestive of impairment of Notch signalling. However, these larger clusters did not display the unchecked growth and invasive behaviour observed when Notch signalling is deregulated in this tissue (Micchelli and Perrimon, 2006; Patel et al., 2015; see also Fig 4A). Moreover, knockdown of *emc* in the ISC/EB compartment led to the loss of many *esg*$^{+}$ cell nests (identified by their expression of Hdc, an RNA binding protein expressed like *esg*; Resende et al., 2017, and Fig S2B) (Fig S2C, compare with Fig 2G). This was due to apoptosis, since co-expression of the caspase inhibitor p35 and *emc*-specific RNAi prevents the loss of *esg*$^{+}$ cells (Fig S2C-D). We conclude that *emc* is necessary for the differentiation and survival of ISCs and EBs.

emc maintains the commitment of EBs

The increased number of Dl^{+} cells in *emc* depletion conditions conflicted with *emc* expression being predominantly in EBs and ECs. To test whether the loss of *emc* in EBs could induce their expression of Dl, we expressed RNAi against *emc* specifically in EBs using *NRE-Gal4*, *UAS-GFP*, *tub-Gal80*^{TS} (Zeng et al., 2010). Indeed, *emc* knockdown led to EBs reacquiring Dl expression, often while displaying reduced GFP signal (Fig 3A-B). This suggested reduced expression of the EB-specific *NRE-Gal4* and that *emc*-depleted EBs were reverting their identity. To verify this, we used a flip-out lineage tracing approach to irreversibly label EBs ("*NRE*^{TS}*-FO*"; Fig 3C) and stained for the mitotic marker phospho-Histone3. In contrast with unmodified EBs, which are post-mitotic and only produce ECs (Micchelli and Perrimon, 2006; Ohlstein and Spradling, 2006; Zeng and Hou, 2015), *emc*-depleted EBs undergo cell division and generate Dl^{+}, proliferative cells (Fig 3D-E). This suggests that *emc*-deficient EBs can not only activate Dl expression but also de-differentiate into functional ISCs, capable of cell division and EE differentiation.

emc acts in parallel to Notch to direct EC differentiation

We next assessed the capacity of *emc* to impose differentiation by overexpressing it with *esg*^{TS}*-FO*. Indeed, all ISCs and EBs differentiated into ECs (Fig 2J), as previously reported (Lan et al., 2018). This establishes Emc as both necessary and sufficient to force differentiation into the absorptive fate.

This phenotype coincides with that of over-activating the Notch pathway, a known regulator of ISC fate (Bardin et al., 2010; Micchelli and Perrimon, 2006; Ohlstein and

Spradling, 2006). This suggests that *emc* might be mediating the effects of Notch signalling. To explore this possibility, we tested whether *emc* could force EC differentiation in the absence of Notch. Knockdown of *Notch* with *esg^{TS}-FO* leads to an invasive expansion of Dl^{+} ISC-like and Pros^{+} EE-like cells (Micchelli and Perrimon, 2006; Patel et al., 2015; see Fig 4A). When we co-expressed *emc* and *N^{RNAi}* together, all ISCs and EBs turned into ECs (Fig 4B), strongly suggesting that *emc* operates downstream of Notch. This phenotype, however, could also be due to the excess of Emc quickly driving cells irreversibly into differentiation before the knockdown of *Notch* resulted in effective clearance of Notch protein. To reduce Notch signalling with the kinetics of protein synthesis, rather than the additive kinetics of RNAi biogenesis, RNA interference, and Notch decay, we overexpressed Hairless (H), the co-repressor associated to the nuclear effector of N^{ICD}, Su(H) (reviewed in Bray and Furriols, 2001), with *esg^{TS}-FO*. As expected, this leads to expansion of Dl^{+} and Pros^{+} cells, with a strong reduction of EC differentiation (Fig 4C, G; see also (Bardin et al., 2010). By contrast, co-expression of *emc* and *H* led to a sharp increase in the levels of EC differentiation (Fig 4D, G). This further indicates that Emc can induce differentiation in the absence of Notch signalling.

Next, we tested whether *emc* mediates the effects of Notch signalling. Activating the pathway by knocking-down *H* leads to an increase in EC differentiation (Bardin et al., 2010; Fig 4E) and to the expression of the NRE reporter in all the progenitor, diploid cells, including those expressing Dl (Fig 4H); this is likely due to an increased baseline expression of the NRE reporter, as expected in the absence of H (Furriols and Bray, 2001). Simultaneously knocking down *H* and *emc* did enhanced the increase in EC differentiation induced by overexpression of *H*-specific RNAi. (Fig 4F-H). Taken together, our data suggest that both Emc and Notch signalling can induce EC differentiation independently of each other.

Emc antagonises self-renewal by disrupting Daughterless homodimers

The molecular function of Emc consists in titrating transcriptional activators. Archetypal targets of this inactivation are the class I and II bHLH factors such as Da or Sc, respectively, to which Emc binds directly in vitro and in vivo (Alifragis et al., 1997; Cabrera et al., 1994; Giot et al., 2003; Shokri et al., 2019; Van Doren et al., 1991). Therefore, the intestinal phenotypes we observed with the loss of *emc* may be due to increased activity of Da and its dimerization partners. Indeed, overexpression of *da*

with *esg^{TS}-FO* precisely recapitulated the effect of *emc* knockdown in *esg^{+}* cell nests. Some of these nests (labelled by *hdc* expression) were expanded (Fig 5A-B, I), and many were lost, their survival rescued by co-expression of p35 (Fig 5B and S3C). Conversely, loss of *da* leads to EC differentiation (Fig S3A-B and 5I), as previously described (Bardin et al., 2010; Lan et al., 2018) and mimics the excess of Emc (Fig 2J). This suggests that the role of Emc is largely to repress one or more Da-containing bHLH dimers, which in turn maintain stemness and survival. Epistasis experiments further confirmed this, as the simultaneous loss of Emc and Da led to the differentiation of ISCs and EBs into ECs (Fig 5E-H).

We sought to determine the identity of the Da partners involved in maintaining stemness. Most *Drosophila* class II bHLH factors (Moore et al., 2000) are expressed in the intestine at very low levels (Fig S3D); several of them are already known not to have an essential role in ISC maintenance (Bardin et al., 2010; Beebe et al., 2015; Hartenstein et al., 2017). However, Da can form homodimers (Cabrera and Alonso, 1991; Cabrera et al., 1994) with demonstrated or suspected functional relevance in several contexts, and often antagonised by Emc (Andrade-Zapata and Baonza, 2014; Bhattacharya and Baker, 2011; D'Rozario et al., 2016; Tanaka-Matakatsu et al., 2014; Tapanes-Castillo and Baylies, 2004; Troost et al., 2015). To test whether Da:Da dimers were involved in preventing differentiation, we overexpressed forced Da homodimers using the *UAS-da:da* tethered construct (Neuhold and Wold, 1993; Tanaka-Matakatsu et al., 2014) with *esg^{TS}-FO*. This blocked the differentiation of ISCs and EBs, which accumulated in larger clusters (Fig 5C-D; Table S2). Expression of Da:Da did not lead to *esg^{+}* cell death (Fig 5D, compare nest density with 5B). This shows that Da homodimers prevent differentiation; it does not rule out that other Da-containing complexes are also necessary to maintain self-renewal. To test this, we took advantage of the *UAS-da^{RNAi}* transgene *P{TRiP.JF02488}*; this targets the 5'UTR of the endogenous *da* transcript, absent in the *UAS-da:da* construct (Fig S3E). Thus, we co-expressed *da:da* and *da^{JF02488}* in ISCs and EBs to evaluate the capacity of Da:Da to prevent differentiation in the absence of endogenous, monomeric Da. The effects were identical to the overexpression of *UAS-da:da* alone (Fig S3F, compare with Fig 5D). This shows that Da can promote ISC self-renewal solely as homodimers, which are antagonised by Emc to allow EC differentiation. Furthermore, our results suggest that Da binds to an unknown partner, to form complexes that are titrated by Emc and whose excess promotes cell death in esg^{+} cells.

Basal levels of Sc in esg$^+$ cells confer ISC properties

While the maintenance of self-renewal corresponds to Da homodimers, the loss of *emc* mimics best the overexpression of monomeric Da, which increases all Da-containing complexes. As *emc* loss induces the elevation of Dl levels, another Da partner, such as Scute, could be mediating this effect. Detectable levels of Sc have recently been found in esg$^+$ cells (Chen et al., 2018; Doupé et al., 2018), and members of the *AS-C* are proposed to promote Dl expression during neurogenesis (Hinz et al., 1994; Kunisch et al., 1994). Therefore, we tested whether the increased levels of Dl observed upon *emc* loss (Fig 2H-I) could be caused by elevated activity of Sc or any of the other bHLH factors belonging to the *AS-C*. We induced *emc* knockdown in MARCM clones, which were either wild-type or homozygous for *sc^{B57}*, a deficiency that removes the entire *AS-C*, and compared their respective levels of Dl expression. Indeed, we observed an average four-fold decrease in Dl levels in *emc*-depleted clones when they were homozygous for *sc^{B57}*, indicating that basal Sc levels (and possibly other members of the *AS-C*) can boost Dl expression in ISCs (Fig 6A-C).

These observations prompted us to consider whether Sc could be contributing to additional stem cell properties in *esg$^+$* cells. Expression of Sc with *esg^{TS}-FO* clearly leads to increased differentiation into Pros$^+$ EEs (Bardin et al., 2010; Chen et al., 2018) and Fig 6D, H). To uncover a role in inducing stem properties, we overexpressed Sc with the EB-specific driver *NRE^{TS}-FO*. Sc-expressing EBs were capable of division, and their offspring included both Pros$^+$ and Dl$^+$ expressing cells, indicating the acquisition of ISC-like properties (Fig 6I, compare with 3D). These functions of Sc depend on the formation of Sc:Da heterodimers, as the overexpression of Sc in *da*-depleted cells prevents the induction of EE differentiation and of Dl expression (Fig S4E).

Sc:Da and Da:Da antagonise each other in the ISC-to-EE transition

We have shown so far that Da homodimers maintain ISC self-renewal, with EC differentiation simply triggered by the Emc-mediated titration of Da. EE differentiation, however, requires the transition from the transcriptional program of Da:Da to that of Da:Sc. Two main scenarios are possible. Sc:Da could initiate a program that silenced, or was epistatic to, that maintained by Da:Da ("succession"). Alternatively, the relative strengths of the two programs could determine the fate outcome for the cell

("antagonism"). To distinguish between these alternatives, we evaluated the ability of an excess of Da monomer to suppress EE differentiation induced by overexpression of Sc with *esg*^{TS}*-FO*. Overexpression of *sc* in ISCs and EBs leads to a quick expansion of Pros^{+} cells (Fig 6D) (Bardin et al., 2010; Chen et al., 2018). These were often mitotic (Fig S4A, D), and the number of GFP^{+} cells generated in only 3 days of lineage tracing was much higher than those generated during normal tissue replacement (Fig 6D, compare with 2A and 5A). This suggests that some of these cells have been trapped into a pre-EE state, where cells express Pros and Dl and are also capable of mitosis (Chen et al., 2018; Zeng and Hou, 2015). However, the co-expression of Da with Sc greatly reduced the number of extra Pros^{+} Dl^{-} cells and led to an increase in Dl^{+} cells (Fig 6E, H), while maintaining mitotic figures (Fig S4B, D). Under these conditions, the amount of Sc:Da is expected to be either similar or higher than with the overexpression of Sc alone (as endogenous *da* is weakly expressed (Bardin et al., 2010), likely in limiting amounts), with a simultaneous increase in Da homodimers. Thus, the reduction of EE differentiation and increase of ISC population is indicative of the antagonism, rather than the succession, of their respective transcriptional programs. If that was true, we would expect that the co-expression of tethered Da:Da and Sc would result in even less EE differentiation, as this construct cannot contribute Da monomers to bind to Sc. This was the case, resulting in an even higher amount of Dl^{+} cells at the expense of Pros^{+} cell formation (Fig 6F, H), with a concomitant increase in mitotic figures (Fig S4C, D).

To further test the capacity of Da homodimers to antagonise EE differentiation, we co-expressed *da:da* and *Notch*^{RNAi} with *esg*^{TS}*-FO*, as the reduction of Notch leads to the formation of abundant Pros^{+} EE-like cells (Bardin et al., 2010; Ohlstein and Spradling, 2006). Indeed, Da:Da opposed the formation of extra Pros^{+} cells under conditions of *Notch* loss (Fig 6G, compare with Fig 4A). This further shows that Da:Da can oppose EE differentiation, and suggests that Sc:Da dimers need to overcome the activity of Da:Da dimers to induce the EE fate.

Da:Da and Da:Sc collaborate to promote ISC identity

Our data indicate that Da:Da and Sc:Da activate distinct and antagonistic transcriptional programs to maintain self-renewal and initiate EE differentiation, respectively. However, during the development of the peripheral nervous system, both Da homodimers and Sc:Da heterodimers seem to be able to direct the same cell fate

(sensory organ precursor) (Troost et al., 2015). To determine the transcriptional signature induced by these factors, we performed RNA-seq analysis of purified ISCs and EBs overexpressing either *da*^{RNAi}, *da*, *da:da* or *sc*.

Each overexpression experiment gave a distinct signature (Fig 7A, Fig S5A and Table S3), with Sc inducing the more widespread changes in gene expression, and Da and Da:Da producing aligned effects (for instance, all genes upregulated by Da:Da were also upregulated by Da, Fig S5B). Meanwhile, the transcriptional effects of *da* knock-down had substantial overlap with that of Da (Fig 7A). Using a panel of 57 cell type-exclusive marker genes for ISCs, EBs, pECs and EEs (Table S4 and Fig S5B), it is clear that both *UAS-da*^{RNAi} and *UAS-da* induced ISC/EBs to reduce expression of key EE-specific genes like the transcription factors *pros* and *dimm*, the signalling peptides *AstA*, *AstB/Mip*, *AstC* and *Dh31* or the secretory fate-promoting factor *phyl* (Fig 7B and S5C). This is likely reflecting, on one hand, the capacity of expressing *UAS-da* to generate excess Da:Da dimers, which in turn prevent EE differentiation (Fig 5C-D and 6G), and on the other, the requirement for Da:Sc to induce EE genes (Li et al., 2017). Consequently, overexpression of Sc induces many EE-specific genes, but also the ISC-specific ones *spdo* and *Dl*. Da:Da and Da also induce some ISC-specific gene expression as shown by upregulation of *mira* and (for Da:Da only) *spdo*. An inspection of the whole panel also shows how EC- and EB-specific genes (*myo31DF*, *nub/pdm1*, α*Try*, β*Try*; *E(spl)m*β*-HLH*, *E(spl)m3-HLH*) tend to be upregulated with *da* knock-down and downregulated in the other three conditions. This is all in agreement with our phenotypic observations.

We next considered whether the genes dis-regulated in these conditions had nearby regulatory elements containing E-boxes (the canonical binding sites of bHLH factors). Indeed, E-boxes and other binding sites for class I and II bHLH factors were overrepresented in predicted regulatory elements close to genes upregulated with overexpression of Da or Da:Da, or downregulated with the knock-down of *da* (Fig 7C). Overexpression of Da and Da:Da led to the downregulation of genes with N-boxes (the binding target of *E(spl)-HLH* genes), maybe reflecting interference with Notch signalling effectors. Intriguingly, Sc overexpression resulted in the downregulation, rather than the upregulation, of genes enriched in E-boxes. The knock-down of *da* led to the upregulation of genes with a marked enrichment in GATA-boxes, likely due to the described role of *GATAe* in ISC differentiation (Okumura et al., 2016).

To characterise the molecular signature induced by Da:Da, Da, Sc and *da* knock-down more systematically, we used Gene Set Enrichment Analysis (GSEA; Subramanian et al., 2005) against previously described transcriptional profiles of midgut cell types (Dutta et al., 2015). The most obvious patterns are with *da* knock-down, which clearly induces genes expressed in the absorptive lineage while reducing the expression of ISC and EE-specific genes, and Sc, which does the opposite (Fig 7D). Sc also induces expression of ISC-specific genes, and does so better than Da:Da in our experimental conditions. This seemed at odds with our observed capacity of Da:Da to prevent differentiation, and the described role of Sc in forcing terminal EE differentiation (albeit at long time scales; Chen et al., 2018). Therefore, we looked at the expression changes in individual genes within the ISC-specific set. While both Da:Da and Sc induce regulatory genes such as *Dl* or *mira*, they have opposite effects on factors involved in DNA replication (*Orc2*, *Mcm2*, *Mcm5*, *Mcm6* and *Mcm7*) or in mitosis (*polo*, *aurB*, *BubR1* and *Ncd80*), which are repressed by Da:Da but induced by Sc. This is in agreement with the described roles of Sc as a mitotic inducer in the intestine (Chen et al., 2018) and of Da:Da as a repressor of cell cycle progression in imaginal discs (Andrade-Zapata & Baonza, 2014), but also highlight the capacity of Sc to regulate key ISC-specific genes, such as *Dl*. This feature is also present in the scRNAseq dataset profiles, where cells classified as ISC/EBs that expressed *scute* had on average much higher levels of *Delta* (Fig S5D).

We wanted to determine whether manipulating *da* and *sc* in our experimental conditions had induced changes in specific cellular processes. Using the Gene List Annotation for Drosophila (GLAD) resource (Hu et al., 2015), we found that Da:Da, Da and Sc overexpression showed enrichment in categories involved in regulatory mechanisms (receptors, kinases, GCPRs or signalling pathways, with Notch singled out) (Fig 7E). Interestingly, they also induced depletion of genes involved in mitochondrial function, oxidative phosphorylation and the ribosome, all gene sets enriched when *da* is knocked-down. Da:Da also downregulated metabolism genes (Fig 7E), while *da* knock-down increased their expression. The GLAD ‘metabolic’ gene set is large (4088 genes) and diverse, so the depletion and enrichment observed with Da:Da misexpression and *da* knock-down respectively could be caused by regulation of different subgroups of genes within the set. However, when we performed GSEA with KEGG metabolic pathways (Kanehisa et al., 2023), the two genetic conditions also show opposing effects on several pathways – biosynthesis of aminoacids and

cofactors, drug and xenobiotic metabolism, and glucolysis/gluconeogenesis, among others (Fig 7F).

These results suggest that Da and Sc induce distinct signatures which actively promote the ISC and EE identities, respectively, but also repress the active metabolism typical from EC function. Moreover, a careful balance of Da and Sc is required to maintain the transcriptional identity of the ISC and its proliferative capacity.

Da:Da is antagonised by Notch and maintains self-renewal in parallel to Esg

Esg, a Zinc finger factor of the Snail family, is required to prevent ISC/EB differentiation, mostly through transcriptionally repressing differentiation genes (Korzelius et al., 2014; Loza-Coll et al., 2014). Da:Da seems to have a similar effect, through the activation of ISC-identity genes. Therefore, we sought to test whether they may be acting in combination to maintain stemness. Knockdown of *esg* in ISCs and EBs leads to differentiation, with an overrepresentation of EEs (Fig S5C and 8F), as expected from the described antagonism between Esg and Sc at the *pros* locus (Li et al., 2017). Simultaneous overexpression of *UAS-esg*^{RNAi} and *UAS-da:da* prevented most differentiation, and specifically reduced EE cell production (Fig 8D and 8F), indicating that Da:Da can maintain ISCs and oppose EE differentiation independently of *esg*. In turn, overexpression of *esg* blocked differentiation irrespective of the presence of *da* (Fig 8C, F and Fig S5B), indicating that Da:Da and Esg largely act independently in ISC maintenance. Furthermore, expression of *da* and *esg* are independent of each other (Fig S6D, E).

We next sought to determine whether Da:Da operates as part of a pathway with the known regulators of ISC fate. Our previous epistatic analysis with *emc* and Notch signalling suggested that Da:Da is required independently of the Notch pathway. Indeed, when we knocked-down both *N* and *da* using *esg*^{TS}*-FO*, all ISCs/EBs became ECs (Fig S5A). Similar results were obtained by simultaneously knocking-down *da* and inhibiting Notch signalling by overexpressing *H* (Fig 8A, compare with Fig 4C, and Fig 8E). However, expressing *UAS-da:da* prevented the differentiation induced by increasing Notch signalling by knockdown of *H* (Fig 8B, compare with Fig 4E, and Fig 8E), suggesting that Da:Da is antagonised by Notch signalling. Together, our data indicate that Da homodimers operate downstream of Notch and in parallel to Esg to maintain ISC self-renewal.

DISCUSSION

Multipotent adult stem cells must be able to change their transcriptional state from self-renewal to several alternative differentiation programs. bHLH transcription factors seem well placed to coordinate these changes molecularly; they have been described to operate synergistically, antagonistically or combinatorially (Amoutzias et al., 2008; Bhattacharya and Baker, 2011; Sallee et al., 2017; Troost et al., 2015); their capacity for dimerization allows reduction in molecular noise (Bundschuh et al., 2003). Here we show that a bHLH factor network controls three alternate fates in the adult *Drosophila* intestine. Class I homodimers (Da:Da) promote the progenitor state of ISCs/EBs. Progenitor cells will acquire the absorptive fate by titration of Da by the HLH factor Emc, or the secretory fate by expression of the class II bHLH Sc, which heterodimerises with Da to initiate EE differentiation. Moreover, Emc is required in EBs to maintain their committed state, while low levels of Sc seem to contribute to boost essential parts of the ISC transcriptional program. Our results argue for a central and conserved role of bHLH factors in the acquisition and maintenance of multiple fates in the fly intestine.

Three fates regulated by a dimerization network

Our data show that Da homodimers maintain ISCs and EBs undifferentiated, but this situation can be easily swayed towards either EC or EE differentiation by Da:Da swapping partners with Emc or Sc (Fig 9). Similar networks involving class I, II and V bHLH factors have been found to regulate the development of the *Drosophila* retina (Bhattacharya and Baker, 2011) and the peripheral nervous system (PNS) (Cubas et al., 1991; Troost et al., 2015; Van Doren et al., 1991). However, in these cases only one new fate is accessed (neural), and Da:Da dimers seem to have a relatively ‘accessory’ role. In the retina, Da:Da dampens the titration by Emc of the bHLH proneural factor Ato (Bhattacharya and Baker, 2011; Li and Baker, 2018). In the PNS, Da:Da cooperate with Da:Ac/Sc heterodimers to overcome Emc inhibition and induce neural fate, which both Da:Ac/Sc and Da:Da can induce (Troost et al., 2015). By contrast, in the adult midgut Da:Da and Da:Sc support distinct fates (progenitor and secretory, respectively) while Emc titrates both dimers to allow EC differentiation, enabling the network to regulate a three-way fate decision. This is possible because Da:Da and Da:Sc induce specific but partially overlapping transcriptional profiles (Fig 7) which yet seem to be able to act antagonistically (Fig 6). By contrast, Emc simply

prevents the DNA binding of either Da or Sc (Cabrera et al., 1994; Van Doren et al., 1991), which suggests that EC is the default fate when Da:Da dimers do not operate (Fig 5).

Interestingly, bHLH transcription factors of classes I-II-V (or E/A/Id) are also active in the mammalian intestine. Class II bHLH factors promote EE differentiation (reviewed in Philpott and Winton, 2014) and ISC maintenance (van der Flier et al., 2009). However, class I factors E2a and Heb are specifically expressed in crypt cells in the mouse intestine (van der Flier et al., 2009). Our work suggests that they may function as class I dimers as well as dimerization partners for class II factors. Meanwhile, Emc homologs Id2 and Id3 are expressed specifically in the intestinal differentiated cells (Wice, 1998), suggesting similar functions. By contrast, Id1 is required in ISCs for regeneration (Zhang et al., 2014) and can act as an oncogene (Wice, 1998), which is the opposite behaviour to that of Emc.

A balance of bHLH factors regulates intestinal stem cell fate

Our observations indicate that additional work on the mechanisms that regulate the transcription of both *emc* and *sc* would be very informative. The *cis*-regulatory region of *emc* seems to be complex and spreads across a large genomic region (Li and Baker, 2019), so it is likely that is subject to multiple regulatory inputs. An obvious candidate for regulation of *emc* transcription is Notch signalling, as shown in other contexts (Adam and Montell, 2004; Baonza and Freeman, 2001; Baonza et al., 2000; Bhattacharya and Baker, 2009; Spratford and Kumar, 2015) and considering the importance of Notch for EB establishment and EC differentiation (Bardin et al., 2010; Micchelli and Perrimon, 2006; Ohlstein and Spradling, 2006). However, our epistasis analyses suggest that Emc and Notch signalling can induce EC differentiation independently of each other, observation that makes Notch an unlikely regulator of *emc* in the midgut. A major regulator of Sc is the transcriptional repressor Ttk69 (Wang et al., 2015), whose stability is in turn regulated by the E3 ubiquitin ligase Sina and its adaptor Phyl (Yin and Xi, 2018). Therefore, Phyl expression could be the trigger to allow the accumulation of Sc and initiate EE differentiation, but how *phyl* is regulated is not known yet. On the other hand, Chen et al. (2018) observed that Sc activates itself as well as *E(spl)m8*, a member of the *Enhancer of split Complex* (which comprises several transcription factors, many of which are often both targets and effectors of the Notch pathway; reviewed in Delidakis et al., 2014). As E(spl)m8 is able to repress Sc,

the authors proposed that a cell-autonomous oscillator involving Sc and *E(spl)-C* members would create pulses of Sc expression which would set the pace of EE production (Chen et al., 2018). However, while we too observe induction of *E(spl)-C* genes upon Sc forced expression, so we do as a result of Da:Da and Da overexpression (Fig S5E). This suggests that the regulation of *E(spl)-C* (and therefore of *sc*) may be more complicated. Additional work will be necessary to determine the sequence and timing of regulatory inputs that initiate fate-inducing levels of Sc or Emc.

Moreover, the Emc/Da/Sc network clearly does not work as a simple boolean switch. The three genes are expressed, and seem to be functionally relevant, in both ISCs and EBs (Bardin et al., 2010; Chen et al., 2018; Doupé et al., 2018; this work). This argues against a mechanism that simply initiates Emc or Sc expression to induce the EC or EE fates. Moreover, the functions of Emc and Sc seem to be more complex than acting as mere fate switches. Adequate levels of Emc are required in the EB, as its depletion induces de-differentiation into ISCs (Fig 3), but its overexpression accelerates EB terminal differentiation into ECs (Fig 2). Intriguingly, it has recently been proposed that the Emc homolog Id1 protects pluripotency specifically during the peri-implantation transition (Malaguti et al., 2019), which suggests that Emc could be part of a timing mechanism in the transition between ISC and EC. As for Sc, our data suggests that low levels of expression in ISCs (Chen et al., 2018; Doupé et al., 2018) induce the expression of ISC-specific genes (Fig 7) without eliciting a conversion into pre-EE cell. Higher levels would be required to initiate EE differentiation. Therefore, the amounts of Da, Sc and Emc must be maintained in balance. In the case of Emc, this is likely to be achieved through post-transcriptional mechanisms. During neurogenesis, Emc is stabilised when bound to Da, and degrades when Da frees itself to bind a proneural factor (Bhattacharya and Baker, 2011; Li and Baker, 2018). This may be the reason why Emc shows higher levels of accumulation in EBs and ECs (Fig 1), where Da is expressed but other bHLH factors are at minimal levels (Fig S3).

Da and Sc cooperate to endow cells with ISC properties

Our data shows a critical requirement for Da:Da in ISC maintenance, but also that Sc can contribute to enhance ISC features (proliferation, expression of Dl). This is likely an active promotion of ISC fate, rather than just prevention of differentiation, as our transcriptomic analysis shows an enrichment of the ISC transcriptional signature when overexpressing Da:Da and Da, but this is only significant with Sc. This active role is

further supported by the capacity of Sc to induce EBs to de-differentiate into ISCs and re- or transdifferentiate into EEs (Fig 6). We also find evidence that mitochondrial function is reduced by expression of Sc, Da and Da:Da (Fig 7), which fits well the observation that ISCs display an 'empty mitochondria' phenotype under the electron microscope (Hung et al., 2020). By contast, knocking-down *da* has the opposite effect, as one would expect when the EC fate is promoted (Fig 7 and S3). However, it has recently been shown that the EGFR pathway stimulates ISC activation of proliferation by inducing mitochondrial biogenesis, β-oxidation and oxidative phosphorylation (Zhang et al., 2022). Therefore, it is not clear whether this part of the transcriptional signature of Da:Da, Da and Sc is due to their simply promoting the ISC fate as opposed to the EC fate, or a more complex scenario where, at least Da:Da, also promotes the quiescent state of the ISC. This would be compatible with our observation that overexpression of Da:Da and Da seems to slow proliferation (Fig 5) and the effect of *da* as repressor of cell cycle progression in imaginal tissues (Andrade-Zapata & Baonza, 2014). Together, this suggests that Da:Da dimers are essential to maintain the ISC undifferentiated, but low levels of Sc:Da dimers also contribute non-critically to the maintenance of the ISC identity and function; we surmise that Emc expression in EBs might be required to dampen this Sc activity.

Regulation of the committed state of the enteroblast

Under normal conditions, EBs will only give rise to mature ECs without further division (Wang et al., 2015; Yin and Xi, 2018; Zeng and Hou, 2015). Depletion of Emc specifically in EBs coupled to lineage tracing results in cells proliferating and expressing Dl (Fig 3), which likely represents a de-differentiation process. This may partially reflect the need to control Sc activity in the EB, as the expression of Sc can induce the same effects (Fig 6). Intriguingly, EB-specific depletion of Ttk69 or Klu or overexpression of Phyl seem to induce trans-differentiation, as these conditions produce Pros^{+} cells (Wang et al., 2015; Yin and Xi, 2018). This difference could be because Emc is also influencing the activity of multiple Da-containing complexes in the EB, or because of the activity of additional targets of Klu and the Phyl/Ttk69 pathway than Sc. The reacquisition of proliferation and Delta expression is also controlled by other factors, such as Sox21a (Zhai et al., 2015), or the global co-repressor Gro, acting downstream of Notch (Guo et al., 2019) (though surprisingly Notch itself does not seem to be required to maintain the EB; Siudeja et al., 2015). Together with our observations, this indicates a previously unappreciated plasticity of the EB, which is

consistent with its similarity to the ISC in transcriptome (Hung et al., 2020) and chromatin accessibility landscape (Aughey et al., 2018). Importantly, this resembles the behaviour of EC precursors in the mammalian intestine, which can dedifferentiate and repopulate the intestinal crypt during regeneration (Tetteh et al., 2016).

MATERIALS AND METHODS

Drosophila culture and genetics

Flies were raised and maintained on standard cornmeal/yeast medium at 25ºC unless specifically indicated. Experiments were conducted in mated females, collected daily and maintained in fresh vials with added yeast until 4-7 days old. For experiments using Gal80^{TS}, flies were reared and aged to gut maturity (~7 days) at 18ºC, then transferred to 29 ºC. For induction of MARCM and flip-out clones, 4-7 days old flies were treated at 37ºC for 60 or 15 min, respectively. Flies were aged for 7 days after induction treatment before dissection, unless otherwise indicated. See Appendix for fly strains.

Immunohistofluorescence, confocal microscopy and image analysis

Stainings were performed as in (Bardin et al., 2010). See Appendix for details on primary and secondary antibodies. DNA dye was Hoechst (Sigma-Aldrich B2261 used 1:5000 from a stock solution at 10 mg/ml). Mounting medium was 4% N-propyl-galate, 80% glycerol. Confocal stacks were obtained in a Zeiss LSM 710 with an EC Plan-Neofluar 40X oil immersion objective (numerical aperture 1.3). All stack positions were acquired in the posterior midgut. Typically, three positions along the anterior-posterior axis of the posterior midgut were acquired for each organ. In MARCM clone experiments, stacks were acquired from all clones in the posterior midgut. Figures were assembled using Adobe Photoshop and Illustrator CS6 and Affinity Designer 2.

Cell counts and statistics

For evaluating the proportion of cell types in GFP^{+} clones or labelled tissue, confocal stacks were maximum-intensity projected using FIJI/ImageJ (Schindelin et al., 2012), cells of the relevant types were counted with the Cell Counter plugin. In the experiments overexpressing *UAS-sc* and *UAS-da:da*, the associated increase in proliferation generated large, highly densely populated cell clusters which could not be counted with single-cell precision. Therefore, for this genotype we estimated the proportion of each cell population in each field of view separately. See the Appendix for details on cell type identification.

Statistical tests were performed in Prism 8 or in R (R Core Team, 2023) using RStudio 2023.3.0.386 (Posit Software PBC, 2022). Statistical significance of changes in proportions of cell types was assessed by binomial logistic regression in base R. We used Firth's bias reduced logistic regression (package *logistf*; Heinze et al., 2023) in experiments with zero observations in one or more of the genotypes considered, to avoid the nonsensical results arising from the 'complete separation' of data (Albert and Anderson, 1984).

RNA-seq

Flies bearing either *UAS-da*, *UAS-da*^{RNAi}, *UAS-da:da* or *UAS-sc* as well as *esg-Gal4*, *UAS-GFP* and *tub-Gal80*^{TS} were reared at 18ºC until 4-7 days old, transferred to 29ºC for 24h and their posterior midguts dissected, then processed as described in (Dutta et al., 2013). Libraries from three biological replicates per condition (except one condition, with two) were prepared in two batches and ~37 million reads (either 50 or 300bp long, for each batch respectively) were generated per library using Illumina technology. Sequencing files are available at the Gene Expression Omnibus (accession number GSE234019).

RNA-seq and scRNA-seq analysis

Fastq files, pre-trimmed by the sequencing provider (Genewiz), were mapped to the release 6.28 of the *Drosophila melanogaster* genome using STAR and bamtools (Barnett et al., 2011; Dobin et al., 2013) and assigned to genes with featureCounts (Liao et al., 2014). Differential gene expression, gene set and DNA motif enrichment analyses were performed with the key R packages DEseq2 (Love et al., 2014), limma (Ritchie et al.), fgsea (Korotkevich et al., 2021) and RcisTarget (Aibar et al., 2017). scRNAseq data was obtained from GEO and https://flycellatlas.org/ and analysed in R using Seurat (Hao et al., 2021) and slingshot (Street et al., 2018). The analysis scripts are available at https://github.com/jdenavascues/Puigetal2023_bioinformatics_scripts.

ACKNOWLEDGEMENTS

We thank Nicholas Baker, Allison Bardin, Antonio Baonza, Sonsoles Campuzano, Sangbin Park, Mike Taylor, Shinya Yamamoto, Alfonso Martínez-Arias, the Bloomington *Drosophila* Stock Center, the Vienna *Drosophila* Resource Center, the *Drosophila* Genetics Resource Center (Kyoto) and the National Institute of Genetics (Japan) for providing fly stocks, and the Developmental Studies Hybridoma Bank (University of Iowa) for supplying antibodies. We also thank Juan Modolell, Sonsoles Campuzano, Catherine Hogan, Fernando dos Anjos-Afonso, Florian Siebzehnrubl and Terrence Trinca for critical comments on the manuscript, and Sonia López de Quinto, Helen White-Cooper, Mike Taylor and Wynand van der Goers van Naters for useful discussions during the development of the project. We would like to acknowledge the the data analysis team of the College of Biomedical and Life Sciences and the Genomics for bioinformatics training and the FACS and Imaging Core Facility at the Max Planck Institute for Biology of Ageing for technical assistance.

This work was supported by funding from Cardiff University to JdN and AP, from the University of Essex and NC3Rs SKT grant NC/W001047/1 to PVW and JdN, a DFG-Grant KO5594/1-1 and an EMBO Long-Term Fellowship to JK, a FAPESP fellowship #2021/00393-9 to VDN, a FAPESP São Paulo Excellence Chair #2019/16113-5 to PVW, and an ERC Advanced Grant no. 268515 to BAE.

JdN dedicates this work to the memory of Juan Modolell Mainou.

AUTHOR CONTRIBUTIONS

Conception of the study and supervision: JdN. Data acquisition: AP, SD, JK and JdN. Analysis and interpretation: AP, VDN, JdN, JK, SN. Training and supervision: JdN, JK, PVW, BAE. Writing the manuscript: JdN, AP. Revising and approval of the manuscript: all authors.

CONFLICT OF INTEREST

The authors declare no competing interest.

FIGURE LEGENDS

Figure 1. *emc* is expressed preferentially in the absorptive lineage.

A-D. Projected confocal stacks showing endogenous Emc:GFP expression. Emc is expressed in EBs (marked with *NRE-lacZ*, arrowheads in A, A') and most ECs (positive for *myoIA-lacZ*, solid arrowheads in B, B') but not all (empty arrowhead in B, B'). *emc* expression can also be detected in some ISCs (Dl^+) and EEs ($Pros^+$) (solid arrowheads and asterisks, respectively, in C-D') but not all (empty arrowheads and asterisks in C-D'). **E, F.** Expression of *emc* (E) and the EC marker *myoIA* (F) by scRNAseq in EEs, ISCs, EBs and ECs from the posterior midgut (pECs). Both genes are enriched in EBs and ECs, with rare events of detection in EEs and ISCs. **G.** Pseudotime trajectories from ISC/EB cells into mature EEs or ECs derived from scRNAseq data. **H.** *emc* expression in single cells declines as these get closer to the EE fate and increases as they progress towards EC differentiation.

Figure 2. *emc* is necessary for differentiation and sufficient for selection of the EC fate.

A-E. Differentiation is impaired in *emc* mutant clones. Control clones display more polyploid Dl^-/$Pros^-$ ECs (asterisks in E, H) than clones homozygous of for the mutant alleles emc^{AP6} (F), emc^{1} (G) and $emc^{LL02590}$ (I), which are enriched in Dl^+ ISCs (solid arrowheads) and diploid Dl^-/$Pros^-$ EBs (empty arrowheads) with respect to controls. **F.** The esg^{TS}-*FO* method allows the temporal control of expression in progenitor cells (esg^+ ISCs and EBs) and their lineage tracing to access their fate potential. *UAS-GFP* and *UAS-FLP* are expressed specifically in ISCs and EBs after induction by temperature switch. This leads to the progenitor-restricted expression of the (potentially) ubiquitous *Actin5C-Gal4* by excision of a transcriptional terminator. *Actin5C-Gal4* maintains *UAS-GFP* expression in the ISC/EB lineage when differentiation occurs and *esg* expression stops. **G-J.** Emc is required for differentiation and sufficient for induce the EC fate. ISCs and EBs in control guts produce both EC (Dl^-/$Pros^-$) and EE (Dl^-/$Pros^+$) GFP^+ cells (B, note the difference in intensity levels between Dl and Pros, respectively solid and empty arrowheads), while differentiation is reduced when ISCs/EBs express *emc*-specific RNAi , shown by increased number of Dl^+/GFP^+ cells (C, D). Reduction of *emc* also induces increased Dl expression: compare Dl-expressing with Pros-expressing cells (respectively solid and empty arrowheads in C, D, compare with those similarly marked in B). Forced expression of *emc* in ISCs/EBs forces differentiation into ECs (E). Scale bar: 20μm in all panels.

Figure 3. *emc* is required for maintenance of the EB committed state.

A-B. The *NRE-Gal4* driver is expressed in EBs at different levels in normal intestines (solid and empty arrowheads in A for higher and lower-expressing EBs), but never co-express Dl. When *emc* is depleted with *NRE-Gal4*, many EBs co-express Dl (empty arrowheads, B) but not all (solid arrowheads, B). Also, EEs develop close to *emc*-depleted EBs (arrow, B). **C.** Use of NRE^{TS}*-FO* allows the simultaneous manipulation and tracing of the EB lineage. *NRE-Gal4, UAS-GFP* is expressed specifically in EBs after induction and activates specifically in EBs the (potentially) ubiquitous *Actin5C-Gal4*, which in turn will maintain *UAS-GFP* expression if differentiation or decommitment occur. **D-E**. Loss of *emc* in EBs leads to their decommitment. Activation of NRE^{TS}*-FO* in control guts leads exclusively to the labelling of EBs and ECs, with no co-expression of GFP and Dl or Pros (solid and empty arrowheads, respectively, in D; some ISCs and EEs in the micrograph seem inside the GFP territory due to the Z-projection of the confocal stack). By contrast, in NRE^{TS}*-FO* guts where *emc* is depleted, EB-derived cells express Dl and are capable of mitosis, as judged by phospho-Histone 3 staining (arrowhead in E). Scale bar: 20μm in all panels.

Figure 4. *emc* acts in parallel to Notch signalling.

A-D. *emc* can direct EC differentiation in the absence of Notch signalling. EC differentiation is abolished or strongly impaired with knockdown of *N* (A) or overexpression of H (C), respectively, with accumulation of Dl^+ or $Pros^+$ cells (solid and empty arrowheads, respectively, in A, C). Simultaneous overexpression of *emc* induces EC differentiation in both cases (arrowheads in B, D), but to a lesser an extent in a background of *H* overexpression (asterisks in D, compare with B). **E-F.** Notch signalling can direct EC differentiation in the absence of *emc*. Loss of EC differentiation caused by depletion of *emc* (see panel H and Fig. 2H-I and S2F) is rescued by simultaneous knockdown of *H* (F). Depletion of *H* can induce EC differentiation on its own (E). **G-H.** Stacked bars plots summarising the cell type composition of esg^{TS}*-FO* tissue. Numbers (total GFP^+ cells / fields of view) are 1810/21 and 1292/27 for *UAS-H* and *UAS-H+UAS-emc*, respectively (G) and 984/15 and 126/16 for $UAS\text{-}H^{RNAi}{}_{HMS}$ and $UAS\text{-}H^{RNAi}{}_{HMS}$+$UAS\text{-}emc^{RNAi}{}_{NIG}$, respectively (H). p-values are <0.05, <0.01 and <0.001 for one, two or three asterisks (see Tables S1 and S2). Scale bar: 20μm in all panels.

Figure 5. Da homodimers prevent differentiation and are antagonised by Emc.

A-D. Da homodimers maintain ISCs/EBs undifferentiated. Solid arrowheads: ISCs/EBs; empty arrowheads: ECs/EEs. Overexpression of Da with *esg^{TS}-FO* prevents formation of ECs, though it allows EE differentiation (B, compare with A). Expression of a forced Da:Da dimer in esg^{+} cells prevents all differentiation, as evaluated with Hdc/NRE-lacZ (C) or Dl (D). **E-H.** *da* is epistatic over *emc*. Solid arrowheads: ECs; empty arrowheads: Dl^{+} ISCs. MARCM clones mutant for *da^{10}* differentiate as ECs (E). This phenotype is recapitulated when these clones simultaneously express *emc*-specific RNAi (F). Likewise, impaired differentiation in *emc^{LL02590}* clones (G) is rescued by simultaneous knockdown of *da*, which leads to EC formation (H). **I.** Stacked bars plots summarising the cell type composition of *esg^{TS}-FO* tissue. Numbers (total GFP^{+} cells / fields of view) are 1135/15, 549/14, 693/13, 803/15 and 1086/19 for control, *UAS-da*, *UAS-da:da*, *UAS-da$^{RNAi}{}_{JF}$* and *UAS-da$^{RNAi}{}_{HMS}$*, respectively. P-values are <0.05, <0.01 and <0.001 for one, two or three asterisks (see Tables S1 and S2). Scale bar: 20μm in all panels.

Figure 6. Sc:Da and Da:Da antagonise each other in EE formation but collaborate in ISC maintenance.

A-C. Emc antagonises Sc in inducing Dl expression. Loss of the *AS-C* using the *sc^{B57}* deficiency reduces the increase in Dl expression observed with knockdown of *emc* (B, compare with A). Dl levels are quantified by comparison with those of Pros (see Appendix). Numbers (Dl^{+} cells/Pros^{+} cells) are 210/146 and 718/210 for *UAS-emc$^{RNAi}{}_{NIG}$* and *UAS-emc$^{RNAi}{}_{NIG}$, Df(1)sc^{B57}*, respectively. P-values are <0.05, <0.01 and <0.001 for one, two or three asterisks (Mann-Whitney test). **D-F.** Antagonism/collaboration between Sc and Da. Solid arrowheads: Dl^{+}/Pros^{+} pre-EEs; empty arrowheads: Dl^{+} ISCs; asterisks: Pros^{+} EEs. Overexpression of Sc leads to a dramatic increase of Dl^{+}/Pros^{+} pre-EEs and Pros^{+} EEs, but maintains a population of Dl^{+} ISCs (D). The population of ISCs progressively increases by co-expression of monomeric Da (E) and dimeric Da:Da (F), at the cost of EE differentiation (E, F) and pre-EE formation (F). Quantified in H. **G.** Da:Da antagonises EE formation. While reduction of Notch signalling leads to an excess of EEs, simultaneous overexpression of Da:Da and *N*-specific RNAi results in very few Pros^{+} cells (solid arrowheads, compare with Fig 4A). **H.** Stacked bars plots summarising the cell type composition of *esg^{TS}-FO* tissue. Numbers (total GFP^{+} cells / fields of view) are 1241/6, 2131/10 and not determined/11 for *UAS-sc, UAS-sc, UAS-da*, and *UAS-sc, UAS-da:da*, respectively. p-values are <0.05, <0.01 and <0.001 for one, two or three asterisks (see Tables S1 and S2). **I.** Sc can induce decommitment of EBs. Expression of Sc with *NRE^{TS}-FO*

leads to reacquisition of ISC properties (Dl expression, mitosis assessed with phospho-Histone 3, empty arrowhead) and pre-EE formation (Pros expression and mitosis, solid arrowhead). Scale bar: 20μm in all panels.

Figure 7. Da and Sc cooperate to induce expression of the ISC signature.
A. Euler diagrams showing the intersections of sets of genes either upregulated (left) or downregulated (right) by *da*^{RNAi}, Da:Da, Da and Sc, with fold changes larger than $\log_2(1.5)$. The area of the sets denotes (approximately) the size of the intersection. **B.** MA plots for the overexpression of *da*^{RNAi}, Da:Da, Da and Sc, respectively. Some genes of interest are identified because of their role or expression pattern. Colours match the cell types where genes are expressed. **C**. Beeswarm plots of the normalised enrichment scores of DNA motifs found in upregulated (left) or downregulated (right) genes in the four conditions analysed, using RcisTarget. Dots represent individual motifs, binned for simplicity of representation, and coloured by the type of transcription factor they are bound by. *da*, *sc*, and *E(spl)-C* gene members are identified. **D-F**. Heatmaps of gene set enrichment scores for sets of genes that are: expressed in specific cell types (Dutta *et al*, 2015) (D), annotated because of their broad molecular function in *Drosophila* (Hu *et al*, 2015) (E) or belonging to annotated KEGG pathways (F). Asterisks indicate the enrichment scores with significant adjusted p-values, with colour indicating the level of significance. The colourmap of NES scores corresponds to different values in each heatmap.

Figure 8. *da* is epistatic over Notch signalling but acts in parallel with *esg*.
Solid arrowheads: ISCs (Hdc$^+$/NRE$^-$ or Dl$^+$/NRE$^-$); empty arrowheads: EEs (Pros$^+$). **A-B.** *da* is epistatic over Notch signalling. Loss of *da* leads to EC and EE differentiation when Notch signalling is impaired by overexpression of *H* (A), while overexpression of *da:da* maintains cell undifferentiated even when Notch signalling is exacerbated by knockdown of *H* (B). **C-D.** *da* and *esg* function in parallel to prevent differentiation. Overexpression of *esg* can prevent differentiation and compensate for the loss of *da* (C), while overexpression of *da:da* can similarly compensate for the loss of *esg*, including prevention of EE formation (D). **E-F.** Stacked bars plots summarising the cell type composition of *esg*^{TS}*-FO* tissue. Numbers (total GFP$^+$ cells / fields of view) are 1510/21, 1810/21, 613/8 and 984/15 for *UAS-H*, *UAS-H+UAS-da*$^{RNAi}_{JF}$, *UAS-H*$^{RNAi}_{HMS}$ and *UAS-H*$^{RNAi}_{HMS}$*+UAS-da:da*, respectively (E) and 850/9, 1236/12, 628/18 and 939/19 for for *UAS-esg UAS-esg+UAS-da*$^{RNAi}_{JF}$, *UAS-esg*$^{RNAi}_{HMS}$ and

UAS-esg$^{RNAi}{}_{HMS}$*+UAS-da:da*, respectively (F). p-values are <0.05, <0.01 and <0.001 for one, two or three asterisks (see Tables S1 and S2 for further details). Scale bar: 20μm in all panels.

Fig 9. Sc, Da and Emc regulate self-renewal and bipotential differentiation in the intestine.

Orange arrows: Da homodimers maintain ISCs and EBs undifferentiated, antagonising EC differentiation and pre-EE formation by inducing expression of ISC-specific genes. Pink arrows: An increase of Sc expression induces formation of Da:Sc heterodimers, which steer the transcriptional program towards pre-EE formation; Sc can also induce expression of the ISC-specific signature, potentially leading to EB decommitment if its activity is not prevented by titration by Emc. Green arrows: Emc increased expression in EBs maintains commitment for EC differentiation and can direct terminal EC differentiation by interruption of Da:Da and Da:Sc transcriptional activities. This occurs in parallel to Esg prevention of differentiation and Notch signalling being upstream and opposing Da:Da to induce EC differentiation.

REFERENCES

Adam, J. C. and Montell, D. J. (2004). A role for extra macrochaetae downstream of Notch in follicle cell differentiation. **131**, 5971–5980.

Aibar, S., González-Blas, C. B., Moerman, T., Huynh-Thu, V. A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* **14**, 1083–1086.

Albert, A. and Anderson, J. A. (1984). On the existence of maximum likelihood estimates in logistic regression models. *Biometrika* **71**, 1–10.

Alifragis, P., Poortinga, G., Parkhurst, S. M. and Delidakis, C. (1997). A network of interacting transcriptional regulators involved in Drosophila neural fate specification revealed by the yeast two-hybrid system. *Proc Natl Acad Sci USA* **94**, 13099–13104.

Amcheslavsky, A., Song, W., Li, Q., Nie, Y., Bragatto, I., Ferrandon, D., Perrimon, N. and Ip, Y. T. (2014). Enteroendocrine Cells Support Intestinal Stem-Cell-Mediated Homeostasis in Drosophila. *CellReports* 1–19.

Amoutzias, G. D., Robertson, D. L., Van de Peer, Y. and Oliver, S. G. (2008). Choose your partners: dimerization in eukaryotic transcription factors. *Trends Biochem. Sci.* **33**, 220–229.

Andrade-Zapata, I. and Baonza, A. (2014). The bHLH Factors Extramacrochaetae and Daughterless Control Cell Cycle in Drosophila Imaginal Discs through the Transcriptional Regulation of the cdc25 Phosphatase string. *PLoS Genet* **10**, e1004233–16.

Aughey, G. N., Estacio Gomez, A., Thomson, J., Yin, H. and Southall, T. D. (2018). CATaDa reveals global remodelling of chromatin accessibility during stem cell differentiation in vivo. *eLife* **7**, 6061.

Baonza, A. and Freeman, M. (2001). Notch signalling and the initiation of neural development in the Drosophila eye. **128**, 3889–3898.

Baonza, A., de Celis, J. F. and Garcia-Bellido, A. (2000). Relationships between extramacrochaetae and Notch signalling in Drosophila wing development. **127**, 2383–2393.

Bardin, A. J., Perdigoto, C. N., Southall, T. D., Brand, A. H. and Schweisguth, F. (2010). Transcriptional control of stem cell maintenance in the Drosophila intestine. **137**, 705–714.

Barnett, D. W., Garrison, E. K., Quinlan, A. R., Strömberg, M. P. and Marth, G. T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics* **27**, 1691–1692.

Beebe, K., Park, D., Taghert, P. H. and Micchelli, C. A. (2015). The Drosophila Prosecretory Transcription Factor dimmed Is Dynamically Regulated in Adult

Enteroendocrine Cells and Protects Against Gram-Negative Infection. *G3 (Bethesda)* **5**, 1517–1524.

Bertrand, N., Castro, D. S. and Guillemot, F. (2002). Proneural genes and the specification of neural cell types. *Nat Rev Neurosci* **3**, 517–530.

Bhattacharya, A. and Baker, N. E. (2009). Developmental Biology. *Developmental Biology* **327**, 288–300.

Bhattacharya, A. and Baker, N. E. (2011). A Network of Broadly Expressed HLH Genes Regulates Tissue-Specific Cell Fates. *Cell* **147**, 881–892.

Bray, S. and Furriols, M. (2001). Notch pathway: Making sense of suppressor of hairless. *Curr Biol* **11**, R217–R221.

Brown, J. B., Boley, N., Eisman, R., May, G. E., Stoiber, M. H., Duff, M. O., Ben W Booth, Wen, J., Park, S., Suzuki, A. M., et al. (2014). Diversity and dynamics of the Drosophila transcriptome. *Nature* **512**, 393–399.

Bundschuh, R., Hayot, F. and Jayaprakash, C. (2003). The role of dimerization in noise reduction of simple genetic networks. *J Theor Biol* **220**, 261–269.

Cabrera, C. V. and Alonso, M. C. (1991). Transcriptional activation by heterodimers of the achaete-scute and daughterless gene products of Drosophila. *EMBO J* **10**, 2965–2973.

Cabrera, C. V., Alonso, M. C. and Huikeshoven, H. (1994). Regulation of scute function by extramacrochaete in vitro and in vivo. **120**, 3595–3603.

Campuzano, S. (2001). Emc, a negative HLH regulator with multiple functions in Drosophila development. *Oncogene* **20**, 8299–8307.

Chen, J., Xu, N., Wang, C., Huang, P., Huang, H., Jin, Z., Yu, Z., Cai, T., Jiao, R. and Xi, R. (2018). Transient Scute activation via a self-stimulatory loop directs enteroendocrine cell pair specification from self-renewing intestinal stem cells. *Nature Cell Biology* **20**, 152–161.

Chintapalli, V. R., Wang, J. and Dow, J. A. T. (2007). Using FlyAtlas to identify better Drosophila melanogaster models of human disease. *Nat Genet* **39**, 715–720.

Cinquin, O. and Demongeot, J. (2002). Positive and negative feedback: Striking a balance between necessary antagonists. *J Theor Biol* **216**, 229–241.

Cinquin, O. and Demongeot, J. (2005). High-dimensional switches and the modelling of cellular differentiation. *J Theor Biol* **233**, 391–411.

Crosnier, C., Stamataki, D. and Lewis, J. (2006). Organizing cell renewal in the intestine: stem cells, signals and combinatorial control. *Nat. Rev. Genet.* **7**, 349–359.

Cubas, P., de Celis, J. F., Campuzano, S. and Modolell, J. (1991). Proneural clusters of achaete-scute expression and the generation of sensory organs in the Drosophila imaginal wing disc. *Genes & Development* **5**, 996–1008.

de Navascués, J., Perdigoto, C. N., Bian, Y., Schneider, M. H., Bardin, A. J., Martinez-Arias, A. and Simons, B. D. (2012). Drosophila midgut homeostasis involves neutral competition between symmetrically dividing intestinal stem cells. *EMBO J* **31**, 2473–2485.

Delidakis, C., Monastirioti, M. and Magadi, S. S. (2014). E(spl): Genetic, Developmental, and Evolutionary Aspects of a Group of Invertebrate Hes Proteins with Close Ties to Notch Signaling. *Current Topics in Developmental Biology: Transcriptional Switches During Development* **110**, 217–262.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.

Doupé, D. P., Marshall, O. J., Dayton, H., Brand, A. H. and Perrimon, N. (2018). Drosophila intestinal stem and progenitor cells are major sources and regulators of homeostatic niche signals. *Proceedings of the National Academy of Sciences* **115**, 12218–12223.

Dutta, D., Dobson, A. J., Houtz, P. L., Gläßer, C., Revah, J., Korzelius, J., Patel, P. H., Edgar, B. A. and Buchon, N. (2015). Regional Cell-Specific Transcriptome Mapping Reveals Regulatory Complexity in the Adult Drosophila Midgut. *CellReports* 1–14.

Dutta, D., Xiang, J. and Edgar, B. A. (2013). RNA Expression Profiling from FACS-Isolated Cells of the Drosophila Intestine. *Curr Protoc Stem Cell Biol* **27**, 2F.2.1–2F.2.12.

D'Rozario, M., Zhang, T., Waddell, E. A., Zhang, Y., Sahin, C., Sharoni, M., Hu, T., Nayal, M., Kutty, K., Liebl, F., et al. (2016). Type I bHLH Proteins Daughterless and Tcf4 Restrict Neurite Branching and Synapse Formation by Repressing Neurexin in Postmitotic Neurons. *CellReports* **15**, 386–397.

Ellis, H. M. (1994). Embryonic expression and function of the Drosophila helix-loop-helix gene, extramacrochaetae. *Mech Dev* **47**, 65–72.

Furriols, M. and Bray, S. (2001). A model Notch response element detects Suppressor of Hairless-dependent molecular switch. *Curr Biol* **11**, 60–64.

Garrell, J. and Modolell, J. (1990). The Drosophila extramacrochaetae locus, an antagonist of proneural genes that, like these genes, encodes a helix-loop-helix protein. *Cell* **61**, 39–48.

Giot, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y. L., Ooi, C. E., Godwin, B., Vitols, E., et al. (2003). A protein interaction map of Drosophila melanogaster. *science (New York, NY)* **302**, 1727–1736.

Graf, T. and Enver, T. (2009). Forcing cells to change lineages. *Nature* **462**, 587–594.

Guo, X., Huang, H., Yang, Z., Cai, T. and Xi, R. (2019). Division of Labor: Roles of Groucho and CtBP in Notch-Mediated Lateral Inhibition that Controls Intestinal Stem Cell Differentiation in Drosophila. *Stem Cell Reports* 1–26.

Guo, Z. and Ohlstein, B. (2015). Bidirectional Notch signaling regulates Drosophila intestinal stem cell multipotency. *science (New York, NY)* **350**, aab0988–aab0988.

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., Zheng, S., Butler, A., Lee, M. J., Wilk, A. J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29.

Hartenstein, V., Takashima, S., Hartenstein, P., Asanad, S. and Asanad, K. (2017). bHLH proneural genes as cell fate determinants of entero-endocrine cells, an evolutionarily conserved lineage sharing a common root with sensory neurons. *Developmental Biology* **431**, 36–47.

Heinze, G., Ploner, M., Jiricka, L. and Steiner, G. (2023). Firth's Bias-Reduced Logistic Regression [R package logistf version 1.25].

Hinz, U., Giebel, B. and CAMPOSORTEGA, J. A. (1994). The Basic-Helix-Loop-Helix Domain of Drosophila Lethal of Scute Protein Is Sufficient for Proneural Function and Activates Neurogenic Genes. *Cell* **76**, 77–87.

Hu, Y., Comjean, A., Perkins, L. A., Perrimon, N. and Mohr, S. E. (2015). GLAD: an Online Database of Gene List Annotation for Drosophila. *J. Genomics* **3**, 75–81.

Hung, R.-J., Hu, Y., Kirchner, R., Liu, Y., Xu, C., Comjean, A., Tattikota, S. G., Li, F., Song, W., Ho Sui, S., et al. (2020). A cell atlas of the adult Drosophila midgut. *Proceedings of the National Academy of Sciences* **117**, 1514–1523.

Jiang, H. and Edgar, B. A. (2012). Intestinal stem cell function in Drosophila and mice. *Curr. Opin. Genet. Dev.* 1–7.

Jiang, H., Patel, P. H., Kohlmaier, A., Grenley, M. O., McEwen, D. G. and Edgar, B. A. (2009). Cytokine/Jak/Stat signaling mediates regeneration and homeostasis in the Drosophila midgut. *Cell* **137**, 1343–1355.

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. and Ishiguro-Watanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **51**, D587–D592.

Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M. N. and Sergushichev, A. (2021). Fast gene set enrichment analysis. *bioRxiv* 060012.

Korzelius, J., Naumann, S. K., Loza-Coll, M. A., Chan, J. S., Dutta, D., Oberheim, J., Gläßer, C., Southall, T. D., Brand, A. H., Jones, D. L., et al. (2014). Escargot maintains stemness and suppresses differentiation in Drosophila intestinal stem cells. *EMBO J* **33**, 2967–2982.

Kunisch, M., Haenlin, M. and Campos-Ortega, J. A. (1994). Lateral inhibition mediated by the Drosophila neurogenic gene delta is enhanced by proneural proteins. *Proc Natl Acad Sci USA* **91**, 10139–10143.

Lan, Q., Cao, M., Kollipara, R. K., Rosa, J. B., Kittler, R. and Jiang, H. (2018). FoxA transcription factor Fork head maintains the intestinal stem/progenitor cell identities in Drosophila. *Developmental Biology* **433**, 324–343.

Lee, T. and Luo, L. (1999). Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron* **22**, 451–461.

Levine, M. and Davidson, E. H. (2005). Gene regulatory networks for development. *Proc Natl Acad Sci USA* **102**, 4936–4942.

Li, K. and Baker, N. E. (2018). Regulation of the Drosophila ID protein Extra macrochaetae by proneural dimerization partners. *eLife* **7**.

Li, K. and Baker, N. E. (2019). Transcriptional and post-transcriptional regulation of extra macrochaetae during Drosophila adult peripheral neurogenesis. *Developmental Biology* **449**, 41–51.

Li, Y., Pang, Z., Huang, H., Wang, C., Cai, T. and Xi, R. (2017). Transcription Factor Antagonism Controls Enteroendocrine Cell Specification from Intestinal Stem Cells. *Sci Rep* **7**, 988.

Liao, Y., Smyth, G. K. and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930.

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550.

Lowe, N., Rees, J. S., Roote, J., Ryder, E., Armean, I. M., Johnson, G., Drummond, E., Spriggs, H., Drummond, J., Magbanua, J. P., et al. (2014). Analysis of the expression patterns, subcellular localisations and interaction partners of Drosophila proteins using a pigP protein trap library. **141**, 3994–4005.

Loza-Coll, M. A., Southall, T. D., Sandall, S. L., Brand, A. H. and Jones, D. L. (2014). Regulation of Drosophila intestinal stem cell maintenance and differentiation by the transcription factor Escargot. *EMBO J* **33**, 2983–2996.

Malaguti, M., Migueles, R. P., Blin, G., Lin, C.-Y. and Lowell, S. (2019). Id1 Stabilizes Epiblast Identity by Sensing Delays in Nodal Activation and Adjusting the Timing of Differentiation. *Dev Cell* 1–22.

Massari, M. E. and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Molecular and Cellular Biology* **20**, 429–440.

Micchelli, C. A. and Perrimon, N. (2006). Evidence that stem cells reside in the adult Drosophila midgut epithelium. *Nature* **439**, 475–479.

Moore, A. W., Barbel, S., Jan, L. Y. and Jan, Y. N. (2000). A genomewide survey of basic helix-loop-helix factors in Drosophila. *Proc Natl Acad Sci USA* **97**, 10436–10441.

Moris, N., Pina, C. and Arias, A. M. (2016). Transition states and cell fate decisions in epigenetic landscapes. *Nat. Rev. Genet.* **17**, 693–703.

Murre, C. (2005). Helix-loop-helix proteins and lymphocyte development. *Nat Immunol* **6**, 1079–1086.

Murre, C., Bain, G., VANDIJK, M. A., Engel, I., Furnari, B. A., Massari, M. E., Matthews, J. R., Quong, M. W., Rivera, R. R. and Stuiver, M. H. (1994). Structure and Function of Helix-Loop-Helix Proteins. *Biochim Biophys Acta* **1218**, 129–135.

Neuhold, L. A. and Wold, B. (1993). HLH forced dimers: Tethering MyoD to E47 generates a dominant positive myogenic factor insulated from negative regulation by Id. *Cell* **74**, 1033–1042.

Ohlstein, B. and Spradling, A. (2006). The adult Drosophila posterior midgut is maintained by pluripotent stem cells. *Nature* **439**, 470–474.

Ohlstein, B. and Spradling, A. (2007). Multipotent Drosophila intestinal stem cells specify daughter cell fates by differential notch signaling. *science (New York, NY)* **315**, 988–992.

Okumura, T., Takeda, K., Kuchiki, M., Akaishi, M., Taniguchi, K. and Adachi-Yamada, T. (2016). GATAe regulates intestinal stem cell maintenance and differentiation in Drosophila adult midgut. *Developmental Biology* **410**, 24–35.

Patel, P. H., Dutta, D. and Edgar, B. A. (2015). Niche appropriation by Drosophila intestinal stem cell tumours. *Nature Cell Biology* **17**, 1182–1192.

Philpott, A. and Winton, D. J. (2014). Lineage selection and plasticity in the intestinal crypt. *Curr Opin Cell Biol* **31**, 39–45.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. Vienna, Austria.

Resende, L. P. F., Truong, M. E., Gomez, A. and Jones, D. L. (2017). Intestinal stem cell ablation reveals differential requirements for survival in response to chemical challenge. *Developmental Biology* **424**, 10–17.

Ritchie, M. E., Phipson, B., Wu, D. I., acids, Y. H. N.2015 limma powers differential expression analyses for RNA-sequencing and microarray studies. *Biometrika*.

Sallee, M. D., Littleford, H. E. and Greenwald, I. (2017). A bHLH Code for Sexually Dimorphic Form and Function of the C. elegans Somatic Gonad. *Curr Biol* **27**, 1853–1860.e5.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–682.

Shokri, L., Inukai, S., Hafner, A., Weinand, K., Hens, K., Vedenko, A., Gisselbrecht, S. S., Dainese, R., Bischof, J., Furger, E., et al. (2019). A Comprehensive Drosophila melanogaster Transcription Factor Interactome. *cell reports* **27**, 955–970.e7.

Simon, C. S., Hadjantonakis, A.-K. and Schröter, C. (2018). Making lineage decisions with biological noise: Lessons from the early mouse embryo. *WIREs Dev Biol* **7**, e319–16.

Siudeja, K., Nassari, S., Gervais, L., Skorski, P., Lameiras, S., Stolfa, D., Zande, M., Bernard, V., Frio, T. R. and Bardin, A. J. (2015). Frequent Somatic Mutation in Adult Intestinal Stem Cells Drives Neoplasia and Genetic Mosaicism during Aging. *Cell Stem Cell* **17**, 663–674.

Spratford, C. M. and Kumar, J. P. (2015). Inhibition of Daughterless by Extramacrochaetae mediates Notch-induced cell proliferation. *Development* **142**, 2058–2068.

Street, K., Risso, D., Fletcher, R. B., Das, D., Ngai, J., Yosef, N., Purdom, E. and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477–16.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102**, 15545–15550.

Tanaka-Matakatsu, M., Miller, J., Borger, D., Tang, W.-J. and Du, W. (2014). Daughterless homodimer synergizes with Eyeless to induce Atonal expression and retinal neuron differentiation. *Developmental Biology* **392**, 256–265.

Tapanes-Castillo, A. and Baylies, M. K. (2004). Notch signaling patterns Drosophila mesodermal segments by regulating the bHLH transcription factor twist. **131**, 2359–2372.

Tetteh, P. W., Basak, O., Farin, H. F., Wiebrands, K., Kretzschmar, K., Begthel, H., van den Born, M., Korving, J., de Sauvage, F., van Es, J. H., et al. (2016). Replacement of Lost Lgr5-Positive Stem Cells through Plasticity of Their Enterocyte-Lineage Daughters. *Cell Stem Cell* **18**, 203–213.

Troost, T., Schneider, M. and Klein, T. (2015). A re-examination of the selection of the sensory organ precursor of the bristle sensilla of Drosophila melanogaster. *PLoS Genet* **11**, e1004911.

van der Flier, L. G., van Gijn, M. E., Hatzis, P., Kujala, P., Haegebarth, A., Stange, D. E., Begthel, H., van den Born, M., Guryev, V., Oving, I., et al. (2009). Transcription factor achaete scute-like 2 controls intestinal stem cell fate. *Cell* **136**, 903–912.

Van Doren, M., Ellis, H. M. and Posakony, J. W. (1991). The Drosophila extramacrochaetae protein antagonizes sequence-specific DNA binding by daughterless/achaete-scute protein complexes. *Development* **113**, 245–255.

Wang, C., Guo, X., Dou, K., Chen, H. and Xi, R. (2015). Ttk69 acts as a master repressor of enteroendocrine cell specification in Drosophila intestinal stem cell lineages. *Development* **142**, 3321–3331.

Wice, B. M. (1998). Forced Expression of Id-1 in the Adult Mouse Small Intestinal Epithelium Is Associated with Development of Adenomas. *Journal of Biological Chemistry* **273**, 25310–25319.

Yin, C. and Xi, R. (2018). A Phyllopod-Mediated Feedback Loop Promotes Intestinal Stem Cell Enteroendocrine Commitment in Drosophila. *Stem Cell Reports* **10**, 43–57.

Zeng, X. and Hou, S. X. (2015). Enteroendocrine cells are generated from stem cells through a distinct progenitor in the adult Drosophila posterior midgut. **142**, 644–653.

Zeng, X., Chauhan, C. and Hou, S. X. (2010). Characterization of midgut stem cell- and enteroblast-specific Gal4 lines in drosophila. *Genesis* **48**, 607–611.

Zhai, Z., Kondo, S., Ha, N., Boquete, J.-P., Brunner, M., Ueda, R. and Lemaitre, B. (2015). Accumulation of differentiating intestinal stem cell progenies drives tumorigenesis. *Nat Commun* **6**, 10219–27.

Zhang, C., Jin, Y., Marchetti, M., Lewis, M. R., Hammouda, O. T. and Edgar, B. A. (2022). EGFR signaling activates intestinal stem cells by promoting mitochondrial biogenesis and β-oxidation. *Curr Biol* **32**, 3704–3719.e7.

Zhang, N., Yantiss, R. K., Nam, H.-S., Chin, Y., Zhou, X. K., Scherl, E. J., Bosworth, B. P., Subbaramaiah, K., Dannenberg, A. J. and Benezra, R. (2014). ID1 Is a Functional Marker for Intestinal Stem and Progenitor Cells Required for Normal Response to Injury. *Stem Cell Reports* **3**, 716–724.

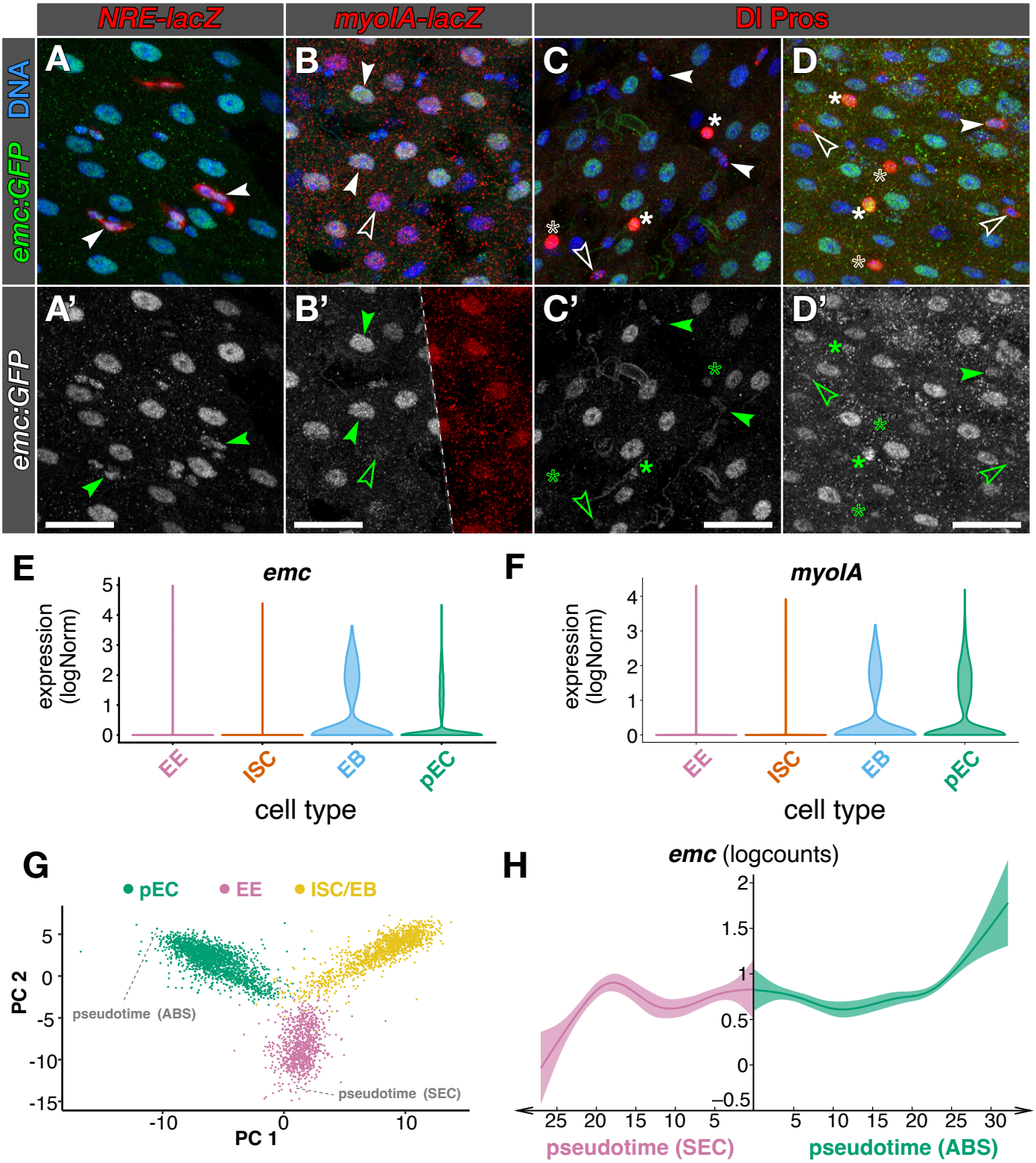


Figure 1. *emc* is expressed preferentially in the absorptive lineage.
A-D. Projected confocal stacks showing endogenous Emc:GFP expression. Emc is expressed in EBs (marked with *NRE-lacZ*, arrowheads in A, A') and most ECs (positive for *myoIA-lacZ*, solid arrowheads in B, B') but not all (empty arrowhead in B, B'). *emc* expression can also be detected in some ISCs (Dl⁺) and EEs (Pros⁺) (solid arrowheads and asterisks, respectively, in C-D') but not all (empty arrowheads and asterisks in C-D'). **E, F.** Expression of *emc* (E) and the EC marker *myoIA* (F) by scRNAseq in EEs, ISCs, EBs and ECs from the posterior midgut (pECs). Both genes are enriched in EBs and ECs, with rare events of detection in EEs and ISCs. **G.** Pseudotime trajectories from ISC/EB cells into mature EEs or ECs derived from scRNAseq data. **H.** *emc* expression in single cells declines as these get closer to the EE fate and increases as they progress towards EC differentiation.

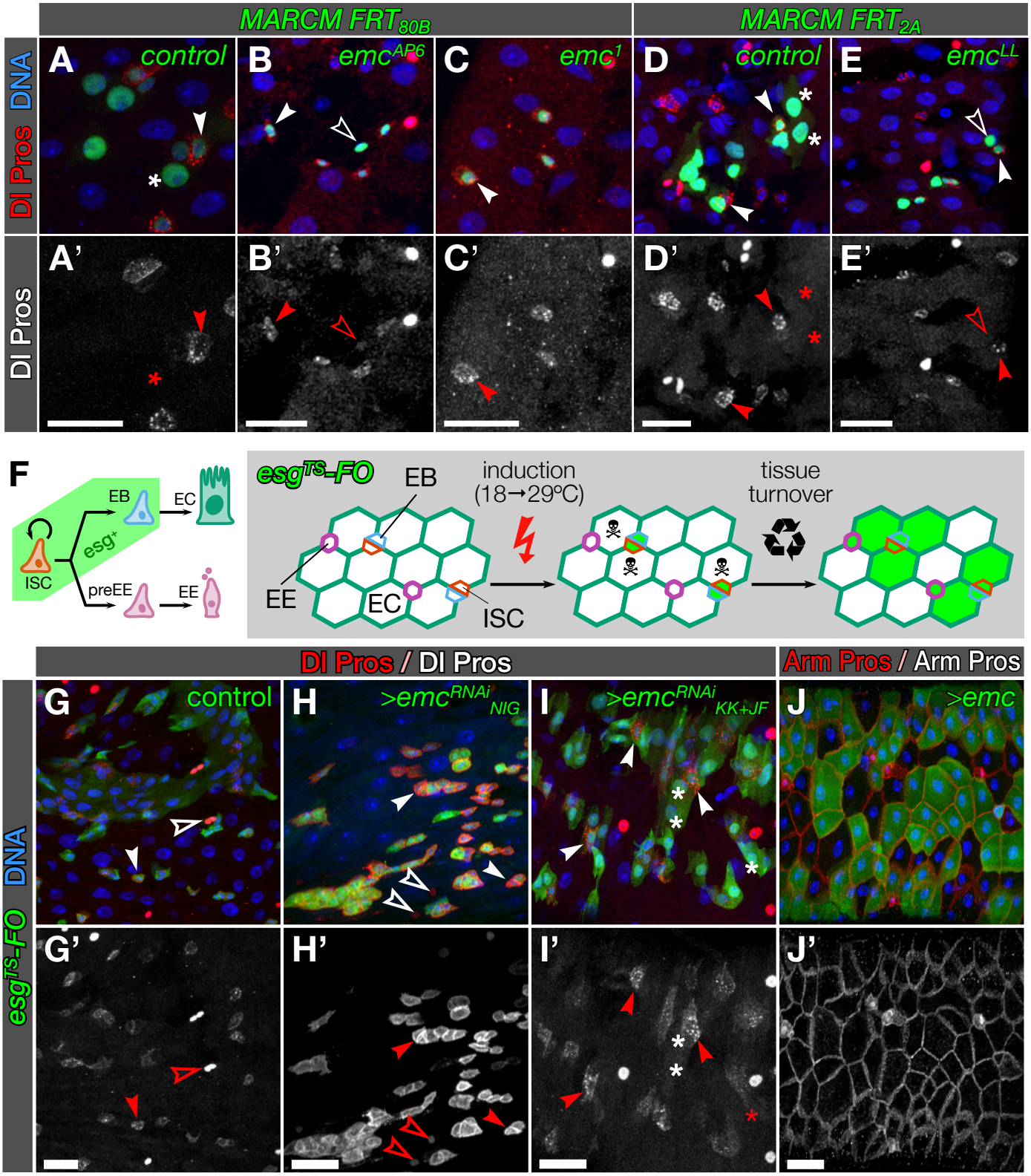


Figure 2. *emc* is necessary for differentiation and sufficient for selection of the EC fate.
A-E. Differentiation is impaired in *emc* mutant clones. Control clones display more polyploid Dl⁻/Pros⁻ ECs (asterisks in E, H) than clones homozygous of for the mutant alleles *emc*^{AP6} (F), *emc*^{1} (G) and *emc*^{LL02590} (I), which are enriched in Dl⁺ ISCs (solid arrowheads) and diploid Dl⁻/Pros⁻ EBs (empty arrowheads) with respect to controls. **F.** The *esg*^{TS}*-FO* method allows the temporal control of expression in progenitor cells (esg⁺ ISCs and EBs) and their lineage tracing to access their fate potential. *UAS-GFP* and *UAS-FLP* are expressed specifically in ISCs and EBs after induction by temperature switch. This leads to the progenitor-restricted expression of the (potentially) ubiquitous *Actin5C-Gal4* by excision of a transcriptional terminator. *Actin5C-Gal4* maintains *UAS-GFP* expression in the ISC/EB lineage when differentiation occurs and *esg* expression stops. **G-J.** Emc is required for differentiation and sufficient for induce the EC fate. ISCs and EBs in control guts produce both EC (Dl⁻/Pros⁻) and EE (Dl⁻/Pros⁺) GFP⁺ cells (B, note the difference in intensity levels between Dl and Pros, respectively solid and empty arrowheads), while differentiation is reduced when ISCs/EBs express *emc*-specific RNAi , shown by increased number of Dl⁺/GFP⁺ cells (C, D). Reduction of *emc* also induces increased Dl expression: compare Dl-expressing with Pros-expressing cells (respectively solid and empty arrowheads in C, D, compare with those similarly marked in B). Forced expression of *emc* in ISCs/EBs forces differentiation into ECs (E). Scale bar: 20μm in all panels.

Figure 3 Puig et al., 2023

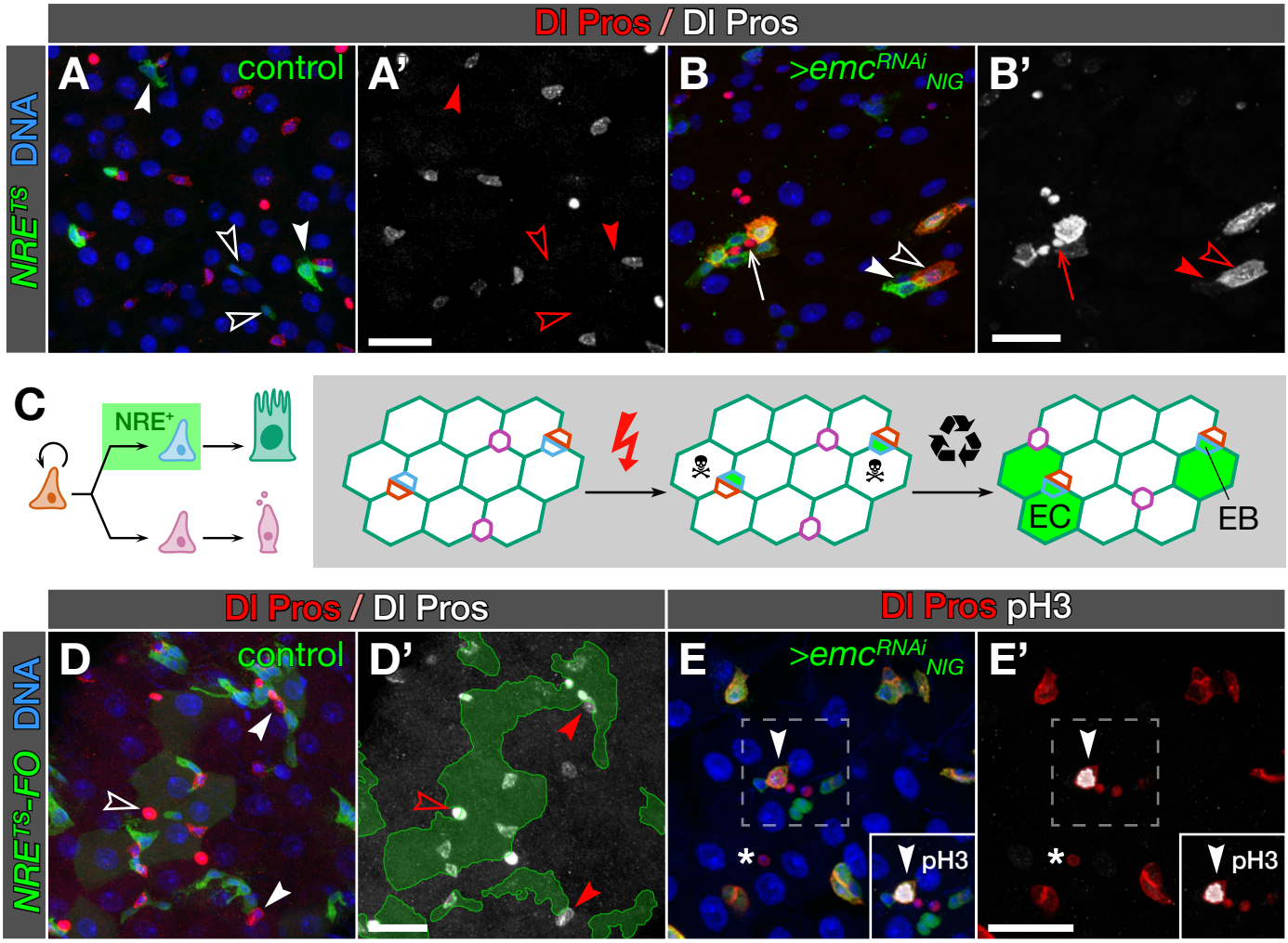
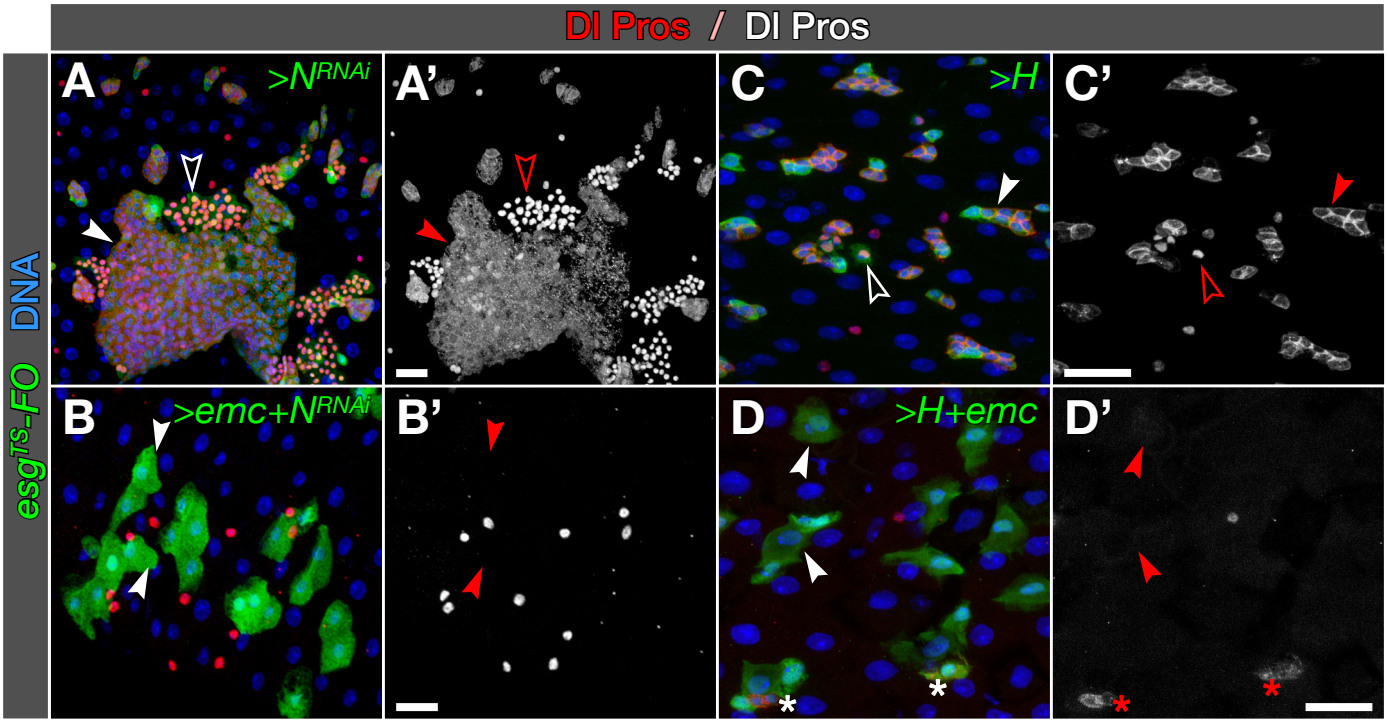
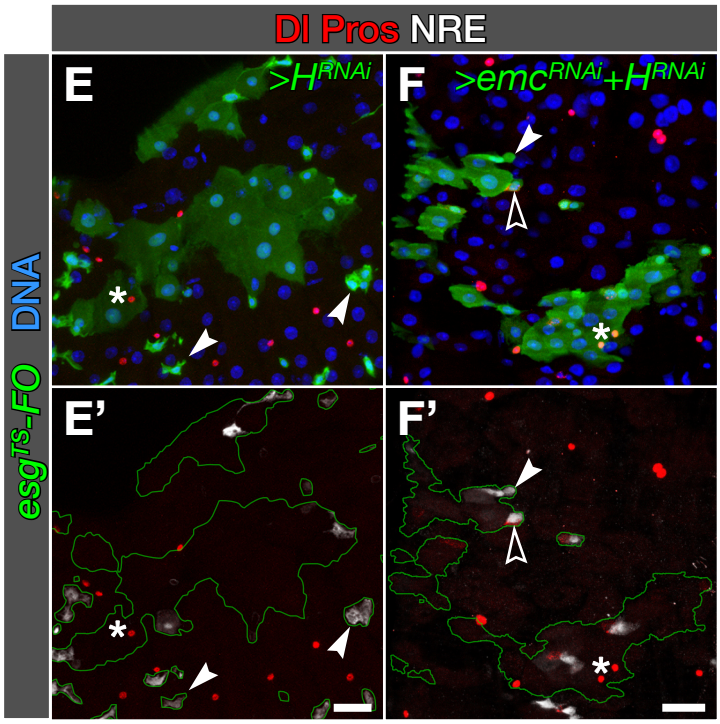
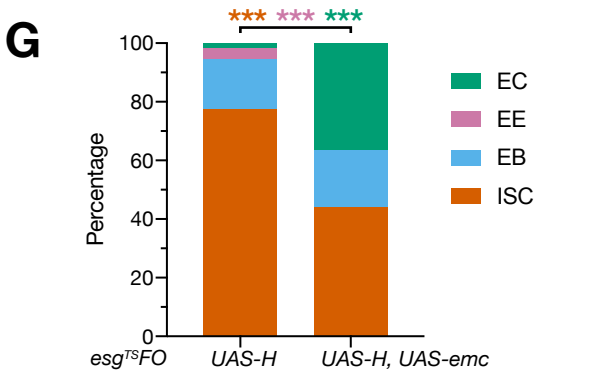


Figure 3. *emc* is required for maintenance of the EB committed state.
A-B. The *NRE-Gal4* driver is expressed in EBs at different levels in normal intestines (solid and empty arrowheads in A for higher and lower-expressing EBs), but never co-express Dl. When *emc* is depleted with *NRE-Gal4*, many EBs co-express Dl (empty arrowheads, B) but not all (solid arrowheads, B). Also, EEs develop close to *emc*-depleted EBs (arrow, B). **C.** Use of *NRE^{TS}-FO* allows the simultaneous manipulation and tracing of the EB lineage. *NRE-Gal4, UAS-GFP* is expressed specifically in EBs after induction and activates specifically in EBs the (potentially) ubiquitous *Actin5C-Gal4*, which in turn will maintain *UAS-GFP* expression if differentiation or decommitment occur. **D-E**. Loss of *emc* in EBs leads to their decommitment. Activation of *NRE^{TS}-FO* in control guts leads exclusively to the labelling of EBs and ECs, with no co-expression of GFP and Dl or Pros (solid and empty arrowheads, respectively, in D; some ISCs and EEs in the micrograph seem inside the GFP territory due to the Z-projection of the confocal stack). By contrast, in *NRE^{TS}-FO* guts where *emc* is depleted, EB-derived cells express Dl and are capable of mitosis, as judged by phospho-Histone 3 staining (arrowhead in E). Scale bar: 20μm in all panels.







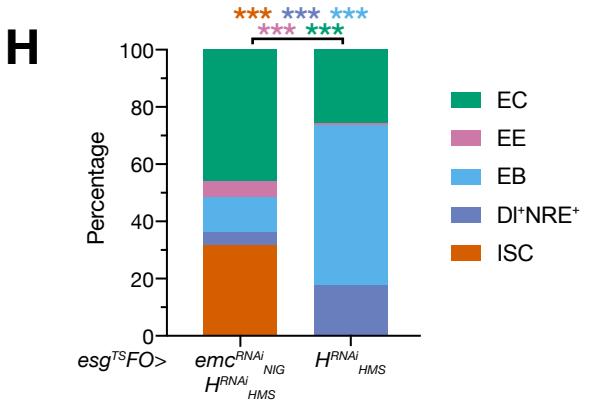
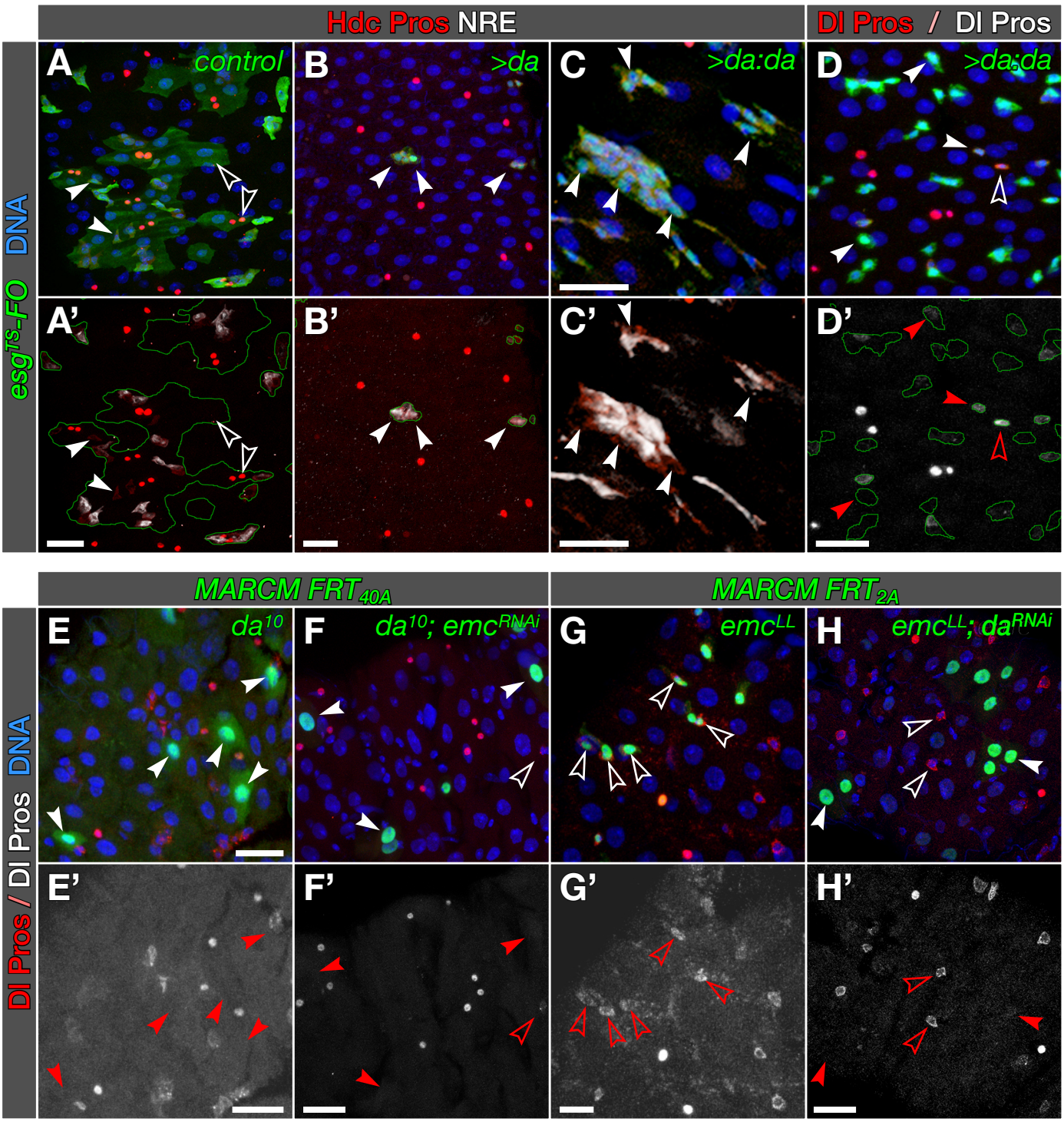


Figure 4. *emc* acts in parallel to Notch signalling.
A-D. *emc* can direct EC differentiation in the absence of Notch signalling. EC differentiation is abolished or strongly impaired with knockdown of *N* (A) or overexpression of H (C), respectively, with accumulation of Dl⁺ or Pros⁺ cells (solid and empty arrowheads, respectively, in A, C). Simultaneous overexpression of *emc* induces EC differentiation in both cases (arrowheads in B, D), but to a lesser an extent in a background of *H* overexpression (asterisks in D, compare with B). **E-F.** Notch signalling can direct EC differentiation in the absence of *emc*. Loss of EC differentiation caused by depletion of *emc* (see panel H and Fig. 2H-I and S2F) is rescued by simultaneous knockdown of *H* (F). Depletion of *H* can induce EC differentiation on its own (E). **G-H.** Stacked bars plots summarising the cell type composition of *esg*^{TS}*-FO* tissue. Numbers (total GFP⁺ cells / fields of view) are 1810/21 and 1292/27 for *UAS-H* and *UAS-H+UAS-emc*, respectively (G) and 984/15 and 126/16 for *UAS-H*$^{RNAi}_{HMS}$ and *UAS-H*$^{RNAi}_{HMS}$*+UAS-emc*$^{RNAi}_{NIG}$, respectively (H). p-values are <0.05, <0.01 and <0.001 for one, two or three asterisks (see Tables S1 and S2). Scale bar: 20μm in all panels.



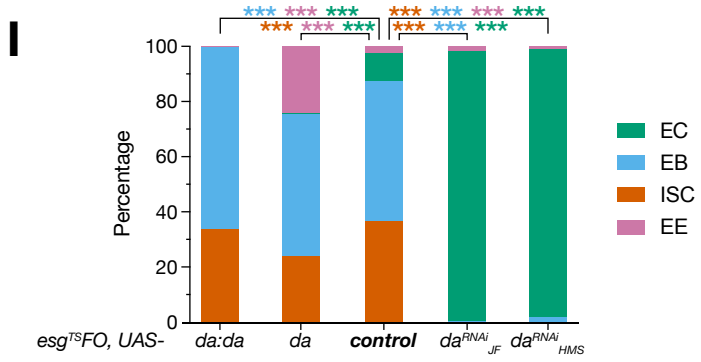
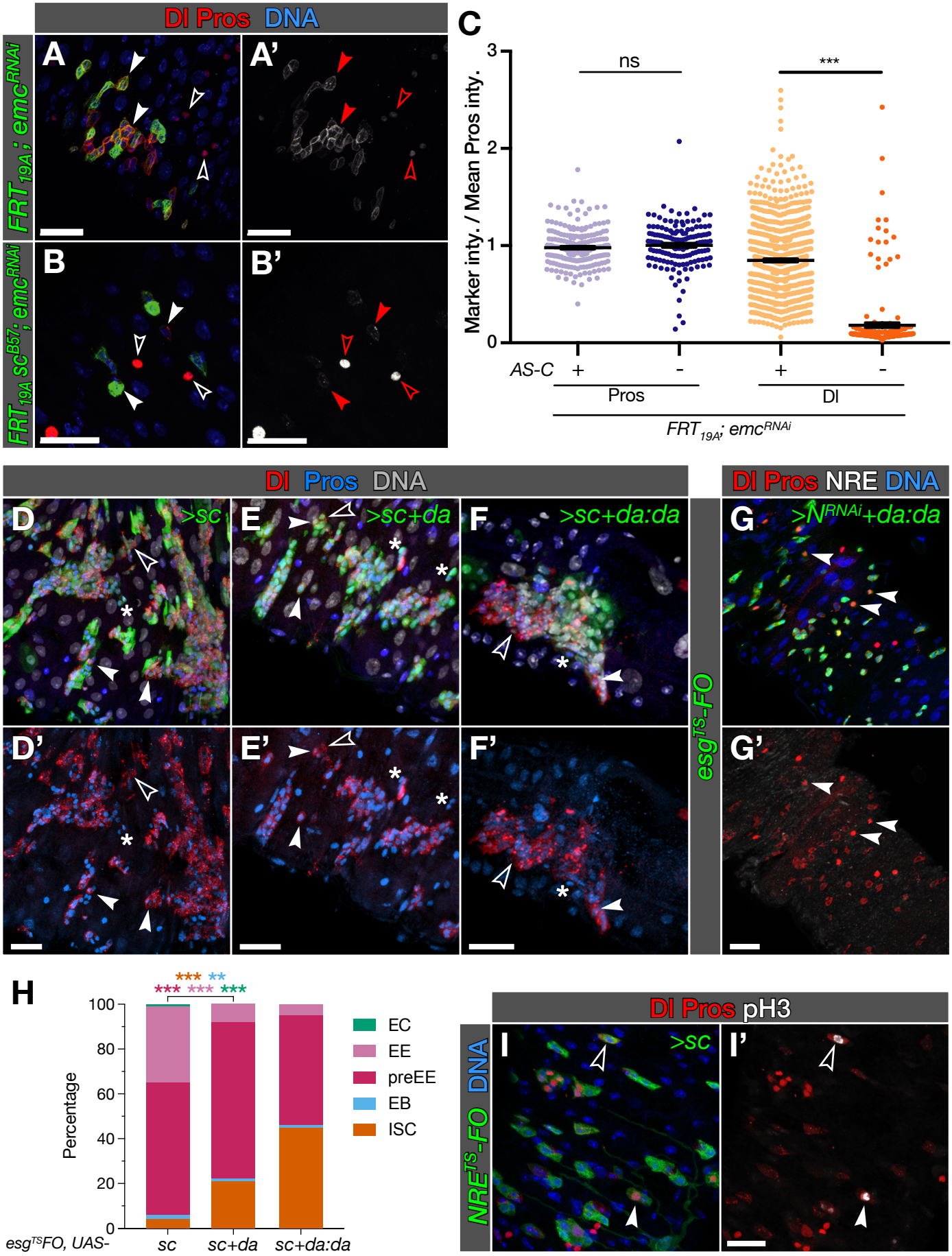


Figure 5. Da homodimers prevent differentiation and are antagonised by Emc.
A-D. Da homodimers maintain ISCs/EBs undifferentiated. Solid arrowheads: ISCs/EBs; empty arrowheads: ECs/EEs. Overexpression of Da with *esg^{TS}-FO* prevents formation of ECs, though it allows EE differentiation (B, compare with A). Expression of a forced Da:Da dimer in esg^{+} cells prevents all differentiation, as evaluated with Hdc/NRE-lacZ (C) or Dl (D). **E-H.** *da* is epistatic over *emc*. Solid arrowheads: ECs; empty arrowheads: Dl^{+} ISCs. MARCM clones mutant for *da^{10}* differentiate as ECs (E). This phenotype is recapitulated when these clones simultaneously express *emc*-specific RNAi (F). Likewise, impaired differentiation in *emc^{LL02590}* clones (G) is rescued by simultaneous knockdown of *da*, which leads to EC formation (H). **I.** Stacked bars plots summarising the cell type composition of *esg^{TS}-FO* tissue. Numbers (total GFP^{+} cells / fields of view) are 1135/15, 549/14, 693/13, 803/15 and 1086/19 for control, *UAS-da*, *UAS-da:da*, *UAS-da$^{RNAi}_{JF}$* and *UAS-da$^{RNAi}_{HMS}$*, respectively. P-values are <0.05, <0.01 and <0.001 for one, two or three asterisks (see Tables S1 and S2). Scale bar: 20μm in all panels.

Figure 6

Puig et al., 2023

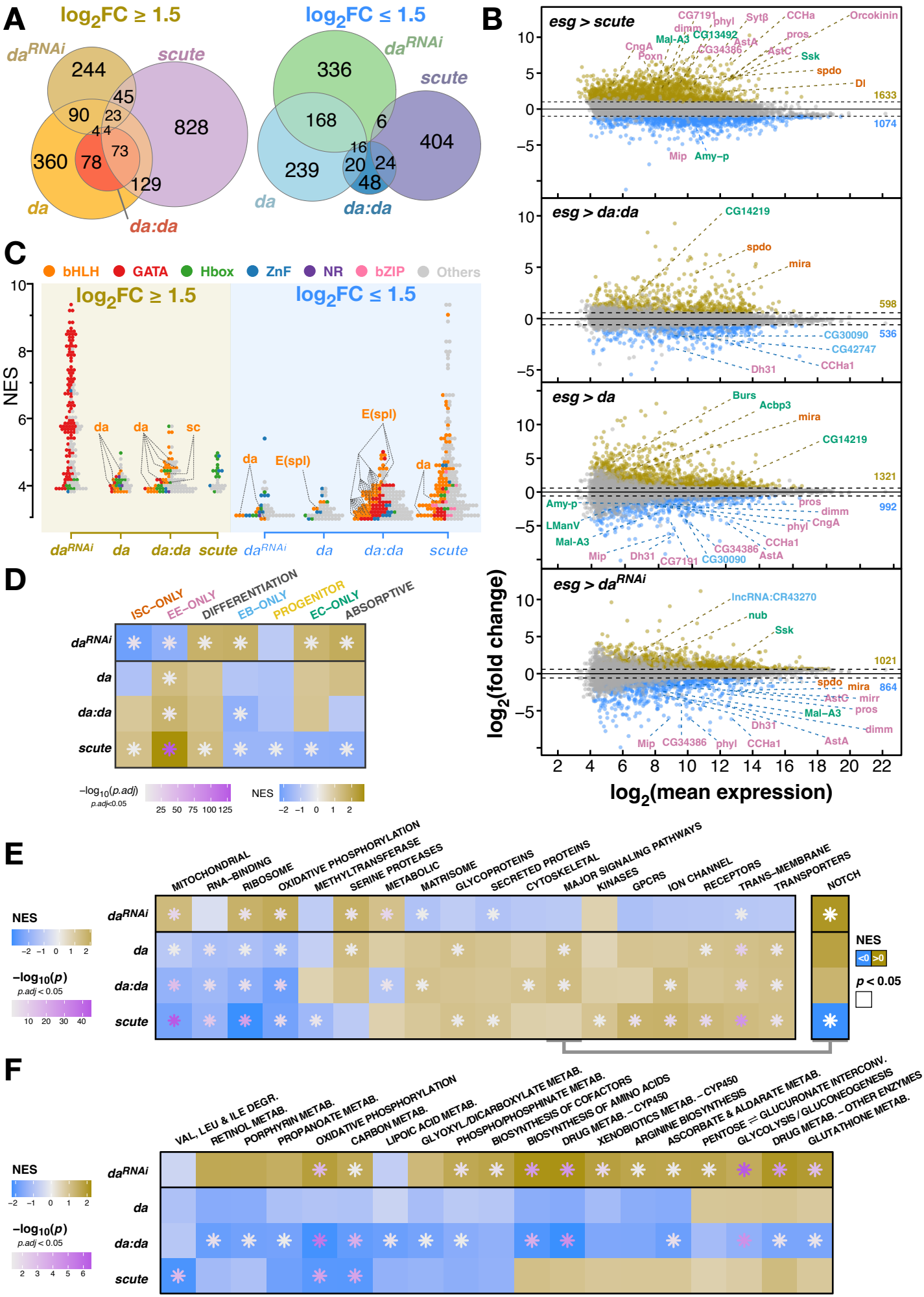


(from the previous page)

Figure 6. Sc:Da and Da:Da antagonise each other in EE formation but collaborate in ISC maintenance.
A-C. Emc antagonises Sc in inducing Dl expression. Loss of the *AS-C* using the sc^{B57} deficiency reduces the increase in Dl expression observed with knockdown of *emc* (B, compare with A). Dl levels are quantified by comparison with those of Pros (see Appendix). Numbers (Dl^+ cells/$Pros^+$ cells) are 210/146 and 718/210 for $UAS\text{-}emc^{RNAi}{}_{NIG}$ and $UAS\text{-}emc^{RNAi}{}_{NIG}$, $Df(1)sc^{B57}$, respectively. P-values are <0.05, <0.01 and <0.001 for one, two or three asterisks (Mann-Whitney test). **D-F.** Antagonism/ collaboration between Sc and Da. Solid arrowheads: Dl^+/$Pros^+$ pre-EEs; empty arrowheads: Dl^+ ISCs; asterisks: $Pros^+$ EEs. Overexpression of Sc leads to a dramatic increase of Dl^+/$Pros^+$ pre-EEs and $Pros^+$ EEs, but maintains a population of Dl^+ ISCs (D). The population of ISCs progressively increases by co-expression of monomeric Da (E) and dimeric Da:Da (F), at the cost of EE differentiation (E, F) and pre-EE formation (F). Quantified in H. **G.** Da:Da antagonises EE formation. While reduction of Notch signalling leads to an excess of EEs, simultaneous overexpression of Da:Da and *N*-specific RNAi results in very few $Pros^+$ cells (solid arrowheads, compare with Fig 4A). **H.** Stacked bars plots summarising the cell type composition of $esg^{TS}\text{-}FO$ tissue. Numbers (total GFP^+ cells / fields of view) are 1241/6, 2131/10 and not determined/11 for *UAS-sc, UAS-sc, UAS-da*, and *UAS-sc, UAS-da:da*, respectively. p-values are <0.05, <0.01 and <0.001 for one, two or three asterisks (see Tables S1 and S2). **I.** Sc can induce decommitment of EBs. Expression of Sc with $NRE^{TS}\text{-}FO$ leads to reacquisition of ISC properties (Dl expression, mitosis assessed with phospho-Histone 3, empty arrowhead) and pre-EE formation (Pros expression and mitosis, solid arrowhead). Scale bar: 20μm in all panels.

Figure 7

Puig et al., 2023



(from the previous page)

Figure 7. Da and Sc cooperate to induce expression of the ISC signature.
A. Euler diagrams showing the intersections of sets of genes either upregulated (left) or downregulated (right) by *da*^{RNAi}, Da:Da, Da and Sc, with fold changes larger than $\log_2(1.5)$. The area of the sets denotes (approximately) the size of the intersection. **B.** MA plots for the overexpression of *da*^{RNAi}, Da:Da, Da and Sc, respectively. Some genes of interest are identified because of their role or expression pattern. Colours match the cell types where genes are expressed. **C**. Beeswarm plots of the normalised enrichment scores of DNA motifs found in upregulated (left) or downregulated (right) genes in the four conditions analysed, using RcisTarget. Dots represent individual motifs, binned for simplicity of representation, and coloured by the type of transcription factor they are bound by. *da*, *sc*, and *E(spl)-C* gene members are identified. **D-F**. Heatmaps of gene set enrichment scores for sets of genes that are: expressed in specific cell types (Dutta *et al*, 2015) (D), annotated because of their broad molecular function in *Drosophila* (Hu *et al*, 2015) (E) or belonging to annotated KEGG pathways (F). Asterisks indicate the enrichment scores with significant adjusted p-values, with colour indicating the level of significance. The colourmap of NES scores corresponds to different values in each heatmap.

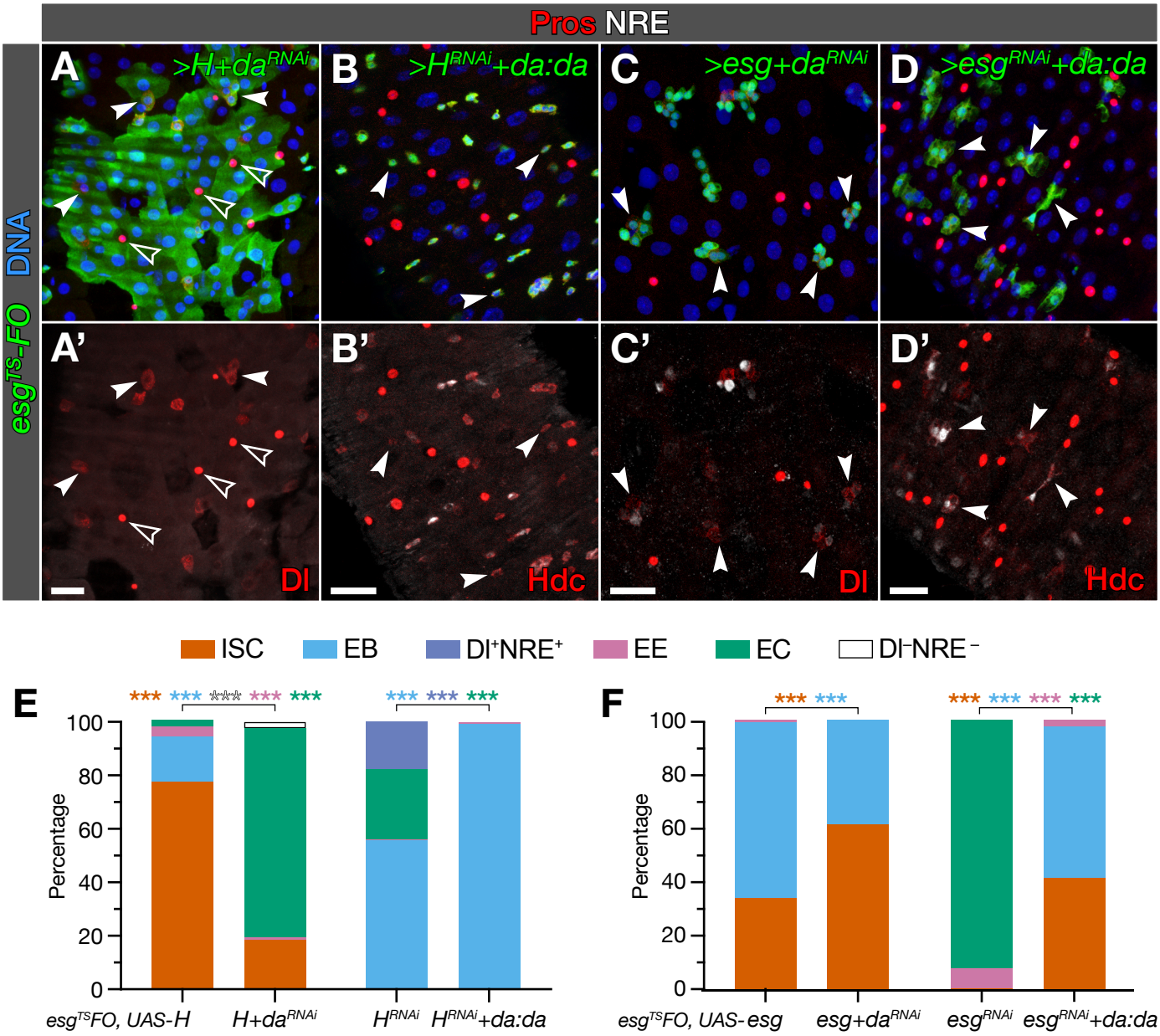


Figure 8. *da* is epistatic over Notch signalling but acts in parallel with *esg*.
Solid arrowheads: ISCs (Hdc⁺/NRE⁻ or Dl⁺/NRE⁻); empty arrowheads: EEs (Pros⁺). **A-B.** *da* is epistatic over Notch signalling. Loss of *da* leads to EC and EE differentiation when Notch signalling is impaired by overexpression of *H* (A), while overexpression of *da:da* maintains cell undifferentiated even when Notch signalling is exacerbated by knockdown of *H* (B). **C-D.** *da* and *esg* function in parallel to prevent differentiation. Overexpression of *esg* can prevent differentiation and compensate for the loss of *da* (C), while overexpression of *da:da* can similarly compensate for the loss of *esg*, including prevention of EE formation (D). **E-F.** Stacked bars plots summarising the cell type composition of esg^{TS}-*FO* tissue. Numbers (total GFP⁺ cells / fields of view) are 1510/21, 1810/21, 613/8 and 984/15 for *UAS-H*, *UAS-H+UAS-da*$^{RNAi}_{JF}$, *UAS-H*$^{RNAi}_{HMS}$ and *UAS-H*$^{RNAi}_{HMS}$*+UAS-da:da*, respectively (E) and 850/9, 1236/12, 628/18 and 939/19 for for *UAS-esg UAS-esg+UAS-da*$^{RNAi}_{JF}$, *UAS-esg*$^{RNAi}_{HMS}$ and *UAS-esg*$^{RNAi}_{HMS}$*+UAS-da:da*, respectively (F). p-values are <0.05, <0.01 and <0.001 for one, two or three asterisks (see Tables S1 and S2 for further details). Scale bar: 20μm in all panels.

Figure 9 Puig et al., 2023

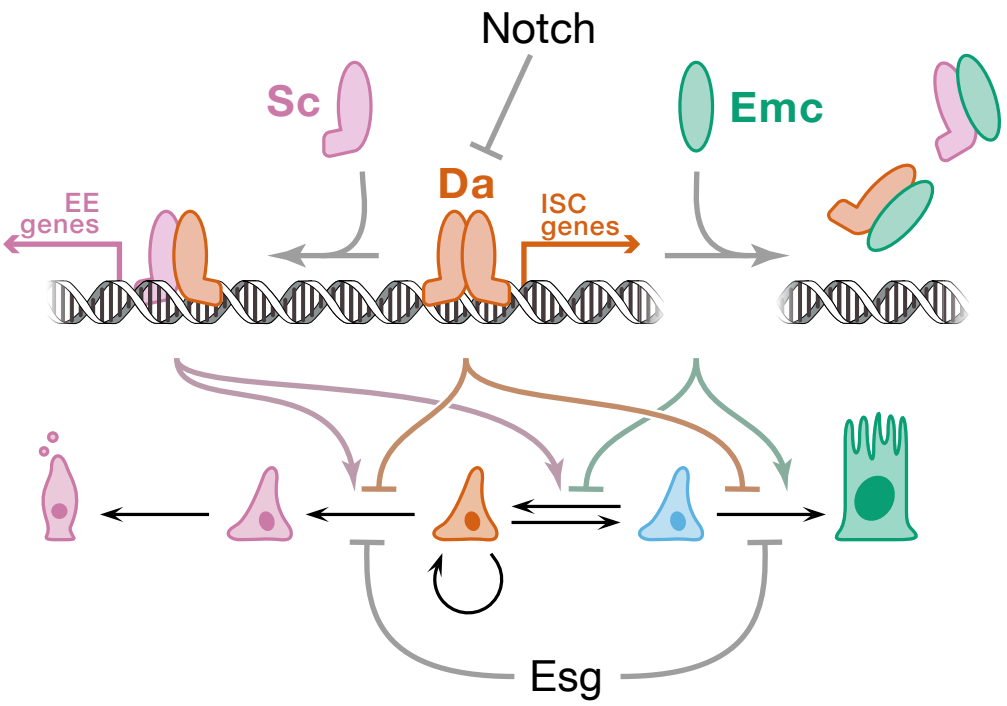


Fig 9. Sc, Da and Emc regulate self-renewal and bipotential differentiation in the intestine. Orange arrows: Da homodimers maintain ISCs and EBs undifferentiated, antagonising EC differentiation and pre-EE formation by inducing expression of ISC-specific genes. Pink arrows: An increase of Sc expression induces formation of Da:Sc heterodimers, which steer the transcriptional program towards pre-EE formation; Sc can also induce expression of the ISC-specific signature, potentially leading to EB decommitment if its activity is not prevented by titration by Emc. Green arrows: Emc increased expression in EBs maintains commitment for EC differentiation and can direct terminal EC differentiation by interruption of Da:Da and Da:Sc transcriptional activities. This occurs in parallel to Esg prevention of differentiation and Notch signalling being upstream and opposing Da:Da to induce EC differentiation.