

Linköping University Medical Dissertation No. 1550

Genetic Mechanisms during Terminal Cell Fate Specification in the *Drosophila* CNS

Johannes Stratmann



Department of Clinical and Experimental Medicine

Faculty of Medicine and Health Sciences

Linköping University, SE-58185

Linköping, Sweden

Linköping 2017

© Johannes Stratmann, 2017

Front cover: Confocal image of the *Nplp1-CRM-GFP* construct.

Back cover: Confocal image of the *Nplp1-CRM-GFP* construct in misexpression background for the *Nplp1* cocktail, resulting in over activation of the enhancer.

Published articles I and II have been reprinted with permission of the copyright holders.

Printed in Sweden by Liu-Tryck, Linköping, Sweden, 2017

ISBN: 978-91-7685-647-5

ISSN: 0345-0082

“There is no such thing as hard bread. No bread-that’s hard”

(My Dad)

Faculty Opponent

Christos Samakovlis

Department of Molecular Biosciences,

The Wenner-Gren Institute

Stockholm University, Sweden

Supervisor

Stefan Thor

Department of Clinical and Experimental Medicine

Faculty of Health Sciences

Linköping University, Sweden

Co-Supervisor

Jan-Ingvar Jönsson

Department of Clinical and Experimental Medicine

Faculty of Health Sciences

Linköping University, Sweden

Abstract

Specification of the many unique neuronal subtypes found in the nervous system depends on spatiotemporal cues and terminal selector cascades, mostly acting in sequential combinatorial codes of transcription factors (TFs) to dictate cell fate. Out of 10,000 cells in the *Drosophila* embryonic ventral nerve cord (VNC), only 28 cells selectively express Nplp1. The Nplp1 neurons in the *Drosophila* VNC can be subdivided into the thoracic ventro-lateral Tv1 and the dorsal-medial dAp neurons. Nplp1 expression in both cell subtypes is activated by the same terminal selector cascade: *col* > *ap/eya* > *dimm* > *Nplp1*. However Tv1 and dAp neurons are generated by different neuronal progenitors (neuroblasts, NB), and depend on different upstream cues to activate the cell specification cascade. The Tv1 cells are generated by NB5-6T, and in these cells the Nplp1 terminal selector cascade is triggered by spatio-temporal input provided by *Antp/hth/exd/lbe/cas*. Our studies identified that NB4-3 gives rise to the dAp cells and that the Nplp1 terminal selector cascade in dAp cells is activated by *Kr/pdm>grn*. I demonstrated how two different spatio-temporal combinations can funnel on a shared downstream terminal selector cascade to determine a highly related cell fate, in different regions of the VNC. I tested this scenario at the molecular level, by identification of cis-regulatory modules (CRMs) for the main factors involved in the Nplp1 terminal selector cascade. Intriguingly, I found that *col* is under control of two separate CRMs, which are controlled by either *Antp/hth/exd/lbe/cas* in the NB5-6T lineage, and *Kr/pdm/grn* in the NB4-3 lineage. In addition, CRISPR deletion of the endogenous *col* CRMs did not result in loss of Col and Nplp1, indicating that *col* might be under control of more, yet unidentified CRMs. Nplp1 is expressed in one out of four cells in the thoracic Apterous cluster (Ap cluster); the Tv1 cell. The allocation of the right cell fate to each of the four Ap cluster cells, is regulated by the sub-temporal cascade including the factors Sqz/Nab/Svp, acting downstream of the temporal factor Cas. The sub-temporal factors have a repressive action on Col and Dimm, and thus on the terminal selector cascade regulating Nplp1 expression in the Tv1 cell. We demonstrated that the late and Tv1 specific expression of the early temporal factor Kr suppresses Svp in the Tv1 cell and allows for the progression of the Nplp1 cell fate specification cascade. Hence, early temporal factors involved in temporal progression of neuronal progenitors, can be re-utilized late and postmitotically to specify cell fate. It is tempting to speculate that similar mechanisms act to generate similar cell fate in different regions of the CNS, as well as the issue of sub-temporal multitasking, are common features both in *Drosophila* and higher organisms.

Populärvetenskaplig sammanfattning

Det centrala nervsystemet (CNS) är samlingsnamnet för hjärnan och ryggmärgen och det är kroppens mest komplexa biologiska struktur. CNS består av miljardtals celler som tillsammans bygger upp ett nätverk som styr kroppen. Alla celler i kroppen har under utvecklingen från embryonalstadiet utgått från så kallade stamceller. Olika stamceller ligger till grund för organismers olika vävnader och organ. Under utvecklingen kommer dessa stamceller att dela sig och bilda allt mer specifika celltyper, för att till slut producera mogna celler som styr kroppens olika funktioner. Det innebär att cellerna vid vissa tidpunkter under utvecklingen kommer att stå inför olika val: till vilken typ av cell de ska utvecklas; hur många gånger de skall dela sig; när delningarna skall ske. All information för att fatta dessa beslut finns i cellens genetiska material, i dess DNA. Eftersom alla celler i kroppen innehåller samma DNA ställs frågan: hur kan då cellerna ta *olika* beslut, och bilda *olika* typer av celler, om alla celler innehåller samma information? Svaret ligger i att cellerna använder olika delar av informationen i DNA under utvecklingens gång. Forskning pågår för att kartlägga och förstå hur cellerna använder specifika delar av sitt DNA och hur informationen i delarna används. Det stora antalet celler i människans CNS gör det svårt att forska på mänskligt material. Istället använder forskare sig av enklare modellorganismer som bananflugan, *Drosophila*. Många gener som är viktiga för utvecklingen av människans nervsystem finns också i bananflugan. Antalet celler i bananflugans CNS är uppskattningsvis 15 000 celler – att jämföra med miljarder i människans. Experimenten som beskrivs i denna avhandling fokuserar på utvecklingen av 28 nervceller i bananflugans CNS. En spännande egenskap hos dessa 28 celler är att de specificerar sig till en likadan celltyp trots att de sitter i olika delar av CNS och därmed uppkommer från olika stamceller. Dessa celler skapas dessutom vid olika tidpunkter vilket betyder att trots att cellerna har olika tillgång till den genetiska informationen som bestämmer var och när cellen skall uppkomma så aktiveras ett gemensamt genetisk program som i slutändan skapar likadana celler. Genom att analysera specifika regioner av bananflugans DNA och deras inverkan på cellutvecklingen hos de 28 nervcellerna kan vi nu bättre förstå hur informationen som finns lagrad i DNA bestämmer cellernas öde. Celler har också inbyggda ”genetiska klockor” som styr deras utveckling. Dessa ”klockgener” startar vissa genetiska program vid specifika tidpunkter och vi har identifierat en av flertalet gener som bestämmer nervcellernas öde. De nya mekanismerna som identifieras i avhandlingen är inte bara viktiga för bananflugans neuronala utveckling, utan troligtvis också för den neuronala utvecklingen hos däggdjur.

Table of Contents

Abstract.....	i
Populärvetenskaplig sammanfattning	ii
List of Papers.....	iii
Abbreviations.....	iv
Introduction	- 1 -
General Aspects of Early Embryonic Development	- 1 -
Molecular Aspects of Development.....	- 3 -
Examples of Molecular Events Involved in Germ Layer Formation	- 3 -
Examples of Molecular Events Involved in Body Patterning	- 4 -
Terminal Cell Fate can be reversed.....	- 5 -
The Fruit Fly <i>Drosophila melanogaster</i> as a Model Organism	- 5 -
<i>Drosophila</i> Embryogenesis.....	- 7 -
Early <i>Drosophila</i> Embryonic Development	- 7 -
Neural Development in <i>Drosophila</i>	- 10 -
From Ectoderm to Neuroectoderm	- 12 -
.....	- 12 -
Anterior Posterior Axis Formation of the Nervous System	- 13 -
Dorsal Ventral Axis formation of the Nervous system: Columnar patterning.....	- 13 -
.....	- 14 -
Segment-polarity patterning.....	- 15 -
Hox Genes in the CNS	- 16 -
Neuroblast delamination, Lateral inhibition.....	- 18 -
Lineage Trees	- 19 -
Apoptosis/PCD	- 21 -
Temporal Cascade and Sub-Temporal Factors	- 22 -
Similar Regulation in Higher Organisms	- 23 -
Model: Apterous cluster, Tv1 and dAp cells	- 25 -
Regulatory Networks	- 26 -
Identification of Candidate Genes and Relations between Different Factors.....	- 28 -
Gene Regulation.....	- 29 -
General Structure of a Gene	- 30 -
Regulation of Gene Expression: Transcriptional Control.....	- 31 -
Results and Discussion	- 36 -

Paper I: Neuronal Cell Fate Specification by the Convergence of Different Spatiotemporal Cues on a Common Terminal Selector Cascade	- 36 -
Background and Aim	- 36 -
Results	- 37 -
Discussion.....	- 39 -
Paper II: Neuronal cell fate diversification controlled by sub-temporal action of Kruppel.....	- 39 -
Background and Aim	- 39 -
Results	- 40 -
Discussion.....	- 43 -
Paper III: Neuronal Cell Fate Specification by the Molecular Convergence of Different Spatio-Temporal Cues on a Common Initiator Terminal Selector Gene	- 44 -
Background and Aim	- 44 -
Results	- 45 -
.....	- 47 -
Discussion.....	- 47 -
Concluding Remarks.....	- 49 -
Methods.....	- 51 -
<i>Gal4/UAS</i> transgenic expression system	- 51 -
Immunohistochemistry (IHC)	- 51 -
Confocal scanning and image analysis.....	- 52 -
Statistics	- 52 -
Transgenic flies	- 52 -
Enhancer identification Enhancer Mutation.....	- 52 -
CRISPR/Cas9 deletion of Enhancers.....	- 52 -
.....	- 52 -
Acknowledgements.....	- 53 -
REFERENCES	- 55 -

List of Papers

This thesis is based on the following papers, which will be referred to in the text by their roman numerals (I-III):

Paper I

Hugo Gabilondo[✉], **Johannes Stratmann**[✉], Irene Rubio-Ferrera, Irene Millán-Crespo, Patricia Contero-García, Shahrzad Bahrampour, Stefan Thor, Jonathan Benito-Sipos.

Neuronal Cell Fate Specification by the Convergence of Different Spatiotemporal Cues on a Common Terminal Selector Cascade. **PLoS Biology**, 2016 May; 14(5) e1002450

[✉] Equal contribution

Paper II

Johannes Stratmann[✉], Hugo Gabilondo[✉], Jonathan Benito-Sipos, Stefan Thor.

Neuronal cell fate diversification controlled by sub-temporal action of Kruppel. **eLife**, 2016 Oct; (5)

[✉] Equal contribution

Paper III

Johannes Stratmann and Stefan Thor

Neuronal Cell Fate Specification by the Molecular Convergence of Different Spatio-Temporal Cues on a Common Initiator Terminal Selector Gene. **Submitted Manuscript**.

Abbreviations

AFT	Air filled trachea
Ap	Apterous
Ap Cluster	Apterous Cluster
BMP	Bone morphogenetic protein
Cas9	CRISPR associated protein 9
CNS	Central nervous system
CRISPR	Clustered regulatory interspaced short palindromic repeat
CRM	Cis-regulatory module
dAp cell	Dorsal apterous cell
DV	Dorsal ventral
ESC	Embryonic stem cell
FFL	Feed forward loop
FMRFa	FMRFamide (Phe-Met-Arg-Phe-NH ₂)
GFP	Green fluorescent protein
GMC	Ganglion mother cell
GRN	Gene regulatory network
NB	Neuroblast
Nplp1	Neuropeptide-like precursor 1
PCD	Programmed cell death
PIC	Preinitiation complex
RA	Retinoic acid
RARE	Retinoic acid response element
RNA Pol II	RNA polymerase II
TF	Transcription factor
TFBS	Transcription factor binding site
Tv cell	Thoracic ventral cell
UAS	Upstream activating sequences
VNC	Ventral nerve cord

Introduction

General Aspects of Early Embryonic Development

Embryonic development in animals is an intricate process, generating specific tissues with correct orientation and dimension, at a specific location and time. In addition, a general body plan, with respect to axis formation; anterior-posterior, dorsal-ventral and medial-lateral, has to be established in bilaterian animals. Since all cells in a multicellular organism share a common progenitor and therefore the same genetic material, as such they possess in theory the possibility to become any cell type of the organism. Yet the cells comprising a multicellular system show an astonishing diversity and specificity. During the first steps of zygotic divisions during vertebrate development, the resulting daughter cells are totipotent up to the 8 cell stage, and after that, in the morula/blastula stage, the cells become pluripotent. At the blastocyst stage the cells from the inner cell mass are considered to be pluripotent, and are known as embryonic stem cells (ESCs). Totipotent cells are defined as being able to develop into a whole organism, whereas pluripotent cells have a slightly limited potential but still can give rise to most cell types or tissues of an organism [1]. Once the embryo enters late gastrulation the three germ layers i.e., ectoderm, mesoderm and endoderm are formed and the cells become multipotent. At this stage the cells are determined to create any cell type within their specific germ layer [2]. As the development of an organism progresses, the cells start to differentiate and specify into specific cell types, which results in a loss of developmental plasticity. As early as in 1892, August Weismann described that during embryonal development, certain “determinants” in the cells are inactivated or even lost during progression of cell fate determination [3]. Today we know that while the zygote undergoes multiple divisions to generate an embryo and later an organism, the gene expression profiles of the cells constituting the organism change over time, by either activation or inactivation of certain genes crucial during specific steps in development and later in cell fate specification. Once cells are terminally differentiated and committed to a certain cell fate, those cells are unlikely, under normal conditions, to revert into progenitor type stem cells or to dedifferentiate into another terminal differentiated cell type. However, cell ablation experiments in the mouse airways demonstrated that fully differentiated epithelial cells could revert back to stable and functional basal stem cells [4]. Studies on heart and limb regeneration present further evidence that under challenging conditions, differentiated cells can revert into progenitor type cells in order to support tissue regeneration [5-7]. Waddington described this sequential restriction of cell fate by using a famous metaphor; a marble rolling down a valley making binary decision between different bifurcations (developmental

pathways), while the ridges between the valleys describe the increasing irreversibility of cell type differentiation (Fig. 1) [8].

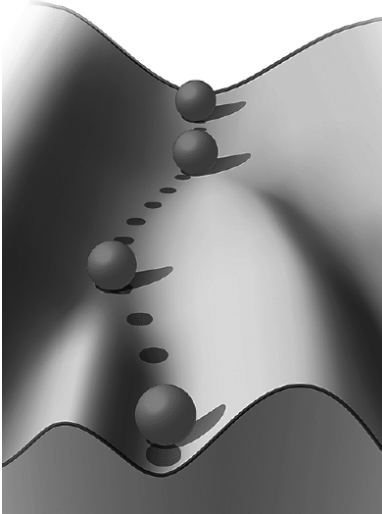


Figure 1: Waddington's Epigenetic Landscape describing the stepwise restriction of developmental plasticity

Changes in gene expression occurring during development are the result of extra-cellular, cell-to-cell, or intracellular signals which in turn both activate and interact with transcriptional programs that lead to cell fate specification [9]. Cell-to-cell signalling could be described as one group of cells acting to affect another group of cells. In this scenario, signal producing cells are termed “inducers” and the cells receiving a certain signal are called “responders” [10]. Signalling pathways involved in development fall into 11 main categories, defined by their ligand and or signal transducers: Notch, fibroblast growth factors (FGF), epidermal growth factor (EGF), Wnt/Wingless (Wg), Hedgehog (Hh), transforming growth factor (TGF- β)/bone morphogenetic protein (BMP), cytokine (nonreceptor tyrosine kinase JAK-STAT), Hippo, Jun kinase (JNK), nuclear factor (NF- κ B) and Nuclear Hormone Receptors. The signal transduction can be mediated either by cell-to-cell contact, via cell surface proteins, or by secreted diffusible differentiation factors [11].

Molecular Aspects of Development

While the previous section described the general aspects behind development, this section will briefly describe the molecular mechanisms involved. The aforementioned signalling molecules, crucial for developmental progression, do not act alone to control cell fate specification. Rather, those molecules often act in concentration dependent gradients, typically establishing spatial identity, and are involved in complex cascades of downstream events, that ultimately result in cell differentiation. Some of the molecular signalling cascades might even regulate each other at some point in development, in order to confer the correct cell type identity. How can those signals translate into cell fate? Signalling molecules can bind to their specific receptors and thereby regulate small numbers of transcription factors (TFs) or transcriptional co-factors. TFs in turn can bind directly to the DNA and act as activators or repressors of transcription, to thereby control cell type specific gene expression. Some of the signalling molecules involved in the early embryo development, with regards to germ layer specification are maternally loaded, which means that those molecules are not induced after fertilization, but instead are already present in the unfertilized oocyte.

Examples of Molecular Events Involved in Germ Layer Formation

Studies of the development of the aquatic frog *Xenopus laevis* showed that maternal determinants like the VegT (T-box) transcription factor, together with TGF- β signalling, are crucial for meso- and endoderm formation. In addition, maternally inherited Ectodermin (Ecto) is pivotal in formation of the ectoderm, and its major role is to attenuate the action of TGF- β in order to ensure an ectoderm identity [12]. BMP signalling plays a critical role in the specification of the ectoderm (Fig. 2). While BMP activates a cascade of downstream regulators, it positively controls epidermal genes via the TFs, Msx1 and Xvent2. Simultaneously, BMP signalling restricts the expression of pro-neural genes, and thereby prevents the specification of the ectoderm towards neuroectoderm fate. In order to facilitate neural specification of the ectoderm, the BMP activity needs to be constrained. Indeed the factors Chordin (Chd), Noggin and Follistatin antagonize BMP signalling via inhibition of the Smad1 molecule, one of the downstream facilitators of the BMP pathway. Simultaneously, FGF acts to activate the pro-neural genes *Sox2* and the neuronal cell-adhesion molecule *Ncam*, thus acting to promote neural induction[13].

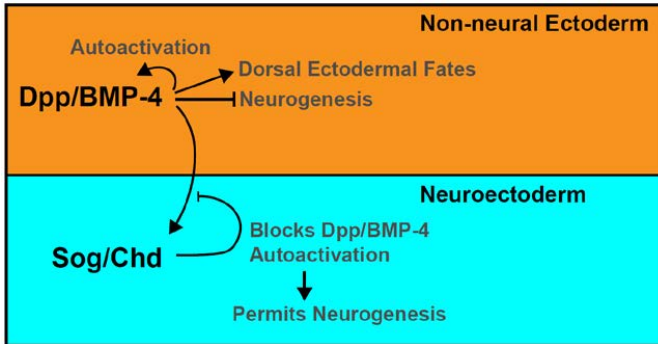


Figure 2: Antagonistic action of Sog/Chd and Dpp/BMP-4 during ectoderm specification in invertebrates and vertebrates (adapted from Bier, 1997).

Examples of Molecular Events Involved in Body Patterning

Later in vertebrate development, during axis formation of the anterior-posterior axis (AP patterning), other signalling molecules and TFs are involved in the patterning of the embryo. Positional identity of embryonic tissues is to some extent controlled by the vitamin A derived signalling molecule retinoic acid (RA), and studies in the chick limb bud suggest involvement in AP axis development. Retinoic acid acts as a ligand for nuclear retinoic acid receptors (RARs, nuclear receptor superfamily), and binding to the receptors changes their role from transcriptional repressors into transcriptional activators, which interact with the DNA in a direct manner [14]. Application of exogenous RA during early embryonic development, caused severe malformations, particularly in the hindbrain and branchial regions of the head [15]. Another critical class of factors involved in the AP patterning are the *Hox* homeotic genes, which encode for a class of nuclear TFs containing a conserved DNA binding homeodomain. *Hox* genes in vertebrates comprise a class of 38 genes, organized in four separate chromosomal clusters, named *HoxA*, *-B*, *-C*, and *-D*. *Hox* genes are expressed in gradients along the AP axis and control spatially restricted gene expression, depending on the location of their specific expression. Studies in cell culture experiments and transgenic mice could also show that the *Hox* genes contain retinoic acid response elements (RAREs) in their gene regulatory regions, to which the activated retinoic acid receptors (RARs) can bind, and thereby control *Hox* gene expression (Fig. 3) [16]. Once activated, the *Hox* TFs can bind the DNA of their target genes and control their expression. DNA binding can result in the activation of cell specification cascades and thereby contribute to terminal cell fate. Establishment and maintenance of a certain cell fate is ultimately controlled by the stable expression of transcription factor networks acting in concert to specify cell identity.

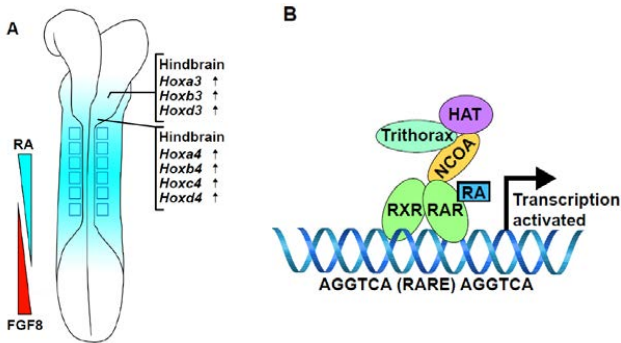


Figure 3: (A) Retinoic acid (RA) gradient in the developing mouse embryo; (B) showing RA interaction with the retinoic acid receptor (RAR), activating transcription (adapted from Cunningham and Duester, 2015).

Terminal Cell Fate can be reversed

Based on the work done in *Xenopus* it was shown that a “locked down” state of terminally differentiated cells could be reversed in nuclear transfer experiment. Nuclei from non-dividing somatic cells injected into unfertilized frog oocytes, whose own nuclei had been destroyed, resulted in a low number of dividing zygotes with DNA replication, progressing into viable tadpoles. This indicates that the environment of the oocyte could reverse the stable restriction of developmental capacity by inducing DNA replication and transcription, as well as cell division [17]. Still, at this point little was known about the specific factors involved and responsible for the induced reprogramming from a somatic cell fate back to a stem cell state. In 2006, Yamanaka and Takahashi showed, that a combinatorial expression of the four TFs Oct3/4, Sox2, c-Myc and Klf4 ("Yamanaka cocktail") was sufficient to reprogram mouse embryonic fibroblasts into an embryonic stem cell state, thus that the terminal differentiated status of somatic cells could be reversed [18].

The Fruit Fly *Drosophila melanogaster* as a Model Organism

For more than a 100 years, *Drosophila* has been used as a model organism to study development and gene regulation [19], and since then has contributed tremendously to the deeper understanding of general biological processes in higher organisms. For instance, the previously described *Hox* genes were first described and characterized in *Drosophila* by Ed Lewis [20] and later on in the lab of Walter Gehring [21]. Other discoveries made in *Drosophila* include the work of Christiane Nüsslein-Volhard and Erik Wieschaus on the genetic mechanisms behind segment polarity and the complex spatial organization of the body plan in higher organisms [22]. In 1995, Lewis, Nüsslein-Volhard and Wieschaus shared the Nobel

Prize for their work. Still today, *Drosophila* is used extensively to study many basic biological mechanisms involved in development, disease and behaviour. Even with the advent of induced pluripotent stem cells, which can be directed into desired cell types [23], together with lab grown human cerebral [24] or intestinal [25] organoids, the information retrieved from the use of model organisms is unmatched. Since induced pluripotency based systems represent highly artificial conditions, which cannot mimic the complex nature of the in vivo context underlying development, those models are limited in their capacity to explain fundamental biological mechanisms. Because *Drosophila* shares more than 75% of disease genes with humans [26] it is widely used as a model to study for instance Alzheimer disease [27] and other neurodegenerative diseases, as well as alcohol addiction [28], haematopoiesis [29] and cancer [30]. Based on the high degree of evolutionary genetic conservation to humans, many genes relevant for development or disease in humans, have orthologues in *Drosophila* [26, 31]. Additionally, many molecular genetic tools have been developed to introduce mutations at certain positions in the genome (*P* elements, TALENs, CRISPR/Cas9), or to integrate human disease genes. More recently, site-specific landing-site transgenesis allows for very precise structure function analysis [32]. Furthermore, introduction of fluorescent gene coupled reporter constructs are used label certain regions or cell types in the developing embryo and the adult *Drosophila*. The rapid generation time (10 days from egg to fertile adult) (Fig. 4), comparably low maintenance costs, absence of ethical constrains, availability of the whole genome sequence, presence of many visible genetic markers, such as stubble bristles, curly wings and different eye colours, make *Drosophila* an ideal model organism for studying diverse aspects of biology.

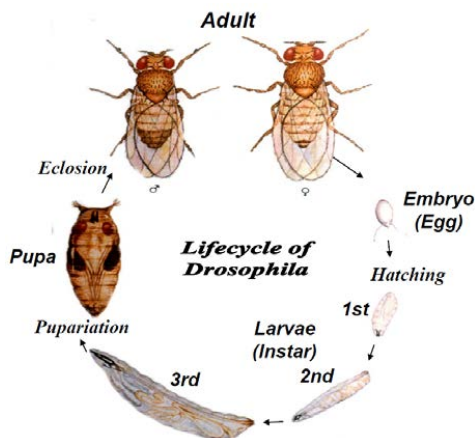


Figure 4: The *Drosophila* life cycle takes approximately 10 days from egg to adult.

***Drosophila* Embryogenesis**

Drosophila embryogenesis is a well-established system for investigating many different aspects of development, due to several practical properties. The development of the *Drosophila* embryo is a rapid process and takes about 22 hours from egg to larvae. In addition, a uniform grading system for developmental stages has been established by Campos-Ortega and Hartenstein [33]. *Drosophila* embryos are transparent, which allows microscopic investigation down to single-cell resolution [34], either by immunohistochemistry on fixed tissue or via live imaging and the use of fluorescent transgenic markers [35]. Large abundance of available embryonic material per experiment, permits high through-put *in situ* hybridization [36] but also high-throughput transcriptomic experiments [37].

Early *Drosophila* Embryonic Development

As the *Drosophila* oocyte is fertilized, it undergoes several rounds of division. However, in contrast to the vertebrate systems, the cleavage of the *Drosophila* embryo is nuclear or superficial, as it is in other insect species (Fig. 5). Because the nuclei divide in a common cytoplasm without cytokinesis, these nuclear divisions result in the formation of a syncytium with an increasing number of nuclei. The cleavage process lasts for two hours and 13 consecutive divisions with an average cell cycle length of 8-10 minutes. The initial nucleus of the oocyte is located centrally in the egg and the first daughter nuclei remain in the centre, for the first five divisions [38]. The nuclei then arrange at the egg surface, while cell membranes are build up around the nuclei. This stage is termed the cellular blastoderm. During cycle 10, located at the posterior position in the embryo, 34-37 pole cells are formed, which are later involved in the germ line formation. During progression of development the cells located at the pole caps and mid ventral part of the blastoderm invaginate. During this process the three germ layers are formed. While the cells which remain at the surface become the ectoderm, the invaginating cells give rise to the meso- and endoderm [39].

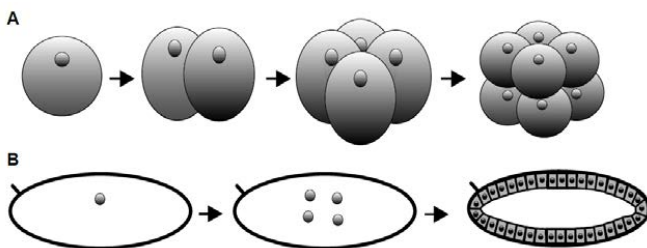


Figure 5: (A) Total (holoblastic) cleavage and (B) superficial nuclear cleavage (insects) of the fertilized egg.

Drosophila embryonic development is initially controlled by maternally inherited factors which build up gradients along the anterior posterior (AP) axis. Those early factors are involved in the subsequent segmentation of the developing embryo and include the genes *bicoid* (*bcd*), *caudal* (*cad*) and *nanos* (*nos*). Bicoid contains a homeodomain and thereby can interact with the DNA of its downstream targets, thus acting as a transcription factor. Its messenger RNA (mRNA) is localized at the anterior part of the early embryo, forms an anterior-posterior gradient, and is involved in the formation of the anterior segmented patterning of the head and thorax [40]. Abdominal formation is under control of Nanos, which acts as a repressor for *hunchback* (*hb*), which in turn is a repressor of posteriorly expressed gap genes. The Caudal protein forms a gradient of reverse polarity to Bicoid and is important in the formation of the posterior segments. The initial action of the maternal genes *bicoid*, *caudal* and *nanos* sets in motion a downstream cascade of transcription factors, which in turn act to further subdivide the embryonic AP axis. In brief, this cascade of hierarchically expressed transcription factors includes the gap genes, pair-rule genes, segment polarity genes and homeotic genes. Sequential expression of those factors subdivides and specifies the embryo into a series of segments (Fig. 6) [41].

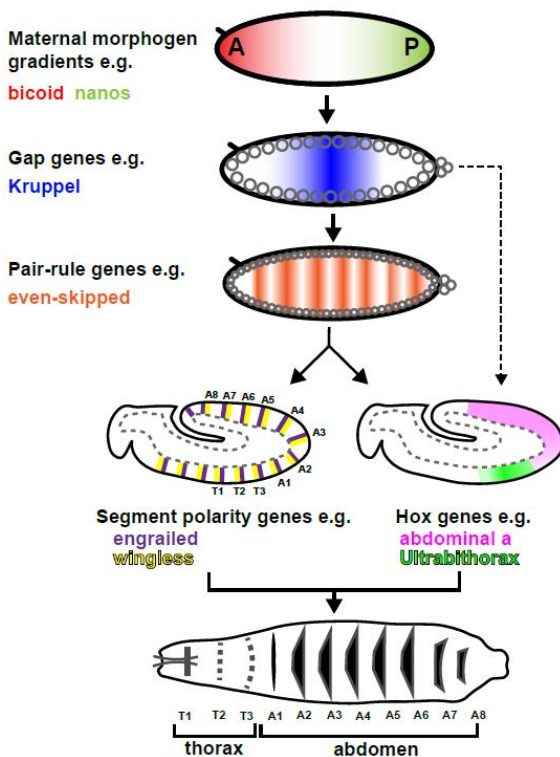


Figure 6: AP patterning in *Drosophila*. Early maternal factors activate the gap genes in discrete compartments in the embryo. Gap genes will then activate the pair-rule genes and to some extent the *Hox* genes. The pair rule genes activate the segment polarity and *Hox* genes. Segment polarity genes control intra segmental identity, while the *Hox* genes control segment fates along the AP axis (adapted from Sanson, 2001).

In addition to patterning along the AP axis, the developing embryo will undergo patterning along the Dorsal-Ventral (DV patterning) axis. The specification of dorsal and ventral regions of the embryo depends on a graded Dorsal (DI) (vertebrate NF- κ B) signal along the DV axis. The *dorsal* mRNA is maternally loaded in the oocyte, translated after fertilization and ubiquitously distributed in the embryo. This broad distribution alone cannot account for the specification of dorsal and ventral regions. Another crucial factor in DV patterning, *gurken* (*grk*) (EGF-like domain) is expressed at the dorsal region of the oocyte. It interacts with the receptor Torpedo (epidermal growth factor receptor, EGFR), which leads to an inhibition of the *pipe* (*pip*) gene in the dorsal region. Activation of *pip* in the ventral region of the embryo, results in the activation of *spatzle* (*spz*). Subsequent cleavage of the Spz protein enables it to bind as a ligand to the Toll receptor, which in turn results in translocation of the DI protein from the cytoplasm into the nucleus (Fig. 7) [42]. Once in the nucleus DI acts as a transcriptional activator for the genes *twist* (*twi*), *snail* (*sna*) and *rhomboid* (*rho*). Both *twi* and *sna* are essential for ventral mesoderm development, while *rho* is involved in the specification of the presumptive neuroectoderm. DI is an activator of mesoderm determinants, but simultaneously acts as a transcriptional repressor for genes crucial in dorsal ectoderm development i.e. *zerknüllt* (*zen*) and *decapentaplegic* (*dpp*) (vertebrate BMP4) [43]. In brief, an intricate interplay of several factors, acting as transcription factors, ligands and receptors, create a gradient of nuclear localized DI protein along the DV axis, with high levels ventrally and low levels dorsally. This results in a gradual subdivision of the embryo into the ventral mesoderm, more dorsally the mesectoderm and neuroectoderm and most dorsal the epidermis and amnioserosa [41].

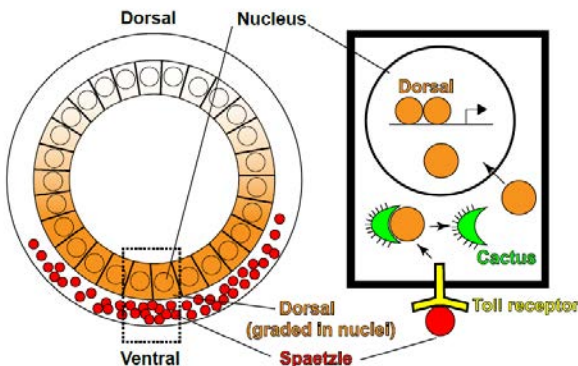


Figure 7: Binding of Spz to the Toll receptor results in segregation of DI and Cactus in the cytoplasm. This leads to subsequent uptake of the liberated DI protein into the nucleus where it acts as a transcription factor.

Neural Development in *Drosophila*

The *Drosophila* central nervous system (CNS) develops from a bilaterally symmetric sheet of neurogenic ectoderm which is located laterally in the developing embryo, separated by the mesoderm. The lateral neuroectoderm later gives rise to the brain and the ventral nerve cord (VNC) [41]. Invagination of the mesoderm during gastrulation joins the lateral neuroectoderm, forming the CNS at the future midline (mesectoderm) of the embryo (Fig. 8) [44].

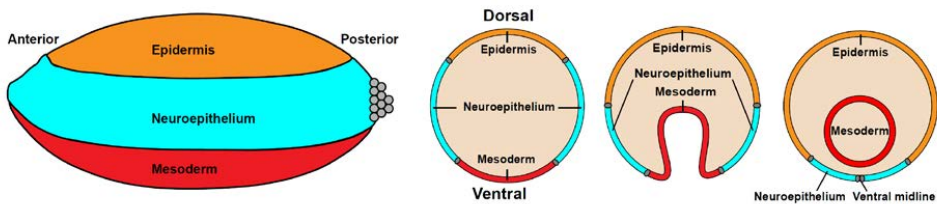


Figure 8: Developing *Drosophila* embryo, prior, during and after invagination of the mesoderm resulting in the formation of the ventral neuroepithelium (adapted from Bier, 2008).

The brain and VNC in the *Drosophila* embryo (Fig. 9A) in principle correspond to the vertebrate brain and spinal cord [45]. The VNC is subdivided along the AP axis into the subesophageal segments (S1-S3) followed by three thoracic segments (T1-T3) and ten abdominal segments (A1-A10). Based on the bilateral constitution of the VNC, each segment can be further subdivided into hemisegments. Each hemisegment gives rise to an almost invariant number of ~30 neuronal stem cells, the neuroblasts (NBs), with each having a unique identity based on position and marker expression [46]. Since the NBs generated during development are arranged in invariant positions in each hemisegment, they form a grid-like pattern of rows and columns. This pattern is comparable to a cartesian coordinate system, therefore NBs are termed according to their fixed position in a particular hemisegment (Fig. 9B). As an example, NB5-6T is located in row 5 and column 6 of a hemisegment, while T denotes that this particular NB is part of a thoracic hemisegment. NB5-6A on the other hand, relates to the equivalent NB in an abdominal location. After neurogenesis is complete, the developing *Drosophila* CNS contains about 5,000 cells in the brain and 800 cells per segment in the VNC, which are the result of NB and daughter cell divisions [47]. Therefore cell numbers

are significantly lower when compared to the human brain, which is estimated to contain about 86 billion neurons [48]. Segmental identity of the *Drosophila* CNS is regulated by similar mechanisms involved in the patterning of the general body plan. The early segmentation factors *bicoid*, *nanos* and *caudal* act on a cascade of downstream factors to achieve AP patterning. Together with opposing gradients of Dorsal and Dpp/BMP4 signalling along the DV axis, this results in the early patterning of the developing *Drosophila* embryo. High concentration of nuclear D1 in the ventral region of the embryo specifies the mesoderm (somatic muscle, heart). Moderate levels of D1 towards the more dorsally regions specifies ectoderm (ventral epidermis, CNS) [49].

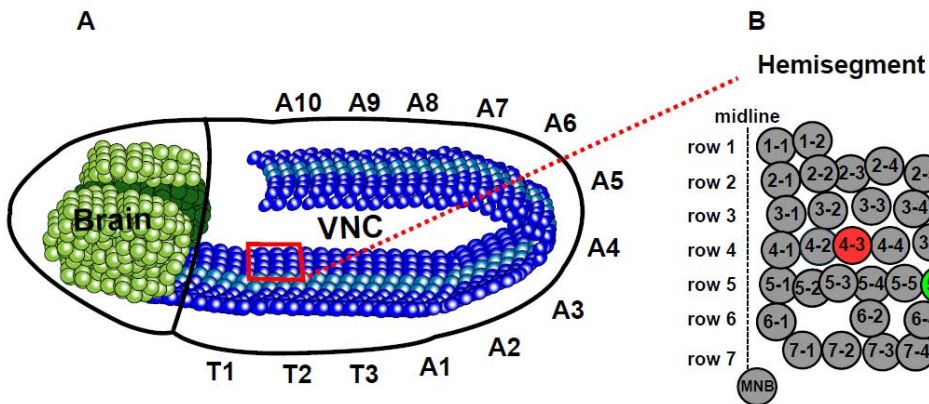


Figure 9: **(A)** The CNS of the developing *Drosophila* embryo can be subdivided into the brain and ventral nerve cord (VNC). The VNC is subdivided at the midline into reiterating hemisegments along the AP axis. In total the brain contains roughly 5,000 cells and the VNC about 10,000 cells. **(B)** One hemisegment, showing the position of the NBs. In red is the NB4-3 which gives rise to the dorsal Apterous cells (dAp) and in green is the NB5-6 which gives rise to the Apterous cluster cells (Ap cluster) in the thoracic region of the VNC. (NB-map adapted from <http://uoneuro.uoregon.edu/doelab/nbmap.html>).

From Ectoderm to Neuroectoderm

Opposing gradients of Dorsal and Dpp/BMP-4 act to subdivide the ectoderm into epidermis and neuronal ectoderm. The Dpp/BMP pathway is conserved and shows similar properties from *Drosophila* to humans. The primary role of Dpp in the early phase of neural tissue specification is to prevent ectodermal cells to acquire a neural fate in the dorsal-most region of the embryo. Dpp expression represses neural genes and maintains the expression of genes crucial for ectodermal specification. Dpp transcriptionally activates the dorsally acting gene *zerknüllt* (*zen*) and *dpp* itself, and represses the proneural genes of the *acheate-scute* complex AS-C, which are involved in neural development [44]. Opposing to *dpp* expression in the dorsal-most region, high levels of Df in the ventral-most region of the early embryo regulate the expression of mesoderm determinants *twi* and *sna*. Moderate levels of Df in the lateral regions of the ectoderm activates the neurogenic genes *rho*, *intermediate neuroblasts defective* (*ind*) and *ventral nervous system defective* (*vnd*). Lower concentrations of Df towards the dorsal region of the embryo, activate the gene *short gastrulation* (*sog*) (vertebrate *chordin*), which attenuates Dpp and allows those cells of the ectoderm to acquire a neural fate [50]. In a simplistic view, high levels of Dpp/BMP signalling in the *Drosophila* epidermis represses expression of neurogenic factors. Simultaneously, repression of Dpp/BMP by antagonists in the neuroectoderm, together with activation of neurogenic factors by Df, permits the ectoderm to acquire a neuroectodermal fate (Fig. 10).

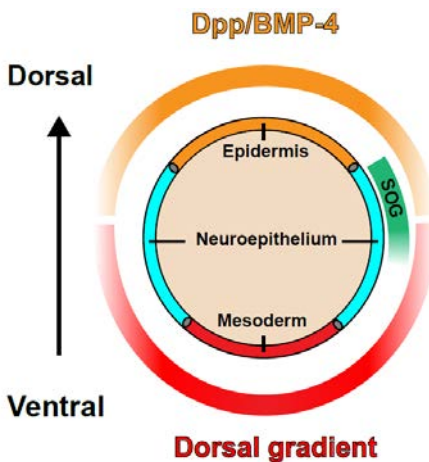


Figure 10: Schematic representation of the antagonistic action of Dorsal and Dpp, controlling the development of the mesoderm, neuroectoderm and epidermis.

Anterior Posterior Axis Formation of the Nervous System

As CNS development progresses, the neurogenic ectoderm is gradually subdivided into segments along the AP axis in order to confer spatial identity. Since the nervous system is of ectodermal origin, patterning in of the future CNS depends on the same regulatory cascade of factors which pattern the general body axis (see Fig. 6) [51]. The expression of the early maternal morphogens such as *bicoid*, *caudal*, *nanos* regulates a downstream cascade of factors involved in the generation of repeated units along AP axis. Maternal morphogens first regulate the expression of the gap genes including *oskar* (*osk*), *hunchback* (*hb*), *Kruppel* (*Kr*), *knirps* (*kni*), *giant* (*gt*), *tailless* (*tll*), *unpaired* (*upd*) and *hopsotch* (*hop*). The gap genes are expressed at defined positions along the AP axis and cross-regulation between these genes results in the establishment of relatively precise spatial domains along the axis. In the segmented regions, another group of genes, the pair rule genes, are activated almost simultaneously with the gap genes. The pair rule genes include the factors *even-skipped* (*eve*), *hairy* (*h*), *run* (*run*), *fushi-tarazu* (*ftz*), *paired* (*prd*), *odd-paired* (*opa*), *odd-skipped* (*odd*) and *sloppy-paired* (*slp 1 and 2*). Expression of the gap genes and pair rule genes is a transient process and declines during gastrulation, and thus can be described as pre-patterning. In order to provide and maintain an intra-segment identity, after gastrulation another group of genes, the segment polarity genes, are expressed within the segmented regions along the AP axis. The segment polarity genes include the factors *engrailed* (*en*), *gooseberry* (*gsb*), *wingless* (*wg*), *armadillo* (*arm*), *cubitus-interruptus* (*ci*), *fused* (*fu*), *hedgehog* (*hh*), *naked* (*nkd*), *patched* and *dishevelled* (*dsh*) [52, 53]. During the progressive subdivision of the AP axis, graded expression of the *Hox* genes along the AP axis also acts to provide segment identity [54].

Dorsal Ventral Axis formation of the Nervous system: Columnar patterning

During neurogenesis, the ectoderm is specified into the neurogenic ectoderm by a graded signal of nuclear D1. Simultaneous expression of the previous described patterning factors subdivides the neuroectoderm into metameric segments along the AP axis. In addition to AP patterning, the neuroectoderm is subdivided along the medial lateral axis into mainly three longitudinal parallel columns. This subdivision of the neuroectoderm is achieved during DV patterning, by the restricted expression of the columnar genes. The columnar genes include the homeobox factors *ventral neuroblast defective* (*vnd*), *intermediate neuroblast defective* (*ind*) and *muscle specific homeobox* (*msh*; *Drop* [Dr]). The ventral neuroectodermal column expresses *vnd*, the

intermediate neuroectodermal column expresses *ind* and the dorsal neuroectodermal column expresses *msh*. While high concentrations of nuclear D_I specifies the presumptive mesoderm, via the mesoderm specific factors *twi* and *sna*, gradually lower concentrations towards the dorsal epidermis activates genes involved neuroectoderm specification. Evidence suggests that D_I can directly activate the columnar genes *vnd* and *ind* in a concentration dependent manner [50], whereas *msh* is suggested to be activated by low levels of ubiquitous Dpp/BMP/TGF- β signalling, as observed in activation of *msh* homologues in zebrafish [55]. Inhibition of *ind* by *vnd*, prevents the ventral column to acquire an intermediate neuroectodermal fate. Both *vnd* and *ind* repress *msh*, which prevents the intermediate and ventral columns to specify into a dorsal neuroectodermal fate, yet *msh* does not repress *vnd* or *ind*. High concentration of Dpp however, results in repression of *msh*. This indicates that the sharp border of *msh* expression is established from bidirectional repression, ventrally by *vnd* and *ind* and dorsally by high levels of Dpp [56]. Thus, repression between the columnar factors which are expressed in their defined domains, helps to establish the sharp boundaries between neuroectodermal columns (Fig. 11).

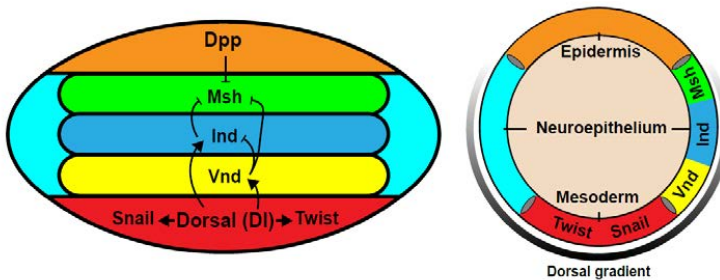


Figure 11: The D_I gradient activates the columnar genes in a concentration dependent fashion. High levels of D_I activates the mesodermal genes *twi* and *sna*. Moderate levels of D_I activates Vnd and Ind. Msh is not activated by D_I but thought to be activated by low levels of TGF- β signaling. Both Vnd and Ind together with high levels of Dpp can repress Msh.

Segment-polarity patterning

DV and AP patterning cues act to subdivide the developing CNS into a segmented pattern. Another group of factors, the segment-polarity genes, act to subdivide each segment along the AP axis. These factors are involved in the specification of different AP rows inside the segments and thereby critical for positional determination of the correct NB identity (Fig. 12). The segment-polarity genes include the factors *engrailed* (*en*), *gooseberry* (*gsb*), *wingless* (*wg*) (Wnt-1 homologue), *armadillo* (*arm*), *cubitus-interruptus* (*ci*), *fused* (*fu*), *hedgehog* (*hh*), *naked* (*nkd*), *patched* (*ptc*) and *dishevelled* (*dsh*). Mutation of segment-polarity genes results in abnormal CNS development, and can also result in NB identity switches evident by altered marker expression in the NB. As an example, *wg* is expressed specifically in row 5 NBs, and secreted Wg protein can be detected in adjacent NBs of row 4 and 6/7. Wg signalling acts in row 4 and specifies the NB4-2 identity, which give rise to RP2 neurons. Mutants for *wg* exhibit problems in NB4-2 specification, and in the case that the NB forms, it acquires an NB3-2 identity according to its marker expression [57]. The segment polarity gene *gsb* is expressed in NBs of row 5 and 6. Mutants of *gsb* show a transformation of NB5-3 into a NB4-2 identity, and hence give rise to RP2 neurons. In *wg* positive NBs of row 5, *gsb* represses the competence of the NB to receive the Wg signal, and thus prevents those NBs from acquiring a row 4 identity. Row 4 NBs on the other hand, express the segment polarity gene *ptc* which acts as a repressor for *gsb*, hence allowing for the *wg* signal to specify NB4-2 identity. This regulation can therefore explain the NB4-2 identity acquisition of NB5-3 in *gsb* mutants, which hence stems from de-repressed *wg* signalling [58].

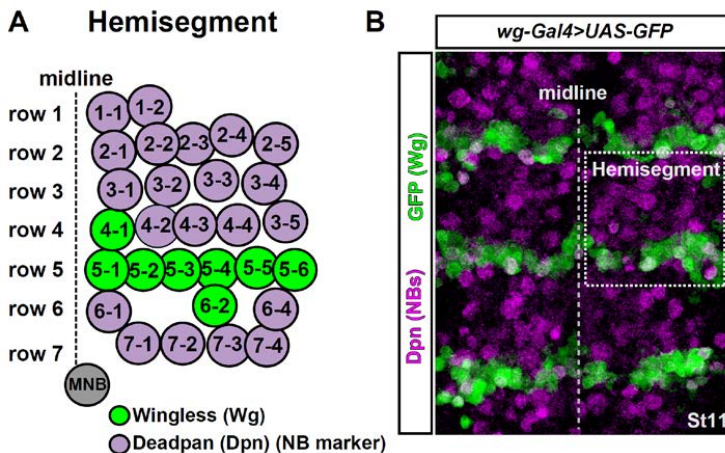


Figure 12: (A) Expression of *wingless* (*wg*) in specific NBs. (B) Immunostaining shows the restricted expression of *wg* in the hemisegments of a *Drosophila* VNC, (green showing *wg-Gal4>UAS-GFP*, magenta showing Deadpan (Dpn) to mark NBs.

Hox Genes in the CNS

Homeotic (Hox) genes or the homeotic (HOM-C) complex, comprise a group of evolutionary conserved transcription factors that play a central role during AP patterning in invertebrates and vertebrates, but also to generate cell type diversity during development [16, 20, 45]. Hox genes were first discovered in *Drosophila* [59] and one of the most well-known Hox genes in *Drosophila* is probably the *Antennapedia* (*Antp*) gene, which when overexpressed in the antenna primordium, turns the antenna into a leg [60]. The HOM-C complex in *Drosophila* contains two separate clusters which are grouped into genes of the *Antennapedia Complex* (*Antp-C*) and the *Bithorax Complex* (*Bx-C*) [61]. The Hox genes of the *Antp-C* in *Drosophila* include *labial* (*lab*), *proboscipedia* (*pb*), *Deformed* (*Dfd*), *Sex combs reduced* (*Scr*) and *Antennapedia* (*Antp*) while the *Bx-C* contains the genes *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*) and *Abdominal-B* (*Abd-B*). In vertebrates, the Hox complex contains 38 genes, which are organized in four different clusters which emerged from an ancestral complex via gene duplication [62]. Hox genes in vertebrates are termed *Hox-a/1*, *Hox-b/2*, *Hox-c/3* and *Hox-d/4* (Fig. 13).

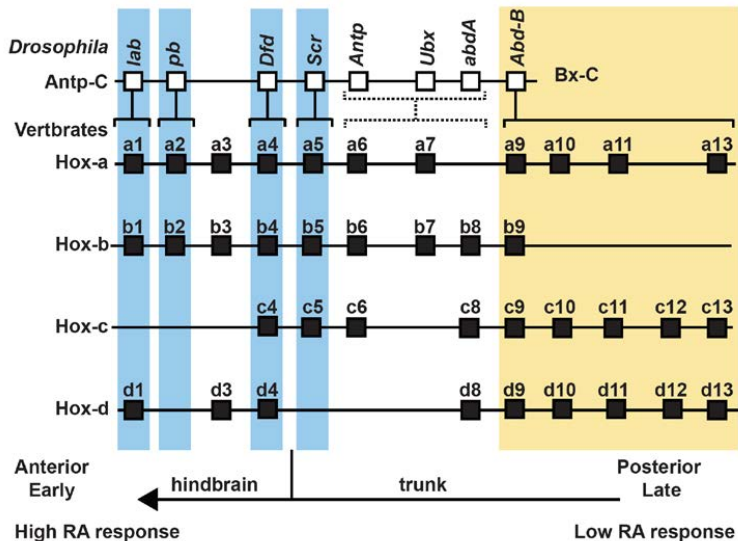
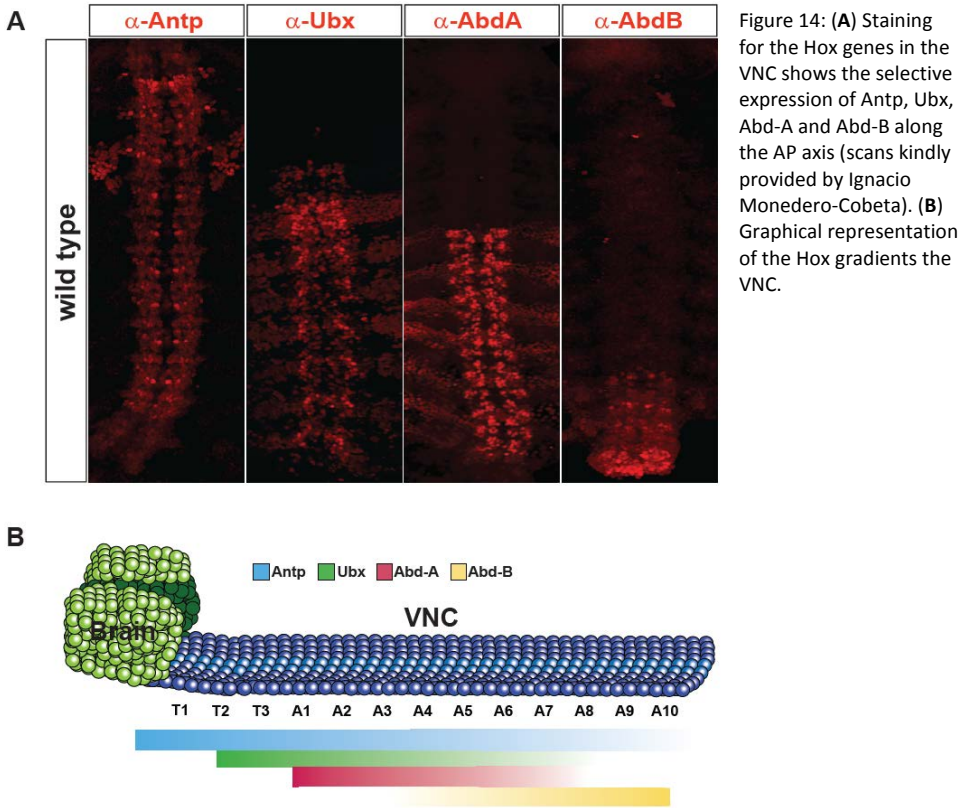


Figure 13: Hox genes in *Drosophila* and vertebrates (adapted from Krumlauf, 1994).

At the molecular level homeotic genes contain a DNA homeobox sequence, which encodes a homeodomain (HD) at the protein level. The homeodomain allows these factors to bind to DNA via sequence specific binding onto TAAT sites (HD core binding motif), to thereby regulate gene expression of their target genes [63]. Furthermore, the Hox factors increase sequence specific DNA recognition by binding cooperatively with cofactors, mainly the HD proteins Extradenticle (Exd) (vertebrate PBX1/4) and Homothorax (Hth) (vertebrate MEIS1/2/3) [64]. In addition to the classical Hox homeotic genes, many other genes, such as *ladybird early (lbe)*, *apterous (ap)*, *POU domain protein 2 (pdm2)*, *paired (prd)*, *tailup/islet (tup, Isl)* and *bicoid (bcd)* [40, 65-68], contain homeobox DNA sequence motifs, thus encoding for a homeodomain. Yet not all factors containing a homeobox DNA motif are homeotic (Hox) genes, and are rather referred to as homeobox genes.

During VNC development Hox genes are expressed in gradients along the AP axis, and contribute to the specification of the segments into morphologically distinct regions (Fig. 14). Their expression is controlled by upstream regulators such as the early segmentation genes and genes of Polycomb/Trithorax group. Hox genes expressed at the posterior-most part of the VNC repress more anterior Hox genes (posterior prevalence). Early Hox gene expression in the neuroectoderm overlaps with the segmental expression of the embryo, but declines after NB delamination. Later, Hox genes are re-expressed in individual cells, NBs and postmitotic neurons, and take part in combinatorial codes, essential to specify cell type identity [45].

Studies performed on the NB5-6 showed the importance of Hox gene influence on NB behaviour regarding the lineage size and specification. Present in both the thoracic and abdominal hemisegments, NB5-6 generates lineages with differences in cell numbers and fates. NB5-6 in the thoracic region (NB5-6T) generates a lineage of ~20 cells [69], of which four cells specify into Apterous neurons (Ap cluster). Ap cluster cells are absent in the lineage produced by the abdominal NB5-6A. Evidence was presented that the truncation of the NB5-6A lineage was the result of an earlier cell cycle exit and apoptosis compared to NB5-6T. This earlier exit was controlled by the Bx-C factors together with the Hox co-factors Exd and Hth. Mutation of the *Bx-C* genes resulted in the formation of larger lineages and generation of Ap cluster cells along the entire AP axis. Thus the NB5-6A acquired a NB5-6T identity, which is controlled by the Bx-C factors [70].



Neuroblast delamination, Lateral inhibition

As outlined above, during CNS development a stereotyped number of NBs form at defined locations and time points. What are the mechanisms that decide if a cell in the neurogenic ectoderm should become an NB or to stay in an epidermoblast state? Studies have shown that this is gated by the *Notch* pathway during a process called lateral inhibition. Expression of the columnar genes *vnd*, *ind* and *msh* in the ventral, intermediate and lateral column of the neuroectoderm, together with the expression of the segment-polarity and pair-rule genes in the transverse rows in each hemisegment directly acts to control the expression the proneural genes. The proneural genes include the factors *achaete* (*ac*), *scute* (*sc*) and *lethal of scute* (*l'sc*) and are summarized to the *Achaete-Scute Complex* (*AS-C*). A combination of these genes are expressed in 10 groups of cells per hemisegment with each group containing 6-8 cells, so called equivalence groups. The AS-C factors activate the expression of the Delta ligand which in turn activates the Notch receptor in the neighbouring cells. Notch activation will result in the

repression of the proneural AS-C factors, via factors of the *Enhancer-of-split* (*E(spl)*) gene complex, which results in inhibition of the NB fate (Fig. 15). Notch pathway mutations result in ectopic production of NBs. Once the NBs are specified, they will delaminate from the neuroectoderm in five subsequent waves termed S1-S5. The first wave of S1 NBs derives from the medial and lateral columns of the neuroectoderm, closely succeeded by S2 and S3 NBs which delaminate from the intermediate and medial or lateral columns respectively. S4 and S5 NBs derive from the intermediate columns of the neuroectoderm [71].

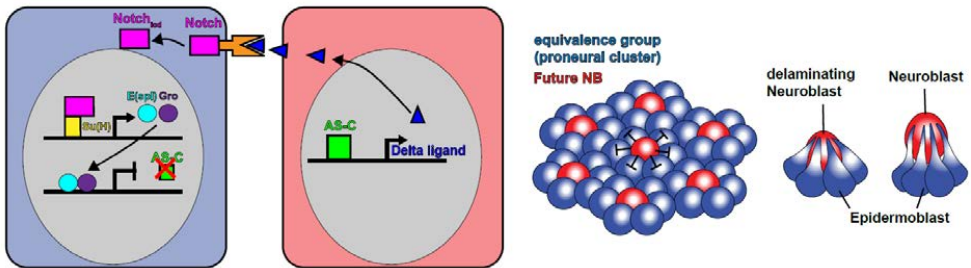


Figure 15: The process of lateral inhibition via Delta and Notch, results in the specification of neuroblasts (NBs).

Lineage Trees

By the time that the NBs delaminate from the neuroectoderm they possess a unique identity based on their position, time of birth and gene expression, which will govern that they give rise to different numbers of progeny, of different neural sub-type fate. NBs delaminating early, during S1-S3, can generate lineages with between 10 to more than 20 daughter cells, while NBs which segregate during S4-S5 produce lineages between 2 to 10 cells. Due to unique NB identity, lineages produced by different NBs show variability regarding the cell types they generate. For instance, the NB4-2 neuroblast generates the RP2 motoneuron, NB7-3 generates a pair of serotonergic neurons, and NB1-1 gives rise to three abdominal glial cells [72]. Each NB undergoes a stereotyped number of asymmetric divisions to renew itself and to generate daughter cell with a reduced developmental potential. In the *Drosophila* VNC, two types of division modes are apparent, type I and type 0. In the type I division mode, the NB will divide to generate a NB and a ganglion mother cell (GMC), which can in turn divide once more to generate two neurons and/or glia cells. In the type 0 division mode, the NB divides and gives rise to a NB and to one directly born neuron. Some NBs generate lineages of daughter cells,

which are born in different division modes, and thus undergo a type I \rightarrow 0 division mode switch [73]. NB5-6T, which delaminates at developmental stage 8 and gives rise to 20 cells, will first divide nine times in the type I mode, and then five times in the type 0 mode (Fig. 16).

The NB asymmetric division is controlled by an extensive set of asymmetric determinants. Proteins which are part of the PAR/aPKC complex (Bazooka [Baz], atypical protein kinase C, DaPKC [DmPAR6]), are distributed to the apical pole (future NB) [74], whereas the scaffolding protein Miranda (Mira) is localized at the basal pole (future GMC) of the dividing NB. Mira is associated with at least three different repressive factors; the translational repressors Staufen and Brain Tumor (Brat) and the transcriptional repressor Prospero (Pros). During NB division, the different protein complexes are distributed asymmetrical into the newly formed cells. In the GMC, Mira gets degraded which results in the release and nuclear localization of Pros. In the nucleus, Pros acts as a repressor for progenitor and cell cycle genes and in turn initiates neuronal differentiation [75].

NB divisions, and consequently lineage size with regards to cell numbers is highly regulated, and thus each NB stops dividing at a defined stage. Studies of the NB5-6T showed that the NB undergoes 14 asymmetric divisions and goes into apoptosis at stage 16 [69]. The cell cycle regulator *dacapo* (*dap*) (mammalian *Cdkn1a,b,c*; p21^{CIP1}/p27^{Kip1}/p57^{Kip2}), plays a critical role to control the division mode switch and cell cycle exit. Mutants for *dap* show an extended type I daughter division mode, and extended NB proliferation. Further studies of NB5-6T demonstrated that the temporal gene *castor* (*cas*) and the Hox gene *Antp*, in combination with the late temporal factor *grainy head* (*grh*), activate Dap and repress the cell cycle factors E2f and Cyclin E, and by that control the type I \rightarrow 0 daughter proliferation switch and NB cell cycle exit [76].

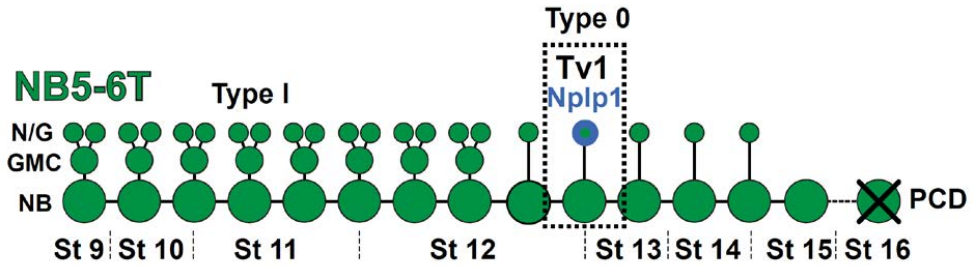


Figure 16: The NB5-6T lineage. The NB divides asymmetrically in a type I division mode, to give rise to a ganglion mother cell (GMC) and to renew itself. The GMCs divide once more to give rise to two neurons and/or glia cells. During lineage progression, the NB undergoes a division mode switch to type 0 and divides to give rise to direct born neurons and to renew itself. At St 15 the NB exits cell cycle, and undergoes apoptosis at St16

Apoptosis/PCD

Apoptosis, or programmed cell death (PCD) is another mechanism for shaping the VNC and controlling lineage size. PCD affects lineage size in three ways: It can act to terminate NB divisions, by killing the NB, by removing early postmitotic cells, or by removing cells that have undergone differentiation [76, 77]. It is estimated that during development 30% of the cells in the CNS undergo PCD and die [78]. The postmitotic PCD can either occur immediately after a cell is born, without overt signs of differentiation, or after a cell has differentiated into a neuron or glia. An example of the first type of PCD stems from studies of NB7-3, where specific new-born cells in this lineage immediately undergo PCD [79]. An example of the latter, stems from studies on the neuropeptidergic dMP2 neurons in the *Drosophila* VNC. This showed that dMP2 neurons are generated in all segments of the VNC, and extend axons that serve a critical axon scaffold function, but are subsequently removed by PCD in all segments anterior to A6-A8. Attenuation of cell death, by mutation of the cell death activators *reaper* (*rpr*) and *grim*, lead to survival of the dMP2 neurons along the entire VNC. This effect was also achieved when the VNC posterior-most expressed Hox factor *Abd-B* was misexpressed in the whole VNC [80]. Taken together, these experiments show the critical role of PCD in controlling lineage size, by stopping NBs divisions, by removing postmitotic cells immediately, or after differentiation.

Temporal Cascade and Sub-Temporal Factors

Asymmetric NB division, cell cycle exit control and PCD act in combination to shape lineage size. However, those mechanisms do not directly contribute to cell fate and neuronal diversity, and do not account for the stereotyped birth order of different neuronal cell types at different stages of lineage progression. Analyses of developing NBs in the *Drosophila* VNC have identified a molecular mechanism of temporal specification, in which a series of TFs is expressed sequentially in the NBs as they progress through development. These TFs are denoted temporal factors, and include Hunchback (Hb), Kruppel (Kr), Pdm (Nubbin/Pdm1 and Pdm2), Castor (Cas) and Grainy head (Grh). Their stereotyped expression patterns are restricted to sequential periods during NB development, and are termed temporal windows or temporal competence windows. Progeny born in a certain temporal competence window inherit or maintain the expression of the temporal factor which contributes to terminal cell fate (Fig. 17A). Most if not all NBs undergo this sequential temporal competence transition irrespective of the time of NB delamination, and even isolated NBs, cultured *in vitro*, undergo a temporal transition [81, 82]. How is the transition from one temporal window into the next one controlled? Studies reveal that the temporal factors cross-regulate each other, by activation of the next direct downstream TF and repression of the previous TF, together with forward repression of the next TF, downstream of the TF which gets activated (Fig. 17B) [81, 83]. The temporal cascade is initiated by the expression of HB, a zinc finger TF, and the down-regulation of Hb allows for the expression of Kr and subsequent progression of the temporal cascade. Misexpression of *hb* in NB7-1 and NB7-3 resulted in the generation of ectopic early born neurons, in the absence of late born neurons [84]. The COUP-TF transcription factor Seven up (Svp), was identified to be critical to attenuate *hb* expression in NB7-1 and NB7-3, and thus being involved for the switch from *hb* to *Kr* expression in order to allow for temporal progression of the NBs. Early misexpression of *svp*, resulted in the loss of early born neurons, which depend on the early temporal factor *hb* [85]. During NB5-6T development the late temporal gene Cas does not only act to activate Grh, but also activates a cascade of factors, which act to subdivide the large Cas temporal window into more restricted competence windows, via the so called “sub-temporal” genes. The sub-temporal cascade includes the factors *squeeze* (*sqz*) a Kruppel C2H2-type zinc-finger protein, and its co-factor *nab*, a NGFI-A-binding protein and Svp (Fig. 17B). During NB5-6T development, this NB generates four direct born Apterous neurons out of which two neurons become neuropeptidergic. Svp is only expressed in the neurons which are non-neuropeptidergic. Targeted misexpression of Svp in

the Apterous neurons could demonstrate that Svp attenuates the neuropeptidergic cell fate [86]. Thus Svp prohibits the two non-neurogenic interneurons to acquire a neuropeptide cell fate. The example of Svp shows that this factor has dual roles; early, to attenuate Hb to ensure proper temporal cascade progression, and late, to attenuate neuropeptide cell fate. In our work we recently discovered that the temporal gene Kr has such a dual role as well, and that late postmitotic re-activation of Kr via Cas, represses the action of Svp to promote neuropeptidergic cell fate. Therefore, Kr can be considered a sub-temporal gene as well. In summary, temporal genes can be described as timing devices, allowing NBs to change their competence to give rise to different daughter cells as the NBs progress through development.

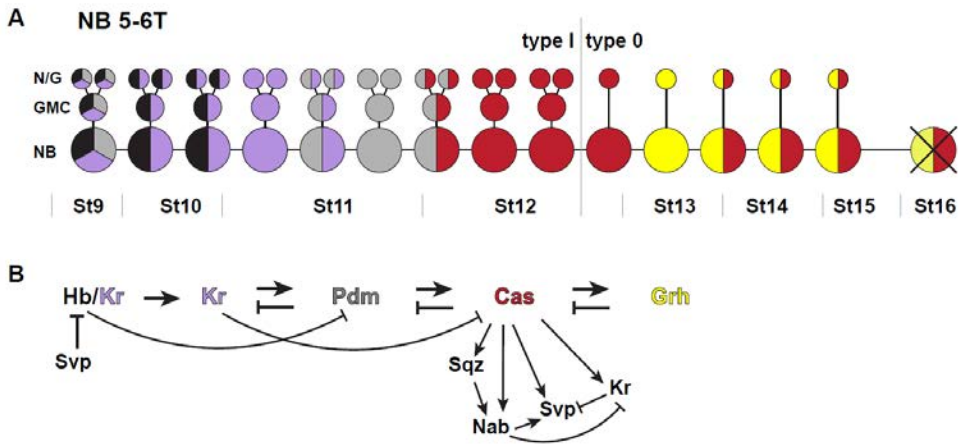


Figure 17: (A) The NB5-6T lineage during temporal progression, indicating that the temporal factors expressed in the NB can be inherited to their progenies. (B) Interactions of the temporal factors during development of the NB. Svp and Kr both show re-onset during later development of the NB5-6T lineage, in the sub-temporal cascade and contribute to cell fate.

Similar Regulation in Higher Organisms

Many of the aforementioned mechanisms are, though not identical, generally conserved into more derived organisms. For instance, the neural ectoderm development is under control of *short gastrulation (sog)*/Chordin which repress *decapentaplegic (dpp)*/BMP-4 activity, in both *Drosophila* and vertebrates. As well as in *Drosophila*, AP patterning in vertebrates is controlled by the Hox genes. However the Hox gene expression in vertebrates is controlled by gradients of retinoic acid, which interacts with retinoic acid response elements (RAREs) in the Hox gene adjacent DNA sequence. Dorsal ventral patterning in *Drosophila* is established by the

expression of *vnd*, *ind* and *msh*. Studies in mouse found the expression of *vnd* homologues *Nkx2.1*, *Nkx2.2* and *Nkx2.9* in the ventral neuroectoderm. The *ind* homologues *Gsh1* and *Gsh2* are expressed in the intermediate neuroectoderm, and the *msh* homologue *Msx* was found to be expressed in the dorsal neuroectodermal domains [87]. In *Drosophila*, neurogenic progenitors are generated by Notch mediated lateral inhibition of the proneural genes in the neuroectoderm. The exact mechanism in vertebrates are less well understood. Studies in *Xenopus* demonstrated that primary neurons are formed by an initial wave of neurogenesis. This process is controlled by proneural gene related factors *neurogenin* (*neuro D3*) and Notch pathway homologous *X-Notch-1* and *X-Delta-1*. Ectopic expression of *neurogenin* results in formation of additional primary neurons, while ectopic expression of *X-Delta-1* reduces or eradicates primary neuron formation [88]. While *Drosophila* NBs express a stereotyped cascade of temporal TFs, Hb>Kr>Pdm>Cas>Grh during development, experiments on mouse retinal neural progenitors revealed a temporal expression of the zinc-finger type factor *Ikaros* and *casz1* (orthologs of *hb* and *cas*). *Ikaros* and *casz1* are expressed in retinal neural progenitors and involved in specification of cell types during early and mid-retinogenesis [89]. Studies of the mammalian cerebral cortex showed that neuronal progenitors first divide symmetrically to expand the progenitor pool. Dividing progenitors can give rise to direct born neurons or daughter that divide once to generate two neurons, but also undergo self-renewing divisions which gives rise to intermediate neural progenitors (INPs). INPs can then divide symmetrical to give rise to two neurons and two more progenitors. An INP-like division mode has been observed in the *Drosophila* brain [90], and hence both mammals and *Drosophila* display three similar modes of daughter divisions: Direct-born neurons (type 0), daughters that divide once (type I), or daughters that divide multiple times (denoted INP in mammals and Type II in *Drosophila*). Apoptosis or PCD during vertebrate neuronal development is a well-established system to eliminate excess cells and it is suggested that 50% of all neurons undergo cell death. PCD in vertebrates is controlled by many different genes, including BMP-4, Wnt signalling and Sonic hedgehog (Shh) [91].

Model: Apterous cluster, Tv1 and dAp cells

The studies on which this work is based upon, focus on the specification of a subset of neurons which when terminally specified express the Neuropeptide-like precursor 1 (Nplp1). In this section I will therefore describe the model used to study the various aspects cell fate specification in more detail, with regards to location of the cells and marker expression.

Nplp1 is selectively expressed in distinct subsets of neurons (28 out of 10,000 neurons) in the *Drosophila* VNC, which makes the Nplp1 expression a powerful marker to address cell fate specification. Cells expressing Nplp1 include the dorsal Apterous cells (dAp) and the Thoracic ventral (Tv1/Ap1) cells of the four cell Ap cluster (Ap1/Tv1/Nplp1, Ap2/Tv2, Ap3/Tv3, Ap4/Tv4/FMRFa) (Fig. 21). The Ap clusters are located ventrolaterally of the VNC in the thoracic segments T1-T3. The dAp cells constitute a set of bilateral dorsal-medial located neurons, distributed along the VNC from T1 to A8. Both the dAp and Tv1 cells project their axons ipsilateral anteriorly and join a common Ap fascicle. The Ap cluster cells are generated by the NB5-6T and are born sequentially between stage 13 and 15 in a direct type 0 division mode in a Cas/Grh mixed temporal window. The dAp cells are generated early in stage 11 by NB4-3 and are born in a type I division mode in a Pdm temporal window.

All four Ap cluster cells and the dAp cells express the LIM-homeodomain TF Apterous (Ap, mammalian Lhx2a/b) and the transcriptional co-factor Eyes absent (Eya, mammalian EYA1-4). By stage AFT (air filled trachea) the Tv1 and dAp cells will also express Collier (Col, mammalian Ebf1-3) and Dimmed (Dimm, mammalian Mist1). Dimm is important to activate *PHM* which is critical to amidate neuropeptides [92]. By stage AFT, the Tv4 cells of the Ap clusters are specified into another neuropeptidergic cell and express FMRFamide. The Tv4 cells also express Ap/Eya/Dimm together with phosphorylated Mothers against dpp (pMad, vertebrate SMAD1-3/5/9, ZNF396), Dachshund (Dac, vertebrate DACH1/2), Squeeze (Sqz vertebrate ZNF384) and the co-repressor Nab (vertebrate Nab1/2), but lack Col. The Tv2/3 cells stay in a “generic” interneuron state, express Ap/Eya/Sqz, Nab (just Tv3) and the COUP-TF Seven-up (Svp, nuclear hormone receptor; vertebrate NR2F1/2/6), but lack Col and neuropeptide expression [69]. Nplp1 expression and therefore Tv1 and dAp specification both depend on the expression of a shared terminal selector cascade involving Col>Ap/Eya>Dimm>Nplp1.

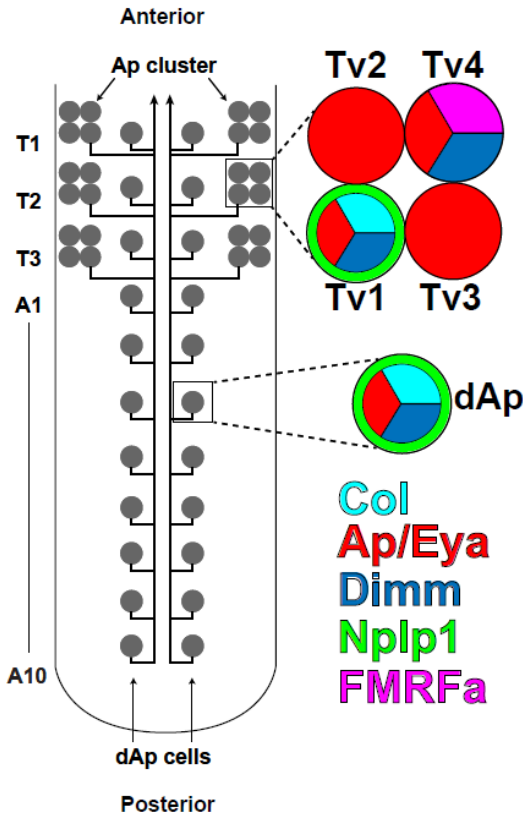


Figure 21: Model of the *Drosophila* VNC, with focus on the Apterous cluster and the dorsal Apterous (dAp) cells. The Ap cluster consist of four neurons out of which the Tv1 cell expresses the neuropeptide Nplp1 together with Col, Ap/Eya and Dimm. The dAp cells express the same marker combination. Both the Tv1 and dAp cells, project their axons ipsilaterally along a common Ap fascicle anterior towards the brain.

Regulatory Networks

On the simplest level, gene regulation can be described by the interaction of one TF which binds to a gene regulatory sequence, to either activate or repress a gene. Factor A regulates factor B. Generation of cellular diversity in metazoans however is a multistep process, which requires more complex regulatory networks (GRNs) to direct cell type specific gene expression. Nevertheless, gene regulatory networks motifs found in bacteria and yeast, are also found in plants and animals. Hence, these regulatory motifs can be transferred from bacteria or yeast, with some variation of the numbers of factors involved, to describe cell type specific gene regulation in higher organisms. Network motifs most commonly used to control gene expression are different versions of feedforward loops (FFLs). An activating FFL could be

described as follows: Factor A activates factor B and in turn factor A and B act in combination to activate factor C; thus factor A feeds forward on factor C. This type of FFL is termed coherent type 1 FFL or C1-FFL. Another commonly occurring regulatory motif is the incoherent type 1 FFL or I1-FFL. In this scenario, factor A activates B, A feeds forward on factor C as an activator, whereas B acts on C as a repressor (Fig. 20) [93].

Coherent type 1 FFL (C1 FFL) **Incoherent type 1 FFL (I1 FFL)**

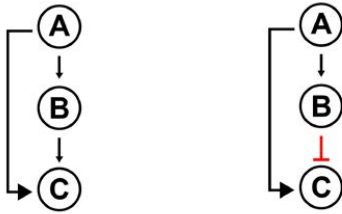


Figure 20: Two types of feedforward loops (FFLs) are most common in gene regulatory networks. The coherent type 1 FFL, which is purely activating, and the incoherent type 1 FFL, which combines activation and repression (adapted from Alon, Cell Snapshot, 2010).

This type of FLL can act to control levels of gene expression or activate a certain gene in one cell but not in another. Transcription factors part of GRNs can therefore act to refine broad input signals and dictate cell type specific gene expression [94, 95]. Numerous studies in *C.elegans*, *Drosophila* and mammals demonstrate that cell specification depends on the combinatorial actions of more than one TFs, so called “combinatorial codes” of TFs to generate a distinct cell fate [96-100]. Hence, unique cell fate identity does not depend solely on single TF input [101]. These combinatorial codes can include purely activating information and resemble a C1-FLL or contain activating and repressing signals and hence resemble an I1-FLL. Studies on the Nplp1 expression in the *Drosophila* VNC, identified combinatorial TF codes of both types, the C1-FLL and I1-FLL, to be involved in cell fate specification but also cell type diversification [69, 102]. Studies of vertebrate sensory neuron specification identified combinatorial codes of the I1-FLL type, to be involved to diversify neuronal subtypes [103].

Based on *Drosophila* development, TFs involved in cell specification can be categorized into spatial selectors, temporal selectors, tissue/cell type selectors and terminal selectors. Spatial selectors include TFs which define regional identity, spatial limits and control developmental programs of multipotent progenitors along the VNC axis, e.g. Hox genes, columnar genes and segment polarity genes. Temporal selectors include TFs which define the temporal limits of developmental capacity of multipotent progenitors inside a cell lineage e.g. temporal factors which are expressed during NB progression. Tissue/cell type selectors,

comprise factors which induce a genomic response in progenitor cells to trigger a tissue or cell type specific fate in daughter cells. While there exist TF networks involved in the tissue specification critical in the compound eye development, there are no TFs functioning as tissue/cell type selectors, necessary and sufficient in the specification of “generic” pan neuronal tissue. Terminal selectors include TFs which are selectively and postmitotically expressed. These factors are critical to direct the expression of effector genes, which determine the final cell type and fate i.e. expression of neurotransmitters, neuropeptides or ion channels [101]. The TFs expressed postmitotically in the Ap cluster cells which activate the expression of *Nplp1* *Col>Ap/Eya> Dimm* are categorized as terminal selectors. Hence, the combinatorial *C1-FLL*, which activates the *Nplp1* expression can be termed a terminal selector cascade.

Identification of Candidate Genes and Relations between Different Factors

Stereotyped postmitotically expressed TFs, have been extensively used as markers to monitor cell fate decision i.e., loss or gain of either one of the markers or combinations of markers. For instance, *FMRFa* is expressed in just six cells in the thoracic region as part of the Ap cluster. In a recent published screen, an *FMRFa-GFP* transgene was used to monitor EMS induced mutations. This led to the characterization of many genes and their alleles with regards to their role in VNC development, among which were the factors *ladybird early (lbe)* and *svp* [104]. The expression of *lbe* in particular is very intriguing because it starts to express in the early NB5-6 along the VNC and is maintained in parts of the NB5-6 postmitotic lineage. Usage of a *lbe*-reporter construct *lbe(K)-GFP/lacZ*, allows to specifically label the NB5-6 [105], and because of the persistence of the reporter protein the whole NB5-6T lineage including the Ap clusters. In addition, *lbe* mutants show a specific loss of *Col* in the NB5-6T and consequently a loss *Ap/Eya/Dimm* and *Nplp1*.

The factors identified during the mutant screen are important for *FMRFa* cell fate specification, hence have a specific role in the transcriptional network of factors acting during cell specification. How can be determined at which position a factor acts in a regulatory network and what role it plays? Usage of the Gal4/UAS system (Materials and Methods) allows misexpression of single factors or combinations of factors in a tissue or cell type specific fashion. This system also allows to reintroduce a factor into its own mutant background to “rescue” its mutant phenotype, or to introduce a factor which is under control of the gene which is mutated to “cross-rescue” its mutant phenotype.

In a scenario in which a mutant $A^{-/-}$ affects the expression of B and C, reintroduction of A into an $A^{-/-}$ mutant rescues the expression of B and C. Introduction of B into an $A^{-/-}$ mutant also restores the expression of C; B can cross rescue mutant $A^{-/-}$. This would indicate that A acts on B and B acts downstream of A to activate C. What if the reintroduction of B into mutant $A^{-/-}$ does not cross-rescue the expression of C? In this scenario it is most likely that A acts on B and further acts in combination with B to activate C. Another possibility to test for the specific role of factors is misexpression. Single misexpression of A could result in increase of B and C. Misexpression of B results in increase of C but not A. This indicates that B acts downstream of A to activate C. If co-misexpression of A and B results in stronger increase of C, than the misexpression of either factor A or B alone, then this could indicate cooperative action of A and B. Therefore, rescue and cross-rescue together with single factor and combinatorial misexpression experiments, are used to identify positions and roles of specific factors inside a gene regulatory network.

Gene Regulation

Embryonic patterning, subdivision of the neuroectoderm, NB delamination and terminal differentiated cell fate are ultimately the result of highly coordinated tissue and cell type specific gene expression. Gene expression can be controlled on multiple levels i.e., transcriptional, posttranscriptional, translational and/or posttranslational. My thesis has focused on transcriptional regulation, and the following sections will therefore address the molecular aspects of gene regulation with regards to transcriptional control. During gene transcription, a decision is made whether to start transcribing a gene or to leave it in a silent state. Since most cells in multicellular organism share the same genome, transcriptional regulation is a critical step and will decide about the initiation of gene expression. In order to achieve cellular diversity in an organism, gene expression needs to be restricted to only a subset of all available genes of the genome in each specific cell type or tissue. While gene expression can be regulated by extrinsic cues such as signalling molecules, for instance Notch mediated lateral inhibition prior NB delamination, or retinoic acid signalling during patterning in vertebrates, it is not the signalling molecule that directly confers cell identity. Signalling molecules rather act as mediators through binding to their receptors. In a few cases, such as for Notch, activated receptors can then act directly as TFs and bind to the DNA to activate or repress a gene, under the condition that the gene regulatory region has the correct binding

motifs. More frequently, signal transduction will affect TF activity in an in-direct, second-messenger manner [53]. In addition, TFs which act in concert to regulate gene expression can also be regulated by intrinsic cues, such as the temporal genes during NB development described above. During NB5-6T development, the late born Apterous cluster cells are generated in a Cas/Grh temporal window. Cas encodes a C2H2 zinc-finger type factor and acts as a TF, activates the expression of a downstream cascade of different TFs which are indispensable to regulate cell fate specification [66, 69, 94, 106]. Consequently, every aspect of a functional cell is linked to the gene expression products in that cell, and the sum of all genes expressed in a cell defines its identity. Transcriptional regulation plays a central role in this differential gene expression.

General Structure of a Gene

In order to better understand the molecular aspects of gene regulation, I will in brief describe the general structure of a gene with focus on the gene regulatory components crucial during transcriptional control. The basic molecular structure of a gene entails: Enhancers, silencers, promoter, exons and introns. The exons contain the coding and non-coding (5' and 3' un-translated) regions of the gene (Fig. 18) [107]. While the introns get removed from the transcribed RNA during further processing steps via splicing, the exons remain part of the mRNA, which gets translated into a protein product. The enhancer, silencer, promoter, introns and 5'-3' un-translated regions will not be translated into proteins, and thus fall into the non-coding part of the DNA. Yet, those elements contain crucial information about, when and where as well as how much transcript should be produced during transcription [108]. Promoters are located around the transcriptional start site of a gene (+1 TSS) and consist of short ~100-200bp long DNA stretches [109]. The promoters contain core promoter sequences, which are involved in the recruitment of RNA polymerase II (Pol II) and components of the pre-initiation complex (PIC) including transcription factor IID (TFIID), which contains the TATA binding protein (TBP) and TBP-associated factors (TAFs). One, almost regarded as universal, core promoter element is the TATA-Box (TATAAA; DNA binding motif), which is located -27bp upstream of the TSS [110]. Yet the TATA-box is present in just about 10-20% of metazoan core promoters. However, promoters contain other core promoter elements, such as the initiator (Inr) or downstream promoter elements (DPEs), which are recognized by different combinations of TFs [111]. Besides promoters, enhancers and silencers play a critical role in gene regulation. Enhancers or cis-regulatory modules (CRMs) are regulatory DNA stretches

which can vary in length between several hundred base pairs (core enhancers) up to several thousand bps [112]. Their main function is to integrate activating signals, mediated by TFs which bind to defined regions in the enhancers (transcription factor binding sites, TFBS) and by that activate the promoters of their corresponding genes. In contrast, silencers can be bound by TFs, which mediate repressive signals. In comparison to promoters which reside close to the TSS of the gene they regulate, enhancers can be scattered across the locus of their corresponding genes and can be found in close proximity or several thousand bp up- or downstream of the TSS, or in intronic regions of genes. In addition, CRMs may not necessarily regulate the closest located genes but can bypass those, to regulate more distant genes. Furthermore, CRMs are independent of orientation, in order to be regulatory active [113]. CRMs can be referred to as integration hubs, which funnel upstream information onto promoters to regulate specific gene expression.

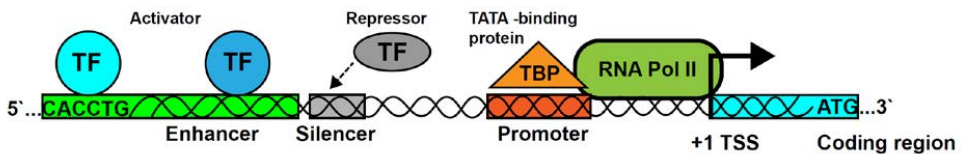


Figure 18: General structure of a gene regulatory region, showing the enhancer bound by activating TFs, a silencer, the promoter bound by the TATA binding protein, which is involved in recruitment of RNA polymerase II (Pol II), and the coding region of the gene.

Regulation of Gene Expression: Transcriptional Control

Gene expression at the transcriptional level is controlled by regulatory DNA elements that are not part of the final protein product. The promoter region is essential for the recruitment of the transcriptional machinery, including the Pol II complex, which will under the right circumstances start to transcribe the corresponding gene. Because it binds widely expressed factors, the promoter region on its own does not typically contribute to spatial or temporal selectivity with regards to gene expression, but will in a simplistic view regulate how much of a transcript should be produced [114]. Therefore, promoters need additional information to regulate their transcriptional activity and subsequent gene expression. Spatio-temporal information is provided by enhancers or cis-regulatory modules (CRMs). The first enhancer was discovered about 30 years ago in the locus of simian virus SV40, a 72 bp DNA sequence, which enhanced reporter gene expression by 200 fold, thus the term enhancer. Since then

enhancers have been subject to extensive studies to understand their regulatory potential during gene expression in development [115-117]. How do enhancers regulate spatio-temporal activation of promoters, and how do enhancers provide information to the promoter?

To answer the first part of this question, I will describe the function of enhancers. Enhancers contain multiple specific DNA elements, so called transcription factor binding sites (TFBS) which are ~6-12 base pairs long motifs [112]. Pioneer transcription factors, for instance GATA [118], Oct4, Sox2, Klf4 and c-Myc, can bind to condensed and nucleosome enriched DNA [119] and by that can open up the condensed chromatin structure in which enhancer elements reside. In this open state, cell or tissue specific TFs can recognize and bind to their specific motifs. TF binding can occur in combination with co-factors, and/or as a combinatorial “cocktails” of different TFs at multiple motifs [120]. Binding of sequence specific TFs can act to prevent repositioning of nucleosomes, thus cell or tissue specific TFs can serve as placeholders and keep the chromatin in an accessible state (Fig. 19A) [112]. However, TFs, in particular pioneer factors, can also bind to inaccessible chromatin and recruit chromatin remodelling complexes, such as the SWI/SNF complex, which depletes the DNA from nucleosomes and thus makes the DNA accessible for other TFs to bind [121]. Besides recruitment of pioneer factors and sequence specific TFs, Pol II and parts of the preinitiation complex (PIC) are recruited to and assembled on the accessible enhancer [122, 123]. Simultaneous recruitment of histone modifying factors such as Histone Acetyltransferases (HATs, CBP/p300), acetylate histones and by that destabilize as well increase the flexibility of the chromatin. Up to this point, the role of cell and tissue specific TFs is to keep the enhancer chromatin in an accessible state, which allows for recruitment of the transcriptional machinery to activate cell or tissue specific gene expression.

To answer the second part of the question; how enhancers deliver their information, I will describe the enhancer-promoter interaction. Since enhancers can be several kb away from the promoter and gene they regulate, they have to engage in physical contact with the promoters. One broadly accepted model is that the destabilized and flexible enhancer/chromatin structure loaded with TFs and the transcriptional machinery can loop over to the promoter region, and by that span large distances. Evidence for enhancer looping was presented by chromosome conformation capture experiments (3C and its variations). Chromosome conformation capture experiments, which detect the 3D organization of the genome, could demonstrate that enhancer and promoter regions, separated by several kb DNA sequence, are located in close proximity after looping has occurred. Connections formed via

proteins between the enhancer and promoter regions, such as cohesin and the Mediator complex, are involved in stabilizing such loops [123, 124]. Experiments on the chromosomal 3D structure identified topology associated domains (TADs), regions in the genome characterized by high frequency of interactions. The regions containing TADs are enriched for CCCTC-binding factor (CTCF) binding sites. Enriched binding sites for CTCF inside TADs is believed to facilitate enhancer looping to the promoter, by CTCF recruitment of cohesin [125]. Furthermore, it is believed that loop formation delivers components of the transcriptional machinery from the enhancer to the promoter, and thus increases the concentration of components critical for transcription at the promoter (Fig. 19B) [113]. Taken together, just in those cells or tissues in which the pioneer factors and sequence specific TFs are present, the enhancer is kept in an open/active state, loops over to the promoter and activates gene expression of a downstream target in a highly specific fashion.

Because of their importance in facilitating cell- and tissue-specific gene transcription, much research has been focused on identification and characterization of enhancer elements with regards to their genomic location, length and TF binding motif content. A recent study in *Drosophila* embryos predicts an abundance of ~41.000 active enhancers, and an average regulation of one protein coding gene by four enhancers. However, the *Drosophila* genome is estimated to harbour in total ~50.000-100.000 developmental enhancers [126]. Enhancer detection studies based on epigenetic modifications estimated an abundance of 0.4-1 million enhancers in the human genome [127].

More detailed enhancer studies focus on DNA binding, demonstrating that TFs can bind to a variety of different DNA motifs. For instance, Hox homeodomain factors with a glutamine at position 50 in the homeodomain (so-called Q50 homeodomains), typically recognize a 6 bp sequence with a TAAT core motif [63, 64]. Hox homeodomain binding is promoted by the Exd/Hth (vertebrate orthologue Pbx/Meis) Hox cofactors. Studies of the COE factor Collier (Col, Knot; vertebrate orthologue Ebf1-3), showed a preferred binding to CCCNNGGG or TCCNNNNNGGA palindromic motifs [128-130]. In addition, evidence points to that the flanking regions of core binding motifs, together with the spacing between the TFBS, are critical for enhancer activity, specificity and robustness. Furthermore, enhancers can contain suboptimal binding sites, which result in imperfect binding of TFs to their specific motifs. Suboptimal, low affinity motifs were demonstrated to increase enhancer specificity at the expense of TF binding affinity. Changing those imperfect binding sites to high affinity motifs, resulted in ectopic enhancer activity [64, 131, 132]. In order to sustain robust gene

expression, enhancers controlling early developmental gene expression, can in addition have a functionally redundant second enhancer, which regulates the same genes as the primary module. These redundant elements are termed “shadow-enhancers” because they reside in the remote “shadow” of their proximal counterpart, and can be located several thousand kb away from the primary enhancer [133-135].

Enhancer elements have been known to exist and to contain valuable information which promotes specific gene expression for over three decades. The recent advent of next generation sequencing (NGS) techniques revealed that enhancers play an even more important role in regulating gene expression, besides just functioning as integration hubs for TFs. As described previously, not only are promoters associated with Pol II, but so are active enhancers, which results in enhancer transcription and generation of eRNAs [136]. During transcription initiation, Pol II is phosphorylated at different serine residues (chiefly Ser2 and Ser5) of the C-terminal domain or CTD. Phosphorylation of Ser5 is mainly associated with promoter-proximal paused Pol II (herein referred to as poised Pol II), whereas a Ser2 phosphorylation is associated with active elongation [137]. First the phosphorylation of the negative elongation factor (NELF) and CTD Ser2 of Pol II results in subsequent elongation. Phosphorylation of NELF and Ser2 is catalysed by the cyclin dependent kinase 9 and 12 (CDK9 and CDK12) [138]. Recently it was shown that also eRNAs can associate with NELF which results in a subsequent release from Pol II, which in turn allows for phosphorylation and active transcription (Fig. 19B) [122].

In many cases, Pol II is recruited to the promoter region, and subsequently escapes the TSS and engages in RNA transcription, without facing any hindrance. However, Pol II can be paused at around the TSS, without commencing in RNA transcription, thus accumulating at the promoter region. Why would a biological system allocate resources to recruit the transcriptional machinery to the promoter of a gene without transcribing it? Recent studies demonstrated that poised Pol II is an important mechanism to prime genes for later gene expression [139, 140]. Studies on the early mesoderm development in *Drosophila*, indicated that poised Pol II is necessary for rapid and simultaneous onset of gene expression in a larger number of cells. “De-novo”, recruitment of Pol II could cause a delay and thus prevent a synchronous onset of gene expression [108].

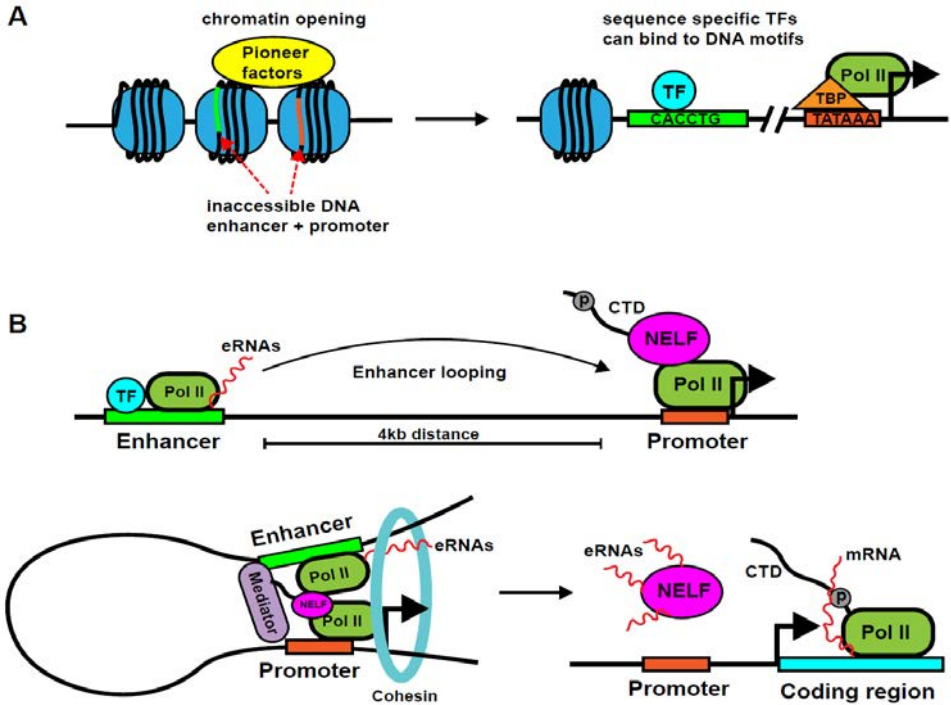


Figure 19: Model of gene regulation and transcription initiation. **(A)** Condensed DNA with inaccessible enhancer and promoter elements. Pioneer factors can open up the chromatin, and sequence specific TFs can bind to their motifs and act to keep the chromatin in an open state. This allows for the recruitment of the general transcriptional machinery including Pol II. **(B)** Enhancers and promoters separated by large distances, engage in physical contact via enhancer looping. The enhancer promoter interaction is stabilized by Cohesin and the Mediator complex. eRNAs promote dissociation of the negative elongation factor NELF from Pol II, allowing for the phosphorylation of the CTD, which results in promoter escape of Pol II and active transcription.

Results and Discussion

Paper I: Neuronal Cell Fate Specification by the Convergence of Different Spatiotemporal Cues on a Common Terminal Selector Cascade

Background and Aim

During nervous system development, daunting number of neuronal subtypes are generated, and it remains a major challenge to understand the mechanisms which determine cell fate. Specification of neuropeptidergic cells in particular is highly intriguing, because cells belonging to the same subtype of neuropeptidergic cells can be specified in different locations of the CNS in both vertebrates and invertebrates [141, 142]. This suggest that neuropeptidergic cells, which express the same neuropeptides, are specified in different ways and most likely will be under control of different spatio-temporal input. In this study we turn to the specification of two highly related neuronal subtypes, the Npl1 neurons (Tv1 and dAp), which are specified in different locations (thoracic-lateral (Tv1) vs. dorsal-medial (dAp)) of the *Drosophila* VNC. The Tv1 cells are generated by NB5-6T in a type 0 division mode and the terminal selector cascade including *col>ap/eya/dimm>Npl1* crucial for cell specification, is partly activated by spatial input via Antp/Hth/Exd and temporal input via Cas onto *col*. Specification of the dAp cells depends on the same terminal selector cascade as the Tv1 specification. However it was unknown which NB gives rise to the dAp cells, if the dAp cell is born in a type I or type 0 division mode and which spatio-temporal input activates the shared terminal selector cascade.

In this study we sought out to decode how two highly related cell subtypes can be generated in different locations and at different time points, and how different temporal spatial inputs regulate a shared terminal selector cascade, ultimately resulting in specification of the same cell. Our analysis included extensive marker stains, mutant background, rescue and cross-rescue, misexpression and co-misexpression experiments. Furthermore, designing a transgenic reporter construct (*col-dAp-GFP*) allowed us to specifically label the dAp lineage to find the NB which generates the dAp cells.

Results

During the course of this work we first tried to identify the NB which gives rise to the dAp cells. By using *Eya* as a marker we found that dAp cells start to express *Eya* at stage 13. Together with a *col-dAp-GFP* enhancer construct we were able to mark the lineage which generates the dAp cells, including its NB. Staining for a set of markers which can identify most if not all NBs [143-145], in combination with the *col-dAp-GFP* and *Eya* staining, we were able to identify the NB which generates dAp cells as NB4-3. The *col-dAp-GFP* construct was also used to analyze the NB4-3 lineage for the expression of temporal factors. Given the early birth of dAp cells compared to Tv1 cells, which are born in a Cas/Grh temporal window we could not detect late Cas expression. Therefore we turned to the earlier temporal factors. *Kr* and *hb* mutants both affect dAp cells, however we were not able to detect their expression in the lineage, which was probably a technical limitation of the *col-dAp-GFP* constructs, which was detectable after the Hb>Kr temporal windows. However we detected the expression of Pdm during the time that dAp cells were generated. In addition we found the expression of the GATA factor *grain* (*grn*, mammalian GATA1-6). Staining the NB4-3 lineage for Pros and phospho-histone 3 (PH3), showed that the dAp cells is born during GMC division which provided evidence that the dAp cell is born in a type I division mode.

Having identified that NB4-3 gives rise to dAp cells, born in an early temporal window and in a type I division mode, we turned to analyze mutants for the temporal factors. We analyzed *hb*, *Kr*, *pdm*, *cas* and *grh* mutants for *Eya* and *Nplp1* expression. While *hb* mutants showed increased numbers of *Eya* and *Nplp1* positive dAp cells, *Kr* and *pdm* resulted in dramatic decrease in Col, *Eya* and *Nplp1* positive cells as well. Mutants for the late temporal factors *cas* and *grh* did not affect the numbers of *Eya* and *Nplp1* positive dAp cells. Attempts to cross-rescue *Kr* mutants with *UAS-pdm*, resulted in partial rescue of Col and *Nplp1*. This shows that Col, and consequently the downstream terminal selector cascade for *Nplp1* specification in dAp cells, depends on the early temporal factors *Kr* and *Pdm*. However, most NBs undergo the stereotyped temporal cascade, hence *Kr* and *Pdm* expression alone do not account for the restricted expression of *col* and finally *Nplp1* in the dAp cells. Analyzing mutants for the GATA factor *grn*, showed a highly specific loss of Col, *Eya* and *Nplp1* expression in the dAp cells, while the Tv1 cells of the Ap cluster were unaffected. Our lineage analysis also suggested that *grn* is expressed early in the NB4-3 lineage and that it acts downstream of *Kr* and *Pdm*. Attempts to cross rescue *Kr* and *pdm* mutants with *UAS-grn* resulted in partial rescue of dAp neurons, whereas cross rescue of *grn* mutants with *UAS-col*,

resulted in ectopic dAp cells expressing *eya* and *Nplp1*. Cross rescue of *Kr* and *pdm* mutants with *UAS-col* resulted in ectopic expression of Eya and Nplp1. Those experiments suggested that dAp cell fate is specified by *Kr/pdm* acting on *grn* which activates the Nplp1 terminal selector cascade starting with *col*, and that *grn* acts as a spatial regulator on which the temporal factors act to specify dAp fate (Fig. 22).

Having identified *grn* as a spatial regulator which restricts the action of temporal factors, we attempted to find a similar factor for the NB5-6T lineage, which generates the Ap cluster and the Tv1/Nplp1 cell. A recent screen identified the homeobox gene *ladybird early* (*lbe*) to be critical for the Ap4/FMRFa cell specification [104]. Indeed, *lbe* is very restrictively expressed in the NB5-6 throughout the VNC [105]. Mutants for *lbe* result in complete loss of all the common Ap cluster markers i.e. Col, Ap/Eya, Dimm and Nplp1. Intriguingly, dAp cell fate is not affected by *lbe* mutants. Col is lost in *lbe* mutants in the NB5-6T, and reciprocal analysis of *lbe* expression in *col* mutants showed that *lbe* is not affected in NB5-6T. Analysis of *lbe* expression also demonstrated that *lbe* expression precedes *col* expression in the NB5-6T. Attempts to cross rescue *lbe* with *UAS-col*, showed that *col* alone could not restore Eya and Nplp1 expression in the Ap clusters. We speculated that Lbe acts both to activate *col* but also downstream *col* in a combinatorial manner to activate the Nplp1 terminal selector cascade. Single misexpression of *UAS-lbe* and *UAS-col* both result in ectopic Eya and Nplp1 expression. Combinatorial misexpression of *UAS-lbe*, *UAS-col*, however resulted in dramatic increase of Eya and Nplp1 positive cells throughout the entire VNC.

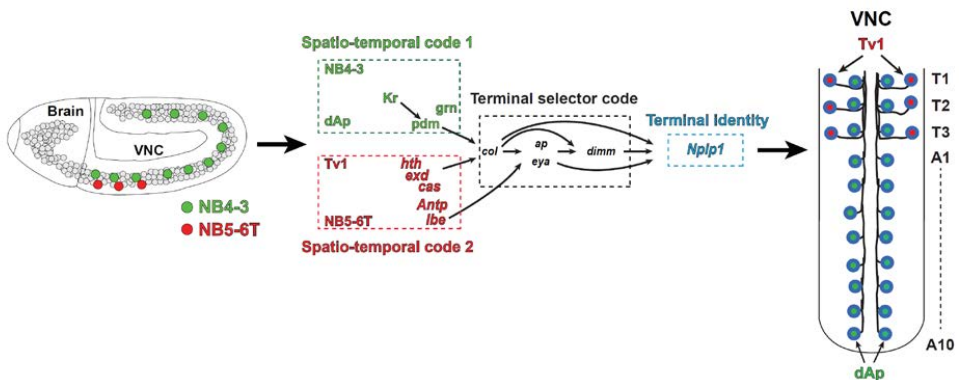


Figure 22: Model summarizing the main findings of Paper I. Two different spatio-temporal codes in two different NBs, activating a shared terminal selector cascade to specify the same terminal cell fate identity in different locations of the VNC (adapted from Gabilondo and Stratmann et. al. 2016).

Discussion

The study presented here, addresses the mechanism which can trigger the expression of the same neuropeptide in two spatio-temporally unrelated cells; the Tv1 and dAp cells. While Nplp1 expression depends on an identical terminal selector cascade of *col>ap/eya>dimm*, the cascade is triggered in different ways. The terminal selector cascade starting with *col* is triggered in the late born Tv1 neurons by spatial input provided by Antp/Exd/Hth/Lbe and by the late temporal factor Cas. Specification of the early dAp cells depends on early temporal input from Kr/Pdm and the early spatial factor Grn, to start the terminal selector cascade. Col expression in NB5-6T is controlled by Antp/Exd/Hth/Lbe/Cas, however Col expression is too broad than that it could explain the selective activation of Ap/Eya/Dimm in the NB5-6T lineage. However, the very selective expression of Lbe and feedforward action with Col can account for selectivity. In NB4-3 the action of Kr and Pdm onto Col is refined by Grn, however Grn is broadly expressed as well. Therefore it is tempting to speculate about the existence of a factor similar to Lbe which feeds forward with Col to confer spatial specificity. In summary, we could envision that this mechanism of two different spatio-temporal inputs, which funnel onto a common terminal selector entry point factor (Col) to generate similar cell fate in different locations of the CNS, is used not just in invertebrates but also in higher organisms.

Paper II: Neuronal cell fate diversification controlled by sub-temporal action of Kruppel

Background and Aim

Expression of the Hb > Kr > Pdm > Cas > Grh temporal cascade changes the competence of the NBs to give rise to different cells in each step of division. Therefore, different NB lineages can display a wide variety of different cell types. However, the instructive capacity of the five temporal factors alone cannot account for the diversity generated. Broad temporal input can be subdivided into smaller windows of competence, by the action of factors acting downstream of temporal genes, the sub-temporal genes [69]. During NB5-6T development, the four last Apterous neurons (Tv1/Nplp1, Tv2/Tv3, Tv4/FMRFa) of this lineage are born in a Cas/Grh mixed temporal window. Cas is essential to activate the terminal selector cascade Col>Ap/Eya>Dimm>Nplp1, crucial for the Nplp1 cell fate specification of the first direct born

Ap1 cell, whereas Grh is involved in specification of the Tv4/FMRFa cell. However not all neurons born in a Cas temporal window will express Nplp1 when terminally specified, and Tv2/3 become “generic” Ap interneurons and Tv4 expresses FMRFa. While Cas activates Col, it also activates the sub-temporal cascade which involves the factors Sqz, Nab and Svp. These factors are involved in repressing Col in the Tv2/3 cells, which prevents these neurons from acquiring the Nplp1 cell fate [69, 86]. However, the precision of Ap cluster cell fate allocation i.e., Nplp1 expression in just Tv1, absence of neuropeptides from Tv2/3 and FMRFa expression in just Ap4, suggest the existence of additional factors, besides the know sub-temporal factors.

In our previous study [66] of factors involved in dAp specification, we found that mutants for the early temporal factor *Kr* showed dramatic impact on dAp specification. Intriguingly, *Kr* mutants displayed another very interesting phenotype in the four cell Ap cluster with regards to Nplp1 expression. We found that in *Kr* mutants, the Nplp1 expression is lost from Tv1 cells, while Eya expression was still detectable. Therefore we aimed in this second project to explore how the early temporal gene *Kr*, expressed in NBs, is involved in the late and postmitotic Nplp1 cell specification. We were intrigued if *Kr* also experiences a late re-onset as it was reported for Svp, and by that integrates with the terminal selector or sub-temporal cascades to confer cell specific identity. Furthermore, was it of great interest to find factors involved in *Kr* regulation, in the case that *Kr* is re-expressed. Previous studies on other NB lineages demonstrated that *Kr* was expressed postmitotically, however without analyzing the biological role of this late expression [146]. To explore the role of *Kr* in Ap cluster specification, we analyzed mutant and misexpression phenotypes for *Kr* and various factors of the sub-temporal cascade. In addition we carried out rescue and cross-rescue experiments together with combinatorial misexpression.

Results

Analyzing *Kr* mutants at stage AFT, we found that all four Ap cluster cells still express Eya, and we found a specific loss of Dimm and Nplp1 in the Tv1 cells, while the Dimm and FMRFa expression in Tv4 was unaffected. Thus, we observed a highly specific loss of just one cell fate in the Ap cluster. Dimm expression in the Tv1 cells depends upon Col, Ap and Eya. Since Eya showed normal expression, this prompted us to turn to Col expression in the Tv1 cell in *Kr* mutants. While the early Col expression was not affected in stage 14 NBs and Tv1 cells, we observed a late loss of Col in the Tv1 cells in *Kr* mutants. It seemed that *Kr* acts to maintain

Col expression. Nab one of the sub-temporal factors, which allows to distinguish between the Col positive Tv1 and Tv2 cells at stage 14, was not affected by the *Kr* mutants. This apparently late action of *Kr* on *col* prompted us to analyze the *Kr* expression itself. As anticipated, in stage 14 NBs, we were not able to detect *Kr* expression given its early temporal role. Turning to stage 14 Tv1 cells, we could not detect any *Kr* expression either. By stage 16 however, we found that *Kr* is robustly expressed in just the Tv1 cell within the Ap cluster. In summary, the wildtype and *Kr* mutant analyses provided evidence that *Kr* is re-expressed late in the Tv1 cell, and that loss of *Kr* just impacts the Nplp1 cell fate, via Col, without further affecting the second neuropeptide cell of the Ap cluster. Targeted and late misexpression of *UAS-Kr* to the Ap clusters using the *ap-Gal4* driver, resulted in elevated numbers of positive Col, Dimm and Nplp1 cells, while FMRFa was lost. We also included Nab expression in our analysis and found that Nab was not affected by *Kr* misexpression. Attempts to rescue of *Kr* mutants with *UAS-Kr* and cross-rescue with *UAS-col*, both resulted in robust reappearance of Dimm and Nplp1 cells in the Ap clusters. To this point we established that late Tv1 cell specific *Kr* expression is important to activate *col* and *dimm* and therefore involved to trigger Nplp1 cell fate. Therefore we were interested in which factors might be involved in *Kr* activation. We tested mutants for factors important for NB5-6T lineage progression and specification, such as *Antp*, *lbe*, *cas*, *col*, *ap*, *eya* as well as *grh*. Strikingly, we found that *Kr* expression is lost from the NB5-6T lineage and the Ap clusters only in *cas* mutants. However, *Cas* spans the latter half of the NB5-6T lineage including the Ap neurons, therefore *cas* expression alone does not explain the precise onset of *Kr* in just the Tv1 cell. It was previously shown that the sub-temporal factors *sqz* and *nab* are activated by *Cas* in a feedforward loop (*cas>sqz>nab*) which creates a temporal delay and restricts the expression of *nab* to the Tv2-Tv4 neurons. In addition it was demonstrated that *sqz* and *nab* act to repress Col in Tv2/Tv3 neurons [69]. This prompted us to investigate if *Kr* plays a role in the sub-temporal FLL. Misexpression of *Kr* as well as mutants for *Kr* did not affect Nab. Thus the sub-temporal FLL is not affected by *Kr*. Turning to mutants for *sqz* and *nab* we observe additional Dimm and *Kr* positive cells in the Ap clusters. Previous studies showed that misexpression of *nab*, results in loss of Col, Dimm and Nplp1. This prompted us to misexpress *nab*, which resulted in the loss of *Kr*. *Sqz* and Nab restrict the expression of *Kr*, whereas *Kr* has no effect on the sub-temporal cascade. Cross rescue of *Kr* with *UAS-col*, resulted in restoration of Nplp1 expression, hence *Kr* is not critical for the *col>ap/eya>dimm>Nplp1* terminal selector cascade. Therefore we turned to the sub-temporal factor *Svp*. A previous study demonstrated that mutants for *svp* show ectopic Col, Dimm and Nplp1 cells in the Ap cluster, while postmitotic *svp* misexpression results in loss of Col, Dimm

and Nplp1. Svp is first expressed in all four Ap cluster cells, however disappears rapidly by stage 15/16 from Tv1 and Tv4, and becomes restricted to only Tv2/3. When we analyzed *sqz* and *nab* mutants, we found that in addition to ectopic *dimm* and *Kr* expression, there was loss of Svp expression. Misexpression of *nab*, resulted in one extra Svp positive cell and loss of Dimm. This implies that *sqz* and *nab* contribute to Ap cluster diversification by ensuring *svp* expression in Tv2 and Tv3. We next tested the connection between *Kr* and *svp*, and found one extra Svp positive Ap cluster cell in *Kr* mutants. Appearance of one extra Svp cell coincided with loss of Dimm and Nplp1, which gives reason to believe that extra Svp is activated in the Tv1 cell. The reciprocal experiment i.e., misexpression of *Kr* in the Ap cluster, resulted in total loss of Svp in the Ap cells. Next we analyzed *svp* mutants and find additional Kr cells in the Ap cluster, which could indicate a repressive role of Svp on Kr, however misexpression of Svp did not impact the wildtype *Kr* expression, while Dimm was repressed. Testing *svp-Kr* double mutants we find extra Dimm and Nplp1 positive Ap cluster cells, similar to *svp* single mutants. These results demonstrated that the main function of *sqz* and *nab* is to ensure maintenance of Svp in Ap2 and Ap3, by repression of *Kr* and that Svp is a potent repressor for *dimm* and *col*, which is necessary but not sufficient to repress Kr. In addition, the double mutant experiment demonstrates that the main role of Kr is to repress Svp in the Tv1 cell, which safeguards the progression of the terminal *col>ap/eya>dimm>Nplp1* feedforward cascade (Fig. 23). In order to test if Kr is a repressor of Svp globally, we misexpressed Kr from *ham-gal4*, which is active in NBs from row 4. Indeed, we found a loss of Svp in row 4 lineages as well. Turning to *Kr* mutants, we also find elevated numbers of Svp positive cells, in the VNC, however not in the NBs from row 4.

In a final experiment we wanted to test the concept that Kr can refine cell fate specification by suppression of Svp. In Paper I, we demonstrated that combinatorial co-misexpression of *UAS-col* and *UAS-lbe* resulted in dramatic increase of Ap, Eya, Dimm and Nplp1 positive cells in the VNC. However not all cells positive for Ap and Eya were Dimm or Nplp1 positive. Therefore we speculated that *UAS-col* and *UAS-lbe* misexpression resulted to some extent in Ap/Eya generic cells without reaching final Nplp1 cell fate, because they also express *svp*. Analysis of Svp positive cells in *UAS-col*, *UAS-lbe* double misexpressing VNCs, indeed revealed dramatic increase of Svp positive cells, overlapping with ectopic Eya cells. Adding *UAS-Kr* to this double misexpression reduced the number of Svp positive cells, while Ap/Eya was still ectopic expressed. Strikingly, we found that Dimm cells numbers were increased in *UAS-col*, *UAS-lbe* double misexpression experiments. However, adding UAS-Kr

to those two factors, we could increase the Dimm cell numbers and Nplp1 expression significantly, which was due to decreased *svp* expression.

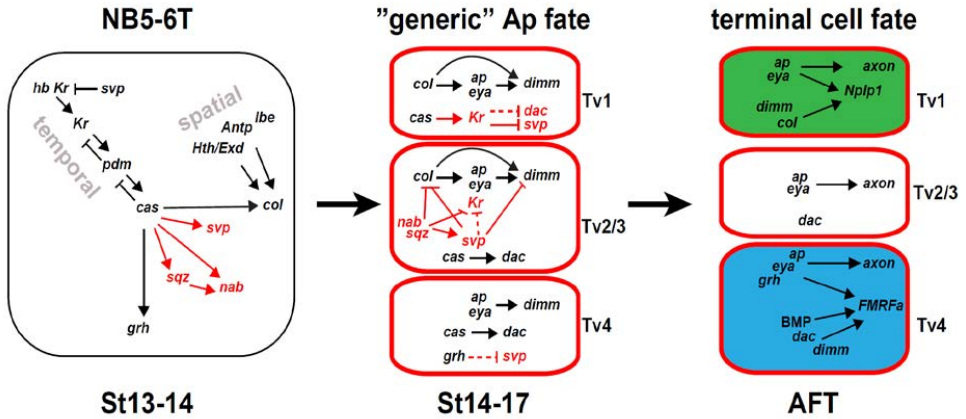


Figure 23: Model summarizing the main findings of Paper II. Cas acts in concert with spatial selectors to activate the terminal cell selector cascade to specify Tv1 cell fate. Simultaneously, Cas activates the sub-temporal cascade which in turn subdivides the four generic Apterous cells and by that allocates the right cell fate to the four Ap cluster cells. Kr is re-expressed late in the postmitotic Tv1 cells and helps to suppress Svp in order to maintain Col and Dimm expression, which supports progression of the terminal cell fate cascade activating Nplp1, in just one out of four Ap cells (adapted from Stratmann and Gabilondo et. al. 2016).

Discussion

The temporal factor Cas acts together with spatial information provided by Lbe/Antp/Hth/Exd to activate Col in the NB5-6T, and thus the feedforward terminal selector cascade to activate Nplp1 in the Tv1 cells. Cas also activates the sub-temporal genes *sqz*, *nab*, *svp* and *Kr* together with the late temporal factor Grh which acts in the specification cascade for the Ap4/FMRFa cells. The Sub-temporal factors *sqz* and *nab* ensure Svp expression in Tv2/3 due to their repressive role of Kr. This allows Svp to repress Col and Dimm, hence keeps the Tv2/3 cells in a generic interneuron fate. Due to delayed expression, *nab* is absent from the Tv1. This allows *Kr* to be expressed, which at the same time attenuates *svp* expression and allows for maintenance of Col to ensure the progression of the Nplp1 specification cascade. Svp however is also absent from the Tv4 cell, which is specified into the neuropeptidergic FMRFa cell, which depends on the terminal selector cascade *ap/eya>dimm>dac/BMP>FMRFa* [69, 99]. Mutants for *grh* show elevated numbers of Svp cells, however misexpression of *grh*, does not result in Svp suppression. Hence we envision that in the Tv4 a similar factor as Kr exists, to

attenuate *Svp*. Analyses to find activators for late *Kr* expression showed that the spatial factors *Lbe* and *Antp*, had no impact on *Kr* or on other factors of the sub-temporal cascade evident by continued *nab* expression. *Cas* is the main factor involved in sub-temporal cascade and *Kr* activation. This points to a mechanism in which the spatial factors are only involved in triggering a coherent type 1 FLL terminal selector cascade, which however is not restricted to one cell, but rather to a whole lineage. Therefore the temporal factors which are more restrictively expressed, in this case *Cas*, take part in activation of the C1-FLL. This restricts the terminal selector cascade to just cells born in a *Cas* temporal window. Simultaneous activation of the sub-temporal factors by *Cas*, triggers the *sqz>nab>svp* FFL, and also activates the *Svp* repressor *Kr*. *Kr* is partly regulated by members of the sub-temporal cascade and represses back on *Svp* which creates a sub-temporal incoherent type1 FLL. This incoherent FLL acts to micromanage the cell specific gene expression, and thereby for cell specific allocation of cell identity in a group of similar cells. It is tempting to speculate that such mechanisms i.e., integration of activating and repressive FLLs, and late repurposing of early temporal factors in order to create cell diversification, is not just used in *Drosophila* but also in higher organisms.

Paper III: Neuronal Cell Fate Specification by the Molecular Convergence of Different Spatio-Temporal Cues on a Common Initiator Terminal Selector Gene

Background and Aim

As described in Paper I, the *Nplp1* neurons in the *Drosophila* ventral nerve cord can be subdivided into the thoracic ventro-lateral *Tv1* and the dorsal-medial *dAp* neurons. *Nplp1* expression in both cell subtypes is activated by the same terminal selector cascade: *col > ap/eya > dimm > Nplp1*. However, *Tv1* and *dAp* neurons are generated by different NBs and depend on different upstream cues to activate the cell specification cascade. In *Tv1* neurons the terminal selector cascade is triggered by *Antp*, *hth/exd*, *lbe*, *cas*, while in *dAp* cells it is triggered by *Kr/pdm* and *grn*. In a molecular approach we aimed to test if the TFs involved in the *Nplp1* specification cascade interact in a direct fashion with the cis-regulatory motifs (CRMs) of their downstream targets. First, we set out to identify the CRMs of the factors which are part of the terminal selector cascade ultimately specifying *Nplp1* cell fate, by building CRM-reporter constructs. Once the CRMs would be identified, we then would try to delete the endogenous CRMs with CRISPR-Cas9 in order to determine if the identified CRMs are specific for the factors involved in the *Nplp1* cascade. In a third approach we wanted to test the

CRM-reporter constructs in flies mutant for the appropriate TF and scored the CRM-reporter constructs for activity loss. In a fourth experimental approach we would generate CRM-reporter constructs with mutated transcription factor binding sites (TFBS) in order to phenocopy the latter results. Finally we overexpressed TFs to determine activation of the specific CRM-reporter constructs.

Results

To identify cell-specific CRMs for the factors which take part in the terminal selector cascade which specifies Tv1 and dAp cells, we analysed a large number of transgenic fly lines containing potential enhancer-reporter constructs (*CRM-GFP/lacZ*). Some of the CRM-reporter constructs were based on previous studies [147], personal communications (Eya-enhancers from Douglas W. Allan), and the large Janelia Gal4 collection [116]. With this large analysis we identified CRMs for all factors involved in the terminal selector cascade and validated their specific enhancer activity in the Ap clusters and dAp cells. Intriguingly, we found two separate CRMs controlling *col* (Fig. 24). One CRM was specific to control *col* in the NB5-6T and the Ap clusters, while the other CRMs showed a strong specificity to control *col* in the dAp cells. We termed those CRMs *col-Tv-CRM* and *col-dAp-CRM*. Since the *col-Tv-CRM* was coupled to a *lacZ* reporter and the *col-dAp-CRM* to a *GFP* reporter, this allowed us to combine those two CRMs to ensure that they showed no overlapping activity. In a next step we deleted the newly identified CRMs with the CRISPR/Cas9 systems from their endogenous genomic loci, to test their importance in controlling their own gene expression, but also their role in the terminal selector cascade to activate Nplp1. Deletion of the *col-Tv-CRM*, had very minor impact on Col expression in the NB5-6T, and no impact on the Eya or Nplp1 expression. As anticipated, deletion of the *col-dAp-CRM*, had no impact in the numbers of Col, Eya and Nplp1 cells in the Ap clusters, while the numbers of Col, Eya, Nplp1 positive dAp cells were significant lower compared to wild type numbers. Deletion of the *eya-CRM*, *aps-CRM*, *dimm-CRM* and *Nplp1-CRM*, all resulted in significant lower numbers of Nplp1 positive dAp and Tv1 cells. Furthermore, each CRM mutation effected its own gene expression significantly.

Then we placed the identified unmodified CRM-reporter constructs into mutant backgrounds for the appropriate TFs. The *col-Tv-CRM-lacZ* transgene was crossed into mutants for *hth*, *Antp*, *lbe* and *cas*, which resulted in reduction of CRM activity in all cases. The *col-dAp-CRM-GFP* construct was placed into mutant backgrounds for *pdm*, *Kr* and *grn*,

which resulted in activity loss of the CRM-reporter construct in the dAp cells in all cases. Analyses of the *eya-CRM-GFP*, *apS2-CRM-LacZ*, *dimm-CRM-GFP* and *Nplp1-CRM-GFP* in mutant backgrounds for their appropriate TFs, all showed a significant reduction or loss of CRM-reporter activity.

In the next step we analysed the CMRs for potential transcription factor binding sites (TFBS) for the appropriate TFs, which we then converted. In the *col-Tv-CRM-lacZ* construct we converted potential TFBS for Hox factors, Cas, Exd and Hth, whereas in the *col-dAp-CRM-GFP* construct we converted potential binding sites for Pdm/POU, Grn/GATA and Kr. Converting or mutating TFBS in the CRMs resulted in reduction of CRM-reporter activity in all cases when compared to the wildtype-CRM. Mutation of TFBS for Hox, Col and Dimm in the *eya-CRM-GFP*, *apS2-CRM-lacZ*, *dimm-CRM-GFP* and *Nplp1-CRM-GFP* constructs led to significant reduction of reporter expression in almost all cases. Mutation of potential Col binding sites in the *dimm-CRM-GFP* construct, resulted in significant decrease of the GFP reporter expression in the dAp cells, while the GFP expression was significantly increased in the Ap cluster cells upon mutation, compared to the wildtype *dimm-CRM-GFP* expression. Mutation of potential Col binding sites in the *Nplp1-CRM-GFP* construct, had no impact on the GFP expression in both dAp or Tv1 cells, compared to the wildtype *Nplp1-CRM-GFP* reporter expression.

Finally we tested the wildtype CRM-reporter constructs in misexpression backgrounds, in order to test if we could activate ectopic reporter expression. It was previously demonstrated that a different combinations (cocktails) of TFs can ectopically activate the different terminal selectors in the *Nplp1* specification cascade [66, 94, 102]. Placing the *apS2-CRM-lacZ* and *eya-CRM-GFP* transgenes into *UAS-lbe*, *UAS-col* misexpression, resulted in dramatic activation and increase of reporter expression. Furthermore, we tested the *dimm-CRM-GFP* transgene in a misexpression background of *UAS-col*, *UAS-ap*, *UAS-eya* (Dimm "cocktail"), and the *Nplp1-CRM-GFP* transgene in a misexpression background of *UAS-col*, *UAS-ap*, *UAS-eya*, *UAS-dimm* (*Nplp1* "cocktail"). In both cases the misexpression of the respective cocktails resulted in dramatic CRM activation and increase in the GFP reporter expression throughout the entire VNC.

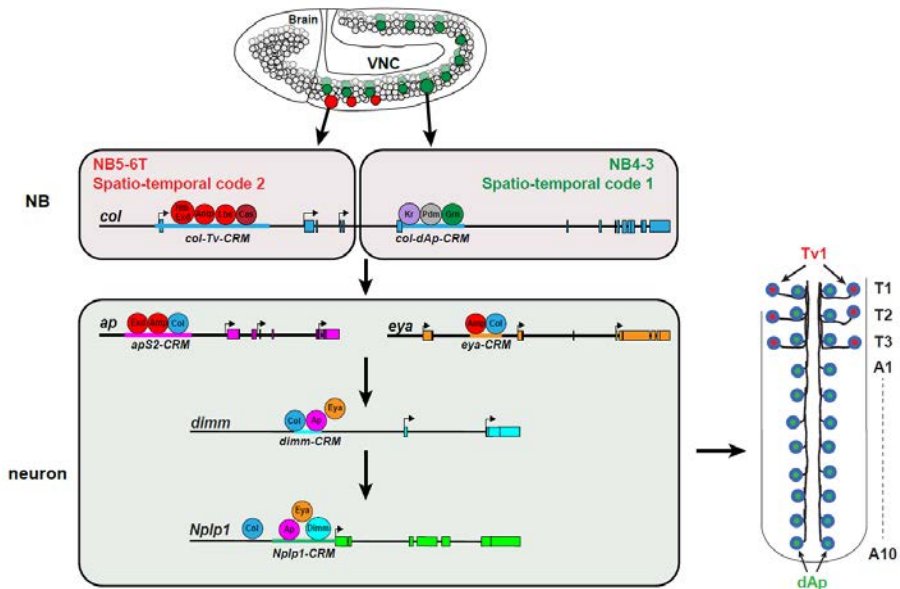


Figure 24: Model summarizing the main findings of Paper III. *col* is under control of two different enhancers, one which is specifically activated in NB5-6T and one which is specifically activated in NB4-3, by different upstream TFs. In the postmitotic neurons Col acts in a feedforward manner on the enhancers of *ap*, *eya*, *dimm* and *Nplp1*. The enhancer elements activating cell type specific gene expression of the factors involved in *Nplp1* activation, are shared between the Ap cluster and dAp cells.

Discussion

It was demonstrated by previous studies that *Nplp1* expression depends on a feedforward cascade of the terminal selectors *col*>*ap/eya*> *dimm*. However little was known about their direct interaction with the CRMs of the genes they act on. Therefore we tried to find and characterize the cis-regulatory modules (CRMs), for the factors involved in the *Nplp1* terminal selector cascade, in order to better understand the molecular regulatory mechanisms during cell fate specification. This is to our knowledge the first study which approaches such an ambitious undertaking, given the fact that cell specific enhancers are notoriously hard to identify in multicellular organisms, based on their scattered distribution in the genome.

Fascinatingly, we could identify two different CRMs regulating *col* expression in a spatially specific fashion. The finding of *col* is regulated by different CRMs is in line with our previous findings that *col* is differently regulated in NB5-6T vs. NB4-3, by different spatial and temporal inputs *Lbe/Antp/Hth/Exd/Cas* vs. *Pdm/Kr/Grn*. The different regulation by different upstream cues, therefore requires two different CRMs which can serve as integration hubs to funnel the different upstream information onto the same gene. This notion is further

supported by mutant background and TFBS mutation analyses of those two different CRMs. However we found that deletion of the 6.3kb long *col-Tv-CRM* had a minor impact on Col expression and further no effect on Eya or Nplp1 expression. Deletion of the *col-dAp-CRM*, resulted in significant loss of Col/Eya and Nplp1 positive dAp cells, however the effects were not as great as a deletion of the *apS-*, *eya-*, *dimm-* or *Nplp1-CRM*, which resulted in almost complete absence of their respective gene expression. Studies on early developmental and broadly expressed genes such as *sog*, *twi* and *sna*, suggested that those genes are controlled by multiple redundant enhancers, or shadow enhancers, in order to sentinel robust gene expression. Therefore it is tempting to speculate that *col* is also regulated by multiple redundant enhancers, given its broad and comparatively early expression.

Mutation of TFBS in CRMs resulted in almost all cases in significant reduction of reporter expression. Bioinformatic analysis of potential Col binding sites, showed that the *eya-CRM* contains two perfectly conserved CCCNNGGG binding sites and the *apSJ2-CRM* one conserved binding site for Col i.e., TCCNNNNGGA. Mutation of these motifs resulted in strongly decreased reporter expression. Bioinformatic analysis of the *dimm-CRM* and *Nplp1-CRM* did not reveal such conserved binding sites. However we decided to mutate non perfect matches of potential Col binding in these enhancers, which could perhaps resemble low affinity sites as it was shown for other factors [64]. Analyses of the *dimm-CRM-GFP* and *Nplp1-CRM-GFP* with mutated potential Col binding motifs, demonstrated no strong effects on the reporter expression. However, previous work suggested that Col is crucial for Dimm and Nplp1 expression. It was shown that maintenance of Col is necessary for Nplp1 expression, and experiments with RNAi against *col* in larvae and adult flies, showed that Nplp1 is lost when Col is absent, while Ap/Eya and Dimm are not affected [148]. Our findings here suggest that Col action upon *dimm* and *Nplp1* could be indirect. It was shown that the mammalian Col homologue Ebf1 can interact with the SWI/SNF chromatin remodeling complex, to keep the chromatin in an accessible state [149]. A recent screen from our lab found the Brahma complex, a central component of the SWI/SNF, to be important in FMRFa cell specification. Mutants for Brahma affect neuropeptide expression while Eya is unaffected, thus indicating a late role. Therefore, it is a possible scenario that Col in the Tv1 cells acts to activate *ap* and *eya* expression in a direct fashion. Later on Col switches its role to interact with chromatin remodelers and acts to keep the chromatin in an open state for Ap and Eya to bind to the Dimm and Nplp1 enhancers.

Concluding Remarks

The work presented in the different studies sheds light on the complex nature of terminal cell fate specification, based upon the model of the neuropeptidergic Nplp1 cells, including the Tv1 cells of the Ap cluster and the dorsal Ap cells. To better understand those intricate mechanisms we combined classical genetic (mutant background, targeted misexpression, rescue and cross-rescue experiments) but also molecular genetic approaches (enhancer mutagenesis, CRISPR/Cas9 deletion), which enabled us to gain deeper understanding of the regulatory logic behind cell fate specification. Combining those different approaches created a powerful “toolbox” with which we were able to elucidate the role of multiple factors and enhancers pivotal in cell determination, from various angles. In addition, many findings on which we built our experimental reasoning and design, are the result of over a decade of hard work on Ap cell fate specification. Based on this ground-work, we were able during the course of this work to find new regulatory mechanisms which determine cell fate and which are likely to be important not just in *Drosophila*, but also in higher organism.

The first main finding is, that two unrelated neuronal progenitors can create cells of the same cell sub-type, by integrating different upstream information. Funnelling two different upstream cues on a common entry point factor, allows to activate a shared terminal selector cascade in different cells, which then specify and express the same neuropeptide Nplp1. Besides expression of the same neuropeptide, these cells also project their axons into a common fascicle and into the same anterior direction. In this scenario the common terminal selector cascade acts on a shared enhancer to activate the neuropeptide. However, another scenario to create the same cell type identity is possible to imagine. Hereby, two unrelated upstream cues may instead act on two unrelated terminal selector cascades, which act on two different enhancers for the same neuropeptide. The specification would therefore happen in two parallel unrelated pathways, by usage of different CRMs for the same effector gene.

The second main finding of this project is, that the early temporal factor Kr has dual roles. First Kr takes part in the temporal progression of the developing NB. Kr then experiences a later re-expression, and acts as part of the sub-temporal cascade to diversify cell fate. The concept of the sub-temporal factors acting to sub-divide larger temporal windows in order to create different cell types was shown before [69], but could not fully account for the differences in cell types. Finding that Kr is part of the sub-temporal FFL, however can logically explain the terminal cell fate specification of the Tv1 cell. Cas being the main activator of late

Kr expression was another intriguing finding, stressing the important role of Cas to act in two main ways, first taking part in a terminal selector cascade, and second being the main activator of a counteracting sub-temporal cascade including Kr.

The third main finding of this project was, that we could identify the enhancer cascade which integrates the molecular action of the TFs involved in the Tv1 and dAp specification. The most intriguing part certainly remains the identification of two separate CRMs for Col, which are under control by different upstream cues. This was a molecular proof-of-concept of the previous genetic findings that Col is activated by two different upstream sets of TFs. This study however included two more highly interesting findings. First, the highly dynamic and broad expression of Col is most likely under control of a redundant second enhancer or a shadow enhancer, to ensure robust expression. Second, Col might not interact with the Dimm or Nplp1 CRMs in a direct fashion, but rather has to be maintained to act together with chromatin remodelling complexes in order to keep the enhancers in accessible states for other TFs.

Besides from the developmental perspective, to better understand the logic behind biological systems, the work presented here may be of interest from a medical perspective. Deeper understanding of gene regulatory networks, is of great interest to gain better knowledge of disease development and progression [150]. In addition, much effort is also focused on regenerative medicine to grow specific tissues or cell types under lab conditions, in order to be able to replace impaired or degenerated tissue or cells. Therefore, deciphering TF networks will improve the ability to direct cells into specific cell types.

Turning to the enhancer project, recent large genome-wide association studies (GWAS) have demonstrated that most of the genetic variants which predispose to cancer, reside inside of putative enhancer elements [117, 151]. Mutations of binding motifs for the MYB TF are associated with increased enhancer activity and involved in T cell acute lymphoblastic leukaemias [152]. However, the abundance of potential enhancer elements in the human genome (~400,000-1million) creates a challenge to identify and functionally characterize these non-coding elements. Therefore it is of great interest to broaden the knowledge of the regulatory potential of enhancers, by understanding their molecular properties.

Methods

***Gal4/UAS* transgenic expression system**

The *Gal4/UAS* system was adopted from yeast, applied to *Drosophila* and is used for targeted gene expression. The *Gal4/UAS* system consists of two components or transgenic flies. One fly carries a construct in which the Gal4 protein is expressed under the control of a gene regulatory sequence (enhancer-promoter sequence). As an example, Gal4 can be expressed under control of the regulatory sequence for the gene *embryonic lethal abnormal vision (elav; elav-Gal4)*, which results in that *Gal4* is constantly expressed in postmitotic neurons of the CNS. Another fly carries a construct which contains the cDNA of a gene of interest (for instance *Col*) coupled the UAS sequence (Upstream Activating Sequence) (*UAS-col*). Crossing those flies together results in the misexpression of *col* in the postmitotic neurons in the entire CNS if the embryos carry both constructs [153].

Immunohistochemistry (IHC)

IHC was applied to label protein/marker expression in the developing embryonic *Drosophila* VNC. In brief, *Drosophila* embryos at the appropriate stage were hand dissected under stereomicroscopes in 1x phosphate buffered saline (PBS) on poly-lysine covered glass slides. PBS was removed and replaced by 4% para-formaldehyde (PFA) and embryonic VNCs were fixed for 20 minutes at room temperature. Subsequently, PFA was washed off the sample with PBS and PBT (PBS + TritonX) and the tissue was incubated in PBTN (phosphate buffered saline + TritonX + Horse serum) for 1 hour at ambient temperature. Primary antibodies (ABs) were diluted in PBTN and samples were covered with AB mixture for further incubation. Incubation was performed at 8°C, over-night. After primary incubation, samples were washed 3 times 20 minutes with PBT. Secondary fluorophore labeled ABs, against primary antibodies were diluted in PBTN and added to the washed samples. Incubation with secondary AB was performed at ambient temperature for 1 hour. Subsequently samples were washed with PBT, 3x 20 minutes. PBT was removed and samples were covered in Vectashield H-1000 (Vector Laboratories, Burlingame CA, USA), and mounted with a glass cover slip. Samples were stored afterwards in 4-8°C for one night prior confocal scanning. For further information regarding ABs, please see publications. The validity of primary antibodies used was tested by staining mutants and misexpression embryos.

Confocal scanning and image analysis

Please see Paper I, II and III for further information.

Statistics

Please see Paper I, II and III for further information.

Transgenic flies

For detailed information about transgenic flies, please see Material and Methods of Paper I, Paper II and Material and Methods plus Supplemental Information 1 in Paper III.

Enhancer identification Enhancer Mutation

For detailed information, please see Supplemental Figure 1 and Supplemental Information 3 and 4 in Paper III.

CRISPR/Cas9 deletion of Enhancers

For detailed information, please see Material and Methods CRIPSR/Cas9 and Supplemental Information 2 in Paper III and Figure 25.

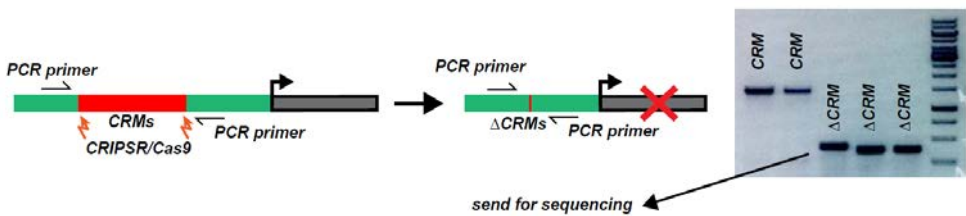


Figure 25: CRISPR/Cas9 method was used to delete CRMs. PCR primes were designed to flank potential CRM deletion sites. To test for successful deletion, PCR was performed to detect shorter PCR fragments. PCR fragments were cloned into plasmid and sent for sequencing to characterize deletion. See supplemental information 2 in Paper III.

Acknowledgements

I would like to express my sincere gratitude to the people who have contributed to this thesis and supported my work. I try to describe their role from a biological perspective.

The **S₁-Thorase** main supervisor enzyme (1), contains a Serine-Threonine (ST) rich hammer-domain, which breaks up the highly condensed chromatin of science and propagates binding of the PhD-stud domain containing TF JSt, to the research promoter. An elegant study revealed that, besides its ST-hammer domain, the Thorase contains the high intelligence (HI), and multiple motivation enhancing subunits (MES), which interact strongly with the JSt PhD-stud domain to increase its function dramatically. Furthermore, it was demonstrated that the Thorase interacts with the co-supervisor enzyme **Jan-Ingvarase** (2), in order to prevent the 3D structure of JSt to undergo non-favourable conformation changes. Proper function of JSt in its environment, also requires constant interactions with its co-factors **ShA-Z**, **Ca²⁺BS**, **ErkGnR**, **NaChO**, **Be²⁺Z**, **JeRC**, **SuB**, **JoSFe³⁺** and **Mg²⁺B** (3). In addition, the chaperones **HeK**, **AnSt** and **Ca²⁺ JS** (4), support proper folding of the JSt 3D structure, and mutations for either one of these factors, has been shown to decrease JSt function. Other lines of evidence were presented, that short range interaction between JSt and the mediator complex **RMcD**, led to interaction of JSt and the Thorase enzyme. In addition, regular long range interactions between the heterodimer **RMcD/JoMhK** (5) and JSt are observed and contribute to its developmental control. Other crucial long-range interactions include the TFs **HuGal4** and **Jo-BeS** (6), both acting in a C1-FFL to activate JSt and increase transcriptional output. Bioinformatic analysis showed binding sites in JSt for the adaptor proteins **JoR**, **SoS**, **JoUB** and **AndC** (7), which all contribute to the function of JSt by providing the correct amount of substrate and by controlling homeostasis of its environment. Co-IP experiments carried out over a decade ago could impressively show that the ying-yang motif containing factors **SebGn** and **JuSp** (8) both act as strong transcriptional co-activators for JSt as well mild repressors. Extensive studies demonstrated however, that the repression usually coincided with higher concentrations of EtOH in the physiological environment. Initial experiments however proved that the first activation of JSt strongly depends on the maternal and paternal inherited factors **MuM1** and **DaD1** which form together with the brother domain containing TF **Rob2** the familiar trimeric complex (9). When not engaged in active transcription, JSt is covalently bound to its co-factor **SuS**, via the HeaRT subdomain (10).

Thanks to all my co-factors!!!

References*

1. **Stefan Thor.** (2012): Functional characterization of the HI and MES group in the Thorase enzyme complex- a positive outcome long term study. *Enzyme* 8(14) 298-356
2. **Jan-Ingvar Jönsson.** (2012): Jan-Ingvarase prevents PhD-stud domain ubiquitination and proteolytic degradation. *Mol.Enz.Inv.* 4(23) 478-489

3. **Shahrzad Bahrampour., Caroline Bivik S., Erika Gunnar, Behzad Salmani, Ignacio Monedero C. (Nacho), Jesus Rodriguez C. Susanne Bauer, Josefine Fernius, Magnus Baumgardt.** Positive outcome on JSt function depends on continuous interaction with its co-factors. (2015): *Cofactors 1(3) 189-234*
4. **Carolin Jönsson, Annika Starkenberg, Helen Ekman.** (2012): 3D structure maintenance of JSt depends on chaperone function and support. *Crystallization 5(8) 981-1002*
5. **Ryan MacDonald, Joanne M Hackett.** (2012): Long range interactions involved in PhD-stud domain activation. *Phys. Chem. 3(10) 2012-2100*
6. **Hugo Gabilondo, Jonathan Benito-Sipos.** (2014): Two TFs act in a combinatorial feedforward loop and increase JSt transcription rate. *Mech. Euka. Transc. 5(5) 190-221*
7. **Jonathan Rakar, Sofie Sundberg, Jonas Ungerbäck, Anders Carlsson.** (2011): In-vitro translation of JSt mRNA in a high caloric, game and beer enriched environment, proves positive outcome on protein function. *Lingon, Lådvín & Långkalsonger 8(12) 165-190*
8. **Sebastian Gnosa, Julian Spies.** (2005): Interactions of SebGn and JuSp on JSt provide highly efficient functional output. *Darmstädter Echo 24(10) 12-19*
9. **Marianne S. Erich S. Robert S.** (1985): Tripartite family complex correlates with formation and stabilization of a highly functional JSt factor. *Curr. Opi. Fam. Interact. 58(56)88-2017*
10. **Susanna L.** (2010): Covalent binding of SuS to JSt via HearT domain is pivotal for optimal function of JSt. *Åbo Underrättelser 8(5) 312-544*

*References are for entertainment purposes only and do not exist in reality. Any resemblance with original literature is purely coincidental.

REFERENCES

1. Thesleff, I. and G.T.J. Huang, *Stem Cells, Craniofacial Development and Regeneration*. 2013, Somerset, UNITED STATES: Wiley.
2. Mueller, W.A., M. Hassel, and M. Greal, *Development and reproduction in humans and animal model species. [Elektronisk resurs]*. 2015: Berlin : Springer, 2015.
3. August, W., *Das Keimplasma, Eine Theorie der Vererbung*. . 1892: New York Scribner
4. Tata, P.R., et al., *Dedifferentiation of committed epithelial cells into stem cells in vivo*. *Nature*, 2013. **503**(7475): p. 218-23.
5. Jopling, C., et al., *Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation*. *Nature*, 2010. **464**(7288): p. 606-9.
6. Kragl, M., et al., *Cells keep a memory of their tissue origin during axolotl limb regeneration*. *Nature*, 2009. **460**(7251): p. 60-5.
7. Lehoczky, J.A., B. Robert, and C.J. Tabin, *Mouse digit tip regeneration is mediated by fate-restricted progenitor cells*. *Proc Natl Acad Sci U S A*, 2011. **108**(51): p. 20609-14.
8. Waddington, C.H., *Organisers & genes*. Reprinted. ed. Cambridge biological studies. 1947, Cambridge: at the University Press. 160 s. 2 tav.
9. Edlund, T. and T.M. Jessell, *Progression from extrinsic to intrinsic signaling in cell fate specification: a view from the nervous system*. *Cell*, 1999. **96**(2): p. 211-24.
10. Spemann, H. and H. Mangold, *Induction of embryonic primordia by implantation of organizers from a different species*. 1923. *Int J Dev Biol*, 2001. **45**(1): p. 13-38.
11. Perrimon, N., C. Pitsouli, and B.-Z. Shilo, *Signaling Mechanisms Controlling Cell Fate and Embryonic Patterning*. *Cold Spring Harbor Perspectives in Biology*, 2012. **4**(8).
12. Dupont, S., et al., *Germ-layer specification and control of cell growth by Ectodermin, a Smad4 ubiquitin ligase*. *Cell*, 2005. **121**(1): p. 87-99.
13. Heasman, J., *Patterning the early Xenopus embryo*. *Development*, 2006. **133**(7): p. 1205-17.
14. Rhinn, M. and P. Dollé, *Retinoic acid signalling during development*. *Development*, 2012. **139**(5): p. 843-858.
15. Brockes, J.P., *Retinoids, homeobox genes, and limb morphogenesis*. *Neuron*, 1989. **2**.
16. Heather Marshall, A.M., Michele Studer, Heike Pöpperl, Robb Krumlauf, *Retinoids and Hox genes*. *The FASEB Journal* **10**(9): p. 969-978.
17. Gurdon, J.B., *The Developmental Capacity of Nuclei taken from Intestinal Epithelium Cells of Feeding Tadpoles*. *Development* 1962(10): p. 622-640.
18. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors*. *Cell*, 2006. **126**(4): p. 663-76.
19. Morgan, T.H., *CHROMOSOMES AND ASSOCIATIVE INHERITANCE*. *Science*, 1911. **36**: p. 636-638.
20. Lewis, E.B., *A gene complex controlling segmentation in Drosophila*. *Nature*, 1978. **276**.
21. W. McGinnis, M.S.L., E. Hafen, A. Kuriowa & W. J. Gehring, *A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes*. *Nature*, 1984. **308**.
22. Wieschaus, C.N.-V.E., *Mutations affecting segment number and polarity in Drosophila*. *Nature* 1980. **287**.
23. Chambers, S.M., et al., *Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling*. *Nat Biotechnol*, 2009. **27**(3): p. 275-80.
24. Lancaster, M.A., et al., *Cerebral organoids model human brain development and microcephaly*. *Nature*, 2013. **501**(7467): p. 373-9.
25. Sato, T., et al., *Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche*. *Nature*, 2009. **459**(7244): p. 262-5.
26. Bier, E., *Drosophila, the golden bug, emerges as a tool for human genetics*. *Nature Reviews Genetics*, 2005. **6**: p. 9-23.

REFERENCES

27. Linda Helfmorsa, Andrea Bomanb, 1, Livia Civitellib, Sangeeta Nathb, Linnea Sandinb, Camilla Janefjordb, Heather McCann, Henrik Zetterbergd, e, Kaj Blennowd, Glenda Hallidayc, Ann-Christin Brorsson, a, 1, Katarina Kågedalb, , 1, , *Protective properties of lysozyme on β -amyloid pathology: implications for Alzheimer disease*. *Neurobiology of Disease*, 2015. **83**: p. 122-133.
28. K.R. Kaun, R.A., Z. Maung, J. Hirsh, and U. Heberlein, *A Drosophila model for alcohol reward*. *Nature Neuroscience* 2011. **14**: p. 612-619.
29. Saikat Ghosh, A.S., 1 Sudip Mandal,2 and Lolitika Mandal1,*, *Active Hematopoietic Hubs in Drosophila Adults Generate Hemocytes and Contribute to Immune Response*. *Dev Cell*, 2015. **33**(4): p. 478–488.
30. Elif Eroglu, T.R.B., 1 Yanrui Jiang,2 Nidhi Saini,2 Catarina C.F. Homem,1 Heinrich Reichert,2 and Juergen A. Knoblich1, *SWI/SNF Complex Prevents Lineage Reversion and Induces Temporal Patterning in Neural Stem Cells*. *Cell*, 2014. **156**: p. 1259–1273.
31. Yamamoto, S., et al., *A drosophila genetic resource of mutants to study mechanisms underlying human genetic diseases*. *Cell*, 2014. **159**(1): p. 200-14.
32. Bischof, J., et al., *An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases*. *Proc Natl Acad Sci U S A*, 2007. **104**(9): p. 3312-7.
33. Weigmann, K., et al., *FlyMove--a new way to look at development of Drosophila*. *Trends Genet*, 2003. **19**(6): p. 310-1.
34. Skeath, J.B. and S.B. Carroll, *Regulation of proneural gene expression and cell fate during neuroblast segregation in the Drosophila embryo*. *Development*, 1992. **114**(4): p. 939-46.
35. Evans, I.R., et al., *Live imaging of Drosophila melanogaster embryonic hemocyte migrations*. *J Vis Exp*, 2010(36).
36. Tomancak, P., et al., *Systematic determination of patterns of gene expression during Drosophila embryogenesis*. *Genome Biol*, 2002. **3**(12): p. RESEARCH0088.
37. Rhee, D.Y., et al., *Transcription factor networks in Drosophila melanogaster*. *Cell Rep*, 2014. **8**(6): p. 2031-43.
38. *Evolutionary Developmental Biology of Invertebrates 5 Ecdysozoa III: Hexapoda*. Vol. 5. 2015: Springer.
39. Hartenstein, V. *Atlas of Drosophila Development*. 1993; Available from: <http://www.sdbonline.org/sites/fly/atlas/02stages3text.htm>.
40. Wolfgang Driever, C.N.-V., *A gradient of bicoid protein in Drosophila embryos*. *Cell*, 1988. **54**(1): p. 83-93.
41. *Organogenetic Gene Networks Genetic Control of Organ Formation*, ed. P.B. James Castelligair Hombría. 2016: Springer.
42. Parthier, C., et al., *Structure of the Toll-Spatzle complex, a molecular hub in Drosophila development and innate immunity*. *Proc Natl Acad Sci U S A*, 2014. **111**(17): p. 6281-6.
43. Gilbert, S.F., *Developmental Biology*. Vol. 6. 2000: Sinauer Associates Inc.
44. Mizutani, C.M. and E. Bier, *EvoD/Vo: the origins of BMP signalling in the neuroectoderm*. *Nat Rev Genet*, 2008. **9**(9): p. 663-77.
45. Estacio-Gomez, A. and F.J. Diaz-Benjumea, *Roles of Hox genes in the patterning of the central nervous system of Drosophila*. *Fly (Austin)*, 2014. **8**(1): p. 26-32.
46. Doe, C.Q., *Molecular markers for identified neuroblasts and ganglion mother cells in the Drosophila central nervous system*. *Development*, 1992. **116**(4): p. 855-863.
47. Jarman, A., *Neurogenesis in Drosophila*. 2013.
48. Azevedo, F.A., et al., *Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain*. *J Comp Neurol*, 2009. **513**(5): p. 532-41.
49. Bier, E. and E.M. De Robertis, *EMBRYO DEVELOPMENT. BMP gradients: A paradigm for morphogen-mediated developmental patterning*. *Science*, 2015. **348**(6242): p. aaa5838.
50. Hong, J.W., et al., *How the Dorsal gradient works: insights from postgenome technologies*. *Proc Natl Acad Sci U S A*, 2008. **105**(51): p. 20072-6.

REFERENCES

51. Y. Zhang, A.U., C. Fresquez, R. Holmgren, *Ectopic expression of either the Drosophila gooseberry-distal or proximal gene causes alterations of cell fate in the epidermis and central nervous system.* Development, 1994. **120**(5): p. 1151-1161.
52. Akam, M., *The molecular basis for metameric pattern in the Drosophila embryo.* Development, 1987.
53. Struhl, P.A.L.a.G., *Morphogens, Compartments, and Pattern Lessons from Drosophila.* Cell, 1996. **85**: p. 951-961.
54. Sprecher, S., *Drosophila Neural Development.* 2012.
55. von Ohlen, T. and C.Q. Doe, *Convergence of dorsal, dpp, and egfr signaling pathways subdivides the drosophila neuroectoderm into three dorsal-ventral columns.* Dev Biol, 2000. **224**(2): p. 362-72.
56. Ramos, C. and B. Robert, *msh/Msx gene family in neural development.* Trends Genet, 2005. **21**(11): p. 624-32.
57. Doe, Q.C.-L.a.C.Q., *Neuroblast Specification and Formation Regulated by wingless in the Drosophila CNS* Science, 1993. **261**: p. 1594-1597.
58. Bhat, K.M., *The patched signaling pathway mediates repression of gooseberry allowing neuroblast specification by wingless during Drosophila neurogenesis.* Development, 1996. **122**(9): p. 2921-2932.
59. Rogers, B.T., *Drosophila Embryo: Homeotic Genes in Specification of the Anterior-Posterior Axis*, in eLS. 2001, John Wiley & Sons, Ltd.
60. Richard S. Mann, M.A., *Hox proteins meet more Partners.* Current Opinion in Genetics & Development, 1998. **8**(4): p. 423-429.
61. Brody, T.B. *The Interactive Fly.* 1996; Available from: <http://www.sdbonline.org/sites/fly/aigfam/antpedia.htm>.
62. Krumlauf, R., *Hox genes in vertebrate development.* Cell, 1994. **78**(2): p. 191-201.
63. Mann, R.S.L., Katherine M. Joshi, Rohit, *Chapter 3 Hox Specificity.* 2009. **88**: p. 63-101.
64. Crocker, J., et al., *Low affinity binding site clusters confer hox specificity and regulatory robustness.* Cell, 2015. **160**(1-2): p. 191-203.
65. Thor, S. and J.B. Thomas, *The Drosophila islet gene governs axon pathfinding and neurotransmitter identity.* Neuron, 1997. **18**(3): p. 397-409.
66. Gabilondo, H., et al., *Neuronal Cell Fate Specification by the Convergence of Different Spatiotemporal Cues on a Common Terminal Selector Cascade.* PLoS Biol, 2016. **14**(5): p. e1002450.
67. van Meyel, D.J., et al., *Chip and Apterous Physically Interact to Form a Functional Complex during Drosophila Development.* Molecular Cell, 1999. **4**(2): p. 259-265.
68. Prakash K1, F.X., Engelberg D, Behal A, Parker CS., *dOct2, a Drosophila Oct transcription factor that functions in yeast.* PNAS, 1992. **89**(15): p. 7080-7084.
69. Baumgardt, M., et al., *Neuronal Subtype Specification within a Lineage by Opposing Temporal Feed-Forward Loops.* Cell, 2009. **139**(5): p. 969-982.
70. Daniel Karlsson, M.B., Stefan Thor*, *Segment-Specific Neuronal Subtype Specification by the Integration of Anteroposterior and Temporal Cues.* PLoS biology, 2010. **8**(5).
71. Hartenstein, V. and A. Wodarz, *Initial neurogenesis in Drosophila.* Wiley Interdiscip Rev Dev Biol, 2013. **2**(5): p. 701-21.
72. Torsten Bossing , G.U., Chris Q. Doe, Gerhard M. Technau, *The Embryonic Central Nervous System Lineages of Drosophila melanogaster.* Developmental Biology, 1996. **179**(1): p. 41-64.
73. Gunnar, E., et al., *sequoia controls the type I>0 daughter proliferation switch in the developing Drosophila nervous system.* Development, 2016. **143**(20): p. 3774-3784.
74. Wodarz, A. and W.B. Huttner, *Asymmetric cell division during neurogenesis in Drosophila and vertebrates.* Mechanisms of Development, 2003. **120**(11): p. 1297-1309.
75. Lai, S.L. and C.Q. Doe, *Transient nuclear Prospero induces neural progenitor quiescence.* Elife, 2014. **3**.

REFERENCES

76. Baumgardt, M., et al., *Global programmed switch in neural daughter cell proliferation mode triggered by a temporal gene cascade*. *Dev Cell*, 2014. **30**(2): p. 192-208.
77. Novotny, T., R. Eiselt, and J. Urban, *Hunchback is required for the specification of the early sublineage of neuroblast 7-3 in the Drosophila central nervous system*. *Development*, 2002. **129**(4): p. 1027-36.
78. Pinto-Teixeira, F., N. Konstantinides, and C. Desplan, *Programmed cell death acts at different stages of Drosophila neurodevelopment to shape the central nervous system*. *FEBS Lett*, 2016. **590**(15): p. 2435-53.
79. Lundell, M.J., et al., *The regulation of apoptosis by Numb/Notch signaling in the serotonin lineage of Drosophila*. *Development*, 2003. **130**(17): p. 4109-21.
80. Miguel-Aliaga, I. and S. Thor, *Segment-specific prevention of pioneer neuron apoptosis by cell-autonomous, postmitotic Hox gene activity*. *Development*, 2004. **131**(24): p. 6093-105.
81. Li, X., Z. Chen, and C. Desplan, *Temporal patterning of neural progenitors in Drosophila*. *Curr Top Dev Biol*, 2013. **105**: p. 69-96.
82. Brody, T. and W.F. Odenwald, *Programmed transformations in neuroblast gene expression during Drosophila CNS lineage development*. *Dev Biol*, 2000. **226**(1): p. 34-44.
83. Grosskortenhaus, R., et al., *Regulation of temporal identity transitions in Drosophila neuroblasts*. *Dev Cell*, 2005. **8**(2): p. 193-202.
84. Takako Isshiki, B.P., Scott Holbrook, Chris Q. Doe, *Drosophila neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny*. *Cell*, 2001. **106**(4): p. 511-521.
85. Kanai, M.I., M. Okabe, and Y. Hiromi, *seven-up Controls switching of transcription factors that specify temporal identities of Drosophila neuroblasts*. *Dev Cell*, 2005. **8**(2): p. 203-13.
86. Benito-Sipos, J., et al., *Seven up acts as a temporal factor during two different stages of neuroblast 5-6 development*. *Development*, 2011. **138**(24): p. 5311-20.
87. Jocelyn A. McDonald, S.H., Takako Isshiki, Joseph Weiss, Chris Q. Doe, and Dervla M. Mellerick, *Dorsoventral patterning in the Drosophila central nervous system the vnd homeobox gene specifies ventral column identity*. *Genes and Development*, 1998. **12**(22): p. 3603-3612.
88. Doe, C.Q., S. Fuerstenberg, and C.Y. Peng, *Neural stem cells: from fly to vertebrates*. *J Neurobiol*, 1998. **36**(2): p. 111-27.
89. Konstantinides, N., A.M. Rossi, and C. Desplan, *Common temporal identity factors regulate neuronal diversity in fly ventral nerve cord and mouse retina*. *Neuron*, 2015. **85**(3): p. 447-9.
90. Homem, C.C. and J.A. Knoblich, *Drosophila neuroblasts: a model for stem cell biology*. *Development*, 2012. **139**(23): p. 4297-310.
91. Yeo, W. and J. Gautier, *Early neural cell death: dying to become neurons*. *Dev Biol*, 2004. **274**(2): p. 233-44.
92. Park, D., et al., *The Drosophila basic helix-loop-helix protein DIMMED directly activates PHM, a gene encoding a neuropeptide-amidating enzyme*. *Mol Cell Biol*, 2008. **28**(1): p. 410-21.
93. Alon, U., *Network motifs: theory and experimental approaches*. *Nat Rev Genet*, 2007. **8**(6): p. 450-61.
94. Stratmann, J., et al., *Neuronal cell fate diversification controlled by sub-temporal action of Kruppel*. *Elife*, 2016. **5**.
95. Erwin, D.H. and E.H. Davidson, *The evolution of hierarchical gene regulatory networks*. *Nat Rev Genet*, 2009. **10**(2): p. 141-8.
96. Wenick, A.S. and O. Hobert, *Genomic cis-regulatory architecture and trans-acting regulators of a single interneuron-specific gene battery in C. elegans*. *Dev Cell*, 2004. **6**(6): p. 757-70.
97. Thor, S., et al., *A LIM-homeodomain combinatorial code for motor-neuron pathway selection*. *Nature*, 1999. **397**(6714): p. 76-80.
98. Allan, D.W., et al., *Regulators acting in combinatorial codes also act independently in single differentiating neurons*. *Neuron*, 2005. **45**(5): p. 689-700.

REFERENCES

99. Allan, D.W., et al., *Specification of neuropeptide cell identity by the integration of retrograde BMP signaling and a combinatorial transcription factor code*. *Cell*, 2003. **113**(1): p. 73-86.
100. Briscoe, J., et al., *A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube*. *Cell*, 2000. **101**(4): p. 435-45.
101. Allan, D.W. and S. Thor, *Transcriptional selectors, masters, and combinatorial codes: regulatory principles of neural subtype specification*. *Wiley Interdiscip Rev Dev Biol*, 2015. **4**(5): p. 505-28.
102. Baumgardt, M., et al., *Specification of Neuronal Identities by Feedforward Combinatorial Coding*. *PLoS Biology*, 2007. **5**(2): p. e37.
103. Lou, S., et al., *Incoherent feed-forward regulatory loops control segregation of C-mechanoreceptors, nociceptors, and pruriceptors*. *J Neurosci*, 2015. **35**(13): p. 5317-29.
104. Bivik, C., et al., *Novel Genes Involved in Controlling Specification of Drosophila FMRFamide Neuropeptide Cells*. *Genetics*, 2015. **200**(4): p. 1229-44.
105. De Graeve, F., et al., *The ladybird homeobox genes are essential for the specification of a subpopulation of neural cells*. *Dev Biol*, 2004. **270**(1): p. 122-34.
106. DESPLAN, F.P.-T.A.C., *Re-utilization of a transcription factor*. *Elife*, 2016. **5**.
107. Griffiths AJF, M.J., Suzuki DT, et al., *An Introduction to Genetic Analysis*. Vol. 7th edition. 2000: W. H. Freeman.
108. Mounia Lagha, 5 Jacques P. Bothma,2,5 Emilia Esposito,1 Samuel Ng,1 Laura Stefanik,3 Chiahao Tsui,1 Jeffrey Johnston,4 and D.S.G. Kai Chen, 3 Julia Zeitlinger,4 and Michael S. Levine1,* , *Paused Pol II Coordinates Tissue Morphogenesis in the Drosophila Embryo*. *Cell*, 2013. **153**: p. 976–987.
109. Zabidi, M.A. and A. Stark, *Regulatory Enhancer-Core-Promoter Communication via Transcription Factors and Cofactors*. *Trends Genet*, 2016.
110. Goodrich, J.A. and R. Tjian, *Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation*. *Nat Rev Genet*, 2010. **11**(8): p. 549-58.
111. Kadonaga, J.T., *Perspectives on the RNA polymerase II core promoter*. *Wiley Interdiscip Rev Dev Biol*, 2012. **1**(1): p. 40-51.
112. Spitz, F. and E.E. Furlong, *Transcription factors: from enhancer binding to developmental control*. *Nat Rev Genet*, 2012. **13**(9): p. 613-26.
113. Pennacchio, L.A., et al., *Enhancers: five essential questions*. *Nat Rev Genet*, 2013. **14**(4): p. 288-95.
114. Pfeiffer, B.D., et al., *Refinement of tools for targeted gene expression in Drosophila*. *Genetics*, 2010. **186**(2): p. 735-55.
115. Julian Banerji, S.R., Walter Schaffner, *Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences*. *Cell*, 1981. **27**(1): p. 299-308.
116. Arnim Jenett, G.M.R., 1,* Teri-T B. Ngo,1 David Shepherd,2 Christine Murphy,1 Heather Dionne,1 Barret D. Pfeiffer,1 Amanda Cavallaro,1 Donald Hall,1 Jennifer Jeter,1 Nirmala Iyer,1 Dona Fetter,1 Joanna H. Hausenfluck,1 Hanchuan Peng,1 Eric T. Trautman,1 Rob Svirskas,1 Eugene W. Myers,1 Zbigniew R. Iwinski,3 Yoshinori Aso,1 Gina M. DePasquale,1 Adrianne Enos,1 Phusun Hulamm,1 Shing Chun Benny Lam,1 Hsing-Hsi Li,1 Todd R. Lavery,1 Fuhui Long,1 Lei Qu,1 Sean D. Murphy,1 Konrad Rokicki,1 Todd Safford,1 Kshiti Shaw,1 Julie H. Simpson,1 Allison Sowell,1 Susana Tae,1 Yang Yu,1 and Christopher T. Zugates, *A GAL4-Driver Line Resource for Drosophila Neurobiology*. *Cell Reports*, 2012. **2**(4): p. 991-1001.
117. Andersson, R., et al., *An atlas of active enhancers across human cell types and tissues*. *Nature*, 2014. **507**(7493): p. 455-61.
118. Zaret, K.S. and J.S. Carroll, *Pioneer transcription factors: establishing competence for gene expression*. *Genes Dev*, 2011. **25**(21): p. 2227-41.
119. Soufi, A., et al., *Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming*. *Cell*, 2015. **161**(3): p. 555-68.

REFERENCES

120. Berndt, A.J., et al., *Gene Regulatory Mechanisms Underlying the Spatial and Temporal Regulation of Target-Dependent Gene Expression in Drosophila Neurons*. PLoS Genet, 2015. **11**(12): p. e1005754.
121. Calo, E. and J. Wysocka, *Modification of enhancer chromatin: what, how, and why?* Mol Cell, 2013. **49**(5): p. 825-37.
122. Schaukowitz, K., et al., *Enhancer RNA facilitates NELF release from immediate early genes*. Mol Cell, 2014. **56**(1): p. 29-42.
123. Lam, M.T., et al., *Enhancer RNAs and regulated transcriptional programs*. Trends Biochem Sci, 2014. **39**(4): p. 170-82.
124. Kagey, M.H., et al., *Mediator and cohesin connect gene expression and chromatin architecture*. Nature, 2010. **467**(7314): p. 430-5.
125. Ong, C.T. and V.G. Corces, *CTCF: an architectural protein bridging genome topology and function*. Nat Rev Genet, 2014. **15**(4): p. 234-46.
126. Kvon, E.Z., et al., *Genome-scale functional characterization of Drosophila developmental enhancers in vivo*. Nature, 2014. **512**(7512): p. 91-5.
127. Li, W., D. Notani, and M.G. Rosenfeld, *Enhancers as non-coding RNA transcription units: recent insights and future perspectives*. Nat Rev Genet, 2016. **17**(4): p. 207-23.
128. Hagman, J., et al., *Cloning and functional characterization of early B-cell factor, a regulator of lymphocyte-specific gene expression*. Genes & Development, 1993. **7**(5): p. 760-773.
129. James Hagman, M.J.G., Haishan Lin and Rudolf Grosschedl, *EBF contains a novel zinc coordination motif and multiple dimerization and transcriptional activation domains*. The EMBO Journal 1995. **14**(12): p. 2907-2916.
130. de Taffin, M., et al., *Genome-Wide Mapping of Collier In Vivo Binding Sites Highlights Its Hierarchical Position in Different Transcription Regulatory Networks*. PLoS One, 2015. **10**(7): p. e0133387.
131. Erceg, J., et al., *Subtle changes in motif positioning cause tissue-specific effects on robustness of an enhancer's activity*. PLoS Genet, 2014. **10**(1): p. e1004060.
132. Emma K. Farley, K.M.O., Wei Zhang, Alexander J. Brandt, Daniel S. Rokhsar, Michael S. Levine, *Suboptimization of developmental enhancers*. Science, 2015. **350**(6258).
133. Cannavo, E., et al., *Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks*. Curr Biol, 2016. **26**(1): p. 38-51.
134. Joung-Woo Hong, D.A.H., Michael S. Levine, *Shadow Enhancers as a Source of Evolutionary Novelty*. Science, 2008. **321**(5894).
135. Barolo, S., *Shadow enhancers: frequently asked questions about distributed cis-regulatory information and enhancer redundancy*. Bioessays, 2012. **34**(2): p. 135-41.
136. Kim, T.K., et al., *Widespread transcription at neuronal activity-regulated enhancers*. Nature, 2010. **465**(7295): p. 182-7.
137. Hsin, J.P. and J.L. Manley, *The RNA polymerase II CTD coordinates transcription and RNA processing*. Genes Dev, 2012. **26**(19): p. 2119-37.
138. Yu, M., et al., *RNA polymerase II-associated factor 1 regulates the release and phosphorylation of paused RNA polymerase II*. Science, 2015. **350**(6266): p. 1383-6.
139. Levine, M., *Paused RNA polymerase II as a developmental checkpoint*. Cell, 2011. **145**(4): p. 502-11.
140. Adelman, K. and J.T. Lis, *Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans*. Nat Rev Genet, 2012. **13**(10): p. 720-31.
141. Bjorklund, A. and S.B. Dunnett, *Dopamine neuron systems in the brain: an update*. Trends Neurosci, 2007. **30**(5): p. 194-202.
142. Park, D., et al., *Mapping peptidergic cells in Drosophila: where DIMM fits in*. PLoS One, 2008. **3**(3): p. e1896.
143. Doe, C.Q., *Molecular markers for identified neuroblasts and ganglion mother cells in the Drosophila central nervous system*. Development, 1992. **116**(4): p. 855-63.

REFERENCES

144. Broadus, J., et al., *New neuroblast markers and the origin of the aCC/pCC neurons in the Drosophila central nervous system*. *Mech Dev*, 1995. **53**(3): p. 393-402.
145. Isshiki, T., M. Takeichi, and A. Nose, *The role of the msh homeobox gene during Drosophila neurogenesis: implication for the dorsoventral specification of the neuroectoderm*. *Development*, 1997. **124**(16): p. 3099-109.
146. Tsuji, T., E. Hasegawa, and T. Isshiki, *Neuroblast entry into quiescence is regulated intrinsically by the combined action of spatial Hox proteins and temporal identity factors*. *Development*, 2008. **135**(23): p. 3859-69.
147. Dubois, L., et al., *Collier transcription in a single Drosophila muscle lineage: the combinatorial control of muscle identity*. *Development*, 2007. **134**(24): p. 4347-55.
148. Eade, K.T., et al., *Developmental transcriptional networks are required to maintain neuronal subtype identity in the mature nervous system*. *PLoS Genet*, 2012. **8**(2): p. e1002501.
149. Gao, H., et al., *Opposing effects of SWI/SNF and Mi-2/NuRD chromatin remodeling complexes on epigenetic reprogramming by EBF and Pax5*. *Proc Natl Acad Sci U S A*, 2009. **106**(27): p. 11258-63.
150. Fuxman Bass, J.I., et al., *Human gene-centered transcription factor networks for enhancers and disease variants*. *Cell*, 2015. **161**(3): p. 661-73.
151. Sur, I. and J. Taipale, *The role of enhancers in cancer*. *Nat Rev Cancer*, 2016. **16**(8): p. 483-93.
152. Vaharautio, A. and J. Taipale, *Cancer. Cancer by super-enhancer*. *Science*, 2014. **346**(6215): p. 1291-2.
153. Duffy, J.B., *GAL4 system in Drosophila: a fly geneticist's Swiss army knife*. *Genesis*, 2002. **34**(1-2): p. 1-15.

Papers

The articles associated with this thesis have been removed for copyright reasons. For more details about these see:

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-133225>