Volodymyr Averkov

Inaugural-Dissertation

Tissue-specific roles of Snail-like transcription factors in the embryogenesis of *Drosophila melanogaster*

Köln

2015

Tissue-specific roles of Snail-like transcription factors in the embryogenesis of *Drosophila melanogaster*

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Volodymyr Averkov

Aus Charkiw (Ukraine)

Berichterstatter: Prof. Dr. Maria Leptin

                            Prof. Dr. Siegfried Roth

Datum der Disputation: 16. Juni 2015

O snail,

Slowly climb Mount Fuji

Up to the peak!

(Kobayashi Issa, 1763 - 1828)

Table of Content

1 Frequently used abbreviations

AEL	after egg lay
bp	base pairs
ChIP	chromatin immunoprecipitation
ChIP-seq	ChIP with following sequence analysis
CNS	central nervous system
CRM	*cis*-regulatory modules
DamID	DNA adenine methyltransferase identification
EMT	epithelial-mesenchymal transition
EtOH	ethanol
fig.	figure
GFP	green fluorescent protein
GMC	ganglion mother cell
GP	guinea pig (source of immune serum)
hpf	hours post fertilization
HRP	horseradish peroxidase
MP2	midline precursor 2: a neuroblast lineage
NaAc	Sodium acetate
Nb	neuroblast
NGS	next generation sequencing
PNS	peripheral nervous system
PWM	position weight matrix
TF	transcription factor
TSS	transcription start site
UTR	untranslated region
V5, HA	oligopeptide epitops

The genes and proteins were abbreviated according to the standard FlyBase abbreviation list.

2 Introduction

Transcription factors (TFs) are key regulators of gene expression. TF binding to chromatin depends on multiple conditions, such as tissue identity, developmental stage or stress conditions to name a few, so called differential binding of TFs. To gain an understanding of the mechanisms of differential binding of TFs to DNA is one of the important aims of present day functional genomics and systems biology (Spitz and Furlong, 2012).

Extensively studied *Drosophila* development provides multiple examples of differential binding of TFs (Shlyueva et al., 2014a; Yáñez-Cuna et al., 2012). Multiple novel assays, developed over the past years, allowed a description of TF binding events on a genomic level (Gilchrist et al., 2009; Shlyueva et al., 2014b). However, the studies performed on *Drosophila* mainly compared TF binding profiles in a single tissue at different developmental stages. The TF Snail provides an opportunity to compare binding profiles of a TF in different tissues without complicated tissue dissection procedures. At stages 4-7 of *Drosophila* embryonic development (2-4 hours post fertilization, hpf at 25°C) Snail is expressed in mesoderm (Leptin and Grunewald, 1990). Later at stages 8-11 (4-6 hpf at 25°C) Snail is expressed in the neuroblasts of the developing ventral nerve cord (Ashraf et al., 1999). The functions of Snail in mesoderm are well known and mesodermal target genes have been described (Kölsch et al., 2007; Leptin, 1991; Macarthur et al., 2009; Rembold et al., 2014), but the functions of Snail in neuroblasts remain obscure. In mesoderm Snail performs its functions without cooperation partners. In neuroblasts Snail cooperates and acts redundantly with the TFs Worniu and Escargot, paralogs of Snail (Ashraf and Ip, 2001; Ashraf et al., 1999; Cai et al., 2001).

In this work I analyze and compare the binding profiles of Snail in mesoderm and neuroblasts, and try to decipher the mechanisms of tissue-specific cooperation of Snail, Worniu and Escargot.

2.1   Binding of TFs to DNA

Each TF specifically binds a particular short DNA sequence (usually in the range of 6 to 15 bp), which are called TF binding motifs. TF binding motifs tolerate certain levels of sequence variability. The TF Snail binds to an extended E-box motif with the core CAGGTG. The CA nucleotides at position 1 and 2 are invariant, i.e. their mutation completely abolishes Snail´s binding, whereas G at the 6th position can be substituted by A without a major impact on binding affinity. The variant sequence CAGGTA was thus described as an "alternative Snail motif" (Mauhin et al., 1993; Rembold et al., 2014).

TF binding motifs can be represented as consensus sequences or as position weight matrices (PWM), calculated by comparison of multiple experimentally proven binding sites for a given TF. A PWM calculates the probability of a given nucleotide at a given position as well as the information content of this position. The information content (calculated in bits from 0 to 2) represents the sequence conservation: positions with a high tolerance for substitutions have low information content in contrast to highly conserved positions. Positions at which only one specific nucleotide is allowed have the maximum information content of 2. The PWM can be visualized with so-called sequence logos, in which the height of the corresponding symbol in a stack shows the relative frequency of this base at this position. The sequence logos of the PWM of Snail and Twist binding motifs are shown in the fig. 1 (Rembold et al., 2014; Stormo, 2000).
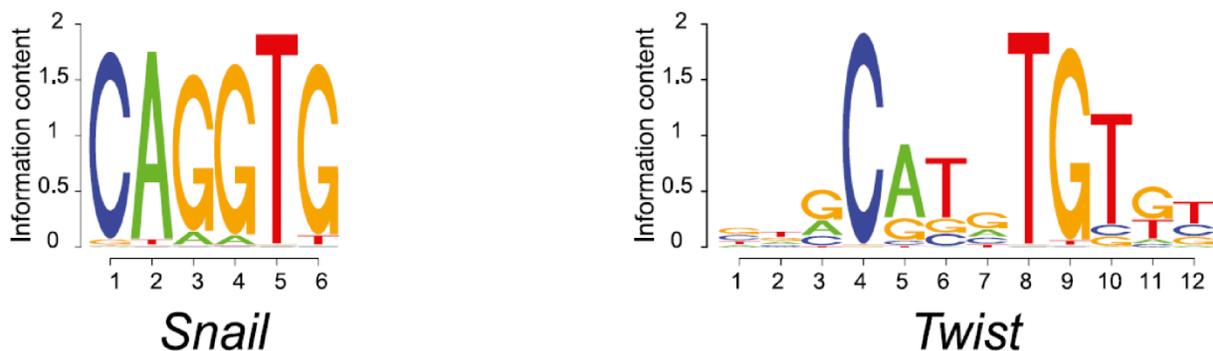
Figure 1. Examples of TF binding motifs. The position weight matrices of the binding motifs for Snail and Twist are represented as sequence logos. The y-axis specifies the information content of a specific position in bits, where 2 indicates that only a single nucleotide can occur at this position. The total height of the stack at one position specifies the information content of this position and the height of a nucleotide symbol within a stack indicates the relative frequency of this nucleotide at this position. Image modified from (Rembold et al., 2014).

The affinity of a TF for to a given DNA fragment depends on two main conditions. First, on the the DNA sequence of the fragment. Second, on post-translational modifications, homo- or hetero-dimerization of the TF. For instance, Twist homodimers activate the transcription of myogenic genes, such as *Mef2*. In contrast, Twist/Daughterless heterodimers recognize a different set of targets and repress somatic myogenic genes (Castanon et al., 2001; Wong et al., 2008).

Often multiple TF binding sites are clustered in a small DNA region into a non-coding DNA sequence, called enhancer. Enhancers are commonly identified as the compact non-coding DNA regions, whose deletion reduce or drastically changes expression of a gene. Enhancers are located within a regulated gene, in the upstream or downstream non-transcribed region of a target gene, or can control expression of a distant gene skipping one or two other genes (Kvon et al., 2014).

The word "enhancer" reflects the "enhancing" effect of this sequence on transcription from a promoter that was first described for a 72-bp fragment of the *SV40* virus genome, which drastically enhanced the transcription of a reporter gene in *HeLa* cells (Banerji et al., 1981). Since enhancers can also mediate repression of a gene the term "cis-regulatory modules" (CRMs) is often used as a synonym or as a replacement for the term "enhancer".

2.1.1 Tissue-specific binding of TFs to enhancers

Even though TFs might be present in cells, they do not necessarily bind to each of their target binding sites: a histone modification or a prior binding of additional TFs is required to make enhancers accessible. Binding of a TF to enhancers that depends on the cellular context is called "context-specific binding" (reviewed in Spitz and Furlong, 2012). Context-specific binding of TFs and context-specific enhancer activation in an organism may occur in a temporarily and spatially controlled manner.

Multiple studies monitored binding of a TF to enhancers in different cell contexts in multiple model organisms at a genome-wide scale: in yeasts under different growth conditions (Buck and Lieb, 2006; Zeitlinger et al., 2003), in *C.elegans* at different developmental stages (Zhong et al., 2010), in *Drosophila* (Sandmann et al., 2007; Wilczyński and Furlong, 2010; Zinzen et al., 2009), in mice and human (Lin et al., 2010; Palii et al.,

2010). Spatiotemporal changes in TF binding to DNA, not the expression of the TF itself, correlate with dynamic patterns of target expression (Wilczyński and Furlong, 2010).

To bind and activate an enhancer a TF often requires co-binding of other TFs (Sandmann et al., 2007; Wilczynski and Furlong, 2010; Wilczynski et al., 2012; Zinzen and Furlong, 2008). In turn, the TFs require the binding motifs included into the enhancer´s sequence to bind it. Thus, the sequence of enhancer, in particular TF binding motifs within the enhancer, specifies the spatio-temporal pattern of enhancer activation. The motif content of an enhancer can therefore be used to predict its activity.

Several recent studies were dedicated to the *in silico* prediction of spatio-temporal patterns of enhancer expression based on the sequence of enhancer elements. Machine learning algorithms have been developed that analyze multiple enhancer elements, and compare their motif composition to data on the of enhancer´s expression pattern. These algorithms are able to effectively predict TF binding profiles and activity of genomic enhancers, mutant enhancers and newly designed enhancers (Yáñez-Cuna et al., 2012; Zinzen et al., 2009).

2.2 Methods that allow the identification of enhancers and TF binding sites

Historically the first enhancers were discovered and characterized using "enhancer trap" procedures. Genomic DNA fragments, usually upstream of the transcriptional start site, were ligated into a plasmid containing a reporter gene. After transgenesis or transient cell transfection the activity of the reporter gene was measured to identify those genomic fragments that displayed enhancer activity. However, testing enhancer properties of multiple DNA sequences surrounding a gene of interest required time and effort and was therefore restricted to a limited number of enhancers (reviewed in M. Hoy, 2013). Later numerous approaches have been developed to allow enhancer identification and characterization at the genomic scale: there are multiple advanced versions of the "enhancer trap" procedure like immunoprecipitation-based assays or predictions based on DNA accessibility in nuclei (reviewed in Shlyueva et al., 2014b).

Active enhancers, as well as other active regions of DNA, are generally more accessible than non-active genomic regions. This characteristic of chromatin is used in

methods such as DNase-seq and FAIRE-seq (formaldehyde-assisted identification of regulatory elements followed by deep sequencing) chromatin is cross-linked and subjected to DNase treatment. Active regions, such as actively transcribed genes, promoters and enhancers are accessible and cut by the enzyme. Tightly packed inactive DNA, in contrast, is protected by proteins and therefore remains intact. Next generation sequencing analysis is then applied to generate a map of the frequently cut regions, i.e. highly active regions of DNA, which include actively transcribed genes, active promoters and enhancers (reviewed in Shlyueva et al., 2014b).

The DamID (DNA adenine methyltransferase identification) technique employs the expression of a protein chimera, consisting of a DNA-binding protein fused to the enzyme domain of bacterial DNA adenine methyltransferase (Dam). Being bound to DNA, the protein chimera methylates adenines within GATC sequences in the neigbouring DNA. This DNA modification is not characteristic for eukaryotes and therefore specifically marks binding sites of the chimeric protein. However, the DamID methods works with an average resolution of ~1 kb (average TF binding site is 5-15 bp long) and therefore cannot be used to identify TF binding sites with precise resolution (van Steensel and Henikoff, 2000).

Numerous bioinformatic scripts were developed to search for the genomic regions that satisfy the criteria specific for enhancers. The programs search features like 'non-protein-coding', 'conserved', 'enriched for TF binding sites'. Other programs analyze the distribution of multiple TFs on DNA, using datasets of TF binding available from the experimental work performed previously, compare TF distribution to the localization of RNA polymerase molecules and to the enrichment of enhancer-specific histone modifications. Despite machine-based methods being quick and effective, they always require subsequent *in vivo* tests of the predicted regions (reviewed in Shlyueva et al., 2014).

2.2.1 ChIP techniques and their modifications

Chromatin immunoprecipitation (ChIP) is a method extensively used today to map the interaction of proteins with DNA with a genome-wide scale. The technique became popular because of its simple design and the high level of precision and resolution. The technique is based on affinity purification of proteins that were crosslinked *in vivo* to the genomic regions they are bound to (Gilchrist et al., 2009; Ho et al., 2011). It has been

successfully applied to study TF binding, histone modifications or chromatin modifying enzymes, epigenetic marks and the location of RNA Polymerase II (Park, 2009). In short, after isolation of nuclei and fragmentation of chromatin, antibodies are applied to selectively precipitate chromatin fragments, cross-linked to the protein or histone modification of interest. The purified DNA fragments are mapped to the genome. ChIP-seq (ChIP with following sequence analysis) employs sequencing techniques to read the precipitated DNA and is able to identify the binding sites with ~10-50 bp resolution (Park, 2009; fig. 2).
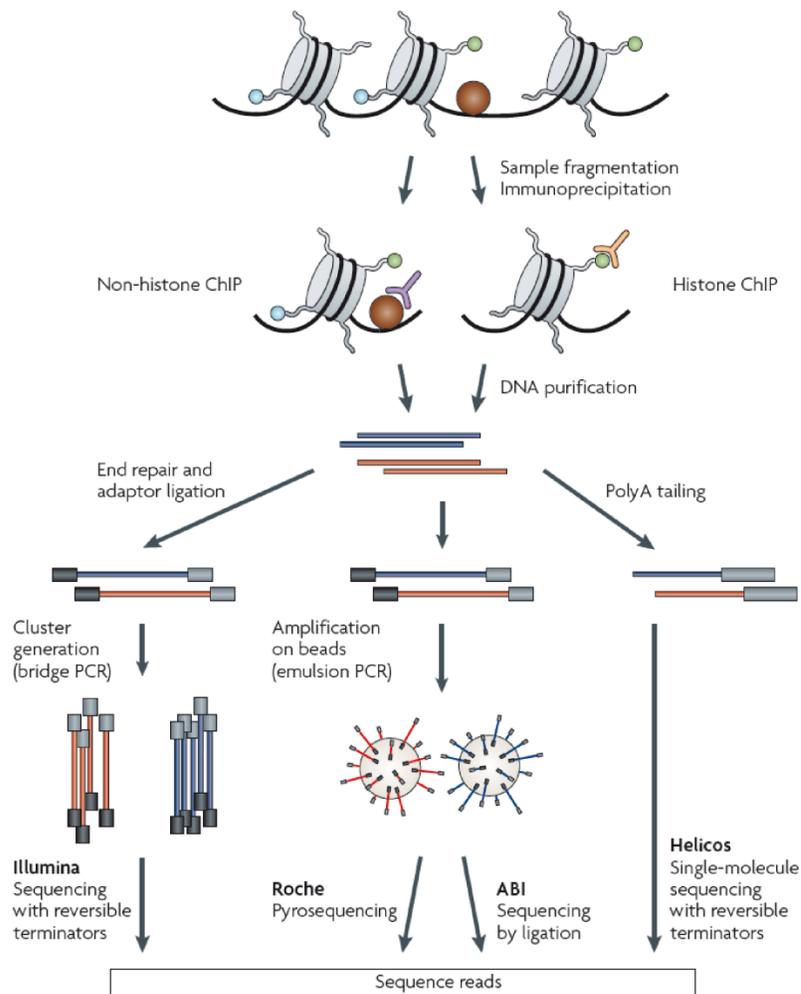
Figure 2. Overviewed ChIP-seq pipeline (from Park, 2009). Main steps of the workflow protocol: chromatin fragmentation, immunoprecipitation, DNA purification, library preparation and following sequence analysis. Several possible DNA sequencing approaches are illustrated: Illumina sequencing with reversible terminators, Roche pyrosequencing, ABI sequencing by ligation, Helicos single-molecule sequencing with reversible terminators. As Illumna technology was used for sequence analysis in the current work, it is going to be explained in the following chapters.

ChIP followed by hybridization to DNA microarrays (ChIP-chip) in contrast, identifies TF-binding sites with a resolution of several hundred bps. Several modifications to the basic technique have been developed. Bits-ChIP for instance allows the analysis of tissue-specific binding or chromatin states. The cells of a single tissue are fluorescently labeled with a fluorescent marker and sorted from the entire body lysate by FACS (Bonn et al., 2012).

2.2.2 Generation of tagged proteins using *FlyFosmid* libraries

The availability of high quality antibodies is crucial for the success of a ChIP experiment. The antibody should recognize the protein of interest with high specificity and show minimal cross-reactivity to other proteins. An alternative method is the use of ChIP-quality antibodies against protein tags, such as V5, HA, FLAG or GFP. Multiple recombination techniques are available to produce tagged versions of TFs. For instance, fosmid recombination allows the insertion of a tag-coding sequence into the gene, while preserving the gene's genomic environment and expression pattern.

Fosmids are bacterial vectors, designed to carry large DNA fragments (up to 50 kb), for instance fragments of genomes. *FlyFosmid* is a library of *Drosophila* genomic fragments cloned into the fosmid vector, specially designed for *Drosophila* transgenesis (*pFlyFos*) (Ejsmont et al., 2009). The library covers ~96% of the genes of *Drosophila*, with an average size of the genomic fragments of 36 kb, the majority ranging between 20 kb and 50 kb. The *pFlyFos* vector contains DNA sequences required for all necessary fosmid manipulations in bacteria (such as an origin of replication and antibiotic resistance gene), as well as the elements used for *Drosophila* transgenesis. The vector contains the *attB* recombination site for φC31-mediated integration into *attP* sites present in *Drosophila* landing lines and a gene encoding red fluorescent protein (*dsRed*) under control of a *Drosophila* eye-specific enhancer as a selection marker (Ejsmont et al., 2009).

FlyFosmid DNA can be easily manipulated. In this work I used the homologous recombination system *Red* derived from λ prophage (Copeland et al., 2001). The *Red* system requires specially developed bacterial strains, such as the *EL250* strain, containing the *red* recombination genes *exo*, *bet*, and *gam* integrated into their genome. The bacteria must be electroporated first with a *FlyFosmid* that carries the gene of interest and later with a linear

DNA molecule, the recombination cassette. Two 50 bp long homology arms at the termini of the recombination cassette direct the recombination event to the identical 50 bp homology regions on the *FlyFosmid*. The expression of *exo*, *bet* and *gam*, normally repressed by temperature-sensitive repressor cI857, is activated upon heat-shock. The recombination system recognizes homology arms on the cassette as double-stranded DNA breaks and starts repair (Lee et al., 2001). The 5'-3' exonuclease Exo creates single-stranded overhangs on the linear DNA of the recombination cassette. Bet protects the overhangs and assists with the recombination process. Gam prevents degradation of linear DNA by inhibiting the *E. coli* RecBCD protein. Bacterial reparation system exchanges fosmid DNA between homology regions by the recombination cassette DNA. Recombination is very efficient and requires homology arms as long as only 30-50 bp (Copeland et al., 2001). Depending on the cassette design, the recombination could be used to insert, delete or modify any piece of *FlyFosmid* DNA.

The recombination cassette usually contains an antibiotic resistance gene, as a selection marker for recombination. The marker is framed by *FRT* sites, so that it can be removed after a successful recombination. The *flippase* gene is present in the genome of the *EL250* strain under control of the arabinose operator.

2.2.3   Sequencing techniques applied in ChIP-seq

In the ChIP-seq workflow a ChIP DNA sample must be sequenced and the sequencing reads must be mapped onto the genome. Conventionally these steps are done using next-generation sequencing (NGS) techniques, designed to sequence a large number of DNA molecules in parallel. Several different high-throughput sequencing techniques are available, of which Roche 454, SOLiD, Illumna sequencing is one of the most commonly used systems (fig. 3). The technique relies on sequencing by synthesis (SBS) with fluorescently labeled terminator bases. The fragments can bend over and the adaptor at the free end basepairs with the primers on the slide (bridge amplification). The DNA is amplified with unlabeled nucleotides to produce local DNA colonies of identical molecules. The sequencing cycle starts with the addition of fluorescently labeled reversible terminator bases (ddA, ddT, ddG, ddC) along with primers and polymerase. After the termination base has been incorporated a microscopic image of the entire slide is taken and stored. The color of the

DNA colony identifies the base pair that has been incorporated. As each DNA molecule within one cluster has incorporated the same nucleotide, the signal becomes strong enough to be recorded. To start the new cycle the reversible terminator base is enzymatically cleaved to allow the incorporation and recording of the next base. The color sequence of each colony on the slide can thus be translated into a nucleotide sequence (fig. 3.; Bennett, 2004; Bentley et al., 2008; Meyer and Kircher, 2010; Shen et al., 2005).
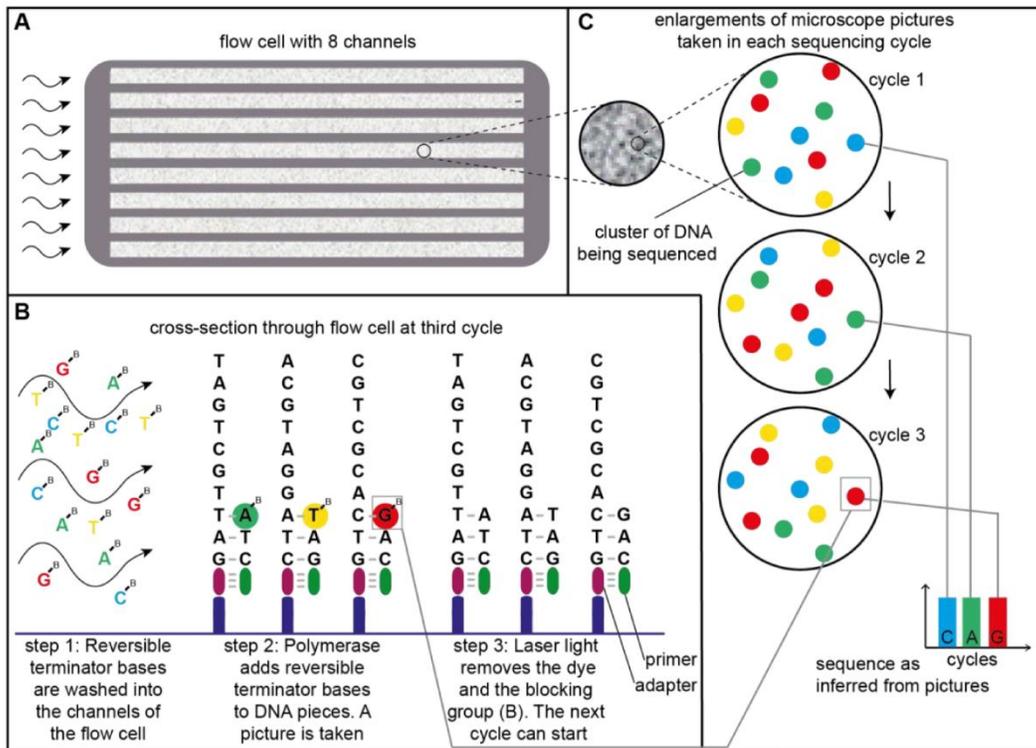
Figure 3. General principles of the Illumina sequencing technology.
a. The Illumna flow cell with eight channels, a platform for the sequencing reaction. DNA fragments are linked to the cell via the DNA linker fragments and replicated to produce colony clusters of the identical DNA fragments. One cluster is represented by a single DNA fragment in b., or by a color circle in c.
b. Schematic representation of the sequencing cycle. One cycle is required to sequence a single base pair. Sequencing of multiple DNA clusters is performed in parallel. First step: DNA polymerase adds a reversible terminator base to the DNA fragment, linked to the flow cell via DNA adaptor. Terminator base blocks following DNA synthesis, therefore only one base pair is added to the growing chain. Each kind of the terminator base (A, T, G, C) is stained with a specific fluorescent dye. Second step: the excess base pairs are washed away and a microscopic image of the cell is taken. Third step: The terminator base is substituted by a usual base and a fluorescent dyes are chemically removed to start a new sequencing cycle.
c. Schematic representation of a flow cell under high enlargement. During each sequencing cycle each DNA colony cluster gains a certain color that indicates the base pair incorporated into the growing DNA chain. Following of the DNA cluster colors throughout all sequencing reaction reveals the sequence of the DNA fragments in the cluster. Figure modified from *tucf-genomics.tufts.edu* with the kind help of Dominik Stappert.

2.2.4   **Bioninformatic approaches used in ChIP-seq**

NGS Illumna allows high throughput sequencing and the data generated (dataset of reads) is in the order of several millions of successfully sequenced DNA fragments. These short-sequence reads (usually 100-300 bp) must be first mapped to the genome (fig. 4). Thereafter the genomic coverage of the reads must be analyzed to evaluate the quality of the sequenced ChIP library: genome coverage, complexity and amount of background. If the ChIP was successful, TF-bound genomic fragments are going to be abundantly represented (enriched) in the library, whereas the majority of genome would be underrepresented (depleted). Since the DNA reads are going to be read from both DNA strands, the enrichment peaks mapped to genome are going to be shifted to the 3´ direction of a correspondent DNA strand.

After the reads are mapped, the decision must be made, which genome regions are going to be selected as "TF bound peaks" and which parts of genome are going to be "background". Model-based analysis of ChIP-seq (MACS): were used to analyze the ChIP data produced in this project (Feng et al., 2011, 2012; Zhang et al., 2008).

MACS workflow pipeline consists of the following steps: initially the program removes redundant reads generated by PCR over-amplification by excluding precisely coincided reads. Taking 1000 highest-enrichment regions the program calculates the base pair shift between the enrichment regions found independently upon + and - DNA strands, calculates the average base pair shifts for genome database, and shifts all reads to average shift/2 distance. After the reads are shifted, the overlapping enriched peaks are merged and the highest pileup of the reads is characterized as the peak summit.

MACS calculates λ_{BG}, a parameter of the Poisson model of reads distribution in the genome. For each given peak summit the program estimates λ_{local}, the maximal value out of (λ_{BG}, λ_{1k}, λ_{5k}, λ_{10k}), where λ_{1k}, λ_{5k}, λ_{10k} are the distribution parameters, calculated for 1 kb, 5 kb and 10 kb fragments, centered in the summit of analyzed peak. Comparing λ_{local} between experiment and control samples the program evaluates the local biases of read distributions and thus calculates p-values of each candidate peak to remove potential false positives. Candidate peaks with p-values below a defined threshold are called (Feng et al., 2011, 2012; Zhang et al., 2008).
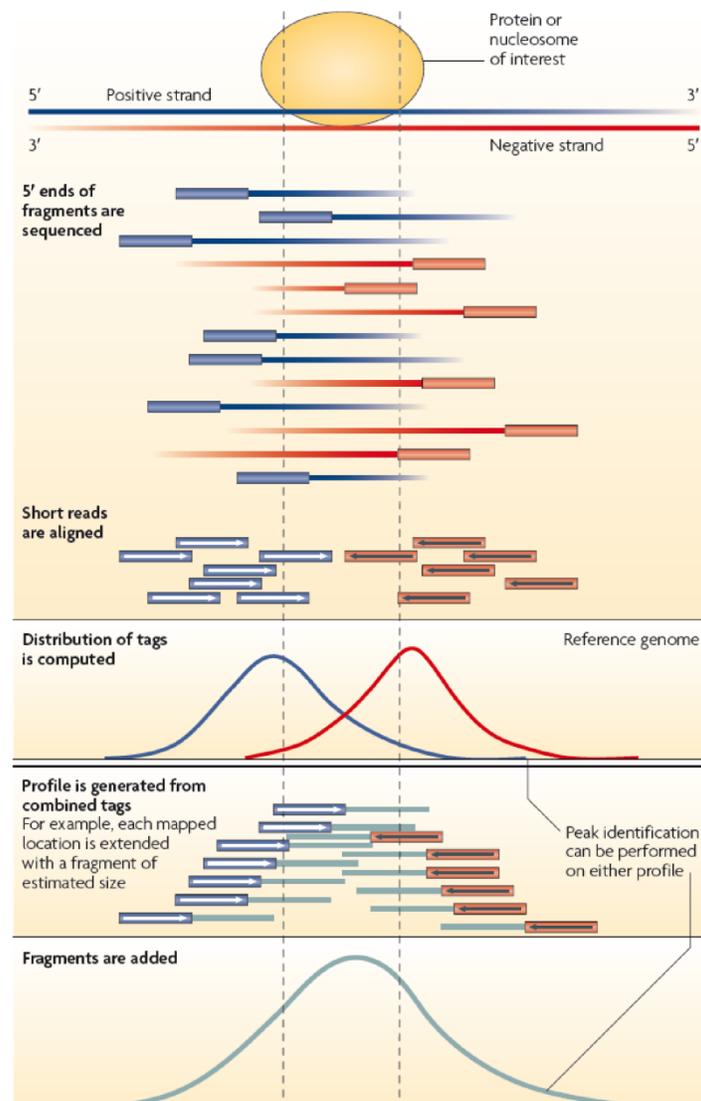
Figure 4. Mapping of the ChIP seq reads to the genome. Strand-specific peak localization within an enriched region (from Park, 2009).

DNA fragments precipitated in ChIP experiment derive from both + and - DNA strands, but both are sequenced from the 5´ end. Thus, the majority of reads are shifted towards the 5´ end of the DNA strand they derived from. Therefore each protein binding peak is composited of two shifted peaks (each peak derives from one DNA strand). The distance between the shifted peaks depends on the average size of the DNA fragments after chromatin sonication. The peaks frame the actual binding location of the precipitated protein. The script that performs peak calling must be able to recognize the strand-specific peaks, calculate average distance between them and transform the double strand-specific peaks into a single protein-binding one.

2.3 Snail/Scratch superfamily

The Snail/Scratch superfamily of TFs is a subgroup in the superfamily of zinc finger TFs.

2.3.1 Discovery of Snail

The founding member of the superfamily was identified in *Drosophila* by the team of C. Nüsslein-Volhard, E. Wieschaus and I. Kluding in the screen for *Drosophila* mutations affecting embryonic development. Mutant embryos displayed a phenotype that resembled a snail and were defective in mesoderm invagination during gastrulation (Alberga et al., 1991; Leptin and Grunewald, 1990; Simpson, 1983; Wieschaus et al., 1984). Later similar proteins were discovered in multiple species throughout the animal kingdom (Barrallo-Gimeno and Nieto, 2009; Hemavathy et al., 2000; Nieto, 2002; Wu and Zhou, 2010). The entire group of TFs was named after the first discovered member: "Snail-like family of TFs" or simply "Snail-like TFs". Together with the closely related Scratch family they form the Snail/Scratch superfamily of zinc-finger TFs.

2.3.2 Structure of Snail-like TFs

A typical Snail-like TF consists of five (sometimes six) C_2H_2 type zinc-finger DNA binding domains on the C-terminus (reviewed in Knight and Shimeld, 2001; Nieto, 2002). The zinc fingers share high levels of homology between organisms (50-70% between fly and human TFs: Hemavathy et al., 2000). The more divergent N-termini of Snail-like TFs contain conserved motifs that mediate interaction with co-repressor proteins: the CtBP (carboxy-terminal binding protein) domain in *Drosophila* and other insects or the SNAG (Snail/Gfi) domain in vertebrates and some invertebrates (Barrallo-Gimeno and Nieto, 2009; Langeland et al., 1998; Nakayama et al., 1998; Nibu et al., 1998a, 1998b; Nieto, 2002; Qi et al., 2008).

2.3.3 The diversity of Snail/Scratch superfamily of TFs

More than 150 genes of the Snail/Scratch superfamily were characterized in 52 species ranging from placozoan *Trichoplax* to *Homo sapiens*. Apart from the founding family member the *Drosophila* genome encodes two additional Snail-like TFs, *worniu* and *escargot*,

as well as three *scratch* genes: *scratch*, *scratch-like1* and *scratch-like2* (Barrallo-Gimeno and Nieto, 2009; Hemavathy et al., 2000; Nieto, 2002; Wu and Zhou, 2010).

The mouse genome contains *Snai1* (*Snail*), *Snai2* (*Slug*), *Snai3* (*Smuc*) and two homologues of *scratch*: *Scrt1* and *Scrt2*. In the human genome the members of *Snail/Scratch* superfamily genes are very similar to the mouse gene: *SNAI1* (*SNAIL*)*, SNAI1P1, SNAI2* (*SLUG*)*, SCRT1* and *SCRT2* (Barrallo-Gimeno and Nieto, 2009; Hemavathy et al., 2000; Nieto, 2002; Wu and Zhou, 2010).

2.3.4   Evolution of Snail/Scratch superfamily

The six genes of the *Drosophila* Snail/Scratch superfamily, which are present in all twelve *Drosophila* species with sequenced genomes, are the result of a tandem duplication of a single ancestral *snail* and a single *scratch* gene. In all other insects, including the closely related *Dipteran* mosquitoes, a full set of *scratch* genes is present, whereas only one *snail* member can be found. Out of three genes of the *Drosophila* Snail-like family, non-Drosophilidae *snail* has highest homology with *escargot*. Therefore *snail* and *worniu* genes can be considered as a product of a tandem duplication of *escargot* (fig. 5; Barrallo-Gimeno and Nieto, 2009).

2.3.5   General functions of Snail-like TFs

Snail-like TFs are best known for triggering epithelial-mesenchymal transition (EMT) and, consequently, for their involvement in multiple cancer types (reviewed in Wu and Zhou, 2010). However, the family is also involved in mammalian gastrulation (Carver et al., 2001), cell cycle progression, apoptosis, control of the neuronal stem cell self-renewal process and pluripotency (reviewed in Barrallo-Gimeno and Nieto, 2009 and Nieto, 2002).
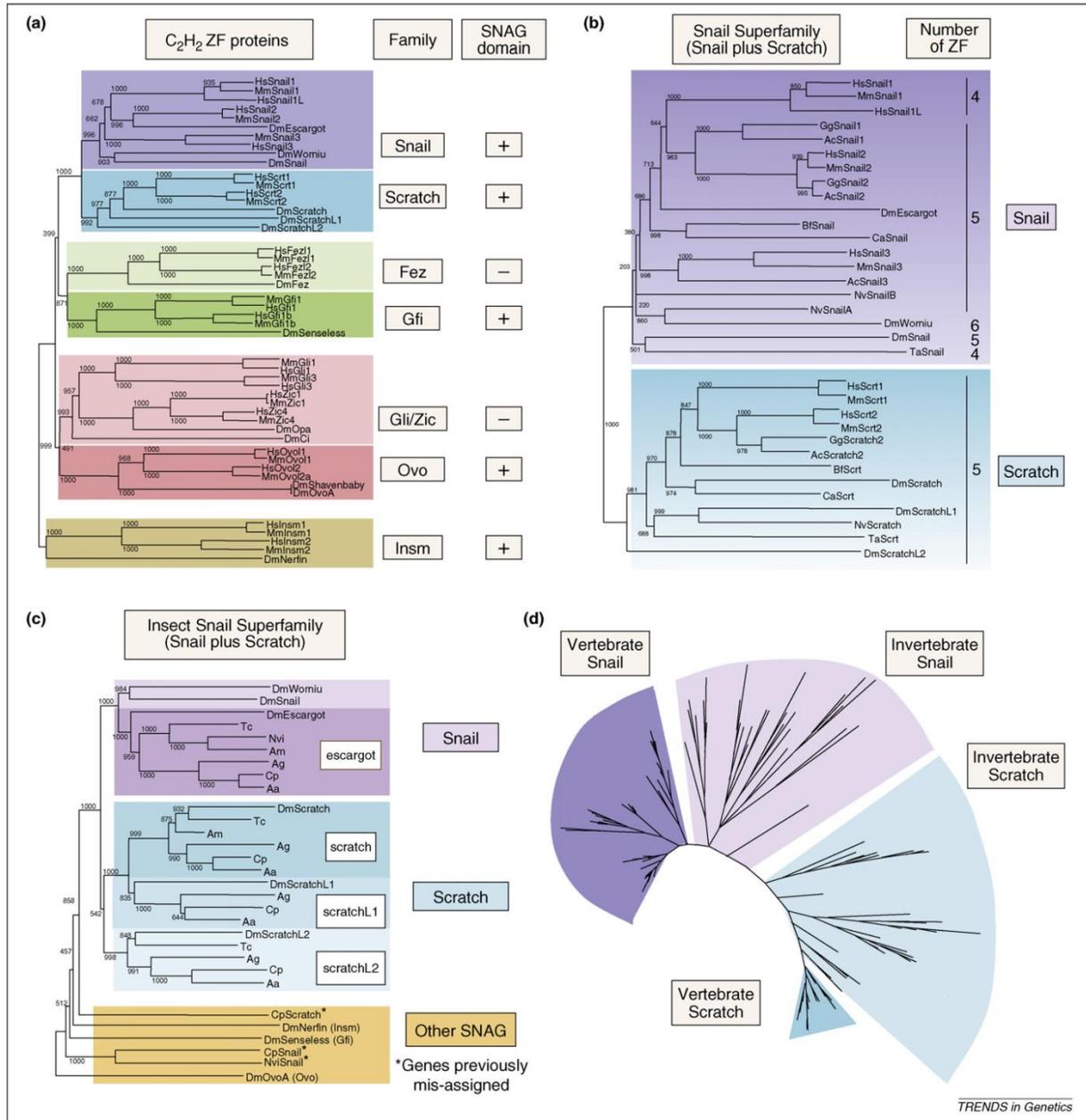
Figure 5. Evolution of Snail/Scratch superfamily (from Barrallo-Gimeno and Nieto, 2009).

a. Snail/Scratch superfamily and its evolutionary relationship to other conserved C_2H_2 TFs. Neighbor-joining (NJ) phylogenetic tree including *Drosophila* and vertebrate (mouse and human orthologues). Additional columns show the name of the family and the presence of the SNAG domain.

b. NJ phylogenetic tree of the Snail/Scratch superfamily. *Drosophila* and vertebrate members are shown. The number of the zinc fingers is given.

c. The Snail/Scratch superfamily in insects. The tandem duplication of the *escargot* gene is unique for *Drosophila*. In contrary, duplication of *scratch* is characteristic for all insect species studied. Some genes annotated as Snail do not belong to the Snail/Scratch superfamily (labeled with asterisk).

d. Phylogenetic tree of the Snail/Scratch superfamily generated by the BI method, 153 sequences from 52 species are included.

As a mammalian proto-oncogene, Snail regulates the expression of matrix metalloproteases (MT1-MMP, MT4-MMP, MMP9), the enzymes that degrade extracellular matrix and increase the motility of a cell (Huang et al., 2009; Shields et al., 2012; Welch-Reardon et al., 2014; Zhang et al., 2011). In mice and chick embryos Snail regulates cell-cycle progression and survival, particularly the components of the early to late G1 transition and the G1/S checkpoint, including repression of Cyclin D2 transcription and the increase of p21/Cip1 expression (Takahashi et al., 2004; Vega et al., 2004). *snail-like* genes are employed in different types of cell migration, such as migration of neural crest cells of *Xenopus* (Aybar et al., 2003), and in the anterior migration of the axial mesendoderm of zebrafish (Blanco et al., 2007). Snail-like TFs trigger migration of multiple cancer tumors (Hu et al., 2008; Usami et al., 2008).

snail-like genes regulate several aspects of stem cell biology. Transient co-expression of Slug, a homolog of Snail, together with Sox9, converts differentiated luminal cells of the mammary gland into mammary stem cells (Guo et al., 2012). The combination of Slug and Sox9 activates an endogenous stem cell auto-regulatory network, which remains stable for many days after silencing of the ectopic TF genes. Moreover, Slug and Sox9 are overexpressed in multiple metastatic cancer tumors (Guo et al., 2012). In the mouse embryonic cortex genetic ablation of Snail-like family leads to p53 dependent cell death of neurons and neuronal precursors. Loss of *snail-like* family causes a decrease in proliferation of neonatal cortical neural precursors and mis-localization as well as mis-specification of cortical neurons (Zander et al., 2014a).

2.3.6   Snail-like TFs in mammalian neurogenesis

Throughout the animal kingdom Snail-like TFs regulate different aspects of neurogenesis. In mammals Snail-like TFs regulate embryonic radial precursor proliferation and survival by regulation of two distinct targets: *p53* and *cdc25b*, while Snail/Scratch superfamily member Scratch regulates delamination of newborn cortical neurons from the cortical apical epithelium by transcriptionally repressing downstream targets like the *neurogenins* and *E-cadherin* (Zander et al., 2014a, 2014b). In the embryonic cortex Snail-like TFs regulate radial precursor numbers and proliferation as well as neuronal localization. Inducible ablation of Snail-like TFs in embryonic precursors and their progeny in the

postnatal cortex leads to deficits of postnatal precursors and neurons (Zander et al., 2014a). Together with Sox10, Snail-like TFs are expressed in neural crescent stem cells and specifically marks undifferentiated cells (Guo et al., 2012; Motohashi et al., 2007; Xu et al., 2014).

Scratch1 and Scratch2 are expressed downstream of pro-neural genes and control the very first step of neuronal migration; namely the cell´s departure from the ventricular surface (Itoh et al., 2013; Roark et al., 1995).

2.3.7   Regulation of the activity of Snail-like TFs

The expression of Snail-like TFs is tightly regulated by several major signaling pathways, such as FGF, EGF, TGFβ and BMP (Ciruna and Rossant, 2001; Haraguchi, 2009; Lu et al., 2003).

Subcellular localization of vertebrate Snail-like family members is tightly controlled. Nuclear transportation of Snail-like TFs is controlled by nuclear karyopherins, among them importin alpha, which imports Snail-like TFs into the nucleus in a zinc-finger dependent manner and exportin 1 together with exportin 5 as major exporters (Muqbil et al., 2014; Yamasaki et al., 2005). Mouse Snail-like TFs, containing a leucin-rich nuclear exclusion signal ($LX_{(1-3)}LX_{(2-3)}LX_{(1-2)}L$), shuttles between cytoplasm and nucleus, while an NES is not present in Slug, whose localization is found to be exclusively nuclear (Lischka et al., 2001). Multiple kinases and phosphatases (e.g. Lats2 and SCP) link nuclear transport of Snail-like TFs to WNT, TGFβ and FGF signaling pathways (Kim et al., 2012; Muqbil et al., 2014).

The stability of Snail-like TFs is regulated by molecular interactions and post-translational modifications. In mammalian cells the catalytic domains of the lysine hydroxylases, LOXL2 and LOXL3, bind to the N-terminal domain of Snail1, but not Snail2 to repress E-cadherin transcription and cause an EMT (reviewed in Peinado et al., 2007). Snail1 and Snail2 use their SNAG domains, bind to the LIM domains of the co-repressor proteins LIMD1 and WTIP, recruit them to the nucleus and enhance repression. Mammalian Snail1 and Snail2 are regulated via phosphorylation by GSK3β which causes their ubiquitination and degradation (reviewed in Sou et al., 2010).

No data about post-translational modifications and regulations is available for *Drosophila* Snail-like TFs.

2.4   Snail/Scratch superfamily in *Drosophila*

2.4.1   Family 1: Snail-like TFs

There are three *snail* genes in *Drosophila* genome. *snail* together with its paralogs *worniu* and *escargot* is located in tandem duplication on the second chromosome (fig. 6), in the cytogenetic region 35D. The C-terminal domains of the proteins consist of five (six in Worniu) zinc-finger domains, which share 75-85% of homology in amino-acid sequence. N-terminal domains contain Ebi and dCtBP-interaction amino-acid motif (Ashraf and Ip, 2001; Ashraf et al., 1999).

Figure 6. Localization of *snail*, *worniu* and *escargot* (red arrows) on the second chromosome of *Drosophila*. Screenshot of the *Dmel* UCSC genome browser. http://genome.ucsc.edu/

2.4.2   Family 2: Scratch

Scratch is a Snail-like pan-neural TF, named after the mild rough eye phenotype characteristic for homozygous mutant flies. *Scratch* together with *snail*, *worniu*, *hunchback* (*hb*) and *deadpan* (*dpn*) form the group of the embryonic pan-neuronal TFs. Unlike the other TFs of this group, *scratch* expression persists in ganglion mother cells (GMCs; Roark et al., 1995). *scratch* mutants do not exhibit mutant embryonic phenotypes, however the combination of *scratch* and *dpn* mutation leads to the drastically reduced number of neuroblasts. Ectopic expression of *scratch* results in the formation of additional neuronal clusters (Roark et al., 1995). At later developmental stages *scratch* expression can be detected in the central nervous system (CNS) and peripheral nervous system (PNS), as well as in the neuronal cells of the imaginal discs. Adult mutant flies show mild rough eyes and impaired bristle polarity.

Two paralogues of *scratch*, namely *scratch-like1* and *scratch-like2*, were annotated in the *Drosophila* genome. Their function is unknown.

2.5 Functions of Snail in *Drosophila* embryogenesis

2.5.1 Snail in mesoderm invagination

Snail is the only Snail-like TF in the *Drosophila* genome that is expressed in the mesoderm and has mesoderm-specific functions. In mesoderm during gastrulation Snail represses ectoderm and mesectoderm determinants, e.g. *rhomboid* (*rho*) and *single minded* (*sim*) (Leptin, 1991; fig. 7). *snail* mutations arrest disassembly of the junctions at sub-apical positions and disable apical constriction of the mesoderm cells (Kölsch et al., 2007; Martin et al., 2010) interfering with gastrulation.

As a repressor, Snail interacts with the co-repressor proteins Ebi and CtBP (Nibu et al., 1998a, 1998b; Qi et al., 2008) and prevents the formation of enhancer-promoter loops, in this manner disrupting target-gene activation (Chopra and Levine, 2009). However, Snail works both as a transcriptional repressor and as a transcriptional activator. Gene expression profiling of *snail* mutant embryos showed that 53% of differently expressed genes are up-regulated, i.e. repressed by Snail in wild type. The remaining 47% are down-regulated, in other words activated by the Snail protein. Detailed analysis for expression profiles of Snail-activated enhancers with wild type and mutated Snail binding sites proved that the transcriptional activation arises from direct binding of Snail to enhancer DNA, not by de-repression of a third party transcriptional activator (Rembold et al., 2014).

In contrast to the majority of mesodermal TFs, which are required for the post-gastrulation differentiation of the mesoderm-derived tissues, Snail is expressed in mesoderm only transiently. At stage 7 mesodermal expression of *snail* ceases. Instead *snail* expression emerges in the neuroblasts, the precursors of the central nervous system (Alberga et al., 1991; Ip et al., 1994; Kosman et al., 1991; Leptin, 1991).
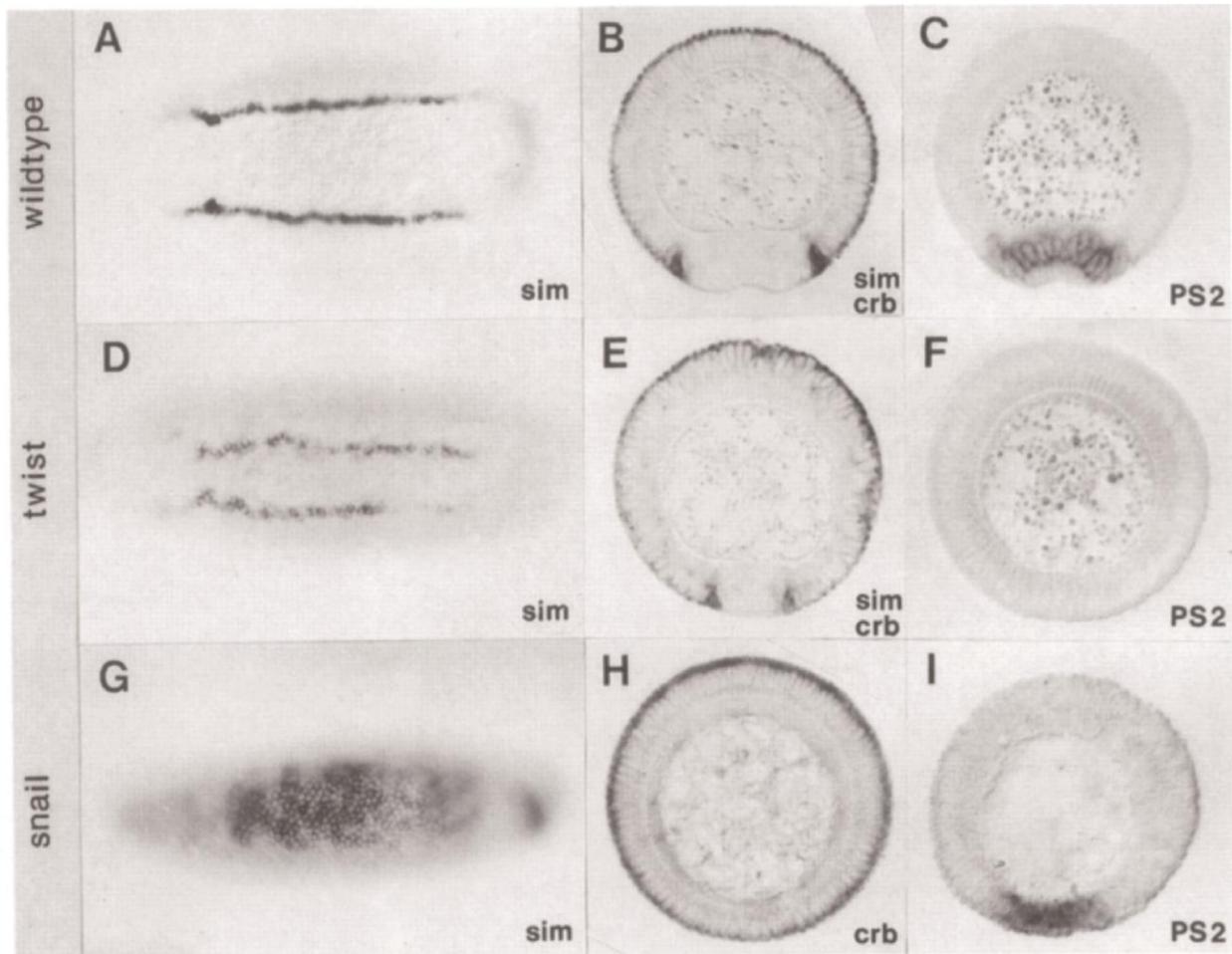
Figure 7. Twist and Snail as a pair of positive and negative transcription regulators of gene expression during *Drosophila* gastrulation. (Figure and legend from from Leptin, 1991).

a., d., g. Ventral views of *in situ* hybridizations of whole-mount embryos at cellular blastoderm; all other panels are transverse sections of embryos during early gastrulation, dorsal side up.

b., c. the ventral furrow has begun to form.

a., d., g. *sim* is expressed in two stripes, each one cell wide, at the edge of the mesoderm in *twist* mutants these lines lie closer together (d.).

g. in *snail* mutants, the gene is expressed over the whole width of the mesodermal territory.

b., e. Embryos hybridized simultaneously with probes for *sim* and *crumbs*. Because *crumbs* RNA is localized only in the apical part of each cell and *sim* is distributed over the whole cytoplasm, their expression domains can be distinguished in sectioned embryos. In *twist⁻* embryos, both the *sim* cells and the edges of the *crumbs* domain lie farther ventrally and closer to each other than in wild type.

h. *snail* mutant hybridized with the *crumbs* probe. *crumbs* is expressed around the whole periphery of the embryo.

c., f., i. Embryos hybridized with a *PS2* probe. *PS2* is not expressed in *twist* mutants (f.) but is expressed normally in *snail* mutants (i.). In embryos mutant for both *twist* and *snail*, the *crumbs* expression pattern looks like that in (h.), and PS2 looks like that in (f.).

Expression of *snail* in mesoderm and in neuroblasts is regulated by several different enhancers (Dunipace et al., 2011; Ip et al., 1994; Perry et al., 2010). Mesoderm expression is driven by two enhancers: a proximal enhancer (within 2.8 kb upstream from the start site) and a distal "shadow" enhancer (located within the exon of the neighbor gene *Tim17b2*). The two enhancers generate similar mesodermal expression patterns. Each enhancer is fully able to recapitulate expression of the gene and a BAC transgene containing either of the enhancers is able to rescue the gastrulation. However, under extreme heat shock conditions only the presence of both enhancers ensures robust gastrulation (Perry et al., 2010). The neurogenic expression of *snail* is controlled by two enhancer elements with two distinct expression patterns. One, proximal, localized within 1 kb upstream the start site, drives expression of *snail* in the peripheral nervous system, while the second, distal (2-3 kb upstream), is required for *snail* expression in the central nervous system (Ashraf and Ip, 2001; Ip et al., 1994; fig. 8).
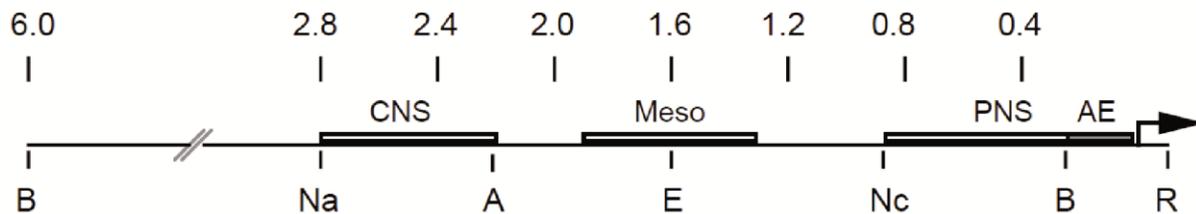
Figure 8. The upstream flanking region of the *snail* gene. The *cis* regulatory modules are represented as horizontal bars and the transcription start site (+1 bp) is indicated by the arrow. CNS denotes the region of the *snail* enhancer required for expression in the central nervous system and includes expression in early neuroectoderm cells, neuroblasts and postmitotic neurons. PNS denotes the region of the *snail* enhancer sufficient for expression in PNS sensory mother precursor cells and in postmitotic neurons. Meso is the enhancer element required for mesoderm expression. (modified from Ip et al., 1994).

2.6 **Early *Drosophila* neurogenesis**

The formation of the embryonic nervous system in *Drosophila* starts with the delamination of single cells – neuroblasts (Nb) – from the neurogenic ectoderm.

At late stage 8 the expression of patterning genes in the ectoderm identifies 10 groups of 6-8 cells per each hemisegment. These cells, defined as the "proneural clusters", express combinations of the three proneural genes: *achaete*, *scute* and *lethal of scute* (Cabrera et al., 1987; Romani et al., 1987; Skeath and Carroll, 1992). Via lateral inhibition, a Notch signaling dependent process, one cell is selected to become a neuroblast. The remaining cells develop into epidermoblasts. The neuroblast cell delaminates from the single cell layer of epidermal ectoderm via cell shape changes, summarized in fig. 9 (from Hartenstein and Wodarz, 2013, originally published in Doe, 1992; reviewed in Hartenstein and Wodarz, 2013; Kohwi and Doe, 2013).

Neuroblasts delaminate following a stereotyped spatiotemporal pattern, repeated in every body segment. The neuroblast segregation occurs in five delamination waves: SI, SII, SIII, SIV and SV (Doe, 1992; Hartenstein and Campos-Ortega, 1984). The location of a neuroblast within the delamination pattern identifies neuroblast lineage characteristics and the future identity of the neuroblast's progeny. Multiple studies were performed to generate a precise maps of neuroblast formation patterns (Broadus and Doe, 1995; Doe, 1992; fig. 10). After delamination the mitotic spindle of the neuroblast rotates to change its orientation from lateral-lateral to apical-basal and the cell starts division in a stem cell mode to produce a new neuroblast and a GMC which divides later again to generate a pair of cells (a neuron and/or a glia cell; Campos-Ortega, 1993, 1994)
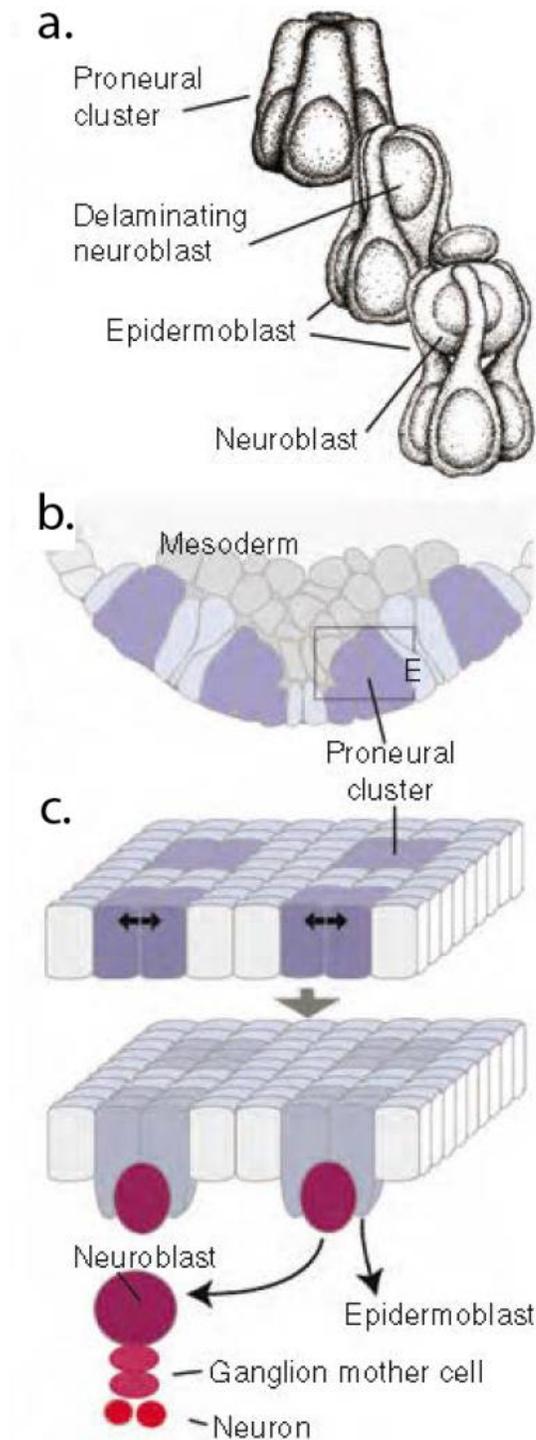
Figure 9. Neuroblast delamination. The figure and the legend are modified from (Hartenstein and Wodarz, 2013).
d. Artistic rendering of proneural cluster before (top), during (middle) and after neuroblast delamination (bottom; Reprinted from Doe, 1992; Copyright 1992 The Company of Biologists).
c. Schematic cross-section of neuroectoderm, indicating pattern of proneural clusters (purple shading).
f. Two-step model of neuroblast specification. Expression of proneural genes defines proneural clusters (top); lateral inhibition in each cluster selects neuroblast (bottom).

Individual number was specified for each neuroblast lineage. The first wave of delamination forms the neuroblasts 1-1, 3-2, 2-5, MP2, 3-5, 5-2, 5-3, 5-6, 7-1 and 7-4. Second wave generates the neuroblasts num. 1-2, 2-2, 4-2, 6-2 and 7-2 (fig. 10). Following delamination waves contribute more lineages to the neuroblast population. Each lineage has its unique spatial localization within a segment as well as an individual set of TFs and cellular markers expressed in the cell. For instance, during the second wave of delamination *wingless* expression is characteristic for the neuroblasts 5-2, 5-3 and 5-6, while the expression of *engrailed* marks the cells 6-2, 7-1, 7-2 and 7-4. Expression of cell-specific markers enables the identification of neuroblast cell identity (Doe, 1992).
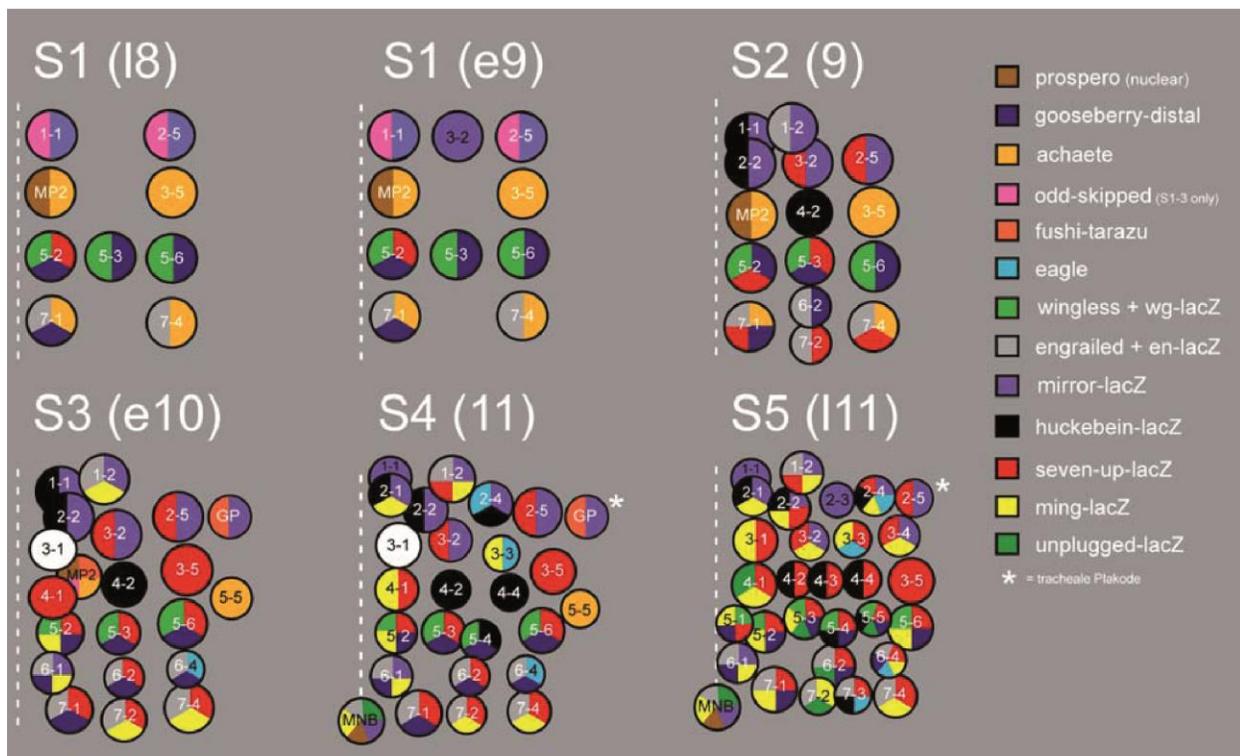
Doe (1992) und Broadus und Doe (1995)

Figure 10. The map of neuroblast delamination according to (Broadus and Doe, 1995; Doe, 1992).
Each delaminating neuroblast expresses a specific combination of the marker genes, which allows precise identification of the neuroblast identity. The expression of the marker genes is represented in a color code. Every cartoon shows one hemisegment; anterior is at the top, dot line indicates the embryonic midline. The asterisk marks the location of the trachea (the segment border).
The figure and the legend are modified from the PhD thesis of Christian Berger (Gerd Technau group, Mainz, 2004).

In addition to the spatial differentiation by the expression of a specific set of TFs and cellular markers, neuroblasts show also a temporal succession of the TFs expression. At the

embryonic developmental stages 9-10 neuroblasts express the transcriptional factor Hb. At the time point that differ for each neuroblast, the cell starts expression of *seven-up* (*svp*), which down-regulates expression of *hb* and starts expression of *kruppel* (*kr*) (Kanai et al., 2005). Later in development expression of *kr* is substituted by *pdm*. Altogether the stereotypic sequence of time-specific TFs expressed in the neuroblasts can be summarized as: *hb -> svp -> kr -> pdm* (Isshiki et al., 2001; Kambadur et al., 1998; Novotny et al., 2002). The neuroblast cells consequently switch TF expression profiles, while all cells of neuroblast's progeny (GMC and neurons) maintain the TF identity of the ancestor neuroblast at the moment of neuroblast division. Thus, the progeny of each single neuroblast is segregated according to the neuroblast's temporal identity (Li et al., 2013).

The mechanisms of such temporal transition are poorly understood and remain mainly speculative. Nevertheless, isolated neural stem cells cultured *in vitro* are able to recapitulate the sequential generation of different neuron types indicating a cell-intrinsic control of TF temporal patterning (Brody and Odenwald, 2000; Grosskortenhaus et al., 2005).

MP2 is a special cell among the neuroblast pool. While all other neuroblast cells divide asymmetrically to produce one GMC and one new neuroblast, MP2 divides symmetrically to generate two neurons (vMP2 and dMP2; Doe et al., 1988; Thomas et al., 1988), i.e. it is a neuroblast which does not possess stem cell identity.

During late embryogenesis the neuroblasts become quiescent to re-enter cell cycle and start a second wave of neurogenesis in larvae.

2.7 Roles of Snail, Worniu and Escargot

2.7.1 Snail-like TFs in development of the central nervous system

In early neurogenesis Snail, Worniu and Escargot act in concert. Their DNA-binding zinc-finger domains are highly similar and show 85% (118/139 amino acids) homology between Escargot and Worniu, and 74% (103/139 amino acids) homology between Snail and Worniu (Ashraf et al., 1999). The DNA binding motif of Snail and Escargot is almost identical (Fuse et al., 1994; Mauhin et al., 1993).

In the embryonic CNS Snail*,* Worniu and Escargot have strong functional redundancy. The deletion of just one gene has no significant effect on the expression of CNS markers. In contrast, the embryos with double deletion of *snail/worniu* or triple deletion of *snail/worniu/escargot* lose expression of the early CNS marker genes *even skipped* and *fushi tarazu*. The expression of *inscuteable* (*insc*), a regulator of asymmetric cell division in neuroblasts, is lost. Thus, the subsequent division and differentiation of neuroblasts in double- and triple-mutant embryos is defective. Surprisingly, the neuroblast delamination from the neuroectoderm was reported to occur in the mutant embryos without major problems. Expression of the early neuroblast markers (*hunchback*, *achaete*, *deadpan*, *scratch*, *lethal of scute*) was reported to be unchanged (Ashraf et al., 1999; Ashraf and Ip, 2001; Cai et al., 2001).

The transgenic expression of either a *snail*, *worniu* or *escargot* transgene under control of the neuroblast-specific *snail* 2.8 kb enhancer partially rescues expression of *fushi tarazu* and *inscuteable* in CNS the neuroblast of *snail/worniu/escargot* triple deletion mutants (Ashraf and Ip, 2001; Ashraf et al., 1999). In contrast, ectopic expression of *worniu* or *escargot* in the mesoderm is not able rescue the *snail* mutant gastrulation phenotype (Ashraf and Ip, 2001; Hemavathy et al., 2004).

The transgenic expression of *worniu*, in the mesoderm under the 2.8-kb *snail* promoter cannot rescue the mesodermal phenotypes of *snail* mutation. However, if the Worniu zinc-fingers are fused to the Snail N-terminus a certain degree of rescue is achieved. Another chimeric construct, the DNA-binding Snail zinc-fingers fused to the Worniu N-terminus is not functional (Hemavathy et al., 2004). That indicates that, (a) Snail and Worniu bind similar, if not the same DNA sequences and (b) that the N-terminus of Snail has unique mesoderm-specific functions that cannot be replaced by Worniu.

In summary, Snail cooperates with it paralogs in a context-dependent way. While it alone is sufficient to ensure mesoderm development and gastrulation, it acts in synergy with Escargot and Worniu during neuroblast development.

2.7.2   Other functions of *worniu*

Expression of *worniu* starts with formation of the first neuroblasts (stage 8) and persists in the neuroblast cells of the developing nerve system (Ashraf and Ip, 2001; Lai et

al., 2012). At stage 13 *worniu* expression is detectable in the developing brain hemispheres and in the ventral nerve cord. At late embryonic stages *worniu* expression is down-regulated and almost gone by stage 16, that correlates with the gradually lower activity of embryonic neuroblasts. During larval development *worniu* expression can be detected in the proliferating centers of brain hemispheres (Ashraf et al., 2004).

worniu is required to maintain stem cell identity of larval neuroblasts (Lai et al., 2012). Expression profiling showed down-regulation of neuroblast-specific genes and up-regulation of neuroblast differentiation genes in *worniu* mutant cells. *worniu* maintains neuroblast self-renewal in two ways: it promotes cell-cycle progression and suppresses the levels of Elav protein to inhibit premature cell differentiation (Lai et al., 2012). RNAi experiments proved that *worniu* is required for brain development in larvae: loss of *worniu* expression (via shRNA injections) results in abnormal elongated brain morphology (Ashraf et al., 2004).

2.7.3   Other functions of *escargot*

Escargot is the third Snail-like TF, which is expressed in *Drosophila* neuroblasts. Neuroblast-specific expression of *escargot* was reported as transient (Ashraf et al., 1999: fig. 11), while *escargot* is actively expressed in the neuroectoderm, a neuroblast-generating tissue (Whiteley et al., 1992).
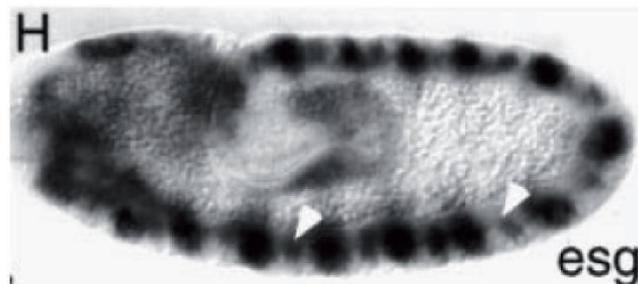
Figure 11. Expression of *escargot* in *Drosophila* embryo at stage 10. Color *in situ* hybridization. The embryo is shown from the lateral side; anterior is to the left. *escargot*-positive CNS precursors are indicated by white arrows. (The figure and the legend are from Ashraf et al., 1999).

In development of the tracheal network, *escargot* mutant tip cells fail to adhere to each other and continue the search for alternative targets, extending longer filopodia. Thus *escargot* is a regulator of DE-cadherin targeting, cell adhesion, and motility in tracheal morphogenesis (Tanaka-Matakatsu et al., 1996). *escargot* maintains stem cell identity and

suppresses differentiation of *Drosophila* intestinal stem cells (Korzelius et al., 2014; Loza-Coll et al., 2014). Depletion of *escargot* in the hub cells of the *Drosophila* testis leads to the differentiation of stem cells into cyst cells, resulting in the complete loss of the hub cells (Voog et al., 2014). In larval optic lobes *escargot* controls EMT-like processes of neural progenitor formation (Apitz and Salecker, 2014). In late eye development *escargot* is required for the survival and maturation of the 2° and 3° interommatidial pigment cells. In the *escargot* mutant clones the actin cytoskeleton of the pigment cells had no obvious defects, yet, the cells failed to undergo apical constriction (Lim and Tomlinson, 2006).

2.7.4 Redundancy of Snail, Worniu and Escargot in eye formation

Apart from embryonic neurogenesis, Snail, Worniu and Escargot were reported to function redundantly in eye development. In this context, the three TFs, regulated by Wingless signaling, promote apoptotic elimination of a distinct subpopulation of ommatidial cells. Only elimination of all three Snail-like TFs induced a mutant eye phenotype: survival of the peripheral ommatidia, which are normally eliminated by apoptosis (Lim and Tomlinson, 2006).

In summary, Snail-like TFs, namely Snail, Worniu and Escargot are required for stem cell identity, proliferation and cell motility in a variety of developmental contexts.

2.8 Datasets of DNA binding sites available for *Drosophila* Snail-like TFs

Data on TF binding sites on genomic DNA is available only for Snail. In total three publications describe Snail binding to DNA. (MacArthur et al., 2009) applied ChIP-on-chip to map the genome-wide binding landscape of 15 TFs, including Snail and Twist. The databases were produced using mesoderm-staged embryos (2-3 hours after egg lay, AEL), majority at stage 4 or early stage 5. The study identified aroximtely 1,000 Snail-bound regions.

(Rembold et al., 2014) described genome-wide Snail binding in mesodermal *Drosophila* cells performing ChIP-on-chip experiments. The embryonic staging was broader: 2-4 h AEL, corresponding to stages 5-7. The authors identified 2021 Snail-bound genomic regions.

In (Southall and Brand, 2009) the chromatin profiling technique DamID (Vogel et al., 2007) was used to map and compare the binding profiles of Snail, Prospero, Deadpan and Asense in embryos at stages 10-11 (4-7 hours AEL).

The available databases provide a useful resource for the comparison and validation of the ChIP-seq data, expected as an outcome of this work.

2.9 Aim of the research project

In my thesis I address two scientific questions.

First, what are the functions of Snail-like TFs during the development of the central nervous system. I aim for the comparative analysis of Snail, Worniu and Escargot target genes and target enhancers in neuroblasts to discover the mechanisms underlying cooperation of the TFs.

Second, which biological mechanisms underlie the tissue-specific functions of the TF Snail. Despite the evident efforts in the field, we still miss a precise high-resolution map of Snail DNA binding in neuroblasts. Comparison of Snail mesoderm-specific binding sites to the neuroblast-specific ones can give insight into the mechanisms of tissue-specific work of TFs.

3 Materials and Methods

3.1 Materials

3.1.1 Computer software

Microsoft Office 2011
Adobe Photoshop CS4
Adobe Illustrator CS4
Adobe Acrobat 9 Pro
Papers 1.9.4
Mendeley
ApE-A plasmid Editor v1.17
Vector NTI Advance 11.5.1
BioRad Quantity One 4.6.5
BioRad CFX Manager 3.0
Leica LAS V3.8
AxioVision
Imaris 7.0.0
Fiji (ImageJ)
NanoDrop 2000

3.1.1.1 Confocal microscope software

Olympus Fluoview 4.1a

3.1.1.2 Online resources

Blast, Basic Logic Alignment Search Tool: blast.ncbi.nlm.nih.gov
PubMed: ncbi.nlm.nih.gov/pubmed
UCSC Genome Bioinformatics: genome.ucsc.edu
FlyTF, fly TF database: FlyTF.org

RedFly: regulatory elements database: redfly.ccr.buffalo.edu

FlyBase: flybase.org

FlyMove: flymove.uni-muenster.de

Ligation calculator: www.insilico.uni-duesseldorf.de/Lig_Input.html

RegRNA 2.0 regrna2.mbc.nctu.edu.tw

3.1.2 Laboratory equipment

3.1.2.1 Water bath

Water bath "Medingen W6"

Dry heat bath "Unitek HB-130"

3.1.2.2 Incubators

Thermo scientific Heraeus function line

Memmert Universal oven

Emmendingen EHRET KBK 4600

Emmendingen EHRET KBK 4200

3.1.2.3 Shakers

Innova 2100

Infors HT Triple Stack Incubator Shaker Type AJ111/6

Infors HT aquatron water bath shaker

3.1.2.4 PCR cyclers

Biometra Professional basic thermocycler

Biometra UNO-thermoblock Thermal Cycler

BioRad CFX96 TouchTM Real-Time PCR Detection System

3.1.2.5 Centrifuges

Eppendorf centrifuge 5416R

Eppendorf centrifuge 5810R

Eppendorf MiniSpin plus

Thermo Scientific Sorvall Evolution RC

3.1.2.6 Microscopes and stereomicroscopes

Carl Zeiss Stemi DRC

Carl Zeiss Stemi 2000

Light Supply Faust Halolux 160

3.1.2.7 Confocal microscope

Olympus FluoView™ FV1000 Confocal Microscope

3.1.2.8 Fluorescent stereomicroscope

Leica M165 FC Fluorescent Stereo Microscope equipped with a Leica DFC490 digital camera

3.1.2.9 Phase contrast microscopes

Carl Zeiss Acioplan equipped with Axio Cam MRc5

3.1.2.10 Blotting mashine

Trans-Blot SD Semi-Dry transfer Cell

3.1.2.11 Electroporator

BioRad GenePulser Xcell Version 3.15

3.1.2.12 Fluorometers

NanoDrop 2000 Spectrofotometer (Thermo Scientific)

Qubit® Fluorometer from Life technologies

3.1.3 Chemicals

Antibiotics and chemicals, if not mentioned otherwise were purchased from AppliChem, Sigma-Aldrich, FisherScientific, Fluka, VWR, Merk, Roche and Bio-Rad

Western Blocking Reagent, Solution (Roche: 11921673001)

Westburg Epicentre CopyControl induction solution (CCIS125)

Phenol:chloroform:isoamylalcohol mix (Ambion: AM9730)

VOLTALEF PCTFE 3S mineral oil (Arkema; Lehmann&Voss&Co)

VOLTALEF PCTFE 10S mineral oil (Arkema; Lehmann&Voss&Co)

3.1.4 Affinity gels

Anti-V5 Agarose Affinity Gel, antibody produced in mouse (Sigma: A7345-1ML)

EZviewTMred Anti-HA Affinity Gel (Sigma: E6779-1ML)

3.1.5 Protein A beads

Pierce® Protein A Agarose (ThermoScientific: 20333)

Protein A Sepharose ® 4B, Fast Flow from *Staphylococcus aureus* (Sigma Aldrich: P9424)

3.1.6 Enzymes

The restriction enzymes, DNA ligase, DNA phosphatase and RNA polymerases were purchased from New England BioLabs

Taq DNA polymerase recombinant (Invitrogen: 10342-053)

MyTaq RedMix 2x (BioLine: BIO-25043)

Proteinase K (Sigma: EC3.4.21.64; P-2308)

RNase A (from Qiagen Large construct kit: 12462)

Protease Inhibitors (Roche: cOmplete Tablets, Mini EDTA-free, EASYpack, 04693159001)

3.1.7 Embedding media

ProLong® Gold antifade reagent (Invitrogen: P36930)

VectaShield medium for fluorescence (Vector laboratories: H-1000)

Fluoromount-G (SouthernBiotech: 0100.01)

3.1.8 Kits

innuPREP mini kit (Analytic Jena: 845-KS 5040250)

JetStar 2.0 Midi kit (Genomed: 210050)

JetSorb Gel extraction kit (Genomed: 110150)

iQ™ SYBR® Green Supermix (Bio-Rad: 170-8880)

MyTaqTM RedMix (BioLine: BIO-25043)

RNeasy Mini kit (Qiagen: 74104)

Vectastain ABC Systems (Vector Laboratories Inc.: PK-6100)

TOPO TA Cloning kit (Invitrogen: 1114817)

DIG, Biotin and Fluorescein RNA labeling mixes (Roche: 11 277 073 910, 11 685 597 910 and 11 685 619 910, respectively)

PerkinElmer Signal Amplification (TSA) system (NEL753001KT)

Kapa Library Preparation Kits (KAPABIOSYSTEMS: KK8241; 50 libraries)

Amsterdam™ ECL™ Western Blotting Detection Reagents (GE healthcare: RPN2209)

Large Construct purification kit (Qiagen: 12462)

Quant-iT™ PicoGreen® dsDNA Assay Kit (Life technologies: P7589)

Qubit® dsDNA HS Assay Kit (Life technologies: Q32854)

Dual Promoter TA Cloning Kit pCRII vector (Invitrogen: Life technologies: 45-0007LT)

QiaQuick PCR Purification Kit (Qiagen: 28104)

3.1.9 Lab plastic

DNA LoBind Tubes 1.5 ml (Eppendorf: 022431021)

Phase Lock Gel Heavy 2 ml (5Prime: 2900309)

Falcon tubes (blue caps; 15, 50 ml; VWR: 210 261; REF 352096)

VWR © Centrifuge tubes, polystyrene (magenta caps; 50 ml; VWR: 21008-242)

Hard-Shell qPCR 96-well plates (BioRad: HSP9601)

0.1 cm BioRad Gene Pulser cuvettes

3.1.10   Glassware

Wheaton Dounce Tissue Grinder, 15 ml

3.1.11   Solutions, buffers, media

3.1.11.1   LB medium

Bacto-trypton	10 g
Yeast extract	5g
NaCl	5g
water up to 1 l	

3.1.11.2   Apple juice agar plates

agar	90 g
water	3 l
mix and autoclave	
apple juice	1 l
sugar	100 g

Mix and autoclave; cool down to 60°C, add 40 ml Nipagin solution (15% in 70% ethanol). Mix agar solution and apple juice solution, distribute on the Petri dishes

3.1.11.3   Heat fixation solution

NaCl	4 g
Triton X-100	300 μl
Water up to 1 l	

3.1.11.4   7% Sodium hypochlorite

14% Sodium hypochlorite (VWR: 27900.365)	One volume
tap water	One volume

3.1.11.5   Formaldehyde fixation mix

37% formaldehyde	90 µl
PBS	810 µl
heptane	900 µl

3.1.11.6   PBS

Na_2HPO_4	14.4 g
KH_2PO_4	2.4 g
NaCl	80 g
KCl	2 g
water up to 1 l	
pH 7.4	

3.1.11.7   PBStw

PBS	99.9%
Tween20	0.1%

3.1.11.8   PBStr

PBS	99.9%
Triton X-100	0.1%

3.1.11.9   BSA-NGS-PBStw

PBS	97.9%
BSA	1%
NGS	1%
Tween20	0.1%

3.1.11.10   DAB solution

PBSTw	500 µl
3,3´-diaminobenzidine (DAB)	50 µl
0.3% H_2O_2	10 µl

3.1.11.11   Pre-hybridization mix

Formamide	5.0 ml
SSC buffer, pH 7.0	20x
torula RNA, 5 mg/ml	200 µl
Salmon sperm DNA, 10 mg/ml	100 µl
heparin 100 mg/ml	5 µl
Tween 20,	10 µl
water	up to 10 ml

3.1.11.12   NBT-staining mix

Nitroblue tetrazolium (75 mg/ml) in 70% dimethylformamide	4.5 µl
5-bromo-4-chloro-indolxylphosphate (50 mg/ml) in dimethylformamide	3.5 µl

3.1.11.13   Staining buffer

TrisHCl, pH 9.5	50 mM
$MgCl_2$	2 mM
Tween-20	0.1%
prepare shortly before the experiment, use fresh	

3.1.11.14   Embryo Lysis Buffer

TrisHCl, pH 7.5	60 mM
NaCl	150 mM
EDTA	6 mM
Triton X-100	2%

3.1.11.15   Western blot sample loading buffer (6x Laemmli buffer)

Tris-HCl pH 6.8	65.8 mM

SDS	2.1%
glycerol	26.3% (w/v)
bromophenol blue	0.01%

3.1.11.16   Crosslinking buffer

EDTA	1 mM
EGTA	0.5 mM
NaCl	100 mM
HEPES pH 8.0	50 mM

3.1.11.17   Stop solution

glycine	125 mM
Triton X-100 in	0.1%
PBS	up to 100%

3.1.11.18   Cell lysis buffer

KCl	85 mM
IGEPAL CA-630	0.5% (v/v)
HEPES, pH 8.0	5 mM

3.1.11.19   Nuclear lysis buffer

EDTA	10 mM
N-lauroylsarcosine	0.5% (w/v)
HEPES pH 8.0	5 mM

3.1.11.20   RIPA buffer

NaCl	140 mM
EDTA	1 mM
Triton X-100	1% (v/v)
SDS	0.1% (w/v)
Sodium deoxycholate	0.1% (w/v)
Tris-HCl pH 8.0	10 mM

3.1.11.21   RIPA 500 buffer

NaCl	500 mM
EDTA	1 mM
Triton X-100	1% (v/v)
SDS	0.1% (w/v)
Sodium deoxycholate	0.1% (w/v)
Tris-HCl pH 8.0	10 mM

3.1.11.22   LiCl buffer

LiCl	250 mM
EDTA	1 mM
IGEPAL CA-630	0.5% (v/v)
sodium deoxycholate	0.5% (w/v)
Tris-HCl, pH 8.0	10 mM

3.1.11.23   TE buffer

EDTA	1mM
Tris-HCl pH 8.0	10 mM

3.1.12   DNA constructs

3.1.12.1   *FlyFosmids* and plasmids

pBluescript SK (+) cloning vector (StrataGene) was used as the basis to assemble recombination cassettes, (laboratory plasmid stock).

FlyFosmids FlyFos015520, *FlyFos017825*, and *FlyFos021004* were kindly provided by Pavel Tomancak (Ejsmont et al., 2009).

Tagging cassettes were designed and generated, according to (Ejsmont et al., 2009).

3.1.12.2   Primers

Number	Name	Sequence

Primers used to assemble the FRT-KanR-FRT-Start-V5-HA and FRT-KanR-FRT-Start-HA-V5 construct

Number	Name	Sequence
1.	*KanNeo_FRT_EcoRI_for*	aagaattctcctcgggaagttcctatt
2.	*KanNeo_FRT_BamHI_rev*	aaggatccaatctgtcgacgaagttcc
3.	*V5_Start_BamHI_Ala_NotI_for*	gatccatgggcaagcccatccccaacccctgctgggcctggatagcaccgctgc
4.	*V5_Start_BamHI_Ala_NotI_rev*	ggccgcagcggtgctatccaggcccagcaggggggttggggatgggcttgcccatg
5.	*HA_NotI_SacI_Ala_for*	ggccgcgtacccatatgatgttccagattacgctgctgcggcggagct
6.	*HA_NotI_SacI_Ala_rev*	ccgccgcagcagcgtaatctggaacatcatatgggtacgc
7.	*HA_Start_Ala_BamHI_NotI_For*	gatccatgtacccatatgatgttccagattacgctgc
8.	*HA_Start_Ala_BamHI_NotI_rev*	ggccgcagcgtaatctggaacatcatatgggtacatg
9.	*V5_Ala_NotI_SacI_for*	gccgccggcaagcccatccccaacccctgctgggcctggatagcaccgctgcgg cggcggagct
10.	*V5_Ala_NotI_SacI_rev*	ccgccgccgcagcggtgctatccaggcccagcaggggggttggggatgggcttgccggc

Primers, used to amplify the recombination cassettes from the template plasmids pBSSK-FRT-KanR-FRT-Start-V5-HA and pBSSK-FRT-KanR-FRT-Start-HA-V5

Primers have 50 bp overhangs (capital letters) complementary to the genome to target the recombination process. The primers used to amplify recombination cassettes for snail *(UPPERCASE used to highlight 50 bp homology regions)*

Number	Name	Sequence
11.	Forward primer	CCATCTCGATCAGTACCGGAAACTAAAACTTAATCACACACACATCAAAA gatatcgaattctcctcgg
12.	Reverse primer to amplify V5-HA tagging cassette	ACGAAGACAATGGGGCGCTTCTTTAGCGGGCAGCTTTTGTAGTTGGCGGC agcgtaatctggaacatcatatgggtacgc
13.	Reverse primer to amplify HA-V5 tagging cassette	ACGAAGACAATGGGGCGCTTCTTTAGCGGGCAGCTTTTGTAGTTGGCGGC cgccgccgcagcggtgctatcc

Primers used to amplify recombination cassettes for worniu *(UPPERCASE used to highlight 50 bp homology regions)*

Number	Name	Sequence
14.	Forward primer	AAATCGAAGATCAGTGATCAGTAACCATAAGGAATAATAACATAGCAACC tgatatcgaattctcctcgg
15.	Reverse primer to amplify V5-HA tagging cassette	ATTATCGGACGCTTCTTGAGCGGGCAGCGGCTGTACTTGAGTTTATC agcagcagcgtaatctggaacatcatatgggta
16.	Reverse primer to amplify HA-V5 tagging cassette	ACCATTATCGGACGCTTCTTGAGCGGGCAGCGGCTGTACTTGAGTTTATC cgccgccgcagcggtgctatcc

Primers used to amplify recombination cassettes for escargot *(UPPERCASE used to highlight 50 bp homology regions)*

Number	Name	Sequence
17.	Forward primer	ACCATTATCGGACGCTTCTTGAGCGGGCAGCGGCTGTACTTGAGTTTATC cgccgccgcagcggtgctatcc
18.	Reverse primer to amplify V5-HA tagging cassette	AGCGGGCACTTGCTGTAGTTTTTCTCCACCAACATGTCTTCCACGGTATG agcgtaatctggaacatcatatgggtacgc
19.	Reverse primer to amplify HA-V5 tagging cassette	AGCGGGCACTTGCTGTAGTTTTTCTCCACCAACATGTCTTCCACGGTATG cgccgccgcagcggtgctatccaggc

Primers to sequence the sites of recombination. Primers amplify ~400 bp around the start codon of the gene

20.	*snail* forward	ccgaactatataactcgcctctcg
21.	*snail* reverse	ctgcggctgatcctgggcaaac
22.	*worniu* forward	tcattcgagcaaggaactcc
23.	*worniu* reverse	catgacttaggtgatcctctg
24.	*escargot* forward	tcagttattcccagactcta
25.	*escargot* reverse	agggattttcctccagaatt

Primers that amplify the binding site of Mef2

28.	Mef2 for	ccccctatgctgaaggatctatt
29.	Mef2 rev	gctcatggatgcgataatgct

Primers that amplify putative binding sites of Snail. The sequences are provided by Robert Zinzen

30.	*rho* for	ggcggaatttcctgattcgcg
31.	*rho* rev	gaatttccccaaaacatgtgc
32.	*kr* for	cttcctgtctgcacccttat
33.	*kr* rev	ataagtacgctctgtctcgg
34.	*osk* for	tgcgaatggtcttcatggaa
35.	*osk* rev	caccgtcaagcagcgtgtac
36.	*sim 1* for	tttgctcatcctacctgtttcc
37.	*sim 1* rev	tcacataaggaggacaatgctg
38.	*rho 1* for	ggaaaaatgggaaaacatgcggtggg
39.	*rho 1* rev	cgcacttgtcttccgcaagttgcgcc
40.	*rho 2* for	ttcctgattcgcgatgccatgagg
41.	*rho 2* rev	cccgtcgcccgggaatttccc
42.	*sog* for	ctatatggctgtatggtgcggg
43.	*sog* rev	ctcggcaccggctggaattctac
44.	*sim 2* for	tgcgtgggaatctttgctcatcctac
45.	*sim 2* rev	ttcacataaggaggacgatg

3.1.13   Antibodies

3.1.13.1   Immunohistochemistry

Guinea pig and rabbit anti-Snail immune serum. From laboratory collection. Immune serum was produced designed and produced by Fiona Ukken, our former PhD student. The antibodies, raised in a rabbit and guinea pig, are directed against an N-terminal fragment of the polypeptide (amino acids 1-200) fused to GST. Dilution 1:1000

Guinea pig anti-Snail immune serum (Qi et al., 2008). Dilution 1:400

Rat anti-HA antibody (Boehringer Mannheim, 1867431). Immunohystochemistry dilution 1:500

Rabbit anti-Bazooka immune serum (Wodarz et al., 1999). Dilution 1:2000

Rabbit anti-Hb immune serum (Tran and Doe, 2008). Dilution 1:1000

Rabbit anti-E-cadherin immune serum (Santa Cruz Biotechnology inc., SC 33743). Dilution 1:100

Rat anti-Worniu immune serum (Lee et al., 2006). Dilution 1:1

Mouse anti-neurotactin immune serum (Corey Goodman lab: Patel et al., 1987). Dilution 1:5000

Mouse anti-actin immune serum (monoclonal; Millipore MAB1501R). Dilution 1:2000

3.1.13.2   *In situ* hybridization

Mouse anti-biotin antibody, horseradish peroxidase (HRP) conjugated (Invitrogen, 03-3720). Dilution 1:2000

Mouse anti-fluorescein antibody HRP-conjugated (Roche, 11 426 320 001). Dilution 1:2000

Sheep anti-DIG antibody HRP-conjugated (Roche, 11 650 300). Dilution 1:500

3.1.13.3   Western blot

Mouse anti-V5 monoclonal antibody (Invitrogen, R960-25). Dilution 1:2000

Rat anti-HA antibody (Boehringer Mannheim, 1 867 431). Dilution 1:1000

Mouse anti-Actin monoclonal antibody (Roche, 1378996). Dilution 1:2000

3.1.13.4 Chromatin immunoprecipitation

Rabbit anti-Mef2 immune serum. Kind gift of Eileen Furlong (Sandmann et al., 2006a)

Guinea pig anti-Snail immune serum. Kind gift of Robert Zinzen (Zinzen et al., 2006)

Guinea pig and Rabbit anti-Snail immune serum (see immunohistochemistry section)

Anti-V5 agarose affinity gel. Sigma, A7345-1ML

EZviewTM Red Anti-HA affinity gel, Sigma, E6779-1ML

3.1.13.5 Secondary antibodies

Rabbit anti-goat antibody Alexa Fluor 488 conjugated. MoBiTec AZA11008

Goat anti-rat antibody Alexa Fluor 568 conjugated. MoBiTec A11077

Goat anti-rabbit antibody Alexa Fluor 568 conjugated. Molecular Probes A-11011

Goat anti-rat antibody Alexa Fluor 488 conjugated. MoBiTec

Goat anti-mouse antibody Alexa Fluor 488 conjugated. Molecular probes A11001

Goat anti-guinea pig antibody Alexa Fluor 488 conjugated. Invitrogen A-11073

Goat anti-rat antibody HRP-conjugated. Jackson ImmunoResearch Laboratories (Dianova) , Inc.; cat. num. 112 035 003

Goat anti-mouse antibody HRP-conjugated. Jackson ImmunoResearch Laboratories (Dianova), Inc.; cat. num. 115 036 006

Goat anti-rabbit antibody HRP-conjugated. Jackson ImmunoResearch Laboratories (Dianova), Inc.; cat. num. 111 035 003.

Secondary antibodies conjugated to Alexa Fluo were used at a dilution of 1:1000. Secondary antibodies conjugated to HRP were diluted 1:2000

3.1.13.6 Bacteria strains

Strain	Purpose	Relevant genotype	Source	reference
GH5α	Molecular cloning, plasmid isolation		Lab stock collection	
EL250	*FlyFosmid* recombination, *FRT* recombination	DH10B [λ*cl*857 (*cro-bio*A) <> *ara*C-P$_{BAD}$*flpe*][b]	Colin Donohoe, Mirka Uhlirova lab, CECAD, Cologne, Germany	(Lee et al., 2001)
EPI3000	Isolation of *FlyFosmid* DNA; Inducible replication of *oriV* at high copy number	Inducible repression of *trfA* gene	Colin Donohoe, Mirka Uhlirova lab, CECAD, Cologne, Germany	TransforMax; Epicentre

3.1.14 Fly stocks used in the experiments

3.1.14.1 Wild types

w^{1118} (abbreviated in the text as *w$^-$*)

OregonR

3.1.14.2 Transgenic *FlyFosmid* lines generated in this work

V5-HA-snail

HA-V5-snail

V5-HA-worniu

HA-V5-worniu

V5-HA-escargot

HA-V5-escargot

3.1.14.3 Balancer lines

w$^-$; Bl/SM1

w$^-$; IF/SM1; MKRS/Tm6B

3.1.14.4   Mutation lines

Genotype	num. in Bloomington stock center
$esg^{35Ce\text{-}3}$ cn^1 bw^1 sp^1/CyO	30475
wor^4 cn^1 bw^1 sp^1/CyO	25170
sna^{18} cn^1 bw^1 sp^1/CyO	2311
$Df(2L)osp29, Adh^{UF}$ $osp29$ pr^1 cn^1/CyO	106824

3.1.14.5   Fly lines carrying attP recombination sites, used for phiC31-mediated transgenesis

y^1 $w\text{-}$ $vas\text{-}phiC31; attP2$	
y^1 sc^1 v^1 $nos\text{-}phiC31; attP2$	
y^1 w^1 $nos\text{-}phiC31; attP40$	

(Bischof et al., 2007; Ejsmont et al., 2009; Groth et al., 2004)

3.2   **Methods**

3.2.1   **Molecular cloning**

Basic molecular cloning procedures (DNA digestion, ligation, competent cell transformation, etc) were performed according to standard protocols (Green and Sambrook, 2012).

3.2.2   **Fly work**

Basic fly work (mating, crossings) was performed according to the standard protocols (Bownes et al., 1990).

3.2.2.1   Embryo collection in small scale

A fly population from at least 8 bottles was put into a "cage": a piece of plastic tube, covered with a metal net on one side. The other side of the tube was covered by a petri dish with apple juice agar. Yeast paste on apple juice agar served as a source of fresh food for the flies. Consuming yeasts, flies laid embryos on the agar. In the morning to stimulate fly embryo production the flies were provided with two fresh apple juice agar plates, each plate was incubated on the cage for one hour. These plates were discarded. The third and following plates were used to collect and fix stage-specific embryo collections.

3.2.2.2   Handling of the large fly populations

Middle-sized and growing fly populations are kept in fly bottles (fig. 12). The population of one fly bottle consists of approximately 0.5 – 1 g of flies. To upscale the fly population I used plastic boxes ("fly chambers") with hermetic plastic lids (1.7l, 25 x 12 x 8 cm, NeoLabs 2-1587). Aeration holes (5-cm in diameter) were cut in the plastic lids, and sealed with the fine metal net, to make the holes impenetrable for flies. One chamber was loaded with ~0.6 l of fly medium (fig. 12). The chambers were populated with embryos, that have been collected on apple juice agar plates. The populated chambers were kept in incubators at 25°C.

Figure 12. Handling large populations of *Drosophila*.
a. A bottle with fly medium, standardly used to handle fly populations. One bottle can yield up to 0.7 g of flies, depending on the handling conditions and fly line genotype.
b. "Fly chambers". One chamber is able to substitute for 10 fly bottles. Yield of the fly material is extremely sensitive to the handling conditions and amount of fly embryos used to populate one chamber. Over-population results into the rapid accumulation of faeces, slowdown of fly development and smaller size of the adult flies.
c. Large fly population cages. One population cage must be populated with 40-50 g of flies. 4 population cages can produce up to 1 g of fly embryos (weight of the live embryos) in 2 hours.

3.2.2.3 Embryo collection in large scale

To collect large amounts (>0.5 g) of tightly staged fly embryos, large fly cages (population cages: app. 40x40x100 cm) must be populated with plenty of flies. Flies collected from 10 fly chambers were used to populate one large fly cage. Cages populated with less than 75 g of flies are under-populated and give very small yield of biological material. A cage with more than 150 g of flies is over-population, flies are overcrowded and

a thick layer of dead fly bodies accumulates on the cage bottom within few days, complicating embryo collection. One cage can be loaded with 5-6 apple juice plates and is able to produce up to 0.5 g of embryos in 2 hours (weight of dry embryos after ChIP formaldehyde fixation, which is usually ½ of the weight of live embryos). In the morning the first two 1 hours embryo lays were discarded, thereafter 2 hours embryo collections could be started.

3.2.2.4 Embryo staging

Temperature	Hours AEL	Staging, purpose
28°C	2-4 hours AEL	mesoderm formation and germ band extension
	1.5-3.5	mesoderm formation
25°C	2-4	timing specific for mesoderm formation
	4-6	neuroblast delamination, ventral nerve cord formation
room temperature (21-22°C)	2-6	mesoderm formation, germ band extension and neuroblast delamination
	2.5-4.5	mesoderm formation; embryo collection for the ChIP (~85% of embryos are at stages 5-8)
	4.5-6.5	neuroblast delamination stages for ChIP (~85% of embryos are at the stages 9-11)

3.2.2.5 *Drosophila* transgenesis

Transgenic fly lines *y¹ w⁻ vas-phiC31;; attP2* and *y¹ w¹ nos-phiC31; attP40* (references in the stock list), used for *FlyFosmid* transgenesis, were taken from the laboratory fly stock collection. Fly transgenesis using microinjections was done according to the standard protocols (Bownes et al., 1990).

Shortly, plasmid DNA, diluted in distilled water to 200 ng/µl (400 ng/µl for *FlyFosmids*) and clarified by centrifugation for 15 min at 13 000 g, was injected through a glass capillary needle into the posterior terminus of an embryo at stages 1-2 of embryonic development (0-30 min post fertilization). Injected embryos were covered with Voltalef mineral oil (3S) until hatching. φC31 recombinase, expressed under control of the germ-line specific promoters of *vas* or *nos* genes, recombine an *attB* site of the vector with an *attP* site in the fly genome and integrate the plasmid into the fly chromosomes in germ line nuclei. After 24 hours, when first instar larvae were fully developed, the larvae were transferred to standard *Drosophila* food medium, where they continued their development. Eclosed flies were collected and crossed to the balancer lines to obtain transgenic flies in the next generation.

Approximately 10-40% of injected embryos developed into fertile adult flies. 0.1-1% of injections produced transgenic flies.

3.2.2.6　　　　Identification of transgenes

Adult flies, developed from the injected embryos, were crossed to males or virgins of the balancer lines *IF/SM1; MKRS/Tm6B* or *Bl/SM1*. Progeny of the cross was screened to find the transgenes.

If the injected construct carried *white* or *mini-white* gene as a marker, transgenic flies were identified by the appearance of colored eye (*attP* fly line and balancer lines are *w$^-$*, i.e. have white eye color, while the eyes of a transgenic fly are red).

FlyFosmid vector carries the gene of red fluorescent protein (*dsRed*) under control of an eye-specific promoter. Potential transgenes were selected for the red fluorescence in the eye using a Leica M165 FC fluorescence stereomicroscope.

3.2.2.7　　　　Fixation of *Drosophila* embryos for immunostainings and *in situ* hybridizations (small scales of embryo collection)

3.2.2.7.1　　　Heat fixation

Embryos were dechorionated with 7% sodium hypochlorite for 30 sec (commercially supplied 14% sodium hypochlorite, diluted with tap water roughly 1:1), collected on a fine nylon sieve and carefully rinsed with tap water. Dechorionated embryos were thrown into a glass beaker with boiling heat fixation buffer (2g NaCl and 150 µl Triton X-100 in 500 ml H$_2$O). After 30 seconds the beaker was put on ice, some pieces of ice were thrown directly into the beaker to rapidly chill down the boiling solution. Embryos from the beaker were transferred into a 2 ml Eppendorf tube. Heat fixation buffer was replaced by 1 ml of methanol combined with 1 ml of heptane. After 30 sec of vigorous shaking (removing vitelline membrane) the embryos were washed 3 times with methanol and stored in methanol at -20°C.

3.2.2.7.2　　　Formaldehyde fixation

Embryos, dechorionated and rinsed with tap water as described in the heat fixation protocol, were collected into 2 ml Eppendorf tubes and supplied with fresh formaldehyde fixation mix: 90 µl 37% formaldehyde, 810 µl PBS, 900 µl heptane. The tube was vigorously

shaken 25 min at 37°C. After shaking, the aqueous (lower) phase was substituted for methanol. After 30 sec of vigorous shaking (to remove the vitelline membrane), methanol/heptane was removed from the tube and the embryos were washed 3 times with methanol and stored in methanol at -20°C.

3.2.3 Immunohistochemistry

3.2.3.1 Fluorescent immunostainings

Fixed embryos stored in methanol were washed with PBStw (PBS with 0.1% Tween 20) 3 times for 10 minutes each, blocked one hour in PBStw with 1% BSA and 1% NGS, and incubated overnight with the primary antibodies diluted in PBStw. On the next day the embryos were washed 3 times for 10 minutes with PBStw and blocked in PBStw-BSA-NGS as before, then incubated 1 hour with the corresponding secondary antibodies diluted in PBStw. After 3 final washes with PBStw the embryos were embedded in VectaSchield mount medium for fluorescence (Vector laboratories, H-1000).

3.2.4 *In situ* hybridizations

3.2.4.1 RNA anti-sense probe preparation

A gene fragment, chosen to be a probe for *in situ* hybridization, was amplified using standard PCR program (reaction was done as PCR amplification of recombination cassette for *FlyFosmid* recombineering) and specifically designed primers. The PCR product was ligated into the *pCRII-TOPO* vector, using the TOPO TA Cloning kit. Being ligated into the *pCRII-TOPO* vector, the PCR product is flanked by two different RNA polymerase promoters (SP6 and T7) and M13 primer binding sites. This cassette was amplified with M13 primers and the PCR fragment served as a template for the synthesis of both sense and antisense RNA probes. DIG- Biotin- and Fluorescin-10x RNA labeling mixes were used to label the probes (Roche, 11 277 073 910, 11 685 597 910 and 11 685 619 910 correspondently).

After NaAc precipitation, the probe was dissolved in 60 µl of water and diluted with 150 µl of pre-hybridization mix (5.0 ml formamide; 20x SSC buffer, pH 7.0; 200 µl torula RNA, 5 mg/ml; 100 µl ssDNA, 10 mg/ml; 5 µl heparin 100 mg/ml; 10 µl Tween 20; water up to 10 ml). The probe was used with the dilution 1:100.

3.2.4.2   **Fluorescent *in situ* hybridization**

Embryos, that had been fixed with formaldehyde and stored in methanol, were rehydrated by washing with 70% methanol/PBSTw, 50% methanol/PBSTw, 30% methanol/PBSTw, and twice with PBSTw (5 min each), post-fixed for 20 min in 3.7% formaldehyde and washed again 5 x 5 min with PBSTw. The embryos were washed once with a 1:1 mix of pre-hybridization solution and PBSTw and once with pre-hybridization solution. The hybridization mix was prepared by diluting the labeled RNA-probe(s) 1:50 in pre-hybridization solution. After heating the mix for 10 min at 80°C, it was cooled on ice for 2 min. The embryos were incubated in hybridization mix at 65°C overnight.

On the second day of the protocol, the embryos were washed extensively at 65°C, first with pre-hybridization mix 3 x 30 min and 3 x 1 hour, next with a dilution series of pre-hybridization: PBSTw: 4:1, 1:1, 1:4 (each wash 20 min) and finally 4 times with PBSTw for 5 min each. After cooling down to room temperature the embryos were washed 2 x 30 min with blocking solution (100 µl Roche Western blocking reagent and 400 µl PBSTw). After blocking, the embryos were incubated overnight at 4°C with horseradish peroxidase (HRP)-conjugated antibodies against either Digoxigenin, Biotin or Fluroescein, diluted 1:2000 in blocking solution. To limit unspecific staining preabsorbed antibodies were used (see protocol for antibody preabsorption).

On the third day, the embryos were rocked once and then washed 6 x 20 min with PBSTw. The fluorescent staining was performed using the Tyramide Signal Amplification (TSA) system kit (PerkinElmer, NEL753001KT). Shortly, the appropriate fluorophore-tyramide was diluted 1:50 in 200 µl of amplification buffer and incubated for 3 min with the embryos. The reaction was rocked with 1 ml of PBSTw and the embryos were washed with PBSTw 6x 20 min.

In the case of double *in situ* hybridization, the peroxidase was inactivated by a 15 min wash with stop solution (10 mM HCl and 0.2% Tween 20). After 2 x 5 min washes with PBSTw, a second deactivation step was performed. The embryos were incubated in 3% H_2O_2 in H_2O. The embryos were washed in PBSTw 3x 10 min and 3x 20 min and blocked 2 x 30 min in blocking reagent.

The second overnight incubation with the primary antibodies against the second RNA *in situ* probe was done exactly as the first one. The second TSA reaction was performed exactly as the first reaction.

After extensive final washes, the embryos were mounted in VectaShield medium for fluorescence (Vector laboratories, Cat. Num. H-1000).

3.2.4.3 Antibody preabsorption

An overnight collection of wildtype embryos (w^{1118}) was fixed in 3.7% formaldehyde in PBS at a heptane interface, devitellinzed, washed and stored in methanol as described above. Approximately 100 µl of embryos were rehydrated by washing 3 x 10 min with PBStw (sterile-filtered). The antibody was diluted 1:10 in 200 µl sterile-filtered PBS and incubated with the embryos at 4°C overnight. The next day, the embryos were spun down for 2 min at 2.000 rcf in a table top centrifuge and the supernatant was stored at 4°C.

3.2.5 Imaging

Fluorescent images were analysed on the Olympus FluoView FV1000 confocal microscope with 10x, 20x air, 20x oil, 40x oil and 60x oil objectives.

Color immunostaining and *in situ* hybridizations were analyzed and imaged with a Zeiss AxioPlan Microscope, supplied with Nomarski phase contrast optics.

Images were analyzed using the up-to-date versions of Volocity (PerkinElmer), Imaris (Bitplane), Fiji and ImageJ tools.

3.2.6 Protein biochemistry

3.2.6.1 Embryo Lysate preparation for Western blot

Drosophila embryos at the appropriate developmental stage (2-4 hours AEL at 25°C for mesoderm samples, 4-6 hours AEL for neuroblast samples) were dechorionated with 7% sodium hypochlorite. Rinsed extensively with tap water, the embryos were washed 2 times in PBSTw and transferred to Eppendorf tubes with 1 ml of Lysis Buffer (60 mM TrisHCl, pH 7.5, 150 mM NaCl, 6 mM EDTA, 2% Triton X-100), supplied with Protease Inhibitors (cOmplete Tablets, Mini EDTA-free, EASYpack, Roche 04693159001). The embryos were homogenized by pulling the lysate 10x through a 27G syringe needle. After

15 minutes incubation on ice, the lysates were centrifuged for 16 min at 16.000 rpm at 4°C. The transparent supernatant was transferred to a new Eppendorf tube and the protein concentration in the lysate was measured using standard Bradford assay. The fresh lysate was flash-frozen in liquid nitrogen and stored at -80°C.

200 μl embryo lysate were mixed with 40 μl of 6x Laemmli buffer, incubated 5 minutes at 95°C, chilled on ice for 1 minute, and centrifuged 5 min at maximal speed at 4°C. The samples were divided into aliquots that contained 100 μg of total protein each. The aliquots were stored at -20°C. Western blotting was performed according to standard protocols (Green and Sambrook, 2012).

3.2.6.2 Immunoprecipitation

Embryonic lysates were prepared as described above and diluted with cell lysis buffer to 1 mg/ml. 40 μl aliquots of anti-V5 or anti-HA affinity gel beads were washed 5 x with 1 ml of PBS. The affinity gel beads were collected after each by 30 sec centrifugation at 16.000 g at room temperature. After the last wash, 100 μl of affinity gel-PBS mix were left in the tube. 500 μl of 1mg/ml embryo lysate (500 μg protein) were added to the washed affinity gels. The tubes were incubated overnight at 4°C on a rotating wheel at 12 rounds per minute. The next day the samples were spun down as before and the supernatant was saved to analyse IP efficiency on a Western blot. After the last of three washes with 1 ml of ice cold PBS, all but 40 μl of affinity gel-PBS mix was removed. 10 μl of 6 x Laemmli buffer was added and the immunprecipitated complex was eluted from the beads by heating to 95°C for 5 min.

3.2.6.3 Western blot

For the following western blot analysis standard protocols, 1% polyacrylamide gel and semi-wet transfer were used (Green and Sambrook, 2012).

3.2.7 Chromatin immunoprecipitation procedures

I followed the ChIP-on-chip protocol that was developed in the lab of Eileen Furlong, EMBL Heidelberg (Sandmann et al., 2006b), with minor modifications.

3.2.7.1 Crosslinking of embryonic proteins to DNA

Large amounts (>0.5 g) of precisely staged embryos were collected, as described above. The embryos were transferred into a glass beaker with 7% sodium hypochlorite and incubated for 2 minutes. Then the suspension of embryos in sodium hypochlorite was filtered through a nylon net and rinsed with a large amount of water. Clean, dechorionated embryos were transferred into a falcon tube filled with 9.5 ml of crosslinking buffer (1 mM EDTA, 0.5 mM EGTA, 100 mM NaCl, 50 mM HEPES, pH 8.0), 485 µl of 37% formaldehyde and 40 mL of heptane. The falcon tube was shaken vigorously (500 rpm on the shaker "Innova2100") for 15 minutes. After crosslinking the embryos were spun down for 1 min at 500 g. The liquid was sucked off from the falcon tube and replaced by 45 ml of Stop solution (125 mM glycine, 0.1% Triton X-100 in PBS). The falcon tube was shaken vigorously for 1 min to quench the formaldehyde crosslinking reaction and centrifuged for 1 min at 500 g. The Stop solution was replaced by ice cold PBStr (PBS with 0.1% of Triton X-100) and the tubes were centrifuged for 1 min at 500 g. The embryos were washed once more with PBStr and the suspension of embryos was filtered through a nylon net. The embryos were dried, weighed, put into an Eppendorf tube and flash-frozen in liquid nitrogen. The cross-linked embryos were stored at -80°C.

3.2.7.2 Embryo lysis and nuclei extraction

All procedures were performed on ice, in the cold room (+4°C) or in the centrifuge chilled down to +2°C.

1.5 g aliquot of cross-linked embryos was resuspended in 15 mL of ice cold PBStr, supplied with Roche Protease inhibitors (cOmplete Tablets, Mini EDTA-free, EASYpack, Roche 04693159001). The embryos were homogenized in a cold 15 mL Dounce homogenizer by applying 20 strokes with a loose-fitting pestle. The lysate was transferred into 15 mL falcon tubes and centrifuged at 400 g for 1 min. The supernatant was transferred into a new tube and centrifuged at 1100 g for 10 min.

The pellet was re-suspended in 15 mL of cold cell lysis buffer (85 mM KCl, 0.5% IGEPAL CA-630 v/v, 5 mM HEPES, pH 8.0), supplied with Roche protease inhibitors, and homogenized in a Dounce homogenizer with 20 strokes with the tight-fitting pestle. The

sample was centrifuged at 2000 g for 4 min to pellet the nuclei. The supernatant was discarded, while the nuclei were frozen in liquid nitrogen and stored at -80°C.

3.2.7.3   Sonication of chromatin

The frozen nuclei were resuspended in 0.5 mL of cold nuclear lysis buffer (10 mM EDTA, 05% N-lauroylsarcosine w/v, 5 mM HEPES, pH 8.0), supplied with Roche protease inhibitors, and were incubated at room temperature for 20 minutes. 0.5 mL of cold nuclear lysis buffer supplemented with protease inhibitors were added to the nuclei and the sonication was started.

Sonication of chromatin was performed using a Branson sonifier 250. During sonication the tube was sitting on ice-cold water. Sonication settings were: output 4 (sonication strength) and duty cycle 50 (0.5 sec sonication pulse and 0.5 sec rest). One sonication cycle consisted of 30 seconds sonication and 30 seconds of rest. 15 sonication cycles, mixing the sample after every 5 cycles resulted in the desired size distribution of chromatin fragments in the range of 100-300 bp. If required, the chromatin samples were frozen in liquid nitrogen and stored at -80°C.

3.2.7.4   Chromatin immunoprecipitation

25 μl of 50% ProtA agarose suspension (Protein A agarose Thermo Scientific Prod Pierce® #20333) per precipitation were washed once with RIPA buffer (140 mM NaCl, 1 mM EDTA, 1% Triton X-100 v/v, 0.1% SDS w/v, 0.1% sodium deoxycholate w/v, 10 mM Tris-HCl, pH 8.0). The beads were pelleted by centrifugation at 1000g for 2 min and resuspended in 100 μl of RIPA buffer.

The aliquots of chromatin were thawed on ice. 500 μl of chromatin solution, diluted with TE to the required chromatin concentration, were mixed with 100 μl of 10% Triton X-100, 100 μl of 1% deoxycholate, 100 μl of 1% SDS, 100 μl of 1.4M NaCl. Each new component was dropped into the Eppendorf tube and the tube was turned over several times to mix the content. The chromatin sample was incubated with the prepared ProtA agarose beads in the cold room on a rotating wheel for 1 hour, to remove all chromatin that unspecifically binds to ProtA.

The beads were pelleted as before. The pre-cleaned chromatin was transferred into a new Eppendorf tube, retaining 10 μl (1% of the input) in a separate tube for the following

input analysis. The chromatin was mixed with an appropriate amount of an immune serum of interest. The tube with the chromatin-antibody mix was left overnight at 4°C on the rotating wheel.

In parallel, 25 μl of ProtA beads (agarose or sepharose) were washed once with RIPA buffer, centrifuged, and washed with 1 mg/ml solution of BSA in RIPA. The beads were left in BSA/RIPA solution overnight, slowly mixing on the rotating wheel at +4°C.

On the next day, the tube with antibody-chromatin mix and the tube with beads in BSA/RIPA were centrifuged. BSA/RIPA solution was discarded and the beads were re-suspended in immune serum-chromatin mix. The sample was incubated at 4°C for at least 3 hours (routinely the incubation lasted 4 hours).

The beads were pelleted by centrifugation and washed with the following solutions: 1 x 10 min with RIPA buffer, 4 x 10 minutes with RIPA500 buffer (composition similar to RIPA, but with 0.5 M of NaCl). 1 x 2 minutes with LiCl buffer (250 mM LiCl, 1mM EDTA, 0.5% IGEPAL CA-630 v/v, 0.5% sodium deoxycholate w/v, 10 mM Tris-HCl, pH 8.0) and 2 x 2 min with TE buffer (1 mM EDTA, 10 mM, Tris-HCl, pH 8.0). During each wash the Eppendorf tube with beads was incubated on a slowly rotating wheel at 4°C, with the exception of the 2 min washes, during which the tubes sat on ice.

After all washes were finished, the beads were resuspended in 100 μl of TE. The 1% inputs, retained from the beginning of the ChIP were diluted with 90 μl of TE to get 100 μl sample. All samples were supplied with 50 μg/mL RNase A and incubated at 37°C for 30 minutes. Then the samples were supplied with 0.5% SDS from a 10% stock (w/v) and with 0.5 mg/ml of proteinase K. After 1 hour of incubation at 37°C the samples were put at 65°C to start overnight reverse crosslinking.

On the next day, QiaQuick PCR purification kit was used to purify DNA from the sample.

The ChIP DNA samples were stored at -80°C.

3.2.7.5   ChIP quality control by real time quantitative PCR

A single qPCR reaction mix contained following compounds:

SYBR Green 2X	5 μl
Primer 1 (10 μM)	0.5 μl
Primer 2 (10 μM)	0.5 μl
Water	2 μl
Template	2 μl

Template-free master mix was prepared, and distributed into the wells of the reaction plate. 2 μl of the template were added to the well.

The CFX_3StepAmp+Melt program of CFX96 RealTimeSystem was used in qPCR experiments:

1. 95°C   3 min
2. 95°C   10 sec
3. 55°C   10 sec
4. 72°C   30 sec, go to the 2 step 39 times
5. 95°C   10 sec
6. 65°C   5 sec
7. 95°C   50 sec

1% input samples were used to prepare the standard curves for each primer pair. 1% input and its 1:10, 1:100, 1:1000 dilutions were given to the PCR program as the standard templates with the relative concentrations as much as 1000, 100, 10 and 1 accordingly. Standard reactions were performed in duplicate. Experiment reactions were performed in dupli-, tripli- or quadruplicate.

The program calculates the starting quantity (Sq) in the experimental sample based on the standard curve. To calculate the "enrichment" of an assayed region, the mean Sq value, calculated for the "positive" primer pair was divided by the Sq value calculated for "negative" primer pair: Fold enrichment = Sq(postive primer)/Sq(negative primer).

qPCR analysis was done by BioRad CFX96 TouchTM Real-Time PCR Detection System and BioRad Quantity One 4.6.5 program.

3.2.7.6   Library preparation and sequencing of ChIP samples

All further processing of the ChIP samples, i.e., DNA quantification, quality control, library preparation and sequencing was done by the NGS CSF in Vienna, Austria (Next Gerenetion Sequencing Campus Science Support Facilities GmbH, Dr. Bohr Gasse 3, Vienna, Austria).

Kapa Library Preparation Kits (KAPABIOSYSTEMS, KK8241; 50 libraries) and truseq adaptors were used to prepare ChIP DNA library.

Next Generetion Sequencing was performed on Illumina platforms (GAIIx and HiSeq).

3.2.7.7   Bioinformatic analysis of ChIP data

The obtained reads were mapped against the Drosophila genome (assembly Apr. 2006, BDGP R5/dm3) using Bowtie (Langmead et al., 2009). MACS2 (Feng et al., 2011; Zhang et al., 2008) was used for the peak calling on individual replicates using a respective input as control at FDR threshold of 1%.

3.2.8   Molecular manipulations with *FlyFosmids*

3.2.8.1   *FlyFosmid* purification

FlyFosmid purification for bacteria electroporation was made using the JetStar Plasmid Purification MIDI kit (210050). The kit gives high contamination of broken *FlyFosmid* and bacterial genomic DNA, and therefore cannot be used to prepare DNA for transgenesis microinjections into fly embryos. *FlyFosmid* purification for *Drosophila* transgenesis was performed using the Qiagen Large Construct kit (12462).

The *EL250 E.coli* strain harbors an L-arabinose-inducible enhanced version of the FLP recombinase gene (FLPe), which can be used to excise *frt*-flanked sequences. The protocol used is a modified version of the protocol described by (Lee et al., 2001).

1. A miniculture with LB-12.5ug/mL chloramphenicol was inoculated with *EL250* cells containing the *FlyFosmid* construct. The culture was set at 28° for overnight shaking.

2. The next day, 500uL from the miniculture were put into two cultures of 20mL LB-12.5ug/mL chloramphenicol and shook for 2-3 hours at 28°.

3. 400uL of sterile 10% L-arabinose were added to the culture and were allowed to continue shaking for 1 hour.

4. All but 1 ml of the culture (this did not involve centrifugation) were removed, and this 1mL was used to inoculate a culture of 20mL LB-12.5ug/mL chloramphenicol into which 400uL of 10% L-arabinose were pipetted. The cultures incubated overnight at 28°

(The protocol is provided by Colin Donohoe, Mirka Uhlirova Lab, Institute for Genetics, University of Cologne/CECAD, Cologne Germany).

3.2.8.2 Electroporation of *FlyFosmids* into *EL250* bacteria strain

A single colony of *EL250* cells was used to inoculate a 3 ml LB culture. The culture was incubated overnight at 28°C with shaking. The next day two 40 ml LB cultures were inoculated with 800 µl of overnight culture (1:50 dilution) and were cultured for approximately 4 hours (OD_{600}=0.6). The cultures were spun in 50 ml conical tubes for 10 minutes at 2500 RPM at 2°C. The supernatant was discarded. The pellets were resuspended with 1 mL of ice cold H_2O and gentle pipetting. Once resuspended, an additional 19 ml of H_2O were added to wash the cells. The cells were spun for 10 minutes at 2500 RPM at 2°C and the supernatant was discarded. The pellets were resuspended again with 1 mL of ice cold H_2O and gentle pipetting. Once resuspended, an additional 19 ml of H_2O were added to wash the cells. The cells were spun for 10 minutes at 2500 RPM at 2°C, and the supernatant was very gently discarded to leave about 2 ml. At this point, the pellet is very soft. The cells were spun for 10 minutes at 2500 RPM at 2°C. The remaining H_2O was removed with a pipette, leaving about 100-200 µl of cells behind in each 50 ml conical tube. 100 µl of cells were transferred to 1.5 ml tubes that were sitting on ice. 5 µg of *FlyFosmid* DNA was added to each tube and mixed with the cells gently. The tubes were left to sit on the ice for 10 minutes. The cell/DNA mixture was transferred to 0.1 cm BioRad Gene Pulser cuvettes and pulsed once with the default *E.coli* program (1 mm, 1.8V) of the BioRad GenePulser Xcell Version 3.15. 1 ml of LB medium was immediately added to the cells after electroporation. The cells were transferred to culture tubes and were allowed to recover for 1 hour at 28°C while shaking. After 1 hour of recovery, the cells were briefly spun at 14000 RPM for 15 seconds. All but 200 µl of LB medium was removed, and the remaining 200 µl of cells were plated onto LB plates with 12.5 µg/ml chloramphenicol

(The protocol is provided by Colin Donohoe, Mirka Uhlirova Lab, Institute for Genetics, University of Cologne/CECAD, Cologne Germany)

3.2.8.3 Preparation of the cassettes for *FlyFosmid* recombination

pBSSK-frt-kanR-frt-Start-V5-HA and *pBSSK-frt-kanR-frt-Start-HA-V5* plasmids were used as the templates for synthesis of recombination cassettes. 50 µl PCR mix was prepared according to the following table:

MyTaq RedMix, 3x (BioLine, Cat.No. BIO-25043)	25 µl
Recombination Primer forward (10 µM)	2.5 µl
Recombination Primer reverse (10 µM)	2.5 µl
DNA template	5 ng
Water	up to 50 µl

The following PCR program was used to amplify the recombination cassette:

Step	Temperature (°C)	Duration (sec)
1	98	30
2	98	10
3	60	30
4	72	60
5	go to step 2, repeat 30 times	
6	72	600
7	2	hold

The PCR product was combined with 5 µL of 3M NaAc and 100 µL of ethanol. After vigorous shaking the tube was incubated at -80°C for >30 min and centrifuged at the maximal speed at +4°C for 30 min. The pellet was carefully washed with ice cold 70% ethanol and centrifuged in the same way. Dry pellet was dissolved in 20 µl of water.

3.2.8.4 Electroporation of recombination cassette into cells competent for recombination

The cassette was electroporated into *EL250* cells containing the *FlyFosmid*. The protocol for electroporation is nearly identical to the one used to electroporate *FlyFosmid* into *EL250* cells.

EL250 cells containing the *FlyFosmid* were cultured in LB with 12.5 μl/ml chloramphenicol. However, after the cells grew to a OD_{600} of 0.6, the culture flasks were incubated for 20 minutes in a 42°C water bath with 200 RPM shaking. After the 42°C heat shock, the cells were handled as in the other protocol. 10 or 20 μl of purified recombination cassette were electroporated into the cells. After recovery the cells were plated onto LB plates containing 12.5 μg/ml chloramphenicol and 10 μg/ml kanamycin incubated overnight at 28°C.

(The protocol is provided by Colin Donohoe, Mirka Uhlirova Lab, Institute for Genetics, University of Cologne/CECAD, Cologne Germany)

3.2.8.5 FLP-recmobinase mediated removal of the *frt*-flanked *kanR* gene from the recombined *FlyFosmid*s

The *EL250 E.coli* strain harbors an L-arabinose-inducible version of the FLP recombinase gene, which can be used to excise *frt*-flanked sequences. The protocol used is a modified version of the protocol described by (Lee et al., 2001).

A mini-culture with LB-12.5 μg/ml chloramphenicol was inoculated with *EL250* cells containing the *frt-kanR-frt* construct. The culture was incubated at 28°C overnight with shaking. The next day, 500 μl from the overnight culture were used to inoculate two cultures of 20 ml LB supplemented with 12.5 μg/ml chloramphenicol and shook for 2-3 hours at 28°C. 400 μl of sterile 10% L-arabinose were added to the culture and were allowed to continue shaking for 1 hour. All but 1 ml of the culture (this did not involve centrifugation) were removed, and this 1 ml was used to inoculate a culture of 20 ml LB-12.5 μg/ml chloramphenicol into which 400 μl of 10% L-arabinose were pipetted. The cultures were incubated overnight at 28°C. 100 μl of a 1:100,000 dilution of the overnight culture were plated onto LB-12.5 μg/ml chloramphenicol plates. 16 colonies from the plate were struck onto LB-12.5 μg/ml chloramphenicol + 20 μg/ml kanamycin agar plates to test for kanamycin sensitivity. The plates were incubated for two days at 28°C.

(The protocol is provided by Colin Donohoe, Mirka Uhlirova Lab, Institute for Genetics, University of Cologne/CECAD, Cologne Germany)

3.3 Generation of the transgenic lines that express V5- and HA-tagged Snail, Worniu and Escargot

Chromatin immunoprecipitation assays require high quality anti-protein antibodies. Anti-Snail antibodies have already been successfully used in ChIP experiments (Macarthur et al., 2009; Rembold et al., 2014; Zeitlinger et al., 2007). Anti-Worniu and anti-Escargot antibodies are available but have never been tested for their applicability in ChIPs (Lai and Doe, 2014; anti-Escargot antibody ab43635, Abcam). The sensitivity of the anti-Worniu immune serum (kind gift of Chris Doe) was low. I decided to produce transgenic flies that express Snail, Worniu or Escargot tagged with standard epitope tags (i.e. V5 and HA), to make the proteins detectable by commercially available anti-tag antibodies.

To that produce the tagged genes *FlyFosmid* recombination was selected, as this method allows gene manipulations preserving all *cis*-regulatory sequences of the gene (Berman et al., 2002; Ejsmont et al., 2009). As DNA-recognition zinc finger domains are located on the C-terminus of the Snail protein, I designed an N-terminal recombination, which add V5 and HA tags to the N terminus of the proteins by inserting V5 and HA coding sequences right behind the start codons of the genes.

3.3.1 Generation of *FRT-KanR-FRT-Start-V5-HA* and *FRT-KanR-FRT-Start-HA-V5* recombination cassettes

I designed *FRT-KanR-FRT-Start-V5-HA* and *FRT-KanR-FRT-Start-HA-V5* recombination cassettes (fig. 13 represents cloning strategy used to produce recombination cassettes). In a cassette the start codon is followed by *V5* and *HA* coding sequences. A triple alanine linker separates the tags and the protein. A kanamycin resistance gene (*KanR*) flanked by *FRT* sites was used as a selection marker.

a. **pR6KGFP**

GFP-tag **FRT** KanR **FRT**

Stop

Primer *1.KanNeo_FRT_EcoRI_for*   Primer *2.KanNeo_FRT_BamHI_rev*

EcoR1 **FRT** rpsL KanR **FRT** BamHI

FRT-KanR-FRT cassette

b.

FRT KanR **FRT**

EcoRI BamHI

pBSSK-FRT-KanR-FRT

c.

Primer *3.V5_BamHI_Ala_NotI_for*

Primer *4.V5_BamHI_Ala_NotI_rev*

ATG

V5 Ala   Oligonucleotide *BamHI-Start-V5-ala-NotI*

BamHI NoI1

d.

ATG

FRT KanR **FRT** V5 Ala

EcoRI BamHI NotI

pBSSK-FRT-KanR-FRT-Start-V5

e.

Primer *5.HA_NotI_SacI_Ala_for*   HA   Oligonucleotide

Primer *6.HA_NotI_SacI_Ala_rev*   Not1 SacI *NotI-HA-SacI*

ATG

FRT KanR **FRT** V5 HA

EcoRI BamHI NotI SacI

pBSSK-FRT-KanR-FRT-Start-V5-HA (*KanR-V5-HA*)

f.

Forward primer with genomic   Reverse primer with genomic
homology region   homology region

ATG

FRT KanR **FRT** V5 HA

Homology arm   Homology arm

5′ V5-HA tagging cassette

Figure 13. Cloning strategy used to produce the tagging cassettes for *FlyFosmid* recombinations.

a. *pR6KGFP* vector (Sarov et al., 2006), a template for 3' GFP tagging cassette. The *kanamycin* resistance gene (*KanR*), framed with *FRT* sequences, was amplified with primers, which encode EcoRI and BamHI restriction sites on the 5' overhangs to produce ат *FRT-KanR-FRT* cassette. Abbreviations: FRT, FLP recognition target; GFP, green fluorescent protein.

b. The *FRT-KanR-FRT* cassette was cloned into the *pBSSK* vector (pBlurscript SK Short et al., 1988) via EcoRI and BamHI to assemble *pBSSK-FRT-KanR-FRT*.

c. Oligonucleotide BamHI-Start-V5-ala-NotI produced by primer annealing. The oligonucleotide encodes a start codon and the V5 oligopeptide, framed by BamHI and NotI restriction sites. "ala" stands for the triple alanine linker.

d. *pBSSK-FRT-KanR-FRT-Start-V5* plasmid. Oligonucleotide *BamHI-Start-V5-ala-NotI*, ligated into *pBSSK-FRT-KanR-FRT* vector via BamHI and NotI restriction sites.

e. *NotI-HA-SacI* oligonucleotide, assembled by primer annealing and ligated into *pBSSK-FRT-KanR-FRT-Start-V5* plasmid via NotI and SacI enzymes to assemble *pBSSK-FRT-KanR-FRT-Start-V5-HA* vector, the final template for *FRT-KanR-FRT-Start-V5-HA* tagging cassette.

f. Amplification of the *FRT-KanR-FRT-Start-V5-HA* recombination cassette. Amplification primers include 50 bp homology arms at the 5´ termini.

The cassettes were constructed in several steps. First an *FRT-KanR-FRT* cassette was amplified with 1. *KanNeo_FRT_EcoRI_for* and 2*. KanNeo_FRT_BamHI_rev* primers taking the *pR6KGFP* vector containing an *EGFP*-tagging cassette (*EGFP-frt-KanR-frt*) (Sarov et al., 2006) as a template. EcoRI and BamHI restriction sites were added to the termini of the cassette. The PCR product was sub-cloned into *pCRII* vector using the TOPO TA Cloning ® kit (Invitrogen) and subsequently cloned into the multiple cloning site of *pBS SK (-)* vector via *EcoRI* and *BamHI* sites (fig. 13)*. The resulting plasmid was named *pBSSK-FRT-KanR-FRT* (fig. 13a., b.).

Second, two primers (3*. V5_BamHI_Ala_NotI_for* and 4*. V5_BamHI_Ala_NotI_rev*) were annealed to produce a "*Start-V5*" oligonucleotide with sticky *BamHI* and *NotI* overhangs. The "*Start-V5*" oligonucleotide was ligated into *pBSSK-FRT-KanR-FRT*, cut by *BamHI* and *NotI* enzymes, resulting in the *pBSSK-FRT-KanR-FRT-Start-V5* vector (fig. 13c., d.).

In the third step, an "*HA*" oligonucleotide with the sticky *NotI* and *SacI* overhangs was assembled by annealing 5*. HA_NotI_SacI_Ala_for* and 6*. HA_NotI_SacI_Ala_rev* oligonucleotides. "HA" oligonucleotide was ligated into *pBSSK-FRT-KanR-FRT-Start-V5*, cut by *NotI* and *SacI* enzymes. The final plasmid was named *pBSSK-FRT-KanR-FRT-Start-V5-HA* or "*KanR-V5-HA*" (fig. 13e.).

In a similar way, *pBSSK-FRT-KanR-FRT-Start-HA-V5* (*KanR-HA-V5*) was generated by ligating "*Start-HA*" and "*V5*" oligonucleotides (annealed primers 7*.

HA_Start_Ala_BamHI_NotI_for + 8. *HA_start_Ala_BamHI_NotI_rev*, and 9. *V5_Ala_NotI_SacI_for* + 10. *V5_Ala_NotI_SacI_rev*) into *pBS SK (-)*, using the same sets of restriction enzymes.

3.3.2 *FlyFosmid* recombineering using fosmid recombination cassettes

FlyFosmid collection (Ejsmont et al., 2009) was searched through to find the *FlyFosmid* clones which carry genomic fragments that cover full length *snail*, *worniu* and *escargot* genes and all known cis-regulatory modules of the genes. *FlyFos021004*, *FlyFos015520* and *FlyFos017825* clones were selected for the subsequent *FlyFosmid* recombinations (Ejsmont et al., 2009; fig. 14).

FRT-KanR-FRT-V5-HA and *FRT-KanR-FRT-HA-V5* tagging cassettes were amplified from the constructed plasmids (*KanR-V5-HA and KanR-HA-V5*), using primers which include 50 bp homology regions at the 5' end (fig. 15; primers 11.-19.). Homology regions are required to direct the recombination to the respective genomic region. The PCR-amplified cassettes were recombined into the *FlyFos021004*, *FlyFos015520* and *FlyFos017825* fosmids to create the following constructs:

(1) *FlyFos015520-FRT-KanR-FRT-V5-HA-snail*;

(2) *FlyFos015520-FRT-KanR-FRT-HA-V5-snail*;

(3) *FlyFos017825-FRT-KanR-FRT-V5-HA-worniu*;

(4) *FlyFos017825-FRT-KanR-FRT-HA-V5-worniu*;

(5) *FlyFos021004-FRT-KanR-FRT-V5-HA-escargot*;

(6) *FlyFos021004-FRT-KanR-FRT-HA-V5-escargot*.
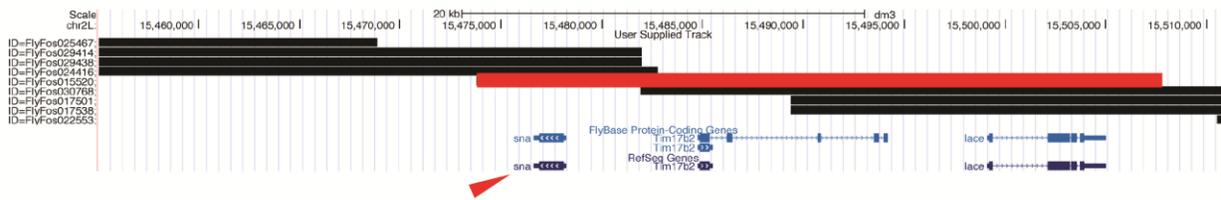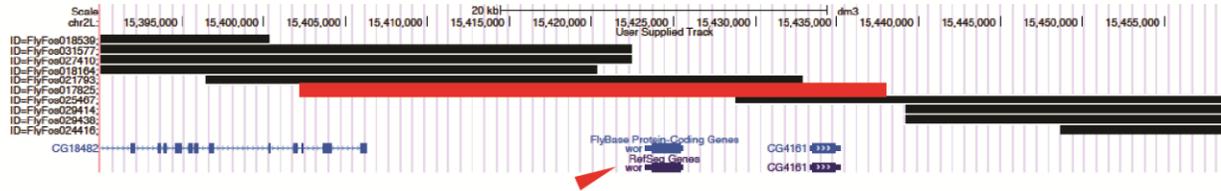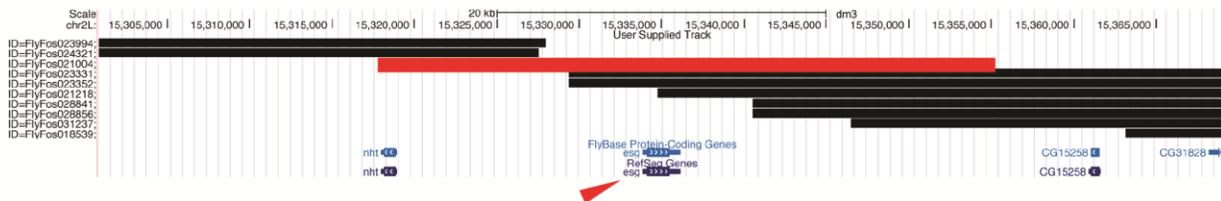
Figure 14. Selection of *FlyFosmid* clones for subsequent tagging recombination. Screen shots of the *Drosophila* UCSC genome browser with annotated RefSeq genes and the uploaded *FlyFosmid* track list. Black bars indicate the genomic region cloned into the *pFlyFos* vector. ID numbers at the left side indicate the numbers clones in the collection.

a. *snail* (red arrow) within its genomic environment and *FlyFosmids* that cover the region around *snail*. *FlyFos015520*, which covers *snail* along with *Tim17b2* and *lace* and includes all described *snail* enhancers (located at -0.8 kb; -1.8 kb; -2.8 kb; -7 kb) was selected for following recombination experiments (highlighted with the red stripe).

b. *worniu* (red arrow) within its genomic environment and *FlyFosmids* that cover the region around *snail*. *FlyFos017825*, which covers *worniu* and the neighboring gene *CG4161* was selected for the following recombination experiments (highlighted with the red stripe).

c. *escargot* (arrow) within its genomic environment and *FlyFosmids* that cover the region around *escargot*. *FlyFos021004*, which covers *escargot* and the neighboring gene *nht* was selected for following recombination experiments (highlighted with the red stripe).

Successful recombinants were selected by plating the bacteria on LB-kanamycin medium. Sequence analysis proved that no point mutations were introduced into the genes, and the recombination products correspond to the designed constructs.

a. *pBSSK-FRT-KanR-FRT-Start-V5-HA (KanR-V5-HA)*

Forward primer with genomic
homology region

Reverse primer with genomic
homology region

b. Homology arm

5′ V5-HA tagging cassette

fragment of fly genome

snail

fragment of fly genome
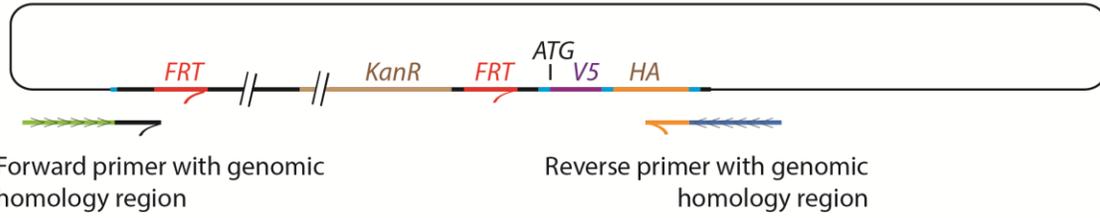
snail region in FlyFosmid vector
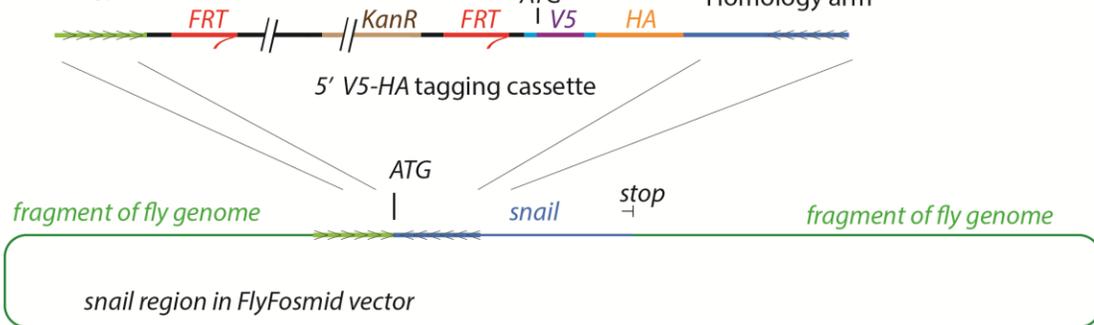
c.

Homology arm

Homology arm

fragment of fly genome

FRT-KanR-FRT; V5-HA-snail in FlyFosmid vector
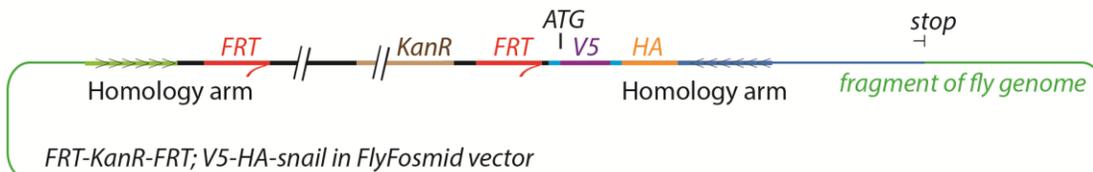
d. *fragment of fly genome*

V5-HA-snail in FlyFosmid vector

e.

enhancer

gene

Snail

V5   HA

ChIP, western blot
immunohistochemistry
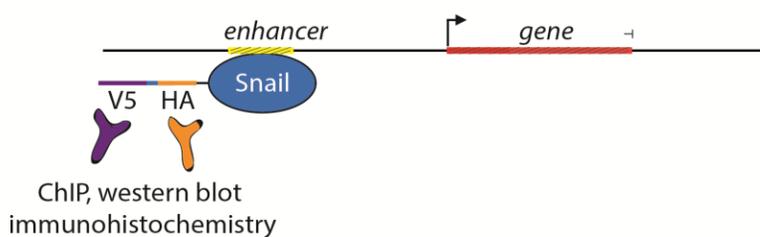
Figure 15. *FlyFosmid* recombination.
a. Amplification of the *FRT-KanR-FRT-Start-V5-HA* tagging cassette with primers that introduce 50 bp homology arms at the 5' termini.
b. Recombination of the *FRT-KanR-FRT-Start-V5-HA* tagging cassette into the *snail* gene contained in a *FlyFosmid*. The PCR-amplified tagging cassette was electroporated into the *FlyFosmid*-carrying bacteria. Reparation machinery, especially available in the genome of recombination strain *EL250*, recognizes homology arms on the tagging cassette as double-strand breaks of DNA. The enzymes find homology regions on the *FlyFosmid* and recombine the cassette into the *FlyFosmid*, in a process resembling DNA repair.
c. *V5-HA-snail* in the *FlyFosmid* after recombination. *V5*- and *HA*- tag sequences are inserted between the *ATG* and the second codon of the *snail* gene. *KanR*, a selective marker for kanamycin resistance, is framed by *FRT* sites.
d. *FlyFosmid* with *V5-HA-snail* after a flippase-mediated excision recombination removed the *KanR* gene. The *FlyFosmid* was prepared for transgenesis.
e. Schematic representation of V5-HA-Snail bound to an enhancer in a nucleus of a transgenic fly.

Bacteria with the recombined *FlyFosmids* were cultured in presence of L-arabinose to induce expression of flippase (encoded in the bacterial genome under control of an arabinose-inducible promoter) in order to remove the *FRT*-flanked antibiotic resistance gene. Single colonies derived from the induced cells were tested for kanamycin sensitivity. If any colony did not grow on LB-kanamycin plates, PCR reactions were run to confirm the absence of *KanR* (fig. 15d.).

The resulting constructs were re-named:

(7) *FlyFos015520-V5-HA-snail*;   (*V5-HA-snail*)

(8) *FlyFos015520-HA-V5-snail*;   (*HA-V5-snail*)

(9) *FlyFos017825-V5-HA-worniu*;   (*V5-HA-worniu*)

(10) *FlyFos017825-HA-V5-worniu*;   (*HA-V5-worniu*)

(11) *FlyFos021004-V5-HA-escargot*,   (*V5-HA-escargot*)

(12) *FlyFos021004-HA-V5-escargot*.   (*HA-V5-escargot*)

The DNA sequences of the region around the start codon of the tagged gene were amplified and verified by sequencing (primers 20.-25.). As an example, the structure and the sequence of the recombined region of the *FlyFos015520* clone, and of the *FlyFos015520-V5-HA-snail* construct after final recombination is shown in fig. 16.

A.

cgaactatataactcgcctctcgatcgccgatctcccgatttacccatctcgatcagtaccggaaactaaaacttaatcacacacacatcaaaa

ATG GCC GCC AAC TAC AAA AGC TGC CCG CTA AAG AAG CGC CCC ATT GTC TTC GTG GAG GAG CGT CTG CCA CAA ACG GAG GCC TTG GCC CTG ACC AAG GAC TCA CAG TTT GCC CAG GAT CAG CCG CAG

B.

*Cgaactatataactcgcctctcgatcgccgatctcccgatttacccatctcgatcagtaccggaaactaaaacttaatcacacacacatcaaaa*t gatatcgaattctcctcgggaagttcctattctcta gaaagtataggaacttc gtcgacagattggatcc

<div align="center">FRT</div>

atg ggc aag ccc atc ccc aac ccc ctg ctg ggc ctg gat agc acc gct gcg gcc gcg tac cca tat gat gtt cca gat tac gct

<div align="center">V5-tag   ala-liner   HA-tag</div>

GCC GCC AAC TAC AAA AGC TGC CCG CTA AAG AAG CGC CCC ATT GTC TTC GTG GAG GAG CGT CTG CCA CAA ACG GAG GCC TTG GCC CTG ACC AAG GAC TCA CAG TTT GCC CAG GAT CAG CCG CAG

Figure 16. The sequence that surrounds the start codon of *snail* before and after *FlyFosmid* recombination.
A: The sequence of *FlyFos015520* around the *snail* start codon.
B: The sequence of *FlyFos015520-V5-HA-snail*, with the region around *snail* start codon after recombination.
Color code: Gray: primers, which were used for sequence analysis of the constructs. Yelow: open reading frame of *snail*. Red: start-codon.

3.3.3 Establishment of the *FlyFosmid* transgenic fly lines

3.3.3.1 Initial injections to ensure *FlyFosmid* transgenesis into *attP2* recombination site

The φC31 integrase, encoded in the genome of *φC31* phage, recognizes the *phage attachment site* (*attP*), present in the phage DNA and the *bacteria attachment site attB* site on the bacterial chromosome and subsequently catalyzes recombination between two DNA sites to insert phage DNA into the bacterial chromosome (Thorpe and Smith, 1998). This system was used to produce φC31-mediated *Drosophila* transgenesis system. Transgenesis vectors (including *pFlyFos* vector) contain an *attB* site (Ejsmont et al., 2009). Various fly lines with transgenic *attP* sites in diverse genome locations and transgenic expression of φC31-integrase under control of different germline-specific promoters are available. *attP40* line, due to its reliable performance and high transgenesis rate most frequently used for transgenesis in our lab, carries the *attP* site on the second chromosome and expresses the integrase in the germline under control of *nanos* regulatory elements (*y¹ w¹ nos-phiC31; attP40*). Since *snail*, *worniu* and *escargot* are located on the second chromosome, insertion of the *FlyFosmids* with the tagged genes into the third chromosome would facilitate subsequent rescue experiments. For this reason I had to dispose *attP40* line from following *FlyFosmid* transgenesis and take *y¹ sc¹ v¹ nos-phiC31; ; attP2* line, which carries the *attP* site

on the third chromosome. This line is generally known for being weaker and less efficient in transgenesis than the *attP40* line. Moreover, the *FlyFosmid* constructs are unusually large and therefore difficult to insert into fly genome. The *pFlyFos* vector is 9,630 bp long, while the average size of the genomic fragment, ligated into vector, is ~36 kb (Ejsmont et al., 2009). The size of *FlyFos015520-V5-HA-snail* construct is 34 kb. DNA constructs of such a big size were never used for transgenesis in our lab.

I therefore performed a number of preliminary control injections to evaluate the general transgenesis rates for (1) the *attP2* line and the transgenesis rates for (2) *pFlyFos* injections. To test *attP2*, *pValium22msni*, an 8.5 kb-long *attB*-containing vector, was injected into the embryos of the *attP2* fly line. As a control the same vector was injected into *attP40* line. To test transgenesis rates of the *pFlyFos* constructs, two of the modified *FlyFosmids* (*V5-HA-snail* and *HA-V5-snail* constructs) were injected into the *attP40* fly line. Table 1 summarizes the number of injected embryos, hatched larvae, crossed flies and transgenic flies from the control injections. The viability seemed higher for the *attP40* line compared to *attP2,* as more embryos hatched and gave rise to adult fertile flies, when *pValium22msni* was injected. But even with the larger *FlyFosmid* vectors transgenic flies were obtained if a sufficient number of *attP2* embryos was injected. Thus both *attP2* fly line and *FlyFosmid* vector successfully passed the preliminary transgenesis test.

Table 1. DNA injections into *Drosophila* embryos to evaluate the transgenesis rates of the *attP2* landing site in the y^1 sc^1 v^1 *nos-phiC31*; ; *attP2* fly line and the transgenesis rate of *FlyFosmids* in y^1 w^1 *nos-phiC31*; *attP40*. The percentages refer to the fraction of injected embryos. G0 males or females were crossed to *IF/SM1; MKRS/Tm6B* balancer line (in the case of *pValium22msni* injection the G0 flies were crossed to the special balancer line w^+v^-; *IF/SM1; MKRS/Tm6B*). The progeny was screened for the presence of v^+ (*pValium22msni* injection) or eye expression of red fluorescent protein dsRed (*pFlyFos* injections, Ejsmont et al., 2009).

Construct	Line	Injected embryos	Hatched larvae	Crossed flies	Sterile crosses	Independent insertions
pValium22msni	*attP2*	1000	167 (16%)	22 (2%)	3 (0.3%)	10 (0.1%)
pValium22msni	*attP40*	487	150 (30%)	76 (15%)	0	3 (0.6%)
V5-HA-snail	*attP40*	2000	684 (34%)	296 (13%)	18 (0.9%)	2 (0.1%)
HA-V5-snail	*attP40*	230	189 (82%)	76 (33%)	5 (2.1%)	0

3.3.3.2   Insertion of *FlyFosmids* into the *Drosophila* genome

Bacteria cells, carrying modified *FlyFosmid* clones were sent to TheBestGene Inc. injection service (*thebestgene.com*). The injections were performed into the y^1 w^- *vas-phiC31; attP2* fly line and surviving flies were crossed to screen their progeny for transgenes. The injections (400 embryos per construct) yielded no transgenic flies.

FlyFosmids were purified with the Qiagen large construct purification kit (see methods) and injected into embryos of the y^1 sc^1 v^1 *nos-phiC31;; attP2* line. Transgenic homozygous lines with *V5-HA-snail*, *HA-V5-snail*, *V5-HA-worniu*, *HA-V5-worniu, V5-HA-escargot,* and *HA-V5-escargot* insertions were established. Table 2 summarizes the number of injected embryos, hatched larvae, crossed flies and independent transgene insertions.

V5-HA-worniu stock appeared homozygous lethal. I therefore used *HA-V5-worniu* for all following experiments.

Table 2. Injections of *pFlyFos* constructs into *Drosophila* embryos of y^1 sc^1 v^1 *nos-phiC31*; ; *attP2* fly line to produce fly transgenes with express the tagged versions of Snail, Worniu and Escargot. The percentages refer to the fraction of injected embryos. G0 males or females were crossed to *IF/SM1; MKRS/Tm6B* balancer. The progeny was screened for expression of red fluorescent protein dsRed in the eye (Ejsmont et al., 2009).

Construct	Injected embryos	Hatched larvae	Crossed flies	Steriles	Independent insertions
V5-HA-snail	953	282 (29%)	91 (9.5%)	1 (0.1%)	1 (0.1%)
HA-V5-snail	1346	628 (46%)	287 (21%)	19 (1.4%)	multiple
V5-HA-worniu	507	239 (47%)	66 (13%)	5 (0.98%)	1 (0.19%) [a]
HA-V5-worniu	385	112 (29%)	32 (8.3%)	6 (1.55%)	1 (0.25)
V5-HA-escargot	1380	573 (41%)	310 (22.4%)	8 (0.58%)	multiple
HA-V5-escargot	1010	550 (54%)	200 (20%)	10 (1%)	9 (0.9%)

[a] The transgene is homozygous lethal.

To test the functionality of the recombinant proteins, the transgenes were crossed into the respective mutant background (*snail^18^*, *worniu^4^* or *escargot^35Ce-3^*) and tested if they can rescue the mutant phenotype to viability. Neither transgene could fully rescue the mutation of the corresponding genes. A subsequent analysis of developmental phenotypes showed that the transgenic *V5-HA-snail* cannot rescue gastrulation in *snail^18^* mutant embryos.

4   Results

4.1   Expression patterns of *snail*, *worniu* and *escargot*

To compare Snail-, Worniu- and Escargot- binding sites during embryogenesis a clear understanding of the spatial and temporal expression patterns of these proteins is required. Although the expression patterns of *snail*, *worniu* and *escargot* have been described (Alberga et al., 1991; Ashraf et al., 1999; Cai et al., 2001; Ip et al., 1992, 1994; Leptin, 1991; Leptin and Grunewald, 1990; Whiteley et al., 1992; Yagi and Hayashi, 1997), a comparative analysis of the spatio-temporal expression patterns is missing.

In this chapter I am going to talk about three types of expression patterns: primarily, expression patterns of the endogenous *snail*, *worniu* and *escargot* genes, detected by RNA *in situ* hybridization. Secondly, expression patterns of endogenous Snail, Worniu and Escargot proteins, detected by specific antibodies. Finally, expression patterns of the *FlyFosmid* transgenic proteins V5-HA-Snail, HA-V5-Worniu and V5-HA-Escargot (described in Materials and Methods//Generation of the transgenic lines that express V5- and HA-tagged Snail, Worniu and Escargot) detected with anti-V5 antibodies.

If gene expression is regulated at the post-transcription level, an expression pattern of a protein does not recapitulate the one of an mRNA. *FlyFosmid* carry large randomly shared fragments of genome that do not necessarily comprise the entire set of the cis-regulatory modules relevant to the gene of interest. *FlyFosmid* manipulations, described in Materials and Methods, leave an FRT site directly in front of the start codon of the tagged gene. Both these facts might cause differences between expression patterns of endogenous Snail, Worniu, Escargot and transgenic V5-HA-Snail, HA-V5-Worniu, V5-HA-Escargot. Therefore, it is important to clearly distinguish between three types of expression patterns. In each part of the chapter I indicate specifically which expression pattern (RNA, endogenous protein or transgenic tagged protein) is currently discussed.

4.1.1 Expression patterns of *snail*, *worniu* and *escargot* during early embryonic development detected with *in situ* hybridizations

At early stages of development (stages 4-7) *snail* is expressed in the mesoderm (Leptin, 1991; fig. 17a., b.). At stage 7 *snail* mRNA can be detected in the invaginated mesoderm and in the posterior midgut (fig. 17a.). At stage 8 expression of *snail* ceases in the mesoderm, while the first stripes of *snail* mRNA can be detected in the neurogenic ectoderm (fig 17c., white arrow) and the first *snail*-positive neuroblasts delaminate (fig. 17c., yellow arrows). At stage 9 *snail* is expressed in neuroblasts and persists in stripes in the neurogenic ectoderm (fig. 17d., e., yellow arrows). At stage 10 the ectoderm loses expression of *snail* and the number of *snail*-positive neuroblasts declines (fig 17f.-h., compare with the numbers of *worniu$^+$* neuroblasts in fig. 18g.).

worniu has been described as a pan-neuroblast marker (Ashraf et al., 1999). The first *worniu*-positive cells (the first neuroblasts) can be detected at stage 8 (fig. 18a., b., arrows). At the stages 9 - 10 *worniu* expression marks delaminated neuroblasts of the ventral nerve cord and procephalon (fig. 18c.-f.). At later stages (st. 11-12, fig. 18g.), the number of *worniu*-positive cells increases (fig. 18f., compare to the fig. 17h.).

The expression of *escargot* is dynamic throughout development (Whiteley et al., 1992; Yagi and Hayashi, 1997). Its expression in the developing CNS has never been analyzed in detail. At stage 7 *escargot* is expressed in the procephalic region and at the dorsal side and in a segmentally repeated pattern in the ventral neuroectoderm, covering the lateral and medial column from which neuroblasts will delaminate, this pattern has been described in detail (Whiteley et al., 1992; Yagi and Hayashi, 1997); fig. 19a. At stage 8 *escargot* is expressed in segmental transversal stripes in neurogenic ectoderm (fig. 19b., c.). Early stage 9 is characterized by segmental stripes of *escargot* expression in the entire ectoderm (fig. 19d.). At stage 10 *escargot* expression comes up and persists in the cells of the ventral midline, while the segmental stripes vanish. At late stage 11 a small number (1 CNS neuroblast and several PNS sells per segment) of *escargot*-positive CNS and PNS neuroblasts could be detected (fig. 19e., f.; arrows highlight *escargot*-positive neuroblasts).
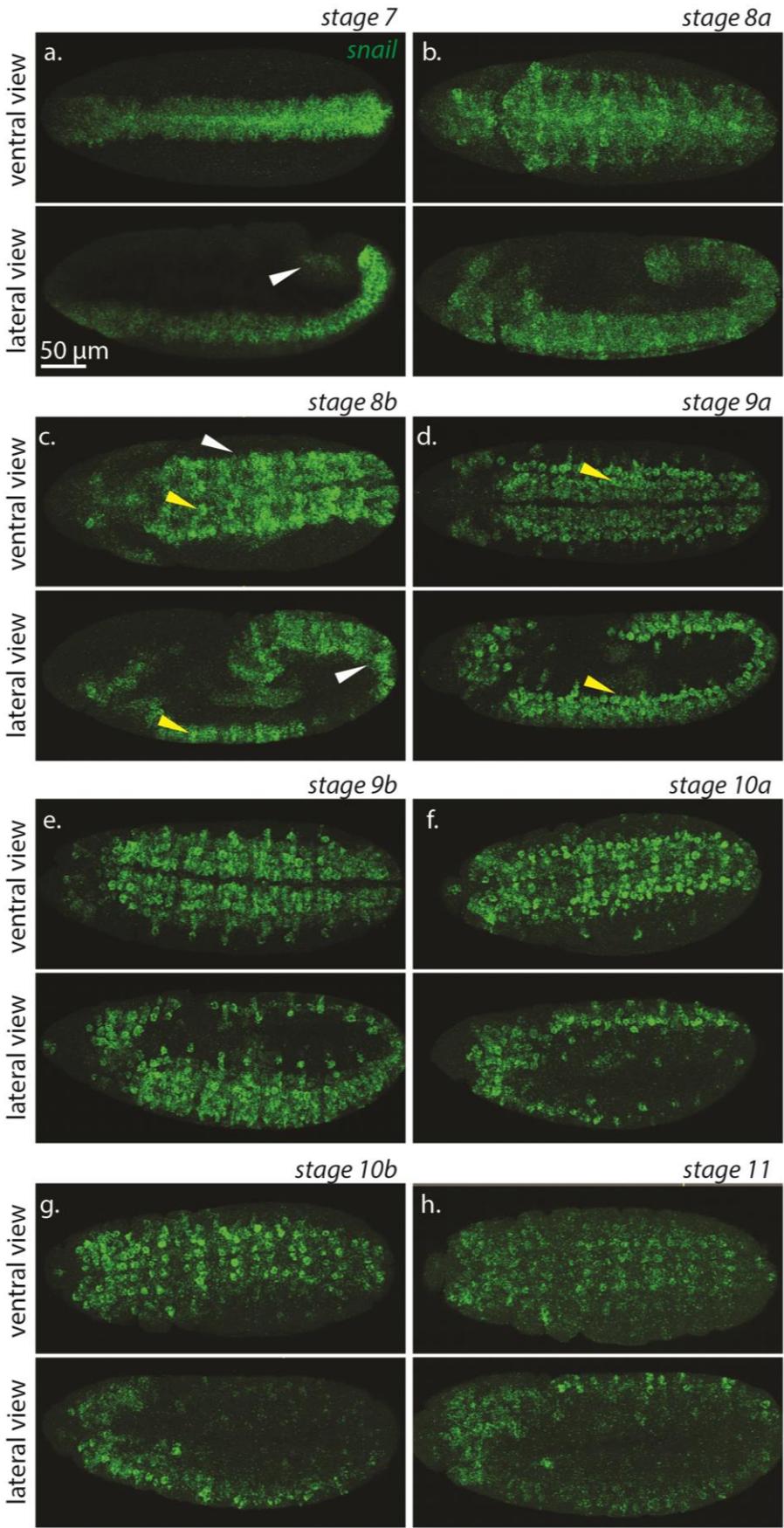
Figure 17. Expression patterns of *snail* during early embryogenesis, stages 7 - 11. Fluorescent *in situ* hybridizations, the expression pattern is shown as maximum intensity projections of a stack of confocal images. Lateral and ventral views, anterior is to the left. The scale bar is 50 µm.

a. Embryo at stage 7: expression of *snail* in mesoderm and posterior midgut (white arrow).

b., c. Embryos at stage 8: *snail* expression ceases from mesoderm and starts in neurogenic ectoderm (white arrow) and in neuroblasts (yellow arrow).

d., e. Embryos at stage 9: massive expression of *snail* in neurogenic ectoderm and in neuroblasts of the segmented germ band (white arrow) and in the head.

f.-h. Embryos at stage 10-11: expression of *snail* in neurogenic ectoderm is down-regulated, the amount of *snail*-positive neuroblasts gradually decreases.

stage 8 — *worniu*
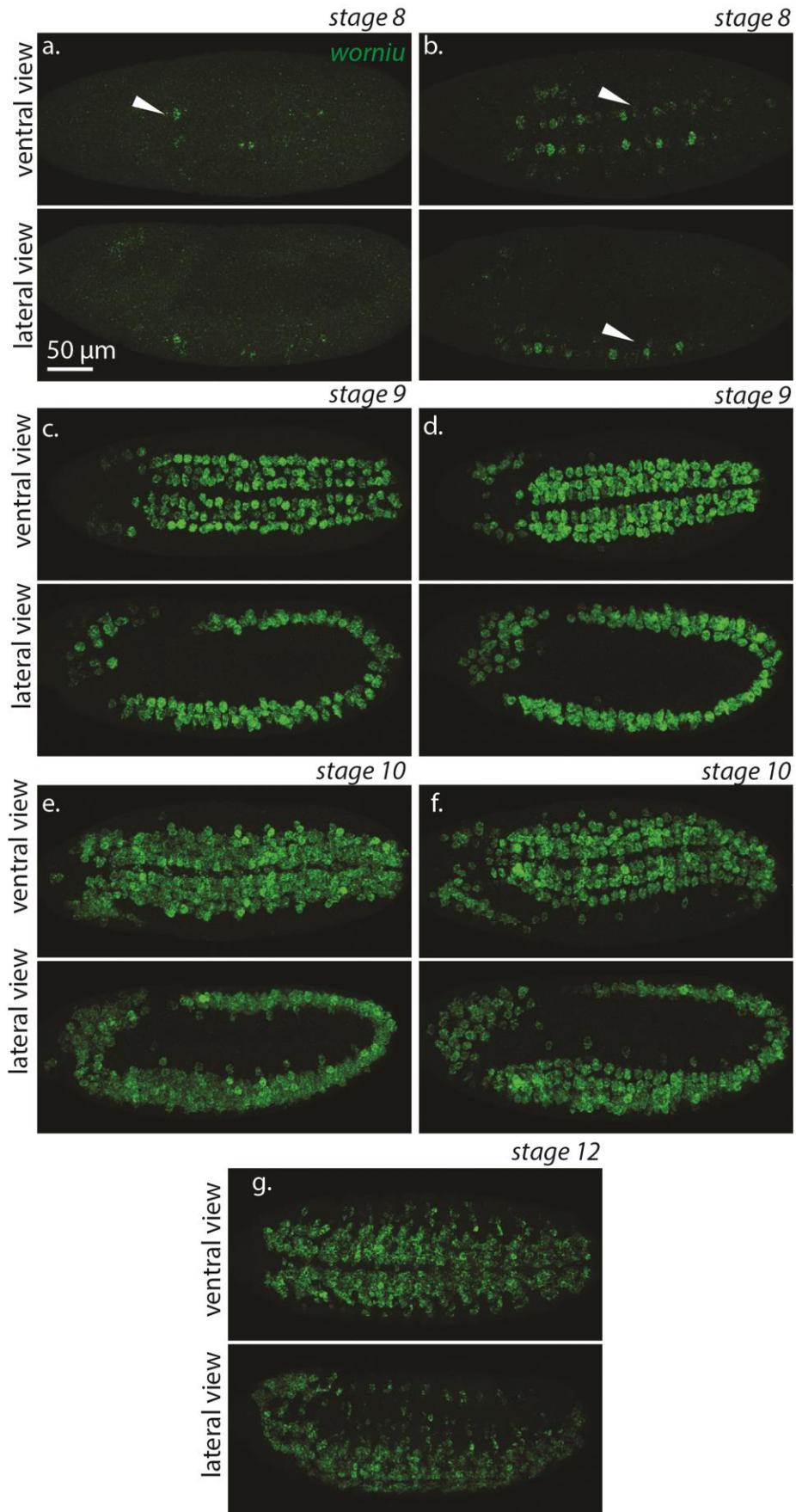
Figure 18. Expression patterns of *worniu* during early embryogenesis, stages 8 - 12. Fluorescent *in situ* hybridizations, the expression pattern is shown as maximum intensity projections of a stack of confocal images. Lateral and ventral views, anterior is to the left. The scale bar is 50 μm.

a., b.Embryos at stage 8: first *worniu*-positive neuroblasts are in the medial neuroblast columns (white arrows).

c., d. Embryos at stage 9: Second and third wave of neuroblast formation. A regular neuroblast pattern is formed in the segmented germband. Also neuroblasts in the embryonic head express *worniu*.

e., f. Embryos at stages 10: neuroblasts formed during the fourth and fifth waves of neuroblast delamination also express *worniu*.

g. Embryo at stage 12: formation of the embryonic CNS.
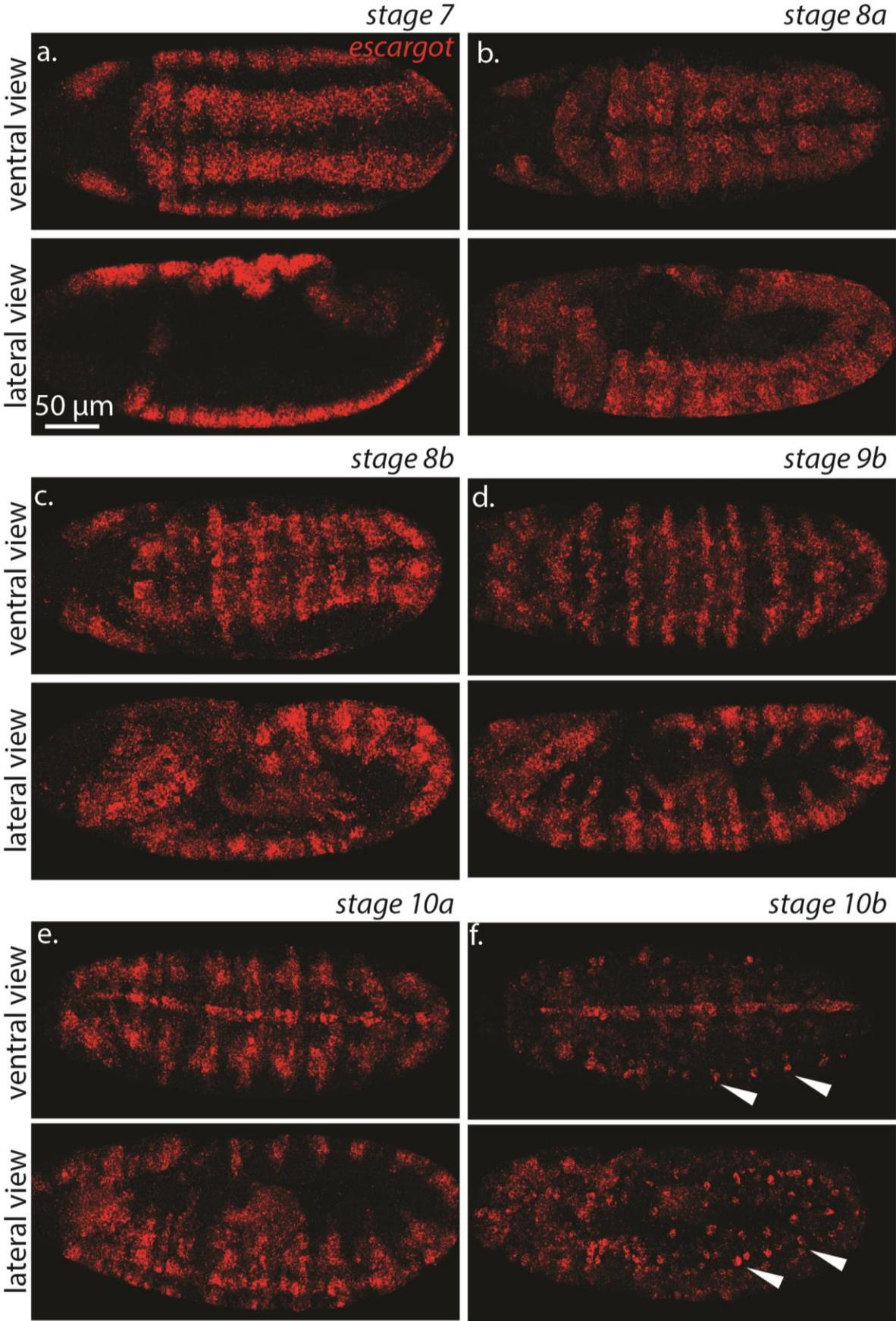
Figure 19. Expression patterns of *escargot* during embryogenesis, stages 7 - 10. Fluorescent *in situ* hybridizations, the expression pattern is shown as maximum intensity projections of a stack of z-stack confocal images. Lateral and ventral views, anterior is to the left. The scale bar is 50 μm.

a. Embryo at stage 7. *escargot* is expressed in longitudinal stripes: one stripe in ventral and one stripe lateral neurogenic ectoderm, one stripe in dorsal ectoderm. Longitudinal stripes of *escargot* expression are sub-divided into transversal segments.

b.Embryo at early stage 8. *escargot* is expressed in neurogenic ectoderm. Expression in the lateral and ventral ectoderm ceased.

c. Embryo at late stage 8. *escargot* is expressed in neurogenic ectoderm in a segmentally repeated pattern.

d. Embryo at stage 9. *escargot* is expressed in neurogenic ectoderm in clear segment stripes.

e. Embryo at early stage 10. *escargot* is expressed in segmental stripes in neurogenic ectoderm; expression starts i the midline.

f. Embryo at late stage 10. *escargot* is expressed in the midline and in single neuroblasts of the peripheral nervous system (white arrows).

4.1.1.1   Comparison of *snail*, *worniu* and *escargot* expression patterns on fluorescent *in situ* hybridization

To collect more information about expression patterns of *snail*, *worniu* and *escargot*, I decided to perform a comparative analysis of expression of these genes. The analysis was performed on the neuroblast cells of the ventral nerve cord. Expression of Snail-like TFs in the head neuroblasts was not analyzed.

4.1.1.2   Comparison between the expression patterns of *snail* and *worniu*

At stages 4-7, while *snail* is expressed in the mesoderm, no *worniu* expression can be detected in embryos. The first *worniu*-positive cells appear at stage 8 (fig. 20f´.). Both *snail* and *worniu* mRNA can be detected in neuroblasts. However, neuroblasts appeared to be sub-divided in three populations: majority of neuroblasts are *worniu*-positive; some neuroblasts are *snail*-positive, and there are few *snail*- and *worniu*-positive neuroblasts (fig. 20e.). As no clear spatial pattern of co-expression could be detected, expression of *snail* and *worniu* was compared in neuroblasts of the segmented germband at different time points of embryonic development. The neuroblast of the 1-3 lineage (neuroblast 1-3) is the most lateral in a hemisegment and therefore can be easily identified (fig. 20d., arrow). It was therefore used to follow the fate of an individual neuroblast on cross-sections of confocal image stacks (fig. 20c., d., e., arrows). The neuroblast 1-3 has not yet been formed at stage 8, as it is part of the third wave of delamination (Doe, 1992; fig. 20f´., arrow). The neuroblast 1-3 delaminates from the neurogenic ectoderm at stage 9. The cell expresses *snail* (fig. 20f´´.). The neuroblast 1-3 starts expression of *worniu* at stage 9b (fig. 20f´´´.) and loses *snail* expression, however remaining *worniu* positive, at early stage 10 (fig 20f´´´´.). Thus, neuroblast 1-3 undergoes a successive expression of TFs. Initially the cell expresses *snail,* later starts expression of *worniu* and after co-expression of *snail* and *worniu* the neuroblast down-regulates and loses expression of *snail.* Later the neuroblast is *worniu* positive*.*

To test whether this temporal succession of *snail* and *worniu* is a cell specific or a general feature of neuroblast development, I followed TF profiles of the other neuroblasts in a comparative analysis of *snail* and *worniu* co-expression on physical cros-sections of *Drosophila* embryos (fig. 21).
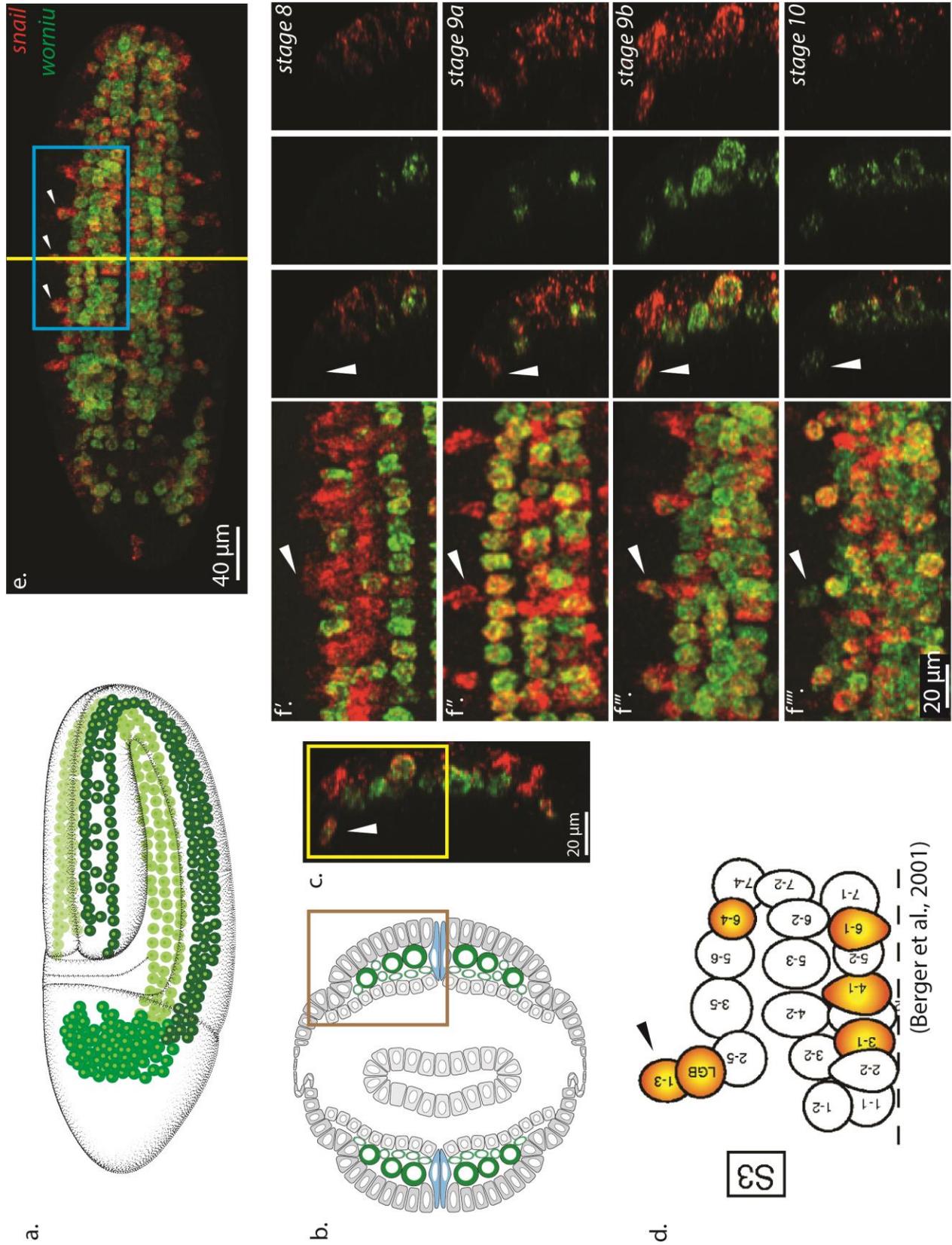
Figure 20. Comparison of *snail* and *worniu* expression in the ventral nerve cord, stages 8-10. Double anti-*snail*, anti-*worniu* fluorescent *in situ* hybridizations. *snail* is red; *worniu* is green.

a. Schematic drawing of neuroblast localization in a *Drosophila* embryo at stage 9. Neuroblasts are visualized as big green circles. Lateral view. Anterior is to the left (Campos-Ortega and Hartenstein, 1997).

b. Schematic drawing of a cross-section through a *Drosophila* embryo at stage 9 (Campos-Ortega and Hartenstein, 1997). The ventral side is to the right. Neuroblast cells are highlighted in green. The yellow rectangle shows the area chosen for analysis in d., right panels of f..

c. Embryo at stage 9. Digital cross-section through a stack of confocal images, the scale bar is 20 μm. The ventral side is to the right. The yellow rectangle shows the area chosen for analysis in f. The white arrow labels the neuroblast 1-3, which was selected for the following analysis of neuroblast fate in f.

d. The schematic drawing of the neuroblast delamination map at stage 9 (Berger et al., 2001). The cells of the third wave of neuroblast delamination are highlighted in yellow. The sells which were formed during previous two waves of neuroblast delamination are white. The black arrow points out the neuroblast 1-3, the most lateral neuroblast in a segment, selected for the analysis of neuroblast fate in f.

e. Embryo at stage 9, ventral view, anterior to the left. Maximum intensity projections of a stack of confocal images. The scale bar is 40 μm.

The blue rectangle highlights the region, which was chosen for the analysis in f. The yellow line represents the plane, which was selected to analyze digital cross-sections of the confocal stacks in f.

f. Time points of neuroblasts delamination in the area highlighted with the blue box in b. The left image shows a maximum intensity projection of a z-stack of confocal images; the other images are digital cross-section through the plane indicated by the yellow line in b. The scale bar is 20 μm.

f'. First wave of neuroblast delamination. The neuroblast 3-1 has not been formed yet. White arrow indicates place where neuroblast is going to delaminate.

f''. Beginning of the third wave of neuroblast delamination. The neuroblast 3-1 is delaminating from the neurogenic ectoderm. The cell is *snail*-positive. No *worniu* expression can be detected.

f'''. The neuroblast 3-1 co-expresses *snail* and *worniu* during delamination.

f''''. The mature neuroblast 3-1: the cell is *worniu*-positive, expression of *snail* ceased.

a.

snail worniu

15μm

neurogenic ectoderm, overview

b.

stages of neuroblast delamination

c.

d.
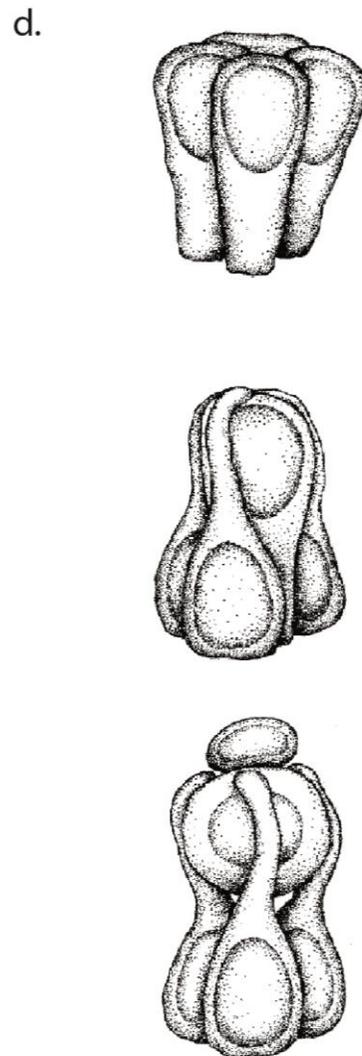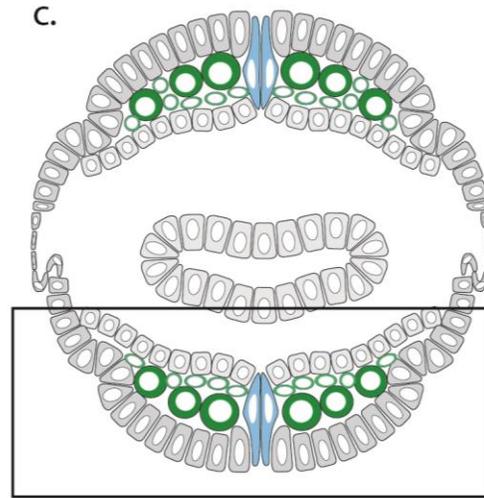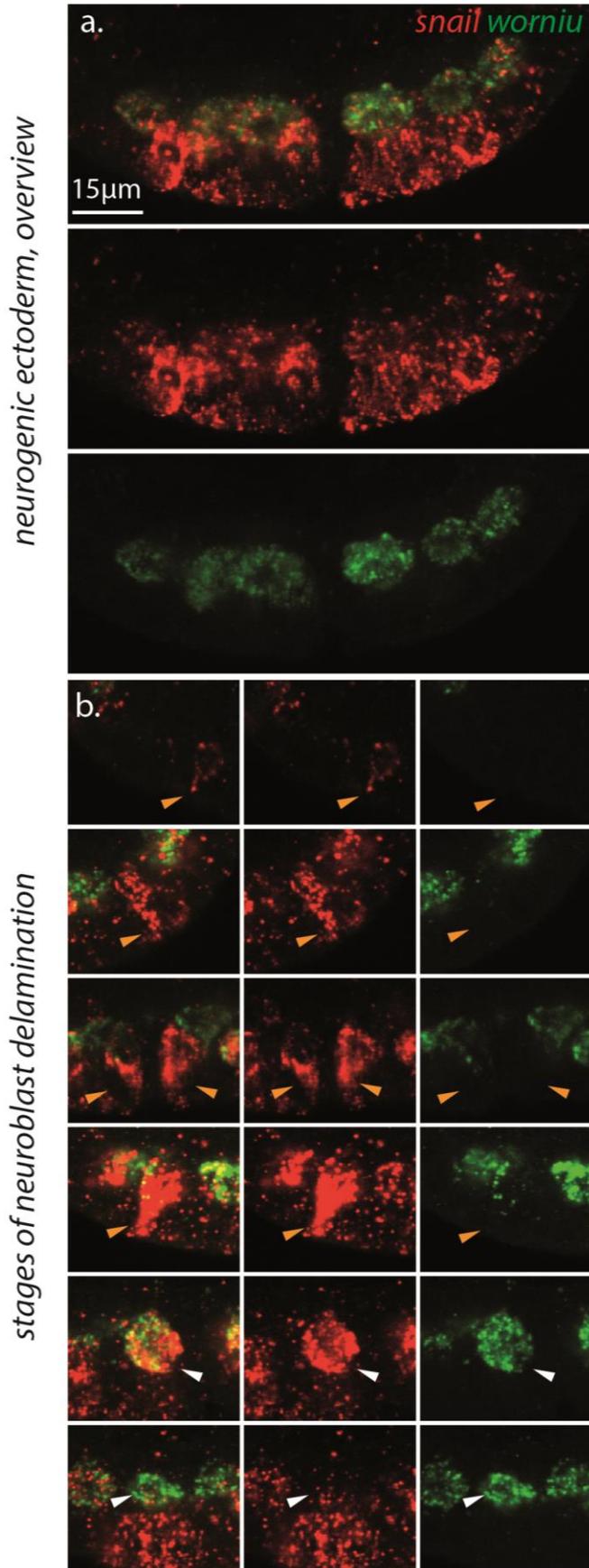
Figure 21. Expression of *snail* and *worniu* during neuroblast delamination. Fluorescent *in situ* hybridizations. *snail* is red; *worniu* is green. Embryonic cross-sections; single confocal planes. Scale bar is 15 µm.

a. Cross-section through the embryonic ventral nerve cord. Neuroectodermal expression of *snail* is clearly seen (stage 9).

b. Selection of neuroblasts at different stages of delamination. Delamination was staged according to the form of neuroblasts. The arrows indicate the cells of interest. Yellow arrows point bottle-shaped neuroblasts during delamination. These cells express *snail* and have only weak traces of *worniu* expression. White arrows point mature fully delaminated neuroblasts, round in shape. These neuroblasts either are *snail-* and *worniu-* positive or express only *worniu*. The images represent multiple neuroblasts from a single fixated embryo at the different stages of delamination.

c. Schematic cross-section of *Drosophila* embryo at stage 9 of development (Campos-Ortega and Hartenstein, 1997). Black rectangle highlights the area analyzed in a. and b.

d. Cell shape changes during delamination of the single neuroblast. First neuroblast cell elongates and cell body moves inwards, into the embryo. The cell acquire specific bottle shape. Later the cell contract thin cell projection and become round in shape (Doe, 1992).

Neuroblasts undergo stereotypic shape changes during their delamination, summarized on the fig. 21d. (modified from Doe, 1992). The neuroblast elongates and acquires the specific bottle shape (middle scheme). Subsequently the cell contracts its thin projection and forms the round shape of the mature neuroblast. Thus, the shape of a neuroblast cell is a reliable index for staging of the neuroblast formation. Additionally, the stage of neuroblast delamination can be precisely identified by the subcellular localization of the Bazooka protein. During delamination, when the neuroblast acquires the bottle shape, Bazooka is localized in the thin cell projection. Later, when formation of the mature neuroblast is finished, Bazooka is localized in a crescent at the apical side of the neuroblast. Therefore I studied expression of *snail* and *worniu*, in combination with anti-Bazooka immunostaining (fig. 22).

At the onset of delamination, when the cell starts elongation (fig. 22a´. and a´´., yellow arrows), the cell expresses *snail*. I could not find any *worniu* positive cell at this stage. The neuroblast cells express *snail* also when the bottle shape is fully formed, and later, when the projection is contracted and the cells become round (fig. 22a´´´. and a´´´´.).

At the "bottle shape" stage of delamination weak traces of *worniu* expression could be detected (fig. 21b.; yellow arrows point to the cell projections of bottle-shaped neuroblasts). *worniu* positive cells are either bottle-shaped (fig. 22b´.) or round, with the fully formed apical Bazooka crescent (fig. 22b´´.-b´´´´.). The majority of the *worniu*-positive cells are fully delaminated (21b., white arrows).
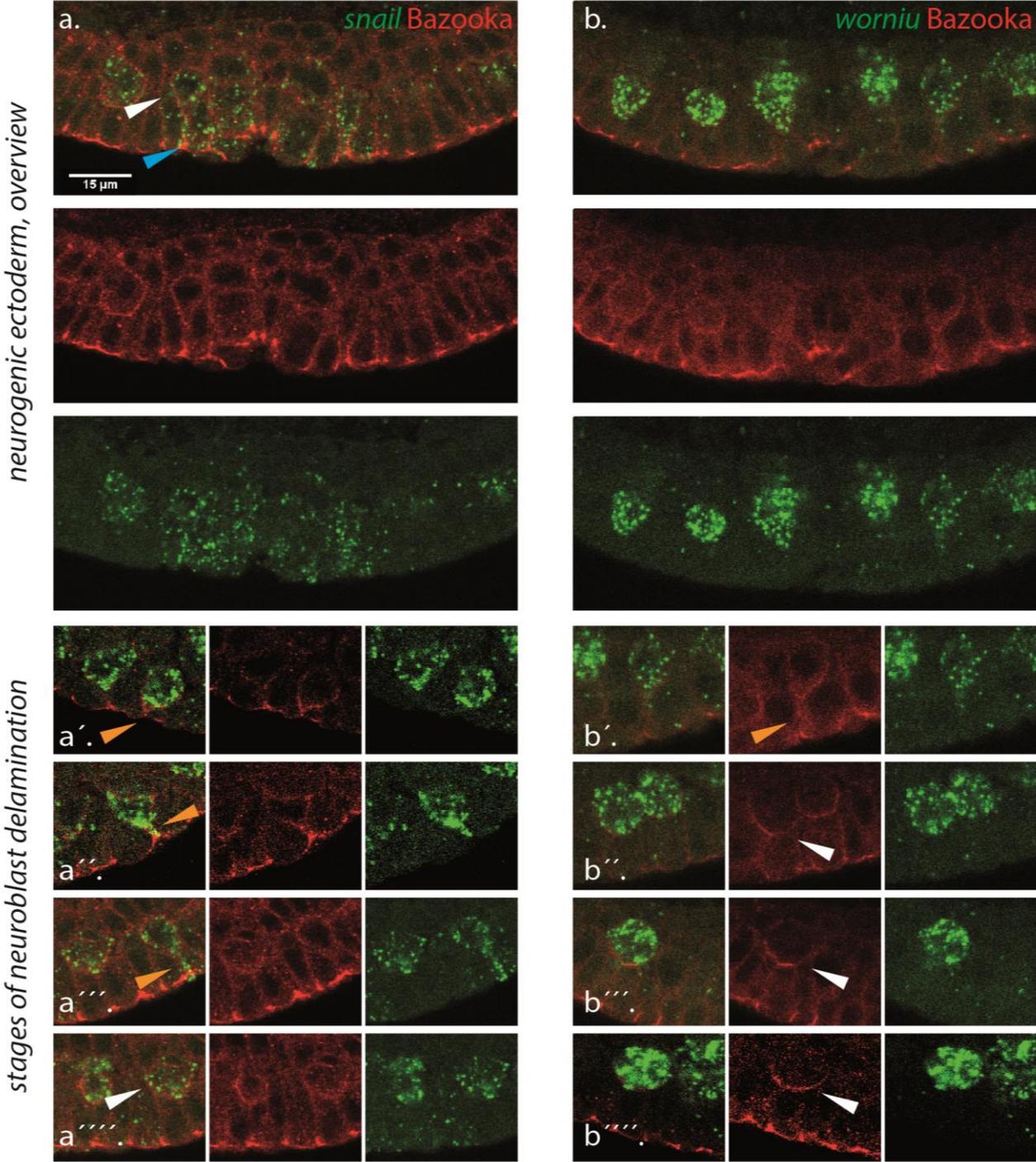
Figure 22. Expression of *snail* and *worniu* during neuroblast delamination (stage 9). Fluorescent *in situ* hybridizations. Bazooka immunostaining is provided as a cell outline and a cell polarity marker. Embryonic cross-sections; single confocal planes. The scale bar is 15 µm.

a. Expression of *snail* is green; Bazooka is red. *snail* can be detected both in neuroblasts (white arrows) and in neurogenic ectoderm (blue arrow).

a′., a′′., a′′′. Expression of *snail* in neuroblasts at different stages of neuroblast delamination. *snail*-positive neuroblasts are still connected to the embryonic surface by a cell projection (yellow arrows). Selection of neuroblasts was picked from a single embryo.

a′′′′. Expression of *snail* in the neuroblasts shortly after delamination (white arrow).

b. Expression of *worniu* is green; Bazooka is red. *worniu* can be detected only in neuroblasts.

b′. Expression of *worniu* in the neuroblasts during delamination. Cell projection has not been contracted yet (white arrow).

b′′.-b′′′′. Expression of *worniu* in the mature neuroblasts. Cell projection has been fully contracted. Bazooka crescent has been formed (white arrows). Each selection of neuroblasts was picked from a one single embryo.

Thus, I conclude that *snail* is expressed in neuroblasts only in a short pulse during delamination. When delamination is finished, expression of *snail* ceases.

4.1.1.3 Comparison of *escargot* expression pattern to expression patterns of *snail* and *worniu*

At developmental stages 6-7 *snail* is expressed in the mesoderm and *escargot* is expressed in one dorsal, one lateral and one ventro-lateral stripe. The ventral stripe of *escargot* expression directly borders the mesodermal *snail*-positive domain. The cells do not co-express *snail* and *escargot*; it has been shown that Snail represses *escargot* in the mesoderm (Yagi and Hayashi, 1997; fig. 23a.). During stage 8 and early stage 9 *snail* and *escargot* are both expressed in transversal stripes in the neurogenic ectoderm. The stripes do not overlap, but alternate with each other. At these stages no *escargot*-positive neuroblasts could be detected (fig. 23b., c.). Intensity projections of the confocal z-stack indicated single embryonic regions as positive for both *snail* and *escargot* (fig. 23c., arrows). However, digital cross-sections reveal that expression is present in the different embryonic tissues, overlaying on the projections (fig. 23c′., arrow). At stage 9 expression patterns of *snail* and *escargot* do not overlap. At later stages, particularly at stage 10, *escargot* is expressed in the midline cells and in two neuroblasts per segment: 1 CNS neuroblast and one neuroblast of the peripheral nervous system (fig. 23d., d′.; arrows highlight *escargot*-positive neuroblasts). In summary, this result confirms mutually exclusive expression of *snail* and *escargot* at two developmental time points: stages 6-7, longitudinal stripes in ventral embryonic region, fig. 23a.; stage 8, transversal stripes in neurogenic ectoderm fig. 23b., c., c′.
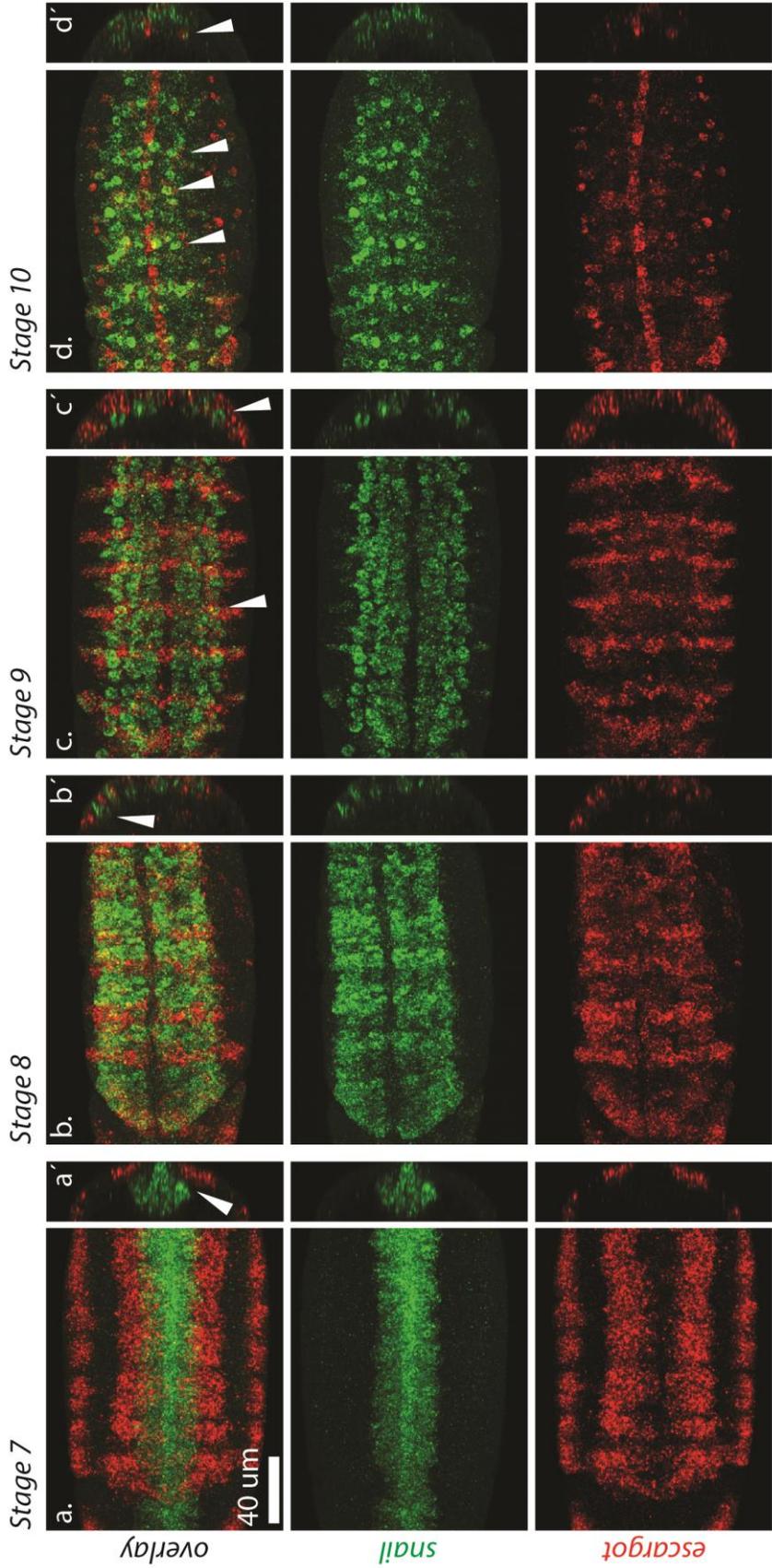
Figure 23. Comparison of *snail* and *escargot* expression during stages 7-10 of embryonic development. Double fluorescent *in situ* hybridization, z-stack confocal image, maximum intensity projection; *snail* is green, *escargot* is red. Embryos are shown from the ventral side. Side panels represent digital cross-sections through the stack of confocal images.

a. Embryo at stage 7. Snail is expressed in the gastrulated mesoderm. Four, out of five, stripes of *escargot* expression are visible on the image. Expression of *escargot* in the medial and lateral columns of the ventral neurogenic ectoderm is seen on either side of the embryo. The neurogenic expression domain of escargot does not overlay with *snail* expression in mesoderm (white arrow).

b. Embryo at stage 8. *snail* and *escargot* are expressed in neurogenic ectoderm in parasegmental stripes.

c. Embryo at stage 9. *snail* is expressed in delaminated neuroblasts. *escargot* is expressed in clear parasegmental stripes in neurogenic ectoderm. Neuroblasts do not co-express *snail* and *escargot* (white arrows).

d. Embryo at stage 10. *snail* is expressed in a subset of neuroblasts. Expression of *escargot* is detected in midline cells and in single neuroblasts. A single neuroblast in a segment co-expresses *snail* and *escargot* (white arrow).

Expression of *worniu* and *escargot* at stages 8 and 9 is also mutually exclusive (fig. 24). Cross-sections of confocal z-stacks clearly show expression of *escargot* mRNA in the neurogenic ectoderm (fig. 24, arrows).

Thus precise analysis of *snail*, *worniu* and *escargot* expression leads me to several conclusions. First, *snail* and *worniu* are expressed during neuroblast development in temporal succession. Second, *escargot* is expressed in the neurogenic ectoderm, but neither in the delaminating neuroblasts, nor in mature neuroblasts at stages 8 and 9. However, expression of *escargot* can be detected in single neuroblast lineages at the later developmental stages (stages 10-11).
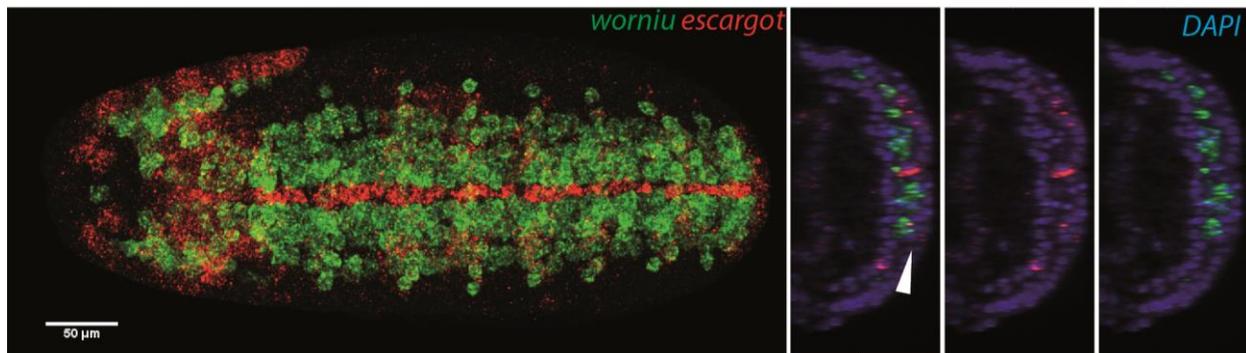
Figure 24. Comparison of *worniu* and *escargot* expression at stage 10 of embryonic development.
z-stack confocal image, maximum intensity projection, *worniu* is green, *escargot* is red, DAPI is blue; the panel to the right shows digital cross-sections of the z-stack confocal image.

No neuroblast (i.e. *worniu*-positive) cells express *escargot*. *escargot* expression is detectable in the neurogenic ectoderm.

4.1.2   Expression of endogenous Snail, Worniu and Escargot;

The embryonic expression of Snail protein has been described (Alberga et al., 1991; Boettiger and Levine, 2013; Ip et al., 1992, 1994; Leptin, 1991; Perry et al., 2010). Generally, expression of Snail protein recapitulates the mRNA pattern, with one notable exception. Snail protein is not present in the *snail*-positive embryonic ectodermal stripes at the stages 8-10 (Ip et al., 1994; fig. 22, 23, 25).

Expression of Worniu protein, detected with anti-Worniu antibodies (Lai and Doe, 2014), showed a neuroblast-specific pattern. There was one striking difference to the expression of the mRNA: neuroblast MP2 is negative for Worniu protein, while it expresses *worniu* mRNA (fig. 26c., d.; arrows). The cell is positive for both Snail protein and mRNA (fig. 26a., b.; arrows). Detection of the expression pattern of the endogenous Escargot protein was impossible, since there were no anti-Escargot antibodies available for the experiments (Abcam ab43635 anti-Escargot antibodies were reported as unspecific).

Expression of Snail and Worniu protein partially recapitulates expression of *snail* and *worniu* mRNA: in each case there is a subset of RNA-positive cells where the protein cannot be detected. Thus, expression of Snail and Worniu is controlled at the post-transcriptional level.
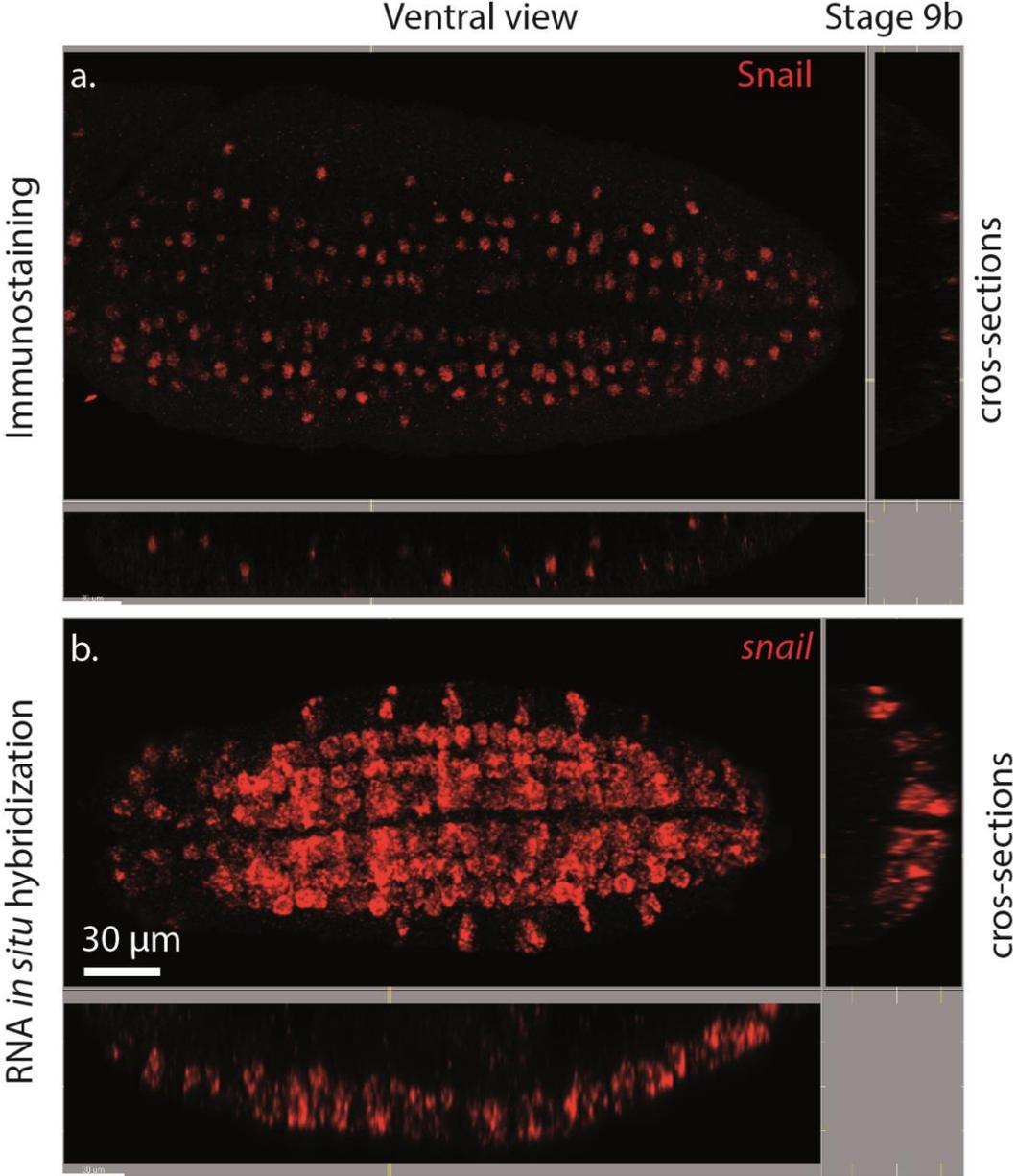
Figure 25. Comparison of transcription of the *snail* gene and translation the Snail protein.

a. anti-Snail fluorescent immunostaining. z-stack confocal image, maximum intensity projection, Snail is red. The side panels show digital xz and yz-sections.

b. anti-*snail* fluorescent *in situ* hybridization. z-stack confocal image, maximum intensity projection, *snail* is red. The side panels show digital xz and yz sections.

Massive transcription of *snail* gene is clearly seen in the neurogenic ectoderm, while no Snail protein can be detected there.
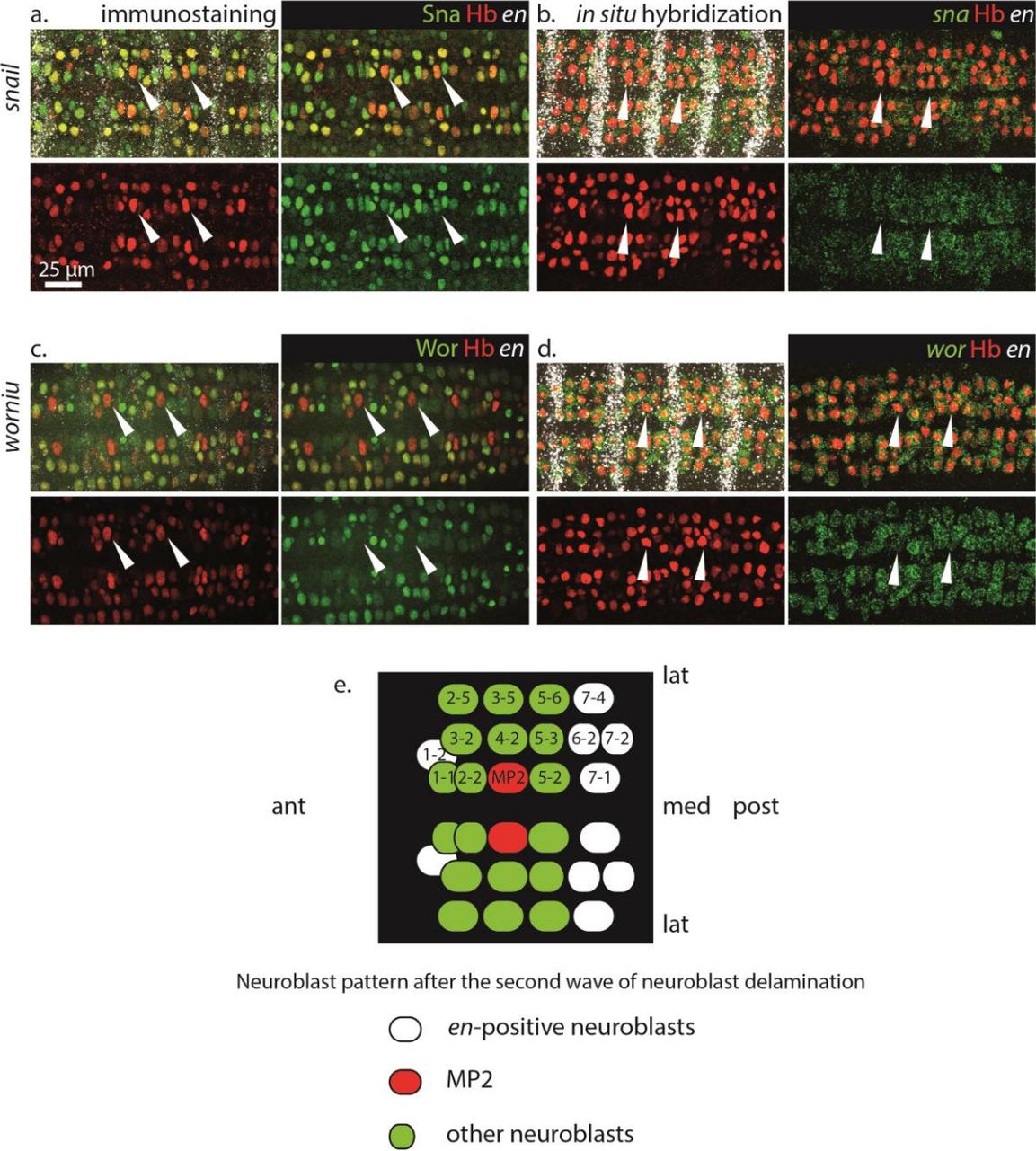
Figure 26. Expression of Snail and Worniu (genes and proteins) in the MP2 neuroblasts.

a-d. Expression of *snail*, *worniu, en* and Hb during the second wave of neuroblast delamination at stage 9. Images represent four embryonic parasegments; ventral view of the midsection of an embryo. Expression of *en* is white. Hb, the neuroblast marker, is red. All images are maximum intensity projections of z-stack confocal images. Scale bar is 25 μm.

a. anti-Snail immunostaining (green). MP2 (white arrows) cells show reduced amounts of Snail.

b. anti-*snail in situ* (green). MP2 (white arrows) cells are *snail*-positive.

c. anti-Worniu immunostaining (green). MP2 (white arrows) cells are Worniu-negative.

d. anti-*worniu in situ* (green). MP2 (white arrows) cells are *worniu*-positive.

e. Schematic map of neuroblasts during the second wave of delamination. MP2 is red; en-positive cells are shown as white circles. Abbreviations: ant: anterior; med: medial; post: posterior; lat: lateral.

4.1.3 Expression of transgenic *V5-HA-snail*, *HA-V5-worniu* and *V5-HA-escargot*

w⁻ and *V5-HA-snail*, transgenic embryos were stained with anti-V5 antibodies. At stage 5 V5-HA-Snail is expressed in mesodermal cells. The expression domain has sharp lateral and posterior borders, at the anterior side expression ceases in a gradient manner (fig. 27a., b.: arrows). The tagged protein localizes in the nucleus (fig. 27a., b.). Neuroblast-specific expression of the V5-HA-Snail recapitulated the expression of endogenous Snail (data not shown). No specific fluorescent signal could be found in *w⁻* embryos stained with anti-V5 (data not shown).
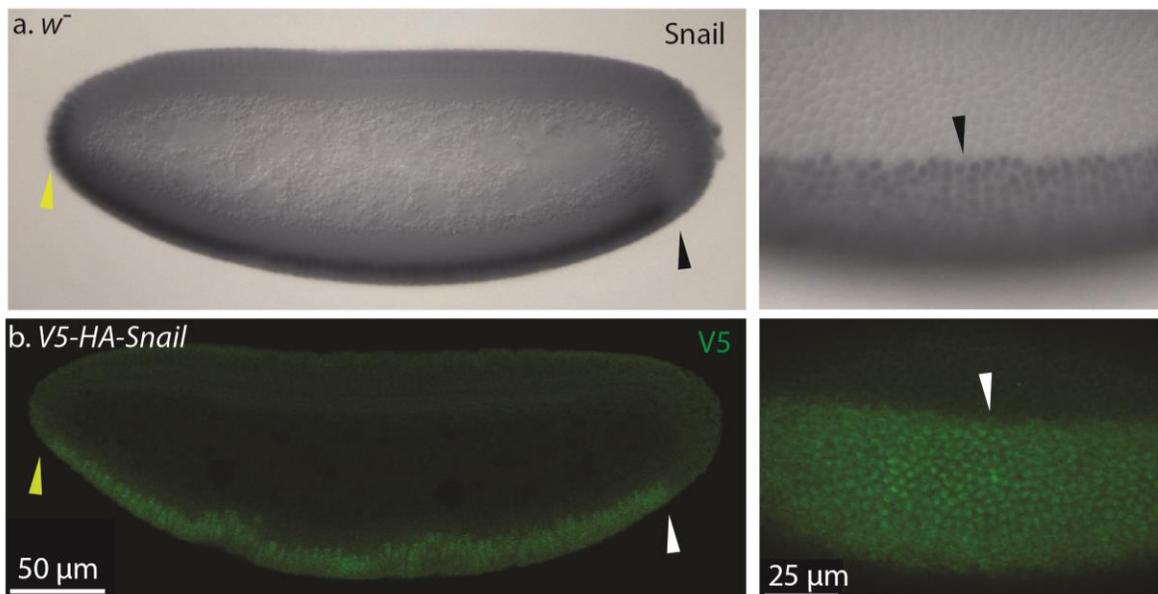
Figure 27. Comparison between expression of the endogenous Snail and the transgenic V5-HA-Snail.
a. Expression of the endogenous Snail in an early wild type embryo (*w⁻*; stage 6). Anti-Snail color immunostaining. Lateral embryonic view and lateral surface of the embryo under high magnification. Snail localizes in nuclei. Expression of Snail has clear posterior and lateral edges (black arrows). At the anterior edge expression of Snail ceases in a gradient manner (yellow arrow).
b. Expression of the V5 epitope in a *V5-HA-snail* transgenic embryo (stage 6). Anti-V5 fluorescent immunostaining. Images represent single confocal planes. Lateral embryonic view and lateral surface of the embryo under high magnification. V5 epitope localizes in nuclei. Similarly to the expression of the endogenous Snail, expression of the V5-tagged protein has sharp posterior and lateral edges (white arrows). At the anterior edge V5 signal ceases in a gradient (yellow arrows).

In *HA-V5-worniu* transgenes the V5 epitope was detected only in nuclei of the cells of ventral nerve cord (fig. 28). In a subset of cells HA-V5-Worniu is found in the cytoplasm. A co-staining with Phospho-Histone H3 indicated that these are dividing cells (data not shown). Overall, the expression pattern of the HA-V5-Worniu corresponds to the

expression pattern of endogenous Worniu protein, identified in the immunostaining, described above (fig. 26). No specific fluorescent signal could be detected in transgenic *V5-HA-escargot* embryos neither anti-V5 nor anti-HA antibodies.
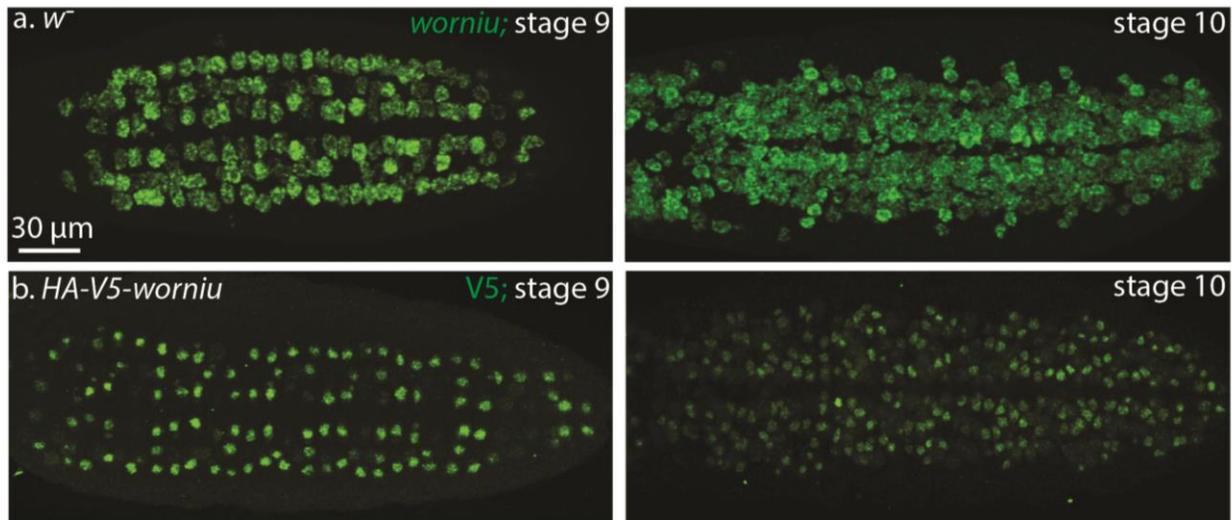
Figure 28. Comparison between expression of the endogenous *worniu* and the transgenic HA-V5-Worniu.
a. Expression of the endogenous *worniu* in early wild type embryos (*w⁻*). Anti-*worniu* fluorescent *in situ* hybridization. Ventral views of embryos at stages 9 and 10. Z-stack confocal image, maximum intensity projection. Worniu is expressed in the neuroblasts; mRNA localizes in cytoplasm.
b. Expression of HA-V5-Worniu in *HA-V5-worniu* transgenic embryos. Anti-V5 fluorescent immunostaining. Ventral views of embryos at stages 9 and 10. Images represent single confocal planes. HA-V5-Worniu shows cell-specific nuclear or cytoplasmic localization. Cytoplasmic localization of HA-V5-Worniu marks dividing cells (data not shown).

4.1.4 Comparison of expression patterns of Snail and HA-V5-Worniu

Since we had limited quantities of anti-Worniu immune serum, and since HA-V5-Worniu accurately recapitulated the expression pattern of endogenous Worniu, I used transgenic *HA-V5-worniu* line for the detailed analysis of the Worniu expression patterns.

I performed a double fluorescent immunostaining with anti-Snail and anti-V5 antibodies on a 2-6 h old collection of *HA-V5-worniu* transgenic embryos. To test whether neuroblasts express Snail and HA-V5-Worniu proteins sequentially, I studied the individual fates of the neuroblasts 6-2 and 7-2. These neuroblasts delaminate during the second wave of delamination as a group of two neighbor cells (fig. 29b., red boxes). At stage 9 the 6-2 and 7-2 cells first express very low levels of Snail (fig. 29c., white box). Later cells become Snail- and Worniu-positive (on the image only 6-2 is double-positive, 7-2 has no Snail protein

anymore fig. 29e.) and finally lose expression of Snail (fig. 29f.). Expression profiles of the other neuroblasts were not traced individually. Overall, the number of Snail-positive neuroblasts decreases over stages 9 and 10, while the number of Worniu-positive neuroblasts increases (fig. 29). Therefore neuroblasts show a tendency to express *snail* to *worniu* sequentially. Sequential expression of the genes can be detected on mRNA as well as on the protein level.

Expression analysis shows distinct expression patterns of *snail*, *worniu* and *escargot* during embryonic neurogenesis. *escargot* is expressed in the neurogenic ectoderm; *snail* is expressed in neuroblast cells during delamination; expression of *worniu* starts during delamination and persists thereafter (fig. 30).
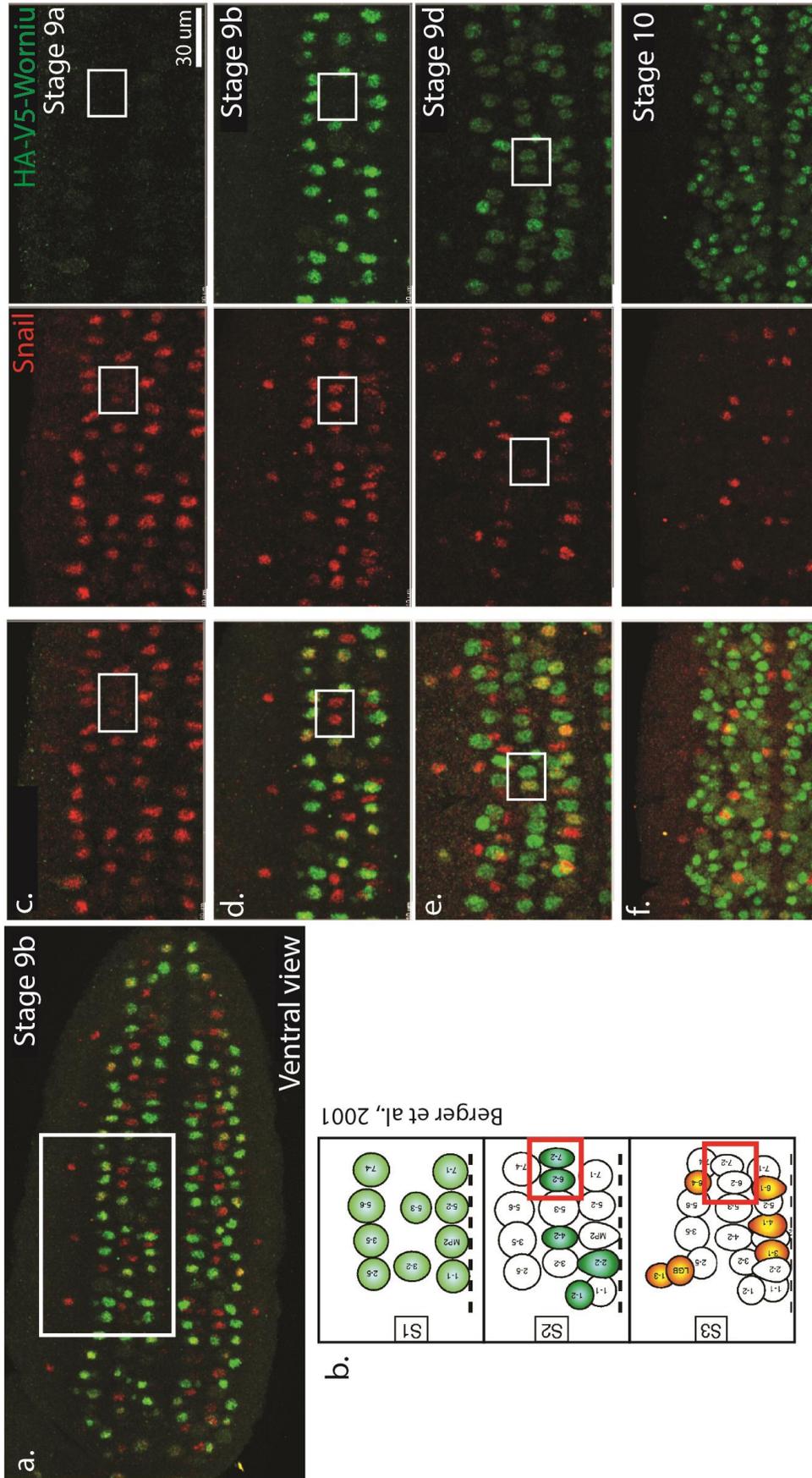
Figure 29. Comparison of Snail and HA-V5-Worniu expression in individual neuroblasts.

a. Expression of Snail and HA-V5-Worniu in a *Drosophila* embryo at stage 9b. Ventral view, double anti-Snail (red) and anti-V5 (green) fluorescent immunostaining. Single confocal plane. White rectangle highlights the area, analyzed in c.-f; ventral view of the midsection of an embryo.

b. The patterns of the first three waves of neuroblast delamination; schematic cartoon (modified from Berger et al., 2001). The pattern is given for a one single parasegment. The pattern repeats in each parasegment of the embryonic body. The red box highlights neuroblasts 6-2 and 7-2, analyzed later on images c.-e. (white box).

c.–f. Expression of Snail and HA-V5-Worniu in the ventral nerve cord during neuroblast delamination, ventral view of the midsection of an embryo the three middle ventral segments (stages 9-10 of embryonic development). The temporal sequence Snail -> Snail + HA-V5-Worniu -> HA-V5-Worniu can be easily observed in middle columns of neuroblasts, in the neuroblasts 6-2 and 7-2, highlighted with the white boxes.

c. Early stage 9. Neuroblasts of the first delamination wave express Snail. Some neuroblasts have weak HA-V5-Worniu signal in the cytoplasm. Neuroblasts 6-2 and 7-2 are not delaminated yet.

d. Middle stage 9. Neuroblasts of the first delamination wave are HA-V5-Worniu-positive; some are HA-V5-Worniu- and Snail-positive. Neuroblasts of the second wave (including the 6-2 and 7-2 cells in the white box) have delaminated. 6-2 and 7-2 are Snail-positive.

e. Late stage 9. The neuroblasts of the second wave have already been formed. 6-2 and 7-2 (white box) are HA-V5-Worniu-positive. 7-2 is Snail-negative, while a weak Snail signal is still present in 6-2.

f. Stage 10. Later waves of delamination. Massive formation of neuroblasts. Large number of Worniu-positive cells and several Snail-positive cells are present in each segment.
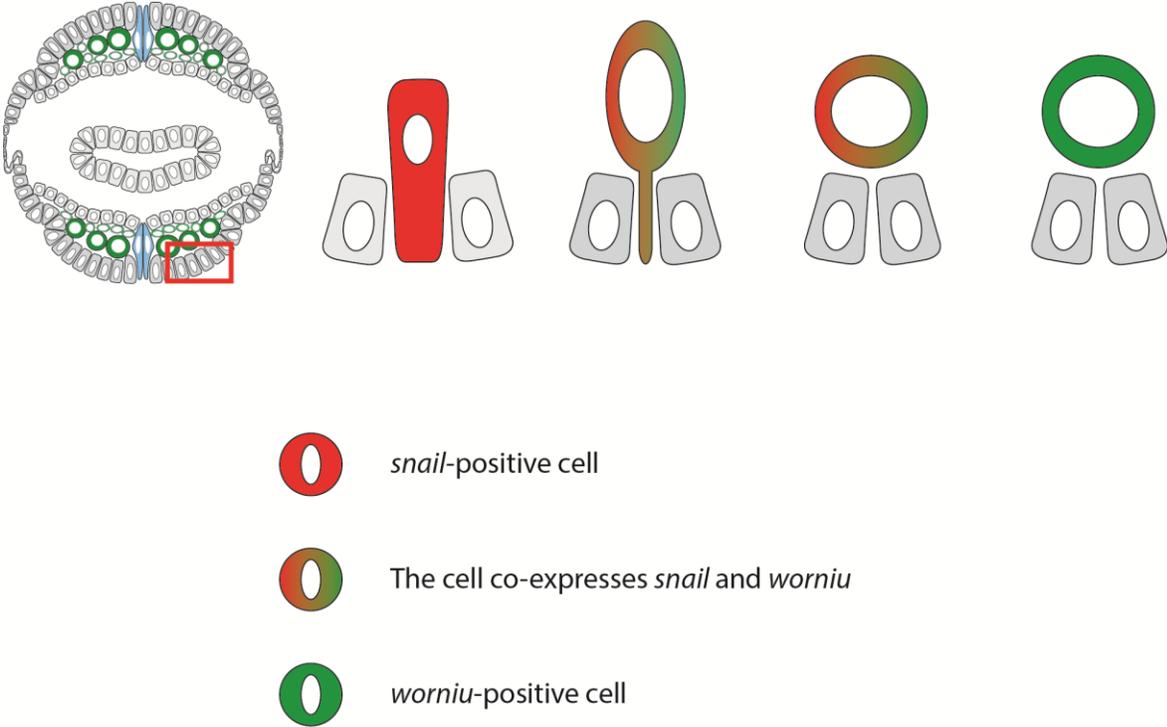
Figure 30. Expression of Snail-like TFs in neuroblast cells during delamination.
Neuroblast starts delamination as a *snail*-positive cell. During delamination the cell starts expression of *worniu*, the cell co-expresses *snail* and *worniu*. After delamination has been finished, expression of *snail* is down-regulated and stopped. From this moment the neuroblast cell is *worniu*-positive.

4.2   Chromatin immunoprecipitation

4.2.1   Detection of V5- and HA-tagged transgenic proteins on western blot

Before starting anti-V5 ChIP on *V5-HA-snail* and *HA-V5-worniu* transgenes I need to validate the transgenic constructs by biochemical characterization of the tagged proteins.

Both anti-V5 and anti-HA antibodies detected specific protein bands on a western blot. As anti-V5 antibodies detected a high amount of unspecific background (data not shown), anti-HA antibodies were selected for all further experiments. In lysates of transgenic embryos 2-6 hours AEL the anti-HA antibody detects a specific band at ~45 kDa for V5-HA-Snail, at ~65 kDa for V5-HA-Worniu and at ~55 kDa for V5-HA-Escargot (fig. 31). These bands correspond to the predicted molecular weights of V5-HA-Snail (43 kDa for endogenous protein, 45.5 kDa recombinant), HA-V5-Worniu (61.9 kDa endogenous protein, 64.4 kDa recombinant) and V5-HA-Escargot (52 kDa endogenous, 55 kDa recombinant). No specific bands were detected in lysates prepared from wild type embryos at the same developmental stage. An unspecific 100 kDa band was present in all embryonic lysates including *w*⁻. As a positive control I used lysate from *Kc167* cells transfected with *pAc5.1-3xHA-snail* (3xHA-Snail is 50 kDa) and *pAc5.1-3xHA-twist* (3xHA-Twist is 80 kDa) (fig. 31): 3xHA-Snail band in the cell lysate corresponds to V5-HA-Snail band in the embryonic lysate.

In summary, V5- and HA- tagged proteins are specifically detected only in embryos carrying the correspondent transgenes. The molecular weights correspond to expected sizes of the tagged Snail, Worniu and Escargot proteins.
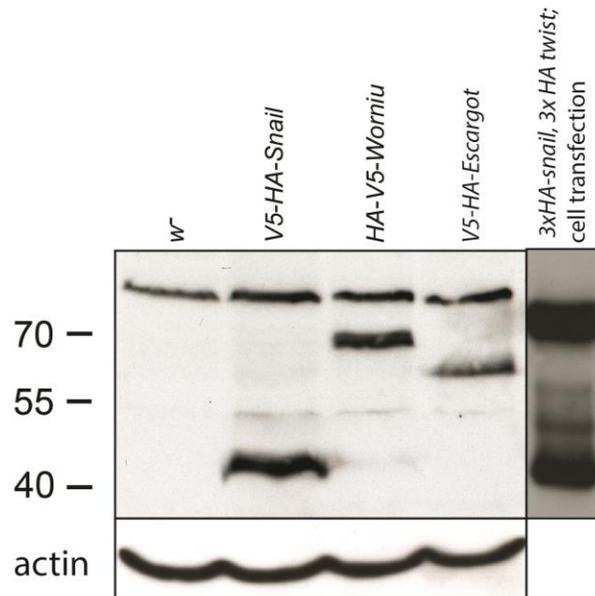
Figure 31. anti-HA western blot detects the tagged proteins in the embryonic lysate of *FlyFosmid* transgenic embryos, not in *w⁻* (2-6 h AEL).

Each lane was loaded with embryonic lysate containing 50 µg of protein. An unpsecific band at 100 kDa is also present in the *w⁻* lysate. A specific band of ~45 kDa is detected in embryonic lysate of *V5-HA-snail* transgenes (predicted MW of wild type Snail: 43 kDa, V5-HA-Snail: 45.5 kDa). *HA-V5-worniu* embryonic lysate contains a ~65 kDa band (predicted MW of wild type Worniu: 61.9 kDa, recombinant HA-V5-Worniu: 64.4 kDa). *V5-HA-escargot* embryonic lysate contains a ~55 kDa band (predicted MW of wild type Escargot: 52 kDa, recombinant V5-HA-Escargot: 55 kDa). Anti-actin was used as loading control. A lysate of *Kc167* cells transfected with *pAc5.1-3xHA-Snail* (50 kDa) + *pAc5.1-3xHA-Twist* (80 kDa) vectors served as positive control for anti-HA antibodies: both 3xHA-Snail and 3xHA-Twist are clearly detectable.

4.2.2 Immunoprecipitation of V5- and HA-tagged transgenic proteins

4.2.2.1 Immunoprecipitation under native conditions

The V5-tagged proteins were designed to be used for chromatin immunoprecipitations and I therefore established conditions for V5 immunoprecipitation experiments. I used anti-V5 antibodies coupled to agarose beads (anti-V5 affinity gel) to precipitate the tagged proteins from native embryonic lysate and subsequently detect it with anti-HA antibodies. Native lysate is a mechanically disrupted embryonic material diluted with PBStr and treated with Roche protease inhibitors; it must not be confused with the nuclear extracts prepared according to the formaldehyde fixation and nuclei purification protocols from ChIP-seq procedure.

As illustrated in fig. 32, all three proteins were successfully precipitated. Specific V5-HA-Snail, HA-V5-Worniu and V5-HA-Escargot bands were detected in the inputs and IP samples, while no specific bands were observed in supernatant.
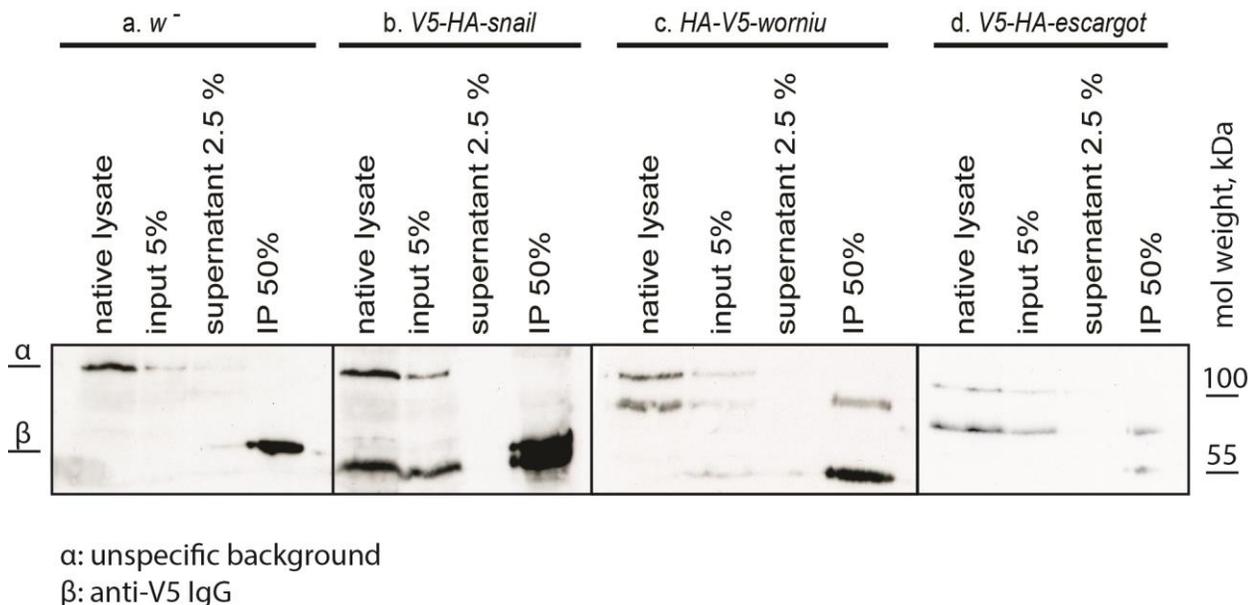
Figure 32. V5-HA-Snail, HA-V5-Worniu and V5-HA-Escargot can be immunoprecipitated with anti-V5 affinity gel. The gel is loaded with the embryonic lysate (100 µg of protein), 5% of input, 2.5% of supernatant and 50% of the IP. Western blot probed with anti HA-antibody.
a. Immunoprecipitation of *w⁻* lysate. Unspecific ~100 kDa band could be detected in the lysate, input supernatant, but not in IP sample. Secondary antibodies cross-react with anti-V5 IgG derived from the affinity gel. IgG band was be detected in IP and supernatant samples.
b. V5-HA-Snail band was detected in the lysate, input and IP, not in the supernatant. V5-HA-Snail band runs very close to IgG band.
c. HA-V5-Worniu band was detected in the lysate, input and IP, not in the supernatant.
d. V5-HA-Escargot band was detected inthe lysate, input and IP, not in the supernatant.

4.2.2.2   Immunoprecipitation of the tagged TFs under ChIP conditions

Next I tested immunoprecipitation under conditions used in the ChIP protocol, i.e. on formaldehyde-crosslinked and subsequently purified and sonicated embryonic chromatin. I performed IP under ChIP conditions and analyzed input, supernatant and precipitant on the western blot. To find an optimal chromatin-to-affinity gel ratio for the anti-V5 ChIP on *V5-HA-snail* transgenic chromatin, I precipitated several different chromatin amounts (ranging from 0.05 g to 0.5 g embryonic collection) with an equal amount of the affinity gel. V5-HA-Snail could not be detected in any supernatant, while the band of precipitated protein was getting enriched when the increased amounts of input was taken (fig. 33). Thus 25 µl of

anti-V5 affinity gel is enough to precipitate the entire pool of V5-HA-Snail from 0.5 g of transgenic embryos.

In summary, the size of the V5- and HA-tagged proteins expressed in the transgenic embryos correspond to the predicted size of the tagged Snail, Worniu and Escargot. The expression patterns of the tagged proteins correspond to those of the endogenous proteins. The proteins can be IP´ed with anti-V5 affinity gel under mild native and experimental ChIP conditions. Therefore, I can proceed to the ChIP experiments and immunoprecipitate V5-HA-Snail-bound DNA fragments from sonicated chromatin.
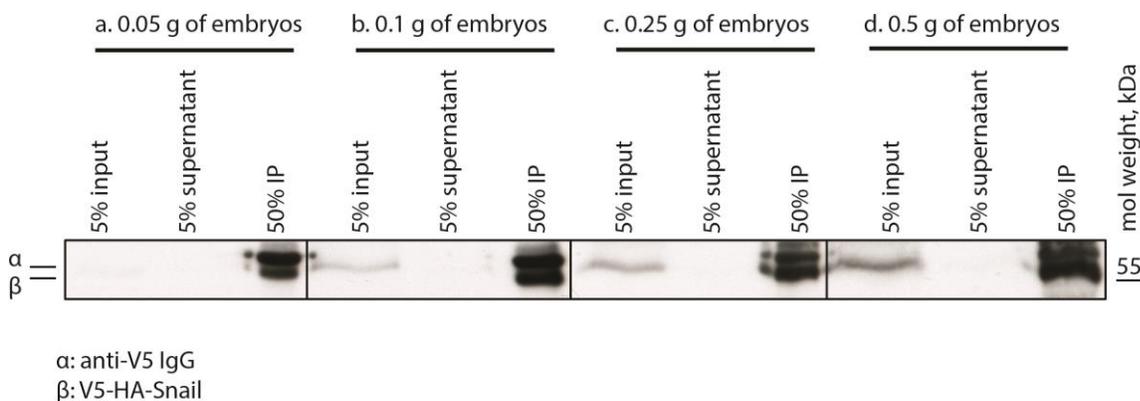
Figure 33. V5-HA-Snail can be IP'ed with anti-V5 affinity gel under ChIP condition.
IP was performed with the lysate of *V5-HA-snail* transgenic embryos, 2-6 hours AEL. The embryos were crosslinked, lysed and sonicated as in the ChIP protocol; western blot probed with anti-HA antibodies.
Each gel is loaded with 5% of input (sonicated nuclei), 5% of supernatant and 50% of final IP (beads diluted in TE and boiled with Laemmli buffer). 25 µl of affinity gel were used in each immunoprecipitation.
a. Immunoprecipitation of V5-HA-Snail from 0.05 g of embryos. No band seen in the input. Clear V5-HA-Snail band in IP sample, not merged with IgG band.
b. Immunoprecipitation of V5-HA-Snail from 0.1 g of embryos.
c. Immunoprecipitation of V5-HA-Snail from 0.25 g of embryos.
d. Immunoprecipitation of V5-HA-Snail from 0.5 g of embryos. V5-HA-Snail band merges with IgG band. Supernatant does not contain a specific V5-HA-Snail band.

4.2.3   Collection of large amounts of tightly staged embryos

One of the major challenges of a ChIP experiment is the collection of sufficient amount of biological material. In the case of *Drosophila* embryos, the protocol, used as a ChIP guideline (Sandmann et al., 2006b), suggests to start with approx. 1.5 g of embryos. To compare stage-specific binding of Snail to DNA, I need tightly staged embryonic collections, with minimal contamination of embryos of other developmental stages.

Snail is detected in the mesoderm first at stage 4; where it persists until stage 7, when it appears in neuroblasts (fig. 17). Ideally, the mesoderm-specific collection should contain embryos at the stages 4, 5 and 6. The first Snail-positive neuroblasts delaminate during late stage 8. Snail is expressed in neuroblasts shortly before being replaced by Worniu (fig. 18, 20). At the end of stage 11 neuroblast delamination is finished, and the number of Snail-positive neuroblasts decreases. Thus, I would like to limit the neuroblast-specific collection to stages 8-10.

The speed of embryonic development depends on multiple factors like temperature or air moisture; therefore precise embryo staging is needed for a given experimental environment. Table 3 summarizes embryo staging made for the environment used in the work described here.

Population cages were loaded with ~40-50 g of flies. Apple juice agar plates with a small drop of yeast paste in the middle were deposited into the cage for egg collection. Overnight plates were supplied with a larger amount of yeast to ensure good nutritional status of the flies. Every morning two one-hour-long pre-lay sets of apple juice plates were provided to fly females to activate embryo production and to remove over-staged embryos. The flies were allowed to lay eggs for two hours. The embryos from the two hour collection were divided into 8 portions. The first portion of embryos was fixed after 2 hours of incubation at room temperature (22°C), to make a sample of 2 h – 4 h old embryos. After 30 minutes the second portion was fixed to produce 2 h 30 min – 4 h 30 min collection. The next portions were fixed every 30 minutes, until the time point 5 hours 30 minutes (5 h 30 min – 7 h 30 min) was reached. The stage distribution of embryos was counted for every time point collection. Collection of embryos 2.5-4.5 hours after egg laying (AEL) at 22°C was selected as a mesoderm-specific (65% of embryos are at developmental stages 4-5). Neuroblast specific collection: 4.5 – 6.5 h AEL at 22°C (73.5% of embryos are at developmental stages 8-10).

Table 3. Embryonic stage distribution in 2 hour collections of embryos after 2, 2.5, 3, 3.5, 4, 4.5, 5 and 5.5 hours of aging at 22°C. Percent of embryos at the correspondent stage in a selected sample and the amount of embryos counted are provided. Time collection, aged for 2h 30 min (2 h 30 min– 4 h 30 min old embryos) was selected as "mesoderm" time point (87% of embryos are at stages 5-8). Collections of 4h 30 min – 6 h 30 min old embryos were selected as the "neuroblast-specific" (92% of embryos are at the stages 9-11).

stage h AEL	4	5	6	7	8	9	10	11	>11	total
2-4	38	43				6	3.6		8.4	83
2.5-4.5	9.3	56		2.3	1	7	4.6		20	86
3-5	4.6	58	4.6	4.6	3	6	3	3	12	65
3.5-5.5	2.7	35.6	15	5.4	9.5	5.4	4	2.7	19	73
4-6		2	15	5.7	38.4	15.3	5.7		17	52
4.5-6.5				1	21.5	51	1		25.4	102
5-7						66.6	2.6	6	24	114
5.5-7.5						58	4	2	35.5	90

4.2.4 Establishment of the ChIP protocol

ChIP-seq is a multistep experiments. Establishment of a ChIP seq in a new laboratory requires adaptation of a general ChIP seq workflow to the particular conditions. If the protocol is not optimized the DNA will be lost during the procedure. The main steps of ChIP seq protocol and the most sensible stages of the workflow are summarized at the fig. 34.

4.2.4.1 Optimisation of chromatin sonication

To perform ChIP, the genomic DNA should be gently sheared into fragments using ultrasound. The average fragment size recommended for ChIP-seq is 200-250 bp (Bonn et al., 2012).

embryos

fixation

cell and nuclei isolation

DNA is degraded:
conditions of sonication

sonication

◄━━━━━━━━━

DNA left unbound:
chromatin to antibody ratio

binding of the TF by antibody

◄━━━━━━━━━

precipitation of the immunocomplexes
with protein A coupled beads

specifically bound DNA washed out
together with unspecific:
was conditions

beads wash

◄━━━━━━━━━

R.C. is not sufficient:
DNA is lost during PhChlor purification

reverse crosslinking

◄━━━━━━━━━

DNA purification

DNA is lost during PhChlor purification

◄━━━━━━━━━

qPCR assay

library preparation and sequencing

Figure 34. The main steps of the western blot workflow procedure. Left panel of the scheme summarizes the main steps of the ChIP protocol. Right panel indicates the moments of the protocol where DNA could be lost during the experiment.

To determine the size distribution of DNA fragments produced by different sonication conditions and to investigate the influence of chromatin abundance on the sonication process I purified the nuclei from 1 g of mesoderm-staged *w⁻ Drosophila* embryos. The material was unequally distributed into two tubes, so that the first tube received four times more nuclei than second tube. I performed 30 cycles of chromatin sonication, taking aliquots of sheared chromatin after every 5 cycles. The DNA was purified as suggested in (Sandmann et al., 2006b), but replacing phenol-chloroform purification with a column-based purification (QiaQuick PCR purification kit). The DNA size distribution in samples after increasing numbers of sonication was determined by gel electrophoresis

(fig. 35). After 5 cycles of sonication the sheared DNA contains a large proportion of high molecular weight DNA (600-1000 bp), while the average size rapidly decreases during the next 10 cycles. After 15 sonication cycles the size of the majority of chromatin fragments ranges between 100 and 400 bp. The difference in DNA size distribution of chromatin samples produced by longer sonication (15, 20, 25, 30 cycles) is minor. To avoid over-warming of the sample and degradation of the chromatin I decided not to extend sonication longer than 15 minutes. The amount of chromatin in a sonication tube does not influence the size distribution of shared DNA fragments.
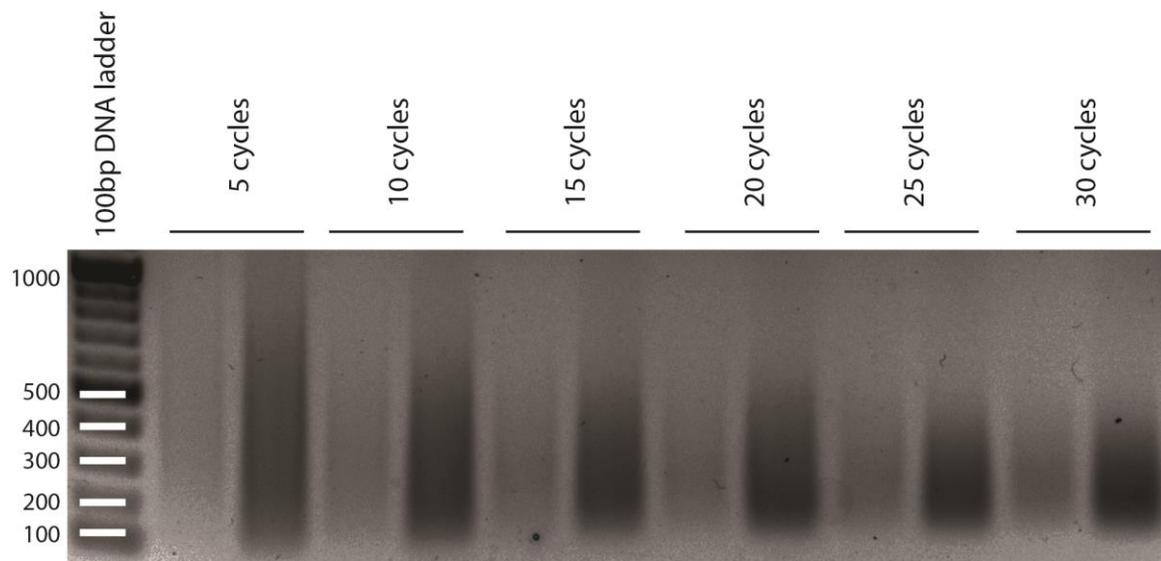
Figure 35. Size distribution of DNA fragments in chromatin samples after different number of sonication cycles. One cycle of sonication consists of 30 sec of sonication (0.5 sec pulse and 0.5 sec rest), followed by 30 sec of rest. To test how abundance of chromatin in a tube influences size distribution, two different amounts of chromatin were sonicated in parallel. Each first lane of sonication sample represents sonication of chromatin extracted from 0.2 g of embryos; second lane represents sonication of chromatin from 0.8 g of embryos.

4.2.4.2   Positive control test for established ChIP procedure

Anti-Mef2 antibodies, kindly provided by Eileen Furlong, have been successfully used in ChIP experiments on 2-4 hours old embryos (Sandmann et al., 2006a). I performed anti-Mef2 ChIP as described in (Sandmann et al., 2006a). To analyze the enrichment of the Mef2-bound DNA regions in the final ChIP samples I amplified the Mef2-bound actin57B enhancer (primers 28. and 29.) and a non-bound *oskar* region (primers 34. and 35.).

In the ChIP samples I found a 1.5-2-fold enrichment of the Mef2-bound region over non-bound control region in the final ChIP samples. No enrichment could be detected in the mock ChIP samples, for which pre-immune serum was used (Table 4).

Table 4. qPCR analysis of the anti-Mef2 ChIP samples. Sample name indicates the antibody used for ChIP. qPCR reactions were performed in triplicate. Bound region: Mef2-bound actin57B enhancer; non-bound region: *oskar.*

Sample	target	Cq mean	Cq Std Dev	Sq mean	Sq Std Dev	Enrichment
anti-Mef2	bound	28.55	0.426	6.56936	1.74944	1.6
anti-Mef2	non-bound	29.00	0.285	4.01042	0.79475	
pre-immune serum	bound	29.38	0.739	3.95696	1.67113	0.7
pre-immune serum	non-bound	28.71	0.383	4.95787	1.24163	

4.2.4.3   Comparison and selection of Snail-bound regions for qPCR ChIP

Having established that the ChIP procedure *per se* works in my hands, I moved on to the anti-Snail ChIP experiments. In the first step I used the guinea-pig anti-Snail antibody provided by Eileen Furlong, (EMBL, Heidelberg), which has been successfully used ChIP (Rembold et al., 2014; Zeitlinger et al., 2007). I followed the protocol for ChIP-on-chip developed and published by the laboratory of Eileen Furlong (Sandmann et al., 2006b) introducing some minor modifications, described in the Materials and Methods section.

Mesoderm-staged embryos were lysed to prepare embryonic chromatin. After reverse-crosslinking and DNA purification the chromatin concentration was calculated and the material was distributed into aliquots containing 40 µg DNA each. 8 µl of the guinea pig anti-Snail immune serum (diluted in glycerol, 1:1; Zeitlinger et al., 2007) and 8 µl of a guinea pig pre-immune serum (diluted in glycerol, 1:1) were used for the Snail ChIP and mock ChIP samples, respectively. The antibody-to-chromatin ratio had been titrated by Robert Zinzen before (Zeitlinger et al., 2007; Zinzen et al., 2006). DNA was diluted in 60 µl of TE, 2 µl of DNA solution were taken for the subsequent qPCR test. Each qPCR reaction was performed in triplicate.

To evaluate enrichment of Snail-bound DNA regions in ChIP samples I performed qPCR. The *oskar* gene was selected as a Snail non-bound DNA region (Rembold et al., 2014; Sandmann et al., 2006b). The primer pair is identical to the one that was used as a non-bound test region in the anti-Mef2 ChIP. Multiple primer pairs that amplify ~100 bp long products from the regions located in *sim*, *rho* and *sog* loci (primers 30.-45.; all three genes

are well-established Snail target genes: Rembold et al., 2014; Zeitlinger et al., 2007) were tested.

qPCR analysis of the target region enrichment over background is summarized in the table 5. Since qPCR reaction was performed on three individual reaction plates, three individual negative controls (*osk* primers) are shown in the table. The *sim* primers (44. and 45.) detected a 2-fold enrichment of this region over the control region in the anti-Snail ChIP sample. The most efficient enrichment (5-fold) enrichment was detected in the qPCR reaction performed with the *rho* primers (num. 30., 31.). This primer pair was selected as the positive primer pair for all subsequent anti-Snail ChIP experiments.

Analysis of the target enrichment in the mock sample (prepared using pre-immune serum) was performed only for *sim* primers (num. 36., 37.) and for *rho* primenrs (num. 30., 31.).

Table 5. qPCR analysis of the proposed Snail-bound target regions. 3 primer pairs that amplify fragments of *rho* region, two primer pairs that align to *sim* one primer pair that align to *sog* region were tested in the experiment. Enrichment value was calculated by comparison to the Sq of non-bound control *oskar* region. Since experiment was performed on four qPCR amplification plates independently, the independent values of control amplification of non-bound *osk* region are given for each plate. Reactions are performed in triplicate. Since *rho* 3 and negative *osk* reactions on the second plate were contaminated, I repeated the reaction on one independent plate and added PCR of mock control ChIP sample to the experiment.

Reaction plate	target	Cq mean	Cq Std Dev	Sq mean	Sq Std Dev	Enrichment
Anti-Snail ChIP Plate 1	36., 37. *sim* 1	27.88	0.240	15.20378	2.53885	1.64739198
//	38., 39. *rho* 1	26.22	0.159	42.45937	4.51110	4.600646947
//	44., 45. *sim* 2	27.52	0.359	19.33960	4.78880	2.09552433
//	40., 41 *rho* 2	26.14	0.097	9.29219	1.17462	1.006846969
//	34., 35. Negative *osk*	28.92	0.422	9.22900	2.45679	
Anti-Snail ChIP Plate 2	42., 43. *sog*	26.92	0.117	37.35621	3.20155	4.13428967
//	30., 31. *rho* 3	25.81	0.000	50.98764	0.00000 *	5.642908088
//	34., 35 Negative *osk*	27.82	0.000	9.03570	0.00000 *	
Anti-Snail ChIP, Plate 3	30., 31. *rho 3*	24.41	0.085	62.32083	4.17879	9.349904221
	44., 45. *sim* 2	26.20	0.149	15.83153	1.60286	2.37518153
	123 osk	27.25	0.317	6.66540	1.43158	
Mock ChIP Plate 4	30., 31. *rho 3*	28.13	0.000	3.25763	0.00098	0.894608233
	44., 45. *sim* 2	28.39	0.167	3.56987	0.40348	0.980355579
	34., 35 Negative *osk*	28.12	0.146	3.64140	0.36166	

• two of three reaction mixes were contaminated, therefore only one Cq and one Sq value were taken for the analysis. Therefore the primers *rho* 3 were tested again on the plate 3.

4.2.4.4 Anti-V5 ChIP on the lysate of *V5-HA-snail* transgenic embryos

Since western blot analysis proved efficient precipitation of V5-HA-Snail from sheared chromatin, and since ChIP procedure was established *per se*, I decided to evaluate efficiency of V5-ChIP on *V5-HA-snail* transgenic embryos. Several experimental conditions were tested to adjust the protocol (Sandmann et al., 2006b) to the current experiment (the modified protocol is given in the Materials and Methods).

For the final experiment I collected 0.3 g of *V5-HA-snail* transgenic embryos and performed ChIP using 20 μl of anti-V5 affinity gel. qPCR analysis was able to detect a 2.2 - fold enrichment of the positive DNA target (*rho* region, primers 30. and 31.) over the negative target (*osk* region, primers 34. and 35.). No enrichment was detected in the mock samples (anti-V5 affinity gel incubated with w^- embryonic lysate). Table 6 summarizes the results of the qPCR.

Table 6. qPCR analysis of anti-V5 ChIP of the chromatin extracted from transgenic *V5-HA-snail* and w^- embryos. Sample name indicates the genotype of analyzed embryos. Cq and Sq values are explained in the text. qPCR reactions were performed in triplicates, thus "mean" and "std dev" values characterize repeatability of the result in one triplicate. Bound region: *rho,* non-bound region: *oskar.*

Sample	target	Cq mean	Cq Std Dev	Sq mean	Sq Std Dev	Enrichment
V5-HA-snail	Bound	27.67	0.32	18.82762	3.63282	2.29
V5-HA-snail	Non-bound	25.27	0.79	8.25238	2.90190	
w^-	Bound	29.79	0.43	7.91099	2.75178	0.9
w^-	Non-bound	28.71	0.17	8.59441	0.97623	

4.2.4.5 Identification of the antibody-to-chromatin ratio for the rabbit anti-Snail immune serum

Rabbit anti-Snail antibody (produced by Fiona Ukken, the laboratory of Maria Leptin) was raised against an N-terminal fragment of Snail (amino acids 1-200) fused to GST. Rabbit antibody has never been validated for use in chromatin immunoprecipitations experiments. The optimal antibody-to-chromatin ratio is a key to successful immuneprecipitation. An excessive amount of immune serum may enhance the unspecific precipitation of chromatin, while an insufficient amount may not suffice to bind all TF-DNA complexes, which will subsequently be discarded with in a supernatant.

I performed immunoprecipitations of an equal amounts of chromatin by a titration series of the rabbit anti-Snail immune serum from 3 to 0.0015 μl, using *V5-HA-snail* embryos (2.5 – 4.5 h AEL), as the V5 and HA tags allow the convenient detection of the precipitated protein on a Westen Blot. I prepared nuclear extract from 0.6 g of mesoderm-staged embryos. The chromatin was sonicated and distributed into 6 equal aliquots. The immunoprecipitations were performed by the applying 3; 1.5; 0.5; 0.15; 0.015 and 0.0015 μl of antibody dilutions to the samples. 25 μl of protein A agarose beads were used to precipitate antigen-antibody complexes. ChIP samples and supernatant solution were analysed on western blot by anti-HA antibodies. 0.5 μl of anti-Snail immune serum were

sufficient to precipitate all V5-HA-Snail protein from 0.1 g of mesoderm-staged embryos (fig. 36). A higher amount of antibody does not precipitate any further protein. Instead unbound antibody can be detected in the supernatant at concentrations higher than 0.5 µl of immune serum. Since the lysate of V5-HA-Snail embryos contains the double amount of Snail (endogenous Snail and transgenic V5-HA-Snail), I conclude that 0.25 µl of the immune serum is needed to precipitate 0.1 g of mesoderm-staged wild type embryos.
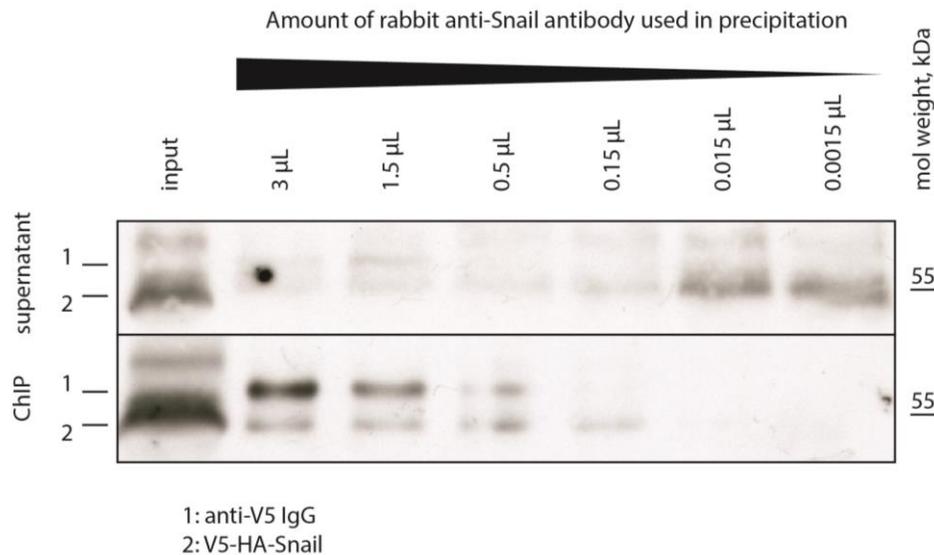
Figure 36. Titration of antibody-to-chromatin ratio to optimize ChIP procedure. Titration experiment is performed with rabbit anti-Snail immune serum and *w⁻* embryos.
Chromatin, extracted from 0.6 g of embryos was distributed into 6 aliquots and immunoprecipitated with the range of antibody volumes (3 µl, 1.5 µl, 0.5µl, 0.15 µl, 0.015 µl, 0.0015µl). The gel at the top shows western blot analysis of ChIP supernatant, i.e. the material which was not precipitated during the experiment. The lower panel shows the western blot analysis of the material precipitated on the beads. V5-HA-Snail in cross-linked input samples can be detected as two ~43 and ~55 kDa bands. ~50 kDa IgG band is seen exactly in between. Implementation of 0.0015 – 0.5 µl of antibody yields progressively brighter bands of V5-HA-Snail in ChIP sample. Increased amount of antibody from 0.5 to 3 µl does not enhance protein yield, while IgG bands becomes progressively brighter.
ChIP with less than 0.5 µl reveals high amount of V5-HA-Snail in supernatant.
In the experiments where chromatin was precipitated with 1.5 or 3 µl IgG band is detectable in supernatant.

4.2.4.6   Comparison of the available of anti-Snail antibodies

To assess available anti-Snail immunosera (rabbit anti-Snail and guinea pig anti-Snail), I used the antibodies to perform parallel ChIPs and test the enrichment of Snail-bound *rho* region (primer num. 30., 31.) versus non-bound *oskar* region (primer num. 32., 33.) in the final ChIP samples. The rabbit anti-Snail immune serum gave 7-fold enrichment; a 4-5-fold enrichment was detected in the ChIP samples, prepared with guinea pig anti-Snail antibodies. No enrichment of positive Snail targets was detected in mock ChIPs,

produced with pre-immune guinea pig or rabbit sera. The results of qPCR analysis are shown in the table 7.

Table 7. qPCR analysis of the ChIP samples prepared with different anti-Snail antisera. Sample name indicate the antibody, used to perform the ChIP. qPCR reactions were performed in duplicate, thus the "mean" and "std dev" values characterize repeatability of the result in one duplicate. Bound region: enhancer of *rho*, non-bound region: *oskar.*

Sample	target	Cq mean	Cq Std Dev	Sq mean	Sq Std Dev	Enrichment
Guinea pig	Bound *rho*	25.68	0.120	22.77985	2.16417	4.9
	non-bound *osk*	27.79	0.380	4.61239	1.18321	
Rabbit	Bound *rho*	24.75	0.137	47.43907	5.13561	7.5
	non-bound *osk*	27.33	0.096	6.25564	0.40957	
GP pre-immune serum	Bound *rho*	28.20	0.139	3.09697	0.34013	0.79
	non-bound *osk*	28.01	0.057	3.90943	0.15162	
Rabbit pre-immune serum	Bound *rho*	27.84	0.279	4.14723	0.90713	0.8
	non-bound *osk*	27.67	0.004	4.92877	0.01251	

4.2.4.7   DNA losses during ChIP-seq

I have shown that I can precipitate Snail protein from sheared chromatin samples by western blotting. qPCR experiments indicated the enrichment of known Snail-bound regions (e.g., up to 7-fold enrichment of *rho*) in the ChIP samples. A problem that hindered the subsequent library preparation and sequencing was a consistently low DNA yield in the final ChIP samples.

In order to solve this problem, I analyzed the ChIP protocol to figure out those steps of the experiment where the loss of DNA occurred. The ChIP workflow is schematically shown in fig. 34. In short, DNA could be lost during the following steps of the protocol:

- DNA-protein complex can be disrupted during chromatin preparation or sonication.
- The de-crosslinking of DNA-protein complex is insufficient and DNA, still bound to the proteins, is lost during phenol-chloroform purification.

- DNA is lost during one of the purification steps: phenol-chloroform purification, sodium acetate (NaAc) DNA precipitation or subsequent wash with Ethanol (EtOH).

4.2.4.8   Comparison of agarose and sepharose beads as a matrix for protein A

The TF-antibody complex can be precipitated with protein A coupled to either agarose or sepharose beads. No preference is specified in (Sandmann et al., 2006b). To detect the difference between agarose and sepharose beads I performed ChIP in parallel with both kinds of Protein A beads. To compare binding capacities of Protein A slurries I tested two amounts of antibody within each experiment.

A collection of mesoderm-staged *OregonR* embryos (0.8 g) was used to prepare four chromatin samples and test four precipitation conditions:

1. 50 µl of agarose protein A beads, 2 µl of rabbit anti-Snail antibody;

2. 50 µl of agarose protein A beads, 10 µl of antibody;

3. 50 µl of sepharose protein A beads, 2 µl of antibody;

4. 50 µ of sepharose protein A beads, 10 µl of antibody.

To monitor and calculate the DNA losses throughout the ChIP procedure I retained 1% of the experimental volume at each step of the experiment. I measured DNA concentrations throughout the experimental procedure, including all waste materials. Table 8 summarizes the results of DNA concentration measurements.

Table 8. DNA loss during the ChIP experiment. DNA concentrations in the retained 1% probes, as well as recalculation for the total DNA amount in the experimental samples at the every step of the ChIP protocol. DNA concentration in the washing solutions was measured only for the Agarose A 2 µl sample. DNA concentrations were determined with Qubit dsDNA HS Assay Kit from Life Technologies.

	Experimental step	Amount of solution analyzed	Measured DNA concentration in purified sample ng/µl	DNA amount in 1% aliquot (ng)	DNA in 100% sample
1	Lysed embryos	1%	18.6	930 ng	93 µg
2	First pellet Vitelline membranes, etc.	1%	0.596	29.8	2.9 µg
3	Supernatant after 2nd centrifugation (harvesting of the cells from lysed embryos)	1%	5.82	291	29 µg
4	Lysed cells	1%	10.3	515	51 µg
5	Nuclei supernatant	1%	0.304	15.2	1.52 µg
6	Nuclei before sonication	1%	8.63	431.5	43.15 µg
7	Nuclei after sonication	1%	2.54	127	12.7 µg
8	Nuclei after Sonication after centrifugation	1%	6.18	309	30.9 µg
At this step the lysate was divided into 4 samples					
9	Agarose 2 Pre-clearing beads	All beads	0.957	47.85	0.04785 µg
10	Agarose 10 Pre-clearing beads	All beads	1.2	60	0.06 µg
11	Sepharose 2 Pre-clearing beads	All beads	4.48	224	0.224 µg
12	Sepharose 10 Pre-clearing beads	All beads	4.45	222.5	0.2225 µg
13	Agarose 2 1% of iput	1%	1.24	62	6.2 µg
14	Agarose 10 1% input	1%	0.565	28	2.8 µg
15	Sepharose 2 1% input	1%	1.34	67	6.7 µg
16	Sepharose 10 1% input	1%	1.2	60	6 µg
17	Agarose 2 supernatant	0.5%	0.44	22	4.4 µg
18	Agarose 10 supernatant	0.5%	0.445	22.25	4.45 µg
19	Sepharose 2 supernatant	0.5%	0.711	35.55	7.11 µg
20	Sepharose 10 supernatant	0.5%	0.133	6.5	1.3 µg
21	RIPA supernatant	1%	Too low		
22	RIPA500 1 supernatant wash 1	1%	<<		
23	RIPA500 supernatant wash 2	1%	<<		

24	RIPA500 supernatant wash 3	1%	<<		
25	RIPA500 supernatant wash 4	1%	<<		
26	LiCl supernatant	1%	<<		
27	TE 1 supernatant	1%	<<		
28	TE 2 supernatant	1%	<<		
29	Agarose 2 ChIP	1%	<<		
30	Agarose 10 ChIP	1%	<<		
31	Sepharose 2 ChIP	1%	<<		
32	Sepharose 10 ChIP	1%	<<		

The experiment shows high DNA losses during entire procedure. Already upon harvesting of the cells after the embryo lysis (step 3) I lose ~30% of DNA in the discarded supernatant. Overall, 70% of the DNA present in the starting material is lost during chromatin preparation: 90 µg of DNA in crude embryonic lysate vs. 30 µg sheared DNA (input). Before addition of the antibodies, the ChIP sample must be pre-cleared with beads, to remove all DNA that can bind to the beads unspecifically. I tested the amount of DNA unspecifically bound to the pre-clearing beads after pre-clearing incubation. Unspecific binding of DNA to the beads during the pre-clearing procedure appeared to be high; sepharose beads unspecifically bind ~5 times more DNA than agarose.

4.2.4.9 De-crosslinking of DNA-protein complex

At the beginning of the ChIP protocol DNA is cross-linked to DNA-bound proteins. Therefore proteins must be digested and de-crosslinked before final DNA purification starts. Otherwise the cross-linked DNA will be discarded together with the bound proteins. The protocol (Sandmann et al., 2006b) suggests overnight incubation of the sample supplied with Proteinase K at 37°C, followed by de-crosslinking at 65°C for 6 hours. The modifications of the protocol, allow one hour of incubation with proteinase K and overnight decrosslinking at 65°C. Comparison of the two protocol variations did not show any impact on qPCR-detected enrichment in the final DNA yield (data not shown). Since the total

experimental procedure becomes one day shorter, if the second variation is applied, I adopted the less time-consuming version of the protocol for subsequent ChIP experiments.

4.2.4.10 DNA precipitation with NaAc

After phenol-chloroform purification, the DNA is precipitated with NaAc, washed with 70% EtOH and re-dissolved in water. NaAc precipitation is a sensitive technique, especially if it works with small amounts of DNA. Therefore I tested which amount of DNA could be lost at this step of the protocol and how protocol variations can improve final DNA yield. According to the (Sandmann et al., 2006b) protocol DNA low binding tubes are used throughout the entire ChIP procedure (without any precise specification for the NaAc precipitation step). NaAc precipitation of DNA at -80℃ is one hour long. I tested NaAc precipitation in the standard Eppendorf tubes and extension of the NaAc precipitation overnight. I treated four equal amounts of sonicated DNA chromatin (0.1 g of mesoderm-staged *w*⁻ embryos were used to prepare each aliquot) with RNAse A and proteinase K. The samples were

- transferred to the normal *or* low binding Eppendorf tubes;
- incubated at -80℃ 1 h *or* overnight.

The results of the experiment are summarized in table 9.

Table 9. DNA yield using different NaAc precipitation conditions from equal amounts of sonicated chromatin. DNA concentrations were determined using Qubit dsDNA HS Assay Kit from Life Technologies.

	Low binding tube	Standard tube
1h @ -80˚C incubation	2.68 ng/µl	1.36 ng/µl
o/n @ -80˚C incubation	3.75 ng/µl	3.45 ng/µl

Overnight incubation indeed yielded more DNA, while the switch from low binding tubes to standard tubes led to the loss of DNA. There was therefore no need to change to the standard tubes, but overnight incubation appeared preferable.

4.2.4.11 Phenol-Chloroform and EtOH/NaAc versus column-based DNA purification

I compared phenol-chloroform based purification procedure and column-based QiaQuick PCR purification kit to find the optimal DNA purification method. I took 0.5 g broadly staged w- embryos (~2-7 h AEL), isolated chromatin and sonicated it in 1 ml total

sonication volume. The lysate was distributed into twelve aliquots with the following chromatin content:

    1. 10 µl of lysate   +   90 µl of TE   in quadruplicate

    2. 50 µl of lysate   +   50 µl of TE   in quadruplicate

    3. 100 µl of lysate   in quadruplicate

After DNA de-crosslinking 2 samples out of each quadruplicate were purified according to the phenol-chloroform purification protocol, while the other two samples were purified using QiaQuick PCR purification kit. DNA concentrations in the final DNA solutions are summarized in the table 10.

Table 10. Comparison of phenol-chloroform DNA purification and QiaQuick PCR purification kit. Total DNA yield in the samples purified by phenol-chloroform purification protocol or QiaQuick PCR purification kit (ng). DNA concentrations were determined using Qubit dsDNA HS Assay Kit from Life Technologies.

Chromatin input	Phenol-cloroform purification (ng)	QiaQuick (ng)
10 µl	318	365.5
	299	366
50 µl	143.5	2075
	1700	2110
100 µl	2375	2650
	1200	3000

In all cases the QiaQuick PCR purification kit yielded higher amounts of DNA. Moreover, unlike phenol-chloroform treated samples, the yield of QiaQuick PCR purification kit was consistent and was increasing along with increasing amount of input. However, this result needs to be verified on ChIP samples. Therefore I used the QiaQuick PCR purification kit to purify DNA from a ChIP sample, to eventually compare DNA yield, as well as the enrichment of Snail-bound DNA regions.

1 g aliquot of mesoderm-staged *OregonR* embryos was lysed and distributed into 4 equal chromatin samples. 1% of input chromatin was retained. Input DNA was purified using the method selected for each sample, DNA yield after purification was calculated. Immunoprecipitation was done using 5 µl of rabbit anti-Snail immune serum and 50 µl of Protein A Agarose beads. DNA yield was calculated (table 11).

Table 11. DNA purification by QiaQuick PCR purification kit compared and by phenol-chloroform purification in anti-Snail ChIP. Four ChIP samples were prepared in parallel using equal amounts of chromatin as an input. The Input and the ChIP DNA were purified using phenol-chloroform purification protocol (2 samples) or QiaQuick PCR purification kit (2 samples). DNA concentrations were determined using Qubit dsDNA HS Assay Kit from Life Technologies.

Sample	Amount of DNA in the retained 1% of input; ng	Amount of DNA in ChIP sample; ng
QiaQuick I	276.5	4.35
QiaQuick II	311.5	3.5
PhChlor I	226	2.8
PhChlor II	53.5	3

Purification of 1% input aliquots with QiaQiuck purification kit detected that each experimental sample contained ~30 µg of dry chromatin DNA. Thus, calculating a proportion I can roughly estimate that one gram of mesoderm-staged embryos yields 120 µg of dry chromatin DNA. Previously I identified that 0.25 µl of rabbit anti-Snail immune serum is sufficient to precipitate Snail protein from 0.1 g of mesoderm-staged wild type embryos. Thus 1 g of mesoderm-staged embryos can produce 120 µg of dry chromatin DNA, which can be precipitated with 2.5 µl of rabbit anti-Snail immune serum.

To compare DNA enrichment in the ChIP samples purified using different purification protocols I distributed 1 g of mesoderm-staged embryos into four equal samples and performed full ChIP experiment. DNA from the two samples was purified according to the phenol-chloroform protocol. DNA of the two other samples was purified with QiaQuick PCR purification kit. I analyzed enrichment of the *rho* region (primers num. 30., 31.) over the non-bound *osk* region (primer num. 34., 35.). Cq values and fold enrichment of the positive DNA region over the non-bound negative region are shown in the table 12. The Sq values of the samples purified with QiaQuick PCR purification kit were consistent between the samples. The samples purified with phenol-chloroform purification showed almost 3-fold difference in Sq values.

Table 12. qPCR test of ChIP samples prepared by different DNA purification methods: phenol-chloroform and QiaQuick PCR purification kit. Abundances of two genomic regions were analyzed: *rho* and *osk*. Sq values represent relative starting quantities recalculated for each sample. qPCR reactions were performed in triplicate. Standard deviations show reproducibility of the reaction value within the triplicate.

Sample	*rho* Sq	*rho* Sq st dev	*osk* Sq	*osk* Sq st dev
QiaQuick I	103.6	7.7	9.2	0.14
QiaQuick II	114.5	17.15	6.3	0.65
PhChlor I	79.3	0.2	6.6	0.8
PhChlor II	237.1	52.4	20.1	2.95

In summary, the replacement of the phenol-chloroform purification step with QiaQuick PCR purification kit enhanced the yield of DNA and produced more consistent enrichment. Therefore, in all subsequent ChIP experiments phenol-chloroform purification step was substituted by DNA purification with QiaQuick PCR purification kit.

4.2.5   Anti-Snail ChIP

After successful modification and establishment of the ChIP procedure I proceeded to the final ChIP experiment.

4.2.5.1   DNA sample preparation and sequencing

~2.5 g of mesoderm-staged and ~4.5 g neuroblast-staged *OregonR* wild type embryos were used to prepare 6 mesoderm-staged chromatin samples (50 μg of chromatin DNA in each sample) and 6 neuroblast-staged samples (90 μg of chromatin DNA in each sample). 5 μl of the guinea pig anti-Snail immune serum and 1.5 μl of the rabbit anti-Snail immune serum were applied in each immunoprecipitation. The antibody complexes were precipitated with 50 μl of Protein A agarose. Immunoprecipitation with the guinea pig immune serum was done in duplicate. Immunoprecipitation with the rabbit immune serum was performed in quadruplicate. ChIP samples prepared with rabbit immune serum were pooled pairwise.

Collections of neuroblast-staged V5-HA-Snail and HA-V5-Worniu transgenic embryos were used to prepare the chromatin samples in duplicate. Each *V5-HA-snail* sample contained 80 μg of chromatin; Each *HA-V5-worniu* sample contained 30 μg of chromatin. All samples were ChIP´ed with 25 μl of anti-V5 affinity beads. Table 13 summarizes the amount of DNA in each of the ChIP samples. Quantitative PCR analysis showed a ~4-fold enrichemnt of the *rho* locus over *osk* locus in the mesoderm sample.

Library preparation and Solexa sequencing using the HiSeq system from Illumina was performed at the CSF NGS Unit at the Vienna Biocenter, Austria. Sequencing analysis revealed between 13 and 32 Mio reads in a sample.

Table 13. Description of the ChIP samples, prepared for sequence analysis. "M" stands for a mesoderm sample.,"Nb" for a neuroblast sample. "GP" indicates the samples prepared with the guinea pig anti-Snail immune serum. "R" indicates the samples prepared with rabbit anti-Snail immune serum. DNA concentration in the ChIP samples was measured with Qubit dsDNA HS Assay Kit. "<<" indicates that the DNA concentration is below the detection limit of the assay (10 pg/µl, according to the kit manual). *rho* is a Snail-bound genomic region, *osk* is not bound by Snail.

Sample	DNA amount in input	DNA yield	Cq mean (*rho/osk*)	Sq mean (*rho/osk*)
MGP i	50 µg	<<		
MGP ii	50 µg	<<		
NbGP i	90 µg	<<		
NbGP ii	90 µg	<<		
MR i	50 µg	<<	21.4/25.3	53.8/3.8
MR ii	50 µg	<<	21.3/25.3	60.7/3.8
MR iii	50 µg	<<	21.8/25.15	43.9/4.2
MR iv	50 µg	<<	21.4/25.9	54.4/4.09
NbR i	90 µg	<<		
NbR i	90 µg	2.85 ng		
NbR iii	90 µg	<<		
NbR iv	90 µg	4.3 ng		
V5-HA-Snail i	80 µg	<<		
V5-HA-Snail ii	80 µg	<<		
HA-V5-Worniu i	30 µg	<<		
HA-V5-Worniu ii	30 µg	<<		

4.2.5.2 Read mapping and peak calling

After next-generation sequencing the reads were computationally mapped to the *Drosophila* genome using Bowtie (Langmead et al., 2009). The mapping and the following bioinformatic analysis of the ChIP data was performed by Christoph Dieterich, Max Planck Institute for Biology of Ageing, Cologne, Germany.

65 - 85% of the reads could be mapped to the genome. All reads that mapped to multiple locations in the genome were discarded, to avoid false enrichment of reads at repetitive elements. Regions of high ChIP enrichment were detected with the MACS2 peak calling algorithm with a false-discovery rate (FDR) cutoff of 1%. (Feng et al., 2011, 2012; Zhang et al., 2008). The peaks were called on individual replicates with the respective input sample as control. Only peaks that were present in both technical replicates were kept and analyzed further.

Only a small number of significant peaks or no peaks at all were called in the ChIP samples prepared with the guinea pig anti-Snail antibody and the samples prepared with anti-V5 antibodies. These samples were therefore excluded from the analysis. The read datasets generated from the samples prepared with the rabbit anti-Snail antibodies were used for the following bioinformatic analysis.

We obtained 1703 Snail-bound regions in the mesoderm and 2501 regions in neuroblasts. The comparison of the two datasets revealed that a considerable fraction of Snail binding is tissue-specific. Out of 1703 Snail-binding sites in the mesoderm, 837 (49.1%) were specific for this stage. In neuroblasts 1633 (65%) out of 2501 Snail-bound regions were occupied specifically in neuroblasts. 866 mesodermal regions and 868 neuroblasts regions were bound at both developmental stages. (In the mesoderm 8 peaks overlapped with two neuroblast peaks simultaneously, in neuroblasts 6 peaks overlapped with two mesodermal peaks simultaneously).

To evaluate the level of confidence, the bound regions were analysed for the presence of a Snail binding motif. The alternative Snail motif with the core CAGGTA, which was identified in Snail-bound regions in the mesoderm, was used (Rembold et al., 2014).

496 (59%) of the 837 mesoderm-specific bound regions and 385 (44%) of the 866 continuously bound regions contained at least one Snail binding motif.

The fraction of Snail-bound regions with the alternative Snail motif was slightly lower in the neuroblast data set. 576 (35%) out of 1633 neuroblast-specific regions and 392 (45%) out of 868 shared regions contained a Snail binding motif.

To evaluate the quality of the tissue-specific peaks I inspected the binding of Snail to the genomic region of the *snail* gene (fig. 37). Strong mesoderm-specific binding of Snail is observed at the primary (-1.6 kb) and shadow (-10 kb) mesodermal *snail* enhancers (Hong et al., 2008). In neuroblasts, Snail binds to the 0.8 kb PNS-specific promoter and to a region within the 6 kb *snail* promoter just upstream of the CNS-specific regulatory element (Ip et al., 1994). The 6 kb promoter element includes mesodermal, PNS as well as CNS regulatory elements.
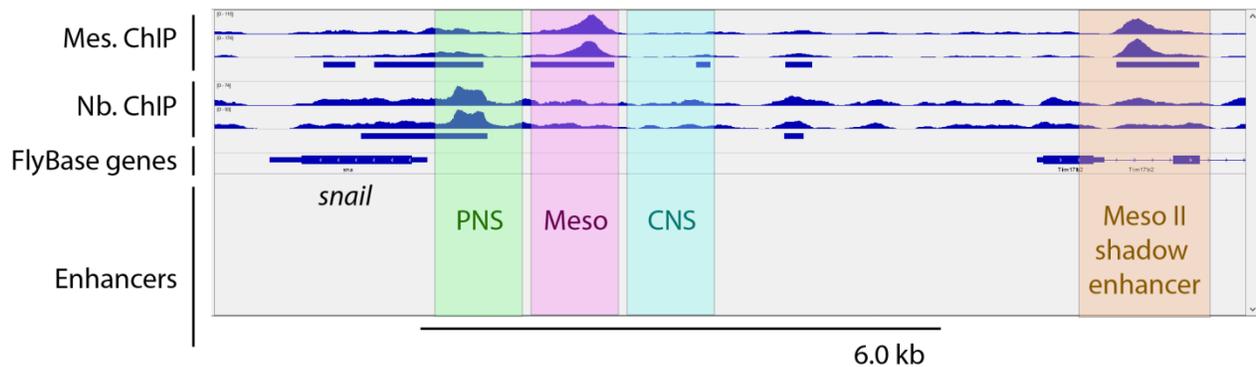
Figure 37. Snail tissue-specifically binds enhancers of *snail*.
Binding sites of the TF Snail in the genomic region of the *snail* gene. Screen shot of the genome browser session.
Mes. ChIP: read pileup coverage (blue peaks) and the regions bound by Snail (blue rectangles) in the mesodermal ChIP samples.
Nb. ChIP: read pileup coverage (blue peaks) and the regions bound by Snail (blue rectangles) in the neuroblast ChIP samples.
FlyBase Genes: the genes annotated in the FlyBase.
Enhancers: approximate location of the *cis*-regulatory modules that control expression of *snail* (Ip et al., 1994; Perry et al., 2010). PNS: enhancer responsible for the gene expression in the peripheral nerve system; CNS: enhancer responsible for the gene expression in central nerve system; Meso: Primery mesoderm-specific enhancer; Meso II: Mesoderm specific shadow enhancer.
6.0: 6 kb promoter fragment which was reported to fully recapitulate wild type expression of *snail*.

4.2.5.3   Tissue-specific transcriptional network of Snail

The peaks were assigned to the closest transcription start sites to generate a list of tissue-specific and shared Snail-bound genes.

To assess the quality of the data I inspected whether it contains known Snail target genes. A subset of Snail-bound genes in mesoderm and in neuroblasts was taken from (Rembold et al., 2014) and (Southall and Brand, 2009), respectively (table 14).

Table 14. Binding of Snail to known target genes.
A subset of genes bound by Snail in mesoderm was taken from (Rembold et al., 2014); a selection of neuroblast targets of Snail was taken from (Southall and Brand, 2009). Binding is indicated '+'.

Gene	Bound by Snail in mesoderm	Bound by Snail in neuroblasts
Mesodermal targets of Snail		
rho	+	
Mef2	+	
wntD	+	
tin	+	
sim	+	+
sog	+	
htl		+
vnd	+	+
Neuroblast targets of Snail		
pros		+
CycE	+	+
lola	+	+
neur		+
sim	+	+
baz	+	
grh	+	+
stg		+
Dl		+
insc	+	+
Snail-like family		
snail	+	+
worniu		+
escargot	+	

The mesodermal target genes contain well-characterized activated and repressed targets of Snail such as *rho*, *Mef2*, *wntD* or *tinman*, demonstrating the quality of the dataset. *heartless*, a known mesodermal target of Snail, is bound by Snail only in neuroblasts, in contrast to earlier results (Rembold et al., 2014). The neuroblast dataset recapitulates binding reported using DamID (Southall and Brand, 2009). *prospero* and *string* are bound specifically in neuroblasts. The expression of the cell cycle regulator *string* (*cdc25*) in neuroblasts is strongly reduced in *snail/worniu/escargot* triple mutants (Ashraf 2001). *bazooka*, which was described as a target of Snail in neuroblasts, is bound only in mesoderm. *CycE* and *sim* are bound in neuroblasts as well as in mesoderm, in agreement with a described or proposed role of Snail in their regulation in both tissues (Kosman et al., 1991; Rembold et al., 2014; Southall and Brand, 2009).

The *snail* gene itself is bound by Snail both in the mesoderm and in neuroblasts but the binding occurs tissue-specifically at different enhancers as described above (fig. 38). This indicates that Snail controls its own expression in both tissues. Binding of Snail to the

escargot and *worniu* genes is tissue-specific as well: Snail binds to a region upstream of the TSS of *escargot* in mesoderm, while a region in the first exon of *worniu* is bound exclusively in neuroblasts. *escargot* is expressed in the mesoderm in the absence of *snail* (Yagi and Hayashi, 1997) and the binding data shows that Snail directly represses *escargot* in the mesoderm. If the expression of *worniu* is changed in *snail* mutant embryos has to be investigated further.
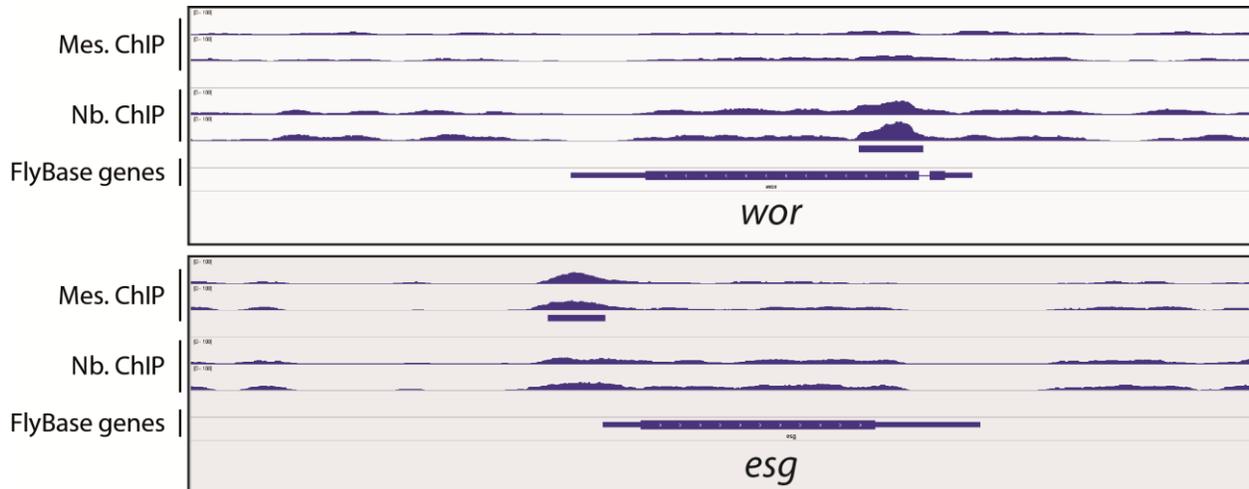
Figure 38. Snail tissue-specifically binds *worniu* and *escargot*.
Binding sites of the TF Snail in the genomic region of the *worniu* and *escargot* gene. Screen shot of the genome browser session.
Mes. ChIP: read pileup coverage (blue peaks) and the regions bound by Snail (blue rectangles) in the mesodermal ChIP samples.
Nb. ChIP: read pileup coverage (blue peaks) and the regions bound by Snail (blue rectangles) in the neuroblast ChIP samples.
FlyBase Genes: the genes annotated in the FlyBase.

5 Discussion

5.1 Expression analysis of *snail*, *worniu* and *escargot*.

5.1.1 Expression of *snail*

snail is expressed in two spatio-temporally distinct domains, regulated by the separate enhancer elements (Ip et al., 1994; Perry et al., 2010): in the mesoderm and the neuroblasts. While the mesodermal expression of *snail* has been studied and annotated several times (Alberga et al., 1991; Ip et al., 1992; Leptin, 1991; Leptin and Grunewald, 1990) the available annotation of the neuroblast expression of *snail* remained incomplete. *snail* was described as a pan-neuroblast neuroblast-specific TF (Ashraf and Ip, 2001; Ashraf et al., 1999; Ip et al., 1994). A detailed, time-resolved analysis confirmed this observation but revealed an unexpected temporal expression profile: *snail* mRNA is detectable in all neuroblasts but only during the short time window of cell delamination. The expression of *snail* starts while a neuroblast cell is still a part of the neurogenic ectoderm, i.e. shortly before delamination, and persists throughout the delamination process. After delamination has finished, the expression of *snail* ceases. This time restriction for expression of *snail* has never been described before.

It is difficult to estimate the duration of *snail*'s expression in a single cell. The first wave of neuroblast delamination occurs during late stage 8, while the neuroblasts of the second wave delaminate at middle stage 9. The third wave of delamination occurs during stage 10. Stage 8 lasts for 30 min (3:10-3:40 hpf); stage 9 lasts for approximately 40 minutes (3:40-4:20 hpf); stage 10 lasts for about one hour (4:20-5:20 hpf; Campos-Ortega and Hartenstein, 1997). Thus, duration of one delamination wave is approximately 30 min. Expression of *snail* starts before delamination, and is finished after cell has been fully delaminated, therefore duration of expression of *snail* can be estimated as ~40-45 min.

It is currently unknown how *snail* expression is controlled. Proneural genes (*achaete, scute* and *daughterless*) are not sufficient to control *snail* expression in the CNS (Ip et al., 1994). Neither the mechanisms of expression ceasing are clear.

5.1.2 Expression of *worniu*

worniu was described as a pan-neuroblast neuroblast-specific marker as well (Ashraf and Ip, 2001; Ashraf et al., 1999; Cai et al., 2001; Hemavathy et al., 2004; Lai and Doe, 2014). Unlike other neuroblast markers (e.g. *hunchback*), which are expressed in neuroblasts and in GMC as well, *worniu* is expressed exclusively in neuroblasts. For the first time I describe a precise moment, when expression of *worniu* starts. Expression of *worniu* starts during delamination. In contrast, expression of *snail* or *hb* starts before delamination, when a neuroblast cell is still a part of neurogenic ectoderm (Doe, 1992). Thus, the neuroblast fate determination and the start of *worniu* expression are separated by a time window. This time window is characterized by extensive rearrangement in cell shape and cell physiology.

The activators of expression of *worniu* are currently unknown. The timing of expression leads to the conclusion that *worniu* might be activated by neuroblast maturation factors, activated in turn by the proneural genes. ChIP-seq data show that Snail binds *worniu* in neuroblasts, not in mesoderm. Thus, Snail can be one of the activators of *worniu* expression.

5.1.3 Post-transcriptional control of *worniu* and *snail*

Transcriptional regulation determines the spatio-temporal expression pattern of a protein to a large extent but not exclusively. Post-transcriptional and post-translational mechanisms provide additional control of a gene and protein expression. Post-transcriptional control mechanisms are diverse and are either mediated by RNA-bindig proteins or by microRNAs that bind to specific elements in the UTR of a transcript to control RNA translation, localization or stability. Like-wise, post-translational mechanisms regulate the localization, stability or activity of a protein (Wu and Zhou, 2010). The comparison of the mRNA expression patterns with the expression of either endogenous Snail, Worniu and Escargot proteins or the V5-tagged transgenic proteins showed a divergence of transcription and protein expression patterns. *snail* mRNA was present in the

neurogenic ectoderm and in neuroblasts, while the Snail protein could be detected only in neuroblasts. *worniu* mRNA was detectable in all neuroblasts, while a subpopulation of neuroblasts – namely the MP2 cells – had a normal pool of *worniu* mRNA, but were negative for Worniu protein. The MP2 cells are the only neuroblasts that do not possess a stem cell identity (Campos-Ortega and Hartenstein, 1997). The cell divides symetrically to produce two dMP2 and vMP2. Thus the Worniu protein is present exclusively in stem cell-like neuroblasts in the embryo.

Both cases suggest regulation at the post-transcriptional level. Two scenarios are possible. The translation of *snail* and *worniu* is cell-specifically inhibited, e.g. by a neuroectoderm-specific microRNA or RNA-binding protein that targets the *snail* mRNA or an MP2-specific translational repressor for *worniu*. Alternatively, Snail and Worniu protein stability is cell-specifically controlled and the proteins are rapidly degraded in those cell types. Vertebrate Snail is extensively controlled at the level of protein stability (Wu and Zhou, 2010) but it is unknown if the same mechanisms act on *Drosophila* Snail. Computer prediction analysis (RegRNA 2.0, regrna2.mbc.nctu.edu.tw) of the 3' UTRs of the *snail* and *worniu* mRNAs predicted binding sites of the Musashi protein, a regulator of stem cell identity (Charlesworth et al., 2006; MacNicol et al., 2011; Nakamura et al., 1994; Okabe et al., 2001; Okano et al., 2002). Musashi binding to the 3' UTR of the *ttk69* mRNA results in the mRNA degradation and subsequent decrease of the *ttk69* translational output keeps stem cell identity of the neuroblast lineage of mechanosensory bristles (Badenhorst, 2001; Okabe et al., 2001). No information is available about the expression of Musashi in embryonic neuroblasts at stages 8-9. Expression analysis as well as RNA binding assays must be performed to verify whether Musashi can be a regulator of translation of *worniu* and *snail*. The 3'UTR of *worniu* but not of *snail* also contains predicted AU-rich elements, which have been implicated in regulation of mRNA turnover and decay (Burow et al., 2015). Mutational analysis of the predicted motifs will be required to address if *worniu* is regulated at the level of RNA stability. Nevertheless, MP2 neuroblast cells were *worniu* positive on *in situ* hybridizations; moreover the level of mRNA was not down-regulated. Therefore targeted *worniu* mRNA decay is unlikely. Nevertheless, MP2-targeted expression of *worniu* mRNA with wild type and with the mutated AREs and Musashi binding elements can give additional insight into the control of neuroblast stem cell maintenance.

Expression of *escargot* at stages 5-8 has been studied in detail and the generation of the complex and dynamic expression pattern is described (Yagi and Hayashi, 1997; Yagi et al., 1998). Expression of *escargot* was also detected in neuroblast cells at stages 8-9 (Ashraf et al., 1999). My examination of *escargot* expression pattern detected expression of the gene in neurogenic ectoderm at stage 8 and in the cells of the ventral midline at stage 9 (fig. 19). Neuroblast-specific expression of *escargot* could be observed only at stages 10-11 and appeared to be lineage specific. Identification and characterization of *escargot*-positive neuroblast lineages can give additional insight into the roles of *escargot* in *Drosophila* development.

In several cases I detected mutually exclusive expression of *snail* and *escargot* (fig. 23): at stages 4-7 *snail*-positive mesoderm is surrounded by two stripes of *escargot*-positive neuroectoderm. At stage 8, when *snail* and *escargot* are expressed in the neurogenic ectoderm, *snail*-positive parasegments alternate with *escargot*-expressing parasegments. A cross-section through the *escargot*-positive parasegment shows that the delaminating neuroblasts are the only *snail*-positive cells in this segment. Snail represses expression of *escargot* in embryonic mesoderm at stages 4-7 (Yagi and Hayashi, 1997). Similar repression might take place in the *snail*-positive neuroblasts which delaminate from the *escargot*-positive neuroectoderm, however in the current ChIP data Snail binding to *escargot* could be detected only in mesoderm, not in the neuroblasts. As Snail protein is detected exclusively in neuroblasts, not in neurogenic ectoderm, the alternation of *snail*- and *escargot*-positive parasegments cannot be attributed to Snail-mediated repression. Although it is formally possible that the protein might be present below the detection limit. The neurogenic expression is likely controlled by the activating and repressive activity of different pair-rule genes of the anterior-posterior patterning system.

5.1.4 Cooperation of *snail*, *worniu* and *escargot*

To summarize, *escargot* mRNA is expressed in parasegment stripes in neurogenic ectoderm; Snail protein is expressed in neuroblasts during delamination; expression of Worniu starts when delamination is almost finished and persists in the neuroblasts throughout their development.

snail, *worniu* and *escargot* cooperate during initial embryonic neurogenesis (Ashraf and Ip, 2001; Ashraf et al., 1999; Cai et al., 2001). A series of experiments demonstrated that neither of the single mutants displays defects in CNS development. Only if all three genes are mutated is neurogenesis affected. Neuroblast cell polarity is disrupted in triple *snail/worniu/escargot* mutant embryos or in double *snail/escargot* mutant embryos treated with *worniu* RNAi. In the triple mutant embryos the neuroblasts fail to localize the cell fate determinant Pros and the adaptor protein Partner of Numb at the basal side of the cell membrane. In wild type these proteins are selectively distributed into the GMC where they promote differentiation and inhibit self-renewal (Wodarz and Huttner, 2003). In *snail/worniu/escargot* triple mutants Pros protein is reduced in GMCs and as a consequence the development of GMCs is affected. The single mutants, as well as the double mutants in any combination localize the cell polarity determinants correctly (Cai et al., 2001).

Expression studies identified a range of genes (*fushi tarazu*, *pdm*, *even skipped*, *inscuteable, string*) that lose their expression in triple *snail/worniu/escargot* deletion mutants, but retained expression in single or double mutant embryos. Expression of these genes was rescued by the tansgenic expression of either *snail*, *worniu* or *escargot* under control of the *snail* promoter (Ashraf and Ip, 2001; Ashraf et al., 1999). In the current ChIP data *ftz*, *eve* and *insc* are bound by Snail in mesoderm, while *eve* and *insc* are bound in neuroblasts as well. String is bound only in neuroblasts.

The redundancy of the Snail-like TFs in neural development can reflect a multilayer complexity of neuroblast specification and cell polarity formation. For instance, the polar localization of cell fate determinants in neuroblasts is controlled by two independent asymmetry mechanisms. One mechanism relies on the *insc* gene. In *insc* mutant embryos the cell fate determinants Pros and Numb are mislocalized during early mitosis; however, in anaphase and telophase both determinants restore their basal localization and segregate during asymmetric division to the GMCs. This effect, called telophase rescue, for the first time indicated the existence of a second *insc*-independent mechanism of asymmetry formation, that involves a number of genes such as the membrane-bound guanylate kinase Dlg (discs large) and its interaction partner Khc-73, a kynesin motor heavy chain (Neumüller and Knoblich, 2009). Both mechanisms depend on Snail-like TFs (Ashraf and Ip, 2001; Cai et al., 2001). In addition, Snail-like TFs control the orientation of the mitotic

spindle. Similar to the different mechanisms involved in neuroblast polarity also Snail-like TFs control different pathways of the cell polarity. Thus only triple *snail/worniu/escargot* mutation is able to destroy all three mechanisms of the cell polarity formation.

The expression of any of the three genes of *snail-like* family could rescue the polarity defects and gene expression defects (to a certain extent). The rescue constructs were expressed under control of a neuroblast-specific promoter of *snail*. Thus, *worniu* and *escargot* were expressed within the expression domain of *snail*. The zinc finger domains of Snail, Worniu and Escargot share high levels of similarity and Escargot and Snail bind very similar motifs, moreover, Snail, Worniu and Escargot might share the binding targets (Cai et al., 2001; Whiteley et al., 1992). Ectopic expression of Worniu and Escargot might force cross-reaction binding of the Snail´s targets and in this way substitute for Snail functions. Comparison of Snail, Worniu and Escargot binding targets in genome will provide additional insight into the mechanisms of the cooperation of Snail-like TFs.

5.1.5 Possible roles of Snail, Worniu and Escargot, suggested by their expression patterns

During gastrulation Snail represses expression of the ectodermal genes in mesoderm (Leptin, 1991). Similarly, Snail might repress expression of ectodermal genes in neuroblasts, when specification of a cell starts. Furthermore, during gastrulation *snail* mutants fail to disassemble adherence junctions at the sub-apical positions (Kölsch et al., 2007), and Snail is needed for the pulsing apical constriction of mesodermal cells (Martin et al., 2009). Thus, in mesoderm cells Snail is also required for the rearrangement of the cytoskeleton. Expression of *snail* takes place in a short time window during neuroblast delamination; therefore Snail might coordinate cell shape changes during delamination. However, neuroblasts delaminate in *snail* mutant embryos and also in the triple deletion *snail/worniu/escargot*. It remains however possible that the delamination process is disturbed or delayed in the mutant embryos. Careful examination of the delamination process, both in wild type and in mutant embryos, will clarify this aspect of Snail´s functions in neuroblasts.

worniu is transcribed in neuroblasts throughout their lifespan. A study of the gene expression profiles of wild type and *worniu* mutant neuroblasts in larvae with RNA-seq

proved that the loss of *worniu* results in the reduced transcription of the cell cycle genes and the increased transcription of the neural differentiation genes (Lai et al., 2012). Mutant cells show a delay of the prophase/metaphase transition and increased expression levels of splicing regulator Elav, the neuronal differentiation factor. In embryonic neuroblasts Worniu might have similar roles, namely maintain the stem cell identity by preventing neuroblast differentiation and promoting cell cycle progression (Lai et al., 2012).

The role of *escargot* in the initial neurogenesis remains obscure. As the gene is expressed only in a small subset of neuroblasts it remains unclear, how it contributes to early neuroblast development. *escargot* is expressed in the neuroectoderm, from which a subset of neuroblasts delaminate and it might have a function at this step. In the intestinal stem cells *escargot* is a critical stem cell determinant that maintains stemness by repressing differentiation-promoting factors, such as Pdm (Korzelius et al., 2014). *escargot* mutant embryos die at late embryonic stages or during pupation (Hayashi et al., 1993; Whiteley et al., 1992). Semi-lethal and viable alleles pupate and form adult flies with poorly differentiated abdomens, twisted legs and defects in the bristle formation, reduced eyes and other defects in the structures, derived from the imaginal discs. No defects in development of an embryonic or adult nervous system have been reported (Hayashi et al., 1993).

5.2   Evolutionary conserved functions of Snail-like TFs

Snail-like TFs have been studied in great detail in *Drosophila* and other model organisms. To better understand of the roles of the Snail-like TFs in *Drosophila* development I would like to briefly overview the known functions of Snail-like TFs in other organisms. Snail-like TFs have been implicated in epthelial-to-mesenchymal transition, cell migration and morphogenesis, in the maintenance of stem cell identity, in asymmetric division and cell survival and in neural development. All of these Snail-controlled mechanisms could also be involved during embryonic neuroblast development. Finally, I will provide evidence that, within the arthropod lineage, the function in neurogenesis is evolutionary older than the role in mesoderm formation.

5.2.1   Snail-like TFs and the epithelial-mesenchymal transition

During EMT epithelial cells lose their contact to the neighboring cells, down-regulate expression of E-cadherin, lose cell polarity, leave epithelial sheet and thus acquire classical mesenchymal treats (Acloque et al., 2009). Snail-like TFs are broadly known as inducers of epithelial-mesenchymal transitions during development and cancer progression (Acloque et al., 2009; Barrallo-Gimeno and Nieto, 2009; Cano et al., 2000; Nieto, 2002; Smith and Odero-Marah, 2012). *escargot* is required for an EMT-like transition of neural progenitor cells in the optic lobe of *Drosophila.* This involves also a down-regulation of E-cad in these cells (Apitz and Salecker, 2014). However, *Drosophila* gastrulation cannot be characterized as a classical EMT event: *snail* expression in mesoderm cells leads to the down-regulation of E-cadherin expression, however this repression is not required for the mesoderm to invaginate. Invagination proceeds normally even if E-cadherin is ectopically expressed in the mesoderm (Schäfer et al., 2014). Although the down-regulation of E-caderin expression by Snail is not crucial for mesoderm invagination in *Drosophila*, Snail functions at the level of adherens junctions also in this context. Snail is required for the change in the localization of adherens junctions from a subapical to the apical position before invagination (Kölsch et al., 2007). The neuroblast delamination is often described as a form of EMT (Ashraf et al., 2004; Lai et al., 2012). Unlike during classical EMT neuroblasts cells do not loose cell polarity. The rearrangements of the cytoskeleton and cell adhesion complexes in a neuroblast cell during delamination, as well as the expression levels of cadherin, have never been studied. Thus, neither in the case of gastrulation, nor in the neuroblast delamination Snail was characterized as a classical initiator of EMT. However, Snail might be required for the efficient EMT-like delamination of the neuroblasts. Therefore careful study of the wild type delamination, as well as the examination of neuroblast delamination in *snail* mutant embryos is needed.

5.2.2   Snail-like TFs and neurogenesis

Members of the Snail/Scratch superfamily of TFs participate in the neurogenesis in arthropods and vertebrates. Murine Scratch1 and Scratch2 are expressed downstream of the proneural genes and control the very first step of the neuronal migration, departure from the ventricular surface (Itoh et al., 2013). Snail-like TFs are involved in neuroblast

proliferation and migration in vertebrates (Zander et al., 2014) And Scratch is expressed in the CNS of mice (Nakakuraet al., 2001). Scratch-induced neuronal migration is mediated by repression of E-cadherin, and therefore resembles an EMT. Snail2 is necessary for neural crest EMT, and it is sufficient to induce generation and delamination of the neural crest cell in cranial regions (del Barrio and Nieto, 2002).

5.2.3   Snail-like TFs and stem cell identity

Multiple members of Snail/Scratch superfamily are involved in determination of the stem cell fate. Mammalian TFs Slug (a member of the Snail-like family) and Sox9, for example, cooperatively determine mammary stem cell fate. and a transient expression of exogenous Slug and Sox9 is sufficient to convert differentiated luminal cells into mammary stem cells (Guo et al., 2012). Murine Snail-like TF regulate the number and proliferation of neural precursor cells (Zander et al., 2014; reviewed in detail in Wu and Zhou, 2010). As mentioned previously, *worniu* maintains the stem cell identity of the ventral nerve cord cells. *escargot* has a similar function in the midgut (Korzelius et al., 2014; Lai et al., 2012). Since expression of *snail* can be detected during initial neurogenesis only at a short time point, *snail* might be involved in the establishment of the stem cell identity, while *worniu* in turn keeps stem cell identity thereafter.

5.2.4   Snail-like TFs in cell survival and asymmetric cell division

In *Caenorhabditis elegans* the Snail-like protein CES-1 is down-regulated in the neuro-secretory motoneuron sister cell to allow apoptosis, and directly controls the enzymatic machinery involved in asymmetric cell division. A gain-of-function mutation of the *ces-1* gene, which results in the gene´s miss- and over-expression, prevents the programmed death of the NSM sister cell (Hatzold and Conradt, 2008). In mice and chicken embryos the expression of Snail-like TFs negatively correlates with the cell death in different developing tissues (Haraguchi, 2009; Vega et al., 2004). Down-regulation of *snail* expression, induced by antisense oligonucleotides increases cell death in colon tumors in a mouse model (Roy et al., 2004). Snail-dependent resistance to apoptosis is tightly linked to the cell cycle progression and stem cell identity maintenance. Since apoptosis is not part of the developmental program of *Drosophila* early embryogenesis, a role as survival factors

cannot be assigned to Snail-like TFs in this context. Similar to the asymmetric division of the MSN neuroblast in *C. elegans*, Snail, Worniu and Escargot are required for asymmetric division of *Drosophila* neuroblasts.

5.2.5 Snail-like TFs in arthropod development

Several studies performed on the non-model arthropoda species provide an evolutionary perspective on the developmental functions of the Snail-like TFs.

5.2.5.1 The functions of Snail-like TFs in development of nervous system in different non-vertebrate species

In spiders (*Cupiennius salei*) neuroblasts are generated from invagination of pro-neuronal cell clusters localized in the neurogenic ectoderm in a stereotypical pattern (Stollewerk et al., 2003; Weller and Tautz, 2003). Every neuroblast expresses *snail* in a short pulse during the initial step of the cell invagination, preceding expression of the pan-neuroblast gene *pros*.

In myriapods (*Glomeris marginata*), Snail-like TF is initially expressed in the presumptive neural precursors; later in all neural precursors in the ventral neuroectoderm and in the ectodermal midline (Pioro and Stollewerk, 2006).

In amphipod crustacean (*Parhyale hawaiensis*) three *snail* homologs (*Ph-sna 1*, *2* and *3*) are expressed in the ventral neuroectoderm with a different spatiotemporal pattern. No expression of *snail-like* genes was found in mesoderm during gastrulation. *Ph-Sna1* is expressed in mesoteloblasts, mesoderm cells which are formed during mesoderm specification, long after gastrulation (Hannibal et al., 2012).

In branchiopoda (*Daphnia magna*) *snail*'s homologue *Dam_sna* is the first neural gene to be expressed in the ventral neuroectoderm (VNE), a tissue which generates neuroblasts. At the start of neuroectoderm specification *snail* is expressed in bilateral longitudinal bands, flanking the midline. Each band is 1 to 6 cells wide. Additionally, *Dam_sna* mRNA is expressed in the intersegmental furrows, the transverse stripes lateral to the VNE. Similarly to the spider *snail*, expression of *Dam_sna* precedes expression of other known neuroblast-specific genes *prospero* and *asense* (Ungerer et al., 2011).

Despite Snail-like TFs being expressed in the developing nervous system in multiple organisms, its expression pattern (and the functions) can be variable. It is expressed in the neuroectoderm, which is able to produce neuroblast cells (*Daphnia*), in a short pulse during neuroblast formation (spiders) or it can be continuously present in all cells of the nerve cord. In *Drosophila*, *snail* underwent a tandem duplication to produce three paralogous genes: *snail*, *worniu* and *escargot*. TFs are expressed at three different time points: *escargot* is expressed in the neurogenic ectoderm; *snail* is expressed in the neurogenic ectoderm and in the neuroblasts during their formation, while *worniu* is continuously expressed in the neuroblasts. Since in non-insect arthropods (spiders, myriapods, crustaceans) Snail-like TFs are not expressed in mesoderm during gastrulation, in the evolution history of arthropods the mesoderm-specific functions of Snail seem to be newly acquired. Neuroblast transcription network of Worniu, as well as the list of Worniu target genes, comparison to mesodermal and neuroblast target genes of Snail will provide further information about evolution of the Snail transcription networks in the neurogenesis of arthropods.

5.3 ChIP-seq

5.3.1 Challenges of the experiment

Since the first description of chromatin immunoprecipitation in 1999 (Kuo and Allis, 1999), the technique has been rapidly adopted by many laboratories and developed as an important technique to map the binding of proteins to genomic DNA. The protocol was adapted to different model organisms and scientific questions and continuously improved. Despite the widespread use of the technique and its seemingly easy setup, it faces the experimenter with many challenges. Some of these challenges and difficulties of ChIP experiments have been described multiple times (Bonn et al., 2012; Park, 2009; Sandmann et al., 2006b).

One major challenge of a ChIP experiment using embryos is the collection of large amounts of biological material. For library preparation and sequencing at least 5 ng of precipitated DNA are recommended. This appears to be a small quantity but is in fact technically challenging. The amount of precipitated DNA depends on many factors. First of all, the quality of the antibody is a crucial factor. It has to be specific for the protein of interest and its affinity has to be high enough to result in efficient precipitation of the protein-DNA complex. A titration has to be performed to determine the correct antibody to chromatin (antigen) ratio. The extent of binding of the protein to DNA also determines how much DNA is precipitated. A TF that is bound only to a handful of loci will bring down less DNA than a protein that is broadly associated with DNA, such as histones. The use of different DNA purification protocols in my hands drastically changed amount DNA in the final ChIP sample, as well as the levels of unspecific background.

For the ChIPs performed on *Drosphila* embryos established protocols recommend to collect at least 1 - 1.5 g of tightly staged embryos (Sandmann et al., 2007). 40 µg of genomic DNA should be used as input for one IP. When the starting material is collected, it should also be considered that a lot of DNA is lost already during the preparation of chromatin from embryos. In my experience, approximately 70% of the chromatin present in mesoderm stage embryos is lost during embryo and nuclear lysis and the sonication of chromatin. Therefore, I required 8 g of tightly-staged embryos for the entire final ChIP

experiment. Totally for the all experiments at least 25 g of tightly staged embryos were used.

The authors of the protocol, which was used as a standard reference in my experiments (Sandmann et al., 2006b), prefer phenol-chloroform DNA purification to column-based techniques, since, according to the authors´ experience, column-based purification can interfere with the subsequent qPCR and sequencing analysis of the sample. In my experience application of the QiaQuick PCR purification kit produced high quality DNA samples applicable in qPCR as well as sequencing analysis, and, on the other hand, removed an extra overnight incubation from the protocol workflow.

5.3.2   Analysis of ChIP-seq datasets

In the course of this project I performed ChIP experiments with two different anti-Snail antibodies and with a V5 antibody on *V5-HA-snail* and *HA-V5-worniu* transgenic embryos. Only the ChIP experiment with the rabbit anti-Snail antibody was successful. The ChIP with the guinea pig anti-Snail antibody enriched known Snail-bound regions in the mesoderm but produced only low enrichment in neuroblasts that was not consistent between the technical replicates. The chromatin was isolated from whole embryos and the fraction of nuclei that express Snail is higher (~25%) in early stages. The mesoderm anlage accounts for approximately a quarter of the embryo, while neuroblasts might not constitute more than 10% of all cells at stages 9 to 10. To compensate for this difference I used 2 times more input chromatin for the neuroblast-stage ChIP. Still, this amount might not have been sufficient to guarantee enrichment of precipitated DNA fragments above background in neuroblasts. The ChIP experiment with the rabbit anti-Snail antibody worked not only in the mesoderm but also in neuroblasts. Two factors could have contributed to this success: a) a higher affinity of the antibody and b) twice as much input was used for the experiments, as the experimental quadruplicates were pooled pairwise in the end. In support of the latter argument is the fact, that the samples produced with the rabbit antibody were the only ones in which I could detect DNA (Table 13).

The anti-V5 ChIP datasets showed no enrichment above input. The V5-HA-Snail ChIP samples did not reveal any enrichment at any of the known Snail binding sites. A major problem of the anti-V5 ChIPs might have been the fact, that the tagged protein was

expressed in a wild type background. Therefore, the anti-V5 antibody could IP only half of the protein bound to DNA. Thus a competition with the endogenous untagged protein might have lowered the efficiency of the IP. To overcome this problem, a novel anti-Worniu antibody was generated. Preliminary tests demonstrated a high specificty of the antibody in immunostainings. The applicability of the antibody to ChIP experiments awaits further testing. The *wor-G4, UAS-CD8-GFP* line, which expresses GFP exclusively in neuroblasts, could be used to sort the embryonic neuroblast cells for a cell type-specific chromatin immunoprecipitation, BiTS-ChIP (Bonn et al., 2012).

5.3.3 Primary analysis of the binding sites of Snail in mesoderm and neuroblasts

The sequencing data prepared with the rabbit anti-Snail antibody was applicable to peak calling and to the analysis of the Snail binding. MACS2 called 1703 peaks in the mesoderm and 2501 peaks in neuroblasts at an FDR cutoff of 1%. The datasets can be improved by filtering out the high credibility peaks. Multiple approaches can be used for this purpose. One of them, selection of the peaks that contain Snail binding motifs, was applied in this work and filtered out >50% of the detected peaks.

Another alternative approach, which might be used for further improvement of the database, includes calculation of the peak reproducibility in the ChIP technical duplicates and selection of highly reproducible peaks. The modENCODE consortium recommends the use of the irreproducible discovery rate (IDR) to improve the parameters for peak calling and distinguish real signal from noise (Landt et al., 2012).

5.3.4 Mesodermal and neuroblast targets of Snail

A preliminary analysis for the presence of known tissue-specific targets of Snail (Rembold et al., 2014; Southall and Brand, 2009) in the newly generated database of Snail target genes shows that Snail is bound to the majority of the inspected mesoderm targets (namely to *rho*, *Mef2*, *wntD*, *tin*, *sim*, *sog* and *vnd*). *htl*, a known mesodermal target of Snail was bound in only neuroblasts. In the neuroblasts Snail binds the known neuroblast-specific targets (*pros*, *CycE*, *lola*, *neur*, *sim*, *grh*, *stg, Dl*). *baz*, also known as a target of Snail

bound in neuroblasts, is bound in the ChIP-seq database exclusively in mesoderm. Thus, the pilot analysis confirms the credibility of the tissue-specific databases of Snail binding.

In the nervous system the expression of *snail* is activated by two separable enhancers, a PNS- and a CNS-specfic element (Ip et al., 1994). The PNS enhancer localizes within 0.8 kb upstream of the TSS and is bound by Snail both in neuroblasts and in the mesoderm (fig. 37). In neuroblasts Snail binding to the enhancer is stronger than in the mesoderm. The CNS element localizes within 2.0 and 2.8 kb upstream of *snail*. It is part of two larger promoter elements that were characterized in detail (Ip et al., 1992, 1994). The 2.8 kb promoter contains all three regulatory elements, PNS-, CNS- and mesodermal, and displays a pattern of expression that is almost identical to the *snail* expression pattern. The 6 kb promoter fragment contains additional distal sequences with an expression pattern identical to the 2.8 kb promoter (Ip et al., 1994).

Neuroblast-specific binding to the CNS-element in the 2.8 kb enhancer was not detected. Instead, Snail binds at approximately -4.0 kb within the -6.0 kb enhancer. Expression analysis of transgenic reporter constructs (Ip et al., 1994) revealed a more robust and stronger expression from the -6.0 kb enhancer compared to the -2.8 kb enhancer. It is possible that additional CNS regulatory elements have been missed.

Snail binds *worniu* and *escargot* in a tissue-specific manner as well. *escargot* is bound exclusively in the mesoderm. This finding is in agreement with previously published reports that Snail represses mesodermal expression of *escargot* (Yagi and Hayashi, 1997). *worniu*, in turn, is bound specifically in neuroblasts. This observation, in combination with the observed sequential expression of *snail* and *worniu* in neuroblasts, feeds the speculation that Snail directly activates expression of *worniu*. Further experiments, in particular the analysis of the expression of reporter genes under control of selected enhancers, are required to verify and prove this hypothesis.

5.4　　　**Future directions**

In summary, I have generated a genome-wide set of tissue-specific Snail-bound regions and potential target genes. This dataset provides the basis to analyze the mechanism of tissue-specific interaction between TFs and enhancers. In particular, future research will address the following questions:

1. A genome-wide comparison of the genes bound by Snail in mesoderm and in neuroblasts will be made to elucidate the mesoderm- and neuroblast-specific gene regulatory network of Snail.

2. The set of tissue-specific Snail-target enhancers will be analysed for their architecture and for the characteristic, tissue-specific elements of the enhancers.

3. Finally, the predictions generated with *in silico* models will be tested with enhancer assays *in vivo*.

6   Bibliography

Acloque, H., Adams, M.S., Fishwick, K., Bronner-Fraser, M., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions: The importance of changing cell state in development and disease. J. Clin. Invest. *119*, 1438–1449.

Alberga, A., Boulay, J.L., Kempe, E., Dennefeld, C., and Haenlin, M. (1991). The snail gene required for mesoderm formation in Drosophila is expressed dynamically in derivatives of all three germ layers. Development *111*, 983–992.

Apitz, H., and Salecker, I. (2014). A region-specific neurogenesis mode requires migratory progenitors in the Drosophila visual system. Nat. Neurosci. *18*, 46–55.

Ashraf, S.I., and Ip, Y.T. (2001). The Snail protein family regulates neuroblast expression of inscuteable and string, genes involved in asymmetry and cell division in Drosophila. Development *128*, 4757–4767.

Ashraf, S.I., Hu, X., Roote, J., and Ip, Y.T. (1999). The mesoderm determinant snail collaborates with related zinc-finger proteins to control Drosophila neurogenesis. EMBO J. *18*, 6426–6438.

Ashraf, S.I., Ganguly, A., Roote, J., and Ip, Y.T. (2004). Worniu, a Snail family zinc-finger protein, is required for brain development in Drosophila. Dev. Dyn. *231*, 379–386.

Aybar, M.J., Nieto, M.A., and Mayor, R. (2003). Snail precedes slug in the genetic cascade required for the specification and migration of the Xenopus neural crest. Development *130*, 483–494.

Badenhorst, P. (2001). Tramtrack controls glial number and identity in the Drosophila embryonic CNS. Development *128*, 4093–4101.

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299–308.

Barrallo-Gimeno, A., and Nieto, M.A. (2009). Evolutionary history of the Snail/Scratch superfamily. Trends Genet. *25*, 248–252.

Del Barrio, M.G., and Nieto, M.A. (2002). Overexpression of Snail family members highlights their ability to promote chick neural crest formation. Development *129*, 1583–1593.

Bennett, S. (2004). Solexa Ltd. Pharmacogenomics *5*, 433–438.

Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Hall, K.P., Evers, D.J., Barnes, C.L., Bignell, H.R., et al. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. Nature *456*, 53–59.

Berger, C., Urban, J., and Technau, G.M. (2001). Stage-specific inductive signals in the Drosophila neuroectoderm control the temporal sequence of neuroblast specification. Development *128*, 3243–3251.

Berman, B.P., Nibu, Y., Pfeiffer, B.D., Tomancak, P., Celniker, S.E., Levine, M., Rubin, G.M., and Eisen, M.B. (2002). Exploiting transcription factor binding site clustering to identify cis-regulatory modules involved in pattern formation in the Drosophila genome. Proc. Natl. Acad. Sci. U. S. A. *99*, 757–762.

Bischof, J., Maeda, R.K., Hediger, M., Karch, F., and Basler, K. (2007). An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. Proc. Natl. Acad. Sci. U. S. A. *104*, 3312–3317.

Blanco, M.J., Barrallo-Gimeno, A., Acloque, H., Reyes, A.E., Tada, M., Allende, M.L., Mayor, R., and Nieto, M.A. (2007). Snail1a and Snail1b cooperate in the anterior migration of the axial mesendoderm in the zebrafish embryo. Development *134*, 4073–4081.

Boettiger, A.N., and Levine, M. (2013). Rapid transcription fosters coordinate snail expression in the Drosophila embryo. Cell Rep. *3*, 8–15.

Bonn, S., Zinzen, R.P., Perez-Gonzalez, A., Riddell, A., Gavin, A.-C., and Furlong, E.E.M. (2012). Cell type-specific chromatin immunoprecipitation from multicellular complex samples using BiTS-ChIP. Nat. Protoc. *7*, 978–994.

Bownes, M., Abrahamssen, N., Gilman, C., Howden, R., Martinez, A., and Slee, R. (1990). Drosophila: A Laboratory Handbook. By Michael Ashburner. New York: Cold Spring Harbor Laboratory. 1989. 1331 pages. Price $180.00. ISBN 0 87969 321 5.Drosophila: A Laboratory Manual. By Michael Ashburner. New York: Cold Spring Harbor Laboratory. 1989. 434.

Broadus, J., and Doe, C.Q. (1995). Evolution of neuroblast identity: seven-up and prospero expression reveal homologous and divergent neuroblast fates in Drosophila and Schistocerca. Development *121*, 3989–3996.

Brody, T., and Odenwald, W.F. (2000). Programmed transformations in neuroblast gene expression during Drosophila CNS lineage development. Dev. Biol. *226*, 34–44.

Buck, M.J., and Lieb, J.D. (2006). A chromatin-mediated mechanism for specification of conditional transcription factor targets. Nat. Genet. *38*, 1446–1451.

Burow, D.A., Umeh-Garcia, M.C., True, M.B., Bakhaj, C.D., Ardell, D.H., and Cleary, M.D. (2015). Dynamic regulation of mRNA decay during neural development. Neural Dev. *10*.

Cabrera, C. V, Martinez-Arias, A., and Bate, M. (1987). The expression of three members of the achaete-scute gene complex correlates with neuroblast segregation in Drosophila. Cell *50*, 425–433.

Cai, Y., Chia, W., and Yang, X. (2001). A family of snail-related zinc finger proteins regulates two distinct and parallel mechanisms that mediate Drosophila neuroblast asymmetric divisions. EMBO J. *20*, 1704–1714.

Campos-Ortega, J.A. (1993). Mechanisms of early neurogenesis in Drosophila melanogaster. J. Neurobiol. *24*, 1305–1327.

Campos-Ortega, J.A. (1994). Genetic mechanisms of early neurogenesis in Drosophila melanogaster. J. Physiol. *88*, 111–122.

Campos-Ortega, J.A., and Hartenstein, V. (1997). The Embryonic Development of Drosophila melanogaster.

Cano, A., Pérez-Moreno, M.A., Rodrigo, I., Locascio, A., Blanco, M.J., del Barrio, M.G., Portillo, F., and Nieto, M.A. (2000). The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. Nat. Cell Biol. *2*, 76–83.

Carver, E.A., Jiang, R., Lan, Y., Oram, K.F., and Gridley, T. (2001). The mouse snail gene encodes a key regulator of the epithelial-mesenchymal transition. Mol. Cell. Biol. *21*, 8184–8188.

Castanon, I., Von Stetina, S., Kass, J., and Baylies, M.K. (2001). Dimerization partners determine the activity of the Twist bHLH protein during Drosophila mesoderm development. Development *128*, 3145–3159.

Charlesworth, A., Wilczynska, A., Thampi, P., Cox, L.L., and MacNicol, A.M. (2006). Musashi regulates the temporal order of mRNA translation during Xenopus oocyte maturation. EMBO J. *25*, 2792–2801.

Chopra, V.S., and Levine, M. (2009). Combinatorial patterning mechanisms in the Drosophila embryo. Brief. Funct. Genomic. Proteomic. *8*, 243–249.

Ciruna, B., and Rossant, J. (2001). FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak. Dev. Cell *1*, 37–49.

Copeland, N.G., Jenkins, N.A., and Court, D.L. (2001). Recombineering: a powerful new tool for mouse functional genomics. Nat. Rev. Genet. *2*, 769–779.

Doe, C.Q. (1992). Molecular markers for identified neuroblasts and ganglion mother cells in the Drosophila central nervous system. Development *116*, 855–863.

Doe, C.Q., Hiromi, Y., Gehring, W.J., and Goodman, C.S. (1988). Expression and function of the segmentation gene fushi tarazu during Drosophila neurogenesis. Science *239*, 170–175.

Dunipace, L., Ozdemir, A., and Stathopoulos, A. (2011). Complex interactions between cis-regulatory modules in native conformation are critical for Drosophila snail expression. Development *138*, 4075–4084.

Ejsmont, R.K., Sarov, M., Winkler, S., Lipinski, K.A., and Tomancak, P. (2009). A toolkit for high-throughput, cross-species gene engineering in Drosophila. Nat. Methods *6*, 435–437.

Feng, J., Liu, T., and Zhang, Y. (2011). Using MACS to identify peaks from ChIP-Seq data. Curr. Protoc. Bioinformatics *Chapter 2*, Unit 2.14.

Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using MACS. Nat. Protoc. *7*, 1728–1740.

Fuse, N., Hirose, S., and Hayashi, S. (1994). Diploidy of Drosophila imaginal cells is maintained by a transcriptional repressor encoded by escargot. Genes Dev. *8*, 2270–2281.

Gilchrist, D.A., Fargo, D.C., and Adelman, K. (2009). Using ChIP-chip and ChIP-seq to study the regulation of gene expression: genome-wide localization studies reveal widespread regulation of transcription elongation. Methods *48*, 398–408.

Green, M., and Sambrook, J. (2012). Molecular Cloning: A Laboratory Manual (Fourth Edition).

Grosskortenhaus, R., Pearson, B.J., Marusich, A., and Doe, C.Q. (2005). Regulation of temporal identity transitions in Drosophila neuroblasts. Dev. Cell *8*, 193–202.

Groth, A.C., Fish, M., Nusse, R., and Calos, M.P. (2004). Construction of Transgenic Drosophila by Using the Site-Specific Integrase from Phage φC31. Genetics *166*, 1775–1782.

Guo, W., Keckesova, Z., Donaher, J.L., Shibue, T., Tischler, V., Reinhardt, F., Itzkovitz, S., Noske, A., Zürrer-Härdi, U., Bell, G., et al. (2012). Slug and Sox9 cooperatively determine the mammary stem cell state. Cell *148*, 1015–1028.

Hannibal, R.L., Price, A.L., Parchem, R.J., and Patel, N.H. (2012). Analysis of snail genes in the crustacean Parhyale hawaiensis: insight into snail gene family evolution. Dev. Genes Evol. *222*, 139–151.

Haraguchi, M. (2009). The role of the transcriptional regulator snail in cell detachment, reattachment and migration. Cell Adh. Migr. *3*, 259–263.

Hartenstein, V., and Campos-Ortega, J.A. (1984). Early neurogenesis in wild-typeDrosophila melanogaster. Wilhelm Roux's Arch. Dev. Biol. *193*, 308–325.

Hartenstein, V., and Wodarz, A. (2013). Initial neurogenesis in Drosophila. Wiley Interdiscip. Rev. Dev. Biol. *2*, 701–721.

Hatzold, J., and Conradt, B. (2008). Control of apoptosis by asymmetric cell division. PLoS Biol. *6*, 771–784.

Hayashi, S., Hirose, S., Metcalfe, T., and Shirras, A.D. (1993). Control of imaginal cell development by the escargot gene of Drosophila. Development *118*, 105–115.

Hemavathy, K., Ashraf, S.I., and Ip, Y.T. (2000). Snail/slug family of repressors: slowly going into the fast lane of development and cancer. Gene *257*, 1–12.

Hemavathy, K., Hu, X., Ashraf, S.I., Small, S.J., and Ip, Y.T. (2004). The repressor function of snail is required for Drosophila gastrulation and is not replaceable by Escargot or Worniu. Dev. Biol. *269*, 411–420.

Ho, J.W.K., Bishop, E., Karchenko, P. V, Nègre, N., White, K.P., and Park, P.J. (2011). ChIP-chip versus ChIP-seq: lessons for experimental design and data analysis. BMC Genomics *12*, 134.

Hong, J.W., Hendrix, D.A., and Levine, M.S. (2008). Shadow enhancers as a source of evolutionary novelty. Sci. (New York, NY) *321*, 1314.

Hu, C.-T., Wu, J.-R., Chang, T.Y., Cheng, C.-C., and Wu, W.-S. (2008). The transcriptional factor Snail simultaneously triggers cell cycle arrest and migration of human hepatoma HepG2. J. Biomed. Sci. *15*, 343–355.

Huang, C.-H., Yang, W.-H., Chang, S.-Y., Tai, S.-K., Tzeng, C.-H., Kao, J.-Y., Wu, K.-J., and Yang, M.-H. (2009). Regulation of membrane-type 4 matrix metalloproteinase by SLUG contributes to hypoxia-mediated metastasis. Neoplasia *11*, 1371–1382.

Ip, Y.T., Park, R.E., Kosman, D., Yazdanbakhsh, K., and Levine, M. (1992). dorsal-twist interactions establish snail expression in the presumptive mesoderm of the Drosophila embryo. Genes Dev. *6*, 1518–1530.

Ip, Y.T., Levine, M., and Bier, E. (1994). Neurogenic expression of snail is controlled by separable CNS and PNS promoter elements. Development *120*, 199–207.

Isshiki, T., Pearson, B., Holbrook, S., and Doe, C.Q. (2001). Drosophila neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny. Cell *106*, 511–521.

Itoh, Y., Moriyama, Y., Hasegawa, T., Endo, T.A., Toyoda, T., and Gotoh, Y. (2013). Scratch regulates neuronal migration onset via an epithelial-mesenchymal transition-like mechanism. Nat. Neurosci. *16*, 416–425.

Kambadur, R., Koizumi, K., Stivers, C., Nagle, J., Poole, S.J., and Odenwald, W.F. (1998). Regulation of POU genes by castor and hunchback establishes layered compartments in the Drosophila CNS. Genes Dev. *12*, 246–260.

Kanai, M.I., Okabe, M., and Hiromi, Y. (2005). seven-up Controls switching of transcription factors that specify temporal identities of Drosophila neuroblasts. Dev. Cell *8*, 203–213.

Kim, J.Y., Kim, Y.M., Yang, C.H., Cho, S.K., Lee, J.W., and Cho, M. (2012). Functional regulation of Slug/Snail2 is dependent on GSK-3β-mediated phosphorylation. FEBS J. *279*, 2929–2939.

Knight, R.D., and Shimeld, S.M. (2001). Identification of conserved C2H2 zinc-finger gene families in the Bilateria. Genome Biol. *2*, RESEARCH0016.

Kohwi, M., and Doe, C.Q. (2013). Temporal fate specification and neural progenitor competence during development. Nat. Rev. Neurosci. *14*, 823–838.

Kölsch, V., Seher, T., Fernandez-Ballester, G.J., Serrano, L., and Leptin, M. (2007). Control of Drosophila gastrulation by apical localization of adherens junctions and RhoGEF2. Science *315*, 384–386.

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

Kosman, D., Ip, Y.T., Levine, M., and Arora, K. (1991). Establishment of the mesoderm-neuroectoderm boundary in the Drosophila embryo. Science *254*, 118–122.

Kuo, M.H., and Allis, C.D. (1999). In vivo cross-linking and immunoprecipitation for studying dynamic Protein:DNA associations in a chromatin environment. Methods *19*, 425–433.

Kvon, E.Z., Kazmar, T., Stampfel, G., Yáñez-Cuna, J.O., Pagani, M., Schernhuber, K., Dickson, B.J., and Stark, A. (2014). Genome-scale functional characterization of Drosophila developmental enhancers in vivo. Nature *512*, 91–95.

Lai, S.-L., and Doe, C.Q. (2014). Transient nuclear Prospero induces neural progenitor quiescence. Elife *3*.

Lai, S.-L., Miller, M.R., Robinson, K.J., and Doe, C.Q. (2012). The Snail family member Worniu is continuously required in neuroblasts to prevent Elav-induced premature differentiation. Dev. Cell *23*, 849–857.

Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B.E., Bickel, P., Brown, J.B., Cayting, P., et al. (2012). ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res. *22*, 1813–1831.

Langeland, J.A., Tomsa, J.M., Jackman, W.R., and Kimmel, C.B. (1998). An amphioxus snail gene: expression in paraxial mesoderm and neural plate suggests a conserved role in patterning the chordate embryo. Dev. Genes Evol. *208*, 569–577.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lee, C.-Y., Robinson, K.J., and Doe, C.Q. (2006). Lgl, Pins and aPKC regulate neuroblast self-renewal versus differentiation. Nature *439*, 594–598.

Lee, E.C., Yu, D., Martinez de Velasco, J., Tessarollo, L., Swing, D.A., Court, D.L., Jenkins, N.A., and Copeland, N.G. (2001). A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. Genomics *73*, 56–65.

Leptin, M. (1991). twist and snail as positive and negative regulators during Drosophila mesoderm development. Genes Dev. *5*, 1568–1576.

Leptin, M., and Grunewald, B. (1990). Cell shape changes during gastrulation in Drosophila. Development *110*, 73–84.

Li, X., Chen, Z., and Desplan, C. (2013). Temporal patterning of neural progenitors in Drosophila. Curr. Top. Dev. Biol. *105*, 69–96.

Lim, H.-Y., and Tomlinson, A. (2006). Organization of the peripheral fly eye: the roles of Snail family transcription factors in peripheral retinal apoptosis. Development *133*, 3529–3537.

Lin, Y.C., Jhunjhunwala, S., Benner, C., Heinz, S., Welinder, E., Mansson, R., Sigvardsson, M., Hagman, J., Espinoza, C.A., Dutkowski, J., et al. (2010). A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. Nat. Immunol. *11*, 635–643.

Lischka, P., Rosorius, O., Trommer, E., and Stamminger, T. (2001). A novel transferable nuclear export signal mediates CRM1-independent nucleocytoplasmic shuttling of the human cytomegalovirus transactivator protein pUL69. EMBO J. *20*, 7271–7283.

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

Lu, Z., Ghosh, S., Wang, Z., and Hunter, T. (2003). Downregulation of caveolin-1 function by EGF leads to the loss of E-cadherin, increased transcriptional activity of beta-catenin, and enhanced tumor cell invasion. Cancer Cell *4*, 499–515.

M. Hoy (2013). Insect Molecular Genetics, 3rd Edition An Introduction to Principles and Applications. U.S.A. Insitute of Food and Agricultural Sciences, University of Florida, Gainesville, ed. p. 840.

Macarthur, S., Li, X., Li, J., Brown, J., Chu, H., Zeng, L., Grondona, B., Hechmer, A., Simirenko, L., Keränen, S. V, et al. (2009). Developmental roles of 21 Drosophila transcription factors are determined by quantitative differences in binding to an overlapping set of thousands of genomic regions. Genome Biol. *10*, R80.

MacArthur, S., Li, X.-Y., Li, J., Brown, J.B., Chu, H.C., Zeng, L., Grondona, B.P., Hechmer, A., Simirenko, L., Keränen, S.V.E., et al. (2009). Developmental roles of 21 Drosophila transcription factors are determined by quantitative differences in binding to an overlapping set of thousands of genomic regions. Genome Biol. *10*, R80.

MacNicol, M.C., Cragle, C.E., and MacNicol, A.M. (2011). Context-dependent regulation of Musashi-mediated mRNA translation and cell cycle regulation. Cell Cycle *10*, 39–44.

Martin, A.C., Kaschube, M., and Wieschaus, E.F. (2009). Pulsed contractions of an actin-myosin network drive apical constriction. Nature *457*, 495–499.

Martin, A.C., Gelbart, M., Fernandez-Gonzalez, R., Kaschube, M., and Wieschaus, E.F. (2010). Integration of contractile forces during tissue invagination. J. Cell Biol. *188*, 735–749.

Mauhin, V., Lutz, Y., Dennefeld, C., and Alberga, A. (1993). Definition of the DNA-binding site repertoire for the Drosophila transcription factor SNAIL. Nucleic Acids Res. *21*, 3951–3957.

Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. *5*.

Motohashi, T., Aoki, H., Chiba, K., Yoshimura, N., and Kunisada, T. (2007). Multipotent cell fate of neural crest-like cells derived from embryonic stem cells. Stem Cells *25*, 402–410.

Muqbil, I., Wu, J., Aboukameel, A., Mohammad, R.M., and Azmi, A.S. (2014). Snail nuclear transport: the gateways regulating epithelial-to-mesenchymal transition? Semin. Cancer Biol. *27*, 39–45.

Nakamura, M., Okano, H., Blendy, J.A., and Montell, C. (1994). Musashi, a neural RNA-binding protein required for Drosophila adult external sensory organ development. Neuron *13*, 67–81.

Nakayama, H., Scott, I.C., and Cross, J.C. (1998). The transition to endoreduplication in trophoblast giant cells is regulated by the mSNA zinc finger transcription factor. Dev. Biol. *199*, 150–163.

Neumüller, R.A., and Knoblich, J.A. (2009). Dividing cellular asymmetry: Asymmetric cell division and its implications for stem cells and cancer. Genes Dev. *23*, 2675–2699.

Nibu, Y., Zhang, H., and Levine, M. (1998a). Interaction of short-range repressors with Drosophila CtBP in the embryo. Science *280*, 101–104.

Nibu, Y., Zhang, H., Bajor, E., Barolo, S., Small, S., and Levine, M. (1998b). dCtBP mediates transcriptional repression by Knirps, Krüppel and Snail in the Drosophila embryo. EMBO J. *17*, 7009–7020.

Nieto, M.A. (2002). The snail superfamily of zinc-finger transcription factors. Nat. Rev. Mol. Cell Biol. *3*, 155–166.

Novotny, T., Eiselt, R., and Urban, J. (2002). Hunchback is required for the specification of the early sublineage of neuroblast 7-3 in the Drosophila central nervous system. Development *129*, 1027–1036.

Okabe, M., Imai, T., Kurusu, M., Hiromi, Y., and Okano, H. (2001). Translational repression determines a neuronal potential in Drosophila asymmetric cell division. Nature *411*, 94–98.

Okano, H., Imai, T., and Okabe, M. (2002). Musashi: a translational regulator of cell fate. J. Cell Sci. *115*, 1355–1359.

Palii, C.G., Perez-Iratxeta, C., Yao, Z., Cao, Y., Dai, F., Davidson, J., Atkins, H., Allan, D., Dilworth, F.J., Gentleman, R., et al. (2010). Differential genomic targeting of the transcription factor TAL1 in alternate haematopoietic lineages. EMBO J. *30*, 494–509.

Park, P.J. (2009). ChIP–seq: advantages and challenges of a maturing technology. Nat. Rev. Genet. *10*, 669–680.

Patel, N.H., Snow, P.M., and Goodman, C.S. (1987). Characterization and cloning of fasciclin III: a glycoprotein expressed on a subset of neurons and axon pathways in Drosophila. Cell *48*, 975–988.

Peinado, H., Olmeda, D., and Cano, A. (2007). Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? Nat. Rev. Cancer *7*, 415–428.

Perry, M.W., Boettiger, A.N., Bothma, J.P., and Levine, M. (2010). Shadow enhancers foster robustness of Drosophila gastrulation. Curr. Biol. *20*, 1562–1567.

Pioro, H.L., and Stollewerk, A. (2006). The expression pattern of genes involved in early neurogenesis suggests distinct and conserved functions in the diplopod Glomeris marginata. Dev. Genes Evol. *216*, 417–430.

Qi, D., Bergman, M., Aihara, H., Nibu, Y., and Mannervik, M. (2008). Drosophila Ebi mediates Snail-dependent transcriptional repression through HDAC3-induced histone deacetylation. EMBO J. *27*, 898–909.

Rembold, M., Ciglar, L., Yáñez-Cuna, J.O., Zinzen, R.P., Girardot, C., Jain, A., Welte, M.A., Stark, A., Leptin, M., and Furlong, E.E.M. (2014). A conserved role for Snail as a potentiator of active transcription. Genes Dev. *28*, 167–181.

Roark, M., Sturtevant, M.A., Emery, J., Vaessin, H., Grell, E., and Bier, E. (1995). scratch, a pan-neural gene encoding a zinc finger protein related to snail, promotes neuronal development. Genes Dev. *9*, 2384–2398.

Romani, S., Campuzano, S., and Modolell, J. (1987). The achaete-scute complex is expressed in neurogenic regions of Drosophila embryos. EMBO J. *6*, 2085–2092.

Roy, H.K., Iversen, P., Hart, J., Liu, Y., Koetsier, J.L., Kim, Y., Kunte, D.P., Madugula, M., Backman, V., and Wali, R.K. (2004). Down-regulation of SNAIL suppresses MIN mouse tumorigenesis: modulation of apoptosis, proliferation, and fractal dimension. Mol. Cancer Ther. *3*, 1159–1165.

Sandmann, T., Jensen, L.J., Jakobsen, J.S., Karzynski, M.M., Eichenlaub, M.P., Bork, P., and Furlong, E.E.M. (2006a). A temporal map of transcription factor activity: mef2 directly regulates target genes at all stages of muscle development. Dev. Cell *10*, 797–807.

Sandmann, T., Jakobsen, J.S., and Furlong, E.E.M. (2006b). ChIP-on-chip protocol for genome-wide analysis of transcription factor binding in Drosophila melanogaster embryos. Nat. Protoc. *1*, 2839–2855.

Sandmann, T., Girardot, C., Brehme, M., Tongprasit, W., Stolc, V., and Furlong, E.E.M. (2007). A core transcriptional network for early mesoderm development in Drosophila melanogaster. Genes Dev. *21*, 436–449.

Sarov, M., Schneider, S., Pozniakovski, A., Roguev, A., Ernst, S., Zhang, Y., Hyman, A.A., and Stewart, A.F. (2006). A recombineering pipeline for functional genomics applied to Caenorhabditis elegans. Nat. Methods *3*, 839–844.

Schäfer, G., Narasimha, M., Vogelsang, E., and Leptin, M. (2014). Cadherin switching during the formation and differentiation of the Drosophila mesoderm - implications for epithelial-to-mesenchymal transitions. J. Cell Sci. *127*, 1511–1522.

Shen, R., Fan, J.B., Campbell, D., Chang, W., Chen, J., Doucet, D., Yeakley, J., Bibikova, M., Garcia, E.W., McBride, C., et al. (2005). High-throughput SNP genotyping on universal bead arrays. Mutat. Res. - Fundam. Mol. Mech. Mutagen. *573*, 70–82.

Shields, M.A., Krantz, S.B., Bentrem, D.J., Dangi-Garimella, S., and Munshi, H.G. (2012). Interplay between β1-integrin and Rho signaling regulates differential scattering and motility of pancreatic cancer cells by snail and Slug proteins. J. Biol. Chem. *287*, 6218–6229.

Shlyueva, D., Stelzer, C., Gerlach, D., Yáñez-Cuna, J.O., Rath, M., Boryń, L.M., Arnold, C.D., and Stark, A. (2014a). Hormone-responsive enhancer-activity maps reveal predictive motifs, indirect repression, and targeting of closed chromatin. Mol. Cell *54*, 180–192.

Shlyueva, D., Stampfel, G., and Stark, A. (2014b). Transcriptional enhancers: from properties to genome-wide predictions. Nat. Rev. Genet. *15*, 272–286.

Short, J.M., Fernandez, J.M., Sorge, J.A., and Huse, W.D. (1988). ?? ZAP: A bacteriophage ?? expression vector with in vivo excision properties. Nucleic Acids Res. *16*, 7583–7600.

Simpson, P. (1983). Maternal-Zygotic Gene Interactions during Formation of the Dorsoventral Pattern in Drosophila Embryos. Genetics *105*, 615–632.

Skeath, J.B., and Carroll, S.B. (1992). Regulation of proneural gene expression and cell fate during neuroblast segregation in the Drosophila embryo. Development *114*, 939–946.

Smith, B., and Odero-Marah, V. (2012). The role of Snail in prostate cancer. Cell Adh. Migr. *6*, 433–441.

Sou, P.W., Delic, N.C., Halliday, G.M., and Lyons, J.G. (2010). Snail transcription factors in keratinocytes: Enough to make your skin crawl. Int. J. Biochem. Cell Biol. *42*, 1940–1944.

Southall, T.D., and Brand, A.H. (2009). Neural stem cell transcriptional networks highlight genes essential for nervous system development. EMBO J. *28*, 3799–3807.

Spitz, F., and Furlong, E.E.M. (2012). Transcription factors: from enhancer binding to developmental control. Nat. Rev. Genet. *13*, 613–626.

Van Steensel, B., and Henikoff, S. (2000). Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. Nat. Biotechnol. *18*, 424–428.

Stollewerk, A., Tautz, D., and Weller, M. (2003). Neurogenesis in the spider: New insights from comparative analysis of morphological processes and gene expression patterns. Arthropod Struct. Dev. *32*, 5–16.

Stormo, G.D. (2000). DNA binding sites: representation and discovery. Bioinformatics *16*, 16–23.

Takahashi, E., Funato, N., Higashihori, N., Hata, Y., Gridley, T., and Nakamura, M. (2004). Snail regulates p21(WAF/CIP1) expression in cooperation with E2A and Twist. Biochem. Biophys. Res. Commun. *325*, 1136–1144.

Tanaka-Matakatsu, M., Uemura, T., Oda, H., Takeichi, M., and Hayashi, S. (1996). Cadherin-mediated cell adhesion and cell motility in Drosophila trachea regulated by the transcription factor Escargot. Development *122*, 3697–3705.

Thomas, J.B., Bastiani, M.J., Bate, M., and Goodman, C.S. (1988). From grasshopper to Drosophila: a common plan for neuronal development. Nature *310*, 203–207.

Thorpe, H.M., and Smith, M.C. (1998). In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. Proc. Natl. Acad. Sci. U. S. A. *95*, 5505–5510.

Tran, K.D., and Doe, C.Q. (2008). Pdm and Castor close successive temporal identity windows in the NB3-1 lineage. Development *135*, 3491–3499.

Ungerer, P., Eriksson, B.J., and Stollewerk, A. (2011). Neurogenesis in the water flea Daphnia magna (Crustacea, Branchiopoda) suggests different mechanisms of neuroblast formation in insects and crustaceans. Dev. Biol. *357*, 42–52.

Usami, Y., Satake, S., Nakayama, F., Matsumoto, M., Ohnuma, K., Komori, T., Semba, S., Ito, A., and Yokozaki, H. (2008). Snail-associated epithelial-mesenchymal transition promotes oesophageal squamous cell carcinoma motility and progression. J. Pathol. *215*, 330–339.

Vega, S., Morales, A. V., Ocaña, O.H., Valdés, F., Fabregat, I., and Nieto, M.A. (2004). Snail blocks the cell cycle and confers resistance to cell death. Genes Dev. *18*, 1131–1143.

Vogel, M.J., Peric-Hupkes, D., and van Steensel, B. (2007). Detection of in vivo protein-DNA interactions using DamID in mammalian cells. Nat. Protoc. *2*, 1467–1478.

Voog, J., Sandall, S.L., Hime, G.R., Resende, L.P.F., Loza-Coll, M., Aslanian, A., Yates, J.R., Hunter, T., Fuller, M.T., and Jones, D.L. (2014). Escargot Restricts Niche Cell to Stem Cell Conversion in the Drosophila Testis. Cell Rep. *7*, 722–734.

Welch-Reardon, K.M., Ehsan, S.M., Wang, K., Wu, N., Newman, A.C., Romero-Lopez, M., Fong, A.H., George, S.C., Edwards, R.A., and Hughes, C.C.W. (2014). Angiogenic sprouting is regulated by endothelial cell expression of Slug. J. Cell Sci. *127*, 2017–2028.

Weller, M., and Tautz, D. (2003). Prospero and Snail expression during spider neurogenesis. Dev. Genes Evol. *213*, 554–566.

Whiteley, M., Noguchi, P.D., Sensabaugh, S.M., Odenwald, W.F., and Kassis, J.A. (1992). The Drosophila gene escargot encodes a zinc finger motif found in snail-related genes. Mech. Dev. *36*, 117–127.

Wieschaus, E., Nüsslein-Volhard, C., and Jürgens, G. (1984). Mutations affecting the pattern of the larval cuticle in Drosophila melanogaster. Wilhelm Roux's Arch. Dev. Biol. *193*, 296–307.

Wilczynski, B., and Furlong, E. (2010). Challenges for modeling global gene regulatory networks during development: Insights from Drosophila. Dev. Biol. *340*, 161–169.

Wilczynski, B., Liu, Y.H., Yeo, Z.X., and Furlong, E.E.M. (2012). Predicting Spatial and Temporal Gene Expression Using an Integrative Model of Transcription Factor Occupancy and Chromatin State. PLoS Comput. Biol. *8*, e1002798.

Wilczyński, B., and Furlong, E.E.M. (2010). Dynamic CRM occupancy reflects a temporal map of developmental progression. Mol. Syst. Biol. *6*, 383.

Wodarz, A., and Huttner, W.B. (2003). Asymmetric cell division during neurogenesis in Drosophila and vertebrates. Mech. Dev. *120*, 1297–1309.

Wodarz, A., Ramrath, A., Kuchinke, U., and Knust, E. (1999). Bazooka provides an apical cue for Inscuteable localization in Drosophila neuroblasts. Nature *402*, 544–547.

Wong, M., Castanon, I., and Baylies, M. (2008). Daughterless dictates Twist activity in a context-dependent manner during somatic myogenesis. Dev. Biol. *317*, 417–429.

Wu, Y., and Zhou, B.P. (2010). Snail: More than EMT. Cell Adh. Migr. *4*, 199–203.

Xu, W., Sun, Y., Zhang, J., Xu, K., Pan, L., He, L., Song, Y., Njunge, L., Xu, Z., Chiang, M.Y.M., et al. (2014). Perivascular Derived Stem Cells with Neural Crest Characteristics are Involved in Tendon Repair. Stem Cells Dev. 1–34.

Yagi, Y., and Hayashi, S. (1997). Role of the Drosophila EGF receptor in determination of the dorsoventral domains of escargot expression during primary neurogenesis. Genes Cells *2*, 41–53.

Yagi, Y., Suzuki, T., and Hayashi, S. (1998). Interaction between Drosophila EGF receptor and vnd determines three dorsoventral domains of the neuroectoderm. Development *125*, 3625–3633.

Yamasaki, H., Sekimoto, T., Ohkubo, T., Douchi, T., Nagata, Y., Ozawa, M., and Yoneda, Y. (2005). Zinc finger domain of Snail functions as a nuclear localization signal for importin beta-mediated nuclear import pathway. Genes Cells *10*, 455–464.

Yáñez-Cuna, J.O., Dinh, H.Q., Kvon, E.Z., Shlyueva, D., and Stark, A. (2012). Uncovering cis-regulatory sequence requirements for context-specific transcription factor binding. Genome Res. *22*, 2018–2030.

Zander, M.A., Cancino, G.I., Gridley, T., Kaplan, D.R., and Miller, F.D. (2014a). The Snail transcription factor regulates the numbers of neural precursor cells and newborn neurons throughout mammalian life. PLoS One *9*, e104767.

Zander, M.A., Burns, S.E., Yang, G., Kaplan, D.R., and Miller, F.D. (2014b). Snail coordinately regulates downstream pathways to control multiple aspects of mammalian neural precursor development. J. Neurosci. *34*, 5164–5175.

Zeitlinger, J., Simon, I., Harbison, C.T., Hannett, N.M., Volkert, T.L., Fink, G.R., and Young, R.A. (2003). Program-specific distribution of a transcription factor dependent on partner transcription factor and MAPK signaling. Cell *113*, 395–404.

Zeitlinger, J., Zinzen, R.P., Stark, A., Kellis, M., Zhang, H., Young, R.A., and Levine, M. (2007). Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the Drosophila embryo. Genes Dev. *21*, 385–390.

Zhang, K., Zhaos, J., Liu, X., Yan, B., Chen, D., Gao, Y., Hu, X., Liu, S., Zhang, D., and Zhou, C. (2011). Activation of NF-B upregulates Snail and consequent repression of E-cadherin in cholangiocarcinoma cell invasion. Hepatogastroenterology. *58*, 1–7.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Zhong, M., Niu, W., Lu, Z.J., Sarov, M., Murray, J.I., Janette, J., Raha, D., Sheaffer, K.L., Lam, H.Y.K., Preston, E., et al. (2010). Genome-wide identification of binding sites defines distinct functions for Caenorhabditis elegans PHA-4/FOXA in development and environmental response. PLoS Genet. *6*.

Zinzen, R.P., and Furlong, E.E.M. (2008). Divergence in cis-regulatory networks: taking the "species" out of cross-species analysis. Genome Biol. *9*, 240.

Zinzen, R.P., Senger, K., Levine, M., and Papatsenko, D. (2006). Computational models for neurogenic gene expression in the Drosophila embryo. Curr. Biol. *16*, 1358–1365.

Zinzen, R.P., Girardot, C., Gagneur, J., Braun, M., and Furlong, E.E.M. (2009). Combinatorial binding predicts spatio-temporal cis-regulatory activity. Nature *462*, 65–70.

7 Abstract

The family of Snail-like transcription factors is well-known for the regulation of epithelial-mesenchymal transition and its participation in the formation of various cancer tumors. The *Drosophila* genome contains three genes that belong to the family of Snail-like transcription factors, namely *snail*, *worniu* and *escargot*. At early stages of embryonic development *snail* is expressed in the mesoderm and acts as one of the key regulators of gastrulation. At later stages of embryonic development *snail, worniu* and *escargot* are expressed in the neuroblasts, the cells of the developing nervous system. The three transcription factors cooperate to regulate neurogenesis. The functions that Snail-like transcription factors fulfill in neuroblasts, as well as the mode of cooperation between the proteins remain obscure.

Chromatin immunoprecipitation followed by sequencing (ChIP-seq) as well as the comparative analysis of expression patterns was used to get a better understanding of the roles of Snail, Worniu and Escargot in neuroblasts and to elucidate the tissue-specific roles of Snail, in mesoderm and neuroblast development. In this work I characterized the successive expression of *snail* and *worniu* during formation of neuroblast cells. *Snail* mRNA is expressed in a short pulse before and during delamination of neuroblasts, while *worniu* expression starts at the end of delamination and persists throughout neuroblast's life. I identified a putative post-transcriptional regulatory mechanism, which assures the presence of Worniu protein exclusively in neuroblast cells that possess stem cell identity.

ChIP-seq analysis of Snail's binding in mesoderm and neuroblasts identified a genome-wide set of putative enhancers and genes that are regulated by Snail in a tissue-specific manner. In total 1703 mesoderm- and 2501 neuroblast-specific Snail-bound regions were identified. 866 regions that are bound by Snail in the mesoderm are also bound in neuroblasts. The established dataset forms the basis for the elucidation of tissue-specific enhancer architecture and the function of Snail in neuroblasts.

8   Zusammenfassung

Die Familie der Snail-ähnlichen Transkriptionsfaktoren ist bekannt für die Regulation der epithelial-mesenchymalen Transition und ihrer Beteiligung in der Entstehung verschiedener Krebstumoren. Das *Drosophiila*-Genom beinhaltet drei Gene, welche zur Familie der Snail-ähnlichen Transkriptionsfaktoren gehören: *snail*, *worniu* und *escargot*. In frühen Phasen der Embryonalentwicklung wird *snail* im Mesoderm exprimiert und ist einer der Hauptregulatoren der Gastrulation. In späteren Phasen der Embryonalentwicklung werden *snail*, *worniu* und *escargot* in den Neuroblasten, den Zellen des sich entwickelnden Nervensystems, exprimiert. Diese drei Transkriptionsfaktoren wirken zusammen an der Regulation der Neurogenese. Die Funktionen, welche Snail-ähnliche Transkriptionsfaktoren in Neuroblasten erfüllen, sowie die Art des Zusammenwirkens der drei Transkriptionsfaktoren sind bisher unbekannt.

Die Chromatin-Immunopräzipitation mit anschließender Sequenzierung (ChIP-seq) sowie der Vergleich der Expressionsmuster wurden für ein besseres Verständnis der Funktion von Snail, Worniu und Escargot in Neuroblasten und für die Aufklärung der gewebsspezifischen Funktion von Snail in Mesoderm- und Neuroblastenentwicklung herangezogen. In dieser Arbeit habe ich die aufeinanderfolgende Expression von *snail* und *worniu* während der Bildung von Neuroblasten charakterisiert. *Snail*-mRNA wird für einen kurzen Zeitraum vor und während der Delamination von Neuroblasten exprimiert, während die Expression von *worniu* am Ende der Delamination beginnt und während der kompletten Lebenszeit des Neuroblasts anhält. Ein vermeintlicher posttranskriptioneller Regulationsmechanismus wurde identifiziert, welcher das Vorhandensein von Worniu-Protein ausschließlich in Neuroblasten mit Stammzellidentität sicherstellt.

Die ChIP-seq-Analyse der DNA-Bindung von Snail im Mesoderm und in Neuroblasten führte zur Identifikation einer genomweiten Gruppe vermeintlicher Enhancer und Genen, die durch Snail spezifisch reguliert werden. Insgesamt wurden 1703 Mesoderm- und 2501 Neuroblasten-spezifische Snail-Bindestellen identifiziert. 866 Regionen, die von Snail im Mesoderm gebunden werden finden sich auch in Neuroblasten. Das erhaltene Datenset bildet eine Basis für die weitere Aufklärung der gewebsspezifischen Enhancer-Struktur und der Funktion von Snail in Neuroblasten.

9 Acknowledgements

10 Declaration

Hereby I declare that I elaborate the entire PhD thesis on my own and that I used only the references listed.

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzen Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildingen -, die anderen Werken im Wortlaut oder im Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist, sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation wurde durch Prof. Dr. Maria Leptin betreut.

Köln, 9. Mai 2015

11 Lebenslauf

Volodymyr Averkov

Володимир Аверков

Persönliche Daten

Name	Volodymyr Averkov
Anschrift	Lochnerstr. 7; 50674 Köln
Tel.	0049 1575 1752755
E-Mail	vldaverkov@gmail.com
Geboren	Charkiw, Ukraine, am 10.09.1988
Staatsangehörigkeit	Ukrainisch

Hochschulausbildung

2011 – heute	<u>Promotion</u> in Molekulargenetik; Institut für Genetik, Universität zu Köln, Deutschland Die Projektbezeichnung: "Tissue-specific roles of Snail-like transcription factors in the embryogenesis of *Drosophila melanogaster*" voraussichtlicher Abschluss August 2015
2008 – 2010	<u>Weiterbildung</u> in Molekularbiologie; Bachelor- und Masterthesis. Institut fürProteinforschung (Russische Akademie der Wissenschaften), Puschtschino Filiale der Staatlichen Lomonossow-Universität Moskow, Stadt Puschtschino, Oblast Moskau, Russland
2005 – 2009	<u>Bachelor (B.Sc.)</u> in Biochemie; Nationale Wassyl-Karasin-Universität Charkiw
2009 – 2010	<u>Master (M.Sc.)</u> in Biochemie; Nationale Wassyl-Karasin-Universität Charkiw

Schulausbildung

2005	Abschluss: „Zertifikat der höheren Schulausbildung"
1998 – 2005	Sekunderstufe: Charkiw Physical and Mathematical Lyceum No.27, Charkiw Ukraine;
1995 – 1998	Primarstufe: Gründschule No.46, Charkiw, Ukraine

Staatsstipendium

2010 – 2013 <u>Doktorandenstipendium</u> der *International Graduate School for Development, Health and Disease,* Ministerium für Innovation, Wissenschaft und Forschung des Landes Nordrhein-Westfalen

Volodymyr Averkov

Köln, Deutschland

Den 9. Mai 2015