Analysis of border cell migration using systematic gain-of-function studies in *Drosophila melanogaster*

Ph.D. Thesis

Szabó Kornélia Ágnes

Supervisor: Dr. Pernille Rorth

Hungarian Academy of Sciences
Biological Research Center
Genetics Department

Szeged, 2003

1. BRIEF INTRODUCTION AND OBJECTIVES

Insertional mutagenesis is a method that allows the experimental modification of selected genomes using tansposons as mutagens (Cooley, Berg, et al. 1989). These kinds of experiments have been carried out successfully in model organisms such as bacteria (Kleckner, Roth, et al. 1977), yeast (Burns, Grimwade, et al. 1994; Garraway, Tosi, et al. 1997), *Arabidopsis* (Sundaresan, Springer, et al. 1995; Bhatt, Page, et al. 1996; Smith, Yanai, et al. 1996), *Drosophila* (Cooley, Kelley, et al. 1988; Torok, Tick, et al. 1993; Rorth 1996), zebrafish (Allende, Amsterdam, et al. 1996; Gaiano, Amsterdam, et al. 1996), and mice (Jaenisch 1988; Gossler, Joyner, et al. 1989; Wurst, Rossant, et al. 1995; {Zambrowicz, Friedrich, et al. 1998).

In *Drosophila* the most frequently used mobile element is the P element, which is one of the best-characterized eucaryotic transposons. Over the years it became a particularly useful genetic tool because it moves with a high frequency and can be controlled by limiting the availability of the element-encoded transposase, the enzyme that is required for this process.

Since the P elements were discovered, much of the knowledge of developmental processes in the fly has been obtained via studying these transposon generated loss-of-function genetic perturbations. These kinds of mutations are produced when the mobile element jumps into a gene, physically disrupts it, and reduces or abolishes its activity.

Single-insert lines turned out to have a long-term value as genetic resources for analysing the *Drosophila* genome. The Berkeley Drosophila Genome Project is currently cataloguing these lines. The ultimate goal is to generate a "library" of stable, single insert strains that together define all the genes in the fruitfly. P elements were previously estimated, however, to mutate only about 30% of all genes, which means

that alternative approaches are needed in order to mutate every single ORFs in the fly genome.

In the 1990's the UAS-Gal4 system was developed for ectopic gene expression in *Drosophila*, which allows the selective activation of certain genes or artificial constructs in a spatially and temporarily tightly regulated manner. This way one can produce gain-of-function phenotypes, by e.g. increasing the activity of a gene, or making it active in inappropriate circumstances.

Gal4 is a sequence-specific transcriptional transactivator in yeast that can activate tissue-specific conditional expression of particular target genes in *Drosophila* (Brand, Perrimon 1993).

The gain-of-function screening system that was developed by P. Rorth combines the advantages of the classical P element based mutagenesis method, and the Gal4 system. Using it we can easily generate thousands of mutant fly stocks, and the phenotypes are generated by the tissue-specific misexpression of random genes using the UAS-Gal4 system (Rorth 1996). The modified transposon that was generated to perform these experiments is P element based, and named as EP element. It contains special enhancer and promoter sequences that can force ectopic expression or misexpression of the genomic sequences right next to the transposon.

In this work I describe a screen we performed using the EP construct, in order to generate a mutant collection containing random, completely unselected transposon insertions on the X, second and third chromosomes of the fruitfly. I will explain the way this screen was done, and also the results of this mutagenesis experiment (Rorth, Szabo, et al. 1998).

The system, and the EP collection were developed in order to use it to study our favourite biological problem, a special cell migration process that happens during Drosophila oogenesis.

The newly forming cysts after leaving the germarium containing two different cell types based on their origin. The 15 nurse cells and the oocyte that have germline origin are surrounded by the monolayer of follicle cells with somatic origin. During mid stages of the oogenesis the follicle cells start to slide on the surface of the cyst posteriorly. In the meantime 6-10 cells from the anterior pole delaminate from the others, and move toward the oocyte but in between the nurse cells. They move until they reach the border between the nurse cells and the oocyte. That is why they are called border cells (King 1969; Spradling 1993). This process has to happen in a spatially and temporally tightly regulated manner, or else the resulting eggs will be defective, and they cannot be fertilized. The females that lay these eggs are sterile.

Certain mutations can disturb this migration process. It has been shown previously that in *slbo* mutant ovaries the border cells migrate late or don't migrate at all depending on the strength of the allele used in the experiment (Montell, Rorth, et al. 1992). The *slbo* (*slow border cells*) locus encodes the *Drosophila C/EBP*, which is a member of the basic region-leucine zipper transcription factor family. We knew, that *C/EBP* is expressed in the migrating border cells where its presence is absolutely necessary. The *slbo* phenotype and the fact that *C/EBP* is a transcription factor suggested that its most likely function is the regulation of certain genes with a possible role during border cell migration. To investigate this possibility we decided to set up different experiments using the EP collection in order to study the border cell migration process.

Controlled overexpression of certain genes can identify important genetic interactions (Fortini, Artavanis-Tsakonas 1994). If increased expression of one gene enhances or suppresses a certain phenotype caused by a mutation of another one, their products are likely to be involved in the same process. This notion is the basis of powerful genetic approaches for example in yeast, and also the main idea behind the so-called interaction screens that we carried out. We decided to mis- or overexpress target genes specifically the migrating border cells on a special, sensitised genetic background. We tried to find answers to our questions, namely identify genes with a role in this process. In the screen itself we were looking for modifiers (enhancers and suppressors) of our pre-existing phenotype.

We overexpressed the panel of EP lines in the migrating border cells in the presence of the *slbo*^{1} mutation. Females carrying the *slbo*^{1} allele are sterile, because the border cells cannot migrate properly in these follicles. We looked for slbo suppressors this way.

We analysed the 60 putative suppressors our screen resulted. In various test experiments showed that the suppression effect was caused by the forced overexpression of the different target genes. We also determined the affected genes by searching the Drosophila sequence database with the genomic sequences we obtained.

One of the slbo suppressors, EP(X)1487, turned out to be a *Drosophila* homologue of the human *RIN1*. We named the *Drosophila* homologue *sprint.* In this work I describe the cloning of this gene, the phenotypic and molecular analysis of it. Biochemical data gathered by analysing the hRIN1 gene, and our experimental results suggest, that sprint might serve as a molecule that liks together signal transduction, and regulation of the actin cytoskeleton.

2. INTRODUCTION

The great tragedy of Science-
The slaying of a beautiful hypothesis
By an ugly fact.

-T. H. Huxley

2.1.1 Transposons

More than 50 years ago, geneticist Barbara McClintock rocked the scientific community with her discovery that maize contained mobile genetic elements, bits of DNA that move about the genome, often causing mutations if they happen to land in functioning genes. Her findings were considered so outlandish that they were at first dismissed as anomalies unique to corn. It was only in the 1970's when the molecular basis of transposition was first appreciated, and subsequently the existence of mobile genetic elements became widely accepted. Over the years, transposons, as the mobile elements are called, have proved to be nearly universal. They have turned up in species ranging from bacteria to mammals, where their movements have been linked to a variety of mutations, including some that cause diseases and others that add desirable diversity to genomes (Moffat, 2000).

Today we know that transposable elements are short DNA sequences that replicate and are inserted around both the prokaryotic and eukaryotic genome. The movement process is called transposition; it can be performed by proteins coded by the transposable element itself, or by proteins coded by genes elsewhere in the genome (P.Winter, I.Hickey, et al. 1998).

Members of the former group are the so-called autonomous elements; they are able to move in different host cells. Non-autonomous elements on the other hand belong to the latter group, they are unable to transpose without the help of certain host-specific cellular factors (B.Alberts, D.Bray, et al. 2002).

Mobile elements have a number of effects on the host genome. Although they move very rarely, there are many types of them, and altogether their movement has a major effect on the variability of the species. More than half of the spontaneous

mutations, for example, examined in Drosophila are caused by transposable elements. Mutations can occur when an element inserts into a gene or when it exits to move somewhere else. One of the most important effects mobile elements can have in the host genome is that they provide lesions in the genes where they insert, physically disrupt them, and cause mutations this way. All known transposable elements cause short target site duplication because of their mechanisms of insertion, and when they exit, they generally leave behind part of this duplication, often with other local sequence changes as well. They can also have a contribution to genome diversity in different ways. When two transposable elements that are jumping using the same transposase integrates into close chromosomal sites, they can jump together and remove the DNA that was originally between them. This provides a particularly effective pathway for the duplication or movement of exons called exon shuffling. This process can lead to the formation of new genes, and have an evolutionary importance. If the transposon moves to a site immediately upstream of a cellular gene, then it might disrupt the efficiency with which DNA-binding proteins are attached to upstream regulatory sequences. Transcription of the affected gene might also be influenced by the presence of promoters and/or enhancers within the transposon, so it becomes subjected to an entirely different regulatory regime. There are also examples where the insertion of a transposon into a gene causes altered splicing patterns. They can also be highly deleterious because of insertion into the coding region of a gene resulting in its inactivation (T.A.Brown 2002).

Mobile elements in *Drosophila* and other eukaryotes make up substantial part of the host genomes. They may be divided into two classes according to their structure and mechanisms of transposition (Figure 1). Class I elements transpose by reverse

transcription of an RNA intermediate, while class II elements move directly from DNA to DNA (B.Alberts, D.Bray, et al. 2002).

The best studied class I elements are those that are structurally similar to integrated retroviruses. These include copia-like elements in *Drosophila*, Ty elements in *Saccharomyces_cerevisiae* and many others. They are characterized by having long direct repeats (LTRs) at their termini, and contain long open reading frames, one of which encodes a reverse transcriptase enzyme. Examples of the second type of class I elements are structurally quite different from retroviruses. These include e.g. L1 elements of mammalian genomes, I, F, G and jockey elements of *Drosophila*. They have no terminal repeats, and have a characteristic A-rich sequence at the 3' end. They also contain long ORFs and reverse transcriptase gene.

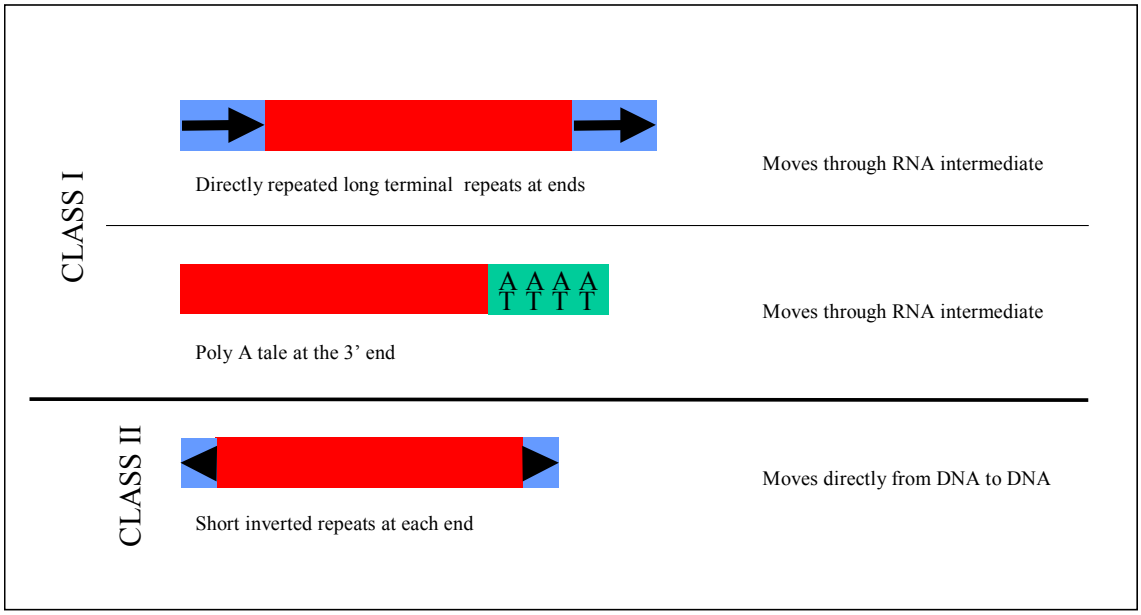

Figure 1. Three major families of transposable elements, based on their structure and the mechanism of transposition

The class II elements are those that transpose by DNA-DNA mediated mechanisms. These include Tn3 and IS1 elements in *E. coli*, P and hobo elements in *Drosophila melanogaster*, Ac/Ds elements in *Zea mays*, Tc1 elements in *Caenorhabditis elegans*, all of which have short inverted repeats at their termini. They encode a transposase enzyme that is required for their movement (B.Alberts, D.Bray, et al. 2002).

2.1.2 Transposons in research

Large amounts of data including whole genome DNA sequences are rapidly accumulating in publicly available databases. It is widely accepted, that even the most complex (to date), increasingly powerful computer analysis methods of integrated databases are insufficient to take us directly from DNA sequence to biological function but, on the other hand, they provide an important foundation for the design of appropriate experiments. Data derived from detailed work on specific genes of model organisms (like *Drosophila melanogaster*, *Caenorhabditis elegans*, *Danio rerio*, *Mus musculus*, *Arabidopsis thaliana*, and *Saccharomyces cerevisiae*) in conjunction with total genome sequences and other database information should significantly enhance our ability to unravel the multilayered networks controlling the different biological processes. This rapidly obtainable knowledge in different model organisms will allow the reduction of most of the ~40,000 individual human genes into a much smaller number of multicomponent core processes with already known biochemical function (Miklos, Rubin 1996).

How is it possible that work e.g. on a small fly can provide us with valuable information concerning even man? Multicellular organisms are proving to have far more in common at the biological level than previously suspected. They share many genes with each other, whose sequences and functions have been conserved. Research on the most tractable models can greatly advance our understanding of what specific genes do, including many that are directly relevant to human biology and medicine. Unlike in humans, any open reading frame within the model organism's genome can be mutated and subjected to detailed functional analysis within the context of an intact organism.

The ability to use transposons as mutagens made it possible to systematically identify new genes by insertional mutagenesis (Cooley, Berg, et al. 1989) in order to experimentally modify selected genomes. Insertional mutagenesis experiments have been carried out successfully in bacteria (Kleckner, Roth, et al. 1977), yeast (Burns, Grimwade, et al. 1994; Garraway, Tosi, et al. 1997), *Arabidopsis* (Sundaresan, Springer, et al. 1995; Bhatt, Page, et al. 1996; Smith, Yanai, et al. 1996), *C. elegans* (Plasterk 1992; Korswagen, Durbin, et al. 1996), *Drosophila* (Cooley, Kelley, et al. 1988; Torok, Tick, et al. 1993; Rorth 1996), zebrafish (Allende, Amsterdam, et al. 1996; Gaiano, Amsterdam, et al. 1996), and mice (Jaenisch 1988; Gossler, Joyner, et al. 1989; Wurst, Rossant, et al. 1995; Zambrowicz, Friedrich, et al. 1998). They allow us to correlate genetic and molecular information because they generate simple, reproducible lesions upon insertion that we can detect much more easily than damages produced by other mutagenes (e.g. chemical mutagenes or ionizing radiation).

Single-insert lines have a long-term value as genetic resources for analysing the *Drosophila* genome. The Berkeley Drosophila Genome Project is currently cataloguing these kinds of lines. The ultimate goal is to generate a "library" of stable, single insert

strains that together define all the genes in the fruitfly (Spradling, Stern, et al. 1995; Spradling, Stern, et al. 1999).

2.1.3 The P element of the *Drosophila melanogaster*

The *Drosophila* P element is one of the best-characterized eucaryotic transposon. It became a particularly useful genetic tool because it moves with high frequency that can be controlled by limiting the availability of the element-encoded transposase. The naturally occurring full-length versions are 2.9 kb in length and encode an 87 kD transposase protein. This is a special site-specific recombinase that mediates the DNA cleavage and strand transfer during transposition. Other factor that is required for the movement is a special ~150 bp sequence at each ends of the P element. These sequences include 31 bp terminal inverted repeats, internal transposase binding sites, and internal 11 bp inverted repeats. Both the functional transposase enzyme and the intact terminal sequences are required for successful transposition (Beall, Rio 1997).

Researchers started to use P elements as insertional mutagens in the 1980's in so-called "hybrid-dysgenic" crosses, when one crosses P strains, isolated from the wild, containing 30 to 50 complete and defective P elements to laboratory females lacking any of these mobile elements (D.B.Roberts 1986). This approach turned out to be very limited, because the newly induced mutations were very unstable, they were frequently lost, and it was difficult to determine which element was responsible for the observed phenotype.

In 1988 Cooley et al. showed that individual, experimentally modified P elements could be mobilized in large genetic screens to generate thousands of stable mutants (Cooley, Kelley, et al. 1988). In the first generation of these engineered

elements the transposase gene was replaced with some kind of marker gene. Among these were genes causing resistance to chemicals (neo^R gene, neomycin selection (Cooley, Kelley, et al. 1988); organophosphate degrading gene, paraoxon selection (Benedict, Salazar, et al. 1995)), or inducing eye color changes (mini-white (Klemenz, Weber, et al. 1987), rosy (Rubin, Spradling 1982)) allowing easy detection of the presence of different constructs from one generation to the next. This also means that scientists had to find a way to provide somehow the lost transposase enzyme in order to mobilize the element, so they constructed transgenic transposase-producing fly stocks (Cooley, Kelley, et al. 1988). These changes subsequently led to greater control over the movement of the mobile element. The new design facilitated the cloning of genomic DNA flanking the insertions and allowing rapid isolation of the affected genes (Cooley, Kelley, et al. 1988). Together these advantages were enough to make P elements one of the most widely used mutagens in *Drosophila*.

At the end of 1980's a new generation of P-element vectors emerged. They were able to identify cis-acting patterning information within the *Drosophila* genome. They contained a *lacZ* reporter gene that could be expressed under the control of local endogenous enhancers, thereby their detectable expression pattern reflected the expression associated with the genes at the site of insertion (Bellen, O'Kane, et al. 1989; Bier, Vaessin, et al. 1989; Wilson, Pearson, et al. 1989).

Nowadays several different variations of this "old theme" are known. The "gene-trap" system (GT), for example, allows the recovery only of fly lines whose genes are inactivated by a P-element insertion, i.e., mutants. In the gene-trap system, the reporter gene expression reflects precisely the spatial and temporal expression pattern of the trapped gene. Flies in which gene trap occurred are identified by a two-step screening process using two independent markers, mini-white and Gal4, each indicating the

integration of the vector downstream of the promoter of a gene (dual tagging). mini-white has its own promoter but lacks a polyadenylation signal. Therefore, mini-white mRNA is transcribed from its own promoter regardless of the vector integration site in the genome. However, the eyes of flies are not orange or red unless the vector is incorporated into a gene enabling mini-white to be spliced to a downstream exon of the host gene and polyadenylated at the 3' end. The promoter-less Gal4 reporter is expressed as a fusion mRNA only when it is integrated downstream of the promoter of a host gene. (Lukacsovich, Yamamoto 2001). The users of the GT element can save the time they would use for selecting out valuable mutants from stocks carrying neutral insertions.

The generation of gain-of-function phenotypes in specialized P element screens by targeted misexpression of different genes provides also a powerful complement to the classical genetic approach based on loss-of-function mutations (Rorth, Szabo, et al. 1998; Toba, Ohsako, et al. 1999) (see further chapters for more details).

2.1.4 Alternative ways of generating mutations

Why do we need newer, more sophisticated P element constructs and methods in order to generate new mutants? P elements were previously estimated to mutate only about 30% of all genes, which means that alternative approaches are needed in order to mutate every single ORFs in the fly genome. There are several different ways how we can possibly overcome this problem (Spradling, Stern, et al. 1999).

It may be possible to disrupt the remaining genes by mobilizing modified P elements with different insertional specificity, although it is still not absolutely clear

what kind of sequences or mechanisms are responsible for that process. Both natural and genetically engineered versions of the P elements have been studied extensively, and it is known that neither inserts randomly throughout the genome, and the insertion is independent of homology between the ends of the elements and the target site. For most P element derivative, this target preference is very broad (although little sequence specificity has been noted). It is possible that this nonrandom target selection is an intrinsic property of the P element transposition mechanism. Since a clear-cut consensus nucleotide sequence is not present at sites of insertion, the molecular explanation for local preferences are not completely clear. However, there is an apparent preference for chromosomal regions that are likely to be accessible in chromatin; euchromatic sites are favoured over heterochromatic sites, interbands appeared to be favoured over bands, and there is a marked tendency to integrate at the 5' end of genes. The most common integration sites are located in the 5’ untranslated region, and a few hundred base pair upstream of the site where transcription initiates (Berg, Spradling 1991; Liao, Rehm, et al. 2000).

Other possibility is to employ different transposable elements with different target specificity whose movements can be controlled similarly to P elements. There were attempts to use e.g. hobo (Kiss Istvan personal communication) and mariner (Garza, Medhora, et al. 1991) elements as a possible alternative. Unfortunately, the jumping rate of these elements are much lower than that of the P element, which makes these screens insufficient, time- and work-consuming. Another drawback in case of the hobo element is that most of the checked common laboratory stocks contain wild type hobo transposons making it very hard (almost impossible) to use the isolated insertional mutants in a wide variety of genetic experiments.

Alternatively, it can also provide solution for the problem if one approaches it from a completely different way e.g. by creating gain-of-function mutations instead of the loss-of-function ones (Garza, Medhora, et al. 1991; Miklos, Rubin 1996).

2.1.5 Generating gain-of-function mutations

In the 1990's a new system, named Gal4 system, was developed for ectopic gene expression in *Drosophila.* It allows the selective activation of certain genes or artificial constructs in a spatially and temporarily tightly regulated manner (Brand, Perrimon 1993). In addition to producing gain-of-function phenotypes, the Gal4 system can be used to interfere with gene expression by transcription of antisense RNA or expression of dominant negative forms of a protein. Other applications of this system include target cell ablation by cell specific expression of cellular toxins or cell death genes (edited by: Sullivan, Ashburner, et al. 2000), or the expression of cellular proteins that enable visualization of cell structure and function (Brand, Perrimon 1993; Phelps, Brand 1998).

Gal4 is a sequence-specific transcriptional transactivator in yeast that can activate tissue-specific conditional expression of particular target genes in *Drosophila*. Gal4 has no known endogen targets in flies, and can be expressed in various cells and tissues of the organism without an observable effect.

The Gal4 system based gain-of-function screening method makes use of the previously observed fact that many genes can cause severe developmental defects when overexpressed. The scheme (Figure 2) consists of the transcriptional activator and its inducible target gene, which are separated into two discrete transgenic lines (Rorth 1996).

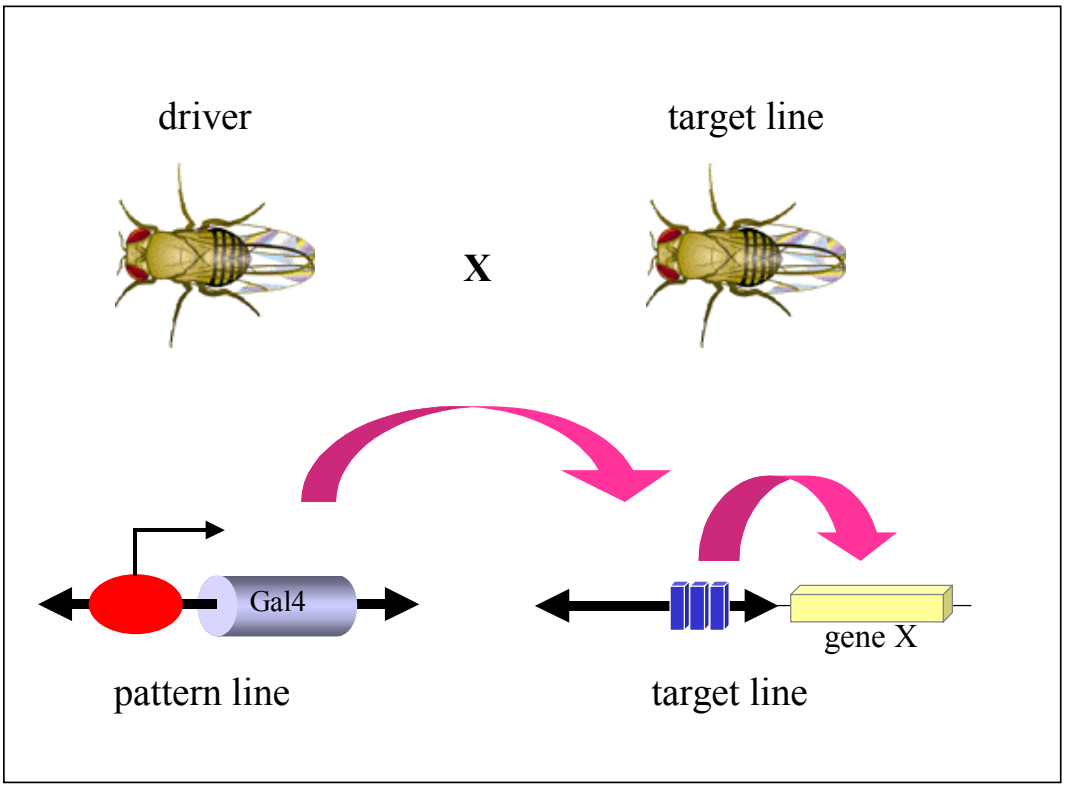

Figure 2. Basic outline of the UAS-Gal4 system; where the transcriptional activator and its inducible target are separated into two discrete transgenic lines. This ensures that the desired phenotype is only generated when the two lines are crossed together

This ensures that the parental lines are viable; in one line the activator protein is present, but it has no target gene to activate, and in a second line the target gene is silent in the absence of the activator. When the two lines are crossed, the target gene is turned on in the progeny where both the activator and the target are present. This allows spatially and temporarily regulated mis- or overexpression of the desired genes (Figure 2).

To study the role of several different genes in a specific tissue, a single Gal4-expressing line can be crossed to an array of different UAS-target gene lines. Conversely, to study the function of a single gene in different cell and tissue types, one

UAS-target line can be crossed to a diverse array of lines expressing Gal4 in distinct cell and tissue-specific patterns (Figure 3).

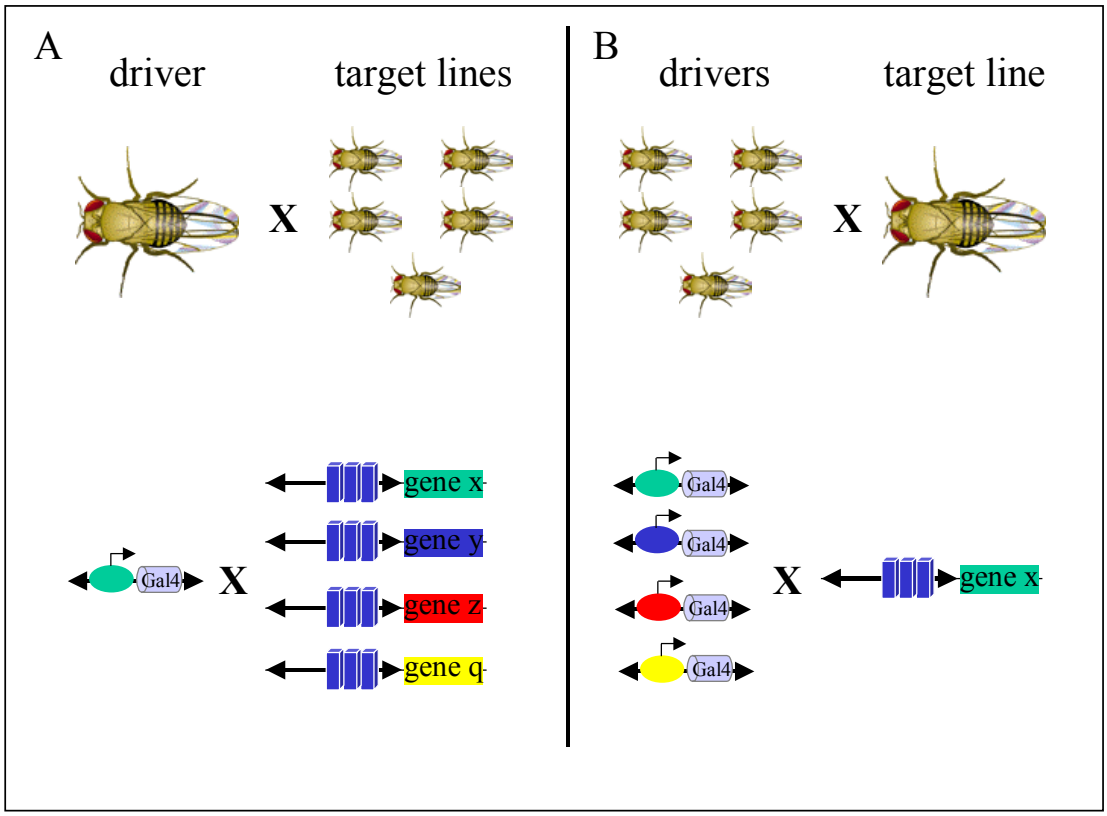

Figure 3. A The same driver can be crossed to a set of target lines in order to study the effect of different genes in the same process. **B** Alternatively, one can also study the effect of one particular gene in different processes by crossing the appropriate target line to different drivers.

2.1.6 Comparing the gain-of-function versus the loss-of-function methods

Much of the knowledge of developmental processes in the fly, worm, mouse, and zebrafish and of the cell biology of yeast has been obtained via loss of function

genetic perturbations (Nusslein-Volhard 1994; Mullins, Hammerschmidt, et al. 1994; Burns, Grimwade, et al. 1994; Brandon, Idzerda, et al. 1995). This information was particularly useful in elucidating complex biological pathways such as the cell cycle (Hartwell 1991), early pattern formation of *Drosophila* embryo, or the formation of the peripheral nervous system. It is known, however, that the classical loss-of-function approach has its limitations, and it cannot provide answers to every single biological problem. What are those limitations?

The number of transcripts appears to exceed the number of mutable loci in several extensively studied segments of the *Drosophila* genome. In yeast and *C. elegans,* the majority of ORFs revealed by genomic sequencing have not been associated with mutant phenotypes even in genomic regions that have been subjected to saturation mutagenesis. This can happen, for example, if the effects we would cause by mutating these "silent" genes are masked by the activity of redundantly related genes. Unrelated genes with similar functions can also take over, and parallel signal transduction pathways can substitute for the missing one (Spradling, Stern, et al. 1995).

Sometimes the resulting loss-of-function mutations are just not readily apparent for casual observation under normal laboratory conditions. In many cases a phenotype in fact can be detected by using more sophisticated and specialized assays.

Even when there is a phenotype, it is not always possible to fully comprehend the phenotypic consequences of a given mutation in multicellular organisms. For that, one has to have a deep knowledge of the different cell types, developmental stages, and cellular processes in which the gene functions.

The classical genetic studies in *Drosophila* and in *Mus* revealed that certain genes affected many aspects of the resulting phenotype. In *Drosophila* pleiotrophy is also a rule rather than an exception. This can cause problems, because sometimes it is

quite difficult to study the phenotypic consequences of a given loss-of-function mutation for example in the development of an adult organ, if the mutation also causes embryonic lethality and kills the animal long before it would reach adulthood.

Such limitations can be partially circumvented by screens that are based on analysing the phenotypes of clones of mutant tissues generated by somatic recombination (Xu, Rubin 1993), or by screens for enhancers or suppressors of a particular mutant phenotype (Simon, Bowtell, et al. 1991). Nevertheless, many important genes might have escaped detection by loss-of-function approach alone.

The gain-of-function screening system complements the classical genetic approach. This type of screen is based on the analysis of phenotypes generated by tissue-specific misexpression of genes using the UAS-Gal4 system (Brand, Perrimon 1993).

One potential drawback of gain-of-function screens is that misexpression of a gene may have an influence on the development of tissues in which that gene is not normally expressed. In some cases it may ectopically affect a signalling pathway that functions in multiple developmental processes. Another concern is that phenotypes may be artificial. For example, the phenotype caused by misexpression of a gene at levels much higher than normal may interfere with the development of an organism even if that gene does not have any function during normal development. This can happen because the overexpressed protein can titrate out other ones important for normal cell functions, enhance or reduce the activity of other proteins by interaction, or it can cause completely non-specific events. It is always important to make sure that the observed phenotype caused by the mis- or overexpression of certain genes is real by looking at the endogenous expression patterns and the loss-of-function phenotype of the affected gene, and genetic interactions with other known components of the studied process.

Taken all these into consideration, the gain-of-function approach seems to be a useful complement to the loss-of-function mutation-based genetic studies, and it can help to overcome some imperfections of the classical method.

2.1.7 EP element

In the early 1990's Rorth developed a novel modular misexpression system combining the benefits of both the Gal4 system, and the classical P element mutagenesis screens (Rorth 1996). The technique allows rapid generation of individual strains in which ectopic expression of random, unselected genes can be directed to different tissues or cell types.

The target what is used in this scheme is called EP, which is an **e**nhancer and **p**romoter-containing specialized P element derivative (Figure 4). It has P element ends (that are necessary for the transposition) making it possible to mobilize the element by simply providing the transposase enzyme, and also a mini-white marker gene allowing the easy selection of flies carrying the construct in their genome. Apart from this at the 3' end there is an enhancer with 14 UAS binding sites (upstream activator sequence) for Gal4 molecules that activate transcription of adjacent sequences regardless of the position of the elements in the genome. The EP element has a basal promoter from the hsp70 gene immediately adjacent to enhancer, so Gal4 activation should induce a transcript starting within the element and activate transcription of an adjacent gene.

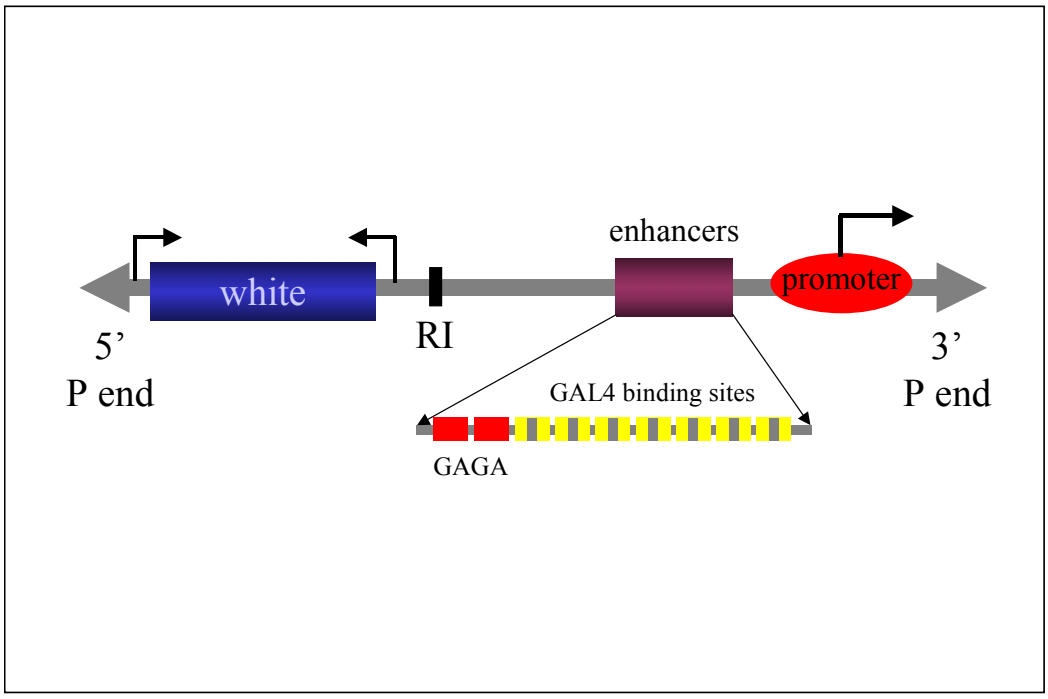

Figure 4. The structure of the EP element

The EP target element can be introduced into the fly genome with P element transformation and mobilized *in vivo* by exposure to a stable source of transposase.

2.2.1 Oogenesis

In the laboratory of Pernille Rorth we generated a collection of EP element containing lines, in order to study a special cell migration process during Drosophila oogenesis. Before I describe the details of the mutant isolation experiment, and the ways the different screens were done, let me introduce the reader the female reproductory system, and the major events that happen during the fly oogenesis.

Patterning and constructing of a new multicellular organism is one of the most fascinating biological processe. It occurs via precise ordering of events in space and time. The entire process is genetically determined; every organism develops according to an inherent pattern generation after generation. Tracking the development of a multicellular organism back in time, one eventually reaches an initial stage, the fertilized egg, yielded by the fusion of the two gametes, the oocyte, and the sperm.

The internal reproductory system of the adult female *Drosophila melanogaster* is located in the abdomen, and comprised of a pair of ovaries, the efferent genital duct system and its accessory organs and glands (see Figure 8).

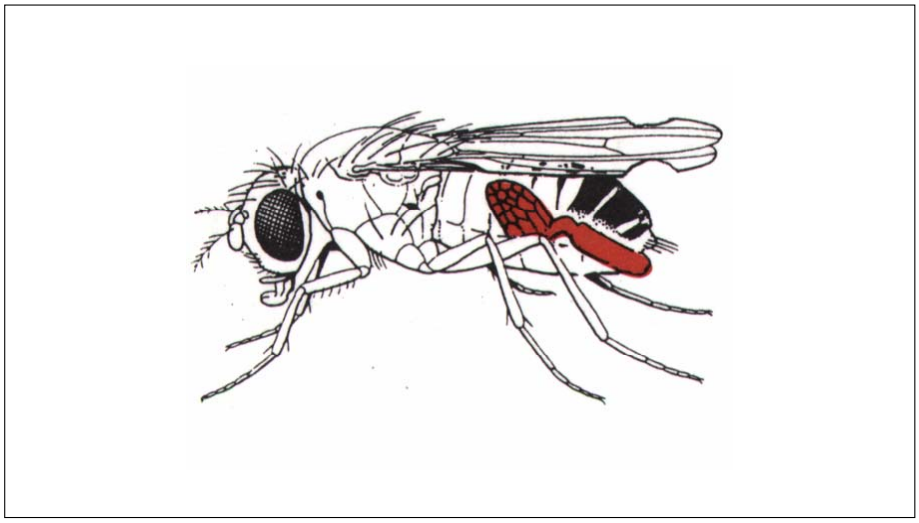

Figure 8: The internal reproductory system (shown in red) of the adult female ***Drosophila melanogaster*** is located in their abdomen.

A typical ovary contains about 16 ovarioles, each representing an independent egg assembly line. The ovarioles open into the lateral oviducts, which unite medially to form a common oviduct. Contractions of the oviduct force the eggs into the uterus, where the fertilization takes place. The oviduct runs into the vulva, which is the opening

on the female body. Eggs are released, and also copulation takes place through this structure.

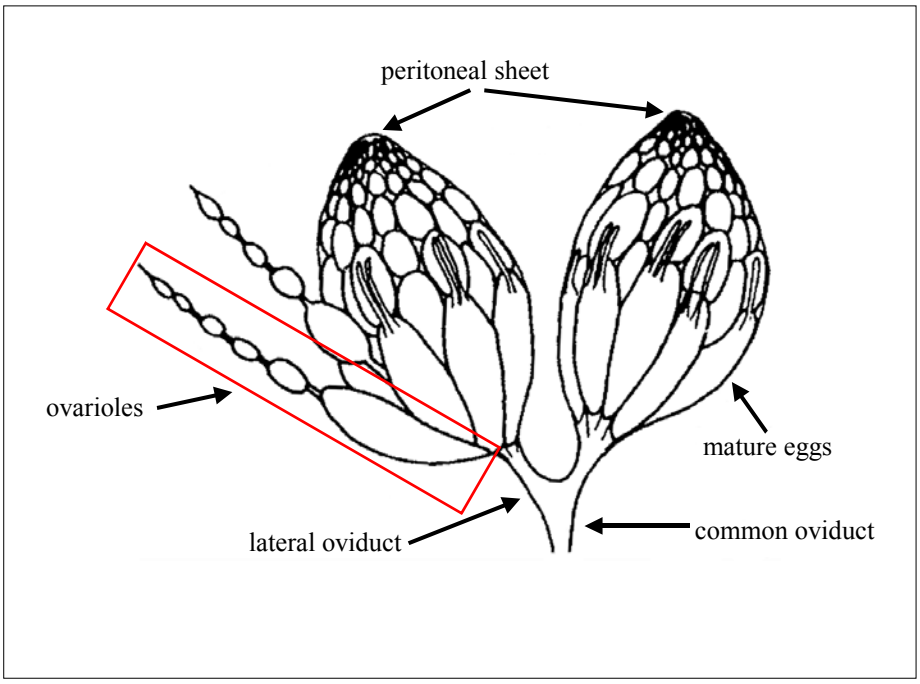

Figure 9: The Drosophila ovary is divided into two parts. They both covered by a peritoneal sheet. Each half contains about 16 ovarioles. They open into the lateral oviducts, which unite medially to form a common oviduct

One can find also different accessory glands, spermathecae, and seminal receptacle attached to the oviduct. These organs have a major role in storing the sperm after copulation, and assisting the fertilization.

Every ovariole can be divided into two parts, the germarium, and the vitellarium (Figure 10). The germarium (most anterior tip of each ovariolum) (Figure 11) contains the germ-line and somatic stem cells that are continuously dividing in order to produce the cysts. The germ-line stem cells divide unequally. One of the resulting cells keeps the stem-cell identity and continues dividing, and the other one becomes a cystoblasts. The new cystoblast immediately divides four more times with incomplete cytokinesis, so the

resulting 16 cells remain in a close syncytium connected by a specialized structure called fusome. Because of the incomplete cytokinesis the remaining cleavage furrows, which are called ring canals, allows the cells to be interconnected with each other. As the 16 cell-cyst moves posteriorly in the

germarium it changes its shape, and makes room for the newly forming ones.

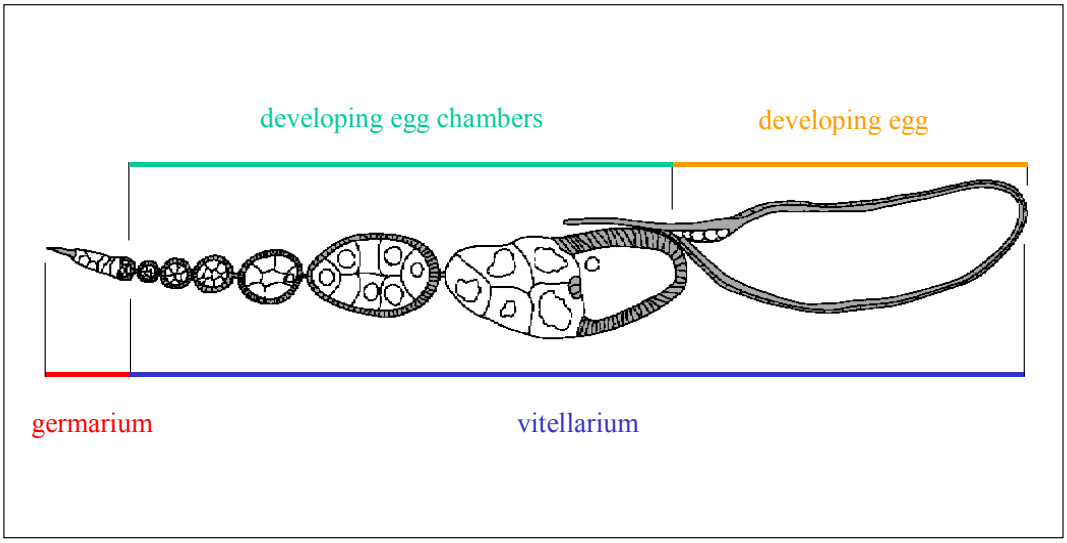

Figure 10: Each ovariole can be divided into two main regions. They are called germarium and vitellarium. Cysts are forming in the germarium region, and they continue to grow and develop in the vitellarium.

In region 2A of the germarium follicle cells start to migrate from the wall of the germarium to separate the germinal cysts. It is estimated that initially 16 follicle cells surround each cyst. This is he first of several major migration events these cells accomplish during oogenesis.

Fully formed stage 1 egg chambers are visible in the most posterior part of the germarium, in region 3. (Definitions of the different stages are based on (King 1979; Spradling 1993). (For reference see the Appendix.)

After leaving the germarium some differences become visible among the 16 germ-line cells. One of them grows as oocyte, which develops later as the egg. The

other 15 cells progresses as nurse cells. Their name reflects their functions; they provide nutrients, and the vast majority of the cytoplasmic constituents of the developing oocyte, even ribosomes and mitochondria.

This is also the time of rapid growth. Egg chambers multiply their volume a couple of times during this period. In the meantime follicle cells divide four to five times to maintain the uniform monolayer around the rapidly growing cyst. By the end of stage 5 they will reach their final number; between 600-800.

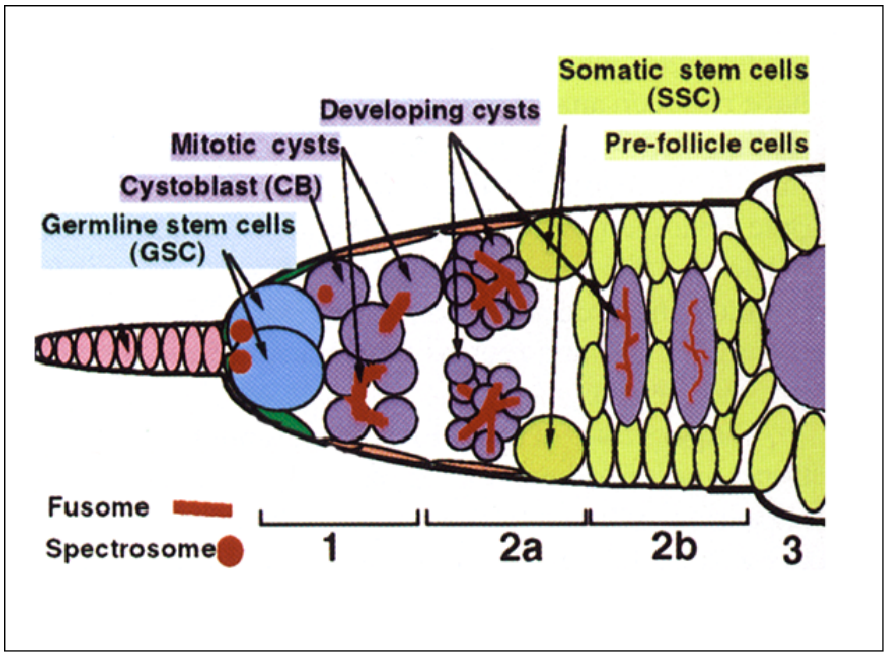

Figure 11: The figure shows the fine cellular structure of the germarium. The description of the different parts can be found in the text.

During mid-stages of oogenesis the major event is still the growth of the egg chamber. At the end of stage 9, however, a major reorganization of the follicle cells becomes visible. They start to move posteriorly in a concerted manner toward the oocyte. Only about 50 cells remain to cover the nurse cells. In order to accomplish that they have to flatten extensively, so the nurse cells remain completely sheltered. The

others, on the other hand, move until they become in close contact with the oocyte, where they take up a cubical shape. Some consider this as the second major migration event of the follicle cells that happen during oogenesis, but we think that this process cannot be classified as "true" cell migration. These cells don't adopt the characteristic polarized morphology one can see when looking at real migrating cells. During these changes, as I mentioned the cells at the posterior end of the oocyte start to become more columnar. Because they remain tightly connected at all times, subsequently they start to move relative to their substrate but not relative to their immediate neighbours. By the end of this “sliding” process the well-known polarized follicle cell layer develops around the egg chambers (Figure 12). It is still not clear, however, whether this movement is a cause or the result of the described cell shape changes.

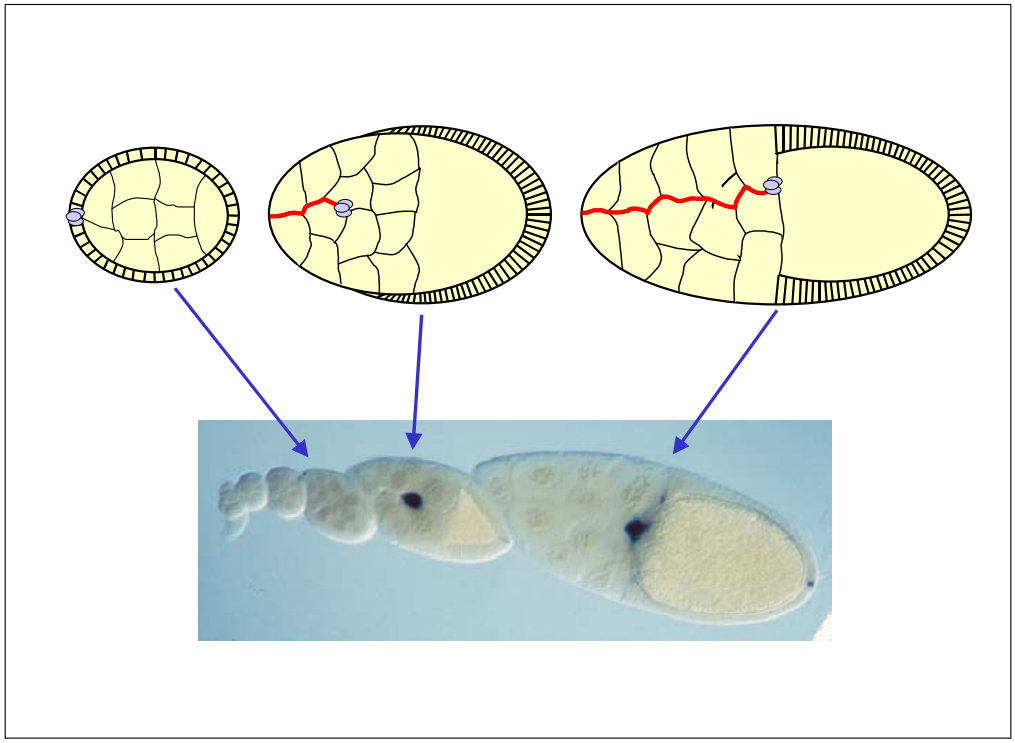

Figure 12: A group of six to ten follicle cells migrate from the anterior pole of the egg chamber to the oocyte. They move between the nurse cells, and called border cells, shown in blue.

A group of six to ten cells at the most anterior pole of the egg chamber undergoes a spectacular migration at the same time as the general movement toward the oocyte takes place. I consider this process as the second major cell migration event that occurs during oogenesis. These cells delaminate from the others at the anterior tip of the cyst, and shortly after they extend processes between the two most anterior nurse cells. They take up a typical migratory shape and begin to move toward the anterior tip of the oocyte. They remain closely associated with each other and, although they pass several nurse cell junctions, normally never choose an incorrect path. After arrival, the anterior follicle cells move a bit dorsally from the centre, but remain at the border between the anterior end of the oocyte and the nurse cells. This is why they are also called border cells.

There are at least two different subpopulations the border cells can be separated to. Initially on the most anterior part of the egg chamber the centrally located anterior polar cells are visible. They develop following a completely different lineage than the other follicle cells from very early on, expressing even different markers. The rest of the migrating anterior follicle cells are called "outer border cells" based on their initial position relative to the anterior polar cells at stage 9.

At the end of stage 10 by the time the border cells finish their migration, the follicle cells that lie in the junction between the oocyte and the nurse cells start to migrate centripetally. They got their name based on this process; this group of cells is called centripetal follicle cells. They close the anterior part of the oocyte separating it from the soon degenerating nurse cells during late stages of oogenesis. This can be considered also as a real cell migration event, because these cells have to partially

delaminate from the other follicle cells that remain covering the nurse cells, and they actively invade in between the nurse cells and the oocyte. This is the third migration process during oogenesis.

Note that all three migration events are very similar to each other in one feature; namely cells with somatic origin move between germ line ones. It would be interesting to investigate in more detail whether common signals drive these processes or the underlying mechanisms are different. We have some data showing that the transcription factor C/EBP is expressed both in the border cells and in the centripetal follicle cells. While its expression is absolutely necessary in the earlier cell type, there is no detectable problem in the migration of the centripetal follicle cells in slbo null mutant egg chambers.

Two previously mentioned subpopulations of cells, the border cells and the centripetal follicle cells, will have a major role in later stages in the formation of a specialized eggshell structure, the micropyle. This cone-shaped canal provides a passage into the egg for the sperm. Micropyle-like structures can be formed in the absence of the border cells, but they completely lack the so-called pore channel, and subsequently they cannot function properly. Egg chambers that have these defective structures, although they look visibly normal, can't be fertilized. The females that lay these eggs show a sterile phenotype (Montell, Rorth, et al. 1992).

3. RESULTS AND DISCUSSION

3.1 MUTAGENESIS SCREEN

3.1.1 Flyscreen

Previously it has been proven in many laboratories that single P element insertional mutagenesis is an efficient method to generate many independent transposon insertion events (Cooley, Kelley, et al. 1988; Bellen, O'Kane, et al. 1989; Bier, Vaessin, et al. 1989; Berg, Spradling 1991; Gaul, Mardon, et al. 1992; Torok, Tick, et al. 1993; Deak, Omar, et al. 1997). We performed such a screen using the EP construct in order to generate a mutant collection containing random, completely unselected transposon insertions on the X, second and third chromosomes (Figure 5) (Rorth, Szabo, et al. 1998). A protocol similar to previous single P element mutagenesis screens was employed with several technical improvements.

The usual scheme relies on two strains called “mutator”, and “jumpstarter”. The mutator was constructed by embryonic transformation of the EP element by P. Rorth (Rorth 1996). We used a second chromosome EP insertion on a dominantly marked CyO chromosome called EP43. In the previous mutant screens X chromosome insertions were often used in the different laboratories (Torok, Tick, et al. 1993; Deak, Omar, et al. 1997; Cooley, Kelley, et al. 1988). In those experiments the jumping events were generated mostly in males, thus the new insertions were recovered only on the autosomes. Our setup, on the other hand, allowed us to isolate mutations on both the autosomes and the X chromosome.

In a series of preliminary experiments (Rorth, personal communication) we made sure that the transposons in the potential jumpstarter lines were intact, and they could move when the transposase enzyme was provided. It has been noted previously

that different insertions of specific mobile elements varied widely in their ability to undergo remobilisation based on the size and the structure of the transposon, and also the chromosomal region where they were inserted (Berg, Spradling 1991). We also tested the possible mutators whether their jumping efficiency was optimal, and based on these results we chose to use the EP43 line.

The jumpstarter is a transposase-producing stock in which a defective P element expresses high levels of the enzyme required for mobilization, but is itself incapable of transposition because of missing or imperfect terminal inverted repeats. We used the publicly available Δ2-3 transposase source in our experiments.

To initiate mutagenesis, the mutator strain was crossed to the jumpstarter (P generation). Progeny with mosaic eyes that inherited both the EP element and the transposase source were individually mated to *white* homozygous flies (G_0 generation). P elements normally excise and transpose only when a source of transposase is supplied, and in case of wild type transposase this happens only in the germline. The germline specificity depends on the presence of the intron between the 2^{nd} and 3^{rd} exons of the transposase gene, which is normally not spliced in somatic cells. The Δ2-3 transposase source we used in our experiments lacks this intron, and as a result it initiates the movement of the P element not only the germline, but also in somatic cells. We can detect this event by the appearance of $white^{+}$ cells on an otherwise w^{-} background in the eye of these animals. These type of eyes are called mosaic eyes.

The transposase catalyzed the movement of the EP element to a new insertion site in small percentage of the germ line cells of these animals. Premeiotic insertion of the transposon can lead to the appearance of the same mutation in multiple progeny forming a mutant cluster. Since in the mutagenesis experiment the G_0 dysgenic flies were crossed individually and single inserts were recovered from each vial, we avoided

the appearance of such clusters. This way every single recovered mutant line represents an independent event in our collection.

It has been proposed previously in the literature that there is a possible difference in the P transposition in male versus female germline (Zhang, Spradling 1993). This might be due to some intrinsic difference between the mechanism of transposition or its efficiency in the different germ line cells. Alternatively, the target sites or the chromosome regions that are permissive for transposition could also vary in the two sexes. Last but not least the level of the different cellular factors that help the movement can also differ. In order to provide a new tool for investigating this question, we decided to recover jumping events induced in both sexes. We selected males and also females with yellow eyes in the G_0 generation.

G_1 males and females having an insertion in their genome were identified after their uniformly coloured yellow eyes due to the presence of the mini-white marker gene in the transposon. Since the second chromosome that carried the original insertion, and the third chromosome having the transposase-producing construct were both dominantly marked, we could avoid the transmission of either of them to the progeny by selecting actively against these two markers. This way we could prevent both an accidental reisolation of the original mutator insertion and any further transposition of the newly inserted EP element.

In our screen approximately 25% of the G_0 males containing both the mutator and jumpstarter elements yielded useful insertions within their progeny. It would have been possible to substitute the EP43 mutator stock with other ones having greater jumping frequency but we decided against it, based on the results of previous mutant screens done in different laboratories (Torok, Tick, et al. 1993). On one hand this could have increased the jumping rate much closer to the maximum of 100% where basically

each G_0 cross would produce a new line, but it would also have resulted in a higher number of stocks carrying multiple insertions. In our case this was disadvantageous, so we chose to use a line with a reasonable, but not too high jumping frequency.

F_1 CyO males carrying the EP element were crossed to w; TM3 / TM6 virgins.

The chromosomal location of each new insertion was mapped, by following the segregation of the yellow eye colour compared to the different chromosomal markers in both sexes in the next generation. Parallel to that the stocks were balanced using the dominantly marked balancer chromosomes.

In subsequent generations we determined if the transposon insertion itself caused a recessive lethal phenotype.

3.1.2 Screen statistics

Part of this work and the screen itself was done in collaboration with the *Berkeley Drosophila Genome Project* as part of its massive gene disruption project that uses P transposable elements to target open reading frames throughout the *Drosophila* genome (Spradling, Stern, et al. 1999). Our plan was to generate a reasonable size, representative collection of random, unselected EP insertions, so we established 2296 individual stocks. As expected, the screen was far from saturation. The *Drosophila* genome is estimated to contain near to 15.000 genes in average. Since the EP element is directional, only 50% of the insertions are oriented in such a way that it directs sense strand expression of the adjacent gene. If we also take the fact into consideration that P element insertions are not completely random in the genome because of the presence of

so-called hot- and cold spots, than we can estimate that the ~2300 EP insertions are targeting less than 10% of the total fly genome.

Based on the results of previous mutagenesis screens we expected that in approximately 40% of the lines the insertions were going to map to the second chromosome. Another 40% would fall to the third, and the remaining 20% would be localized to the X (Cooley, Berg, et al. 1988; Torok, Tick, et al. 1993). The basis of these predictions is the observation, that the size of the second and third chromosomes is comparable, while the size of the X is approximately half of the autosomes. There are cold- and hot spots, but no preferred or excluded chromosomes with regard to P-transposition in the fly genome. The results verified our prognosis; in 413 stocks the insertion was mapped to the X, in 939 lines to the second, and in almost an equal number of cases (944) to the third chromosome (Figure 6).

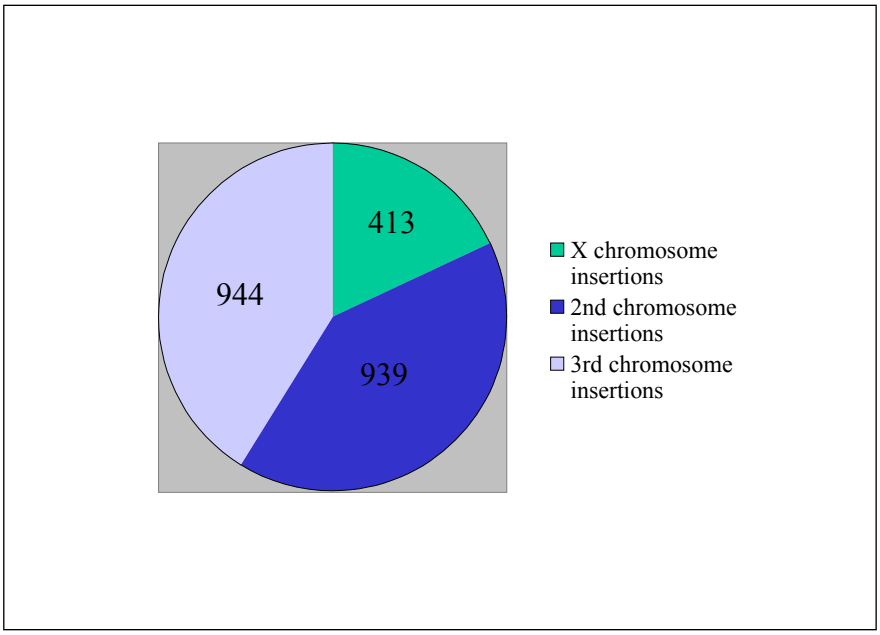

Figure 6. The distribution of the generated EP insertions on the X, second and third chromosome.

Insertion of a transposable element to an essential locus may cause a recessive lethal phenotype. It is expected to happen in approximately 10% of the resulting lines. This data is estimated based on the results of both P element (Cooley, Berg, et al. 1988) and chemical mutagenesis experiments (Gans, Audit, et al. 1975). After mapping and balancing the insertions we determined whether they caused such a phenotype. In the EP collection the ratio of the lethal stocks was higher than previously predicted among the third chromosome insertions (~25%), which suggests that there might have been a recessive lethal background mutation present on some of the starting chromosomes at the time of the mutagenesis.

The majority of the stocks carry a single EP element insertion. There are multiple inserts in about 4% of the lines. This estimation was made by the Berkeley Drosophila Genome Project. They cytologically mapped every single EP inserts by inverse PCR, and made this estimation based on their results.

3.2 MUTANT SCREENS

3.2.1 How to use the EP lines?

Our collection allows systematic misexpression of target genes that are adjacent to the EP element by crossing the lines to different Gal4 drivers.

There are two ways to study one's favourite biological problem using the EP system (Figure 7). The more "traditional" way is to choose a suitable driver that can direct ectopic expression of target genes in the required place at the right time. This way it is possible to study the direct phenotypic consequences of this event on an otherwise normal, wild type background. This is the so-called dominant gain-of-function method.

Controlled overexpression of certain genes can identify important genetic interactions (Fortini, Artavanis-Tsakonas 1994). If increased expression of one gene enhances or suppresses a certain phenotype caused by a mutation of another one, their products are likely to be involved in the same process. This notion is the basis of powerful genetic approaches for example in yeast (Bender, Pringle 1989; Ramer, Elledge, et al. 1992), and also the main idea behind the so-called interaction screens (Rogge, Karlovich, et al. 1991; Simon, Carthew, et al. 1992). In this kind of experiments we mis- or overexpress target genes using different drivers on a special, sensitised genetic background (see Figure 7). This usually means a pre-existing mutation that is present in the system. In this case we can isolate modifiers (enhancers and suppressors) of our pre-existing phenotype, allowing the genetic dissection of certain signal transduction pathways, for example.

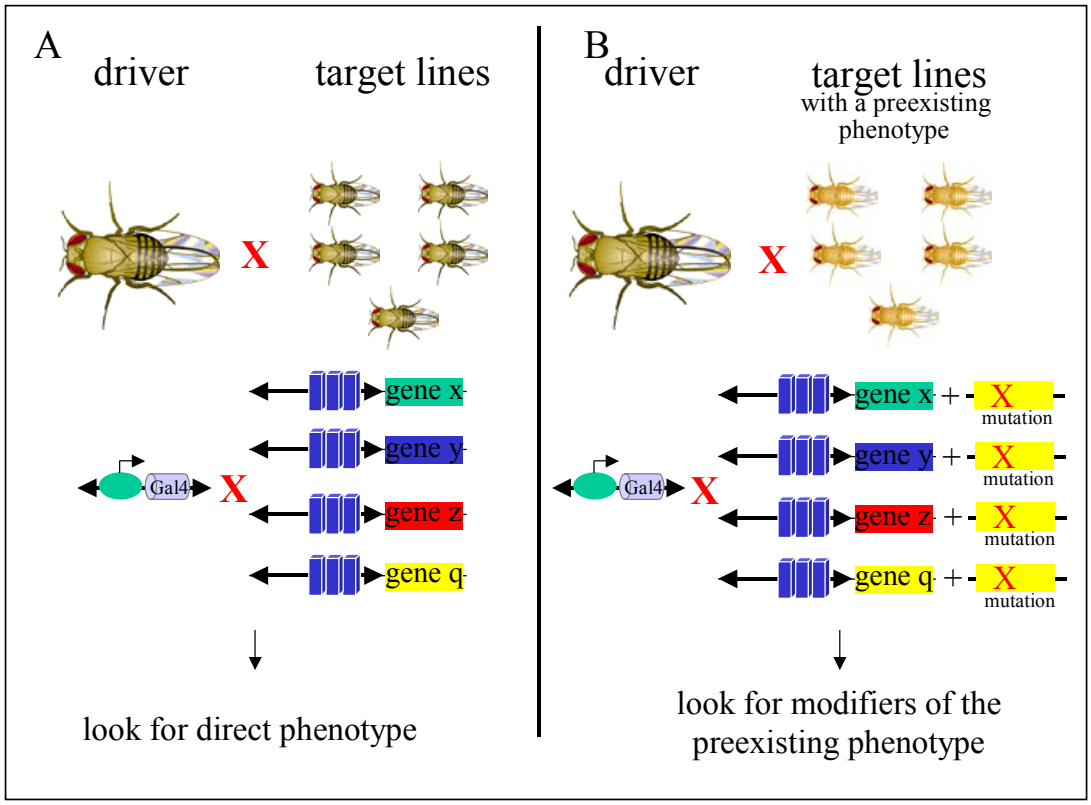

Figure 7. Systematic presentation of how to set up a (**A**). dominant gain-of-function, or an (**B**.) interaction screen using the EP system. For more detailed description see the text.

In the laboratory of Pernille Rorth we have decided to analyse a special cell migration process during *Drosophila* oogenesis using both of these methods.

3.2.3 The role of *slbo* in oogenesis

It has been shown previously that in *slbo* mutant ovaries the border cells migrate late or don't migrate at all depending on the strength of the allele used in the experiment. The *slbo* (*slow border cells*) locus encodes the *Drosophila C/EBP*, which is a member of the basic region-leucine zipper transcription factor family. It has been shown

previously, that *C/EBP* is expressed in the migrating border cells where its presence is absolutely necessary (Montell, Rorth, et al. 1992). It is also present in centripetal follicle cells.

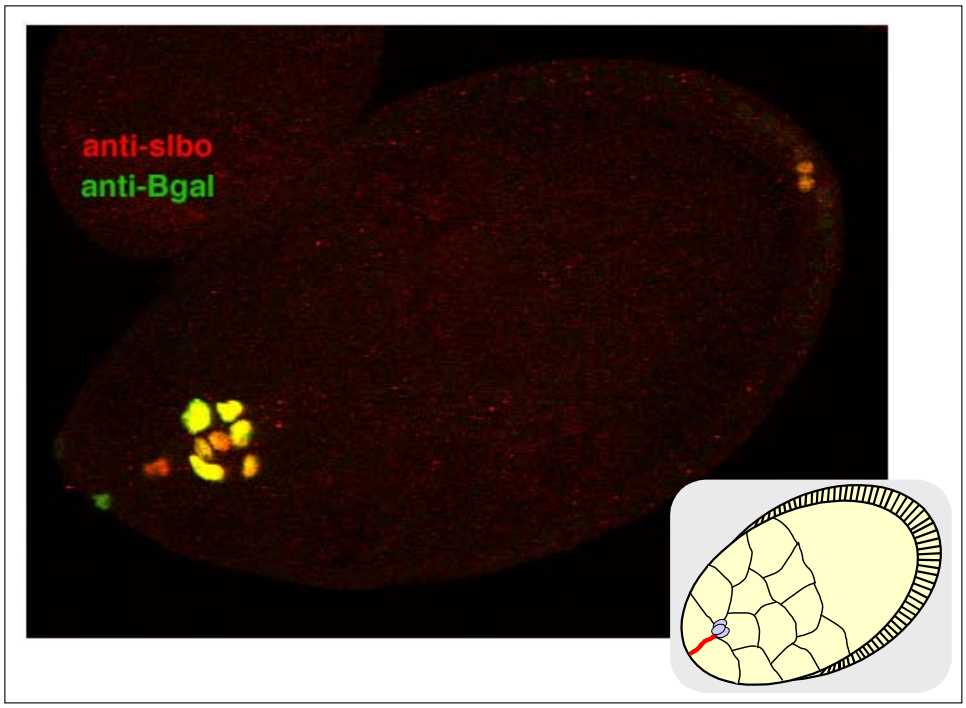

Figure 13: *slbo* protein can be detected in border cells before and as they migrate by immunostaining.

Experiments however clearly show, that the migration of this latter cell type is perfectly normal even in the absence of *C/EBP* function, so it's role in these cells is not clear yet (Rorth, Szabo, et al. 2000).

slbo mRNA and protein is detected in border cells before and as they migrate (Figure 13) both in the centrally located anterior polar cells and also in the outer border cells. Detailed clonal analysis using *slbo* null mutant alleles in border cells revealed some interesting facts (Rorth, Szabo, et al. 2000). The presence of *C/EBP* protein is absolutely necessary in the outer border cells in order to become migratory. Even if one

or two of the outer border cells are *slbo* null mutant the others will force them to move together, although they always locate at the rear of the group. This shows that there are special factors that hold the group selectively together, and the "normal" cells are dragging along the mutant ones. If the majority of the group is mutant and only one or two outer cells are wild type, they would initiate migration but they cannot move very far. The anterior polar cells are not migratory cells by themselves, they seem to be moving also because the other outer border cells force them to do so. Their clonal origin is completely different than the rest of the follicle cells, and they even have distinct shape. It is possible that they have a role in the recruitment of the outer border cells.

It is clear that the border cells form a tight little cluster. The factors that are responsible for this selective adhesion within the group remain to be unknown.

After migration, the level of *C/EBP* protein in border cells is rapidly decreasing, and subsequently it disappears completely.

As a phenotypic consequence, in the homozygote *slbo* mutant females the border cells fail to migrate properly, and the result of it is that the border cells cannot reach their destination on time. The pore channel of the micropyle will be missing, and the sperms cannot get to the oocyte during fertilization. Although the eggs laid by the *slbo* mutant mothers look completely normal, they show a female sterile phenotype (Montell, Rorth, et al. 1992).

3.2.4 Introducing the different screen types that can be used using the EP system

The *slbo* phenotype and the fact that *C/EBP* is a transcription factor suggest that its most likely function is the regulation of certain genes with a possible role during

border cell migration (Rorth, Szabo, et al. 1998). To investigate this possibility we decided to set up both an interaction and a direct screen, because we expected to find different set of genes with the two methods (Figure 14).

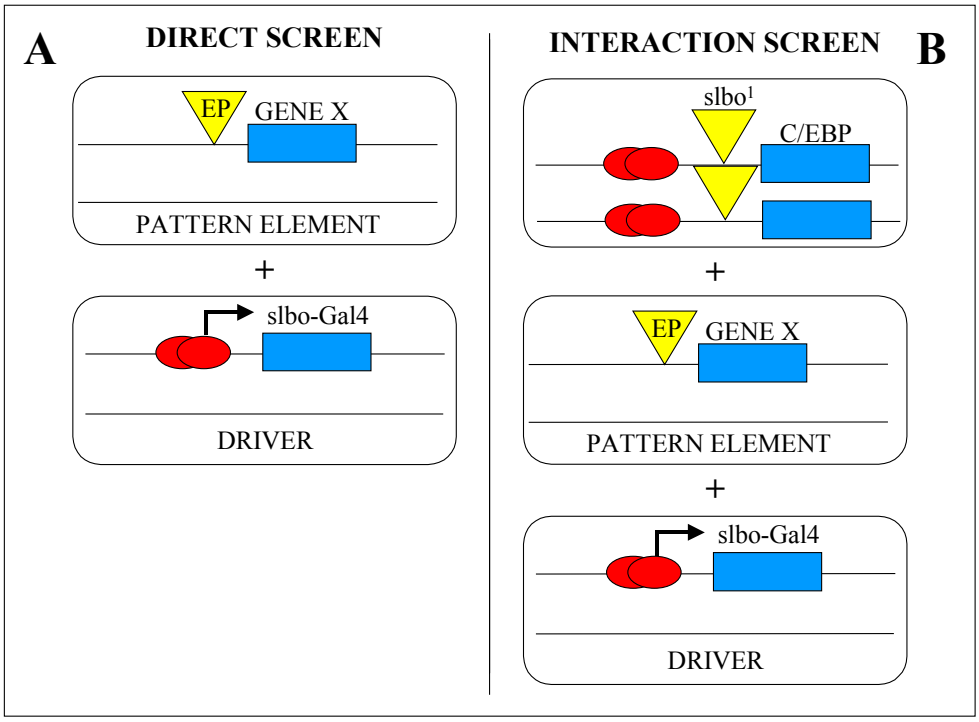

Figure 14: In the so-called direct screen one can overexpress random genes on an otherwise normal, wild type background (shown in figure A). In the interaction screen, however, one uses a sensitized background, by introducing a special mutation (figure B).

With the interaction screen we tried to target genes that have positive effect on the migration of our favourite cell type. In this case we drove the expression of random loci in the anterior follicle cells on a special background, where their migration was disturbed. We tried to isolate genes that help this movement. Detailed description of the experimental setup and the results will follow in the next chapters (Rorth, Szabo, et al. 1998).

Using the direct approach, however, we wanted to get to know the players that work against this migration process. For this we overexpressed random genes on an otherwise wild type background, and looked for the ones that disturb our studied event. See chapter 2.6 for details (Rorth, Szabo, et al. 2000).

3.2.5 slbo-Gal4 driver

In both screen types the basic idea is the overexpression of genes using special Gal4 drivers on different genetic backgrounds. In order to do that specifically in border cells, P. Rorth constructed a suitable vector (Rorth, Szabo, et al. 2000). Genomic DNA units upstream of the *C/EBP* coding region were cloned in front of a *lacZ* gene. She looked for a construct where the *lacZ* gene expression mimicked the endogenous *C/EBP* expression pattern (described in chapter 3.2.3.). The same fragment was then placed in front of a *Gal4* gene (driven by a hsp70 minimal promoter) in order to generate a driver for the further experiments. It is visible on Figure 15 that slbo-Gal4 drives high expression of the reporter gene in border cells before and throughout their migration.

Test experiments showed that slbo-Gal4 also successfully rescued the female sterility of the homozigote *slbo* mutant females when combined with UAS-C/EBP, suggesting that the driver is working in the right place, at the right time.

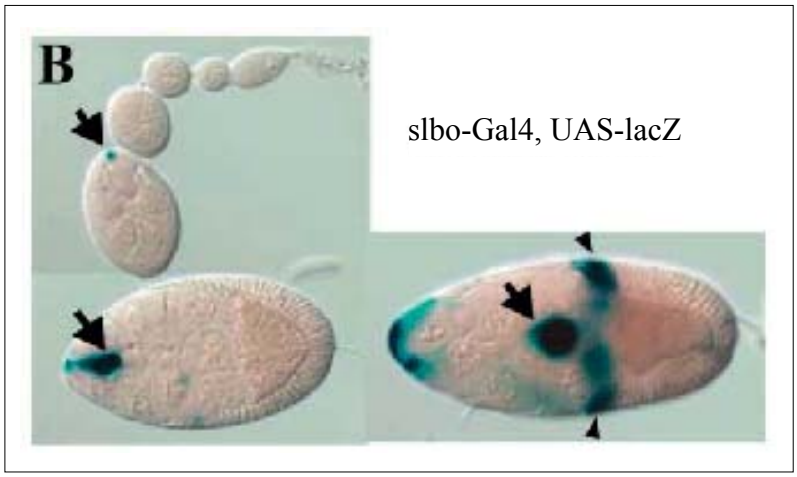

Figure 15: slbo-Gal4 was combined with UAS-lacZ. The ovaries of the females carrying both elements were dissected out, and they were stained with b-galactosidase. The border cells are shown in blue, suggesting that the slbo-Gal4 driver drives expressionin these cells before and as they migrate. This mimics the endogene C/EBP expression.

3.3 INTERACTION SCREEN

To identify *slbo* suppressors, 2082 EP lines were crossed into the $slbo^{1}$ background in the presence of the slbo-Gal4 driver (Rorth, Szabo, et al. 2000). In the next generation females carrying both elements were tested for fertility (see Materials and Methods for detailed description of the experiment). We analysed 90.7% of the generated insertions. In the remaining ~10% the EP inserted in the close proximity of the *slbo* locus and the generation of the recombinant chromosomes containing both the EP elements and the driver would have required a much greater effort.

The degree of Gal4 expression is temperature sensitive with increased expression at higher temperature, similarly to the rate of border cell migration. In order to standardize the conditions we kept all of these crosses at room temperature. At this temperature the homozygote $slbo^{1}$ mutant females were completely sterile.

4.2% of the total number of tested EP lines (88 out of 2082) suppressed the sterility caused by the *slbo* mutation in our preliminary screen. These lines were retested and the rate of border cell migration was determined. After dissecting out the ovaries of the partially fertile females we X-gal stained them making use of the fact that the *slbo*^{1} allele is an enhancer trap. Migration was quantified after visual observation as a percentage of stage 10 egg chambers where we could detect any migration of the anterior follicle cells. We artificially defined several categories (as shown in Figure 16) based on the distance the cells travelled, in order to give us a finer picture on how well they were moving, and to make the comparison of the different stocks easier.

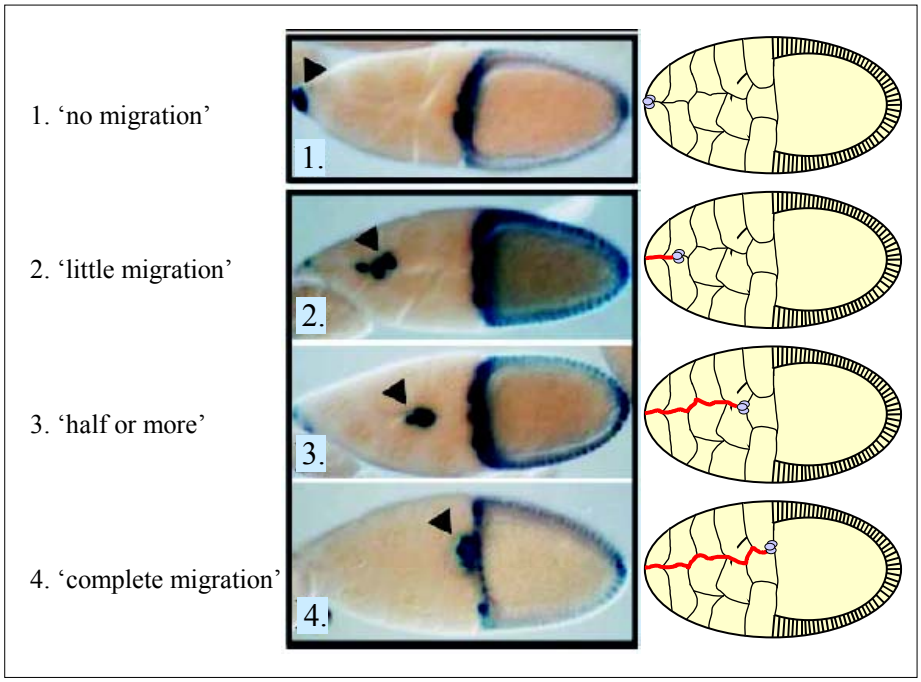

Figure 16: Migration was quantified after visual observation as a percentage of stage 10 egg chambers where we could detect any migration of the anterior follicle cells. We artificially defined 4 categories, as shown in this figure, based on the distance the cells travelled, and counted the number of follicles that fell into the different classes.

This way we could even distinguish between cases where we could detect, for example, little movement in most egg chambers from where we saw long-range movement in smaller percent of the follicles. This kind of information could give us some information on the finer mechanisms of the migratory process.

slbo mutant females carrying no or inert EP inserts show some migration in approximately 3.4% of the stage 10 eggs at room temperature, which was not sufficient to make these females fertile. 60 out of the 88 putative suppressors produced migration in at least 9% of the stage 10 egg chambers (an arbitrary cut off).

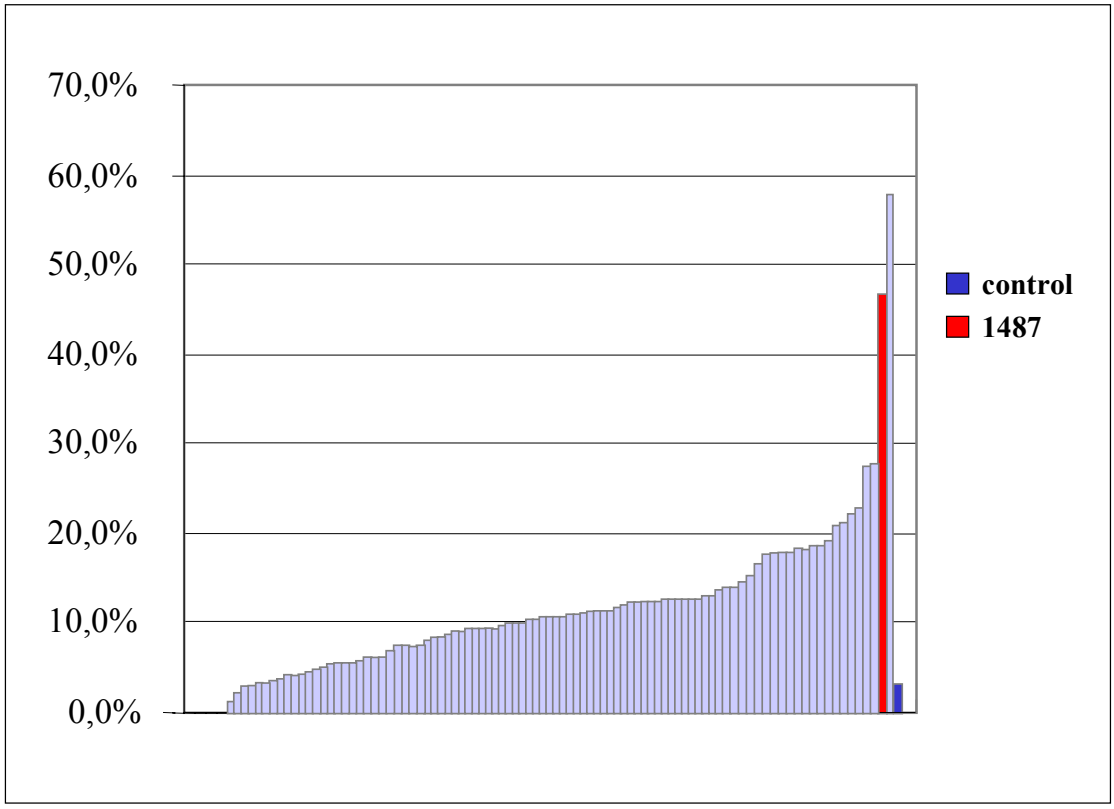

Figure 17: The majority of the 88 putative suppressor lines (showed in light blue) showed a moderate but reproducible 3- to 5-fold increase in the rate of migration relative to the control (shown in dark blue). We also found a few stocks that showed a much more robust effect e.g. EP1487, shown in red.

As expected the majority of lines showed a moderate but reproducible 3- to 5-fold increase in the rate of migration relative to the control, and we also found a few stocks that showed a much more robust effect (as shown in Figure 17).

3.3.1 Testing whether the *slbo* suppression effect was Gal4 dependent

After isolating the potentially interesting stocks we checked if the mutant phenotype, originally associated with the line, was in fact caused by the Gal4-driven overexpression of the affected locus.

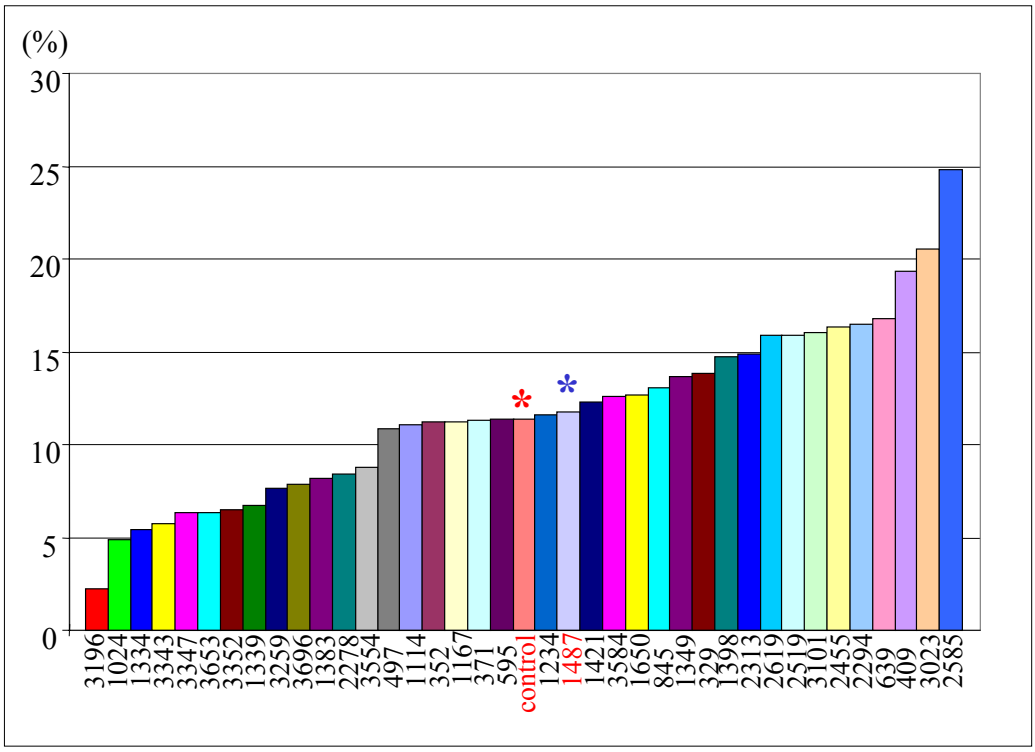

Figure 18: The majority of the EP inserts did not increase the rate of border cell migration when the driver was not present. One can conclude that by comparing the rate of cell migration in the different lines compared to the control (marked with a red star). This is true for the line EP1487 (marked also with a blue star).

For this we crossed the majority of our 60 suppressor lines to homozygote *slbo*^{1} mutant background again, but this time in the absence of the slbo-Gal4 driver (Figure 18). After dissecting and X-gal staining the ovaries of such females we determined the rate of border cell migration in the previously described way.

The majority of the EP inserts (56 out of 60) did not increase significantly the movement of the anterior follicle cells indicating that the suppression of the *slbo* phenotype was indeed caused by the forced overexpression of the different target genes (Figure 18).

3.3.2. Does the EP insertion itself has any effect on the border cell migration in the different lines?

60 potential suppressors were crossed to a *slbo* mutant background in this experiment, but instead of the null allele called *slbo*^{1} (that we used in the suppressor screen) we used an intermediate allele named *e14* this. This time there was no slbo-Gal4 driver present in the flies and the EP insertions were in homozygote form.

This way we tried to test if the EP insertion itself has any effect on the border cell migration. In other words whether the transposon causes a loss-of-function mutation in the affected locus resulting any phenotypic consequence. We expected to find 3 different categories.

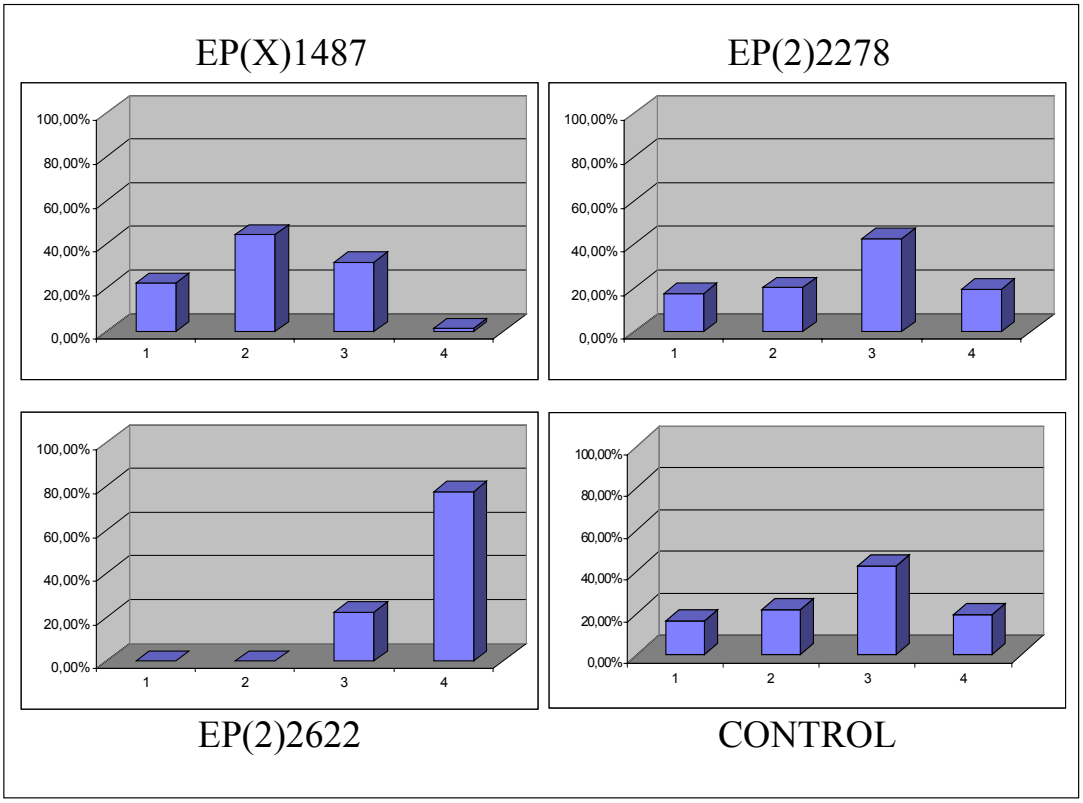

Figure 19: The insertion of the transposon itself can cause loss-of-function mutations in some cases. The affected gene itself in wild type ovaries under physiological conditions can have either positive (EP(X)1487), or negative effect (EP(2)2622) on border cell migration, and in some cases it has no effect at all (EP(2)2278). We could show that on a special, sensitised slbo mutant background.

When the insertion does not cause a loss-of-function mutation, the rate of border cell migration should be the same as in the control (e.g in EP(2)2278). If the EP insertion leads to a mutation in the affected locus, and the gene itself has a positive effect on the cell migration in wild type under physiological conditions, than we should detect reduced rate of migration in this experiment (e.g. EP(X)1487). If, however the gene normally inhibits the migration of the anterior follicle cells during oogenesis, the mutation should result increased level of the rate of migration in this experimental setup (e.g. EP(2)2278) (Figure19).

3.4 MOLECULAR ANALYSIS

At this point we had 60 promising stocks that we categorized as potential *slbo* suppressors. We showed that this phenotype was due to the overexpression of certain genes, so we started to determine the insertional sites in these lines.

3.4.1 Determining the site of EP insertions in the different *slbo* suppressors

The EP element contains special sequences: a bacterial replication origin, dominant selectable bacterial marker (kanamycin resistance), and special restriction enzyme sites (Figure 20.). Altogether these make the isolation of the genomic regions immediately next to the transposon possible (both at the 5' and the 3' ends of the P-element) by plasmid rescue. This way it is easy to find out exactly which is the mutated gene in the different stocks. Briefly, we used a unique EcoRI site to obtain a genomic piece at the 3' end of the transposon, which contained the overexpressed sequences. In case the resulting fragment was too small or we did not get anything at all, we repeated the experiment using an SstI site. Finally, we obtained DNA fragments in all the 60 cases.

Before determining the sequence of them, we wanted to make sure that we had the right fragments, so we performed Southern experiments. Genomic DNA from all the suppressor stocks was digested with the appropriate enzyme, the same that we used to obtain the rescued fragments. We prepared the blots, probed them with radiolabelled transposon-specific probes, and compared the size of the labelled bands to the size of the plasmid rescue fragments. If they matched, we started the further analysis.

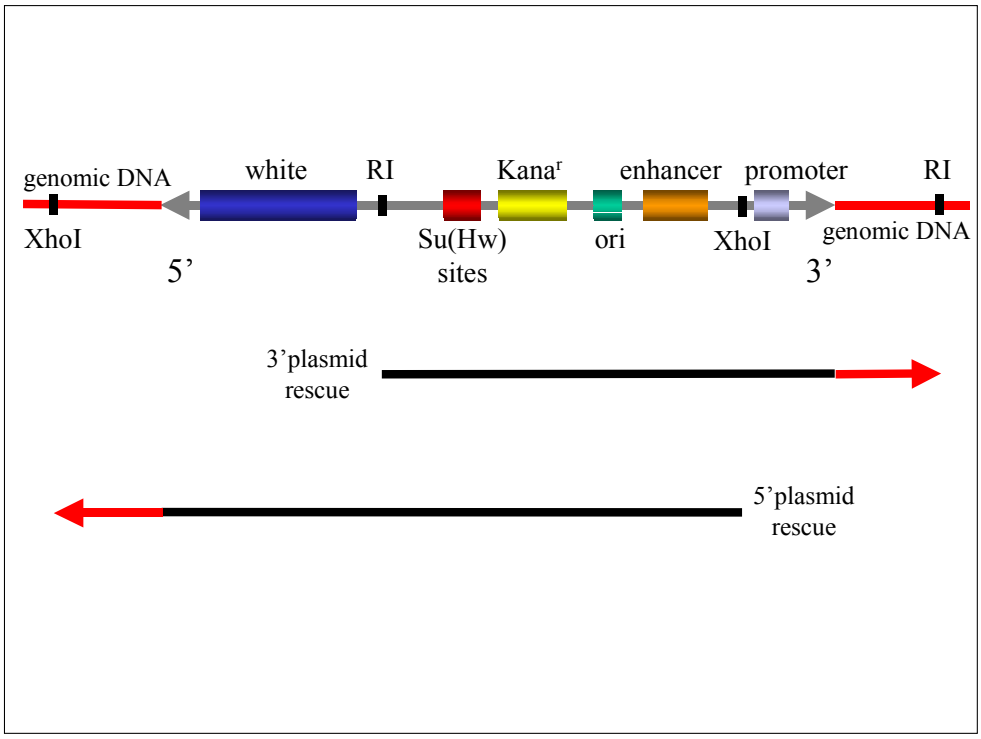

Figure 20: Because of the special structure of the EP element it is possible to obtain the genomic sequences from both sides of it using a technique called plasmid rescue. Using special restriction sites it is possible to obtain fragments containing a genomic piece, and sequences that allow the circularised fragment to grow in bacteria.

We sequenced the ends of the verified rescued genomic pieces in all of the 60 potential *slbo* suppressor lines, and searched the *Drosophila* sequence database to determine the site of transposon insertion. (The results are shown in Table 1)

In most cases the element inserted in previously identified genes or predicted open reading frames. In a few cases, however, it was not possible to determine so far the affected loci even after the complete sequence of the *Drosophila* genome became available, because there was no gene in the expected distance that could be overexpressed in these lines at the site of insertion. This is the case for example in EP(2)0639.

In some lines the EP element inserted in such way that it landed in other mobile elements. Many cases of similar nested transposable element insertions are known from the literature. Sometimes this observation is being interpreted as a way to direct mobile elements outside of gene coding regions to reduce the damage inflicted to the host by their mobilization (Kidwell, Lisch 1997). This is what happened in the following lines; EP(2)0550, EP(3)1064, EP(3)3213, EP0324.

3.4.2 Can the EP element generate antisense transcripts?

The sequencing and the determination of the insertional sites and the affected genes revealed some interesting facts. Because of the asymmetric structure of the EP element, Gal4- induced transcription is initiated only at the 3' end of it. Thus, depending on the insertion site and orientation of this transposon relative to the transcription start site of the target gene, the EP element can direct the formation of either sense or antisense transcripts.

When sense transcripts are being formed, the P elements' tendency to preferentially insert in the 5' regions of the different genes is particularly useful. This allows the formation of full- or near full-length transcripts upon overexpression. In the majority of the closely examined cases, the EP element was indeed inserted either upstream or in the first part of the coding region, and directed the formation of sense transcripts (Abdelilah-Seyfried, Chan, et al. 2001; Rorth, Szabo, et al. 1998; Pena-Rangel, Rodriguez, et al. 2002; Kraut, Menon, et al. 2001; Tseng, Hariharan 2002).

Antisense transcripts, on the other hand, can produce Gal4-dependent loss-of-function phenotypes (Kraut, Menon, et al. 2001; Rorth 1996; Rorth, Szabo, et al. 1998). In some of the slbo suppressor lines the EP element inserted downstream of a

transcription start site in an orientation where overexpression would produce an antisense transcript. We analysed one line called EP(2)2585 in more detail. Our tests indicated that the insertion itself caused a mild loss-of-function phenotype in this case; it weakly suppressed the *slbo* phenotype even in the absence of the slbo-GAL4. The suppression effect was much greater, however, in the presence of the driver, indicating that the induced antisense transcript caused a more severe loss-of-function effect (Rorth, Szabo, et al. 1998).

In other screens, performed by different laboratories, other potential "antisense" lines were isolated. In the dominant eye and wing screen a line called EP(2)0683 was found. The insertion mapped to a gene called *escargot*, and in this particular gene Gal4-driven overexpression resulted also in an antisense transcript. This region seems to be a hot spot for P element insertions, and in the same screen 4 more "sense" lines were found. Interestingly, the single "antisense" line produced different phenotype than the 4 other ones when they were overexpressed with the same drivers, further strengthening the idea that the underlying mechanisms causing the mutations were also different (Rorth, Szabo, et al. 1998).

All these results suggest that using this system it is possible to generate antisense transcripts *in vivo* in flies, and this effect can even generate conditional loss-of-function phenotypes. One has to note although that the frequency of this phenomenon seems to be quite low.

3.4.3 Testing the cloned genes

Several of the EP suppressors were inserted in such way, that they would direct essentially full-length transcripts of different genes. (E.g. EP(3)0329-*Hsp27*; EP(3)3101-*Abelson tyrosine kinase*; EP(2)2519-*α-Adaptin*; EP(2)2278-*big brain*.) We obtained and cloned the corresponding cDNAs into the pUAST vector (Brand, Perrimon 1993). We tried to prove that by cloning these cDNAs into an expression vector and overexpressing them exactly the same way as we did with the original EP lines in the *slbo* screen, we can reproduce the same phenotype. In other words, we wanted to show that the suppression effect we saw was due to the Gal4-driven overexpression of the identified genes.

We were able to reproduce the original results using the transgenic lines that carried these direct cDNA fusions. I would like to note that in every single case we detected stronger suppressor effects when we directly overexpressed the cDNAs than in case of the corresponding original EP lines. This can be explained by the fact that in case of the direct cDNA fusions the promoter sequences lay much closer to the transcription start site, whereas in the corresponding EP lines this distance can be as big as a couple of hundred base pairs. These genomic pieces might also contain regulatory elements that can interfere with transcription.

The results nonetheless confirmed our prediction; we really did identify and clone the genes that can suppress the migration defect in *slbo* mutant border cells, when we overexpressed them using a suitable driver.

3.4.4 Some of the known genes identified in the screen

In case of several lines it was relatively easy to interpret the suppression effect based on previously known features of the affected genes.

3.4.4.1 Syndapin, EP(3)0409

Recovery of plasma membrane after stimulated exocytosis is commonly referred to as compensatory endocytosis. This is a clathrin-mediated endocytotic process that helps to recycle the membrane proteins and membrane fragments at the presynaptic nerve terminal. Known players of this process are dynamin, clathrin, and AP2 (heterotetrameric adaptor complex protein 2).

Syndapin was originally described as a protein that interacts via its SH3 domain with the proline-rich C terminus of dynamin. Dynamin binds to the collar region of the clathrin coated pits, and helps the pinching off process. Recruitment of dynamin to the forming pits seems to be dependent on the interaction between syndapin and dynamin. Furthermore, syndapin also interacts with another major nerve terminal enriched protein called neuronal Wiskott-Aldrich syndrome protein (N-WASP). There is strong evidence that N-WASP plays a major role in regulating the actin cytoskeleton, by activating the Arp2/3 complex (Miki, Miura, et al. 1996; Miki, Sasaki, et al. 1998). Thus, syndapin might serve to link the dynamin-mediated processes of endocytosis to actin cytoskeleton rearrangements (Qualmann, Roos, et al. 1999).

Endocytosis plays an important role not only in compensatory endocytosis but also allows eucaryotic cells to regulate for example certain signal transduction pathways by receptor internalisation. Activated receptor molecules become trapped into

endocytotic vesicles. The receptor molecules after complex processes will either loose their ligand and will be recycled and presented again on the membrane of the cells, or they can be targeted for lysosomal degradation. It has been noted that, for example, the attenuation of the EGFR signalling is commonly mediated by internalisation and degradation of the activated receptor molecules (Tall, Barbieri, et al. 2001). It is not too difficult to imagine that this kind of processes can have important role in the regulation of border cell migration, or in the pathfinding-process in response to chemical clues.

Last, but not least syndapin is very similar to the chicken FAP52, an ubiquitously expressed SH3 domain-containing protein (Merilainen, Lehto, et al. 1997). Whereas syndapin colocalizes with dynamin to presynaptic nerve terminals, FAP52 seems to be localized to focal adhesions in cultured chicken fibroblasts. It is well known that dynamic establishment and reorganisation of the adhesion sites is another factor that is critical for the proper movement of the different migratory cells, and the same seems to be true for the growing nerve growth cones. Growth cones express various cell adhesion molecules (CAMs) that recognize localized guidance clues present on neighboring cells or in the extracellular matrix and translate them into a directed axonal extension (Tessier-Lavigne, Goodman 1996). Spatially localized actin polymerization/depolymerization and actin-myosin interactions generate retrograde flow of actin filaments (Mitchison, Cramer 1996), which then produces a traction force to pull the growth cone forward (Suter, Forscher 1998). This requires not only the CAM-cytoskeletal linkage, but also a gradient of adhesive interactions with its environment (strong adhesion at the leading edge of the growth cone and weak adhesion at the rear) (Lauffenburger, Horwitz, 1996). CAM-mediated adhesion can be regulated by the dynamic regulation of CAM density in the different regions of the moving cells, which can be achieved by preferential internalization of these molecules at some region of a

cell followed by recycling into a different region. This is another example suggesting that there must be a link between endocytosis and the dynamic changes of the actin cytoskeleton, and the family of syndapin/FAP52/PACSIN possibly play a key role.

As dynamic actin remodelling is one of the major cytoplasmic events that happen during any kind of cell movement, syndapin can easily be important during border cell migration. Data that points in this direction show that overexpression of full-length syndapin modulates the cortical actin cytoskeleton *in vivo* to induce filopodia formation in cell culture (Qualmann, Kelly 2000).

3.4.4.2 α-adaptin, EP(2)2519

Not only syndapin, but also other members of the endocytotic pathway seems to have a role in border cell migration, based on their *slbo* suppression phenotype. We isolated a stock called EP(2)2519 in our suppressor screen, where the insertion turned out to be in a gene called α-adaptin. Adaptins are important components of clathrin-coated endocytotic vesicles, transporting ligand-receptor complexes from the plasma membrane or from the trans-Golgi network to lysosomes. The adaptin family of proteins is composed of four classes of molecules named α, β-, β'- and γ adaptins. They, together with medium and small subunits, form a heterotetrameric complex called an adaptor, whose role is to promote the formation of clathrin-coated pits and vesicles (Hirst, Robinson 1998).

The affected gene encodes the alpha-adaptin subunit of the adaptor protein 2 (AP-2) complex. The complex is part of the protein coat on the cytoplasmic face of coated vesicles that links clathrin to receptors in vesicles.

3.4.4.3 Hsp27, EP(3)0329

One of our suppressors turned out to be the small heat shock protein HSP27. The role of these proteins is not well understood. They thought to have important function in protecting the cells from stress e.g. heat or certain chemicals, but more recent studies indicate that they are also important during normal development. Lately, tissue culture experiments showed that HSP27 itself were ubiquitous targets of phosphorylation upon cell stimulation with a variety of growth factors and agents that affected cellular differentiation and actin filament dynamics. Results also illustrated that HSP27 was enriched at the leading edge of polarized fibroblasts, and was localized in lamellipodia and membrane ruffles where usually most actin polymerization occurs (Lavoie, Hickey, et al. 1993). Further experimental evidence suggests that HSP27 can act as an actin-polymerizing factor during these processes. The ability of HSP27 overexpression to suppress the *slbo* migration phenotype can easily be explained based on these functions of the protein. It is quite intriguing to speculate that HSP27 might serve as a link between receptor tyrosine kinase signalling and the cytoskeletal remodelling required for cell motility.

Traditional loss-of-function based genetic analysis of the function of these proteins in *Drosophila* was quite difficult because of the presence of several closely related family members with partially overlapping functions. The fact that we were able to specifically identify the gene Hsp27 in our *slbo* suppressor screen was important in our case. It clearly demonstrated that using gain-of-function genetics we could really overcome the problem of genetic redundancy and also gain some insight of what these proteins can possibly do during normal conditions.

3.5 DETAILED ANALYSIS OF EP(X)1487

At the time these experiments were done, the complete sequence of the *Drosophila* genome was not available. In the majority of our potential suppressors we could not determine the exact site of the insertion only by searching the databases with the genomic sequence we obtained by plasmid rescue. This was the case with a line called EP(X)1487. It was not only one of our best suppressors, but it also behaved as we expected in our different tests. We could show that the suppression effect we saw in our primary tests was highly reproducible, and it was also Gal-4 dependent. We decided to learn more about the affected gene in this line.

3.5.1 Cytological localization

All we knew at that point was that the EP insertion genetically mapped to the X chromosome. In order to determine the exact cytological position of the EP insertion, Ann-Marie Voie performed a salivary gland in situ hybridization experiment using an element-specific probe in the EP(X)1487 stock. The results indicated that in this particular line a single EP element was inserted at the 9D3-4 cytological position on the X chromosome. (Later, these results were also confirmed by database search.)

Now we knew the approximate region where our transposon landed, but did not know what the affected gene was, and what was its possible function.

3.5.2 cDNA library screen

A cDNA library is a collection of plasmid clones or phage lysates, which contain recombinant DNA molecules, and can be screened for the presence of the desired clone if a suitable nucleic acid probe is available. The most common and reliable method of screening cDNA libraries for a clone of interest is based on the nucleic acid hybridisation technique. As a probe one can use a synthetic, radioactively labelled single-stranded oligonucleotide that corresponds to the region of interest.

We decided to identify the gene that was affected in the EP1487 stock using the 8 kb long radioactively labelled genomic plasmid rescue fragment as a probe on different cDNA libraries.

We screened a lambda-ZAP ovary library (a generous gift of Allan Spradling). Looking at approximately 500.000 plaques we found a single cDNA clone what overlapped at its 5' end with 111 bp of the plasmid rescue sequence. We named it as *sprint-a* for reasons I will explain in later chapters (Genbank accession number: AF312692). Sequencing of this clone revealed that unfortunately it did not have a poly-A tail, indicating that it was not a full-length one.

In order to find a complete cDNA clone, we tried to screen two other cDNA libraries with the same probe. One of them was an embryonic library. We got a number of positives screening a couple of hundred thousand plaques with the same probe as before, but sequencing revealed that the clones were most likely genomic contaminations. From the disc library (gift from Stephen Cohen) we did not obtain any positives.

The next thing we did was to isolate additional cDNA clones from the ovary library, but this time using *sprint-a* as a probe. We reasoned that the RNA should be present in the ovary during oogenesis, but it is probably quite rare. We concluded that based on the fact that we found only a single partial clone screening through 500.000 plaques. In the second round we identified several further potential candidates after screening through a few hundred thousand plaques. Sequencing revealed that the majority of them were unfortunately different shorter or identical versions of the original *sprint-a* clone (Figure 21).

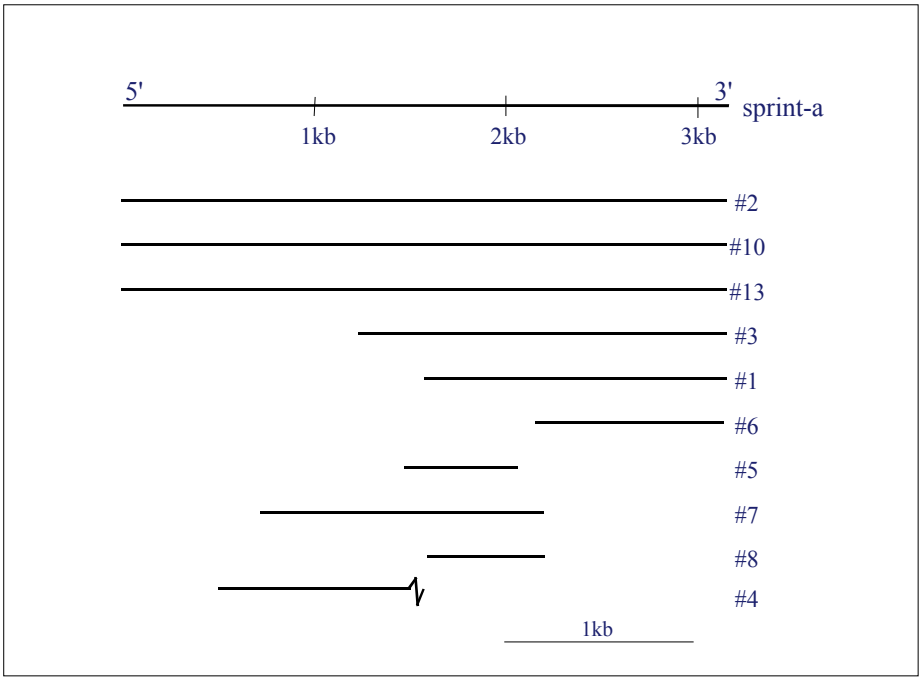

Figure 21: Structure of the different cDNA clones that were isolated from the lambda-Zap ovary cDNA library using the *sprint-a* DNA.

There was one clone, however, that seemed to be promising, as it was much longer than the others. We named it as *sprint-b* (Genbank accession number: AF312693). Sequencing of this clone showed interesting things. It contained a poly-a tail suggesting that it was indeed a full-length version, but its sequence was not

completely identical with the *sprint-a* sequence; they had different 5' ends. After this about 1 kb long variable region, they shared an approximately 2 kb long matching piece (Figure 22).

There was no sequence overlap between the genomic plasmid rescue piece and *sprint-b*, but this explained why we could isolate this clone only when we probed the ovary library with the *sprint-a* DNA.

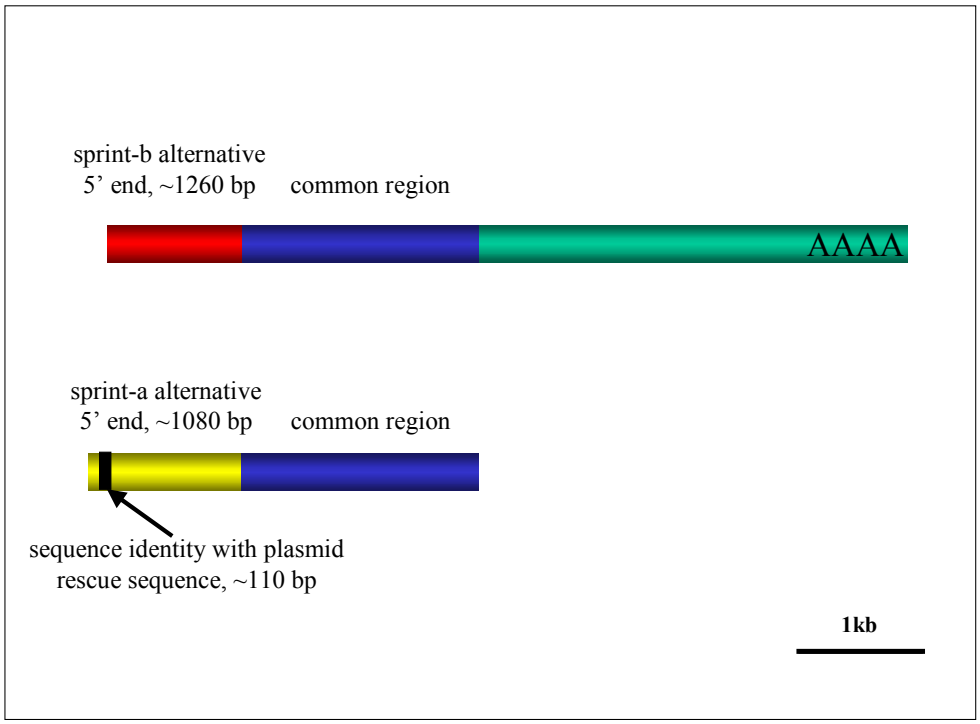

Figure 22: Sprint-a and sprint-b had different, alternative 5' ends. After this about 1 kb long variable region, they shared an approximately 2 kb long matching piece. Sprint-b did have a poly-A tail, while the original sprint-a was a partial clone. In the sprint-a alternative 5' end there was a region that matched to a piece in the genomic plasmid rescue fragment.

We could not explain all these interesting features at this point, only after the detailed structure of the sprint locus became known, but all these data suggested two possibilities. Either the *sprint-b* clone was some kind of cloning artefact that was created while the cDNA library was generated, or the different 5' ends could have

suggested the possibility of alternative splicing events on this end of the gene. In order to distinguish between these two probabilities we decided to carry out a Northern blot analysis.

3.5.3 Northern analysis

The first thing we wanted to prove was whether it was really the *sprint* locus that we overexpressed in the EP(X)1487 stock. For this we prepared poly $(a)^+$ mRNA from 0 to 16 hour old embryos in which the overexpression of the affected locus was forced using the 32B-Gal4 embryonic driver. As a control we isolated poly $(a)^+$ mRNA from same age embryos that contained the driver construct in the absence of the EP element, and also from wild type ovary samples. To probe this Northern we hybridised to a probe from the complete *sprint-b* cDNA, most of which overlapped with *sprint-a*. The results are shown in Figure 23.

We detected a quite strongly induced, about 8 kb long mRNA that was only present in the EP(X)1487 embryos. Multiple, smaller size bands were visible in all three samples, corresponding to the endogenous *sprint* mRNA isoforms. This suggested the existence of different alternatively spliced variants even in wild type animals during normal developmental conditions.

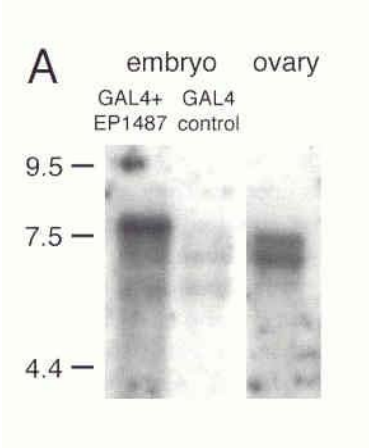

Figure 23: Northern blot analysis of *sprint* mRNA expression. *sprint* is induced by GAL4 activation of EP(X)1487. PolyA+ mRNA from 0 to 16h EP1487/+; 32B-GAL4/+ embryos, +/+; 32B-GAL4/+ embryos, and wild type ovaries hybridised to a probe from the complete *sprint-b* cDNA, most of which overlaps with *sprint-a*.

After that we made another Northern blot where we isolated poly $(a)^+$ mRNA samples from wild type ovaries, and different developmental stage embryos (Figure 24).

When a *Drosophila* embryo forms, almost up to the cellular blastoderm stage development depends largely on stocks of maternal mRNA and protein that accumulated in the eggs before fertilization. After cellularization, zygotic transcription increases dramatically. In order to distinguish between these two states we prepared poly $(a)^+$ mRNA from early (0-2 hour old) embryos to check if the *sprint* mRNA was maternally deposited to the eggs, and also from older embryos to look for the presence of zygotic transcripts.

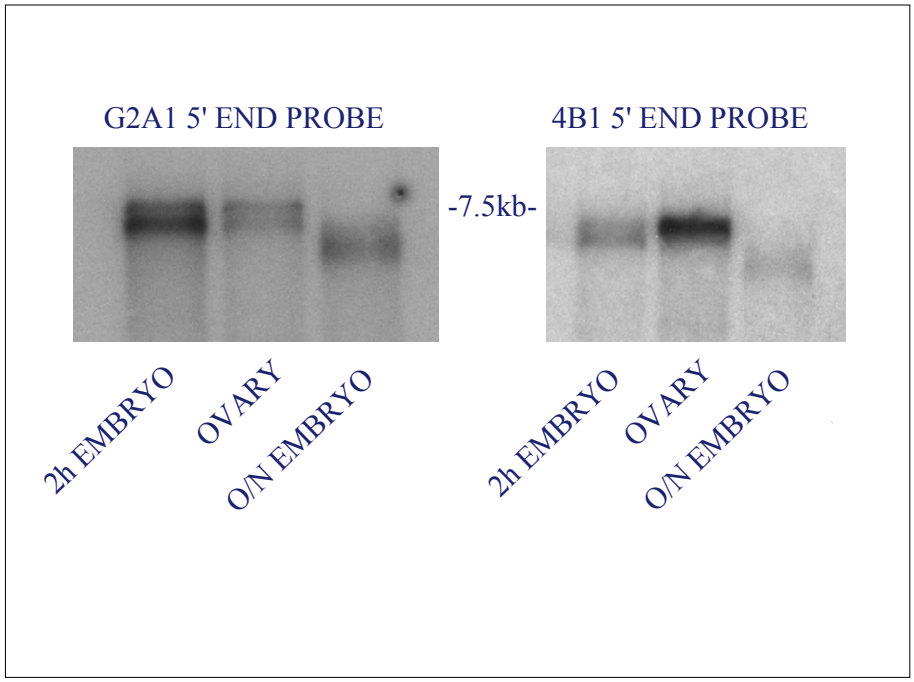

Figure 24: Northern blots were probed using the unique 5' regions of the *sprint-a* and *sprint-b* clones. Results indicated that both of the alternative 5' ends are used to generate a full-length message. Possibly alternative-splicing events both on the 5' and the 3' end of the gene lead to the formation of the different size transcript in the various lanes.
The mRNAs are present in the adult ovary, deposited to the eggs as a maternal contribution (0-2 hour embryos), and later they are zygotically expressed in older embryos.

As I mentioned above, based on our previous results we suspected that alternative splicing events caused the appearance of multiple bands we detected in the previous experiment. We also had two potential cDNA isoforms (*sprint-a* and *sprint-b*) that we isolated in the cDNA library screens, where *sprint-a* was only a partial one. So far we did not have any clear experimental proof that would confirm that either of them was only some kind of cloning artefact. In order to find an answer to these questions, we probed our Northern blots using the unique 5' regions of the *sprint-a* and *sprint-b* clones separately. The *sprint-a* mRNA was expected to be 200 bp longer than the spri*nt-b*. This was estimated based on the size of the different 5' ends.

Our results indicated that both of the alternative 5' ends were indeed used to generate a full-length message, because both probes labelled mRNAs with comparable size ranging from 6 to 7.5 kb (as one can see in Figure 24).

We detected multiple transcripts in every analysed developmental stage with both probes again. This suggested that the gene we were looking for was very complex, and probably alternative-splicing events both on the 5' and the 3' end of it led to the formation of the different size transcript in the various lanes. We had direct evidence now that at least 2 different alternative 5' ends were used to generate mRNAs, and later it was shown that full-length *sprint–b* cDNAs had at least two different 3' non-coding UTRs, and additional splice variants existed (P. Rorth personal communication).

The mRNAs were present in the adult ovary, deposited to the eggs as a maternal contribution (0-2 hour embryos), and later they were zygotically expressed in older embryos. This we judged based on the size difference of the labelled mRNAs in the 0-2 hour versus the older embryo samples.

The intensity of the signals of the different mRNA isoforms varied in the discrete developmental stages indicating that the expression pattern of the *sprint* gene is very dynamic and probably under tight regulation (Szabo, Jekely, et al. 2001).

3.5.4 *sprint* gene

Translation of the full-length *sprint-a* and *sprint-b* isoforms result proteins of 1790 and 1776 amino acid long, respectively, where the C-terminal 1523 amino acids are identical (Figure 25).

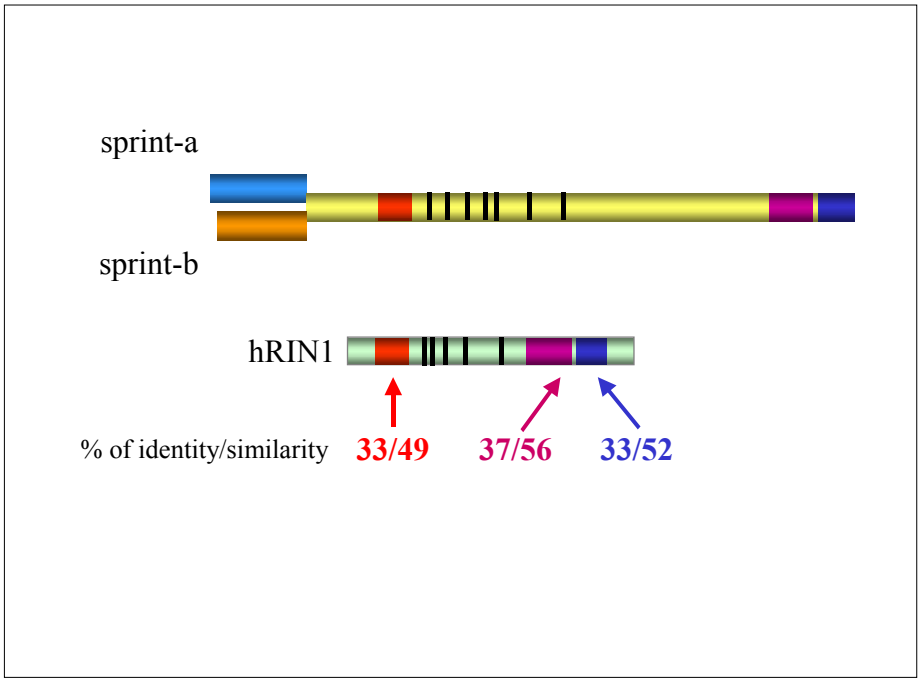

Figure 25: Structure of the Sprint protein, (Sprint-a and Sprint-b isoforms) and homology to the human RIN1 protein. The percentages of amino acid identity and similarity between the conserved domains of the two proteins are indicated on the figure.

Database search indicated that the Sprint protein was very similar to a mammalian protein named RIN1 (Han, Colicelli 1995).

The human version of RIN1 was isolated among several proteins that can cause interference with the Ras signal transduction pathway in the yeast *Saccharomyces cerevisiae*. The Ras genes encode signal-transducing guanine nucleotide-binding proteins that are involved in various processes during eucaryote differentiation. Ras proteins are known to make physical contact with a variety of proteins classified either as regulators of Ras activity or signalling effectors. The regulators of Ras include the GTPase-activating proteins (GAPs) and guanine nucleotide exchange factors (GEFs). These can directly stimulate the intrinsic properties of Ras that control its guanine nucleotide-bound state and hence its signalling activity. The effectors on the other hand propagate signals from Ras (B.Alberts, D.Bray, et al. 2002).

RIN1 was found to suppress an activated RAS2 allele ($RAS2^{V19}$) when it was overexpressed in yeast. Later Han and Coliccelli (Han, Colicelli 1995) showed that this protein could interact directly *in vitro* and also *in vivo* with the yeast Ras allele Ras2p, and the human H-Ras, when it is in the activated (GTP bound) form. Based on these features they named this protein RIN1 (Ras interaction/interference protein).

Some characteristics of the interaction between the activated Ras and RIN1 were found to be very similar to the interaction between Ras and its well-known effector Raf1. For this reason it was suspected, that RIN1 was a new Ras effector (Han, Colicelli 1995; Han, Wong, et al. 1997). RIN 1 and Raf actually competed with each other in binding experiments. Later it had also been shown, that this interaction happened between the effector domain of Ras and the carboxyl-terminal of RIN1, further supporting this idea. Unfortunately, it was not possible for quite a long time to reveal the actual biological role of this newly found molecule in mammalian cells.

Knowing all these facts our results seemed even more interesting. We isolated the *Drosophila* homologue of RIN1 based on a phenotypic assay, so we immediately had very strong indication of what kind of biological role it might have. We suspected that it probably functioned during *Drosophila* oogenesis, but at this point it was just an assumption. Our screen was based on gain-of-function studies, so we needed further experimental evidence to prove that the endogen protein was really involved in border cell migration during normal conditions.

As a rin (rasputin) abbreviation was already used in Drosophila (Pazman, Mayes, et al. 2000) we named the new gene *sprint* (spri) for reasons I will explain in the next chapter (Szabo, Jekely, et al. 2001) .

3.5.5 Domain structure of *sprint*

We looked at the sequences of *hRIN1* and *sprint* and compared them to each other (Figure 26). Several different known domains were identified and studied in detail earlier in the case of *hRIN1* (Han, Colicelli 1995; Han, Wong, et al. 1997). Our results indicated that all of the known motives were not only present in *sprint*, but also their relative position to each other were the same, and there was a good homology between the corresponding human and *Drosophila* ones. (The percentage of amino acid identity and similarity between the different domains of RIN1 and *sprint* are shown in Figure 25.)

What kind of domains could we identify? First of all the Ras binding domain (RBD) that was important for the interaction with the activated Ras was close to the C-terminal end of the protein. Immediately adjacent to it we found a region with homology to VPS9 (Burd, Mustol, et al. 1996; Hama, Tall, et al. 1999) and RABEX (Horiuchi, Lippe, et al. 1997), which are exchange factors for the small GTPase Rab5. The human version, RIN1, has been proven to function as a GEF for Rab5 (Tall, Barbieri, et al. 2001).

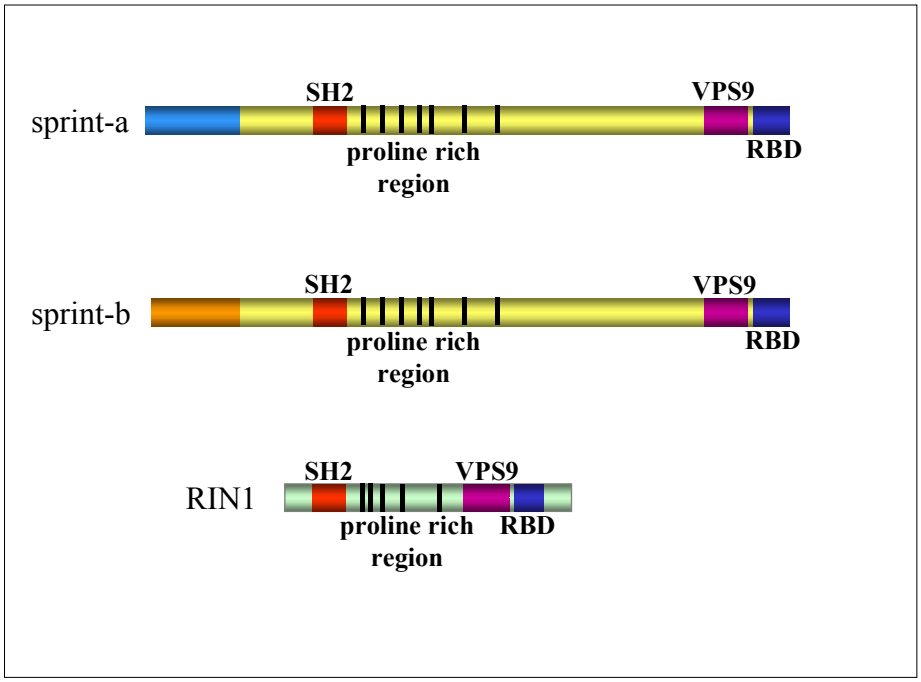

Figure 26: Domain structure of Sprint and RIN1. The description of the conserved regions can be found in the text.

Sprint also contained an SH2 domain (important for protein-protein interaction) near the N-terminus.

Finally the protein had proline rich stretches that can mediate interactions with the SH3 domain (mediates also protein-protein interaction) of the Abelson non-receptor tyrosine kinase (ABL) (Han, Wong, et al. 1997).

The name of the human gene, RIN1, is coming from its domain structure and function, for SH2, poly-proline containing Ras interactor.

The sequence seemed much less conserved outside of the above-mentioned motives between the two species. Note that in case of the *Drosophila sprint* gene the domains were farther from each other with more, presumably, spacer sequences between them.

No recognizable domains were found in the different alternative 5' regions of the Sprint isoforms.

We had also independent indications suggesting that the important domains were present in the common region. We cloned both the *sprint-a* and *–b* isoforms into pUAST expression vector. Transgenic lines were made, and then we overexpressed them with the slbo-GAL4 driver on a usual slbo mutant background. In both cases we could see the suppression of the border cell phenotype.

3.5.6 What does the different domains do in *hRIN1* and *sprint*?

Receptor endocytosis is facilitated, in part, by members of the small GTP-binding protein superfamily, including Ras, (Bar-Sagi, Feramisco 1986; Li, D'Souza-Schorey, et al. 1997; Barbieri, Kohn, et al. 1998), Rac/Rho, (Lamaze, Chuang, et al. 1996; Ellis, Mellor 2000), and Rab (Gorvel, Chavrier, et al. 1991; Bucci, Parton, et al. 1992). Rab5 is shown to be responsible for mediating membrane trafficking events early in the endocytotic pathway (Bucci, Parton, et al. 1992; Gorvel, Chavrier, et al. 1991). Cargo bearing early endocytotic vesicles fuse with early endosomes in a Rab5-dependent manner. Overexpression of Rab5 has also been shown to stimulate fluid phase endocytosis and EGFR uptake (Barbieri, Roberts, et al. 2000), and Ras overexpression has a similar effect (Bar-Sagi, Feramisco 1986). Though the exact mechanism by which this happens is not known, Ras stimulation seems to occur upstream of the action of Rab5 (Barbieri, Kohn, et al. 1998). *Sprint* seems to fit nicely into this model as it links biochemically the activities of Ras and Rab5, because in one

hand it can function as a downstream effector of the activated Ras, and on the other hand thorough the VPS9 domain it might function in the regulation of Rab5 activity by guanine nucleotide exchange.

I described in earlier chapters how and why endocytotic processes could have an important role during border cell migration. Experimental data presented by different laboratories based on tissue culture studies indicate that when the EGF is bound to its receptor, the ligand-receptor complexes are rapidly endocytosed (Haigler, McKanna, et al. 1979; Carpenter, Cohen 1990; Lemmon, Schlessinger 1994; Riese, Stern 1998; Wells 1999). When NR6 fibroblasts are transfected with RIN1 (the human homologue of *sprint*) the EGF uptake is markedly increased, whereas overexpression of the RIN1Δ inhibited this process (RIN1Δ is a naturally occurring splice variant of the RIN1 protein, containing an internal deletion, lacking completely the GEF activity). It has been proposed that the activity of RIN1 might be specifically coupled to certain receptors. The observation, that RIN1 expression does not effect the endocytosis of the transferring receptor seems to support this model (Tall, Barbieri, et al. 2001).

Duchek and Rorth (Duchek, Rorth 2001) showed that EGF receptor signaling is important during the second (dorsal) part of border cell migration, during the pathfinding process. It is very intriguing to speculate that that the Drosophila *sprint* could play a role in this process.

Further clues to the in vivo function of *sprint* will come from a greater understanding of its other signaling domains and the mediated interactions. In addition to its VPS domain (Burd, Mustol, et al. 1996; Hama, Tall, et al. 1999) *sprint* contains a proline rich region that mediates interaction with the tyrosine kinase Abelson (Abl) (P. Rorth personal communication). Note that the human RIN1 also shows preferential binding to the ABL-SH3 domain (Han, Wong, et al. 1997).

Drosophila Abelson has been shown to decrease axonal outgrowth when the activity of the *disabled* gene is also reduced (Gertler, Bennett, et al. 1980), and these proteins seems to have a conserved role in cell migration and axonal pathfinding, because disabled mutant mice also shows defects in neuronal migration (Howell, Gertler, et al. 1997). We have direct evidence that Abelson might have an important function during the migration of the anterior follicle cells, as one of our suppressors that we isolated in the original slbo interaction screen turned out to be an insertion in the Drosophila Abelson gene (*EP(3)3101*).

Another fact that supports this idea is that an additional protein called Enabled seems to show a slbo suppression effect, when we overexpress it with slboGal4 (unpublished data).

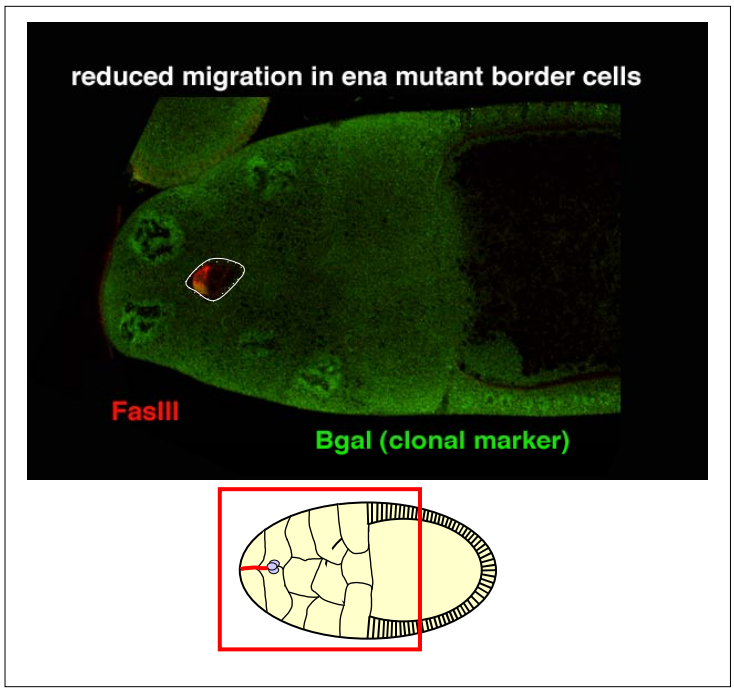

Figure 27: The ena mutant border cells (outlined with white, and also indicated by the lack of the green staining) are severely delayed in stage 10 mutant egg chambers. This indicates that wild type enabled has a clear role in border cell migration, and normally it promotes border cell migration.

We could also show that the cell migration is reduced in *ena* mutant border cells (Figure 27). Ena is a member of the VASP protein family. It can biochemically and genetically interact with Abl. Directly phosphorylated by it, and has been shown to be a suppressor of the Abelson tyrosine kinase (Gertler, Comer, et al. 1995; Gertler, Doctor, et al. 1990). Enabled has been implicated in the actin-polymerization driven movement of the Listeria cells (Laurent, Loisel, et al. 1999).

Actin cytoskeleton remodelling is known to be important during cell migration, and it can be a key procedure when moving cells interpret the different guidance cues and determine their direction based on this information. We know that EGF receptor signalling is key in part of the pathfinding process, and have indirect evidence that receptor mediated endocytosis can play a role in the regulation and recycling of this receptor.

All these data points towards a novel signal transduction pathway, in which Ras regulates endocytotic events by controlling the GEF activity of *sprint* and thus the activation of Rab5. (In case of the human homologue, RIN1, it has been proven that the interaction with the activated Ras through its binding domain facilitates the GEF activity of the protein (Tall, Barbieri, et al. 2001). In the meantime, through the interaction with the Abelson tyrosine kinase it presents a link toward the regulation of the actin cytoskeleton.

Further interactions with so far unknown partners can make the picture even more difficult and detailed trough interactions with the SH2 domain of *sprint*.

3.5.7 Functional homology

Earlier we showed that hRIN1, and the Drosophila *sprint* contained the same functional domains, and the sequence homology on the amino acid level was relatively high. We wanted to try if the human *RIN1* was a true, functional homologue of Drosophila *sprint*. To test this idea we obtained the full-length human cDNA, and cloned it into pUAST expression vector (Brand, Perrimon 1993). We generated transgenic flies carrying this construct, and overexpressed it on the usual $slbo^{1}$ background in the migrating border cells.

We were able to suppress -in some extent- the cell migration defect caused by the slbo mutation when we overexpressed the human cDNA. This further suggested that *sprint* was a real functional homologue of *hRIN1*, and the similarity was not restricted to the sequence level, but they could partially substitute for each other.

3.5.8 Genomic organization of the *sprint* locus

When the complete sequence of the Drosophila genome was published, we compared *sprint-a* and *sprint-b* cDNA sequences to the genomic ones in order to determine the genomic organization of the *sprint* locus. It turned out to be very complex, and encompasses 90 kb of the *Drosophila* genome (Figure 28).

sprint-a initiated close to the EP(X)1487 insertion, while another class of the transcripts (*sprint-b*) started at least 30 kb downstream. Different exons of the *sprint* gene had been predicted by sequence analysis to be separate genes previously by the Genome Project, and two additional short genes with no known homology seemed to be

present in different introns. They are both in the same orientation as *sprint* itself, and may represent alternatively spliced exons.

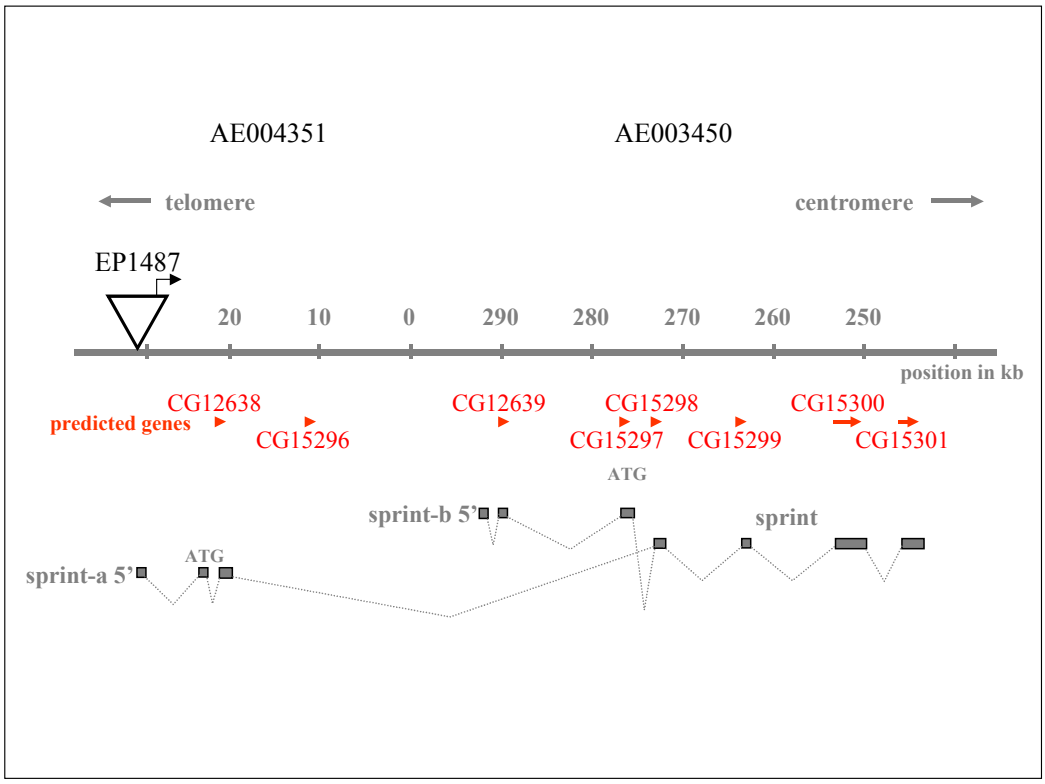

Figure 28: Genomic organisation of the sprint locus. It was derived from the published Drosophila genome sequence. All predicted genes in the region according to annotations from Berkeley Drosophila Genome project are indicated. Except for CG15296 and CG12639, each of these is included in the *sprint* cDNAs identified.

2.5.9 Antibodies

All these data I have presented so far were based on gain-of-function studies. Formally at this point one could not rule out the possibility without any doubt that the suppression of the *slbo* phenotype was due to some non-specific effects we caused by overexpressing EP(X)1487. The only data that argued against it was the fact that we

could detect the presence of the endogen *sprint* mRNA in wild type ovary samples by Northern analysis (see in chapter 2.4.3).

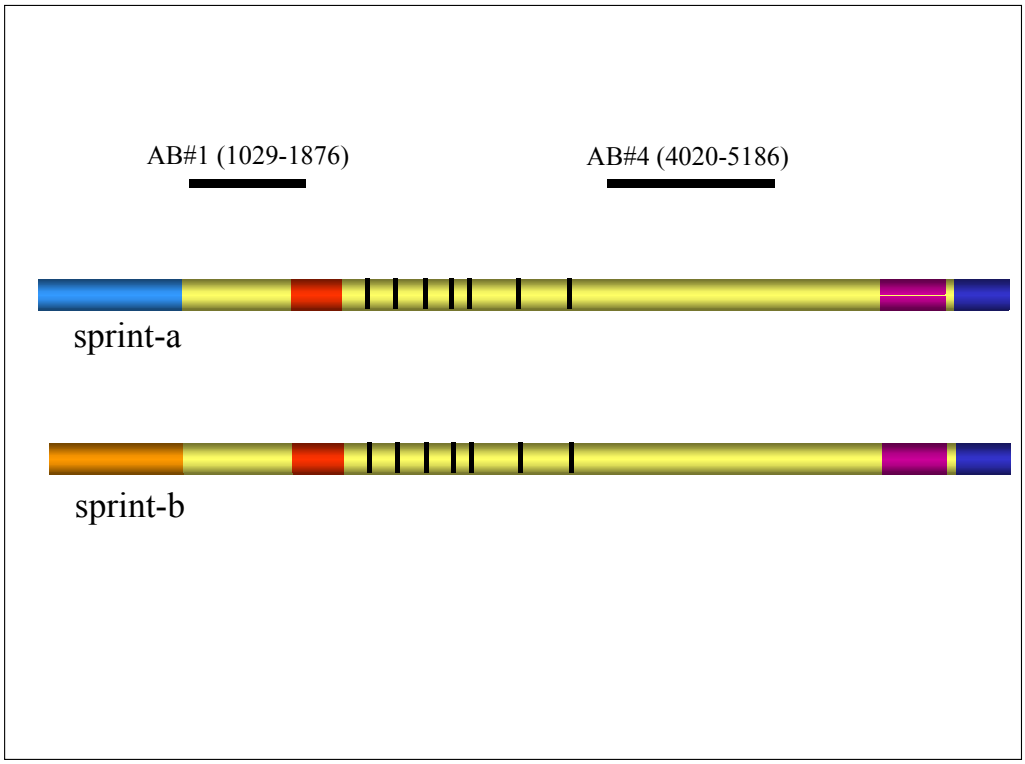

Figure 29: Polyclonal antibodies were raised against two different parts of the Sprint protein. Approximately 1 kb long pieces were chosen from the common region, in order to raise antibodies that can recognize both major Sprint isoforms.

In order to gain more information about the endogen sprint protein we decided to raise polyclonal antibodies against distinct parts of it (Figure 29). First of all we had to choose which parts of the gene we wanted to use to make the antibodies. We targeted four, approximately 1 kb long pieces, all of which were present in the common region of *sprint-a* and *sprint-b*. Potentially this allowed us to recognize both of the isoforms with all the resulting antibodies.

We recovered the chosen pieces either by polymerase chain reaction or by cutting them out from the *sprint-b* cDNA using the appropriate restriction sites, and cloned all of them into pGEX-4T1 inducible expression vector (Pharmacia). The recombinant proteins were expressed as fusion proteins together with the 26 kDa glutathione S-transferase (GST).

Because the Ptac promoter drove the expression of the cloned fusions in this particular vector, it was chemically inducible in E. coli host cells by the addition of IPTG to the bacterial cultures. The resulting GST fusions were relatively easy to purify from bacterial lysates using Glutathione Sepharose beads.

Fusion number one contained a region very close to the 5' end of the *sprint-b* clone (from base 1029 to 1876). We were able to successfully induce this fragment by IPTG. Test purification showed the induced protein was partially soluble, but part of it was closed into inclusion bodies in the cells. Because the soluble sample looked cleaner after purification, we ended up using this fraction for immunizing mice. (I will refer to the antibody we got using this sample as AB#1.)

In case of two other fragments (from base 4430 to 5536 and from 5186 to 6500 respectively) we were not able to induce their expression in E. coli. This can happen for example when the resulting protein is toxic for the bacteria.

Finally the last region we chose was closer to the 3' end of the *sprint-b* clone (from base 4020 to 5186). After successful induction we isolated the fusion protein from the spheroplast fraction in a relatively clean form. (I will refer to the antibody we made using this sample as AB#4.)

We injected 4 mice with each of the samples. After the initial immunization all of the animals got 3 more boosting injections in order to raise the antibody level in their blood. (This work was done by the Transgenic Animal Facility of

the EMBL, Heidelberg.) A week after the final immunization all of the animals were harvested, and their blood was taken. We prepared serum from each sample and used them to test if they contained antibodies that recognized the sprint protein by immunfluorescent microscopy.

3.5.10 Cellular localization

First we tested if the antibodies recognized the overexpressed Sprint protein. For this we stained ovaries of EP(X)1487 females where we drove the expression of the *sprint* locus using the slbo-Gal4 driver. One out of the 4 serum samples behaved the way we expected in case of AB#1 and AB#4 too. They strongly labelled the border- and also the follicle cells, and showed very little non-specific background staining. We could draw the conclusion based on these results that we had 2 independent polyclonal antibodies (one for each of the different regions of the sprint protein we targeted) that recognized the overexpressed protein.

The next thing we wanted to know if these antibodies could recognize the endogen Sprint protein, and whether it was present in the egg chambers. In order to test this we stained the ovaries of wild type (Oregon-R) females. We detected nice, clear staining in the migrating follicle cell populations; in the border- and centripetal follicle cells (Figure 30).

Detailed analysis of the stained samples revealed interesting facts. The endogen sprint protein was present in the above-mentioned cell types throughout their whole cytoplasm, and it was not excluded from the nucleus either.

Figure 30: We tested if the antibodies could recognize the endogen Sprint protein, and whether it was present in the egg chambers. Wild type ovaries from Oregon-R females were stained. We detected nice, clear staining in the migrating follicle cell populations; in the border- and centripetal follicle cells. The figure shows an example of a stained stage 10 egg chamber.

In some cases one could see stronger staining at the "business end" of the migrating cells, near the cortical region. Here, at the so-called leading edge, occasional accumulation of the staining was visible in membrane protrusions during the movement of the cells (Figure 31, 32).

These results were very encouraging for us. They provided further evidence that the wild type endogen protein was not only present in the right place at the right time (in the migrating cells of the Drosophila egg chamber during their movement), but also it showed an interesting subcellular localization.

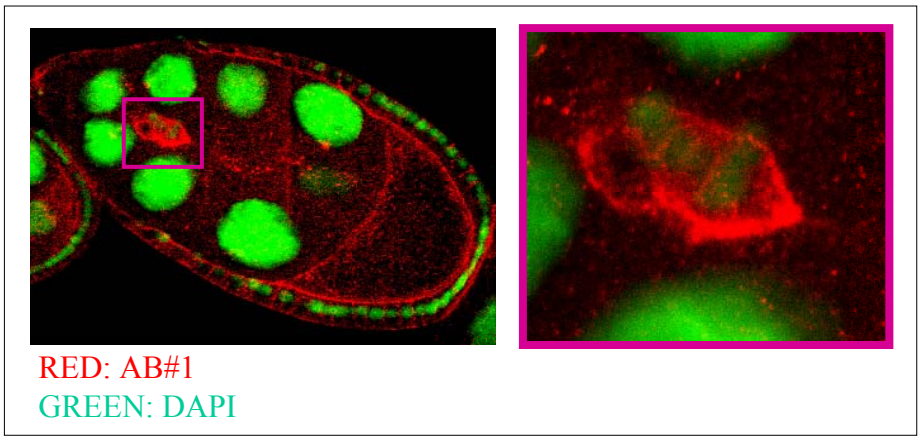

Figure 31: The endogene Sprint protein is present in the migrating border cells in wild type stage 9 egg chambers. The protein is present throughout the whole Cytoplasm, and not excluded entirely from the nucleus either.

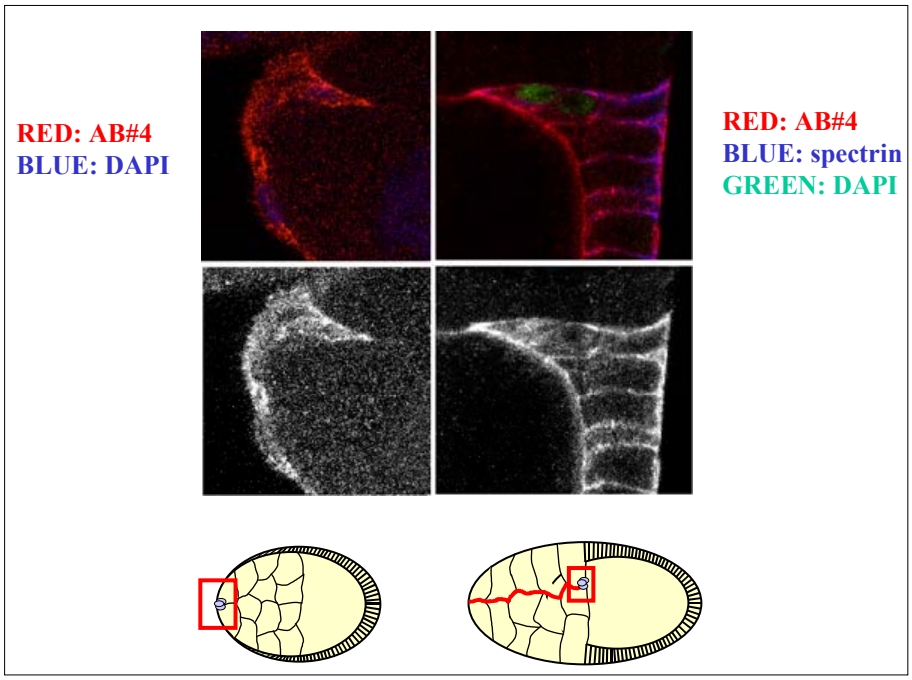

Figure 32: Nice, clear staining can be detected in the migrating follicle cell populations; in the border- and centripetal follicle cells in wild type egg chambers when stained with AB#4. Occasional accumulation of the staining at the so-called leading edge is visible in membrane protrusions during the movement of the cells.

3.5.11 Trying to generate imprecise excicions

Unlike mutations generated by chemical mutagenes or radiation, single P element insertions allow the rapid generation of new alleles of the affected genes by imprecise excision of the original element. Studying a range of mutant alleles that include nulls is frequently important for understanding gene function.

We tried to generate excision lines soon after we started the detailed analysis of the EP(X)1487 line. I mentioned earlier that at that point the complete sequence of the *Drosophila* was not available, so we did not have any information about the genomic organization of the *sprint* locus. Because of that, we followed the well-known general strategy for generating an excision allele.

We remobilised the EP element by crossing the jumpstarter line we originally used in our mutant screen to generate the EP(X)1487 flies. When the transposon leaves the locus where it was inserted, a break on the chromosome will be generated, which is repaired by the cellular repair machinery. In some percent at the end of this process the original sequence of the locus will be restored. This way we can generate so-called precise excisions. In other cases, however, the leaving transposon will remove part of the flanking genomic sequences from either side of it. In these cases the formation of different size deletions can be detected. The size of this deletion can vary form a couple of bases to even kilobases, and this causes the desired mutations.

We could detect the movement of the EP element after crossing the transposon to a Δ2-3 transposase source by looking at the eye colour of the progeny. We saw the appearance of mosaic-eye-flies. Crossing them to white background we could find white-eye-progeny in the next generation, where the transposon left the *sprint* locus.

When a deletion occurs in a certain locus, as a consequence of this event a phenotype can be generated. Depending on the biological function of our gene this can be for example lethality, male or female sterility. All the previously known genes with a role in border cell migration had lethal loss-of-function phenotypes, so we checked if any of our potential excision lines showed lethality. In some lines this was the case, but further analysis showed that none of them carried a mutation in the *sprint* gene itself.

We tried to generate imprecise excisions by remobilizing the EP(X)1487 insertion in 3 different experiments. During the first test we collected 118 white-eye males after the transposon was mobilized. We established stocks using them, and checked if any of the lines showed a homozygote lethal phenotype, but none of them did.

Next time we scaled up the experiment, and generated almost 300 white-eye males. 5 of the stocks showed lethality, but genomic Southern analysis proved that deletions were not affecting the sprint locus.

We got the same results when analysed 1021 excision males. They were unexpected and surprising at first, but we could easily interpret them as soon as the complete sequence of the Drosophila genome, and subsequently the organization of the sprint locus was published. Even if we delete the sequences next to the EP element, we only affect the *sprint-a* isoforms. *Sprint-b* initiates more than 30 kb downstream of the insertion site, and the two isoforms have overlapping expression patterns. Even if we knock out one of them the other can still function, and possibly no phenotype will be generated. The gene encompasses almost 90 kb of the Drosophila genome, so it is too big to delete the entire locus at once.

Now knowing a lot more about this gene, new, more fruitful experimental setups can be developed in order to isolate loss-of-function alleles.

3.6 DOMINANT SLBO SCREEN

As it was depicted above in Chapter 2.2.4, there are two major ways one can set up experiments using the EP system. In the last few chapters I described the results and our experiences of the *slbo* interaction screen. The other possibility was to perform a so-called direct screen. The two methods allow us to answer different questions. With the interaction screen we were searching for suppressors of the *slbo* mutant phenotype, genes that would make the mutant border cells move better. In the direct screen, on the other hand, we looked for genes that would stop border cell migration when they are overexpressed.

In the dominant *slbo* screen our collection was crossed to slbo-Gal4 driver in the presence of one copy of $slbo^{1310}$ enhancer trap. We performed the screen by checking the egg production of the females carrying all these elements. When the border cells fail to migrate properly, a channel-like structure called micropyle will not form properly during egg development. Although these eggs look perfectly normal, they cannot be fertilized, and as a result further development will not take place.

We allowed the females with the right genotype to lay eggs on apple juice plates in the presence of excessive amount of food and control males with normal fertility. The adults were removed after a day, and we placed the plates to a 25°C incubator. At this temperature normally developing eggs hatch in 24 hours after they were laid, so after this incubation period we placed them to a refrigerator in order to inhibit further development, and to slow down the crawling of the first instar larvae. The cold plates were scored using a dissecting scope, and by counting the number of intact eggs versus the empty eggshells we determined the rate of hatching.

In the second round of screening we retested all the lines where the ratio of unhatched eggs were greater than 30% using the same method.

Finally the ovaries of the selected females were dissected and X-gal stained, and we visually inspected the border cell migration in them using a dissecting scope. Out of the approximately 2000 tested lines only one single line, *EP(3)3519* showed a clear, reproducible phenotype (Rorth, Szabo, et al. 2000).

3.6.1 *tribbles*, a gene that was identified in the dominant *slbo* screen

The site of the EP insertion was determined by plasmid rescue and sequencing. It turned out that the corresponding gene was already cloned and named *tribbles* (Mata, Curado, et al. 2000). We verified our results by cloning the tribbles cDNA into a UAS vector (Brand, Perrimon 1993), and overexpressing the construct using the slboGal4 driver (Rorth, Szabo, et al. 1998). We could reproduce the same effect this way, which meant that indeed, we cloned the right gene. The effect was somewhat stronger than what we detected in case of the original EP line; by overexpressing of the cDNA we could cause the complete arrest of the border cells. This effect was unique to them; we did not detect any deleterious effect on the migrating centripetal follicle cells.

Further analysis done by Pernille Rorth showed that the Slbo protein level is under very tight control under physiological conditions. This regulation is accomplished both on transcriptional and posttranslational level. If the spatially and temporarily tightly regulated transcription is disturbed by forced overexpression of Slbo, it has deleterious effects in most tissues; independent of whether it is normally expressed in it at all. This can be explained by the elevated expression of the target genes, or by the

activation of inappropriate genes. Slbo has been shown to be quite promiscuous, meaning that it can bind to suboptimal or low affinity target sites if present in higher than normal concentrations. This is also true for other C/EBP family transcription factors.

Overexpression of tribbles caused a dramatic decrease in Slbo protein levels, so this explained nicely the effect we saw in the dominant screen; overexpression of this locus caused a complete stop of the migration of the anterior follicle cells. It was also shown that border cells that are mutant for tribbles have slightly elevated Slbo protein levels, indicating that tribbles might downregulate slbo during normal development.

Our results indicated that tribbles is a negative regulator of Slbo expression in migrating border cells, but the sequence of this gene does not suggest how this happens biochemically. They also suggest an unexpected importance of the tight regulation of the transcriptional factor levels (Rorth, Szabo, et al. 2000).

4. DISCUSSION

4.1 Model organisms

Why do we use model organisms such as Drosophila or C. elegans to study different very complex biological problems? Is there any near-term contribution from these invertebrate systems to the deeper understanding of human biology? How will the information that various Genome Projects gather help us?

While the human genome may contain approximately 37.000 genes (Xuan et al. 2003) they encode the components of only a couple of hundred biological processes, for example amino acid biosynthesis, protein synthesis, protein secretion, cell cycle regulation, signal transduction, cell-cell interaction, cell-substrate adhesion and so on. Of all the invertebrate model organisms the Drosophila shows one of the highest degree of structural conservation in its gene system to that of humans. Because of that, it seems to be quite reasonable to expect that most of the components of the above mentioned processes, the way they interact with one another, and also the developmental processes are conserved between flies and humans. Experimental data seems to confirm these expectations (Miklos, Rubin 1996).

But there are other reasons for using even invertebrate systems to gather more information that can be used in human studies. We have excellent experimental tools we can utilize in model organisms such as Saccharomyces, mice, Drosophila or C. elegans, but not in humans. They can greatly facilitate the discovery of new genes and their basic biological role, and also they can help us to organize them into pathways. Together with the sequence analysis of the human genome we can transfer this information to human biology, and reduce this 37.000 individual components encoded by the human genome

into a much smaller number of core processes of known biochemical function. For the precise knowledge of how these processes are working in man, and what kind of perturbations would lead to a development of different diseases we need, however, vertebrate or human studies.

4.2 EP collection

We established the first large collection of unselected P element insertions in collaboration with the Berkeley Drosophila Genome Project (BDGP) as part of its effort to understand gene function (Spradling, Stern, et al. 1995; Spradling, Stern, et al. 1999). Previously made similar collections of insertions were usually biased in some respect, because the lines were selected based on their phenotype (Bier, Vaessin, et al. 1989; Cooley, Kelley, et al. 1988; Deak, Omar, et al. 1997; Nusslein-Volhard, Wieschaus 1980; Torok, Tick, et al. 1993). In our experiments no selection other than the ability of the EP element to express the dominant eye colour marker it carries was applied. These features make it possible to statistically examine many properties of completely random P element insertions.

Gain-of-function screening based on misexpression phenotypes is an alternative to a loss-of-function screening approach to discover new genes. Based on the experiences of the previous P element screens, we know that it is impossible to mutate every single gene in the Drosophila genome because of several reasons I described in Chapter X, but with the above-described new system some of the previous drawbacks can be easily eliminated (Bier, Vaessin, et al. 1989; Cooley, Kelley, et al. 1988; Deak, Omar, et al. 1997; Nusslein-Volhard, Wieschaus 1980; Spradling, Stern, et al. 1995). I

think that one of the most useful features of the EP system is that we can detect and analyse now redundant genes, with partially overlapping function. Considering that the P elements are frequently inserted upstream of transcription start sites, some traditional P insertions might be close to a gene but remain silent, meaning that they do not disrupt its function. These kinds of mutations can also be potentially detected with the gain-of-function method, because ectopic expression of genes in various tissues would increase the probability of producing a phenotype, and the site preference of the P element insertions is in favour of causing misexpression of the entire gene.

I also showed an example of how the EP system can identify genes that expressed normally at very low levels, in a few cells, or only transiently during the fly development.

The EP element, we used in our gain-of-function studies turned out to be a well designed, and efficient vector, and behaved exactly as we expected. We were able to discover new genes based on their phenotype as for example in case of *sprint*, establish a specialized role for a highly redundant gene as the small heat shock gene *Hsp27*, and also to identify a new biological role for an otherwise known gene as for example in case of *α-adaptin* or *big brain*.

This new transposon seems to be quite versatile. Using it we can create not only classical loss-of-function and inducible gain-of-function mutants, but in some, however small percent of all cases it is also possible to generate inducible loss-of-function mutations as I showed in previous chapters.

The collection of EP lines that are currently publicly available is thought to account for <10% of the genes in the entire genome. In our screen, 88 of the 2082 EP lines generated some phenotype. If one were to assume that ~50% of the EP lines were oriented with the promoter toward the immediately adjacent ORF, this would imply that

8.5% of appropriately oriented insertions (88 of 1041) would have an effect on border cell migration. This is a rough estimate of the percentage of genes that can generate this phenotype since there are instances where more than one EP line has an insertion in the same locus. The genome sequencing effort identified 13,647 protein coding genes in the entire genome (Data from BDGP, Release 3.1, 2003 January). On the basis of the number of lines identified in our screen, we would predict that a screen of the entire genome would identify on the order of 1100 genes that could generate this phenotype

Naturally the EP element behaves also as a physical landmark in the genome, it allows the rapid isolation and identification of the mutated genes using different molecular biological techniques. The obtained insertion lines can serve as starting materials for loss-of function studies on the newly identified genes by creating imprecise excisions and new alleles of them.

In other words, the EP vector kept all of the useful features of previous P-element versions, but took one step further, and added some very useful new features that make it even more adaptable to special needs.

4.3 Comparison of the different screens

After screening 2082 EP lines in our slbo screen, we identified 88 lines (3.4%) that resulted in some rescue of the slbo mutant phenotype. Screens using other *GAL4* driver lines have elicited phenotypes in 2–7% of the lines screened (Abdelilah-Seyfried, Chan, et al. 2001; Huang, Rubin 2000; Kraut, Menon, et al. 2001; Pena-Rangel, Rodriguez, et al. 2002; Rorth, Szabo, et al. 1998; Tseng, Hariharan 2002). Thus, the number of lines identified in our screen is comparable to the numbers identified in these other ones.

It was very interesting to look at and compare the list of stocks that were identified in the various screens. By looking at the GAL4 drivers the different laboratories used, we can see that the cellular processes they are looking at are fairly diverse. Because of that, we would expect to find different set of genes in the various screens with relatively small overlap where the studied biological problems are dissimilar. The results seem to meet our expectations. As you can see in Table 2 there is practically no overlap for example between the results of the slbo screen (Rorth, Szabo, et al. 1998) compared to the other ones. The reason is that in our experiments we were looking at a cell migration process, rather than growth and patterning than in most of the other tests. On the other hand, we can see higher level of overlap for example between the results of the screens done with the sca-GAL4 driver (Abdelilah-Seyfried, Chan, et al. 2001) (to disrupt external sensory cell formation (SOP)) compared to the experiment done using pnr-GAL4 (Pena-Rangel, Rodriguez, et al. 2002) (looking at thorax formation). This was expected too, since the previous screen focused on SOP formation, a structure also present in the developing thorax, in other words the thorax formation is a more complex process that involves different aspects that are covered by the SOP screen too.

I think these results are very encouraging, and exactly as we expected. These data proves that the effects we are looking at are specific. What are the different factors we need to get specific results using the EP system?

First of all we possibly do not generate such high levels of ectopic transcripts that the cells cannot tolerate, and start to generate unspecific problems. For this one needs an inducible construct, which causes only moderate overexpression of the different genes. The EP vector turned out to be right for this purpose. In order to design a successful experiment using this system, one also needs to have a suitable driver that

works in the right cells at the right time. One also has to design the experiments very carefully, especially in case of the interaction screens. In the slbo screen, for example, we tried to suppress a mutation caused by the reduced levels of a transcription factor by overexpressing potential effector genes. It might be easier then to suppress a mutation caused by another effector gene. The fact that we used an intermediate slbo mutant allele in our scheme might be also important. slbo^{1} caused only the reduction of the border cell migration, but did not stop the cells. It is probably easier to make the already moving cells to migrate better, than to make them move at all, as the initiation of migration is a completely different process.

4.4 *sprint*

I showed, that using the EP system we were able to identify and clone a gene that turned out to be a *Drosophila* homologue of the human *RIN1*. Since the discovery of *RIN1* in 1995 by Han and Colicelli (Han, Colicelli 1995) in a yeast two-hybrid screen. Over time there was a nicely growing amount of biochemical data describing the interactions with its different binding partners, but no data on its real physiological function *in vivo* was discovered.

Because the *Drosophila* homologue, that we named *sprint*, was discovered in a mutant screen based on its slbo suppression phenotype, we immediately had a strong indication of what its function was. Based on the gathered biochemical data presented by other laboratories,

and the results of our own experiments, we could build up a possible, very intriguing model describing the exact role of the gene during border cell migration. Further, more specific experiments can be designed in order to try to prove the predictions.

4.5 EP versus GS

Assuming that the orientation of the element is completely random upon insertion, statistically in only about 50% of all cases will the EP sit in the right orientation with regard to the orientation of the affected genes. This problem seemed to be solved soon after the EP element was published, because a Japanese group developed a construct with a strikingly similar design (Toba, Ohsako, et al. 1999). They named their new vector GS (gene search) element, and the only difference compared to the EP vector is that in their construction the promoter and enhancer sequences are present at both sides of the transposon. This special modification can raise the probability of inducing a forced phenotype. Normally one would expect approximately two-fold increase in the percentage of detected mutants using the GS transposon, but interestingly the Japanese group reported an unexpectedly robust effect. They detected an extremely high frequency of Gal4 dependent phenotypes; in case of the GS element an average 10 fold higher using the same drivers than what we saw testing the EP insertions. For example they saw visible eye phenotypes in 38% of all their lines with sev-Gal4 driver, compared to 4% in case of the EP lines. Likewise, 32% of the GS inserts showed wing phenotype with dpp-Gal4, while only 2% of the EP inserts had phenotypes with the same driver.

What they claim is that practically every third insertion shows some kind of induced mutations, no matter which driver they used in the experiments. This is hard to

imagine, because in some percentage the transposon probably does not even land in a right position or distance where there is a gene or ORF in either direction to be overexpressed.

In case of the slbo suppressors I checked the position of the EP transposon within the affected genes by database search. In 22% of the examined cases the EP landed upstream of a gene or ORF. In 9.1% it was in the first exon, in 8% in the first intron, and in another 9.1% landed in another transposable element. 2.3% of all lines contained EP insertions downstream of coding sequence; in these cases most likely antisense transcript was induced. Altogether these lines account for 50.5% of all the slbo suppressor lines. The remaining 45.5% contains stocks where there is either no gene in a 5 kb distance in either direction of the transposon; not entirely sure what is the affected gene because the insertion can have an effect on more than one ORF because of its special position; or only the affected gene is known but so far the transposon is not placed onto the graphical map I searched. Based on our results we can see that in case of the EP element the average distance between the inducible promoter and the overexpressed gene or ORF was never greater than ~1kb in the lines where we detected phenotypes.

	omb	dpp	sev	slbo	ey	pnr	sca
	163 (7.1%)	47 (2%)	98 (3%)	60 (4.2%)	53 (2.3%)	172 (7.5%)	104 (4.5%)
overlap omb	163/163	28/163	48/163	2/163	-	-	-
overlap dpp	28/47	47/47	28/47	0/47	-	-	-
overlap sev	48/98	28/98	98/98	2/98	-	-	-
overlap slbo	2/60	0/60	2/60	2/60	1/60	9/60	3/60
overlap ey	-	-	-	1/53	53/53	18/53	22/53
overlap pnr	-	-	-	9/172	18/172	172/172	41/172
overlap sca	-	-	-	3/104	22/104	41/104	104/104

Table 2. Comparing the results of the different gain-of-function screens.

If we keep this in mind and look at the results again, they look even more striking. Could it be possible, that almost every second or third randomly tested gene would have a real physiological role in processes like eye development or wing pattern formation?

The difference between the two sets of results is even bigger if we only look at the induced lethals. For example 0.3% for dpp-Gal4 in case of the EP element compared to 14% of all the tested lines in the GS screen (Table 2).

How can we explain these differences? The high frequency of mutant phenotypes can only be attributed to the mirror symmetric structure of the GS vector. The two promoters and enhancers are within a very short distance; so there might be some cross talk between them. The enhancer can work in either direction meaning that the enhancer on the 5' end of the GS element will possibly have an effect on the promoter sequences both on the 3' and the 5' end of the vector. As a result of that elevated efficiency of forced expression can be detected.

Whether this strong effect we can see using the GS system is useful or not is questionable. The very high occurrence of lethal mutations using different drivers suggest that this transposon might drive so high levels of the affected genes that it causes unspecific problems within the cells. It is also difficult to understand how and why a restricted overexpression of different genes in non-essential organs like the wing or the eyes would lead to the death of the animals in approximately 14% of all tested cases. One possible explanation is that extremely high levels of certain proteins might cause unspecific interactions that would never happen under physiological conditions. This lead to cell death, abnormal signalling to otherwise unaffected tissues, which would finally lead to severe developmental problems and subsequently the death of the animal. Formally it is also possible that the GS element is more promiscuous, meaning

that the level of background transcription is considerable in cell types where the element should not work according to the driver that was used for a specific experiment.

The fact I wanted to point out in this chapter is that in case of the GS element using a regular driver one can possibly induce so high expression levels that are possibly deleterious for the organism. Further tests would be necessary to show weather this is true or not, but if so, this system lacks one of the most useful features the EP element was designed for. Namely, we can induce misexpression of different genes in a very restricted manner, e.g. during a short period of the animals' development, or in a well-defined set of cells, and look at the phenotypic consequences of this relatively small perturbation in the target tissue. In the slbo screen I described earlier, we generated induced mutations only in six to ten cells, but the rest of the animal remained "unaffected". This way we could study the effect of otherwise lethal genes in this process. Also, an induced lethal phenotype does not give too much extra information about a certain gene's role in a specific process So even though the GS system seems to cure one of the disadvantages of the EP system, it also looses one of the most useful features of it. Mild overexpression of different genes probably leads to less unspecific interactions and subsequently smaller perturbations in the cells. As a result the phenotypic consequences will be more specific, and they give more detailed and relevant information about the studied processes.

Please note that the comparisons of the different screen results I made in this chapter are arbitrary. Although the same drivers were used in the different experiments, the tested collections and the experimental conditions were different. The basis of my comparison is that both collections contain completely random, unselected transposon insertion events that were tested with the same drivers.

4.6 Other use of the EP collection

It is known for quite a long time that P elements insert nonrandomly, but the factors that influence this specificity were not known. We knew the existence of so called cold and hot spots that are avoided and preferred by this transposon, respectively. There is a known preference for sites that are more accessible; euchromatic sites seems to be favoured over heterochromatic, and there is a tendency to integrate at the 5'-end of genes (Berg, Spradling 1991; Liao, Rehm, et al. 2000). This is not very surprising; these are the same chromosomal regions that must be accessed by the transcriptional control machinery.

Until the EP lines were made, it was relatively difficult to study the true target preference of the P element. The previous, publicly available, large-scale mutant collections were not unselected, meaning that the stocks were usually chosen based on some phenotype. As I mentioned before, we applied no selection other than the stock had to express the yellow eye colour marker the EP element carries. Because of the size and the properties of our set of mutants, data obtained from these lines allows the application of statistical methods to gain more insight about these processes.

New data gathered by analysing the EP insertions supports the previously suspected idea that the nonrandom target preference of P elements is caused by DNA structural properties and chromatin accessibility rather than primary DNA sequence (Liao, Rehm, et al. 2000). The preferred target sites (mostly within a few hundred base pairs of the transcription start site) are the same regions that must be accessed by the transcriptional control machinery, and even within these open regions of chromatin P insertion does not appear to be random. There is a significantly higher than average GC

content, and also several measures of DNA physical properties (e.g. DNA bendability, B-DNA twist, protein-induced deformability) show specific patterns at the target sites.

It has been shoved previously that P element insertion causes an 8-bp target site duplication (O'Hare, Rubin 1983). Although there are base preferences at each position, these are not strong enough to generate a clear consensus sequence. There is on the other hand a 14-bp palindrome pattern of hydrogen donor and acceptor sites (containing the 8-bp target duplication and 3 bases at both sides). This can be significant, because it is known, that many protein-DNA interactions occur through hydrogen bonding between amino acid side-chains and sites in the major groove of the DNA double helix.

All these data verifies, that the EP collection allows not only the identification and characterization of new genes, but it can be useful to gain more insight into the exact mechanism and properties of P element insertion and its target preference.

5. CONCLUSIONS

In this present thesis I described a work in which we wanted to look for genes playing important role in border cell migration. In order to gain a better understanding of this process we made a set of transposon-containing mutant fly-stocks using a new tool that was just developed at that time by Pernille Rorth. We set up a screen in which we could successfully isolate a panel of genes that we suspected to play important roles during the movement of these cells. During all this studies we proved that the EP system works as it was expected, and it is capable to identify new genes or new function of well-known old ones. We performed a more detailed analysis of the few, most interesting looking genes, cloned some of them that were unknown to that date, and described one, we named *sprint*, in more detail. We showed that gain-of-function studies, indeed, are useful, and can solve problems that are difficult to address with purely loss-of-function analysis. By making the stocks publicly available through the Berkeley Dropsophila Genome Project we provided a good-size stock collection for the entire fly-community in order to perform similar gain-of-function studies. Comparing the results of different laboratories we were able to gather further experimental data that shows that our EP system gives specific, biologically relevant data when it is used in a carefully designed experimental setup.

In this, so-called, post genomic era the main goal of the biologists is to understand complex biological networks, regulatory mechanisms rather than studying single genes and their functions. Some says that one way to address new problems in a not too distant future is that using the gathered tools in the model organisms now we can introduce even multiple modifications and look at the effects on the level of the whole organism. Our work is a nice example of this idea. We tried to find unknown

players that act during border cell migration by genetic interactions, in other words introduced random modifications on an already mutant genetic background, and studied the effect of these multiple alterations.

I believe that this work, among a lot of others, confirms that studies made on different animal models are useful, and will always stay very valuable in biomedical research.

6. SUMMARY

1. During the course of this work we established the first large-scale mutant collection using the EP element that allows systematic overexpression of the genomic regions neighboring the vector. We introduced several technical improvements compared to previous screens. Because of our setup, we were able to generate a large number of X chromosome insertions that were underrepresented in earlier experiments. Also we induced the movement of the EP element in both sexes, in order to allow comparative studies on the mechanisms of transposition in male versus female germline.
2. We generated 2296 new transposon insertions, established the stocks, mapped the insertions into chromosomes, and determined if the EP element itself resulted a recessive lethal phenotype. In 413 stocks the transposon mapped to the X, in 939 lines to the second, and in almost an equal number of cases (944) to the third chromosome.
3. We set up two different types of screens in order to study the migration of the anterior follicle cells during Drosophila oogenesis, and to find genes that have a role in this process.
4. In the slbo interaction screen we looked for genes that can suppress the female sterility, caused by the homozygote $slbo^{1}$ mutation, when we overexpressed them in the migrating border cells. In the primary screen we identified 88 potential slbo suppressors, what we further analyzed. We determined the rate of the border cell migration in all 88 stocks by lacZ staining the ovaries of the females, and counting the percentage of stage 10 egg chambers where we could detect any border cell migration. The majority of the lines (60 out of 88) showed a reproducible increase in the rate of migration relative to the control.

5. We made sure that the suppression effect was due to the Gal4 dependent overexpression of the different target genes in various test experiments. Also checked if the transposon insertion itself had any effect on the border cell migration in the absence of the driver, suggesting the role of the gene in this process under physiological conditions.
6. We obtained genomic DNA fragments flanking the EP element in the 60 potential suppressors, sequenced them, and searched the Drosophila sequence database in order to determine the site of the transposon insertions, and the affected genes.
7. By analyzing the data we got, we were able to draw a couple of conclusions:

   -using the EP system we can detect and analyze redundant genes with partially overlapping function

   -it can identify genes that are expressed normally at very low levels, in a few cells, or only transiently during the fly development

   -although with a relatively low frequency, it is possible to create conditional loss-of-function mutations using this system.
8. We cloned a gene that was affected in the EP(X)1487 stock by screening the lambda-ZAP ovary cDNA library. There were two different cDNA isoforms that we identified, and named as sprint-a and sprint-b.
9. We proved that it was really the sprint locus that was overexpressed in the EP(X)1487 stocks by Northern analysis. The analysis of wild type ovary samples showed that both sprint-a and sprint-b were present in them. We detected multiple transcripts in every analyzed developmental stage, that suggested that the sprint gene were very complex, and alternative splicing events led to the formation of multiple bands in the different lanes.

10. Database search proved that the Sprint protein was very similar to a mammalian protein named RIN1. Using the human RIN1 cDNA we were able to partially rescue the border cell migration defects caused by the homozygote slbo mutation. This suggested, that the human RIN1 was not only sequence, but also a functional homologue of the sprint gene.
11. Performing sequence comparisons we showed, that all the previously identified domains of the mammalian RIN1 protein were present also in the Drosophila Sprint.
12. We raised polyclonal antibodies against two different regions of the Sprint protein. It turned out that both serum samples indeed contained antibodies that recognized both the overexpressed and the wild type endogen Sprint proteins. Sprint was not only present in the ovaries of the females, but also its subcellular localization seemed to be very interesting. It was present in the migrating border cells, and we could detect an occasional accumulation of the staining at the protrusions of the leading edge.
13. In the dominant slbo screen we identified a gene called tribbles, and showed that this gene is a negative regulator of Slbo expression in migrating border cells.

7. MATERIALS AND METHODS

7.1 Fly stocks and their maintenance

For maintaining the *Drosophila* stocks we used standard fly food, containing cornmeal, dried yeast, agar, and sugar. The stocks were kept under controlled humidity either at 18°C, room temperature or at 25°C, depending on the actual experiment.

The detailed description of the different markers, mutations, and balancer chromosomes we used for our mutagenesis screen can be found in the work of Lindsley and Zimm {Lindsley, Zimm 1992).

7.2 Embryonic transformation of Drosophila embryos

Certain stocks (e.g. slbo-Gal4 [made by Pernille Rorth], hRIN1-pUAST, sprint-a-pUAST, sprint-b-pUAST) were generated by embryonic transformation of the fly embryos using artificially generated constructs. The purified plasmids were mixed with Δ2-3 helper plasmid (1:1 ratio) in injecton buffer (5mM KCl, 0.1mM K-phosphate, pH=6.8). The females were well-fed before the egg collection. The w^{1118} females were kept at room temperature in a dark drawer, and the injection was done at 18°C.

Eggs were collected for one hour. The embryos were washed with distilled water, and their chorion was removed using diluted commercial bleach (1:1 ratio). They were dried in a desiccator, and covered with halocharbon oil to prevent further drying.

After injecting the DNA, the eggs were placed in wet chamber at 25°C, and the hatching larvae were collected, and placed onto regular flyfood in the presence a drop of wet yeast.

The progeny was individually crossed to w^{1118} flies (Voie & Cohen 2003).

7.3 Interaction screen

2082 EP lines were crossed into the *slbo*^{1} background in the presence of the slbo-Gal4 driver. In the next generation females carrying both the driver and the EP element in the presence of the slbo mutation were tested for fertility using the following method. These females were placed into fresh vials in the presence of yeast and a few sibling males with normal fertility. They were kept at room temperature, and about 10 days later we visually inspected them. In a small percentage of the vials we could detect living larvae moving in the food. These lines were subjected to further analysis.

7.4 Dominant *slbo* screen

In the dominant slbo screen our EP collection was crossed to slboGal4 driver in the presence of one copy of slbo^{1310} enhancer trap. We performed the screen by checking the egg production of the females carrying all these elements. Females were laying eggs on apple juice plates in the presence of excessive amount of food and sibling males with normal fertility. The adults were removed after a day, and the plates were placed to a 25°C incubator for 24 hours to let the eggs hatch. After this incubation period the plates were placed to a refrigerator in order to inhibit further development, and to slow down the crawling of the first instar larvae. The cold plates were scored using a dissecting scope, and by counting the number of intact eggs versus the empty eggshells the rate of hatching was determined.

7.5 β-galactosidase staining of ovaries

The ovaries of female flies were dissected out in EBR. (The samples were kept on ice until further processing.) After the EBR was carefully removed, the ovaries were placed into 100 μl (500 μl) of fixation buffer for 10 minutes. Later they were washed in 100 μl (500 μl) PT for 10 minutes. After the PT was taken off, 100 μl (500 μl) staining solution was added. The samples were incubate either @37 °C for 1-3 hours, or at room temperature overnight. When the staining was finished, the ovaries were mounted in 100 μl (500 μl) 50% Glycerol.

7.6 Genomic DNA isolation

30-50 flies were grinded in 400ml DNA extraction buffer. The samples were placed into a microcentrifuge briefly. An extra 400ml of extraction buffer was added, and the samples were grinded again until they were well homogenized. The tubes were incubated at 65°C for 30 minutes. After the incubation 120ml 8M K-acetate was added to each sample. The tubes were placed on ice for 30 minutes, and then centrifuged at top speed in the microcentrifuge for 30 minutes. The supernatant was carefully removed, and the DNA was precipitated using 0.7 volume isopropanol. After 10 minutes of centrifugation the DNA was washed with 300 μl ice cold 70% ethanol. The DNA was resuspended in an appropriate volume of TE buffer. Optionally RN-ase can add to the TE buffer to get rid of the RNA contamination.

7.7 Plasmid rescue

Plasmid rescue was performed as it is described in *Drosophila Protocols* (p.429) (Edited by: Sullivan, Ashburner, et al. 2000).

7.8 Southern blot

The gel was cut to the desired size after electroforesis, and photographed using a ruler. Acid depurinarion followed in 0.25N HCl for 30 minutes. The gel was rinsed with dH_2O, and it was denatured 1x denaturation buffer twice, for 15 minutes. Neutralization followed in 0.025M $NaHPO_4$, 3 times for 10 minutes. During these various washing steps the nylon blotting membrane and the blotting paper was cut to the desired size. The membrane was presoaked in 0.025M $NaHPO_4$. The gel was transfer to GeneScreen in 5xSSC overnight using 500 ml transfer buffer, but one can use also 0.025M $NaHPO_4$ as transfer buffer with identical results.

DNA was crosslinked to the membrane on the next day. (This step is absolutely necessary for the high concentration SDS incubations.

7.9 Northern blot

The gel was cut to the desired size after electroforesis, photographed using a ruler, then it was rinsed with DEPC treated dH_2O briefly. During these washing steps the nylon blotting membrane and the blotting paper was cut to the desired size. The gel was transferred to GeneScreen paper in 10xSSC overnight using 500 ml transfer buffer.

DNA was crosslinked to the membrane on the next day. (This step is absolutely necessary for the high concentration SDS incubations.) Optionally the blot can be stained with methilene-blue solution after the crosslinking.)

7.10 Southern and Northern hybridisation experiments

After the crosslinking of the membrane it was placed into hybridisation tubes in the presence of 10 ml Church-Gilbert prehybridisation solution. At 65°C it was incubated at least for half an hour. (Prehybridisation can be longer, up to a couple of days if the hybridisation didn't work for the first time.)

After half an hour the labelled radionucleotide was added, and hybridize at 65°C with shaking overnight.

On the next day the membrane was washed in Buffer 1, twice, for 15 minutes at 65°C. Then it was washed in Buffer 2, up to five times, each time for 5 minutes at 65°C. Finally the membrane was blotted, and exposed.

7.11 Nucleic acid labelling experiments

All of the nucleic acid labeling experiments were performed using the Prime-It® Random Primer Labeling kit from Stratagene.

7.12 Salivary gland in situ hybridisation

Salivary gland in situ hybridisation experiments were done by special facilities at the EMBL Developmental Biology Programme.

7.13 cDNA library screening

Plating the Bacteriophage plaques

Dish size	Dish volume, bottom agar	Volume of bacterial lawn	Volume of bacterial lawn	Number of plaques per plate
90 mm	25 ml	0.2 ml	3 ml	10,000
150 mm	80 ml	0.6 ml	10 ml	40,000

On the first day an appropriate amount of bacteria lawn and bacteriophage particles were mixed in 15 ml snap cap tubes, and they were incubate at 37^{o}C for 20 minutes. Prevarmed molten (50^{o}C) top agarose (0.7%) was added carefully to the infected bacterial mixture, and was immediately poured onto prewarmed and dried agar plates. The surface of the bottom agar was covered. The agarose was allowed to harden and then the plates were inverted and incubated at 37^{o}C for 10-12 hrs. They were chilled the next day at 4^{o}C for 1 hr to allow top agarose to harden.

The plates were evenly covered using dry nitrocellulose circles, and they were marked asymmetrically using needles and black ink.

After 1 minute, using a blunted-ended forceps, the filters were peeled off the surface of the plate and was placed its DNA side up on wet 3MM paper that has been soaked in denaturing solution (1.5 M NaCl, 0.5 NaOH) for 1 minutes. The circles were transfered to neutralizing solution (1.5 M NaCl, 0.5 M Tris.Cl[pH 8.0]) for 5 minutes, and finally to 2X SSPE for 2 minutes. At the end they were placed on 3MM Whatman paper to dry.

After all the filters were dry, they were wrapped between sheets of Whatman 3MM paper. The DNA was fixed to the filter by baking for 1 hour at 80^{o}C in a vacuum oven.

Hybridisation was performed as it is described in chapter 6.10.

7.14 Cloning of hRIN1 into pUAST vector

The full-length hRIN1 cDNA was cloned into a pCDNA3 vector (Invitrogen). According to the map of this construct it was possible to release the cDNA using EcoRI and XbaI restriction enzymes. In the next step the fragment was directionally cloned into pUAST vector.

7.15 Immunstaining of Drosophila ovaries

The fly ovaries were dissected in EBR. The ovarioles should be teased apart and the epithelial sheath should be removed to allow better penetration of the antibodies. For fixing the them we used at least 200 μl devitellinizing buffer, and incubated for 10 minutes. Brief washes followed 3 times in order to wash out the fix with PT, and then we incubated the samples in PT for 30 minutes. To avoid non-specific reactions the ovaries were kept in PBT+N for 30 minutes. The staining with the primary antibody was performed by incubation at least for 4 hours (up to overnight) in the primary antibody diluted in PBT+N.

After this step the samples were washed in several changes of PBT for 2 hours. Non-specific binding was blocked again by incubating the samples in PBT+N. The ovaries were then placed to a mix of fluorescently labelled secondary antibody diluted in PT, and incubated for 4-6 hours. After a 2-hour long was in PT (in order to remove all unincorporated secondary antibodies) the samples can be mounted in mounting media, and examined under fluorescent microscope.

7.16 Generating polyclonal antibodies

For making the AB#1 the sprint-b cDNA was cut with SalI and HindIII restriction enzymes. Both of these enzymes recognize unique sites in the sprint-b clone at bases 1029 and 1876 respectively. An ~0.8 kb long piece was released, and was cloned into a pET-23a fusion vector. The fragment was released from it using SalI and NotI enzymes, and cloned into a pGEX-4T1vector. After checking by restriction enzyme digest that the cloning worked, we tried if the construct was inducible.

For making the AB#4 the sprint-b cDNA was cut with EcoRI and PstI restriction enzymes. An ~1.2 kb long piece was released, and was cloned into a pBluescript vector. The fragment was released from it using BamHI and EcoRI enzymes, and cloned into a pGEX-4T1vector. After checking by restriction enzyme digest that the cloning worked, we tried if the construct was inducible.

7.17 Checking the inducibility of the fusion constructs

For this we inoculated 15 ml liquid LB media using overnight 2 ml bacterial cultures. We grew the culture for about 2 hours, when the density reached approximately OD_{600}=0.6. At this point uninduced samples were taken, and also 15ml IPTG (1M) was added to the rest of the culture. After at least an hour of induction we took another, induced sample. We harvested both samples, and run them parallel on a 12% polyacrylamide gel in order to see if the checked constructs were inducible, and whether they resulted the expected size fusion proteins.

7.18 Purifying the fusion proteins

Bacterial cells can handle the foreign fusion proteins different ways according to the physical and chemical properties of the resulting molecule. They can be water-soluble and remain in the cytoplasm of the cells in soluble form. Insoluble or toxic proteins are sometimes packed into small membrane particles, and this way they cannot interfere with the cellular biochemical reactions. Naturally the methods one uses for the purification of these different samples are different, and there is no method that never fails. I describe a basic recipe that can be modified based on the individual properties of the different samples.

On the first day 10 ml bacterial cultures were started in antibiotic containing LB media. The next morning 100 ml bacterial cultures were started by diluting the overnight ones to a hundred fold, and they were grown at 37°C up to OD600=0.6. At this point we took 1 ml uninduced sample as a control. The rest of the culture was induced with 100 μl (1M) IPTG, and we let the culture to grow for at least another 2 hours at 37°C. After taking 1 ml induced sample the rest of the cultures were centrifuged at 6000 rpm for 10 minutes (4°C). The pellet was resuspended in 1 ml resuspension solution I (100 mM NaCl, 1mM EDTA, 50 mM Tris pH=8.0). We harvested the cells by adding some lysozyme, and keeping the samples at room temperature for 20 minutes. MgCl2 was added to a final concentration of 8 mM, and DNase1 to a final concentration of 10 μg/ml. The tubes were incubated then at 4C with occasional mixing until the viscosity disappears. The inclusion bodies were removed from the suspension by centrifugation, at 10.000 g for 10 minutes. If, however the fusion proteins seem to be soluble, the supernatant fraction can be used after the centrifugation to perform further purifications if necessary.

7.19 Sequence comparisons, database searches

Flybase: www.ebi.ac.uk:7081

Berkeley Drosophila Genome Project: www.fruitfly.org

GadFly: www.fruitfly.org/annot/index.html

PubMed, Blast, Entrez: www.ncbi.nlm.nih.gov

7.20 Solutions

Devitellinizing buffer:

1 volume Buffer B

3 volumes dH_2O

2 volumes 16% Formald. EM grade

Buffer B:

100 mM KH_2PO_4/K_2HPo_4 (pH=6.8)

450 mM KCl

150 mM NaCl

20 mM $MgCl_2$

Denaturation buffer for Southern(1l)

160 ml 5M NaCl

16g NaOH

1x EBR

1x PBS plus 0.1% Triton X-100

130 mM NaCl

5 mM KCl

2 mM $CaCl_2$

Fixation buffer for β-galactosidase staining

1 volume Buffer B

5 volume dH_2O

1/50 volume Glutaraldehyde (25% stock)

0.25N HCl

21 ml cc HCl/liter

PT

1x PBS plus 0.1% Triton X-100

PBT

PBS plus 0.2% BSA

0.1% Triton X-100

PBT+N

PBT plus 5% heat inactivated normal

goat serum

Store @ 4˚C

8% X-Gal in DMSO

40 mg/500 µl (stock diluted in staining solution)

Staining solution

10 mM NaH2PO4/Na2HPO4 (pH7.2)

150

mMNaCl

1 mM MgCl2

3.1 mM K4[FeII(CN)6]

3.1 mM K3[FeIII(CN)6]

0.3% TX-100

add Xgal to 0.2% before use

8. LIST OF REFERENCES

Abdelilah-Seyfried,S., Chan,Y.M., Zeng,C., Justice,N.J., Younger-Shepherd,S., Sharp,L.E., Barbel,S., Meadows,S.A., Jan,L.Y., and Jan,Y.N. (2001). A Gain-of-Function Screen for Genes That Affect the Development of the Drosophila Adult External Sensory Organ. Genetics *157*, 455-456.

Allende,M.L., Amsterdam,A., Becker,T., Kawakami,K., Gaiano,N., and Hopkins,N. (1996). Insertional mutagenesis in zebrafish identifies two novel genes, pescadillo and dead eye, essential for embryonic development. Genes Dev. *10*, 3141-3155.

Ashburner, M. Drosophila, A Laboratory Manual. 1989. Cold Spring Harbor, Cold Spring Harbor Laboratory Press.

B.Alberts, D.Bray, J.Lewis, M.Raff, K.Roberts, and J.D.Watson (2002). Molecular Biology of the Cell. Garland Publishing).

Bar-Sagi,D. and Feramisco,J.R. (1986). Induction of membrane ruffling and fluid-phase pinocytosis in quiescent fibroblasts by ras proteins. Science *233*, 1061-1068.

Barbieri,M.A., Kohn,A.D., Roth,R.A., and Stahl,P.D. (1998). Protein kinase B/akt and rab5 mediate Ras activation of endocytosis. J. Biol. Chem. *273*, 19367-19370.

Barbieri,M.A., Roberts,R.L., Gumusboga,A., Highfield,H., Alvarez-Dominguez,C., Wells,A., and Stahl,P.D. (2000). Epidermal growth factor and membrane trafficking. EGF receptor activation of endocytosis requires Rab5a. J. Cell Biol. *151*, 539-550.

Beall,E.L. and Rio,D.C. (1997). Drosophila P-element transposase is a novel site-specific endonuclease. Genes Dev. *11*, 2137-2151.

Bellen,H.J., O'Kane,C.J., Wilson,C., Grossniklaus,U., Pearson,R.K., and Gehring,W.J. (1989). P-element-mediated enhancer detection: a versatile method to study development in Drosophila. Genes Dev. *3*, 1288-1300.

Bender,A. and Pringle,J.R. (1989). Multicopy suppression of the cdc24 budding defect in yeast by CDC42 and three newly identified genes including the ras-related gene RSR1. Proc. Natl. Acad. Sci. U. S. A *86*, 9976-9980.

Benedict,M.Q., Salazar,C.E., and Collins,F.H. (1995). A new dominant selectable marker for genetic transformation; Hsp70-opd. Insect Biochem. Mol. Biol. *25*, 1061-1065.

Berg,C.A. and Spradling,A.C. (1991). Studies on the rate and site-specificity of P element transposition. Genetics *127*, 515-524.

Bhatt,A.M., Page,T., Lawson,E.J., Lister,C., and Dean,C. (1996). Use of Ac as an insertional mutagen in Arabidopsis. Plant J. *9*, 935-945.

Bier,E., Vaessin,H., Shepherd,S., Lee,K., McCall,K., Barbel,S., Ackerman,L., Carretto,R., Uemura,T., Grell,E., and . (1989). Searching for pattern and mutation in the Drosophila genome with a P- lacZ vector. Genes Dev. *3*, 1273-1287.

Brand,A.H. and Perrimon,N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development *118*, 401-415.

Brand,A.H. and Dormand,E.L. (1995). The GAL4 system as a tool for unravelling the mysteries of the Drosophila nervous system. Curr. Opin. Neurobiol. *5*, 572-578.

Brandon,E.P., Idzerda,R.L., and McKnight,G.S. (1995). Knockouts. Targeting the mouse genome: a compendium of knockouts (Part I). Curr. Biol. *5*, 625-634.

Bucci,C., Parton,R.G., Mather,I.H., Stunnenberg,H., Simons,K., Hoflack,B., and Zerial,M. (1992). The small GTPase rab5 functions as a regulatory factor in the early endocytic pathway. Cell *70*, 715-728.

Burd,C.G., Mustol,P.A., Schu,P.V., and Emr,S.D. (1996). A yeast protein related to a mammalian Ras-binding protein, Vps9p, is required for localization of vacuolar proteins. Mol. Cell Biol. *16*, 2369-2377.

Burns,N., Grimwade,B., Ross-Macdonald,P.B., Choi,E.Y., Finberg,K., Roeder,G.S., and Snyder,M. (1994). Large-scale analysis of gene expression, protein localization, and gene disruption in Saccharomyces cerevisiae. Genes Dev. *8*, 1087-1105.

Carpenter,G. and Cohen,S. (1990). Epidermal growth factor. J. Biol. Chem. *265*, 7709-7712.

Cooley,L., Kelley,R., and Spradling,A. (1988). Insertional mutagenesis of the Drosophila genome with single P elements. Science *239*, 1121-1128.

Cooley,L., Berg,C., and Spradling,A. (1988). Controlling P element insertional mutagenesis. Trends Genet. *4*, 254-258.

Cooley,L., Berg,C., Kelley,R., McKearin,D., and Spradling,A. (1989). Identifying and cloning Drosophila genes by single P element insertional mutagenesis. Prog. Nucleic Acid Res. Mol. Biol. *36*, 99-109.

D.B.Roberts (1986). Drosophila: A Practical Approach., O.IRL Press, ed., p. p.59.

Deak,P., Omar,M.M., Saunders,R.D., Pal,M., Komonyi,O., Szidonya,J., Maroy,P., Zhang,Y., Ashburner,M., Benos,P., Savakis,C., Siden-Kiamos,I., Louis,C., Bolshakov,V.N., Kafatos,F.C., Madueno,E., Modolell,J., and Glover,D.M. (1997). P-element insertion alleles of essential genes on the third chromosome of Drosophila melanogaster: correlation of physical and cytogenetic maps in chromosomal region 86E-87F. Genetics *147*, 1697-1722.

Duchek,P. and Rorth,P. (2001). Guidance of cell migration by EGF receptor signaling during Drosophila oogenesis. Science *291*, 131-133.

Edited by: Sullivan, W., Ashburner, M., and Hawley, R. S. Drosophila Protocols. 2000. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press.

Ellis,S. and Mellor,H. (2000). Regulation of endocytic traffic by rho family GTPases. Trends Cell Biol. *10*, 85-88.

Fortini,M.E. and Artavanis-Tsakonas,S. (1994). The suppressor of hairless protein participates in notch receptor signaling. Cell *79*, 273-282.

Gaiano,N., Amsterdam,A., Kawakami,K., Allende,M., Becker,T., and Hopkins,N. (1996). Insertional mutagenesis and rapid cloning of essential genes in zebrafish. Nature *383*, 829-832.

Gans,M., Audit,C., and Masson,M. (1975). Isolation and characterization of sex-linked female-sterile mutants in Drosophila melanogaster. Genetics *81*, 683-704.

Garraway,L.A., Tosi,L.R., Wang,Y., Moore,J.B., Dobson,D.E., and Beverley,S.M. (1997). Insertional mutagenesis by a modified in vitro Ty1 transposition system. Gene *198*, 27-35.

Garza,D., Medhora,M., Koga,A., and Hartl,D.L. (1991). Introduction of the transposable element mariner into the germline of Drosophila melanogaster. Genetics *128*, 303-310.

Gaul,U., Mardon,G., and Rubin,G.M. (1992). A putative Ras GTPase activating protein acts as a negative regulator of signaling by the Sevenless receptor tyrosine kinase. Cell *68*, 1007-1019.

Gertler,F.B., Bennett,R.L., Clark,M.J., and Hoffmann,F.M. Drosophila abl tyrosine kinase in embryonic CNS axons: a role in axonogenesis is revealed through dosage-sensitive interactions with disabled.

Gertler,F.B., Doctor,J.S., and Hoffmann,F.M. (1990). Genetic suppression of mutations in the Drosophila abl proto-oncogene homolog. Science *248*, 857-860.

Gertler,F.B., Comer,A.R., Juang,J.L., Ahern,S.M., Clark,M.J., Liebl,E.C., and Hoffmann,F.M. (1995). enabled, a dosage-sensitive suppressor of mutations in the Drosophila Abl tyrosine kinase, encodes an Abl substrate with SH3 domain-binding properties. Genes Dev. *9*, 521-533.

Gorvel,J.P., Chavrier,P., Zerial,M., and Gruenberg,J. (1991). rab5 controls early endosome fusion in vitro. Cell *64*, 915-925.

Gossler,A., Joyner,A.L., Rossant,J., and Skarnes,W.C. (1989). Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. Science *244*, 463-465.

Gruneberg, H. The Genetics of the Mouse. 1952. The Hague, Netherlands, Martinus Nijhoff.

Haigler,H.T., McKanna,J.A., and Cohen,S. (1979). Direct visualization of the binding and internalization of a ferritin conjugate of epidermal growth factor in human carcinoma cells A-431. J. Cell Biol. *81*, 382-395.

Hama,H., Tall,G.G., and Horazdovsky,B.F. (1999). Vps9p is a guanine nucleotide exchange factor involved in vesicle- mediated vacuolar protein transport. J. Biol. Chem. *274*, 15284-15291.

Han,L. and Colicelli,J. (1995). A human protein selected for interference with Ras function interacts directly with Ras and competes with Raf1. Mol. Cell Biol. *15*, 1318-1323.

Han,L., Wong,D., Dhaka,A., Afar,D., White,M., Xie,W., Herschman,H., Witte,O., and Colicelli,J. (1997). Protein binding and signaling properties of RIN1 suggest a unique effector function. Proc. Natl. Acad. Sci. U. S. A *94*, 4954-4959.

Hartwell,L.H. (1991). Twenty-five years of cell cycle genetics. Genetics *129*, 975-980.

Hirst,J. and Robinson,M.S. (1998). Clathrin and adaptors. Biochim. Biophys. Acta *1404*, 173-193.

Horiuchi,H., Lippe,R., McBride,H.M., Rubino,M., Woodman,P., Stenmark,H., Rybin,V., Wilm,M., Ashman,K., Mann,M., and Zerial,M. (1997). A novel Rab5 GDP/GTP exchange factor complexed to Rabaptin-5 links nucleotide exchange to effector recruitment and function. Cell *90*, 1149-1159.

Howell,B.W., Gertler,F.B., and Cooper,J.A. (1997). Mouse disabled (mDab1): a Src binding protein implicated in neuronal development. EMBO J. *16*, 121-132.

Huang,A.M. and Rubin,G.M. (2000). A misexpression screen identifies genes that can modulate RAS1 pathway signaling in Drosophila melanogaster. Genetics *156*, 1219-1230.

Jaenisch,R. (1988). Transgenic animals. Science *240*, 1468-1474.

Kidwell,M.G. and Lisch,D. (1997). Transposable elements as sources of variation in animals and plants. Proc. Natl. Acad. Sci. U. S. A *94*, 7704-7711.

King, R. C. Ovarian Development in Drosophila melanogaster. New York, London, Academic Press.

Kleckner,N., Roth,J., and Botstein,D. (1977). Genetic engineering in vivo using translocatable drug-resistance elements. New methods in bacterial genetics. J. Mol. Biol. *116*, 125-159.

Klemenz,R., Weber,U., and Gehring,W.J. (1987). The white gene as a marker in a new P-element vector for gene transfer in Drosophila. Nucleic Acids Res. *15*, 3947-3959.

Korswagen,H.C., Durbin,R.M., Smits,M.T., and Plasterk,R.H. (1996). Transposon Tc1-derived, sequence-tagged sites in Caenorhabditis elegans as markers for gene mapping. Proc. Natl. Acad. Sci. U. S. A *93*, 14680-14685.

Kraut,R., Menon,K., and Zinn,K. (2001). A gain-of-function screen for genes controlling motor axon guidance and synaptogenesis in Drosophila. Curr. Biol. *11*, 417-430.

Lamaze,C., Chuang,T.H., Terlecky,L.J., Bokoch,G.M., and Schmid,S.L. (1996). Regulation of receptor-mediated endocytosis by Rho and Rac. Nature *382*, 177-179.

Lauffenburger,D.A. and Horwitz,A.F. (1996). Cell migration: a physically integrated molecular process. Cell *84*, 359-369.

Laurent,V., Loisel,T.P., Harbeck,B., Wehman,A., Grobe,L., Jockusch,B.M., Wehland,J., Gertler,F.B., and Carlier,M.F. (1999). Role of proteins of the Ena/VASP family in actin-based motility of Listeria monocytogenes. J. Cell Biol. *144*, 1245-1258.

Lavoie,J.N., Hickey,E., Weber,L.A., and Landry,J. (1993). Modulation of actin microfilament dynamics and fluid phase pinocytosis by phosphorylation of heat shock protein 27. J. Biol. Chem. *268*, 24210-24214.

Lemmon,M.A. and Schlessinger,J. (1994). Regulation of signal transduction and signal diversity by receptor oligomerization. Trends Biochem. Sci. *19*, 459-463.

Li,G., D'Souza-Schorey,C., Barbieri,M.A., Cooper,J.A., and Stahl,P.D. (1997). Uncoupling of membrane ruffling and pinocytosis during Ras signal transduction. J. Biol. Chem. *272*, 10337-10340.

Liao,G.C., Rehm,E.J., and Rubin,G.M. (2000). Insertion site preferences of the P transposable element in Drosophila melanogaster. Proc. Natl. Acad. Sci. U. S. A *97*, 3347-3351.

Lindsley, D. L. and Zimm, E. H. The Genome of Drosophila melanogaster. 1992. San Diego, Academic Press, San Diego.

Lukacsovich,T. and Yamamoto,D. (2001). Trap a gene and find out its function: toward functional genomics in Drosophila. J. Neurogenet. *15*, 147-168.

Mata,J., Curado,S., Ephrussi,A., and Rorth,P. (2000). Tribbles coordinates mitosis and morphogenesis in Drosophila by regulating string/CDC25 proteolysis. Cell *101*, 511-522.

Merilainen,J., Lehto,V.P., and Wasenius,V.M. (1997). FAP52, a novel, SH3 domain-containing focal adhesion protein. J. Biol. Chem. *272*, 23278-23284.

Miki,H., Miura,K., and Takenawa,T. (1996). N-WASP, a novel actin-depolymerizing protein, regulates the cortical cytoskeletal rearrangement in a PIP2-dependent manner downstream of tyrosine kinases. EMBO J. *15*, 5326-5335.

Miki,H., Sasaki,T., Takai,Y., and Takenawa,T. (1998). Induction of filopodium formation by a WASP-related actin- depolymerizing protein N-WASP. Nature *391*, 93-96.

Miklos,G.L. and Rubin,G.M. (1996). The role of the genome project in determining gene function: insights from model organisms. Cell *86*, 521-529.

Mitchison,T.J. and Cramer,L.P. (1996). Actin-based cell motility and cell locomotion. Cell *84*, 371-379.

Moffat,A.S. (2000). Genetics. Transposons help sculpt a dynamic genome. Science *289*, 1455-1457.

Montell,D.J., Rorth,P., and Spradling,A.C. (1992). slow border cells, a locus required for a developmentally regulated cell migration during oogenesis, encodes Drosophila C/EBP. Cell *71*, 51-62.

Mullins,M.C., Hammerschmidt,M., Haffter,P., and Nusslein-Volhard,C. (1994). Large-scale mutagenesis in the zebrafish: in search of genes controlling development in a vertebrate. Curr. Biol. *4*, 189-202.

Nusslein-Volhard,C. and Wieschaus,E. (1980). Mutations affecting segment number and polarity in Drosophila. Nature *287*, 795-801.

Nusslein-Volhard,C. (1994). Of flies and fishes. Science *266*, 572-574.

O'Hare,K. and Rubin,G.M. (1983). Structures of P transposable elements and their sites of insertion and excision in the Drosophila melanogaster genome. Cell *34*, 25-35.

P.Winter, I.Hickey, and H.Fletcher (1998). Instant notes in Genetics. Springer Verlag.

Pazman,C., Mayes,C.A., Fanto,M., Haynes,S.R., and Mlodzik,M. (2000). Rasputin, the Drosophila homologue of the RasGAP SH3 binding protein, functions in ras- and Rho-mediated signaling. Development *127*, 1715-1725.

Pena-Rangel,M.T., Rodriguez,I., and Riesgo-Escovar,J.R. (2002). A misexpression study examining dorsal thorax formation in Drosophila melanogaster. Genetics *160*, 1035-1050.

Phelps,C.B. and Brand,A.H. (1998). Ectopic gene expression in Drosophila using GAL4 system. Methods *14*, 367-379.

Plasterk,R.H. (1992). Reverse genetics of Caenorhabditis elegans. Bioessays *14*, 629-633.

Qualmann,B., Roos,J., DiGregorio,P.J., and Kelly,R.B. (1999). Syndapin I, a synaptic dynamin-binding protein that associates with the neural Wiskott-Aldrich syndrome protein. Mol. Biol. Cell *10*, 501-513.

Qualmann,B. and Kelly,R.B. (2000). Syndapin isoforms participate in receptor-mediated endocytosis and actin organization. J. Cell Biol. *148*, 1047-1062.

Ramer,S.W., Elledge,S.J., and Davis,R.W. (1992). Dominant genetics using a yeast genomic library under the control of a strong inducible promoter. Proc. Natl. Acad. Sci. U. S. A *89*, 11589-11593.

Riese,D.J. and Stern,D.F. (1998). Specificity within the EGF family/ErbB receptor family signaling network. Bioessays *20*, 41-48.

Rogge,R.D., Karlovich,C.A., and Banerjee,U. (1991). Genetic dissection of a neurodevelopmental pathway: Son of sevenless functions downstream of the sevenless and EGF receptor tyrosine kinases. Cell *64*, 39-48.

Rorth,P. (1996). A modular misexpression screen in Drosophila detecting tissue-specific phenotypes. Proc. Natl. Acad. Sci. U. S. A *93*, 12418-12422.

Rorth,P., Szabo,K., Bailey,A., Laverty,T., Rehm,J., Rubin,G.M., Weigmann,K., Milan,M., Benes,V., Ansorge,W., and Cohen,S.M. (1998). Systematic gain-of-function genetics in Drosophila. Development *125*, 1049-1057.

Rorth,P., Szabo,K., and Texido,G. (2000). The level of C/EBP protein is critical for cell migration during Drosophila oogenesis and is tightly controlled by regulated degradation. Mol. Cell *6*, 23-30.

Rubin,G.M. and Spradling,A.C. (1982). Genetic transformation of Drosophila with transposable element vectors. Science *218*, 348-353.

Sambrook, Fritsch, and Maniatis. Molecular Cloning, A Laboratry Manual. 1989. Cold Spring Harbor, Cold Spring Harbor Laboratory Press.

Simon,M.A., Bowtell,D.D., Dodson,G.S., Laverty,T.R., and Rubin,G.M. (1991). Ras1 and a putative guanine nucleotide exchange factor perform crucial steps in signaling by the sevenless protein tyrosine kinase. Cell *67*, 701-716.

Simon,M.A., Carthew,R.W., Fortini,M.E., Gaul,U., Mardon,G., and Rubin,G.M. (1992). Signal transduction pathway initiated by activation of the sevenless tyrosine kinase receptor. Cold Spring Harb. Symp. Quant. Biol. *57*, 375-380.

Smith,D., Yanai,Y., Liu,Y.G., Ishiguro,S., Okada,K., Shibata,D., Whittier,R.F., and Fedoroff,N.V. (1996). Characterization and mapping of Ds-GUS-T-DNA lines for targeted insertional mutagenesis. Plant J. *10*, 721-732.

Spradling, A. C. Developmental genetics of oogenesis. Bate, M. and Martinez, A. The Development of Drosophila melanogaster. 1-70. 1993. Cold Spring Harbour, Cold Spring Harbour Laboratory Press.

Spradling,A.C., Stern,D.M., Kiss,I., Roote,J., Laverty,T., and Rubin,G.M. (1995). Gene disruptions using P transposable elements: an integral component of the Drosophila genome project. Proc. Natl. Acad. Sci. U. S. A *92*, 10824-10830.

Spradling,A.C., Stern,D., Beaton,A., Rhem,E.J., Laverty,T., Mozden,N., Misra,S., and Rubin,G.M. (1999). The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes. Genetics *153*, 135-177.

Sundaresan,V., Springer,P., Volpe,T., Haward,S., Jones,J.D., Dean,C., Ma,H., and Martienssen,R. (1995). Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. Genes Dev. *9*, 1797-1810.

Suter,D.M. and Forscher,P. (1998). An emerging link between cytoskeletal dynamics and cell adhesion molecules in growth cone guidance. Curr. Opin. Neurobiol. *8*, 106-116.

Szabo,K., Jekely,G., and Rorth,P. (2001). Cloning and expression of sprint, a Drosophila homologue of RIN1. Mech. Dev. *101*, 259-262.

T.A.Brown (2002). Genomes 2. John Wiley, Sons.

Tall,G.G., Barbieri,M.A., Stahl,P.D., and Horazdovsky,B.F. (2001). Ras-activated endocytosis is mediated by the Rab5 guanine nucleotide exchange activity of RIN1. Dev. Cell *1*, 73-82.

Tessier-Lavigne,M. and Goodman,C.S. (1996). The molecular biology of axon guidance. Science *274*, 1123-1133.

Toba,G., Ohsako,T., Miyata,N., Ohtsuka,T., Seong,K.H., and Aigaki,T. (1999). The gene search system. A method for efficient detection and rapid molecular identification of genes in Drosophila melanogaster. Genetics *151*, 725-737.

Torok,T., Tick,G., Alvarado,M., and Kiss,I. (1993). P-lacW insertional mutagenesis on the second chromosome of Drosophila melanogaster: isolation of lethals with different overgrowth phenotypes. Genetics *135*, 71-80.

Tseng,A.S. and Hariharan,I.K. (2002). An overexpression screen in Drosophila for genes that restrict growth or cell-cycle progression in the developing eye. Genetics *162*, 229-243.

Voie, AM and Cohen, S. Cell Biology: A Laboratory Handbook. Germ-Line Transformation of Drosophila melanogaster. 2003. Academic Press.

Wells,A. (1999). EGF receptor. Int. J. Biochem. Cell Biol. *31*, 637-643.

Wilson,C., Pearson,R.K., Bellen,H.J., O'Kane,C.J., Grossniklaus,U., and Gehring,W.J. (1989). P-element-mediated enhancer detection: an efficient method for isolating and characterizing developmentally regulated genes in Drosophila. Genes Dev. *3*, 1301-1313.

Wurst,W., Rossant,J., Prideaux,V., Kownacka,M., Joyner,A., Hill,D.P., Guillemot,F., Gasca,S., Cado,D., Auerbach,A., and . (1995). A large-scale gene-trap screen for insertional mutations in developmentally regulated genes in mice. Genetics *139*, 889-899.

Xu,T. and Rubin,G.M. (1993). Analysis of genetic mosaics in developing and adult Drosophila tissues. Development *117*, 1223-1237.

Xuan, Z., Wang, J., and Zhang M. Q. (2003). Computational comparison of two mouse draft genomes and the human golden path. Genome Biology 4(1)

Zambrowicz,B.P., Friedrich,G.A., Buxton,E.C., Lilleberg,S.L., Person,C., and Sands,A.T. (1998). Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. Nature *392*, 608-611.

Zhang,P. and Spradling,A.C. (1993). Efficient and dispersed local P element transposition from Drosophila females. Genetics *133*, 361-373.

ÖSSZEFOGLALÁS

Több mint 50 évvel ezelőtt Barbara McClintock teljesen felforgatta a tudományos közéletet azzal a kijelentésével, hogy a kukorica úgynevezett mozgékony genetikai elemeket tartalmaz. Ez a megfigyelés szöges ellentétben állt a korábbi dogmával, mely szerint a gének a kromoszómák jól meghatározható helyein foglalnak helyet. McClintock felfedezéseit sokáig a kukoricára kizárólag jelemző egyedi sajátosságnak tekintették. Az elkövetkező években azonban a mozgékony elemek létezését számos más fajban is bebizonyították, és összefüggésbe hozták őket mutációkkal, melyek egy része betegségek kialkulásához vezet, mások hozzájárulnak a fajon belüli evolúcióhoz.

Ma a mozgékony genetikai elemeket széles körben használjuk a tudományos kutatásban. A transzpozonoknak az a sajátossága hogy mutagénként viselkedhetnek lehetővé teszi új gének azonosítását az un. inszerciós mutagenesis módszerrel, mellyel mesterségesen módosíthatjuk egyes kiválasztott fajok genomját. Ilyen kisérleteket sikerrel hajtottak már végre különböző model organizmusokban mint például baktériumokban (Kleckner, Roth, et al. 1977), élesztőben (Burns, Grimwade, et al. 1994; Garraway, Tosi, et al. 1997), *Arabidopsisban* (Sundaresan, Springer, et al. 1995; Bhatt, Page, et al. 1996; Smith, Yanai, et al. 1996), *Drosophilában* (Cooley, Kelley, et al. 1988; Torok, Tick, et al. 1993; Rorth 1996), zebrahalban (Allende, Amsterdam, et al. 1996; Gaiano, Amsterdam, et al. 1996), vagy egérben mice (Jaenisch 1988; Gossler, Joyner, et al. 1989; Wurst, Rossant, et al. 1995; {Zambrowicz, Friedrich, et al. 1998). A transzpozonok egyik leglényegesebb tulajdonsága hogy olyan változásokat hoznak létre a genomban amleyek sokkal egyszerűbben detektálhatóak mint a más mutagének (pl. ionizáló sugárzás, kémiai mutagének) használata esetén létrejövőek.

A Drosphilában előforduló P-elemek a legelterjedtebben felhasznált eukarióta transzpozonok közé tartoznak, és gyakran alkalmazzuk őket mutagenezis kisérletekben. Aktivitása jól kontrollálható az elem mozgásához szükséges transzpozáz enzim szintjének szabályozásával, és az ugrási gyakorisága is magas.

A kutatók az 1980-as években kezdték a genetkailag módosított P-elemeket nagyméretű genetikai kisérletekben használni, melyek során nagyszámú mutáns Drosophila törzset hoztak létre. Ezzel a módszerrel un. funkcióvesztéses mutációk kialakulását okozzuk, hiszen a transzpozon fizikailag elrontja a géneket amelyekbe beépül, miáltal azok aktivitása csökken, vagy teljesen meg is szünik. Az egyetlen P-elemet tartalmazó vonalak jól használhatóak a légygenom, és annak egyes génjeinek megismeréséhez. Mára a fejlődésbiológiai ismereteink zömét ilyen P-elem által létrehozott mutáns ecetmuslica vonalak tanulmányozásával szereztük.

A P-elemes törzsek számának, és az iránt mutatott kereslet folyamatos növekedésével szükségessé vált ezek összegyűjtése, és mindenki számára hozzáférhetővé tétele. A Berkeley Drosophila Genome Project vállalta ezt a feladatot, akik egy olyan gyűjtemény létrehozásán dolgoznak, ami minden Drosophila génre tartalmaz transzpozon beépülés okozta mutáns vonalat. Ez azonban nem egyszerű feladat. Adataink alapján az ecetmulsica génjeinek csak mintegy 30%-a mutáltatható P-elemekkel, ami azt jelenti, hogy alternatív módszerek szükségesek a fennmaradó mintegy 70% tanulmányozásához.

Az 1990-es években egy új módszert fejlesztettek ki ektopikus gén expresszáltatáshoz Drosophilában, amit UAS-Gal4 rendszernek neveztek el. A módszer lehetővé teszi bizonyos gének vagy mesterséges konstrukciók kifejezését időben és térben szabályozott módon, ami által funkciónyeréses mutációkat kapunk.

A Gal4 egy szekvencia specifikus transzkripcionális transzaktivátor az élesztőben, és a Drosophila rendszerben lehetővé teszi target gének szövetspecifikus aktivációját (Brand, Perrimon 1993).

A P. Rorth által létrehozott funkciónyeréses mutagenezis módszer lehetővé teszi a klasszikus P-elem alapú mutagenezis kisérletek, valamint a Gal4 rendszer előnyeinek kombinálását. Alkalmazásával nagyszámú P-elemet tartalmazó törzset hozhatunk létre, de a mutáns fenotipust a különböző gének szövetspecifikus expresszáltatása okozza a UAS-Gal4 rendszer felhasználásával (Rorth 1996). A kisérlet során kidolgozott módosított transzpozon speciális enhanszer és promóter elemket tartalmaz, és emiatt EP elemnek nevezték el. Alkalmazása lehetővé teszi a mozgékony elem közvetlen környezetében levő szekvenciák ektopikus kifejeztetését.

Az EP vektor segítségével transzpozon inszerciókat hoztunk létre az ecetmuslica X, a második, és a harmadik kromoszómáján. Munkánk során jónéhány technikai módosítást vezettünk be korábban elvégzett hasonló kisérletekkel összehasonlítva (Rorth, Szabo, et al. 1998). Dolgozatomban részletesen bemutatom ezen kisérletek jellemzőit, és eredményeit.

A mutáns kollekciót laboratóriumunkban Drosophila oogenezise során végbemenő sejtvándorlási folyamat tanulmányozására használtuk fel.

Az újonnan formálódó petekamra kialakulása után kétféle eredetű sejttípust tartalmaz. A 15 dajkasejt és a petesejt csíravonal eredetű, ezeket a szomatikus eredetű follikulus sejtek egysejtű rétege veszi körül. Az oogenezis során a follikulus sejtek poszteror irányba kezdenek elmozdulni. Ezzel egyidőben az anterior póluson 6-10 sejt elválik a többitől, és a dajkasejtek között a petesejt irányába indul. Egészen addig vándoronak amíg el nem érik a dajkasejtek és a petesejt közötti határt, ezért határ sejteknek is nevezik őket (King 1969; Spradling 1993). A vándorlási folyamat térben és

időben szigorúan szabályozott, ellenkező esetben a kialakuló peték nem termékenyíthetők meg. A nőstények amelyek ilyen hibás fejlődésű petéket raknak sterilek.

Ismertek olyan mutációk melyekben a határsejtek vándorlása zavart szenved. Az un. slbo mutáns ováriumokban például a határ sejtek egyáltalán nem mozognak, vagy legalábbis késnek, a slbo allél erősségétől függően. A slbo (slow border cells) lókusz a C/EBP nevű transzkripciós faktort kódolja. Korábban már igazolták, hogy a C/EBP gén kifejeződik a határ sejtekben, és a géntermék nélkülözhetetlen a sejtel vándorlásának ideje alatt. A slbo fenotipus, és az a tény, hogy C/EBP egy transzkripciós faktor azt sugallja, hogy valószínűleg olyan géneket szabályoz, melyeknek szerepe van a határsejtek vándorlásában. Ennek a lehetőségnek a vizsgálatára az EP gyűjtemény segítségével az anterior follikulus sejtek vándorlását vizsgáltuk.

Ha egy gént túlexpresszáltatva erősíthetjük vagy szuppresszálhatjuk egy másik gén meglevő fenotipusát, akkor a két génről elmondhatjuk, hogy valószínűleg ugyanabban a folyamatban játszanak szerepet, vagyis interakció van közöttük (Fortini, Artavanis-Tsakonas 1994). Ez az alapja az általunk elvégzett interakciós vizsgálatoknak, melynek során különöző géneket túlhajtottunk a vándorló határsejtekben egy különleges, érzékenyített háttéren.

Az EP vonalakat az anterior follikulus sejtekben expresszáltattunk túl slbo[1] mutáns allél jelenlétében. Az olyan nőstények amelyek hordozzák ezt az allélt sterilek, a határsejtek vándorlásának súlyos zavara miatt. Mintegy 60 olyan törzset találtunk, melyek slbo szupresszorként visekedtek a tesztjeink során. Bizonytottuk, hogy ezekben a vonalakban a mutációt szenvedett gének Gal4 függő kifejeztetése okozta a nőstény steril fenotípus menekítését. A mutáns géneket különböző molekuláris módszerekkel azonosítottunk.

Az EP(X)1487 jelű törzsben a transzpozon beépülése a humán RIN1 nevű gén Drosophila homológjában történt, melyet sprintnek neveztük el. Dolgozatomban bemutom hogyan, és milyen módszerekkel történt a gén klónozása, illetve milyen egyéb adatokat sikerült megtudni róla a fenotipusos és molekuláris vizsgálatok során. Saját, és a rokon RIN1 génen végzett biokémiai kisérletek eredményei arra utalnak, hogy sprint feltehetően egy olyan gén, amely szerepet játszik a Ras génen keresztül haladó jelátviteli útban, és emellett feladata még az aktin citoszkeleton felépítésének és szerveződésének szabályozása.