DRESDEN UNIVERSITY OF TECHNOLOGY
FACULTY OF MATHEMATICS AND NATURAL SCIENCES

TECHNISCHE UNIVERSITÄT DRESDEN

A recombineering pipeline for functional genomics applied to *Caenorhabditis elegans*

PhD thesis of Mihail Sarov

Biotec Centre
TU Dresden

INTERNATIONAL MAX PLANCK RESEARCH SCHOOL
MAX PLANCK INSTITUTE OF MOLECULAR CELL BIOLOGY AND GENETICS

CBG
Max Planck Institute of Molecular Cell Biology and Genetics

Acknowledgements

I would like to start by thanking my thesis supervisor Francis Stewart. I remember him once saying "You work with me, not for me". True to this principle, he always treated me as equal and gave me complete freedom to plan and perform my work independently.

I am grateful to my TAC members Tony Hyman and David Drechsel, who shared my enthusiasm for methods development and always had some useful advice when I needed it.

I am much indebted to Youming Zhang, who made even the most impossible recombineering task seem as the most simple, easy and natural thing to do.

I thank Susan Schneider and Andrei Pozniakovski for their help and advice with cloning, Susanne Ernst for bombarding so many worms, Ina Poser for helping me with the BAC transgenesis in mammalian cells, Assen Roguev and Jean Karim Heiche for the bioinformatics support. Thanks to Senming Zhao and the MPI-CBG Protein Expression Facility for their help in purifying and testing the new TAP cassettes and to the MPI-CBG MS facility for the mass spectrometry analyses.

I thank Michelle Meredith, Konstantinos Anastassiadis and Eugeni Entchev for the critical reading of the manuscript.

I thank all past and present members of the Stewart group, for all the fun in the lab. Thanks to all my friends for reminding me that life outside the lab is also fun and full of excitement. Thanks to my family who always trusted and supported me in everything for as long as I can remember.

And above all, I thank Marta for all that she is.

Publications

Parts of this work have been published:

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

Wang, J., **Sarov, M.**, Rientjes, J., Fu, J., Hollak, H., Kranz, H., Xie, W., Stewart, A.F. & Zhang, Y. An Improved Recombineering Approach by Adding RecA to lambda Red Recombination. *Molecular Biotechnology* 32, 43-54 (2006).

Sarov, M. & Stewart, A.F. The best control for the specificity of RNAi. *Trends in Biotechnology* 23, 446-448 (2005).

I have also contributed to the following papers:

Pasheva, E., **Sarov, M.**, Bidjekov, K., Ugrinova, I., Sarg, B., Lindner, H. & Pashev, I.G. *In vitro* acetylation of HMGB-1 and -2 proteins by CBP: the role of the acidic tail. *Biochemistry* 43, 2935-2940 (2004).

Schaft, D., Roguev, A., Kotovic, K.M., Shevchenko, A., **Sarov, M.**, Shevchenko, A., Neugebauer, K.M. & Stewart, A.F. The histone 3 lysine 36 methyltransferase, SET2, is involved *in trans*criptional elongation. *Nucleic Acids Research* 31, 2475-2482 (2003).

Summary

Protein tagging with fluorescent or affinity tags provides a generic way to describe protein expression and localization patterns and protein-protein interactions. The genome wide application of this approach in *Saccharomyces cerevisiae* has resulted in a comprehensive picture of the core proteome of a simple, well-studied model system. Extending this approach to more complex multicellular model organisms would require efficient protein tagging methods and new high performance tags.

Here we present a generic protein tagging approach for the model nematode *Caenorhabditis elegans*. The method is based on recombination mediated DNA engineering of genomic BAC clones into tagged transgenes for integrative transformation. We apply a novel DNA engineering strategy, based on a pipeline of serial recombineering steps in liquid culture. This approach is fast, straightforward and facilitates simultaneous processing of multiple samples in parallel.

We show that the BAC derived transgenes can be used for stable, integrative transformation in C. *elegans*. We show that the tagged transgene can take over the function of its endogenous counterpart. Using florescent reporter, we reproduce known and document new expression patterns.

The second part of the thesis describes a project that we undertook to develop improved double affinity cassettes for protein purification. We evaluated the performance of 5 new double tag combinations *in vitro* and in mammalian culture cells. All of the new cassettes performed well and present a valuable tool for protein interaction studies in higher model systems.

1 General introduction

1.1 Protein tagging as a function discovery tool

Genome sequencing and annotation projects have already provided the complete set of genes for nearly all model organisms. The advance of the functional genomics methods has given us a way to assign biological function to thousands of previously uncharacterized genes. Among these methods, protein tagging stands out as very versatile tool. Tagging with fluorescent proteins is a generic strategy for description of protein expression and localization patterns. Tag based affinity purification under native conditions has been established as arguably the most powerful technique for studding physical protein interactions. In this approach, a sequence coding for an affinity tag is fused in frame to the gene of interest, resulting in the production of a tagged protein. The protein is purified under native conditions and the copurified proteins are identified by mass spectrometry. The development of cassettes based on multiple affinity tags unleashed the full power of this approach. The tandem affinity purification (TAP)[1] results in much cleaner preparations and facilitates the direct analysis of the retrieved protein entities. The TAP approach has been particularly efficient in yeast, whereas in higher model systems with more complex proteomes, it is more challenging and its application has been limited. In an attempt to improve the TAP performance, we tested a set of new double affinity cassettes (Chapter 3).

The high throughput application of GFP localization[2] and TAP purification[3,4] in *Saccharomyces cerevisiae* has produced a comprehensive picture of the core proteome of a simple, well studied model system[5,6]. In combination with the wealth of genetic information already available, these data sets provide plausible predictions of the functional properties of thousands of previously unstudied proteins and reveal new functional pathways[7,8]. Extending this approach to more complex, multicellular model organisms such as *Caenorhabditis elegans* or *Drosophila* would allow us to place protein function onto a 4 dimensional space-

time map, and will drastically increase our understanding of the complex processes of development and differentiation[9].

Generation of tagged proteins on a large scale requires a high throughput approach. Unlike yeast in most higher model systems homologous gene targeting is too inefficient for large scale application. Transgene based approaches have been used instead. The best transgenic constructs to date are based on large genomic DNA clones such as Bacterial Artificial Chromosomes (BACs), phage P-1 derived Artificial Chromosomes (PACs) or fosmids[10]. They can be big enough to contain all the endogenous regulatory sequences and are likely to result in highly accurate levels and patterns of expression. Manipulation of such large constructs is only possible through *in vivo* homologous recombination mediated cloning (recombineering). In the past recombineering has been restricted to exploiting the natural homologous recombination potential of *Saccharomyces cerevisiae*. More recently, the development of recombineering in *E. coli*[11-17] has made it possible to tap in on the huge resources of mapped genomic clones that are currently available.

1.2 Red/ET recombination

In *E. coli* the classical homologous recombination mediated double strand break repair pathway involves the strand annealing protein RecA and the exonuclease RecBCD[18]. To be efficient, RecA mediated homologous recombination requires long regions of homology and that has limited the practical application of this pathway for DNA engineering in *E. coli*[19,20]. In contrast the homologous recombination pathway mediated by the phage lambda Red proteins is efficient with regions of homology of just 30-50bp[16,21-24]. In this pathway, the exonuclease Redα and the strand annealing protein Redβ drive the homologous recombination reaction, while Redγ specifically inhibits the action of RecBCD to prevent degradation of the free DNA ends[18,24].

The potential of this pathway for recombineering in *E. coli* was first utilized[11] with the Rac prophage homologs of Redα and Redβ - RecE and RecT, hence the name Red/ET recombination. Expression *in trans* of the red operon from a plasmid can

transfer the recombination potential to any *E. coli* strain, which is particularly useful for modification of large constructs such as BACs[11,12,25]. Using this approach virtually any modification can be precisely introduced into the BAC in its original host strain (for review see references [13,16,21-23]). More complex modifications, such as generation of tagged transgenes, usually involve multiple recombination steps. Combined with the need to verify the correct recombination at each step, generation of a single construct may require several weeks. In Chapter 2, we show that the high fidelity of Red/ET recombination permits a new way to engineer DNA using sequential steps in liquid culture without cloning or checking until the final product. Based on that, we have established a pipeline for generation of transgenic constructs for protein tagging in C. *elegans*.

1.3 C. *elegans* as a model organism

C. *elegans* has powerful tools for mapping gene expression, a very well annotated and compact genome, good comparative genomics resources, a simple and well-understood anatomy and pattern of development, and it is easy to maintain in the laboratory.

Due to its small size and well studied, almost invariant cell lineage, C. *elegans* is an excellent model for documenting gene expression at the single cell level throughout development. Algorithms have now been developed that permit automatic 4D protein localization studies using fluorescent reporters[26,27].

C. *elegans* has the most thoroughly annotated metazoan genome. A large amounts of functional data have been accumulated trough scores of genome wide studies: transcriptome analysis by Serial Analysis of Gene Expression (SAGE)[28] and DNA microarrays[29]; phenotypic screens by RNAi[30-35], chemical[36] or transposon mutagenesis[37-39]; genetic interaction screens by combinatorial RNAi[40]; protein interaction analysis by yeast two hybrid screening[41]. All this information is well organized and readily accessible through the Wormbase database[42].

The C. *elegans* genome is very compact, and almost a fifth of the sequence codes for protein. Protein-coding genes usually have relatively short introns compared to higher species. In most cases the neighboring genes are within 5 kbp apart[43].

Therefore, relatively small regions of genomic DNA (10-15 kbp) can contain all the regulatory sequence of an average C. *elegans* gene and can complement in many cases loss of function mutations.

Finally, C. *elegans* is easy to culture and to grow for biochemical experiments such as protein complex purification and chromatin immunopurification (ChIP). Protein purification methods based on generic tandem affinity tags have been successfully adapted to C. *elegans*. A project aimed at obtaining the crystal structures of all C. *elegans* proteins is now underway[44,45]. Proteomic approaches based on high-resolution two-dimensional electrophoresis or multi-dimensional liquid chromatography coupled with mass spectrometry are in development[46].

In summary, C. *elegans* offers unique advantages for function discovery through protein tagging. However, the protein tagging technologies for C. *elegans* are much less developed than those available in other model systems.

Here we describe an efficient recombineering pipeline for the generation of tagged transgenes from genomic DNA clones and show that such transgenes can be used for integrative transformation in C. *elegans*.

2 Protein tagging in C. *elegans*

2.1 Introduction

Genome scale application of protein tagging in C. *elegans* requires a method that can be applied to any gene and is easy to scale up for high throughput tagging. Importantly, the expression level of the tagged gene should correctly represent that of the endogenous counterpart. This is best achieved by homologous targeting of the gene of interest, but in C. *elegans* homologous gene targeting although possible, works only at a very low frequency, which is impractical for routine application[47,48]. Transgene based approaches are typically used instead. Securing a correct gene expression is dependent both on the quality of the transgenic construct and on the transformation method. There are two commonly used methods for DNA transformation in C. *elegans:* microinjection and microparticle bombardment. In the first method the transgene is coinjected together with a selectable phenotypic marker into the gonad of an adult hermaphrodite[49]. The two molecules recombine together to form a large extrachromosomal array containing thousands of copies of the transgene. These arrays resemble free chromosomal duplications lacking a centromere. They are replicated and can be maintained for many generations if selected for. However, they segregate randomly during mitosis, leading to mosaic expression. Expression levels vary depending on the transgene copy number, but are typically higher than endogenous. Due to the repetitive nature of these arrays they are subject to transcriptional silencing, especially in the germline[50] but also in the somatic cells[51]. This process is mediated by RNAi and chromatin modification and serves to protect the genome from expression of repetitive exogenous sequences such as transposable elements[52-54].

An alternative ballistic transgene delivery method has recently been adapted to C. *elegans*[55-57]. In this technique the transgene is bound to inert micron sized particles, which are shot at high velocity onto a layer of worms. This method is less technically demanding and is easier to scale up. The most important

advantage however is the fact that bombardment transformation often integrated into the genome. Typically, only a few copies of the transgene get integrated, resulting in reliable, near endogenous expression levels. The low copy transgene integration requires that a selectable marker is present on the same DNA molecule.

Ideally, transgene expression should be driven and regulated by the endogenous control elements. However, the practical difficulties in generation of such transgenes by traditional restriction-ligation cloning have led to the widespread use of simpler transgenes based on cDNAs or ORFs. In the most common approach the coding sequence is cloned in a vector carrying the tag. The endogenous 5′ and 3′ noncoding sequences can either be cloned alongside the cDNA or can be substituted by generic promoters and 3′UTR, directly included in the vector. Transgene generation was greatly simplified by the creation of the C. *elegans* ORF-eome[58-61] and Promoterome libraries[62]. Using MultiSite Gateway Recombination the promoter and the ORF can be moved from these libraries into a tag containing vector[62,63]. This method can be automated for high throughput tagging. However the Gateway generated transgenic constructs are inherently artificial. The cDNA derived ORF clones do not recapitulate the endogenous regulation through alternative splicing. The promoter fragments in the Promoterome library are limited to 2kbp and can lack important *cis* regulatory elements. Finally, this method does not allow cloning of the endogenous 3′ downstream sequences. Hence, these constructs can result in incorrect level and pattern of gene expression.

In contrast, protein tagging by recombineering of genomic DNA clones allows for seamless tag insertion in the natural genomic context. In C. *elegans* this approach was first applied for GFP tagging of the *dsh*-2 gene using homologous recombination in *Saccharomyces cerevisiae*[64]. In a further development of this approach, yeast recombineering was used to subclone and tag genes from C. *elegans* yeast artificial chromosome (YAC) clones[65]. This approach however is slow (up to 6 weeks) and the reported recombineering efficiency is very low. Furthermore, yeasts are constitutively recombinogenic. Hence, cloned DNA

sequences, especially large ones, are continuously prone to rearrange and need to be carefully monitored.

The more efficient *E. coli* based recombineering approaches were until recently hampered by the lack of a suitable genomic clone resource. With the generation of BAC and fosmid clone libraries, recombineering of large genomic clones became possible. Here, we describe an application of recombineering in *E. coli* that allows generation of a tagged transgene from a genomic DNA clone for an arbitrary worm gene. The strategy that we applied has two steps: tagging of the gene in its natural genomic context and subcloning in a vector carrying a selection marker for integrative ballistic transformation. Both tagging and subcloning are done by recombination and are independent of the presence of restriction sites or the size of the gene. We show that such transgenes can be used for ballistic transformation in C. *elegans* and that the tagged transgene can take over the function of the endogenous gene. This approach combines the advantages of authentic regulation with a new application of recombineering, which is simple, fast and can be easily scaled up for automated large scale tagging.

2.2 Results

2.2.1 Evaluation of the available genomic libraries

2.2.1.1 YACs and cosmids

Sequencing of the C. *elegans* genome was initiated in the early 80s based on cosmid libraries and was finished using a yeast artificial chromosome (YAC) library. When we started this project, these were the only available mapped genomic clone resources for C. *elegans*.

Recombineering of YAC clones in yeast is possible and such approach has recently been described[65]. However cloning in yeast is slow and inefficient compared to recombineering in bacteria. For these reasons we decided to use Red/ET recombination in *E. coli* to engineer cosmid clones into tagged transgenes.

Our preliminary test with C. *elegans* cosmids showed that they are not suitable for routine gene tagging due to the high level of spontaneous rearrangements that occur even in the absence of Red activity. Furthermore growing the cosmids in the presence of the Red/ET expression plasmid was difficult and significantly reduced the efficiency of Red/ET recombination. The problems were not related to the host strain as they persisted, even after we moved the cosmids to the common cloning strains DH10B or HS996. Both the pJB8 and Lorist vectors used for the generation of the cosmid libraries are based on the pBR322 origin of replication. It is known that in the absence of RecBCD activity plasmids with the pBR322 origin can switch to rolling circle mode of replication[66], leading to gradual loss of the plasmid. RecBCD activity is inhibited by red gamma in order to prevent end degradation of the transformed linear DNA in a typical Red/ET reaction. Even without red operon induction, the basal level of red gamma expression can result in reduced RecBCD activity.

These results led us to search for an alternative clone resource based on modern single copy vectors such as BACs and fosmids.

2.2.1.2 Caenorhabditis briggsae BAC library

Instead of pursuing the issues with cosmid recombineering further we decided to use the available *Caenorhabditis briggsae* BAC library[67] to establish the conditions for recombination based transgenics in C. *elegans*. For most C. *elegans* genes an obvious, well-conserved ortholog can be found in the C. *briggsae* genome[67] and often in these cases, the C. *briggsae* genes can rescue C. *elegans* mutants.

The C. *briggsae* BAC library was generated to facilitate sequencing of the C. *briggsae* genome, and was partially end sequenced. At the time we initiated this work, there were no publicly available data mapping the clones to the genome. We generated an interactive clone map by aligning the end sequence pairs to the CB25 C. *briggsae* genome assembly[67]. Several quality criteria were built into the mapping algorithm to reduce the risk of incorrect mapping due to repetitive sequences or low quality end sequence reads (See materials and methods).

A total of 225 Mbp were mapped in 2769 clones, which is approximately two-fold coverage of the predicted genome size[67]. Almost 90% of the clones have insert sizes between 50 and 150 Kbp with a mean of 81 kbp (Figure 1).

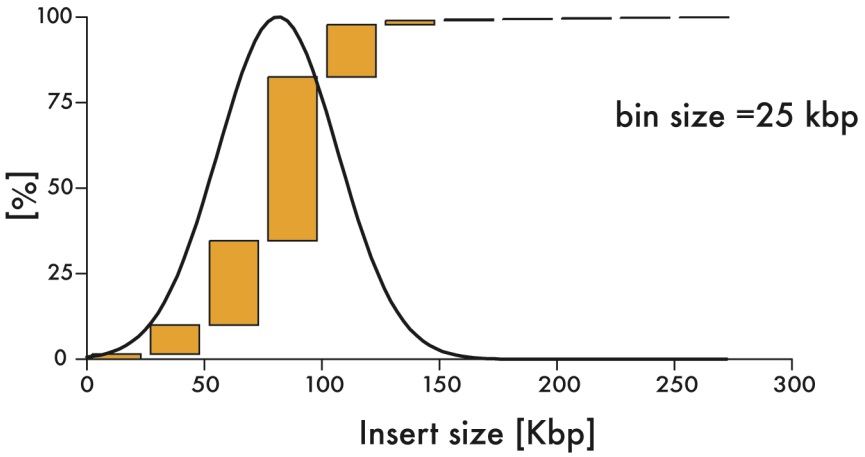


Figure 1. BAC clone size distribution (uninterrupted line) and cumulative histogram at 25 Kbp bins.

The BAC clone map was generated in the Wormbase external annotation format[42] to allow graphic representation within the genome browser (Figure 2). The map is fully interactive: clicking on the bar representing a BAC clone refocuses the browser to the region spanned by the clone insert. The relative orientation to the BAC vector is also indicated. The correct mapping was confirmed by restriction digest analysis of more than 30 clones (data not shown).

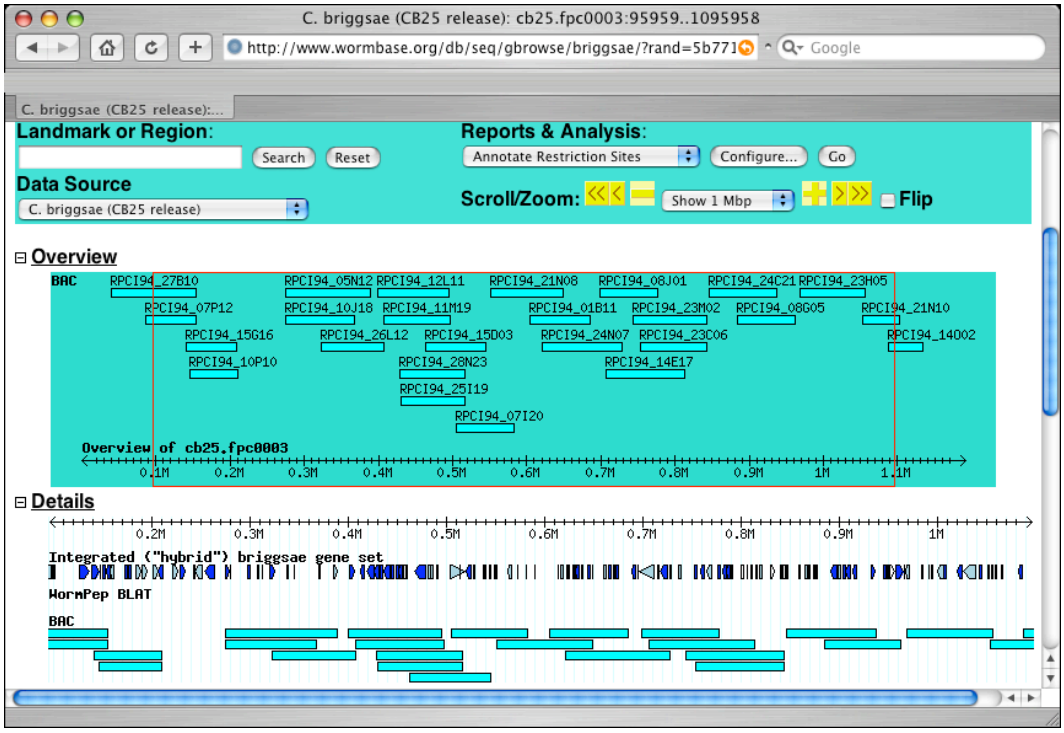


Figure 2. Graphic representation of the BAC clones in the Wormbase browser

2.2.1.3 C. *elegans* fosmid library

Recently a large complex fosmid library (5.7-fold coverage) was constructed and mapped onto the genome by paired end-sequencing[68]. The library covers about 85% of the genome.

The fosmid library became available only after most of the work on this project was finished using C. *briggsae* BACs. However, we confirmed that the protocol that we established using the C. *briggsae* BAC clones can be applied to the fosmid clones with only minor modifications (see Chapter 2.2.7).

2.2.2 Recombineering toolkit

2.2.2.1 Subcloning vectors

Subcloning of the gene of interest from the BAC clone into a vector carrying a transformation marker for integrative transformation in C. *elegans* can be achieved in a single step by homologous recombination mediated gap repair. In this approach a linear vector with regions of homology in the ends recombines with the region of interest on the BAC clone to generate a closed circular construct[14,17]. However, the commonly used pAZ series of vectors[55], based on the *unc–119* transformation marker, are too large for recombineering based subcloning. We generated and tested a set of new transformation vectors specifically optimized for gap repair subcloning (Figure 3). Linear subcloning cassettes can be generated from these template plasmids by PCR amplification with primers containing 50 nucleotide long homology arms to the region of interest.

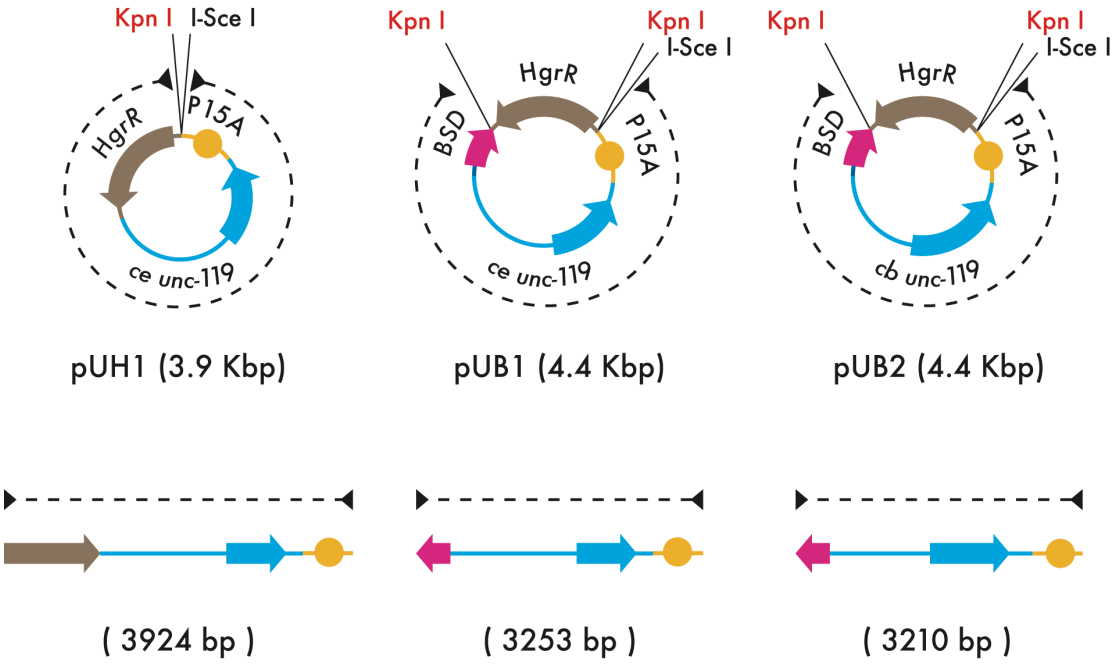


Figure 3. Minimal subcloning vectors with the *unc-119* transformation marker. The dotted line depicts the extent of the PCR from the template plasmid (above) for generation of the subcloning cassettes (below).

The first generation vector consisted of the low copy P15A origin of replication, a C. *elegans unc-119* promoter::cDNA fusion[69] and a hygromycin resistance gene (pUH1 for *unc-119* / hygromycin). A unique Kpn I site allows for linearization of the vector before PCR amplification. In addition a site for the rare cutting intron encoded endonuclease I–SceI[70] was included to allow linearization of the final construct prior to bombardment. In the second version of the plasmid (pUB1 for *unc-119* / blasticidin), we added a blasticidin selection marker (BSD) followed by a second KpnI site. The blasticidin resistance gene provides better selection and is significantly smaller than the hygromycin resistance gene. Bombardment transformation of these two constructs in *unc-119* mutant (ED3) worms rescued the severe uncoordinated phenotype as expected (Table 1). However, the number of integrated lines was low and some of the other phenotypes typical for the *unc-119* mutants – slow growth, dumpy and dauer larva formation defective were often not rescued.

Table 1. Ballistic transformation efficiencies with the new subcloning vectors.

		unc-119 rescue	
construct	screened plates	total	integrated
pAZ132	40	18	6
pUH1	40	20	4
pUB1	80	52	6
pUB2	80	56	22

In the third generation construct (pUB2) the C. *elegans unc–119* promoter::cDNA fusion was replaced by the genomic region containing the C.*briggsae unc-119* ortholog. The 2.1 kb genomic region that we used has been shown to completely rescue *unc-119* knockout in C. *elegans*[71]. In a side by side comparison the construct based on the genomic rescue region performed better. After screening 80 plates for each construct, the number of integrated lines for pUB2 was higher and the number of lines with partial rescue of *unc-119* phenotypes was reduced. The final size of the PCR amplified subcloning cassette from the optimal pUB2 vector is just 3.2 kb, which makes it suitable for gap repair subcloning.

2.2.2.2 Tagging cassettes

We generated a set of carboxy-terminal tagging cassettes consisting of a tag (eGFP, eYFP, or the TAP-tag) and a kanamycin resistance (KmR) marker (Figure 4). Linear tagging cassettes for Red/ET mediated insertion in place of the stop codon can be generated from these template plasmids by PCR amplification with primers containing 50 nucleotide long homology arms. The homology arms are selected to correspond to the 50 bp upstream (for the forward tagging primer) and downstream (for the reverse tagging primers) of the stop codon.

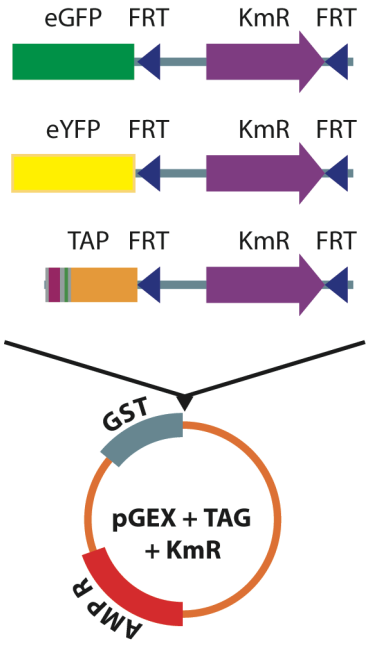


Figure 4. Cassettes for protein tagging

The kanamycin resistance marker is flanked by two Flp recombinase target (FRT) sites and can be removed by Flp mediated site-specific recombination. Removal of the kanamycin resistance gene is required to avoid any potential negative effect on gene regulation upon transformation in C. *elegans*. This extra step restores the natural genomic context and improves the chances of securing correct gene expression.

2.2.3 Proof of principle and strategy optimization

2.2.3.1 The test gene: *lin-59*

To test the recombination based cloning with BAC clones we applied it to tag the C. *briggsae* homolog of *lin-59*. *Lin-59* is the closest worm homolog of the *Drosophila's* Trithorax group gene *ash1*. *Ash1* has a characteristic domain structure of a PHD finger domain, a bromo adjacent domain and the catalytic histone methyltransferase SET domain. In *Drosophila ash1* has been shown to regulate hox gene expression through histone tail methylation. The C. *elegans* homolog was initially identified in a screen for defects in hindgut morphology and development [72]. *Lin-59* appears to be an essential gene. Several weak mutants have been identified and all of them have developmental abnormalities including various defects in the morphology and development of the hermaphrodite vulva, the male tail and the nervous system and hindgut in both sexes. Expression of the hox genes *mab-5, egl-5* and *lin-39* is downregulated in *lin-59* mutants.

Lin-59 is a relatively large protein of about 1000 aa, which makes it a good test gene for recombineering based cloning. The two orthologs from C. *elegans* and C. *briggsae* are well conserved (Figure 5). The genomic region containing the C. *briggsae* ortholog and all the surrounding noncoding regions is almost 12 kbp, which is significantly larger than the average gene size for both species.

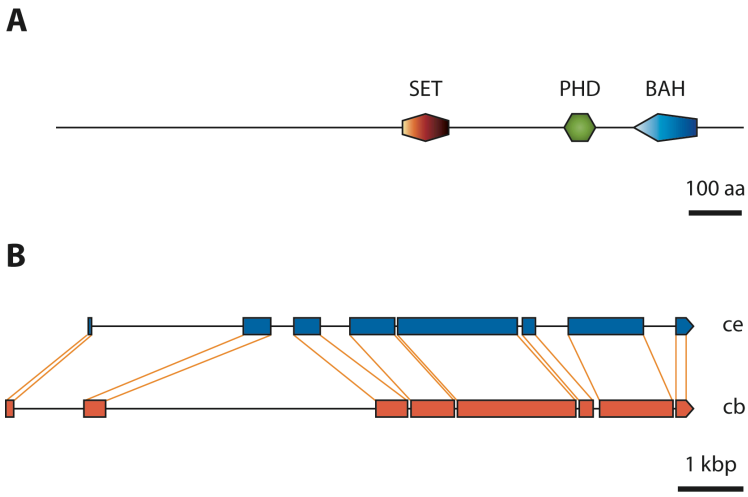


Figure 5. *lin-59* A. Protein domain architecture. B Gene structure of the C. *elegans* (ce) and C. *briggsae (cb)* orthologs.

2.2.3.2 Initial cloning strategy

Our initial strategy, outlined in Figure 6, was to first subclone the region of interest from the BAC clone into the *unc-119* containing vector pUB1 and to then insert the tag in the subcloned gene by a combination of Red/ET recombination and Flp mediated site-specific recombination. The Red/ET and Flp/FRT recombination potentials were transferred to the cells by transformation of the expression plasmids pSC101BADgbaA[25] and pSC101CIFlpe[73]. Both plasmids have the temperature sensitive pSC101 replication origin. They replicate normally at 30°C but inefficiently at 37°C and can be conveniently removed from the cells by overnight growth at 37°C.

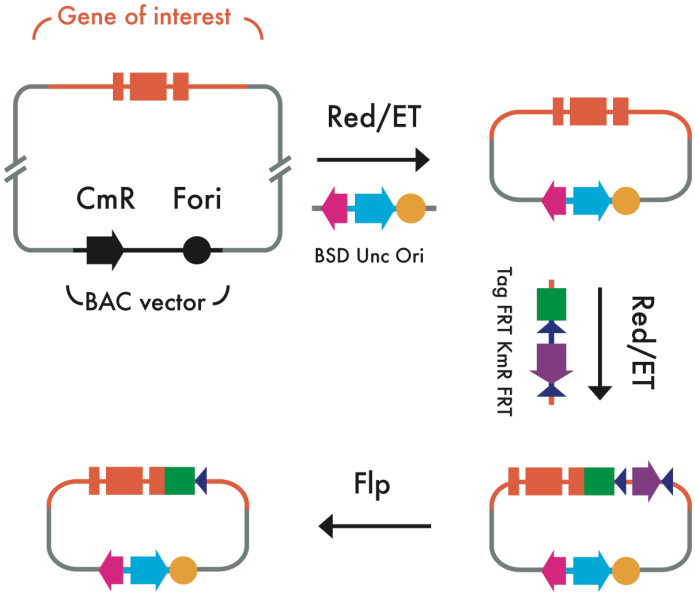


Figure 6. Initial tagging strategy. See text for detailed description; CmR: chloramphenicol resistance gene, Fori: BAC replication origin, Unc: *unc-119* transformation marker for C. *elegans*; KmR: kanamycin resistance gene; Tag (green) eGFP, eYFP or the TAP tag.

Subcloning of the *lin-59* genomic region (from the closest upstream to the closest downstream gene stop or start codon) by Red/ET recombination mediated gap repair was straightforward. All checked colonies had the expected restriction pattern (Figure 7 and data not shown). Traces of the BAC were present in all

plasmid preps. After retransformation of the purified plasmid, the contaminating BAC was lost.

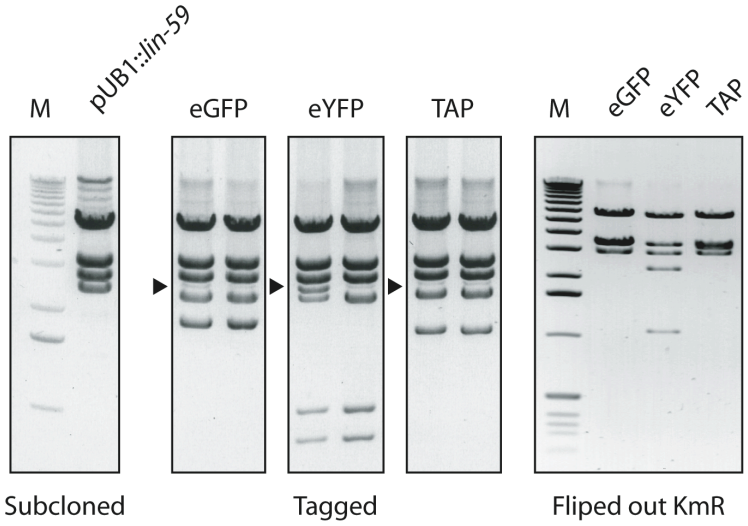


Figure 7. *Lin-59* tagging with the initial strategy. Pst I digest of the product of each recombineering step. The arrows indicate contaminating band coming from the unmodified plasmid. See also the supplementary data in Chapter 5.

Insertion of the tagging cassettes by Red/ET recombination was very efficient: hundreds of clones were obtained per transformation, and the expected restriction pattern was confirmed for all checked clones (Figure 7 and data not shown). However, the untagged plasmid was also present in all checked clones. Plating again on selection and growing a new culture did not solve the problem. Only after retransformation the untagged plasmid was lost and we could proceed to the next step - removal of the kanamycin marker. This was achieved by transformation of the pSC101CIFlpe[73] plasmid and heat shock inducible expression of Flpe recombinase. The loss of the kanamycin resistance gene was confirmed by the absence of growth on selection and restriction digest analysis (Figure 7 and data not shown).

2.2.3.3 Improved strategy

The major problem with the initial strategy was the presence of the BAC after the subcloning step and the untagged plasmid after the tagging step. This problem is caused by carryover of unmodified target molecules within the same cell.

The F replicon machinery ensures correct segregation of one copy of the BAC to each daughter cell, and can maintain the BAC for many generations even without selection. For the subcloning vector we selected the low to moderate (5 to 30) copy number P15A replication origin, to avoid problems with insert rearrangements that often occur with large constructs based on high copy origins. This however did not remove the problem of carryover of unrecombined plasmids commonly observed with recombineering of multicopy target molecules. Since distribution of the plasmid to the daughter cells is stochastic and there is no specific pressure against the cells that still contain unmodified copies, significant time might be required to lose the plasmid.

To overcome this obstacle we tried an alternative strategy in which the two-step tagging was performed on the single copy BAC and the region of interest was then subcloned (Figure 8A). With this improved strategy, unmodified BAC was never observed after the tagging step (Figure 8B and data not shown). Flipout efficiency was near 100% when preformed on the single copy BAC. We checked 48 single clones after flipout by replica plaiting on either chloramphenicol or kanamycin/chloramphenicol plates, and all clones had lost the kanamycin resistance (data not shown).

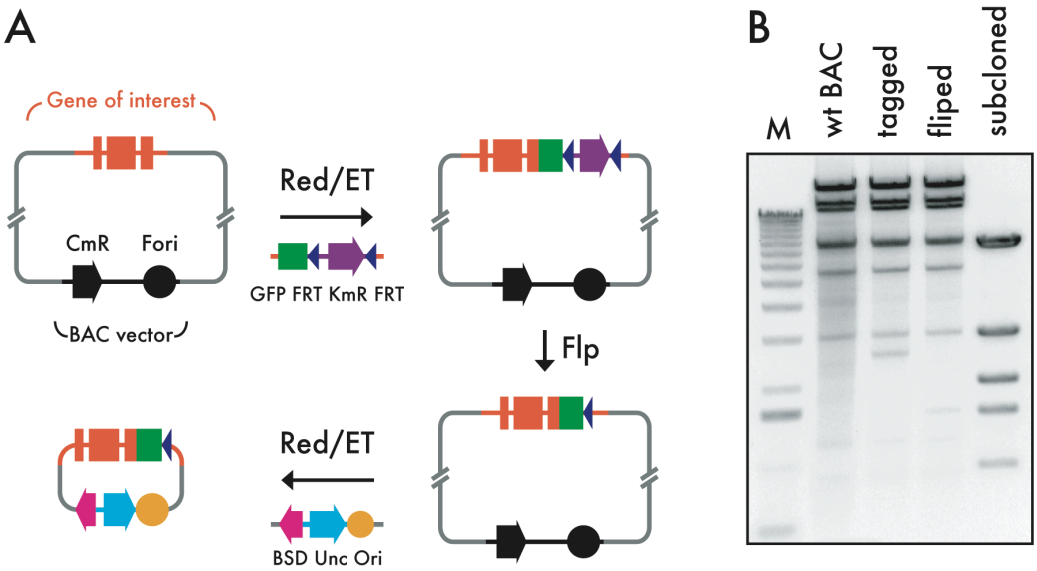


Figure 8. Improved tagging strategy. A cloning scheme. B Age I digest of the product from each step. See also the supplementary data in Chapter 5.

2.2.4 High throughput capable tagging method

In the initial protocol optimization we used the previously described pSC101BADgbaA[25] and pSC101CIFlpe[73] expression plasmids for inducible expression of the recombination proteins. The whole process has only three principal steps: tagging, flip out of the kanamycin resistance marker and subcloning. However, the need to transform different recombination plasmids for the Red/ET and Flp recombination steps slows down the procedure. Furthermore plating and picking clones at each step requires manual handling, and are hard to directly scale up for genomic applications.

In a follow up study we investigated the possibility for scale up of the tagging protocol[74]. We identified the need to use two different expression vectors for Red/ET and Flp recombination and the manual colony picking as the two bottlenecks of the tagging protocol. To address the first problem we generated and tested dual expression vectors with two independently inducible promoters for regulated expression of both recombination potentials (Chapter 2.2.4.1). The manual plating and colony picking can be avoided using specialized robots, at significant cost. Alternatively, solid medium plating can be avoided altogether and all steps could be performed in liquid culture, which is much easier to scale up using automated liquid handling devices. In Chapter 2.2.4.3 we show that after some optimization the efficiency of the protocol is high enough to permit continuous non-clonal selection throughout the whole procedure.

2.2.4.1 Dual expression plasmids

We tested several combinations of promoters in order to achieve independently regulated expression of both the Red operon and Flp recombinase from the same promoter (Figure 9). Transcription termination sites flanking the expression cassettes were included to prevent the RNA polymerase from running into neighboring genes. All constructs have the pSC101 origin and can be removed from the cells by temperature shift in the absence of selection, so that the final product is not contaminated by the expression plasmid.

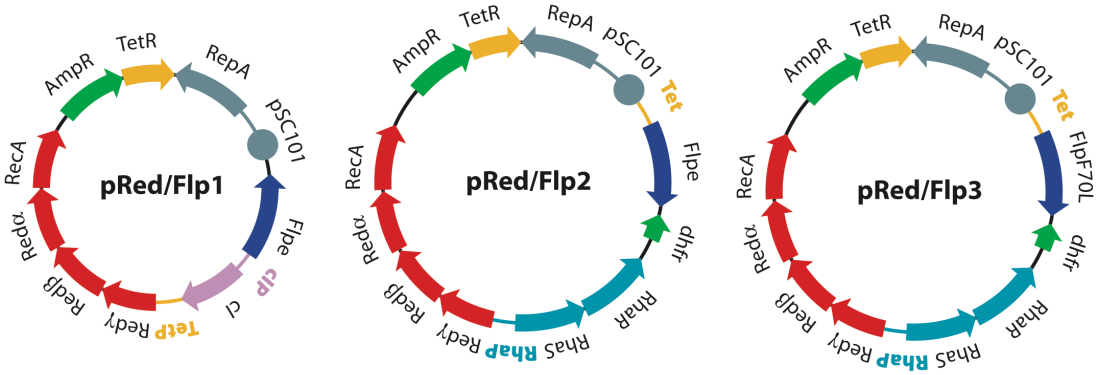


Figure 9. Dual expression plasmids for inducible Red/ET and Flp/FRT recombination. See the text for detailed description.

In the first vector design that we tried (Figure 9; pRed/Flp1) the Red proteins were under the control of the tetracycline inducible promoter[75], and Flpe was under the control of the CI578 thermosensitive promoter from the pSC101CIFlpe plasmid[73]. The CI578 promoter is active between 37°C and 42°C and is turned off by the cI repressor at temperatures under 32°C. To avoid premature Flpe expression the Red/ET steps had to be performed at 30°C, which turned out to be too inefficient (data not shown).

In the second version of the plasmid (Figure 9; pRedFlp2) we put Flpe under the tetracycline inducible promoter[75] and the red operon under the control of the L–rhamnose inducible promoter and the regulatory operon consisting of the RhaS and RhaR genes[76]. Red/ET recombination efficiency with pRed/Flp2 was similar to that of the L-arabinose inducible pSC101BADgbaA vector (data not shown). Flpe expression was tightly repressed in uninduced conditions as judged by western blot with antibody against Flpe (Figure 10).

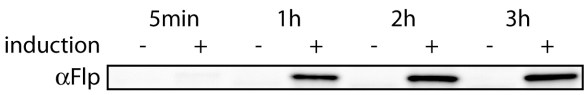


Figure 10. Timecourse of Flpe expression from pRed/Flpe2. Western blot analysis with anti Flp antibody at different time points of anhydrotetracycline induction.

To functionally test the kanamycin cassette excision we transformed pRedFlp2 in the strain containing the eGFP FRT KmR FRT tagged *lin-59* BAC. Single clones were grown without induction to saturation and 1µl was spotted on either

chloramphenicol (the BAC vector resistance) or kanamycin/chloramphenicol. All clones grew on chloramphenicol but none of them grew on kanamycin, showing that even the undetectable amounts of Flpe present in the cells in the absence of induction provide sufficient basal activity to drive premature excision of the FRT KmR FRT cassette.

Flpe is a highly active and thermostable version of the Flp recombinase generated by several rounds of directed evolution[73]. In terms of ensuring strict control over the recombination window the high activity of Flpe turned out be a disadvantage. In the next version of the plasmid (Figure 9; pRedFlp3) we replaced Flpe with a debilitated mutant of Flp (F70L)[77]. That essentially eliminated the problem of the basal Flp activity (Figure 11). The functional test showed no signs of premature cassette excision. At the same time, FlpF70L was sufficiently active to drive nearly complete KmR gene removal upon induction. The tight control of both the Red operon and FlpF70L expression form pRedFlp3 was confirmed by western blot.

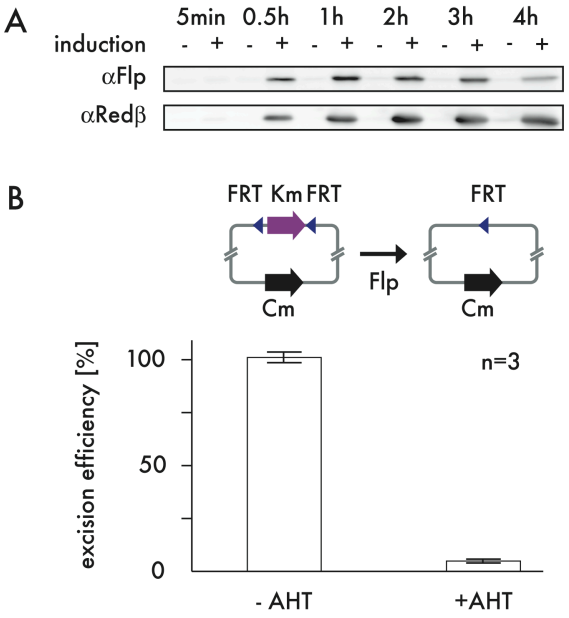


Figure 11. Controlled expression of recombinases from pRedFlp3 (A) Western blot analysis with anti Flp or anti Red beta antibody at different time points of induction with either anhydrotetracyline (AHT) or L-rhamnose. (B) Cells containing EGFP-FRT-KmR-FRT tagged BAC and the pRedFlp3 plasmid were grown in liquid culture in the absence or the presence of AHT to saturation and were plated in serial dilutions to obtain single colonies on plates with either chloramphenicol or kanamycin. The excision efficiency is presented as the number of kanamycin resistant colonies for 100 chloramphenicol resistant colonies.

2.2.4.2 R6K origin based tagging cassette template

To further facilitate the liquid culture recombineering we moved the GFP FRT KmR FRT cassette to a vector with the R6K origin of replication (Figure 12).

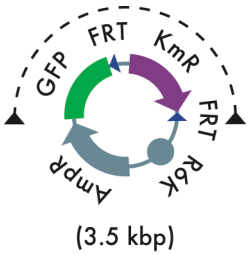


Figure 12. pR6KGFP. The interupted line depicts the extent of the PCR product generated from this template.

The R6K origin requires *in trans* supply of the product of the *pir*-encoded π protein. Plasmids with this origin can be propagated in strains carrying the *pir* gene, but not in normal cloning strains such as the BAC host DH10B[25,78]. As a result, the PCR generated targeting cassette can be used directly, without removing the template plasmid. This completely eliminates the background at the tagging step, and streamlines the procedure.

2.2.4.3 Liquid culture recombineering pipeline

After finding a configuration that permits both Red/ET and Flp/FRT recombination to be done using a single expression vector, we explored the applicability of liquid culture selection for multistep recombineering. We repeated the tagging of *lin-59* with EGFP using 1ml liquid cultures in 1,5 ml Eppendorf tubes as described in Figure 13. The plating steps were replaced with direct inoculation of a fresh 1ml culture supplemented with the appropriate antibiotics. The inoculation volume was 20 µl from overnight culture and 50 µl after the transformation steps. The culture temperature after transformation of pRed/Flp3 was 30°C, except at the Red/ET steps when growth at 37°C was required for optimal Red expression and recombination. Chloramphenicol selection for the BAC and ampicillin/trimethoprim double selection for the pRed/Flp3 plasmid were kept until

the final subcloning step when the cells were cultured at 37°C overnight only in the presence of blasticidin in order to lose the expression vector and the BAC.

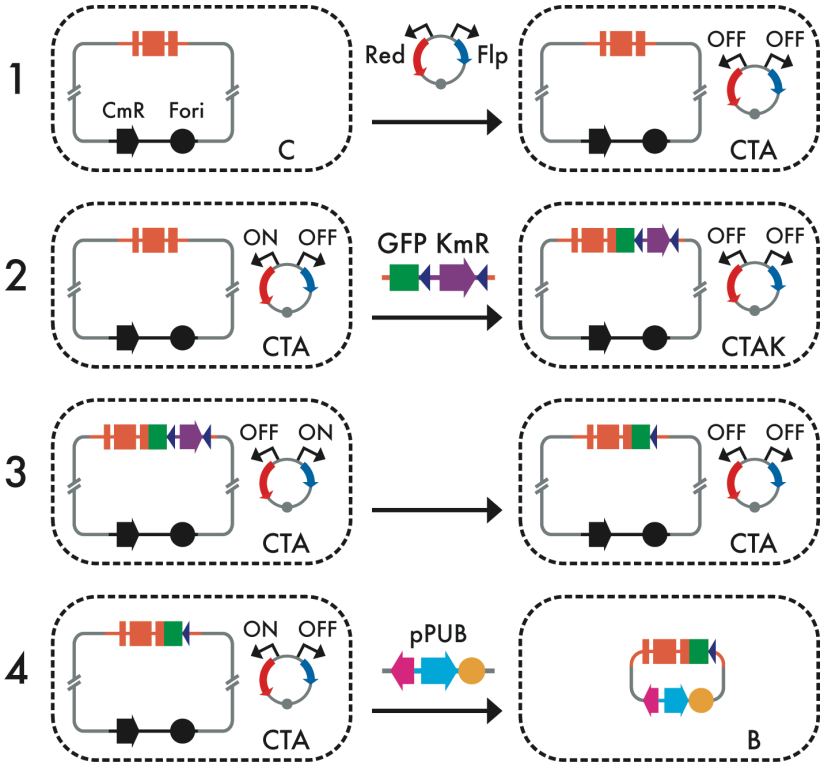


Figure 13. Liquid culture recombineering pipeline. Suitable BAC clone for the gene of interest (depicted in orange) is chosen and all recombineering steps are performed in the original BAC host cells (illustrated with dotted line). In Step 1, the host is transformed with the dual Red/Flp plasmid. In Step 2, expression of the Red operon (depicted in red) is induced with rhamnose and the cells are then electroporated with the EGFP – kanamycin resistance (KmR) cassette. In Step 3, expression of Flp recombinase is induced with anhydrotetracycline. Flp binds to the FRT sites (depicted as blue triangles) and excises the KmR gene. In Step 4, expression of the Red operon is again induced with rhamnose and the cells are then electroporated with the pUB2 subcloning vector, followed by selection for blasticidin, temperature shift to 37°C and omission of all other antibiotics. The inclusion of antibiotics is indicated within each state; C – chloramphenicol, T - trimethoprim, A - ampicillin, K – kanamycin, B – blasticidin;

The cells grew under selection as expected. Notably no growth was observed in the uninduced controls at the Red/ET steps. At the final step 50 µl were transferred either to liquid culture or to plates to get single colonies. To evaluate the efficiency of the cloning procedure 24 individual clones were picked, grown in liquid and spotted on either blasticidin, chloramphenicol or kanamycin selection All clones had lost kanamycin resistance, indicating near complete Flp recombination. Five of the clones appeared to have lost the parent BAC, as

indicated by the loss of chloramphenicol resistance. (Figure 14A). After restriction digest analysis only one clone showed an incorrect recombinant (lane 12 in Figure 14B). In agreement with the chloramphenicol resistance test, bands from the parent BAC were present in some of the clones. No traces of the pRedFlp3 plasmid were observed.

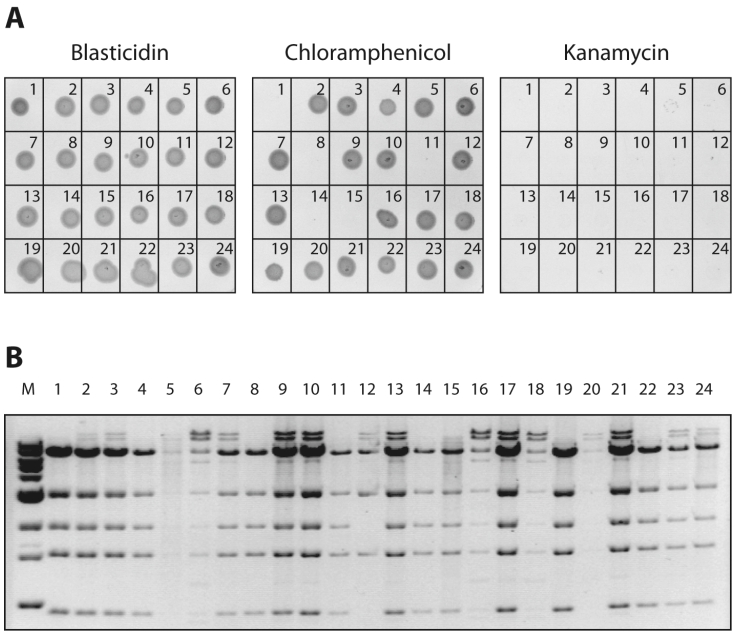


Figure 14. Recombineering outcome with individual clones. 24 colonies were taken for (A) replating onto either blasticidin, chloramphenicol or kanamycin selection only (B) growth and mini-preps for Age I digestion. 21 clones appear correct. Only lane 12 shows an incorrect recombinant, lanes 5 and 20 are probably failed mini-preps.

2.2.4.4 Parallel processing of multiple genes

The liquid culture recombineering pipeline was designed to permit simultaneous processing of multiple genes. To evaluate the efficiency of the protocol we processed 12 genes in parallel in two independent experiments, using 1ml cultures in Eppendorf tubes. The genomic region from the closest upstream gene stop or start codon to the closest downstream gene stop or start codon was subcloned for all genes, unless the tagged gene is part of an operon, in which case the whole operon was taken in the same way (Figure 15).

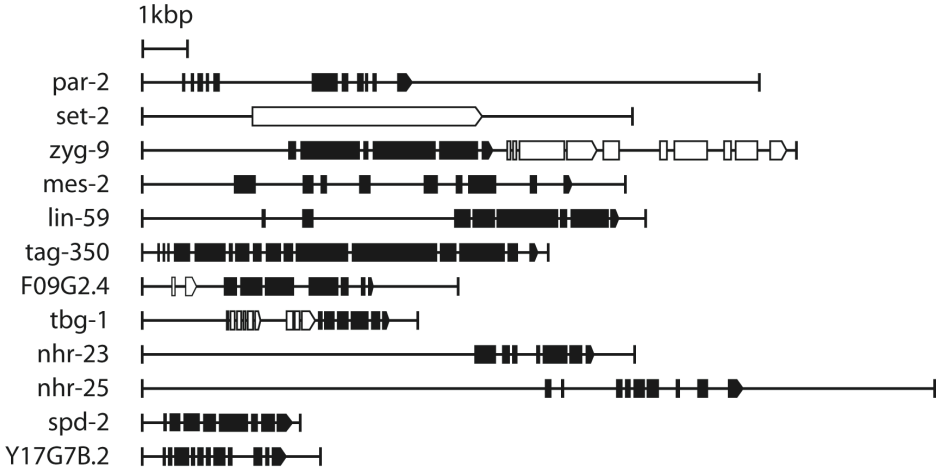


Figure 15. Gene models of the subcloned regions for the 12 genes tested in the recombineering pipeline. The target gene exons are depicted as filled boxes. Exons of additional genes involved in the same operon are depicted as empty boxes, except for set–2, where the whole gene is a single empty box because the exon structure is not known.

Nine of the genes were successfully tagged and subcloned in both experiments (Figure 16A). Tagging of one of the genes failed in one of the experiments, but was successful in the other (*nhr*-23). Two genes failed in both experiments at the tagging step (Y17G7B.2), or at the subcloning step (*spd*-2). The final cultures were streaked onto plates and two colonies for each product were examined for correct recombination by restriction digest. All checked clones were correct (Figure 16B and data not shown).

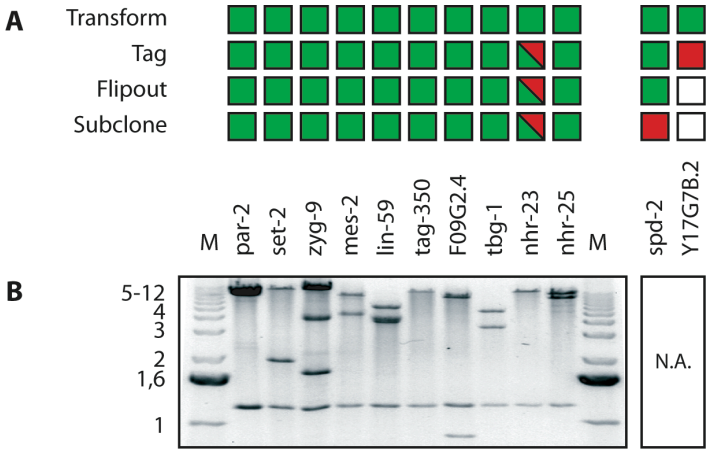


Figure 16. Parallel tagging of multiple genes in liquid culture. (A) Each box represents the outcome of the steps in two independent experiments. Green depicts growth in selection as expected, red means no growth and half red, half green box indicates that this step was successful in only one of the two experiments. (B) An agarose gel of Dre I restriction digestions to evaluate the success of the pipeline, with the names of the C. *elegans* orthologs of the genes given on top. See also the supplementary data in Chapter 5.

2.2.5 Generation of transgenic worm lines

2.2.5.1 Transformation efficiency

Transgenic worm lines were generated by ballistic transformation as previously described[55] with some minor modifications (See Materials and Methods). Single unc rescued worms were transferred to new plates. Lines that resulted in only wild type phenotype worm populations for at least 3 generations were considered integrated. The observed transformation efficiency was close to that of the commonly used *unc-119* based pAZ132 vector (Table 2).

Table 2. Ballistic transformation with the transgenes generated in the recombineering pipeline.

		unc-119 rescue		GFP positive
construct	screened plates	total	integrated	(integrated)
pUB2::lin-59::GFP	80	34	16	6
pUB2::lin-59::GFP I-SceI cut	80	42	8	2
pUB2::par-2::GFP	40	ND	5	0
pUB2::set-2::GFP	40	ND	2	0
pUB2::zyg-9::GFP	40	ND	3	0
pUB2::mes-2::GFP	40	ND	3	0
pUB2::tag-350::GFP	40	ND	0	0
pUB2::F09G2.4::GFP	40	ND	5	3
pUB2::tbg-1::GFP	40	ND	3	1
pUB2::nhr-23::GFP	40	ND	4	1
pUB2::nhr-25::GFP	40	ND	5	2

After screening 40 plates for each construct stably integrated GFP positive lines were obtained for 5 of the genes. The transient lines were not taken into account. Using the pUB2::*lin-59::GFP* transgene we tested whether I-Sce I linearization of the construct prior to bombardment could increase the number of integrated and/or GFP positive lines, but no significant improvement was observed.

2.2.5.2 GFP expression patterns

2.2.5.2.1 Lin-59

All GFP positive lines showed a nuclear signal, first observed at the mid-blastula stage, and then expressed in many cells, throughout development and into adulthood. The GFP signal is stronger in the neurons, hypodermal cells and some unidentified cells in the head and the tail (Figure 17). This pattern of cb*lin-59::GFP*

expression is similar to that of a previously published *lin-59* promoter::*GFP* reporter transgenic line.

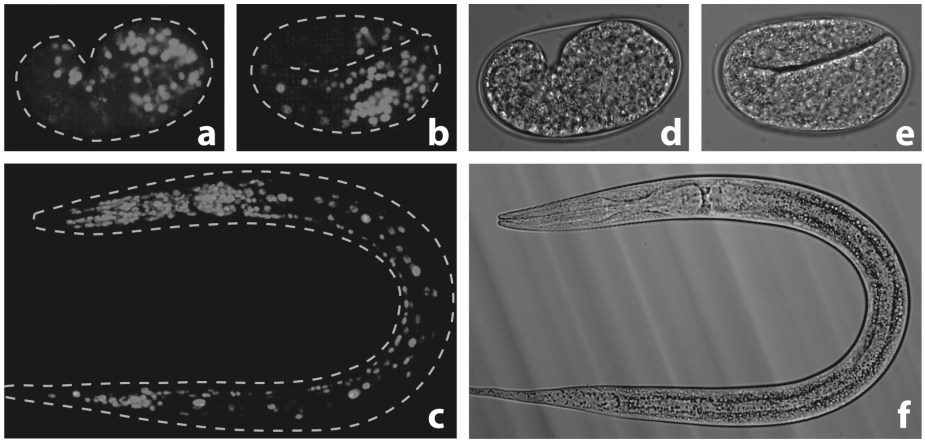


Figure 17. Expression pattern of the cb*lin-59*::*GFP* transgene. GFP fluorescence (a-c) and the corresponding bright field image (d-f) at different stages of development. a,d onset of morphogenesis; b,e elongation; c,f larva.

2.2.5.2.2 F09G2.4

F09G2.4 is the C. *elegans* ortholog (79% homology) of the 100 kDa subunit of the human RNA cleavage and polyadenylation specificity factor complex (CPSF). The C. *elegans* and C. *briggsae* orthologs are almost identical on protein level. In 3 independently generated lines, weak nuclear GFP expression was observed in body wall muscles and head muscles (Figure 18).

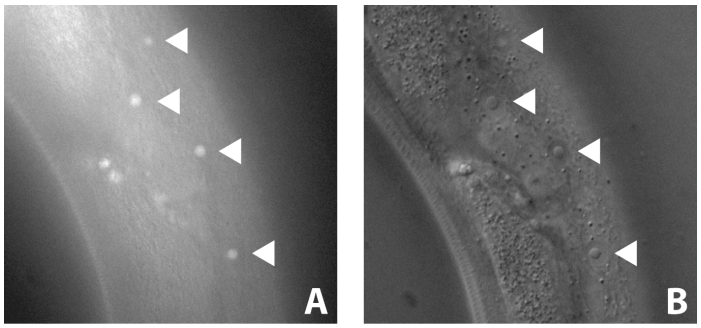


Figure 18. Expression pattern of F09G2.4. (A) GFP fluorescence; (B) DIC. Arrowheads point to the characteristically shaped body wall muscle nuclei.

This tissue specificity is somewhat surprising because cpsf2 is an essential subunit of the CPSF complex[79], and is therefore expected to be ubiquitously expressed. RNAi of F09G2.4 in C. *elegans* is embryonic lethal, and various other phenotypes have been reported in the surviving worms. Microarray and SAGE analysis data show that the message for F09G2.4 is present in all stages.
However, the same expression pattern was reported by the BC C. *elegans* Gene Expression Consortium using a promoter::*GFP* reporter construct[28]. It is possible that expression in other cell types is simply under the detection limit.

2.2.5.2.3 *Tbg-1*

The expression pattern of gamma tubulin (*tbg-1*) has been well characterized. It is expressed ubiquitsly, and shows a characteristic cell cycle dynamics. Gamma tubulin is detected in all cells as diffuse staining and rapidly localizes to the two centromeres during mitosis[80]. This pattern was exactly reproduced by the transgene (Figure 19). Gamma tubulin is the third gene in an operon together with F58A4.9 and *ubc-7* and in this case, the whole operon was subcloned. This result shows that the strategy is applicable to genes that are part of an operon.

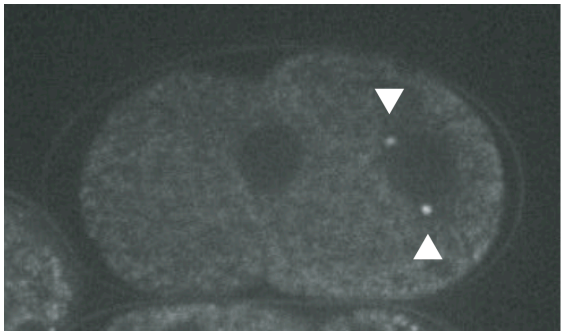


Figure 19. Subcellular localization of *tbg-1::GFP*. Gamma tubulin correctly localizes to the two centrosomes of the male pronucleus in the fertilized zygote.

2.2.5.2.4 Nhr-23

Nhr-23 is an orphan nuclear hormone receptor, required for proper epidermal development[81,82]. It is broadly expressed in the early embryo, and later becomes restricted to the hypodermal lineage[81]. The pUB2::*nhr-23::GFP* transgene correctly reproduced this pattern (Figure 20A).

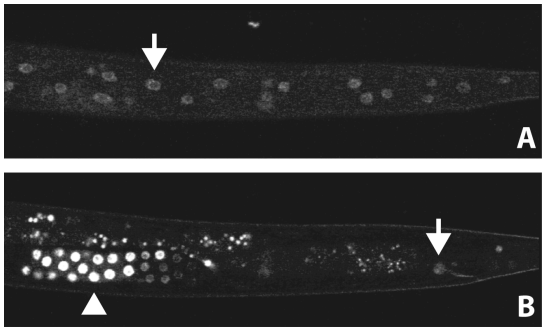


Figure 20. Expression pattern of *nhr-23*. Two optical sections of the same animal: A at the body surface, B midsection. Arrows point to hypodermal cell nuclei and arrowheads to germline nuclei.

In addition, we observed strong germline expression (Figure 20B), which was not detected in a previous study[81] using a promoter::*nhr–23*(truncated)::*GFP* reporter construct, probably because of germline silencing of the extrachromosomal transgenic array. The *nhr-23* mRNA is indeed expressed in the germline as shown by in situ hybridization[81].

2.2.5.2.5 Nhr-25

Nhr-25 encodes a nuclear hormone receptor orthologous to Drosophila Ftz-F1. It is required for embryogenesis, molting, vulval and gonadal development, and hypodermal development. *Nhr-25* is first expressed around the 100 cell stage and is strongly expressed throughout development in the hypodermal cells[83,84]. The pUB2::*nhr-25*::*GFP* transgene correctly reproduced the known expression pattern (Figure 21).

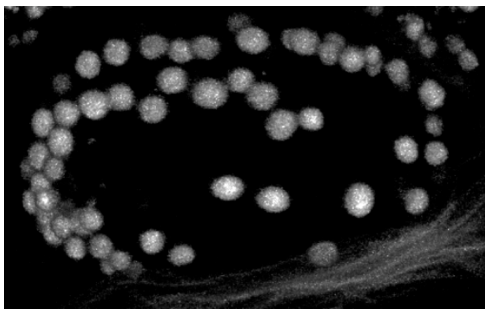


Figure 21. Expression pattern of *nhr-25* in the coma stage embryo. 3D reconstruction of the GFP signal from confocal sections.

2.2.6 Transgene rescue of RNAi induced phenotype

To prove that a transgene is functional, it must be able to rescue known mutant phenotypes. The use of cross species transgenes presents the opportunity to use RNAi to knockdown the endogenous gene[85]. Most C. *elegans* genes have an obvious C. *briggsae* ortholog, and are well conserved at the protein level. However, at the DNA level they are likely to be too diverse to have long stretches of completely identical nucleotides. Therefore, dsRNA targeted to the C. *elegans* gene is expected to have no effect on the C. *briggsae* transgene (Figure 22).

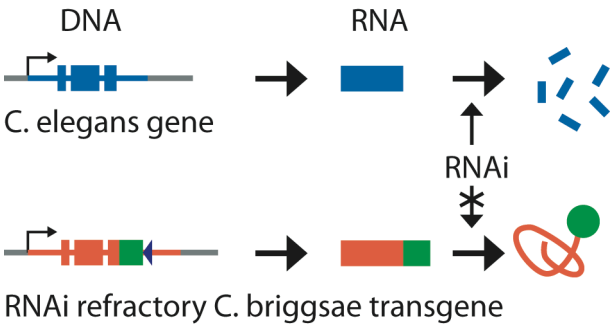


Figure 22. Third allele strategy for C. *elegans*. The host alleles are knocked-down by RNAi, however the sequence differences in the cross species ortholog preclude RNAi knock-down. Thereby the transgenic third allele can replace the endogenous protein.

RNAi inactivation of the endogenous gene makes the transgene the primary expressed copy or what we call a "third allele". A similar approach has been applied in tissue culture cells with cDNA[86] or BAC[87] based transgenes.

2.2.6.1 Rescue of *lin-59* RNAi phenotype by cross-species third allele

DNA sequence alignment of the coding sequence of the C. *briggsae* and the C. *elegans lin-59* orthologs showed that while the two genes share good overall homology no single stretch of 20 identical nucleotides could be found. This allowed us to use RNAi to specifically knock-down the endogenous gene. We used the RNAi by the feeding method in which the worms are grown on a lawn of dsRNA producing bacteria[88,89].

As previously shown[34], feeding wild type C. *elegans* with *lin-59* dsRNA resulted in no obvious phenotype, however various phenotypes were observed in the RNAi hypersensitive strain *rrf-3*, including defects in vulva development (Figure 23A).

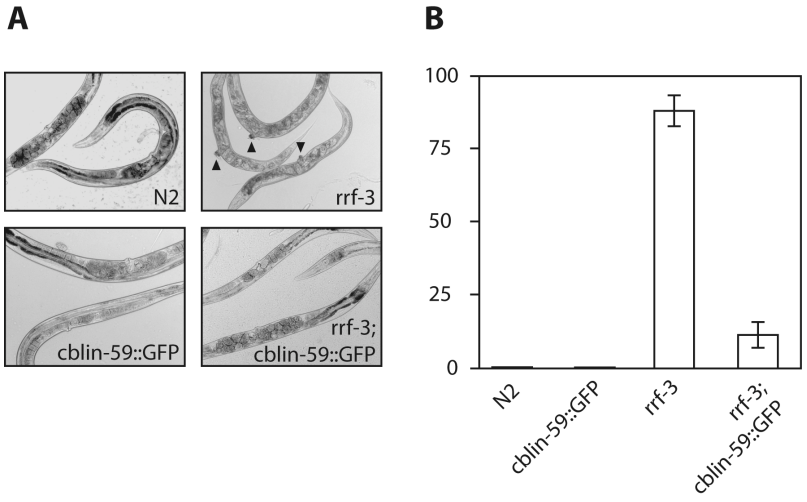


Figure 23. Rescue of an RNAi induced phenotype by a cross species transgene. (A) Knock-down of *lin-59* produces defects in vulva formation (arrowheads) in the RNAi hypersensitive *rrf-3* strain, but not in wt N2, cb*lin-59::GFP* or *rrf-3*;cb*lin-59::GFP* strains. (B) Penetrance of the defective vulva phenotype in three independent experiments. The error bars depict the highest and lowest values.

We crossed the transgenic line to *rrf-3* and compared the effect of *lin-59* RNAi in the presence or the absence of the transgene (Figure 23A). In three independent experiments the number of animals with vulva defects after *lin-59* RNAi was reduced from almost 90% for *rrf-3* to just above 10% in the *rrf-3*;cb*lin-59::GFP* strain (Figure 23B). As expected the level of GFP expression in the *rrf-3;*cb*lin-59::GFP* was not obviously affected by RNAi (Figure 24).

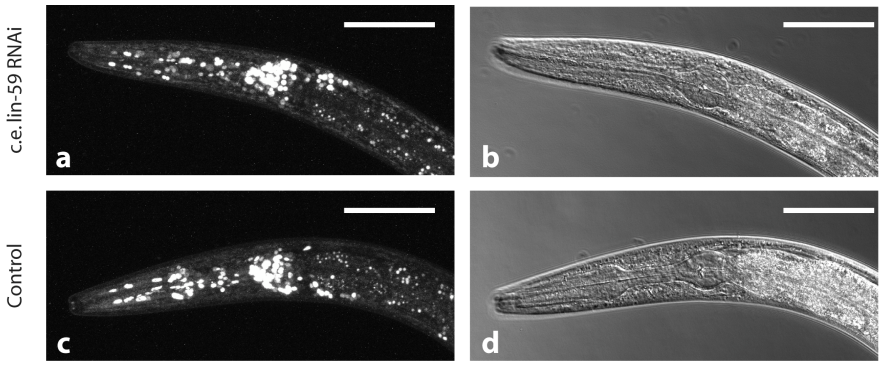


Figure 24. RNAi against C. *elegans lin-59* has no effect on GFP expression from the cross species transgene. The same level of GFP expression is observed in the *rrf-3;*cb*lin-59::GFP* worms fed with bacteria expressing either dsRNA against the C. *elegans lin-59* (a,b) or carrying the empty expression vector as a control (c,d).; (a,c) Maximum intensity projection of confocal stacks; (b,d) Bright field image; Scale bar 50 µm.

2.2.7 Liquid culture pipeline for the C. *elegans* fosmid library

The availability of the C. *elegans* fosmid library (introduced in Chapter 2.2.1.3) made it possible to apply the transgene based tagging approach to C. *elegans* genes. Since this library was introduced after most of the experiments were completed we only tested the performance of these clones in liquid culture recombineering.

Along with the F factor replicon the fosmid vector pCC1FOS contains an inducible high copy number oriV origin of replication. This origin is dependent on the product of the trfA gene, which is supplied by the fosmid host strain EPI300. Expression of TrfA is inducible with arabinose (Wild et al., 2002). In the absence of induction, the clones are maintained at single copy to reduce the likelihood of rearrangement. Induction of trfA leads to high copy replication from oriV, facilitating recovery of large amounts of DNA. The TrfA gene was introduced into the genome of the fosmid host strain EPI300 by transposon insertion along with a trimethoprim resistance marker.

2.2.7.1 Liquid culture recombineering with fosmids

Due to the presence of a trimethoprim resistance gene in the host genome, we could not use this selection marker for maintenance of the pRedFlp plasmid. Our preliminary results showed that continuous selection for pRedFlp in liquid culture only with ampicillin was not possible. Ampicillin resistance is driven by the beta lactamase enzyme, which is exported to the periplasmic space and soon accumulates in the culture medium, leading to complete hydrolysis of all the ampicillin. Growth in the absence of selection led to loss of pRedFlp even at increased ampicillin concentrations (to 200 µg/ml). Exchange of ampicillin for the more stable analog carbencillin did not solve this problem. To overcome this issue we exchanged the ampicillin selection marker for hygromycin, resulting in a new dual expression plasmid pRedFlp4 (Figure 25).

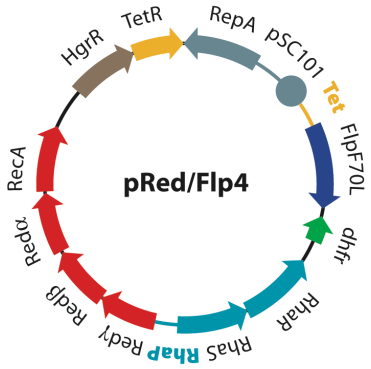


Figure 25. pRedFlp4. Hygromycin resistant vector for Red/ET and Flp recombination

We tested the performance of the fosmids in liquid culture recombineering by tagging C. *elegans lin-59*. In addition, we tested a new recombineering strategy in which the subcloning is avoided and the *unc-119* selection gene is directly introduced into the fosmid vector backbone (Figure 26). Using the hygromycin resistant pRedFlp4 plasmid, the liquid culture protocol worked as expected for both strategies: subcloning and inserting the *unc-119*-BSD cassette in the fosmid vector. The final products were plated on blasticidin selection, and the correct recombination was confirmed by restriction analysis (data not shown). As with the C. briggsae BACs the preps of the subcloned *lin-59* contained traces of the parent fosmid. In contrast, the fosmids modified with the new strategy appeared to be 100% the expected product without carryover from the previous steps.

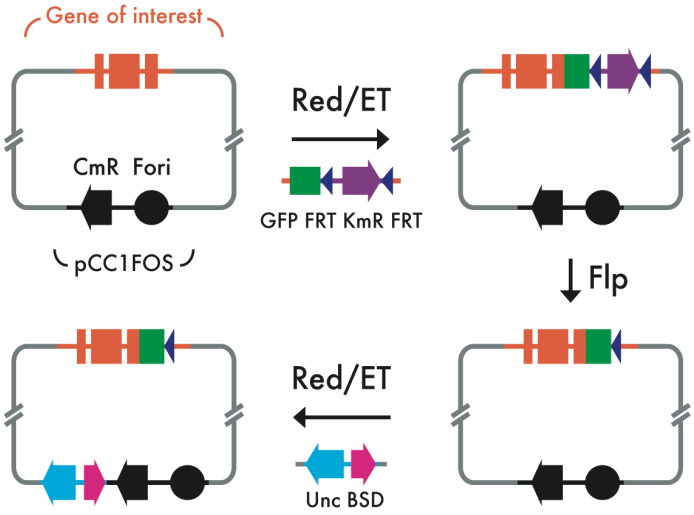


Figure 26. Generic fosmid cloning strategy

2.2.7.2 Web based clone and oligo selection tool

The new strategy for fosmid modification by direct insertion of the *unc-119* gene in vector backbone requires careful selection of the fosmid clone. The likelihood of successful genomic cointegration of the tagged gene and the *unc-119* marker decreases as the distance between them increases. We designed an algorithm for automatic selection of a suitable clone (Figure 27).

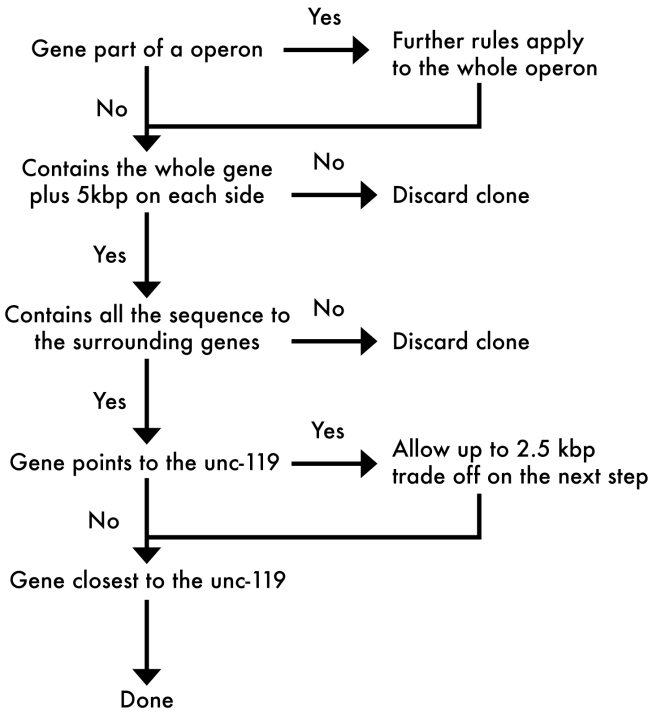


Figure 27. Clone selection algorithm

First, clones that contain the gene plus the sequence at least up to the neighboring genes, but not less than 5kbp are selected. The optimal clone is then chosen in a tradeoff process whereby the distance between the target and the *unc-119* insertion point can be increased by up to 2.5 kbp if the target gene runs towards *unc-119*. In this way we can select for clones that are likely to result in successful transgenesis, yet will contain enough surrounding sequence to ensure correct regulation. A web-based interface of the algorithm is available at:

http://www.mitocheck.org/cgi-bin/FOSfinder

The program also finds the oligonucleotides for generation of the tagging cassette and allows bach processing of multiple genes.

2.3 Discussion

2.3.1 Strategy optimization

Our goal was to establish an optimal strategy for generation of transgenes from genomic BAC clones with minimal disturbance to the natural genomic context. In order to ensure correct regulation from the endogenous 3'UTR it is important that only the tag and no other exogenous sequences such as selection marker genes are left after insertion of the tag coding sequence in front of the stop codon. This can be achieved solely by homologous recombination in two steps, first inserting a selection-counterselection cassette, and then replacing this cassette with the tag, using homologous recombination and counterselection pressure[90,91]. This approach results in completely seamless insertion and is the method of choice for generation of precise point mutations. However, all counterselection approaches suffer from high background, which can result from one or more of the following: spontaneous or PCR introduced mutations leading to inactivation of the counterselection gene; spontaneous or Red induced rearrangements leading to deletion of the counterselection gene; carryover of unmodified molecules from the previous step. All of these events lead to a molecule that is resistant to the counterselection pressure, but is not the expected recombination product. The false positive background is always an issue with counterselection strategies, especially when high throughput applications are considered. For these reasons we decided to insert the tag as a cassette with a FRT flanked selection marker that is subsequently removed by Flp/FRT mediated site-specific recombination. We used Flp/FRT and not Cre/Loxp for this step because the BAC vector already contains two lox sites. In *E. coli* Flp/FRT recombination is highly efficient and can proceed to near completion even without counterselection[77,92]. This approach is much more reliable, and practically eliminates the background at this step. This efficiency comes at the cost of leaving a "scar" of a single 34 bp FRT site in the insertion point. While this is a problem when precise mutations have to be seamlessly introduced in the coding sequence, for our purpose - insertion of a large fluorescent or affinity tag at the end of the gene this is not an issue.

Selection for stable integration of the transgene by ballistic transformation requires that the selectable marker is present on the same molecule. The first approach in which we achieved this was by subcloning the region of interest into a vector containing the *unc-119* marker gene. Subcloning the whole gene and all noncoding sequences up to the neighboring genes makes it likely that all regulatory sequences will be present in the transgene. At the same time by subcloning only the gene of interest, we isolate the functional effects of the transgene expression from the other genes present in the BAC clone. For the subcloning step, we generated a series of vectors suitable for gap repair mediated subcloning. Previous application of gap repair subcloning have identified self-closure of the vector as the major source of background in this approach[14]. Placing the origin and the selection gene at the opposite ends of the subcloning cassette reduces the probability of such event because intramolecular recombination would produce a molecule that cannot propagate under selection in *E. coli* and will be lost.

When generating the plasmids the focus was on minimizing their size to make it possible to add the homology arms to the subcloning cassette directly by PCR. This was achieved by using a very short cassette for selection in *E. coli* (BSD) and shortening of the C. *elegans* transformation marker *unc-119*. The previously published *unc-119* genomic fragment used as a rescue marker for transgenesis is 5.7 kbp long[55]. As an alternative, shorter rescuing constructs we tested an *unc-119* promoter::cDNA fusion[69] and the genomic region containing the C. *briggsae unc–119* ortholog, which codes for an almost identical protein but has much smaller introns. In our experiments the construct based on the rescuing genomic region performed better, perhaps due to the lack of intron-encoded regulatory sequences in the cDNA based construct. These observations are in agreement with a recent study[93], in which that an transformation of an intronless C. *elegans* unc-119 transgenic construct results only in partial phenotype rescue.

The preliminary experiments showed that the optimal sequence of recombineering steps is to first tag the gene on the single copy BAC clone, and then to subclone the region of interest in the vector for ballistic transformation. This practically

eliminated the problem of carried-over unmodified plasmid, which was always observed when the tagging was performed on the pUB subclone. The p15A origin plasmids are maintained at a low (5-30) copy number, however Red/ET recombination is a rare event and is unlikely to happen with all target molecules present in the cell. The unmodified molecules in the same cell are therefore always in excess, and since there is no specific pressure against them it may take a significant time until they disappear from the plasmid pool. In contrast, when the target molecule is a BAC with the single copy F origin, complete recombination is achieved, both at the Red/ET and at the Flp/FRT steps.

As an alternative to subcloning, we introduced the *unc-119* marker for bombardment in the genomic clone vector backbone. The major advantage of this approach is that a single *unc-119* insertion cassette is used for all clones, which further streamlines the protocol and results in significant cost reduction. The better coverage and the smaller size of the fosmid clones makes it possible to find a clone in which the gene is relatively close to the *unc-119* insertion point in the vector backbone.

2.3.2 Liquid culture recombineering for high throughput applications

To date, multistep DNA engineering by either conventional or recombineering approaches has involved clonal selection on plates and screening of individual clones after each step. Previous studies have established the high fidelity of Red/ET recombineering[11-13,16,23]. That is, the frequency of illegitimate recombination is much less than the intended, homologous event mediated by Red or RecET proteins. Here we show that this high fidelity permits a new way to engineer DNA using sequential steps in liquid culture without cloning or checking until the final product. Liquid culture cloning is faster and much easier to scale up for parallel processing of multiple genes.

To facilitate the multistep recombineering procedure we generated a double expression vector for both Red/ET homologous recombination and Flp/FRT site-specific recombination. By introduction of an expression plasmid, all steps were performed in the original BAC host strain. This avoids the more difficult isolation and retransformation of the BAC into a special recombination proficient host such

as YZ2000[14] or DY380[17], and the concomitant need to recheck the integrity of the BAC. The continuous presence of the plasmid over several rounds of recombineering steps requires strictly controlled expression of the recombination potentials. This was achieved using the tightly controlled tetracycline inducible[75] and rhamnose inducible promoters[76]. We placed the Red operon under rhamnose induction because our previous utilizations[11,25] of arabinose for Red operon induction would have been incompatible with the C. *elegans* fosmid library[68]. These fosmids contain an arabinose inducible, copy number amplification circuit[94,95] that causes them to switch to multicopy mode of replication. So far the rhamnose system performs as well as, if not better than, the arabinose system for regulated expression of the Red operon and consequent control over the recombinogenic window. In the preliminary tests the recombination efficiency with fosmid clones was similar to that with the C. *briggsae* BACs.

2.3.3 Integrative transformation with BAC derived transgenes

We demonstrated that the transgenes generated in the recombineering pipeline could be used for bombardment based transformation with good transformation frequencies. We showed that even very large genes (*nhr-23*, *nhr-25*, *lin-59*) or genes part of an operon (*tbg-1*) could be successfully transformed. The observed GFP expression patterns matched previously described ones. However GFP expression was not always detectable. After screening 40 plates no GFP positive lines were found for the genes *par-2*, *set-2*, *zyg-9*, *mes-2*, and *tag-350*. It is possible that for some genes the GFP expression level is simply below the detection limit. A more likely explanation is that the transgene breaks down during the transformation procedure. Integration of the circular transgene occurs through random linearization. In order for the transgene to be correctly expressed the integrated fragment has to contain the whole gene along with the *unc-119* selection marker. In the case of *tag-350,* for example, the promoter and the 3′ noncoding regions are relatively small compared to the gene itself and this event would be very unlikely. Subcloning larger regions, or using whole fosmids as transgenes, may improve the rate of GFP positive integrative events.

Linearization of the transgene at a defined position prior to bombardment may increase the probability of the desired cointegration of the tagged gene and the *unc-119* selection marker. However, in our initial test linearization through the unique I-Sce I site in the subcloning vector backbone did not result in the expected increase of the number of GFP positive lines. It is possible that the free DNA ends increase the rate of unwanted recombination between the transgene molecules leading to array formation. Further tests with more genes and different transgene concentrations will be required to evaluate the effect of linearization on the rate of integrative transformation.

2.3.4 Third allele strategy

By using a cross-species C. *briggsae* transgene, we were able to selectively knock down the endogenous C. *elegans* gene by RNAi. Hence the transgene became the primary expressed copy of the gene. Because the transgene should include all of the regulatory elements required for appropriate expression, this "third allele" strategy presents a convenient methodological alternative to homologous recombination of the endogenous gene. It allows for quick evaluation of the functionality of the transgene when the endogenous gene has an obvious phenotype. For example, we showed that a C. *briggsae lin-59* transgene rescues a developmental defect caused by C. *elegans lin-59* knockdown. In addition, knockdown of the endogenous gene might be used to promote inclusion of the tagged protein into the native protein complexes or to counter transgene silencing.

This approach is not without limitations. The obvious concern is always that the cross species gene may not function exactly like the endogenous, or that the produced protein, despite the good conservation may not completely substitute for its counterpart. To distinguish such problems from more trivial technical reasons such as transgene rearrangements several independently obtained lines would have to be tested for phenotype rescue.

Another concern is that RNAi against the endogenous gene may target the transgene as well. This may occur through short regions of completely identical DNA sequences, but may as well be triggered by partially homologous sequences

containing a few mismatches. Off target RNAi has an additional implication – the observed phenotype can be a result (at least partially) from off-target effects on genes that cannot be rescued by the transgene. Off target effects in mammalian culture cells have been demonstrated with as little as 7nt of homology[96]. A recent computational analysis of 30 RNAi screens in *Drosophila* cells suggests that as many as 40% of the used dsRNAs may have off-targets, based on a 19 nt perfect match threshold[97]. The extent of this problem in C. *elegans* has not been thoroughly studied. Careful selection of the targeted region can help overcome such issues. Several algorithms now exist that can help in the design of RNAi experiments[98-100].

2.3.5 Comparison with other protein tagging methods for C. *elegans*

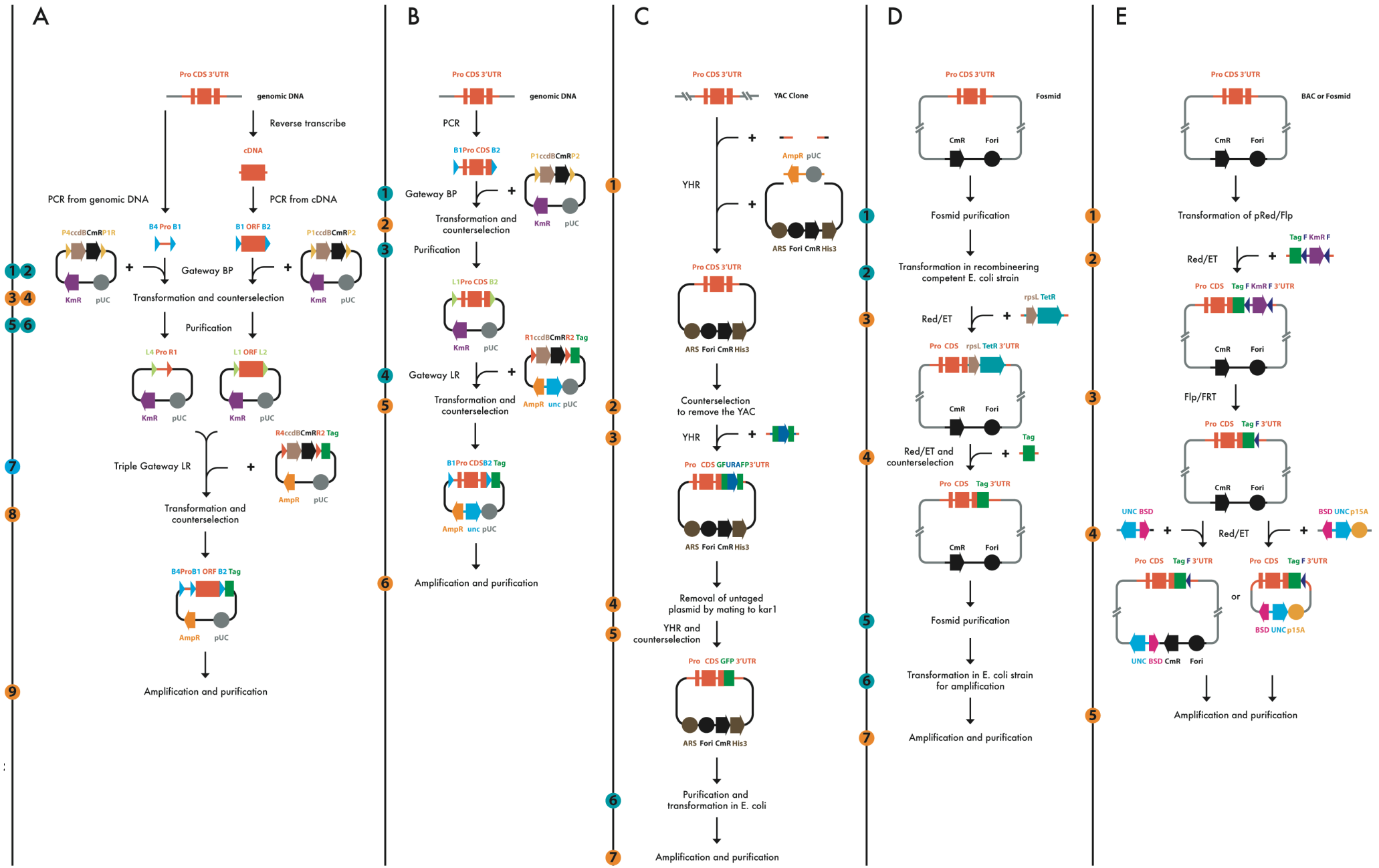
In this chapter we compare the recombineering pipeline application that we developed with the currently available recombination based methods for protein tagging in C. *elegans* (outlined in Table 3 and Figure 28).

Instead of giving comprehensive description of all approaches, we focus on several factors important for scale up to high throughput application: cloning efficiency, quality of the transgene, and cost.

Table 3. Comparison of the recombination mediated methods for protein tagging applied to C. *elegans*

	Dupuy et al.	Polanowska et al.	Sassi et al.	Dolphin and Hope	Sarov et al.
Recombinereeing host	in vitro	in vitro	*S. cerevisiae*	*E. coli*	*E. coli*
Recombinereeing method	Gateway	Gateway	YHR	Red/ET	Red/ET and Flp/FRT
promoter	YES (partial?)	YES (partial?)	YES	YES	YES
coding sequence	ORF	CDS (exons/introns)	CDS (exons/introns)	CDS (exons/introns)	CDS (exons/introns)
3'UTR	NO	NO	YES	YES	YES
applicable to operons	NO	NO	YES	YES	YES
PCR introduced mutations	YES	YES	NO	NO	NO
Experimental steps	9	6	7	7	5
Time for generation of single clone	6 days	5 days	40 days	8 days	5 days
Automation of all steps possible	YES	YES	NO	NO	YES
Transformation method	Transient	Integrative	Transient	Transient	Integrative
Expression level	Overexpression	Near endogenous	Overexpression	Overexpression	Near endogenous

Figure 28 (on the right). Overview of the recombination mediated methods for protein tagging applied to C. *elegans*. A. Duppy et al. 2004, B. Polanowska et al. 2004, C. Sassi et al 2005, D Dolphin et al 2006, E Sarov et al 2006; Methods A and B are based on Gateway recombination *in vitro;* Methods C-E use *in vivo* recombineering in S. cerevisiae (C) or E. coli (D,E). The experimental steps are numbered in blue (*in vitro* step) or yellow (*in vivo* step) circles.


A
Pro CDS 3'UTR
genomic DNA
Reverse transcribe
cDNA
PCR from genomic DNA
PCR from cDNA
P4ccdBCmRP1R
B4 Pro B1
B1 ORF B2
P1ccdBCmRP2
KmR
pUC
Gateway BP
Transformation and counterselection
Purification
L4 Pro R1
L1 ORF L2
Triple Gateway LR
R4ccdBCmRR2 Tag
AmpR
B4ProB1 ORF B2 Tag
Amplification and purification
B
PCR
B1Pro CDS B2
Gateway LR
R1ccdBCmRR2 Tag
AmpR unc pUC
B1Pro CDSB2 Tag
C
YAC Clone
YHR
AmpR pUC
ARS Fori CmR His3
Counterselection to remove the YAC
Pro CDS GFURAFP3'UTR
Removal of untaged plasmid by mating to kar1
YHR and counterselection
Pro CDS GFP 3'UTR
Purification and transformation in E. coli
D
Fosmid
CmR
Fori
Fosmid purification
Transformation in recombineering competent E. coli strain
Red/ET
rpsL TetR
Pro CDS rpsL TetR 3'UTR
Red/ET and counterselection
Tag
Pro CDS Tag 3'UTR
Transformation in E. coli strain for amplification
E
BAC or Fosmid
Transformation of pRed/Flp
Tag F KmR F
Pro CDS Tag F KmR F 3'UTR
Flp/FRT
Pro CDS Tag F 3'UTR
UNC BSD
BSD UNC p15A
UNC BSD CmR Fori
or

2.3.5.1 Cloning efficiency

Direct comparison of cloning efficiencies is difficult, but we can consider the factors that affect overall performance for each of the methods. This includes not only the absolute efficiency in terms of ratio of intended versus unintended recombination, but also the time necessary for tagging of a single gene and the ease of automation of each step. So far Gateway cloning methods have been the preferred for high throughput applications, because they are simple and easy to automate. The *in vitro* cloning steps can be performed in 96 well plates. Transformation in a 96 well format is also straightforward. Plate selection and clone size verification is required for each step, but this can also be automated with specialized plating and picking robotics. Not taking into account the initial steps required for generation of the ORFeome and the Promoterome libraries 9 experimental steps are required to produce a clone ready for transformation in C. *elegans* (Figure 28A). Using automation and parallel processing these steps can be accomplished in less than a week. The Gateway cloning approach however suffers from several inherent problems that currently have no solution. First, it is dependent on PCR amplification of the promoter and the coding sequence from genomic DNA or the ORF from cDNA (Figure 28 A,B and Table 3). These steps are prone to PCR introduced mutations and require time and cost consuming sequence verification. This problem increases with the size of the genes. Even with the relatively compact C. *elegans* genome a fraction of the genes will simply be too big for PCR amplification. Furthermore Gateway cloning efficiency decreases sharply with increased size and the success rate drops down from 84% for 1kbp to 59% with 2 kbp fragments[62].

Benefiting from the very high fidelity of the host repair machinery the *in vivo* recombination methods are essentially error free and do not require additional costs for sequence verification. Furthermore, these methods have no practical size limit and are applicable to very large genes and genes part of an operon. However, the *in vivo* recombineering strategies published so far have been relatively slow and inefficient compared to the Gateway approach. The yeast recombineering method of Sassi et al. (Figure 28 C), requires up to 6 weeks for

the cloning of a single gene and involves multiple plating and screening steps. In addition the reported efficiency for the yeast recombineering steps is very low and up to 90% of the tested clones in some cases were false positives. Additional steps were also required to remove the unmodified products at each step. The *E. coli* based method of Dolphin and Hope (Figure 28D) is limited to tagging, but still requires 8 days for a single gene, due to the need for repeated isolation and retransformation of the fosmid into specialized host strains. Both methods rely on selection-counterselection strategies for the tag insertion and are therefore prone to all the problems inherent in this approach.

In contrast, the method that we present here is both fast and highly efficient. By performing all the steps in the original host strain, we cut down the number of required experimental steps and avoid the risk of rearrangements due to retransformation of the clone. The improvements in overall efficiency that we introduced allowed us to use liquid culture cloning. As result, our method is faster than all of the previously published methods including the *in vitro* Gateway cloning. At the same time, it retains all the advantages of an *in vivo* cloning approach. The liquid culture protocol can be easily scaled up for high throughput application. All steps require only simple pipetting can be fully automated.

2.3.5.2 Transgene quality

The transgene quality can be defined as the likelihood of reproducing the endogenous level and pattern of gene expression. In this respect, the methods based on *in vivo* recombineering of genomic clones have significant advantages. As previously discussed, tagging in the natural genomic context is much more likely to result in correct expression pattern.

The Gateway based cloning approaches result in artificial constructs lacking important *cis* acting control elements. Both Gateway based methods (Figure 28 A,B;), produce a transgene lacking the endogenous 3'UTR. The transport, stability, and the rate of translation are all subject to regulation through proteins and noncoding RNAs binding to the 3'UTR. Furthermore, the reduction in cloning efficiencies with longer PCR products has led to a 2kbp size restriction of the PCR

amplified promoter regions in the Promoterome library[62]. For many genes however, this length may not be sufficient to cover all upstream regulatory elements.

Finally, even the best transgene, when delivered by an inadequate transformation method, can result in incorrect expression. As previously discussed microinjection based transformation results in the formation of transgene arrays, which can contain thousands of copies of the gene of interest and inevitably result in overexpression. Of all previously published methods, only that of Polanowska et al. (Figure 28B) uses a vector containing the *unc-119* marker and integrative transformation by ballistic bombardment. All other methods rely on coinjection with the transformation marker *rol-6*. The method we present here is the first *in vivo* recombineering method that results in a transgene carrying the *unc-119* marker for integrative transformation.

2.3.5.3 Cost

Cost efficiency is an important factor when high throughput applications are considered. Methods based on Gateway recombination have high initial costs for generation of the ORF and promoter clones. These involve not only the cost of PCR amplification and cloning, but importantly the high cost of sequence verification. However, these projects are well underway and an ORF clone is now available for about 70% of the C. *elegans* genes. Cloning of the promoterome is in an earlier phase (coverage about 30%) but will probably have a higher success rate, as amplification from genomic DNA is much simpler than from a cDNA library. The availability of these resources does not completely remove the expense for generation of the clones as the current ORFeome collection costs about € 11000 and the promoterome set cost another € 7000. In addition, each Gateway cloning step cost about € 10.

In contrast, the whole fosmid library set covering about 85% of the genome costs only € 5000. Fosmid based tagging using the generic *unc-119* marker insertion protocol will cost only about € 25 per gene (the price of two 70 nt oligonucleotides for recombineering). This makes the *in vivo* recombineering

pipeline costs comparable with the Multisite Gateway reaction approach. By inserting the tag in the natural genomic context, most splice isoforms are going to be expressed as fusion proteins (except for those genes with alternative last exon). If all the splice isoforms are to be covered by an ORF clone the total number of ORFs will be several fold larger than the number of protein coding genes. Based on a conservative estimate of about 3 isoforms per gene up to 60000 individual ORFs may need to be cloned for the two methods to be fairly comparable.

2.3.6 Summary and future prospects

In conclusion we have shown that recombineering of genomic clones into transgenes is simple and efficient enough for routine tagging of proteins in C. *elegans*. We have shown that such transgenes can be used for integrative transformation. We have reproduce known and document new expression patterns, and we showed that an example transgene can take over the function of its endogenous counterpart.

The method that we have developed compares favorably with all currently available alternatives both in terms of efficiency and transgene quality. For the first time we apply liquid culture cloning for multiple Red/ET recombineering steps. This is particularly important when high throughput applications are considered, as it offers significant advantages in scale up and automation. The liquid culture recombineering pipeline we have developed here is directed towards making GFP tagged transgenes for ballistic transformation in C. *elegans* from an indexed BAC library. Applications other than protein tagging are clearly possible, such as transgenes carrying site directed mutations or deletions aimed at mapping of *cis* regulatory elements. Different applications will require alterations to the pipeline strategy and different expression plasmids. For example Red/ET recombination can be coupled with site specific recombinases other than Flp, such as Cre or Dre[101] or rare cutting homing endonucleases[102] such as I-Sce I, I-Ceu I or PI-Sce I for inducible *in vivo* linearization[103,104]. Similar pipelines can be developed for any other model system that permits transgenesis and have a mapped genomic clone resource.

The future developments of our method include evaluation of the transformation efficiency with fosmid clones. Increasing the efficiency of transformation is an important next step towards high throughput application of the method. Even the integrated transgenes can suffer from artifacts such as position effects. A future improvement of the method would be the development of Recombination Mediated Cassette Exchange (RMCE)[105] approach for transgene integration similar to that applied in culture cells[106,107] and Drosophila[108,109]. Another line of development is related to attempts in improving the rate of homologous recombination in C. *elegans* using transposon excision generated double stranded breaks[110,111]. As more and more transposon integration mutants are being mapped, this approach may become feasible for large scale application. A liquid culture recombineering pipeline for generation of targeting constructs can easily be designed. Using this approach the size of the homology arms used for targeting can be increased significantly, which may lead to improved targeting efficiency.

3 New double affinity tags for proteomic exploration

3.1 Introduction

The purpose of this project was to develop new double affinity tag combinations with improved performance in comparison with the original TAP cassette. The tandem affinity approach (Figure 29) was developed initially for use in *S. cerevisiae*, and its direct application in higher model systems has been challenging. Since then several other double affinity cassettes have been described for use in various model systems[86,112-120].

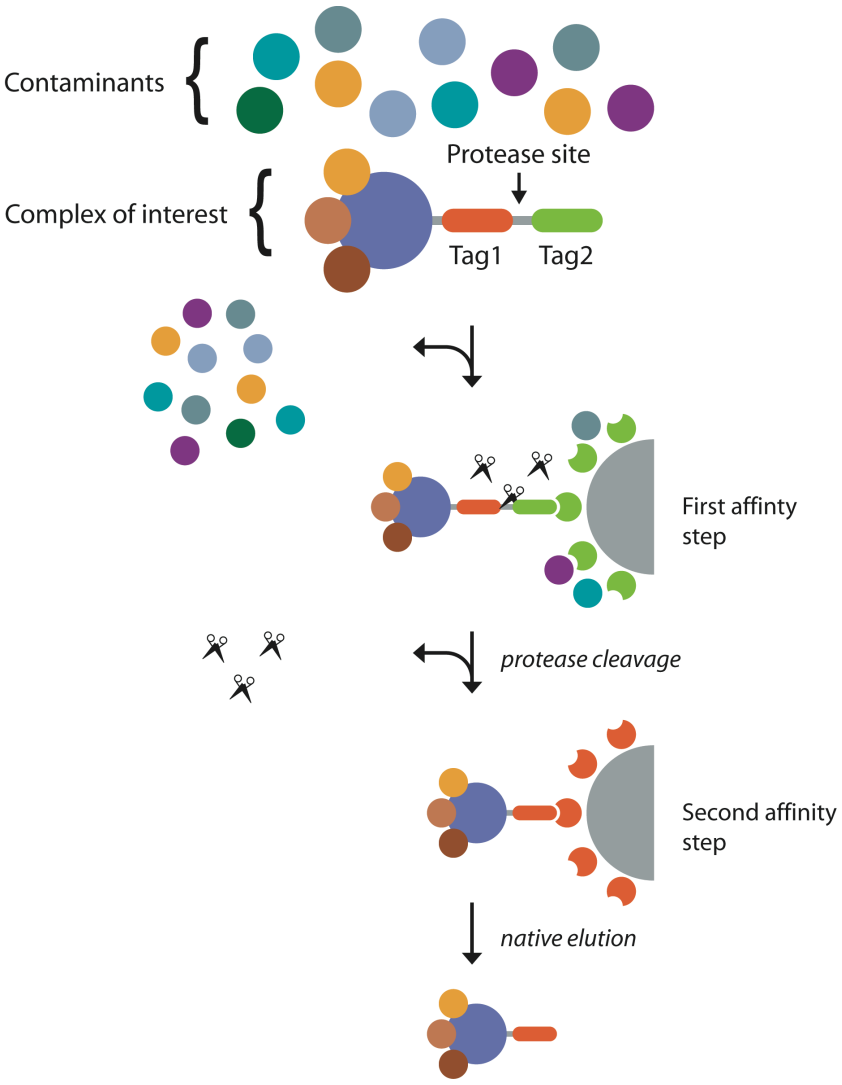


Figure 29. The tandem affinity method (modified from reference 1).

The original TAP-tag cassette consists of two copies of a modified IgG binding domain of protein A (Z domains), followed by a tobacco etch virus (TEV) protease[121] cleavage site and a calmodulin binding peptide (CalBD). On the first step of the TAP purification the Z domain is bound to immobilized IgG. This interaction is very strong and protein elution under native conditions is only possible through TEV protease cleavage. This further increases the purity of the preparation, as only the tagged protein is specifically eluted. The eluate is then bound to calmodulin in the presence of Ca^{+2}. The interaction of CalBD with calmodulin is very strong (Kd in the nanomolar range) but can be broken by removal of the Ca^{+2} with chelators such as EGTA. This allows the elution of the retrieved protein complex under mild conditions. However, EGTA elution can affect the function of some metalloproteins. Furthermore, for some proteins EGTA elution can be problematic and may require conditions of high salt and/or detergents.

The TAP tag was developed for use in S. cerevisiae, and one problem that we expected in mammalian cells was the unwanted interaction of endogenous IgG and calmodulin with the tag. To address this potential problem we replaced the ZZ tag and the CalBD with other tags with very high reported affinity, at least in the nanomolar range (Table 4).

Table 4. Tested affinity tags

Tag	Abbreviation	Affinity ligand	Reported affinity	Peptide	Elution
2 IgG binindg Z domains	ZZ	IgG heavy chain	nM	12,5 kDa	Protease cleavage
Calmodulin binding domain	CalBD	Calmodulin	nM	2,5 kDa	EGTA
Chitin binding domain	ChiBD	Chitin	irreversible	5,8 kDa	Protease cleavage
Streptavidin binding domain	Sbp	Streptavidn	nM	4,3 kDa	biotin
S-peptide	S	S-protein	nM	1,7 kDa	Protease cleavage
SNAP tag	Snap	O6-Alkylguanin	covalent	19,3 kDa	Protease cleavage

As a first tag we tested the Chitin binding domain (ChiBD), the S peptide and the SNAP tag. The first two tags are short peptides attractive due to their small size and high affinity. The SNAP tag on the other hand is a small protein, a variant of the O6-Alkylguanin-DNA Alkyltransferase enzyme, which was engineered to bind to free alkylated guanine. This reaction can tolerate addition of different adducts

to the alkylated substrate, which permits the specific covalent attachment of fluorescent of affinity tags or immobilization on inert surfaces. An attractive advantage of the SNAP tag is that it can be used both for purification and for localization.

For the second affinity step we needed a tag, which like CalBD can be eluted at mild conditions. The streptavidin binding peptide[122] (Sbp) proved to be an excellent replacement for CalBD as a second tag. This peptide was developed by guided *in vitro* evolution from a random peptide library[123,124] to bind streptavidin with very high affinity (Kd 10^{-9}). At the same time, SBP can be specifically eluted at very mild conditions (2mM biotin), allowing recovery of active proteins at native conditions.

Finally, in all of the new cassettes we replaced the TEV cleavage site with the cleavage site for PreScission protease. PreScission is an engineered version of the human rhinovirus 3C protease[125]. Unlike TEV, PreScission is highly active at low temperatures, allowing the entire TAP purification to be carried out at 4°C, to prevent non-specific proteolysis. In addition it has an increased specificity through a longer recognition site[126].

3.2 Results

3.2.1 Test of the new tandem affinity purification cassettes *in vitro*

3.2.1.1 Cloning and purification of recombinant TAP cassettes

We generated a series of new double affinity cassettes by either conventional or recombinational cloning as described in materials and methods. Flexible linkers were introduced between the affinity domains and the protease cleavage sites. All cassettes were cloned in pGEX vectors in fusion with the GST coding sequence (Figure 30A).

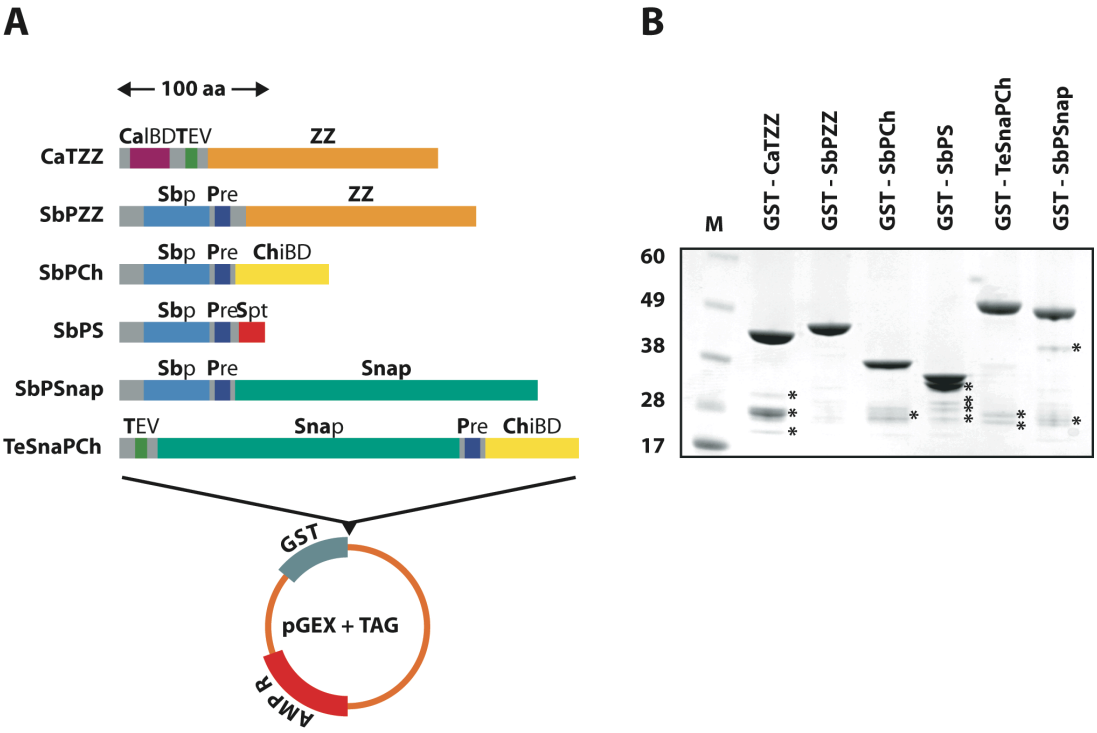


Figure 30. Cloning and purification of recombinant TAP cassettes. (A) In scale depiction of the new cassettes, cloned as GST fusion protein. (B) Purification of the recombinant GST-TAP proteins; stars depict degradation products; The names are derived from the first letters of the cassette elements (given in bold above the cartoon).

Recombinant GST-TAP fusion proteins were expressed and purified from bacteria (Figure 30B). Some of the constructs required further optimization in the flexible linkers to overcome problems with proteolysis during expression in *E. coli*,

probably as a result of misfolding. Some truncated products were always present, most notably in the case of SbPS. However the proteins were not further degraded upon incubation with mammalian extracts in the absence of protease inhibitors (data not shown), indicating that this was a problem related to overexpression in *E. coli*, rather than to the inherent proteolytic stability of the cassettes.

3.2.1.2 Test for background of endogenous proteins binding the tag.

One potential problem for all affinity tags is the interference of endogenous proteins interacting with the tag and preventing it from binding to the affinity matrix. To address this question we did GST pull downs from mammalian tissue extract (Figure 31). No protein, GST only or the recombinant GST-TAP proteins were bound to GST affinity beads and were incubated with extracts from either brain or kidney. In all of the pull downs two bands of the size of the GST were purified (Figure 31B, bands 3 and 4). They were identified by mass spectrometry as the endogenous GST (data not shown).

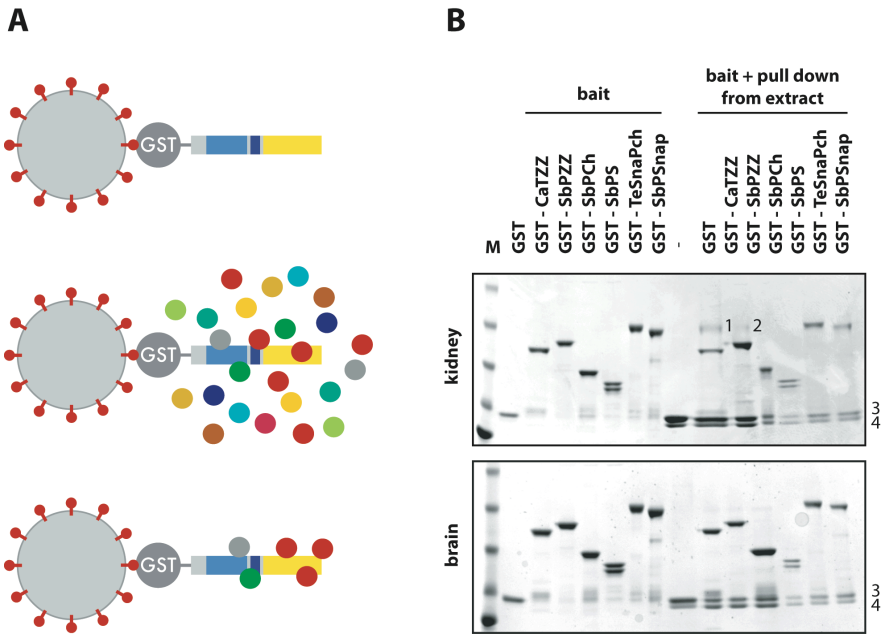


Figure 31. Background from proteins binding to the tag. Bands identified by mass spectrometry: 1,2 IgG heavy chain; 3,4 GST.

An additional band of approximately 50 kDa was pulled down from kidney extract with the two tags containing the Z domain from protein A. They were identified by mass spectrometry as mouse IgG heavy chain. These bands were weaker or completely absent in the brain extract, reflecting the lower amount of blood present in the brain. No other major contaminants were identified. The pull downs were performed in the presence of 2 mM EDTA, which explains why endogenous calmodulin was not bound to the Calmodulin Binding Domain of the CaTZZ tag.

3.2.1.3 Purification of the recombinant GST-TAPs from mammalian extracts.

To evaluate the performance of the tags we did a simulated TAP pull down experiment by spiking a known amount of recombinant GST-TAP into 5 mg/ml mouse brain extract (Figure 32).

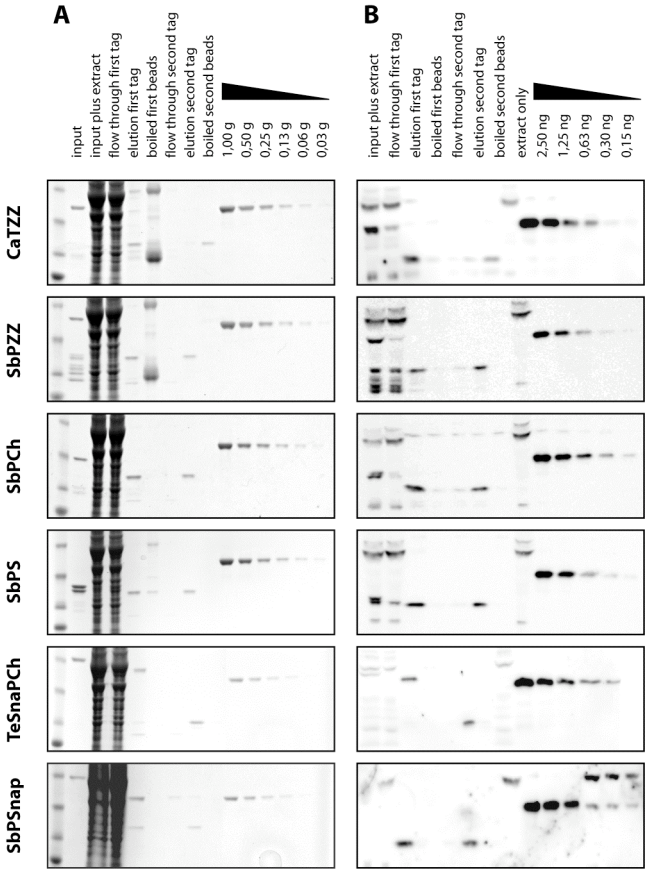


Figure 32. Tandem Affinity Purification of the recombinant GST-TAP cassettes from mouse brain extract. A Coomasie blue stained gels; 20 µg/ml recombinant GST-TAP protein. B Western blots; 100 ng/ml recombinant GST-TAP.

Two different concentrations of recombinant protein were purified: 20 µg/ml or 1:250 dilution to the total protein in the extract (corresponding to relatively abundant protein) and 100 ng/ml or 1:50000 dilution (corresponding to a scarce protein). Samples were taken at each step of the purification and were analyzed by either Coomassie blue staining or Western blot. The same volume was kept at all steps to allow quantitative comparison. The amount of purified tag was determined by comparison with a serial dilution of a control GST tagged protein with known concentration (Figure 33).

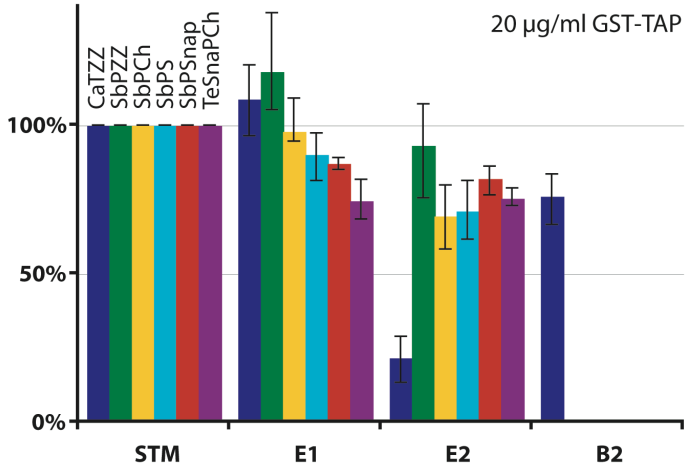


Figure 33. Quantification of the GST-TAP purifications. STM - starting material, E1 elution from the first beads, E2 elution from the second beads, B2 boiled second beads.

A problem with the original TAP cassette (CaTZZ) was revealed. Elution from the calmodulin beads was problematic, and most of the bait was retained on the beads. In contrast, all new tags performed well, and the differences in bait recovery were within the margin of error. At the low bait concentration, the results were essentially the same, although exact quantification by western blot was not possible.

3.2.1.4 Effect of formaldehyde crosslinking on the tag performance

Next, we examined the effect of formaldehyde crosslinking on the tags performance. Crosslinking can fix weak or transient interactions, and is required

for applications such as chromatin immunoprecipitation (ChIP). We repeated the experiments from Figure 32A after an initial 5 minute fixation with 1% formaldehyde (Figure 34). Crosslinking did not significantly decrease the performance of the CaTZZ, SbPZZ, SbPCh of SbPS tags. However binding of the Snap tag in both SbPSnap and TeSnaPCh was severely affected. The same result was observed when the crosslinking was performed in buffer, which indicates that the problem was not related to crosslinking with proteins present in the extract.

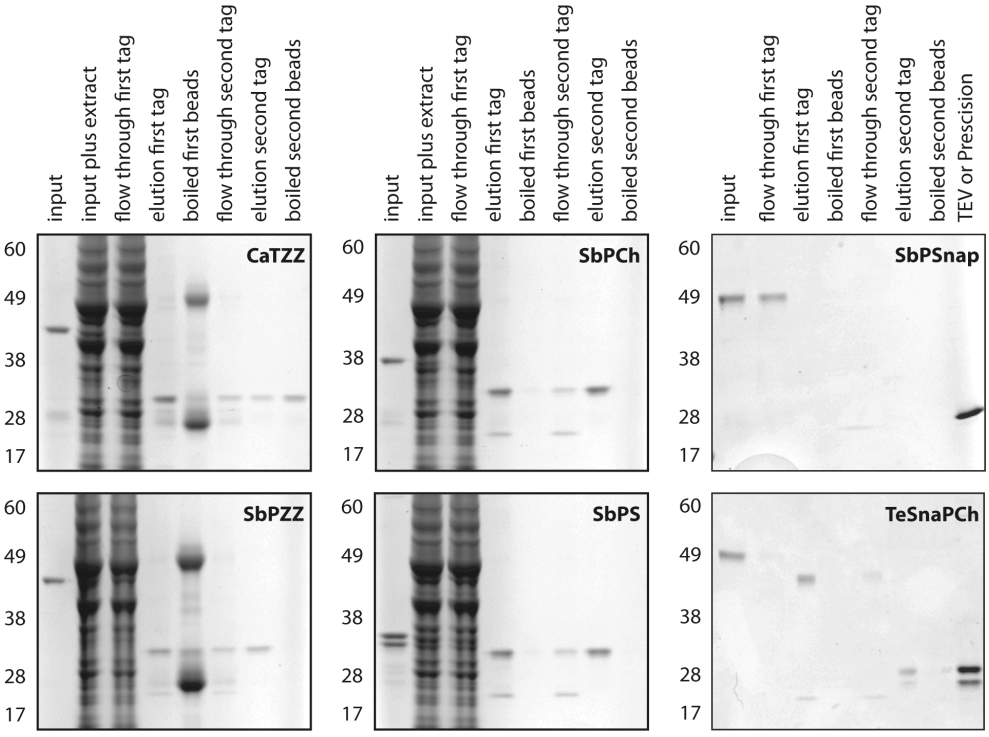


Figure 34. Effect of formaldehyde crosslinking on the tags performance. SbPSnap and TeSnaPCh were purified in the absence of extract (see the text for details).

3.2.2 Protein tagging in tissue culture cells

3.2.2.1 Generation of transgenic cell lines

Based on the *in vitro* tests we selected the SbPCh and the SbPSNAP for further protein localization and purification tests in cultured mammalian cells. For generation of transgenic cell lines we used a BAC transgene based approach [87]. A selection cassette consisting of the SV40 virus internal ribosome entry site, the minimal *E. coli* promoter EM7 and the neomycin gene for selection in both *E. coli*

and mammalian cells (IRES-EM7-*neo*), was inserted by Red/ET recombination just after the stop codon of the tags in the pGEX vector.

We tagged two genes of interest: the Aurora B kinase and the heterochromatin protein1 beta. Mouse genomic BAC containing the gene of interest were selected and the TAG-IRES-EM7-neo cassette was inserted in front of the stop codon of the gene by Red/ET recombination. The correct insertion was confirmed by PCR. The tagged BACs was transfected into HeLa S3 cells and stably integrated transgenic cell lines were established as described in Materials and Methods. Expression of the tagged protein was confirmed by RT-PCR.

3.2.2.2 Protein localization with the SbPSNAP tag

To evaluate the performance of SbPSNAP for protein localization studies we fused it to the mouse Aurora B kinase. Aurora B has a characteristic dynamic subcellular localization. In the interphase nuclei Aurora B localizes to the centromeres and during cytokinesis, at the metaphase–anaphase transition it translocates from the centromeres to the central spindle. The transgenic AuroraB::SbPSNAP cells were stained *in vivo* with cell permeable benzyl-guanine-tetramethylrhodamine (TMR-Star), which binds to the SNAP tag. The known subcellular localization was correctly reproduced (Figure 35).

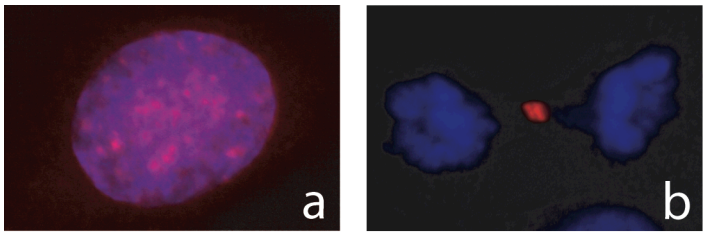


Figure 35. Aurora B SbPSNAP localization in (A) interphase and (B) cytokinesis; Red BG-TMR Star (binds to the SNAP tag); Blue DAPI stained DNA.

3.2.2.3 Protein purification with the SbPCh tag

We tested the SbPCh for purification of the heterochromatin protein 1β. In the initial tests, we used liquid culture to obtain enough cells and purified the protein as in the *in vitro* tests above. A single band of the expected size was recovered

after the second step from the HP1::SbPCh cells but not from wild type HeLa S3 cells (data not shown).

In the further improvement of the protocol, we applied a magnetic affinity media for SbPCh TAP purification. The cells from 5 confluent 10 cm plates (200 µl cell pellet) of either wild type or SbPCh tagged HeLa S3 cells were homogenized by sonication in 1ml of lysis buffer, and the purification was performed using magnetic Chitin and Streptavidin beads directly from crude extracts (Figure 36).

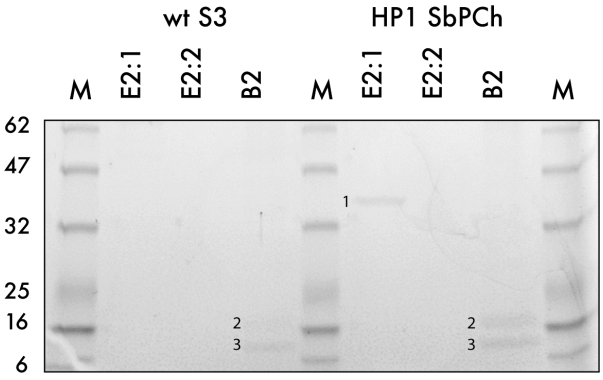


Figure 36. Rapid magnetic Tandem Affinity Purification of SbPCh tagged HP1 from crude extract; E2:1, E2:2: first and second elution from the second affinity step; B2 boiled streptavidin beads; Bands indicated: 1 HP1; 2,3 - streptavidin.

Only 30 minutes of incubation was used for binding to the beads at both the affinity binding steps and the elution steps. The whole procedure took less than 3 hours. A band of the expected molecular weight was recovered from the HP1-SbPCh cells but not in the mock IP. Just 50 µl of biotin elution buffer were required for complete recovery of the bait protein from the Streptavidin beads; no protein was detectable in a second 50 µl elution or after boiling the beads.

3.3 Discussion

By using purified recombinant GST-TAP proteins we could easily set comparable experiments to evaluate the performance of the new double affinity cassettes. Problems with the tag design leading to misfolding were immediately revealed and corrected. We could compare the tags directly in the pull downs from complex mammalian extract. These experiments revealed two potential problems of the original TAP cassette. First as expected endogenous IgG present in the extract can bind and block the Z domain. Blood and hence IgGs are present in all tissues and in some cases, depending on the expression level of the target protein, can be sufficient to completely block the tag. Unlike the interaction of CBD with endogenous calmodulin, which can be prevented by adding chelating agents to the extraction buffer, the IgG binding to the Z domain is essentially irreversible and cannot be specifically inhibited. The second problem was the issue with poor elution from the calmodulin beads[127,128]. The interaction of the CBD with calmodulin shows a cooperative effect when the bait protein is dimeric, as it is in the case of GST, and can result in resistance to elution by EGTA.

In contrast, all of the new tags did not suffer from such issues and performed well under native conditions. In cross-linking conditions however, the cassettes containing the SNAP tag performed poorly. Unlike the other tags, which are all relatively small affinity peptides, the SNAP tag has to be enzymatically active to bind to the matrix, and hence is more sensitive to crosslinking.

In conclusion, we have generated a set of double affinity cassettes, which are comparable in performance with the original TAP tag and avoid some of its limitations such as the background interactions with endogenous proteins, and the poor elution from the calmodulin beads. The choice of tag still depends on the model system, for example SbPZZ is not optimal for retrieval from protein extracts rich in IgG and SbPCh is not suitable for use in model organisms that produce chitin such as C. *elegans* or *Drosophila*.

Using BAC transgenesis we further tested the SbPCh tag in tissue culture cells. We developed a rapid two-step magnetic purification procedure, which has several

advantages. First, as it employs magnetic phase separation it can be applied directly to the crude extract, without prior clarification by centrifugation and the related lost of material. Second, magnetic beads have low non-specific binding, resulting in very clean preparations. Third, complete elution can be achieved in very small volume, so that the eluate can be directly analyzed without the need for further concentration. Finally, the magnetic IP is much faster and more suitable for automation[129].

4 Materials and methods

4.1 Reagents

4.1.1 Plasmids

A brief description of the used plasmids is given below. Detailed maps and the complete sequences are available upon request.

Name	Description	Source
pGEX 6P1	GST expression vector	GE Healthcare
pGEX CaTZZ	GST-TAP expression plasmid	This study
pGEX SbPZZ	GST-TAP expression plasmid	This study
pGEX SbPCh	GST-TAP expression plasmid	This study
pGEX SbPS	GST-TAP expression plasmid	This study
pGEX SbPSNAP	GST-TAP expression plasmid	This study
pSC101BADgbaA	Red/ET expression plasmid	Ref. [25]
pSC101CIFlpe	Flpe expression plasmid	Ref. [73]
pRedFlp1	Red/ET and Flp expression plasmid with pSC101 ori	This study
pRedFlp2	Red/ET and Flp expression plasmid with pSC101 ori	This study
pRedFlp3	Red/ET and Flp expression plasmid with pSC101 ori	This study
pRedFlp4	Red/ET and Flp expression plasmid with pSC101 ori	This study
pRedFlp5	Red/ET and Flp expression plasmid with pSC101 ori	This study
pGEX6P1-EGFP-FRT-KmR-FRT	EGFP tagging cassette template	This study
pGEX6P1-EYFP-FRT-KmR-FRT	EYFP tagging cassette template	This study
pGEX6P1-CaTZZ-FRT-KmR-FRT	CaTZZ tagging cassette template	This study
pR6KGFP	zero background EGFP template plasmid	This study
pUH	ce *unc-119* promoter:cDNA based subclonig plasmid	This study
pUB1	ce *unc-119* promoter:cDNA based subclonig plasmid	This study
pUB2	cb *unc-119* based subclonig plasmid	This study
L4440	ds RNA empty expression vector	Ref. [130]
L4440 *lin-59*	*lin-59* ds RNA expression vector	GeneService

4.1.2 *E. coli* strains

Name	Description	Source
DH10B	BAC host strain	Invitrogen
HS996	DH10B fhuA::IS2; phage T1-resistant	Invitrogen
EPI300	HS996 trfA	Epicentre
YZ2005	YZ2000 endA, rpsL	Stewart lab
YZ2005 pir	YZ2000 plus the pir gene under the lac promoter	Stewart lab
HT115(DE3)	RNAse III deficient strain for dsRNA production	Ref. [88]

4.1.3 C. elegans strains

Name	Description	Source
N2	wild type C. *elegans*	Ref. [131]
unc-119(ED3)	severe unc phenotype	Ref. [69]
rrf-3	increased sensitivity to RNAi	Ref. [132]
unc-119[cb*tbg-1::EGFP*; *unc-119*(+)]	*tbg-1::EGFP* transgenic strain	This study
unc-119[cb*nhr-23::EGFP*; *unc-119*(+)]	*nhr-23::EGFP* transgenic strain	This study
unc-119[cb*nhr-25::EGFP*; *unc-119*(+)]	*nhr-25::EGFP* transgenic strain	This study
unc-119[cb*F09G2.4::EGFP*; *unc-119*(+)]	*cpsf2::EGFP* transgenic strain	This study
unc119[cb*lin-59::EGFP*; *unc-119*(+)]	*lin-59::EGFP* transgenic strain	This study
rrf-3; *unc-119*[cb*lin-59::EGFP*; *unc-119*(+)]	*lin-59::EGFP* transgenic strain crossed to *rrf-3*	This study

4.1.4 Primers

Given below are the primers for amplification of the tagging and the subcloning cassettes used in the experiments in Chapter 2.3.2.

Tagging primers:

F par2 tag	CAACGGCTCGTATTCCTACTGCTATTCCTACTTCTATTCCTACCACATCAAGCTCAGGAGGTAGCGGCATGGTGAGCAAGGGCGAG
R par2 tag	TATAAAAATTATATATGGGGGATTCTGAATATCCTGAATACAGTAGTCCTTTAGGCAGATCGTCAGTCAG
F set2 tag	ACAAAATCGACTGTCTCTGTGGTGCCAAATCGTGTCGAGGATACCTCAATAGCTCAGGAGGTAGCGGCATGGTGAGCAAGGGCGAG
R set2 tag	TGTATAAAAAACTAGAAGAGATCAAAAAGGCACAGAGTACAATGTACCAACTAGGCAGATCGTCAGTCAG
F zyg9 tag	TGAGCCGAGAACAACACGAGGAGCTCAGGAACCGTCTTCAACAAGCAAAGAGCTCAGGAGGTAGCGGCATGGTGAGCAAGGGCGAG
R zyg9 tag	GCAGAACATTTGAAAACGAGCGAGCGAGGAAATGGGATTTATTGTCCGAAGGCAGATCGTCAGTCAG
F mes2 tag	CTTCGAAATTTGGTCACTATGATCCGGATAACGATAATCATTTGTATTTTAGCTCAGGAGGTAGCGGCATGGTGAGCAAGGGCGAG
R mes2 tag	AATTTAAAAAATGTAACTGAAATTTAGAGGAGATAGACAAAAAAGAAGAATCAGGCAGATCGTCAGTCAG
F lin59 tag	AAATCGAAAAAGTTCTGAACCGCCTTGTTTCTTTAAATTCTAAAAAAACTAGCTCAGGAGGTAGCGGCATGGTGAGCAAGGGCGAG
R lin59 tag	ATAGAAACGAGATGAGATGAGGGGATTGTTTTGAAGTGATATACAAACTAGGCAGATCGTCAGTCAG
F tag350 tag	ATAAGGTCCCATGTCTTTGCGGAGCACCGAATTGTGTTAAATGGATGAATAGCTCAGGAGGTAGCGGCATGGTGAGCAAGGGCGAG
R tag350 tag	CACACACACAATCGTACAAGAAGTTGTGAAAACAAAACCTAAACAAAACATTAGGCAGATCGTCAGTCAG
F F09G2.4 tag	AAGACTTTTACAAACTTCGGAAGTTGTTCTACGACCAGTTTGCTGTTTTGAGCTCAGGAGGTAGCGGCATGGTGAGCAAGGGCGAG
R F09G2.4 tag	TTTACTCGAAAAGTGTAAAGAAATCTGTTAAAATTATATGAGATCATCATGGCAGATCGTCAGTCAG
F tbg1 tag	TTGACGAGTACAAAGCAGTGGTTCAAAAGGATTACCTTACGAGAGGACTAAGCTCAGGAGGTAGCGGCATGGTGAGCAAGGGCGAG
R tbg1 tag	AATAGAGATCCTGCCAAAAGAGAATAAGAATCCCGCCAAGAATAAGTGAAGGCAGATCGTCAGTCAG
F nhr23 tag	CGGACCTGTACAAAGAGCTATTCACAGCGGATCGACTTTCACCAACGATAAGCTCAGGAGGTAGCGGCATGGTGAGCAAGGGCGAG
R nhr23 tag	GAATGGGAAGAAGGAGAAGGAAATTTGGTTAAAACTATTTTTGGGGACGAGGCAGATCGTCAGTCAG
F nhr25 tag	CAGTACCTCTGCAACAAACCACATACGCTCCTGTAGTATATATGACATCTAGCTCAGGAGGTAGCGGCATGGTGAGCAAGGGCGAG
R nhr25 tag	CTCAAACTAAAGGTATTTCATAAATGATTTGTTGATTTTTTTCAATCCGGGGCAGATCGTCAGTCAG
F spd2 tag	AAACTTTTGTAAATGACGTCACGATGGTTCCAAGACATACGAAACTAATGAGCTCAGGAGGTAGCGGCATGGTGAGCAAGGGCGAG
R spd2 tag	TTTAAACAAGAATTAGGCTTTCCAGTGGAATCAATGAAAATGTTGGTTTGCTAGGCAGATCGTCAGTCAG
F Y17G7B.2 tag	TTCTGAATTACCCTGCTGTCAAGAAAGAAATCAAGGAAGAGGACATGGAGAGCTCAGGAGGTAGCGGCATGGTGAGCAAGGGCGAG
R Y17G7B.2 tag	AGGCAAAGAGCAAAACAATGAGATAAAGTGACAGCGGAGATTCGGTCTAATTAGGCAGATCGTCAGTCAG

Subcloning primers:

F par2 sub	TTGAATCCTATGATCTAAAGCAATAAAAAATTCCAGAATATGCAAAAACCCTAGGGATAACAGGGTAATTG
R par2 sub	TTTCGTTTTTATCGAATTTTCGCATTTTTTTAAAGTAAAAAATTTTCCAGTTAGCCCTCCCACACATAACC
F set2 sub	ATGGAGCCAAAGGGCGAAACAAAAAAAAAACTGCTTCCTTTCCTTTTTCAGCTAGGGATAACAGGGTAATTG
R set2 sub	TTAGCTGTTCGAATGGCTGTTTATTAGAAATTGAAAAAAAAAAATTATCGTTTAGCCCTCCCACACATAACC
F zyg9 sub	TTATTATTTATATAATTTATTTCGGTGGCATAAAAAGTATGATTGTATTCCTAGGGATAACAGGGTAATTG
R zyg9 sub	TAAAATTTTGAAATTCCAACATTTATTACTTCATTTTCGCTTCCACAAAATTAGCCCTCCCACACATAACC
F mes2 sub	CGCAAAAATTTCACTTTCGGAAGTTATCAATCGCTTCAAAATGCCATAAACTAGGGATAACAGGGTAATTG
R mes2 sub	ACCGGCTCTAAATTTAAAAATATATAATTAAAAAAGGAAAATAGAAATATTTAGCCCTCCCACACATAACC
F lin59 sub	AATGCAAAATTACCCCAGCTAGTGAAACCCATTAGACCAAGATTCGTCTAGGGATAACAGGGTAATTG
R lin59 sub	GATTTCAACACGAATACTGTGAATAACTTGAAAGAAAACGAAGAACTTTAGCCCTCCCACACATAACC
F tag350 sub	AACATGTATTAAATTTGAAACCTGAGAAACAGAATATATGTAAGGTAAAACTAGGGATAACAGGGTAATTG
R tag350 sub	GTGAATAAAATTTCAGAAAAGTCTCACTTGGGGAAGTTTTTATACAAAGGTTAGCCCTCCCACACATAACC
F F09G2.4 sub	TAAAACGAGAGAACGGTGGTCATTTGTCTCCACAGTCACGACGGAGAAACCTAGGGATAACAGGGTAATTG
R F09G2.4 sub	AACCAAAGGAATTATTGCTTTTTGATACACATAACACAATTATCCAAGTGTTAGCCCTCCCACACATAACC
F tbg1 sub	AGAAATATTGAGTAATATATAGAAATCTAAAGAAGACGATGAAGAACGGACTAGGGATAACAGGGTAATTG
R tbg1 sub	TCGAAATCAAAAAAAAAAAAAACTTACACTTCCTTCTTGATTTTGATCTCTTTAGCCCTCCCACACATAACC
F nhr23 sub	AGTACATTTTGTTTCTATAAATTCGAAGAATATATTCCAGTTATTCGAAACTAGGGATAACAGGGTAATTG
R nhr23 sub	TTTGGCTCAAAGTCAACTGTTGAAAAAACTGAATTTGAAAGAGAAGTTATTTAGCCCTCCCACACATAACC
F nhr25 sub	CGTTGTACGAGAAGCACCGGGAAATGAATAGAACAATCATTAAAGAATAATCTAGGGATAACAGGGTAATTG
R nhr25 sub	TATACATGAACAGAAATCGTATAAAATTCCAATAAAATTTTCACAACAAATTAGCCCTCCCACACATAACC
F spd2 sub	CGCTTTAATTTGTATTTGAATAAGTTATAGCTAAATAGTACAAAATGAGACTAGGGATAACAGGGTAATTG
R spd2 sub	CTTATGTGTTTTTTGAAAATTTCGAAACTTGTGGAATGACAGGCCCGCCGTTAGCCCTCCCACACATAACC
F Y17G7B.2 sub	CGCTTCCGCCGCTCCAGAATAATAAAATACTATTTTTTTAAAATTTACAGCTAGGGATAACAGGGTAATTG
R Y17G7B.2 sub	ACTTTTGTGTTGCTTTCTCGAAACCCAACCCCCTTTTCAAAGAAATCAGATTAGCCCTCCCACACATAACC

4.1.5 Antibodies

Antigen	Raised in	Type	Working Dilution	Source
GST	Goat	Polyclonal	1:10000	GE Healthcare
Red Beta	Rabbit	Polyclonal	1:5000	Ref. [25]
Flpe	Rabbit	Polyclonal	1:5000	Ref. [77]

4.2 General methods

4.2.1 DNA methods

4.2.1.1 Plasmid, Fosmid or BAC DNA purification for analytical purposes

A modified alkaline lysis protocol was used for quick purification of DNA from *E. coli* for restriction analysis:

1. Pick a single colony and inoculate in 2ml of LB + chloramphenicol (15µg/ml) + kanamycin (15 µg/ml). Incubate overnight (but for no more than16 h!) at 37°C with vigorous shaking.
2. Spin down the cultures for 1 min at 11,000 rpm.
3. Discard the supernatant and resuspend the cell pellet in 200 µl buffer P1 with RNAse.
4. Add 200 µl of buffer P2 and mix by inverting the tube several times.
5. Add 200 µl of buffer P3 and mix by inverting the tube several times.
6. Spin down the white precipitate at highest speed for 15 min.
7. Transfer the clear supernatant into a new 1.5ml Eppendorf tube and add 0.50 ml of 2-propanol.
8. Mix by inverting the tube and spin down the DNA at highest speed for 10 min.
9. Discard the supernatant and add 1 ml of 70% ethanol to rinse the pellet.
10. Clean the inner wall of the tube with a piece of tissue or cotton stick.
11. Dry the pellet at room temperature for 5 to 10 min.
12. Add 10-20µl of 5mM Tris-HCl pH=8 and let the DNA dissolve for 30 min at room temperature.

4.2.1.2 Purification of BAC or fosmid DNA for transformation.

Highly pure large molecular weight DNA for trnasfection in HeLa cells or transformation in C. *elegans* was obtained using the Nucleobond BAC Maxi Kit (Clontech K3008-1), following the manufacturer's protocol.

4.2.1.3 Purification of total genomic DNA from HeLa Cells

Proteinase K lysis buffer: 50 mM Tris-HCl pH 8,0,100 mM EDTA,100 mM NaCl, 0.5 % SDS, 0.5 mg/ml Proteinase K.

1. Add 500 µl lysis buffer to the cell pellet from one 10cm and incubate at 55°C for 2h.
2. Add 1 volume of phenol : chloroform : Isoamyl alcohol (25 : 24 : 1) and incubate on a rotating wheel for 1 h at RT.
3. Centrifuge for 10 min at 12000 g.
4. Using a cut tip transfer the upper aqueous phase to a new tube.
5. Add 1 volume of chloroform : isoamyl alcohol (24:1) and incubate on a rotating wheel for further 1 h at RT.
6. Centrifuge for 10 min at 12000 g.
7. Using a cut tip transfer the upper aqueous phase to a new tube.
8. Add 25 µl 3.0 M potassium acetate, pH 5.5 and 0.7 volumes isopropanol.
9. Mix well and centrifuge for 10 min at 12000 g.
10. Remove the supernatant and wash the pellet twice in 70 % ethanol
11. Remove all the ethanol and let the pellet dry at RT for 10 min.
12. Dissolve the DNA pellet in 50 µl TE Buffer.

4.2.1.4 Purification of total genomic DNA from C. *elegans*

Purification of genomic DNA from C. *elegans* was as from cultured cells. For PCR analysis the Protease K digested DNA was used directly, without further purification.

4.2.1.5 Polymerase chain reaction (PCR)

Polymerase chain reactions were performed using the Eppendorf Mastercycler EP PCR cycler. Thin wall PCR tubes were used and the volume of the reaction was kept up to 50 µl to improve the temperature exchange. The concentration of the template, primers and dNTPs were varied depending on the application. A typical reaction is shown below:

Reagent	Volume (µl)
H_20	40
10X Buffer	5
10 mM dNTPs	1
10 mM Forward primer	1,25
10 mM Reverse primer	1,25
1 ng/µl template	2
Triplemaster polymerase	0,15
Total volume	50

The number of PCR cycles was adjusted depending to the template, so that the reaction would not proceed beyond saturation, to reduce the risk of PCR introduced mutations. Annealing temperatures were adjusted depending on the primer used and extension times were varied depending on the length of the template.

4.2.1.6 Concentration and desalting of PCR products by LiCl precipation

PCR products were precipitated with LiCl, which is very soluble in ethanol and leaves the pellet essentially salt free. This is important when the PCR product is to be used for electroporation.

1. For each 100 µl PCR reaction add 5 µl 5 M LiCl and 300 µl 100% ethanol.
2. Mix well and precipitate for 30 min at -20°C.
3. Spin down the DNA at maximal speed for 15 min.
4. Carefully wash the pellet once with 1ml 70% ethanol.
5. Dry the pellet at room temperature for 5 -10 min.
6. Resuspend in 10 µl 10mM Tris-HCl, pH 8.0 (0.2 -0.5 µg/µl).

When the PCR product was used for further enzymatic reactions such as ligation the Qiagen PCR purification kit was used instead, according to the product manual.

4.2.1.7 Agarose gel electrophoresis

10x Loading buffer: 100 mM Tris pH=8, 100mM EDTA, 25% Ficoll, 0,5% Orange G.
TBE Buffer: 0.89 M Tris, 0.89 M Boric acid, 0.02 M EDTA Na salt

DNA was separated on TBE agarose gels with 0,5 to 2% agarose depending on the size of the analyzed fragments. Ethidium bromide was added to the gels at 100ng/ml. DNA was visualized and documented with Molecular Imager ChemiDoc XRS System (Bio-Rad).

4.2.1.8 Restriction digest

All restriction enzymes were from New England Biolabs (NEB). The manufacturer's recommendations on buffer conditions and temperature were followed. The volume of enzyme per reaction was kept to less than 10% to prevent star activity. Typically, 1 unit of enzyme was used for each microgram of DNA, and incubation time was 2h to overnight.

4.2.1.9 Cloning by ligation

Ligation was performed with T4 ligase (NEB) in the provided ATP containing buffer at 16°C for 2h to overnight. Typical reaction volume was 20 µl. The molar ratio vector:insert was typically 1:3, and the total free ends concentration was kept around 1µM.

4.2.2 RNA methods

4.2.2.1 RNA extraction

RNA was extracted from HeLa cells using the Trisol reagent (Invitrogen) as described below:
1. Estimate the volume of the cell pellet. If the pellet was frozen let it thaw completely on ice before proceeding.
2. Add 10 times the pellet volume Trisol to the pellet.
3. Add 1 pellet volume Chloroform and vortex to dissolve the pellet.
4. Centrifuge for 15 min at 13000 x g at 4°C.
5. Transfer the supernatant to a new tube and add 1 volume isopropanol.
6. Incubate 15 min on ice.
7. Centrifuge for 15 min at 13000 x g at 4°C.
8. Remove the supernatant and wash the RNA pellet twice with 80 % ethanol.
9. Let the pellet dry completely at RT.
10. Dissolve in diethylpyrocarbonate (DEPC) treated water.

RNA concentration was measured at OD_{260} using a UV/VIS spectrophotometer at a dilution of 1:500 in DEPC water (at OD_{260} = 1, RNA concentration is 40 µg/ml).

4.2.2.2 cDNA synthesis and RT PCR

RNA was isolated (as described above) and reverse transcribed from wild type and tagged HeLa S3 cells in parallel. 1 µg of total RNA was treated with RQ1 RNAse-free DNAse (Promega) for 1h, the DNAse was inactivated by 20 min incubation at 65°C. The RNA was reverse transcribed in 20µl reaction containing 1 µg RNA, 1 µg Oligo (dT)15 Primer, 4 µl 5 X bbuffer, 0.5 µl Rnasin (Promega), 2 µl 10 mM dNTPs, 1 µl Reverse Transcriptase (Promega) in DEPC treated water. Reaction was carried out at 37°C for 1 hour.
RT PCR was performed with 1 µl of cDNA and the reaction was stoped in the linear range (25-30 cycles).

4.2.3 Protein methods

4.2.3.1 Separation of proteins on Polyacrylamide gels.

Stacking gel: 5% Acrylamide, 125mM Tris pH 6.8, 0.1% SDS, 0.1% APS, 0.1% TEMED
Running gel: 5-15% Acrylamide, 375mM Tris pH 8.8, 0.1% SDS, 0.1% APS, 0.04% TEMED
3X protein loading buffer: 250mM Tris, 25% Glycerol, 5% SDS, 0.25% Bromphenolblue
10X Running buffer: Tris base 30.3 g, Glycine 144 g SDS 10 g, add dH2O to 1l.

1. Prepare a gel of suitable concentration for the desired separation range.
2. Add 0.25 voulmes of 3X loading buffer to the sample and incubate for 5 min at 95°C.
3. Load the samples and run the gel at 150V for 1-2h depending on the protein size.

4.2.3.2 Staining of Polyacrylamide gels with Coomassie Blue

GelCode BlueStain reagent from PIERCE (Colloidal Coomassie Blue)
Fixing solution: 50% methanol and 7% acetic acid

1. Incubate the gel in fixing solution with gentle agitation for 15 min.
2. Wash 5 times for 10 min with ultrapure water to remove fixing solution.
3. Swirl the GelCode solution gently for 5-10 min on a shaker.

4. Add enough of the GelCode solution to cover the gel, and incubate for 2h with gentle agitation.
5. Remove the stain and add ultrapure water. Exchange the water several times and incubate overnight to fully develop the bands.

4.2.3.3 Western blot analysis

Bjerrum and Schafer-Nielsen Transfer buffer, pH 9.2: 48 mM Tris, 39 mM glycine, 20% methanol, 0.375% SDS.
PBST: 0.1% Tween 20 in PBS.
Blocking buffer: 5% non-fat dry milk in PBST

Transfer to nitrocellulose membrane
Protein separated on SDS polyacrylamide gels were transferred to nitrocellulose membrane (Protran, Schleicher & Schnell) by semi-dry blot in Bjerrum and Schafer-Nielsen transfer buffer using the Trans-Blot SD Semi-Dry Electrophoretic Transfer Cell (Bio-Rad) for 45 min at 15V.
Western blot
Membranes were blocked overnight with Blocking buffer; incubated with primary antibody for 1h; washed 5x10 min with PBST; incubated with horseradish peroxidase labeled secondary antibody for 1h and washed again as above. The blots were developed with ECL Plus Western Blotting Detection Reagent (Amersham).

4.2.3.4 Expression and purification of GST tagged proteins from *E. coli*

Buffer E: 150 mM NaCl, 20 mM HEPES pH=8.0, 10% glycerol, 0.1% Tween 20

1. Transform the pGEX expression construct into E. coli BL21(DE3) cells.
2. Inoculate a single colony in 50 ml of LB + 100 µg/ml ampicillin and grow for 8h.
3. Dilute 1:50 with fresh LB + 100 µg/ml ampicillin and grow at 30°C until OD600 of 0.5.
4. Add IPTG to final concentration 1mM to induce expression.
5. Grow for further 2h at 30°C with vigorous shaking.
6. Collect the cells by centrifugation at 5000g for 10 minutes.
7. Remove the supernatant and add 50 ml of Buffer E plus protease inhibitors (Roche).
8. Sonicate in ice-water bath at maximum settings for a total of 90 seconds at 10 second intervals with 60 seconds cooling.
9. Spin at 10000g for 10 minutes.
10. Transfer the supernatant to a new tube and spin at 1000 000g for 30 minutes.
11. Transfer the supernatant to a 50 ml Falcon tube and add 2 ml of glutathione beads (equilibrated in lysis buffer).
12. Incubate on a wheel for 1h and pour into a 20 ml disposable Econopack column.
13. Let the column drain by gravity flow and wash with 5x20 ml Buffer E.
14. Elute with 10 times 0.5 ml of Buffer E + 10 mM glutathione (collect 0.5 ml fractions).
15. Check the amount of protein by Bradford reaction and pool the peak fractions.
16. Dialyze overnight to Buffer E and freeze in aliquots in liquid nitrogen. Store at -80°C.

4.2.3.5 Protein extraction from HeLa cells

Buffer E: 150 mM NaCl, 20 mM HEPES pH=8.0, 10% glycerol, 0,5%NP40

1. Thaw the cell pellets and add 4 pellet volumes Buffer E and protease inhibitors (Sigma).
2. Suspend the pellet by pipetting up and down.
3. Dounce at least 50 times with type B pestle.
4. Spin at 10000g for 10 minutes.
5. Transfer the supernatant to a new tube and spin again at 10000g for 15 min.
6. The supernatant contains the soluble protein extract.

4.2.3.6 Tandem affinity purification from mammalian cell extracts.

All tags were purified using Buffer E for all steps, except for the calmodulin step of CaTZZ purification where calmodulin binding buffer (CBB) and calmodulin elution buffer (CEB) were used (CBB), as in the original TAP protocols[1,86].
Buffer E: 150 mM NaCl, 20 mM HEPES pH=8.0, 10% glycerol, 0.1% Tween 20.
SEB (Streptavidin Elution Buffer): Buffer E+2 mM biotin.
CBB: 10 mM Tris-Cl, pH 8, 10 mM beta-mercaptoethanol, 150 mM NaCl, 1 mM Mg-acetate, 1 mM Imidazole, 2 mM CaCl2, 0.1 % NP40
CEB: 10 mM Tris-Cl, pH 8, 10 mM beta-mercaptoethanol, 150 mM NaCl, 1 mM Mg-acetate, 1 mM Imidazole, 10 mM EGTA, 0.1 % NP40

Tag	First Beads	First elution	Second beads	Second elution
CaTZZ	IgG sepharose	TEV	Calmoduin	CEB
SbPZZ	IgG sepharose	PreScission	Steptavidin	SEB
SbPCh	Chitin	PreScission	Steptavidin	SEB
SbPS	S-Protein agarose	PreScission	Steptavidin	SEB
SbPSnap	BG sepharose	PreScission	Steptavidin	SEB
TeSnaPCh	Chitin	PreScission	BG sepharose	TEV

1. Add 100µl of First beads to the clarified protein extract.
2. Incubate for 2 hours at 4˚C using a rotating wheel.
3. Pour into disposable 20 ml Econopack column.
4. Wash 5 times with 20 ml Buffer E
5. Let the column drain.
6. Add 1 ml of Buffer E and 10U of TEV or PreScission.
7. Incubate overnight at 4˚C using a rotating wheel.
8. Collect the eluate and wash the beads with 5 ml of Buffer E (or SBB for CaTZZ). Combine the eluate and the wash.
9. (CaTZZ only) Add 3 µl/ml
10. Add 100 of second beads.
9. Incubate for 2 hours at 4˚C using a rotating wheel.
10. Pour into disposable 10 ml Econopack column.
11. Wash 5 times with 20 ml Buffer E (or SBB for CaTZZ).
12. Elute with 1 ml of SEB or CEB.

4.2.3.7 Mass spectrometry

The mass spectrometry analysis was performed by the MPI-CBG MS facility.

4.3 Recombinational cloning

4.3.1 Preparation of cassettes for Red/ET recombination by PCR

Linear cassettes for homologous recombination are generated by PCR primed by oligonucleotides with additional 30-50bp of "homology arms". The template plasmid is removed by Dpn I treatment. Dpn I is a restriction endonuclease that works only on methylated DNA. It will digest template DNA purified from methylatioin proficient *E. coli* strain leaving only the PCR product. Use of Dpn I is not required for the generation of the tagging cassette if the pR6K origin plasmid is used as a template.

1. Set up 100 µl PCR reaction with proofreading polymerase. Do not use more cycles than necessary to reduce the risk of PCR introduced mutations.
2. Check 3 µl of the reaction on a gel to ensure the reaction was successful.
3. Add 1 µl (10 units) of DpnI and incubate at 37°C for 1 hour.
4. Desalt and concentrate the PCR product by LiCl precipitation (see chapter 4.2.1.6).

4.3.2 Preparation of competent cells and electroporation.

1. Grow a fresh 1ml bacterial culture to OD600 of 0.4.
2. Centrifuge for 30 seconds at 11,000 rpm in a cooled benchtop centrifuge (at 2°C).
3. Discard the supernatant by quickly tipping out the supernatant twice, and place the pellet on ice.
4. Resuspend the pellet with 1 ml chilled ddH2O, pipetting up and down three times to mix the suspension.
5. Repeat the centrifugation and resuspend the cells in 50 µl of ddH2O.
6. Add 1 µl DNA to your cell pellet. Mix briefly. Transfer 50µl of the cell suspension from the tube to the chilled electroporation cuvette.
7. Electroporate at 1350 V, 10µF, 600 Ohms. Settings apply to an Eppendorf Electroporator 2510 with 1 mm gap electroporation cuvette. For other devices follow manufacturer recommendations.
8. Resuspend the electroporated cells in 1 ml LB medium without antibiotics and return them to the Eppendorf tube.

4.3.3 Liquid Culture Recombineering Pipeline:

Before starting, inspect the integrity of the BAC clone by DNA preparation and restriction digest using the protocol in Chapter 4.2.1.1. The pipeline has 4 steps each of which takes 1 day. Detailed protocol is given below:

Day 1. Transformation of pRedFlp3

1. Start a 1 ml culture of the BAC clone of interest in LB plus chloramphenicol. Grow to OD600 of 0.4. The most convenient way is to start from fresh overnight culture. If you inoculate with 20 µl it will take 2 hours at 37°C to reach the required OD.
2. Prepare competent cells and electroporate with 100 ng of pRedFlp as described (Chapter 4.3.2.)
3. After electroporation incubate for 1 hour at 30°C with vigorous shaking in 1ml of LB without antibiotics.
4. Transfer 100 µl to a new tube containing 1 ml of LB plus chloramphenicol, ampicillin and trimethoprim.
5. Grow at 30°C with vigorous shaking until the culture is saturated (takes about 20h).

Day 2. Insertion of the GFP FRT KmR FRT cassette:

1. Transfer 20 µl of the saturated culture to a new tube with 1 ml of LB plus chloramphenicol, ampicillin and trimethoprim.
2. Grow at 30°C for 2h with vigorous shaking to OD600 of 0.2.
3. Induce Red expression by adding 20 µl of 25% L-rhamnose.
4. Grow for 1h at 37°C with vigorous shaking to OD600 of 0.4.
5. Prepare competent cells and electroporate 1µg of the tagging cassette as described in Chapter 4.3.2.
6. Grow for 1h at 37°C with vigorous shaking in LB without antibiotics.
7. Transfer 100 µl of the culture to new tube with 1 ml of LB plus chloramphenicol, ampicillin, trimethoprim and kanamycin.
8. Grow at 30°C with vigorous shaking until the culture is saturated (takes 20h-30h).

Day 3. Flipout of the KmR gene:

1. Transfer 10 µl of the saturated culture to a new tube with 1ml of LB plus chloramphenicol, ampicillin, trimethoprim and 200 nM anhydrotetracycline.
2. Grow at 30°C with vigorous shaking until the culture is saturated (overnight).

Day 4. Subcloning in pUB:

1. Transfer 20 µl of the saturated culture to a new tube with 1 ml of LB plus chloramphenicol, ampicillin and trimethoprim.
2. Grow at 30°C for 2h with vigorous shaking to OD600 of 0.2.
3. Induce Red expression by adding 20 µl of 25% L-rhamnose.
4. Grow for 1h at 37°C with vigorous shaking to OD600 of 0.4.

5. Prepare competent cells and electroporate 1µg of the subcloning cassette as described in Chapter 4.3.2.
6. Grow for further 1h at 37°C with vigorous shaking in LB without antibiotics.
7. Plate 100 µl of the culture to LSLB agar plate plus blasticidin.
8. Incubate overnight at 37°C.
9. Pick single colonies and analyze them by DNA mini preps and restriction digest.

4.4 C. *elegans* specific methods

4.4.1 Maintaining worms on solid media

Nematode growth medium (NGM): 3 g NaCl, 17 g agar, and 2.5 g peptone Add 975 ml H2O. Autoclave for 50 min. Cool flask in 55°C water bath for 15 min. Add 1 ml 1 M CaCl2, 1 ml 5 mg/ml cholesterol in ethanol, 1 ml 1 M $MgSO_4$ and 25 ml 1 M KPO_4 buffer. Swirl to mix well. Pour plates.

Seeding NGM plates

Grow an overnight culture of *E. coli* OP50 in LB. Dispense 100 µl to each 10cm or 50µl to each 4cm NGM plate. Spread the drop with sterile glass rod to obtain a thin lawn covering most of the plate, but not touching the walls. Let grow overnight at room temperature (12-16h). Seal in plastic bag and store at 4°C for up to 1 month.

Transferring worms

Once a week transfer 3-5 young adults to a new 10cm plate. Grow the worms at 20°C in the dark.

4.4.2 Cleaning and synchronizing worm populations.

Occasionally the plates will get contaminated with bacteria molds or yeast. To obtain a clean population the worms are treated with alkaline hypochlorite solution, which dissolves the worms and all contaminating bacteria but not the embryos, which are protected by the egg shell.

Alkaline hypochlorite solution: 1 ml of household bleach, 2.5 ml 1M NaOH, 1.5 ml H_2O.

1. Collect the worms from one 10 cm plate with several washes with 500 µl H_2O into 1,5 ml Eppendorf tube.
2. Let settle for 5 min and remove the supernatant, which contains bacteria and debris.
3. Add 1ml of alkaline hypochlorite solution. Put on a rotating wheel for about 3-5 min.
4. Inspect a droplet of the solution every 2 minutes under dissecting microscope. When all the adults are dissolved spin down at 1000g for 1 min.
5. Remove the supernatant leaving about 30-50 µl and suspend the eggs in the remaining solution.
6. Distribute the eggs over 2-3 new 10 cm plates.
7. The next day the eggs will hatch and the worms will crow away onto the bacterial lawn. The new population will be roughly synchronous.
8. Check for contamination after 2 days and if necessary repeat the bleaching.

4.4.3 Freezing and thawing worm stocks

S-Buffer:129 ml 0.05 M K_2HPO_4, 871 ml 0.05 M K_2HPO_4, 5.85 g NaCl

Soft Agar Freezing Solution: 0.58 g NaCl, 0.68 g K_2HPO_4, 30 g glycerol, 0.56 ml 1 M NaOH, 0.4 g agar, H_2O to 100 ml (autoclave)

Freezing:

1. Melt Soft Agar Freezing Solution in autoclave or microwave and place in 50°C water bath for at least 15 minutes.
2. Use one large, 2-3 medium, or 5-6 small NGM plates that have lots of freshly starved L1-L2 animals. Wash the plates with 0.6 ml S Buffer for each vial you will freeze. Collect liquid in a covered sterile test tube and place in ice for 15 minutes.
3. Add an equal volume of Soft Agar Freezing Solution to the test tube. Mix well.
4. Aliquot 1 ml of mixture into 1.8 ml cryovials labelled with strain name and date.
5. Pack the cryovials in a small styrofoam box with slots for holding microtubes.

6. Place the box in a -80°C freezer overnight (or for at least 12 hours).
7. The next day transfer the vials to their permanent freezer locations.

Thawing:
1. Remove a vial from -80°C freezer and transfer to a small styrofoam box with slots for microtubes.
2. Flame a small scoop or spatula and use it to remove 1/4 - 1/3 ml of the frozen solution. Place solution on a NGM plate with E. coli OP50 lawn.
3. Return vial to -80°C freezer as quickly as possible.
4. After 2-3 days, transfer 10-15 animals individually to separate plates. Allow the animals to reproduce for one generation and score the progeny for correct phenotypes

4.4.4 Crossing worms

Male *rrf-3* worms were used to set a genetic cross to the *lin-59*:GFP line. Single F1 hermaphrodites were left to self fertilize. Single F2 hermaphrodites were picked and their progeny was analyzed by PCR to find *rrf-3* homozygotes. Lines that transmitted 100% GFP positive were assumed *lin-59*:GFP homozygote.
Crossing plates: 2 cm NGM plate with small OP50 lawn (~5 mm diameter)

1. Pick a single L4 hermaphrodite and 5-10 young males.
2. Let them mate overnight on a crossing plate and transfer them to a new plate to avoid food depletion.
3. Pick single L2-L3 stage F1 worms and transfer to new plates.

4.4.5 Ballistic transformation

Preparation of the microcarier beads.
Weigh out 30 mg of golden microparticles (Chempur 0.3-3 micron particles) into a 1.5 ml microfuge tube. Add 1 ml of 70% ethanol (v/v). Vortex vigorously for 3–5 minutes. Allow the particles to soak in 70% ethanol for 15 minutes. Pellet the microparticles by spinning for 5 seconds in a microfuge. Remove and discard the supernatant. Repeat the following wash steps three times:
• Add 1 ml of sterile water.
• Vortex vigorously for 1 minute.
• Allow the particles to settle for 1 minute.
• Pellet the microparticles by briefly spinning in a microfuge.
• Remove the liquid and discard.
After the third wash, add 500 µl sterile 50% glycerol to bring the microparticle concentration to 60 mg/ml (assume no loss during preparation). The microparticles can be stored at room temperature for up to two weeks or at 4 °C for 2 months.
Coating the microcarriers with DNA
10 µl of DNA at least 1 mg/ml, 16.7 µl gold solution, 150 µl CaCl2, 60µl 0.1M spermidine (Sigma-Aldrich) were mixed together. After adding each component the mixture was vortexed on lowest setting for 1 minute and after adding spermidine for 30 minutes. After spinning down, the supernatant was removed and the mixture was washed first with 70% ethanol then with 100% ethanol and resuspended in 100% ethanol.
Preparation of worms for bombardment
For each bombardment *unc-119*(ED3) worms were spread onto 15 peptone plates seeded with C600 bacteria and were grown for 7 days. Before bombardment the worms were washed 2 times with 1X M9 buffer, and were spread on NGM plates (sitting on ice).
Performing a bombradment
Sterilize the bombardment chamber walls with 70% ethanol. Load sterile rupture disk into sterile retaining cap. Secure retaining cap to end of gas acceleration tube and tighten with torque wrench. Suspend the DNA coated microparticles and load them onto a sterile macrocarrier. Load macrocarrier and stopping screen into microcarrier launch assembly. Place microcarrier launch assembly and target cells in chamber and close door. Evacuate chamber, hold vacuum at desired level (minimum 5 inches of mercury). Bombard sample: hold the Fire button continuously depressed until rupture disk bursts and helium pressure gauge drops to zero.

Screening for rescue
After bombardment collect the worms and distribute them onto 20 NGM plates. Let the worms grow for 2 weeks at 25°C. Inspect under dissecting microscope for unc rescued dauer larvae. Pick single worms and transfer to fresh 4cm NGM plates. Keep transferring for 3 generations; the populations that contain 100% wt worms after 3 generations are considered integrated.

4.4.6 RNAi by feeding with dsRNA

The worms are grown on a lawn of bacteria transformed with the plasmid for dsRNA production or with the empty vector. Production of dsRNA is induced by IPTG. The *E. coli* strain HT115 lacks double-strand-specific RNase III and accumulates large amount of the produced dsRNA.
Reagents:
NGM plates with carbencillin 25 ug/ml.
IPTG 2.1M (0,5 g/ml)
E. coli HT115(DE3) transformed with dsRNA producing vector.
Protocol:
1. Inoculate a single colony in 15 ml LB + carbencillin 50 µg/ml.
2. Grow for 8 to 16 h.
3. Spin and dissolve in 1.5 ml LB + carbencillin 50 µg/ml.
4. Add 75 µl 2.1 M IPTG (final concentration 100 mM).
5. Plate 50 µl on 4 cm plate with NGM + carbencillin 25 µg/ml.
6. Let grow ON.
7. Bleach adults and put eggs on the plate.
8. After 2 days pick a single worm and transfer to a new plate.
9. Score the progeny for phenotypes.

4.5 Cell culture

4.5.1 Maintenance of HeLa S3 cells

4.5.1.1 Adherent culture on plates

HeLa S3 and HeLa TDS cells were grown on plates in DMEM media supplemented with Penicillin/Streptomycin and 10% FCS. Cell were passaged at about 80% confluency.

4.5.1.2 Liquid culture of Hela S3 cells in spinner flasks

Spinner medium	**[for 1000 ml]**
SD medium	860 ml
FBS	100 ml
L-glutamine	10 ml
Penicillin Streptomycin	10 ml
HEPES	20 ml
G418 50 mg/ml (if appropriate)	10 ml

1. Trypsinize the cells from 3-5 T175 flasks grown to 60-80% confluency.
2. Suspend them in 100 ml of Spinner medium and determine the density.
3. Dilute the cells to $3x10^5$ and move them to the Spinner flask.
4. Put on a magnetic stirrer set at 130 rpm at 37°C.
5. Grow to ~$7x10^5$ before and dilute again to $3x10^5$.
6. Keep expanding until sufficient cells are obtained
7. Collect by centrifugation at 1200 g for 5min.
8. Remove the media and suspend the pellet in ice cold PBS.
9. Repeat steps 7,8 3 times.
10. Aliquot and freeze the cell pellet in liquid nitrogen.
11. Store at - 80°C.

4.5.1.3 Freezing and thawing cells

Freezing medium	[for 5 ml]
DMEM+10%FCS	2.5 ml
FCS	1.5 ml
DMSO	1 ml

Freezing:

1. Trypsinize the cells and resuspend in culture medium (DMEM+10%FCS).
2. Pipette 0.5 ml of freezing medium into labeled cryovials and put them on ice.
3. Mix the cell suspension and add 0.5 ml to each cryovial.
4. Close the lid and mix by inverting the tubes twice.
5. Store the cryovials at –80°C for overnight before moving them to liquid nitrogen for long term storage.

Thawing:

1. Prepare a 15 ml Falcon tube with 5 ml of culture medium for each vial to be thawed.
2. Quickly thaw the cells in a water bath 37°C.
3. Wash the tube with 70% ethanol.
4. Transfer the cell suspension to the Falcon tube and mix with the culture medium.
5. Centrifuge at 1000g for 5 min.
6. Aspirate the supernatant.
7. Resuspend the pellet in 1 ml of fresh culture media and transfer to a culture dish containing 9 ml of media.
8. Let the cells attach overnight and change the media.

4.5.2 Establishment of transgenic cell lines

4.5.2.1 Transfection

BAC DNA was transfected into HeLa cells using the Effecten transfection reagent (Qiagen). Control untransfected cells were always processed in parallel.

2. Mix 10 µl of 0.1 µg/µl BAC DNA with 140 µl DNA condensation buffer and 8 µl Enhancer buffer and 25 µl Effecten transfection reagent and incubate at room temperature for 10 min to allow DNA complex formation.
3. During the incubation wash the cells with PBS and add 4ml of fresh culture medium.
4. Add the DNA complexes drop-wise to 1ml of culture media, mix well and immediately transfer to the plate with the cells, gently swirling to ensure even distribution.
5. Let the cells grow under normal conditions.

4.5.2.2 Selection

1. Exchange the culture media after 24h incubation of the cells with the transfection complexes and let the cells grow for further 12-18h.
2. Transfer the cells to 10 cm plates and grow them in medium with increasing concentrations of G418 as follows: Day 1 500 µg/ml, Day 3 800 µg/ml, Day 5 1000 µg/ml, Day 7 1000 µg/ml. Keep changing the medium with 1000 µg/ml G418 until all cells on the control plate are dead.
3. Trypsinise the resistant clones and expand them as a clone pool.

4.5.2.3 Single cell sorting by FACs

1. Trypsinize the cells from one 60% confluent 10 cm clone pool plate.
2. Resuspend in PBS and filter through FACS filter.
3. Prepare 96 well plate containing 100 µl of 24h media supernatant of wild type Hela S3 cells.
4. FACS sort single cells into each 96 well.
5. After 2 day add 100 µl of fresh media with 1600 µg/ml G418.
6. Expand the cells to 48 well plates.
7. Select clones by RT PCR
8. Expand and freeze positive clones.

4.6 Imaging

4.6.1 Gel documentation

4.6.1.1 Imaging

Coomassie stained gels and chemiluminescent signal from western blots were scanned using the LAS 3000 CCD imaging system (FUJIFILM). Exposure was adjusted to prevent saturation of more abundant bands.

4.6.1.2 Image analysis

Bands were quantified with the AIDA image analyzer software (Raytest). Automatic peak and baseline determination were used. For quantitative analysis band intensity was compared to serial dilution of standards with known concentration, adjusted for the differences in molecular weight.

4.6.2 Microscopy

4.6.2.1 Preparation of agarose pads for mounting of worms:

1. Spot a drop of melted 2% agarose on a microscope slide
2. Immediately place a second slide on top and let the agarose solidify for 5 min.
3. Separate the two slides; the agarose pad should stick to one of them.

4.6.2.2 Mounting

1. Spot 10 µl of M9 supplemented with NaN_3 on a coverslip.
2. Using a wormpick collect 10-15 worms and place them in the drop of M9.
3. If embryos are to be observed cut the adult worms with syringe needles.
4. Place an agarose pad covered microscope slide on top of the drop; the coverslip should stick to the agar pad.

4.6.2.3 Imaging

Images were taken with Zeis axioplan II microscope equipped with DIC and fluorescent optics and Hamamatsu digital camera. Confocal images were taken using Zeis META 510 Confocal Laser Scanning Microscope.

4.6.2.4 Image processing

Image processing was performed with the Zeis LSM software The MetaMorph Imaging System (Molecular Devices) and the OsiriX software [133]. Figures were prepared using Adobe Photoshop and Adobe Illustrator (Adobe).

4.7 Bioinformatics

4.7.1 BAC clone map

To map the BAC clones we used BLAST[134] to align each end sequence to the CB25 C. *briggsae* genome assembly. Only high-scoring hits, longer than 300 bp, mapping to a single position were used. BAC pairs less than 10 and more than 300 kbp apart were considered false positive and were excluded.
The algorithm was implemented using Pearl script by Assen Roguev (Stewart group, Genomics TU Dresden). The source code is available upon request. The clone map is available online[74].

4.7.2 Fosmid selection tool

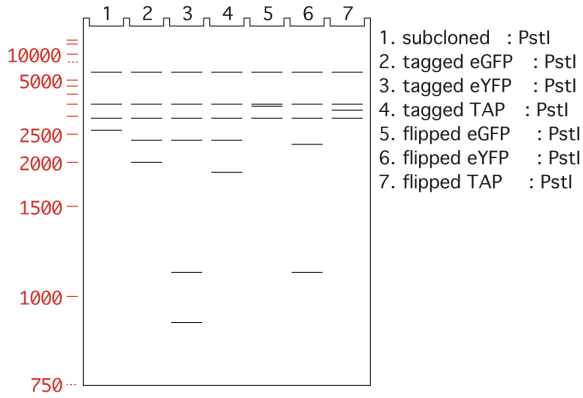
The algorithm was implemented by Jean Karim Heirche (Wellcome Trust Sanger Institute). The source code is available upon request.

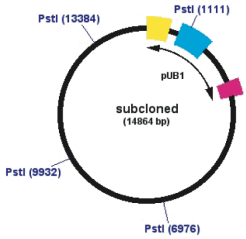
5 Supplementary data

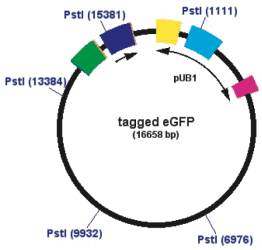
This section lists supplementary information required for understanding some of the figures but would have taken too much space in the main text.

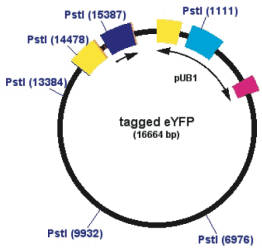
Supplementary data to Figure 7.

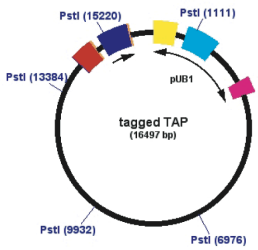
Plasmid maps and theoretical Pst I digest of the constructs from Figure 7.

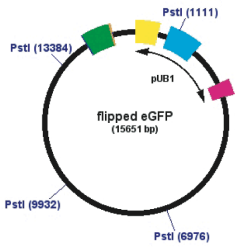


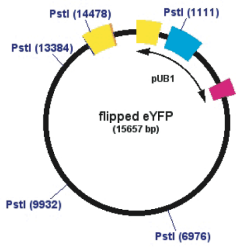


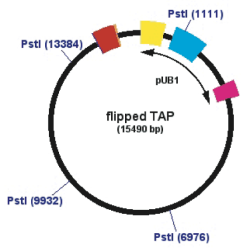






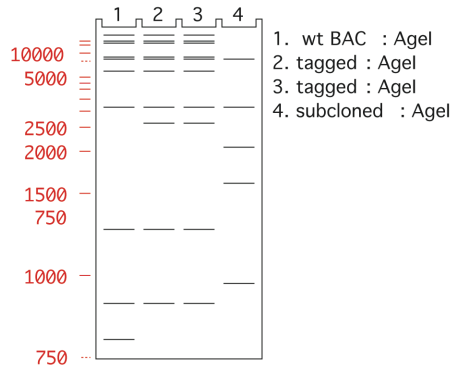


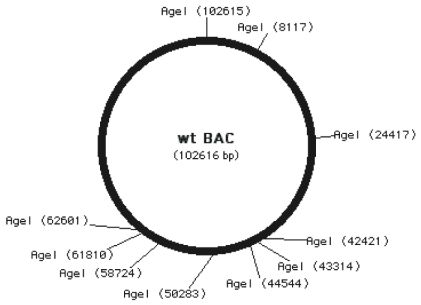


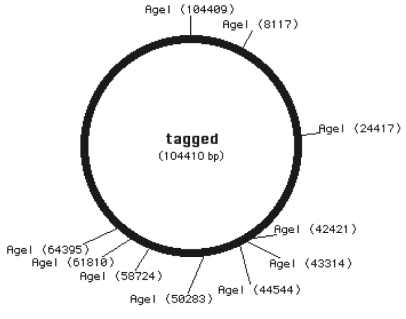


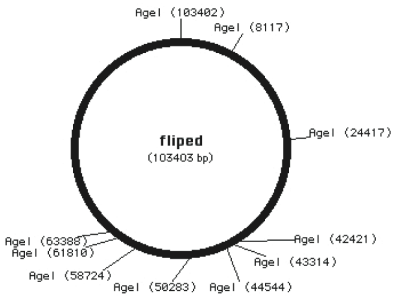
Supplementary data to Figure 8.

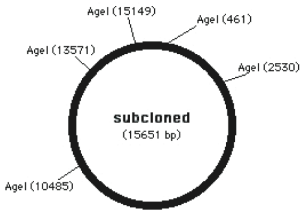
Plasmid maps and theoretical Age I digest of the constructs from Figure 8.





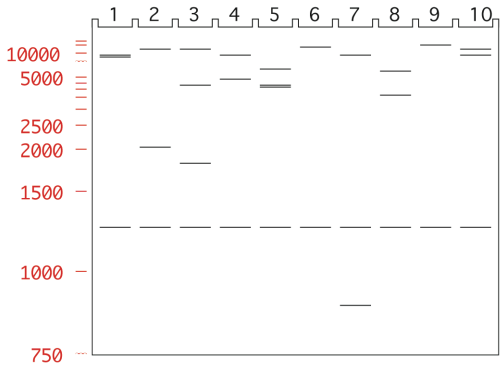




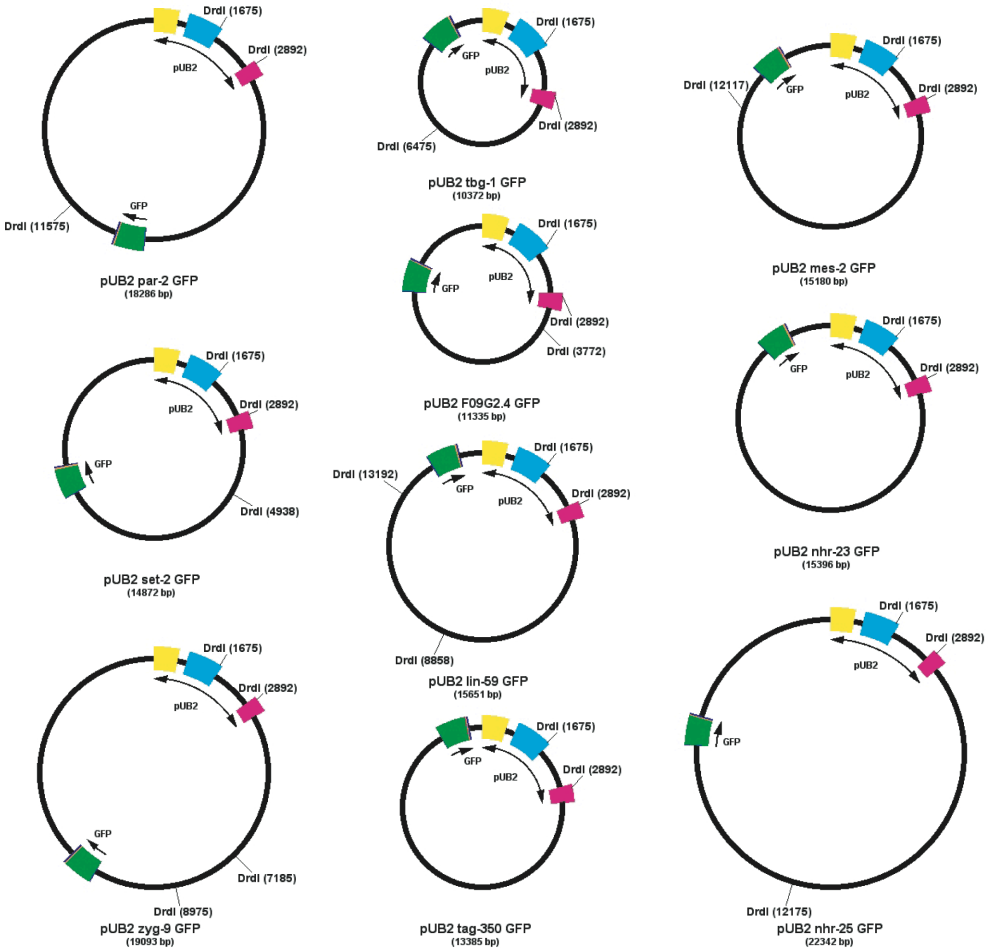


Supplementary data to Figure 16.

Plasmid maps and theoretical Drd I digest of the constructs from Figure 16.



1. pUB2 par-2 GFP : DrdI
2. pUB2 set-2 GFP : DrdI
3. pUB2 zyg-9 GFP : DrdI
4. pUB2 mes-2 GFP : DrdI
5. pUB2 lin-59 GFP : DrdI
6. pUB2 tag-350 GFP : DrdI
7. pUB2 F09G2.4 GFP : DrdI
8. pUB2 tbg-1 GFP : DrdI
9. pUB2 nhr-23 GFP : DrdI
10. pUB2 nhr-25 GFP : DrdI



6 References:

1. Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. & Seraphin, B. A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**, 1030-1032 (1999).

2. Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S. & O'Shea, E.K. Global analysis of protein localization in budding yeast. *Nature* **425**, 686-691 (2003).

3. Gavin, A.C. et al. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141-147 (2002).

4. Ho, Y. et al. Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* **415**, 180-183 (2002).

5. Gavin, A.C. et al. Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440**, 631-636 (2006).

6. Krogan, N.J. et al. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature* **440**, 637-643 (2006).

7. Bader, G.D., Heilbut, A., Andrews, B., Tyers, M., Hughes, T. & Boone, C. Functional genomics and proteomics: charting a multidimensional map of the yeast cell. *Trends Cell Biol* **13**, 344-356 (2003).

8. Suter, B., Auerbach, D. & Stagljar, I. Yeast-based functional genomics and proteomics technologies: the first 15 years and beyond. *Biotechniques* **40**, 625-644 (2006).

9. Hillier, L.W., Coulson, A., Murray, J.I., Bao, Z., Sulston, J.E. & Waterston, R.H. Genomics in C. elegans: so many genes, such a little worm. *Genome Res* **15**, 1651-1660 (2005).

10. Ristevski, S. Making better transgenic models: conditional, temporal, and spatial approaches. *Mol Biotechnol* **29**, 153-163 (2005).

11. Zhang, Y., Buchholz, F., Muyrers, J.P. & Stewart, A.F. A new logic for DNA engineering using recombination in Escherichia coli. *Nat Genet* **20**, 123-128 (1998).

12. Muyrers, J.P., Zhang, Y., Testa, G. & Stewart, A.F. Rapid modification of bacterial artificial chromosomes by ET-recombination. *Nucleic Acids Res* **27**, 1555-1557 (1999).

13. Muyrers, J.P., Zhang, Y. & Stewart, A.F. ET-cloning: think recombination first. *Genet Eng (N Y)* **22**, 77-98 (2000).

14. Zhang, Y., Muyrers, J.P., Testa, G. & Stewart, A.F. DNA cloning by homologous recombination in Escherichia coli. *Nat Biotechnol* **18**, 1314-1317 (2000).

15. Yu, D., Ellis, H.M., Lee, E.C., Jenkins, N.A., Copeland, N.G. & Court, D.L. An efficient recombination system for chromosome engineering in Escherichia coli. *Proc Natl Acad Sci U S A* **97**, 5978-5983 (2000).

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

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

18. Kuzminov, A. Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda. *Microbiol Mol Biol Rev* **63**, 751-813, table of contents (1999).

19. Hamilton, C.M., Aldea, M., Washburn, B.K., Babitzke, P. & Kushner, S.R. New method for generating deletions and gene replacements in Escherichia coli. *J Bacteriol* **171**, 4617-4622 (1989).

20. Yang, X.W., Model, P. & Heintz, N. Homologous recombination based modification in Escherichia coli and germline transmission in transgenic mice of a bacterial artificial chromosome. *Nat Biotechnol* **15**, 859-865 (1997).

21. Muyrers, J.P., Zhang, Y., Benes, V., Testa, G., Rientjes, J.M. & Stewart, A.F. ET recombination: DNA engineering using homologous recombination in E. coli. *Methods Mol Biol* **256**, 107-121 (2004).

22. Muyrers, J.P., Zhang, Y. & Stewart, A.F. Techniques: Recombinogenic engineering--new options for cloning and manipulating DNA. *Trends Biochem Sci* **26**, 325-331 (2001).

23. Court, D.L., Sawitzke, J.A. & Thomason, L.C. Genetic engineering using homologous recombination. *Annu Rev Genet* **36**, 361-388 (2002).

24. Poteete, A.R. What makes the bacteriophage lambda Red system useful for genetic engineering: molecular mechanism and biological function. *FEMS Microbiol Lett* **201**, 9-14 (2001).

25. Wang, J., Sarov, M., Rientjes, J., Fu, J., Hollak, H., Kranz, H., Xie, W., Stewart, A.F. & Zhang, Y. An Improved Recombineering Approach by

Adding RecA to lambda Red Recombination. *Mol Biotechnol* **32**, 43-54 (2006).

26. Bao, Z., Murray, J.I., Boyle, T., Ooi, S.L., Sandel, M.J. & Waterston, R.H. Automated cell lineage tracing in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **103**, 2707-2712 (2006).

27. Boyle, T.J., Bao, Z., Murray, J.I., Araya, C.L. & Waterston, R.H. AceTree: a tool for visual analysis of Caenorhabditis elegans embryogenesis. *BMC Bioinformatics* **7**, 275 (2006).

28. McKay, S.J. et al. Gene expression profiling of cells, tissues, and developmental stages of the nematode C. elegans. *Cold Spring Harb Symp Quant Biol* **68**, 159-169 (2003).

29. Reinke, V. Functional eXploration of the C. elegans genome using DNA microarrays. *Nature Genetics* **32 Suppl**, 541-546 (2002).

30. Colaiacovo, M.P., Stanfield, G.M., Reddy, K.C., Reinke, V., Kim, S.K. & Villeneuve, A.M. A targeted RNAi screen for genes involved in chromosome morphogenesis and nuclear organization in the Caenorhabditis elegans germline. *Genetics* **162**, 113-128 (2002).

31. Gonczy, P. et al. Functional genomic analysis of cell division in C. elegans using RNAi of genes on chromosome III. *Nature* **408**, 331-336 (2000).

32. Lee, S.S., Lee, R.Y., Fraser, A.G., Kamath, R.S., Ahringer, J. & Ruvkun, G. A systematic RNAi screen identifies a critical role for mitochondria in C. elegans longevity. *Nat Genet* **33**, 40-48 (2003).

33. Hamilton, B., Dong, Y., Shindo, M., Liu, W., Odell, I., Ruvkun, G. & Lee, S.S. A systematic RNAi screen for longevity genes in C. elegans. *Genes Dev* **19**, 1544-1555 (2005).

34. Simmer, F., Moorman, C., van der Linden, A.M., Kuijk, E., van den Berghe, P.V., Kamath, R.S., Fraser, A.G., Ahringer, J. & Plasterk, R.H. Genome-wide RNAi of C. elegans using the hypersensitive rrf-3 strain reveals novel gene functions. *PLoS Biol* **1**, E12 (2003).

35. Sonnichsen, B. et al. Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. *Nature* **434**, 462-469 (2005).

36. Liu, L.X. et al. High-throughput isolation of Caenorhabditis elegans deletion mutants. *Genome Res* **9**, 859-867 (1999).

37. Martin, E., Laloux, H., Couette, G., Alvarez, T., Bessou, C., Hauser, O., Sookhareea, S., Labouesse, M. & Segalat, L. Identification of 1088 new transposon insertions of Caenorhabditis elegans: a pilot study toward large-scale screens. *Genetics* **162**, 521-524 (2002).

38. Granger, L., Martin, E. & Segalat, L. Mos as a tool for genome-wide insertional mutagenesis in Caenorhabditis elegans: results of a pilot study. *Nucleic Acids Res* **32**, e117 (2004).

39. Williams, D.C., Boulin, T., Ruaud, A.F., Jorgensen, E.M. & Bessereau, J.L. Characterization of Mos1-mediated mutagenesis in Caenorhabditis elegans: a method for the rapid identification of mutated genes. *Genetics* **169**, 1779-1785 (2005).

40. Lehner, B., Crombie, C., Tischler, J., Fortunato, A. & Fraser, A.G. Systematic mapping of genetic interactions in Caenorhabditis elegans identifies common modifiers of diverse signaling pathways. *Nat Genet* **38**, 896-903 (2006).

41. Li, S. et al. A map of the interactome network of the metazoan C. elegans. *Science* **303**, 540-543 (2004).

42. Schwarz, E.M. et al. WormBase: better software, richer content. *Nucleic Acids Res* **34**, D475-478 (2006).

43. Genome sequence of the nematode C. elegans: a platform for investigating biology. *Science* **282**, 2012-2018 (1998).

44. Johnson, D.H., Tsao, J., Luo, M. & Carson, M. SGCEdb: a flexible database and web interface integrating experimental results and analysis for structural genomics focusing on Caenorhabditis elegans. *Nucleic Acids Res* **34**, D471-474 (2006).

45. Huang, R.Y., Boulton, S.J., Vidal, M., Almo, S.C., Bresnick, A.R. & Chance, M.R. High-throughput expression, purification, and characterization of recombinant Caenorhabditis elegans proteins. *Biochem Biophys Res Commun* **307**, 928-934 (2003).

46. Paik, Y.K., Jeong, S.K., Lee, E.Y., Jeong, P.Y. & Shim, Y.H. C. elegans: an invaluable model organism for the proteomics studies of the cholesterol-mediated signaling pathway. *Expert Rev Proteomics* **3**, 439-453 (2006).

47. Broverman, S., MacMorris, M. & Blumenthal, T. Alteration of Caenorhabditis elegans gene expression by targeted transformation. *Proc Natl Acad Sci U S A* **90**, 4359-4363 (1993).

48. Berezikov, E., Bargmann, C.I. & Plasterk, R.H. Homologous gene targeting in Caenorhabditis elegans by biolistic transformation. *Nucleic Acids Res* **32**, e40 (2004).

49. Mello, C.C., Kramer, J.M., Stinchcomb, D. & Ambros, V. Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. *Embo J* **10**, 3959-3970 (1991).

50. Pirrotta, V. Silence in the germ. *Cell* **110**, 661-664 (2002).

51. Cui, M., Kim, E.B. & Han, M. Diverse chromatin remodeling genes antagonize the Rb-involved SynMuv pathways in C. elegans. *PLoS Genet* **2**, e74 (2006).

52. Sijen, T. & Plasterk, R.H. Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. *Nature* **426**, 310-314 (2003).

53. Robert, V.J., Vastenhouw, N.L. & Plasterk, R.H. RNA interference, transposon silencing, and cosuppression in the Caenorhabditis elegans germ line: similarities and differences. *Cold Spring Harb Symp Quant Biol* **69**, 397-402 (2004).

54. Kim, J.K. et al. Functional genomic analysis of RNA interference in C. elegans. *Science* **308**, 1164-1167 (2005).

55. Praitis, V., Casey, E., Collar, D. & Austin, J. Creation of low-copy integrated transgenic lines in Caenorhabditis elegans. *Genetics* **157**, 1217-1226 (2001).

56. Wilm, T., Demel, P., Koop, H.U., Schnabel, H. & Schnabel, R. Ballistic transformation of Caenorhabditis elegans. *Gene* **229**, 31-35 (1999).

57. Jackstadt, P., Wilm, T.P., Zahner, H. & Hobom, G. Transformation of nematodes via ballistic DNA transfer. *Mol Biochem Parasitol* **103**, 261-266 (1999).

58. Lamesch, P. et al. C. elegans ORFeome version 3.1: increasing the coverage of ORFeome resources with improved gene predictions. *Genome Res* **14**, 2064-2069 (2004).

59. Wei, C., Lamesch, P., Arumugam, M., Rosenberg, J., Hu, P., Vidal, M. & Brent, M.R. Closing in on the C. elegans ORFeome by cloning TWINSCAN predictions. *Genome Res* **15**, 577-582 (2005).

60. Vaglio, P., Lamesch, P., Reboul, J., Rual, J.F., Martinez, M., Hill, D. & Vidal, M. WorfDB: the Caenorhabditis elegans ORFeome Database. *Nucleic Acids Res* **31**, 237-240 (2003).

61. Reboul, J. et al. C. elegans ORFeome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression. *Nat Genet* **34**, 35-41 (2003).

62. Dupuy, D. et al. A first version of the Caenorhabditis elegans Promoterome. *Genome Res* **14**, 2169-2175 (2004).

63. Hope, I.A., Stevens, J., Garner, A., Hayes, J., Cheo, D.L., Brasch, M.A. & Vidal, M. Feasibility of genome-scale construction of promoter:reporter gene fusions for expression in Caenorhabditis elegans using a multisite gateway recombination system. *Genome Res* **14**, 2070-2075 (2004).

64. Hawkins, N.C., Garriga, G. & Beh, C.T. Creating precise GFP fusions in plasmids using yeast homologous recombination. *Biotechniques* **34**, 74-78, 80 (2003).

65. Sassi, H.E., Renihan, S., Spence, A.M. & Cooperstock, R.L. Gene CATCHR--gene cloning and tagging for Caenorhabditis elegans using yeast homologous recombination: a novel approach for the analysis of gene expression. *Nucleic Acids Res* **33**, e163 (2005).

66. Silberstein, Z. & Cohen, A. Synthesis of linear multimers of OriC and pBR322 derivatives in Escherichia coli K-12: role of recombination and replication functions. *J Bacteriol* **169**, 3131-3137 (1987).

67. Stein, L.D. et al. The genome sequence of Caenorhabditis briggsae: a platform for comparative genomics. *PLoS Biol* **1**, E45 (2003).

68. Perkins, J., Wong, K., Warren, R., Schein, J.E., Stott, J., Holt, R., Jones, S., Marra, M.A. & Moerman, D. in **15th International *C. elegans* Conference**University of California, Los Angeles; 2005).

69. Maduro, M. & Pilgrim, D. Identification and cloning of unc-119, a gene expressed in the Caenorhabditis elegans nervous system. *Genetics* **141**, 977-988 (1995).

70. Monteilhet, C., Perrin, A., Thierry, A., Colleaux, L. & Dujon, B. Purification and characterization of the in vitro activity of I-Sce I, a novel and highly specific endonuclease encoded by a group I intron. *Nucleic Acids Res* **18**, 1407-1413 (1990).

71. Maduro, M. & Pilgrim, D. Conservation of function and expression of unc-119 from two Caenorhabditis species despite divergence of non-coding DNA. *Gene* **183**, 77-85 (1996).

72. Chamberlin, H.M. & Thomas, J.H. The bromodomain protein LIN-49 and trithorax-related protein LIN-59 affect development and gene expression in Caenorhabditis elegans. *Development* **127**, 713-723 (2000).

73. Buchholz, F., Angrand, P.O. & Stewart, A.F. Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nat Biotechnol* **16**, 657-662 (1998).

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

75. Skerra, A. Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in Escherichia coli. *Gene* **151**, 131-135 (1994).

76. Cardona, S.T. & Valvano, M.A. An expression vector containing a rhamnose-inducible promoter provides tightly regulated gene expression in Burkholderia cenocepacia. *Plasmid* **54**, 219-228 (2005).

77. Buchholz, F., Ringrose, L., Angrand, P.O., Rossi, F. & Stewart, A.F. Different thermostabilities of FLP and Cre recombinases: implications for applied site-specific recombination. *Nucleic Acids Res* **24**, 4256-4262 (1996).

78. Penfold, R.J. & Pemberton, J.M. An improved suicide vector for construction of chromosomal insertion mutations in bacteria. *Gene* **118**, 145-146 (1992).

79. Giaever, G. et al. Functional profiling of the Saccharomyces cerevisiae genome. *Nature* **418**, 387-391 (2002).

80. Bobinnec, Y., Fukuda, M. & Nishida, E. Identification and characterization of Caenorhabditis elegans gamma-tubulin in dividing cells and differentiated tissues. *J Cell Sci* **113 Pt 21**, 3747-3759 (2000).

81. Kostrouchova, M., Krause, M., Kostrouch, Z. & Rall, J.E. Nuclear hormone receptor CHR3 is a critical regulator of all four larval molts of the nematode Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **98**, 7360-7365 (2001).

82. Brooks, D.R., Appleford, P.J., Murray, L. & Isaac, R.E. An essential role in molting and morphogenesis of Caenorhabditis elegans for ACN-1, a novel member of the angiotensin-converting enzyme family that lacks a metallopeptidase active site. *J Biol Chem* **278**, 52340-52346 (2003).

83. Chen, Z., Eastburn, D.J. & Han, M. The Caenorhabditis elegans nuclear receptor gene nhr-25 regulates epidermal cell development. *Mol Cell Biol* **24**, 7345-7358 (2004).

84. Silhankova, M., Jindra, M. & Asahina, M. Nuclear receptor NHR-25 is required for cell-shape dynamics during epidermal differentiation in Caenorhabditis elegans. *J Cell Sci* **118**, 223-232 (2005).

85. Sarov, M. & Stewart, A.F. The best control for the specificity of RNAi. *Trends Biotechnol* **23**, 446-448 (2005).

86. Forler, D., Kocher, T., Rode, M., Gentzel, M., Izaurralde, E. & Wilm, M. An efficient protein complex purification method for functional proteomics in higher eukaryotes. *Nat Biotechnol* **21**, 89-92 (2003).

87. Kittler, R., Pelletier, L., Ma, C., Poser, I., Fischer, S., Hyman, A.A. & Buchholz, F. RNA interference rescue by bacterial artificial chromosome transgenesis in mammalian tissue culture cells. *Proc Natl Acad Sci U S A* **102**, 2396-2401 (2005).

88. Timmons, L., Court, D.L. & Fire, A. Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans. *Gene* **263**, 103-112 (2001).

89. Kamath, R.S., Martinez-Campos, M., Zipperlen, P., Fraser, A.G. & Ahringer, J. Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in Caenorhabditis elegans. *Genome Biol* **2**, RESEARCH0002 (2001).

90. Muyrers, J.P., Zhang, Y., Benes, V., Testa, G., Ansorge, W. & Stewart, A.F. Point mutation of bacterial artificial chromosomes by ET recombination. *EMBO Rep* **1**, 239-243 (2000).

91. Dolphin, C.T. & Hope, I.A. Caenorhabditis elegans reporter fusion genes generated by seamless modification of large genomic DNA clones. *Nucleic Acids Res* **34**, e72 (2006).

92. Buchholz, F., Angrand, P.O. & Stewart, A.F. A simple assay to determine the functionality of Cre or FLP recombination targets in genomic manipulation constructs. *Nucleic Acids Res* **24**, 3118-3119 (1996).

93. Materi, W. & Pilgrim, D. Novel Caenorhabditis elegans unc-119 axon outgrowth defects correlate with behavioral phenotypes that are partially rescued by nonneural unc-119. *Genesis* **42**, 104-116 (2005).

94. Wild, J., Hradecna, Z. & Szybalski, W. Conditionally amplifiable BACs: switching from single-copy to high-copy vectors and genomic clones. *Genome Res* **12**, 1434-1444 (2002).

95. Wild, J. & Szybalski, W. Copy-control pBAC/oriV vectors for genomic cloning. *Methods Mol Biol* **267**, 145-154 (2004).

96. Lin, X., Ruan, X., Anderson, M.G., McDowell, J.A., Kroeger, P.E., Fesik, S.W. & Shen, Y. siRNA-mediated off-target gene silencing triggered by a 7 nt complementation. *Nucleic Acids Res* **33**, 4527-4535 (2005).

97. Kulkarni, M.M., Booker, M., Silver, S.J., Friedman, A., Hong, P., Perrimon, N. & Mathey-Prevot, B. Evidence of off-target effects associated with long dsRNAs in Drosophila melanogaster cell-based assays. *Nat Methods* **3**, 833-838 (2006).

98. Henschel, A., Buchholz, F. & Habermann, B. DEQOR: a web-based tool for the design and quality control of siRNAs. *Nucleic Acids Res* **32**, W113-120 (2004).

99. Naito, Y., Yamada, T., Matsumiya, T., Ui-Tei, K., Saigo, K. & Morishita, S. dsCheck: highly sensitive off-target search software for double-stranded RNA-mediated RNA interference. *Nucleic Acids Res* **33**, W589-591 (2005).

100. Naito, Y., Yamada, T., Ui-Tei, K., Morishita, S. & Saigo, K. siDirect: highly effective, target-specific siRNA design software for mammalian RNA interference. *Nucleic Acids Res* **32**, W124-129 (2004).

101. Sauer, B. & McDermott, J. DNA recombination with a heterospecific Cre homolog identified from comparison of the pac-c1 regions of P1-related phages. *Nucleic Acids Res* **32**, 6086-6095 (2004).

102. Belfort, M. & Roberts, R.J. Homing endonucleases: keeping the house in order. *Nucleic Acids Res* **25**, 3379-3388 (1997).

103. Jamsai, D., Orford, M., Nefedov, M., Fucharoen, S., Williamson, R. & Ioannou, P.A. Targeted modification of a human beta-globin locus BAC clone using GET Recombination and an I-Scei counterselection cassette. *Genomics* **82**, 68-77 (2003).

104. Tischer, B.K., von Einem, J., Kaufer, B. & Osterrieder, N. Two-step red-mediated recombination for versatile high-efficiency markerless DNA manipulation in Escherichia coli. *Biotechniques* **40**, 191-197 (2006).

105. Baer, A. & Bode, J. Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes. *Curr Opin Biotechnol* **12**, 473-480 (2001).

106. Long, Q., Shelton, K.D., Lindner, J., Jones, J.R. & Magnuson, M.A. Efficient DNA cassette exchange in mouse embryonic stem cells by staggered positive-negative selection. *Genesis* **39**, 256-262 (2004).

107. Cobellis, G. et al. Tagging genes with cassette-exchange sites. *Nucleic Acids Res* **33**, e44 (2005).

108. Wimmer, E.A. Insect transgenesis by site-specific recombination. *Nat Methods* **2**, 580-582 (2005).

109. Oberstein, A., Pare, A., Kaplan, L. & Small, S. Site-specific transgenesis by Cre-mediated recombination in Drosophila. *Nat Methods* **2**, 583-585 (2005).

110. Barrett, P.L., Fleming, J.T. & Gobel, V. Targeted gene alteration in Caenorhabditis elegans by gene conversion. *Nat Genet* **36**, 1231-1237 (2004).

111. Plasterk, R.H. & Groenen, J.T. Targeted alterations of the Caenorhabditis elegans genome by transgene instructed DNA double strand break repair following Tc1 excision. *Embo J* **11**, 287-290 (1992).

112. Cheeseman, I.M. & Desai, A. A combined approach for the localization and tandem affinity purification of protein complexes from metazoans. *Sci STKE* **2005**, pl1 (2005).

113. Gottschalk, A., Almedom, R.B., Schedletzky, T., Anderson, S.D., Yates, J.R., 3rd & Schafer, W.R. Identification and characterization of novel nicotinic receptor-associated proteins in Caenorhabditis elegans. *Embo J* **24**, 2566-2578 (2005).

114. Tagwerker, C., Flick, K., Cui, M., Guerrero, C., Dou, Y., Auer, B., Baldi, P., Huang, L. & Kaiser, P. A tandem affinity tag for two-step purification under fully denaturing conditions: application in ubiquitin profiling and protein complex identification combined with in vivocross-linking. *Mol Cell Proteomics* **5**, 737-748 (2006).

115. Schimanski, B., Nguyen, T.N. & Gunzl, A. Highly efficient tandem affinity purification of trypanosome protein complexes based on a novel epitope combination. *Eukaryot Cell* **4**, 1942-1950 (2005).

116. Drakas, R., Prisco, M. & Baserga, R. A modified tandem affinity purification tag technique for the purification of protein complexes in mammalian cells. *Proteomics* **5**, 132-137 (2005).

117. Li, Q., Dai, X.Q., Shen, P.Y., Cantiello, H.F., Karpinski, E. & Chen, X.Z. A modified mammalian tandem affinity purification procedure to prepare functional polycystin-2 channel. *FEBS Lett* **576**, 231-236 (2004).

118. Geng, J. & Carstens, R.P. Two methods for improved purification of full-length mammalian proteins that have poor expression and/or solubility using standard Escherichia coli procedures. *Protein Expr Purif* **48**, 142-150 (2006).

119. Rohila, J.S., Chen, M., Cerny, R. & Fromm, M.E. Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. *Plant J* **38**, 172-181 (2004).

120. Zhou, D., Ren, J.X., Ryan, T.M., Higgins, N.P. & Townes, T.M. Rapid tagging of endogenous mouse genes by recombineering and ES cell complementation of tetraploid blastocysts. *Nucleic Acids Res* **32**, e128 (2004).

121. Parks, T.D., Leuther, K.K., Howard, E.D., Johnston, S.A. & Dougherty, W.G. Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. *Anal Biochem* **216**, 413-417 (1994).

122. Keefe, A.D., Wilson, D.S., Seelig, B. & Szostak, J.W. One-step purification of recombinant proteins using a nanomolar-affinity streptavidin-binding peptide, the SBP-Tag. *Protein Expr Purif* **23**, 440-446 (2001).

123. Keefe, A.D. & Szostak, J.W. Functional proteins from a random-sequence library. *Nature* **410**, 715-718 (2001).

124. Wilson, D.S., Keefe, A.D. & Szostak, J.W. The use of mRNA display to select high-affinity protein-binding peptides. *Proc Natl Acad Sci U S A* **98**, 3750-3755 (2001).

125. Walker, P.A., Leong, L.E., Ng, P.W., Tan, S.H., Waller, S., Murphy, D. & Porter, A.G. Efficient and rapid affinity purification of proteins using recombinant fusion proteases. *Biotechnology (N Y)* **12**, 601-605 (1994).

126. Cordingley, M.G., Callahan, P.L., Sardana, V.V., Garsky, V.M. & Colonno, R.J. Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro. *J Biol Chem* **265**, 9062-9065 (1990).

127. Peersen, O.B., Madsen, T.S. & Falke, J.J. Intermolecular tuning of calmodulin by target peptides and proteins: differential effects on Ca2+ binding and implications for kinase activation. *Protein Sci* **6**, 794-807 (1997).

128. Barth, A., Martin, S.R. & Bayley, P.M. Specificity and symmetry in the interaction of calmodulin domains with the skeletal muscle myosin light chain kinase target sequence. *J Biol Chem* **273**, 2174-2183 (1998).

129. Safarik, I. & Safarikova, M. Magnetic techniques for the isolation and purification of proteins and peptides. *Biomagn Res Technol* **2**, 7 (2004).

130. Timmons, L. & Fire, A. Specific interference by ingested dsRNA. *Nature* **395**, 854 (1998).

131. Brenner, S. The genetics of Caenorhabditis elegans. *Genetics* **77**, 71-94 (1974).

132. Sijen, T., Fleenor, J., Simmer, F., Thijssen, K.L., Parrish, S., Timmons, L., Plasterk, R.H. & Fire, A. On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107**, 465-476 (2001).

133. Rosset, A., Spadola, L. & Ratib, O. OsiriX: an open-source software for navigating in multidimensional DICOM images. *J Digit Imaging* **17**, 205-216 (2004).

134. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J Mol Biol* **215**, 403-410 (1990).

I herewith declare that I have produced this paper without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This paper has not previously been presented in identical or similar form to any other German or foreign examination board.

The thesis work was conducted from October 2001 to October 2006 under the supervision of Prof Francis Stewart at the Department of Genomics, Biotec Centre, Technische Universitat Dresden.

Dresden, 04.10.2006,

/Mihail Sarov/