Generation of a conditional *lima1a* allele in zebrafish using the FLEEx switch technology

Peggy Jungke[§], Stefan Hans^{§*}, Mansi Gupta, Anja Machate, Daniela Zöller, Michael Brand*

Biotechnology Center and Center for Regenerative Therapies Dresden, Dresden University of Technology, Tatzberg 47-49, 01307 Dresden, Germany

[§] Contributed equally.

* Corresponding authors:

michael.brand@biotec.tu-dresden.de

stefan.hans@biotec.tu-dresden.de

Running Title: FLEEx technology in zebrafish

Key words: conditional alleles; FLEEx switch; Lima1a; zebrafish

Abstract

Gene trapping has emerged as a valuable tool to create conditional alleles in various model organisms. Here we report the FLEEx-based gene trap vector SAGFLEEx that allows the generation of conditional mutations in zebrafish by gene-trap mutagenesis. The SAGFLEEx gene trap cassette comprises the rabbit β -globin splice acceptor and the coding sequence of GFP, flanked by pairs of inversely oriented heterotypic target sites for the site-specific recombinases Cre and Flp. Insertion of the gene-trap cassette into endogenous genes can result in conditional mutations that are stably inverted by Cre and Flp, respectively. To test the functionality of this system we performed a pilot screen and analyzed the insertion of the gene-trap cassette into the *lima1a* gene locus. In this *lima1a* allele, GFP expression faithfully recapitulated the endogenous *lima1a* expression and resulted in a complete knockout of the gene in homozygosity. Application of either Cre or Flp was able to mediate the stable inversion of the gene trap cassette and showed the ability to conditionally rescue or reintroduce the gene inactivation. Combined with pharmacologically inducible site specific recombinases the SAGFLEEx vector insertions will enable precise conditional knockout studies in a spatial- and temporal-controlled manner.

Introduction

Limitations of germline gene inactivation studies have been bypassed by the use of conditional alleles that allow the analysis of gene functions in a spatial manner at various developmental and adult stages. In this context site-specific recombinases (SSR) such as Cre and Flp have emerged as an indispensable tool for the precise manipulation of the mammalian genome (Branda and Dymecki, 2004). Traditionally, to generate a conditional allele homologous recombination is applied to insert SSR target sites into the locus of interest resulting in gene deletions, insertions or inversions in the presence of Cre or Flp, respectively (Schnutgen et al., 2003; Branda and Dymecki, 2004). Recently, we and others have shown that Cre and Flp are also highly efficient in developing and adult zebrafish (Boniface et al., 2009; Hans et al., 2009; Hans et al., 2011; Knopf et al., 2011; Kroehne et al., 2011) but the targeted generation of conditional alleles is impeded due to a lack of embryonic stem cell technology in zebrafish. To circumvent the limitations posed by germline mutations in mouse, gene trap vectors have been developed that allow conditional gene inactivation based on a strategy for directional site-specific recombination termed flip-excision or FLEEx (Schnutgen et al., 2005). Gene trapping provides a fast, easy, unbiased and high-throughput way to generate transgenic animals that accurately report endogenous gene expression (Stanford et al., 2001). In general, a gene trap vector consists of a splice acceptor (SA), a promoterless selectable marker gene and a transcriptional termination sequence or polyadenylation sequence (pA). The FLEEx configuration is provided by pairs of inversely oriented heterotypic target sites for Cre and Flp (Fig. 1). Upon integration into a transcribed gene the selectable marker becomes spliced to the endogenous upstream exon and results in a premature stop of the endogenous protein due to the polyA signal (Fig. 1A). Depending on the position within the trapped locus the insertion might facilitate the knockout of the endogenous protein function. Application of Flp results in an inversion of the cassette that restores normal splicing and rescues the mutational situation because the SA is no longer recognized (Fig. 1B). Due to the arrangement of the inversely oriented heterotypic target sites for Flp, the inversion event places two homotypic sites into a direct orientation which leads to an excision event that locks the recombination product against reinversion (Fig. 1C). Subsequent application of Cre results in the same sequence of events (inversion and excision) and restores the mutational situation in a Cre-dependent manner (Fig. 1D,E). General functionality of the FLEEx

technology in zebrafish was shown previously (Boniface et al., 2009; Maddison et al., 2014). Several gene-trap mutagenesis approaches have been reported, including the “gene-breaking transposon” (Clark et al., 2011) and the “Fliptrap” vector (Trinh le et al., 2011). The latter creates fluorescent tags of full-length proteins at their endogenous loci by the insertion of the Citrine coding sequence and simultaneously produces conditional mutants from a single integration. The Fliptrap represents an improved approach that provides protein expression information at cellular and subcellular resolutions. However, the incorporation of a Citrine-tag into a full-length protein at a random position can interfere with protein function resulting in a hypomorphic allele prior to the Cre-mediated flipping event. In order to generate conditional alleles that are not subjected to this disadvantage, a mCherry-tagged FLEEx-based conditional gene-trap mutagenesis approach was applied (Ni et al., 2012). Trapping of the *supv3l1* locus resulted in a complete knockout and embryonic lethality at 8 dpf. Application of site-specific recombinases was able to completely inactivate the mutagenized locus and rescue the mutant phenotype (Ni et al., 2012). To expand the pool of existing conditional alleles we generated the GFP-tagging gene trap vector SAGFLEEx and conducted a small pilot screen. In one isolated allele (*Tg(FLEEx-lima1a:GFP)*) the gene trap cassette was inserted into the *lima1a* locus. Lima1a is a novel tumor suppressor gene, also known as epithelial protein lost in neoplasm (EPLIN), that was initially identified as a transcriptionally down-regulated gene in oral cancer cells (Chang et al., 1998). In vitro studies revealed a role in cell growth, cytoskeletal organization and motility where it functions by cross linking F-actin and stress filaments. Furthermore, it has been shown to bind to the E-cadherin-catenin complex through α -catenin (Taguchi et al., 2011), and to inhibit branching nucleation of actin filaments via the Arp2/3- complex (Maul et al., 2003). In this study, we report the generation of a conditional *lima1a* allele, analyzed the expression pattern and tested the tightness and functionality of our SAGFLEEx gene trap cassette. Taken together, we show that the integration of the SAGFLEEx gene trap cassette results in a complete knockout of *lima1a*, that the functionality of the gene trap cassette operates as predicted and that this approach can be used for the generation of conditional alleles in zebrafish.

Material and Methods

Plasmid construction. To generate the gene trap vector SAGFLE_x, the heterotypic FRT/F3 and lox511/loxP recombinase target sequences for Flp and Cre recombinase were amplified from the pFlipROSA β geo plasmid (Schnutgen et al., 2005) using the following primers flanked by the indicated restriction sites: Flip-5'-for (EcoNI) 5'-AAAGCCTCTCACAGGTCGAGAATTCGAAGTTCCTATTCCGAAGTTCC-3'; Flip-5'-rev (ApaI) 5'-CTCGGGGGCCCAATTCGATAACTTCGTATAATGTGTACTATAC-3'; Flip-3'-for (BamHI) 5'-CACTGGGATCCGATAACTTCGTATAATGTATGC-3'; Flip-3'-rev (AscI) 5'-CCGGGGCGCGCCGTTGCTGAAGTTCCTATTCCGAAGTTCC-3'. Resulting amplicons were successively introduced up- and downstream to the GFP coding sequence of the Tol2 transposon-based gene trap vector T2KSAG (Kawakami et al., 2004). The full-length coding sequence of *lima1a* was amplified by PCR from cDNA of wild-type animals with the following primers and the indicated restriction sites: *lima1a*-for (ClaI) 5'-ATATATCGATAGGTACCAGAAGGCAGCAGAAG-3' and *lima1a*-wt-rev (NheI) 5'-ATATGCTAGCGACCACATCTTCATCATCCTCA-3'. The amplicon was subsequently inserted into a CS2+ GFP vector yielding a C-terminal in-frame wild-type Lima1a-GFP fusion construct. To obtain the truncated Lima1a-GFP fusion construct, mimicking the gene trap in *Tg(FLE_x-lima1a:GFP)*, the truncated *lima1a* gene was amplified by PCR from cDNA of wild-type animals with *lima1a*-for and *lima1a*-mut-rev (NheI) 5'-AGCTGCTAGCATTACTATGAAGGGTCTCTCCT-3' and inserted into the CS2+ GFP vector.

Zebrafish husbandry and germ line transformation. Zebrafish embryos were obtained by natural spawning of adult AB wild-type fish maintained at 28.5°C on a 14-hr light, 10-hr dark cycle and staged in hours post fertilization (hpf) and as previously described (Kimmel et al., 1995). For germ line transformation, plasmid DNA and transposase mRNA were injected into fertilized eggs (F0), raised to adulthood and crossed to AB wild-type fish as previously described (Kawakami et al., 2004). To identify transgenic carriers F1 embryos were examined under a fluorescent microscope at various developmental stages (1-5 dpf) and GFP-positive embryos were raised. This way twelve independent F0 were identified and *Tg(FLE_x-lima1a:GFP)* was investigated further and will be available to the research community upon acceptance of the manuscript. The CreER^{T2}-driver line

Tg(pax2a:CreERT2)tud102 which drives expression of CreER^{T2} in rhombomeres 3 and 5 of the developing hindbrain has been described previously (Hans et al., 2009).

***in situ* hybridization and mRNA synthesis.** Exon 9 of *lima1a* was PCR-amplified using cDNA from 24 hpf wild-typed embryos (lima1a-E9-for: 5'-GACGCCCCGAAGGAGTTGACG-3'; lima1a-E9-rev: 5'-ATGGAGGGCTGTTATCCGGCGT-3'). The amplicon was cloned into pCR®2.1-TOPO (Invitrogen). For probe synthesis the vector and pEGFP-N1 (Clontech) were linearized with Not1 and transcribed with SP6. Probe synthesis and *in situ* hybridization was performed essentially as previously described (Westerfield, 2000). Stainings were analyzed using a Zeiss Axiophot 2 or an Olympus MVX10 microscope. *In vitro* mRNA synthesis of wild-type and mutant Lima1-GFP as well as Flipase (Flp) and mCherry were performed using the mMACHINE SP6 Kit (Ambion). For mRNA injections, Flp (Sarov et al., 2006) was delivered into the cytoplasm of Tg(FLEx-lima1:GFP) embryos at the one-cell stage. mCherry mRNA was used for mock injections. Embryos were analyzed using a Zeiss Axiophot 2 microscope and a Zeiss LSM780.

Pharmacological treatments and heat induction. For tamoxifen (Sigma, St. Louis, MO;T5648) treatments, a 50mM stock solution was made in 100% DMSO and stored at -20°C. For embryo treatments, tamoxifen was diluted in embryo medium to 5 µM. At mid-gastrulation (75% epiboly or 6 hpf), embryos, still in their chorions, were transferred into petri dishes containing the treatment solution. For control treatments, sibling embryos were incubated in corresponding dilutions of DMSO. All incubations were conducted in the dark.

5' RACE and inverse PCR (iPCR). RNA was isolated from 10-15 embryos using Trizol (Ambion, Life Technologies) according to the manufacturer's protocol. 5'RACE was performed according to the manufacturer's protocol of the SMARTer™ RACE cDNA Amplification Kit (clontech) (GFP-r2 5'-CTTGCCGTAGGTGGCATCGCCCTC-3'; Nested-GFP r4 5'-GATGGGCACCAACCCCGGTGA-3'). For inverse PCR, 0.5-1µg purified genomic DNA was digested using HaeIII and MspI for 4h and self-ligated (T4 DNA Ligase, NEB) at room temperature in a total volume of 400µl. Ligated DNA was precipitated by adding 40µl 3M sodium acetate and 1ml ice-cold 100% EtOH, incubated at -80°C for 1h, centrifuged and washed with 750µl 70% EtOH. The dried pellet was resuspended in 50µl TE-buffer. 2.5µl of the resuspended DNA was used

as template for PCR (1st PCR: HaeIII/MspI-for 5'-CTCAAGTACAATTTTAATGGAGTAC-3'; HaeIII-rev 5'-TCCTGCAGTGCTGAAAAGC-3'; MspI-rev 5'-TGCCTTCTTCTCTTTCCTACAGC-3'; 2nd PCR: HaeIII/MspI-(nested)-for 5'-ACTCAAGTAAGATTCTAGCCAGA-3'; HaeIII-(nested)-rev 5'-GAAGTTCCTATTCTCTAGAAAGTATA-3'; MspI-(nested)-rev 5'-CATTTTGGCAAAGAATTCCTCGA-3'). PCR parameters: 30sec 95°C, 30sec 55°C, 45sec 72°C for 35 cycles followed by 7min 72°C.

Allele-specific PCR. Genomic DNA was extracted from 10-15 embryos of the respective condition and isolated as previously described (Jungke et al., 2015). 1µl of the resuspended DNA was used as a template for touchdown PCR (E2-for 5'-GCAGGACAGTATGTGCTGGGCAC-3'; GFP-rev 5'-GATGGGCACCACCCCGGTGA-3'; pA-for 5'-CAGGGGGAGGTGTGGGAGGT-3'). PCR parameters: 30sec 95°C, 30sec 68°C, 55sec 72°C for 5 cycles followed by 30sec 95°C, 30sec 60°C, 55sec 72°C for 20 cycles and 7min 72°C.

Quantitative real-time PCR (qRT-PCR). cDNA was generated using Superscript III First Strand Synthesis System (Invitrogen). qRT-PCR was performed as described (Kizil et al., 2009) using SYBR Green Supermix (BIORAD). Primer pairs used: E9-for 5'-GGCCTCCAGAGGGCGACACT-3'; E9-rev 5'-TCTCTTGGGTTGAGCCACGGG-3'; GFP-for 5'-CGCGCCGAGGTGAAGTTCGA-3'; GFP-rev 5'-TGCCGTCCTCGATGTTGTGGC-3'.

Whole mount immunohistochemistry. Antibody staining was carried out as described (Westerfield, 2000). Antibodies were used in the following dilutions: α-GFP (Molecular Probes), 1:500; goat α-rabbit Alexa Fluor 488 (Molecular Probes), 1:500. Embryos were analyzed using a Zeiss Axiophot 2 microscope.

Results

***Tg(FLEx-lima1a:GFP)* recapitulates endogenous *lima1a* expression**

In order to generate the gene trap vector SAGFLEx, we inserted the pairs of inversely oriented heterotypic target sites for Cre and Flp recombinases from the FlipRosa β geo vector (Schnutgen et al., 2005) into the T2KSAG vector, containing a rabbit β -globin splice acceptor, the coding sequence of GFP and a SV40 polyadenylation signal (Kawakami et al., 2004). In a subsequent pilot screen 12 transgenic lines were isolated (data not shown) and the insertion into the *lima 1a* gene locus *Tg(FLEx-lima1a:GFP)* was chosen as a proof-of-principle to test the functionality of the FLEx switch. Zebrafish *lima1a* is located on chromosome 23 and comprises nine exons. The trapping event occurred in intron 2, which was identified using inverse PCR (Fig. 2A and data not shown). To corroborate that the observed fluorescent GFP (Fig. 2B-F) recapitulates the endogenous *lima1a* expression, we performed *in situ* hybridization comparing GFP reporter and *lima1a* expression. At 80% epiboly, *lima1* is expressed ubiquitously throughout the embryo (data not shown) and becomes restricted to the fore-, mid and hindbrain primordia at 20 ss (Fig. 2L,M). This pattern is maintained during subsequent stages with the exception of the telencephalon which is devoid of *lima1a* at 48 hpf (Fig. 2N-P). The resemblance in expression is in particular striking in a dorsal view at 30 hpf when *gfp* and *lima1a* are expressed in the developing hindbrain primordium which was corroborated by double *in situ* hybridization showing a complete overlap of *gfp* and *lima1a* expression (Fig. 2Q-S). Taken together, *gfp* expression in *Tg(FLEx-lima1a:GFP)* mimics *lima1a* expression at all stages examined indicating that the insertion of the SAGFLEx vector faithfully recapitulates the endogenous *lima1a* expression.

Tg(FLEx-lima1a:GFP)* represents a null allele of *lima1a

Because the trapping event occurred in intron 2, we hypothesized that the polyadenylation sequence of the gene trap cassette entails transcriptional termination and the production of truncated *lima1a* transcripts. Thus, mRNA downstream to the insertion is not translated and expected to be degraded due to nonsense-mediated-decay. To test this hypothesis, heterozygous *Tg(FLEx-lima1a:GFP)* animals were

crossed and the resulting offspring were genotyped based on GFP fluorescence (Fig. 3B). Consistent with Mendelian inheritance 25% of the offspring showed no fluorescence, 50% medium and 25% strong GFP fluorescence. Subsequently, the three subgroups (no GFP, GFP+, GFP++) were analyzed separately by *in situ* hybridization using an exon 9 probe which hybridizes to *lima1a* transcripts downstream to the insertion site (Fig. 3A). Specimen with no and medium GFP fluorescence displayed a regular *lima1a* staining with a clearly, reduced expression in the latter (Fig. 3C,D). In sharp contrast, embryos with strong GFP fluorescence did not show any *lima1a* expression (Fig. 3E). To corroborate reduced levels of *lima1a* transcripts in *Tg(FLEX-lima1a:GFP)* embryos, we performed quantitative real-time PCR (qRT-PCR) on the three subgroups (no GFP, GFP+, GFP++). Compared to embryos with no GFP fluorescence, siblings with medium and strong GFP fluorescence showed a three- and nine-fold decrease in *lima1a* expression, respectively (Fig. 3F). Because Lima1 has been shown to localize to the cell cortex due to interaction with the E-cadherin-catenin complex (Taguchi et al., 2011) and that this cortical localization is abolished after removal of either the N- or C-terminal region (Song et al., 2002; Abe and Takeichi, 2008), we analyzed if the truncated Lima1a protein caused by the gene trap integration resulted in a changed Lima1a distribution. To this end we generated a full-length Lima1a-GFP construct and a truncated Lima1a-GFP variant which mimics the condition in *Tg(FLEX-lima1a:GFP)*. Injection of mRNA coding for the full-length Lima1a-GFP at the one-cell stage shows that the fusion protein localizes mostly to the cell membrane (Fig. 3G). In contrast, the truncated Lima1a-GFP fusion protein is distributed throughout the cytoplasm and also within the nuclei with no obvious enrichment at the cell cortex (Fig. 3H). Taken together, these results show that insertion of the SAGFLEX vector into the *lima1a* locus causes a strong reduction of *lima1a* expression and results in a truncated and incorrectly distributed Lima1a protein. However, homozygous *Tg(FLEX-lima1a:GFP)* animals are viable and do not show any overt phenotype compared to heterozygous or wild-type siblings indicating the existence of another protein compensating Lima1a function. Indeed, *lima1b* has been recently identified on chromosome 11 and shows a basal expression throughout the embryo with increased levels in the anlagen of the developing inner ear, kidney and posterior lateral line (Supporting Information Fig1).

Flp-mediated inversion of the gene trap cassette

To demonstrate the functionality of the SAGFLE_x vector, heterozygous *Tg(FLE_x-lima1a:GFP)* embryos were injected with Flp mRNA at the one-cell stage and analyzed for Flp-mediated inversion. In comparison to un- or mock-injected *Tg(FLE_x-lima1a:GFP)* embryos, injection of Flp mRNA resulted in a reduced (Fig. 4A,B) but not complete loss of fluorescent GFP indicating only a partial inversion of the gene trap cassette. To corroborate the Flp-mediated inversion, we applied allele-specific PCRs on genomic DNA extracted from Flp mRNA injected and uninjected *Tg(FLE_x-lima1a:GFP)* embryos. In the original insertion, an exon 2 forward primer (E2-for) and a GFP reverse primer (GFP-rev) yielded a 1018 bp fragment that could be detected in both Flp mRNA injected and uninjected *Tg(FLE_x-lima1a:GFP)* embryos (Fig. 4C). As expected, the combination of E2-for with a polyadenylation sequence forward primer (pA-for) never resulted in a PCR product in uninjected *Tg(FLE_x-lima1a:GFP)* embryos, because both primers anneal in the same orientation on the same DNA strand. In contrast, Flp-mediated inversion leads to primer annealing on opposite DNA strands in close proximity, yielding an 803 bp fragment, which could be detected in Flp mRNA injected *Tg(FLE_x-lima1a:GFP)* embryos (Fig. 4C). Subsequent cloning and sequencing of the 803 bp fragment confirmed the Flp-mediated inversion and excision event (data not shown). To determine if the inverted orientation of the gene trap cassette has a deleterious effect on the expression of *lima1a*, we raised Flp mRNA injected *Tg(FLE_x-lima1a:GFP)* embryos to adulthood, outcrossed them to wild-type and identified progeny carrying the Flp-inverted gene trap cassette *Tg(FLE_x-lima1a:GFP)^{inv}* based on the absence of GFP fluorescence and the before-mentioned allele-specific PCR. Subsequently, heterozygous *Tg(FLE_x-lima1a:GFP)^{inv}* animals were crossed and the resulting offspring were analyzed by *in situ* hybridization using the exon 9 probe which hybridizes to *lima1a* transcripts downstream to the insertion site. We found that all progeny displayed a regular *lima1a* staining indicating that embryos homozygous for the inverted GFP allele are indistinguishable from wild-type siblings (data not shown). In summary, these data demonstrate the Flp-mediated inversion of the gene trap cassette which has no deleterious effect on the expression of the trapped locus.

Restoration of the gene trap in a Cre-dependent manner

To investigate if the original orientation of the gene trap cassette can be restored in a Cre-dependent manner, *Tg(FLEX-lima1a:GFP)^{inv}* carrying the inverted orientation of the gene trap cassette was crossed with *Tg(pax2a:CreER^{T2})tud102* which expresses *CreER^{T2}* in rhombomeres 3 and 5 of the developing hindbrain (Hans et al., 2009). The resulting progeny were treated with tamoxifen to elicit recombination from late gastrulation stages and examined for GFP expression under a fluorescent dissecting microscope at 30hpf. In comparison to untreated or single transgenic embryos, GFP expression was observed in rhombomeres 3 and 5 of double-transgenic embryos indicating a successful re-inversion of the gene trap in a Cre-dependent manner (Fig. 5A,B). To corroborate this finding, we reiterated the allele-specific PCRs (see above) on genomic DNA extracted from GFP-positive and GFP-negative embryos. The primer combination E2-for and pA-for resulted in a PCR product of 803 bp. It thereby identified the FLP-inverted gene trap cassette in both GFP-positive and GFP-negative embryos (Fig. 5C). In contrast, usage of the primer combination E2-for and GFP-rev resulted in a 900 bp PCR product exclusively in GFP-positive, double transgenic embryos but not in GFP-negative embryos carrying the FLP-inverted gene trap cassette only (Fig. 5C). Subsequent cloning and sequencing of the 900 bp fragment confirmed the Cre-mediated inversion and excision event (data not shown). Taken together, these results show the successful inversion of the gene trap cassette in a tissue-specific, Cre-dependent manner.

Discussion

In the present study, we adapted gene trap mutagenesis with the site-specific recombination strategy (FLEEx) to create GFP-tagged conditional knockouts in zebrafish. This FLEEx approach enables the directional inversion of an inserted gene trap cassette. Thereby mutations, introduced by the gene trap vector can be inverted and re-induced using site-specific recombinases such as Cre or Flp. In general, the application of gene trap vectors has become a great tool for studying gene function in model organisms that lack homologous recombination strategies such as zebrafish. Compared to other gene trap mutagenesis methods applied in the recent past such as the gene-breaking transposon (Clark et al., 2011) or the Fliptrap (Trinh le et al., 2011), the FLEEx-based technology allows generation of both conditional rescue and conditional knockout alleles from a single integration event. Mutagenic disruption of the host gene is marked by the expression of a fluorescence reporter protein and thereby allows the identification of mutagenic integrations. Ni et al. reported the FLEExTrap vector FT1 comprising of an mCherry reporter exon downstream of a strong splice acceptor and upstream of five repeats of a transcriptional stop and polyadenylation sequence derived from the BGH gene (Ni et al., 2012). To increase the pool of differently tagged conditional alleles, we generated the SAGFLEEx gene trap vector consisting of a rabbit β -globin SA, GFP as a selectable marker and a SV40 polyA signal flanked by pairs of inversely oriented heterotypic target sites for the site-specific recombinases Cre and Flp and conducted a pilot screen. Out of 12 transgenic lines reporting a trapped gene locus, we tested the transgenic line *Tg(FLEEx-lima1a:GFP)* for functionality of our FLEEx cassette. We show that our SAGFLEEx gene trap vector inserts successfully into active loci in the zebrafish genome and that integration into intron 2 of the *lima1a* locus in *Tg(FLEEx-lima1a:GFP)* faithfully recapitulates the spatiotemporal expression pattern of the endogenous *lima1a* as shown by *in situ* hybridization. Furthermore, *in situ* analysis using an exon 9 probe which hybridizes to *lima1a* transcripts downstream to the insertion site revealed degradation due to nonsense-mediated-decay. This was corroborated using qRT-PCR analysis of *Tg(FLEEx-lima1a:GFP)* embryos with no GFP fluorescence compared to siblings with medium and strong GFP fluorescence that showed a three- and nine-fold decrease in *lima1a* expression, a decrease which is consistent with other FLEEx-based approaches (Ni et al., 2012). Finally, the truncated Lima1a-GFP fusion protein is not correctly localized to the cell cortex as the full-length Lima1a-

GFP fusion protein, indicating that the insertion of the SAGFLEEx gene trap vector generated *lima1a* null allele. Functionality of the SAGFLEEx using FLP recombinase could be shown and revealed an incomplete inversion of the cassette which can be explained by lower efficiency of FLP compared to Cre recombinase (Schnutgen et al., 2006). However, successful germline transmission of the inverted configuration (*Tg(FLEEx-lima1:GFP)^{inv}*) could be easily identified using an allele-specific PCR. Finally, combination of *Tg(FLEEx-lima1:GFP)^{inv}* with the CreER^{T2}-driver line *Tg(pax2a:CreER^{T2})tud102* which drives expression of CreER^{T2} in rhombomeres 3 and 5 of the developing hindbrain demonstrated tissue-specific Cre-dependent restoration of the cassette, proving the SAGFLEEx to be an excellent tool for conditional mutagenesis.

Expression of *lima1a* has been reported in zebrafish (Maul et al., 2001; Thisse, 2001), but functional data are missing. In cell culture experiments, Lima1 has been indicated as a cytoskeletal-associated protein involved in the regulation of actin dynamics and cell motility (Han et al., 2007) and localizes to the cell cortex due to interaction with the E-cadherin-catenin complex (Taguchi et al., 2011). Moreover knockdown of *lima1* results in the disassembly of adherens junctions, structurally distinct actin remodeling, and activation of β -catenin signaling (Zhang et al., 2011) and cortical Lima1 localization is abolished after removal of either the N- or C-terminal region (Song et al., 2002; Abe and Takeichi, 2008). Here, we confirm that a full-length Lima1a-GFP fusion protein is localized to the cell membrane also *in vivo* and that presence of the N-terminal region is not sufficient for proper cortical localization. However, despite the severe degradation due to nonsense-mediated-decay and the observed mislocalization of the truncated Lima1a-GFP fusion protein caused by the integrated SAGFLEEx gene trap vector into the *lima1a* locus, we did not observe any overt phenotype in homozygous *Tg(FLEEx-lima1a:GFP)* animals compared to heterozygous or wild-type siblings through development or adulthood. This indicates that another protein might compensate Lima1a function and indeed the closely related *lima1b* has been identified recently and is expressed similarly to *lima1a*. Loss of both, *lima1a* and *lima1b* will be necessary to elucidate any phenotypic consequences and *Tg(FLEEx-lima1:GFP)* might be useful for these future studies.

Recently conducted gene trapping approaches have provided a wealth of useful conditional alleles in zebrafish (Clark et al., 2011; Trinh le et al., 2011; Ni et al., 2012). Similarly, we show that our FLEEx-based gene trap vector SAGFLEEx can be applied in the successful generation of conditional knockouts. However, despite the fact that gene trapping approaches proved to be a fast and efficient strategy, insertion of the gene trap within the transcribed region is random and requires the detailed analysis of the trapping event to constitute a hypo- or amorphic allele. In mouse, defined conditional alleles are generated by the targeted insertion of loxP sites flanking a critical exon that leave the gene non-functional after Cre-mediated recombination (Skarnes et al., 2011). With the recent advent of novel genome editing tools such as TALEN (transcription activator-like effector nucleases) or CRISPR (clustered regularly interspaced short palindromic repeats)/Cas the targeted insertion of loxP sites is now also feasible in zebrafish (Auer and Del Bene, 2014). The nucleases create target-specific DNA double-strand breaks at predefined locations in the zebrafish genome which can be exploited to introduce loxP sites by homology-directed repair and result in the generation of conditional alleles similar to the mouse model (Bedell et al., 2012). In another approach, transgenic expression of Cas9 has been recently demonstrated to allow temporal and spatial controlled gene knockout in zebrafish (Ablain et al., 2015; Yin et al., 2015). Either method will enable the analysis of gene function in a temporal and spatial manner which will leave large gene trapping screens dispensable.

Acknowledgements

We thank past and present members of the Brand lab for generous support and helpful discussions as well as M. Fischer and J. Michling for excellent zebrafish care.

Disclosure Statement

No competing financial interests exist.

Funding

This project is supported by the Deutsche Forschungsgemeinschaft (BR 1746/3-1 and SFB 655, project A3) and the European Union (ZF Health F-002339-535-000-2301201). In addition, Michael Brand is supported by funding of the Excellence Initiative by the German Federal and State Governments (Institutional Strategy, measure "support the best"). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Figure legends

Fig. 1: Conditional gene inactivation by the SAGFLEEx gene trap vector. (A) The SAGFLEEx gene trap vector comprises the rabbit β -globin splice acceptor (SA), the coding sequence of GFP including a polyadenylation signal (pA) flanked by pairs of inversely oriented heterotypic target sites for Cre (loxP; lox511) and Flp (FRT; F3). Upon integration into a transcribed gene, GFP becomes spliced to the endogenous upstream exon and results in a premature stop of the endogenous protein due to the polyadenylation signal. (B) Application of Flp inverts the cassette into a non-mutagenic orientation using either FTR or F3 sites. (C) Following a subsequent excision event of two heterotypic Flp target sites, the cassette is locked against reinversion. (D-E) Cre-mediated recombination reinverts the cassette back into the original mutagenic situation by the same sequence of events using the Cre target sites. (Modified from Schnutgen et al., 2005.)

Fig. 2: *Tg(FLEEx-lima1a:GFP)* recapitulates endogenous *lima1a* expression. (A) Schematic representation of SAGFLEEx integrated into intron 2 of the *lima1a* locus in *Tg(FLEEx-lima1a:GFP)*. Native GFP fluorescence (B-F) and *gfp* transcription (G-K) in *Tg(FLEEx-lima1a:GFP)* compared to *lima1a* expression in wild-type embryos (L-P). Lateral views at 12somite stage (15hpf), 20somite stage (19hpf), 24hpf, 30hpf and 48hpf. Dorsal views of single and double *in situ* hybridizations for *lima1a* and *gfp* at 30hpf (Q-S).

Fig. 3: Functional consequences of the integration of SAGFLEEx into the *lima1a* locus. (A) Schematic representation of SAGFLEEx integrated into intron 2 of the *lima1a* locus in *Tg(FLEEx-lima1a:GFP)*. The positions of the primers used for qRT-PCR analysis (GFP-for/GFP-rev, E9-for/E9-rev) and the *lima1a* exon 9 probe are indicated. (B) Mating of heterozygous *Tg(FLEEx-lima1a:GFP)* animals results in offspring with no, medium and strong GFP fluorescence. (C-E) Expression of *lima1a* in embryos with no, medium and strong GFP fluorescence (no GFP, GFP+, GFP++) using an exon 9 probe. Lateral and dorsal views at 28hpf. (F) qRT-PCR analysis of *lima1a* exon 9 expression in embryos with medium and strong GFP fluorescence compared to siblings with no GFP fluorescence. ***p-value $\leq 0,001$. (G,H) Cellular

localization of GFP fusions with wild-type Lima1a encoded by exon 1-9 (E1-E9) and truncated Lima1a encoded by exon 1-2 (E1-E2).

Fig. 4: Flp-mediated inversion and excision of SAGFLEx. (A) Schematic representation of the events of SAGFLEx integrated into intron 2 of the *lima1a* locus in *Tg(FLEx-lima1a:GFP)* after Flp mRNA injection. The position of the primers used for the allele-specific PCR (E2-for/pA-for/GFP-rev) are indicated. (B) Native GFP fluorescence in uninjected and Flp mRNA injected *Tg(FLEx-lima1a:GFP)* embryos. Lateral and dorsal views at 28hpf. (C) Allele-specific PCR of uninjected and Flp mRNA injected *Tg(FLEx-lima1a:GFP)* embryos with the primers indicated above.

Fig. 5: Cre-mediated re-inversion of SAGFLEx. (A) Schematic representation of the events of SAGFLEx integrated into intron 2 of the *lima1a* locus in *Tg(FLEx-lima1a:GFP)* after Cre expression. The positions of the primers used for the allele-specific PCR (E2-for/pA-for/GFP-rev) are indicated. (B) Expression of Cre in rhombomeres 3 and 5 (r3/5) of the developing hindbrain results in the restoration of native GFP fluorescence. (C) Allele-specific PCR of GFP-positive and negative *Tg(FLEx-lima1a:GFP)^{inv}* embryos with the primers indicated above.

Supporting Information Fig. 1: Expression of lima1b. Lateral views of wild-type embryos at 12somite stage (15hpf), 24hpf and 30hpf.

References

- Abe, K. and Takeichi, M. (2008) 'EPLIN mediates linkage of the cadherin catenin complex to F-actin and stabilizes the circumferential actin belt', *Proc Natl Acad Sci U S A* 105(1): 13-9.
- Ablain, J., Durand, Ellen M., Yang, S., Zhou, Y. and Zon, Leonard I. (2015) 'A CRISPR/Cas9 Vector System for Tissue-Specific Gene Disruption in Zebrafish', *Developmental Cell* 32(6): 756-764.
- Auer, T. O. and Del Bene, F. (2014) 'CRISPR/Cas9 and TALEN-mediated knock-in approaches in zebrafish', *Methods* 69(2): 142-150.
- Bedell, V. M., Wang, Y., Campbell, J. M., Poshusta, T. L., Starker, C. G., Krug li, R. G., Tan, W., Penheiter, S. G., Ma, A. C., Leung, A. Y. H. et al. (2012) 'In vivo genome editing using a high-efficiency TALEN system', *Nature* 491(7422): 114-118.
- Boniface, E. J., Lu, J., Victoroff, T., Zhu, M. and Chen, W. (2009) 'FlEx-based transgenic reporter lines for visualization of Cre and Flp activity in live zebrafish', *Genesis* 47(7): 484-91.
- Branda, C. S. and Dymecki, S. M. (2004) 'Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice', *Dev Cell* 6(1): 7-28.
- Chang, D. D., Park, N. H., Denny, C. T., Nelson, S. F. and Pe, M. (1998) 'Characterization of transformation related genes in oral cancer cells', *Oncogene* 16(15): 1921-30.
- Clark, K. J., Balciunas, D., Pogoda, H. M., Ding, Y., Westcot, S. E., Bedell, V. M., Greenwood, T. M., Urban, M. D., Skuster, K. J., Petzold, A. M. et al. (2011) 'In vivo protein trapping produces a functional expression codex of the vertebrate proteome', *Nat Methods* 8(6): 506-15.
- Han, M. Y., Kosako, H., Watanabe, T. and Hattori, S. (2007) 'Extracellular signal-regulated kinase/mitogen-activated protein kinase regulates actin organization and cell motility by phosphorylating the actin cross-linking protein EPLIN', *Mol Cell Biol* 27(23): 8190-204.
- Hans, S., Freudenreich, D., Geffarth, M., Kaslin, J., Machate, A. and Brand, M. (2011) 'Generation of a non-leaky heat shock-inducible Cre line for conditional Cre/lox strategies in zebrafish', *Dev Dyn* 240(1): 108-15.
- Hans, S., Kaslin, J., Freudenreich, D. and Brand, M. (2009) 'Temporally-Controlled Site-Specific Recombination in Zebrafish', *PLoS ONE* 4(2): e4640.
- Jungke, P., Hammer, J., Hans, S. and Brand, M. (2015) 'Isolation of Novel CreER^{T2}-Driver Lines in Zebrafish Using an Unbiased Gene Trap Approach', *PLoS ONE* 10(6): e0129072.
- Kawakami, K., Takeda, H., Kawakami, N., Kobayashi, M., Matsuda, N. and Mishina, M. (2004) 'A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish', *Dev Cell* 7(1): 133-44.
- Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. and Schilling, T. F. (1995) 'Stages of embryonic development of the zebrafish', *Dev Dyn* 203(3): 253-310.
- Kizil, C., Otto, G. W., Geisler, R., Nusslein-Volhard, C. and Antos, C. L. (2009) 'Simplex controls cell proliferation and gene transcription during zebrafish caudal fin regeneration', *Dev Biol* 325(2): 329-40.
- Knopf, F., Hammond, C., Chekuru, A., Kurth, T., Hans, S., Weber, C. W., Mahatma, G., Fisher, S., Brand, M., Schulte-Merker, S. et al. (2011) 'Bone regenerates via dedifferentiation of osteoblasts in the zebrafish fin', *Dev Cell* 20(5): 713-24.
- Kroehne, V., Freudenreich, D., Hans, S., Kaslin, J. and Brand, M. (2011) 'Regeneration of the adult zebrafish brain from neurogenic radial glia-type progenitors', *Development* 138(22): 4831-41.
- Maddison, L., Li, M. and Chen, W. (2014) Conditional Gene-Trap Mutagenesis in Zebrafish. in M. F. Ochs (ed.) *Gene Function Analysis*, vol. 1101: Humana Press.
- Maul, R. S., Sachi Gerbin, C. and Chang, D. D. (2001) 'Characterization of mouse epithelial protein lost in neoplasm (EPLIN) and comparison of mammalian and zebrafish EPLIN', *Gene* 262(1-2): 155-60.
- Maul, R. S., Song, Y., Amann, K. J., Gerbin, S. C., Pollard, T. D. and Chang, D. D. (2003) 'EPLIN regulates actin dynamics by cross-linking and stabilizing filaments', *J Cell Biol* 160(3): 399-407.
- Ni, T. T., Lu, J., Zhu, M., Maddison, L. A., Boyd, K. L., Huskey, L., Ju, B., Hesselson, D., Zhong, T. P., Page-McCaw, P. S. et al. (2012) 'Conditional control of gene function by an invertible gene trap in zebrafish', *Proceedings of the National Academy of Sciences* 109(38): 15389-15394.

Sarov, M., Schneider, S., Pozniakovski, A., Roguev, A., Ernst, S., Zhang, Y., Hyman, A. A. and Stewart, A. F. (2006) 'A recombineering pipeline for functional genomics applied to *Caenorhabditis elegans*', *Nat Methods* 3(10): 839-44.

Schnutgen, F., De-Zolt, S., Van Sloun, P., Hollatz, M., Floss, T., Hansen, J., Altschmied, J., Seisenberger, C., Ghyselinck, N. B., Ruiz, P. et al. (2005) 'Genomewide production of multipurpose alleles for the functional analysis of the mouse genome', *Proc Natl Acad Sci U S A* 102(20): 7221-6.

Schnutgen, F., Doerflinger, N., Calleja, C., Wendling, O., Chambon, P. and Ghyselinck, N. B. (2003) 'A directional strategy for monitoring Cre-mediated recombination at the cellular level in the mouse', *Nat Biotechnol* 21(5): 562-5.

Schnutgen, F., Stewart, A. F., von Melchner, H. and Anastassiadis, K. (2006) 'Engineering embryonic stem cells with recombinase systems', *Methods Enzymol* 420: 100-36.

Skarnes, W. C., Rosen, B., West, A. P., Koutsourakis, M., Bushell, W., Iyer, V., Mujica, A. O., Thomas, M., Harrow, J., Cox, T. et al. (2011) 'A conditional knockout resource for the genome-wide study of mouse gene function', *Nature* 474(7351): 337-342.

Song, Y., Maul, R. S., Gerbin, C. S. and Chang, D. D. (2002) 'Inhibition of anchorage-independent growth of transformed NIH3T3 cells by epithelial protein lost in neoplasm (EPLIN) requires localization of EPLIN to actin cytoskeleton', *Mol Biol Cell* 13(4): 1408-16.

Stanford, W. L., Cohn, J. B. and Cordes, S. P. (2001) 'Gene-trap mutagenesis: past, present and beyond', *Nat Rev Genet* 2(10): 756-68.

Taguchi, K., Ishiuchi, T. and Takeichi, M. (2011) 'Mechanosensitive EPLIN-dependent remodeling of adherens junctions regulates epithelial reshaping', *J Cell Biol* 194(4): 643-56.

Thisse, B., Pflumio, S., Fürthauer, M., Loppin, B., Heyer, V., Degraeve, A., Woehl, R., Lux, A., Steffan, T., Charbonnier, X.Q. and Thisse, C. (2001) 'Expression of the zebrafish genome during embryogenesis', *ZFIN Direct Data Submission* (<http://zfin.org>).

Trinh le, A., Hochgreb, T., Graham, M., Wu, D., Ruf-Zamojski, F., Jayasena, C. S., Saxena, A., Hawk, R., Gonzalez-Serricchio, A., Dixon, A. et al. (2011) 'A versatile gene trap to visualize and interrogate the function of the vertebrate proteome', *Genes Dev* 25(21): 2306-20.

Westerfield, M. (2000) 'The zebrafish book. A guide for the laboratory use of zebrafish (*Danio rerio*).', 4th ed., Univ. of Oregon Press, Eugene.

Yin, L., Maddison, L. A., Li, M., Kara, N., LaFave, M. C., Varshney, G. K., Burgess, S. M., Patton, J. G. and Chen, W. (2015) 'Multiplex Conditional Mutagenesis Using Transgenic Expression of Cas9 and sgRNAs', *Genetics* 200(2): 431-441.

Zhang, S., Wang, X., Osunkoya, A. O., Iqbal, S., Wang, Y., Chen, Z., Muller, S., Chen, Z., Josson, S., Coleman, I. M. et al. (2011) 'EPLIN downregulation promotes epithelial-mesenchymal transition in prostate cancer cells and correlates with clinical lymph node metastasis', *Oncogene* 30(50): 4941-52.

Jungke_Fig2

Jungke_Fig3

Jungke_Fig4

Jungke_Supporting Information Fig1