LETTER

doi:10.1038/nature14580

Genetic compensation induced by deleterious mutations but not gene knockdowns

Andrea Rossi[1]*, Zacharias Kontarakis[1]*, Claudia Gerri[1], Hendrik Nolte[1]†, Soraya Hölper[1], Marcus Krüger[1]† & Didier Y. R. Stainier[1]

Cells sense their environment and adapt to it by fine-tuning their transcriptome. Wired into this network of gene expression control are mechanisms to compensate for gene dosage. The increasing use of reverse genetics in zebrafish, and other model systems, has revealed profound differences between the phenotypes caused by genetic mutations and those caused by gene knockdowns at many loci[1–3], an observation previously reported in mouse and *Arabidopsis*[4–7]. To identify the reasons underlying the phenotypic differences between mutants and knockdowns, we generated mutations in zebrafish *egfl7*, an endothelial extracellular matrix gene of therapeutic interest, as well as in *vegfaa*. Here we show that *egfl7* mutants do not show any obvious phenotypes while animals injected with *egfl7* morpholino (morphants) exhibit severe vascular defects. We further observe that *egfl7* mutants are less sensitive than their wild-type siblings to Egfl7 knockdown, arguing against residual protein function in the mutants or significant off-target effects of the morpholinos when used at a moderate dose. Comparing *egfl7* mutant and morphant proteomes and transcriptomes, we identify a set of proteins and genes that are upregulated in mutants but not in morphants. Among them are extracellular matrix genes that can rescue *egfl7* morphants, indicating that they could be compensating for the loss of Egfl7 function in the phenotypically wild-type *egfl7* mutants. Moreover, *egfl7* CRISPR interference, which obstructs transcript elongation and causes severe vascular defects, does not cause the upregulation of these genes. Similarly, *vegfaa* mutants but not morphants show an upregulation of *vegfab*. Taken together, these data reveal the activation of a compensatory network to buffer against deleterious mutations, which was not observed after translational or transcriptional knockdown.

Interfering with a gene's function is a widely used strategy to decipher its role. Several different approaches have been developed over the years to achieve this goal. Yet, despite having the same goal of functional inactivation, different strategies, namely knockdown (via antisense) and knockout (via genetic inactivation), often lead to different phenotypes. These discrepancies could be caused by off-target effects of the knockdown reagents, the generation and use of hypomorphic alleles, or other and more fundamental reasons. The recent development of new genome engineering techniques, such as TAL effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeats (CRISPRs), is allowing the facile generation of mutations and has revived concerns over the lack of specificity of knockdown reagents[1–3]. In several cases, toxicity due to off-target effects, induction of *p53* (also known as *tp53*) transcription, interferon response, engagement of toll-like receptors and/or saturation of the RNA interference machinery can lead to phenotypes unrelated to the silencing of the target gene[8,9]. To investigate further whether toxicity effects are the main reason for the differences between genetic mutation and gene knockdown phenotypes, we analysed the *EGF-like-domain, multiple 7* (*egfl7*) gene. The *egfl7* gene is a good candidate to address this question because of the lack of obvious phenotypes in the mouse mutants[10,11] and the severe vascular tube formation defects observed in knockdown experiments in zebrafish, frogs and human cells[12–14].

We first generated *egfl7* mutants using TALENs[15] targeting exon 3, which encodes part of the EMI domain (Fig. 1a and Extended Data Fig. 1). This domain precedes other domains critical for Egfl7 activity, including EGF domains and the leucine–valine-rich carboxy (C) terminus (Fig. 1a)[16]. We identified several deletion alleles including a *Δ3* and a *Δ4* (Fig. 1b). The *egfl7 Δ3* (hereafter *egfl7*^{s980}) allele encodes a protein that lacks a non-conserved proline at position 50 (p.P50del) while the *egfl7 Δ4* allele (hereafter *egfl7*^{s981}) is predicted to encode a truncated polypeptide containing a stretch of 29 incorrect amino acids starting with a Gln to Leu substitution at position 49 (p.Gln49Leufs*30) (Fig. 1b). To investigate the severity of these mutant alleles, we first examined *egfl7* transcript levels by quantitative PCR (qPCR). The premature stop codon in *egfl7*^{s981} led to a decrease of approximately 50% in transcript levels compared with wild-type (WT) and *egfl7*^{s980} mutant embryos, indicating an increased messenger RNA (mRNA) degradation rate (Extended Data Fig. 2a). To characterize the different *egfl7* mutant alleles further, we cloned the *egfl7* WT, *s980* and *s981* complementary DNAs (cDNAs) in a mammalian expression vector and transfected HUVEC cells. Unlike Egfl7 WT and Egfl7^{s980}, the Egfl7^{s981} protein was mostly absent in the medium or the cells, suggesting that this truncated polypeptide is rapidly degraded and/or poorly translated and secreted (Extended Data Fig. 2b). Altogether, these data indicate that *egfl7*^{s981} is a severe mutant allele, possibly even a null.

To analyse Egfl7 function during vascular development, the *egfl7*^{s981} mutant fish were crossed into the *Tg(kdrl:HRAS-mCherry)*[17] and *Tg(kdrl:GFP)*[18] backgrounds. We also developed a robust method based on high-resolution melt analysis to identify the different genotypes (Extended Data Fig. 1b). Surprisingly, no differences in gross morphology were evident between *egfl7* WT and mutant animals. However, a sporadic onset of brain haemorrhage was observed in fewer than 5% of the mutant animals at 72 hours post-fertilization (hpf) (Fig. 1c, d). Besides the haemorrhagic foci, no obvious abnormalities were detected in vasculogenesis, angiogenesis or circulation in any region of the brain or the rest of the body (Fig. 1e, f and Extended Data Fig. 3). Moreover, *egfl7*^{s981} mutant animals survive to become fertile adults. In summary, while *egfl7* morphants exhibit severe vascular defects[12], *egfl7* mutants exhibit very mild, if any, phenotypes.

This discrepancy between mutant and morphant phenotypes could be explained by several reasons including morpholino (MO) off-target effects. We thus sought to assess the specificity and toxicity of *egfl7* MO. First, to evaluate the effectiveness of the *egfl7* MO, we engineered the *egfl7* locus through the co-injection of TALENs and a single-stranded DNA (ssDNA) donor encoding a Myc-tag (Extended Data Fig. 4), and generated a stable transgenic line. We then injected Egfl7 Myc-tag embryos with 1 ng of *egfl7* MO and analysed protein levels by western blot at 24 hpf. The relative expression of Egfl7 Myc-tag was reduced by approximately 80% in the morphants compared with

[1]Max Planck Institute for Heart and Lung Research, 61231 Bad Nauheim, Germany. †Present address: Institute for Genetics and CECAD, University of Cologne, 50931 Cologne, Germany.
*These authors contributed equally to this work.

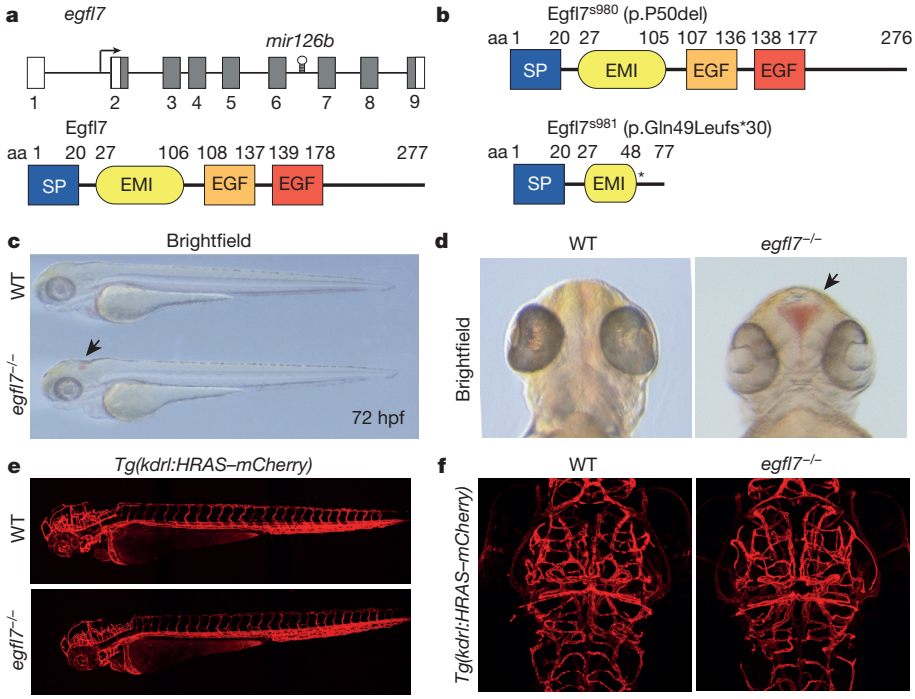


Figure 1 | Generation of zebrafish *egfl7* mutant alleles and sporadic brain haemorrhage in mutant larvae. **a**, Top: *egfl7* consists of 9 exons and *mir126b* is embedded in intron 6. The protein is encoded by exons 2–9 (grey boxes). **a**, Bottom: Egfl7, 277 amino acids (aa) long, contains a signal peptide (blue), an EMI domain (yellow), an EGF domain that contains a Delta-Serrate-LAG-2 (DSL) motif (orange) and a Ca^{2+}-binding EGF domain (red). **b**, Top: the $egfl7^{s980}$ lesion (Δ3) leads to the deletion of proline at position 50 (p.P50del). **b**, Bottom: the $egfl7^{s981}$ allele (Δ4) encodes a truncated 77-amino-acid-long polypeptide (p.Gln49Leufs*30) that contains a signal peptide (blue) and a partial EMI domain (yellow) followed by a frameshift leading to a premature stop codon. **c**, **d**, Brightfield micrographs of 72 hpf WT and $egfl7^{s981/s981}$ larvae in lateral and ventral views. Arrows point to area of haemorrhage. **e**, **f**, Confocal micrographs of 72 hpf *Tg(kdrl:HRAS-mCherry)* WT and $egfl7^{s981/s981}$ larvae in lateral and dorsal views.

uninjected embryos, revealing the ability of the MO to inhibit *egfl7* mRNA translation. A widely recognized MO off-target effect is the transcriptional activation of *p53* (ref. 9). We thus measured *p53* expression by qPCR and observed no significant difference between embryos injected with 1 ng of MO and uninjected embryos. However, *p53* expression was clearly induced in embryos injected with 2 or 4 ng of MO (Extended Data Fig. 5). We next reasoned that if the *egfl7* MO did not induce off-target effects, it should not cause defects in *egfl7* null mutants. Thus, we injected embryos obtained from $egfl7^{s981/+}$ incrosses with 1 ng of *egfl7* MO. We subsequently selected and genotyped 32 embryos that showed a vascular phenotype, namely intersegmental vessel defects, reduced circulatory loop and/or pericardial oedema (Fig. 2a). Notably, we found that these embryos did not follow the Mendelian pattern observed in controls: 17 embryos were WT, 12 heterozygous and only 3 mutant (Fig. 2b, c), suggesting that the *egfl7* mutants were less sensitive than WT to MO injections. Confocal micrographs of WT, heterozygous and mutant embryos injected with 1 ng of *egfl7* MO (Fig. 2a) support this hypothesis. To investigate why some mutant embryos showed a phenotype when injected with *egfl7* MO, we repeated this experiment using a lower MO dose (0.5 ng). This experiment resulted in a clear reduction of the number of mutants in the selected population (1 mutant, 20 heterozygous and 21 WT fish out of 42 selected for vascular abnormalities, $P < 0.0001$). In the same experiment, out of ten WT-looking embryos, eight were mutant and two heterozygous for $egfl7^{s981}$ ($P = 0.0003$) (data not shown), supporting the observation that the mutant fish are less sensitive to *egfl7* MO and indicating that the *egfl7* MO has minimal off-target effects at these concentrations.

To investigate the differences between the mutant and morphant phenotypes further, we used an alternative knockdown approach and took advantage of the recently developed CRISPR interference (CRISPRi) technology[19] to inhibit *egfl7* transcript elongation. We designed two guide RNAs (gRNAs) to target the non-template strand of the 5′ untranslated region (UTR) and exon 2 of *egfl7* as well as one gRNA targeting the template strand of exon 2 (negative control) (Extended Data Fig. 6a). The relative *egfl7* expression levels were then quantified at 20 hpf using qPCR on pools of ten embryos injected with gRNAs and catalytically inactive (dead) CAS9 (dCAS9). Non-template gRNAs were able to inhibit *egfl7* transcript levels by approximately 60% compared with uninjected or template gRNA-injected embryos (Extended Data Fig. 6b). *Tg(kdrl:GFP)* embryos injected with gRNAs and dCAS9 exhibited different degrees of vascular abnormalities at 48 hpf, including intersegmental vessel defects, reduced circulatory loop and pericardial oedema after non-template but not template gRNA injections (Extended Data Fig. 6c). Altogether, these data show that transcriptional or translational knockdown of *egfl7* can lead to severe cardiovascular phenotypes while a severe genetic lesion does not.

To identify molecules underlying the different phenotypes observed in mutants versus morphants, we performed mass spectrometry and RNA profiling analyses in *egfl7* WT, homozygous mutant ($egfl7^{s981}$) and morphant embryos at 24 hpf. We assessed the proteomes by 4 h 'single shot' liquid chromatography–tandem mass spectrometry (LC–MS/MS) and identified more than 6,000 proteins with high reproducibility ($r > 0.90$ for biological and technical replicates between mutants and WT; Extended Data Fig. 7). To identify significant differences in individual protein expression, we used randomization-based false detection rate (FDR) estimation for multiple-testing correction and identified only one protein differentially expressed between mutants and WT (Fig. 3a). Strong upregulation (more than fivefold) was found for Emilin3a, suggesting its possible role in compensation. Additionally, we found no significant upregulation of Emilin3a in morphants compared with WT (Fig. 3b; Extended Data Fig. 8a), further highlighting Emilin3a as a possible compensating protein. Moreover, RNA sequencing (RNA-seq) and qPCR analyses indicated that not only *emilin3a* but also *emilin3b* and *emilin2a* were upregulated in mutants but not in morphants or CRISPRi injected embryos (Fig. 3c; Extended Data Fig. 8b). Interestingly, all these proteins contain an EMI domain, one of the key units of Egfl7 function[16], and, like Egfl7, can regulate elastogenesis[20,21]. We then reasoned that if Emilins are able to functionally replace Egfl7, they might rescue *egfl7* morphants. Embryos were injected with *egfl7* MO or co-injected with *egfl7* MO and *egfl7*, $egfl7^{s981}$, *Emilin2* or *Emilin3* mRNA and screened for circulatory loop defects at 48 hpf. Similarly to *egfl7* mRNA, *Emilin2* and *Emilin3* mRNAs were both able to rescue the circulatory defects in a significant proportion of *egfl7* morphants, while $egfl7^{s981}$ mRNA was not (Fig. 4). These results support the hypothesis that the upregulation of *emilin* genes in $egfl7^{s981}$ mutants is at least partly responsible for their lack of phenotype. To test whether the transcriptional changes we identified in mutants but not morphants were a peculiarity of the *egfl7* locus, we generated TALEN mutants for *vegfaa* (data not shown).

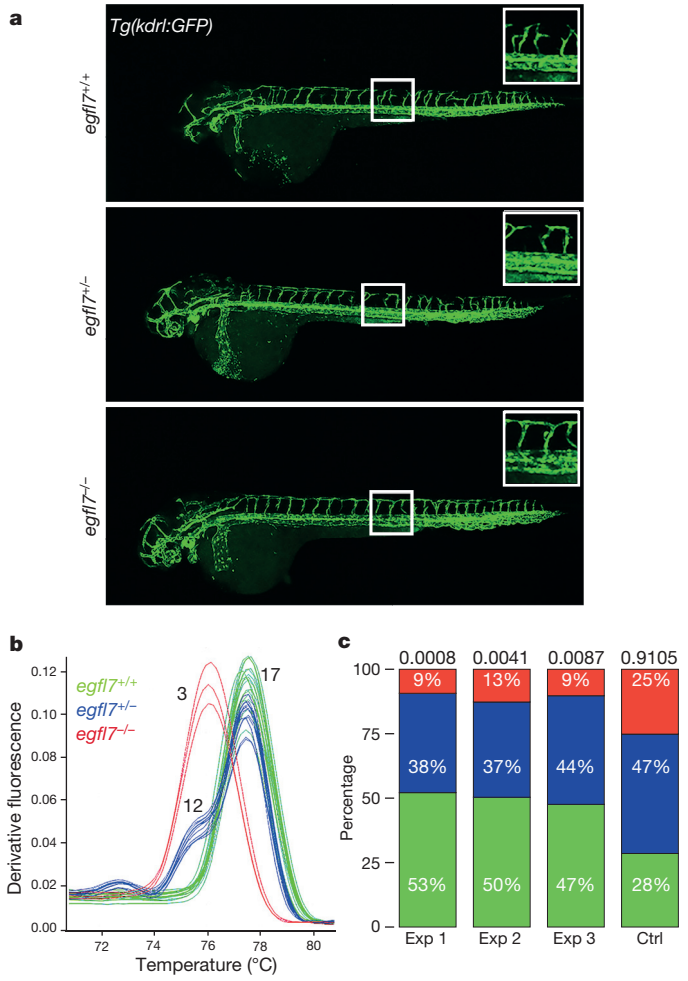


Figure 2 | Zebrafish *egfl7* mutant embryos are less sensitive to *egfl7* morpholino injections. **a**, Confocal micrographs of *Tg(kdrl:GFP)* WT, $egfl7^{s981/+}$ and $egfl7^{s981/s981}$ 48 hpf embryos injected with 1 ng of *egfl7* MO (AS_{-47}) in lateral views. **b**, High-resolution melt analysis genotyping example of 32 embryos (from an $egfl7^{s981/+}$ incross) selected for vascular defects at 48 hpf, showing the melting curves of 17 *egfl7* WT (green), 12 $egfl7^{s981/+}$ (blue) and 3 $egfl7^{s981/s981}$ (red) embryos. **c**, Genotype distribution (at 48 hpf) of $egfl7^{s981/+}$ incross progeny injected with 1 ng of *egfl7* MO at the one-cell stage and subsequently selected for the vascular phenotypes (independent experiments (Exp) 1, 2 and 3) or randomly selected (Ctrl). The population of randomly selected embryos follows the expected Mendelian ratio, but the phenotype-selected populations show significant skewing towards *egfl7* WT. *P* value represents two-tailed value for χ^2 test with two degrees of freedom; $n = 32$ embryos genotyped in each experiment. Note that the $egfl7^{s981/+}$ embryos are also underrepresented in the phenotype-selected populations (corresponding *P* values for experiments 1, 2 and 3 are 0.0033, 0.0066 and 0.032, respectively).

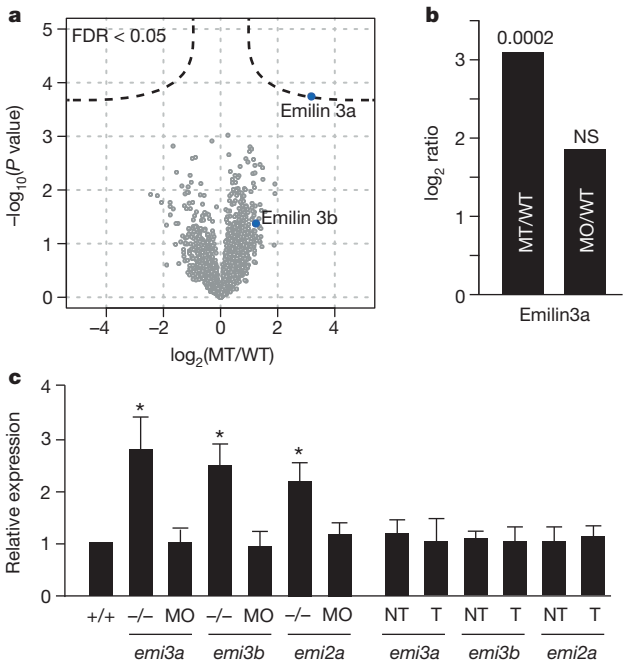


Figure 3 | Emilin3a is upregulated in mutant but not in morphant or CRISPRi embryos. **a**, Volcano plot showing significantly dysregulated proteins between 24 hpf *egfl7* WT and $egfl7^{s981}$ mutant embryos using label-free quantification. Emilin3a and Emilin3b are highlighted in blue. **b**, Morphants did not show a significant upregulation of Emilin3a in unbiased mass-spectrometry-based proteomics comparing mutants, WT and morphants. A two-sided *t*-test was used to assess *P* values and FDR was controlled by a randomization-based SAM approach. **c**, mRNA expression of *emilin3a*, *emilin3b* and *emilin2a* in *egfl7* WT, mutant, morphant and CRISPRi (template and non-template strand) embryos at 20 hpf; qPCR data, pools of 20–30 embryos each, expression normalized to *gapdh* (WT expression set at 1 for each gene). The *emilin* genes were upregulated in $egfl7^{s981}$ mutants but not after translational or transcriptional inhibition. $^*P \leq 0.05$.

Interestingly, qPCR analysis showed that *vegfab*, a paralogue of *vegfaa*, was upregulated in mutants but not morphants (Extended Data Fig. 9a). Additionally, blocking Vegfaa function using a dominant negative approach also failed to trigger *vegfab* upregulation, placing the signal triggering compensation upstream of protein function (Extended Data Fig. 9b).

Concerns have been raised over the use of antisense reagents, including MOs, as they may cause off-target effects and lead to aberrant conclusions. This debate was recently revived by the generation of mutations in many genes whose function was previously studied using MOs; strikingly, a majority of the resulting mutants exhibit a different phenotype from the one reported for the corresponding morphants; in fact, most often the mutants exhibit no obvious phenotype[1]. In our study, we show that, at least for some genes, the phenotypic differences between mutants and morphants can be due to the activation of genetic compensation in the former but not the latter. We show here that the upregulation of Emilins can compensate for the loss of Egfl7, but anticipate that other genes are involved in this process. On the basis of our data, we propose two additional recommendations for using MOs: the first is to use doses that do not induce *p53* expression as in many cases this induction indicates off-target effects; the second is to titrate the MO dose so that it does not cause additional phenotypes in a null mutant background, as such phenotypes would be due to non-specific effects. The mechanisms underlying the compensation observed in mutants but not in morphants are likely to be complex and so will their investigation. Interestingly, we observed no upregulation of the *emilin* genes in the *Δ3* (*s980*) allele, suggesting that a non-deleterious genomic lesion is not sufficient to trigger this response. On the other hand, we observed *emilin* gene upregulation in embryos injected with *egfl7* TALENs, indicating that a deleterious mutation does not need to go through the germline to trigger this response. We also detected *emilin* gene upregulation in embryos carrying only one $egfl7^{s981}$ mutant allele (data not shown). This observation might explain the partial protection of heterozygous embryos from *egfl7* MO injections (Fig. 2).

In summary, our data show that, for *egfl7*, one can identify a dose of MO that has no effect in most *egfl7* mutant embryos but causes clear vascular defects in WT, indicating that these morphant phenotypes are not due to off-target effects. Further, *egfl7* mutants show no phenotypes but exhibit a clear upregulation of several members of the *emilin* gene family. These Emilin proteins share an important functional domain with Egfl7, and, probably with additional proteins, can compensate for the loss of Egfl7 function. Notably, a recent study of the Icelandic population identified individuals with homozygous loss-of-function mutations in *EGFL7* (ref. 22), indicating that compensation

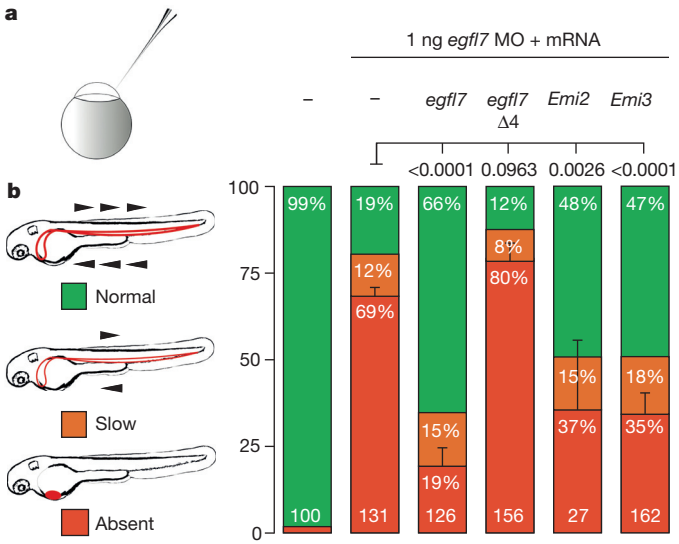


Figure 4 | *Emilin2* and *Emilin3* can rescue *egfl7* morphants. a, Design of the rescue experiment. One nanogram of *egfl7* MO was injected in WT embryos, alone or together with 400 pg of mRNA (*egfl7* WT, *egfl7 Δ4*, *Emilin2* or *Emilin3*). **b**, Injected embryos were sorted according to their circulatory loop phenotype into three classes: normal, slow and absent circulation. Injection of *egfl7* MO resulted in 69% of embryos lacking circulation at 48 hpf. This percentage was reduced to 19% when co-injecting *egfl7* mRNA, and to 37% and 35% when co-injecting *Emilin2* and *Emilin3* mRNA, respectively. In contrast, mRNA from the *egfl7 Δ4* mutant allele did not rescue (80% of embryos lacked circulation). Uninjected siblings are shown for comparison (99% normal). Number at the bottom of each bar is the total number of embryos from two independent experiments. Error bars, s.e.m. for the 'absent circulation' class. *P* value represents two-tailed value for Fisher's exact test.

for severe lesions at this locus might also be at play in humans. It will be interesting to determine whether the upregulation of *EMILIN* genes is also present in these individuals. Of course, detailed studies will be needed to determine whether such compensation is the reason for the phenotypic differences between mutants and morphants for other genes. More importantly, our study illustrates the power of comparing mutants and morphants to identify modifier genes, a goal that remains a major challenge in the field of human genetics.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 29 December 2014; accepted 22 May 2015.
Published online 13 July 2015.

1. Kok, F. O. *et al.* Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. *Dev. Cell* **32,** 97–108 (2015).
2. Stainier, D. Y., Kontarakis, Z. & Rossi, A. Making sense of anti-sense data. *Dev. Cell* **32,** 7–8 (2015).
3. Law, S. H. & Sargent, T. D. The serine-threonine protein kinase PAK4 is dispensable in zebrafish: identification of a morpholino-generated pseudophenotype. *PLoS ONE* **9,** e100268 (2014).
4. Daude, N. *et al.* Knockout of the prion protein (PrP)-like *Sprn* gene does not produce embryonic lethality in combination with PrP(C)-deficiency. *Proc. Natl Acad. Sci. USA* **109,** 9035–9040 (2012).
5. De Souza, A. T. *et al.* Transcriptional and phenotypic comparisons of *Ppara* knockout and siRNA knockdown mice. *Nucleic Acids Res.* **34,** 4486–4494 (2006).
6. Smart, N. & Riley, P. R. Thymosin β4 in vascular development, response to research commentary. *Circ. Res.* **112,** 29–30 (2013).
7. Gao, Y. *et al.* Auxin binding protein 1 (ABP1) is not required for either auxin signaling or *Arabidopsis* development. *Proc. Natl Acad. Sci. USA* **112,** 2275–2280 (2015).
8. Anon. RNA interference on target. *Nature Methods* **3,** 659 (2006).
9. Robu, M. E. *et al.* p53 activation by knockdown technologies. *PLoS Genet.* **3,** e78 (2007).
10. Schmidt, M. *et al.* EGFL7 regulates the collective migration of endothelial cells by restricting their spatial distribution. *Development* **134,** 2913–2923 (2007).
11. Kuhnert, F. *et al.* Attribution of vascular phenotypes of the murine *Egfl7* locus to the microRNA *miR-126*. *Development* **135,** 3989–3993 (2008).
12. Parker, L. S. *et al.* The endothelial-cell-derived secreted factor Egfl7 regulates vascular tube formation. *Nature* **428,** 754–758 (2004).
13. Charpentier, M. S. *et al.* CASZ1 promotes vascular assembly and morphogenesis through the direct regulation of an EGFL7/RhoA-mediated pathway. *Dev. Cell* **25,** 132–143 (2013).
14. Huang, C. *et al.* VE-statin/Egfl7 siRNA inhibits angiogenesis in malignant glioma *in vitro*. *Int. J. Clin. Exp. Pathol.* **7,** 1077–1084 (2014).
15. Cermak, T. *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **39,** e82 (2011).
16. Nichol, D. & Stuhlmann, H. EGFL7: a unique angiogenic signaling factor in vascular development and disease. *Blood* **119,** 1345–1352 (2012).
17. Chi, N. C. *et al.* Foxn4 directly regulates *tbx2b* expression and atrioventricular canal formation. *Genes Dev.* **22,** 734–739 (2008).
18. Jin, S. W., Beis, D., Mitchell, T., Chen, J. N. & Stainier, D. Y. Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. *Development* **132,** 5199–5209 (2005).
19. Larson, M. H. *et al.* CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols* **8,** 2180–2196 (2013).
20. Zanetti, M. *et al.* EMILIN-1 deficiency induces elastogenesis and vascular cell defects. *Mol. Cell. Biol.* **24,** 638–650 (2004).
21. Lelievre, E. *et al.* VE-statin/egfl7 regulates vascular elastogenesis by interacting with lysyl oxidases. *EMBO J.* **27,** 1658–1670 (2008).
22. Sulem, P. *et al.* Identification of a large set of rare complete human knockouts. *Nature Genet.* **47,** 448–452 (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank H.-B. Kwon and other members of the laboratory, past and present, as well as K. Sampath, D. Wainstock, C. Moens and M. Grether, for discussions, comments on the manuscript and/or reagents, and the Max Planck Society, Packard foundation, and EMBO for funding.

Author Contributions All authors were involved in the experimental design, data analysis and writing. Experiments were performed by all except M.K. and D.Y.R.S., who also supervised the project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.Y.R.S. (didier.stainier@mpi-bn.mpg.de).

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment, except for the data shown in Fig. 2 where the inherent design of the experiment includes a blinding component.

Zebrafish handling. All zebrafish husbandry was performed under standard conditions in accordance with institutional (UCSF and MPG) and national ethical and animal welfare guidelines.

Confocal microscopy. An LSM 700 confocal laser scanning microscope (Zeiss) was used for live imaging. Embryos and larvae were anaesthetized with a low dose of tricaine, placed in a glass-bottomed Petri dish (MatTek) with a layer of 1.2% low melt agarose and imaged using Plan-Apochromat 10×/0.45 and LCI Plan-Neofluar 25×/0.8 objective lenses. Vessel integrity and permeability were analysed using micro-angiography. Fluorescein isothiocyanate (FITC)-dextran, 2,000 kDa (Sigma) was injected into the posterior cardinal vein at 48 or 72 hpf and imaged after 10 min.

Plasmids. Total RNA extraction was performed using TRIZOL (Life Technologies) and used for cDNA synthesis using SuperScript second strand (Life Technologies). cDNAs encoding the Egfl7 WT, Egfl7^{s980} and Egfl7^{s981} proteins were PCR-amplified using whole embryo cDNA as template. PCR fragments were ligated into the mammalian expression vector pcDNA3.1 myc-HIS tag between EcoRI and XhoI. All constructs were verified by sequencing. pCMV-6 plasmids containing mouse elastin microfibril interfacer 2 (*Emilin2*) or mouse elastin microfibril interfacer 3 (*Emilin3*) cDNAs were purchased from Origene. pcDNA3.1 and pCMV-6 plasmids were respectively linearized using SmaI and AgeI and *in vitro* transcribed using the mMESSAGE mMACHINE T7 kit (Life Technologies).

Cell culture and transfection. Authenticated HEK293FT (human embryonic kidney) (Life Technologies, R700-07) and HUVEC (human umbilical vein endothelial cells) cells were cultured at 37 °C in 5% CO_2, 95% air in appropriate medium containing 10% fetal bovine serum, 100 units per millilitre penicillin and 100 μg ml^{-1} streptomycin. HEK293FT cells were used for biochemical studies because they are easy to grow and transfect, and have been used widely for cell biology research. All cell lines are routinely tested for mycoplasma in our facilities, and only mycoplasma-free cells were used in this study. Cells were transfected with cDNAs in antibiotic-free medium 12–24 h before protein extraction, using FuGene HD (Roche) at a 3:1 ratio (μl:μg nucleic acid) and 0.18 μg DNA per square centimetre. Cells were lysed in RIPA buffer and extracellular proteins precipitated with TCA (final 20%). Samples were resuspended in Laemmli buffer before gel electrophoresis.

Genome editing. TALENs were designed targeting *egfl7* (Extended Data Fig. 1) using TALEN targeter[23] (https://tale-nt.cac.cornell.edu/) and constructed using Golden Gate assembly[15]. Zebrafish embryos were injected into the cell at the one-cell stage with 100 pg total TALEN RNA.

Genome engineering was performed as previously described[24]. The C terminus *egfl7* TALEN recognition sites are TGCTGGTAGACATCATC and TTGCAGTA GTGACTAGT. Between the binding sites is a 17-bp spacer with the 'tag' stop codon underlined (aggaaaactagacgatc). A ssDNA oligonucleotide (Supplementary Table 1) was designed to target the spacer sequence between the cutting sites. A Myc-tag sequence, flanked by XhoI restriction sites, was introduced in the centre of the oligonucleotide resulting in 25-base homology arms on the 5′ and 3′ ends. Polyacrylamide gel electrophoresis (PAGE)-purified oligonucleotides were obtained from Sigma. One-cell-stage embryos were injected with 100–200 pg TALEN mRNA and 75 pg ssDNA donor. Screening for founders was conducted using PCR followed by XhoI restriction enzyme digest and subsequently by sequencing.

CRISPR interference. gRNA and CAS9 plasmids[25] were purchased from Addgene. Dead Cas9 was generated using the zebrafish-codon-optimized WT CAS (pT3TS–nls–zCas9–nls; nls, nuclear localization signal) as a template. The D10A and H840A mutations were generated using the primers in Supplementary Table 1. Site-directed mutagenesis was performed using PfuUltra Fusion HS (Agilent). The PCR reaction protocol was 95 °C for 1 min, then 18 cycles of 95 °C for 50 s, 60 °C for 50 s and 68 °C for 1 min per kilobase of plasmid length, then 68 °C for 7 min and 4 °C hold. Dpn1 (1 μl) was added to the PCR reaction and incubated at 37 °C for 1 h to digest parental DNA and transform into competent cells. CRISPR gRNAs were designed using CRISPR design (http://crispr.mit.edu/) (Zhang laboratory). Oligonucleotides were annealed in a thermo block at 90–95 °C for 5 min followed by a slow cooling to room temperature (~20 °C) and cloned in gRNA plasmid between BsmbI sites. All constructs were verified by sequencing. To make nls–zCas9–nls RNA, the template DNA (pT3TS–nls–zCas9–nls) was linearized by XbaI digestion and purified using a QIAprep column (Qiagen). Capped nls–zCas9–nls RNA was synthesized using a mMESSAGE mMACHINE T3 kit (Life Technologies) and purified using an RNA Clean and Concentrator kit (Zymo Research). To make gRNA, the template DNA was linearized by BamHI digestion and purified using a QIAprep column. gRNA was generated by *in vitro* transcription using a T7 RNA polymerase MEGA short script T7 kit (Life Technologies). After *in vitro* transcription, the gRNA (approx 100 nucleotides long) was purified using RNA clean and concentrator (Zymo Research). dCAS9 mRNA (100–400 pg) and gRNA (50–100 pg) were co-injected in the cell at the one-cell stage and at least ten pooled embryos were used to evaluate the expression level of the targeted genes by qPCR. Initial experiments were performed with gRNAs targeting the *tnnt2a* gene. In general, a substantial increase in knockdown efficiency was observed when combining multiple guides, indicating a synergistic effect. The *egfl7* gene was knocked down by using two to four gRNAs (Supplementary Table 1).

Microinjection of morpholinos. The ATG morpholinos *egfl7* (5′-CAGGTGT GTCTGACAGCAGAAAGAG-3′), *vegfaa* (5′-GTATCAAATAAACAACCAA GTTCAT-3′) and *tp53* (5′-GCGCCATTGCTTTGCAAGAATTG-3′), were purchased from GeneTools and injected at the indicated amounts (0.5, 1, 2 or 4 ng for *egfl7*, 2 ng for *vegfaa* and 1 ng for *tp53*). To identify the potential effects of *p53* induction in *egfl7* morphants, we compared the phenotype of embryos co-injected with *egfl7* and *p53* MO or *egfl7* MO alone (1 ng for each MO), and did not detect any obvious differences. The experiments testing the *egfl7* MO effect on *egfl7* mutants were blind (injection into fertilized eggs from an incross of heterozygous fish followed by phenotyping and then genotyping). The *egfl7* morphant rescue experiments were not blind. Sample sizes for these and other experiments were determined on the basis of previous experience.

Genotyping. Embryos or fin-clips were placed in PCR tubes, with 50 μl of elution buffer (10 mM Tris-Cl, pH 8.5) and 1 mg ml^{-1} proteinase K added to each well and then incubated at 55 °C for 2 h. The samples were then heated to 95 °C for 10 min to inactivate proteinase K. Primers were designed using primer3: http://biotools.umassmed.edu/bioapps/primer3_www.cgi. An Eco Real-Time PCR System (Illumina) was used for the PCR reactions and high-resolution melt analysis. DyNAmo SYBR green (Thermo Fisher Scientific) was used in these experiments. PCR reaction protocols were 95 °C for 15 s, then 40 cycles of 95 °C for 2 s, 60 °C for 2 s and 72 °C for 2 s. Following the PCR, a high-resolution melt curve was generated by collecting SYBR-green fluorescence data in the 65–95 °C range. The analyses were performed on normalized derivative plots.

Electrophoresis. Laemmli SDS–PAGE gels consisted of a 4–20% running gel and 3% stacking gel or Tricine SDS–PAGE 12%. For immunoblots, membranes were blocked with 5% non-fat milk and incubated at 4 °C overnight with mouse (SC-40) or rabbit (SC-789) anti-Myc antibody (Santa Cruz Biotechnology), or anti-α-tubulin (T9026, Sigma). Membranes were then rinsed, incubated for 1 h with horseradish-peroxidase-conjugated anti-rabbit immunoglobulin-G (IgG) or anti-mouse IgG (Santa Cruz Biotechnology), rinsed extensively, and labelled proteins were detected using the Clarity Western substrate (Biorad).

Mass spectrometry. Embryos (*egfl7* WT, *egfl7*^{s981} mutants and morphants) at 20–24 hpf were lysed in 6 M urea and 2 M thiourea (in HEPES buffer pH = 8.5). The lysate was clarified by centrifugation and proteins were subjected to in-solution digestion. In brief, proteins were reduced (0.1 M DTT, 30 min at room temperature) and alkylated (55 mM IAA, 30 min at room temperature in the dark). Lys-C was added in a 1:100 enzyme:protein ratio and incubated for 3 h. Urea concentration was diluted to 2 M using 50 mM ammonium bicarbonate, and trypsin (Promega) was added in a 1:100 enzyme:protein ratio. After incubation for 18 h, generated peptides were de-salted using the Stop and Go Extraction tip technology before mass spectrometric analysis. All WT and *egfl7*^{s981} mutant experiments were performed at least in technical and biological duplicates. For morphant embryos, we measured protein changes in technical triplicates after pooling more than 20 embryos, thus reducing biological variability. The instrumentation for LC–MS/MS analysis consisted of a nano LC 1000 (Proxeon, now Thermo Scientific) coupled via a nano-electrospray ionization source to a quadrupole-based benchtop QExactive Plus or QExactive mass spectrometer. Separation of peptides according to their hydrophobicities was achieved on a 50 cm in-house packed column (internal diameter 75 μm, C18 Beads (Dr. Maisch) diameter 1.8 μm) using a binary buffer system: (A) 0.1% formic acid in H_2O and (B) 0.1% formic acid in 80% acetonitrile. A linear gradient within 220 min from 8% to 38% of B, followed by an exponential increase to 90% B and a re-equilibration step to 5% B within 20 min, was used for peptide elution. Mass spectra were acquired at a resolution of 70,000 (200 m/z) using an AGC target of 1E6 and a maximal injection time of 20 ms. A top ten method was applied for subsequent acquisition of high-energy collision-induced dissociation (HCD) fragmentation MS/MS spectra of the ten most intense peaks. Resolution was set to 17,500 at 200 m/z and 5E5 ions (AGC target) were collected in the C-trap within a maximal injection time of 60 ms using an isolation window of 1.3 thomsons (Th) (1 Th = 1.036426×10^{-8} kg C^{-1})[26]. Raw files were processed using MaxQuant 1.4.1.2 (ref. 27) and the implemented

Andromeda search engine[28]. For peptide assignment, MS/MS HCD fragmentation spectra were correlated to the Uniprot *Danio rerio* database (2014). A list of common contaminants was included in the searches that were performed with tryptic specificity. Default settings were used for MS and MS/MS mass tolerances and peptide length. The FDR was set to 1% on protein and peptide levels and estimated by the implemented decoy algorithm. Oxidation of methionine residues and acetylation on the protein N-term were set as variable modifications, and carbamidomethyl at cysteine residues was defined as a fixed modification. The match-between-runs, label-free quantification and re-quantification were enabled. Statistical analysis and data visualization were done in the environment R.

The package *siggenes* from Bioconductor was used to determine significance of proteome changes at a FDR cutoff of less than 0.05 (ref. 29). Note that the protein list was filtered for at least 50% quantification over all experiments. In label-free protein quantification, a common problem is that low abundance proteins are likely to be not quantifiable, leading to a right-shifted normal distribution. Thus, missing values were replaced along a Gaussian distribution using a $\log_2$ downshift of 1 and a width of 0.4. Imputation was inspected by histograms to mimic a Gaussian distribution for the complete data set (columnwise) to avoid too high a frequency of low-intensity values. Significance was assessed by a two-sided *t*-test of $\log_2$ intensity values. Note that to compare morphants with WT, we used technical replicates of pooled embryos against all experiments for WT embryos. Five hundred randomizations were used to estimate FDR, using a cutoff of 0.05 while s0 (the fudge factor) was defined as 0.1. Protein ratios were calculated by subtracting the average of the respective groups. Data are shown in Supplementary Table 2.

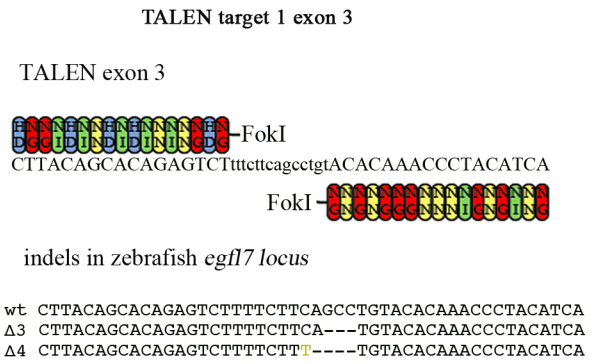
RNA profiling. Total RNA from *egfl7* WT, *egfl7*^{s981} mutants and morphants at 24 hpf was prepared using TRIzol (Life Technologies). RNA profiling was performed by ZF-screens using an Illumina HiSeq 2500 ultra-high-throughput sequencing system. Data are shown in Supplementary Table 3.

qPCR. An Eco Real-Time PCR System (Illumina) was used for qPCR experiments. Gene expression was normalized relative to *gapdh*. All reactions were performed in technical triplicates; the results represent biological triplicates (unless otherwise stated) including the s.e.m. Supplementary Table 1 shows the primers used for these experiments.

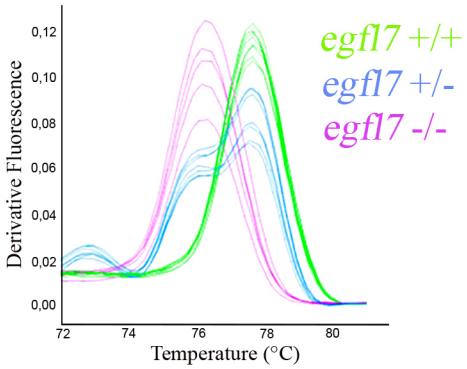
Additional data. Mature miR126 levels were quantified in WT and *egfl7*$^{s981/s981}$ embryos using the miRNA QRT–PCR Detection Kit (Agilent). No significant changes were observed at 48 or 72 hpf. Maternal zygotic *egfl7*^{s981} mutant embryos were generated by incrossing homozygous mutant adults; they exhibited no additional phenotypes compared with zygotic mutant embryos. In addition, we observed no evidence of maternal *egfl7* mRNA by RT–PCR. Mutant samples for proteomics, RNA-seq and qPCR analyses were MZ mutants.

23. Doyle, E. L. *et al.* TAL effector specificity for base 0 of the DNA target is altered in a complex, effector- and assay-dependent manner by substitutions for the tryptophan in cryptic repeat -1. *PLoS ONE* **8,** e82120 (2013).
24. Bedell, V. M. *et al.* In vivo genome editing using a high-efficiency TALEN system. *Nature* **491,** 114–118 (2012).
25. Jao, L. E., Wente, S. R. & Chen, W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc. Natl Acad. Sci. USA* **110,** 13904–13909 (2013).
26. Nolte, H. *et al.* Global protein expression profiling of zebrafish organs based on in vivo incorporation of stable isotopes. *J. Proteome Res.* **13,** 2162–2174 (2014).
27. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnol.* **26,** 1367–1372 (2008).
28. Cox, J. *et al.* Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10,** 1794–1805 (2011).
29. Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl Acad. Sci. USA* **98,** 5116–5121 (2001).

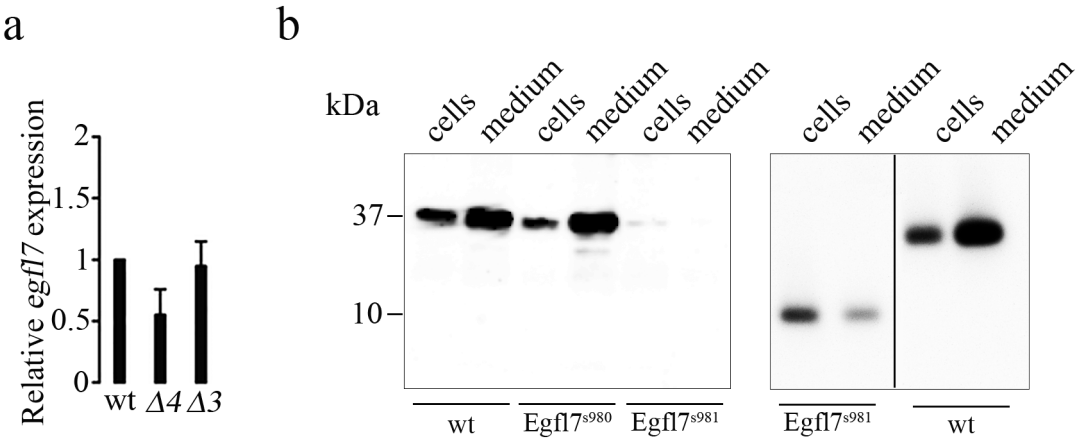
a



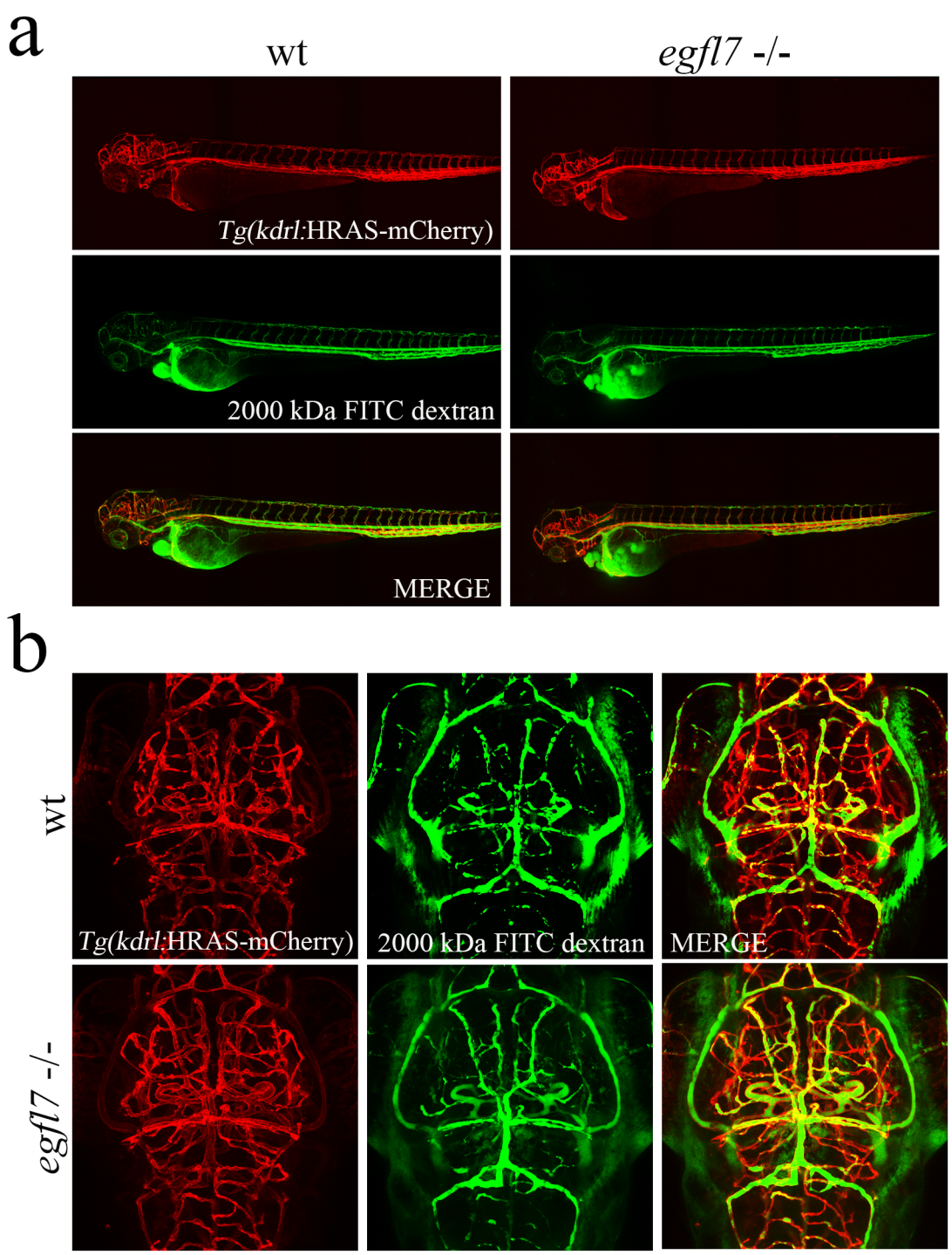
b



Extended Data Figure 1 | Generation and identification of zebrafish *egfl7* mutant alleles. **a**, TALENs were designed to target exon 3 of *egfl7* which encodes part of the EMI domain. Sequence alignment of part of exon 3 from the WT, *egfl7*^{s980} and *egfl7*^{s981} alleles shows TALEN indels: *Δ3/s980* (three nucleotide deletion) and Δ4/s981 (five nucleotide deletion), and one nucleotide insertion (yellow). **b**, Genotyping example of single embryos sampled from a population of *egfl7* WT, *egfl7* $^{s981/+}$ and *egfl7*$^{s981/s981}$ fish using high-resolution melt analysis. The green curve corresponds to the WT allele and the red one to the *egfl7*^{s981} allele. Heterozygous embryos have both alleles and thus the melting profile (in blue) is a composition of the WT and mutant curves.

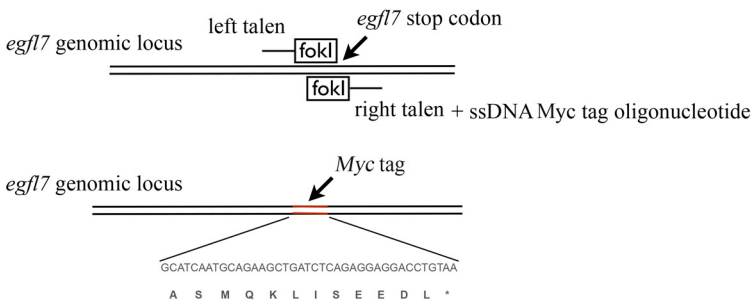


Extended Data Figure 2 | The *egfl7*^{s981} mutation leads to *egfl7* mRNA degradation, reduced protein expression and impaired secretion. a, The *egfl7*^{s981} mutation leads to *egfl7* mRNA degradation: *egfl7* mRNA expression in 24 hpf WT, *egfl7*$^{s981/s981}$ and *egfl7*$^{s980/s980}$ embryos. Expression normalized to *gapdh*. **b**, The *egfl7*^{s981} (p.Gln49Leufs*30) mutation leads to strongly reduced protein expression. Western blot analyses of Egfl7-Myc-tag expression in transfected HUVEC cells. Egfl7 WT and Egfl7^{s980} protein expression was strongly detected in the medium whereas the Egfl7^{s981} isoform was strongly reduced in the cells and very poorly secreted (right), or undetectable in both (left). Furthermore, Egfl7^{s981} shares high similarity to the truncated protein produced in the original *Egfl7* mutant mouse in which the protein was not detectable using an Egfl7 antibody[10].

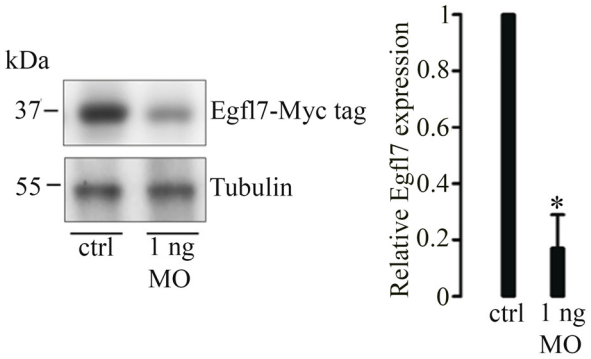


Extended Data Figure 3 | Vessel integrity and permeability do not appear to be affected in *egfl7*$^{s981/s981}$ larvae. A fluorescent molecule (2000 kDa FITC-dextran) was injected directly into the circulation of 72 hpf *Tg(kdrl:HRAS:mCherry)* larvae that previously showed haemorrhage, which was mostly localized around the hindbrain ventricle. Confocal micrographs of 72 hpf *Tg(kdrl:*HRAS:mCherry*)* expression, FITC-dextran and MERGE of WT and *egfl7*$^{s981/s981}$ larvae in (**a**) lateral and (**b**) dorsal views. The FITC-dextran did not accumulate to the sites of haemorrhage, suggesting that these sites had clotted and vascular integrity had been restored after the initial blood leakage.

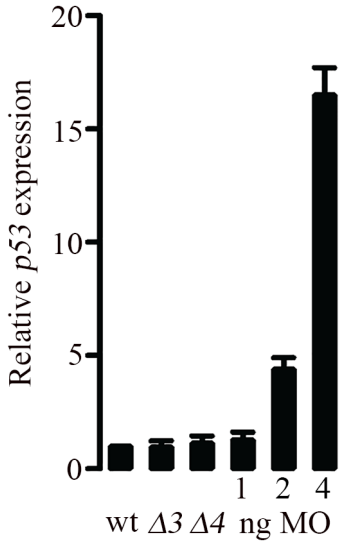
a



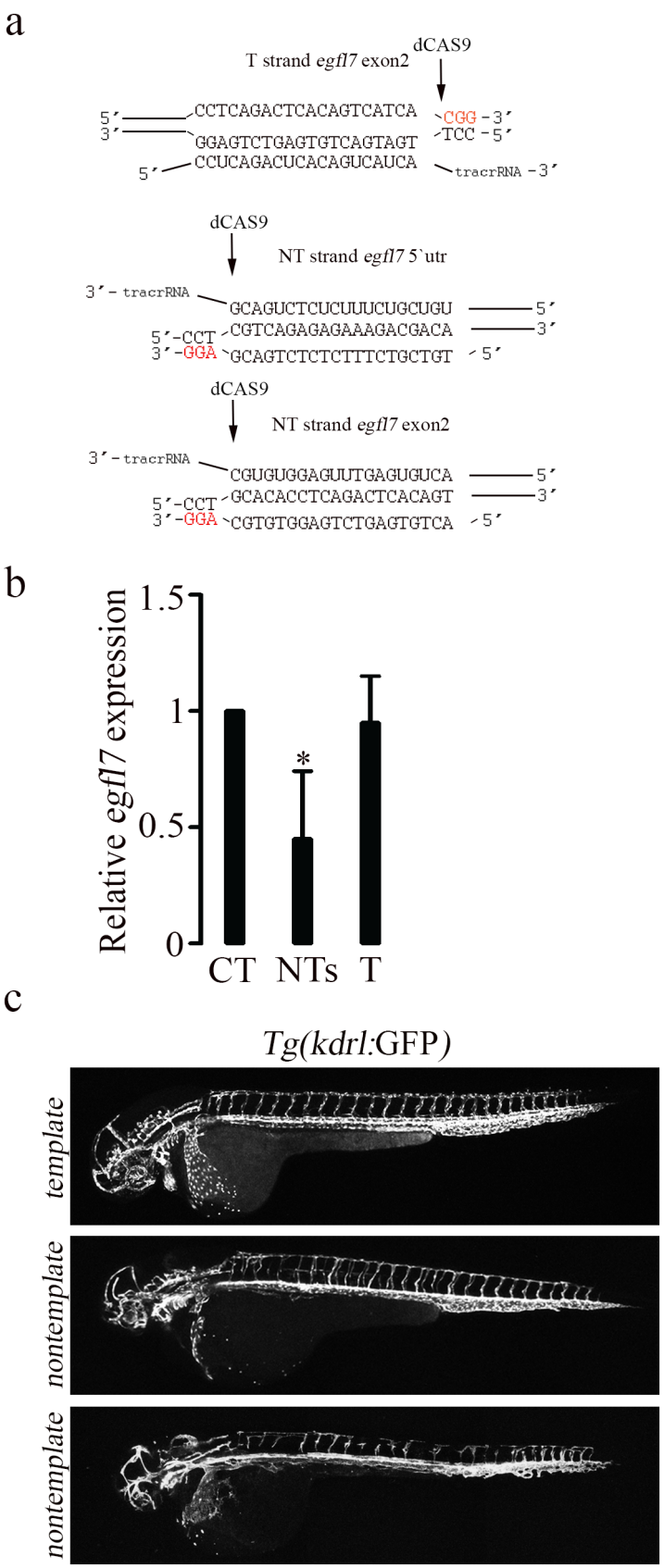
b



Extended Data Figure 4 | *In vivo* genome editing: *Myc*-tag introduction in the *egfl7* endogenous locus. **a**, TALENs targeting the *egfl7* stop codon created double-stranded breaks in the chromosomal DNA. Homology-directed repair precisely incorporated the Myc tag exogenous sequence (ssDNA) at the cut site. **b**, Western blot analysis of Egfl7-Myc-tag expression in 24 hpf control and morphant embryos. Egfl7 Myc-tag signal was reduced by around 80% in morphants (1 ng *egfl7* MO) compared with uninjected. Expression normalized to tubulin ($P \leq 0.05$). Error bars, s.e.m. ($n = 3$).

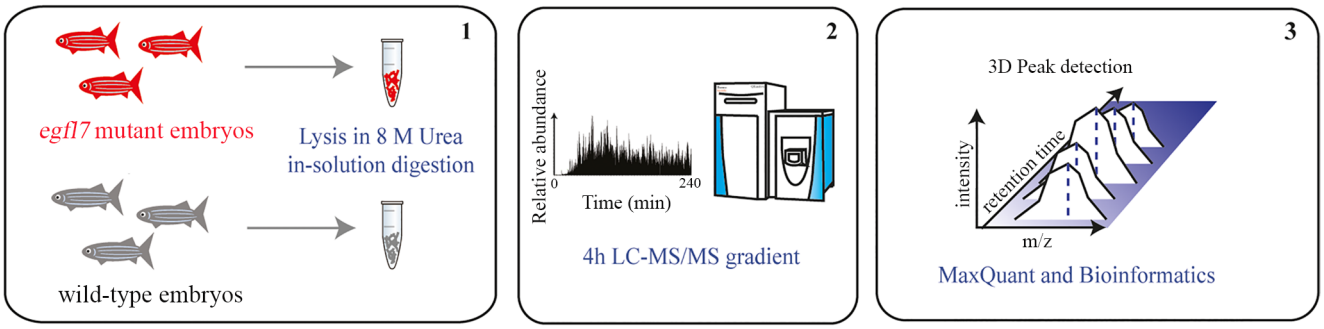


Extended Data Figure 5 | The *egfl7* morpholino does not significantly affect *p53* mRNA expression at 1 ng per embryo but it does so at higher doses. mRNA expression of *p53* in 24 hpf WT, *egfl7 Δ3* (*egfl7*^{s980}) and *egfl7 Δ4* (*egfl7*^{s981}) mutant, and morphant (1, 2 and 4 ng injected) embryos. Expression normalized to *gapdh*. Error bars, s.e.m. of technical triplicates.

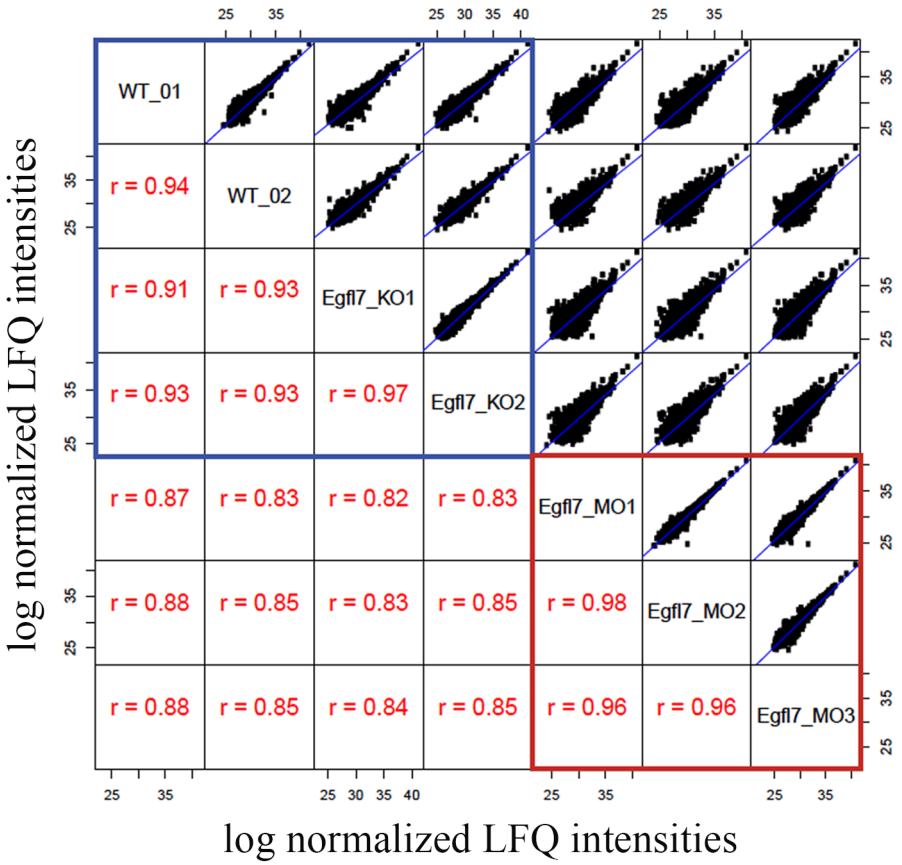


Extended Data Figure 6 | The *egfl7* transcript elongation inhibition causes a phenotype similar to the one seen in morphants. a, gRNAs of *egfl7* targeting the template (T) strand in exon 2 and non-template (NT) strand in the 5′ UTR and exon 2. **b**, Expression of *egfl7* in non-template (NT) gRNA and template (T) gRNA-injected embryos relative to uninjected (CT) siblings at 20 hpf. qPCR data, pools of ten embryos each, expression normalized to *gapdh* ($P \leq 0.05$). Error bars, s.e.m. ($n = 3$). **c**, Lateral view confocal micrographs of 48 hpf *Tg(kdrl:GFP)* embryos injected with *egfl7* template and non-template CRISPRi. Template CRISPRi (top) embryos are indistinguishable from non-injected siblings, while non-template CRISPRi embryos exhibit different degrees of vascular defects (middle: mild; bottom: severe).

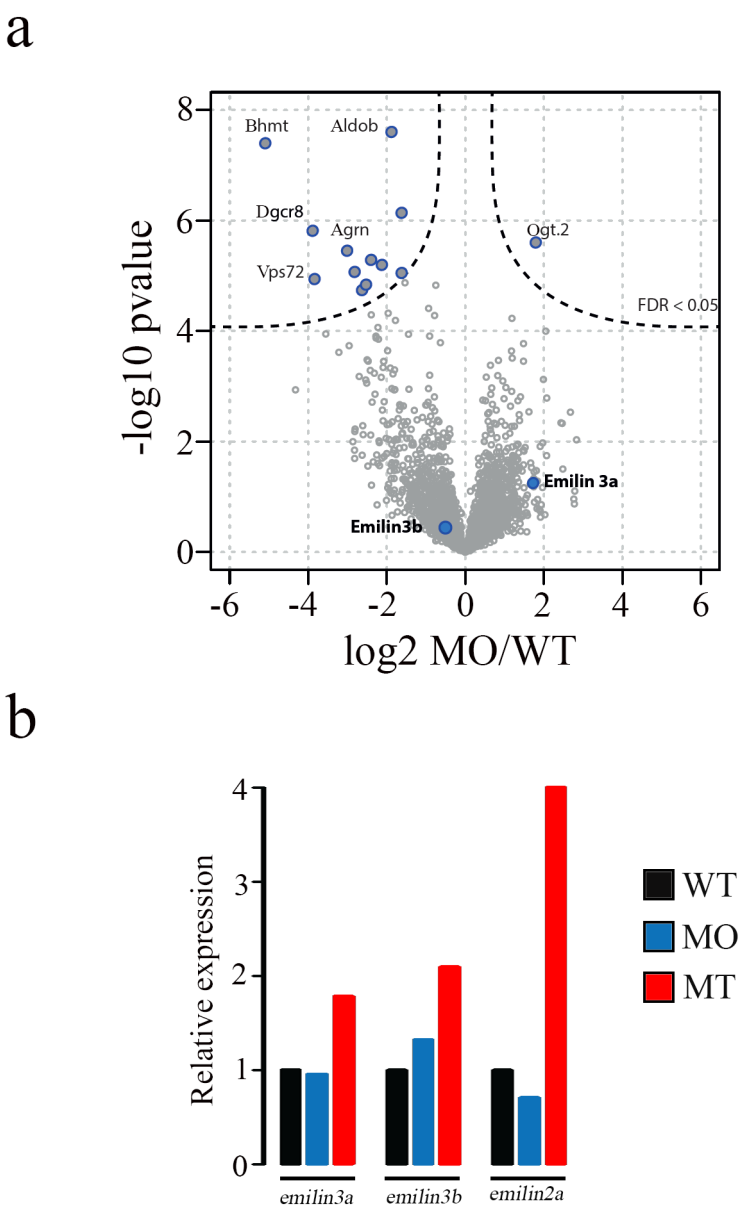
a



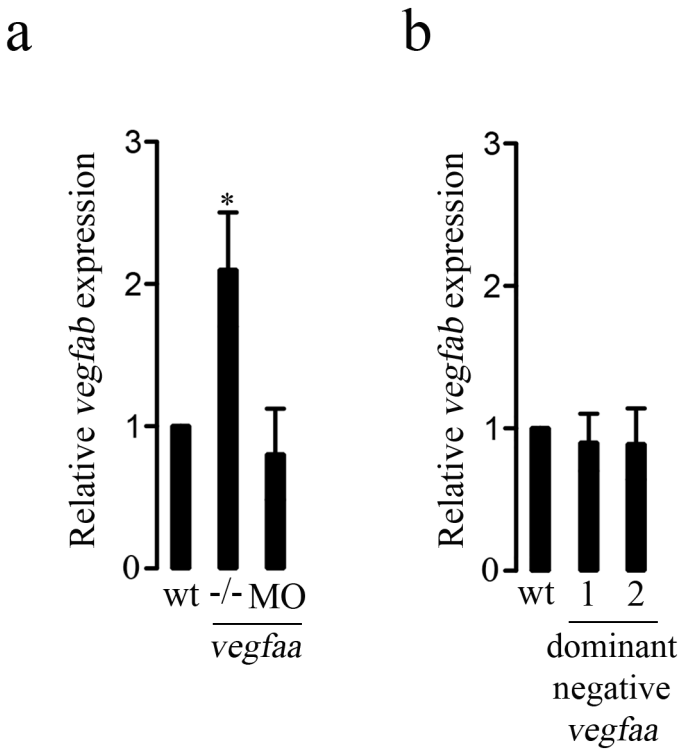
b



Extended Data Figure 7 | Single-shot proteomics to assess changes between WT and *egfl7*^{s981} mutant embryos. **a**, Schematic visualization of proteomic workflow. Embryos were lysed in urea buffer, and proteins were digested in-solution using trypsin and measured on a QExactive bench top instrument. Acquired spectra were analysed against the Uniprot zebrafish database (2014) using MaxQuant. **b**, Scatter plot matrix shows high correlation between biological replicates. Reproducibility was determined by a Pearson correlation coefficient.



Extended Data Figure 8 | Emilin3a expression is upregulated in mutant but not morphant embryos. **a**, Volcano plot showing significantly dysregulated proteins between *egfl7* morphant and WT embryos at 24 hpf using label-free quantification. Emilin3a (blue) levels were not significantly different between morphant and WT embryos. Emilin3b is also highlighted in blue. **b**, Bar plot showing upregulation of *emilin* family members in 24 hpf *egfl7* mutants compared with WT and morphants, as assessed from RNA-seq data (WT expression set at 1 for each gene).



Extended Data Figure 9 | Expression of *vegfab* is upregulated in *vegfaa* mutant embryos but not in morphants, or *vegfaa* dominant negative-injected embryos; qPCR data, pools of ten embryos each, expression normalized to *gapdh* ($P \leq 0.05$). Error bars, s.e.m. ($n = 5$). **a**, mRNA expression of *vegfab* in 24 hpf *vegfaa* WT, mutant and morphant embryos. **b**, mRNA expression of *vegfab* in 24 hpf *vegfaa* WT and *vegfaa* dominant negative-injected embryos (two different dominant negatives were injected).