eLife

Reviewed Preprint
v1 • June 9, 2026
Not revised

For correspondence:
rsemple@exseed.ed.ac.uk
e.patton@ed.ac.uk

Competing interests: RKS has received speaker fees and consulted for Novartis. RRM has received consulting fees from Nested Therapeutics (Cambridge, U.S.) and serves on the Scientific Advisory Board of CLOVES Syndrome Community. EEP is funded by QBiotics Group (Brisbane, Australia) for research not related to CLOVES.

Funding: See page 22

Reviewing editor: Suk-Won Jin, Gwangju Institute of Science and Technology, Republic of Korea

PIK3CA-related overgrowth spectrum (PROS) zebrafish models reveal pan-lineage developmental dysregulation

Hannah Brunsdon[1,2,3], Nuoya Wang[4], Micha Sam Brickman Raredon[4], Ralitsa R Madsen[5], Robert K Semple[1,3] ✉, E Elizabeth Patton[1,2] ✉

[1]MRC Human Genetics Unit, Institute of Genetics and Cancer, The University of Edinburgh, Edinburgh, United Kingdom • [2]Edinburgh Cancer Research, CRUK Scotland Centre, Institute of Genetics and Cancer, The University of Edinburgh, Edinburgh, United Kingdom • [3]Centre for Cardiovascular Science, The University of Edinburgh, Edinburgh, United Kingdom • [4]Department of Anesthesiology, Vascular Biology and Therapeutics, Program in Translational Biomedicine, Yale School of Medicine, New Haven, United States • [5]MRC Protein Phosphorylation and Ubiquitylation Unit, School of Life Sciences, University of Dundee, Dundee, United Kingdom

eLife Assessment

This is an **important** study that establishes a zebrafish model of PIK3CA-related overgrowth syndrome. The imaging characterization of the mesodermal, particularly vascular, lesions of the model is **compelling**. The scRNA-Seq analysis is **convincing**, revealing key perturbations in the PIK3CA-mutation model, although deeper investigation of the exact mechanism leading to the lesions, as well as validation at different time points, could further strengthen the findings. This work will be of interest to medical biologists working on PROS, and potentially to a broader audience interested in non-cell-autonomous signaling of PIK3CA and its implications in other diseases.

https://doi.org/10.7554/eLife.110896.1.sa4

Abstract

Post-zygotic gain-of-function *PIK3CA* mutations arising during embryonic development cause disorders collectively known as the *PIK3CA*-related overgrowth spectrum (PROS). This ranges from minor, localized overgrowth to devastating multi-tissue overgrowth. Disease severity is widely attributed to a combination of PIK3CA genotype, affected cell type, and developmental timing of mutation acquisition. However, in PROS this explains neither the biased pattern of overgrowth - disproportionately affecting mesoderm and neuroectoderm-derived tissues - nor the typical low mutation burden within areas of extensive tissue overgrowth. Indeed, growing evidence suggests PROS mutations cause overgrowth non-cell-autonomously, although mechanisms of this are poorly understood. Here, we develop mosaic zebrafish models of PROS with overexpression of orthologous hotspot *pik3ca* mutations (*pik3ca*^{PROS}) to visualize and examine the effects of mutated cells on early development, in whole live animals. Reminiscent of PROS, we observe a spectrum of embryonic vasculature malformations (VMs), accompanied by larval muscle and bone overgrowth. Surprisingly, VMs only rarely expressed *pik3ca*^{PROS} in constituent endothelial cells, with *pik3ca*^{PROS}-expressing cells often closely abutting malformations instead. Single-cell transcriptomics of *pik3ca*^{PROS}mosaic zebrafish prior to VM onset revealed that most *pik3ca*^{PROS} cells are relatively immature and developmentally inert or constitute a small minority of mesodermal-derived cell types. Despite this constriction, global changes to cell fate were evident, alongside pervasive, pan-lineage abnormalities of gene expression, and rewiring of predicted ligand-receptor communication between lineages. We propose that targeting signals that indirectly propagate overgrowth through non-cell autonomous mechanisms - as well as PI3K activation itself - is worthy of therapeutic investigation.

Significance statement

Patchy, or mosaic, activating mutations in *PIK3CA* cause asymmetric human overgrowth due to aberrant hyperactivation of phosphoinositide 3-kinase (PI3K) signaling. Overgrowth prominently affects blood vessels and may be severely debilitating. While causation by *PIK3CA* mutations is clear, the explanation for the extent and pattern of associated overgrowth is not. Leveraging the power of zebrafish models for observation of early development, we now provide evidence for widely pervasive developmental effects of PIK3CA mutations extending beyond transcriptionally and phenotypically affected cells and lineages. This suggests potential therapeutic value of targeting secondary effects of PIK3CA activation as well as the activated PI3K itself.

Introduction

Coordination of lineage determination and growth rates within and among cell lineages is central to healthy human development. Class 1 phosphoinositide 3-kinases (PI3Ks) are one crucial group of signaling enzymes that mediate such coordination. PI3Ks transduce numerous upstream growth cues into widely ramifying downstream effects on cell growth, metabolism and other behaviors (1). They do this in a cell type, developmental stage and tissue specific manner. These growth cues include hormones and growth factors which activate receptor tyrosine kinases. The dominant class 1A PI3K subtypes regulating cell growth and metabolism in non-hematopoietic cell lineages include the p110α catalytic subunit encoded by *PIK3CA*, acting in concert with one of several regulatory subunits (2, 3). The heterodimeric enzyme is typically referred to as PI3Kα.

Hyperactivation of PI3Kα is common in solid human cancers, most often attributable to somatic activating *PIK3CA* mutations. The mutations with the greatest hyperactivity fall in 'hotspot' regions at the C terminal of the kinase domain (Histidine 1047) or in the helical domain (Glutamates 542 or 545) (4). These same mutations, if acquired during human development, also cause a wide range of asymmetric overgrowth disorders. Many different disease labels are used for these based on phenotypic descriptors, and the term *PIK3CA*-related overgrowth spectrum (PROS) is now used as an umbrella term for the whole range, reflecting its shared etiopathogenesis (5, 6). PROS ranges from very localized overgrowth, for example a circumscribed vascular malformation or single digit overgrowth, to severe and sometimes life-threatening multisystem disorders such as CLOVES syndrome (Congenital Lipomatous Overgrowth, Vascular Malformations, Epidermal Nevus, Spinal/Skeletal anomalies/Scoliosis) (5, 7). Vascular complications such as thromboembolism, hemorrhage, mass effects or high output heart failure are major causes of morbidity and mortality in PROS.

The wide phenotypic spectrum of PROS has most commonly been rationalized in terms of the timing and cell of origin of the founder *PIK3CA* mutation, which determines the anatomical site(s) of overgrowth, with the *PIK3CA* genotype (hotspot *vs* non hotspot) an important accessory factor determining the rate of overgrowth of affected tissues. This formulation fails to provide a complete account of the observed phenotypic heterogeneity, however. One unexplained observation is that not all tissues are equally affected in PROS. Mesoderm-derived veins, capillaries, lymphatics, adipose tissue, skeletal muscle, and bone are commonly and severely affected, for example, as are ectoderm-derived brain and peripheral nerves. In contrast, mesoderm-derived arteries and hematopoietic cells, and endoderm-derived organs are rarely affected (4, 7). Given the presumed stochastic nature of spontaneous *PIK3CA* mutagenesis, this implies that mutations are excluded from, not tolerated in, or not phenotypically penetrant in some lineages (4).

Another important observation is that the burden of *PIK3CA* mutation is variable, but often in the range of <1-10% in tissue biopsies, even from macroscopically severely affected regions, meaning that the bulk of sampled cells are genetically wild-type (WT) (4, 8–10). This will partly reflect sampling of a genetically mosaic tissue, but the extent of this mismatch also suggests that *PIK3CA* mutations may exert both cell-autonomous and non-cell-autonomous effects on tissue growth. Work emerging using other model systems has shown that non-blood vessel cells, such as PIK3CA-mutant lymphatic endothelial cells or Schwann cells, can increase PI3K activation in neighboring

endothelia to worsen vascular lesions and drive proliferation of WT cells, through recruitment of macrophages or secretion of pro-inflammatory eicosanoids and cytokines to drive proliferation of WT cells (11–14). Beyond its conceptual interest, identifying mediators of such non-cell-autonomous mechanisms may be of translational value if these secondary drivers of pathologic growth can be targeted. This is highly relevant in PROS as available potent PI3K inhibitors have limited efficacy in only a subset of patients in studies so far (15, 16), underscoring the need for better classification and understanding of the wider signaling landscape within VMs to ensure optimal treatment strategies (17).

Dissecting how the presence of *PIK3CA*-mutated cells corrupts normal development is ideally interrogated *in situ* in developing animals. Numerous animal models of PROS have indeed been reported to date, using strategies based on endogenous (18–20) or transgenic overexpression of *PIK3CA* hotspot variants (21–27), or a strongly activating but synthetic *Pik3ca* allele (28, 29). However, although such strong genetic *Pik3ca* activation has been shown to drive overgrowth in numerous lineages, the early events in overgrowth *in utero* are not visible in mammals. We now present a zebrafish model of PROS circumventing this limitation. The genetic tractability and transparency of zebrafish from their first cell divisions into adulthood means that critical stages in disease onset and evolution hidden *in utero* can be visualized and followed at cellular resolution from disease onset to adulthood in an intact organism (30). The large clutch size and rapid development of zebrafish also offer high experimental throughput to enable rapid testing of multiple PROS alleles in distinct cellular lineages.

We use this new model to investigate the cellular and transcriptional landscape at the onset of PROS vascular overgrowth. We show that overexpression of *pik3ca* hotspot alleles (henceforth collectively "*pik3ca*^{PROS}") causes extensive vascular malformations, despite the majority of *pik3ca*^{PROS}cells localizing adjacent to, rather than within lesions. Overexpression limited to the mesoderm had widespread effects on subsequent development, manifesting as aberrant partitioning of mutant cells among mesoderm-derived lineages. Single cell transcriptomics suggested more pervasive effects on gene expression and signaling in non-mesodermal lineages, with potential loosening of the normal pattern of developmental signal crosstalk among key embryonic structures. These findings consolidate evidence for developmental effects of *pik3ca*^{PROS} mosaicism that extend far beyond affected cells. This aligns with observations made in human PROS, presents a versatile animal model for further mechanistic interrogation, and opens new possibilities for accessory therapeutic targets beyond PI3K itself.

Results

Mosaic *pik3ca* hotspot transgenes cause vascular malformations in zebrafish embryos

Zebrafish Pik3ca protein has 91% homology with human PIK3CA, and AlphaFold predicts closely similar higher order structure to the experimentally determined structure of *PIK3CA*, evidenced by nearly identical positioning of PROS-related hotspot variants in both species (Figure 1A). To develop a zebrafish *pik3ca-PROS* model, we generated Tol2-based vectors permitting *β-actin* promoter-driven (ubiquitous) expression of zebrafish *pik3ca*; either wild-type (WT) or harboring one of the hotspot mutations, namely *H1048R*, *H1048L* (orthologous to human *H1047R* and *H1047L* respectively), *E542K*, or *E545K* (Figure 1B). A downstream IRES-*mScarlet* sequence enabled co-expression of the mScarlet reporter with mutant *pik3ca*. Additionally, the presence of an uncoupled secondary *cmcl2:GFP* cardiac reporter was included to provide a secondary readout of transgene integration. The *pik3ca*^{H1048R}*-mScarlet* construct was introduced into WT embryos to produce F0 mosaic transgenic embryos, which were raised and screened for transgene integration and overgrowth.

Many *pik3ca*^{H1048R}*-mScarlet*-injected embryos appeared morphologically normal, but 15% developed vascular lesions, observed as blood pooling and poor circulation most evident at 72 hours post-fertilization (hpf) (Figure 1B). To investigate this phenotype further, we repeated *pik3ca*^{H1048R}*-mScarlet* transgene injections into the *Tg(fli1:eGFP)* endothelial reporter line,

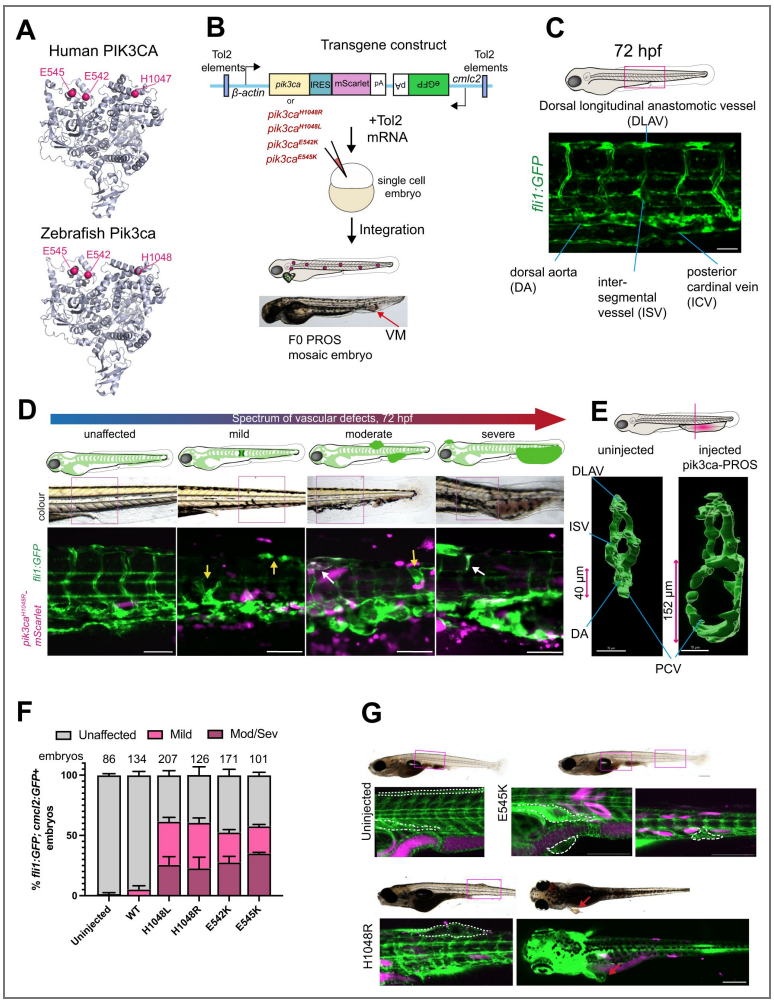


Figure 1. Mosaic *pik3ca* hotspot transgenes cause vascular malformations in zebrafish embryos

A) The structure of human PIK3CA and the predicted structure of the zebrafish Pik3ca orthologue is remarkably conserved, with one additional amino acid in zebrafish Pik3ca shifting the kinase domain amino acids by 1. The position of E542, E545 and H1047 residues found mutated in PROS (and the orthologous E542, E545 and H1048) are labelled in magenta and are predicted to occur in very similar structural regions of the protein in both human and fish. B) Schematic representation of the transgenic *pik3ca*^{PROS} constructs generated for this study, containing *β-actin:pik3ca*^{PROS}*-mScarlet* and the cardiac reporter *cmlc2:GFP* flanked by Tol2 response elements, and is co-injected with *Tol2* mRNA into 1-cell zebrafish embryos. Embryos were screened for successful integration of the transgene from 48 hpf onwards, indicated by mosaic cardiac *cmlc2:GFP*, and *mScarlet* expression. At 72 hpf, some PROS mosaic embryos had tail vascular malformations. C) Maximum Z-projection of a 72 hpf *Tg(fli1:eGFP);casper* tail region, showing typical endothelial vessel structures marked by *fli1:eGFP* fluorescence (green). Scale bar = 50 µm D) Representative diagrams and fluorescent microscopy images of the spectrum of vascular phenotypes seen in *Tg(fli1:eGFP)* embryos injected with *pik3ca*^{H1048R}*-mScarlet*. These ranged from unaffected and mild, to moderate and severe. *Fli1:eGFP* expression in unaffected embryos is as uninjected. Embryos scored 'mild' displayed expansion and mispatterning of isolated inter somitic vessels that was only visible when viewing eGFP fluorescence. Moderate and severe scores were applied to embryos where blood pooling was evident under brightfield, particularly obvious in the PCV. Yellow arrows indicate mis-patterned vessels in 'mild' embryos, and white arrows indicate overlap of *fli1:eGFP* (green) and *mScarlet* (magenta) expression in moderate and severe examples. Quantification of this and numbers screened can be seen in Figure 3. Scale bars = 100 µm. E) Cross sections of trunk blood vessels at 72 hpf, showing the dramatic expansion of the PCV. F) Quantification of these vascular malformation phenotypes for each construct, scored based on severity, as a percentage of the total numbers injected. Numbers indicate *fli1:eGFP; cmcl2:GFP+* embryos screened for vessel malformations at 72 hpf, representative of three biological repeats per construct. Error bars represent S.E.M. G) Representative images of juvenile *Tg(β-actin:pik3ca*^{E545K}*-mScarlet; fli1:eGFP); casper* (top) and *Tg(β-actin:pik3ca*^{H1048R}*-mScarlet; fli1:eGFP* fish that harboured no detectable vascular anomalies at 72 hpf, but were found to have vascular anomalies when imaged two weeks later (*E545K*-injected – 6/16, *H1048R*-injected – 3/9). Dotted lines encompass vessel malformations in PROS animals, and equivalent regions in control

enabling co-visualization of blood vessels and transgene expression (31). *Fli1:eGFP* is first detectable during somitogenesis, during which angioblasts migrate from the lateral plate mesoderm (LPM) between the 10 and 18 somite stage (~16 hpf) towards the midline, with the main embryonic artery and vein being formed at 24 hpf, and trunk circulation being completed at 48 hpf (31, 32). In the trunk of zebrafish embryos at 72 hpf, arteries and veins are well established, with regularly-spaced intersegmental vessels (ISVs) connecting the ventral major aortic and venous vessels with the dorsal longitudinal anastomotic vessel (DLAV) (31, 33) (Figure 1C). Consistent with the variability observed in PROS, the severity of vascular malformations resulting from transgene injection ranged from mild to severe, the latter obvious from dramatic enlargement of the major blood vessels, causing blood pooling in the tail or head, with some vascular overgrowths large enough to compromise swimming (Figure 1D).

In the same system, we next tested the propensity of other PROS hotspot variants to cause vascular malformations, injecting WT, *H1048R, H1048L, E542K* or *E545K pik3ca*^{PROS} constructs into *Tg(fli1:eGFP)* embryos. Transgenic expression of *pik3ca*^{WT}*-mScarlet* did not induce overgrowth (Supplementary Figure 1A). However, all four hotspot variants, whether affecting the kinase (*H1048R/L*) or helical domains (*E542K/E545K*) of Pik3ca, reproducibly caused a spectrum of vascular anomalies, from mild mis-patterning of *fli1:eGFP*+ ISVs only evident by fluorescent transgene expression pattern, to severe balloon-like overgrowth, of the major vessels in the tail (Figure 1D,E, Supplementary Figure 1A). Severity scoring of malformations for each of the four PROS alleles and of WT or uninjected controls showed that transgenic expression of mosaic *pik3ca*^{PROS}variants in zebrafish caused vascular malformations at similar levels of severity (Figure 1F).

To test whether our model could recapitulate other overgrowth elements seen in PROS, we selected 72 hpf embryos displaying a high level of *pik3ca*^{PROS}transgene integration but appearing otherwise normal, and grew them to larvae stages (to 14 days). Transgene integration was assessed either by high numbers of *mScarlet*+ cells or by the surrogate marker of expression of the co-injected heart-targeted *cmcl2:GFP* reporter in most cardiac cells. We imaged these fish as they developed into larvae. Despite no visible vascular abnormality in the embryos, these fish still developed vascular malformations by larval stages, seen as thickening of vessels, or as vascular protrusions from the skin (Figure 1G). Aside from vascular malformations, some *pik3ca*^{PROS} zebrafish larvae also showed appreciable *mScarlet*+ expression and overgrowth in the musculature, as well as premature mineralization in the notochord, fusing and thickening of vertebrae, and abnormal patterning of hypural cartilage in the tail (Supplementary Figure 1B-D). These findings suggest that in the fish model, as in human PROS, the trajectory of malformations is not exclusively determined in embryos, and overgrowth can sometimes manifest or progress post-embryonically, notably in tissues of mesodermal origin.

Evidence for cell-autonomous and non-cell-autonomous mechanisms of *pik3ca*^{PROS} vascular overgrowth

Surprisingly, we observed that only a minority of endothelial cells within vascular malformations expressed both mScarlet and eGFP fluorophores, proxies for *pik3ca*^{PROS} and *fli1* expression, respectively. We used confocal microscopy to build a 3D image of the vascular lesions in higher resolution to investigate this further (Figure 2A; Movie 1). In one example of a mild VM, we observed thickening of the DLAV without mScarlet cells in the lesion itself, but with *pik3ca-PROS-mScarlet*+ cells abutting the overgrown region (Figure 2A; Movie 2). In other moderate and severe vascular malformations, *pik3ca*^{PROS}*-mScarlet*+ cells were again most often found adjacent to malformations rather than comprising abnormal vessel walls (Figure 2A; Movie 3). Indeed, GFP and mScarlet double positive cells accounted for <5% of total lesion volume on average (Figure 2B). This provides circumstantial evidence for the hypothesized non-cell autonomous mechanisms in PROS. Specifically, it argues that unidentified *pik3ca*^{PROS} cells in or close to the vascular niche may exert non-cell-autonomous effects upon WT endothelial cells.

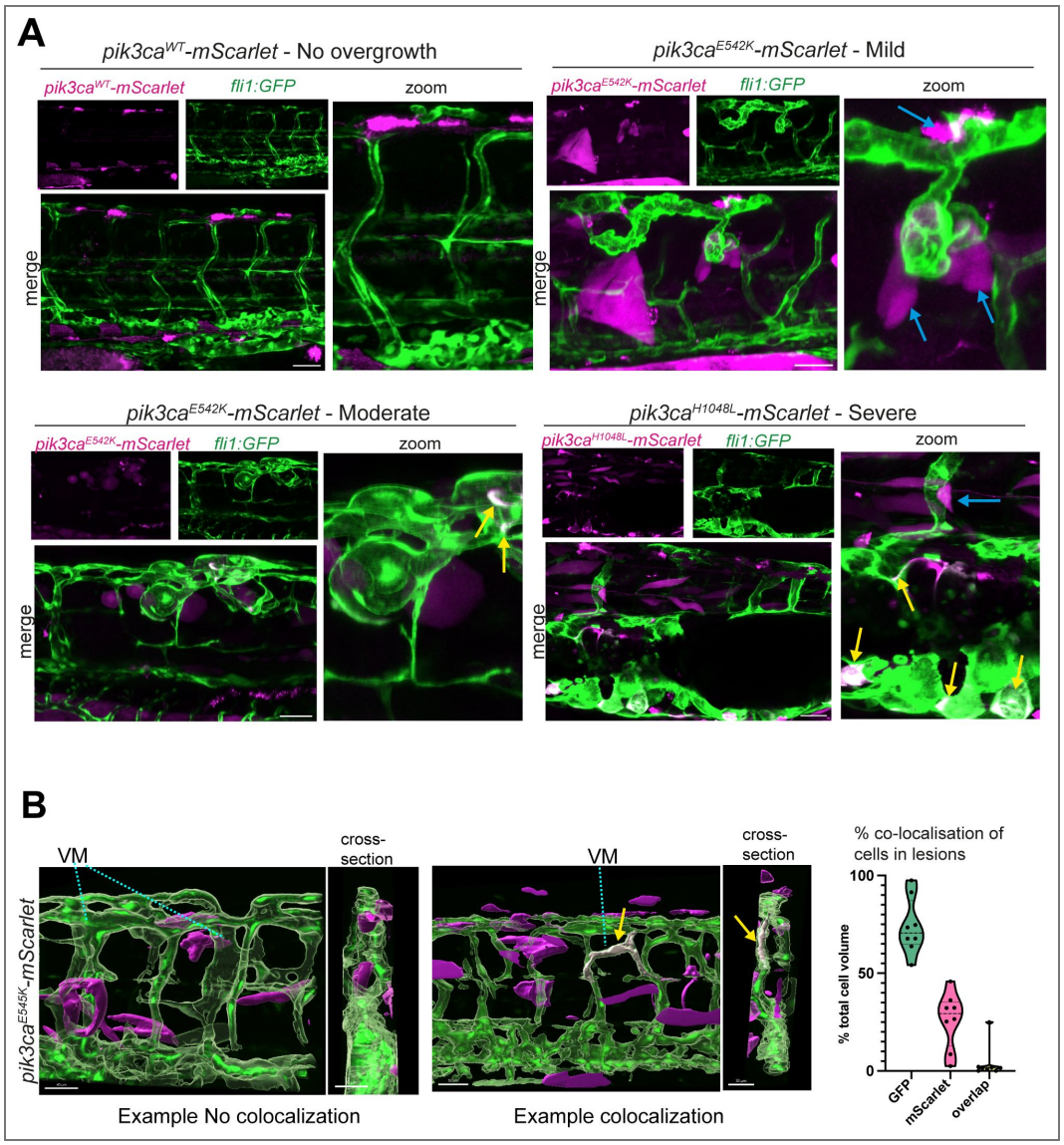


Figure 2. Evidence for cell-autonomous and non-cell-autonomous mechanisms of *pik3ca*^{PROS} vascular overgrowth

A) Representative maximum projection confocal images of 72 hpf *Tg(fli1:eGFP);casper* embryos expressing either wildtype *pik3ca-mScarlet* (no overgrowth) or a *pik3ca*^{PROS} hotspot variant and exhibiting a mild, moderate or severe vascular phenotype. For each, separated *mScarlet* fluorescence (magenta), *fli1:eGFP* channels (green) and merged channels are shown, with a zoomed area to the right. Blue arrows indicate *mScarlet* cells in close proximity to deformed vessels. Yellow arrows indicate white cells co-expressing *fli1:eGFP* and *mScarlet*. Scale bars = 50 μm B) Quantification of *mScarlet* cells by volume within GFP+ endothelial cells. Imaris was used to visualize confocal Z stacks of vessel overgrowth at 72 hpf. 3D structures of each channel, plus a co-localization channel were converted into surfaces, enabling volume calculations for vessels (green), mScarlet+ cells outside vessels (purple) or areas of colocalization (white). Volumes of each were expressed as percentage of total cell volume in 20x magnification view, dots indicate individual images (8 lesions were analysed in total).

Early mesodermal *pik3ca*^{PROS} transgene expression causes endothelial mis-patterning

Many tissues prominently affected in PROS are mesoderm-derived, indicating biases in cell lineage contribution to overgrowth, and the possibility of spatial and temporal windows of especial vulnerability to PROS mutations during development (4). We therefore sought to investigate the consequences of disease onset in the early mesoderm, which has hitherto not been possible to observe *in vivo* in mammalian systems, by swapping our ubiquitous *β-actin* promoter for that of the pan-mesodermal marker *tbxta* (Figure 3A, B). The activation of *tbxta* prior to endothelial specification would also address the possibility that *mScarlet*+ endothelial cells had been present during vasculogenesis but had either silenced transgene expression or died prior to imaging and scoring at 72 hpf. As all tested *pik3ca*^{PROS} alleles gave rise to VMs with the same frequency and severity (Figure 1F), we focused on one hotspot variant for subsequent work, *pik3ca*^{E545K}, for studying mesodermal PI3K gain-of-function.

pik3ca^{E545K}*-mScarlet*+ expressed from the *tbxta* promoter was as effective as the *β-actin* construct in inducing the *pik3ca*^{PROS}vascular phenotype in *Tg(fli1:eGFP)* embryos (Supplementary Figure 2). *tbxta:pik3ca*^{PROS}*-mScarlet* expression was visible from early gastrulation, before the earliest plausible stage of vascular malformation onset, as angioblasts migrate from the lateral plate mesoderm (LPM) towards the midline during somitogenesis (10-22 hpf) (31). At 18 hpf, *mScarlet* expression in injected controls was predominantly in the notochord and myotome, but co-expressed with very few GFP+ endothelial cells, as expected from known expression patterns of *tbxta* at this developmental stage (34) (Figure 3C). In controls, very occasional ectopic *fli1:eGFP* cells were observed dorsal to the developing DA and unattached to the nascent vasculature, thus seemingly independent of normal ISV sprouting, which occurs from 20 hpf (Figure 3C, Movie 4). In embryos overexpressing *pik3ca*^{PROS} however, the frequency of such ectopic *fli1:eGFP* cells was markedly increased (Figure 3D). These ectopic cells persisted through ISV sprouting and subsequent branching to form the DLAV at ~22 hpf. Occasionally, at this time point, we observed rare examples of *eGFP* and mScarlet double positive cells dying (Figure 3C, orange arrow). By 28 hpf, we observed DLAV cells forming premature dilated vessels (Movie 5).

We thus demonstrate that *pik3ca*^{PROS} expression in mesoderm alone disrupts the earliest stages of vascular development, causing dysregulation of endothelial cell sprouting and branching. Occasional *mScarlet*-expressing endothelial cells were present, some of which were observed dying during aberrant development, which may indicate heightened sensitivity of endothelial cells to high levels of *pik3ca*^{PROS} transgene overexpression. However, as in later vascular development, vascular defects in this model appear to be largely a non-cell autonomous effect of *pik3ca*^{PROS}expression on *fli1:GFP*+, *mScarlet*+ endothelial cells.

Mosaic mesodermal *pik3ca*^{PROS} expression induces pan-lineage dysregulation

Given this evidence for non-cell autonomous effects of mesodermal *pik3ca*^{PROS} expression on early vascular development, we next sought to interrogate more agnostically the extent of the developmental disruption. We used FACS to sort live cells from *tbxta:pik3ca*^{PROS}*-mScarlet* and *tbxta:mScarlet* injected controls at 19 hpf and subjected these to single-cell RNA sequencing (scRNA-seq) (Figure 4A). Using Seurat, we identified 40 distinct cell populations common to 'control' and 'PROS mosaic' embryos. These were named using a combination of marker gene expression patterns and projection of two previously published scRNA-seq datasets onto the data (Figure 4B, Supplementary Figure 3A, B) (35, 36). This confirmed the presence of all cell lineages expected at this developmental stage, including endothelial cells, CNS, epithelial, lateral plate and presomitic mesoderm (LPM, PSM), and notochord (Figure 4B, C).

One clear difference between control and *pik3ca*^{PROS} embryos lay in so-called cluster '0', which could not be mapped to a uniquely annotated lineage, instead featuring weak signatures of many different cell types, hinting at a "confused" lineage identity (Supplementary Figure 3C). We

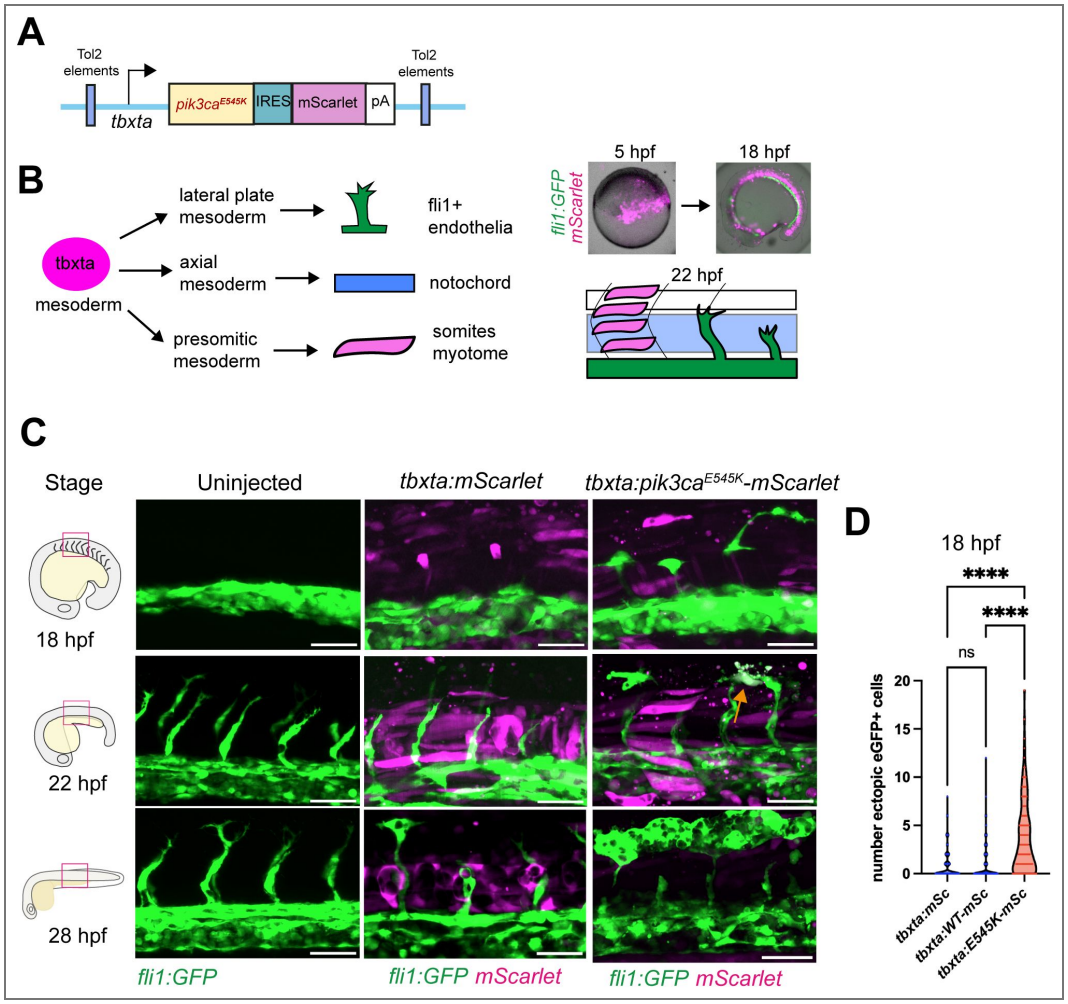


Figure 3. Early mesodermal *pik3ca*^{PROS}transgene expression causes endothelial mis-patterning

A) Plasmid design of the transgenic construct *tbxta:pik3ca*^{E545K}*-mScarlet*. B) Simplified diagram of a *tbxta* progenitor cell which may differentiate into the lateral plate mesoderm, the notochord, or presomitic mesoderm which gives rise to endothelia, the notochord or muscle respectively. Mosaic *tbxta:mScarlet* is visible at 5 hpf (representative images) before onset of *fli1:eGFP* reporter expression. By 18 hpf, *tbxta* is most highly expressed in the notochord. C) Representative maximum projection confocal images of embryos from a range of stages between somitogenesis (~18 hpf) and tailbud (~28 hpf) to visualise the start of *fli1:eGFP* expression and endothelial development. In uninjected embryos, lateral stripes of *fli1:eGFP* are present from 16hpf, with sprouting of ISVs occurring from anterior to posterior from around 19 hpf onwards. In *tbxta:pik3ca*^{E545K}*-mScarlet* injected embryos, ectopic *fli1:eGFP* cells can be seen dorsal to the initial lateral plate mesoderm stripe. Scale bars = 50 μm D) Quantification of ectopic *fli1:eGFP*+ cells in injected *tbxta:pik3ca*^{E545K}*-mScarlet,* or injected with a similar construct lacking the *pik3ca-ires* element (unattached to the lateral mesoderm stripes of *fli1:eGFP*). 3 biological replicates shown with each dot representing 1 embryo. Significance was calculated with a two-tailed Mann Whitney test. ****=p<0.001.

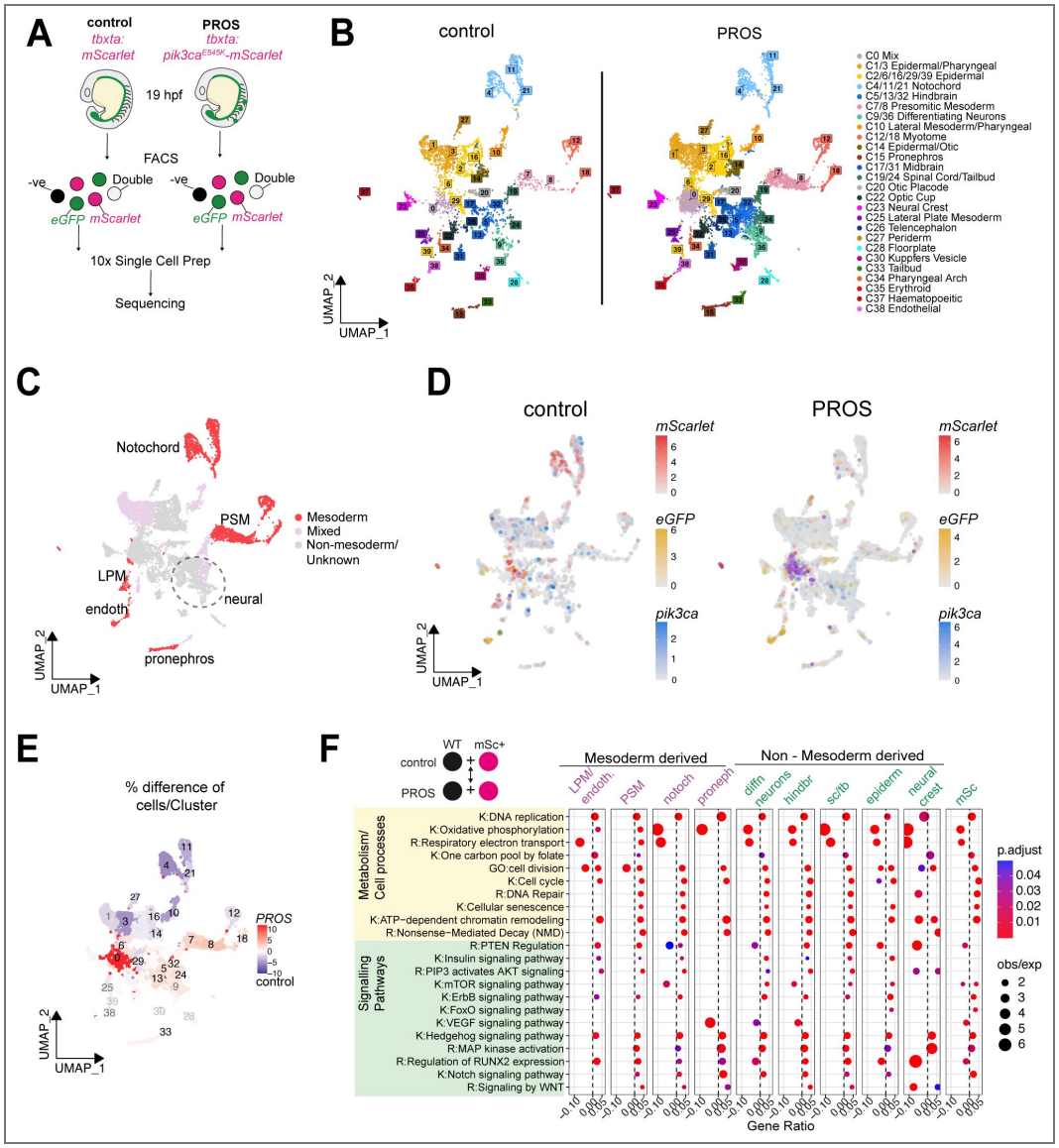


Figure 4. Mosaic mesodermal *pik3ca*^{PROS} expression induces pan-lineage dysregulation

A) Schematic of experimental workflow for scRNA-seq. *Tg(fli1:eGFP)* embryos were injected with *tbxta:mScarlet* (control) or *tbxta:pik3ca*^{E545K}*-mScarlet* (PROS) and grown until 19 hpf, dissociated, and all live cells sorted by FACS before processing for 10x sequencing. B) UMAP showing 40 cell clusters found the aggregated dataset, split by experimental condition. C) UMAP of aggregated dataset coloured by mesodermal or other germ layer origin D) FeaturePlots split by experimental condition, showing expression levels of *mScarlet* (red scale), *eGFP* (yellow scale) and *pik3ca* (blue scale) E) The numbers of cells per cluster was expressed as a percentage of total cells for each experimental condition, and then the difference between percentages in PROS and control clusters visualised in a UMAP heat plot. Blue indicates a higher percentage of control cells relative to PROS, and red indicates a higher proportion of PROS relative to control cells. F) Deseq2 DE analysis and gene enrichment analyses using GO (G:), KEGG (K:) and Reactome (R:) databases were performed to compare PROS mosaic vs control cells in various cluster groups, or in *mScarlet*+ cells across the dataset.

confirmed upregulation of the *pik3ca* transcript itself in this cluster, largely co-segregating with *mScarlet* in PROS mosaic animals (Figure 4D). Velocity analysis showed a lack of clear developmental directionality of transcription in PROS mosaic embryos compared to the small corresponding cluster in control animals. In contrast, velocity in the PSM and neural lineages was increased, suggesting non-cell autonomous perturbations of lineage transitions in both mesodermal (*tbxta*-driven) and non-mesodermal (i.e. neural) lineages (Supplementary Figure 3D).

The changes in relative sizes of different cell clusters between experimental conditions was assessed by expressing the number of cells within each cluster as a percentage of the total cell number for control and PROS datasets, respectively. After calculating the differences in these percentages, a dramatic expansion of cluster 0 was observed in PROS mosaic embryos (Figure 4E; Supplementary Figure 3E). Some clusters of mesodermal origin also showed expansion in PROS mosaics, as may be expected from mesodermal activation of PI3K signaling, but, surprisingly, this was not true of all mesoderm-derived clusters, and ectoderm-derived neural clusters were also enriched in PROS mosaics (Figure 4D, E; Supplementary Figure 3E). This suggests that *pik3ca*^{PROS} cells have developmental effects extending beyond mesodermal lineages themselves, unbalancing cell fate decisions and thereby cell type proportions in the embryo.

In PROS mosaic embryos, 319 genes were significantly upregulated in mScarlet+ PROS cells relative to wildtype in PROS mosaics, including, as expected, *pik3ca, tbxta* and *mScarlet*, but fewer were upregulated in PROS mosaics compared to control embryos (Supplementary Figure 4A). Analysis of curated transcriptomic "footprints" of pathway activity indicated that PI3K-coupled programs were significantly enriched in *mScarlet*+ cells in PROS mosaics relative to their WT *mScarlet*- counterparts, indicating increased PI3K activity as well as *pik3ca*^{PROS} overexpression in transgenic cells (Supplementary Figure 4B). Interestingly, the same analysis revealed enrichment for signatures of apoptosis, DNA damage repair, G2/M checkpoint activation and p53 activity, suggesting that *pik3ca*^{PROS} expression may ultimately be toxic. Such negative selection of *pik3ca*^{PROS} has previously been invoked as a possible explanation for the relative lack of PIK3CA mutations in some human cell lineages. Surprisingly, few genes were significantly upregulated in PROS mosaic cluster 0 relative to control, although footprint analysis did suggest significant enrichment of MYC-Targets and Apical Junction components (Supplementary Figure 4C), although the uncertain nature of cluster 0 in control embryos complicates interpretation of this comparison.

In other cluster groups, we found widespread enrichment of Kegg or Reactome pathway components directly related to PI3K signaling in clusters from the PROS zebrafish, including modules related to Pten, Insulin, PIP3, mTor and Foxo1 (Figure 4F). Other gene sets were also frequently enriched, including Notch, Wnt and Mapk pathways, as well as shared responses to cell metabolism and cell division processes. Strikingly, these transcriptional changes extended to clusters in the PROS zebrafish of non-mesoderm origin that expressed almost no mScarlet cells (Figure 4F). Moreover, these gene expression patterns persisted when we removed the PROS *mScarlet*+ cells from the analysis entirely, indicating that otherwise wild type mesodermal and non-mesodermal lineages are responding transcriptionally to PROS cells even without expressing *pik3ca*^{PROS} themselves (Supplementary Figure 4D).

pik3ca^{PROS} elicits widespread transcriptional changes in ligand-receptor expression patterns

Finally, we sought evidence from single cell transcriptomic data implicating potential mediators of the observed non-cell autonomous effects of genetic PI3K hyperactivation. We did this using CellChat and NICHES as analytic tools to infer communication among cell lineages using the proxy of expression patterns of ligand/receptor pairs (37, 38). Analyses of the top 10% predicted cluster interactions by CellChat suggested major transcriptional changes in signalling mediators in PROS embryos. In WT embryos the transcriptomic landscape appeared dominated by reciprocal

expression of ligand receptor pairs (loosely "signaling traffic") among the notochord, and mixed epithelial/pharyngeal lineages, while in PROS embryos a more distributed pattern was seen, with presomitic mesoderm and various neural lineages much more prominent (Figure 5A).

To assess potential specific signaling pathways driving these predicted changes, we used the NICHES algorithm, which analyses ligand-receptor mRNA expression in one-to-one pairs of cells to infer communication (38). We subsetted our data to include the clusters found most active by CellChat and ran NICHES to convert the single-cell gene expression dataset to a single-cell signalling dataset based on LR signaling pattern genes, grouped by sending clusters, receiving clusters, or individual pair-wise interactions (Figure 5B). To identify potential changes in cell signalling capacity between lineages in PROS mosaic animals relative to control, we performed DE analysis on 'sending' and 'receiving' transcriptional signatures for each cluster (Supplementary Figure 4E). To gain a comprehensive overview of signaling changes, we simplified the outputs by grouping and averaging ligand and receptor fold changes respectively (Figure 5C). We observed a marked increase in Notch, Wnt and Ephrin ligand-receptor pairs in both mesodermal and non-mesodermal lineages, but again cluster 0 exhibited general downregulation of sending and receiving signaling components. This is in keeping with the non-cell autonomous responses to small numbers of *pik3ca*^{PROS}cells in the PROS zebrafish, and the relatively inert state of C0 (Figure 5C; Supplementary Figure 3 C,D).

We next calculated 'module scores' by averaging gene expression of all ligand-receptor pairs within a pathway (e.g. all combinations of *jag*/*dll*/*dld-notch1a/1b/2/3* for notch signalling) and compared the intensities of this score on cell-to-cell UMAPs, split by experimental condition (Figure 5D). Comparing the connectivity of PROS mosaics to control, we observed a reduction in extent and intensity of mRNA expression of *egf* and *pdgf* ligand-receptor pairs in PROS, further highlighting disruption of normal signaling through Pik3ca throughout the embryo (Figure 4F, 5D). In contrast, we observed widespread upregulation of pairs of *notch*, *wnt* and *ephrin* signaling genes (Figure 5D). Focusing on individual ligand-receptor pairs, we found evidence for transcriptional "rewiring" of these pathways between lineages. For example, *dla-notch3, wnt4-fzd5-lrp6* and *efnb3b-epha4a* ligand-receptor pair expression is increased for neural, PSM, myotome and Epi/Ph nodes (Figure 5D).

We conceptualize these data to indicate that *pik3ca*^{PROS}has significant effects on cell fate, behavior and likely signaling which ramify across many lineages, indicating pervasive and disruptive non-cell-autonomous effects on early development. Overall, these analyses highlight transcriptional changes in ligand and receptor gene expression that may contribute to the non-cell-autonomous phenotypes observed and serve as potential drug targets worthy of further investigation.

Discussion

Over ten years since mosaic, activating PIK3CA mutations were identified as the cause of a group of asymmetric overgrowth disorders (39, 40) much has been learned. The wide human clinical spectrum has been delineated, the natural history described, and nomenclature harmonized (5, 6). Furthermore, case series and registry based open label studies have established that the PI3Kα inhibitor alpelisib is of some clinical utility, leading to licensing in the USA (16, 29) and establishment of an international randomized clinical trial still underway (NCT04589650) (41). However, despite this major progress, important fundamental and translational questions about PROS remain. Specifically, the explanation for the skewed appearance of PROS-related pathology in some tissues but not others, and the involvement and extent of non-cell-autonomous mechanisms of overgrowth in PROS pathogenesis have eluded comprehensive analysis. The latter issue is of translational importance as it may yield novel diagnostic and therapeutic targets; given current evidence that alpelisib has only partial efficacy in only some patients with PROS (15), this is of great value.

Fundamental to translational progress is the fidelity of models to human pathology. A wealth of murine models has advanced our understanding of PROS, by showing definitive causal links between acquisition of PROS mutations and vascular malformations (22, 42), identifying critical time windows during which endothelia are most sensitive to excessive PI3K signaling (25), and

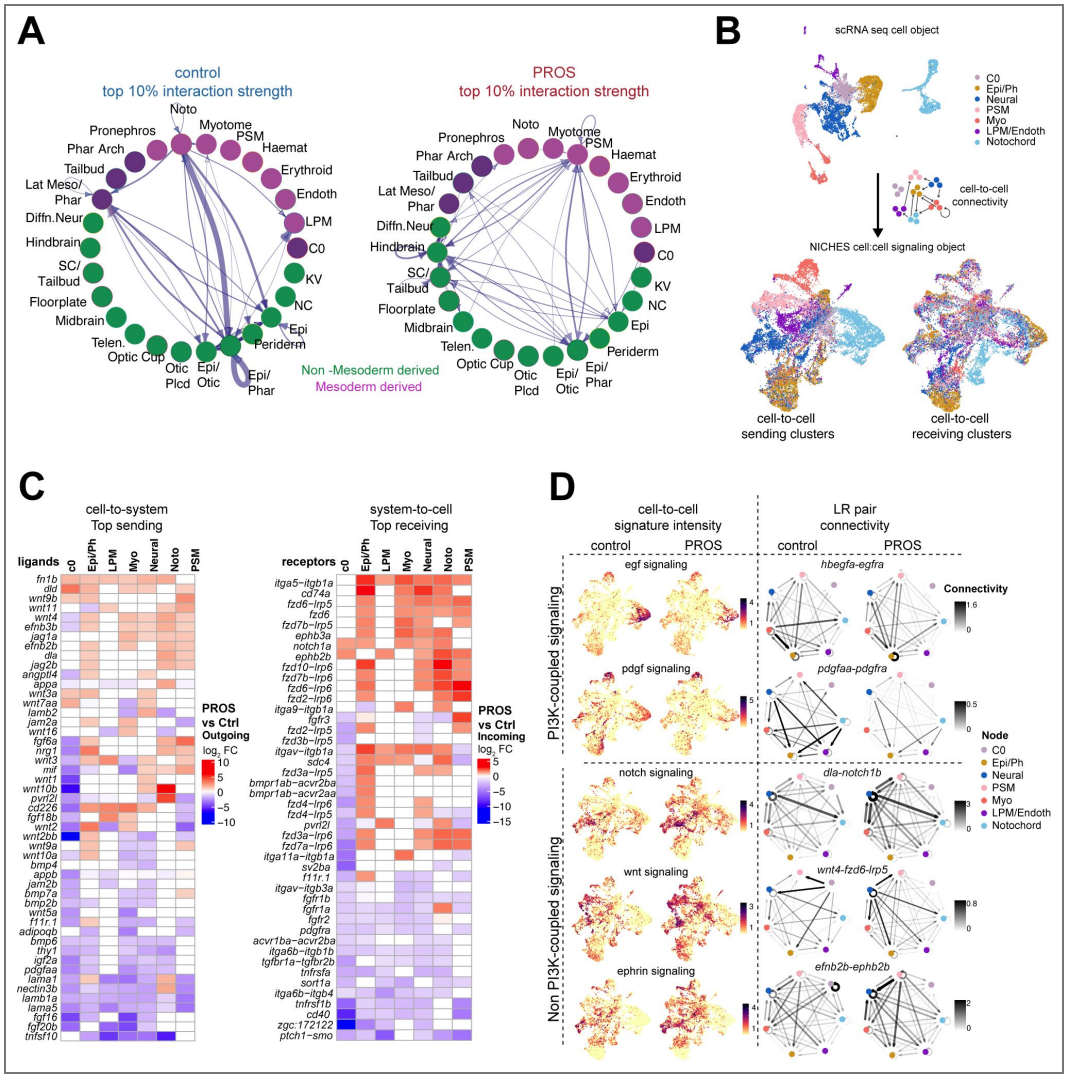


Figure 5. ***pik3ca*^{PROS} promotes widespread transcriptional changes in ligand-receptor expression patterns**

A) CellChat was used to identify potential interactions between ligand and receptor pairs in clusters. The top 10% strongest pathway interactions of 'control' cells relative to 'PROS' and vice versa shown, normalised to cell number. Blue arrows indicating whether a cluster is a sender or a receiver of the signal, with arrow thickness proportional to connection strength. Abbreviations: LPM=lateral plate mesoderm, PSM, presomitic mesoderm, Noto = notochord, Endoth=endothelial, KV = kuppfer's vesicle, NC=neural crest, Epi=epiderm, Epi/Phar epidermal/pharyngeal, Otic Plcd=otic placode, Telen=telencephalon, SC=spinal cord, Diffn. Neuron = differentiating neurons, B) UMAP showing merged data to illustrate the methodology of NICHES, whereby a subset of clusters from Figure 4 is reclustered (Top UMAP). Ligand-receptor interactions between each group were quantified using NICHES, converting it into a signaling object of cell:cell interactions (Bottom UMAPs, grouped by either sending cluster type (left) or receiving cluster type (right). Colours in interaction maps were retained from the original clusters in the top UMAP. Epi/Ph = epiderm/pharyngeal mix, PSM = presomitic mesoderm, Myo=myotome, LPM/Endothelial = grouped lateral plate mesoderm and endothelial clusters. C) Heatmap depicting $\log_2$ fold change in expression of the top differentially expressed "sending" ligand (left) and "receiving" receptor (right) genes in PROS relative to control embryos, clustered by row. Ligand-receptor mechanisms were aggregated by common ligand or common receptor (unaggregated data are shown in Supplemental Figure 4C). Only differences for which p.adj<0.05 by Wilcoxon test between PROS and WT are shown. D) Examples of pathways showing dysregulated expression of ligand-receptor (LR) pairs. The expression values of LR pairs for each pathway were averaged as a module score. FeaturePlots (left) show intensity of this module for each cell:cell interaction, split by experimental condition. Examples of circus plots for each pathway are shown on the right. Signal direction between clusters (nodes) is shown by arrows, with darker and thicker arrows corresponding to greater connectivity strength. Connectivity scale intensity is matched between control and PROS plots for each LR pair.

demonstrating that PI3K and related pathway inhibition can mitigate or rescue important disease manifestations (14, 21–23, 42, 43). More recently, mouse PROS or cancer models have implicated signaling from non-endothelial cell types as drivers of vascular overgrowth (11, 25, 27), but otherwise few animal studies have focused on non-cell-autonomous contributors to PROS pathology to build a holistic view of multi-lineage interactions during early development.

Overexpression of an activating PIK3CA allele has been shown to recapitulate several aspects of PROS-related overgrowth in zebrafish (28), and the successful usage of zebrafish as a preclinical model of lymphatic overgrowth or neurovascular malformations (44, 45) encouraged us to develop patient-led zebrafish models to interrogate the complex interplay of Pik3ca genotype and lineage context, and to examine overgrowth dynamics at early stages of development. We now present a zebrafish model with mosaic overexpression of naturally occurring hotspot *pik3ca*^{PROS} variants. This enables observation of the effects of PROS mutations *in vivo*, from onset, at cellular resolution for the first time. Our model recapitulates a spectrum of overgrowth severity affecting different cell lineages, as in PROS, with vasculature most prominently affected, while scRNA-seq suggests that *pik3ca*^{PROS} expressing cells appeared 'stuck' in a less differentiated cell state. This is reminiscent of iPSC models of *PIK3CA*^{H1047R} activation, where homozygous mutant cells show transcriptional remodeling and impaired differentiation (46). Our model also allows comparison of four different *pik3ca* hotspot gain-of-function variants in parallel, which proved to cause vascular overgrowth with similar frequency and severity. This argues against major differences in phenotypic expression across different hotspot Pik3ca variants, whether affecting helical or kinase domains, although the high level of expression in our model leaves open the possibility that differences may be unmasked at endogenous expression levels.

The non-cell autonomous influences of *pik3ca*^{PROS} cells found in our *pik3ca*^{PROS} model is in keeping with mounting evidence that paracrine signals from *PI3KCA*^{PROS} cells corrupt neighboring wildtype cells. For example, eicosanoids secreted from *PIK3CA* mutant breast cancer cells induce proliferation in WT cells (13), paracrine VEGF and IL-6 signaling from mammary duct culture lead to dysfunction in neighboring innervating vasculature *in vitro* (12). Moreover, *in vivo* activating Akt mutations in murine mammary tissue cause hyperplasia of nearby WT cells in a paracrine manner (47), and lymphatic capillary endothelial cells expressing *Pik3ca*^{H1047R} recruit VEGFC-producing macrophages which cause immune activation and worsening of vascular malformations in a mouse lymphoma model (11). Lastly, driving *Pik3ca*^{H1047R}in murine pericyte progenitors induced widespread capillary and venous malformations as well as hypertrophy of Cre-targeted connective tissues (48). Paracrine signals also have a role in normal endothelial and hematopoietic development, as a subtype of endothelial cells derived from the paraxial mesoderm have been shown to support induction of hematopoietic cells from neighboring hemogenic endothelial cells in zebrafish (49). It is important to note that cell-autonomous and non-cell-autonomous contributions to VMs in our model are not mutually exclusive, and we did consistently observe minorities of *fli1:eGFP*+ endothelia co-expressing *mScarlet* indicative of a mixture of signaling mechanisms occurring in VMs. Likewise, gain-of-function *PIK3CA* mutations are found through genotyping endothelial cells in PROS patients, and are thought to be a driver of overgrowth in their wildtype counterparts. Therefore, both cell-autonomous and non-cell-autonomous signals likely underpin PROS pathology.

A limitation of our model, as for many other published models, is that it relies on transgenic *pik3ca* overexpression, although the variants we overexpressed were orthologues to human disease-causing mutations. The extent of overactivation through transgenesis is likely higher than seen in PROS where *PIK3CA* is regulated endogenously, meaning the lack of *mScarlet* in cell types most sensitive to PI3K levels, such as the endothelia, may be due to an intolerably high levels of PI3K signaling causing cell death, as we observed in some co-expressing cells, and enrichment of apoptotic transcriptional signatures in *mScarlet*+ cells in PROS mosaic embryos. Furthermore, the timing of p110α activation is important for phenotype severity, with prenatal but not postnatal induction of *Pik3ca*^{E545K} causing increased brain size, and timed induction of *Pik3ca*^{H1047R} in endothelial cells revealing critical windows of PI3K sensitivity during development (25, 26). Further temporal resolution of *pik3ca*^{PROS} during zebrafish development, and to an extent the

'dosage' of PI3K activation would be possible using a drug-inducible version of our transgene. Finally, further testing of our transgenic construct in older fish, perhaps with more restricted spatial expression of PROS mutations, is required to see if adipose tissue and other tissue defects observed in PROS occur, and to the same penetrance.

Our novel zebrafish *pik3ca*^{PROS} model has enabled *in vivo* observation of early stages of evolution of PROS-like vascular malformations at a cellular level. It has moreover offered evidence for surprisingly pervasive developmental effects of mosaic *pik3ca* activation, including extensive non-cell autonomous effects ramifying beyond the affected lineage. This complements and extends existing models of PROS and provides a tractable platform for more in-depth interrogation of specific candidate mediators of such remote effects of *pik3ca* activation. Therapeutic targeting of these may in principle be a valuable adjunct or alternative to PI3K inhibition in future.

Materials and methods

Structural comparison of zebrafish and human PIK3CA

Protein structures for human PIK3CA (AF-P27986-F1-v4) and the predicted structure of zebrafish Pik3ca (AF-F1QAD7-F1-v4) were visualized and annotated using PyMOL (Schrodinger).

Fish husbandry, fish lines

Zebrafish were maintained in accordance with UK Home Office regulations, UK Animals (Scientific Procedures) Act 1986, amended in 2013, and European Directive 2010/63/EU under project licenses P8F7F7E52 and PP7317786. All experiments were approved by the Home Office and AWERB (University of Edinburgh Ethics Committee). Fish stocks used were wild-type AB, *Tg(fli1:eGFP)* (50), *casper* (51). Combined transgenic lines were generated by crossing. Adult fish were maintained at ~28.5°C under 14:10 light-dark cycles. Embryos were kept at 28.5°C and staged according to the reference table provided by (52) and (53).

Pik3ca^{PROS} transgene plasmid construction

First, the *eGFP* sequence in the p3E Gateway entry vector IRES-eGFP-pA (54) was swapped for the *mScarlet* CDS (Thermo Fisher Scientific) with the HiFi DNA assembly kit (NEB) to make the p3E entry vector p3E-IRES-mScarlet-pA. The wild type zebrafish *pik3ca* CDS was introduced into pDonr221 (Invitrogen) using Gateway™ BP Clonase™ II (ThermoFisher Scientific) to make pME-*pik3ca*^{WT}. *pik3ca* hotspot mutation sites were introduced into pME-*pik3ca*^{WT} sequence by designing back-to-back mutagenesis primers incorporating hotspot mutations *H1048R/L*, *E542K* or *E545K*, phosphorylating them using T4 PNK (NEB) according to manufacturer's instructions, and amplifying using Q5 polymerase (NEB). T4 ligase (NEB) was used to re-ligate the mutated amplicon, and the original template digested using Dpni (NEB). LR Clonase II+ (ThermoFisher Scientific) was used to combine p5E entry vectors p5E-*β-actin* (54), p5E-*fli1ep* (50), p5E- *tbxta* (this study, based off (55), with a *pik3ca-PROS* pME, and p3E-IRES-mScarletpA (this study, modified from p3E-IRES-GFP) into pDestTol2CG2 (54) which carries a *cmlc2:eGFP* reporter. Sanger sequencing was used to confirm successful mutation of pME-*pik3ca*^{WT} to pME-*pik3ca*$^{H1048/H1048L/E542K/E545K}$and correct integration of all components. Sequences of primers used for plasmid construction are found in Supplementary Table 1.

Generation of mosaic *pik3ca*^{PROS} fish

Approx. 2 nl of 30 ng/µl transgenesis construct was co-injected with 35 ng/µl Tol2 mRNA into zebrafish embryos at the 1 cell stage. Embryos were *Tg(fli1:eGFP)* to enable visualization of blood vessels, in either a wildtype or *casper* (unpigmented) genetic background – the latter to more easily image embryos beyond stages of melanocyte development. Embryos were screened for successful PROS transgene integration from 30 hpf onwards by assessment of *cmlc2:GFP* cardiac fluorescence and/or *mScarlet* fluorescence. Morphologically abnormal embryos at this stage were

discarded, as vascular development is disrupted due to defects in somite and heart development, for example (56). Scoring of vascular malformations was performed at 72 hpf. For studies on older fish, embryos were raised to juvenile stages.

Imaging

Embryos and juvenile fish were anaesthetized with MS:222 and if necessary, placed on 3% methylcellulose. Fluorescent mesoscope images were acquired using a Leica M165FCA fluorescence stereomicroscope fitted with a 0.63x and 1x Plan-Apochromat objective, and a Leica LMT260 XY scanning stage (Leica Microsystems, Germany). Image acquisition was performed using a Leica K8 monochrome cMOS camera in Leica Application Suite software (LASX; version 3.7). Color images were acquired as above but using a Leica KC3 color cMOS camera (Leica, Microsystems, Germany). Images were acquired in Leica Application Suite software (LASX; version 3.7). Movies were generated using Imaris (v9.6.0). Images were processed using FIJI (v2.9.0).

Confocal images were acquired using a 20X/0.75 lens on the multimodal Imaging Platform Dragonfly (Andor technologies, Belfast UK) equipped with 488 and 561 lasers built on a Nikon Eclipse TI-E inverted microscope body with Perfect focus system (Nikon Instruments, Japan. Data were collected in Spinning Disk 25 µm pinhole mode on the iXon 888 EMCCD camera using a Bin of 1 and no frame averaging, and using Andor Fusion acquisition software. Z stacks were collected using the Nikon TiE focus drive. Movies were generated using Imaris (v9.6.0). Images were processed using FIJI (v2.9.0).

Bone and Cartilage staining

To stain bone and cartilage in juvenile fish, the Juvenile No Acid Bone & Cartilage Stain was followed as described (*https://zfin.atlassian.net/wiki/spaces/prot/pages/352157857*) using a 0.02% Alcian Blue 8GX (Merck)/10 mM $MgCl_2$ solution to stain cartilage and then a 0.01% Alizarin red S (Merck) solution to stain mineralized bone.

FACS and library preparation

Tg(fli1:eGFP) embryos were injected with either *tbxta:pik3ca*^{E545K}*-mScarlet* or *tbxta:mScarlet* and >300 embryos per condition raised to 19 hpf. Each batch was dissociated to single cell suspensions using the methods described in (57). Cells from stage matched wild-type embryos were used as negative controls to gate fluorescence. Live cells were sorted using a FACSAria SORP instrument (BD Biosciences UK), with GFP (525/50, BP/488 nm laser), RFP (582/15 BP, 561 nm laser), and DAPI (450/20BP, 405 nm laser) filters into 2% BSA in PBS and immediately processed using the Chromium platform (10x Genomics). Libraries were prepared using v3.1 Standard Throughput kits according to 10x protocols (10x Genomics), and quality control and quantification assays performed using High Sensitivity DNA kits on a Bioanalyzer (Agilent). Libraries were initially sequenced on an Illumina NovaSeq SP Flow cell with 100 cycles with two lanes per sample. The imbalances in cell number/sequencing depth between the two samples necessitated re-sequencing of the PROS sample as before. This resulted in an average of 120,316 reads/cell after normalization.

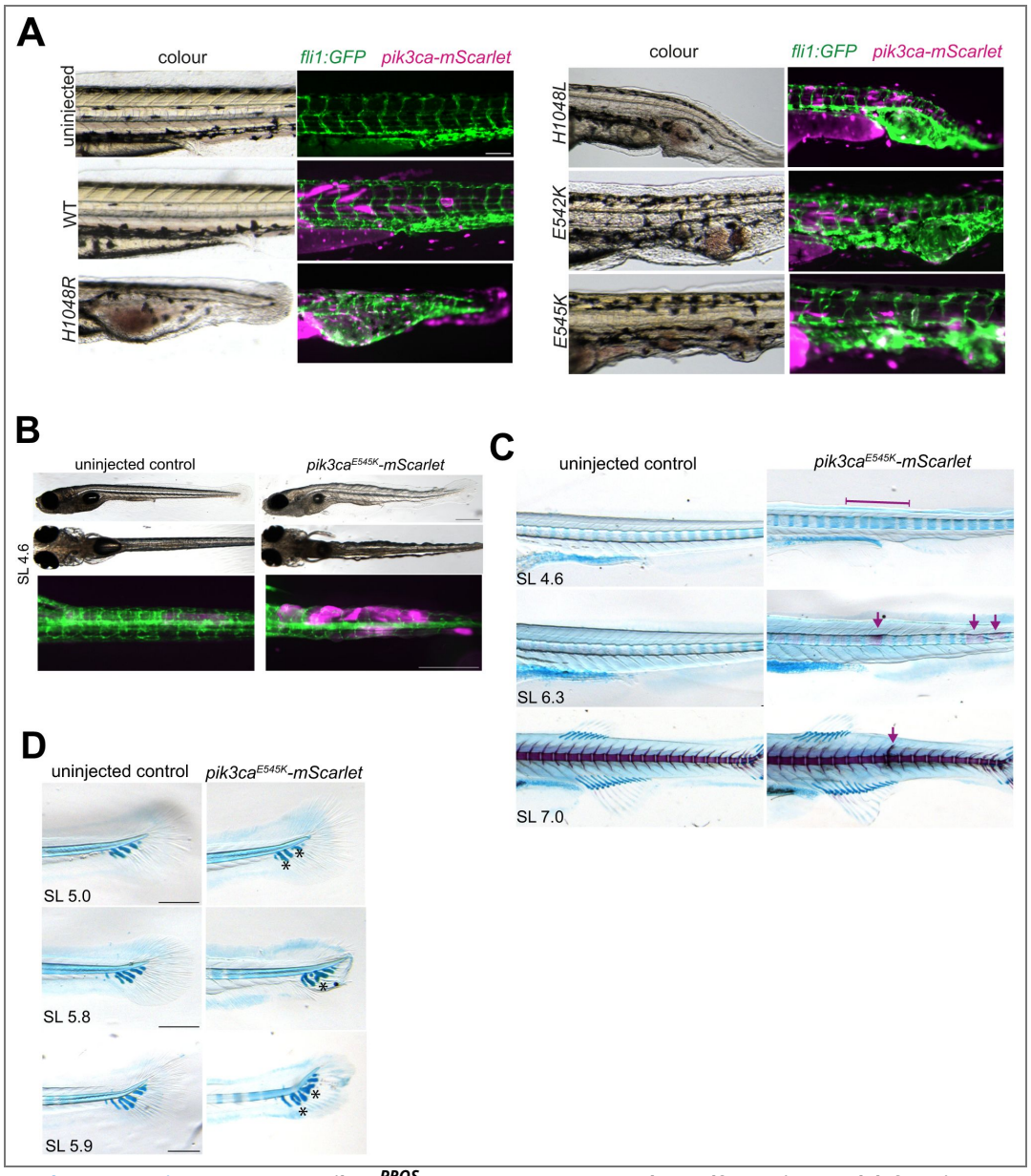
Bioinformatics

FastQ files were aligned using the CellRanger (v8.0.1, 10x Genomics) mkref, mkgtf and aggr pipelines to custom zebrafish STAR genome index using annotations from Ensembl GRCz11 release 112 with manual addition of *eGFP* and *mScarlet*. The resultant h5 matrices were read into R (v4.4.2) and subjected to quality control using Seurat (58) guidelines (v5.2.1, features >200, <13000 and mitochondrial gene < 5%). This resulted in 10306 cells from PROS embryos and 3505 from control. Data were scaled and regressed to cell cycle stage before performing PCA, finding the dimensionality of the data by elbow plot (dims=50) to plot 40 clusters on UMAP plots (SeuratExtend (v1.1.4 (59)). SingleR (60) (v2.8.0) was used with previously published datasets GEO: GSE112294 (35) and NCBI SRA PRJNA940501 (36) to call clusters. DE analyses were performed using zinbwave (v 1.28.0 (61)) and DESeq2 (v1.46.0(62)), selecting genes with padj<0.05 before using clusterProfiler (v4.14.3 (63)) to perform gene enrichment analyses searching Reactome,

GO:BP and KEGG libraries, filtering by p.adj <0.05 (raw data are in Supplementary Table 2). Additional footprint enrichment analyses were conducted using DecoupleR (v. 2.12.0 (64)) using the inbuilt Progeny databases and zebrafish Hallmark datasets imported from MSigdDB according to package guidelines, with Seurat's FindAllMarkers with wilcox testing to determine changes (p.adj < 0.1, p<0.05) to pathway enrichment. Proportions of cells within cluster for PROS and control data were found by dividing the number of cells per cluster/total cell number per experimental condition, with Pearson's chi-squared test (two-tailed) used to determine significance (raw data are in Supplementary Table 3). RNA velocity analysis was performed on loom files for control and PROS data generated with velocyto (v. 0.17 (65)), and velocity computed using scVelo's (v.0.2.5 (66)) dynamical modelling approach, and visualized using heatmaps and stream plots.

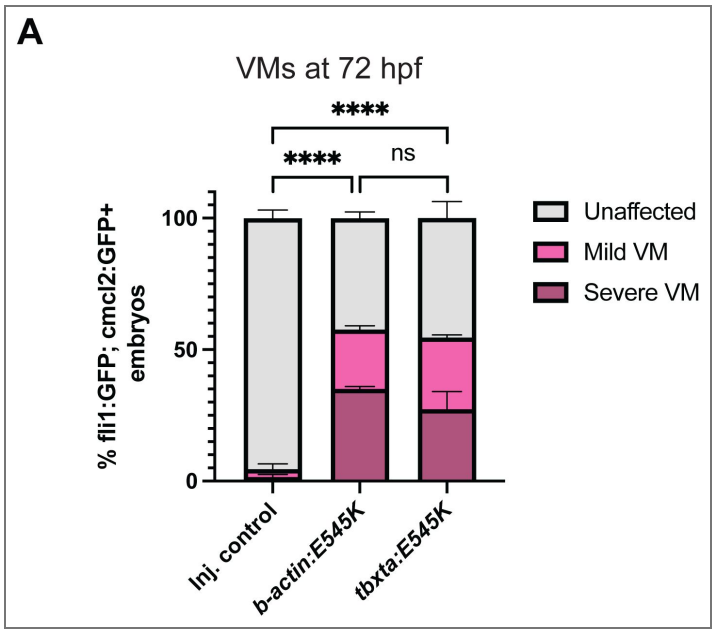
CellChat (v. 2.1.2 (37)) was used to infer potential signaling events between clusters according to relevant package guidelines. NICHES (v. 1.0.0 (38, 67)) was used to perform ligand-receptor expression analysis at single cell resolution, with differential expression analysis within Seurat, using the MAST test, performed to identify significant changes to LR gene expression (raw data are in Supplementary Table 4). Additionally, circuit diagrams were generated using code from *https://github.com/RaredonLab/NICHESMethods*. Plots were generated using Seurat, CellChat, NICHES, Complex Heatmap (v. 2.22.0 (68)) and ggplot2 (v. 4.0.0).

Supplementary figures



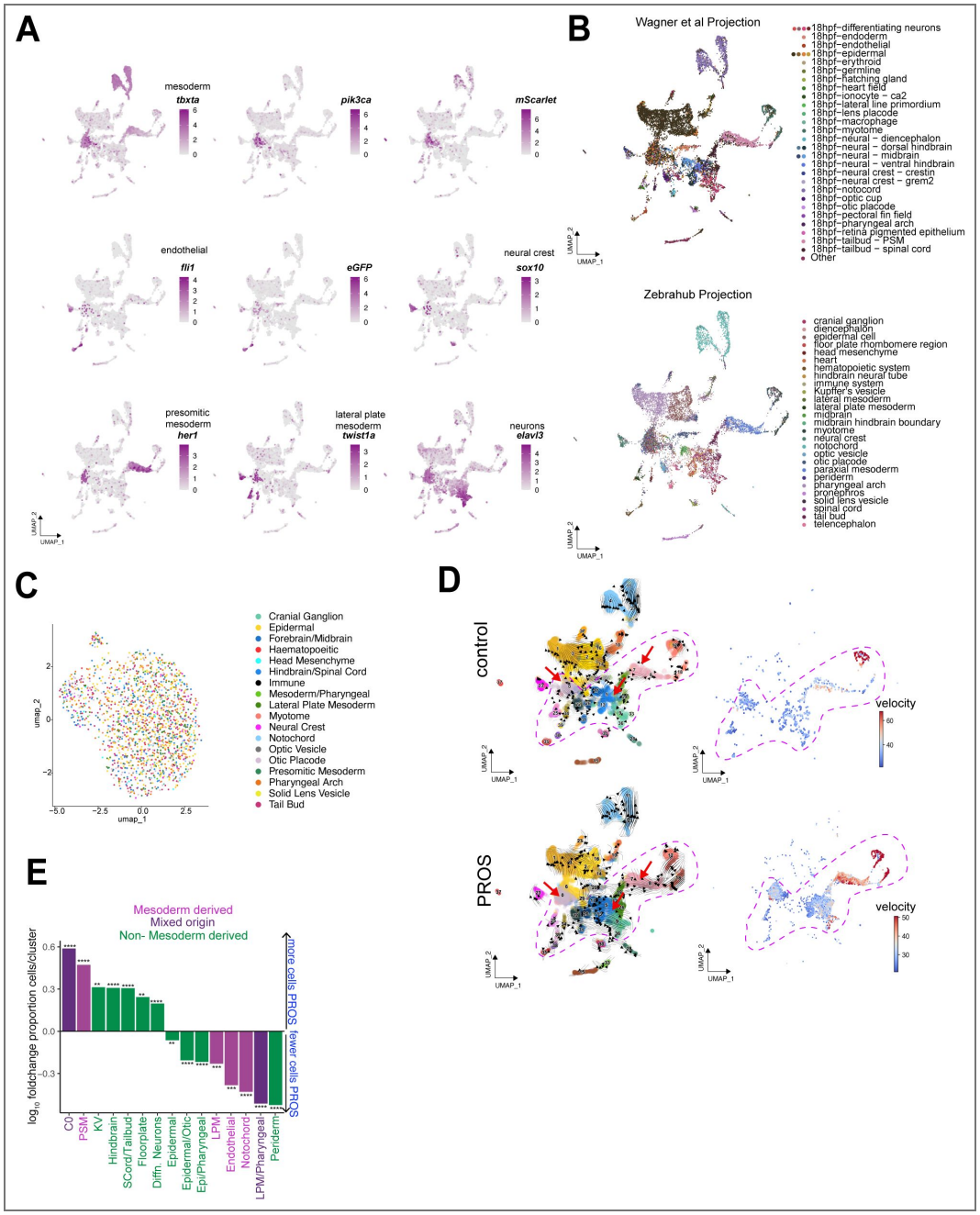
Supplementary Figure 1. Hotspot *pik3ca*^{PROS} transgenes cause vascular malformations and defects in muscle, cartilage and skeletal patterning in larval fish A) Representative images of severe vascular malformations found in 72 hpf *Tg(fli1:eGFP)* embryos previously injected with either *β-actin:pik3caH1048R*, *H1048L, E542K* or *E545K.* Uninjected embryos, and those injected with WT *pik3ca* showed no phenotype at this timepoint. In *pik3ca*^{PROS} embryos, colour panels on the left show blood pooling in malformed vessels, with matched merged fluorescence images of the same area on the right showing enlarged *fli1:eGFP*+ vessels (green), and variable numbers of mosaic *pik3ca*^{PROS}*-mScarlet* cells (magenta). Scale bar = 100 μm. B)*Tg(fli1:eGFP);casper* fish injected with *β-actin:pik3ca*^{E545K}*-mScarlet* but negative for vascular malformations at 72 hpf developed a 'rippled' appearance by two weeks of age (SL4.6) caused by protrusions of *pik3ca*^{PROS}*-mScarlet* cells in the embryo flank (10/25). Scale bar = 100 μm. C) Patterning defects in the developing spine of staged matched *Tg(β-actin:pik3ca*^{E545K}*-mScarlet; fli1:eGFP); casper* injected embryos, showing irregular spacing of cartilage centra during early spine development (3/9 fish in which pattern had developed, prior to bone deposition), premature deposition of bone in patches posterior to stage-matched controls (1/3 fish at a stage where bone deposition is just starting in anterior and so therefore premature deposition is noticeable), and thickened vertebra structure (1/6 fish in which vertebrae were formed) D) Comparison of stage-matched uninjected and *Tg(β-actin:pik3ca*^{E545K}*-*

mScarlet; fli1:eGFP); casper injected fish during tail fin development with patterning defects, revealed by alcian blue staining for cartilage and alizarin red co-staining for bone. 3/9 cartilage hypural patterning defects, 5/7 tail shape defects.



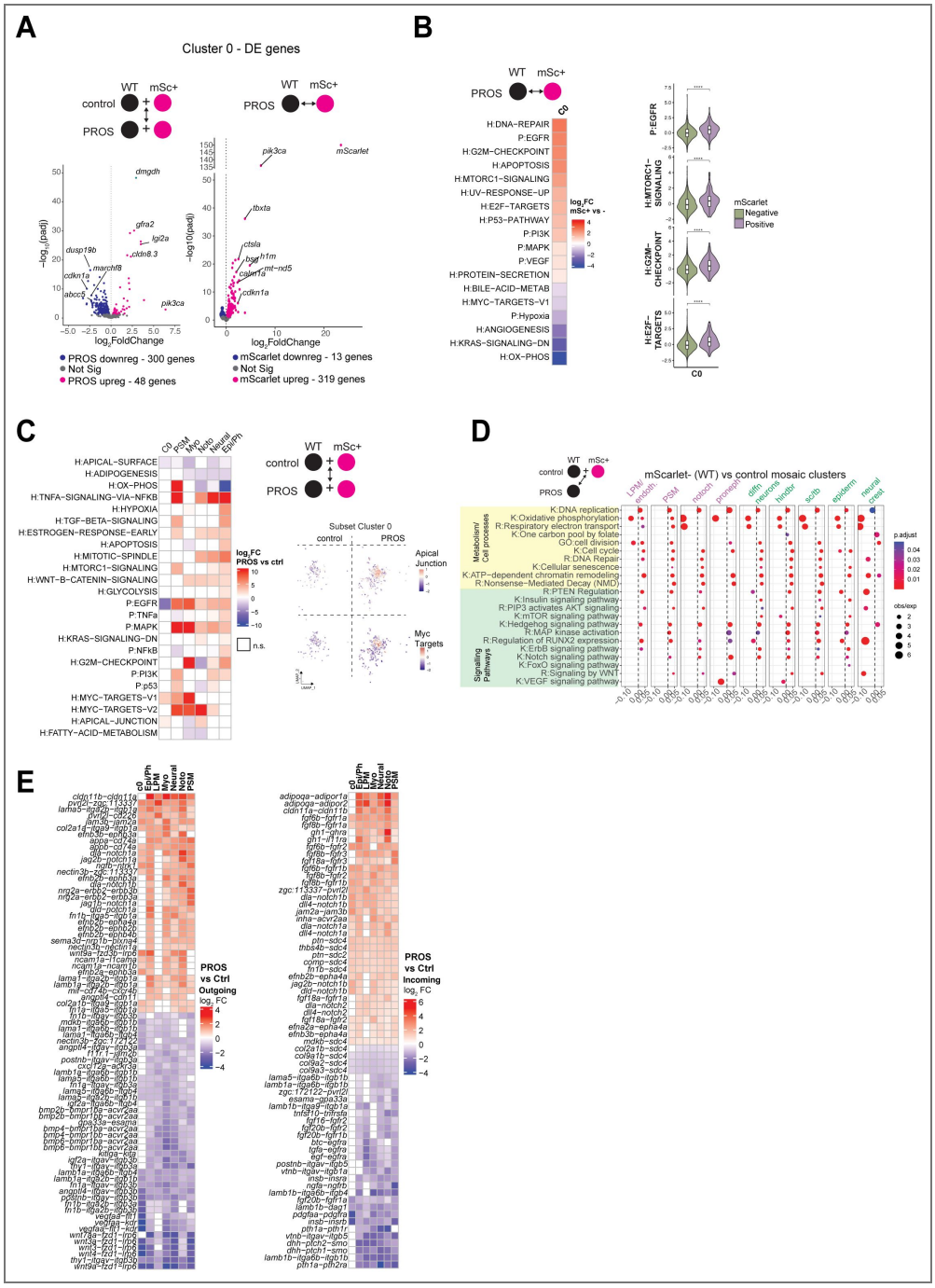
Supplementary Figure 2. Ubiquitous and mesodermal expression of *pik3ca*^{PROS} causes vascular malformations

A) Quantification of vascular phenotypes in 72 hpf Tg(*fli1:eGFP*) embryos injected with *pik3ca*^{E545K}*-mScarlet* constructs driven by *β-actin* or *tbxta*, or an injected no-*pik3ca* control (*tbxta:mScarlet*). Represents 3 biological replicates. Error bars are S.E.M. p-values are from chi:squared tests on raw counts, * p<0.05, ***p<0.001, **** p<0.0001.



Supplementary Figure 3. scRNA seq analysis of *pik3ca*^{PROS} mosaic embryos

A) UMAP FeaturePlots showing expression intensity (purple) of transgenesis or cell lineage markers. B) Mapping of 18hpf (top) and 19 hpf (bottom) published dataset identities onto these data to call clusters. C) UMAP of re-clustered cluster 0, with identities from the zebrahub 19 hpf dataset mapped as in B. D) Single cell velocity analysis was performed on PROS and control data, and the vector field overlaid on a UMAP projection, with arrow indicating velocity dynamics and strength across clusters. The greatest variance in arrow length and density between PROS and control is found in the cluster groups outlined (pink dashed). The velocity length of this subset is displayed on a heatmap, with cells showing the highest abundances of unspliced mRNA in red. E) As Figure 4E, but to show the extent of the statistically significant changes in the proportion of cells within cluster groups. Two-sided Pearson's chi-squared proportion test. ** p<0.01, ***p<0.001, ****p<0.0001



Supplementary Figure 4. Mosaic *pik3ca*^{PROS} causes pan-lineage transcriptional changes, including to ligand-receptor signalling genes.

A) Volcano plots showing significantly differentially expressed genes after DE analysis of cells within cluster 0. Left – comparison of all cluster 0 PROS mosaics vs control, Right – mScarlet+ vs mScarlet-cells in PROS mosaics only. B) Heatmap showing fold change enrichment (Seurat wilcox test p<0.05, p.adj<0.1) of Hallmark (H:) and Progeny (P:) pathways in PROS mosaic cluster 0 mSc+ vs mSc-cells. Violin plots confirm significant upregulation of selected Hallmark signatures in mScarlet+ cells (wilcox test, **** = p<0.001) C) Heatmap showing fold change enrichment (Seurat wilcox test p<0.05, p.adj<0.1) of Hallmark (H:) and Progeny (P:) pathways in PROS mosaics relative to control N.S = non-significant, and Cluster 0 UMAP subset showing expression intensity of Hallmark terms, split by experimental condition. D) As in Figure 4F, but comparing only *mScarlet*- (ie, WT) cells between experimental conditions. E) Heatmaps depicting $\log_2$ fold change in expression of the top differentially expressed "sending" (left) and "receiving" (right) ligand-receptor gene pairs in PROS relative to control embryos, clustered by row. Only differences for which p.adj<0.05 by Wilcoxon test between PROS and WT are shown.

Data availability

All data used to prepare figures including code has been deposited at *https://osf.io/q7kpy/* or is included as Supplementary Tables referenced in this manuscript, with links to access available upon request. single cell RNA seq data has been deposited at GEO and will become public upon publication.

Acknowledgements

This work is supported by funding from the CLOVES Syndrome Community [c-13277479] and the UK Medical Research Council [MR/Z506321/1]. We are grateful to Kristen Davies and Seth Haddix from the CLOVES Syndrome Community for discussions about the experimental aims and design of this work, and to Mihaly Badonyi and Joe Marsh (MRC Human Genetics Unit) for assistance with Pymol. RKS is also supported by the BHF Centre for Research Excellence Award III [RE/18/5/34216]. RRM is funded by a UKRI Future Leaders Fellowship [MR/Y017439/1]. MSBR and NW are funded by T32GM086287 from the NIGMS, and laboratory startup funds from the Yale School of Medicine and the Yale Department of Anesthesiology. EEP is funded by the Medical Research Council (MC_UU_00035/13) and the Cancer Research UK Scotland Centre (CTRQQR-2021\100006).

Additional information

Author Contributions

Conceptualization: HB, RKS, EEP; Data curation: HB; Code checking: RRM, NW, MSBR; Formal analysis: HB; Funding acquisition: RRM, MSBR, RKS, EEP. Investigation: HB; Methodology: HB, RRM, MSBR, EEP; Project administration: RKS, EEP; Supervision: MSBR, RKS, EEP; Visualization: HB; Writing – original draft: HB, RRM, RKS, EEP; Writing – review & editing: All authors

Funding

Funder	Grant reference number	Author
CLOVES Syndrome Community	c-13277479	Hannah Brunsdon
UK Medical Research Council	MR/Z506321/1	Hannah Brunsdon Ralitsa R Madsen
British Heart Foundation	Centre for Research Excellence Award III RE/18/5/34216	Robert K Semple
UK Research and Innovation/institution	Future Leaders Fellowship MR/Y017439/1	Ralitsa R Madsen
National Institute of General Medical Sciences	T32GM086287	Nuoya Wang Micha Sam Brickman Raredon
UKRI | Medical Research Council (MRC)	MC_UU_00035/13	E Elizabeth Patton
Cancer Research UK Scotland Centre	CTRQQR-2021\100006	E Elizabeth Patton

Author ORCID iDs

Hannah Brunsdon: https://orcid.org/0000-0002-8363-6500
E Elizabeth Patton: https://orcid.org/0000-0002-2570-0834

Additional files

Supplementary Table 1. Primers used for plasmid construction.

Supplementary Table 2. Raw results from Deseq2 DE analyses used to generate Figure 4F.

Supplementary Table 3. Raw values used to generate proportion plots in Figures 4E and Supplementary Figure 3E.

Supplementary Table 4. Raw results from DE analyses on NICHES signalling objects used to generate Figure 5C.

Movie 1. Maximum projection movie of Figure 4A of uninjected *Tg(fli1:eGFP)* embryo showing the arrangement of vessels around the central notochord.

Movie 2. Maximum projection movie of a Mild vascular lesion in a *Tg(fli1:eGFP)* embryo injected with *pik3ca*^{E545K}*-mScarlet* (see Figure 4B) showing first separated *pik3ca*^{PROS} *mScarlet* expression (magenta) and then merged with *fli1:eGFP*.

Movie 3. Maximum projection movie of a Moderate vascular lesion in a *Tg(fli1:eGFP)* embryo injected with *pik3ca*^{E545K}*-mScarlet* (see Figure 4B) showing first separated *pik3ca*^{PROS} *mScarlet* expression (magenta) and then merged with *fli1:eGFP.* Cells co-expressing *GFP* and *mScarlet* are white.

Movie 4. Timelapse of injected mosaic mosaic *tbxta:mScarlet Tg(fli1:eGFP)* control embryo between 16-28.5 hpf, showing normal development of blood vessels, with ISV sprouting and DLAV formation from anterior (left) to posterior (right).

Movie 5. Timelapse of injected mosaic *pik3ca*^{PROS} *Tg(fli1:eGFP)* embryo between 16-28.5 hpf showing ectopic *fli1:eGFP*+ endothelial cells contributing to premature formation of the DLAV, and defects in ISV sprouting.

References

1. **Madsen R. R.,** Toker A. (2023) PI3K signaling through a biochemical systems lens. *Journal of Biological Chemistry* **299**:105224 https://doi.org/10.1016/j.jbc.2023.105224 | PubMed
2. **Fruman D. A.,** et al. (2017) The PI3K Pathway in Human Disease. *Cell* **170**:605-635 https://doi.org/10.1016/j.cell.2017.07.029 | PubMed
3. **Sanchez-Vega F.,** et al. (2018) Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* **173**:321-337.e310 https://doi.org/10.1016/j.cell.2018.03.035 | PubMed
4. **Madsen R. R.,** Vanhaesebroeck B., Semple R. K. (2018) Cancer-Associated PIK3CA Mutations in Overgrowth Disorders. *Trends in Molecular Medicine* 856-870 https://doi.org/10.1016/j.molmed.2018.08.003 | PubMed
5. **Keppler-Noreuil K. M.,** et al. (2014) Clinical delineation and natural history of the PIK3CA-related overgrowth spectrum. *American Journal of Medical Genetics, Part A* **164**:1713-1733 https://doi.org/10.1002/ajmg.a.36552 | PubMed
6. **Keppler-Noreuil K. M.,** et al. (2015) PIK3CA-related overgrowth spectrum (PROS): Diagnostic and testing eligibility criteria, differential diagnosis, and evaluation. *American Journal of Medical Genetics, Part A* **167**:287-295 https://doi.org/10.1002/ajmg.a.36836 | PubMed
7. **Martinez-Lopez A.,** et al. (2017) CLOVES syndrome: review of a PIK3CA-related overgrowth spectrum (PROS). *Clinical Genetics* **91**:14-21 https://doi.org/10.1111/cge.12832 | PubMed
8. **De Bortoli M.,** et al. (2024) Somatic Loss-of-Function PIK3R1 and Activating Non-hotspot PIK3CA Mutations Associated with Capillary Malformation with Dilated Veins (CMDV). *J Invest Dermatol* **144**:2066-2077.e2066 https://doi.org/10.1016/j.jid.2024.01.033 | PubMed
9. **Nriagu B. N.,** et al. (2024) Microcystic lymphatic malformations in Turner syndrome are due to somatic mosaicism of PIK3CA. *Am J Med Genet A* **194**:64-69 https://doi.org/10.1002/ajmg.a.63385 | PubMed
10. **Li D.,** et al. (2023) Genomic profiling informs diagnoses and treatment in vascular anomalies. *Nat Med* **29**:1530-1539 https://doi.org/10.1038/s41591-023-02364-x | PubMed
11. **Petkova M.,** et al. (2023) Immune-interacting lymphatic endothelial subtype at capillary terminals drives lymphatic malformation. *J Exp Med* **220** https://doi.org/10.1084/jem.20220741 | PubMed

12. **Kutys M. L.,** et al. (2020) Uncovering mutation-specific morphogenic phenotypes and paracrine-mediated vessel dysfunction in a biomimetic vascularized mammary duct platform. *Nat Commun* **11**:3377 https://doi.org/10.1038/s41467-020-17102-x | PubMed

13. **Koundouros N.,** et al. (2020) Metabolic Fingerprinting Links Oncogenic PIK3CA with Enhanced Arachidonic Acid-Derived Eicosanoids. *Cell* **181**:1596-1611.e1527 https://doi.org/10.1016/j.cell.2020.05.053 | PubMed

14. **Venot Q.,** et al. (2025) PIK3CA gain-of-function mutation in Schwann cells leads to severe neuropathy and aerobic glycolysis through a non-cell autonomous effect. *Proc Natl Acad Sci U S A* **122**:e2424867122 https://doi.org/10.1073/pnas.2424867122 | PubMed

15. **European Medicines Agency** (2023) EMA/492147/2023 EMEA/H/C/005468 Withdrawal of application for the marketing authorisation of Vijoice (alpelisib). European Medicines Agency.

16. **Canaud G.,** et al. (2023) Alpelisib for treatment of patients with PIK3CA-related overgrowth spectrum (PROS). *Genet Med* **25**:100969 https://doi.org/10.1016/j.gim.2023.100969 | PubMed

17. **Seront E.,** Queisser A., Boon L. M., Vikkula M. (2024) Molecular landscape and classification of vascular anomalies. *Hematology Am Soc Hematol Educ Program* **2024**:700-708 https://doi.org/10.1182/hematology.2024000598 | PubMed

18. **Castillo S. D.,** et al. (2016) Somatic activating mutations in Pik3ca cause sporadic venous malformations in mice and humans. *Science Translational Medicine* **8** https://doi.org/10.1126/scitranslmed.aad9982 | PubMed

19. **Kinross K. M.,** et al. (2015) Ubiquitous expression of the Pik3caH1047R mutation promotes hypoglycemia, hypoinsulinemia, and organomegaly. *FASEB J* **29**:1426-1434 https://doi.org/10.1096/fj.14-262782 | PubMed

20. **Berenjeno I. M.,** et al. (2017) Oncogenic PIK3CA induces centrosome amplification and tolerance to genome doubling. *Nat Commun* **8**:1773 https://doi.org/10.1038/s41467-017-02002-4 | PubMed

21. **Roy A.,** et al. (2015) Mouse models of human PIK3CA-related brain overgrowth have acutely treatable epilepsy. *eLife* **4** https://doi.org/10.7554/elife.12703 | PubMed

22. **Castel P.,** et al. (2016) Somatic PIK3CA mutations as a driver of sporadic venous malformations. *Sci Transl Med* **8**:332ra342 https://doi.org/10.1126/scitranslmed.aaf1164 | PubMed

23. **di Blasio L.,** et al. (2018) PI3K/mTOR inhibition promotes the regression of experimental vascular malformations driven by PIK3CA-activating mutations. *Cell Death Dis* **9**:45 https://doi.org/10.1038/s41419-017-0064-x | PubMed

24. **Sabata H.,** et al. (2025) Context-dependent response of endothelial cells to *PIK3CA* mutation. *bioRxiv* https://doi.org/10.1101/2025.02.25.640041

25. **Martinez-Corral I.,** et al. (2020) Blockade of VEGF-C signaling inhibits lymphatic malformations driven by oncogenic PIK3CA mutation. *Nat Commun* **11**:2869 https://doi.org/10.1038/s41467-020-16496-y | PubMed

26. **Torii S.,** et al. (2025) Embryological cellular origins and hypoxia-mediated mechanisms in PIK3CA-driven refractory vascular malformations. *EMBO Mol Med* **17**:1289-1324 https://doi.org/10.1038/s44321-025-00235-1 | PubMed

27. **Kraft M.,** et al. (2025) Angiopoietin-TIE2 feedforward circuit promotes PIK3CA-driven venous malformations. *Nat Cardiovasc Res* **4**:801-820 https://doi.org/10.1038/s44161-025-00655-9 | PubMed

28. **Blum N.,** Harris M. P. (2023) Localized heterochrony integrates overgrowth potential of oncogenic clones. *DMM Disease Models and Mechanisms* **16**:dmm049793 https://doi.org/10.1242/dmm.049793 | PubMed

29. **Venot Q.,** et al. (2018) Targeted therapy in patients with PIK3CA-related overgrowth syndrome. *Nature* **558**:540-546 https://doi.org/10.1038/s41586-018-0217-9 | PubMed

30. **Patton E. E.,** Zon L. I., Langenau D. M. (2021) Zebrafish disease models in drug discovery: from preclinical modelling to clinical trials. *Nat Rev Drug Discov* **20**:611-628 https://doi.org/10.1038/s41573-021-00210-8 | PubMed

31. **Lawson N. D.,** Weinstein B. M. (2002) In vivo imaging of embryonic vascular development using transgenic zebrafish. *Dev Biol* **248**:307-318 https://doi.org/10.1006/dbio.2002.0711 | PubMed

32. **Zhou W.,** et al. (2024) Akt is a mediator of artery specification during zebrafish development. *Development (Cambridge, England)* **151** https://doi.org/10.1242/dev.202727 | PubMed

33. **Greenspan L. J.,** Weinstein B. M. (2021) To be or not to be: endothelial cell plasticity in development, repair, and disease. *Angiogenesis* **24**:251-269 https://doi.org/10.1007/s10456-020-09761-7 | PubMed

34. **Thisse B. P.,** Fürthauer S., Loppin M., Heyer B., Degrave V., Woehl A., Lux R., Steffan A., Charbonnier T., Thisse C. (2001) Expression of the zebrafish genome during embryogenesis (NIH R01 RR15402). ZFIN Direct Data Submission. in *ZFIN Direct Data Submission*. http://zfin.org

35. **Wagner D. E.,** et al. (2018) Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**:981-987 https://doi.org/10.1126/science.aar4362 | PubMed

36. **Lange M.,** et al. (2024) A multimodal zebrafish developmental atlas reveals the state-transition dynamics of late-vertebrate pluripotent axial progenitors. *Cell* **187**:6742-6759.e6717 https://doi.org/10.1016/j.cell.2024.09.047 | PubMed

37. **Jin S.,** et al. (2021) Inference and analysis of cell-cell communication using CellChat. *Nat Commun* **12**:1088 https://doi.org/10.1038/s41467-021-21246-9 | PubMed

38. **Raredon M. S. B.,** et al. (2023) Comprehensive visualization of cell-cell interactions in single-cell and spatial transcriptomics with NICHES. *Bioinformatics* **39** https://doi.org/10.1093/bioinformatics/btac775 | PubMed

39. **Lindhurst M. J.,** et al. (2012) Mosaic overgrowth with fibroadipose hyperplasia is caused by somatic activating mutations in PIK3CA. *Nat Genet* **44**:928-933 https://doi.org/10.1038/ng.2332 | PubMed

40. **Riviere J. B.,** et al. (2012) De novo germline and postzygotic mutations in AKT3, PIK3R2 and PIK3CA cause a spectrum of related megalencephaly syndromes. *Nat Genet* **44**:934-940 https://doi.org/10.1038/ng.2331 | PubMed

41. **Madsen R. R.,** Semple R. K. (2022) PIK3CA-related overgrowth: silver bullets from the cancer arsenal?. *Trends Mol Med* **28**:255-257 https://doi.org/10.1016/j.molmed.2022.02.009 | PubMed

42. **Castillo S. D.,** et al. (2016) Somatic activating mutations in Pik3ca cause sporadic venous malformations in mice and humans. *Sci Transl Med* **8**:332ra343 https://doi.org/10.1126/scitranslmed.aad9982 | PubMed

43. **Rodriguez-Laguna L.,** et al. (2019) Somatic activating mutations in PIK3CA cause generalized lymphatic anomaly. *J Exp Med* **216**:407-418 https://doi.org/10.1084/jem.20181353 | PubMed

44. **Li D.,** et al. (2019) ARAF recurrent mutation causes central conducting lymphatic anomaly treatable with a MEK inhibitor. *Nat Med* **25**:1116-1122 https://doi.org/10.1038/s41591-019-0479-2 | PubMed

45. **Martin-Valiente E.,** et al. (2025) Reversal of cerebrovascular anomalies in a zebrafish model of vein of Galen aneurysm. *Nat Cardiovasc Res* **4**:773-789 https://doi.org/10.1038/s44161-025-00659-5 | PubMed

46. **Madsen R. R.,** et al. (2019) Oncogenic PIK3CA promotes cellular stemness in an allele dose-dependent manner. *Proceedings of the National Academy of Sciences of the United States of America* **116**:8380-8389 https://doi.org/10.1073/pnas.1821093116 | PubMed

47. **Lindhurst M. J.,** et al. (2019) A mouse model of Proteus syndrome. *Hum Mol Genet* **28**:2920-2936 https://doi.org/10.1093/hmg/ddz116 | PubMed

48. **Marechal E.,** et al. (2022) Multiple congenital malformations arise from somatic mosaicism for constitutively active Pik3ca signaling. *Frontiers in Cell and Developmental Biology* **10** https://doi.org/10.3389/fcell.2022.1013001 | PubMed

49. **Sahai-Hernandez P.,** et al. (2023) Dermomyotome-derived endothelial cells migrate to the dorsal aorta to support hematopoietic stem cell emergence. *eLife* **12** https://doi.org/10.7554/elife.58300 | PubMed

50. **Villefranc J. A.,** Amigo J., Lawson N. D. (2007) Gateway compatible vectors for analysis of gene function in the zebrafish. *Dev Dyn* **236**:3077-3087 https://doi.org/10.1002/dvdy.21354 | PubMed

51. **White R. M.,** et al. (2008) Transparent adult zebrafish as a tool for in vivo transplantation analysis. *Cell Stem Cell* **2**:183-189 https://doi.org/10.1016/j.stem.2007.11.002 | PubMed

52. **Kimmel C. B.,** Ballard W. W., Kimmel S. R., Ullmann B., Schilling T. F. (1995) Stages of embryonic development of the zebrafish. *Dev Dyn* **203**:253-310 https://doi.org/10.1002/aja.1002030302 | PubMed

53. **Parichy D. M.,** Elizondo M. R., Mills M. G., Gordon T. N., Engeszer R. E. (2009) Normal table of postembryonic zebrafish development: staging by externally visible anatomy of the living fish. *Dev Dyn* **238**:2975-3015 https://doi.org/10.1002/dvdy.22113 | PubMed

54. **Kwan K. M.,** et al. (2007) The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev Dyn* **236**:3088-3099 https://doi.org/10.1002/dvdy.21343 | PubMed

55. **Harvey S. A.,** Tumpel S., Dubrulle J., Schier A. F., Smith J. C. (2010) no tail integrates two modes of mesoderm induction. *Development* **137**:1127-1135 https://doi.org/10.1242/dev.046318 | PubMed

56. **Stratman A. N.,** Weinstein B. M. (2021) Assessment of Vascular Patterning in the Zebrafish. *Methods Mol Biol* **2206**:205-222 https://doi.org/10.1007/978-1-0716-0916-3_15 | PubMed

57. **Manoli M.,** Driever W. (2012) Fluorescence-activated cell sorting (FACS) of fluorescently tagged cells from zebrafish larvae for RNA isolation. *Cold Spring Harb Protoc* **2012** https://doi.org/10.1101/pdb.prot069633 | PubMed

58. **Hao Y.,** et al. (2024) Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol* **42**:293-304 https://doi.org/10.1038/s41587-023-01767-y | PubMed

59. **Hua Y.,** Weng L., Zhao F., Rambow F. (2024) SeuratExtend: Streamlining Single-Cell RNA-Seq Analysis Through an Integrated and Intuitive Framework. *bioRxiv* https://doi.org/10.1101/2024.08.01.606144

60. **Aran D.,** et al. (2019) Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol* **20**:163-172 https://doi.org/10.1038/s41590-018-0276-y | PubMed

61. **Risso D.,** Perraudeau F., Gribkova S., Dudoit S., Vert J. P. (2019) Publisher Correction: A general and flexible method for signal extraction from single-cell RNA-seq data. *Nat Commun* **10**:646 https://doi.org/10.1038/s41467-019-08614-2 | PubMed

62. **Love M. I.,** Huber W., Anders S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**:550 https://doi.org/10.1186/s13059-014-0550-8 | PubMed

63. **Wu T.,** et al. (2021) clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* **2**:100141 https://doi.org/10.1016/j.xinn.2021.100141 | PubMed

64. **Badia I. M. P.,** et al. (2022) decoupleR: ensemble of computational methods to infer biological activities from omics data. *Bioinform Adv* **2**:vbac016 https://doi.org/10.7490/f1000research.1119067.1

65. **La Manno G.,** et al. (2018) RNA velocity of single cells. *Nature* **560**:494-498 https://doi.org/10.1038/s41586-018-0414-6 | PubMed

66. **Bergen V.,** Lange M., Peidli S., Wolf F. A., Theis F. J. (2020) Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol* **38**:1408-1414 https://doi.org/10.1038/s41587-020-0591-3 | PubMed

67. **Greaney A. M.,** et al. (2024) Engineered Whole Lungs for Tissue Biology. *bioRxiv* https://doi.org/10.1101/2024.10.02.616240

68. **Gu Z.,** Eils R., Schlesner M. (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**:2847-2849 https://doi.org/10.1093/bioinformatics/btw313 | PubMed

Wagner DE, Weinreb C, Collins ZM, Megason SG, Klein AM (2018) Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. NCBI Gene Expression Omnibus. ID GSE112294 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE112294

Lange M et al (2024) A multimodal zebrafish developmental atlas reveals the state-transition dynamics of late-vertebrate pluripotent axial progenitors. NCBI BioProject. ID PRJNA940501 https://www.ncbi.nlm.nih.gov/bioproject/PRJNA940501

Peer reviews

Reviewer #1 (Public review):

Summary:

Brunsdon et al. present a zebrafish model of mosaic PIK3CA activation to investigate mechanisms underlying PIK3CA-related overgrowth spectrum (PROS), with a particular focus on non-cell-autonomous mechanisms of tissue overgrowth. The study is timely and addresses an important gap in the understanding of how mosaic activation of PI3K signaling leads to tissue-specific developmental abnormalities.

Using a Tol2-based mosaic expression system combined with single-cell transcriptomics, the authors provide evidence suggesting that mutant PIK3CA-expressing cells influence surrounding wild-type tissues through indirect signaling mechanisms, contributing to vascular malformations and tissue overgrowth.

Overall, the work presents an interesting and potentially impactful model for studying mosaic PIK3CA-driven overgrowth and non-cell-autonomous signaling mechanisms. However, several aspects require clarification, additional controls, and improved presentation to strengthen the mechanistic conclusions and overall impact of the study.

Strengths:

This study addresses an important and timely question by investigating the mechanisms underlying mosaic PIK3CA activation in the context of PROS, a condition for which developmental mechanisms remain poorly understood. The use of a mosaic zebrafish model is particularly appropriate, as it closely reflects the mosaic nature of PIK3CA mutations observed in patients and allows the investigation of non-cell-autonomous effects.

Another major strength of the study is the integration of single-cell transcriptomics, which provides valuable insight into potential signaling pathways involved in indirect tissue overgrowth and offers a rich dataset for hypothesis generation. The authors also propose an interesting conceptual framework in which PI3K-activated cells influence surrounding tissues through paracrine signaling, which could have broader implications beyond PROS and contribute to understanding mosaic developmental disorders more generally.

Finally, the work has potential translational relevance, as identifying mechanisms driving mosaic PI3K activation and non-cell-autonomous signaling could inform future therapeutic strategies for PROS and related conditions.

Weaknesses:

Despite these strengths, several aspects of the study require clarification and additional experimentation.

Major comments:

(1) The Tol2-based system results in mosaic overexpression of mutant PIK3CA in the presence of endogenous wild-type PIK3CA, making it difficult to determine how co-expression of WT

and mutant proteins influences the observed phenotypes. While mosaic expression is relevant to PROS, a complementary approach in which endogenous PIK3CA is knocked out prior to introducing mutant variants would allow clearer interpretation of mutant-specific effects.

(2) The authors do not clearly describe the validation of editing or integration efficiency. It would be important for the authors to clarify whether sequencing was performed to confirm integration, to quantify the proportion of mosaic expression, and to measure transgene expression levels. These controls would strengthen confidence in the model and interpretation of the results.

(3) The manuscript would benefit from rescue experiments to strengthen causal conclusions. It remains unclear whether the phenotypes induced by PIK3CA PROS variants can be rescued, either through expression of wild-type PIK3CA, pharmacological inhibition of PI3K signaling, or assessment of developmental reversibility. Such experiments would strengthen the link between PI3K activation and the observed phenotypes.

(4) The authors propose candidate signaling molecules mediating non-cell-autonomous effects downstream of PI3K hyperactivation; however, these conclusions remain speculative, as no functional validation is provided. Testing selected candidate mediators identified in the RNA-seq dataset would significantly strengthen the mechanistic conclusions.

https://doi.org/10.7554/eLife.110896.1.sa3

Reviewer #2 (Public review):

In this manuscript, Burnsdon et al. aim to study PIK3CA-related overgrowth spectrum (PROS) by establishing a mosaic zebrafish model with overexpression of pik3ca carrying hotspot mutations, coupled with an mScarlet+ reporter. Using fluorescence microscopy, the authors demonstrated that overexpression of pik3ca with a number of hotspot mutations led to mesodermal and particularly vascular malformations in the zebrafish model. Interestingly, they found a paucity of mScarlet+ mutant cells in the vascular lesions, consistent with the finding of low PIK3CA mutation burden in PROS tissue. Such data suggest a non-cell-autonomous effect of PIK3CA mutation. Following this logic, the authors performed single-cell RNA-Sequencing on zebrafish overexpressing WT pik3ca and mutant pik3ca at 19 hpf, and demonstrated widespread transcriptomic perturbations across multiple lineages, including lineage frequencies, key cell pathways, and cell-cell interactions. Importantly, they demonstrate that mScarlet+ cells carrying mutant pik3ca cluster separately from other cell types, do not demonstrate clear lineage identity, and have a general downregulation in signaling components.

Overall, the conclusions in the manuscript are well-supported by the presented data. The imaging studies are particularly convincing. The transcriptomic analysis generated a list of potential pathways to further investigate and potentially target with future therapeutic interventions. Importantly, this study provides a valuable in vivo model of PROS that: 1) recapitulates key features of PROS (e.g., multiple mesodermal defects, paucity of mutation burden in lesions suggesting non-cell-autonomous interactions); 2) is scalable; and 3) offers direct visualization of lesion development, compatible with time-course live imaging. This model will be valuable to further understand PROS and potentially study other diseases where the PIK3CA pathway is altered (e.g., certain cancers).

The following are not necessarily weaknesses of the data, but rather suggestions where the manuscript could be further strengthened:

(1) The model recapitulates the variability of mesodermal lesions in PROS. It would be valuable to utilize this model to further study factors that are associated with the

development of more severe lesions (e.g., by comparing samples with more severe lesions to those unaffected despite carrying the mutations, Figure 1F).

(2) ScRNA-seq analysis could be enriched with a comparison between cells overexpressing mutant pik3ca vs. those overexpressing WT pik3ca.

(3) In the scRNA-Seq analysis, it is curious that the C0 cluster, enriched with mScarlet+ cells, is found to have downregulated signaling interactions (Fig. 5C), yet exerts a widespread non-cell-autonomous effect. Meanwhile, there is also a noticeable loss of certain lineages (e.g., notochord, Figure 4E) and related cell-cell interactions (e.g., notochord-related interaction, Figure 5A). A deeper exploration of the basis of the non-cell-autonomous effect would be valuable.

(4) The scRNA-Seq analysis was performed at one time point (19 hpf). Additional analysis (not necessarily by scRNA-Seq) at other time points to study whether findings at 19 hpf are persistent throughout development or undergo dynamic changes (e.g., cell fate/state of mSc+ mutant cells) would be helpful.

(5) The scRNA-Seq analysis provides a valuable list of perturbed interactions that could be targeted by future therapeutic approaches. Validation of the scRNA-Seq findings with protein-level analysis, and studying the effect of targeting some of the pathways on the disease phenotype, would offer valuable data for the community.

https://doi.org/10.7554/eLife.110896.1.sa2

Reviewer #3 (Public review):

Summary:

The study "PIK3CA-related overgrowth spectrum (PROS) zebrafish models reveal pan-lineage developmental dysregulation" presents important findings that extend significantly beyond a single subfield, bridging developmental biology, vascular medicine, and cancer-related PI3K signalling. By developing mosaic zebrafish models of PROS and combining live imaging with single-cell transcriptomics, the authors provide compelling evidence for a non-cell-autonomous mechanism of tissue overgrowth, a conceptual shift with meaningful therapeutic implications.

Strengths:

The evidence is overall convincing, with methodology appropriate and well-validated relative to the current state of the art; the integration of multiple approaches (in vivo modelling, scRNA-seq, ligand-receptor inference) strengthens the central claims. However, some aspects of the proposed non-cell-autonomous signalling mechanisms remain partly correlative, and direct functional validation of the rewired ligand-receptor interactions would further consolidate the conclusions.

Weaknesses:

The transgenic overexpression approach chosen by the authors represents a well-established and effective strategy for generating mosaic models in zebrafish. However, this approach introduces notable limitations: the lack of control over transgene dosage and unknown integration sites may generate non-physiological effects, potentially confounding the interpretation of key findings.

The authors are certainly aware that alternative approaches (though technically more demanding) could be considered in future studies to further strengthen the model. For instance, a CRISPR/Cas9-mediated knock-in of the pik3ca-PROS allele at the endogenous locus

(retaining upstream native regulatory elements with only a minimal promoter in the construct, co-expressed with a fluorescent reporter via P2A) could allow even more physiological, lineage-restricted expression while enabling direct visualisation of mutant cells. Mesodermal specificity could potentially be further refined by driving mosaic Cas9 expression under a pan-mesodermal tbx promoter, restricting editing to the relevant lineage while simultaneously marking mutant cells fluorescently, thus even more closely mimicking the post-zygotic mutational events characteristic of PROS. As a complementary strategy, blastula transplantation experiments using pik3ca-PROS donor cells (ideally co-expressing a distinct fluorescent marker such as mCherry) into fli1:GFP transgenic hosts could provide a powerful and technically consolidated approach to directly visualise and quantify non-cell-autonomous effects on host vasculature, with precise control over mutant cell burden. This combinatorial framework, separating donor mutant cells from host tissue in a two-colour imaging setup, could be particularly compelling for validating the ligand-receptor rewiring predicted by single-cell transcriptomics in future investigations.

These reflections are offered in the spirit of prospective methodological development and do not diminish the value of the current work, which opens a valuable new avenue for therapeutic investigation, suggesting that targeting indirect overgrowth-propagating signals, alongside PI3K inhibition, deserves serious consideration.

https://doi.org/10.7554/eLife.110896.1.sa1

Author response:

> ***eLife Assessment***
>
> *This is an important study that establishes a zebrafish model of PIK3CA-related overgrowth syndrome. The imaging characterization of the mesodermal, particularly vascular, lesions of the model is compelling. The scRNA-Seq analysis is convincing, revealing key perturbations in the PIK3CA-mutation model, although deeper investigation of the exact mechanism leading to the lesions, as well as validation at different time points, could further strengthen the findings. This work will be of interest to medical biologists working on PROS, and potentially to a broader audience interested in non-cell-autonomous signaling of PIK3CA and its implications in other diseases.*

We are delighted that the Editors and Reviewers consider the work of value and that it is interesting to a broad audience. We also appreciate and take on board the areas that the reviewers identify for improvement, and their suggestions on how this could be achieved.

There are two major pieces of work suggested by the reviewers which we plan to carry out for this manuscript. The first of these is an additional scRNA-seq experiment at a later developmental stage when vascular malformations are established. Through comparison between pik3caPROS, pik3caWT and no-pik3ca injected controls, this would help answer if the global lineage and transcriptional dysregulation observed at 19 hpf persists over time, and if the largely inert 'C0' cluster of PROS mScarlet^{+} cells changes during development (Reviewer 2 comment 3).

Secondly, we are already optimising rescue experiments with the specific Pik3ca inhibitor alpelisib, which is currently used as a therapy for PROS. Some troubleshooting has been required for the best delivery method and concentration for this to rescue vascular malformations in embryos, and to cause measurable decreases in PI3K signalling at the protein level through Akt and S6 pathways.

> ***Public Reviews:***
>
> ***Reviewer #1 (Public review):***

> *Summary:*
>
> *Brunsdon et al. present a zebrafish model of mosaic PIK3CA activation to investigate mechanisms underlying PIK3CA-related overgrowth spectrum (PROS), with a particular focus on non-cell-autonomous mechanisms of tissue overgrowth. The study is timely and addresses an important gap in the understanding of how mosaic activation of PI3K signaling leads to tissue-specific developmental abnormalities.*
>
> *Using a Tol2-based mosaic expression system combined with single-cell transcriptomics, the authors provide evidence suggesting that mutant PIK3CA-expressing cells influence surrounding wild-type tissues through indirect signaling mechanisms, contributing to vascular malformations and tissue overgrowth.*
>
> *Overall, the work presents an interesting and potentially impactful model for studying mosaic PIK3CA-driven overgrowth and non-cell-autonomous signaling mechanisms. However, several aspects require clarification, additional controls, and improved presentation to strengthen the mechanistic conclusions and overall impact of the study.*

We thank Reviewer 1 for their support of our work, and constructive and helpful comments.

> *Strengths:*
>
> *This study addresses an important and timely question by investigating the mechanisms underlying mosaic PIK3CA activation in the context of PROS, a condition for which developmental mechanisms remain poorly understood. The use of a mosaic zebrafish model is particularly appropriate, as it closely reflects the mosaic nature of PIK3CA mutations observed in patients and allows the investigation of non-cell-autonomous effects.*
>
> *Another major strength of the study is the integration of single-cell transcriptomics, which provides valuable insight into potential signaling pathways involved in indirect tissue overgrowth and offers a rich dataset for hypothesis generation. The authors also propose an interesting conceptual framework in which PI3K-activated cells influence surrounding tissues through paracrine signaling, which could have broader implications beyond PROS and contribute to understanding mosaic developmental disorders more generally.*
>
> *Finally, the work has potential translational relevance, as identifying mechanisms driving mosaic PI3K activation and non-cell-autonomous signaling could inform future therapeutic strategies for PROS and related conditions.*
>
> *Weaknesses:*
>
> *Despite these strengths, several aspects of the study require clarification and additional experimentation.*
>
> *Major comments:*
>
> *(1) The Tol2-based system results in mosaic overexpression of mutant PIK3CA in the presence of endogenous wild-type PIK3CA, making it difficult to determine how co-expression of WT and mutant proteins influences the observed phenotypes. While mosaic expression is relevant to PROS, a complementary approach in which endogenous PIK3CA is knocked out prior to introducing mutant variants would allow clearer interpretation of mutant-specific effects.*

PROS/CLOVES patients co-express endogenous wild-type and mutant PIK3CA in affected cells, which in turn constitute only a small proportion of cells in affected tissues (Madsen et al.

2018). As our intent was strictly to model human PROS/CLOVES (an aim informed by support from and close collaboration with the CLOVES Syndrome Community, a key patient advocacy group), we designed our model to reflect this as closely as possible. It is not clear to us what translational end would be served by expressing mutants in a null background, interesting though this may be. Given our transgenic strategy, we did experiment with overexpressing wildtype pik3ca as a control for some experiments to test whether overexpression of pik3ca itself drives overgrowth phenotypes, without the presence of hotspot PROS mutations (Figure 3D, Supplementary Figure 1A). We found that ubiquitous or mesodermal overexpression of pik3caWT did not cause vascular malformations or cause the ectopic fli1:eGFP endothelial cell phenotype observed when overexpressing pik3caPROS variants. While not precisely addressing the reviewer's comment, this adds to evidence that increased expression of wildtype pik3ca does not confound the observed gain of function phenotype in the PROS model.

> *(2) The authors do not clearly describe the validation of editing or integration efficiency. It would be important for the authors to clarify whether sequencing was performed to confirm integration, to quantify the proportion of mosaic expression, and to measure transgene expression levels. These controls would strengthen confidence in the model and interpretation of the results.*

We used secondary transgenesis markers, such as the cardiac reporter cmlc2:GFP, as a visual readout of integration efficiency and confirmation of integration – for example, embryos with >50% of GFP^+ heart cells indicates that Tol2 transgenesis has occurred efficiently and so these would be included in an experiment, whereas the presence of only 1 or 2 green cardiac cells would suggest the levels of transgene in the embryo would be negligible and so this would be excluded from the experiment. Independently of this reporter, we showed an upregulation of pik3ca transcript in PROS mosaics compared to control by scRNA-seq (Figure 4D, Supplementary Figure 4A) confirming the transgene produces a measurable upregulation of pik3ca.

We agree that it would be optimal to quantify the transgene expression and copy number for each individual embryo. However, for experiments where phenotypes are scored, hundreds of embryos are injected each time. Therefore, although it would be valuable to quantify the transgene expression and transgene copy number in terms of finding its correlation to phenotype severity, it is not feasible to do this at this scale. In the future, we would like to refine our model to include more sophisticated inducible transgenic models, with stable integration sites to control for integration site/copy number variation. However, for this manuscript, the priority as set out by our charity funders was to generate and characterise a pik3caPROS model that could rapidly test different patient hotspot alleles as well as tissue-specific promoter drivers. Thus, we chose this simpler model for now, but we would be very interested in continuing this work with a more refined model for one or two mutations (See Reviewer comment 1).

This heterogeneity in transgene dosage and expression levels will inevitably have introduced 'noise' into our data. We can account for this somewhat by large numbers of embryos injected per experiment and reproducibility across populations of zebrafish between experiments. We also note that this strategy reflects the heterogeneity in human PROS, with disease mosaicism, presentation, and severity being highly variable from person to person. Therefore, we don't necessarily see this as a drawback for our current approach.

> *(3) The manuscript would benefit from rescue experiments to strengthen causal conclusions. It remains unclear whether the phenotypes induced by PIK3CA PROS variants can be rescued, either through expression of wild-type PIK3CA, pharmacological inhibition of PI3K signaling, or assessment of developmental reversibility. Such experiments would strengthen the link between PI3K activation and the observed phenotypes.*

We agree this is an exciting direction and a great next step for this research to take. This work is currently ongoing, using the specific Pik3ca inhibitor alpelisib, and optimizing treatment conditions to ensure our experimental readouts are meaningful. Through phenotype scoring we do see a significant rescue in the severity of vascular malformations in PROS mosaic embryos. However, we didn't feel this work was ready for the initial submission because (1) the concentrations we must add to the zebrafish medium by immersion are far higher than the doses needed for inhibition of PI3K signalling in human cell lines and (2) we do not see an obvious decrease in pAkt or pS6 levels by western blot analyses of embryos at alpelisib doses of up to 100 μM, for either short or long term exposure. This drug is poorly soluble in water, and so we are also experimenting with introducing it to embryos intravenously.

> *(4) The authors propose candidate signaling molecules mediating non-cell-autonomous effects downstream of PI3K hyperactivation; however, these conclusions remain speculative, as no functional validation is provided. Testing selected candidate mediators identified in the RNA-seq dataset would significantly strengthen the mechanistic conclusions.*

We thank the reviewer for this suggestion, and it is indeed a long-term aim of our work to find better treatments for PROS by combining inhibition of PI3K signalling with other candidate mediators to treat overgrowth. Our scRNA-seq experiments suggest that Notch, Wnt and Ephrin signalling pathway components may contribute to disease, and so a lot of potential for treatment strategies. After we have optimised treatment with alpelisib to rescue our disease phenotype in line with current mammalian models (see response to Comment 3 above), then we will start to look at other candidate mediators alone or in conjunction with alpelisib. However, given the challenges we are facing with the alpelisib treatment, we may need to develop this work in a subsequent study.

> ***Reviewer #2 (Public review):***
>
> *In this manuscript, Brunsdon et al. aim to study PIK3CA-related overgrowth spectrum (PROS) by establishing a mosaic zebrafish model with overexpression of pik3ca carrying hotspot mutations, coupled with an mScarlet+ reporter. Using fluorescence microscopy, the authors demonstrated that overexpression of pik3ca with a number of hotspot mutations led to mesodermal and particularly vascular malformations in the zebrafish model. Interestingly, they found a paucity of mScarlet+ mutant cells in the vascular lesions, consistent with the finding of low PIK3CA mutation burden in PROS tissue. Such data suggest a non-cell-autonomous effect of PIK3CA mutation. Following this logic, the authors performed single-cell RNASequencing on zebrafish overexpressing WT pik3ca and mutant pik3ca at 19 hpf, and demonstrated widespread transcriptomic perturbations across multiple lineages, including lineage frequencies, key cell pathways, and cell-cell interactions. Importantly, they demonstrate that mScarlet+ cells carrying mutant pik3ca cluster separately from other cell types, do not demonstrate clear lineage identity, and have a general downregulation in signaling components.*
>
> *Overall, the conclusions in the manuscript are well-supported by the presented data. The imaging studies are particularly convincing. The transcriptomic analysis generated a list of potential pathways to further investigate and potentially target with future therapeutic interventions. Importantly, this study provides a valuable in vivo model of PROS that: 1)*

> *recapitulates key features of PROS (e.g., multiple mesodermal defects, paucity of mutation burden in lesions suggesting non-cell-autonomous interactions); 2) is scalable; and 3) offers direct visualization of lesion development, compatible with time-course live imaging. This model will be valuable to further understand PROS and potentially study other diseases where the PIK3CA pathway is altered (e.g., certain cancers).*

We thank Reviewer 2 for their careful reading and support of our manuscript, and their helpful suggestions.

> *The following are not necessarily weaknesses of the data, but rather suggestions where the manuscript could be further strengthened:*
>
> *(1) The model recapitulates the variability of mesodermal lesions in PROS. It would be valuable to utilize this model to further study factors that are associated with the development of more severe lesions (e.g., by comparing samples with more severe lesions to those unaffected despite carrying the mutations, Figure 1F).*

This is a very interesting question, and something that we have wondered ourselves. The clinical observation that PROS mutations cause pathology in mesodermal-derived tissues suggests that there is a lineage permissivity of PROS mutations. We plan to perform additional scRNA-seq experiments on later stage embryos (aligned with Figure 1) and hope to incorporate comparison of embryos with more severe lesions to those unaffected despite carrying pik3caPROS mutations.

> *(2) ScRNA-seq analysis could be enriched with a comparison between cells overexpressing mutant pik3ca vs. those overexpressing WT pik3ca.*

The scRNA-seq experiment presented in this paper was limited by funding constraints at the time, and so we focussed on choosing samples that were likely to yield the most meaningful data. Ideally, we would have included a WT overexpression control in addition to an injected no-pik3ca control, however as we did not observe any phenotypes associated with mosaic pik3caWT transgenic embryos (Supplementary Figure 1A, Figure 3D), we chose to not include this condition. We are grateful for subsequent funding that will allow us to perform a scRNAseq experiment at a later timepoint, detailed below, where we plan to include this control.

> *(3) In the scRNA-Seq analysis, it is curious that the C0 cluster, enriched with mScarlet+ cells, is found to have downregulated signaling interactions (Fig. 5C), yet exerts a widespread noncell-autonomous effect. Meanwhile, there is also a noticeable loss of certain lineages (e.g., notochord, Figure 4E) and related cell-cell interactions (e.g., notochord-related interaction, Figure 5A). A deeper exploration of the basis of the non-cell-autonomous effect would be valuable.*

Thank you for this important comment. We agree that this finding is very interesting and warrants further investigation, although a definitive answer may be too difficult for this current revision. Using conventional differential expression analyses on our scRNA-seq data (such as was used in Figure 4), we could not find significant upregulation of many genes and pathways, and CellChat and NICHES analyses did suggest that signalling between C0 and other clusters was weak. Nevertheless, using the Decoupler package, we did find significant upregulation of some footprint signatures enriched in mScarlet^{+} vs - cells in PROS mosaics (Supplementary Figure 4B) including PI3K and EGFR (as one would expect), but also apoptosis and UV response suggesting that overexpression of pik3caPROS may cause cellular stress. Using NICHES, we also found Myc, Notch, Wnt and Ephrin ligand-receptor pairs to be upregulated in PROS mosaic C0 sending and receiving interactions compared to controls, which would be candidates for validating in subsequent studies (Supplementary Figure 4C). We will be interested to determine if C0 like cells are present in older embryos in our scRNA-seq analysis, and if they have similar signalling activity.

> *(4) The scRNA-Seq analysis was performed at one time point (19 hpf). Additional analysis (not necessarily by scRNA-Seq) at other time points to study whether findings at 19 hpf are persistent throughout development or undergo dynamic changes (e.g., cell fate/state of mSc+ mutant cells) would be helpful.*

We agree that the inclusion of a later timepoint in our scRNA-seq experiment would be valuable in answering a lot of our questions about the fate of C0 cells and the persistence of the transcriptional dysregulation, including non-cell autonomous interactions that we see at 19 hpf. As mentioned above, we were constrained by time and funding for the original experiment but are now in a position to add to this work and address this point.

> *(5) The scRNA-Seq analysis provides a valuable list of perturbed interactions that could be targeted by future therapeutic approaches. Validation of the scRNA-Seq findings with proteinlevel analysis, and studying the effect of targeting some of the pathways on the disease phenotype, would offer valuable data for the community.*

Thank you for this comment. We agree that this an essential next step to take and is also a priority for our patient advocates. As mentioned above (Reviewer 1, point 4), we would like to be confident that alpelisib is on-target in our system first, and then we very much want to identify new therapeutic venues to explore in this pre-clinical space.

> ***Reviewer #3 (Public review):***
>
> *Summary:*
>
> *The study "PIK3CA-related overgrowth spectrum (PROS) zebrafish models reveal panlineage developmental dysregulation" presents important findings that extend significantly beyond a single subfield, bridging developmental biology, vascular medicine, and cancerrelated PI3K signalling. By developing mosaic zebrafish models of PROS and combining live imaging with single-cell transcriptomics, the authors provide compelling evidence for a noncell-autonomous mechanism of tissue overgrowth, a conceptual shift with meaningful therapeutic implications.*

We thank Reviewer 3 for their time and thoughtful comments considering our work.

> *Strengths:*
>
> *The evidence is overall convincing, with methodology appropriate and well-validated relative to the current state of the art; the integration of multiple approaches (in vivo modelling, scRNAseq, ligand-receptor inference) strengthens the central claims. However, some aspects of the proposed non-cell-autonomous signalling mechanisms remain*

> *partly correlative, and direct functional validation of the rewired ligand-receptor interactions would further consolidate the conclusions.*
>
> *Weaknesses:*
>
> *The transgenic overexpression approach chosen by the authors represents a well-established and effective strategy for generating mosaic models in zebrafish. However, this approach introduces notable limitations: the lack of control over transgene dosage and unknown integration sites may generate non-physiological effects, potentially confounding the interpretation of key findings.*

Thank you for this important comment. We agree that there are limitations in our current model, and we are working to refine it such that we have temporal as well as spatial control over the expression of pik3caPROS.

Our funding for the start of this study came from the CLOVES Syndrome community charity, and in collaboration with them, we decided that for this work, our priority was to understand more about the disease mechanisms at disease onset, and also to be able to test multiple pik3ca hotspot mutations that affect patients. One question for families is if the pik3ca hotspot mutations contribute differently to patient overgrowths. Our data here suggests that all mutations are able to promote overgrowth equally, and that differences between disease presentation in patients likely reflects the timing and cellular origins of the mutation.

As a side note, together with CLOVES Syndrome community, we also felt that we wanted to test actual patient mutations, rather than artificial hyperactivated variants of Pik3ca such as the widely used p110a* allele (Hu et al. 1995; Venot et al. 2018), which can inform important mechanisms about pathway dysregulation, but less about actual patient-specific disease mutations.

> *The authors are certainly aware that alternative approaches (though technically more demanding) could be considered in future studies to further strengthen the model. For instance, a CRISPR/Cas9-mediated knock-in of the pik3ca-PROS allele at the endogenous locus (retaining upstream native regulatory elements with only a minimal promoter in the construct, co-expressed with a fluorescent reporter via P2A) could allow even more physiological, lineage-restricted expression while enabling direct visualisation of mutant cells. Mesodermal specificity could potentially be further refined by driving mosaic Cas9 expression under a pan-mesodermal tbx promoter, restricting editing to the relevant lineage while simultaneously marking mutant cells fluorescently, thus even more closely mimicking the postzygotic mutational events characteristic of PROS. As a complementary strategy, blastula transplantation experiments using pik3ca-PROS donor cells (ideally co-expressing a distinct fluorescent marker such as mCherry) into fli1:GFP transgenic hosts could provide a powerful and technically consolidated approach to directly visualise and quantify non-cell-autonomous effects on host vasculature, with precise control over mutant cell burden. This combinatorial framework, separating donor mutant cells from host tissue in a two-colour imaging setup, could be particularly compelling for validating the ligand-receptor rewiring predicted by single-cell transcriptomics in future investigations.*
>
> *These reflections are offered in the spirit of prospective methodological development and do not diminish the value of the current work, which opens a valuable new avenue for therapeutic investigation, suggesting that targeting indirect overgrowth-propagating signals, alongside PI3K inhibition, deserves serious consideration.*

Thank you for these excellent suggestions and feedback. We are keen to try to generate fish that more closely align with what is happening in patients. Two challenges we have faced include:

(1) In our hands, the pik3ca promoter itself is not strong enough to drive fluorophore expression to an extent that we can observe fluorescent PROS cells in zebrafish. As a control, after we saw no fluorescence attempting to knock-in fluorophores at the 5' end of endogenous pik3ca, we tried making a transgenic using various lengths of pik3ca promoter regions driving GFP expression. Despite having stable integration of the transgene shown by a secondary transgene reporter inherited through to F1 generation, we could not visualise GFP/mNeonGreen expression at any stage of development.

(2) A drawback of the IRES approach we used here is that the fluorophore expression levels will be lower than using a short cleavable peptide sequence such as P2A. Unfortunately, the critical kinase region (and location of the orthologous hotspot codon 1048) is located only a few amino acids from the stop codon, and we found that the function of Pik3ca was likely impeded by the addition of several extra amino acids after the P2A cleaves itself.

Despite these challenges, we hope to be able to generate models in future with more precise control over mutant cell burden.

References

Hu Q, Klippel A, Muslin AJ, Fantl WJ, Williams LT. 1995. Ras-dependent induction of cellular responses by constitutively active phosphatidylinositol-3 kinase. Science 268: 100102.

Madsen RR, Vanhaesebroeck B, Semple RK. 2018. Cancer-Associated PIK3CA Mutations in Overgrowth Disorders. in Trends in Molecular Medicine, pp. 856-870. Elsevier Ltd.

Venot Q, Blanc T, Rabia SH, Berteloot L, Ladraa S, Duong JP, Blanc E, Johnson SC, Hoguin C, Boccara O et al. 2018. Targeted therapy in patients with PIK3CA-related overgrowth syndrome. Nature 558: 540-546.

https://doi.org/10.7554/eLife.110896.1.sa0