Title
Novel Developmental Hemato-Vascular Endothelial Genes and Identification of Direct Genomic Targets of Ets Related Protein/Ets Variant 2 in Zebrafish

Author
Gomez, Gustavo Adolfo

Publication Date
2013

UNIVERSITY OF CALIFORNIA

Los Angeles

Novel Developmental Hemato-Vascular Endothelial Genes and Identification of Direct Genomic Targets of Ets Related Protein/Ets Variant 2 in Zebrafish

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Molecular, Cell, and Developmental Biology

by

Gustavo Adolfo Gomez Jr

2013

ABSTRACT OF THE DISSERTATION

Novel Developmental Hemato-Vascular Endothelial Genes and Identification of Direct Genomic Targets of Ets Related Protein/Ets Variant 2 in Zebrafish

by

Gustavo Adolfo Gomez Jr

Doctor of Philosophy in Molecular, Cell, and Developmental Biology

University of California, Los Angeles, 2013

Professor Shuo Lin, Chair

Seminal studies in the zebrafish mutant line *cloche*, which is embryonically lethal due to a deficiency in blood and vascular lineages, resulted in the identification of a critical transcription factor Ets Related Protein/ Ets Variant 2 (Etsrp/Etv2) that was necessary and sufficient for vasculogenesis and myelopoiesis. The mammalian homolog was subsequently identified and found to play functionally conserved roles. Because the forced expression of Etv2 alone is sufficient to ectopically induce vascular and myelopoietic genes, we interrogated the transcriptome of Etv2 overexpressing embryos with microarray and deep sequencing approaches and identified several novel vascular and myelopoietic genes downstream of Etv2 in zebrafish. Nevertheless, the number of genes directly induced by Etv2 to specify angioblasts from the lateral plate mesoderm, and the location of such binding would provide valuable information to understand how Etv2 regulates vasculogenesis and hematopoiesis. We therefore approached these questions with ChIP-sequencing and RNA-sequencing technologies to identify the direct genetic targets of Etv2 in zebrafish embryos. These datasets were independently validated, and will provide a framework from which to understand how Etv2 directs the genetic networks that initiate vasculogenesis and myelopoiesis in zebrafish.

The dissertation of Gustavo Adolfo Gomez Jr is approved.

Jau-nian Chen

Alvaro Sagasti

Yibin Wang

Shuo Lin, Committee Chair

University of California, Los Angeles

2013

Dedicated to my grandmother Ofelia Gomez and my parents Gustavo and Gloria Gomez

Table of Contents

Acknowledgments

Chapter Three is a version that is in preperation for publication. The following are co-authors (and contributions): Matthew B. Veldman (Contributed Reagents, Edited Manuscript), Jing Lu (Bioinformatics), Xi Ren (Performed experiments), Matteo Pellegrini (Bioinformatics), Shuo Lin (Project Director)

The work presented in this dissertation is partly funded by a fellowship from the National Institutes of Health (NIH) F31HL091713

I would like to thank my mentor, Shuo Lin, and committee members Jau-nian Chen, Alvaro Sagasti, and Yibin Wang for their guidance and advice. This also includes Kathleen M. Sakamoto who was in my committee but moved her lab to Stanford University half way through my graduate studies. I would also like to acknowledge my previous mentors Wilfred F. Denetclaw Jr and Maria Elena DeBellard, as well as my rotation mentors, Patty Phelps and Guoping Fan, and Ellen Carpenter for advice during journal club. Special thanks to Pamela Hurley for her counsel and help with organizing all matters regarding my graduate studies. Besides the co-authors mentioned above, the work presented in this dissertation has been made possible by extensive contribution from other collaborators who are co-authors on the reprinted articles, thank you. I also thank friends who gave me advice on troubleshooting technical problems which includes Tasuku Kitada, Suhua Feng, Roberto Ferrari and Scott Doherty, who helped with some of the benchwork. I'm also grateful to Haigen Huang, Anqi Liu, Hong Jiang, and Yuan Dong for technical support and assistance with animal husbandry. Finally I would like to thank my family and personal friends who have provided ample moral support.

Biographical Sketch

Education and training:

2005-PRESENT	University of California, Los Angeles Molecular Cell & Developmental Biology Mentor: **Shuo Lin** Research Topic: Molecular mechanisms underlying vascular and hematopoietic development mediated by the transcription factor, Etsrp/Etv2.
2003-2005	San Francisco State University M.S. Molecular Cell Biology Mentor: **Wilfred F. Denetclaw Jr**. Research Topic: Myogenic regulation by Ectoderm-Dermomyotome signaling via lipid rafts
2001-2003	California Institute of Technology Research Technician Mentor: **Maria Elena DeBellard** in **Marianne Bronner Lab**
1995-2000	University of California, Santa Barbara B.S. Biopsychology and B.S. Zoology

Publications:

Veldman MB, Zhao C, **Gomez GA**, Lindgren AG, Yang H, Yao S, Martin BL, Kimelman D, Lin S. Transdifferentiation of fast skeletal muscle into functional endothelium in vivo by the Transcription Factor Etv2. *PLoS Biol* 11,e1001590

Giovannone D, Reyes M, Reyes R, Correa L, Martinez D, Ra H, **Gomez G**, Kaiser J, Ma L, Stein MP, deBellard ME. 2012. Slits affect the timely migration of neural crest cells via robo receptor. *Developmental Dynamics* 241:1274-1288

Liu H*, **Gomez G***, Lin S, Lin S, Lin C. 2012. Optogenetic control of transcription in zebrafish. *PLoS ONE* 7:e50738

Gomez G, Lee J-H, Veldman MB, Lu J, Xiao X, Lin S. 2012. Identification of Vascular and hematopoietic genes downstream of etsrp by deep sequencing in zebrafish. *PLoS ONE* 7:e31658.

Ren X, **Gomez GA**, Zhang B, Lin S. 2010. Scl isoforms act downstream of etsrp to specify angioblasts and definitive hematopoietic stem cells. *Blood* 115:5338-5346.

Gomez GA*, Veldman MB*, Zhao Y, Burgess S, Lin S. 2009. Discovery and characterization of novel vascular and hematopoietic genes downstream of etsrp in zebrafish. *PLoS ONE* 4:e4994.

Lee D, Park C, Lee H, Lugus JJ, Kim SH, Arentson E, Chung YS, **Gomez G**, Kyba M, Lin S, Janknecht R, Lim D-S, Choi K. 2008. ER71 acts downstream of BMP, Notch, and Wnt signaling in blood and vessel progenitor specification. *Cell Stem Cell* 2:497-507.

Sumanas S, **Gomez G**, Zhao Y, Park C, Choi K, Lin S. 2008. Interplay among

Etsrp/ER71, Scl, and Alk8 signaling controls endothelial and myeloid cell formation. *Blood* 111:4500-4510.

* equal contribution

In Preparation:

Gomez GA, Veldman MB, Ren X, Lu J, Pellegrini M, Lin S. The Direct Genomic Targets of Etsrp/Etv2 in Zebrafish Hemato-Vascular Development.

Teaching:

2007	MCDB 138- Developmental Biology, Teaching Assistant
2006	MCDB 138- Developmental Biology, Teaching Assistant

Awards and funding:

2008-2013	NIH F31 NRSA Fellowship
2003-2005	NIH R25 Bridges Fellowship

Talks

July 11, 2013	8th European Zebrafish Meeting
November 18, 2006	UCLA, Molecular Cell and Developmental Biology Retreat

General Introduction

The development of blood and vessels has been historically linked by their close anatomical proximal origin, which was first reported by Florence Sabin in 1917 (Sabin, 1917). Since then there have been many studies dissecting the molecules, cells and developmental mechanisms by which these cells form the tissues crucial for circulation of nutrients and oxygen, which are essential for life. The close derivation of these cells promotes the idea that they are descendants of a common progenitor known as the hemangioblast (Murray, 1932), of which there is evidence in several systems including cultured embryonic stem cells (Choi et al., 1998; Wang et al., 2004) and established model organisms like mouse (Huber et al., 2004) and zebrafish (Vogeli et al., 2006).

Genetic evidence has also hinted at the common derivation of blood and vasculature by knockout mice for several molecules expressed in both of these tissues including the VEGF receptor, *fetal liver kinase 1*, *flk1* (Shalaby, 1997; Shalaby et al., 1995) and the transcription factor *stem cell leukemia*, *scl* (Robb et al., 1995; Visvader et al., 1998), as reviewed in Park et al., 2005. Stronger genetic evidence exists in the zebrafish *cloche* mutant in which all lineages of blood and vascular endothelial cells are deficient from the earliest stages of development, and therefore die embryonically (Stainier et al., 1995). However, the identification of the gene responsible for the *cloche* phenotype remains partially discovered at best. A *cloche* candidate was identified in the *lysocardiolipin acyltransferase*, *lycat,* gene which is an endoplasmic reticulum localized acyltransferase, that when knocked down in zebrafish and in *in vitro* embryonic stem cell differentiation models results in blood and vascular deficiencies (Wang et al.; Xiong et al.). However, *lycat* can not be the long sought after gene responsible for the *cloche* phenotype because although the reintroduction of *lycat* in *cloche* mutants rescues blood

development, it does not rescue the endothelial lineage. Furthermore, while *lycat* is deleted in the *cloche*^{m39} spontaneous mutant allele, it is not affected in other ENU generated alleles.

Several studies examined the transcriptome of *cloche* by microarray approaches (Qian et al., 2005; Sumanas et al., 2005; Weber et al., 2005). Through these studies various genes were discovered as novel to the developing blood and vasculature of zebrafish embryos. A particularly important gene identified in these studies was the DNA binding ETS domain containing transcription factor initially referred to as *ets related protein, etsrp* (Sumanas et al., 2005), that was found to be necessary and sufficient for the specification of vasculogenesis and primitive myelopoiesis (Pham et al., 2007; Sumanas et al., 2008; Sumanas and Lin, 2006). But embryos in which *etsrp* was knocked down retained primitive blood formation, suggesting that *etsrp* is not a marker for all hemangioblasts in zebrafish. However the knockout of the mouse ortholog, *ets variant 2* (*etv2*), results in a similar phenotype to *flk1* nulls, in which blood and vasculature are lacking, ultimately resulting in embryonic lethality (Lee et al., 2008), and in culture both blood and vascular programs have been rescued by the reintroduction of *etv2* in *etv2* null derived stem cells (Kataoka et al., 2011). This suggests that *etv2* is a good marker for the hemangioblast lineage in mammals, but in zebrafish it is only a marker for hemangioblasts in the anterior lateral mesoderm from which endothelial and primitive myeloid cells emerge, and not the posterior lateral mesoderm from which trunk endothelial and primitive erythrocytes are derived (Sumanas et al., 2008).

Since *etsrp/etv2* is sufficient to ectopically induce the expression of vascular and myelopoietic lineages when overexpressed in zebrafish (Sumanas et al., 2008; Sumanas and Lin, 2006), we reasoned that we could identify novel genes expressed during the development of these cell lineages in zebrafish by genomic approaches. Our main aim was to identify the direct

genetic targets, and more specifically obtain the comprehensive set of genomic cis-elements bound by *etsrp/etv2* from whole embryos since this will reflect the binding by *etsrp/etv2* in vivo.

Zebrafish is an ideal organism for these studies because of the high reproductive rate per pair of adults, and because the genome has been sequenced. Also there are standard methods available to reveal and characterize gene expression and function. In the first chapter of this dissertation we performed a microarray experiment in *etsrp/etv2* overexpressing embryos to identify a dozen genes previously unknown to be expressed in the vasculature and two in the primitive myeloid lineage. As time progressed and less biased approaches to interrogate transcriptomes were established in higher throughput next generation RNA-sequencing (RNA-seq), we revisted the *etsrp/etv2* overexpression profiles. This work is presented in chapter two, in which we identified 39 more genes previously unknown to be expressed in the developing zebrafish vasculature. In the third chapter we present data revealing the identity of the direct genomic targets of *etsrp/etv2* and more importantly, the specific locations at which *etsrp/etv2* induces these genes. These data were obtained by combining chromatin immunoprecipitation-high throughput sequencing data with RNA-seq profiles from *etsrp/etv2* gain and loss of function embryos.

Chapter 1

OPEN ACCESS Freely available online

PLoS one

Discovery and Characterization of Novel Vascular and Hematopoietic Genes Downstream of *Etsrp* in Zebrafish

Gustavo A. Gomez[1☯], Matthew B. Veldman[1☯], Yan Zhao[1], Shawn Burgess[2], Shuo Lin[1]*

1 Department of Molecular, Cell and Developmental Biology, University of California Los Angeles, Los Angeles, California, United States of America, **2** Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, United States of America

Abstract

The transcription factor Etsrp is required for vasculogenesis and primitive myelopoiesis in zebrafish. When ectopically expressed, *etsrp* is sufficient to induce the expression of many vascular and myeloid genes in zebrafish. The mammalian homolog of *etsrp*, *ER71/Etv2*, is also essential for vascular and hematopoietic development. To identify genes downstream of *etsrp*, gain-of-function experiments were performed for *etsrp* in zebrafish embryos followed by transcription profile analysis by microarray. Subsequent in vivo expression studies resulted in the identification of fourteen genes with blood and/or vascular expression, six of these being completely novel. Regulation of these genes by *etsrp* was confirmed by ectopic induction in *etsrp* overexpressing embryos and decreased expression in *etsrp* deficient embryos. Additional functional analysis of two newly discovered genes, *hapln1b* and *sh3gl3*, demonstrates their importance in embryonic vascular development. The results described here identify a group of genes downstream of *etsrp* likely to be critical for vascular and/or myeloid development.

Citation: Gomez GA, Veldman MB, Zhao Y, Burgess S, Lin S (2009) Discovery and Characterization of Novel Vascular and Hematopoietic Genes Downstream of *Etsrp* in Zebrafish. PLoS ONE 4(3): e4994. doi:10.1371/journal.pone.0004994

Editor: Annarosa Leri, Harvard Medical School, United States of America

Received September 26, 2008; **Accepted** February 24, 2009; **Published** March 24, 2009

Funding: HIN. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: shuolin@ucla.edu

☯ These authors contributed equally to this work.

Introduction

The cardiovascular system of vertebrates provides a means to transport nutrients to and waste away from cells throughout the organism. As such, this system is critical for the survival of the organism. The cardiovascular system is made up of the vasculature and the blood that flows through it. It has been proposed that a common cell type, known as hemangioblasts, contribute to the development of vascular and blood cells [1,2,3,4,5]. This idea is supported by experimental evidence from several different organisms, notably the derivation of both blood and vascular endothelial cells from cultured mouse embryonic stem cells [6,7,8], and through in vivo lineage tracing studies in zebrafish and chick/quail chimeras [9,10]. The zebrafish mutant line *cloche* provides genetic evidence since homozygous mutants lack both hematopoietic and vascular endothelial cells but organogenesis is otherwise normal [11]. As bi-potential precursor cells, hemangioblasts produce cells with more restricted potentials, angioblasts and hematopoietic progenitors [3]. Angioblasts subsequently produce the endothelial cells that line the vasculature during both vasculogenesis and angiogenesis, while hematopoietic progenitors differentiate into distinct blood lineages through hematopoiesis. Although some of the molecular factors that orchestrate these processes are known, many remain to be identified and studied.

Microarray studies were previously used to identify multiple novel hematopoietic and vascular genes misexpressed in the *cloche* mutant [12,13,14]. A particularly important gene identified in these studies was the transcription factor, Ets1-related protein (Etsrp), which is both necessary and sufficient to direct the development of the vascular and primitive myeloid lineages [13,15,16]. In zebrafish, hematopoietic/vascular development is separated into two distinct anatomical locations, the anterior lateral mesoderm (ALM) and posterior lateral mesoderm (PLM). The ALM gives rise to both primitive myeloid cells as well as the endothelial cells of the head vasculature. The PLM gives rise to primitive erythroid cells and the endothelial cells of the trunk and tail. *Etsrp* is expressed in both the ALM and PLM. Loss of *etsrp* function by morpholino antisense or genetic mutation, termed *y11*, results in loss of primitive myeloid cells and disrupted vasculogenesis and angiogenesis in the head and trunk. The defective vasculature in *etsrp* knockdown or mutant fish appears to be due to altered gene expression and cell behavior and not simply a loss of cells as the number of *fli:gfp* transgene positive cells is similar to control animals [15]. In contrast, the complete loss of *pu.1* staining in *etsrp* knockdown and mutants suggests that primitive myeloid cells are never specified [16,17]. Interestingly, the primitive erythroid population in the PLM appears relatively normal when *etsrp* function is blocked. Thus, *etsrp* is critical for primitive myeloid and endothelial development from the ALM and endothelial development from the PLM but not the primitive erythroid cell population of the PLM.

A mammalian homolog of *etsrp*, *ER71/Etv2*, has recently been identified. *ER71/Etv2* knockout mice exhibit loss of vasculature and primitive erythrocytes suggesting it functions in the developing hemangioblast [18]. Additionally, overexpression of *ER71/Etv2* in zebrafish embryos causes the ectopic induction of the hemangioblast marker *scl/tal1* and the *flk1:gfp* transgene, identical to the results of *etsrp* overexpression. Therefore, both *etsrp* and *ER71/Etv2* play an evolutionarily conserved role in hematovascular development.

As members of the Ets transcription factor gene family, Etsrp and ER71/Etv2 have a conserved DNA binding domain and presumably act as transcriptional activators. Limited *in situ* hybridization analysis of *etsrp* morpholino gene knockdown or *y11* mutant embryos demonstrate that *etsrp* is necessary for the expression of *flk1*, *scl*, *fli1*, *pu.1* and a few other known genes specific to vascular and hematopoietic cells [15,19]. Additionally, overexpression of *etsrp* in zebrafish embryos can ectopically induce vascular and myeloid gene expression [16]. We therefore decided to search for novel blood and vascular related genes downstream of *etsrp* by analyzing expression profiles of zebrafish embryos overexpressing *etsrp*. Here we report the identification of 14 genes with little or no prior investigations that are expressed in blood and/or vascular endothelial cells and demonstrate the function of two of these genes, *hapln1b* and *sh3gl3*, in vascular development using morpholino antisense in transgenic zebrafish embryos.

Results

In order to identify novel blood or vascular related genes downstream of *etsrp*, the expression profiles of embryos ectopically expressing *etsrp* at late gastrulation stages, 80% epiboly to tail bud were compared to control embryos by microarray analysis. The *flk1:gfp* transgenic line was used to identify embryos successfully expressing *etsrp* since ectopic *etsrp* induces *flk:gfp* expression. Neither endogenous or transgenic *flk1:gfp* is normally expressed at 80% to tail-bud stage [20,21], while early ectopic expression of *flk1:gfp* is evident in embryos injected with 75 pg of synthetic *etsrp* mRNA (Figure 1A). To ensure that the ectopic expression of *flk1:gfp* is induced by *etsrp* specifically, we injected up to 500 pg of synthetically transcribed mRNA encoding RFP tagged histone H2B, *H2B-RFP1*, as a control. Similar to uninjected controls, *flk1:gfp* expression is not induced in *H2B-RFP1* injected embryos (Figure 1C and 1D).

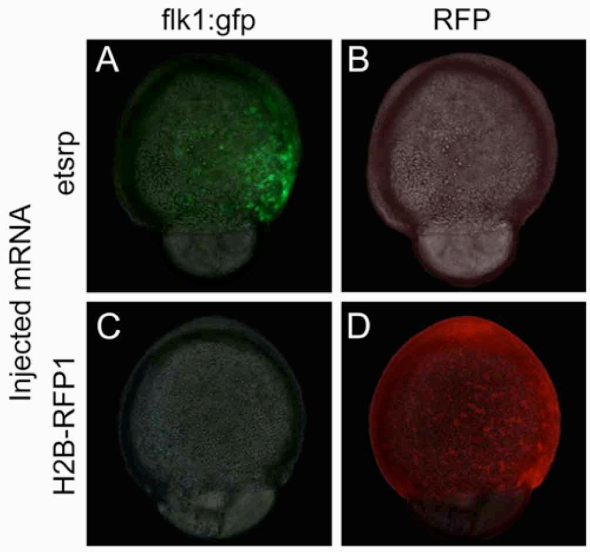


Figure 1. Early ectopic expression of transgenic *flk1:gfp* is induced by *etsrp* over-expression. (A) Injection of 75 pg of *etsrp* mRNA at one-cell stage results in the ectopic induction of *flk1:gfp* before the end of gastrulation. (B) No red fluorescence is observed in *etsrp* mRNA injected *flk1:gfp* transgenic embryos. (C) Injection of 500 pg of control *H2B-RFP1* mRNA does not result in the early induction of *flk1:gfp* similar to uninjected embryos (not shown). (D) Uniform H2B-RFP1 expression in a *H2B-RFP1* mRNA injected embryo. Panels A and C are composite images of light transmitted images merged with the green fluorescent channel, while panels B and D are composites of light transmitted images merged with the red fluorescent channel.
doi:10.1371/journal.pone.0004994.g001

Prior to microarray analysis total RNA was extracted from pools of 70 embryos per treatment group and the change in expression of genes known to be induced by *etsrp* overexpression (OE) were assessed by quantitative RT-PCR. Both *fli1a* (11×) and *scl* (7.1×) were significantly induced by *etsrp* OE, while overexpression of *H2B-RFP1* resulted in no difference from uninjected controls (Figure S1). Gene expression analysis was then performed using a microarray chip containing 34,647 oligos representing approximately 20,000 genes. 686 genes were identified as being induced greater than two-fold by ectopic *etsrp* expression. Although some of the genes affected are not restricted to the circulatory system (Figure S2), a great number of previously identified hematopoietic and vascular genes were induced, validating the microarray strategy used in this study. Table 1 lists the 59 genes with the greatest fold induction as measured on the microarray.

From the genes affected by *etsrp* OE, we set up four criteria to select candidates for further studies: 1. They have not been well studied in hematopoiesis and/or vasculogenesis; 2. They are induced by ectopic *etsrp* in embryos as validated by whole mount *in situ* hybridization (WISH); 3. They are expressed in hematopoietic and/or vascular tissues; and 4. Expression is reduced by *etsrp* deficiency.

To identify genes with few or no prior studies for further analysis, the list of genes induced by *etsrp* OE was compared to the data sets for *cloche* mutants [13,14]. We reasoned that genes identified in both *etsrp* OE and *cloche* mutant microarrays were highly likely to be involved in blood and/or vascular development. From this group of overlapping genes we screened for genes that had not been well studied using the PubMed database at the NCBI website. This resulted in the selection of 31 genes (Table 2), which were cloned and analyzed further.

To validate the results obtained with the *etsrp* OE microarray, WISH was performed on embryos at the same developmental stage as was done for the microarray, 80% epiboly to tailbud. To ensure that the OE samples examined were indeed ectopically expressing *etsrp*, embryos were injected with plasmid DNA encoding an Etsrp-mcherry fusion protein. This fusion protein was functionally active since it efficiently induced ectopic expression of *flk:gfp* when injected into transgenic embryos (data not shown). Embryos expressing Etsrp-mcherry and *flk:gfp* were selected to test for ectopic induction for 23 of the 31 genes identified in Table 2. 21 of 23 genes were clearly induced ectopically by *etsrp* OE. Figure 2 shows the ectopic induction of the 14 genes identified below to be relevant to blood and/or vascular expression. There was ubiquitous endogenous expression in 2/23 genes, *arhgef9* and *znf385*, prohibiting any qualitative discernment of their changes in expression by *etsrp* OE (data not shown). Overall, this demonstrates that the microarray data is low in false positives and reliable.

To screen the 31 selected genes for blood and/or vascular expression, their expression patterns were examined in wild type embryos by WISH at 24–30 hours post fertilization (hpf). Of the genes examined, two are expressed in blood lineages (Figure 3), and twelve in vasculature (Figure 4), while sixteen other genes demonstrate tissue specificity but are irrelevant to blood or vessels (Figure S2).

The dependence of the new blood and vascular specific genes on *etsrp* was then examined by comparing their expression in *etsrp* morphants to uninjected control embryos (Figure 5). In *etsrp* morphants, the expression of *krml2* (Figure 5A) and *lrrp33*

Table 1. Top 59 genes induced by overexpression of *etsrp*.

GB accession	Unigene ID	Location	Gene name	Fold induction
BI980779	Dr.6315	chr14	Cdc42 guanine nucleotide exchange factor (GEF) 9	55.92
AI965222	Dr.45597	chr14	Cadherin 5 (cdh5)	54.96
DQ021472.1	Dr.47591	chr16	Ets1b/etsrp Ψ	47.01
AW233044	Dr.24158		Serglycin	39.62
NM_131626	Dr.81423		HAND2	34.52
AF045432	Dr.75812	chr22	Tal1/SCL	34.35
NM_152964	Dr.17548	chr6	Gastrulation brain homeo box 2 (Gbx2)	27.73
NM_131768	Dr.3499	chr3	Claudin l	26.36
AW420632	Dr.12726	chr21	Yes-relayed kinase	25.02
NM_131530	Dr.5729	chr11	Homeo box C6b	24.88
AI601685	Dr.5365	chr22	Dual specificity phosphatase 5 (dusp5)	24.52
BI673802	Dr.113768	chr11	Fgd5	23.26
AI477500	Dr.4847	chr11	Wnt7A	20.14
BI880801	Dr.83871	chr17	Rasgrp3	20.10
AW115759	Dr.23967	chr17	Complement receptor C1qR-like	19.35
AI477956	Dr.76040	chr8	Nexilin f actin binding	19.32
AW595297	Dr. 105109	chr8	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 3	18.21
Y14538	Dr.5730	chr19	Homeo box A9a	17.82
BI842764	Dr.14186	chr17	Si:ch211-214p16.1	17.50
AI793554	Dr.104443	chr16	Cellular retinoic acid binding protein 2a	15.98
BI882214	Dr.13039	chr15	Solute carrier family 43, member 2	14.62
BM181739	Dr.8453	chr2	PLA2-like otoconin	14.23
AI959344	Dr.28288	chr7	Matrix metalloproteinase precursor	13.34
BM036580	Dr.84866	chr16	Similar to hemicentin	12.93
AB055670	Dr.30464	chr23	Wnt1	12.77
Y14532	Dr.75789	chr12	Homeo box B6b	12.71
AI626374	Dr.2644		Dehydrogenase/reductase (SDR family) member 3	12.60
AI477237	Dr.77024	chr13	Similar to Olig3 protein	12.31
AW279740	Dr.22713	chr23	Zinc finger protein 385	12.18
BI982208	Dr.11052		Ornithine decarboxylase antizyme 2, like	11.42
BM103177	Dr.24950		Creatine kinase CKM3	11.27
AW076731	Dr.23392	chr25	integrin alpha, VLA protein type	10.63
AW233713	Dr.81043		Dusp2	10.56
BG892475	Dr.11064	chr2	amyloid precursor-like protein 2 (aplp2)	10.31
AA605884	Dr.38472		cDNA clone IMAGE:7137566	10.22
Y14527	Dr.5721	chr17	Homeo box A10b	10.15
NM_131359	Dr.75781		Bone morphogenetic protein 2a	9.90
NM_131419	Dr.272		NK2 transcription factor related 7	9.81
AI626636	Dr.45596		Adrenomedullin receptor (admr)	9.76
NM_131070	Dr.20373		Midkine-related growth factor	9.68
AW233616	Dr.121119		Calpain 8	9.52
AI522349	Dr.44232		Rhoub	9.46
NM_131000	Dr.20912		Activated leukocyte cell adhesion molecule	9.29
BG985561	Dr.8617	chr2	Hyaluronan and proteoglycan link protein 1b	9.21
NM_131767	Dr.26572		Claudin h (cldnh)	9.09
NM_131101	Dr.5572		Homeo box B5a	8.81
AF071249	Dr.28647	chr17	Homeo box A9b	8.47
BG727645	Dr.11977	chr16	Leucine rich repeat 33	8.00
BG985673	Dr.24989	chr10	Myosin light chain 2 like	7.80
Y13948	Dr.21032	chr9	Homeo box D3a	7.71

Table 1. cont.

GB accession	Unigene ID	Location	Gene name	Fold induction
AI601355	Dr.77902		Similar to Fraser syndrome 1 isoform 1	7.61
NM_131419_1	Dr.272		NK2 transcription factor related 7	7.60
BM025541	Dr.3373	chr9	Collagen type XVIII, alpha 1	7.56
AF071568	Dr.20962	chr19	Homeo box B2a	7.43
AY028584	Dr.18294		Prostaglandin-endoperoxide synthase 1	7.38
AI601770	Dr.21376	chr23	Plexin precursor	7.27
AI721522	Dr.8674		Septin 9b	6.78
BG728550	Dr.81575	chr15	Cytochrome b5 reductase 4	6.58
AI964237	Dr.76050		ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit 1	5.68

Note: Genebank accession number and unigene ID correspond to the probes on the array. The fold change column is the average ratio of eight measurements. Ψ The large induction ratio of *etsrp* reflects array probe hybridization to both the exogenously injected RNA as well as the induced expression of endogenous *etsrp*, (See Supplementary Figure 1).

doi:10.1371/journal.pone.0004994.t001

(Figure 5B) was abolished in myeloid cells, while the remaining 12 vascular genes were also significantly reduced in expression, most noticeably in the trunk region containing the dorsal aorta, posterior cardinal vein, and intersomitic vessels (Figure 5, C–N). Note that gene expression in the cranial vasculature was less affected in *etsrp* morphants but still reduced. Gene expression in tissues outside of blood and vessels such as somites for *krml2* (Figure 5A), neurons for *rasgrp3* (Figure 5D) and *ldb2* (Figure 5L), and tailbud for *tem8* (Figure 5E) were not affected by *etsrp* knockdown. This demonstrates that *etsrp* is not only sufficient but also required for normal expression in primitive myeloid and/or vascular cells in the 14 genes identified here.

Genes downstream of *etsrp* with endogenous expression in blood

Krml2. Krml2 is a member of the Maf family of basic leucine zipper transcription factors that have high evolutionary conservation with orthologs existing in chicken, *Xenopus*, mice, and humans [22,23]. The expression of *krml2* in zebrafish was previously found in somites, reticulospinal oculomotor neurons, and lens [24]. In this study we have also detected its expression in primitive myeloid cells and the vasculature (Figure 3A). The mouse and avian orthologs of *krml2* have also been detected in macrophages [25], and it has been demonstrated to regulate hematopoiesis by repressing erythropoiesis in favor of the myeloid fate through an interaction with Ets1 [26].

LRRP33. Leucine rich repeat containing protein 33 is a member of the leucine-rich repeat protein family. Lrrp33 is a single pass transmembrane protein that contains 21 leucine-rich repeats, and has orthologs in mice, rats, and humans, which remain to be characterized. The leucine-rich repeat family of proteins consists of over 60,000 members with diverse biological activities, but a common feature shared by proteins with LRR motifs is their propensity to associate with other proteins through LRRs [27]. Prominent members of this family with characterized roles in inflammatory and innate immunity are the toll-like receptors, which are comprised of ten members in humans, with homologs for all encoded in the zebrafish genome [28,29]. *Lrrp33* is excluded from this subfamily because it lacks the cytosolic toll-interleukin (TIR) receptor domain through which TLRs transduce intracellular signals [30,31]. In this study we have detected *lrrp33* expression in primitive myeloid cells and the caudal hematopoietic tail region (Figure 3B).

Genes downstream of *etsrp* with endogenous expression in vasculature

RAS guanyl releasing protein 3 (RASgrp3). RASgrp3 is a member of the RAS family of GTPases which link cell surface receptor activation and RAS activation by switching its state of activity between inactive GDP and active GTP bound states [32]. With orthologs in mice and humans, zebrafish *RASgrp3* encodes a protein with 708 residues containing an N-terminal and catalytic RhoGEF domains, followed by a pair of calcium binding EF-hand domains and a zinc dependent DAG/phorbol-ester binding domain. *RASgrp3* is expressed at low levels in both axial and cranial vasculature, and the caudal vascular plexus region, as well as the neural tube and hindbrain (Figure 4A). A role for *RASgrp3* in endothelial cells was initially identified in an embryonic stem cell gene trap assay [33]. In endothelial cells *RASgrp3* is up-regulated by VEGF signaling, and functions as a phorbol ester receptor in both B-cells and endothelial cells, where it has been ascribed with forming a disorganized angiogenic vasculature in disease states [33,34]. There are currently four *RASgrp* genes encoded in mammals, *RASgrp-1*, *-2*, *-3*, and *-4*, while zebrafish only have three, *RASgrp-1*, *-2*, and *-3*. The mouse ortholog is expressed in the vasculature during developmental stages, but the knockout is viable [33].

Similar to tumor endothelial marker 8 (tem8)/Anthrax Receptor 1 (antxr1). *Tem8* was initially identified as a gene up-regulated in tumors, and is one of two anthrax toxin receptors, the other being *capillary morphogenesis gene 2 (CMG2)/ANTXR2* which together recognize and internalize anthrax toxins [35,36,37]. Zebrafish *tem8* encodes a 552 amino acid transmembrane protein with a von Willebrand factor type A domain, an extracellular Ig anthrax toxin receptor domain and the cytoplasmic c-terminal anthrax receptor domain. It is expressed at low levels in neurons with higher expression levels in the tailbud as well as the axial and cranial vascular endothelial cells (Figure 4B). Two *cmg2* orthologs have been identified in zebrafish, *cmg2a* and *cmg2b*, but their expression remains to be examined [38]. *Tem8* regulates cell adhesion and migration by coupling extracellular stimuli to changes in the actin cytoskeleton [39,40]. Regulation of cell

Table 2. Genes screened for blood and/or vascular expression.

GB Accession	Unigene ID	Gene location	Gene name	Fold Induction	WISH for *etsrp* OE	down in *cloche*
BI980779	Dr.6315	chr14	Cdc42 guanine nucleotide exchange factor GEF 9 (arhgef9)	55.92	X	X
AW171479	Dr.2433	chr12	Carboxylesterase 2-like(cbe2)	3.32	X	X
AF109780	**Dr.81287**		**Kreisler maf-related leucine zipper homolog 2 (krml2)**	**4.82**	**X**	**X**
AF071255	Dr.117289	chr12	Hoxb8b	3.57	X	X
BI880801	**Dr.83871**	**chr17**	**RAS guanyl releasing protein 3 (rasrgrp3)**	**20.51**	**X**	
BG727645	**Dr.11977**	**chr16**	**Leucine rich repeat containing 33 (lrrp33)**	**8.00**	**X**	**X**
AW279740	Dr.22713	chr23	Zinc finger protein 385 (znf385)	12.18	X	
BI673729	Dr.14541	chr6	Similar to leucine rich repeat containing 51 (lrrp51)*	6.08	X	
AI544475	Dr.33183	chr13	SPARC related modular calcium binding 2 (smoc2)	6.68	X	
BI867572	Dr.9564	chr21	Similar to testis-expressed sequence 2(tex2)	4.91	X	
AW421939	Dr.5562	chr20	Apolipoprotein B (apolipo b)	3.07	X	
BI430050	Dr.418		Protocadherin2 gamma20 (pcdhg)	3.50		
BI672391	Dr.14449		weakly similar to XP_422395.1 (novel 2)	5.61	X	
BC093143	Dr.20125	chr1	hypothetical LOC550501	6.80		
BI864031	**Dr.13409**	**chr12**	**Anthrax toxin receptor 1/similar to tumor endothelial marker 8 (antxr1/tem8)**	**7.29**	**X**	
BM095233	Dr.17582	chr23	C1orf192 homolog	4.47		
BI671403	**Dr.120608**	**chr24**	**PTPL1-associated RhoGAP 1 (arhgap29)**	**5.65**	**X**	
BI673802	**Dr.113768**	**chr11**	**Fyve rhogef and ph domain containing zinc finger fyve domain containing protein 23 (Fgd5)**	**23.26**	**X**	**X**
AY028584	Dr.18294		Prostaglandin-endoperoxide synthase 1 (ptgs1)	7.38		
BG985561	**Dr.8617**	**chr2**	**Hyaluronan and proteoglycan link protein 1 precursor (hapln1B)**	**9.21**	**X**	
BI842967	Dr.76442	chr8	hypothetical LOC555375	3.68		
BI880784	Dr.12941		Similar to testican 3	3.81		
AW420632	**Dr.12726**	**chr21**	**Yes-relayed kinase (yrk)**	**25.02**	**X**	**X**
BM036580	**Dr.84866**	**chr16**	**Similar to hemicentin**	**12.93**	**X**	
BI983776	**Dr.12218**	**chr25**	**SH3-domain GRB2-like 3 (sh3gl3) ***	**2.76**	**X**	
AI884043		**chr24**	**Similar to immune costimulatory protein ***	**3.34**	**X**	
X60095	Dr.510	chr23	Hoxc3a	3.37		
BI428793			**Lim domain binding 2 (ldb2)**	**4.21**	**X**	
AI721944	**Dr.21605**	**chr3**	**Est- AI721944**	**2.92**	**X**	
AW019729		**chr8**	**Est- AW019729**	**2.68**	**X**	
AA542593	Dr.75525	chr23	C20.orf.112	5.09		

Note: Genebank accession number and unigene ID correspond to the probes on the array. The fold change column is the average ratio of eight measurements. (*) ectopic induction was observed for lrrp51 but no expression was detected between 24–30 hpf. Genes in bold indicate those that are expressed in blood or vascular endothelial cells at 24–30 hpf.

doi:10.1371/journal.pone.0004994.t002

adhesion and migration by *tem8* plays a role in angiogenesis and is relevant to tumor angiogenesis, anthrax toxin responses, and potentially circulatory system development.

RhoGTPase-activating protein 29 (arhgap29). Arhgap29 was identified in a screen for proteins that interact with the protein-tyrosine phosphatase PTPL1 (which is expressed in many tissues), and has the alternative name PTPL1-Associated RHOGAP1 (PARG1) [41]. Zebrafish arhgap29 is a guanine activating protein of RhoGTPases that is 1365 amino acids long with a protein kinase C domain and a RhoGAP domain which are conserved in orthologs in multiple species including humans. The expression of *ARHGAP29* has been detected in the hearts of mouse embryos among other tissues [42], and in zebrafish it is expressed in the cranial vasculature, dorsal aorta, cardinal vein, intersomitic vessels, caudal vascular plexus, and hatching gland at 24 hpf (Figure 4C).

FYVE, rhoGEF and PH domain containing protein 5 (fgd5). Fgd5 is a member of the Dbl homology (DH) protein family that functions as guanine nucleotide exchange factors for the RhoGTPases Rho, Rac and Cdc42 [43]. There are six paralogs of this gene in mammals. The family's name is derived from the founding member, *fgd1*, which causes faciogenital dysplasia in humans when mutated [44]. Fgd5 has been widely conserved through evolution and, in zebrafish, Fgd5 is 1603 amino acids long with the common domains of all DH family members, which includes a DH domain followed by two plextrin homology PH domains that flank a FYVE containing zinc finger domain. *Fgd5* is expressed in the endocardium of mice during development

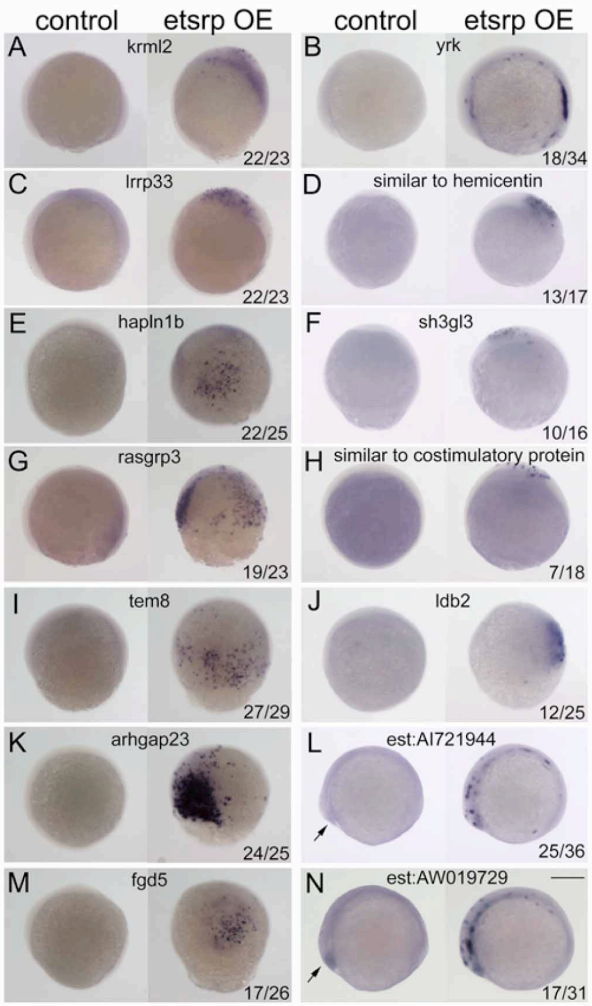


Figure 2. Ectopic induction of microarray identified genes by *etsrp* overexpression. Ectopic gene expression was examined at 80% epiboly to tailbud stages in *flk1:gfp* transgenic embryos injected with 30 pg *etsrp*-mcherry DNA at the one cell stage. Embryos exhibiting both red and green fluorescence were selected for analysis. Representative control embryos are on the left and *etsrp* overexpressing (OE) embryos are on the right side of each panel. Represented genes: (A) *krml2*; (B) *yrk*; (C) *lrrp33*; (D) *similar to hemicentin*; (E) *hapln1b*; (F) *sh3gl3*; (G) *rasgrp3*; (H) *similar to costimulatory protein*; (I) *tem8*; (J) *ldb2*; (K) *arhgap23*; (L) *est:AI721944*; (M) *fgd5*; and (N) *est:AW019729*. Note that there is a low level of endogenous expression in the control embryos for the two EST's, L and N, at the polster (arrows). Ratios in bottom right hand corner in panels represent the number of embryos with ectopic induction of total embryos processed and scored in the injected groups; control embryos never displayed ectopic induction. All embryos are in lateral view, and those at tail bud stage are oriented with anterior to the left. Scale bar: 250 μm.
doi:10.1371/journal.pone.0004994.g002

[45], while in zebrafish it is expressed throughout the vascular system, including the heart (Figure 4D).

Yes-related kinase (yrk). *Yes-related kinase (yrk)* is a homolog of the *yes* oncogene, and is a member of the *src* non-receptor tyrosine kinase family. It is distinct from other *src* members including *fyn* and *yes*, and is expressed in the cerebellum, spleen, lung, and skin in adult chicken tissues [46]. The protein is composed of 528 amino acids with single SH2, SH3 and tyrosine kinase domains. Unlike its ortholog in chicken, the zebrafish *yrk* is expressed exclusively in vascular endothelial cells where its function remains to be investigated (Figure 4E).

Similar to hemicentin. Hemicentin is an extracellular matrix ECM protein of the immunoglobulin superfamily identified originally in *c.elegans* and found to facilitate tissue organization and cell migration [47]. The human ortholog has been associated with age related macular degeneration [48] and in mice its expression has been observed at the pericellular ECM of various epithelial cells including embryonic trophectoderm, skin, tongue, in addition to the ECM of some blood vessels [49]. Distinct from hemicentin, zebrafish similar to hemicentin differs in size, structure and possibly function. While hemicentin contains a von willberand A domain, 48 tandem Ig domains, multiple tandem epidermal growth factor repeats (EGFs), and a single fibulin-like C-terminal domain, zebrafish similar to hemicentin is 264 amino acids long, has a single Ig domain and a transmembrane domain. Its specific expression in the vascular system suggests a potential role for this gene in endothelial cells (Figure 4F).

SH3 domain Growth factor Receptor Bound protein 2-like 3 (sh3gl3). *Sh3gl3* encodes for a protein that is 386 amino acids long which contains a Bin-Amphiphysin-Rvs (BAR) domain implicated in synaptic vesicle internalization, actin regulation, differentiation, and cell survival, and a src homology 3 (SH3) domain [50]. The mammalian SH3gl3, also referred to as endophilin A3 and EEN-B2, is a member of a small family of SH3 containing proteins known as the SH3GL family that includes three members, EEN, EEN-B1 and EEN-B2 [51]. In adult mice, *EEN-B2* is expressed in the brain, testis and spinal cord, while its embryonic expression is ubiquitous and there appears to be a marked transient expression in the vascular system [52]. In zebrafish, *sh3gl3* expression is vascular specific at the stage examined here (Figure 4G). Besides *sh3gl3*, paralogs in zebrafish include *sh3gl1a*, *sh3gl1b*, *sh3gl2*, *sh3gl2b, and zgc:158742*.

Similar to immune costimulatory protein. *Similar to immune costimulatory protein* is a novel gene that encodes a 372 amino acid long transmembrane protein with two tandem Ig-like domains at the N-terminus followed by a transmembrane domain. It is expressed in the cranial vasculature, the caudal vascular plexus, and is restricted to the artery in the axial vessels. It is also expressed in a spotted pattern throughout the ectoderm (Figure 4H).

LIM domain binding protein 2 (ldb2). LIM domain binding protein 2 (ldb2) is a LIM homeodomain co-activator that is one of four *ldb* genes found in zebrafish. *Ldb2* has been reported to be induced in human vascular endothelial cells following stimulation with VEGF [53]. The expression of *ldb2* in zebrafish has been described in the central nervous system and vasculature [54], which we confirmed in Figure 4I. We have also found that *ldb2* is silenced in the axial vasculature of *etsrp* morphants (Figure 5L).

Novel Expressed Sequence Tags (ESTs). The overexpression of *etsrp* resulted in the identification of two ESTs that could not be linked to any annotated genes but which are clearly induced ectopically (Figure 2L and 2N), demonstrate vascular specific expression, including cranial, axial, caudal and intersomitic vessels (Figure 4J and 4K), and whose expression is absent in the axial and intersomitic vessels in *etsrp* morphants (Figure 5M and 5N). Their accession numbers are: AI721944 and AW019729.

Hapln1b. Hyaluronan and proteoglycan link protein 1b is a member of the vertebrate hyaluronan and proteoglycan link protein family which consists of four members in mice and

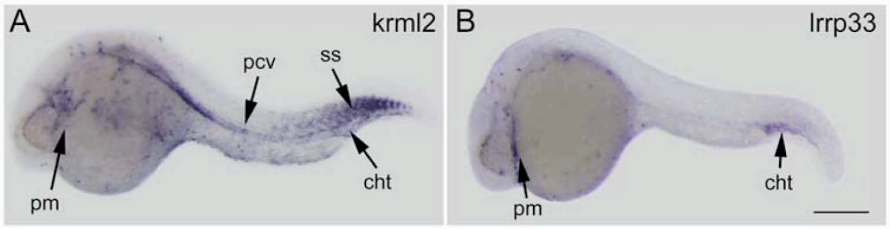


Figure 3. Expression of genes identified in primitive myeloid blood cells. Gene expression was examined at 24–28 hours post fertilization by whole mount in situ hybridization in wildtype embryos. (A) *krml2* is expressed in primitive myeloid cells (dispersed cells labeled throughout yolk and head) in the posterior cardinal vein, caudal hematopoietic tail region, and somites. (B) *Lrrp33* is expressed in primitive myeloid cells and in the caudal hematopoietic tissue Abbreviations: pm, primitive myeloid; pcv, posterior cardinal vein; cht, caudal hematopoietic tail region; ss, somites. Scale bar: 250 μm.
doi:10.1371/journal.pone.0004994.g003

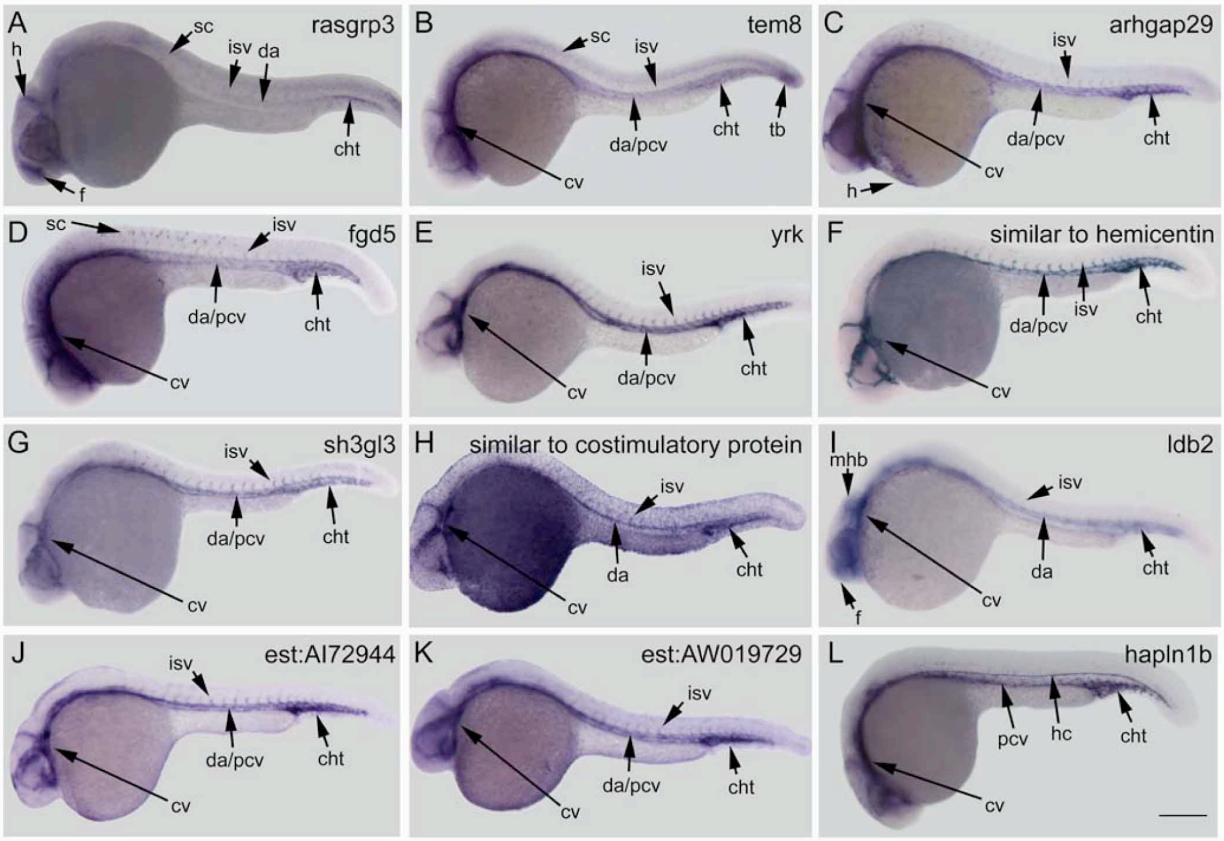


Figure 4. Expression of genes identified in vascular endothelial cells. Gene expression was examined at 24–28 hours post fertilization by whole mount *in situ* hybridization in wildtype embryos. (A) *rasgrp3* is expressed weakly in the forebrain, hindbrain, spinal cord, intersomitic vessels, dorsal aorta, and caudal hematopoietic tail region. (B) *tem8* is expressed highly in the cranial vasculature, the dorsal aorta and posterior cardinal vein, tailbud and weakly in the intersomitic vessels and spinal cord. (C) *arhgap29* is expressed in the cranial vessels, dorsal aorta, posterior cardinal vein, intersomitic vessels, caudal hematopoetic tail region, and hatching gland. (D) *fgd5* is expressed in the cranial vasculature, dorsal aorta, posterior cardinal vein, intersomitic vessels, and caudal hematopoetic tail region. (E, F, and G) *yrk*, *similar to hemicentin*, and *sh3gl3* are all expressed in the cranial vasculature, dorsal aorta, posterior cardinal vein, intersegmental vessels, and caudal vascular hematopoietic tail region. (H) *similar to costimulatory protein* is expressed in the cranial vasculature, dorsal aorta, intersomitic vessels, and caudal hematopoietic tail region; additional expression is present in the ectoderm layer throughout the embryo. (I) *ldb2* is expressed in the forebrain, midbrain-hindbrain boundary, dorsal aorta, intersomitic vessels, and caudal hematopoietic tail region. (J and K) Both *est:AI721944* and *est:AW019729* are expressed in the cranial vasculature, dorsal aorta, posterior cardinal vein, intersomitic vessels, and caudal hematopoietic tail region. (L) *hapln1b* is expressed in the cranial vasculature, posterior cardinal vein, caudal hematopoietic tail region, and hypochord. Abbreviations: f, forebrain; h, hindbrain; sc, spinal cord; cv, cranial vasculature; da, dorsal aorta; pcv, posterior cardinal vein; isv, intersomitic vessels; cht, caudal hematopoietic tail region; tb, tail bud; h, hatching gland; and hc, hypochord. Scale bar: 250 μm.
doi:10.1371/journal.pone.0004994.g004

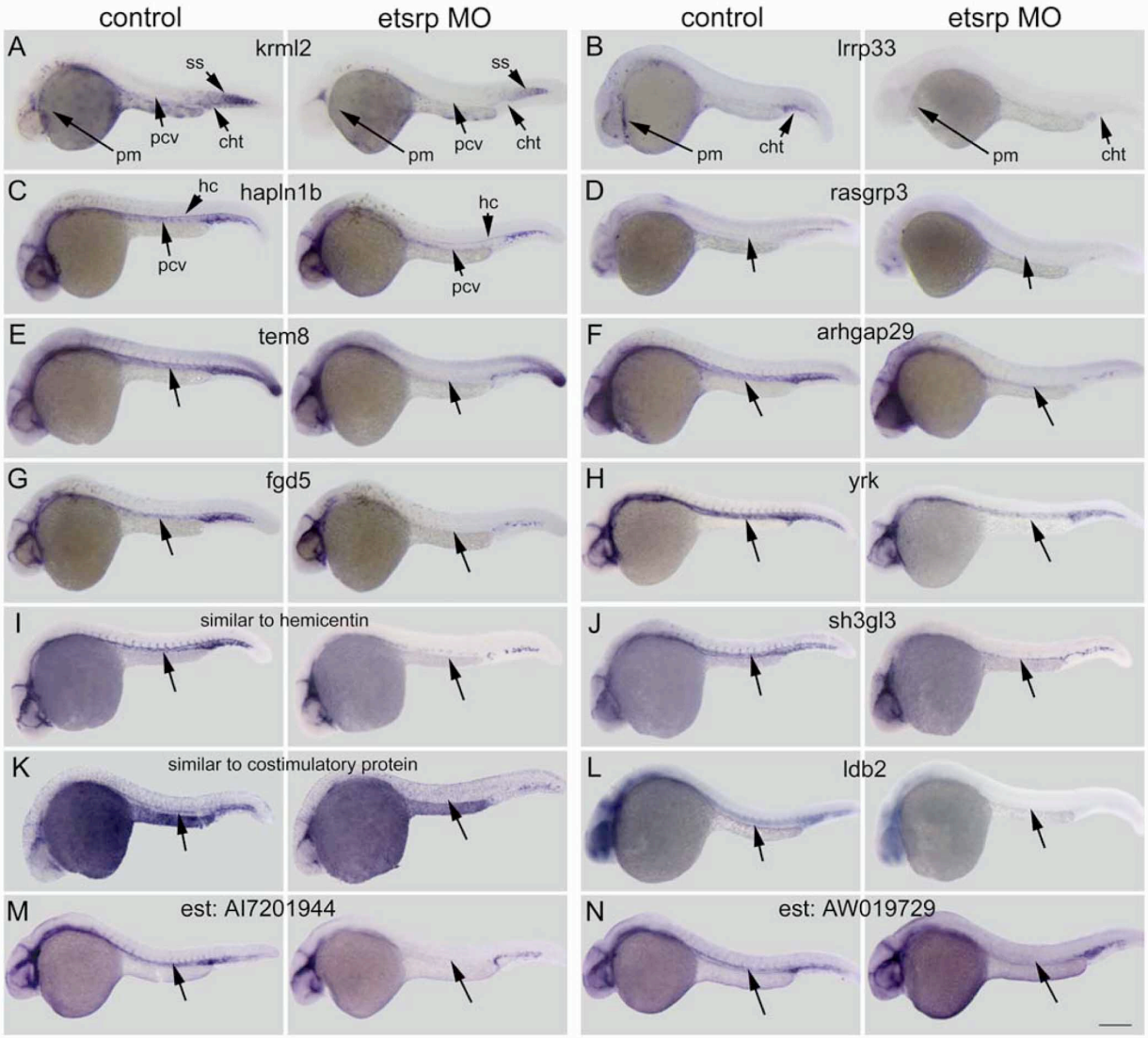


Figure 5. Morpholino knockdown of *etsrp* reduces expression of genes identified by *etsrp* overexpression microarray analysis. Embryos were injected with a combination of two *etsrp* targeting morpholinos (4 ng each) and gene expression was examined by in situ hybridization at 24–26 hours post fertilization. Representative control embryos are on the left of each panel and *etsrp* morpholino (MO) injected embryos are on the right. (A) *krml2* is reduced in primitive myeloid cells, posterior cardinal vein, and caudal hematopoietic tail region by *etsrp* knockdown. (B) *lrrp33* expression is absent in myeloid cells and reduced in the caudal hematopoietic tail region. (C) *hapln1b* expression is lost in the posterior cardinal vein when *etsrp* is knocked down. (D–N) *Etsrp* knockdown causes a significant decrease of gene expression, most pronounced in the trunk (unlabeled arrows). (D) *rasgfp3*; (E) *tem8*; (F) *arhgap29*; (G) *fgd5*; (H) *yrk*; (I) *similar to hemicentin*; (J) *sh3gl3*; (K) *similar to costimulatory protein*; (L) *ldb2*; (M) *est:AI7201944*; and (N) *est:AW019729*. Note that non-hematovascular tissues such as somites in (A) and hypochord in (C) are not affected by *etsrp* knockdown. Abbreviations: pm, primitive myeloid; pcv, posterior cardinal vein; cht, caudal hematopoietic tail region; ss, somites; and hc, hypochord. Scale bar: 250 μm.
doi:10.1371/journal.pone.0004994.g005

humans, *HAPLN-1,-2,-3,-4* [55]. Hapln1, also known as cartilage link protein, is a secreted extracellular matrix protein that stabilizes aggrecan-hyaluronan complexes in cartilage and other tissues, with the knockout in mice resulting in dwarfism and craniofacial abnormalities [56]. Zebrafish also contain four *hapln* genes, *hapln-1a*, *-1b*, *-2*, and *-3*. During early stages of zebrafish development, *hapln1b* is expressed in the hypochord, cranial vasculature, posterior cardinal vein, and caudal hematopoietic tail region (Figure 4L and [57]).

Functional Analysis of *hapln1b* and *sh3gl3*

To determine if the genes highlighted here might play a functional role in vascular development we used antisense morpholino (MO) oligos to knockdown *hapln1b* and *sh3gl3* in *flk:gfp* transgenic fish. *Hapln1b* knockdown using 2 ng of morpholino results in the arrested angiogenic sprouting of many intersomitic vessels (asterisks in Figure 6B, 6D and 6F), preventing them from forming stereotypical dorsolateral anastamosing vessels (dlavs) by 48 hpf (arrow in Figure 6D). Furthermore, the vascular

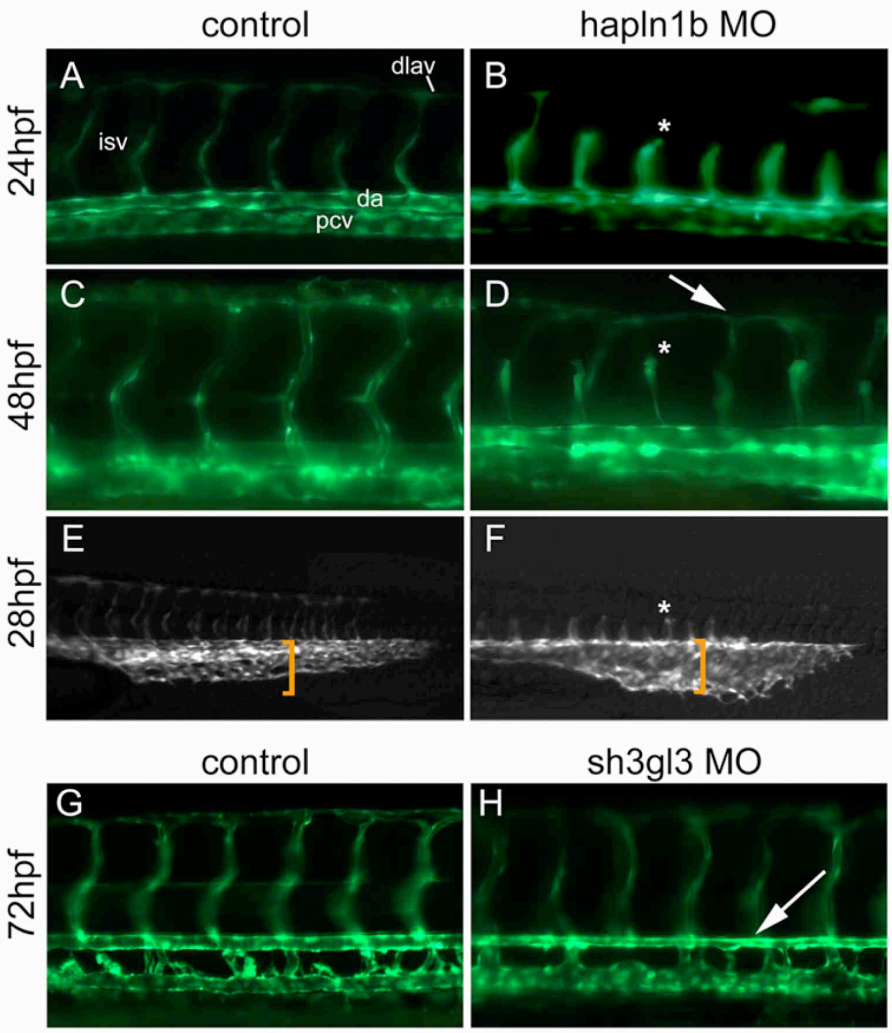


Figure 6. *Hapln1b* and *sh3gl3* are mediators of vascular development. (A–F) *Flk1:gfp* transgenic embryos injected with 2 ng of a translation blocking morpholino (MO) targeting *hapln1b* arrests angiogenic sprouting of the intersomitic vessels (asterisks in panels B, D, F; compare to their counterparts in A, C, and E respectively), resulting in delayed and improper dorsal longitudinal anastomotic vessel formation (arrow in D). Furthermore, the caudal vascular plexus is dilated relative to wild-type controls at 28 hours post fertilization (hpf) (compare the identical sized bracket in panels E and F). (G and H) Injection of 3 ng of translation blocking MO targeting *sh3gl3* results in a thinner dorsal aorta (arrow in H) relative to wild type controls (G) at 72 hpf. Abbreviations: dlav, dorsal longitudinal anastomotic vessels; isv, intersomitic vessels; da, dorsal aorta; and pcv, posterior cardinal vein.
doi:10.1371/journal.pone.0004994.g006

remodeling that occurs at the caudal vascular plexus around 28–30 hpf is severely defective in *hapln1b* morphants (compare bracketed region in Figure 6E and 6F). These defects were observed in 87 out of 92 (92.6%) embryos injected with *hapln1b*-MO while standard control MO (5 ng) injected embryos never exhibited this phenotype (0 out of 70 embryos). *Sh3gl3* morphants (3 ng) had a less severe phenotype than *hapln1b* morphants. *Sh3gl3*-MO treated embryos had relatively normal intersomitic vessel development, but displayed reduced circulation as visualized by red blood cell movement in the axial vasculature when compared to controls (data not shown). A closer examination of the axial vessels at 72 hpf revealed a significant reduction in the diameter of the dorsal aorta with a concomitant increase in the diameter of the posterior cardinal vein in morphants (arrow in Figure 6H compared to 6G). This constriction of the dorsal aorta was likely the cause of the observed decrease in circulation. 61 of 83 (73.5%) embryos injected with *sh3gl3*-MO displayed the shrunken dorsal aorta/reduced circulation phenotype while none of the standard control MO injected (5 ng, 0 out of 70 embryos) or uninjected control embryos did. The demonstration that *hapln1b* or *sh3gl3* knockdown results in distinct vascular phenotypes suggest that other genes identified in this gene expression screen will be functionally important and should be studied in further detail.

Discussion

The transcription factor Etsrp is a critical mediator of vasculogenesis, angiogenesis, and primitive myelopoiesis in zebrafish. In this study we have used gene expression profiling to identify a group of genes that are induced by *etsrp* OE during early

development. Among these genes are many previously identified vascular and myeloid genes, supporting the previous identified function of *etsrp* in vascular and myeloid development. The array data was validated by WISH in *etsrp* OE embryos. Gene expression was then examined at 24 hpf for expression in vascular or blood lineages. Genes that met these criteria were then assayed for expression in embryos with *etsrp* knocked down by morpholino antisense. Fourteen genes were found to be both ectopically induced by *etsrp* OE and down regulated in *etsrp* knockdown embryos. Of these genes two, *krml2* and *lrrp33*, are found in myeloid cells and the remaining twelve are vascular. Any expression of these genes outside of hematopoietic or vascular tissue was unaffected by *etsrp* knockdown demonstrating the specificity of *etsrp*'s function. Two newly identified vascular genes, *hapln1b* and *sh3gl3*, were then independently knocked down by morpholino antisense. Each knockdown resulted in distinct vascular phenotypes demonstrating the utility of the screening method used and the importance of the genes identified. Together, these studies have identified genes downstream of *etsrp* that are likely to mediate its observed role in primitive myeloid and early vascular development in zebrafish.

The *ER71/Etv2* mammalian homolog of *etsrp* has been shown to mediate vascular development in mice. Given that *etsrp* has a similar function in zebrafish, it is likely that many of the genes identified in this microarray screen will have mammalian homologs functionally important for angiogenesis and/or vasculogenesis. Of the fourteen genes we focused on, ten have confirmed mammalian homologs. The remaining four cannot yet be determined due to uncertainty in the assembly of the zebrafish genomic sequence. However, it is expected that mammalian homologs for these genes will be found in the future.

Interestingly, *ER71/Etv2* is critical for primitive erythropoiesis in mice, while *etsrp* knockdown or mutation in zebrafish does not affect primitive erythropoiesis. The data presented here supports the previous finding that *etsrp* regulates primitive myelopoiesis and vasculogenesis but not erythropoiesis since genes related to the former group are highly induced by *etsrp* OE while no erythroid genes are induced. So it appears that *ER71/Etv2* and *etsrp* have evolutionarily divergent functions in the context of primitive erythropoiesis. One possible explanation for this difference is that zebrafish primitive erythropoiesis may not progress through an endothelial cell intermediate while mammalian primitive erythropoiesis in the yolk sac blood islands does. Alternatively, the evolutionary function of *etsrp* in primitive erythrocytes may have been co-opted by another ets transcription factor in zebrafish, possibly due to a historical genome duplication event [58]. Another possibility is that the observed difference is due to differential timing of primitive erythrocyte cell death in *etsrp/ER71/Etv2* null fish and mammals. In *etsrp* mutant fish, primitive erythroid cells marked by *gata1* expression are initially specified but then later lost through apoptosis [15]. In *ER71* null mice, no primitive erythrocytes are observed but an increase in apoptosis was noted [18]. This suggests the possibility that primitive erythrocytes are initially specified in *ER71* null mice, but quickly die due to compromised vasculature as happens in fish.

An issue arising from these studies is that *etsrp* OE results in the induction of many nonhematopoietic and nonvascular genes (Figure S2). It is likely that *etsrp* OE directly mis-targets genes normally regulated by other Ets family members. When overexpressed at abnormally high levels it might also overcome repressive mechanisms and thereby transactivate non-specific genes. Included in the list of genes ectopically induced are many *hox* genes that may represent a secondary response to *etsrp* OE. We have found that *krml2* is expressed in myeloid cells and axial vasculature. It was reported that the expression of *hoxb3* was induced by *krml2* overexpression [59]. In this study, not only is *krml2* induced, but *hoxb3* is also induced about five-fold. Ectopic induction of nonvascular/nonmyeloid target genes may also be the result of secondary effects of *etsrp* overexpression. A possible mechanism is that *etsrp* induces premature differentiation of vascular and myeloid cells and these cells signal to adjacent cells to differentiate into additional nonvascular/nonmyeloid cell types. The strong ectopic induction of genes in the early embryo emphasizes that Etsrp is a very potent, early acting transactivator whose roles in gene regulation must be tightly controlled.

It should also be noted that the overexpression of *etsrp* alone is sufficient to induce many other transcription factors such as *scl* and *fli1a*. The overexpression of *scl* has also been demonstrated to induce other blood and vascular specific genes alone [60]. Recently, it was proposed that *fli1a* appears to be at the top of the transcriptional network driving blood and vascular development. Constitutively active *fli1a*, in which the artificial transactivation domain of the viral VP16 protein was fused to full-length *fli1a* coding sequences, can ectopically induce blood and vascular genes [61]. However, overexpression of *fli1a* alone is not sufficient to induce blood or vascular markers as demonstrated here for *etsrp*. This argues that *etsrp* is more likely the top regulator of hemangioblast development.

The two *etsrp* target genes, *hapln1b* and *sh3gl3*, examined by morpholino gene knockdown each displayed interesting vascular phenotypes. Hapln1b is a member of a protein family that functions as extracellular linkers for hyaluronans and proteoglycans. *Hapln1b* MO treated embryos display truncated intersomitic vessels and disorganized caudal vascular plexus. This phenotype is potentially due to decreased angioblast cell migration or adhesion due to extracellular matrix problems. Future studies will be necessary to establish the exact cellular defect responsible for the vascular phenotype observed in *hapln1b* morphants. *Sh3gl3* is an intracellular protein whose homologs have been suggested to regulate endocytosis and intracellular membrane trafficking [62]. The phenotype observed in the *sh3gl3* morphant, reduced dorsal aorta and increased posterior cardinal vein diameter, is suggestive of disrupted notch signaling as notch signaling mediates artery/vein specification and relative size [63]. Indeed endocytotic mechanisms are critical for proper notch signaling [64]. We hypothesize that *sh3gl3* may regulate artery/vein size and notch signaling through receptor endocytosis.

In summary, we have identified a set of genes induced by *etsrp* OE in zebrafish embryos. Within this set we have identified multiple genes that have restricted expression in the vascular or myeloid lineages and are reduced by morpholino knockdown of *etsrp*. These genes are likely to play important roles in the development of their respective tissues. Additionally, this data can be used in the future to clarify gene hierarchies necessary for hemangioblast differentiation by comparing with mammalian stem cell or zebrafish mutant, morphant, or overexpression gene array data, as we have already done with *cloche*. The results presented here provide a basis for future studies of the hemangioblast.

Materials and Methods

Microarray

flk1:gfp transgenic embryos were injected with 75 pg of synthetic *etsrp* mRNA, transcribed with Ambion's mMessage mMachine T3 polymerase kit from XbaI linearized *etsrp*-T3TS as previously described [19], and harvested at ~80% epiboly in pools of 70 embryos per group. The transcription profile of uninjected transgenic embryos was used as controls. Total RNA was isolated

with Trizol reagent (Invitrogen), according to manufacturer's instructions. The detailed protocol for microarray hybridization and normalization procedures were described by Horak et al [65] and are available upon request. First-strand cDNA probes were generated by aminoallyl dUTP incorporation and then coupled to either Cy3 or Cy5. The resulting cDNA probes were purified, concentrated, and then hybridized to arrays. Each chip contains 34,647 printed oligo elements (Compugen, Operon, and MWG) designed from zebrafish EST assemblies and representing approximately 20,000 genes, approximately 60% of the total predicted genes according to the public Ensembl database. After hybridization, the slides were washed, dried, and scanned using an Agilent DNA microarray scanner (Agilent Technologies) at 635 nm (Cy5) and then at 532 nm (Cy3). Fluorescence intensities were quantified using Agilent feature extraction software (Agilent Technologies), and changes in expression levels were determined by comparing the signal intensities between *etsrp* OE and uninjected control groups. Hybridizations were performed four times, switching the fluorescent labeling to eliminate biases caused by the labeling process. Samples were normalized using the Lowess calculations and cutoffs for significance were set at a two-fold change in either direction. Oligo sequences were mapped to multiple databases, including RefSeq, UniGene, Ensembl, TIGR, and genomic coordinates to maximally determine gene identity and function. Data were deposited into searchable FilemakerPro and Excel databases for analysis.

Quantitative RT-PCR

Total RNA was purified with Trizol reagent (Invitrogen) from equal numbers of embryos between experimental groups, and single strand cDNA was synthesized with Superscript II reverse transcriptase (Invitrogen). Real-time PCR was performed using the iCycler iQ Real-Time PCR Detection System (BioRad, Hercules, CA) with iQ SYBR Green Supermix (BioRad). Gene expression levels were measured by the $\Delta\Delta$ C_t method, comparing *etsrp* OE or H2B-mRFP injected groups relative to uninjected controls, with *β-actin* used as the reference gene. Pre-microarray validation was performed from the same samples used for the microarray experiment, and biological triplicates were examined for the injection control experiment reported in Figure S1. Quantitative RT-PCR primers are listed in Table S1.

Gene cloning and RNA whole mount *in situ* hybridization (WISH)

Genes were cloned into the pCRII Topo vector using the TOPO TA cloning kit (dual promoter) according to manufacturer's instructions (Invitrogen), after amplifying genes of interest from a 24 hpf wild type embryo cDNA library. The primers used are noted in Table S2. Positive clones were analyzed by restriction digest and confirmed by sequencing. WISH was performed as described in [66]. DIG labeled RNA probes were generated by linearizing TOPO-cloned genes with restriction endonuclueases (New England Biolabs), and transcribing with SP6 or T7 RNA polymerase (Promega).

Embryo injections

Injections were performed with borosilicate glass microcapillaries (Harvard apparatus) that were pulled with a flaming/brown micropipette puller, model P-97 (Sutter Instrument Co) on a PLI-90 microinjector (Harvard apparatus). For microarray validation by WISH, an *etsrp-mcherry* construct was generated by sub-cloning *mcherry* from *mcherry-tol2* (a gift from Hong Zhang) into pSPORT6 between XbaI and XhoI, and *etsrp* was subsequently subcloned into this vector at the 5′end of *mcherry* between BamHI and XbaI sites. To test the requirement of *etsrp* for the expression of the blood/vascular related genes a mixture of *etsrp* morpholinos, MO1 and MO2 (4 ng each) described in [19] was injected into embryos at 1–2 cell stage, and harvested for WISH at 24–28 hpf. Translation blocking morpholinos targeting the first codon of *hapln1b* and *sh3gl3* were synthesized by Gene-Tools, LLC. *Hapln1b* MO sequence: 5′-GAGCAGGAAGGTCATCTTGATTTAT-3′; *sh3gl3* MO sequence: 5′-TTAAACCCGGCCACTGA-CATCCTTC-3′; Standard control MO sequence: 5′-CCTCTTACCTCAGTTACAATTTATA-3′.

Image acquisition and processing

In situ stained embryos were further processed by a serial dehydration in ethanol, followed by rehydration into 1×PBS. Embryos were imaged either in 1×PBS on agarose coated dishes, or in 2% methylcellulose in depression microscope slides. Images were captured with a color digital CCD camera (Axiocam, Zeiss) mounted on a dissecting microscope (Stemi SV11, Zeiss) with Openlab 4.0.2 software (Improvision, Lexington, MA). Fluorescence images were obtained with the same equipment but on an Axioskop2 plus compound microscope with 10× and 20× objections. In several cases separate images were captured per embryo and areas in focus were compiled with Adobe Photoshop CS version 8.

Supporting Information

Figure S1 Quantitative RT-PCR of etsrp overexpression and H2B-mRFP injection controls for selected genes. Relative gene expression levels were calculated for etsrp or H2B-mRFP1 expressing as compared to uninjected control embryos at 80 percent epiboly to gastrulation stages for the genes indicated on the x-axis. The induction of endogenous etsrp (3′UTR amplicon from injections encoding etsrp lacking the 3′UTR), fli1a, scl, yrk, and atxr1/tem8 are statistically significant ($p<0.05$, Student's t-test). Although slightly increased hapln1b and sh3gl3 are not significantly different from controls.
Found at: doi:10.1371/journal.pone.0004994.s001 (2.50 MB TIF)

Figure S2 Genes induced by etsrp overexpression but not expressed in blood or vessels. (A) arhgef9; (B) pchdg; (C) novel2; (D) LOC555375 (expression is in ventral somites not vasculature); (E) cbe2; (F) smoc2; (G) LOC550501; (H) similar to testican; (I) hoxb8b; (J) tex2; (K) c1orf192; (L) c20.orf112; (M) znf385; (N) apolipoB; (O) ptgs1; and (P) hoxc3a. Scale bar: 250 um.
Found at: doi:10.1371/journal.pone.0004994.s002 (6.74 MB TIF)

Table S1 Quantitative RT-PCR primers
Found at: doi:10.1371/journal.pone.0004994.s003 (0.03 MB DOC)

Table S2 Primers used to clone in situ hybridization probes
Found at: doi:10.1371/journal.pone.0004994.s004 (0.09 MB DOC)

Acknowledgments

We thank Hong Jiang and Anqi Liu for assistance with in situ hybridization and animal husbandry.

Author Contributions

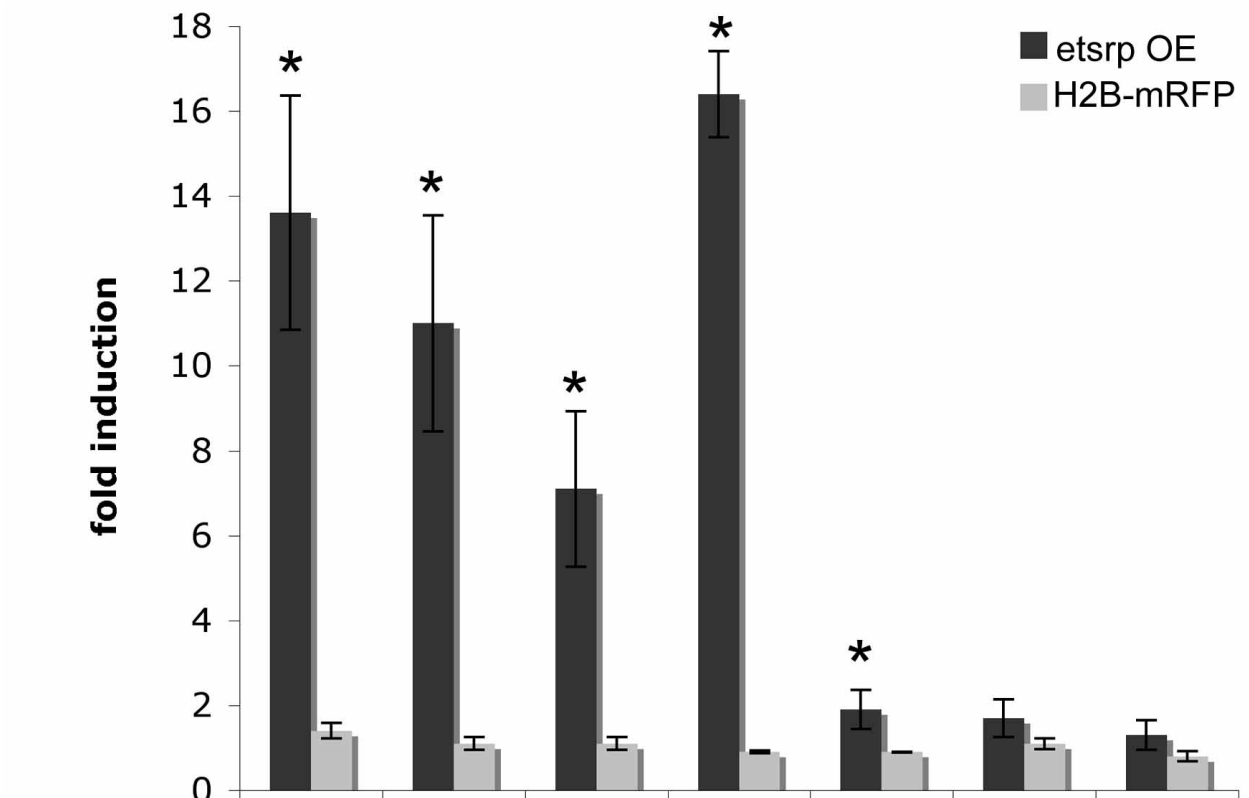
Conceived and designed the experiments: SL. Performed the experiments: GAG MBV YZ SB. Analyzed the data: SL GAG MBV. Wrote the paper: SL GAG MBV.

References

1. Ciau-Uitz A, Walmsley M, Patient R (2000) Distinct origins of adult and embryonic blood in Xenopus. Cell 102: 787–796.
2. Eichmann A, Corbel C, Nataf V, Vaigot P, Breant C, et al. (1997) Ligand-dependent development of the endothelial and hemopoietic lineages from embryonic mesodermal cells expressing vascular endothelial growth factor receptor 2. Proc Natl Acad Sci U S A 94: 5141–5146.
3. Schmidt A, Brixius K, Bloch W (2007) Endothelial precursor cell migration during vasculogenesis. Circ Res 101: 125–136.
4. Shalaby F, Rossant J, Yamaguchi TP, Gertsenstein M, Wu XF, et al. (1995) Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. Nature 376: 62–66.
5. Walmsley M, Ciau-Uitz A, Patient R (2002) Adult and embryonic blood and endothelium derive from distinct precursor populations which are differentially programmed by BMP in Xenopus. Development 129: 5683–5695.
6. Nishikawa SI, Nishikawa S, Hirashima M, Matsuyoshi N, Kodama H (1998) Progressive lineage analysis by cell sorting and culture identifies FLK1+VE-cadherin+ cells at a diverging point of endothelial and hemopoietic lineages. Development 125: 1747–1757.
7. Choi K, Kennedy M, Kazarov A, Papadimitriou JC, Keller G (1998) A common precursor for hematopoietic and endothelial cells. Development 125: 725–732.
8. Palis J, Robertson S, Kennedy M, Wall C, Keller G (1999) Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. Development 126: 5073–5084.
9. Dieterlen-Lievre F, Le Douarin NM (2004) From the hemangioblast to self-tolerance: a series of innovations gained from studies on the avian embryo. Mech Dev 121: 1117–1128.
10. Vogeli KM, Jin SW, Martin GR, Stainier DY (2006) A common progenitor for haematopoietic and endothelial lineages in the zebrafish gastrula. Nature 443: 337–339.
11. Stainier DY, Weinstein BM, Detrich Z, Li F, Mc (1995) Cloche, an early acting zebrafish gene, is required by both the endothelial and hematopoietic lineages. Development 121: 3141–3150.
12. Qian F, Zhen F, Ong C, Jin SW, Meng Soo H, et al. (2005) Microarray analysis of zebrafish cloche mutant using amplified cDNA and identification of potential downstream target genes. Dev Dyn 233: 1163–1172.
13. Sumanas S, Jorniak T, Lin S (2005) Identification of novel vascular endothelial-specific genes by the microarray analysis of the zebrafish cloche mutants. Blood 106: 534–541.
14. Weber GJ, Choe SE, Dooley KA, Paffett-Lugassy NN, Zhou Y, et al. (2005) Mutant-specific gene programs in the zebrafish. Blood 106: 521–530.
15. Pham VN, Lawson ND, Mugford JW, Dye L, Castranova D, et al. (2007) Combinatorial function of ETS transcription factors in the developing vasculature. Dev Biol 303: 772–783.
16. Sumanas S, Gomez G, Zhao Y, Park C, Choi K, et al. (2008) Interplay between Etsrp/ER71, scl and alk8 signaling controls endothelial and myeloid cell formation. Blood 111: 4500–4510.
17. Liu F, Patient R (2008) Genome-wide analysis of the zebrafish ETS family identifies three genes required for hemangioblast differentiation or angiogenesis. Circ Res 103: 1147–1154.
18. Lee D, Park C, Lee H, Lugus JJ, Kim SH, et al. (2008) ER71 acts downstream of BMP, Notch, and Wnt signaling in blood and vessel progenitor specification. Cell Stem Cell 2: 497–507.
19. Sumanas S, Lin S (2006) Ets1-related protein is a key regulator of vasculogenesis in zebrafish. PLoS Biol 4: e10.
20. Jin SW, Beis D, Mitchell T, Chen JN, Stainier DY (2005) Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. Development 132: 5199–5209.
21. Liao W, Bisgrove BW, Sawyer H, Hug B, Bell B, et al. (1997) The zebrafish gene cloche acts upstream of a flk-1 homologue to regulate endothelial cell differentiation. Development 124: 381–389.
22. Coolen M, Sii-Felice K, Bronchain O, Mazabraud A, Bourrat F, et al. (2005) Phylogenomic analysis and expression patterns of large Maf genes in Xenopus tropicalis provide new insights into the functional evolution of the gene family in osteichthyans. Dev Genes Evol 215: 327–339.
23. Moens CB, Cordes SP, Giorgianni MW, Barsh GS, Kimmel CB (1998) Equivalence in the genetic control of hindbrain segmentation in fish and mouse. Development 125: 381–391.
24. Schvarzstein M, Kirn A, Haffter P, Cordes SP (1999) Expression of Zkrml2, a homologue of the Krml1/val segmentation gene, during embryonic patterning of the zebrafish (Danio rerio). Mech Dev 80: 223–226.
25. Eichmann A, Grapin-Botton A, Kelly L, Graf T, Le Douarin NM, et al. (1997) The expression pattern of the mafB/kr gene in birds and mice reveals that the kreisler phenotype does not represent a null mutant. Mech Dev 65: 111–122.
26. Sieweke MH, Tekotte H, Frampton J, Graf T (1996) MafB is an interaction partner and repressor of Ets-1 that inhibits erythroid differentiation. Cell 85: 49–60.
27. Kobe B, Kajava AV (2001) The leucine-rich repeat as a protein recognition motif. Curr Opin Struct Biol 11: 725–732.
28. Meijer AH, Gabby Krens SF, Medina Rodriguez IA, He S, Bitter W, et al. (2004) Expression analysis of the Toll-like receptor and TIR domain adaptor families of zebrafish. Mol Immunol 40: 773–783.
29. Takeda K, Kaisho T, Akira S (2003) Toll-like receptors. Annu Rev Immunol 21: 335–376.
30. Kopp EB, Medzhitov R (1999) The Toll-receptor family and control of innate immunity. Curr Opin Immunol 11: 13–18.
31. Xu Y, Tao X, Shen B, Horng T, Medzhitov R, et al. (2000) Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. Nature 408: 111–115.
32. Rebhun JF, Castro AF, Quilliam LA (2000) Identification of guanine nucleotide exchange factors (GEFs) for the Rap1 GTPase. Regulation of MR-GEF by M-Ras-GTP interaction. J Biol Chem 275: 34901–34908.
33. Roberts DM, Anderson AL, Hidaka M, Swetenburg RL, Patterson C, et al. (2004) A vascular gene trap screen defines RasGRP3 as an angiogenesis-regulated gene required for the endothelial response to phorbol esters. Mol Cell Biol 24: 10515–10528.
34. Aiba Y, Oh-hora M, Kiyonaka S, Kimura Y, Hijikata A, et al. (2004) Activation of RasGRP3 by phosphorylation of Thr-133 is required for B cell receptor-mediated Ras activation. Proc Natl Acad Sci U S A 101: 16612–16617.
35. St Croix B, Rago C, Velculescu V, Traverso G, Romans KE, et al. (2000) Genes expressed in human tumor endothelium. Science 289: 1197–1202.
36. Scobie HM, Rainey GJ, Bradley KA, Young JA (2003) Human capillary morphogenesis protein 2 functions as an anthrax toxin receptor. Proc Natl Acad Sci U S A 100: 5170–5174.
37. Bradley KA, Mogridge J, Mourez M, Collier RJ, Young JA (2001) Identification of the cellular receptor for anthrax toxin. Nature 414: 225–229.
38. Bolcome S, Se Z, R B, ApC, Rj C (2008) Anthrax lethal toxin induces cell death-independent permeability in zebrafish vasculature. Proc Natl Acad Sci U S A 105: 2439–2444.
39. Werner E, Kowalczyk AP, Faundez V (2006) Anthrax toxin receptor 1/tumor endothelium marker 8 mediates cell spreading by coupling extracellular ligands to the actin cytoskeleton. J Biol Chem 281: 23227–23236.
40. Hotchkiss KA, Basile CM, Spring SC, Bonuccelli G, Lisanti MP, et al. (2005) TEM8 expression stimulates endothelial cell adhesion and migration by regulating cell-matrix interactions on collagen. Exp Cell Res 305: 133–144.
41. Saras J, Franzen P, Aspenstrom P, Hellman U, Gonez LJ, et al. (1997) A novel GTPase-activating protein for Rho interacts with a PDZ domain of the protein-tyrosine phosphatase PTPL1. J Biol Chem 272: 24333–24338.
42. Miller RA, Christoforou N, Pevsner J, McCallion AS, Gearhart JD (2008) Efficient array-based identification of novel cardiac genes through differentiation of mouse ESCs. PLoS ONE 3: e2176.
43. Hall A (2005) Rho GTPases and the control of cell behaviour. Biochem Soc Trans 33: 891–895.
44. Zheng Y, Fischer DJ, Santos MF, Tigyi G, Pasteris NG, et al. (1996) The faciogenital dysplasia gene product FGD1 functions as a Cdc42Hs-specific guanine-nucleotide exchange factor. J Biol Chem 271: 33169–33172.
45. Narumiya H, Hidaka K, Shirai M, Terami H, Aburatani H, et al. (2007) Endocardiogenesis in embryoid bodies: novel markers identified by gene expression profiling. Biochem Biophys Res Commun 357: 896–902.
46. Martins-Green M, Bixby JL, Yamamoto T, Graf T, Sudol M (2000) Tissue specific expression of Yrk kinase: implications for differentiation and inflammation. Int J Biochem Cell Biol 32: 351–364.
47. Vogel BE, Hedgecock EM (2001) Hemicentin, a conserved extracellular member of the immunoglobulin superfamily, organizes epithelial and other cell attachments into oriented line-shaped junctions. Development 128: 883–894.
48. Schultz DW, Klein ML, Humpert AJ, Luzier CW, Persun V, et al. (2003) Analysis of the ARMD1 locus: evidence that a mutation in HEMICENTIN-1 is associated with age-related macular degeneration in a large family. Hum Mol Genet 12: 3315–3323.
49. Xu X, Dong C, Vogel BE (2007) Hemicentins assemble on diverse epithelia in the mouse. J Histochem Cytochem 55: 119–126.
50. Peter BJ, Kent HM, Mills IG, Vallis Y, Butler PJ, et al. (2004) BAR domains as sensors of membrane curvature: the amphiphysin BAR structure. Science 303: 495–499.
51. Giachino C, Lantelme E, Lanzetti L, Saccone S, Bella Valle G, et al. (1997) A novel SH3-containing human gene family preferentially expressed in the central nervous system. Genomics 41: 427–434.
52. So CW, Sham MH, Chew SL, Cheung N, So CK, et al. (2000) Expression and protein-binding studies of the EEN gene family, new interacting partners for dynamin, synaptojanin and huntingtin proteins. Biochem J 348 Pt 2: 447–458.
53. Yang S, Toy K, Ingle G, Zlot C, Williams PM, et al. (2002) Vascular endothelial growth factor-induced genes in human umbilical vein endothelial cells: relative roles of KDR and Flt-1 receptors. Arterioscler Thromb Vasc Biol 22: 1797–1803.
54. Toyama R, Kobayashi M, Tomita T, Dawid IB (1998) Expression of LIM-domain binding protein (ldb) genes during zebrafish embryogenesis. Mech Dev 71: 197–200.
55. Spicer AP, Joo A, Bowling (2003) A hyaluronan binding link protein gene family whose members are physically linked adjacent to chondroitin sulfate proteoglycan core protein genes: the missing links. J Biol Chem 278: 21083–21091.
56. Watanabe H, Yamada Y (1999) Mice lacking link protein develop dwarfism and craniofacial abnormalities. Nat Genet 21: 225–229.

57. Kudoh T, Tsang M, Hukriede NA, Chen X, Dedekian M, et al. (2001) A gene expression screen in zebrafish embryogenesis. Genome Res 11: 1979–1987.
58. Woods IG, Kelly PD, Chu F, Ngo-Hazelett P, Yan YL, et al. (2000) A comparative map of the zebrafish genome. Genome Res 10: 1903–1914.
59. Manzanares M, Cordes S, Kwan CT, Sham MH, Barsh GS, et al. (1997) Segmental regulation of Hoxb-3 by kreisler. Nature 387: 191–195.
60. Gering M, Rodaway AR, Gottgens B, Patient RK, Green AR (1998) The SCL gene specifies haemangioblast development from early mesoderm. EMBO J 17: 4029–4045.
61. Liu F, Walmsley M, Rodaway A, Patient R (2008) Fli1 acts at the top of the transcriptional network driving blood and endothelial development. Curr Biol 18: 1234–1240.
62. Ringstad N, Gad H, Low P, Di Paolo G, Brodin L, et al. (1999) Endophilin/SH3p4 is required for the transition from early to late stages in clathrin-mediated synaptic vesicle endocytosis. Neuron 24: 143–154.
63. Kim YH, Hu H, Guevara-Gallardo S, Lam MT, Fong SY, et al. (2008) Artery and vein size is balanced by Notch and ephrin B2/EphB4 during angiogenesis. Development 135: 3755–3764.
64. Le Borgne R (2006) Regulation of Notch signalling by endocytosis and endosomal sorting. Curr Opin Cell Biol 18: 213–222.
65. Horak CE, Lee JH, Elkahloun AG, Boissan M, Dumont S, et al. (2007) Nm23-H1 suppresses tumor cell motility by down-regulating the lysophosphatidic acid receptor EDG2. Cancer Res 67: 7238–7246.
66. Thisse C, Thisse B (2008) High-resolution in situ hybridization to whole-mount zebrafish embryos. Nat Protoc 3: 59–69.

Chapter 1 Supplementary Material



	etsrp 3′UTR	*fli1a*	*scl*	*yrk*	*atxr1*	*hapln1b*	*sh3gl3*
etsrp OE	*13.6*	*11*	*7.1*	*16.4*	*1.9*	*1.7*	*1.3*
H2B-mRFP	*1.4*	*1.1*	*1.1*	*0.9*	*0.9*	*1.1*	*0.8*

Figure S1. Quantitative RT-PCR of *etsrp* overexpression and H2B-mRFP injection controls for selected genes. Relative gene expression levels were calculated for *etsrp* or H2B-mRFP1 expressing as compared to uninjected control embryos at 80 percent epiboly to gastrulation stages for the genes indicated on the x-axis. The induction of endogenous *etsrp* (3′UTR amplicon from injections encoding etsrp lacking the 3′UTR), *fli1a, scl, yrk,* and *atxr1/tem8* are statistically significant ($p<0.05$, Student's t-test). Although slightly increased *hapln1b* and *sh3gl3* are not significantly different from controls. doi:10.1371/journal.pone.0004994.s001

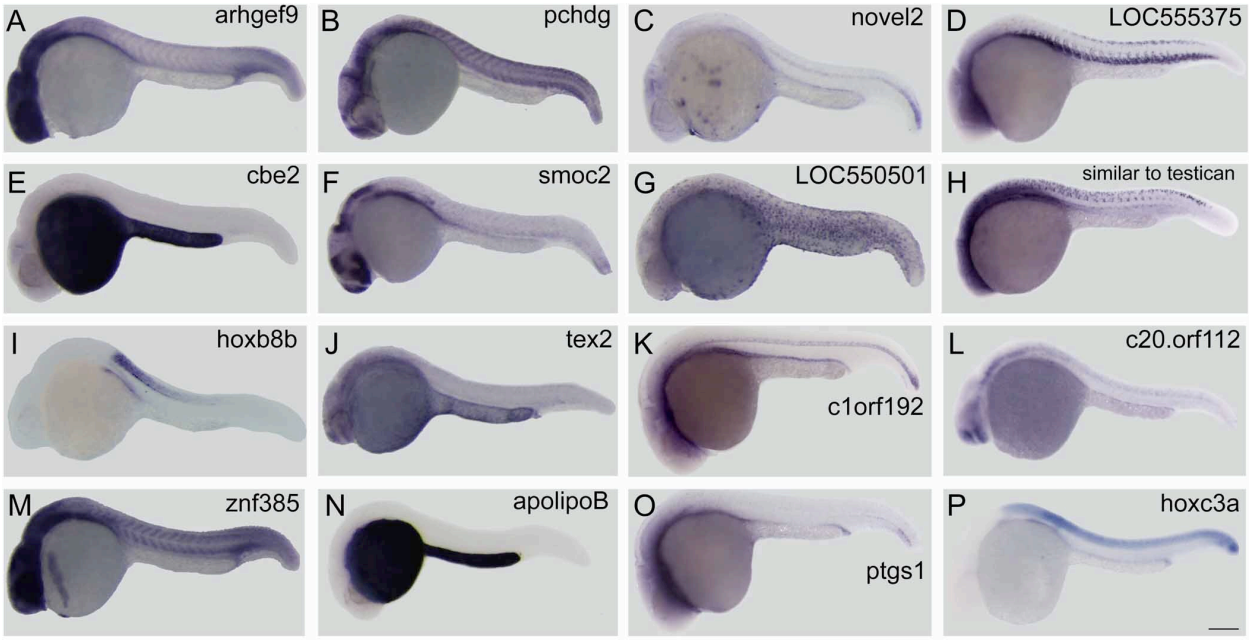


Figure S2. Genes induced by *etsrp* overexpression but not expressed in blood or vessels. (A) *arhgef9*; (B) *pchdg*; (C) *novel2*; (D) *LOC555375* (expression is in ventral somites not vasculature); (E) *cbe2*; (F) *smoc2*; (G) *LOC550501*; (H) *similar to testican*; (I) *hoxb8b*; (J) *tex2*; (K) *c1orf192*; (L) *c20.orf112*; (M) *znf385*; (N) *apolipoB*; (O) *ptgs1*; and (P) *hoxc3a*. Scale bar: 250 um.
doi:10.1371/journal.pone.0004994.s002

Table S1. Quantitative RT-PCR primers

Gene	Forward primer	Reverse primer
flk1	5'- CATGTTGGTGGGACACTCAC-3'	5'-CTGCAGAGCAGTTGAGGATG-3'
scl	GGAGATGCGGAACAGTATGG	GAAGGCACCGTTCACATTCT
β-actin	TGTTTTCCCCTCCATTGTTG	ACATACATGGCAGGGGTGTT
etsrp.3UTR	GAGGAATTCTCGAAGGATTGG	TGGTTTTCTAAAGGCACCTAGC
fli1A	CCGAGGTCCTGCTCTCACAT	GGGACTGGTCAGCGTGAGAT
yrk	CTGAAGCCTTCCTGGATGAG	CTAGCCAGACCGAAATCTGC
atxr1/tem8	GTCGGGGGAATTAAACGAAT	GGTAGTCTGGTGGAGGTGGA
hapln1b	GATGGGCTTCCACAAAAAGA	GGTTTGGTGATGGGGTATTG
sh3gl3	GAGCCTTCATCGATCCTCTG	CACTGTGCTCGCAACTGATT

Table S2. Primers used to clone in situ probes for selected genes

CLONE # PRIMER	PRIMER SEQUENCE	PRODUCT SIZE (Kb)
Arhgef9.fwd	5'- GGTCAACGTGATTCTCATATTCTTC -3'	< 5
Arhgef9.rvs	ACATCCCATGGTTTCTTACATTG	
cbe2.fwd	GGAAAAGCAGCCATGATAAAAAG	1.7
cbe2.rvs	TTCAGGCTGTGGGAGGTG	
Krml2.fwd	GACGACAAACAGGCTATAGTTTAAGAG	1.5
Krml2.rvs	TGAAAATGTTACAACTGCTTCGA	
Hoxb8b.fwd	GCGCATCCAATGATACAAGG	1.1
Hoxb8b.rvs	TTGAATACATGATGCAACAGCC	
Rasgrp3.fwd	TAGAAGTTACACGTCGGATTTGC	3.3
Rasgrp3.rvs	CCTATCGGCTCAAAGACTCACT	
Lrrp33.fwd	ATGTCAAGCACAAGAGAGACCA	2.3
Lrrp33.rvs	GAAATGAGATACATTCACATATTTGCA	
Znf385.fwd	AGCGCGGAGACCAACTTG	2.4
Znf385.rvs	TCAGTACGGTGAGAACAAGATTG	
Lrrp51.fwd	GTCTTGACTGTAACTTTTCTATGGTGG	1.0

Lrrp51.rvs	GTATGTCACGTTACACTGAAAAAAATG	
Smoc2.fwd	AAACTGAATTTCTGACCAGTGTGTT	1.0
Smoc2.rvs	ATGAAGAACTTTCAACAAAATGATCTG	
Pcdhg.fwd	TCATTCTCTTGGATTTACAGAATTTTC	6.2
Pcdhg.rvs	TATTAGCATATAAAATGCAAGTAGCAC	
Novel2.fwd	AGACTAACACACATTAGCATCCGA	0.5
Novel2.rvs	TAATGACAAAGAATGCCCACAT	
LOC550501.fwd	TAACAGGACCTTCACTGCTATAAGG	1.9
LOC550501.rvs	GAAAATTGCAGACATTACAAATGAA	
Tex2.fwd	CACATGTTTCCTTTGGTGTCTAG	3.3
Tex2.rvs	ACATGTCCTCGAGAGATTCTTTAGG	
apolipoB.fwd	TCTGAAGGCTTCAGCCACTT	1.5
apolipoB.rvs	CAGAACAAGGATGACAATAGGTTG	
Tem8.fwd	GAGCCGAAGAAGATGTGGAG	1.4
Tem8.rvs	CTGAGCTTTGGCTGCTTCTT	
C1orf192.fwd	AAGTCGCAAAAGCTCCAGAA	1.1
C1orf192.rvs	GTGGACCATCACATCCACAA	
Arhgap29.fwd	GTGGCTCTTTGAGCTCCATC	1.2
Arhgap29.rvs	TCGCACTGAAACACAAGCTC	
Fgd5.fwd	CTGGATCATGTGGGAGTGTG	1.1
Fgd5.rvs	TGAACAACAGCAGATGATGG	
Ptgs1.fwd	GCTGAAGTGGACGGTGATTT	1.1
Ptgs1.rvs	AGTGGGTGCCAGTGGTAAAG	
Hapln1b.fwd	CACAGCACAACCCAAAGTGA	0.9
Hapln1b.rvs	ACGCCGTAGAGTTTGTGCTT	
Loc555375.fwd	GACAATTACCAACCCGGAGA	0.6
Loc555375.rvs	TGGTCTCCACATGCTCAGTC	
Sim.to testican3 fwd	TGAAGATGCTTGTGCTGGTC	1.1
Sim.to testican3 rvs	CGTCTCCCATTTCCTCTTCA	
Yrk.fwd	GGAAGGCCAGTAAAGCTGTG	1.3
Yrk.rvs	ATGCCGAAAGACCATACGTC	
Sh3gl3.fwd	GGCCGGGTTTAAGAAACAGT	1.0
Sh3gl3.rvs	GAATGATGATGTCGCCCTCT	
Ldb2.fwd	CACGACCCCTTCTACTCGTC	1.1
Ldb2.rvs	GCTGTTCCAGGGACTGTTGT	
Sim.hemicentin.fwd	TTGGGAAAGCCTCACTTTTG	1.1
Sim.hemicentin.rvs	TCACATCCGACCAATCAGAA	
Sim.costim.fwd	TTCCTGTGGACTTTGCCAGT	1.2
Sim.costim.rvs	AGGGCCTTTTCCTGTTGAAT	
EST:AI721944.fwd	TGCAGCCCAATTAAGTGACA	0.5
EST:AI721944.fwd	GTGCACCTTGAGTTTTGCAC	
EST:AW019729.fwd	TGCAAGTGTTGCCTTTAATGTT	0.6
EST:AW019729.rvs	CCACAAATATTCCTCATAATCCTTTT	
Hoxc3a.fwd	GCCACCACGTGACTACATTG	1.4
Hoxc3a.rvs	TGCCTTGGGAAAAGTAGCAT	
C20.orf112.fwd	TTTTTGATATATAGCACACGGTTTT	0.4
C20.orf112.rvs	AAATATTGGCAACCCATTTCC	

OPEN ACCESS Freely available online

PLoS one

Identification of Vascular and Hematopoietic Genes Downstream of *etsrp* by Deep Sequencing in Zebrafish

Gustavo Gomez[1], Jae-Hyung Lee[2], Matthew B. Veldman[1], Jing Lu[1], Xinshu Xiao[2], Shuo Lin[1]*

1 Department of Molecular, Cell and Developmental Biology, University of California Los Angeles, Los Angeles, California, United States of America, **2** Department of Integrative Biology and Physiology, University of California Los Angeles, Los Angeles, California, United States of America

Abstract

The transcription factor *etsrp/Er71/Etv2* is a master control gene for vasculogenesis in all species studied to date. It is also required for hematopoiesis in zebrafish and mice. Several novel genes expressed in vasculature have been identified through transcriptional profiling of zebrafish embryos overexpressing *etsrp* by microarrays. Here we re-examined this transcriptional profile by Illumina RNA-sequencing technology, revealing a substantially increased number of candidate genes regulated by *etsrp*. Expression studies of 50 selected candidate genes from this dataset resulted in the identification of 39 new genes that are expressed in vascular cells. Regulation of these genes by *etsrp* was confirmed by their ectopic induction in *etsrp* overexpressing and decreased expression in *etsrp* deficient embryos. Our studies demonstrate the effectiveness of the RNA-sequencing technology to identify biologically relevant genes in zebrfish and produced a comprehensive profile of genes previously unexplored in vascular endothelial cell biology.

Citation: Gomez G, Lee J-H, Veldman MB, Lu J, Xiao X, et al. (2012) Identification of Vascular and Hematopoietic Genes Downstream of *etsrp* by Deep Sequencing in Zebrafish. PLoS ONE 7(3): e31658. doi:10.1371/journal.pone.0031658

Editor: Henry H. Roehl, University of Sheffield, United Kingdom

Received October 28, 2011; **Accepted** January 17, 2012; **Published** March 16, 2012

Funding: This work was supported by grants from the National Institutes of Health (NIH) (F31HL091713 to GAG and DK54508 to SL). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: shuolin@ucla.edu

Introduction

The cardiovasacular system, which includes the heart, vessels and blood, function together to deliver oxygen and nutrients to cells throughout the body and remove metabolic waste. Understanding the development of this system is instrumental to the advancement of both basic and clinical sciences. The zebrafish, *Danio rerio*, is an excellent model organism for such studies due to embryo transparency, high fecundity, and fast development of organogenesis. In particular, the cardiovascular system is formed within one day of birth [1,2]. Through genetic, cellular and molecular studies in zebrafish, a great deal of knowledge regarding the molecular components and cellular events that establish this system has been obtained. It is notable that many of the key molecular players and events that drive organogenesis in zebrafish are evolutionarily and functionally conserved with other organisms, including mammals.

A critical transcription factor required for the development of vascular endothelial cells is *ets related protein*, *etsrp*, which was originally discovered through transcriptional profiling of the cardiovascular mutant line, *cloche*, that lacks blood and vasculature [3] and independently by a mutagenesis screen [4]. It was subsequently found that *etsrp* is both necessary and sufficient to induce both the vascular endothelial and primitive myelopoietic program in zebrafish [5,6]. Its significance was further underscored by the identification and characterization of the functional homolog *ets variant 2*, *etv2*, in mammals [7,8] and Xenopus [9,10]. The *etsrp* transgenic fish line has been useful to reveal the cellular events that establish the cranial vasculature and myelopoiesis [11], and it is possible that *etsrp* might be associated with the initiation of the definitive hematopoietic program [12].

The zebrafish genome is now in its ninth version and remains incompletely annotated. Nonetheless bioinformatic approaches have been used to identify important genes encoding transcription factors containing the ETS box DNA binding domain in hematopoietic and endothelial development [13]. The use of microarrays by the zebrafish research community has increased the identification of genes expressed in the developing cardiovascular system [3,14,15,16,17,18,19]. The arrival of higher throughput next generation sequencing has expanded the possibilities to deepen our understanding and identification of novel genes, and has already proven its utility for studies in zebrafish. Although the majority of the published studies combining RNA-sequencing and zebrafish have focused on immunity [20,21,22], the use of high throughput sequencing to study other biological subjects in zebrafish has begun to increase [23,24,25]. The ability of this technological approach to examine the zebrafish transcriptome at greater depth than microarrays without an *a priori* bias prompted us to re-examine the transcriptional profile of embryos overexpressing *etsrp*. In this study we present a panel of 39 more genes that are expressed in the developing zebrafish vasculature that were identified by this approach.

Results

As previously demonstrated, the overexpression of *etsrp* results in the induction of vascular related genes during gastrulation stages of development, before the onset of angioblast specification [5]. In this study *etsrp* RNA was injected into one cell *flk1-gfp* transgenic embryos and their transcriptome profiles were examined at later stages of gastrulation, 70–90% epiboly, when ectopic induction of *flk1-gfp* is detected (Figure 1). Total RNA was extracted from pools

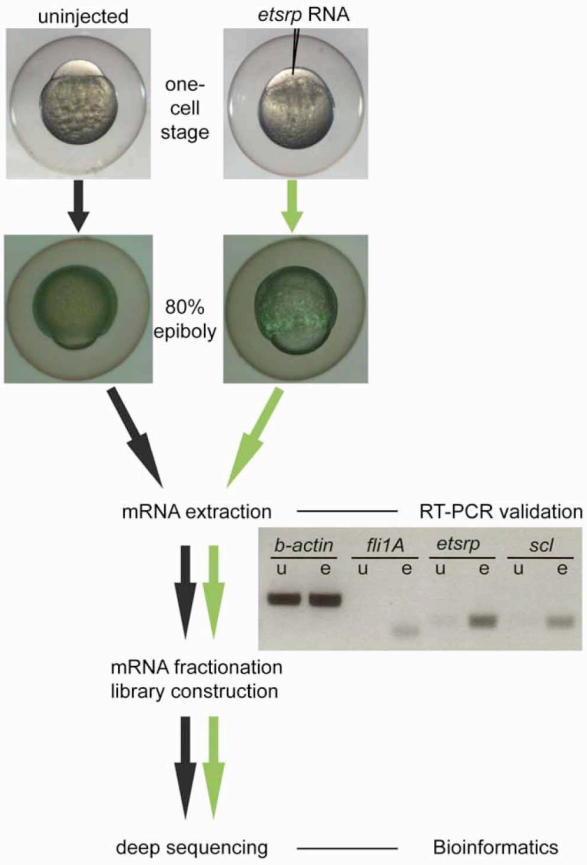


Figure 1. Workflow diagram. Embryos were either uninjected or injected with *etsrp* RNA at the one cell stage, then raised until the late gastrulation stages when the *flk1-gfp* transgenic reporter is induced ectopically. Equal paired groups of embryos were pooled for mRNA extraction, library construction, and solexa mRNA-sequencing. Samples were pre-validated prior to sequencing by RT-PCR with *fli1A*, the 3′UTR of *etsrp*, as well as *scl*, which were all preferentially induced as expected by *etsrp* overexpression (e) relative to uninjected controls (u).
doi:10.1371/journal.pone.0031658.g001

of 125 embryos, and the induction of known hemangioblast markers, *fli1a*, *scl* and *etsrp*, was verified in the *etsrp* overexpressed group by RT-PCR before library construction and sequencing. Two separate sets of RNA-seq libraries were constructed from independent samples that were pre-validated using the same validation scheme as demonstrated in the workflow diagram on Figure 1.

A total of 123 million single-end or paired-end reads were obtained and 35% of the reads were mapped onto the unigene transcript sequence database (Build #117) [26] using the Burrows Wheeler Aligner (BWA) alignment program [27]. Table S1 lists the number of reads in each control and *estrp* overexpressed (*estrp* oe) samples and the mapping results. About 43 million reads were uniquely mapped in total, which covered 77% of the unigene database (39,784 out of 51,481). The expression level difference between the control and *estrp* oe samples were compared, resulting in the upregulation of 849 unigene entries above a 1.9-fold cutoff value in the *etsrp* oe group with a corrected p-value$<$0.05, while 726 entries were downregulated using the same parameters. A sample of the top 35 hits of annotated upregulated genes obtained in this dataset is listed in Table 1.

In order to determine the biological relevance of the gene pool obtained here, a gene ontology analysis was performed with the Database for Annotation, Visualization and Integrated Discovery Gene Ontology (DAVID GO) program, as demonstrated in the identification of novel genes expressed in myeloid cells in zebrafish [28]. Of the 849 upregulated unigene hits examined by functional annotation clustering, 797 were queried successfully, but because some unigene entries are associated with more than one gene, and some genes are represented by more than one unigene identifier, a total of 787 entries were evaluated resulting in 148 clusters. The remaining 52 entries were not identified either because they represent expressed sequence tags not associated with a characterized gene, they are unique to zebrafish, or their orthologs were unidentifiable by DAVID GO gene conversion tools. Conversely, the 726 downregulated hits resulted in the successful query of 306 entries divided among 50 clusters, but relative to the induced clusters, these had rather low fold-enrichment scores.

The top two clusters from the upregulated data contained genes associated with "vascular development", "angiogenesis," and other terms related to vascular endothelial cell biology. These data were combined resulting in 68 genes after duplicate deletion, and examined individually for involvement in vascular development. 18 genes were excluded because although several are expressed in mesoderm or heart, their specific expression or function within endothelial cells remains to be demonstrated. We did examine one of these, *xirp2a*, which we found to have endothelial expression. These were noted as a subset of potential vascular genes in Table S2, which also includes 12 that were removed because there is no indication that they are associated with endothelial cells. By searching the list generated containing 849 unigene entries, 15 more vascular related genes were identified (Table 2), bringing the total number of genes associated with endothelial cells in this dataset to 53.

In addition to endothelial specific genes, there were 4 genes of the myeloid lineage that were preferentially induced by *etsrp* oe. They are *lplastin*, *csf3r*, *cebp1*, *hsd3b7*. This observation is consistent with what has been previously reported by our group [6]. Overall using the DAVID GO program and visual inspection of the RNA-seq dataset we observed that a substantial number of genes (6.2%) are associated with developing endothelial biology, confirming the biological relevance of the upregulated dataset. Contrarily, DAVID GO analysis of the downregulated genes did not result in a distinct biological category of genes, as genes associated with the development of all three germ layers including endoderm, mesoderm, and ectoderm were repressed.

In order to further biologically validate the data and identify new genes expressed in endothelial cells, several genes were selected from the upregulated dataset for analysis by RNA whole mount in situ hybridization (WISH) at the 80% epiboly stages, when they are predicted to be upregulated by overexpressing *etsrp* (Figure 2). For gene selection, the potential of the top 350 hits to be expressed in endothelial cells was evaluated thoroughly, while the remaining portion was assessed at random. Genes examined were selected based on a PUBMED literature search using the following criteria: 1. Not previously shown to be expressed in vascular endothelial cells. 2. Previously detected in endothelial cells but evidence is limited to *in vitro* studies. 3. Has not been examined in developing vasculature. One of the top ranked genes (highest fold-induction), *similar to Src homology 2 domain containing E*, *she*, was used as a positive control since its expression in the vasculature has been reported previously [19]. The gene set examined is listed on Table 3. This assay resulted in the clear ectopic induction by *etsrp* overexpression in 49 of the 50 genes examined (Figure 2). The high percentage of validated genes induced by *etsrp* overexpression

Table 1. Top 35 annotated genes induced by *etsrp*.

Unigene ID	Gene ID	Gene Description	Fold Change	p-value
Dr.24158	*srgn*	*serglycin*	1920.8	7.7E-53
Dr.153782	*LOC100151241*	*similar to loc559821 protein*	1899.9	4.5E-26
Dr.83594	*crabp1a*	*cellular retinoic acid binding protein 1a*	917.4	5.1E-17
Dr.141228	*LOC558783*	*similar to cell death activator cide-a*	452.8	1.1E-05
Dr.149118	*LOC798186*	*similar to cdc42 gtpase-activating protein*	441.5	1.9E-21
Dr.118570	*LOC561493*	*similar to src homology 2 domain containing e*	315.3	1.8E-196
Dr.40434	*ms4a17a.11*	*membrane-spanning 4-domains, subfamily a, member 17a.11*	300.4	3.5E-06
Dr.133475	*LOC100151130*	*similar to rasip1 protein*	254.8	3.6E-10
Dr.37870	*ckmb*	*creatine kinase, muscle b*	231.6	3.8E-07
Dr.92232	*efcab2*	*ef-hand calcium binding domain 2*	231.5	8.4E-03
Dr.84343	*zgc:153721*	*zgc:153721*	225.9	1.1E-09
Dr.85502	*mrc1*	*mannose receptor c1-like protein*	225.7	4.0E-20
Dr.47591	*etv2*	*ets variant gene 2*	219.0	4.9E-237
Dr.110840	*LOC100006361*	*similar to preprogalanin 1b*	217.2	4.9E-03
Dr.103328	*zgc:162298*	*zgc:162298*	198.7	2.0E-05
Dr.86665	*nhlh2*	*nescient helix loop helix 2*	180.6	2.1E-08
Dr.114377	*zgc:173594*	*zgc:173594*	175.7	2.8E-03
Dr.110713	*LOC100149611*	*similar to loc495463 protein*	165.4	3.8E-08
Dr.84866	*LOC100005159*	*similar to hemicentin 1*	159.4	9.6E-04
Dr.111220	*gpr183*	*g protein-coupled receptor 183*	151.4	1.9E-04
Dr.113631	*LOC568486*	*similar to regulating synaptic membrane exocytosis 1*	136.1	2.5E-02
Dr.82184	*LOC569386*	*platelet endothelial cell adhesion molecule-like*	123.4	6.2E-06
Dr.124732	*LOC100148619*	*similar to myelin associated glycoprotein*	113.9	1.4E-02
Dr.152900	*LOC572378*	*similar to transforming growth factor-b type ii receptor*	110.1	4.9E-03
Dr.142587	*csf3r*	*colony stimulating factor 3 receptor (granulocyte)*	104.9	6.5E-05
Dr.100658	*LOC100004843*	*hypothetical loc100004843*	104.1	2.8E-03
Dr.149641	*LOC558126*	*similar to hcg2040171*	103.1	4.9E-03
Dr.151919	*si:ch211-10e8.4*	*si:ch211-10e8.4*	102.9	8.4E-03
Dr.81739	*kcnh2*	*potassium voltage-gated channel, subfamily h, member 2*	101.5	2.0E-09
Dr.78041	*xirp2*	*xin actin-binding repeat containing 2*	95.0	7.6E-56
Dr.118849	*zgc:63958*	*zgc:63958*	91.4	1.7E-03
Dr.102595	*LOC568153*	*similar to chemokine receptor-like 1*	86.4	2.8E-03
Dr.115871	*LOC569038*	*similar to cholecystokinin a receptor*	85.0	2.5E-02
Dr.83127	*LOC100001838*	*similar to loc100001772 protein*	79.5	1.4E-02

Several unigene hits identified in the top list correspond to unannotated expressed sequence tags, est's, and these were removed from this table. The dataset containing the data in its entirety including the est entries that were removed from this table are on Table S7.
doi:10.1371/journal.pone.0031658.t001

by WISH supports the validity of both the deep sequencing and bioinformatic methods used to examine the transcriptional profiles of upregulated genes. The genes predicted to be suppressed by this approach was also evaluated by quantitative RT-PCR, but only 2of 8 genes examined were validated (Table S3). This suggests that this approach is more amenable to the examination of gene induction by *etsrp*. Nevertheless, the complete dataset containing downregulated unigene entries is listed in Table S8.

To determine whether the 50 selected ectopically induced genes are expressed in the developing embryonic zebrafish vasculature, embryos were processed for WISH at a developmental stage when most of the primitive vasculature has formed, 24-hour post fertilization (24hpf). Marked expression in the vasculature was noted for 39 of the 50 genes (Figure 3). To confirm that these genes are functionally downstream of *etsrp* in the vasculature, we examined their expression in *etsrp* morphants, in which case the expression in the axial vasculature is prominently reduced as demonstrated in Figure 4. Sparse expression in myeloid cells was only detected for one of these genes, *myo1F* (Figure 3L). Although there is some ubiquitous expression of some of these genes, the more pronounced expression in the vasculature is clearly demonstrated in higher magnification images of the trunk regions (Figure S1), where the axial vascular expression in morphants is clearly reduced in all genes. Of the 39 genes with vascular expression only *fhl3* (Figure 3 AD′ and Figure S1AD′) and *acsbg2* (Figure 3AL′ and Figure S1AL′) are preferentially expressed in the axial venous system at this stage.

In the course of gene selection and WISH probe generation the Ensembl databases were utilized together with those at NCBI to examine gene structure and evolutionary ontology. Several of the

Table 2. Known vascular genes induced by *etsrp* in RNA-seq dataset.

Unigene ID	Gene ID	Gene Description	Fold Change	p-value
Dr.118570	*LOC561493*	*similar to src homology 2 domain containing e*	315.3	1.8E-196
Dr.133475	***LOC100151130***	***similar to rasip1 protein***	**254.8**	**3.62E-10**
Dr.85502	*mrc1*	*mannose receptor c1-like protein*	225.7	4.0E-20
Dr.47591	*etv2*	*ets variant gene 2*	219.0	4.9E-237
Dr.84866	*LOC100005159*	*similar to hemicentin 1*	159.4	9.6E-04
Dr.152900	***LOC572378***	***similar to transforming growth factor-b type ii receptor***	**110.1**	**4.90E-03**
Dr.107226	*sc:d0254*	*sc:d0254*	37.9	5.0E-10
Dr.78408	*fli1a*	*friend leukemia integration 1a*	37.6	3.0E-61
Dr.86352	*zgc:113016*	*zgc:113016*	36.3	1.5E-09
Dr.36543	*aqp8a*	*aquaporin 8a*	32.6	4.9E-237
Dr.83871	*rasgrp3*	*ras guanyl releasing protein 3 (calcium and dag-regulated)*	32.6	2.6E-08
Dr.89996	***egfl7***	***egf-like-domain, multiple 7***	**21.6**	**1.5E-10**
Dr.79866	*yrk*	*yes-related kinase*	20.9	7.0E-151
Dr.103153	*plxnd1*	*plexin d1*	20.1	4.9E-237
Dr.118013	*cdh5*	*cadherin 5*	19.1	3.5E-18
Dr.80968	*fli1b*	*friend leukemia integration 1b*	16.3	1.8E-07
Dr.77989	*dusp5*	*dual specificity phosphatase 5*	14.8	1.2E-16
Dr.76027	***ker18***	***keratin 18***	**12.5**	**2.1E-05**
Dr.75812	*tal1*	*t-cell acute lymphocytic leukemia 1*	12.4	6.9E-08
Dr.150623	***f10***	***coagulation factor x***	**9.8**	**7.2E-04**
Dr.89035	*tmem88a*	*transmembrane protein 88 a*	9.2	7.2E-16
Dr.75094	*kdrl*	*kinase insert domain receptor like*	7.8	1.6E-36
Dr.80363	*clec14a*	*c-type lectin domain family 14, member a*	6.3	4.0E-06
Dr.151971	*sox7*	*sry-box containing gene 7*	6.1	4.4E-02
Dr.74559	***scarf1***	***scavenger receptor class f, member 1***	**5.6**	**1.4E-02**
Dr.75958	***robo4***	***roundabout homolog 4***	**5.4**	**3.7E-14**
Dr.87001	***cldn5b***	***claudin5b***	**5.1**	**1.6E-15**
Dr.83306	*mcam*	*melanoma cell adhesion molecule*	4.7	4.0E-39
Dr.52827	*zfpm2b*	*zinc finger protein, multitype 2b*	4.4	1.5E-05
Dr.5660	*crip2*	*cysteine-rich protein 2*	4.4	9.4E-11
Dr.107483	***sb:cb911***	***sb:cb911***	**4.3**	**3.1E-03**
Dr.75385	*LOC563577*	*similar to novel apoptosis-stimulating protein of p53*	4.2	3.1E-06
Dr.132454	***smox***	***spermine oxidase***	**3.2**	**4.8E-55**
Dr.599	*ldb2a*	*lim-domain binding factor 2a*	3.2	3.7E-03
Dr.81683	*rbpms2*	*rna binding protein with multiple splicing 2*	3.1	1.2E-09
Dr.81298	*flt4*	*fms-related tyrosine kinase 4*	3.0	2.1E-29
Dr.79626	***ildr2***	***immunoglobulin-like domain containing receptor 2***	**2.9**	**1.4E-21**
Dr.132331	***nrp1b***	***neuropilin 1b***	**2.9**	**2.5E-15**
Dr.78142	***acvrl1***	***activin a receptor type ii-like 1***	**2.8**	**6.7E-19**
Dr.135121	*stab2*	*stabilin*	2.7	4.5E-10
Dr.76054	*tpm4*	*tropomyosin 4*	2.6	1.9E-37
Dr.79413	***jam2***	***junctional adhesion molecule 2***	**2.6**	**1.5E-21**
Dr.91385	*kdr*	*kinase insert domain receptor (a type iii receptor tyrosine kinase)*	2.5	2.5E-29
Dr.88777	*pdlim4*	*pdz and lim domain 4*	2.5	2.3E-02
Dr.76395	*C8orf4*	*chromosome 8 open reading frame 4*	2.4	1.0E-13
Dr.37960	*fgfrl1b*	*fibroblast growth factor receptor-like 1b*	2.4	2.1E-07
Dr.75409	*gapdhs*	*glyceraldehyde-3-phosphate dehydrogenase, spermatogenic*	2.3	2.7E-03
Dr.22604	***amot***	***angiomotin***	**2.3**	**2.4E-08**
Dr.80539	*elovl1b*	*elongation of very long chain fatty acids-like 1b*	2.3	2.7E-09
Dr.82429	*LOC563907*	*similar to tumor endothelial marker 8*	2.1	6.5E-03

Table 2. Cont.

Unigene ID	Gene ID	Gene Description	Fold Change	p-value
Dr.104822	*si:dkey-261h17.1*	*si:dkey-261h17.1*	2.1	4.6E-107
Dr.78553	*micall2*	*mical-like 2*	2.0	5.6E-18

Genes in bold text were identified within the *etsrp* overexpressed dataset by visual inspection of the data. Those not in bold were identified by DAVID cluster analysis.
doi:10.1371/journal.pone.0031658.t002

genes probed for were noted to have redundant gene names as of the current genome build, Zv9 in Ensembl. These paralogs have been noted in their description below with an asterisk and further depicted in Table S4 with chromosomal location and identity. Following is a brief description of 10 selected genes in the order that they are presented in Figures 3 and 4. Descriptions for the remaining 29 genes are continued in Text S1.

A. *lyg2*

The goose type lysozyme ortholog *lyg2* was originally identified in goose egg white, where it functions in the catalysis of bacterial cell wall break down, and is expressed ubiquitously in flounder [29]. In chicken, it is expressed in lung and non-adherent fraction of bone marrow cells, possibly reflecting expression within myelocytes where it exhibits an innate immune function [30].

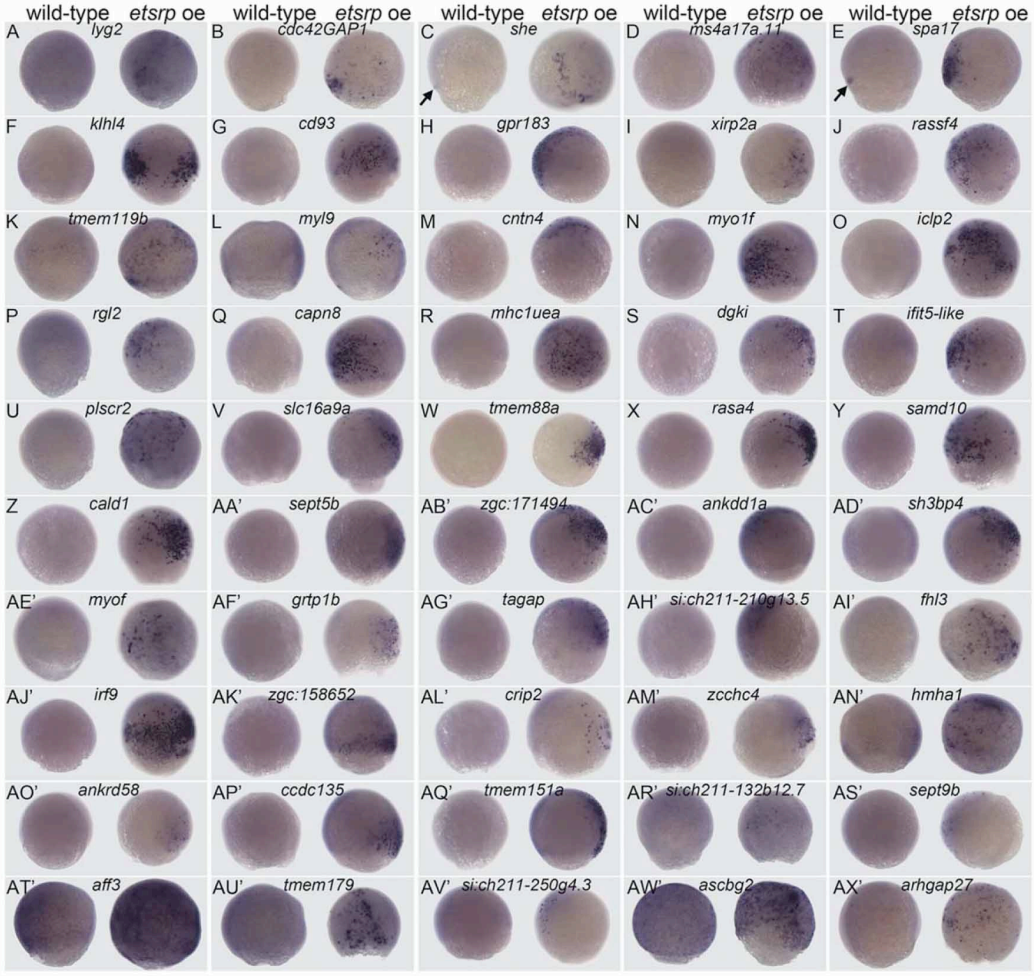


Figure 2. Verification of RNA-seq data. Wild type and *etsrp* injected, *etsrp* oe, embryos were collected at late stages of gastrulation and processed by WISH. Ectopic induction was detected as random positively labeled cells in embryos injected with *etsrp* RNA at the one cell stage. Endogenous expression was observed in the presumptive forerunner cells (arrows) of wild-type uninjected embryos in (C) *she* and (E) *spa17*. Endogenous expression was also observed in (K) *tmem119b* (L) *myl9* (AD′) *sh3bp4* (AN′) *hmha1* (AT′) *aff3* and (AW′) *acsbg2*. Ectopic induction is clearly detected in 49 of the 50 genes examined, but not (AT′) *aff3*, which has a relatively high level of endogenous expression at this stage. Wild type uninjected embryos are positioned on the left column with their *etsrp* oe counterparts on the right for each gene, anterior is facing left where possible. Scale bar: 250 μm.
doi:10.1371/journal.pone.0031658.g002

Table 3. Genes examined from the RNA-seq *etsrp* overexpression dataset.

Unigene ID	Gene ID	Gene Description	Fold Change	p-value
Dr.153782	***lyg2***	***lysozyme G like 2***	**1899.9**	**4.5E-26**
Dr.149118	***cdc42GAP1***	***similar to Cdc42 GTPase-activating protein***	**441.5**	**1.9E-21**
Dr.118570	***she***	***similar to Src homology 2 domain containing E***	**315.3**	**1.8E-196**
Dr.40434	*ms4a17a.11*	*membrane-spanning 4-domains, subfamily A, member 17A.11*	300.4	3.5E-06
Dr.84343	***spa17***	***sperm auto antigen 17***	**225.9**	**1.1E-09**
Dr.103328	***klhl4***	***kelch like 4***	**198.7**	**2.0E-05**
Dr.110713	***LOC100149611***	***similar to LOC495463 protein ; ''CD93''***	**165.4**	**3.8E-08**
Dr.111220	***gpr183***	***g protein-coupled receptor 183***	**151.4**	**1.9E-04**
Dr.78041	***xirp2***	***xin actin-binding repeat containing 2***	**95.0**	**7.6E-56**
Dr.134501	***rassf4***	***ras association (RalGDS/AF-6) domain family member 4***	**51.4**	**1.4E-02**
Dr.91332	***tmem119b***	***transmembrane protein 119***	**50.7**	**1.8E-42**
Dr.89765	***myl9***	***myosin, light chain 9, regulatory***	**50.1**	**5.7E-113**
Dr.115399	***cntn4***	***contactin 4***	**38.4**	**1.4E-02**
Dr.84022	***myo1E***	***similar to myosin IE***	**37.9**	**8.4E-03**
Dr.75719	***iclp2***	***invariant chain-like protein 2***	**32.2**	**1.3E-16**
Dr.88587	***rgl2***	***ral guanine nucleotide dissociation stimulator-like 2***	**28.2**	**1.9E-36**
Dr.76656	***capn8***	***calpain 8***	**23.5**	**1.1E-05**
Dr.11010	***mhc1uea***	***major histocompatibility complex class I UEA gene***	**21.8**	**2.4E-21**
Dr.77849	***dgki***	***similar to diacylglycerol kinase, iota***	**17.4**	**4.4E-04**
Dr.117215	***ifit5***	***similar to Interferon-induced protein with tetratricopeptide repeats 5***	**15.1**	**2.1E-03**
Dr.134371	*plscr2*	*similar to phospholipid scramblase 2*	10.6	1.5E-47
Dr.7340	*slc16a9a*	*solute carrier family 16 (monocarboxylic acid transporters), member 9a*	9.3	2.1E-23
Dr.89035	***tmem88a***	***transmembrane protein 88a***	**9.2**	**7.2E-16**
Dr.100033	***rasa4***	***ras p21 protein activator 4-like***	**7.8**	**2.3E-04**
Dr.96217	***samd10***	***similar to sterile alpha motif domain containing 10***	**7.2**	**1.2E-02**
Dr.114623	***cald1***	***caldesmon 1***	**6.8**	**1.6E-22**
Dr.91020	***sept5b***	***septin 5b***	**6.7**	**1.3E-05**
Dr.119058	***zgc:171494***	***zgc:171494***	**6.4**	**2.8E-02**
Dr.84654	***ankdd1a***	***ankyrin repeat and death domain containing 1A***	**6.4**	**2.5E-09**
Dr.91634	***sh3bp4***	***novel protein similar to vertebrate SH3-domain binding protein 4***	**6.0**	**2.0E-06**
Dr.82145	***myof***	***similar to fer-1-like 3, myoferlin***	**5.7**	**4.1E-32**
Dr.84960	***grtp1b***	***growth hormone regulated TBC protein 1b***	**5.6**	**1.4E-02**
Dr.85673	***tagap***	***t-cell activation GTPase activating protein***	**5.5**	**2.8E-03**
Dr.124255	*plc-l2*	*similar to Inactive phospholipase c-like protein 2*	5.2	2.9E-05
Dr.80073	***fhl3***	***four and a half LIM domains 3***	**5.1**	**5.4E-37**
Dr.133138	***irf9***	***interferon regulatory factor 9***	**4.9**	**9.4E-12**
Dr.18530	*FAM166B*	*family with sequence similarity 166, member B*	4.4	1.8E-04
Dr.5660	*crip2*	*cysteine-rich protein 2.*	4.4	9.4E-11
Dr.135601	*zcchc4*	*zinc finger, CCHC domain containing 4*	4.0	4.4E-03
Dr.77065	***hmha1***	***histocompatibility (minor) HA-1***	**4.0**	**9.1E-12**
Dr.47691	***ankrd58***	***ankyrin repeat domain-containing protein 58-like***	**3.8**	**3.1E-06**
Dr.92393	***ccdc135***	***similar to coiled-coil domain-containing protein 135***	**3.6**	**5.2E-05**
Dr.83578	***tmem151b***	***similar to transmembrane protein tmem151B***	**3.5**	**1.9E-03**
Dr.81141	*si:ch211-132b12.7*	*hypothetical protein LOC564531*	3.5	1.4E-02
Dr.78875	*sept9b*	*septin 9b*	3.3	1.3E-05
Dr.83297	***aff3***	***af4/fmr2 family, member 3; LAF***	**2.4**	**1.5E-10**
Dr.90997	*tmem179*	*transmembrane protein 179*	2.0	2.8E-09
Dr.34190	***si:ch211-250g4.3***	***si:ch211-250g4.3***	**2.0**	**4.5E-04**
Dr.107097	***acsbg2***	***Acyl-CoA synthetase bubblegum family member 2***	**1.9**	**2.8E-64**
Dr.77920	***arhgap27***	***similar to Rho GTPase-activating protein 27***	**1.9**	**9.1E-07**

Bold are expressed in vascular endothelial cells at the time examined (24 hpf).
doi:10.1371/journal.pone.0031658.t003

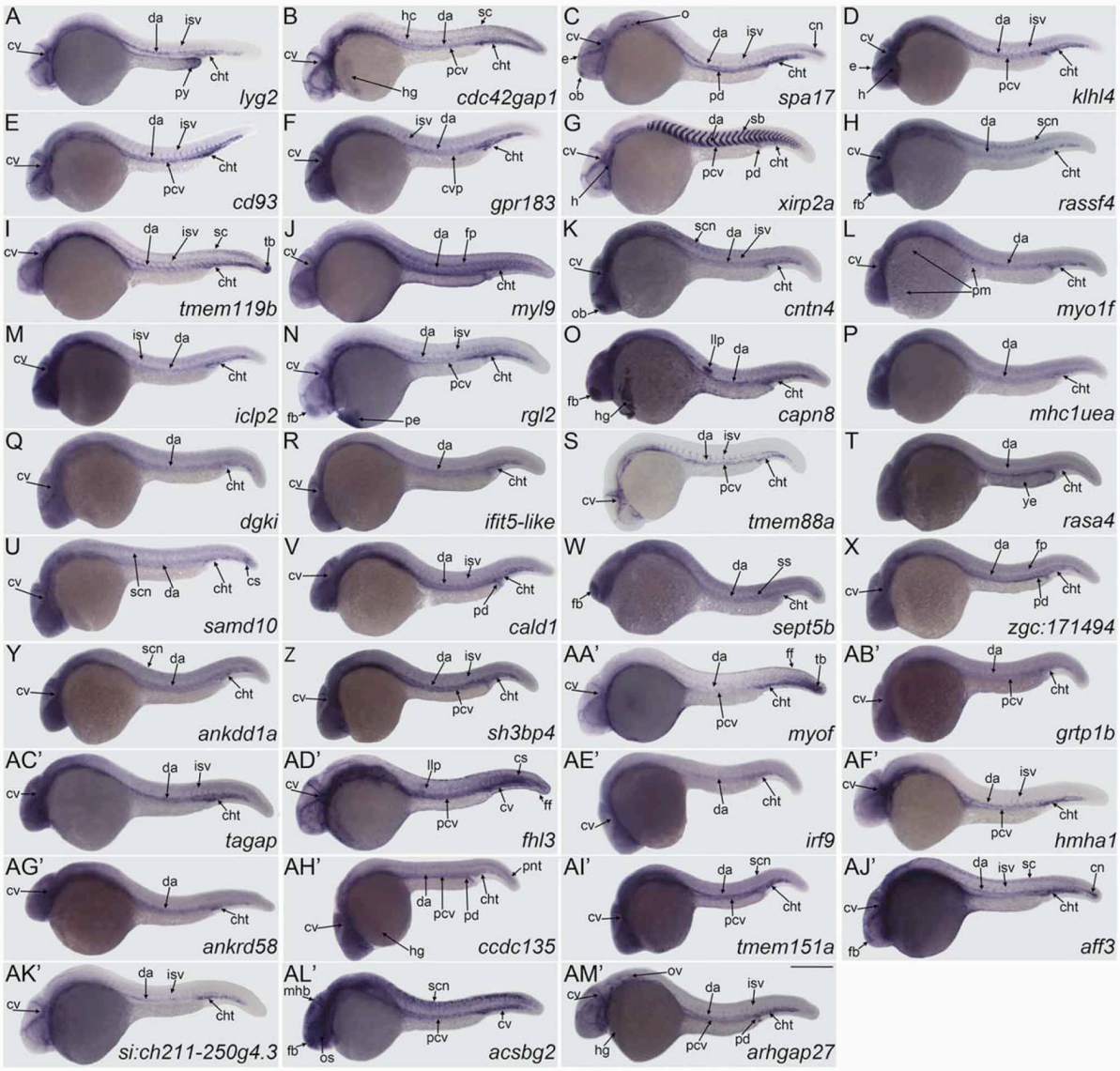


Figure 3. Genes with expression in vascular endothelial cells. The following genes are restricted to the vasculature and are expressed in the cranial vasculature, dorsal aorta, caudal hemtopoietic tail region, posterior cardinal vein, and intersegmental vessels: (E) *cd93*; (F) *gpr183*; (S) *tmem88a*; (AF′) *hmha1*, while the following are expressed in the same structures but exclude the posterior cardinal vein: (A) *lyg2* is also expressed in the posterior yolk extension; (AC′) *tagap*; (AK′) *si:ch211-250g4.3*. These genes are expressed in the cranial vasculature, dorsal aorta, caudal hematopoietic tail region, posterior cardinal vein and other tissues as noted: (B) *cdc42gap1*, is also expressed in the hypochord, hatching gland, and sparsely in spinal cord; (D) *klhl4* is also expressed in the heart, intersegmental vessels, and epiphysis; (G) *xirp2a* is also expressed in the heart, somite boundaries and pronephric duct; (N) *rgl2*, is also expressed in the intersegmental vessels, forebrain and primitive erythorcytes; (AA′) *myof* is also expressed in the fin fold, and tailbud; (AM′) *arhgap27* is also expressed in the intersegmental vessels, otic vesicles, pronephric duct, and hatching gland. The following genes are expressed in the cranial vasculature, dorsal aorta, caudal hematopoietic tail region, and intersegmental vessels as well as (C) *spa17*, the olfactory bulb, epiphysis, otoliths, pronephric duct, and caudal notocord; (H) *rassf4* was not detected in the intersegmental vessels but is observed in the forebrain, floor plate of neural tube, and spinal cord neurons; (I) *tmem119b* is also expressed in spinal cord, and tailbud; (AJ′) *aff3* is also expressed in the forebrain, spinal cord, and caudal notocord. (AD′) *fhl3* is expressed in the cranial vasculature, posterior cardinal vein, caudal vein, fin fold, lateral line primordium, and lightly in caudal somites. (AL′) *acsbg2* is expressed in the posterior cardinal vein, caudal vein, spinal cord neurons, optic stalk, forebrain, and midbrain-hindbrain boundary. The following genes have basal levels of ubiquitous expression with prominent labeling of the cranial vasculature, dorsal aorta, caudal hematopoietic tail region and other structures and tissues where indicated: (J) *myl9* is also expressed in the floor plate; (K) *cntn4*, is also expressed in the intersegmental vessels, olfactory bulb, and spinal cord neurons; (L) *myo1f* is also expressed in primitive myeloid cells; (M) *iclp2*, also has slight expression in anterior intersegmental vessels; (Q) *dgki*; (R) *ifit5-like*; (U) *samd10* is also expressed in the caudal somites, and neurons of the anterior spinal cord; (V) *cald1* is also expressed in the intersegmental vessels and pronephric duct; (X) *zgc:171494* is also expressed in the floor plate, and pronephric duct; (Y) *ankdd1a* is also expressed in the spinal cord neurons; (Z) *sh3bp4* is also expressed in the posterior cardinal vein, and intersegmental vessels; (AB′) *grtp1b* is also expressed in the posterior cardinal vein; (AE′) *irf9*; (AG′) *ankrd58*; (AH′) ccdc135 is also expressed in the posterior cardinal vein, pronephric duct, posterior neural tube, and slightly in the hatching gland; (AI′) *tmem151a* is also expressed in the posterior cardinal vein, spinal cord neurons and intersegmental vessels (see Figure S1AI′). Basal ubiquitous expression with darker

staining of the dorsal aorta and caudal hematopoietic tail region is noted in: (O) *capn8* which includes expression in the forebrain, hatching gland, and lateral line primordium; (P) *mhc1uea*; (T) *rasa4* is also expressed in the yolk extension; (W) *sept5b* is also expressed in the forebrain and somites. Embryos are positioned laterally with the anterior facing left. Abbreviations: cht, caudal hematopoietic tail region; cn, caudal notochord; cs, caudal somites; cv, cranial vasculature; da, dorsal aorta; e, epiphysis; ff, fin fold; fp, floor plate; h, heart; hg, hatching gland; isv, intersegmental vessels; llp, lateral line primordium; mhb, midbrain hindbrain boundary; o, otoliths; ob, olfactory bulb; os, optic stalk; ov, otic vesicle; pd, pronephric duct; pe, primitive erythroid cells; pm, primitive myeloid cells; py, posterior yolk extension; sb, somite boundaries; sc, spinal cord; scn, spinal cord neurons; ss, somites; tb, tailbud; ye, yolk extension. Scale bar: 250 μm.
doi:10.1371/journal.pone.0031658.g003

Contrary to mammalian *lysozyme G*, the *lysozyme G2* in flounder and zebrafish lack a signal peptide and a conserved cystein catalytic triad, suggesting an intracellular role with a potentially distinct function [31]. There are 2 other paralogs for this gene. *lyg1* (*zgc:92608*), is expressed in macrophages [32], while the expression of *zgc:162941* has not been reported.

B. *cdc42GAP1/ARHGAP1*

Originally cloned in mice, *cdc42GAP1* catalyzes the hydrolysis of GTP bound by the Rho GTPases cdc42 and rac1, inactivating them [33]. In mice and humans it is expressed ubiquitously but exhibits marked expression in heart and lungs [34]. Mice knockouts result in generally reduced organ size relative to wild-type siblings due to increased rates of basal apoptosis, and display reduced viability [35]. The few survivors experience premature aging associated with heightened genomic instability [36]. In vitro evidence suggests *cdc42GAP1* negatively modulates angiogenesis [37]. Although it is expressed ubiquitously in mammals, it is predominantly expressed in the developing vasculature and heart in zebrafish at the stage examined here (Figure 3B).

C. *spa17/zgc:153721*

The protein is homologous with the human *sperm protein antigen 17*, *spa17*, and both contain a cAMP dependent protein kinase regulatory subunit domain and an IQ motif that binds proteins with EF-hand motifs, but *zgc:153721/spa17* of zebrafish has a glutamine rich region between these motifs that is absent in *spa17* in humans and mice. Although in mice it is expressed ubiquitously and is expressed in embryos, in rabbits, it is most highly expressed in the tissue from which it was originally identified, testis [38,39]. Increased expression of *spa17* has been detected in esophageal, ovarian, and cervical cancers and has been used as a clinical marker for cancer in these tissues [40,41,42]. In zebrafish, *spa17* is also expressed in the pronephros (Figure 3C).

D. *klhl4*

kelch like homolog like 4, *klhl4* is a member of the kelch motif containing family of proteins [43]. The founding member, *kelch*, was cloned from Drosophila, and mediates cytoplasmic streaming through ring canals in the developing oocyte [44]. In human fetal tissue, *klhl4* shows ubiquitous expression, and the protein is structurally similar to the founding member, *kelch1* [45]. As in humans, the *klhl4* of zebrafish has a BTB/POZ like domain and a BACK domain at the N-terminus, but while human *klhl4* has 6 kelch domains at the carboxy terminus, zebrafish has 3. The function of this protein has not been characterized.

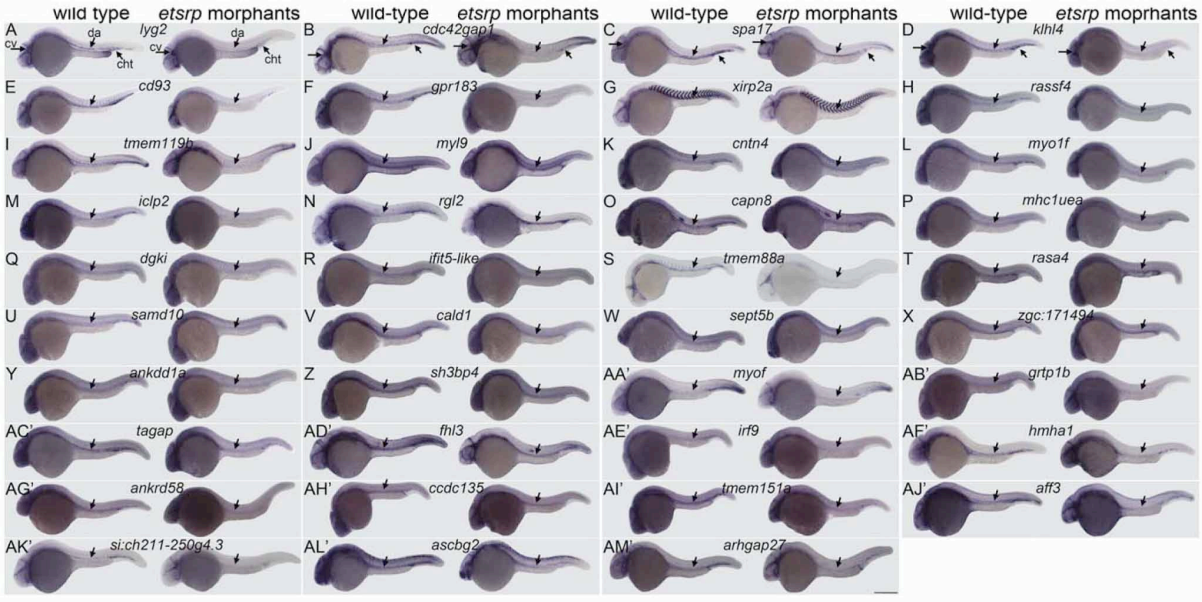


Figure 4. Analysis of gene expression in *etsrp* morphants. *Flk1-gfp* embryos were injected at the one cell stage with a mixture of translation blocking morpholinos for *etsrp* and analyzed by WISH at the 24hpf stage. Marked reduction is apparent in all genes examined in the axial vasculature, which includes the dorsal aorta and posterior cardinal vein, and is marked with a downward facing arrow in all images. The primitive myeloid cells stained in (L) *myo1f* wild-type controls are absent in their *etsrp* morphant counterparts. The staining in the axial trunk region of *rgl2* in *etsrp* morphants marks primitive erythrocytes that are trapped due to lack of circulation. The expression of non-vascular structures is not affected otherwise in *etsrp* morphants. Embryos are positioned laterally with the anterior facing left. Abbreviations: cht, caudal hematopoietic tail region; cv, cranial vasculature; da, dorsal aorta. Scale bar: 250 μm.
doi:10.1371/journal.pone.0031658.g004

E. *cd93*

Other aliases of *cd93* include *complement component C1q receptor like* (*C1qrl*), *lymphocyte antigen 68* (*ly68*), and *AA4.1*. *cd93* is a widely conserved member of the c type lectin family whose expression in endothelial and other cell types has been documented in mice and humans [46,47]. In previous reports the name *complement receptor C1qR-like* (*crl*) has been used for a different gene in zebrafish, which is the likely homolog of *c-type lectin domain family 14, member a* (*clec14a*) [3,18,19]. Although both of these genes are members of the c-type lectin family, they are paralogs of each other that deserve distinction, and we note that both are expressed in the vasculature, since *clec14a* [3] and *cd93* shown here, are vascular specific at the stage examined (Figure 3E).

F. *gpr183**

Epstein bar induced 2, *ebi2*, or *g protein coupled receptor 183*, *gpr183*, is an orphan receptor that was originally identified as a gene induced by Epstein Bar Virus infection in Buritt lymphoma cells [48]. Knockout mice are viable without any notable gross phenotype, however they experience defects in B-cell mobility within lymphoid tissues, and consequently have difficulty mounting rapid antibody responses [49,50]. Northern blot evaluation of its expression was originally limited to peripheral blood mononuclear cells (PBMCs), lymphocytic tissues and lungs [48]. However, in separate assays it has been detected in the aorta, atria, gastrointestinal tract, but not in the hematopoietic bone marrow or fetal liver [51]. At the stage examined, *gpr183* is clearly expressed in the developing vasculature of zebrafish embryos (Figure 3F), but not in the heart.

G. *xirp2a*

xin actin binding repeat containing 2a, *xirp2a*, is one of three orthologs for the mammalian counterpart, *xirp1β* [52]. The paralogs in zebrafish are *xirp2b*, which is restricted to the myotome borders during the larval stages of development (ZFIN), and *xirp2c*, which has not been characterized. Functionally, xin repeat motifs cross-link f-actin with β-catenin at adherens junctions [53]. *Xirp2a* expression in the myoseptum, pronephric duct, heart, otic vesicle, head mesoderm, and liver has been demonstrated in developing zebrafish larvae [32]. In this study the expression in cranial and axial vasculature in the trunk becomes more evident. The most likely explanation for the differences observed might be a longer in situ color developing time used here.

H. *rassf4*

ras association (RalGDS/AF-6) domain family member 4, *rassf4*, is a member of Ras association domain family of proteins, which contain a Ras association domain and in the case of *rassf4*, a SARAH domain (reviewed in [54]). *rassf4* is expressed ubiquitously in humans, but is downregulated in tumorous cells due to the hypermethylation of its promoter [55]. The idea that *rassf4* is a tumor suppressor is supported by the apoptotic death of tumor cells following forced expression of *rassf4* [55]. Besides expression in the vasculature of 1 dpf zebrafish embryos, *rassf4* is also highly expressed in spinal cord neurons (Figure 3H, and Figure S1H).

I. *tmem119b*

Neither the protein encoded by *tmem119b* nor its paralog, *tmem119a*, have been characterized. *tmem119b* encodes a single pass type I transmembrane protein while *tmem119a* encodes a multipass transmembrane protein with two transmembrane domains. As a conserved gene in vertebrates, it is currently unclear whether *tmem119b* or *tmem119a* is the ortholog of their counterpart in mice, *osteoblast induction factor*, *obif*, which promotes osteoblast differentiation [56]. Although *obif* and *tmem119b* share a structural topological profile, *tmem119a* has a negligible higher amount of amino acid conservation with *obif*. It is also currently unknown whether *tmem119/obif* plays the same roles in humans.

J. *myl9*

myosin light chain 9 regulatory, *myl9*, is a calcium-regulated protein that regulates the contraction of myosin heavy chains in non-skeletal muscle cells that was originally cloned from human umbilical artery [57]. Currently there is evidence that the expression of *myl9* may be directly regulated by *junb* [58], and *runx1* [59]. In zebrafish there is a close paralog, *myosin light polypeptide 9-like* (*zgc: 77916*), with 95% gene identity to *myl9* that also remains to be characterized.

Discussion

In this report we provide further support for the utility of combining the relatively unbiased approach of high throughput next generation RNA-sequencing to analyze whole embryo transcriptomes in zebrafish. The high correlation between genes induced by *etsrp* as predicted by RNA-seq and their confirmation by WISH reinforces the effectiveness of the methods used to interrogate transcriptional profiles in zebrafish. Of the genes selected from the dataset obtained, we identified a very high percentage of genes expressed in the vasculature of embryos, validating this approach to identify biologically relevant genes. However there may be non-vascular genes in this data that are induced due to the induction of off-target genes by the artificially elevated level of *etsrp*. This issue can be addressed by WISH using wild type embryos at a proper embryonic stage. Additionally, this method does not discriminate direct and indirect targets induced by *etsrp*. Further analysis by ChIP-Seq should reveal those genes that are transcriptionally regulated by *etsrp*.

Both the ectopic induction of these genes by *etsrp* overexpression and down-regulation by *etsrp* deficiency in vascular endothelial and primitive myeloid cells highlights their genetic epistatic relationship to *etsrp* and show that *etsrp* is a potent transcriptional activator whose expression and activity must be tightly regulated. Although several genes exhibit basal levels of ubiquitous expression their level in the vasculature is still higher. These genes deserve as much attention as those only specific to the vascular lineage, as several studies have shown that the down-regulation of such genes resulted in the disruption of vascular development. For instance, knockdown of *nrarp-a* and *nrarp-b* [60,61], *lpa$_1$* [62], and *dep1a/dep1b* [63] respectively, resulted in defects in intersegmental vessels, lymphangiogenesis, and arterial specification. We obtained approximately 850 genes that were induced 1.9 fold or more by *etsrp* expression. Based on previous annotation and studies we can classify 6.2% of them as vascular genes. Experimental studies identified 39 more vascular genes and further increased the percentage to 10.8%. This leaves over 700 unigene entries in the dataset, some of which may be expressed and actively involved in the development of the zebrafish vasculature, and are potentially worth examining. The complete dataset is available in Table S5 and we hope this will benefit research on the development of the vascular system. The panel of genes found here remain to be characterized further and will require subsequent functional studies.

Materials and Methods

Embryo injection

Flk1-gfp embryos were injected with 75 pg *etsrp* RNA at the one cell stage as described [5], and collected at the late gastrulation

stages, 70–90% epiboly, when ectopic *flk1-gfp* expression is notable in injected embryos. Total RNA was extracted with the PerfectPure RNA Tissue kit (5 Prime). For pre-sequencing validation, cDNA was made with Oligo(dT) primers and Superscript II reverse transcriptase (Invitrogen). The following primers pairs were used for verification by RT-PCR: *β-actin*: 5′ -TGTTTTCCCCT-CCATTGTTG-3′, 5′-ACATACATGGCAGGGGTGTT-3′; *etsrp 3′UTR*: 5′-GAGGAATTCTCGAAGGATTGG-3′;TGGTTTT-CTAAAGGCACCTAG-3′; *Fli1a*: 5′-CCGAGGTCCTGCTCT-CACAT-3′;5′-GGGACTGGTCAGCGTGAGAT-3′; *scl*: 5′-GG-AGATGCGGAACAGTATGG-3′, 5′-GAAGGCACCGTTCA-CATTCT-3′. To test gene expression in *etsrp* morphants, *flk1-gfp* embryos were injected with a mixture of 4 ng each of translation blocking *etsrp* morpholinos as described in [5] at the one cell stage, and embryos were harvested at 24hpf for WISH.

Cloning and Whole Mount In Situ Hybridization (WISH)

Different segments of the genes selected for evaluation were cloned into the pCRII Topo vector with the TOPO TA cloning kit (dual promoter) according to the manufacturer's instructions (Invitrogen). Amplicons were obtained from an *etsrp* overexpression cDNA library at late stages of gastrulation (70%–90% epiboly). Primers used for amplification are noted in Table S5. Positive plasmids were confirmed by sequencing. WISH was performed as described [64]. DIG labeled RNA probes were generated by linearizing TOPO-cloned genes with restriction endonucleases (New England Biolabs), and transcribing with SP6 or T7 RNA polymerase (Promega).

Image acquisition and processing

In situ stained embryos were further processed by a serial dehydration in ethanol, followed by rehydration into 1XPBS. Embryos were imaged in 2% methylcellulose in depression microscope slides. Images were captured with a color digital CCD camera (Axiocam, Zeiss) mounted on an upright microscope (Axioskop2 plus, Zeiss) with Openlab 4.0.2 software (Improvision, Lexington, MA). Serial images were combined, merged and processed with Adobe Photoshop CS5.

Deep sequencing library construction and bioinformatics

mRNA-seq libraries were constructed with the Tru-Seq RNA prep kit (Illumina) according to manufacturer's instructions. After this project was initiated, the Illumina sequencer has gone through a series of improvements. As a result, we obtained both single-end and paired-end reads with different lengths (76, 51, and 46 bp) using the Illumina GA II sequencer and its earlier versions. The obtained reads were mapped to 51,481 Unigene transcript sequences (Build #117, October 2009) [26] using BWA [27] allowing up to 4 mismatches. In case of the paired-end reads, we determined whether the mapped results support correct pairing of the reads according to the unigene annotation. If one read in a pair is mapped to a gene, the other one should be mapped to the same gene. Since each read in a pair may map to multiple locations in the genome, all possible combinations of their mappings were examined for correct pairing. The pair of reads is considered uniquely mapped only if one unique pair of mapped locations was identified.

To determine expression levels of genes and exons, we used the variable RPKM (reads per kilobase of exon per million mapped reads) defined by Mortazavi et al. 2008 [65]. Analysis of differential gene expression was carried out to obtain the differentially expressed genes between control sample and *estrp* oe sample. The number of uniquely mapped read-pairs for each gene in each sample was stored. The total number of mapped reads in each lane was normalized using the total mapped reads in each lane. Fisher's exact test was then performed using the above read counts for each gene. The resulting p-values were corrected via the Benjamini and Hochberg method as implemented in R. Finally, differentially expressed genes were defined as those with changes of at least 1.9-fold between a pair of samples at a false discovery rate (FDR) of 5%. Table S7 lists all 849 upregulated unigene hits in the *estrp* oe sample, while the 726 downregulated hits are tabulated in Table S8.

Quantitative PCR (qPCR)

qPCR was performed and analyzed exactly as stated in [12], with three independent biological replicates of either uninjected or *etsrp* oe groups of embryos harvested at late gastrulation stages. Primers used are listed in Table S6.

Supporting Information

Figure S1 Higher magnification of Figure 4. The axial trunk vasculature of embryos displayed in figure 4 were imaged at higher magnification to highlight the changes observed. Wild-type embryos are on the left half of each column with their *etsrp* morphant counterparts on the right for each gene. Embryos were positioned with anterior facing left.
(TIF)

Table S1 Total RNA-Seq reads obtained in control and *estrp* oe samples and mapping results.
(XLS)

Table S2 DAVID GO derived gene subset without clear vascular expression or not expressed in developing vasculature.
(XLS)

Table S3 Quantitative PCR of unigene hits predicted to be suppressed by *etsrp* oe via RNA-seq.
(XLS)

Table S4 Genes with double entries in Ensembl.
(XLS)

Table S5 Primers used for cloning to make WISH probes.
(XLS)

Table S6 Primers used to test downregulated gene set following *etsrp* overexpression by qPCR.
(XLS)

Table S7 Complete *etsrp* mRNA-seq dataset (induced greater than or equal to 1.9 fold by *etsrp* overexpression).
(XLS)

Table S8 Downregulated mRNA-seq dataset (reduced greater than or equal to 1.9 fold by overexpression of *etsrp*).
(XLS)

Text S1 Supplementary Text. A continuation of the results section that describes information regarding the genes examined in this study.
(DOC)

Acknowledgments

We thank Scott Doherty for excellent technical support and Anqi Liu for help with animal husbandry.

Author Contributions

Conceived and designed the experiments: GAG SL. Performed the experiments: GAG MBV. Analyzed the data: GAG J-HL MBV JL. Contributed reagents/materials/analysis tools: XX. Wrote the paper: GAG.

References

1. Stainier DY, Fishman MC (1994) The zebrafish as a model system to study cardiovascular development. Trends Cardiovasc Med 4: 207–212.
2. Weinstein BM (2002) Plumbing the mysteries of vascular development using the zebrafish. Semin Cell Dev Biol 13: 515–522.
3. Sumanas S, Jorniak T, Lin S (2005) Identification of novel vascular endothelial-specific genes by the microarray analysis of the zebrafish cloche mutants. Blood 106: 534–541.
4. Pham VN, Lawson ND, Mugford JW, Dye L, Castranova D, et al. (2007) Combinatorial function of ETS transcription factors in the developing vasculature. Dev Biol 303: 772–783.
5. Sumanas S, Lin S (2006) Ets1-related protein is a key regulator of vasculogenesis in zebrafish. PLoS Biol 4: e10.
6. Sumanas S, Gomez G, Zhao Y, Park C, Choi K, et al. (2008) Interplay among Etsrp/ER71, Scl, and Alk8 signaling controls endothelial and myeloid cell formation. Blood 111: 4500–4510.
7. Lee D, Park C, Lee H, Lugus JJ, Kim SH, et al. (2008) ER71 acts downstream of BMP, Notch, and Wnt signaling in blood and vessel progenitor specification. Cell Stem Cell 2: 497–507.
8. Kataoka H, Hayashi M, Nakagawa R, Tanaka Y, Izumi N, et al. (2011) Etv2/ER71 induces vascular mesoderm from Flk1+PDGFR{alpha}+ primitive mesoderm. Blood.
9. Neuhaus H, Müller F, Hollemann T (2010) Xenopus er71 is involved in vascular development. Dev Dyn 239: 3436–3445.
10. Salanga MC, Meadows SM, Myers CT, Krieg PA (2010) ETS family protein ETV2 is required for initiation of the endothelial lineage but not the hematopoietic lineage in the Xenopus embryo. Dev Dyn 239: 1178–1187.
11. Proulx K, Lu A, Sumanas S (2010) Cranial vasculature in zebrafish forms by angioblast cluster-derived angiogenesis. Dev Biol 348: 34–46.
12. Ren X, Gomez GA, Zhang B, Lin S (2010) Scl isoforms act downstream of etsrp to specify angioblasts and definitive hematopoietic stem cells. Blood 115: 5338–5346.
13. Liu F, Patient R (2008) Genome-wide analysis of the zebrafish ETS family identifies three genes required for hemangioblast differentiation or angiogenesis. Circulation Research 103: 1147–1154.
14. Mathavan S, Lee SG, Mak A, Miller LD, Murthy KR, et al. (2005) Transcriptome analysis of zebrafish embryogenesis using microarrays. PLoS Genet 1: 260–276.
15. Covassin L, Amigo JD, Suzuki K, Teplyuk V, Straubhaar J, et al. (2006) Global analysis of hematopoietic and vascular endothelial gene expression by tissue specific microarray profiling in zebrafish. Dev Biol 299: 551–562.
16. Qian F, Zhen F, Ong C, Jin S-W, Meng Soo H, et al. (2005) Microarray analysis of zebrafish cloche mutant using amplified cDNA and identification of potential downstream target genes. Dev Dyn 233: 1163–1172.
17. Weber GJ, Choe SE, Dooley KA, Paffett-Lugassy NN, Zhou Y, et al. (2005) Mutant-specific gene programs in the zebrafish. Blood 106: 521–530.
18. Gomez GA, Veldman MB, Zhao Y, Burgess S, Lin S (2009) Discovery and characterization of novel vascular and hematopoietic genes downstream of etsrp in zebrafish. PLoS ONE 4: e4994.
19. Wong KS, Proulx K, Rost MS, Sumanas S (2009) Identification of vasculature-specific genes by microarray analysis of Etsrp/Etv2 overexpressing zebrafish embryos. Dev Dyn 238: 1836–1850.
20. Hegedus Z, Zakrzewska A, Agoston VC, Ordas A, Rácz P, et al. (2009) Deep sequencing of the zebrafish transcriptome response to mycobacterium infection. Mol Immunol 46: 2918–2930.
21. Ordas A, Hegedus Z, Henkel CV, Stockhammer OW, Butler D, et al. (2010) Deep sequencing of the innate immune transcriptomic response of zebrafish embryos to Salmonella infection. Fish Shellfish Immunol.
22. Stockhammer OW, Rauwerda H, Wittink FR, Breit TM, Meijer AH, et al. (2010) Transcriptome analysis of Traf6 function in the innate immune response of zebrafish embryos. Mol Immunol 48: 179–190.
23. Aanes H, Winata CL, Lin CH, Chen JP, Srinivasan KG, et al. (2011) Zebrafish mRNA sequencing deciphers novelties in transcriptome dynamics during maternal to zygotic transition. Genome research.
24. Soares AR, Pereira PM, Santos B, Egas C, Gomes AC, et al. (2009) Parallel DNA pyrosequencing unveils new zebrafish microRNAs. BMC Genomics 10: 195.
25. Aday AW, Zhu LJ, Lakshmanan A, Wang J, Lawson ND (2011) Identification of cis regulatory features in the embryonic zebrafish genome through large-scale profiling of H3K4me1 and H3K4me3 binding sites. Dev Biol.
26. Wheeler DL, Church DM, Federhen S, Lash AE, Madden TL, et al. (2003) Database resources of the National Center for Biotechnology. Nucleic Acids Res 31: 28–33.
27. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760.
28. Zakrzewska A, Cui C, Stockhammer OW, Benard EL, Spaink HP, et al. (2010) Macrophage-specific gene functions in Spi1-directed innate immunity. Blood 116: e1–11.
29. Hikima J, Minagawa S, Hirono I, Aoki T (2001) Molecular cloning, expression and evolution of the Japanese flounder goose-type lysozyme gene, and the lytic activity of its recombinant protein. Biochim Biophys Acta 1520: 35–44.
30. Nakano T, Graf T (1991) Goose-type lysozyme gene of the chicken: sequence, genomic organization and expression reveals major differences to chicken-type lysozyme gene. Biochim Biophys Acta 1090: 273–276.
31. Irwin DM, Gong Z (2003) Molecular evolution of vertebrate goose-type lysozyme genes. J Mol Evol 56: 234–242.
32. Thisse B, Heyer V, Lux A, Alunni V, Degrave A, et al. (2004) Spatial and temporal expression of the zebrafish genome by large-scale in situ hybridization screening. Methods Cell Biol 77: 505–519.
33. Lamarche-Vane N, Hall A (1998) CdGAP, a novel proline-rich GTPase-activating protein for Cdc42 and Rac. J Biol Chem 273: 29172–29177.
34. Tcherkezian J, Triki I, Stenne R, Danek EI, Lamarche-Vane N (2006) The human orthologue of CdGAP is a phosphoprotein and a GTPase-activating protein for Cdc42 and Rac1 but not RhoA. Biol Cell 98: 445–456.
35. Wang L, Yang L, Burns K, Kuan C-Y, Zheng Y (2005) Cdc42GAP regulates c-Jun N-terminal kinase (JNK)-mediated apoptosis and cell number during mammalian perinatal growth. Proc Natl Acad Sci USA 102: 13484–13489.
36. Wang L, Yang L, Debidda M, Witte D, Zheng Y (2007) Cdc42 GTPase-activating protein deficiency promotes genomic instability and premature aging-like phenotypes. Proc Natl Acad Sci USA 104: 1248–1253.
37. Engelse MA, Laurens N, Verloop RE, Koolwijk P, van Hinsbergh VWM (2008) Differential gene expression analysis of tubule forming and non-tubule forming endothelial cells: CDC42GAP as a counter-regulator in tubule formation. Angiogenesis 11: 153–167.
38. Richardson RT, Yamasaki N, O'Rand MG (1994) Sequence of a rabbit sperm zona pellucida binding protein and localization during the acrosome reaction. Dev Biol 165: 688–701.
39. Wen Y, Richardson RT, Widgren EE, O'Rand MG (2001) Characterization of Sp17: a ubiquitous three domain protein that binds heparin. Biochem J 357: 25–31.
40. Gupta G, Sharma R, Chattopadhyay TK, Gupta SD, Ralhan R (2007) Clinical significance of sperm protein 17 expression and immunogenicity in esophageal cancer. Int J Cancer 120: 1739–1747.
41. Li FQ, Liu Q, Han YL, Wu B, Yin HL (2010) Sperm protein 17 is highly expressed in endometrial and cervical cancers. BMC Cancer 10: 429.
42. Straughn JM, Shaw DR, Guerrero A, Bhoola SM, Racelis A, et al. (2004) Expression of sperm protein 17 (Sp17) in ovarian cancer. Int J Cancer 108: 805–811.
43. Adams J, Kelso R, Cooley L (2000) The kelch repeat superfamily of proteins: propellers of cell function. Trends Cell Biol 10: 17–24.
44. Xue F, Cooley L (1993) kelch encodes a component of intercellular bridges in Drosophila egg chambers. Cell 72: 681–693.
45. Braybrook C, Warry G, Howell G, Arnason A, Bjornsson A, et al. (2001) Identification and characterization of KLHL4, a novel human homologue of the Drosophila Kelch gene that maps within the X-linked cleft palate and Ankyloglossia (CPX) critical region. Genomics 72: 128–136.
46. Fonseca MI, Carpenter PM, Park M, Palmarini G, Nelson EL, et al. (2001) C1qR(P), a myeloid cell receptor in blood, is predominantly expressed on endothelial cells in human tissue. J Leukoc Biol 70: 793–800.
47. Petrenko O, Beavis A, Klaine M, Kittappa R, Godin I, et al. (1999) The molecular characterization of the fetal stem cell marker AA4. Immunity 10: 691–700.
48. Birkenbach M, Josefsen K, Yalamanchili R, Lenoir G, Kieff E (1993) Epstein-Barr virus-induced genes: first lymphocyte-specific G protein-coupled peptide receptors. J Virol 67: 2209–2220.
49. Gatto D, Paus D, Basten A, Mackay CR, Brink R (2009) Guidance of B cells by the orphan G protein-coupled receptor EBI2 shapes humoral immune responses. Immunity 31: 259–269.
50. Pereira JP, Kelly LM, Xu Y, Cyster JG (2009) EBI2 mediates B cell segregation between the outer and centre follicle. Nature 460: 1122–1126.
51. Rosenkilde MM, Benned-Jensen T, Andersen H, Holst PJ, Kledal TN, et al. (2006) Molecular pharmacological phenotyping of EBI2. An orphan seven-transmembrane receptor with constitutive activity. J Biol Chem 281: 13199–13208.
52. Grosskurth SE, Bhattacharya D, Wang Q, Lin JJ-C (2008) Emergence of Xin demarcates a key innovation in heart evolution. PLoS ONE 3: e2857.
53. Choi S, Gustafson-Wagner EA, Wang Q, Harlan SM, Sinn HW, et al. (2007) The intercalated disk protein, mXinalpha, is capable of interacting with beta-catenin and bundling actin filaments [corrected]. J Biol Chem 282: 36024–36036.
54. Avruch J, Xavier R, Bardeesy N, Zhang X-F, Praskova M, et al. (2009) Rassf family of tumor suppressor polypeptides. J Biol Chem 284: 11001–11005.
55. Eckfeld K, Hesson L, Vos MD, Bieche I, Latif F, et al. (2004) RASSF4/AD037 is a potential ras effector/tumor suppressor of the RASSF family. Cancer Res 64: 8688–8693.

56. Kanamoto T, Mizuhashi K, Terada K, Minami T, Yoshikawa H, et al. (2009) Isolation and characterization of a novel plasma membrane protein, osteoblast induction factor (obif), associated with osteoblast differentiation. BMC Dev Biol 9: 70.
57. Kumar CC, Mohan SR, Zavodny PJ, Narula SK, Leibowitz PJ (1989) Characterization and differential expression of human vascular smooth muscle myosin light chain 2 isoform in nonmuscle cells. Biochemistry 28: 4027–4035.
58. Licht AH, Nübel T, Feldner A, Jurisch-Yaksi N, Marcello M, et al. (2010) Junb regulates arterial contraction capacity, cellular contractility, and motility via its target Myl9 in mice. J Clin Invest 120: 2307–2318.
59. Jalagadugula G, Mao G, Kaur G, Goldfinger LE, Dhanasekaran DN, et al. (2010) Regulation of platelet myosin light chain (MYL9) by RUNX1: implications for thrombocytopenia and platelet dysfunction in RUNX1 haplodeficiency. Blood 116: 6037–6045.
60. Topczewska JM, Topczewski J, Szostak A, Solnica-Krezel L, Hogan BLM (2003) Developmentally regulated expression of two members of the Nrarp family in zebrafish. Gene Expr Patterns 3: 169–171.
61. Phng LK, Potente M, Leslie JD, Babbage J, Nyqvist D, et al. (2009) Nrarp coordinates endothelial Notch and Wnt signaling to control vessel density in angiogenesis. Dev Cell 16: 70–82.
62. Lee SJ, Chan TH, Chen TC, Liao BK, Hwang PP, et al. (2008) LPA1 is essential for lymphatic vessel development in zebrafish. FASEB J 22: 3706–3715.
63. Rodriguez F, Vacaru A, Overvoorde J, den Hertog J (2008) The receptor protein-tyrosine phosphatase, Dep1, acts in arterial/venous cell fate decisions in zebrafish development. Dev Biol 324: 122–130.
64. Thisse C, Thisse B (2008) High-resolution in situ hybridization to whole-mount zebrafish embryos. Nat Protoc 3: 59–69.
65. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5: 621–628.

Chapter 2 Supplementary Material

Table S1. RNA-seq reads obtained in control and *etsrp* oe samples and mapping results

control		estrp OE		Total
Paired-end	Single-end	Paired-end	Single-end	
38,595,718	32,937,423	20,418,736	31,710,235	123,663,112
7,941,282 (21%)	10,288,710 (32%)	10,927,174 (54%)	13,504,539 (43%)	42,761,705 (35%)

Table S2. DAVID GO derived gene subset without clear vascular expression or not expressed in developing vasculature

Unigene ID	Gene ID	Gene Description	Fold Change	p-value
Dr.83594	*crabp1a*	*cellular retinoic acid binding protein 1a*	917.4	5.10E-17
Dr.37870	*ckmb*	*creatine kinase, muscle b*	231.6	3.80E-07
Dr.81739	*kcnh2*	*potassium voltage-gated channel, subfamily h (eag-related), member 2*	101.5	2.00E-09
Dr.78041	***xirp2a****	***xin actin-binding repeat containing 2a***	**95**	**7.60E-56**
Dr.76040	*nexn*	*nexilin (f actin binding protein)*	33	4.20E-02
Dr.266	***mef2a***	***myocyte enhancer factor 2a***	**9.8**	**7.20E-04**
Dr.31448	*mdka*	*midkine-related growth factor*	7.9	1.00E-149
Dr.86752	***nr6a1b***	***nuclear receptor subfamily 6, group a, member 1b***	**7.1**	**9.30E-04**
Dr.10326	***junb***	***jun b proto-oncogene***	**6.1**	**5.70E-03**
Dr.117443	***bag2***	***bcl2-associated athanogene 2***	**6.1**	**7.10E-04**
Dr.33995	*obscn*	*obscurin, cytoskeletal calmodulin and titin-interacting rhogef*	5.8	9.00E-03
Dr.11428	***svil***	***supervillin***	**4**	**1.10E-25**
Dr.78258	*LOC561356*	*similar to loc494988 protein*	3.8	2.10E-02
Dr.91799	***wt1b***	***wilms tumor 1b***	**3.5**	**2.00E-02**
Dr.84618	***clic2***	***chloride intracellular channel 2***	**3.1**	**1.30E-04**
Dr.81606	***ehbp1l1a***	***eh domain binding protein 1-like 1a***	**3**	**5.00E-03**
Dr.77543	*pklr*	*pyruvate kinase, liver and rbc*	2.9	2.30E-02
Dr.20912	***alcam***	***activated leukocyte cell adhesion molecule***	**2.8**	**1.90E-30**
Dr.3182	***mef2ca***	***myocyte enhancer factor 2ca***	**2.8**	**1.10E-03**
Dr.89477	***tgfb2***	***transforming growth factor, beta 2***	**2.8**	**2.10E-02**
Dr.83682	*sox10*	*sry-box containing gene 10*	2.8	3.10E-02
Dr.14787	*s100b*	*s100 calcium binding protein, beta (neural)*	2.7	4.20E-02
Dr.77184	*slc2a2*	*solute carrier family 2 (facilitated glucose transporter), member 2*	2.5	1.40E-03
Dr.100995	***scn12aa***	***sodium channel, voltage gated, type xii, alpha a***	**2.4**	**9.00E-07**
Dr.132574	***bin1***	***bridging integrator 1***	**2.4**	**2.80E-03**
Dr.132281	***hsp90a.2***	***heat shock protein 90-alpha 2***	**2.3**	**5.30E-98**
Dr.78802	***strbp***	***spermatid perinuclear rna binding protein***	**2**	**5.70E-08**
Dr.81423	***hand2***	***heart and neural crest derivatives expressed transcript 2***	**2**	**4.20E-06**
Dr.78147	***got2b***	***glutamic-oxaloacetic transaminase 2b, mitochondrial (aspartate aminotransferase 2)***	**1.9**	**2.30E-13**
Dr.75914	*ar*	*androgen receptor*	1.9	4.40E-02

Genes in bold may be vascular because they are either expressed in the mesoderm or heart of embryos, but it is currently not clear whether the expression in the heart is specific to the endocardium or myocardium. **xirp2a* was examined here and found to be expressed in the vasculature.

Table S3: quantitative PCR of unigene hits predicted to be suppressed by *etsrp* oe via RNA-seq.

Unigene ID	Gene ID	Gene Description	Average Fold-Change	Standard Deviation	Fold-Change by RNA-seq
Dr.133216	***sox32****	***SRY-box containing gene 32***	**0.38**	**0.24**	**0.51**
Dr.81305	***bon****	***similar to bonnie and clyde***	**0.48**	**0.08**	**0.51**
Dr.36017	*myoD*	*myogenic differentiation 1*	1.11	0.41	0.46
Dr.78792	*nkx2.1b*	*nk2 homeobox 1b*	1.38	0.74	0.43
Dr.117206	*ctsl1b*	*similar to cathepsin L, 1 b*	0.82	0.16	0.39
Dr.75761	*rx1*	*retinal homeobox gene 1*	0.61	0.37	0.24
Dr.152568	*fez2d*	*FEZ family zinc finger 2d*	0.78	0.44	0.09
Dr.81131	*dpf3*	*D4, zinc and double PHD fingers, family 3*	1.22	0.26	0.01
Dr.78408	***fli1a****	***friend leukemia integration 1a***	**53.08**	**30.92**	**37.6**

* Statistically significant, determined by student t-test $p<0.05$. *fli1a* was included as a control.

Table S4: Genes with double entries in Ensembl.

Gene probed for here				Other paralog(s) with same gene name in Ensembl Zv9		
Unigene ID	Gene	Ensembl Gene ID	Chromosome #: Physical Location	Ensembl Gene ID	Gene Name	Chromosome #: Physical Location
Dr.111220	***gpr183***	**ENSDARG00000010317**	**9: 32,138,697-32,141,593**	ENSDARG00000074302	*gpr183 (2 of 3)*	1:29539773-29543235:1
				ENSDARG00000069756	*gpr183 (3 of 3)*	1:29610140-29611225:1
Dr.85673	***tagap***	**ENSDARG00000043475**	**20: 13,563,255-13,637,154**	ENSDARG00000002353	*tagap (1 of 2)*	17:45572274-45578169:1
Dr.115399	***cntn4***	**ENSDARG00000019405**	**6: 45,430,455-45,778,132**	ENSDARG00000062880	*cntn4 (2 of 2)*	23:11775565-11967621:1
Dr.117215	***ifit5-like***	**ENSDARG00000090977**	**12: 18,054,322-18,057,848**	ENSDARG00000008098	*ifit5*	12:18100491-18107181:-1
Dr.96217	***samd10***	**ENSDARG00000094377**	**23: 13,236,075-13,444,853**	ENSDARG00000079978	*samd10 (2 of 2)*	8:23830247-23860192:1
Dr.119058	***zgc:171494/LASS3***	**ENSDARG00000036337**	**18: 6,772,525-6,789,193**	ENSDARG00000078541	*lass3 (1 of 2)*	7:10744380-10783500:1
Dr.80073	***fhl3***	**ENSDARG00000034643**	**16: 35,969,551-36,050,492**	ENSDARG00000059158	*fhl3 (2 of 2)*	19:4885107-4913988:-1
Dr.124255	*si:ch211-210g13.5*	ENSDARG00000062575	3: 16,990,399-17,166,893	ENSDARG00000075247	*plcl2*	16:52371158-52506557:1
Dr.114623	***cald1***	**ENSDARG00000070314**	**4: 6,523,006-6,573,658**	ENSDARG00000086391	*cald1 (2 of 2)*	25:20651846-20707283:1
Dr.82145	***myoferlin***	**ENSDARG00000017128**	**13: 29,462,452-29,512,846**	ENSDARG00000006112	*myof (2 of 2)*	12:11596559-11646274:-1
Dr.150612	***hmha1***	**ENSDARG00000062049**	**22: 17,835,923-17,876,627**	ENSDARG00000052496	*hmha1 (2 of 2)*	2:57410229-57483531:-1
Dr.90997	*tmem179*	ENSDARG00000013292	20: 27,423,260-27,426,177	ENSDARG00000090064	*tmem179 (2 of 2)*	17:46670381-46676560:1
Dr.83578	***tmem151a***	**ENSDARG00000079908**	**21: 25,023,938-25,025,734**	ENSDARG00000077986	*tmem151a (1 of 2)*	7:21040162-21050530:1

Genes in bold are expressed in developing zebrafish vasculature at time examined (24 hpf).

Table S5: Primers used to clone selected genes and make WISH probes

Unigene ID	Gene Symbol	Gene Name (Description)	Forward primer	Reverse primer	Product Size (bp
Dr.153782	***lyg2***	***lysozyme G like 2***	**TATTTGCACTCAGTCTAGTGACAACA**	**TGGCATAATCACGGCCTG**	**576**
Dr.149118	***cdc42GAP1***	***similar to Cdc42 GTPase-activating protein***	**TGATAGCCACTGTGGAGAATTTC**	**ATAGACCTCCTTTGTGGCTGATAG**	**2251**
Dr.118570	***she***	***similar to Src homology 2 domain containing e***	**GTTCTGACAGTCGGTACTGGTTC**	**CTTTTCCAGTGTGTAGCCACAC**	**1101**
Dr.40434	*ms4a17a.11*	*membrane-spanning 4-domains, subfamily A, member 17A.11*	GCTCTTGGACCCTTCTCACCTCGT	GCAGAACTGCCACTCTGCCCA	361
Dr.84343	***spa17***	***sperm auto antigen 17***	**CCCTTCTCCAACACACACCT**	**CGTCTGTCTCTACTCCGTCTGTT**	**1951**
Dr.103328	***klhl4***	***kelch like 4***	**GGAGGAGGGTGGAACAGGCAGA**	**AGCTCCAGGACTCCAGTATATGCAA**	**962**
Dr.110713	***LOC100149611***	***similar to LOC495463 protein; CD93***	**AGCTGCACGGCTTCAGGTGG**	**TCTGTGGGGTTCGGGACGCA**	**1012**
Dr.111220	***gpr183***	***g protein-coupled receptor 183***	**TGCTGGGAAATGTGCTGGCT**	**AGCCTTCGGGGGAGGTCCTG**	**885**
Dr.78041	***xirp2***	***xin actin-binding repeat containing 2***	**CACAGTCTGTAACGGAGCACTT**	**CATTGCTTGTTTCGCTGG**	**1051**
Dr.134501	***rassf4***	***ras association (RalGDS/AF-6) domain family member 4***	**GCGCCTTGCAAAACAGCTGGG**	**TGGACCATGCAGTAGGCGAGA**	**965**
Dr.91332	***tmem119b***	***transmembrane protein 119b***	**AGAACTCTAAAGTATTGAGTCACAGGA**	**ATTTAAAATAGCGATTCAAAACAGATT**	**1101**
Dr.89765	***myl9***	***myosin, light chain 9, regulatory***	**CTCTACCGTTTCCAGCACATT**	**ATGGCTTCATATATTTATCTTTGTGTT**	**701**
Dr.115399	***cntn4***	***contactin 4***	**TCTCTGTGGGCCTCCTCCGC**	**TGCGCCAGGAGATCGAGCCT**	**911**
Dr.84022	***myo1E***	***similar to myosin IE***	**TGGCGCTGCAGATTGTCGCT**	**GCTGCTCTGCTCCACGCCAA**	**1137**
Dr.75719	***iclp2***	***invariant chain-like protein 2***	**GGAGTGACTCTGCTGGCTGGGA**	**CGATCTCTGTTAGCTCTCCATCATCCA**	**879**
Dr.88587	***rgl2***	***ral guanine nucleotide dissociation stimulator-like 2***	**GAAGGAGACGGAGTTCTCCTG**	**TCAGCTCATATTTGGAGGCTTC**	**1951**
Dr.76656	***capn8***	***calpain 8***	**ACCGGCTGCCGACCAGAGAT**	**AGGCCGAGTTTGCCACTGCC**	**1315**
Dr.11010	***mhc1uea***	***major histocompatibility complex class I UEA gene***	**TTACTGGCGCACATCCGGTC**	**TCCCAGACACCTGTGATCCCCA**	**968**
Dr.77849	***dgki***	***similar to diacylglycerol kinase, iota***	**AGGGAGGCCCCAGAGATGCG**	**ACTCGGCCGTGGCAAACCTG**	**1160**
Dr.117215	***ifit5***	***similar to Interferon-induced protein with tetratricopeptide repeats 5 (IFIT-5)***	**TGCACGAGCGCAGTGCTCAA**	**TCGCGAGCCTCATCTACCTCCA**	**795**
Dr.134371	*plscr2*	*similar to phospholipid scramblase 2*	CATTTAGAGTAATAACTGCCGGATT	TTATATTAAACATTATGAGCACTGGCC	1331
Dr.7340	*slc16a9a*	*solute carrier family 16 (monocarboxylic acid transporters), member 9a*	GCATGTGCTGGGCCAATGCG	CCAGCAAGGCAGTGCAGCCA	882
Dr.89035	***tmem88a***	***transmembrane protein 88a***	**GCCTGAGGAGATACTACAGGTGG**	**ACAGGGGTTTATCCAGTAAAAAGC**	**1057**
Dr.100033	***rasa4***	***ras p21 protein activator 4-like***	**AGGATCGCAACGGGGCCTCT**	**GCAAACGCACCATCCACGGC**	**1128**
Dr.96217	***samd10***	***similar to sterile alpha motif domain containing 10***	**GACACCCGTGGGCATGGCAG**	**TGCCGGGAAGTCGAGAGAGC**	**628**
Dr.114623	***cald1***	***caldesmon 1***	**TGATTTCGAGCGCCGCAGGG**	**GCGAATGCCGTCCGTCACCA**	**1264**
Dr.91020	***sept5b***	***septin 5b***	**GGGAAAGCGGCCGGGATGAC**	**CGGTCCTGCTGCTTCGGCTC**	**716**
Dr.119058	***zgc:171494***	***zgc:171494***	**GCAGTGCCACTCGCTCAGGC**	**GCGGAGGATGAGGGAGGCCC**	**795**
Dr.84654	***ankdd1a***	***ankyrin repeat and death domain containing 1A***	**AGGCGCTGCGTTTGCTGCTA**	**TCACGCCGTGCAGCCAGATG**	**1152**
Dr.91634	***sh3bp4***	***novel protein similar to vertebrate SH3-domain binding protein 4***	**CTTGGCAAAGTGGCCCGCCT**	**GTCCTAGGTGCATGTGGGTCTTG**	**1244**
Dr.82145	***myof***	***similar to fer-1-like 3, myoferlin***	**GGTTCCGTCCTGTTCCTCA**	**CTGCTGGCAAATAATCAAATCC**	**1493**
Dr.84960	***grtp1b***	***growth hormone regulated tbc protein 1b***	**GGCACTGGAAAAATGGAAAA**	**TCGCAGATTTCTGGGAGACT**	**716**
Dr.85673	***tagap***	***t-cell activation GTPase activating protein***	**TGCACAGCAGCCGATTGGGG**	**GGGCTCAGAGCTTCGTCGGC**	**976**
Dr.124255	*plc-l2*	*similar to Inactive phospholipase c-like protein 2*	CAAACCGACGAGGTGGGGGC	GTGGCAGGAGAGGCGAAGCG	696
Dr.80073	***fhl3***	***four and a half lim domains 3***	**TTGAATGAAACTTTGTTCGAAAAG**	**TAATTAAGAATGCAAATTAGGCAACTT**	**2351**
Dr.133138	***irf9***	***interferon regulatory factor 9***	**CGACTGATACTGCACAGGACCG**	**GCTGGGTTGGCGTGCCTGAT**	**1091**
Dr.18530	*FAM166B*	*family with sequence similarity 166, member B*	TCAAAACGGTCTGCGGCCACT	GCAGAGATGCCGGACCTGCC	723
Dr.5660	*crip2*	*cysteine-rich protein 2.*	GCAGGAAGCCACGCAGAGCA	GTCACAAACACCCTTGTTCCCATGT	1104
Dr.135601	*zcchc4*	*zinc finger, CCHC domain containing 4*	AGGACGTCACAGTTCAGCGGC	ACACCAGGCCGTTCGCGAGG	1284
Dr.77065	***hmha1***	***histocompatibility (minor) ha-1***	**GATGTGAATCCCCTCATGTACG**	**GATCCTCTGTAAACCCTTGGAG**	**1286**
Dr.47691	***ankrd58***	***ankyrin repeat domain-containing protein 58-like***	**GAGGGGAACTGCAACGGGGC**	**ACCTCCACTGCCCCTCAGGT**	**717**
Dr.92393	***ccdc135***	***similar to Coiled-coil domain-containing protein 135***	**TGGTGCTCTCGGGAAGCCGT**	**TCGGGCCTGGAAAGGAGCCA**	**1340**
Dr.83578	***tmem151b***	***similar to Transmembrane protein tmem151B***	**AGAAACTGTTTGGCTTTGAACATGGCA**	**CCTGTGTCAATCCCATGCCCCC**	**649**
Dr.81141	*si:ch211-132b12.7*	*hypothetical protein LOC564531*	TGCTGAAGCAGCCTGGGCAGT	AGGAGATGGGGGCGCAACCT	1049
Dr.78875	*sept9b*	*septin 9b*	ACACCTGTGGCACGGAACGA	GGGTGATCGGTCCAACGCTGA	1270
Dr.83297	***aff3***	***af4/fmr2 family, member 3; laf***	**TTCACACACAAGCCTGAAGC**	**GCTCTTTAGTCTCATAGCGTATCTTAT**	**2001**
Dr.90997	*tmem179*	*transmembrane protein 179*	GGCCGTCGATAATTTTCTCTT	CAGAATTTATATGTCACGTCTGAGTG	1076
Dr.34190	***si:ch211-250g4.3***	***si:ch211-250g4.3***	**AGAAGGAGGAACAGCCTAGAAAA**	**TTCCACCTCCCTGTGCTC**	**3324**
Dr.107097	***acsbg2***	***acyl-CoA synthetase bubblegum family member 2***	**AAAGCCGCCATCATTTTG**	**CTGAGCTTCGTCTTGCAGC**	**1676**
Dr.87964	***arhgap27***	***similar to Rho GTPase-activating protein 27***	**AAAATCTCCATTGCCTGCAG**	**CTTTCATAATGTTAAAAACATGTTGAT**	**1801**

Bold have vascular expression at time examined (24 hpf)

Table S6: Primers used to test downregulated gene set following *etsrp* overexpression by qPCR

Unigene ID	Gene Symbol	Gene Description	Forward primer	Reverse primer	Product Size (bp)
Dr.81131	*dpf3*	*D4, zinc and double PHD fingers, family 3*	AGCGCTCGGGGACCAGTTCT	ACGCTGCGCTCTGCACACAA	81
Dr.36017	*myoD*	*myogenic differentiation 1*	GCCCAAAGTGGAGATTCTGA	GCCCATAAAATCCATCATGC	169
Dr.152568	*fez2d*	*FEZ family zinc finger 2*	CGGCTGCACCCAAGTCCCTG	CCTTCGGATCGCTCCTCCGC	94
Dr.133216	*sox32*	*SRY-box containing gene 32*	AGTCCCGAAGCGAAAGCGCC	CGGGTTTAGCTGCGCAAGGC	105
Dr.81305	*bon*	*similar to bonnie and clyde*	TAGGGCCTGTGACTGTGTGAGC	ACTCTGTCCAGGAGGGCCAAG	80
Dr.75761	*rx1*	*retinal homeobox gene 1*	TCCTTCAACCGCCCGTCCAT	GATAGAGGCGAGGGCAGCCA	86
Dr.117206	*ctsl1b*	*similar to cathepsin L, 1 b*	GCGTTGCGACAAAGGCCAGC	TGCAGAGAAGGCTCCAGTGACCT	57
Dr.78792	*nkx2.1b*	*NK2 homeobox 1b*	ACTGACCAGCACACCTCGCA	GAACCGCCGCTCAAGCTCAT	76
Dr.78408	*fli1a*	*friend leukemia integration 1a*	CCGAGGTCCTGCTCTCACAT	GGGACTGGTCAGCGTGAGAT	87
Dr.35143	*b-actin*	*beta1-actin*	TGTTTTCCCCTCCATTGTTG	ACATACATGGCAGGGGTGTT	315

Table S7: Complete *etsrp* mRNA-seq dataset (induced greater than or equal to 1.9 fold)

Unigene ID	OE reads	Con Reads	p-value	Fold Change	Gene ID	Gene Description
Dr.24158	223	0	7.70E-53	1920.81	srgn	Serglycin
Dr.153782	112	0	4.50E-26	1899.93	LOC100151241	Similar to LOC559821 protein
Dr.83594	75	0	5.10E-17	917.38	crabp1a	Cellular retinoic acid binding protein 1a
Dr.118567	56	0	9.90E-13	740.5		Transcribed locus, weakly similar to XP_001332649.1 PREDICTED: similar to prostate stem cell antigen precursor-like [Danio rerio]
Dr.141228	26	0	1.10E-05	452.82	LOC558783	Similar to cell death activator CIDE-A
Dr.149118	93	0	1.90E-21	441.49	LOC798186	Similar to Cdc42 GTPase-activating protein
Dr.151699	16	0	1.70E-03	329.33		Transcribed locus
Dr.118570	833	2	1.80E-196	315.3	LOC561493	Similar to Src homology 2 domain containing E
Dr.40434	28	0	3.50E-06	300.42	ms4a17a.11	Membrane-spanning 4-domains, subfamily A, member 17A.11
Dr.80407	11	0	2.50E-02	282.49		Transcribed locus
Dr.19589	18	0	5.70E-04	260.02		Transcribed locus, moderately similar to NP_877959.2 TAFA2 protein [Mus musculus]
Dr.126663	13	0	8.40E-03	257.54		Transcribed locus
Dr.133475	45	0	3.60E-10	254.84	LOC100151130	Similar to Rasip1 protein
Dr.115354	19	0	3.30E-04	254.27		Transcribed locus, strongly similar to NP_998128.1 DIP2 disco-interacting protein 2 homolog B [Danio rerio]
Dr.37870	32	0	3.80E-07	231.58	ckmb	Creatine kinase, muscle b
Dr.92232	13	0	8.40E-03	231.5	efcab2	EF-hand calcium binding domain 2
Dr.84343	43	0	1.10E-09	225.88	zgc:153721	Zgc:153721
Dr.85502	87	0	4.00E-20	225.67	mrc1	Mannose receptor C1-like protein
Dr.154004	14	0	4.90E-03	222.84		Transcribed locus
Dr.117579	10	0	4.20E-02	219.25		Transcribed locus, weakly similar to XP_001919893.1 PREDICTED: similar to pol polyprotein [Danio rerio]
Dr.47591	4339	15	0.00E+00	218.98	etv2	Ets variant gene 2
Dr.110840	14	0	4.90E-03	217.23	LOC100006361	Similar to preprogalanin 1B
Dr.103328	25	0	2.00E-05	198.68	zgc:162298	Zgc:162298
Dr.142539	15	0	2.80E-03	194.41		Transcribed locus, moderately similar to XP_418909.1 PREDICTED: similar to succinic semialdehyde dehydrogenase [Gallus gallus]
Dr.86665	38	0	2.10E-08	180.63	nhlh2	Nescient helix loop helix 2
Dr.55528	13	0	8.40E-03	180.51		Transcribed locus, weakly similar to XP_001922528.1 PREDICTED: similar to transposase domain-containing protein [Danio rerio]
Dr.28772	15	0	2.80E-03	176.29		Transcribed locus
Dr.114377	15	0	2.80E-03	175.73	zgc:173594	Zgc:173594
Dr.110713	37	0	3.80E-08	165.39	LOC100149611	Similar to LOC495463 protein
Dr.84866	17	0	9.60E-04	159.43	LOC100005159	Similar to hemicentin 1
Dr.151955	10	0	4.20E-02	157.14		Transcribed locus, strongly similar to NP_991168.1 HSPC049 protein-like [Danio rerio]
Dr.111220	20	0	1.90E-04	151.35	gpr183	G protein-coupled receptor 183
Dr.113631	11	0	2.50E-02	136.05	LOC568486	Similar to regulating synaptic membrane exocytosis 1
Dr.82184	27	0	6.20E-06	123.37	LOC569386	Platelet endothelial cell adhesion molecule-like
Dr.152176	10	0	4.20E-02	115.44		Transcribed locus, weakly similar to XP_001344258.2 PREDICTED: similar to Uromodulin precursor (Tamm-Horsfall urinary glycoprotein) (THP) [Danio rerio]
Dr.124732	12	0	1.40E-02	113.88	LOC100148619	Similar to myelin associated glycoprotein
Dr.152900	14	0	4.90E-03	110.08	LOC572378	Similar to transforming growth factor-b type II receptor
Dr.142587	23	0	6.50E-05	104.88	csf3r	Colony stimulating factor 3 receptor (granulocyte)

Note: Only few entries included in sample, table is long

Table S8: Downregulated mRNA-seq dataset (reduced greater than or equal to 1.9 fold by overexpression of *etsrp*)

Unigene ID	OE reads	Con Reads	p-value	Fold Change	Gene ID	Gene Description
Dr.132051	365	525	1.62E-19	0.526		Transcribed locus
Dr.86155	50	72	0.004458765	0.526	si:busm1-265n4.4	Si:busm1-265n4.4
Dr.133043	43	62	0.009262994	0.526	zgc:158614	Zgc:158614
Dr.134954	93	134	2.70E-05	0.526		Transcribed locus
Dr.138538	29	42	0.048438932	0.524		Transcribed locus
Dr.151247	132	191	1.86E-07	0.524		Transcribed locus, strongly similar to NP_775382.1 laminin, beta 1 [Danio rerio]
Dr.79919	29	42	0.048438932	0.524	mmp23a	Matrix metallopeptidase 23a
Dr.76174	89	129	3.27E-05	0.524		Transcribed locus
Dr.40630	38	55	0.016958396	0.524		Transcribed locus, strongly similar to XP_001922646.1 PREDICTED: similar to zinc finger protein 335 [Danio rerio]
Dr.74540	378	546	1.39E-20	0.524	si:ch211-133n4.4	Si:ch211-133n4.4
Dr.154333	40	58	0.012001711	0.521		Transcribed locus
Dr.134004	285	414	6.51E-16	0.521	LOC799498	Similar to Tudor domain-containing protein 6 (Antigen NY-CO-45) (Cancer/testis antigen 41.2) (CT41.2)
Dr.103609	33	48	0.031587364	0.521		Transcribed locus
Dr.132371	35	51	0.022261443	0.521		Transcribed locus
Dr.79560	40	58	0.012001711	0.521		Transcribed locus
Dr.150604	135	196	1.06E-07	0.521		Transcribed locus
Dr.150494	31	45	0.034489362	0.521		Transcribed locus, strongly similar to XP_001918516.1 PREDICTED: hypothetical protein [Danio rerio]
Dr.92708	31	45	0.034489362	0.521		Transcribed locus
Dr.540	437	634	3.48E-24	0.521	rx3	Retinal homeobox gene 3
Dr.137654	50	73	0.00341522	0.518		Transcribed locus
Dr.75659	240	351	1.07E-13	0.518	zgc:73311	Zgc:73311
Dr.87327	52	76	0.002403868	0.518	chrna5	Cholinergic receptor, nicotinic, alpha 5
Dr.128789	41	60	0.010026041	0.518		Transcribed locus
Dr.150399	83	121	5.70E-05	0.518		Transcribed locus
Dr.143137	30	44	0.032017144	0.515		Transcribed locus
Dr.153302	34	50	0.020444485	0.515		Transcribed locus
Dr.152446	34	50	0.020444485	0.515	zgc:158653	Zgc:158653

Note: Only few entries included in sample, table is long

Supplementary Text

Text S1

Genes marked with an asterisk have two identifiers in the current Ensembl build, Zv9. To clarify the identity of the gene examined here, their chromosomal locations have been noted along with their un-examined paralog(s) in Table S3.

K. *cntn4/big2**

contactin4 is a member of the subset of immunoglobulin superfamily of proteins expressed in the nervous system (Shimoda and Watanabe, 2009), whose function is crucial for axon guidance in the formation of the olfactory bulbs of mice (Kaneko-Goto et al., 2008). In vitro evidence also suggests *contactin4* is involved in axonal migration behaviors during retinotectal development (Osterfield et al., 2008). Robust expression has also been detected in human testis and at lower levels in various other organs (Zeng et al., 2002), but *cntn4* disruption has only been associated with 3p deletion syndrome and autism (Fernandez et al., 2004; Roohi et al., 2009). In zebrafish there is strong *cntn4* expression in the olfactory bulb, spinal chord neurons, and in both axial and cranial vasculature (Figure 3K).

L. *myo1f*

myo1f is an unconventional myosin that was initially identified from cochlear tissues of mice and has also been detected in eye, brain, liver, lungs and other tissues (Crozet et al., 1997). In a separate study, the expression of *myo1F* was only detected in tissues of the immune system, in particular the knockout mice display defects in neutrophil migration due to an inability of neutrophils to exocytose granules that prevent neutrophil adhesion to integrin containing substrates (Kim et al., 2006). Of clinical relevance, an association between mutations in *myo1f* and hearing pathologies has been noted (Zadro et al., 2009). The expression of several

unconventional myosins in zebrafish has previously been published (Sittaramane and Chandrasekhar, 2008), and we note that *myo1f* has a low level of ubiquitous basal expression with a more robust expression in vasculature and in myeloid cells (Figure 3L).

M. *iclp2*

The mammalian invariant chain proteins function in antigen presentation by immune cells, however there are two separate genes in zebrafish, *iclp1* and *iclp2*. The latter lacks a fragment that is required for antigen presentation, suggesting that in zebrafish, *iclp2* may have a distinct function from the mammalian invariant chain and this alternate role is believed to be enzymatic in nature (Yoder et al., 1999).

N. *rgl2*

ral guanine nucleotide dissociation stimulator-like 2, *rgl2*, is a member of the ral-Guanine Dissociation Stimulator, ral-GDS, family of proteins, which are effectors of ras (Ferro and Trabalzini, 2010), whose induction by overexpression of *etsrp* was detected previously (Wong et al., 2009b). *Rgl2* expression in humans is reported to be ubiquitous, as is the mouse ortholog, *rlf* (Isomura et al., 1996; Peterson et al., 1996). However, *rlf* presents higher expression levels in heart and lung (Post et al., 2002). The activation of *rgl2* by ras proteins is dependent on specific modifications such as phosphorylation by PKA (Ferro et al., 2008), and one function of rgl2 bound by R-ras is the activation of ralA, which promotes exocytosis of endosomes (Takaya et al., 2007). Of clinical relevance, *rgl2*, promotes hypertrophic growth of terminally differentiated cardiomyocytes (Post et al., 2002), and its expression is increased in pancreatic cancer cells, where it promotes cancer cell growth through ral-dependent and ral-independent activation (Vigil et al., 2010). In zebrafish, *rgl2* is expressed throughout the developing vascular endothelial system, primitive erythrocytes and the forebrain (Figure 3N and Figure S1N).

O. *capn8*

calpain8 is a member of the calpain family of intracellular calcium regulated cystein proteases, some of whose expression and evolution in developing zebrafish embryos have been previously examined (Lepage and Bruce, 2008). In mammals *calpain8* is orthologous with *nCL-2*, which is specifically expressed and functional in the stomach of mice (Hata et al., 2006). The ortholog in xenopus, *xcl2*, is expressed in non-mesodermal and non-vascular tissues during developmental stages and its knockdown results in gastrulation defects (Cao et al., 2001). Nonetheless, we previously detected the upregulation of *calpain8* by overexpressing *etsrp* using a microarray approach (Gomez et al., 2009), and therefore examined its expression here. *Calpain8* of zebrafish has basal ubiquitous levels of expression, with pronounced expression in the axial vasculature, forebrain, lateral line primordium, and hatching gland (Figure 3O).

P. *mhc1uea*

major histocompatibility class 1 uea of zebrafish is orthologous with MHC1 Related protein MR1 of humans, which is well conserved and ubiquitously expressed (Huang et al., 2008). A current hypothesis is that MR1 presents antigens to immunogenic cells of mucosa associated invariant T cells (mait), the innate T cells resident within the intestine that can mount a rapid immune response, but the antigens presented to mait cells by MR1 are unknown (Huang et al., 2009b). As previously reported, *mhc1uea* is induced by *etsrp* and expressed ubiquitously (Wong et al., 2009b), and we note here that the expression within the trunk vasculature is reduced in *etsrp* morphants (Figure 4P and Figure S1P).

Q. *dgki*

diacylglycerol kinase iota, *dgki*, is one of nine members of diacylglyerol kinases in mammals that are divided into five subtypes, that metabolize and thereby dampen diacylglycerol signaling

by phosphorylating it to phosphatidic acid (Reviewed in (Luo et al., 2004)). The human *dgki* was originally cloned from retinal and brain libraries, and detected in both cytoplasmic and nuclear fractions (Ding et al., 1998). *Dgki* knockout mice are viable and appear normal but harbor reduced ras signaling and are disposed to defects in metabotropic glutamate receptor-dependent long-term depression (Regier et al., 2005; Yang et al., 2011).

R. *ifit5-like**

interferon-induced protein with tetratricopeptide repeats 5-like, may be the ortholog of the mammalian *ifit5*, which was identified in a promyelocytic leukemia cell line, NB4, through differential hybridization as a gene that is induced directly by interpheron alpha and indirectly by retinoic acid (Niikura et al., 1997). Although knowledge on this gene is limited to the report on its identification, *ifit5* was classified as a member of the interferon inducible family of proteins owing to the means of its induction and the presence of tetratricopeptide repeat (TPR) motifs characteristic of this family. TPR motifs mediate protein-protein interactions, but their functional roles in *ifit5* remains obscure. While human *ifit5* contains 9 TPR motifs (Niikura et al., 1997), a pfam domain scan of the zebrafish *ifit5-like* reported here is predicted to contain 4 TPR motifs.

S. *tmem88a*

transmembrane protein 88a is a multipass transmebrane protein with two transmembrane regions. It was previously listed as a gene induced by *etsrp* (Wong et al., 2009b) and its paralog, *tmem88b* has the same topology. While humans have a single ortholog, *tmem88*, mice have two, *tmem88* and *tmem88b*, both of which remain uncharacterized. Like other vascular specific genes, *tmem88a* is downregulated in the axial vasculature of *etsrp* morphants (Figure 3S).

T. *rasa4*

ras p21 protein activator 4, *rasa4*, or *Ca2+-promoted ras inactivator*, *Capri*, is a well conserved ras gtpase member of the gap1 family that negatively regulates ras signaling by switching off the Ras-MAPK signaling pathway in response to calcium signaling (Lockyer et al., 2001). By northern blot *rasa4* was found to have ubiquitous expression in humans (Nagase et al., 1998), but knockout mice are viable yet experience an impaired ability to mount innate immune responses to pathogens because it functions as an adaptor for cdc42 and rac during FcR-mediated phagocytosis (Zhang et al., 2005). Through an RNAi screen, *rasa4* was also identified as one of several hundred genes associated with necropotosis, emphasizing its involvement in tumor signaling inhibition (Hitomi et al., 2008). Although we note a broad distribution of expression at the stage examined, there is bolder expression within the axial vasculature in the trunk of zebrafish embryos that is reduced in *etsrp* morphants (Figure 3T and Figure S1T).

U. *samd10**

The only notable domain in this gene is the sam domain, which is about 70 amino acids long and is heavily credited with mediating protein-protein interactions in a diverse array of proteins and tissues including several proteins that are involved in developmental processes (Schultz et al., 1997). There is also evidence that sam domains bind mRNA, which increases the possible functional role of the encoded protein (Schultz 2003). We note the expression of *samd10* in both the cranial and axial vasculature and budding somites, but not in intersegmental vessels (Figure 3U).

V. *cald1**

Caldesmon is a protein that binds actin, myosin, tropomyosin and calmodulin. In humans there are two isoforms generated by alternative splicing, the high molecular weight form, *hcald*, is smooth muscle specific and the low molecular weight form, *lcald*, is expressed in non-muscle

cells (Hayashi et al., 1992). The main function associated with *caldesmon* is the regulation of cell motility, and is a putative repressor of cancer cell invasion (Li et al., 2004; Yoshio et al., 2007). *Lcald* has been identified in cancerous gliomas and in endothelial cells of breast, lung, kidney, colon, stomach, ovary, uterus, prostate, thyroid and liver cancers, but never in the normal vasculature of adults (Zheng et al., 2005a). The *caldesmon1*, *cald1*, of zebrafish is the ortholog of *lcald*, and its knockdown by morpholinos results in heart formation defects but the effects on vascular development remains to be examined (Zheng et al., 2009). The *cald1* gene examined here is expressed in the vasculature but not in the heart and is located on chromosome 4, while Zheng et al knocked down the cald1 gene located on chromosome 25 (Table S3).

W. *sept5b*

Initially identified as *cell division cycle related 1*, *CDCrel1*, and *peanut like 1*, *PNUTL1*, in a megakaryocytic cell line and endothelial cells (Kelly et al., 1994; McKie et al., 1997; Zieger et al., 1997), *septin5* expression was also identified in human fetal brain and heart but only in cell bodies and dendrites of adult neurons, and platelets (Bläser et al., 2002; Caltagarone et al., 1998). *CDCrel1/sept5* knockout mice appear normal, which may be due to compensation by other septins, of which there are many (Peng et al., 2002), but they exhibit social and cognitive disorders (Suzuki et al., 2009). Although there is precedent for the expression of *septin5* within huvecs (Bläser et al., 2006), in zebrafish we also note faint but definite expression that outlines the dorsal aorta in the trunk vasculature (Figure 3W and Figure S1W), as well as the forebrain and somites. The nearest evolutionary paralog, *septin5a* has not been examined.

X. *zgc:171494**

The orthologous gene in higher vertebrates is *ceramide synthase 3*, *cers3*, which adds acyl chains to catalyze the synthesis of ceramides containing C18 and longer acyl chains, is expressed in

testis and at low levels in the skin of mice (Mizutani et al., 2006). Although there is some slight ubiquitous expression, there is marked expression of *zgc:171494* in the axial vasculature (Figure 3X) and the expression in the cranial vasculature becomes more apparent in *etsrp* morphants (Figure 4X).

Y. *ankdd1a*

ankyrin repeat and death domain containing 1a, *ankdd1a* encodes a protein that is 489 amino acids long and is predicted to contain five ankyrin repeats and a c-terminal death domain. Ankyrin repeats mediate protein-protein interactions, and the death domain associated with apoptosis is also found in proteins involved in non-apoptotic molecular functions (Reed et al., 2004; Weber and Vincenz, 2001). Although highly conserved, *ankdd1a* has not been characterized and while mice and humans both have a putative paralog, *ankdd1b*, this paralog in zebrafish has currently not been identified or annotated.

Z. *sh3bp4*

sh3 binding protein 4, *sh3bp4*, or *transferring receptor trafficking protein*, *TTP*, was cloned from human cornea fibroblasts and is also expressed in pancreas, heart, placenta, kidney, skeletal muscle, liver (Dunlevy et al., 1999) and retina (Khanobdee et al., 2004). Although the protein encoded by *sh3bp4* has been localized at the plasma membrane by two independent groups, there is debate on its localization at the nucleus (Khanobdee et al., 2004; Tosoni et al., 2005). Functionally, *sh3bp4* is classified as an accessory endocytic protein that specifically internalizes the transferring receptor (Tosoni et al., 2005). Several interaction partners of *Sh3bp4* have been discovered by yeast two hybrid screens including *Ciz1*, *Plekha1*, and *Ttc1*, which regulate signaling and stress responses (Thalappilly et al., 2008).

AA. *myof* *

myoferlin is induced by *etsrp* overexpression (Wong et al., 2009b), and while it is commonly associated with muscle development, its involvement in endothelial cell biology has also been revealed (Bernatchez et al., 2007; Bernatchez et al., 2009). We demonstrate here that its expression in endothelial cells is conserved in zebrafish with signal detected in the cranial and axial vasculature of the trunk, as well as the somitic mesoderm in the tail of 24 hpf embryos (Figure 3AA').

AB. *grtp1b*

growth hormone regulated TBC protein 1b, was originally identified in the transcriptional profiling of cardiomyocytes in response to growth hormone signaling (Lu et al., 2001). Widely conserved, its expression in mice is highest in testis, with moderate expression in kidney and liver and lowest in lung, intestine and stomach. Possession of a TBC domain suggests it functions as an activator of the GTPase, Rab.

AC. *Tagap**

T-cell activation GTPase activating protein, *tagap*, is a Rho gap that was identified as a gene that is regulated during T cell activation (Mao et al., 2004). In mice *tagap* appears to be expressed ubiquitously with an apparent prominence in sperm and although knockouts are viable, they have transmission ratio distortion defects (Bauer et al., 2005). Nevertheless, of clinical significance, *tagap* is downregulated in T cells of patients with down syndrome (Sommer et al., 2008), and is associated with celiac disease and type 1 diabetes (Hunt et al., 2008; Romanos et al., 2009; Smyth et al., 2008).

AD. *fhl3**

four and a half lim domains 3, *fhl3*, contains four LIM domains that mediate protein-protein interactions, and has been localized in both the cytoplasm and the nucleus (Coghill et al., 2003;

Takahashi et al., 2005). In the cytoplasm it regulates α-actinin mediated actin bundling where it enhances cell spreading and stress fiber disassembly (Coghill et al., 2003). In the nuclei of certain blood lineages, *fhl3* is a transcriptional co-repressor of the beta chain of the IgE receptor (Takahashi et al., 2005), while in myogenesis it attenuates *myoD* expression (Cottle et al., 2007). *Fhl3* is also classified as a co-activator in satellite cells where it cooperates with *sox15* in the activation of *foxk1* (Meeson et al., 2007). Regarding disease, *fhl3* together with paralogs *fhl1* and *fhl2* are downregulated in liver cancer cells where they function as tumor suppressors through tgf-β dependent and independent signaling by modulation of *smad* activity (Ding et al., 2009).

AE. *irf9*

interferon regulatory factor 9, *irf9*, is part of the heterotrimeric transcription factor, *interferon stimulated gene factor 3*, *isgf3*, which includes the binding partners *stat1* and *stat2* (Gerber and Pober, 2008). In normoxic endothelial cells, a direct target of *isgf3* and therefore interferon signaling is *hypoxia inducible factor 1α*, *hif1α*, which results in the inhibition of proliferation (Gerber and Pober, 2008). *irf9* is highly expressed in breast and uterine cancer cells, and forced expression of *irf9* regulates interferon response genes independently of interferon signaling, resulting in resistance to anti-microtubule agents and cell survival (Luker et al., 2001). However, there is little evidence demonstrating that *irf9* can activate transcription when not in the *isg3* complex (Kraus et al., 2003). Nonetheless, *etsrp* overexpression ectopically induced the expression of *irf9*, which shows marked expression in the axial (Figure 3AE') and cranial vasculature of *etsrp* morphants (Figure 4AE'). Neither *stat1* nor *stat2* appeared on the dataset of genes induced by *etsrp*.

AF. *hmha1**

histocompatibility (minor) HA-1, located on chromosome 22 encodes a protein with a rhoGAP domain, a protein kinase C-like phorbol ester/diacylglerol binding domain, and an FCH domain. The mammalian ortholog is expressed in hematopoietic cells and some tumorous as well as noncancerous lung and hepatic tissues (Fujii et al., 2002; Klein et al., 2002). Although minor histocompatibility antigens are clinically relevant in stem cell and organ transplantation in the mediation of transplant rejection/acceptance (Spierings, 2008), the expression and presentation by vascular endothelial cells is currently not believed to activate immunologic responses (Bolinger et al., 2008). The function of *hmha1* in the developing endothelial cells of zebrafish remains to be determined.

AG. *ankrd58*

This gene has a modest level of basal expression but has robust axial vascular expression in the trunk (Figure 3AG'). Based on Ensembl predictions, it is conserved and the mammalian ortholog is *ankrd58*. While zebrafish is predicted to contain two ankyrin repeats, the mouse ortholog also has two, while in humans there are three. Ankyrin repeats are involved in protein-protein interactions (Mosavi et al., 2004). This gene awaits further characterization.

AH. *ccdc135*

coiled-coil domain containing 135, *ccdc135*, or *spermatogenesis related gene like*, *srgl*, was first cloned from spermatocytes of mice and also detected in spleen, kidney, ovary and thymus (Ma et al., 2006). *Ccdc135* was subsequently identified as a gene that is preferentially expressed in ciliated cells such as olfactory sensory neurons (McClintock et al., 2008), and more recently it was reported that the ortholog in *Drosophila melanogaster* is specifically localized to axonemes where it regulates sperm motility (Yang et al., 2011). In 24 hpf zebrafish embryos there is slight

basal expression throughout the embryos with pronounced expression in the cranial and axial trunk vasculature, pronephros and tailbud (Figure 3AH').

AI. *tmem151a**

tmem151a encodes a protein with four transmembrane domains that is well conserved in vertebrates. There is a second gene in Ensembl with the same name as well as 4 paralogs called *tmem151b*. *tmem151a* remains to be characterized further.

AJ. *aff3*

AF4/FMR2 family member 3, *aff3*, was originally identified as a lymphoid nuclear protein with transactivation potential that is related to the AF-4 family of genes (Ma and Staudt, 1996). In mice *aff3* is only expressed in thymus, spleen, brain and lungs (Ma and Staudt, 1996), while in humans it is expressed in the heart, brain, and placenta of adults, as well as fetal tissues (Hiwatari et al., 2003). The non-vertebrate ortholog of drosophila, *lilliputian*, controls cell growth and differentiation (Wittwer et al., 2001), suggesting *aff3* may regulate cell growth. Consequently, some breast cancer tumor cells have increased expression of *aff3*, which is otherwise not expressed in these organs (To et al., 2005). Genetically, *aff3* has been associated with acute lymphoblastic leukemia (Hiwatari et al., 2003), and a predisposition to rheumatoid arthritis (Barton et al., 2009).

AK. *si:ch211-250g4.3*

si:ch211-250g4.3 encodes a protein with a death like-domain at the N-terminus followed by a FYVE zinc finger domain, a prefoldin domain and three trompomyosin domains. The evolutionary paralog with highest conservation is an uncharacterized gene called CR385063.1 in the Ensembl database. The main vertebrate ortholog associated with *si:ch211-250g4.3* is *ninein*, however zebrafish also contain both a *ninein* gene as well as a *ninein-like* gene that have higher

sequences conserved with human and mouse *ninein* than *si:ch211-250g4.3*. *si:ch211-250g4.3* has not been characterized, but is expressed specifically in the developing vasculature of zebrafish (Figure 3AK').

AL. *acsbg2*

acyl-CoA synthetase bubblegum-related 2, *acsbg2* is an enzyme that preferentially activates oleic and linoleic fatty acids (Pei et al., 2006). Although there is a consensus that *acsbg2* is expressed in testis, Zheng et al., 2005b (Zheng et al., 2005b) also detected it in human pancreas, liver, small intestine, heart, and kidney, while Fraisl et al., 2006 (Fraisl et al., 2006), only detected in testis, and Pei et al., 2006 (Pei et al., 2006) detected it in testis, and motor neurons in the medulla oblongata and cervical spinal cord. In zebrafish *acsbg2* is expressed in neurons of the spinal cord, as well as the diencephalon, hindbrain and axial vasculature (Figure 3AL').

AM. *arhgap27*

arhgap27, also known as *CIN85 associated multi-domain containing RhoGAP1*, *camgap1*, was initially identified as a binding partner of cin85 by a yeast two hybrid screen, is an active GAP of *cdc42* and *Rac1*, and is implicated in receptor mediated endocytosis (Sakakibara et al., 2004). *Arhgap27* is expressed in various tissues in mice including the heart (Sakakibara et al., 2004), and humans express two splice variants due to alternative splicing (Katoh and Katoh, 2004). In zebrafish, arhgap27 is vascular specific at 24 hpf (Figure 3AM').

References in Supplementary Text

1. Shimoda Y, Watanabe K (2009) Contactins. Cell Adhesion & Migration 3: 64-70.
2. Kaneko-Goto T, Yoshihara S-I, Miyazaki H, Yoshihara Y (2008) BIG-2 mediates olfactory axon convergence to target glomeruli. Neuron 57: 834-846.
3. Osterfield M, Egelund R, Young LM, Flanagan JG (2008) Interaction of amyloid precursor protein with contactins and NgCAM in the retinotectal system. Development 135: 1189-1199.

4. Zeng L, Zhang C, Xu J, Ye X, Wu Q, et al. (2002) A novel splice variant of the cell adhesion molecule contactin 4 (CNTN4) is mainly expressed in human brain. J Hum Genet 47: 497-499.
5. Fernandez T, Morgan T, Davis N, Klin A, Morris A, et al. (2004) Disruption of contactin 4 (CNTN4) results in developmental delay and other features of 3p deletion syndrome. Am J Hum Genet 74: 1286-1293.
6. Roohi J, Montagna C, Tegay DH, Palmer LE, DeVincent C, et al. (2009) Disruption of contactin 4 in three subjects with autism spectrum disorder. J Med Genet 46: 176-182.
7. Crozet F, el Amraoui A, Blanchard S, Lenoir M, Ripoll C, et al. (1997) Cloning of the genes encoding two murine and human cochlear unconventional type I myosins. Genomics 40: 332-341.
8. Kim SV, Mehal WZ, Dong X, Heinrich V, Pypaert M, et al. (2006) Modulation of cell adhesion and motility in the immune system by Myo1f. Science 314: 136-139.
9. Zadro C, Alemanno MS, Bellacchio E, Ficarella R, Donaudy F, et al. (2009) Are MYO1C and MYO1F associated with hearing loss? Biochim Biophys Acta 1792: 27-32.
10. Sittaramane V, Chandrasekhar A (2008) Expression of unconventional myosin genes during neuronal development in zebrafish. Gene Expr Patterns 8: 161-170.
11. Yoder JA, Haire RN, Litman GW (1999) Cloning of two zebrafish cDNAs that share domains with the MHC class II-associated invariant chain. Immunogenetics 50: 84-88.
12. Ferro E, Trabalzini L (2010) RalGDS family members couple Ras to Ral signalling and that's not all. Cell Signal 22: 1804-1810.
13. Wong KS, Proulx K, Rost MS, Sumanas S (2009) Identification of vasculature-specific genes by microarray analysis of Etsrp/Etv2 overexpressing zebrafish embryos. Dev Dyn 238: 1836-1850.
14. Isomura M, Okui K, Fujiwara T, Shin S, Nakamura Y (1996) Isolation and mapping of RAB2L, a human cDNA that encodes a protein homologous to RalGDS. Cytogenet Cell Genet 74: 263-265.
15. Peterson SN, Trabalzini L, Brtva TR, Fischer T, Altschuler DL, et al. (1996) Identification of a novel RalGDS-related protein as a candidate effector for Ras and Rap1. J Biol Chem 271: 29903-29908.
16. Post GR, Swiderski C, Waldrop BA, Salty L, Glembotski CC, et al. (2002) Guanine nucleotide exchange factor-like factor (Rlf) induces gene expression and potentiates alpha 1-adrenergic receptor-induced transcriptional responses in neonatal rat ventricular myocytes. J Biol Chem 277: 15286-15292.
17. Ferro E, Magrini D, Guazzi P, Fischer TH, Pistolesi S, et al. (2008) G-protein binding features and regulation of the RalGDS family member, RGL2. Biochem J 415: 145-154.
18. Takaya A, Kamio T, Masuda M, Mochizuki N, Sawa H, et al. (2007) R-Ras regulates exocytosis by Rgl2/Rlf-mediated activation of RalA on endosomes. Mol Biol Cell 18: 1850-1860.
19. Vigil D, Martin TD, Williams F, Yeh JJ, Campbell SL, et al. (2010) Aberrant overexpression of the Rgl2 Ral small GTPase-specific guanine nucleotide exchange factor promotes pancreatic cancer growth through Ral-dependent and Ral-independent mechanisms. J Biol Chem 285: 34729-34740.
20. Lepage SE, Bruce AEE (2008) Characterization and comparative expression of zebrafish calpain system genes during early development. Dev Dyn 237: 819-829.

21. Hata S, Koyama S, Kawahara H, Doi N, Maeda T, et al. (2006) Stomach-specific calpain, nCL-2, localizes in mucus cells and proteolyzes the beta-subunit of coatomer complex, beta-COP. J Biol Chem 281: 11214-11224.
22. Cao Y, Zhao H, Grunz H (2001) XCL-2 is a novel m-type calpain and disrupts morphogenetic movements during embryogenesis in Xenopus laevis. Dev Growth Differ 43: 563-571.
23. Gomez GA, Veldman MB, Zhao Y, Burgess S, Lin S (2009) Discovery and characterization of novel vascular and hematopoietic genes downstream of etsrp in zebrafish. PLoS ONE 4: e4994.
24. Huang S, Gilfillan S, Kim S, Thompson B, Wang X, et al. (2008) MR1 uses an endocytic pathway to activate mucosal-associated invariant T cells. J Exp Med 205: 1201-1211.
25. Huang S, Martin E, Kim S, Yu L, Soudais C, et al. (2009) MR1 antigen presentation to mucosal-associated invariant T cells was highly conserved in evolution. Proc Natl Acad Sci USA 106: 8290-8295.
26. Luo B, Regier DS, Prescott SM, Topham MK (2004) Diacylglycerol kinases. Cell Signal 16: 983-989.
27. Ding L, Traer E, McIntyre TM, Zimmerman GA, Prescott SM (1998) The cloning and characterization of a novel human diacylglycerol kinase, DGKiota. J Biol Chem 273: 32746-32752.
28. Regier DS, Higbee J, Lund KM, Sakane F, Prescott SM, et al. (2005) Diacylglycerol kinase iota regulates Ras guanyl-releasing protein 3 and inhibits Rap1 signaling. Proc Natl Acad Sci USA 102: 7595-7600.
29. Yang Y, Cochran DA, Gargano MD, King I, Samhat NK, et al. (2011) Regulation of flagellar motility by the conserved flagellar protein CG34110/Ccdc135/FAP50. Mol Biol Cell 22: 976-987.
30. Niikura T, Hirata R, Weil SC (1997) A novel interferon-inducible gene expressed during myeloid differentiation. Blood Cells Mol Dis 23: 337-349.
31. Lockyer PJ, Kupzig S, Cullen PJ (2001) CAPRI regulates Ca(2+)-dependent inactivation of the Ras-MAPK pathway. Curr Biol 11: 981-986.
32. Nagase T, Ishikawa K, Miyajima N, Tanaka A, Kotani H, et al. (1998) Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro. DNA Res 5: 31-39.
33. Zhang J, Guo J, Dzhagalov I, He Y-W (2005) An essential function for the calcium-promoted Ras inactivator in Fcgamma receptor-mediated phagocytosis. Nat Immunol 6: 911-919.
34. Hitomi J, Christofferson DE, Ng A, Yao J, Degterev A, et al. (2008) Identification of a molecular signaling network that regulates a cellular necrotic cell death pathway. Cell 135: 1311-1323.
35. Schultz J, Ponting CP, Hofmann K, Bork P (1997) SAM as a protein interaction domain involved in developmental regulation. Protein Sci 6: 249-253.
36. Hayashi K, Yano H, Hashida T, Takeuchi R, Takeda O, et al. (1992) Genomic structure of the human caldesmon gene. Proc Natl Acad Sci USA 89: 12122-12126.
37. Li Y, Lin JLC, Reiter RS, Daniels K, Soll DR, et al. (2004) Caldesmon mutant defective in Ca(2+)-calmodulin binding interferes with assembly of stress fibers and affects cell morphology, growth and motility. J Cell Sci 117: 3593-3604.

38. Yoshio T, Morita T, Kimura Y, Tsujii M, Hayashi N, et al. (2007) Caldesmon suppresses cancer cell invasion by regulating podosome/invadopodium formation. FEBS Lett 581: 3777-3782.
39. Zheng P-P, van der Weiden M, Kros JM (2005) Differential expression of Hela-type caldesmon in tumour neovascularization: a new marker of angiogenic endothelial cells. J Pathol 205: 408-414.
40. Zheng P-P, Severijnen L-A, Willemsen R, Kros JM (2009) Caldesmon is essential for cardiac morphogenesis and function: in vivo study using a zebrafish model. Biochem Biophys Res Commun 378: 37-40.
41. Kelly MD, Essex DW, Shapiro SS, Meloni FJ, Druck T, et al. (1994) Complementary DNA cloning of the alternatively expressed endothelial cell glycoprotein Ib beta (GPIb beta) and localization of the GPIb beta gene to chromosome 22. J Clin Invest 93: 2417-2424.
42. McKie JM, Sutherland HF, Harvey E, Kim UJ, Scambler PJ (1997) A human gene similar to Drosophila melanogaster peanut maps to the DiGeorge syndrome region of 22q11. Hum Genet 101: 6-12.
43. Zieger B, Hashimoto Y, Ware J (1997) Alternative expression of platelet glycoprotein Ib(beta) mRNA from an adjacent 5' gene with an imperfect polyadenylation signal sequence. J Clin Invest 99: 520-525.
44. Bläser S, Jersch K, Hainmann I, Wunderle D, Zgaga-Griesz A, et al. (2002) Human septin-septin interaction: CDCrel-1 partners with KIAA0202. FEBS Lett 519: 169-172.
45. Caltagarone J, Rhodes J, Honer WG, Bowser R (1998) Localization of a novel septin protein, hCDCrel-1, in neurons of human brain. Neuroreport 9: 2907-2912.
46. Peng X-R, Jia Z, Zhang Y, Ware J, Trimble WS (2002) The septin CDCrel-1 is dispensable for normal development and neurotransmitter release. Mol Cell Biol 22: 378-387.
47. Suzuki G, Harper KM, Hiramoto T, Sawamura T, Lee M, et al. (2009) Sept5 deficiency exerts pleiotropic influence on affective behaviors and cognitive functions in mice. Hum Mol Genet 18: 1652-1660.
48. Bläser S, Röseler S, Rempp H, Bartsch I, Bauer H, et al. (2006) Human endothelial cell septins: SEPT11 is an interaction partner of SEPT5. J Pathol 210: 103-110.
49. Mizutani Y, Kihara A, Igarashi Y (2006) LASS3 (longevity assurance homologue 3) is a mainly testis-specific (dihydro)ceramide synthase with relatively broad substrate specificity. Biochem J 398: 531-538.
50. Reed JC, Doctor KS, Godzik A (2004) The domains of apoptosis: a genomics perspective. Sci STKE 2004: re9.
51. Weber CH, Vincenz C (2001) The death domain superfamily: a tale of two interfaces? Trends Biochem Sci 26: 475-481.
52. Dunlevy JR, Berryhill BL, Vergnes JP, SundarRaj N, Hassell JR (1999) Cloning, chromosomal localization, and characterization of cDNA from a novel gene, SH3BP4, expressed by human corneal fibroblasts. Genomics 62: 519-524.
53. Khanobdee K, Kolberg JB, Dunlevy JR (2004) Nuclear and plasma membrane localization of SH3BP4 in retinal pigment epithelial cells. Mol Vis 10: 933-942.
54. Tosoni D, Puri C, Confalonieri S, Salcini AE, De Camilli P, et al. (2005) TTP specifically regulates the internalization of the transferrin receptor. Cell 123: 875-888.
55. Thalappilly S, Suliman M, Gayet O, Soubeyran P, Hermant A, et al. (2008) Identification of multi-SH3 domain-containing protein interactome in pancreatic cancer: a yeast two-hybrid approach. Proteomics 8: 3071-3081.

56. Bernatchez PN, Acevedo L, Fernandez-Hernando C, Murata T, Chalouni C, et al. (2007) Myoferlin regulates vascular endothelial growth factor receptor-2 stability and function. J Biol Chem 282: 30745-30753.
57. Bernatchez PN, Sharma A, Kodaman P, Sessa WC (2009) Myoferlin is critical for endocytosis in endothelial cells. Am J Physiol, Cell Physiol 297: C484-492.
58. Lu C, Kasik J, Stephan DA, Yang S, Sperling MA, et al. (2001) Grtp1, a novel gene regulated by growth hormone. Endocrinology 142: 4568-4571.
59. Mao M, Biery MC, Kobayashi SV, Ward T, Schimmack G, et al. (2004) T lymphocyte activation gene identification by coregulated expression on DNA microarrays. Genomics 83: 989-999.
60. Bauer H, Willert J, Koschorz B, Herrmann BG (2005) The t complex-encoded GTPase-activating protein Tagap1 acts as a transmission ratio distorter in mice. Nat Genet 37: 969-973.
61. Sommer CA, Pavarino-Bertelli EC, Goloni-Bertollo EM, Henrique-Silva F (2008) Identification of dysregulated genes in lymphocytes from children with Down syndrome. Genome 51: 19-29.
62. Hunt KA, Zhernakova A, Turner G, Heap GAR, Franke L, et al. (2008) Newly identified genetic risk variants for celiac disease related to the immune response. Nat Genet 40: 395-402.
63. Romanos J, Barisani D, Trynka G, Zhernakova A, Bardella MT, et al. (2009) Six new coeliac disease loci replicated in an Italian population confirm association with coeliac disease. J Med Genet 46: 60-63.
64. Smyth DJ, Plagnol V, Walker NM, Cooper JD, Downes K, et al. (2008) Shared and distinct genetic variants in type 1 diabetes and celiac disease. N Engl J Med 359: 2767-2777.
65. Coghill ID, Brown S, Cottle DL, McGrath MJ, Robinson PA, et al. (2003) FHL3 is an actin-binding protein that regulates alpha-actinin-mediated actin bundling: FHL3 localizes to actin stress fibers and enhances cell spreading and stress fiber disassembly. J Biol Chem 278: 24139-24152.
66. Takahashi K, Matsumoto C, Ra C (2005) FHL3 negatively regulates human high-affinity IgE receptor beta-chain gene expression by acting as a transcriptional co-repressor of MZF-1. Biochem J 386: 191-200.
67. Cottle DL, McGrath MJ, Cowling BS, Coghill ID, Brown S, et al. (2007) FHL3 binds MyoD and negatively regulates myotube formation. J Cell Sci 120: 1423-1435.
68. Meeson AP, Shi X, Alexander MS, Williams RS, Allen RE, et al. (2007) Sox15 and Fhl3 transcriptionally coactivate Foxk1 and regulate myogenic progenitor cells. EMBO J 26: 1902-1912.
69. Ding L, Wang Z, Yan J, Yang X, Liu A, et al. (2009) Human four-and-a-half LIM family members suppress tumor cell growth through a TGF-beta-like signaling pathway. J Clin Invest 119: 349-361.
70. Gerber SA, Pober JS (2008) IFN-alpha induces transcription of hypoxia-inducible factor-1alpha to inhibit proliferation of human endothelial cells. J Immunol 181: 1052-1062.
71. Luker KE, Pica CM, Schreiber RD, Piwnica-Worms D (2001) Overexpression of IRF9 confers resistance to antimicrotubule agents in breast cancer cells. Cancer Res 61: 6540-6547.

72. Kraus TA, Lau JF, Parisien J-P, Horvath CM (2003) A hybrid IRF9-STAT2 protein recapitulates interferon-stimulated gene expression and antiviral response. J Biol Chem 278: 13033-13038.
73. Fujii N, Hiraki A, Ikeda K, Ohmura Y, Nozaki I, et al. (2002) Expression of minor histocompatibility antigen, HA-1, in solid tumor cells. Transplantation 73: 1137-1141.
74. Klein CA, Wilke M, Pool J, Vermeulen C, Blokland E, et al. (2002) The hematopoietic system-specific minor histocompatibility antigen HA-1 shows aberrant expression in epithelial cancer cells. J Exp Med 196: 359-368.
75. Spierings E (2008) Minor histocompatibility antigens: targets for tumour therapy and transplant tolerance. Int J Immunogenet 35: 363-366.
76. Bolinger B, Krebs P, Tian Y, Engeler D, Scandella E, et al. (2008) Immunologic ignorance of vascular endothelial cells expressing minor histocompatibility antigen. Blood 111: 4588-4595.
77. Mosavi LK, Cammett TJ, Desrosiers DC, Peng Z-Y (2004) The ankyrin repeat as molecular architecture for protein recognition. Protein Sci 13: 1435-1448.
78. Ma Q, Wang H, Guo R, Wang H, Ge Y, et al. (2006) Molecular cloning and characterization of SRG-L, a novel mouse gene developmentally expressed in spermatogenic cells. Mol Reprod Dev 73: 1075-1083.
79. McClintock TS, Glasser CE, Bose SC, Bergman DA (2008) Tissue expression patterns identify mouse cilia genes. Physiol Genomics 32: 198-206.
80. Ma C, Staudt LM (1996) LAF-4 encodes a lymphoid nuclear protein with transactivation potential that is homologous to AF-4, the gene fused to MLL in t(4;11) leukemias. Blood 87: 734-745.
81. Hiwatari M, Taki T, Taketani T, Taniwaki M, Sugita K, et al. (2003) Fusion of an AF4-related gene, LAF4, to MLL in childhood acute lymphoblastic leukemia with t(2;11)(q11;q23). Oncogene 22: 2851-2855.
82. Wittwer F, van der Straten A, Keleman K, Dickson BJ, Hafen E (2001) Lilliputian: an AF4/FMR2-related protein that controls cell identity and cell growth. Development 128: 791-800.
83. To MD, Faseruk SA, Gokgoz N, Pinnaduwage D, Done SJ, et al. (2005) LAF-4 is aberrantly expressed in human breast cancer. Int J Cancer 115: 568-574.
84. Barton A, Eyre S, Ke X, Hinks A, Bowes J, et al. (2009) Identification of AF4/FMR2 family, member 3 (AFF3) as a novel rheumatoid arthritis susceptibility locus and confirmation of two further pan-autoimmune susceptibility genes. Hum Mol Genet 18: 2518-2522.
85. Pei Z, Jia Z, Watkins PA (2006) The second member of the human and murine bubblegum family is a testis- and brainstem-specific acyl-CoA synthetase. J Biol Chem 281: 6632-6641.
86. Zheng Y, Zhou Z-M, Min X, Li J-M, Sha J-H (2005) Identification and characterization of the BGR-like gene with a potential role in human testicular development/spermatogenesis. Asian J Androl 7: 21-32.
87. Fraisl P, Tanaka H, Forss-Petter S, Lassmann H, Nishimune Y, et al. (2006) A novel mammalian bubblegum-related acyl-CoA synthetase restricted to testes and possibly involved in spermatogenesis. Arch Biochem Biophys 451: 23-33.
88. Sakakibara T, Nemoto Y, Nukiwa T, Takeshima H (2004) Identification and characterization of a novel Rho GTPase activating protein implicated in receptor-mediated endocytosis. FEBS Lett 566: 294-300.

89. Katoh Y, Katoh M (2004) Identification and characterization of ARHGAP27 gene in silico. Int J Mol Med 14: 943-947.

Chapter 3. The Direct Genomic Targets of Etsrp/Etv2 in Zebrafish Hemato-Vascular Development

Abstract

Etsrp/Etv2 is a conserved master control gene for vascular and hematopoietic development. Although the importance of this transcription factor is well established in multiple organisms, the direct target genes within the genome are not. In order to identify these target genes, we have combined ChIP-Seq datasets with gain of function and loss of function RNA-Seq datasets generated in vivo using zebrafish embryos. Through genome wide analysis by ChIP-Seq technology, 10,029 genomic sites neighboring 6,950 genes were found to uniquely bind Etv2. For gain of function, we performed RNA-Seq analysis of embryos overexpressing Etv2 and identified 1907 genes induced over 2-fold, from which 752 genes are associated with ChIP-Seq peaks. For loss of function, we used RNA-Seq data from three cloche alleles that lack vasculature and are deficient in Etv2. This study resulted in the identification of 780 down-regulated genes from which 164 contained Etv2-ChIP peaks. Analysis of the intersection from the three datasets generated a list of 82 genes, likely representing the direct targets positively regulated by Etv2. We functionally validated six Etv2 binding sites at these gene loci as vascular-specific enhancers using transgenic zebrafish. We also examined genes repressed by Etv2 overexpression and identified 111 genes with ChIP-Seq peaks. Among them, 33 of 57 with documented expression are associated with myogenesis. This suggests that Etv2 not only activates vascular and hematopoietic genes but also actively represses myogenic cell fates. The results presented here broaden our understanding of the molecular mechanisms of Etv2 action in vascular and hematopoietic cell specification.

Introduction

Vertebrates that fail to develop a hematopoietic and vascular system die embryonically, so it is crucial to understand the gene regulatory network that direct the initial stages of vascular development. In the zebrafish, initial vasculogenesis occurs by the coalescence of angioblasts derived from the lateral plate mesoderm (Jin et al., 2005; Proulx et al., 2010). Through a previous transcriptome study of *cloche* mutants that are defective in the initiation of vasculogenesis and hematopoiesis (Stainier et al., 1995), we identified a critical transcription factor *ets related protein*, *etsrp*, that is necessary and sufficient for vasculogenesis, primitive myelopoiesis and hematopoietic stem cell development (Ren et al., 2010; Sumanas et al., 2008; Sumanas et al., 2005; Sumanas and Lin, 2006). Independently, *etsrp* was also identified by mapping of an ENU-induced mutant with vascular defects, and suggested to cooperate with other ETS family members for regulating vasculogenesis (Pham et al., 2007).

The mammalian ortholog of *etsrp*, initially named *Er71* when identified in the testis of adult mice (Brown and McKnight, 1992), shows syntenic association with zebrafish, and is able to ectopically induce the vasculogenic program when overexpressed in zebrafish (Sumanas et al., 2008). Er71 is also named *Ets variant 2 (Etv2)*. To remain consistent with the current nomenclature we hereafter refer to *etsrp*/*Etv2*/*Er71* as *etv2*/Etv2. *Etv2* null mice are embryonically lethal and fail to form vasculature and hematopoietic cells (Ferdous et al., 2009; Lee et al., 2008), while the Xenopus ortholog is also required during the initial stages of vasculogenesis (Neuhaus et al., 2010; Salanga et al., 2010), which provides further evidence for the functional conservation of this protein during evolution. Additionally, several molecular signaling pathways involving BMP, Notch, Wnt and VEGF have been implicated in *etv2* regulation (Kataoka et al., 2011; Lee et al., 2008; Rasmussen et al., 2012). Other regulators of

etv2 include CREB, which was found to interact with a cis binding element upstream of the transcription start site of *Etv2* (Yamamizu et al., 2012) and NKX2-5, which induces *Etv2* expression in endocardial cells (Ferdous et al., 2009). In zebrafish, the lateral plate mesoderm is split into separate bilateral anterior and posterior domains, and *etv2* expression in these tissues is regulated by distinct transcription factors. *Gata4*, *gata5* and *gata6* lie upstream of *etv2* in the anterior mesoderm (Peterkin et al., 2009) and *foxc1a*/*foxc1b* directly regulate *etv2* in the posterior mesoderm (Veldman and Lin, 2012).

Because the forced expression of Etv2 alone is sufficient to induce vasculogenic and myelopoietic genes in zebrafish embryos, we and others have carried out transcriptional profile studies of Etv2 overexpressing embryos by microarray and deep RNA sequencing approaches to more fully characterize the genes expressed in these tissues (Gomez et al., 2012; Gomez et al., 2009; Wong et al., 2009a). But those methods do not offer the means to identify the direct genetic targets of the Etv2 protein, which is highly desirable for understanding the early vasculogenic gene regulatory network. Nonetheless, Etv2 has been inferred to bind an evolutionarily conserved composite FOX:ETS motif which contains a non-canonical FOX DNA binding element bound by the mammalian transcription factor, Foxc2, that is directly adjacent to an ETS binding cis element, and is highly represented near endothelial genes (De Val et al., 2008). Evidence supporting the binding of this FOX:ETS site by Etv2 in zebrafish has been noted at a site near an endocardial enhancer for *nfatc1* (Palencia-Desai et al., 2011). In mammals, ETV2 has been demonstrated to directly induce genes including *Flk1* (Ishitobi et al., 2011; Lee et al., 2008), *Mef2c* (De Val et al., 2008), *Lmo2* (Koyano-Nakagawa et al., 2012), *Scl* (Kataoka et al., 2011; Wareing et al., 2012a), *Fli1* (Kataoka et al., 2011), *Gata2* (Kataoka et al., 2011), *Cdh5* (Liu et al., 2012) and *Tie2* (Ferdous et al., 2009; Lee et al., 2011). While in

Xenopus, Etv2 directly induces the endothelial genes *flk1* and *aplnr* as well as the hematopoietic genes, *pu1* and *runx1* (Salanga et al., 2010), which suggests that Etv2 can modulate the expression of many hemato-vascular genes.

As a master regulator, a comprehensive understanding of Etv2's direct targets will be highly useful for elucidating the genetic basis of vasculogenesis and hematopoiesis. Here we aimed to identify the comprehensive set of direct in vivo genomic binding targets of Etv2 in zebrafish embryos by Chromatin Immunoprecipitation followed by high throughput sequencing (ChIP-Seq) through a gain of function approach with a stable transgenic line with inducible expression of an epitope tagged Etv2 protein. Since overexpression is likely to overestimate the true Etv2 genomic binding sites, the dataset was narrowed down to targets near biologically relevant genes by subtraction with complementary transcriptome profiles generated from embryos with Etv2 gain and loss of function by RNA-Seq. In addition to the identification of directly induced genes, Etv2 was also associated with the direct down-regulation of genes associated with skeletal muscle and neurons. These results clarify the molecular mechanisms by which Etv2 specifies vascular and hematopoietic lineages in zebrafish embryos.

Methods

ChIP-Seq

Binding of the mCherry antibody (rabbit anti-DS Red polyclonal, Clontech) to Etv2-mCherry was tested by whole mount immuno-fluorescence before performing ChIP-Seq experiments on progeny from *hsp70l:etv2-mCherry* heterozygous transgenic zebrafish (Veldman et al., 2013) out-crossed to wild type, producing half a clutch with the inducible *etv2-mCherry* transgene and

half without. The primary antibody was tested at 1:500 followed by 1:1000 of the Alexa Fluor 488 anti-rabbit secondary IgG (A11008 Life Technologies).

For ChIP, homozygous heat shock *hsp70l:etv2-mCherry* fish were out-crossed with wild-type fish, exclusively producing embryos with ubiquitous heat inducible C-terminal mCherry tagged Etv2, and equal numbers of wild type embryos were processed in parallel as controls. Transgenic expression was stimulated by floating embryos in Petri dishes at 38.5° C for 45 minutes, followed by incubation for 1 hour at 28.5° C before harvesting. ChIP was performed with the Protein A ChIP kit (Millipore) according to the manufacturer's instructions with the following modifications prior to sonication. Cells from 200 embryos per ChIP were first de-yolked with Ringer's solution containing 0.5M EDTA with 2 cycles of centrifugation at 1,000g between changes of Ringer's buffer before fixation in 1% formaldehyde for 10 minutes. Samples were quenched with 0.125M Glycine, followed by 3 washes in PBS with protease inhibitors (Roche) and stored at -80°C. Cells were lysed in cell lysis buffer (10mM Hepes pH7.9, 0.5% NP40, 1.5mM MgCl2, 10mM KCl, 0.5mM DTT), followed by a high salt wash (50mM Hepes pH7.9, 0.1% SDS, 1% Triton X-100, 0.1% deoxycholate, 1mM EDTA, 500mM NaCl), and a final lysis in nuclear lysis buffer (1% SDS, 50mM Tris-HCl pH8, 20mM EDTA). Each lysis step was performed by passing samples through a 200 µl pipette tip and a 10 minute incubation on ice followed by a 10 minute centrifugation step (2,000g) at 4°C. Thereafter, lysates were sonicated with a Bioruptor (Diagenode) at high frequency setting for 2 cycles, each cycle lasting 10 minutes with each minute divided into 30 seconds of sonication followed by 30 seconds of rest. Ice was changed between cycles to prevent overheating and denaturation of samples, producing a final sheared protein bound-DNA range of 200-500bp. The anti-DS Red

polyclonal antibody (Clontech) was used at 5 ug per ChIP. A total of 1200 embryos were used for each ChIP-Seq library. 2ng of each sample was used to generate each library with the Ovation Ultralow IL Multiplex System (Nugen) according to manufacturer's instructions. The final libraries were gel extracted with mini-elute columns (QIAGEN) at a range of 300-550 bp. Libraries were processed on an Illumina Hi-seq 2000 sequencer to generate single end 50bp reads.

Each ChIP sample was resuspended in 40µl elution buffer (QIAGEN) after chloroform extraction, and 2µl of each was tested by qPCR with SYBR green (Roche) using the primers listed in Supplementary Methods Table 1. Enrichment was determined by the ΔΔ Ct method and standardized with *rhodopsin* primers that were published previously (Wardle et al., 2006).

RNA-Seq

For Etv2 overexpression, pools of 100 embryos injected at the one cell stage with Etv2 RNA or uninjected controls were collected at 90% epiboly for RNA-Seq analysis. Similarly, 100 embryos of either wild type controls or *hsp70l-etv2-mCherry* embryos were heat shocked for 45 minutes, and re-incubated for 6 hours at 28°C followed by RNA-Seq analysis. *Cloche* alleles *m39* plus *m378* containing the *gata1*-GFP transgene and *la1164* containing the *flk1*-GFP transgene were all manually sorted by the absence of transgenic expression in either blood or angioblasts at 15 hpf, together with stage matched wild type controls. Two wild type datasets were generated. One was collected together with and used to standardize *la1164*, while the other was collected and processed along with *m39* and *m378*. Total RNA for all samples were harvested with TRIzol (Life Technologies), and further purified with the RNeasy kit (QIAGEN).

RNA-Seq libraries were constructed with the TRUseq preparation kit according to manufacturer's instructions (Illumina).

qPCR of hsp70l-etv2-mCherry samples before RNA-Seq library preparation (Supp Fig 3) was performed as described (Ren et al., 2010) with primers for *scl*: (F) GGAGATGCGGAACAGTATGG (R) GAAGGCACCGTTCACATTCT and *fli1a*: (F) CCGAGGTCCTGCTCTCACAT (R) GGGACTGGTCAGCGTGAGAT

Enhancer assay

Genomic targets of Etv2-mCherry and negative control regions were amplified from wild type genomic DNA using the primers listed on Supplementary Methods Table 2. Both the forward and reverse primers were preceded by a 6 bp leader sequence, followed by a Bam H1 restriction site (forward primer) and a Sal I site (reverse primer). Restriction sites straddle the genome specific complementary sequences. Amplicons were amplified with hot start HiFi Taq Polymerase (Roche), cloned into the pCR4 Topo TA vector (Invitrogen) and sequenced with M13F primers to verify cloning of intended target sequences. Cloned inserts were double digested with Bam H1 and Sal I (Schoenebeck et al.) and ligated into a Tol2 vector upstream of the *gata2* minimal promoter and eGFP (Veldman and Lin, 2012). To generate transient and germline transgenic fish, 50pg plasmid DNA and 20pg tol2 RNA was injected into single cell zebrafish embryos.

For enhancer mutagenesis, the plasmid containing the enhancer, *fli1a. B,* was amplified with primers (L) 5'-ATACCCATAGATTGAGACACCCT-3' and (R) 5'-GCCCTCTGTATAAAATAGCACAGC-3' in opposite directions of the ETS binding sites with

the Phusion HiFi polymerase (Schoenebeck et al.). The amplicon was purified with a gel extraction mini-elute column (QIAGEN), then ligated overnight with T4 ligase (Schoenebeck et al.), deleting the ETS binding sites and thereby generating the construct *fli1a. Bdel*.

Images were captured on a digital CCD camera (Axiocam, Zeiss) mounted on an upright microscope (Axioskop2 plus, Zeiss), with Openlab 4.0.2 software (Imprvision, Lexington, MA). Images were processed to generate figures on Adobe Photoshop CS5. Schematics were assembled in Adobe Illustrator CS5. RNA-Seq and ChIP-Seq binding profiles (WIG file images) were generated with the UCSC genome browser and annotated on Illustrator and Photoshop.

Whole Mount In Situ Hybridization (WISH)

WISH was performed as described (Gomez et al., 2009). Probes were cloned with the following primers for *wnt3a:* (F) AGGGAGGGAATCAGAGGACG (R) GTGGCATTTCTCTTTGCGCT and *gsx1*: (F) AGGGAGAACAGCGAAAAGGG (R) GGTGTGGCGTACAGAGTCTT.

Bioinformatics

Detailed description of data processing is attached in the Supplemental Bioinformatics Methods file.

Results

I. Etv2 binds 10,029 sites in the zebrafish genome

We raised zebrafish specific Etv2 antibodies twice for the purpose of identifying the direct genomic targets of Etv2 in zebrafish embryos by ChIP-Seq but to no avail. As an

alternative, we performed ChIP-Seq for Etv2 from embryos of a stable transgenic zebrafish line *hsp70l*:*etv2-mCherry* (Veldman et al., 2013) in which the temporal induction of Etv2 can be carefully regulated, and the ubiquitously induced Etv2 protein can be specifically detected with a commercially available antibody for mCherry (Supplementary Fig 1). Pools of wild type control or *hsp70l*:*etv2-mCherry* embryos were heat shocked and processed for ChIP at 15 hours post fertilization, a stage when the endogenous protein is highly transcribed and likely active in angioblasts (Fig 1A). Following bioinformatics processing, relative to the ChIP-Seq profile of the wild type group, *etv2-mcherry* was predicted to bind 10,029 genomic sites neighboring 6,950 genes (Supplementary Table 1). Gene Ontology (GO) analysis of the dataset with the ChIPpeakAnno package (Zhu et al., 2010a) identified genes associated with vascular development and blood vessel development as highly significantly enriched categories (Supplementary Table 2), suggesting that the dataset indeed contains biologically relevant genomic occupancy information. Moreover, when compared with previously published ChIP-Seq maps of the epigenetic enhancer mark, H3K4me1, and regions of actively transcribed chromatin, H3K4me3 (Aday et al., 2011), the Etv2-mCherry dataset was found to preferentially overlap with enhancer sites at a higher rate (53.6%) than with actively transcribed DNA (36.6%), as expected of a transcription factor (Fig 1B and Supplementary Table 3). The most statistically significant enriched motif in the dataset contained the conserved core ETS binding site 5'-GGAA-3' as previously observed for the mammalian ortholog (Brown and McKnight, 1992) (Fig 1C). Globally, Etv2 bound ChIP-Seq peaks are distributed throughout the genome with some genes containing several peaks of enrichment. While peak distribution is not confined to a particular location relative to the transcription start site of a gene, a frequency distribution plot of genomic binding by Etv2 results in a bimodal binding distribution at the Transcription Start

Sites, TSS, of genes (Fig 1D), similar to the binding profiles observed by other transcription factors (Soleimani et al., 2012). Altogether, these observations of the comprehensive dataset suggest that it reflects authentic genomic occupancy information for Etv2.

The dataset was evaluated by independent ChIP-qPCR assays on stage and condition matched pools of embryos at seven predicted sites bound by Etv2 with a randomly selected unbound negative control region per selected peak, as demonstrated for a pair near *fli1a* (Fig 2A). Binding at all sites was normalized to a region in the promoter of *rhodopsin*, *rho*, which is not expressed at the stages examined, or according to the ChIP-Seq data, bound by Etv2-mCherry. Compared to wild types, we found that the Etv2-mCherry group occupies all 7 ChIP-Seq binding sites examined, at statistically significant levels compared with no enrichment at negative control sites (Fig 2B). Five of the Etv2 binding sites tested (near *scl*, *mefcb*, *notch1a*, *arid1ab*, *fli1a*) contain the FOX:ETS co-binding site, which provides further evidence in support of the notion that such an evolutionarily conserved motif is indeed bound by Etv2 in vivo, and our identification of this motif further supports the validity of the dataset obtained here. However, we noted that this FOX:ETS site is not present at all peaks within the vicinity of genes expressed in endothelial cells, and by examining the expression of various genes at ZFIN (Thisse, 2004), we found that Etv2 binds next to genes that are never expressed in endothelial cells. This is a reasonable result, since Etv2 is normally not expressed ubiquitously or at such elevated levels, which may result in ectopic binding to DNA where the endogenous proteins is not likely to bind in angioblasts. For example, another ETS family protein, Etv4/Pea3, is expressed in the tail bud, midbrain-hindbrain boundary and other neural tissues around 15 hpf and directly activates *dusp6* by binding directly upstream of its start site (Znosko et al., 2010). In our data there is a prominent ChIP-Seq peak at the exact location where Etv4/Pea3 binds, which

suggests that the high concentration of Etv2-mCherry protein can result in ectopic binding to the promoters and enhancers bound by other ETS family members. Regardless, the independent validation of the ChIP-Seq data suggests that the ChIP-Seq dataset offers a reliable map of genomic binding by Etv2. In particular, the validation of Etv2 binding near *fli1a* and *scl* is of biological significance since these two genes are well known transcriptional regulators of the vascular and hematopoietic lineages.

II. Etv2 ChIP-seq identifies vascular enhancer elements.

To examine whether any of the genomic Etv2 targets are bona fide vascular enhancers we randomly selected 10 sites bound by Etv2 near 8 genes that we previously found to be downstream of *etv2,* namely *fgd5*, *she*, *cdh5*, *arhgap31*, *sox18*, *fli1a*, *yrk* and *crip2,* to test for enhancer activity in a transient transgenic assay in zebrafish larvae. An average of 500 bases surrounding the predicted Etv2 bound peaks were cloned into an enhancer reporter plasmid, as exemplified for a binding site within an intron of *she* (Fig 3A). Two unbound negative control regions with the same average length and an empty vector were also analyzed. The enhancer tested for *she* resulted in 35 of 40 (87%) embryos with definite expression in the vasculature (Fig 3B). In a similar fashion, we noted a clear vascular expression pattern in 6 of the Etv2 targets examined (red bars, Fig 3C), while the negative controls, and predicted Etv2-mCherry binding sites not expressed in the vasculature displayed random mosaic expression patterns, as illustrated by embryos injected with the empty vector (lower panel in Fig 3B). The only exception was the *crip2* negative control, which overlaps with an H3K4me1 enhancer peak, and in which over 95% of the embryos display marked GFP reporter expression in cardiomyocytes (Supplementary Fig 2). Although not all Etv2 ChIP-seq binding sites tested are functional in the vasculature, we did

not test every locus bound by Etv2 for the genes tested. However the specificity observed in those that are active in the vasculature argue that this dataset offers a useful approximation of functional vascular ets box enhancer cis elements.

Although there is controversy over whether ETV2 directly regulates *fli1* in mice (Kataoka et al., 2011; Wareing et al., 2012b), we have confirmed the direct binding of Etv2 within the *fli1a* locus in zebrafish, and generated transgenic zebrafish with one of the two enhancers studied (labeled *fli1a. B* in Fig 3C). To demonstrate that Etv2 or other ETS factors are involved in the regulation of this enhancer, the ETS binding sites were deleted from this construct, *fli1a. Bdel*, which resulted in the elimination of the vascular specific expression noted with the full-length enhancer (Fig 3D). Overall the enhancer data validates the ChIP-Seq dataset by demonstrating that some of the sites bound by Etv2-mCherry are active enhancers in vivo.

III. Combining Etv2 gain and loss of function RNA-Seq with ChIP-Seq datasets identify direct Etv2 targets in vivo

In order to identify the biologically relevant direct targets of Etv2 within the ChIP-Seq dataset, we compared the transcriptional profiles of RNA-Seq from embryos with gain- and loss-of-function of Etv2. For gain of function, we examined Etv2 overexpression at two developmental stages, 90% epiboly after injection of Etv2 mRNA at one cell stage, and 6 hours post heat activation of *hsp70l*:*etv2-mCherry* at 15hpf, since this was determined to be the stage of maximal induction of *fli1a* in a time-course measurement by qPCR (Supplementary Fig 3). This resulted in 1,029 genes induced above a 2-fold cutoff at 90% epiboly (Supplementary Table 4, Sheet 1), and 959 induced above 2-fold by heat shock (Supplementary Table 4, Sheet 2). Less than 15% (138) of the genes overlapped between these two datasets, which is most likely due to

the differences in developmental stages. ChIP-Seq peaks were compared with the sum of genes induced at both time points (1907), resulting in 752 genes induced within the ChIP-Seq dataset (Supplementary Table 4, Sheet 3).

For loss of function we examined the RNA-Seq profiles of embryos deficient in Etv2. Since it is possible that other ETS domain containing transcription factors expressed in the zebrafish vasculature might compensate for the absence of Etv2 either in morphant or mutant embryos (Pham et al., 2007), we examined the transcript profiles of *cloche* mutants that are devoid of all hemangioblast lineages and therefore all vascular specific ETS factors (Liu and Patient, 2008). Three different *cloche* alleles, *m39*, *m378*, and *la1164* were independently sorted by the absence of either *gata1:EGFP* or *kdrl:EGFP* transgenes at the same stage when *hsp70l*:*etv2-mCherry* embryos were heat shocked for ChIP-Seq, 15 hpf, and processed for RNA-Seq profiling. Relative to stage matched wild-type controls, this resulted in the down-regulation of 403 genes in *m39*, 453 genes in *m378*, and 190 genes in *la1164*, with *etv2* being reduced in all three alleles, as expected (Supplementary Fig 4). Interestingly, although all three alleles are deficient in the vascular and hematopoietic lineages, the datasets generated in all 3 mutants are not completely overlapping. Nevertheless, the comparison between the sum of genes reduced in all three *cloche* alleles (780) with the Etv2 ChIP-Seq dataset resulted in 164 of these containing ChIP peaks (Supplementary Table 4, Sheet 4). Since Etv2 is necessary and sufficient for the expression of endothelial and primitive myeloid genes, the genes that meet both of these criteria in our RNA-Seq datasets are most likely directly induced by Etv2. We therefore examined the intersection between the genes reduced greater than 2-fold in all three *cloche* mutants, with those induced greater than 2-fold by Etv2 overexpression generated in this study resulting in 82 genes being directly induced by Etv2. In addition, we considered vascular and myeloid expressed

genes that we previously identified through microarray of Etv2 overexpression (Gomez et al., 2009). All together, this resulted in a total of 101 genes (Fig 4A and Table 1), which are bound by Etv2-mCherry at 204 locations (Supplementary Table 4, Sheet 5). An example of a gene in this list is *she*, wherein RNA-Seq profiles show reduction in *cloche*, induction by Etv2-mCherry after heat shock, and a ChIP-Seq peak in the Etv2-mCherry group but not wild type controls (Fig 4B). Furthermore, by whole mount in situ hybridization (WISH), ectopic induction is apparent in Etv2-mCherry embryos following heat shock, while vascular expression is severely reduced in Etv2 morphants (Fig 4C). The list of direct targets was then examined by Database for Annotation Visualization and Integrated Discovery Gene Ontology (DAVID GO) analysis (Huang et al., 2009a), resulting in the top category obtained belonging to genes associates with vascular development, and contains many well characterized and studied vasculogenic molecular players that includes transcription factors, adaptor proteins and receptors.

IV. Forced expression of Etv2 directly represses genes expressed in skeletal muscle and neuronal cells

In addition to the activator function ascribed to Etv2 in the vasculogenic program, its propensity to facilitate gene repression has also been demonstrated in cardiomyocytes (Liu et al., 2012; Palencia-Desai et al., 2011; Rasmussen et al., 2011; Schoenebeck et al., 2007). Moreover, a cohort of genes associated with muscle lineages are upregulated in *Etv2* null cells in mice (Kataoka et al., 2011; Rasmussen et al., 2011), while the forced expression of Etv2 in zebrafish results in the repression of several skeletal myogenic genes and the transformation of skeletal myocytes to functional vascular endothelial cells (Veldman et al., 2013). We previously observed the broad repression of genes expressed in all three germ layers in Etv2 overexpression

at 80% epiboly following the ectopic expression of Etv2 from the one cell stage (Gomez et al., 2012), and we obtained the same results here with RNA-Seq data from embryos obtained with similar experimental conditions at 90% epiboly. Ectopic Etv2 induction at 6 hours post heat shock (hphs) at 15 hpf results in the repression of 798 genes relative to wild type controls (Supplementary Table 5, Sheet 1), with a preponderant enrichment of genes associated with cardiomyocytes, skeletal myocytes and neurons. Of the 798 genes, 111 contain ChIP-Seq peaks, and although the expression patterns of only 57 of these have been documented at the developmental stages studied here (between 15 and 21 hpf), 24 are expressed in somites, 19 are expressed in the nervous system, and 9 are expressed in both somites and the nervous system (Fig 5A and Supplementary Table 5, Sheet 2). *Myod*, *myog, myf5, mylpfa* and *tnnt3a* are part of myogenic genes reduced by Etv2-mCherry following heat shock, as shown in Veldman et al., 2013. Notably, the master regulator of myogenesis, *myod*, and a component of the SWI/SNF chromatin modifying complex postulated to modulate myogenenesis after initiation, *smarcd3a* (Ochi et al., 2008), are repressed myogenic genes associated with ChIP-Seq peaks.

Since the repression of neural genes by the forced expression of Etv2 has not been investigated, we examined the expression of *wnt3a* and *gsx1*, which are putative direct targets of Etv2, following heat shock by WISH. Compared to wild type embryos, the expression of *wnt3a* is severely reduced in the midbrain (Fig 5B), while both *wnt3a* and *gsx1* are nearly abolished in the hindbrain. Concomitant with this repression, the normally vascular restricted expression of *she* in the head of wild types is ectopically induced throughout the forebrain, midbrain, and hindbrain following Etv2-mCherry overexpression. Despite the similar effects observed between the myogenic and neural lineages resulting from forced Etv2-mCherry expression, we have not detected the transformation of neuronal cells to functional vascular endothelial cells as observed

in skeletal myocytes (Veldman et al., 2013), but the activation of vascular genes in the neural cells by forced induction of Etv2 can potentially provide insights into clinically relevant diseases like cancer.

Discussion

The critical requirement of Etv2 for the specification of hemangioblasts in the lateral plate mesoderm is undisputable, but the mechanism remains incompletely understood. In mice ETV2 has been proposed to initiate vasculogenesis by directly inducing a core hemangioblast network consisting of *Scl*, *Fli1* and *Gata2*, with the sole re-introduction of *Scl* or *Fli1* being able to restore both vasculogenic and hematopoietic programs in *Etv2* null embryonic stem cells (Kataoka et al., 2011). While others found that only *Scl* is able to rescue the hematopoietic program, and establish the recursive loop with *Fli1* and *Gata2* (Wareing et al., 2012b), perhaps indicating that ETV2 initiates distinct networks in cells fated to become endothelial cells than those fated for the hematopoietic fate. We previously obtained analogous findings in zebrafish, whereby both *scl* and *fli1a* are individually able to partially rescue endothelial specification and subsequent vasculogenesis in Etv2 deficient embryos (Ren et al., 2010). *Scl* and *fli1a* are direct targets of Etv2 in our dataset, and while *gata2a* is both induced by Etv2 and associated with a ChIP-Seq peak, by RNA-seq it is not reduced below 2-fold in any of the *cloche* mutants at the stages examined, but is apparently reduced in *etv2* mutants at later developmental stages (Pham et al., 2007). Therefore, our data indicates that in zebrafish Etv2 directly induces only two of three components of the feed-forward loop, as well as other hemato-vascular transcription factors like the paralog of *fli1a*, *fli1b*, the co-factor of *scl*, *lmo2*, as well as *sox7* and *sox18* which are sox factors required redundantly for artery-vein specification (Cermenati et al., 2008).

The genetic relationship between Etv2 and the VEGF receptor, *Kinase insert domain receptor/Fetal liver kinase1*, *Kdr/Flk1* has also been contended. *Etv2* has been proposed to lie upstream of *Flk1* because FLK1 is severely reduced in *Etv2* null mice, a cis-regulatory element upstream of *Flk1* is directly bound by ETV2 (Lee et al., 2008), and the conditional deletion of *Etv2* in *Flk1* cells results in viable mice (Wareing et al., 2012b). However, independent studies have revealed a rather mild reduction of FLK1 expression in *Etv2* null embryos, and instead shown that VEGF signaling induces *Etv2* through FLK1 to specify the hemangioblasts from primitive mesoderm (Kataoka et al., 2011; Rasmussen et al., 2012). In zebrafish, there are two orthologs of the mammalian *Flk1* resulting from a genome duplication event, *kdr* and *kdrl*, where the sequence of the former has a higher similarity to the mammalian *Flk1*(Bussmann et al., 2008), but the latter appears to play a more highly conserved biological role (Covassin et al., 2006). However, both paralogs are required for artery but not vein formation, since the chemical inhibition of all VEGF signaling results in an artery specific deficit coupled with an expanded vein (Covassin et al., 2006). Similarly, VEGF knockdown depletes *etv2* expression in the dorsal aorta and intersegmental vessels but it continues to be expressed in the vein, which appears expanded (Sumanas and Lin, 2006). The persistent expression of vascular markers after VEGF signal inhibition indicates that in zebrafish the initial expression of *etv2* might be independent of VEGF signaling, but this possibility cannot be completely excluded because expression data has revealed that *kdr* is expressed before *etv2* (Bussmann et al., 2007). Nevertheless, the induction of *kdrl* by the forced expression of VEGF is unsustainable in the absence of *etv2* (Sumanas and Lin, 2006), arguing that Etv2 directly induces VEGF receptors, and that the VEGF-Etv2-Kdr/Kdrl feedback loop is conserved between mice and zebrafish. Indeed, *Kdrl* was identified as

a direct target in this study, while *kdr* only failed to be included because it is only mildly reduced in all 3 *cloche* alleles by RNA-seq at the stages examined.

The role of Etv2 in hematopoietic development varies across the model organisms studied to date. Etv2 deficiency is dispensable for hematopoiesis in Xenopus (Neuhaus et al., 2010; Salanga et al., 2010), dispensable for primitive erythropoiesis but required for primitive myelopoiesis and definitive hematopoiesis in zebrafish (Ren et al., 2010; Sumanas et al., 2008), and required for the full spectrum of hematopoietic development in mouse embryos (Kataoka et al., 2011; Koyano-Nakagawa et al., 2012; Lee et al., 2008) and hematopoietic stem cell (HSC) maintenance in adult mice (Lee et al., 2011). In previous studies, we have found that overexpression of Etv2 is sufficient to induce several myelopoietic genes including the transcription factors *spi1*/*pu1* and *krml2* (Gomez et al., 2012; Gomez et al., 2009; Sumanas et al., 2008). Although in Xenopus the overexpression of Etv2 might directly induce *pu1* (Salanga et al., 2010), we did not see a ChIP-Seq peak associated with either of the two paralogs of *pu1* in zebrafish, *spi1* or *spi1l*, arguing that in zebrafish Etv2 only regulates myelopiesis indirectly, most likely by directly inducing *scl*. Both spatiotemporal expression comparisons between *etv2*, *scl* and *pu1*, as well as epistasis experiments have previously suggested this to be the case (Sumanas et al., 2008). We did note the direct induction of a couple of genes expressed in myeloid cells, *krml2*/*mafbb* and *myo1f*, but these genes are also expressed in the vasculature, so the ChIP-Seq peaks associated with them might reflect their induction by Etv2 in the vascular and not the myeloid lineage. Nevertheless, Etv2 overexpression is sufficient to induce both isoforms of *scl* expressed in zebrafish (Ren et al., 2010), and as with myelopoiesis, Etv2 is likely to direct the definitive hematopoietic program by directly inducing *scl* (Qian et al., 2007). We have also observed a ChIP-Seq peak near *runx1,* which is essential for definitive hematopoiesis (Burns et

al., 2005; Kalev-Zylinska et al., 2002), at a region distinct from the promoter used in a transgenic line previously used to lineage trace the emergence of definitive HSCs from the hemogenic endothelium in the dorsal aorta (Lam et al., 2010). According to the transcriptome profiles generated here, Etv2 overexpression is not sufficient to induce *runx1* expression, however the requirement of Etv2 for the formation of endothelial cells, plus the report that in Xenopus Etv2 might directly activate *runx1* (Salanga et al., 2010) suggest that the binding of Etv2 to *runx1* in zebrafish might reflect the activation of *runx1* expression by Etv2 in vivo.

The *hsp70l-etv2-mCherry* zebrafish line used here was also instrumental in the discovery that the ectopic induction of Etv2 can transform skeletal myocytes to functional endothelial cells (Veldman et al., 2013). Although this might occur by several means such as competition for myogenic co-factors, induction of myogenic microRNAs, alteration of epigenetic landscapes, this study presents the complete catalog of genes down-regulated by Etv2 overexpression and we detected an Etv2 ChIP-Seq peak within *myod*, suggesting that one possible mechanism involved in the myogenic transformation may be the direct downregulation of the master regulator of myogenesis. Among the set of genes down-regulated by Etv2, the histone modifying *smarcd3a* is also an interesting target because it has been proposed to maintain the skeletal myogenic program by modulating the access of *myod* to myogenic genes at the chromatin level (Ochi et al., 2008). It is also interesting that a large portion of the genes down-regulated by Etv2 are expressed in neurons. One possible explanation for the observed repression in these cells is that ETS domain family transcriptional repressor(s) might be expressed specifically in the myogenic and/or neural lineages in zebrafish, and the forced expression of etv2 in these cells can bind genomic repressor elements bound by those repressors, but we did not identify any such factors by searching ZFIN or the literature. Although we do not see the transformation of neurons to

functional vascular endothelial cells, we have observed the ectopic induction of vascular genes in those cells. Although the work presented here carries the caveats that the data was obtained by forced ubiquitous expression of Etv2, which results in binding to sites not normally bound by Etv2, the transcriptome profiles and GO analysis of these profiles do reflect the biologically relevant hemato-vascular fate. Therefore, these results provide an accurate prediction of the sites bound by Etv2 within the embryo and demonstrate that as in mice, in zebrafish, Etv2 initiates the vascular fate by inducing a key transcriptional network loop as well as other cellular proteins including receptors, adaptors, etc. that define angioblasts and certain hematopoietic lineages.

Figures

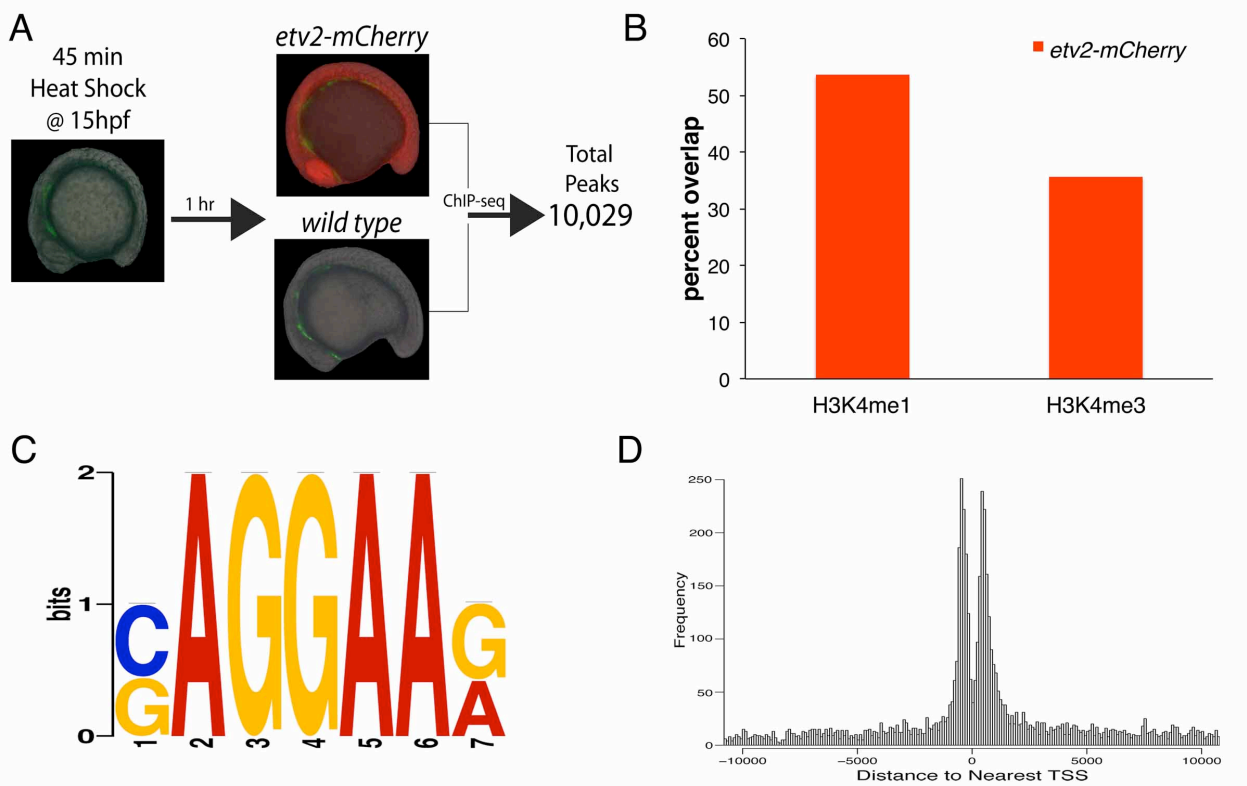


Figure 1: Etv2 ChIP-Seq in vivo. (A) Both *hsp70l-etv2-mCherry* and wild type embryos were heat shocked with the conditions indicated in the schematic and harvested for ChIP-Seq at 1 hphs when mCherry expression (Soleimani et al.) is observed in the *hsp70l-etv2-mCherry* group. Relative to chromatin immunoprecipitated from wild type embryos 10,029 peaks were enriched in the *etv2-mCherry* group. (B) *etv2-mCherry* ChIP-Seq genomic occupancy data overlap with enhancers (H3K4me1) and areas of actively transcribed DNA (H3K4me3). (C) The most statistically significant position weight matrix logo for *etv2-mCherry* binding across the genome (e-value <3.2e-1120) is identical to the ETS consensus binding site. (D) Frequency distribution plot of *etv2-mCherry* ChIP-Seq peaks relative to the transcription start site of all genes bound by *etv2-mCherry*.

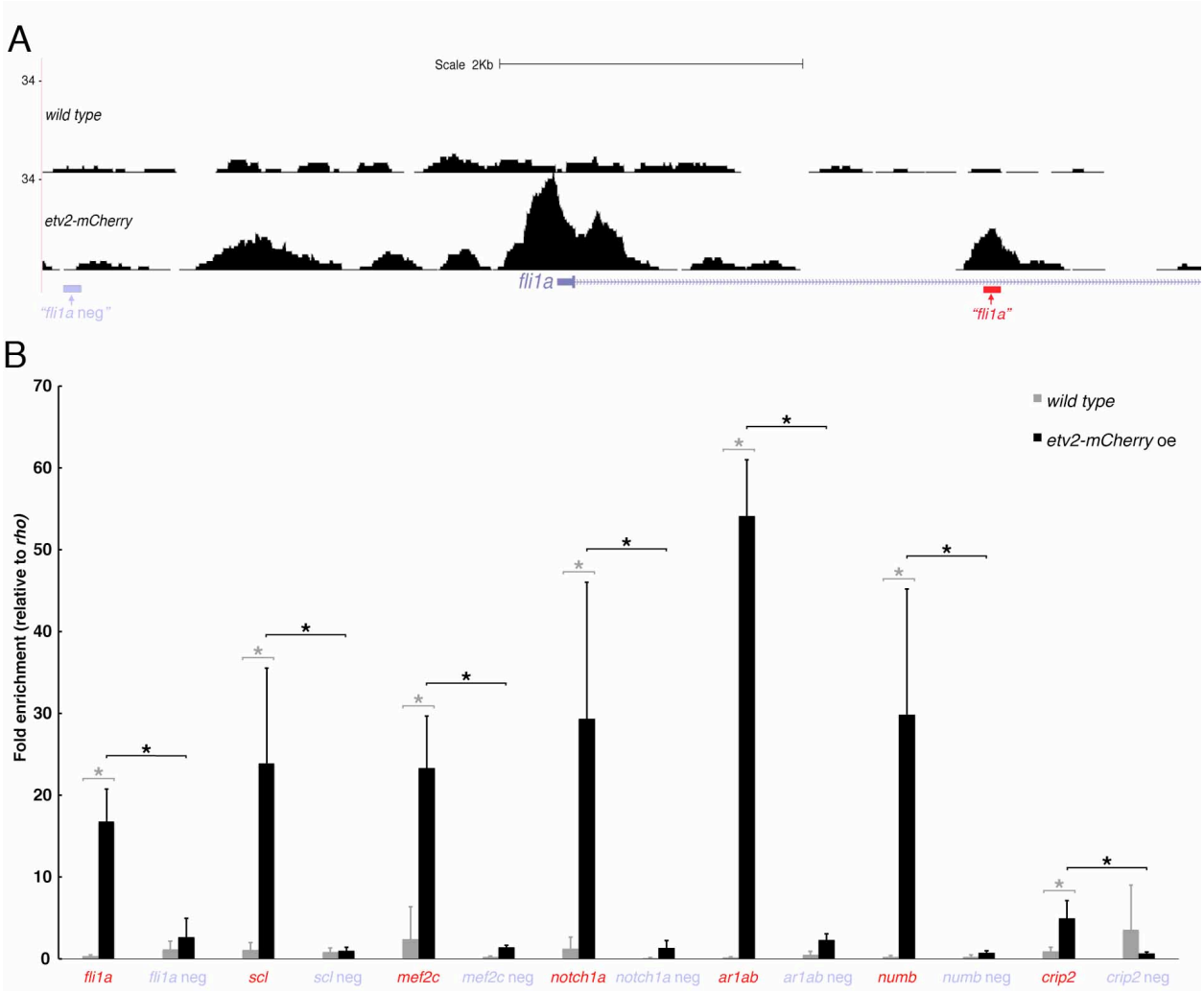


Figure 2. *etv2-mCherry* ChIP-Seq validation. (A) ChIP-Seq binding profiles of wild type control (top track) and *etv2-mCherry* (bottom track) at the *fli1a* locus. Validation was performed at the two areas marked in the gene map where *fli1a* qPCR primers bind at the region marked by the red rectangle that is positive for binding, and *fli1a* negative primers amplify the unbound region marked by the lavender rectangle. (B) ChIP-qPCR was performed at 7 enrichment sites predicted by the ChIP-Seq data with 7 neighboring unbound negative control sites tested per gene, as indicated for the *fli1a* pair in panel A. Binding at all sites was normalized by binding to the promoter of *rhodopsin*, *rho*. The results obtained for *fli1a* at the sites pointed to in panel A are included as the first pair in the bar graph. Gray and black asterisks indicate statistical significance of enrichment at $p<0.05$ (*t-test*), with the gray asterisk comparing binding near a given gene between Etv2-mCherry and wild type embryos, and the black asterisk comparing Etv2-mCherry binding between the predicted binding site and a negative control site.

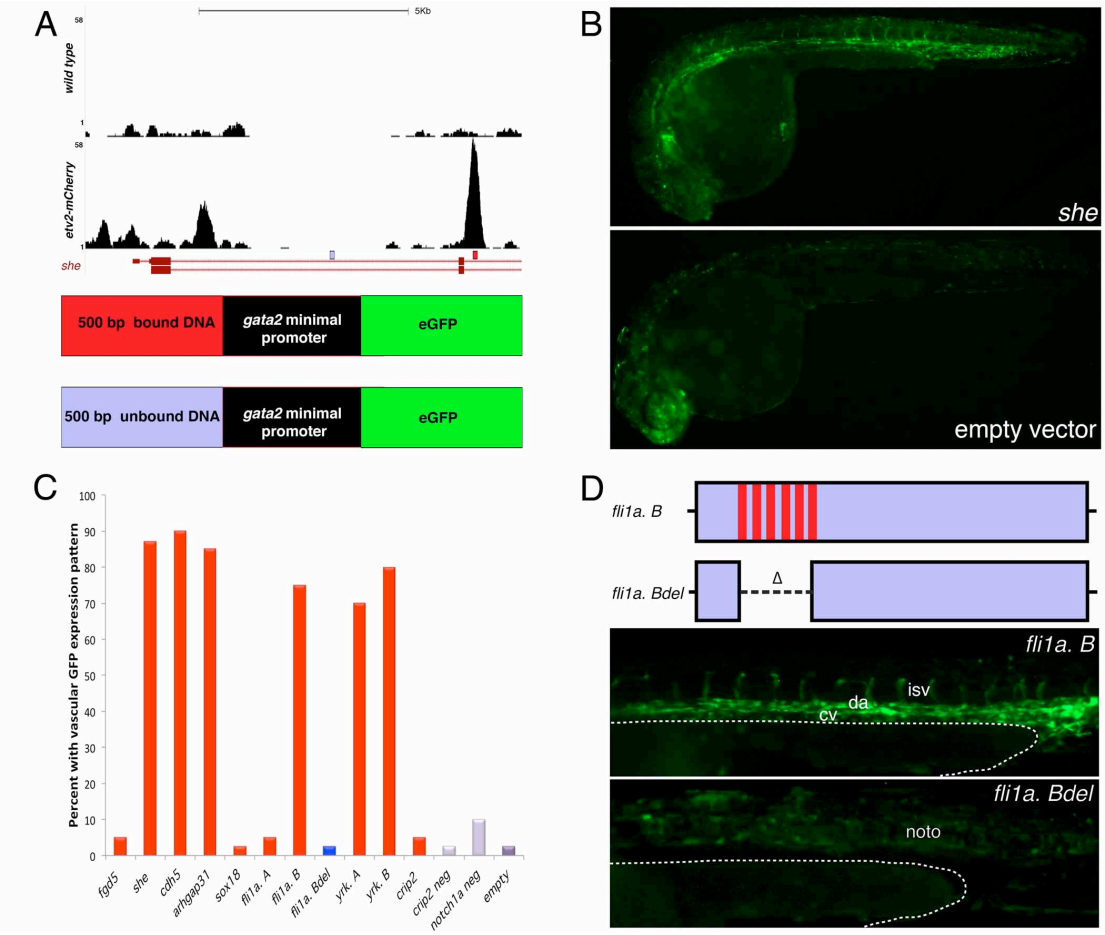


Figure 3. Etv2 binding sites function as vascular enhancer elements. (A) Representative schematic of enhancer assay at *she*. Approximately 500 bp spanning Etv2-mCherry ChIP seq peaks (red rectangle) or sample of signals at negative control regions (lavender rectangle) were cloned upstream of the *gata2* minimal promoter followed by eGFP in a Tol2 transgene plasmid. (B) Representative images of eGFP reporter expression in embryos with enhancers expressed in developing vascular endothelial cells as demonstrated by the enhancer for *she* (top panel), and the random mosaic expression observed in negative controls, as demonstrated by the empty vector (bottom panel). (C) Vascular specific eGFP expression is observed in 6 of 10 Etv2 bound regions tested (red bars). Two separate peaks were tested near *fli1a* and *yrk*, these are labeled with capitals as A and B, and fli1a. B corresponds to the site of Etv2 enrichment in Fig 2A. Negative control regions are represented by lavender bars, and the mutagenized *fli1A. Bdel* site, placed next to the full-length *fli1a. B,* is marked by the blue bar. A sample size of 40 embryos was examined for all constructs. (D) Schematic of *fli1a. B* enhancer with red vertical bars representing the 6 ETS binding sites within the cloned enhancer, which were deleted in *fli1a. Bdel*. Images are of representative embryos injected with these constructs. The expression of the *fli1a. Bdel* construct resulted in random mosaic expression without a predominant expression pattern. The embryo in the lower panel shows expression in the notochord, noto. All embryos were imaged laterally with anterior facing left, and the white dashed line in D outlines the yolk extension. Abbreviations: isv, intersegmental vessels; da, dorsal aorta; cv, caudal vein; noto, notocord.

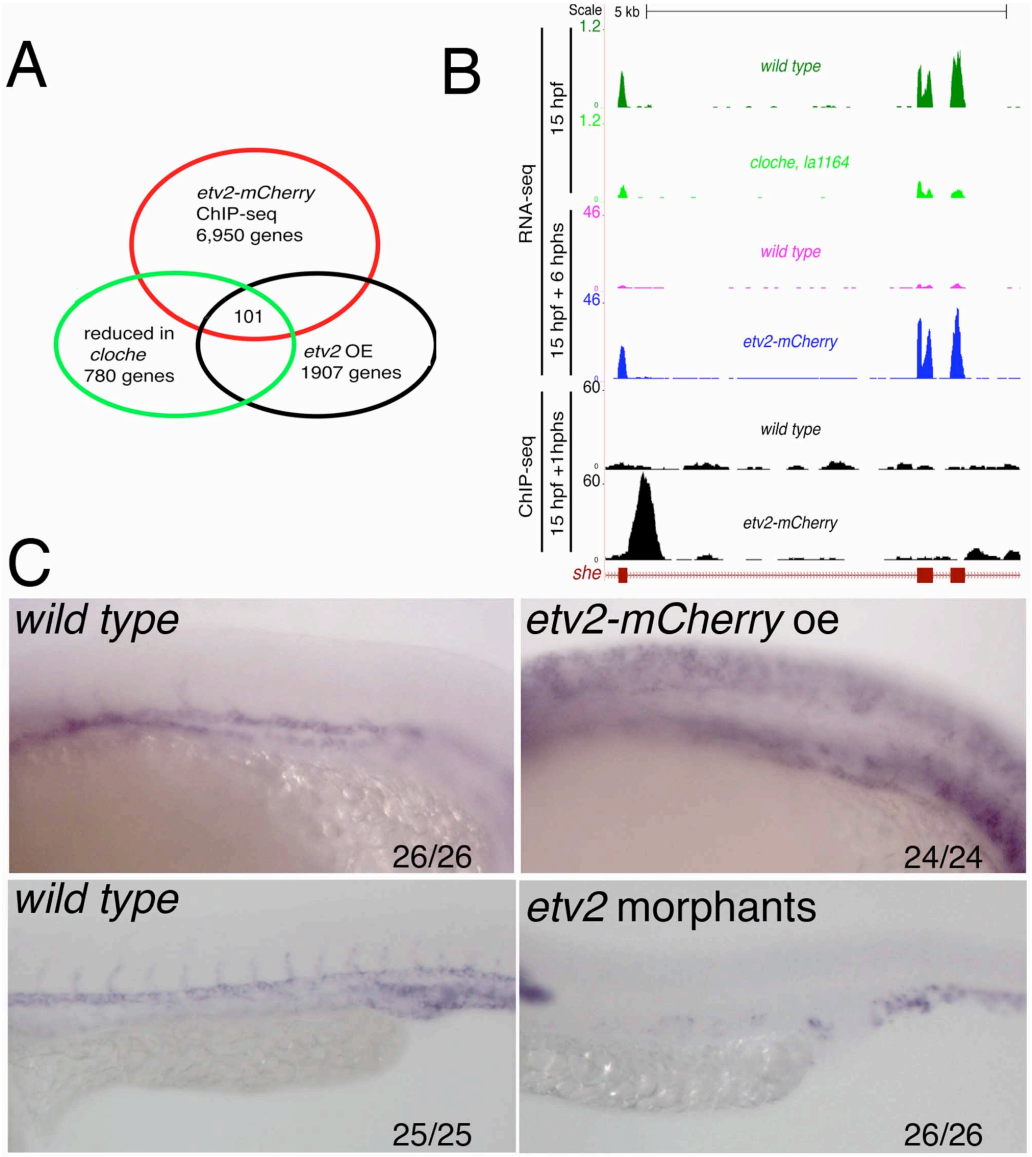


Figure 4. Identification of direct Etv2 target gene by overlap of ChIP-Seq with RNA-seq gain of function, and loss of function datasets. (A) Venn diagram showing overlap between genes containing ChIP-seq peaks and RNA-seq data for etv2 loss of function as represented by the sum of genes reduced below two-fold in three *cloche* alleles, and Etv2 gain of function, which includes genes induced two-fold and above at two developmental stages. Note that the total number of genes (101) includes 26 genes that we previously identified through transcriptome studies. (B) Representative RNA-Seq profiles in the top four tracks (dark green, *wild type*; light green, *la1164*; magenta, *wild type*; blue, *etv2-mCherry*) and ChIP-Seq profiles in the bottom two tracks, at one of the direct target genes, *she*. Gene structure of *she* is shown at the bottom with exons 3, 4, and 5, from left to right. (C) WISH images of ectopic *she* expression at the trunk following Etv2 overexpression (top right) and the reduction in the vasculature after Etv2 knockdown (bottom right). Embryo images are lateral views with anterior facing left, and values indicate the number of embryos showing the expression pattern shown over the number of embryos examined.

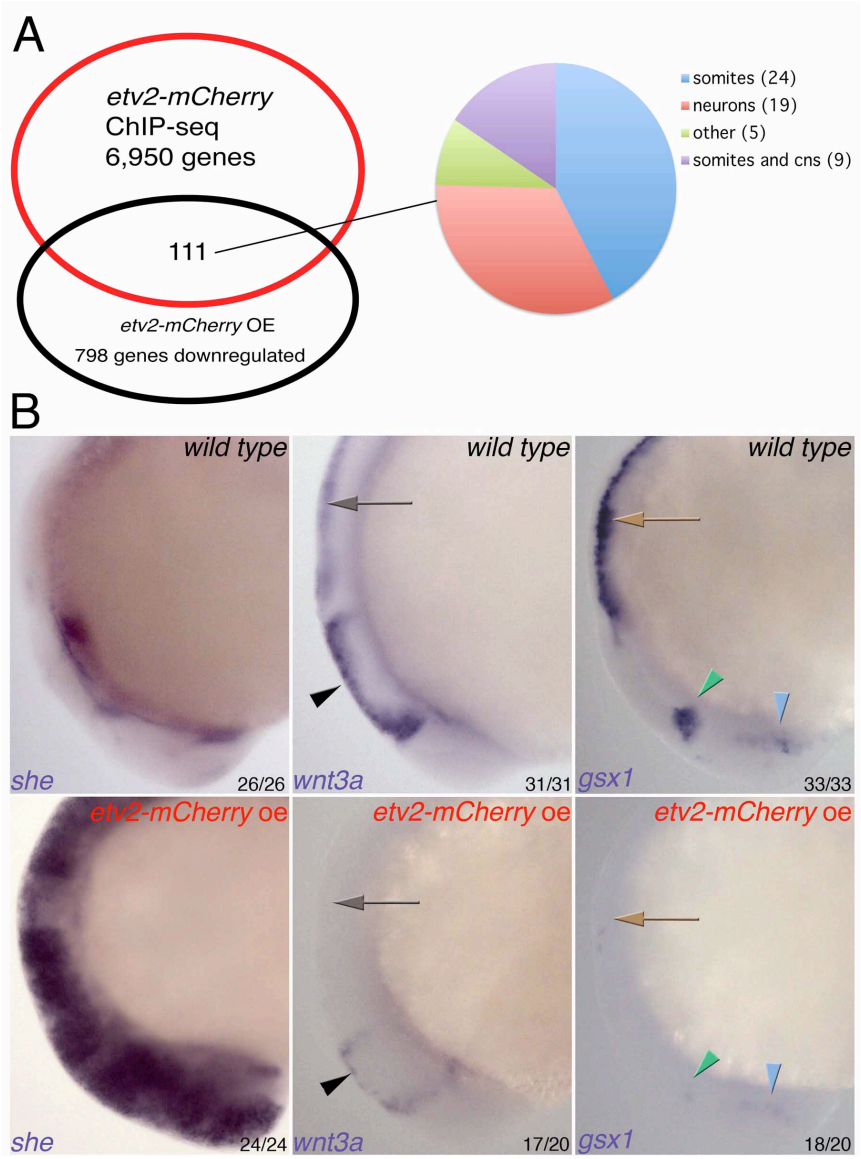


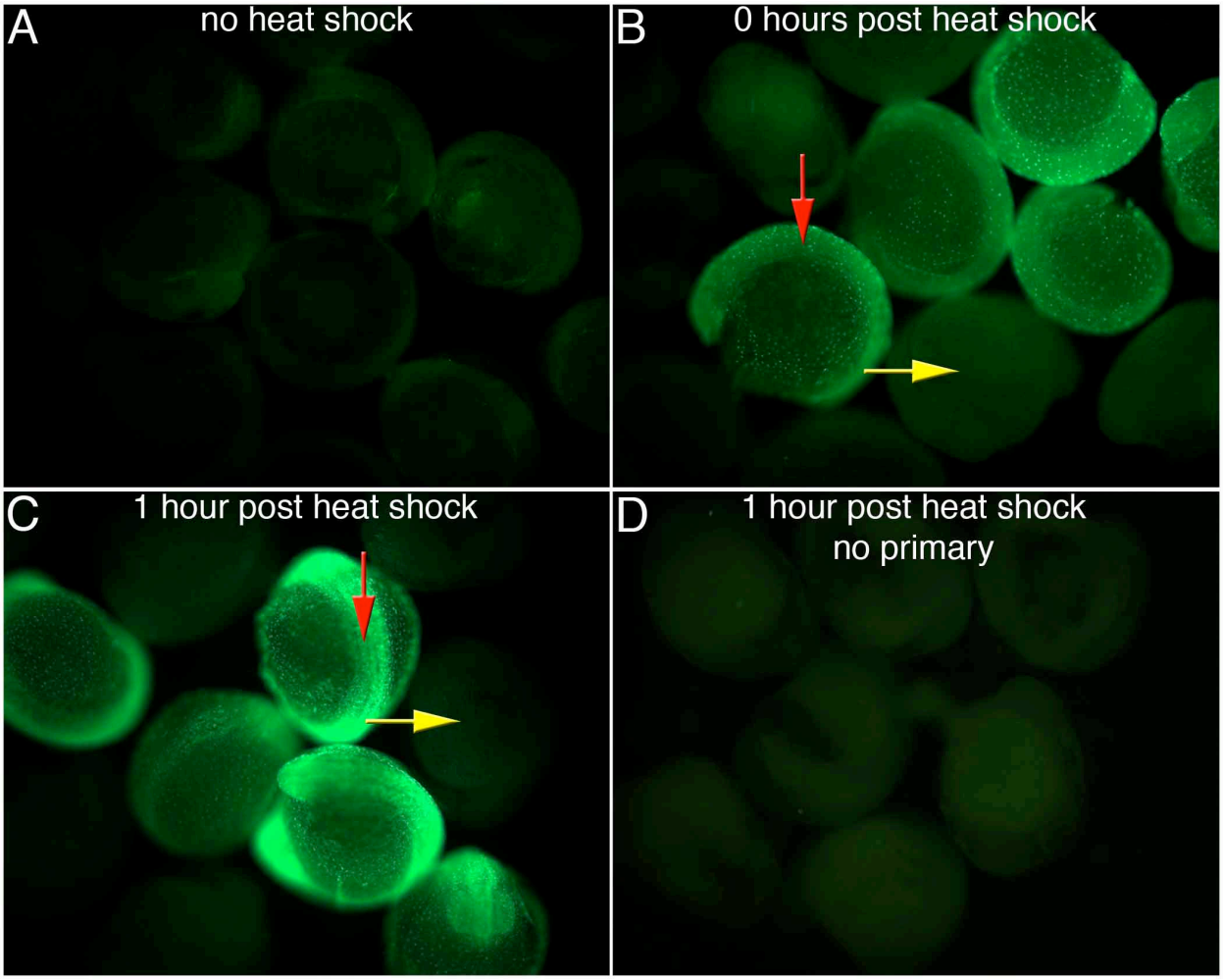
Figure 5. Forced expression of Etv2-mCherry represses expression of some genes in the nervous system. (A) 111 of 798 genes reduced below a 2-fold cutoff in the RNA-Seq profile at 6 hphs in the *etv2-mCherry* OE group are associated with Etv2-mCherry ChIP-Seq peaks. The expression of 57 of the 111 genes are available in ZFIN and pubmed databases, and these are broken down as indicated in the pie chart. (B) WISH for *she* (left column), *wnt3a* (center column) and *gsx1* (right column) at condition and stage matched embryos corresponding to RNA-Seq. Compared to wild types (top row), the expression of *she* in *etv2-mCherry* OE (bottom row) is induced ectopically in the nervous system, while the expression of *wnt3a* is abolished in the hindbrain (gray arrow), and severely reduced in the midbrain (black arrowhead). Similarly, *gsx1* expression is abolished in the hindbrain (brown arrow), the anterior tegmentum (green arrowhead), and severely reduced in the ventral diencephalon (light blue arrowhead). Embryos were imaged in lateral view with the dorsal-anterior end facing left, and values indicate the number of embryos showing the expression pattern shown over the number of embryos examined.

Table 1. The direct targets of Etv2 in zebrafish

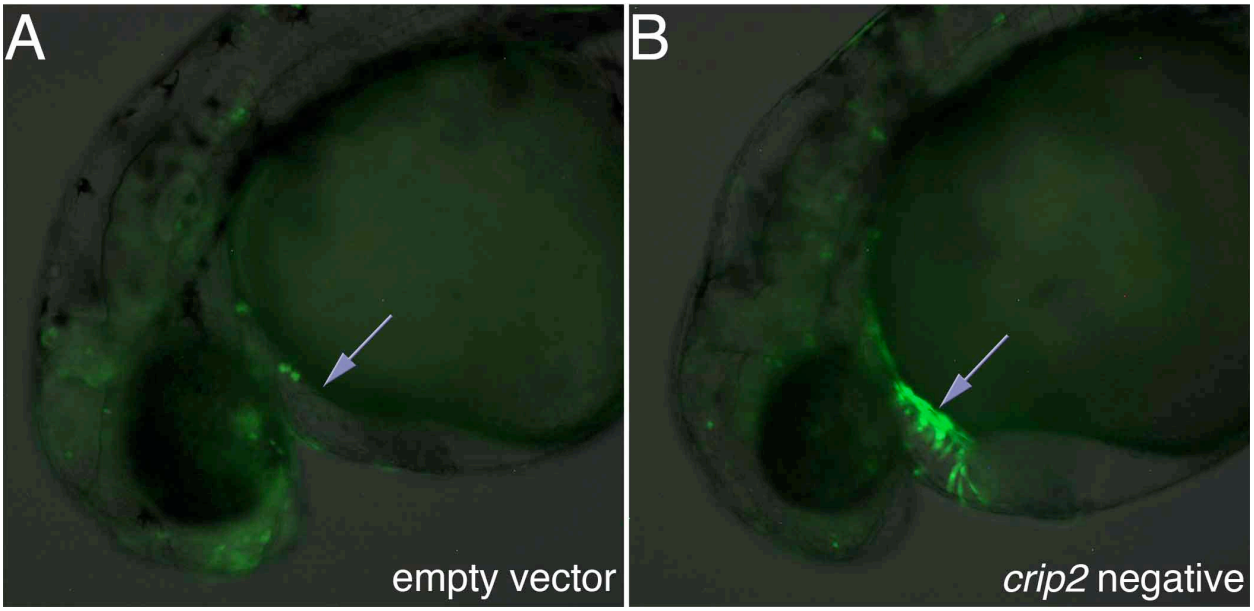
Ensembl Gene ID	Gene Symbol	Gene Name
ENSDARG00000000837	snx9	sorting nexin 9
ENSDARG00000004094	**acsbg2**	**acyl-CoA synthetase bubblegum family member 2**
ENSDARG00000004105	tie1	endothelium-specific receptor tyrosine kinase 1
ENSDARG00000004378	**yrk ***	**Yes-related kinase**
ENSDARG00000004433	btk	Bruton agammaglobulinemia tyrosine kinase
ENSDARG00000005414	grapa	GRB2-related adaptor protein a
ENSDARG00000005785	elovl7b	ELOVL family member 7, elongation of long chain fatty acids (yeast) b
ENSDARG00000007302	**sh3gl3**	**SH3-domain GRB2-like 3**
ENSDARG00000007349	dmrt1	doublesex and mab-3 related transcription factor 1
ENSDARG00000007727	**rgl2**	**ral guanine nucleotide dissociation stimulator-like 2**
ENSDARG00000008414	exoc3l2a	exocyst complex component 3-like 2a
ENSDARG00000009123	sele	selectin E
ENSDARG00000012281	zgc:65851	zgc:65851
ENSDARG00000012460	**rassf4**	**Ras association (RalGDS/AF-6) domain family member 4**
ENSDARG00000013110	epb49	erythrocyte membrane protein band 4.9 (dematin)
ENSDARG00000013653	eltd1	EGF, latrophilin and seven transmembrane domain containing 1
ENSDARG00000014059	cldn5b	claudin 5b
ENSDARG00000015312	klf12a	Kruppel-like factor 12a
ENSDARG00000015717	flt4	fms-related tyrosine kinase 4
ENSDARG00000015815	kdrl	kinase insert domain receptor like
ENSDARG00000016531	**klhl4 ***	**kelch-like 4**
ENSDARG00000019307	dusp5	dual specificity phosphatase 5
ENSDARG00000019371	flt1	fms-related tyrosine kinase 1 (vascular endothelial growth factor/vascular permeability factor receptor)
ENSDARG00000019405	**cntn4**	**contactin 4**
ENSDARG00000019930	tal1	T-cell acute lymphocytic leukemia 1
ENSDARG00000023053	fam129ab	family with sequence similarity 129, member Ab
ENSDARG00000024017	mamdc1	MAM domain containing 1
ENSDARG00000025902	plekhg5a	pleckstrin homology domain containing, family G (with RhoGef domain) member 5a
ENSDARG00000026039	cyp1a	cytochrome P450, family 1, subfamily A
ENSDARG00000026751	clec14a	C-type lectin domain family 14, member A
ENSDARG00000028663	tie2	endothelium-specific receptor tyrosine kinase 2
ENSDARG00000029372	**rasa4**	**RAS p21 protein activator 4**
ENSDARG00000030125	sox7	SRY-box containing gene 7
ENSDARG00000030782	exoc3l2 (2 of 2)	exocyst complex component 3-like 2
ENSDARG00000031751	npr1a	natriuretic peptide receptor 1a
ENSDARG00000034643	**fhl3**	**four and a half LIM domains 3**
ENSDARG00000034895	tgfb1b	transforming growth factor, beta 1b
ENSDARG00000035552	agtr2	angiotensin II receptor, type 2
ENSDARG00000036031	**sept5b**	**septin 5b**
ENSDARG00000036616	gpr182	G protein-coupled receptor 182
ENSDARG00000040065	cx45.6	connexin 45.6
ENSDARG00000040080	fli1b	friend leukemia integration 1b
ENSDARG00000043271	morc3b	MORC family CW-type zinc finger 3b
ENSDARG00000043376	egfl7	EGF-like-domain, multiple 7
ENSDARG00000044975	krt23 (6 of 15)	keratin 23 (histone deacetylase inducible)
ENSDARG00000045141	aqp8a.1	aquaporin 8a, tandem duplicate 1
ENSDARG00000045748	stab2	stabilin 2
ENSDARG00000052694	micall2a	mical-like 2a
ENSDARG00000053405	sord	sorbitol dehydrogenase
ENSDARG00000053868	etv2	ets variant gene 2
ENSDARG00000054632	fli1a	friend leukemia integration 1a
ENSDARG00000056125	myct1	myc target 1
ENSDARG00000056643	slc22a7b	solute carrier family 22 (organic anion transporter), member 7b
ENSDARG00000056920	**tmem88a**	**transmembrane protein 88 a**
ENSDARG00000057927	rps6ka3b	ribosomal protein S6 kinase, polypeptide 3b
ENSDARG00000058598	sox18	SRY-box containing gene 18
ENSDARG00000058940	C8H9orf116	chromosome 9 open reading frame 116
ENSDARG00000059363	TGFBR2 (2 of 2)	transforming growth factor, beta receptor II (70/80kDa)
ENSDARG00000059472	**arhgap31 ***	**Rho GTPase activating protein 31**
ENSDARG00000059534	jph3 (2 of 2)	junctophilin 3
ENSDARG00000059567	CU570691.1	Uncharacterized protein
ENSDARG00000060012	f2r	coagulation factor II (thrombin) receptor
ENSDARG00000060893	col8a2	collagen, type VIII, alpha 2
ENSDARG00000061747	C9HXorf36	chromosome X open reading frame 36
ENSDARG00000061798	**arhgap27 (1 of 2)**	**Rho GTPase activating protein 27**
ENSDARG00000062049	**hmha1**	**histocompatibility (minor) HA-1**
ENSDARG00000063054	shank3b	SH3 and multiple ankyrin repeat domains 3b
ENSDARG00000068516	**hapln1b**	**hyaluronan and proteoglycan link protein 1b**
ENSDARG00000070314	**cald1l**	**caldesmon 1 like**
ENSDARG00000070357	cx39.4	connexin 39.4
ENSDARG00000070542	**mafbb**	**v-maf musculoaponeurotic fibrosarcoma oncogene family, protein B, duplicate b**
ENSDARG00000070653	**FGD5 (1 of 2) ***	**FYVE, RhoGEF and PH domain containing 5**
ENSDARG00000071113	**xirp2a**	**xin actin-binding repeat containing 2a**
ENSDARG00000071491	**lrrc33 ***	**leucine rich repeat containing 33**
ENSDARG00000074378	junba	jun B proto-oncogene a
ENSDARG00000074434	**cd93 ***	**CD93 molecule**
ENSDARG00000074485	**ankdd1a**	**ankyrin repeat and death domain containing 1A**
ENSDARG00000074829	rasip1	Ras interacting protein 1
ENSDARG00000075273	thsd1	thrombospondin, type I, domain containing 1
ENSDARG00000077039	esama	endothelial cell adhesion molecule a
ENSDARG00000077285	scarf1	scavenger receptor class F, member 1
ENSDARG00000077618	rin3	Ras and Rab interactor 3
ENSDARG00000077864	**rasgrp3**	**RAS guanyl releasing protein 3 (calcium and DAG-regulated)**
ENSDARG00000078734	**myo1f**	**myosin IF**
ENSDARG00000079008	si:dkey-7f3.9	si:dkey-7f3.9
ENSDARG00000079575	**aff3**	**AF4/FMR2 family, member 3**
ENSDARG00000079862	kl	klotho
ENSDARG00000086057	plxnd1	plexin D1
ENSDARG00000086141	CU929396.1	Uncharacterized protein
ENSDARG00000086207	grin2da	glutamate receptor, ionotropic, N-methyl D-aspartate 2D, a
ENSDARG00000086499	rftn1 (1 of 2)	raftlin, lipid raft linker 1
ENSDARG00000087237	CU463231.1	Uncharacterized protein
ENSDARG00000087377	lbh	limb bud and heart development homolog (mouse)
ENSDARG00000087457	ecscr	endothelial cell-specific chemotaxis regulator
ENSDARG00000087956	**she ***	**Src homology 2 domain containing E**
ENSDARG00000088283	si:ch73-248e21.5	si:ch73-248e21.5
ENSDARG00000088546	myct1 (2 of 2)	myc target 1
ENSDARG00000088801	BX005175.3	Uncharacterized protein
ENSDARG00000092599	si:ch211-112k3.2	si:ch211-112k3.2
ENSDARG00000095019	lmo2	LIM domain only 2 (rhombotin-like 1)
ENSDARG00000095804	si:ch211-14c17.6	si:ch211-14c17.6

Note: genes in **bold** are genes we previously identified by transcriptional profiling of Etv2 OE using microarray and RNA-seq profiling. Bold and asterisks are genes identified as direct targets by RNA-seq data generated in this manuscript and also by previous transcriptional profile studies of Etv2 OE.

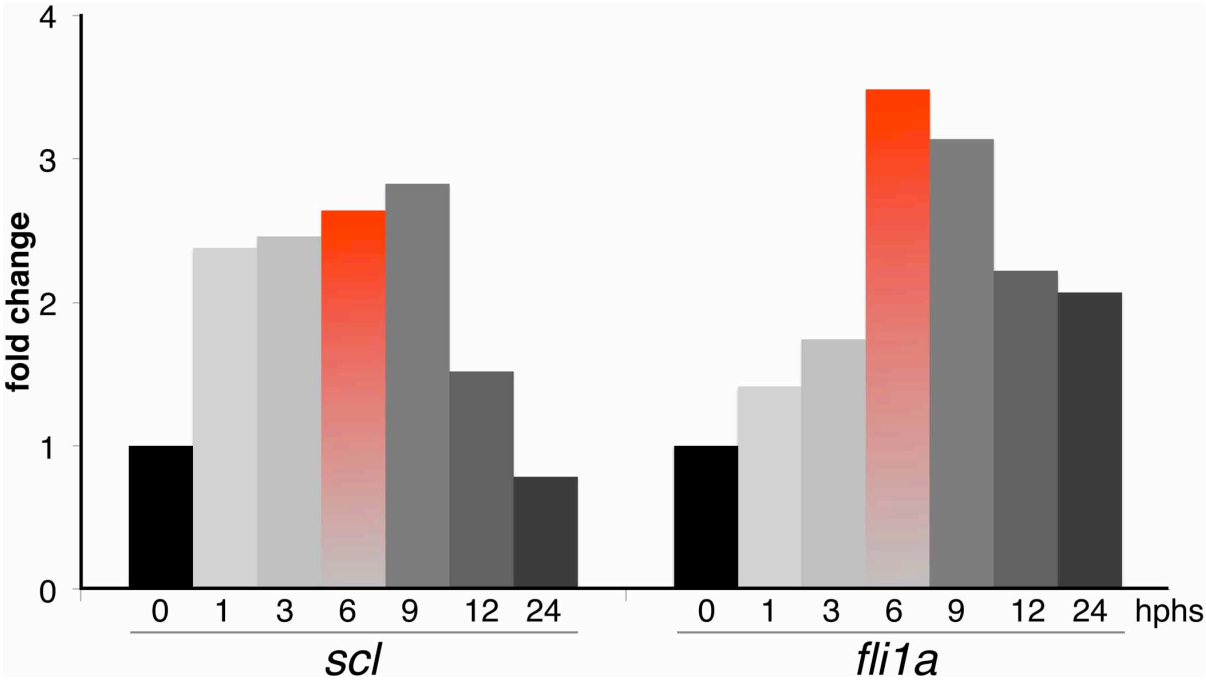
Chapter 3 Supplementary Material



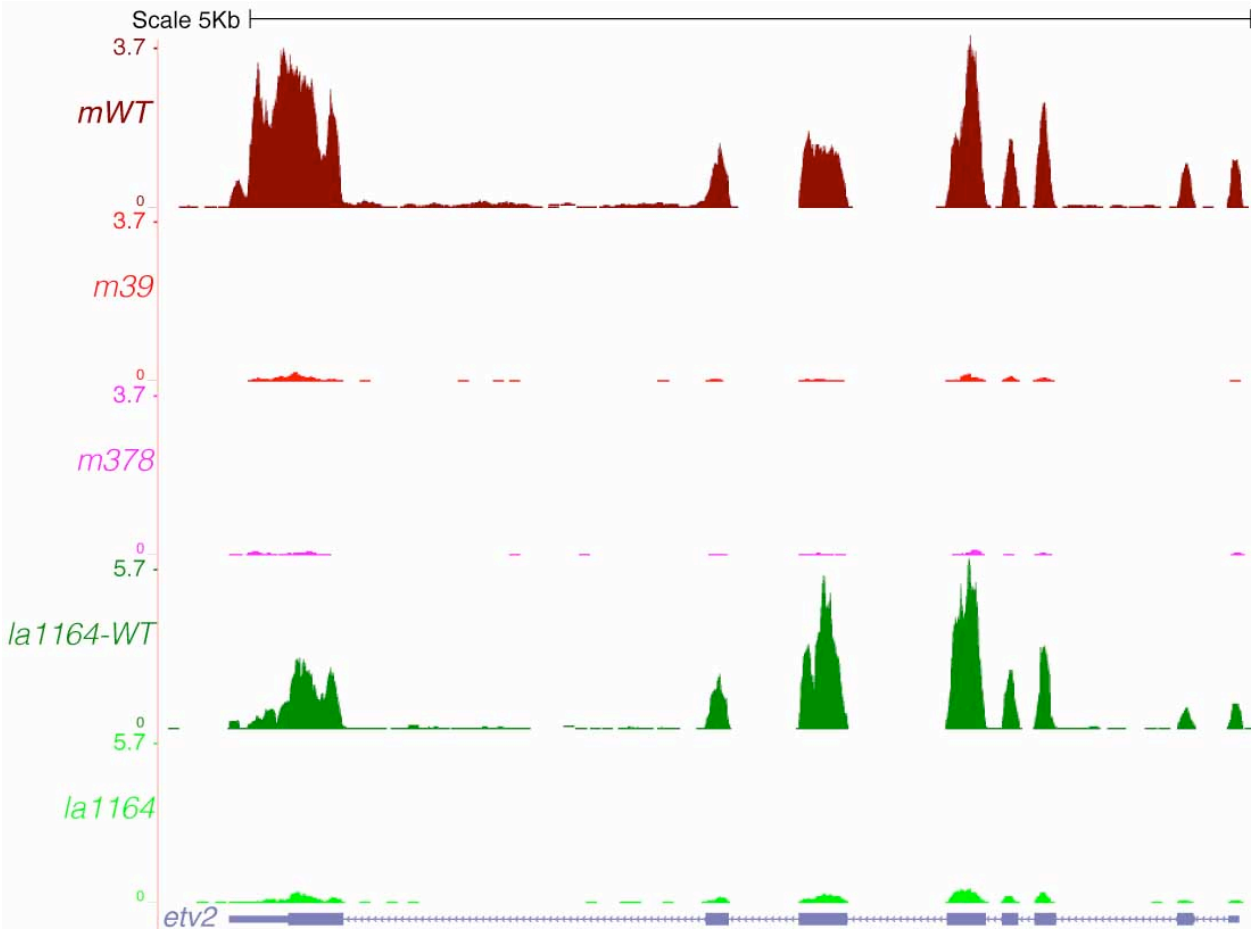
Supplementary Figure 1. The antibody used for ChIP specifically detects Etv2-mCherry. Whole mount fluorescent immunohistochemistry shows specific detection of Etv2-mCherry at 0hphs and 1hphs (downward facing red arrows), but not in wild types (rightward facing yellow arrows). Note the increased intensity observed in positive stained embryos at 1hphs in panel C relative to 0hphs in panel B and the absence of staining in no heat shock controls in panel A and the no primary control in panel D. Half of the embryos in panels B, C, and D displayed red Etv2-mCherry expression at the time of embryo fixation, while none of the embryos in panel A did.



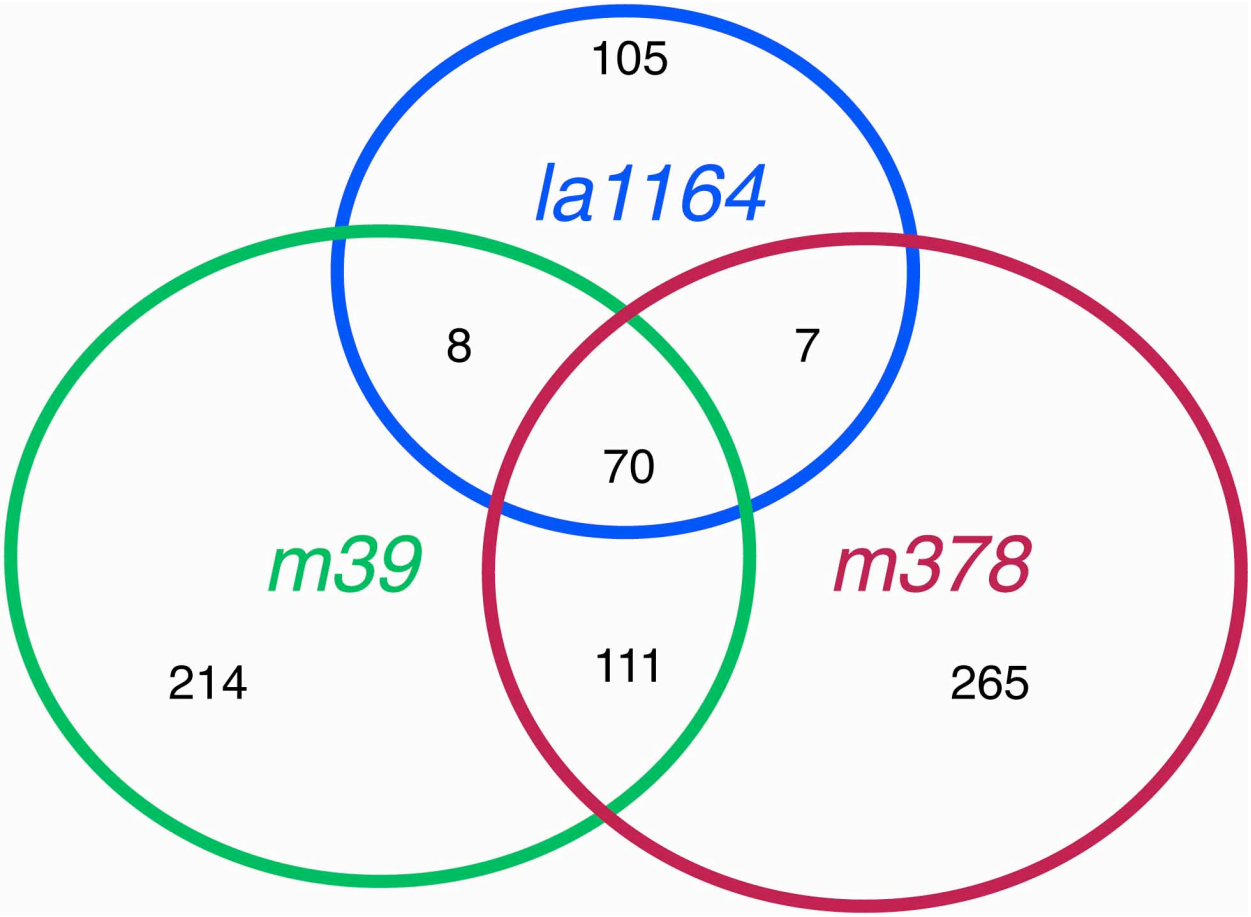
Supplementary Figure 2. The *crip2* negative enhancer is cardiac specific. Representative images of embryos injected with (A) empty vector control, and (B) the *crip2* negative enhancer. Purple arrow points to cardiac region. 38 of 40 wild type embryos injected with the *crip2* negative enhancer display marked GFP expression in cardiomyocytes at 30hpf.



Supplementary Figure 3. Gene expression changes of these genes were examined following heat shock of *hsp70l-etv2-mCherry* embryos at 15 hpf at 1, 3, 6, 9, 12 and 24 hours post heat shock (hphs). The largest induction of *fli1a* was detected at 6hphs, and subsequent RNA-Seq profiles were generated for embryos at this stage.



Supplementary Figure 4. *etv2* is downregulated in 3 *cloche* alleles. RNA-Seq profile at the *etv2* locus with the profile color-coded for each condition. mWT control (top row) was used to standardize expression for *cloche* alleles *m39* and *m378*, while the fourth row, *la1164*-WT, is the wild type control used to standardize expression for the *cloche* allele *la1164* (see methods). This is also reflected by the amplitude values in the y-axis.



Supplementary Figure 5. The transcript profiles of *cloche* alleles are distinct. Genes differentially repressed below a 2-fold cutoff by RNA-Seq profiling are labeled in black numbers.

Supplementary Methods: Bioinformatics

RNA-seq

The Illumina HiSeq 2000 platform yielded approximately 150 ~ 215 million single end short reads (50bp) per sample (Tables SM1 and SM2). Bowtie (version 0.12.8) (Langmead et al., 2009) was used to align the reads to the zebrafish genome Zv9 (danRer7) with up to 2 mismatches allowed. Each sample yielded mappable reads in excess of 80% (Tables SM1 and SM2), however, only uniquely mapped reads were used in subsequent gene differential expression analysis. Read counts per gene were summarized by the R package, GenomicRanges coupled with the UCSC danRer7 gene annotation table ensGene. A standard pair-wise comparison of two samples was performed by DESeq (Anders and Huber, 2010) with default parameters.

Table SM1

Sample	Total raw reads	Total uniquely mapped reads	Total multiple mapped reads	Total reads not mapped
hsp70l-Etv2-mCherry wild type	214,241,173	138,342,349 (64.57%)	19,552,429 (9.13%)	56,346,395 (26.30%)
***hsp70l-Etv2-mCherry* OE**	217,217,421	140,770,907 (64.81%)	17,955,469 (8.27%)	58,491,045 (26.93%)
90% Epiboly *wild type*	141,472,333	84,910,582 (60.02%)	17,178,622 (12.14%)	39,383,129 (27.84%)
90% Epiboly *Etv2* OE	161,511,110	90,997,864 (56.34%)	19,443,252 (12.04%)	51,069,994 (31.62%)

Table SM2

Sample	Total raw reads	Total uniquely mapped reads	Total multiple mapped reads	Total reads not mapped
wild type* for *m39*, *m378	166,083,322	98,042,913 (59.03%)	14,364,299 (8.64%)	53,676,110 (32.32%)
m39	141,547,867	89,900,715 (63.51%)	12,653,724 (8.94%)	38,993,428 (27.55%)
m378	140,179,310	87,543.275 (62.45%)	12,688,254 (9.05%)	39,947,781 (28.50%)
la1164	171,810,545	113,893,845 (66.29%)	15,418,548 (8.97%)	42,498,152 (24.74%)
wild type* for *la1164	199,804,780	112,974,973 (56.54%)	16,231,546 (8.12%)	70,598,261 (35.33%)

ChIP-seq

ChIP-seq DNA samples from wild type and *hsp701:etv2-mCherry* heterozygous zebrafish were sequenced separately in 2 lanes on the Illumina Hiseq 2000 platform. Overall, 82,311,932 short reads (50bp) were obtained from the wild type control sample and 55,560,402 short reads were obtained from *etv2-mCherry*. The R package, FastQC, was utilized to perform a sequencing quality check. The first 4 bases corresponding to the inline barcodes were trimmed from the raw reads before mapping. Bowtie(--sam --best –n 2 –m 1) (Langmead et al., 2009) was used to map the trimmed raw reads to the zebrafish genome Zv9. Approximately 17 million and 13 million uniquely mapped reads for wild type and etv2-mCherry samples were identified respectively (Table SM3). MACS (version 1.4) (Zhang et al., 2008) was then used to identify the corresponding read enriched regions on the genome with the following parameters: -t etv2_uniquly_mapped_reads -c control_uniquely_mapped_reads -g 1.5e9 -s 46 -w --single-profile --verbose 1. 10,029 peak regions were obtained with p-values < 1.00e-5.

Table SM3

Samples	Total raw reads	Total uniquely mapped reads	Total multiple mapped reads	Total reads not mapped
***Etv2-mCherry* ChIP-seq**	55,560,402	13,407,569 (24.13%)	7,694,800 (13.85%)	34,458,033 (62.02%)
***wild type* ChIP-seq**	82,311,932	16,905,309 (20.54%)	11,068,679 (13.45%)	54,337,944 (66.01%)

The R Package ChIPpeakAnno (Zhu et al., 2010b) and biomaRt were used to annotate the peaks regions. Ensembl (version 69) annotated zebrafish Transcription Start Sites(TSSs) were utilized in ChipPeakAnno to perform the annotation step. 6,950 Ensembl zebrafish TSSs were found to be neighboring the aforementioned 10,029 peak regions (Supplementary Table ##1). Among these, 9,687 peak regions were within 10Kb of the nearest TSS, and 6,856/6,950 TSSs were within 10 KB of a peak region. A +/-10Kb flanking distance distribution to these TSSs indicated that most Etv2-mCherry binding sites were within several hundred base pairs of the TSSs, and a bimodal distribution flanking these TSSs can also be observed (Fig 1D).

We also performed a GO terms enrichment analysis by ChIPpeakAnno coupled with GO gene mapping package org.Dr.eg.db (2.8.0). GO Biological Process terms GO:0001568 (blood vessel development) and GO:0001944 (vasculature development) showed up at very significant levels (p-value 1.93e-08 , Benjamini adjusted p-value 3.57e-06 and p-value 4.54e-09 , Benjamini adjusted p-value 1.08e-06 respectively).

Genomic sequences +/- 100bp from the summits of peaks enriched by MACS were extracted and submitted to DREME (Bailey, 2011) to investigate possible motifs in our Etv2-mCherry ChIP-seq data.

Correlation analysis with H3K4me1 and H3K4me3

We compared the 10,029 Etv2-mCherry ChIP-seq peaks generated by MACS with previously published ChIP-seq profiles of histone H3K4me1 and H3k4me3 binding sites by the Lawson lab (Aday et al., 2011). The published H3K4me1 and H3K4me3 binding site Zv7 coordinates were downloaded from (http://www.sciencedirect.com/science/article/pii/S0012160611001564), and converted to Zv9 coordinates with UCSC's batch coordinate conversion tool, liftOver. The coordinates of 37,244 out of 41,478 H3K4me1 binding sites and 24,750 out of 29,033 H3K4me3 binding sites were converted from the Zv7 to the Zv9 version. Overlapping peak regions were identified by the function findOverlappingpeaks from the R package ChipPeakAnno (Zhu et al., 2010a) with default parameters.

Supplementary Methods References

Aday, A.W., Zhu, L.J., Lakshmanan, A., Wang, J., and Lawson, N.D. (2011). Identification of cis regulatory features in the embryonic zebrafish genome through large-scale profiling of H3K4me1 and H3K4me3 binding sites. Dev Biol *357*, 450-462.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol *11*, R106.

Bailey, T.L. (2011). DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics *27*, 1653-1659.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhu, L.J., Gazin, C., Lawson, N.D., Pagès, H., Lin, S.M., Lapointe, D.S., and Green, M.R. (2010). ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. In BMC Bioinformatics, pp. 237.

Supplementary Tables

Supplementary Methods Table 1. Verification primers (ChIP-qPCR)

Fox:ets site	Gene	Sequence	Product	Location
yes	scl	TGAAATCCGAGCAATTTCCGC	117	
		GATGCTCTCCTGTCTCCGC		
no	scl neg	TGAACAAGAACTCGCCCACT	87	3.3 Kb downstream
		CTGGGCACTGGTCAAGTAGA		
yes	mef2cb	TTGGTTTGGTGTGTGTGTGC	115	
		GTGAAGGGGAAGCTGTCAGG		
no	mef2cb neg	AGGCTTCTGGAGATTTGGCA	109	2.3 Kb downstream
		TGTCATGCTGAGGGGTCAGA		
yes	notch1a	CATGCTGCCGTACACAGGTA	178	
		TCCCTCTGCGTTTCCTCCTA		
no	notch1a neg	ATCAGAGCAGGAAAGGCGTC	146	7 Kb downstream
		TCACACATTCATCTGGGGGC		
yes	arid1ab	GCTGTGCCGTTGCAGTATTC	98	
		CAACAGCCCCTGAACCTGAA		
no	arid1ab neg	AGGCCAAAAGAGGCAATCCA	92	4.3 Kb downstream
		GGGATGTGATCAGGCTCTGG		
no	numb	ACCGCTAGTGTTTCCTGCTC	164	
		TCACTACAATAGCCGTCCGC		
no	numb neg	TGCGTTTGTCATCTTGAGGC	112	15.7 Kb downstream
		TCGAGCTCTCCGCATAATCA		
yes	fli1a	TCGACAATGAGAAGGGTGTCT	107	

		GAGGGCATTCCTGTCTTTTG		
no	fli1a neg	TGTTTGGTTAAGGCCGGGCGAGA	92	6.1 Kb upstream
		ACCAAACTGTCTGGCCTCCTGCTC		
no	crip2	TGGCGGGAAAGGGTTGCCAT	94	
		TCTCTGGCCCTTCCGCTCCC		
no	crip2 neg	TCCAGCGCTCCTCCTCAGAGA	96	31 Kb upstream
		ACAATGCGTAAGCACACCGGCT		

Last column location is in reference to binding site tested

Supplementary Methods Table 2. Primers used to clone and test Etv2 bound sites by transient enhancer assay

Gene	Primer	PCR product
fgd5 (2 OF 2)	gagacgGGATCCCCTCACATTCCGTTTTCTGACA	539
	agacgcGTCGACGTCATTTTGACACAGTGTGACTT	
she	gagacgGGATCCAGTCACCCTAACGCTTTCTCA	504
	agacgcGTCGACGCGTTTCTCGATTTATGCACC	
cdh5	gagacgGGATCCGTAGCAGACCAGCCAGTGTC	553
	agacgcGTCGACACGTCTACAACTAGAACCACACA	
arhgap31	gagacgGGATCCCTCGTGTTTGACAGTAATGCAG	413
	agacgcGTCGACAGGAGGAGATGAAATGACTATCAC	
sox18	gagacgGGATCCGCGCGTGTAGGAGTGTTACT	524
	agacgcGTCGACATTAAAGCCCACCCTCACCC	
fli1a. A	gagacgGGATCCCGTAGTTCCAGCACCCAGAG	618
	agacgcGTCGACTAAAACATGGCTTGCGGCAG	
fli1a. B	gagacgGGATCCTCCTTCACAATTTGACACTGATACC	506
	agacgcGTCGACGTGCAGATTCTTGGTACATTGAGG	
crip2	gagacgGGATCCACCTGCTGTCACCCTAAATGT	559
	agacgcGTCGACCTTGTAAAGCTGACCGGCCT	
yrk. A	gagacgGGATCCCAGGCCGTTAATGAACGCAG	502
	agacgcGTCGACGCAGCATTGCATACTGGACTAA	
yrk. B	gagacgGGATCCTCAAATGTGCAGGTTGCTCTC	168
	agacgcGTCGACTTCCCGCTGTCCTATCATTCC	
crip2 neg	gagacgGGATCCACTTGTCCACAGTGGCAGTTA	450
	agacgcGTCGACCCATTCAGTTGACCGAAAGCA	
notch1a neg	gagacgGGATCCCTGTCTCCACTGCAGCATGA	625
	agacgcGTCGACCAACCCCCTGAGTATGAGCG	

Supplementary Table 1: Total set of genomic sites bound by etv2-mCherry

chr	start	end	length	summit	tags	#NAME?	fold_enrich	FDR(%)	peak_id	feature	Feature_start_pos	feature_end_pos	insideFeatur	distancetoFeature	shortestDistance	fromOverlappingOrNearest	symbol	genename
chr21	35524174	35525037	864	550	94	780.53	58.61	0	MACS_peak_5640	ENSDARG00000008057	35427567	35440218	downstream	96606	83955	NearestStart	ndfip1l	Nedd4 family interacting protein 1, like
chr9	15239095	15240397	1303	640	133	1058.56	45.16	0	MACS_peak_9743	ENSDARG00000034555	14842567	15484361	inside	245267	243964	NearestStart	LOC1000049	uncharacterized LOC100004906
Zv9_NA917	19284	20470	1187	580	84	558.18	44.33	0	MACS_peak_104	ENSDARG00000077748	3900	14023	upstream	-5260	5260	NearestStart	LOC1003316	solute carrier family 35 member E2-like
chr22	32638995	32640222	1228	576	113	636.71	41.38	0	MACS_peak_5947	ENSDARG00000019307	32601001	32612547	upstream	-26447	26447	NearestStart	dusp5	dual specificity phosphatase 5
chr18	6502089	6503147	1059	563	62	369.06	41.02	0	MACS_peak_3720	ENSDARG00000061937	6455368	6543712	inside	41624	40565	NearestStart	myo7aa	myosin VIIAa
chr20	28135642	28136692	1051	412	67	421.27	40.19	0	MACS_peak_5171	ENSDARG00000070425	28274771	28283631	upstream	-139130	138079	NearestStart	dll4	delta-like 4 (Drosophila)
chr11	40360316	40361607	1292	636	105	629.62	38.18	0	MACS_peak_1288	ENSDARG00000044573	40353953	40381846	inside	6362	6362	NearestStart	cdc42	cell division cycle 42
chr10	7379010	7380196	1187	612	77	439.87	37.3	0	MACS_peak_630	ENSDARG00000040535	7319768	7381450	inside	2441	1254	NearestStart	LOC793671	chondroitin sulfate N-acetylgalactosaminyltransferase 1-like
chr19	45125470	45126564	1095	552	75	405.97	36.33	0	MACS_peak_4431	ENSDARG00000093453	45083353	45144251	inside	18782	17687	NearestStart	NA	NA
chr19	45190723	45191612	890	415	68	470.01	35.3	0	MACS_peak_4433	ENSDARG00000048378	45151185	45224525	inside	33803	32913	NearestStart	yrk	Yes-related kinase
chr5	4305633	4306414	782	436	57	386.07	34.33	0	MACS_peak_7768	ENSDARG00000069013	4305932	4325985	overlapStart	-300	300	NearestStart	prdx4	peroxiredoxin 4
chr25	2907480	2908589	1110	508	71	401.18	33.08	0	MACS_peak_6729	ENSDARG00000090511	2859446	2860390	upstream	-47089	47089	NearestStart	NA	NA
chr12	6500308	6501350	1043	439	82	529.14	32.37	0	MACS_peak_1381	ENSDARG00000081921	6391777	6427874	upstream	-72433	72433	NearestStart	plce1	phospholipase C, epsilon 1
chr3	40612466	40613805	1340	326	120	804.89	32.16	0	MACS_peak_7320	ENSDARG00000059680	40571229	40584271	downstream	41236	28194	NearestStart	fscn1	fascin homolog 1, actin-bundling protein (Strongylocentrotus purpur
chr3	53480676	53481591	916	572	62	398.63	32.02	0	MACS_peak_7390	ENSDARG00000077915	53506261	53521409	upstream	-25586	24670	NearestStart	NA	NA
chr2	41524193	41525025	833	602	65	382.48	31.4	0	MACS_peak_4817	ENSDARG00000014692	41536204	41544585	upstream	-12012	11179	NearestStart	znf622	zinc finger protein 622
chr21	18997072	18998186	1115	463	95	542.64	31.2	0	MACS_peak_5520	ENSDARG00000044688	18941345	18948418	upstream	-48653	48653	NearestStart	dusp4	dual specificity phosphatase 4
chr11	18936750	18937328	579	244	58	401.1	30.95	0	MACS_peak_1069	ENSDARG00000057940	18955331	18977629	downstream	40880	18003	NearestStart	dido1	death inducer-obliterator 1
chr3	53706422	53707528	1107	573	68	340.03	30.79	0	MACS_peak_7394	ENSDARG00000042338	53721216	53730531	downstream	24110	13688	NearestStart	zgc:101053	zgc:101053
chr4	2070956	2071901	946	467	46	242.71	30.69	0.28	MACS_peak_7492	ENSDARG00000079143	2068247	2080596	inside	9641	2708	NearestStart	LOC1003342	uncharacterized LOC100334203
chr19	44866391	44868213	1823	1321	93	461.96	30.68	0	MACS_peak_4426	ENSDARG00000016350	44846525	44867685	overlapStart	1295	528	NearestStart	cap1	Cap1 CAP, adenylate cyclase-associated protein 1
chr25	18908778	18909855	1078	573	68	312.11	30.52	0	MACS_peak_6863	ENSDARG00000021110	18854268	18922771	inside	13994	12916	NearestStart	poc1b	POC1 centriolar protein homolog B (Chlamydomonas)
chr17	1729244	1730344	1101	542	57	312.34	30.14	0	MACS_peak_3306	ENSDARG00000087627	1718071	1726897	upstream	-2346	2346	NearestStart	NA	NA
chr5	18568765	18569906	1142	565	59	265.39	30.07	0	MACS_peak_7834	ENSDARG00000023267	18496590	18511731	upstream	-57033	57033	NearestStart	atad1a	ATPase family, AAA domain containing 1a
chr7	76777702	76778188	487	218	61	436.26	29.99	0	MACS_peak_9218	ENSDARG00000079771	76781759	76783743	upstream	-4058	3571	NearestStart	LOC557745	astrocytic phosphoprotein PEA-15-like
chr24	4266574	4267345	772	431	98	656.57	29.95	0	MACS_peak_6426	ENSDARG00000094547	4232679	4234119	downstream	33894	32454	NearestStart	NA	NA
chr16	52005040	52006052	1013	693	55	296.53	29.86	0	MACS_peak_3252	ENSDARG00000088308	51985714	52007871	inside	2832	1819	NearestStart	notch1l	notch homolog, like
chr19	36866215	36867437	1223	599	90	460.56	29.77	0	MACS_peak_4373	ENSDARG00000028533	36588317	36888387	inside	22173	20950	NearestStart	macf1	microtubule-actin crosslinking factor 1
chr8	16031755	16032770	1016	598	84	467.47	29.66	0	MACS_peak_9335	ENSDARG00000075624	16020966	16037550	inside	5796	4780	NearestStart	NA	NA
chr14	19851963	19853113	1151	410	142	811.36	29.56	0	MACS_peak_2245	ENSDARG00000015717	19717038	19856641	inside	4679	3528	NearestStart	fit4	fms-related tyrosine kinase 4
chr15	44554325	44555260	936	542	49	275.08	29.56	0	MACS_peak_2853	ENSDARG00000077766	44486393	44534234	upstream	-20888	20088	NearestStart	NA	NA
chr2	43489296	43490834	1539	453	99	475.78	29.44	0	MACS_peak_4837	ENSDARG00000027290	43570359	43633838	upstream	-81064	79525	NearestStart	nrp1b	neuropilin 1b

Note: Only few entries included in sample, table is long

Supplementary Table 2: Geno Ontology (GO) analysis of total etv2-mCherry ChIP-seq dataset

go.id	go.term	Ontology	pvalue	count.InDa	count.InGe	totaltermIr	totaltermIr	BH.adjuste	EntrezID
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	415213
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	554230
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	79375
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	373867
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	100148408
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	406513
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	402973
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	450059
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	794088
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	58073
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	394203
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	556993
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	403049
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	550161
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	402843
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	571773
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	30364
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	58071
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	30424
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	84037
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	566470
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	30266
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	406364
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	325221
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	30412
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	768121
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	563920
GO:0001944	vasculature development	BP	4.54E-09	114	221	80323	246145	1.08E-06	30619
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	494081
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	386772
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	555766
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	406513
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	373867
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	57936
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	402875
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	79375
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	563421
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	170446
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	30746
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	794088
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	403049
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	30766
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	544667
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	352910
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	445184
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	569971
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	405860
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	394127
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	565914
GO:0001568	blood vessel development	BP	1.94E-08	103	199	80323	246145	3.57E-06	571773

Note: Only few entries included in sample, table is long
Generated with the ChIP Peak Anno Package (Referenced in text)

Supplementary Table 3. Etv2-mCherry overlap with H3K4me1

etv2 Binding	chr	mmc1	mmc1_start	mmc1_end	strand	etv2 Binding	etv2 Binding	strand1	overlapFeature	shortestDistar
MACS_peak_	1	mmc1_2368	227678	229243	+	228036	229082	+	inside	161
MACS_peak_	1	mmc1_2369	273465	274890	+	272453	273345	+	upstream	120
MACS_peak_	1	mmc1_2799	2017293	2019729	+	2016783	2017571	+	overlapStart	278
MACS_peak_	1	mmc1_2794	2086919	2087846	+	2086759	2091754	+	includeFeature	160
MACS_peak_	1	mmc1_1970	2797904	2799065	+	2797557	2798797	+	overlapStart	268
MACS_peak_	1	mmc1_2412	2910664	2911458	+	2910865	2911678	+	overlapEnd	201
MACS_peak_	1	mmc1_2412	2909353	2910146	+	2910865	2911678	+	downstream	719
MACS_peak_	1	mmc1_2083	4104880	4106098	+	4104871	4105884	+	overlapStart	9
MACS_peak_	1	mmc1_2076	4165734	4167101	+	4165664	4166798	+	overlapStart	70
MACS_peak_	1	mmc1_2075	4171848	4172720	+	4170309	4171439	+	upstream	409
MACS_peak_	1	mmc1_2074	4172901	4177763	+	4178704	4179860	+	downstream	941
MACS_peak_	1	mmc1_2071	4186592	4188381	+	4187026	4187891	+	inside	434
MACS_peak_	1	mmc1_685	4796362	4797739	+	4796876	4797790	+	overlapEnd	51
MACS_peak_	1	mmc1_688	4838402	4840212	+	4839249	4840258	+	overlapEnd	46
MACS_peak_	1	mmc1_2842	4908123	4909584	+	4908333	4909395	+	inside	189
MACS_peak_	1	mmc1_2913	7166075	7169302	+	7167052	7169068	+	inside	234
MACS_peak_	1	mmc1_2907	7609581	7610818	+	7609746	7610628	+	inside	165
MACS_peak_	1	mmc1_2900	7724975	7727889	+	7725155	7727216	+	inside	180
MACS_peak_	1	mmc1_2899	7728746	7732670	+	7727308	7728372	+	upstream	374
MACS_peak_	1	mmc1_2900	7724975	7727889	+	7727308	7728372	+	overlapEnd	483
MACS_peak_	1	mmc1_2886	8220609	8222288	+	8220800	8221873	+	inside	191
MACS_peak_	1	mmc1_2885	8228958	8230213	+	8228985	8230310	+	overlapEnd	27
MACS_peak_	1	mmc1_2884	8230644	8232005	+	8228985	8230310	+	upstream	334
MACS_peak_	1	mmc1_2867	8558373	8560893	+	8561362	8562526	+	downstream	469
MACS_peak_	1	mmc1_2973	8632973	8634898	+	8633244	8633886	+	inside	271
MACS_peak_	1	mmc1_2951	9587249	9588762	+	9587539	9588459	+	inside	290
MACS_peak_	1	mmc1_2949	9822750	9823497	+	9822764	9823317	+	inside	14
MACS_peak_	1	mmc1_3081	10289575	10291066	+	10290644	10291381	+	overlapEnd	315
MACS_peak_	1	mmc1_3000	11024800	11026438	+	11024989	11025861	+	inside	189
MACS_peak_	1	mmc1_3001	11029429	11031868	+	11029898	11031161	+	inside	469
MACS_peak_	1	mmc1_2980	11573348	11576017	+	11574268	11575234	+	inside	783
MACS_peak_	1	mmc1_2977	11644476	11646062	+	11643808	11644898	+	overlapStart	422
MACS_peak_	1	mmc1_3036	12849514	12853693	+	12851314	12853480	+	inside	213
MACS_peak_	1	mmc1_3038	12862836	12864909	+	12863339	12864441	+	inside	468
MACS_peak_	1	mmc1_3048	13232505	13233999	+	13232657	13233651	+	inside	152
MACS_peak_	1	mmc1_3051	13289676	13292271	+	13290568	13291498	+	inside	773
MACS_peak_	1	mmc1_3053	13338134	13339375	+	13338747	13339793	+	overlapEnd	418

Note: Only few entries included in sample, table is long

Supplementary Table 4, Sheet 1. Genes induced 2-fold and above by Etv2 OE at 90% epiboly

ensembl_id	wtbase	mtbase	foldCh	log2FoldCh	pval	padj	gname	gdesp
ENSDARG00000053868	36.8951216	7685.81285	208.315152	7.70262397	2.63E-27	6.90E-23	etv2	ets variant gene 2 [Source:ZFIN;Acc:ZDB-GENE-050622-14]
ENSDARG00000077069	5.59016994	1112.66743	199.04	7.63691458	3.45E-19	1.81E-15	srgn	serglycin [Source:ZFIN;Acc:ZDB-GENE-030131-7127]
ENSDARG00000040565	1.11803399	205.718254	184	7.52356196	1.54E-08	1.16E-05	ckmb	creatine kinase, muscle b [Source:ZFIN;Acc:ZDB-GENE-040426-2128]
ENSDARG00000075549	8.94427191	1636.80176	183	7.51569984	4.12E-21	5.41E-17	cdh5	cadherin 5 [Source:ZFIN;Acc:ZDB-GENE-040816-1]
ENSDARG00000041828	1.11803399	174.413302	156	7.28540222	1.41E-07	8.23E-05	si:dkeyp-28d	si:dkeyp-28d2.4 [Source:ZFIN;Acc:ZDB-GENE-091204-68]
ENSDARG00000088801	5.59016994	859.544531	153.76	7.26453643	5.50E-17	1.81E-13	FGD5 (2 of 3	FYVE, RhoGEF and PH domain containing 5 [Source:HGNC Symbol;Acc:19117]
ENSDARG00000054632	11.1803399	1471.33273	131.6	7.04001568	1.90E-19	1.25E-15	fli1a	friend leukemia integration 1a [Source:ZFIN;Acc:ZDB-GENE-980526-426]
ENSDARG00000063054	1.11803399	135.058506	120.8	6.91647664	2.98E-06	0.00106995	shank3b	SH3 and multiple ankyrin repeat domains 3b [Source:ZFIN;Acc:ZDB-GENE-041210-74]
ENSDARG00000045141	20.1246118	2178.82464	108.266667	6.75844532	2.37E-20	2.08E-16	aqp8a.1	aquaporin 8a, tandem duplicate 1 [Source:ZFIN;Acc:ZDB-GENE-040912-106]
ENSDARG00000019861	1.11803399	115.381108	103.2	6.68929916	1.55E-05	0.00461253	fgl2	fibrinogen-like 2 [Source:ZFIN;Acc:ZDB-GENE-030131-9506]
ENSDARG00000068036	16.7705098	1356.84605	80.9066667	6.33818668	2.94E-17	1.10E-13	tmem119b	transmembrane protein 119b [Source:ZFIN;Acc:ZDB-GENE-070410-49]
ENSDARG00000016939	1.11803399	67.9764665	60.8	5.92599942	0.0012237	0.12260827	ITGB2	integrin, beta 2 (complement component 3 receptor 3 and 4 subunit) [Source:HGNC Symbol;Acc:6155]
ENSDARG00000071113	26.8328157	1573.29743	58.6333333	5.87364917	1.85E-16	5.39E-13	xirp2a	xin actin-binding repeat containing 2a [Source:ZFIN;Acc:ZDB-GENE-040108-7]
ENSDARG00000042041	1.11803399	61.7154762	55.2	5.78659636	0.00227495	0.18898639	tal2	T-cell acute lymphocytic leukemia 2 [Source:ZFIN;Acc:ZDB-GENE-040115-1]
ENSDARG00000009123	2.23606798	119.853244	53.6	5.7441611	3.23E-05	0.00764196	SELP	selectin P (granule membrane protein 140kDa, antigen CD62) [Source:HGNC Symbol;Acc:10721]
ENSDARG00000073928	6.70820393	350.615459	52.2666667	5.70781925	1.27E-09	1.15E-06	mrc1a	mannose receptor, C type 1a [Source:ZFIN;Acc:ZDB-GENE-090915-4]
ENSDARG00000009782	11.1803399	533.078606	47.68	5.57531233	2.53E-11	3.50E-08	myh11a	myosin, heavy polypeptide 11, smooth muscle a [Source:ZFIN;Acc:ZDB-GENE-050531-1]
ENSDARG00000070266	8.94427191	421.275207	47.1	5.55765515	3.14E-10	3.58E-07	spock3	sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 3 [Source:ZFIN;Acc:ZDB-GENE-090311-24]
ENSDARG00000045926	14.5344419	680.659092	46.8307692	5.54938483	2.45E-12	4.02E-09	crabp1a	cellular retinoic acid binding protein 1a [Source:ZFIN;Acc:ZDB-GENE-020320-3]
ENSDARG00000019930	31.3049517	1460.5996	46.6571429	5.54402606	3.46E-15	8.25E-12	tal1	T-cell acute lymphocytic leukemia 1 [Source:ZFIN;Acc:ZDB-GENE-980526-501]
ENSDARG00000078437	1.11803399	48.2990683	43.2	5.43295941	0.00887346	0.45406868	SH3TC1 (1 of	SH3 domain and tetratricopeptide repeats 1 [Source:HGNC Symbol;Acc:26009]
ENSDARG00000026751	10.0623059	433.797188	43.1111111	5.42998784	3.88E-10	4.15E-07	clec14a	C-type lectin domain family 14, member A [Source:ZFIN;Acc:ZDB-GENE-030131-5536]
ENSDARG00000040701	1.11803399	47.4046411	42.4	5.40599236	0.00973058	0.48286852	CR388160.1	Uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:F1QQ53]
ENSDARG00000077285	3.35410197	140.425069	41.8666667	5.38773015	1.61E-05	0.00465849	scarf1	scavenger receptor class F, member 1 [Source:ZFIN;Acc:ZDB-GENE-030131-8778]
ENSDARG00000056819	15.6524758	622.521325	39.7714286	5.31366048	1.96E-11	3.02E-08	hoxa9b	homeo box A9b [Source:ZFIN;Acc:ZDB-GENE-000823-2]
ENSDARG00000019307	11.1803399	422.169634	37.76	5.23878686	1.12E-09	1.05E-06	dusp5	dual specificity phosphatase 5 [Source:ZFIN;Acc:ZDB-GENE-010625-1]
ENSDARG00000092653	1.11803399	42.038078	37.6	5.23266076	0.01697876	0.67137995	si:dkey-1h24	si:dkey-1h24.6 [Source:ZFIN;Acc:ZDB-GENE-100922-103]
ENSDARG00000079371	1.11803399	41.1436508	36.8	5.20163386	0.01863947	0.70913728	LRRC3C (1 of	leucine rich repeat containing 3C [Source:HGNC Symbol;Acc:40034]
ENSDARG00000077039	12.2983739	445.424741	36.2181818	5.17864222	8.68E-10	8.44E-07	esama	endothelial cell adhesion molecule a [Source:ZFIN;Acc:ZDB-GENE-090629-1]
ENSDARG00000035997	26.8328157	942.726259	35.1333333	5.13476856	1.55E-12	2.71E-09	zfpm2b	zinc finger protein, multitype 2b [Source:ZFIN;Acc:ZDB-GENE-060130-5]
ENSDARG00000044973	2.23606798	76.0263112	34	5.08746284	0.00135278	0.13104025	zgc:171226	zgc:171226 [Source:ZFIN;Acc:ZDB-GENE-070822-8]
ENSDARG00000087956	237.023206	7568.64289	31.9320755	4.99693442	5.48E-16	1.44E-12	she	Src homology 2 domain containing E [Source:ZFIN;Acc:ZDB-GENE-090915-6]
ENSDARG00000010317	2.23606798	69.7653209	31.2	4.96347412	0.00237843	0.19511323	gpr183	G protein-coupled receptor 183 [Source:ZFIN;Acc:ZDB-GENE-070615-28]
ENSDARG00000035552	6.70820393	205.718254	30.6666667	4.93859946	1.99E-06	0.00077968	agtr2	angiotensin II receptor, type 2 [Source:ZFIN;Acc:ZDB-GENE-050306-56]
ENSDARG00000044365	3.35410197	102.859127	30.6666667	4.93859946	0.0002884	0.04159745	angptl3	angiopoietin-like 3 [Source:ZFIN;Acc:ZDB-GENE-010817-3]

Note: Only few entries included in sample, table is long

Supplementary Table 4, Sheet 2. Genes induced 2-fold and above by etv2-mCherry OE

ensembl_id	wtbase	mtbase	foldCh	log2FoldCh	pval	padj	gname	gdesp
ENSDARG00(	14.1832241	11357.3612	800.760189	9.64522644	1.72E-08	0.0004617	jak3	Janus kinase 3 (a protein tyrosine kinase, leukocyte) [Source:ZFIN;Acc:ZDB-GENE-070725-4]
ENSDARG00(	445.758473	160003.958	358.947654	8.48762966	1.38E-07	0.00185601	NEFM	neurofilament, medium polypeptide [Source:HGNC Symbol;Acc:7734]
ENSDARG00(	2.02617488	687.995897	339.554056	8.40749745	3.63E-06	0.00974143	gnrhr2	gonadotropin releasing hormone receptor 2 [Source:ZFIN;Acc:ZDB-GENE-090128-3]
ENSDARG00(	4.05234975	1069.00941	263.799887	8.04330014	2.51E-06	0.00852725	CH211-57M13.15	
ENSDARG00(	2.02617488	511.308285	252.351508	7.9792909	1.12E-05	0.01596791	PCBP3 (1 of 2	poly(rC) binding protein 3 [Source:HGNC Symbol;Acc:8651]
ENSDARG00(	55.7198091	11472.8498	205.902532	7.68581776	9.80E-07	0.00525537	si:dkey-37g1	si:dkey-37g12.1 [Source:ZFIN;Acc:ZDB-GENE-090312-145]
ENSDARG00(	2.02617488	353.375224	174.405096	7.44629839	4.68E-05	0.04161288	si:dkey-80c2	si:dkey-80c24.3 [Source:ZFIN;Acc:ZDB-GENE-061207-71]
ENSDARG00(	5.06543719	876.528488	173.041034	7.43497038	8.25E-06	0.0158015	si:dkey-189e	si:dkey-189e1.3 [Source:ZFIN;Acc:ZDB-GENE-091204-205]
ENSDARG00(	5.06543719	816.316508	161.154206	7.33229803	1.07E-05	0.01596791	si:dkey-37g1	si:dkey-37g12.2 [Source:ZFIN;Acc:ZDB-GENE-090313-309]
ENSDARG00(	1.01308744	144.113918	142.252201	7.15230717	0.00069195	0.16376866	CU459092.1	Uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:F1QPK9]
ENSDARG00(	26.3402734	3506.11395	133.108488	7.05645876	5.08E-06	0.01134542	si:dkey-88l16	si:dkey-88l16.3 [Source:ZFIN;Acc:ZDB-GENE-030131-6015]
ENSDARG00(	3.03926231	394.832652	129.910686	7.0213763	5.85E-05	0.04482183	CR556710.1	Uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:F1R9D5]
ENSDARG00(	1.01308744	131.281857	129.585909	7.01776504	0.00099047	0.19519542	CT027715.1	Uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:E7F7N3]
ENSDARG00(	1.01308744	121.411041	119.842608	6.90499711	0.00133299	0.22053535	opn1mw2	opsin 1 (cone pigments), medium-wave-sensitive, 2 [Source:ZFIN;Acc:ZDB-GENE-030728-5]
ENSDARG00(	6.07852463	697.866713	114.808569	6.84308651	2.82E-05	0.03285991	AL954695.5	Uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:F1Q7K9]
ENSDARG00(	1.01308744	113.514388	112.047967	6.80797266	0.00171591	0.25692669	SLC12A2 (3 c	solute carrier family 12 (sodium/potassium/chloride transporters), member 2 [Source:HGNC Symbol;Acc:10911]
ENSDARG00(	17.2224864	1916.91253	111.302891	6.79834725	1.15E-05	0.01596791	pard6ga	par-6 partitioning defective 6 homolog gamma A (C. elegans) [Source:ZFIN;Acc:ZDB-GENE-050923-1]
ENSDARG00(	106.374181	11237.9244	105.645226	6.72308376	7.83E-06	0.0158015	gfra2	glial cell line derived neurotrophic factor family receptor alpha 2 [Source:ZFIN;Acc:ZDB-GENE-040630-1]
ENSDARG00(	5.06543719	525.127427	103.668728	6.69583695	5.37E-05	0.04231408	BX908759.3	Uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:F1R1V5]
ENSDARG00(	7.09161207	719.582509	101.469525	6.66490269	3.59E-05	0.03696418	si:ch211-284	si:ch211-284e13.13 [Source:ZFIN;Acc:ZDB-GENE-060526-159]
ENSDARG00(	1.01308744	88.8373468	87.6897131	6.4543357	0.00418699	0.39080106	si:ch211-160	si:ch211-160b11.4 [Source:ZFIN;Acc:ZDB-GENE-081104-142]
ENSDARG00(	2.02617488	174.713449	86.2282179	6.43008816	0.00070889	0.16376866	opn1mw3	opsin 1 (cone pigments), medium-wave-sensitive, 3 [Source:ZFIN;Acc:ZDB-GENE-030728-6]
ENSDARG00(	3.03926231	256.641224	84.4419459	6.39988792	0.00030033	0.10084479	CR788231.1	Uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:E7EYU7]
ENSDARG00(	19.2486613	1581.30477	82.1514154	6.36021353	3.00E-05	0.03325275	vtg3	vitellogenin 3, phosvitinless [Source:ZFIN;Acc:ZDB-GENE-991019-2]
ENSDARG00(	1.01308744	82.914857	81.8437322	6.35480003	0.005332	0.43044615	si:dkey-47k2	si:dkey-47k20.2 [Source:ZFIN;Acc:ZDB-GENE-060503-735]
ENSDARG00(	8.1046995	651.473876	80.382237	6.32880482	6.99E-05	0.04999875	OMG (1 of 2)	oligodendrocyte myelin glycoprotein [Source:HGNC Symbol;Acc:8135]
ENSDARG00(	1.01308744	79.9536121	78.9207418	6.30233261	0.00604541	0.46030986	si:ch211-132	si:ch211-132b12.6 [Source:ZFIN;Acc:ZDB-GENE-050420-179]
ENSDARG00(	3.03926231	236.899591	77.9464116	6.2844107	0.00040609	0.11830577	TRIM47 (76 (	tripartite motif containing 47 [Source:HGNC Symbol;Acc:19020]

Note: Only few entries included in sample, table is long

Supplementary Table 5, Sheet 2: Genes reduced by etv2-mCherry OE below 2-fold that contain ChIP-seq peaks

Ensembl Gene ID	Gene Symbol	Description	Expressed in:	
ENSDARG00000001712	CR855338.1	Uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:E7FCW4]		no expression
ENSDARG00000002204	hspb11	heat shock protein, alpha-crystallin-related, b11 [Source:ZFIN;Acc:ZDB-GENE-030131-5148]	somites	
ENSDARG00000002768	pvalb2	parvalbumin 2 [Source:ZFIN;Acc:ZDB-GENE-000322-4]	skeletal muscle, somites	
ENSDARG00000003684	obsl1a	obscurin-like 1a [Source:ZFIN;Acc:ZDB-GENE-060503-649]		no expression but human homolog: OIV predict:cardiac and sk mm
ENSDARG00000003732	mitfa	microphthalmia-associated transcription factor a [Source:ZFIN;Acc:ZDB-GENE-990910-11]	neural crest cells	
ENSDARG00000004078	acsl4a	acyl-CoA synthetase long-chain family member 4a [Source:ZFIN;Acc:ZDB-GENE-040426-1565]	cns and ysl	
ENSDARG00000004082	cx39.9	connexin 39.9 [Source:ZFIN;Acc:ZDB-GENE-040406-2]	somites	
ENSDARG00000005173	pax7a	paired box gene 7a [Source:ZFIN;Acc:ZDB-GENE-990415-201]	skeletal muscle, trunk neural crest cells and neural tube	
ENSDARG00000006760	slc24a3	solute carrier family 24 (sodium/potassium/calcium exchanger), member 3 [Source:ZFIN;Acc:ZDB-GENE-090312-67]		no expression
ENSDARG00000008790	ACTR3B	ARP3 actin-related protein 3 homolog B (yeast) [Source:HGNC Symbol;Acc:17256]		no expression
ENSDARG00000008796	her5	hairy-related 5 [Source:ZFIN;Acc:ZDB-GENE-990415-90]	at stage examined, midbrain hindbrain boundary	
ENSDARG00000009564	syt8	synaptotagmin VIII [Source:ZFIN;Acc:ZDB-GENE-060303-4]		no expression
ENSDARG00000012311	myoz2a	myozenin 2a [Source:ZFIN;Acc:ZDB-GENE-040426-1880]	skeletal muscle	
ENSDARG00000013607	sema3gb	semaphorin 3gb [Source:ZFIN;Acc:ZDB-GENE-050513-4]	cns	
ENSDARG00000014196	myl1	myosin, light chain 1, alkali; skeletal, fast [Source:ZFIN;Acc:ZDB-GENE-030131-8109]	skeletal muscle	
ENSDARG00000014248	pdlim3b	PDZ and LIM domain 3b [Source:ZFIN;Acc:ZDB-GENE-060130-104]	skeletal muscle	
ENSDARG00000014321	eng1a	engrailed 1a [Source:ZFIN;Acc:ZDB-GENE-980526-216]	skeletal muscle	and midbrain hindbrain boundary
ENSDARG00000014479	ptf1a	pancreas specific transcription factor, 1a [Source:ZFIN;Acc:ZDB-GENE-030616-579]	cns at stages examined	
ENSDARG00000016141	slc6a2	solute carrier family 6 (neurotransmitter transporter, noradrenalin), member 2 [Source:ZFIN;Acc:ZDB-GENE-110408-4]		no expression, predicted to be expressed in cns (neurotransmitter transporter in humans (OMIM)
ENSDARG00000017799	tgm1	transglutaminase 1, K polypeptide [Source:ZFIN;Acc:ZDB-GENE-030131-9032]		no expression
ENSDARG00000018303	pea3	ETS-domain transcription factor pea3 [Source:ZFIN;Acc:ZDB-GENE-990415-71]	tailbud, somite, segmental plate, pharyngeal arch, notochord and midbrain hindbrain boundary	
ENSDARG00000018321	lbx1a	ladybird homeobox 1a [Source:ZFIN;Acc:ZDB-GENE-040724-40]	cns	
ENSDARG00000019125	kbtbd5b	kelch repeat and BTB (POZ) domain containing 5b [Source:ZFIN;Acc:ZDB-GENE-060227-1]		no expression
ENSDARG00000022109	kcnip1a	Kv channel interacting protein 1 a [Source:ZFIN;Acc:ZDB-GENE-080721-17]		no expression
ENSDARG00000023261	scrt1a	scratch homolog 1, zinc finger protein a (Drosophila) [Source:ZFIN;Acc:ZDB-GENE-080219-30]	cns	
ENSDARG00000028664	ahsa1	AHA1, activator of heat shock protein ATPase homolog 1 (yeast) [Source:ZFIN;Acc:ZDB-GENE-030131-1869]	myotome (somites)	
ENSDARG00000028804	ankrd9	ankyrin repeat domain 9 [Source:ZFIN;Acc:ZDB-GENE-041114-3]		no expression
ENSDARG00000029710	lrrc30	leucine rich repeat containing 30 [Source:ZFIN;Acc:ZDB-GENE-050506-25]		expression not clear, may be ubiquitous?
ENSDARG00000029832	slc26a1	solute carrier family 26 (sulfate transporter), member 1 [Source:ZFIN;Acc:ZDB-GENE-070112-2182]	anterior part of pronephric duct	
ENSDARG00000030094	mal	mal, T-cell differentiation protein [Source:ZFIN;Acc:ZDB-GENE-050417-175]		expression not clear, may be ubiquitous?
ENSDARG00000030110	myod1	myogenic differentiation 1 [Source:ZFIN;Acc:ZDB-GENE-980526-561]	skeletal muscle	
ENSDARG00000030349	cryba2a	crystallin, beta A2a [Source:ZFIN;Acc:ZDB-GENE-040625-8]		no expression at stages examined by RNA-seq
ENSDARG00000031126	NOTUM (1 of 2)	notum pectinacetylesterase homolog (Drosophila) [Source:HGNC Symbol;Acc:27106]		no expression
ENSDARG00000035327	ckma	creatine kinase, muscle a [Source:ZFIN;Acc:ZDB-GENE-980526-109]	skeletal muscle and heart	
ENSDARG00000035735	gsx1	GS homeobox 1 [Source:ZFIN;Acc:ZDB-GENE-041008-135]	cns	
ENSDARG00000037030	casz1	castor zinc finger 1 [Source:ZFIN;Acc:ZDB-GENE-060130-108]	somite, heart primordium, and cns	
ENSDARG00000037558	asb12b	ankyrin repeat and SOCS box-containing protein 12b [Source:ZFIN;Acc:ZDB-GENE-040822-25]		no expression
ENSDARG00000038786	smarcd3a	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 3a [Source:ZFIN;Acc:ZDB-GENE-070912-491]	somites	
ENSDARG00000038990	wnt3	wingless-type MMTV integration site family, member 3 [Source:ZFIN;Acc:ZDB-GENE-081003-1]	neural tube	
ENSDARG00000039051	hhatla	hedgehog acyltransferase-like, a [Source:ZFIN;Acc:ZDB-GENE-040426-1865]	myotomes, somites	
ENSDARG00000039256	speg	SPEG complex locus [Source:ZFIN;Acc:ZDB-GENE-030131-3230]	cns	
ENSDARG00000039929	ckmt2	creatine kinase, mitochondrial 2 (sarcomeric) [Source:ZFIN;Acc:ZDB-GENE-040426-1654]	myotome, somites	
ENSDARG00000040266	sox19b	SRY-box containing gene 19b [Source:ZFIN;Acc:ZDB-GENE-010111-1]	spinal chord, other cns include midbrain, hindbrain, forebrain	
ENSDARG00000040816	si:dkey-27n14.1	si:dkey-27n14.1 [Source:ZFIN;Acc:ZDB-GENE-100922-123]		no expression
ENSDARG00000040925	wnt10b	wingless-type MMTV integration site family, member 10b [Source:ZFIN;Acc:ZDB-GENE-980526-524]	cns	
ENSDARG00000042857	TCF15	transcription factor 15 (basic helix-loop-helix) [Source:HGNC Symbol;Acc:11627]		no expression
ENSDARG00000043210	nfic	nuclear factor I/C [Source:ZFIN;Acc:ZDB-GENE-080305-2]		no expression
ENSDARG00000043328	fip1l1a	FIP1 like 1a (S. cerevisiae) [Source:ZFIN;Acc:ZDB-GENE-041014-8]		no expression
ENSDARG00000044155	mafa	v-maf musculoaponeurotic fibrosarcoma oncogene homolog A (avian) [Source:ZFIN;Acc:ZDB-GENE-010605-3]	somites and few spinal cord neurons at about 24hpf	
ENSDARG00000044550	hif1al2	hypoxia-inducible factor 1, alpha subunit, like 2 [Source:ZFIN;Acc:ZDB-GENE-050211-2]		no expression
ENSDARG00000045144	HBM	hemoglobin, mu [Source:HGNC Symbol;Acc:4826]		no expression
ENSDARG00000045219	dkk1b	dickkopf 1b [Source:ZFIN;Acc:ZDB-GENE-990708-5]	cns	
ENSDARG00000045687	pdzrn4	PDZ domain containing ring finger 4 [Source:ZFIN;Acc:ZDB-GENE-041210-237]	cns, spinal cord and forebrain midbrain boundary	
ENSDARG00000045695	myca	myelocytomatosis oncogene a [Source:ZFIN;Acc:ZDB-GENE-990415-162]	somite and cns	
ENSDARG00000051836	si:dkeyp-19e1.3	si:dkeyp-19e1.3 [Source:ZFIN;Acc:ZDB-GENE-081104-56]		no expression
ENSDARG00000052154	LIX1	Lix1 homolog (chicken) [Source:HGNC Symbol;Acc:18581]		no expression
ENSDARG00000052846	fsta	follistatin a [Source:ZFIN;Acc:ZDB-GENE-990714-11]	at stages examined in skeletal muscle and central nervous system	
ENSDARG00000053476	lipca	lipase, hepatic a [Source:ZFIN;Acc:ZDB-GENE-040426-1361]		no expression
ENSDARG00000053493	aldh1a2	aldehyde dehydrogenase 1 family, member A2 [Source:ZFIN;Acc:ZDB-GENE-011010-3]	somites	
ENSDARG00000053535	lmo7b	LIM domain only 7b [Source:ZFIN;Acc:ZDB-GENE-060825-242]	somites	
ENSDARG00000055151	smyd2a	SET and MYND domain containing 2a [Source:ZFIN;Acc:ZDB-GENE-050320-126]	heart and muscle	set domain- histone methylation
ENSDARG00000055294	atoh1a	atonal homolog 1a [Source:ZFIN;Acc:ZDB-GENE-990415-17]	cns	
ENSDARG00000055618	acta1b	actin, alpha 1b, skeletal muscle [Source:ZFIN;Acc:ZDB-GENE-030131-55]	somites	
ENSDARG00000056028	slc22a7a	solute carrier family 22 (organic anion transporter), member 7a [Source:ZFIN;Acc:ZDB-GENE-050506-52]	neural crest	
ENSDARG00000056650	zgc:113276	zgc:113276 [Source:ZFIN;Acc:ZDB-GENE-050522-7]	somites	
ENSDARG00000057936	lhx5	LIM homeobox 5 [Source:ZFIN;Acc:ZDB-GENE-980526-484]	cns	
ENSDARG00000058543	lama5	laminin, alpha 5 [Source:ZFIN;Acc:ZDB-GENE-030131-9823]	cns, and other tissues like notocord, epidermis and pronephros	
ENSDARG00000059002	CABZ01080098.1	Uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:F1QR67]		no expression
ENSDARG00000061416	C2CD4A	C2 calcium-dependent domain containing 4A [Source:HGNC Symbol;Acc:33627]		no expression
ENSDARG00000061471	sema5b	sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5B [Source:ZFIN;Acc:ZDB-GENE-0709		no expression
ENSDARG00000062662	LMOD3	leiomodin 3 (fetal) [Source:HGNC Symbol;Acc:6649]		no expression
ENSDARG00000063257	PLD5 (1 of 2)	phospholipase D family, member 5 [Source:HGNC Symbol;Acc:26879]		no expression
ENSDARG00000068397	TENC1 (2 of 2)	tensin like C1 domain containing phosphatase (tensin 2) [Source:HGNC Symbol;Acc:19737]		no expression
ENSDARG00000068719	GMNC	geminin coiled-coil domain containing [Source:HGNC Symbol;Acc:40049]		no expression
ENSDARG00000069054	glula	glutamate-ammonia ligase (glutamine synthase) a [Source:ZFIN;Acc:ZDB-GENE-030131-688]	somites, and also in a few cns tissues and blood.	
ENSDARG00000069660	chrnb1	cholinergic receptor, nicotinic, beta 1 (muscle) [Source:ZFIN;Acc:ZDB-GENE-070821-4]	somites	
ENSDARG00000069763	etv5a	ets variant 5a [Source:ZFIN;Acc:ZDB-GENE-030131-9572]	several structures including spinal chord neurons and pronephros	
ENSDARG00000069843	kctd12.1	potassium channel tetramerisation domain containing 12.1 [Source:ZFIN;Acc:ZDB-GENE-030902-5]	at time examined, notochord	
ENSDARG00000070157	tgm2a	transglutaminase 2, C polypeptide A [Source:ZFIN;Acc:ZDB-GENE-040912-78]	somites	
ENSDARG00000070546	msgn1	mesogenin 1 [Source:ZFIN;Acc:ZDB-GENE-030722-1]	skeletal muscle, paraxial mesoderm	
ENSDARG00000070818	pax7b	paired box gene 7b [Source:ZFIN;Acc:ZDB-GENE-080917-54]	somites and neural crest	
ENSDARG00000071445	myoz1b	myozenin 1b [Source:ZFIN;Acc:ZDB-GENE-040718-146]	myotome, somites	
ENSDARG00000071496	zic6	zic family member 6 [Source:ZFIN;Acc:ZDB-GENE-061207-1]		not enough images
ENSDARG00000073732	myh14	myosin, heavy chain 14, non-muscle [Source:ZFIN;Acc:ZDB-GENE-100921-1]		no expression
ENSDARG00000075524	CABZ01044023.1	Uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:F1QKB9]		no expression
ENSDARG00000075986	RIC3 (2 of 2)	resistance to inhibitors of cholinesterase 3 homolog (C. elegans) [Source:HGNC Symbol;Acc:30338]		no expression
ENSDARG00000076571	sp5	Sp5 transcription factor [Source:ZFIN;Acc:ZDB-GENE-030430-2]		not enough images
ENSDARG00000077295	akap6	A kinase (PRKA) anchor protein 6 [Source:ZFIN;Acc:ZDB-GENE-120207-2]		no expression
ENSDARG00000077826	CCDC48	coiled-coil domain containing 48 [Source:HGNC Symbol;Acc:25692]		no expression
ENSDARG00000077938	cd248b	CD248 molecule, endosialin b [Source:ZFIN;Acc:ZDB-GENE-050309-46]	anterior somites	
ENSDARG00000078227	cspg4	chondroitin sulfate proteoglycan 4 [Source:ZFIN;Acc:ZDB-GENE-050519-1]		no expression
ENSDARG00000078701	prdm13	PR domain containing 13 [Source:ZFIN;Acc:ZDB-GENE-080616-1]	spinal chord and neurons	in set domain family - histone methylation
ENSDARG00000078842	TNS1 (2 of 2)	tensin 1 [Source:HGNC Symbol;Acc:11973]		no expression
ENSDARG00000079198	usp13	ubiquitin specific peptidase 13 (isopeptidase T-3) [Source:ZFIN;Acc:ZDB-GENE-080724-7]		no expression
ENSDARG00000086647	chrng	cholinergic receptor, nicotinic, gamma [Source:ZFIN;Acc:ZDB-GENE-030131-3805]	somites	
ENSDARG00000088157	SMOC2 (1 of 2)	SPARC related modular calcium binding 2 [Source:HGNC Symbol;Acc:20323]		no expression
ENSDARG00000088484	KCNJ1 (6 of 7)	potassium inwardly-rectifying channel, subfamily J, member 1 [Source:HGNC Symbol;Acc:6255]		no expression
ENSDARG00000088881	si:dkeyp-69b9.3	si:dkeyp-69b9.3 [Source:ZFIN;Acc:ZDB-GENE-100922-275]		no expression
ENSDARG00000089102	CABZ01079873.2	Uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:E7F815]		no expression
ENSDARG00000090337	pprc1	peroxisome proliferator-activated receptor gamma, coactivator-related 1 [Source:ZFIN;Acc:ZDB-GENE-030131-9858]	somites and cns	
ENSDARG00000090371	BX664721.5	Uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:E7EZC5]		no expression
ENSDARG00000090974	si:dkey-114e9.1	si:dkey-114e9.1 [Source:ZFIN;Acc:ZDB-GENE-100922-20]		no expression
ENSDARG00000091085	lepa	leptin a [Source:ZFIN;Acc:ZDB-GENE-081001-1]		not enough images
ENSDARG00000091260	MYLK4 (1 of 2)	myosin light chain kinase family, member 4 [Source:HGNC Symbol;Acc:27972]		no expression
ENSDARG00000091660	fgfbp1	fibroblast growth factor binding protein 1 [Source:ZFIN;Acc:ZDB-GENE-081106-1]	ubiquitous	
ENSDARG00000092033	si:dkey-239h2.3	si:dkey-239h2.3 [Source:ZFIN;Acc:ZDB-GENE-081104-383]		no expression
ENSDARG00000092638	ptprua	protein tyrosine phosphatase, receptor type, U, a [Source:ZFIN;Acc:ZDB-GENE-060503-232]		poor image, not enough information
ENSDARG00000093678	si:dkey-102m7.3	si:dkey-102m7.3 [Source:ZFIN;Acc:ZDB-GENE-091204-389]		no expression
ENSDARG00000094379	ascl1b	achaete-scute complex-like 1b (Drosophila) [Source:ZFIN;Acc:ZDB-GENE-980526-174]	cns	
ENSDARG00000094561	si:ch211-69i14.10	si:ch211-69i14.10 [Source:ZFIN;Acc:ZDB-GENE-070705-187]		no expression
ENSDARG00000095885	si:ch211-158m24.12	si:ch211-158m24.12 [Source:ZFIN;Acc:ZDB-GENE-110411-14]		no expression

RED = expressed in nervous system, BLUE = expressed in skeletal muscle or somites, LAVENDER = expressed in somites and nervous system, GREEN = expressed in other tissues, NO COLOR = no expression available

Concluding Remarks

The work presented here provides tools to further understand the development of the circulatory system, in particular in zebrafish, which has become a valuable vertebrate model organism in which to study vascular, cardiac, and hematopoietic development. In chapters 1 and 2 we presented the identification of 50 genes that are expressed in the early stages of vascular development in zebrafish and 3 genes that are expressed in primitive myeloid cells. Various signaling pathways and mechanisms by which the vascular system is formed have been published in the literature, and it is our hope that through future research other investigators will identify functional relevance for some of these genes. On our part we have knocked down ten (*dgki*, *gpr183*, *cntn4*, *myo1e*, *she*, *rgl2*, *tmem119*, *cdc42gap1*, *camgap1*, *klhl4*) with morpholino oligomers but did not identify obvious phenotypic defects in the architectural formation of the vasculature for any of these. This might be due to redundancy with other factors or because we did not look at other cell biology attributes or physiological consequences. However our efforts have produced a substantial number of potential genes that we hope will serve the community to continue to unravel the biological mechanisms by which the circulatory system is developed.

In chapter 3 by subtracting the excess set of genes bound by Etv2 following Etv2-mCherry overexpression with RNA-seq datasets, we found that the majority of genes directly induced by Etsrp/Etv2 are vascular. Chapter 3 provides several points worth emphasizing. In previous work performed in our lab we found that Etv2 autonomously regulates the myeloid lineage, but the absence of Etv2 expression in myeloid derivatives created a gap in our understanding by which Etv2 induces this lineage (Sumanas et al., 2008). Here we show that Etv2 indirectly induces the myeloid lineage by inducing Scl, which subsequently specifies the myeloid lineage by inducing Pu1 (Patterson et al., 2005). Interestingly by lineage tracing of

Etv2-GFP transgenic cells, GFP positive myeloid cells have been identified as descendants of Etv2-GFP progenitors (Proulx et al., 2010). Combined, this shows that in zebrafish Etv2 induces the myeloid lineage by activating Scl within hemangioblasts and subsequently remains asymetrically associated with angioblasts and then with differentiated endothelial cells. A second point worth mentioning is that similar to Etv2's indirect regulation of the myeloid lineage, Etv2 is likely to indirectly regulate the definitive hematopoietic lineage through Scl in the hemogenic endothelium, as alluded to in a separate publication we produced (Ren et al., 2010). The final point worth noting is that besides the induction of Scl and Fli1a, which can partially rescue endothelial specification in Etv2 deficient embryos (Ren et al., 2010), Etv2 directly induces a wide breadth of molecules, but not every single gene expressed in the developing vasculature.

In future work it will be interesting to see whether VEGF signaling is also required for the induction of Etv2 in zebrafish as observed in mice and as mentioned in the discussion of Chapter 3. It will also be useful to generate a zebrafish specific Etv2 ChIPable antibody to identify the direct targets of the endogenous protein, and thereby determine the accuracy of the binding sites predicted by the data presented here. Since we deduced a list of direct genetic targets based on a distance of Etv2 ChIP-seq peaks within 10 Kb of the transcription start site of genes induced by Etv2 overexpression and downregulated in the absence of Etv2, we might have missed many long range Etv2 bound distal enhancers. This can be addressed in the future by chromosome conformation capture experiments. It will also be helpful to obtain Etv2 ChIP-seq data in mammals to understand the evolutionary similarities and differences of Etv2 binding at cis regulatory elements between these organisms. We hope that the data generated here will provide valuable information in future investigations of blood and vascular gene regulation.

Bibliography

Aday, A.W., Zhu, L.J., Lakshmanan, A., Wang, J., and Lawson, N.D. (2011). Identification of cis regulatory features in the embryonic zebrafish genome through large-scale profiling of H3K4me1 and H3K4me3 binding sites. Dev Biol *357*, 450-462.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol *11*, R106.

Bailey, T.L. (2011). DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics *27*, 1653-1659.

Barton, A., Eyre, S., Ke, X., Hinks, A., Bowes, J., Flynn, E., Martin, P., Consortium, Y., Consortium, B., Wilson, A.G.*, et al.* (2009). Identification of AF4/FMR2 family, member 3 (AFF3) as a novel rheumatoid arthritis susceptibility locus and confirmation of two further pan-autoimmune susceptibility genes. Hum Mol Genet *18*, 2518-2522.

Bauer, H., Willert, J., Koschorz, B., and Herrmann, B.G. (2005). The t complex-encoded GTPase-activating protein Tagap1 acts as a transmission ratio distorter in mice. Nat Genet *37*, 969-973.

Bernatchez, P.N., Acevedo, L., Fernandez-Hernando, C., Murata, T., Chalouni, C., Kim, J., Erdjument-Bromage, H., Shah, V., Gratton, J.-P., McNally, E.M.*, et al.* (2007). Myoferlin regulates vascular endothelial growth factor receptor-2 stability and function. J Biol Chem *282*, 30745-30753.

Bernatchez, P.N., Sharma, A., Kodaman, P., and Sessa, W.C. (2009). Myoferlin is critical for endocytosis in endothelial cells. Am J Physiol, Cell Physiol *297*, C484-492.

Bläser, S., Jersch, K., Hainmann, I., Wunderle, D., Zgaga-Griesz, A., Busse, A., and Zieger, B. (2002). Human septin-septin interaction: CDCrel-1 partners with KIAA0202. FEBS Lett *519*, 169-172.

Bläser, S., Röseler, S., Rempp, H., Bartsch, I., Bauer, H., Lieber, M., Lessmann, E., Weingarten, L., Busse, A., Huber, M.*, et al.* (2006). Human endothelial cell septins: SEPT11 is an interaction partner of SEPT5. J Pathol *210*, 103-110.

Bolinger, B., Krebs, P., Tian, Y., Engeler, D., Scandella, E., Miller, S., Palmer, D.C., Restifo, N.P., Clavien, P.-A., and Ludewig, B. (2008). Immunologic ignorance of vascular endothelial cells expressing minor histocompatibility antigen. Blood *111*, 4588-4595.

Brown, T.A., and McKnight, S.L. (1992). Specificities of protein-protein and protein-DNA interaction of GABP alpha and two newly defined ets-related proteins. Genes Dev *6*, 2502-2512.

Burns, C.E., Traver, D., Mayhall, E., Shepard, J.L., and Zon, L.I. (2005). Hematopoietic stem cell fate is established by the Notch-Runx pathway. Genes Dev *19*, 2331-2342.

Bussmann, J., Bakkers, J., and Schulte-Merker, S. (2007). Early endocardial morphogenesis requires Scl/Tal1. PLoS Genet *3*, e140.

Bussmann, J., Lawson, N., Zon, L., Schulte-Merker, S., and Committee, Z.N. (2008). Zebrafish VEGF receptors: a guideline to nomenclature. PLoS Genet *4*, e1000064.

Caltagarone, J., Rhodes, J., Honer, W.G., and Bowser, R. (1998). Localization of a novel septin protein, hCDCrel-1, in neurons of human brain. Neuroreport *9*, 2907-2912.

Cao, Y., Zhao, H., and Grunz, H. (2001). XCL-2 is a novel m-type calpain and disrupts morphogenetic movements during embryogenesis in Xenopus laevis. Dev Growth Differ *43*, 563-571.

Cermenati, S., Moleri, S., Cimbro, S., Corti, P., Del Giacco, L., Amodeo, R., Dejana, E., Koopman, P., Cotelli, F., and Beltrame, M. (2008). Sox18 and Sox7 play redundant roles in vascular development. Blood *111*, 2657-2666.

Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J.C., and Keller, G. (1998). A common precursor for hematopoietic and endothelial cells. Development *125*, 725-732.

Coghill, I.D., Brown, S., Cottle, D.L., McGrath, M.J., Robinson, P.A., Nandurkar, H.H., Dyson, J.M., and Mitchell, C.A. (2003). FHL3 is an actin-binding protein that regulates alpha-actinin-mediated actin bundling: FHL3 localizes to actin stress fibers and enhances cell spreading and stress fiber disassembly. J Biol Chem *278*, 24139-24152.

Cottle, D.L., McGrath, M.J., Cowling, B.S., Coghill, I.D., Brown, S., and Mitchell, C.A. (2007). FHL3 binds MyoD and negatively regulates myotube formation. J Cell Sci *120*, 1423-1435.

Covassin, L.D., Villefranc, J.A., Kacergis, M.C., Weinstein, B.M., and Lawson, N.D. (2006). Distinct genetic interactions between multiple Vegf receptors are required for development of different blood vessel types in zebrafish. Proc Natl Acad Sci USA *103*, 6554-6559.

Crozet, F., el Amraoui, A., Blanchard, S., Lenoir, M., Ripoll, C., Vago, P., Hamel, C., Fizames, C., Levi-Acobas, F., Depétris, D.*, et al.* (1997). Cloning of the genes encoding two murine and human cochlear unconventional type I myosins. Genomics *40*, 332-341.

De Val, S., Chi, N.C., Meadows, S.M., Minovitsky, S., Anderson, J.P., Harris, I.S., Ehlers, M.L., Agarwal, P., Visel, A., Xu, S.-M.*, et al.* (2008). Combinatorial regulation of endothelial gene expression by ets and forkhead transcription factors. Cell *135*, 1053-1064.

Ding, L., Traer, E., McIntyre, T.M., Zimmerman, G.A., and Prescott, S.M. (1998). The cloning and characterization of a novel human diacylglycerol kinase, DGKiota. J Biol Chem *273*, 32746-32752.

Ding, L., Wang, Z., Yan, J., Yang, X., Liu, A., Qiu, W., Zhu, J., Han, J., Zhang, H., Lin, J.*, et al.* (2009). Human four-and-a-half LIM family members suppress tumor cell growth through a TGF-beta-like signaling pathway. J Clin Invest *119*, 349-361.

Dunlevy, J.R., Berryhill, B.L., Vergnes, J.P., SundarRaj, N., and Hassell, J.R. (1999). Cloning, chromosomal localization, and characterization of cDNA from a novel gene, SH3BP4, expressed by human corneal fibroblasts. Genomics *62*, 519-524.

Ferdous, A., Caprioli, A., Iacovino, M., Martin, C.M., Morris, J., Richardson, J.A., Latif, S., Hammer, R.E., Harvey, R.P., Olson, E.N.*, et al.* (2009). Nkx2-5 transactivates the Ets-related protein 71 gene and specifies an endothelial/endocardial fate in the developing embryo. Proc Natl Acad Sci USA *106*, 814-819.

Fernandez, T., Morgan, T., Davis, N., Klin, A., Morris, A., Farhi, A., Lifton, R.P., and State, M.W. (2004). Disruption of contactin 4 (CNTN4) results in developmental delay and other features of 3p deletion syndrome. Am J Hum Genet *74*, 1286-1293.

Ferro, E., Magrini, D., Guazzi, P., Fischer, T.H., Pistolesi, S., Pogni, R., White, G.C., and Trabalzini, L. (2008). G-protein binding features and regulation of the RalGDS family member, RGL2. Biochem J *415*, 145-154.

Ferro, E., and Trabalzini, L. (2010). RalGDS family members couple Ras to Ral signalling and that's not all. Cell Signal *22*, 1804-1810.

Fraisl, P., Tanaka, H., Forss-Petter, S., Lassmann, H., Nishimune, Y., and Berger, J. (2006). A novel mammalian bubblegum-related acyl-CoA synthetase restricted to testes and possibly involved in spermatogenesis. Arch Biochem Biophys *451*, 23-33.

Fujii, N., Hiraki, A., Ikeda, K., Ohmura, Y., Nozaki, I., Shinagawa, K., Ishimaru, F., Kiura, K., Shimizu, N., Tanimoto, M.*, et al.* (2002). Expression of minor histocompatibility antigen, HA-1, in solid tumor cells. Transplantation *73*, 1137-1141.

Gerber, S.A., and Pober, J.S. (2008). IFN-alpha induces transcription of hypoxia-inducible factor-1alpha to inhibit proliferation of human endothelial cells. J Immunol *181*, 1052-1062.

Gomez, G., Lee, J.-H., Veldman, M.B., Lu, J., Xiao, X., and Lin, S. (2012). Identification of vascular and hematopoietic genes downstream of etsrp by deep sequencing in zebrafish. PLoS ONE *7*, e31658.

Gomez, G.A., Veldman, M.B., Zhao, Y., Burgess, S., and Lin, S. (2009). Discovery and characterization of novel vascular and hematopoietic genes downstream of etsrp in zebrafish. PLoS ONE *4*, e4994.

Hata, S., Koyama, S., Kawahara, H., Doi, N., Maeda, T., Toyama-Sorimachi, N., Abe, K., Suzuki, K., and Sorimachi, H. (2006). Stomach-specific calpain, nCL-2, localizes in mucus cells and proteolyzes the beta-subunit of coatomer complex, beta-COP. J Biol Chem *281*, 11214-11224.

Hayashi, K., Yano, H., Hashida, T., Takeuchi, R., Takeda, O., Asada, K., Takahashi, E., Kato, I., and Sobue, K. (1992). Genomic structure of the human caldesmon gene. Proc Natl Acad Sci USA *89*, 12122-12126.

Hitomi, J., Christofferson, D.E., Ng, A., Yao, J., Degterev, A., Xavier, R.J., and Yuan, J. (2008). Identification of a molecular signaling network that regulates a cellular necrotic cell death pathway. Cell *135*, 1311-1323.

Hiwatari, M., Taki, T., Taketani, T., Taniwaki, M., Sugita, K., Okuya, M., Eguchi, M., Ida, K., and Hayashi, Y. (2003). Fusion of an AF4-related gene, LAF4, to MLL in childhood acute lymphoblastic leukemia with t(2;11)(q11;q23). Oncogene *22*, 2851-2855.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009a). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Huang, S., Gilfillan, S., Kim, S., Thompson, B., Wang, X., Sant, A.J., Fremont, D.H., Lantz, O., and Hansen, T.H. (2008). MR1 uses an endocytic pathway to activate mucosal-associated invariant T cells. J Exp Med *205*, 1201-1211.

Huang, S., Martin, E., Kim, S., Yu, L., Soudais, C., Fremont, D.H., Lantz, O., and Hansen, T.H. (2009b). MR1 antigen presentation to mucosal-associated invariant T cells was highly conserved in evolution. Proc Natl Acad Sci USA *106*, 8290-8295.

Huber, T.L., Kouskoff, V., Fehling, H.J., Palis, J., and Keller, G. (2004). Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. Nature *432*, 625-630.

Hunt, K.A., Zhernakova, A., Turner, G., Heap, G.A.R., Franke, L., Bruinenberg, M., Romanos, J., Dinesen, L.C., Ryan, A.W., Panesar, D*., et al.* (2008). Newly identified genetic risk variants for celiac disease related to the immune response. Nat Genet *40*, 395-402.

Ishitobi, H., Wakamatsu, A., Liu, F., Azami, T., Hamada, M., Matsumoto, K., Kataoka, H., Kobayashi, M., Choi, K., Nishikawa, S.-I*., et al.* (2011). Molecular basis for Flk1 expression in hemato-cardiovascular progenitors in the mouse. Development *138*, 5357-5368.

Isomura, M., Okui, K., Fujiwara, T., Shin, S., and Nakamura, Y. (1996). Isolation and mapping of RAB2L, a human cDNA that encodes a protein homologous to RalGDS. Cytogenet Cell Genet *74*, 263-265.

Jin, S.-W., Beis, D., Mitchell, T., Chen, J.-N., and Stainier, D.Y.R. (2005). Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. Development *132*, 5199-5209.

Kalev-Zylinska, M.L., Horsfield, J.A., Flores, M.V.C., Postlethwait, J.H., Vitas, M.R., Baas, A.M., Crosier, P.S., and Crosier, K.E. (2002). Runx1 is required for zebrafish blood and vessel development and expression of a human RUNX1-CBF2T1 transgene advances a model for studies of leukemogenesis. Development *129*, 2015-2030.

Kaneko-Goto, T., Yoshihara, S.-I., Miyazaki, H., and Yoshihara, Y. (2008). BIG-2 mediates olfactory axon convergence to target glomeruli. Neuron *57*, 834-846.

Kataoka, H., Hayashi, M., Nakagawa, R., Tanaka, Y., Izumi, N., Nishikawa, S., Jakt, M.L., Tarui, H., and Nishikawa, S.-I. (2011). Etv2/ER71 induces vascular mesoderm from Flk1+PDGFRα+ primitive mesoderm. Blood *118*, 6975-6986.

Katoh, Y., and Katoh, M. (2004). Identification and characterization of ARHGAP27 gene in silico. Int J Mol Med *14*, 943-947.

Kelly, M.D., Essex, D.W., Shapiro, S.S., Meloni, F.J., Druck, T., Huebner, K., and Konkle, B.A. (1994). Complementary DNA cloning of the alternatively expressed endothelial cell glycoprotein Ib beta (GPIb beta) and localization of the GPIb beta gene to chromosome 22. J Clin Invest *93*, 2417-2424.

Khanobdee, K., Kolberg, J.B., and Dunlevy, J.R. (2004). Nuclear and plasma membrane localization of SH3BP4 in retinal pigment epithelial cells. Mol Vis *10*, 933-942.

Kim, S.V., Mehal, W.Z., Dong, X., Heinrich, V., Pypaert, M., Mellman, I., Dembo, M., Mooseker, M.S., Wu, D., and Flavell, R.A. (2006). Modulation of cell adhesion and motility in the immune system by Myo1f. Science *314*, 136-139.

Klein, C.A., Wilke, M., Pool, J., Vermeulen, C., Blokland, E., Burghart, E., Krostina, S., Wendler, N., Passlick, B., Riethmüeller, G.*, et al.* (2002). The hematopoietic system-specific minor histocompatibility antigen HA-1 shows aberrant expression in epithelial cancer cells. J Exp Med *196*, 359-368.

Koyano-Nakagawa, N., Kweon, J., Iacovino, M., Shi, X., Rasmussen, T.L., Borges, L., Zirbes, K.M., Li, T., Perlingeiro, R.C.R., Kyba, M.*, et al.* (2012). Etv2 is expressed in the yolk sac hematopoietic and endothelial progenitors and regulates Lmo2 gene expression. Stem Cells *30*, 1611-1623.

Kraus, T.A., Lau, J.F., Parisien, J.-P., and Horvath, C.M. (2003). A hybrid IRF9-STAT2 protein recapitulates interferon-stimulated gene expression and antiviral response. J Biol Chem *278*, 13033-13038.

Lam, E.Y.N., Hall, C.J., Crosier, P.S., Crosier, K.E., and Flores, M.V. (2010). Live imaging of Runx1 expression in the dorsal aorta tracks the emergence of blood progenitors from endothelial cells. Blood *116*, 909-914.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Lee, D., Kim, T., and Lim, D.-S. (2011). The Er71 is an important regulator of hematopoietic stem cells in adult mice. Stem Cells *29*, 539-548.

Lee, D., Park, C., Lee, H., Lugus, J.J., Kim, S.H., Arentson, E., Chung, Y.S., Gomez, G., Kyba, M., Lin, S.*, et al.* (2008). ER71 acts downstream of BMP, Notch, and Wnt signaling in blood and vessel progenitor specification. Cell Stem Cell *2*, 497-507.

Lepage, S.E., and Bruce, A.E.E. (2008). Characterization and comparative expression of zebrafish calpain system genes during early development. Dev Dyn *237*, 819-829.

Li, Y., Lin, J.L.C., Reiter, R.S., Daniels, K., Soll, D.R., and Lin, J.J.C. (2004). Caldesmon mutant defective in Ca(2+)-calmodulin binding interferes with assembly of stress fibers and affects cell morphology, growth and motility. J Cell Sci *117*, 3593-3604.

Liu, F., Kang, I., Park, C., Chang, L.-W., Wang, W., Lee, D., Lim, D.-S., Vittet, D., Nerbonne, J.M., and Choi, K. (2012). ER71 specifies Flk-1+ hemangiogenic mesoderm by inhibiting cardiac mesoderm and Wnt signaling. Blood *119*, 3295-3305.

Liu, F., and Patient, R. (2008). Genome-Wide Analysis of the Zebrafish ETS Family Identifies Three Genes Required for Hemangioblast Differentiation or Angiogenesis. Circulation Research *103*, 1147-1154.

Lockyer, P.J., Kupzig, S., and Cullen, P.J. (2001). CAPRI regulates Ca(2+)-dependent inactivation of the Ras-MAPK pathway. Curr Biol *11*, 981-986.

Lu, C., Kasik, J., Stephan, D.A., Yang, S., Sperling, M.A., and Menon, R.K. (2001). Grtp1, a novel gene regulated by growth hormone. Endocrinology *142*, 4568-4571.

Luker, K.E., Pica, C.M., Schreiber, R.D., and Piwnica-Worms, D. (2001). Overexpression of IRF9 confers resistance to antimicrotubule agents in breast cancer cells. Cancer Res *61*, 6540-6547.

Luo, B., Regier, D.S., Prescott, S.M., and Topham, M.K. (2004). Diacylglycerol kinases. Cell Signal *16*, 983-989.

Ma, C., and Staudt, L.M. (1996). LAF-4 encodes a lymphoid nuclear protein with transactivation potential that is homologous to AF-4, the gene fused to MLL in t(4;11) leukemias. Blood *87*, 734-745.

Ma, Q., Wang, H., Guo, R., Wang, H., Ge, Y., Ma, J., Xue, S., and Han, D. (2006). Molecular cloning and characterization of SRG-L, a novel mouse gene developmentally expressed in spermatogenic cells. Mol Reprod Dev *73*, 1075-1083.

Mao, M., Biery, M.C., Kobayashi, S.V., Ward, T., Schimmack, G., Burchard, J., Schelter, J.M., Dai, H., He, Y.D., and Linsley, P.S. (2004). T lymphocyte activation gene identification by coregulated expression on DNA microarrays. Genomics *83*, 989-999.

McClintock, T.S., Glasser, C.E., Bose, S.C., and Bergman, D.A. (2008). Tissue expression patterns identify mouse cilia genes. Physiol Genomics *32*, 198-206.

McKie, J.M., Sutherland, H.F., Harvey, E., Kim, U.J., and Scambler, P.J. (1997). A human gene similar to Drosophila melanogaster peanut maps to the DiGeorge syndrome region of 22q11. Hum Genet *101*, 6-12.

Meeson, A.P., Shi, X., Alexander, M.S., Williams, R.S., Allen, R.E., Jiang, N., Adham, I.M., Goetsch, S.C., Hammer, R.E., and Garry, D.J. (2007). Sox15 and Fhl3 transcriptionally coactivate Foxk1 and regulate myogenic progenitor cells. EMBO J *26*, 1902-1912.

Mizutani, Y., Kihara, A., and Igarashi, Y. (2006). LASS3 (longevity assurance homologue 3) is a mainly testis-specific (dihydro)ceramide synthase with relatively broad substrate specificity. Biochem J *398*, 531-538.

Mosavi, L.K., Cammett, T.J., Desrosiers, D.C., and Peng, Z.-Y. (2004). The ankyrin repeat as molecular architecture for protein recognition. Protein Sci *13*, 1435-1448.

Murray, P. (1932). The Development in vitro of the Blood of the Early Chick Embryo. Proceedings of the Royal Society of London *111*, 497-521.

Nagase, T., Ishikawa, K., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N., and Ohara, O. (1998). Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro. DNA Res *5*, 31-39.

Neuhaus, H., Müller, F., and Hollemann, T. (2010). Xenopus er71 is involved in vascular development. Dev Dyn *239*, 3436-3445.

Niikura, T., Hirata, R., and Weil, S.C. (1997). A novel interferon-inducible gene expressed during myeloid differentiation. Blood Cells Mol Dis *23*, 337-349.

Ochi, H., Hans, S., and Westerfield, M. (2008). Smarcd3 regulates the timing of zebrafish myogenesis onset. J Biol Chem *283*, 3529-3536.

Osterfield, M., Egelund, R., Young, L.M., and Flanagan, J.G. (2008). Interaction of amyloid precursor protein with contactins and NgCAM in the retinotectal system. Development *135*, 1189-1199.

Palencia-Desai, S., Kohli, V., Kang, J., Chi, N.C., Black, B.L., and Sumanas, S. (2011). Vascular endothelial and endocardial progenitors differentiate as cardiomyocytes in the absence of Etsrp/Etv2 function. Development *138*, 4721-4732.

Patterson, L.J., Gering, M., and Patient, R. (2005). Scl is required for dorsal aorta as well as blood formation in zebrafish embryos. Blood *105*, 3502-3511.

Pei, Z., Jia, Z., and Watkins, P.A. (2006). The second member of the human and murine bubblegum family is a testis- and brainstem-specific acyl-CoA synthetase. J Biol Chem *281*, 6632-6641.

Peng, X.-R., Jia, Z., Zhang, Y., Ware, J., and Trimble, W.S. (2002). The septin CDCrel-1 is dispensable for normal development and neurotransmitter release. Mol Cell Biol *22*, 378-387.

Peterkin, T., Gibson, A., and Patient, R. (2009). Common genetic control of haemangioblast and cardiac development in zebrafish. Development *136*, 1465-1474.

Peterson, S.N., Trabalzini, L., Brtva, T.R., Fischer, T., Altschuler, D.L., Martelli, P., Lapetina, E.G., Der, C.J., and White, G.C. (1996). Identification of a novel RalGDS-related protein as a candidate effector for Ras and Rap1. J Biol Chem *271*, 29903-29908.

Pham, V.N., Lawson, N.D., Mugford, J.W., Dye, L., Castranova, D., Lo, B., and Weinstein, B.M. (2007). Combinatorial function of ETS transcription factors in the developing vasculature. Dev Biol *303*, 772-783.

Post, G.R., Swiderski, C., Waldrop, B.A., Salty, L., Glembotski, C.C., Wolthuis, R.M.F., and Mochizuki, N. (2002). Guanine nucleotide exchange factor-like factor (Rlf) induces gene expression and potentiates alpha 1-adrenergic receptor-induced transcriptional responses in neonatal rat ventricular myocytes. J Biol Chem *277*, 15286-15292.

Proulx, K., Lu, A., and Sumanas, S. (2010). Cranial vasculature in zebrafish forms by angioblast cluster-derived angiogenesis. Developmental Biology *348*, 34-46.

Qian, F., Zhen, F., Ong, C., Jin, S.-W., Meng Soo, H., Stainier, D.Y.R., Lin, S., Peng, J., and Wen, Z. (2005). Microarray analysis of zebrafish cloche mutant using amplified cDNA and identification of potential downstream target genes. Dev Dyn *233*, 1163-1172.

Qian, F., Zhen, F., Xu, J., Huang, M., Li, W., and Wen, Z. (2007). Distinct functions for different scl isoforms in zebrafish primitive and definitive hematopoiesis. Plos Biol *5*, e132.

Rasmussen, T.L., Kweon, J., Diekmann, M.A., Belema-Bedada, F., Song, Q., Bowlin, K., Shi, X., Ferdous, A., Li, T., Kyba, M.*, et al.* (2011). ER71 directs mesodermal fate decisions during embryogenesis. Development *138*, 4801-4812.

Rasmussen, T.L., Shi, X., Wallis, A., Kweon, J., Zirbes, K.M., Koyano-Nakagawa, N., and Garry, D.J. (2012). VEGF/Flk1 Signaling Cascade Transactivates Etv2 Gene Expression. PLoS ONE *7*, e50103.

Reed, J.C., Doctor, K.S., and Godzik, A. (2004). The domains of apoptosis: a genomics perspective. Sci STKE *2004*, re9.

Regier, D.S., Higbee, J., Lund, K.M., Sakane, F., Prescott, S.M., and Topham, M.K. (2005). Diacylglycerol kinase iota regulates Ras guanyl-releasing protein 3 and inhibits Rap1 signaling. Proc Natl Acad Sci USA *102*, 7595-7600.

Ren, X., Gomez, G.A., Zhang, B., and Lin, S. (2010). Scl isoforms act downstream of etsrp to specify angioblasts and definitive hematopoietic stem cells. Blood *115*, 5338-5346.

Robb, L., Lyons, I., Li, R., Hartley, L., Köntgen, F., Harvey, R.P., Metcalf, D., and Begley, C.G. (1995). Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene. Proc Natl Acad Sci USA *92*, 7075-7079.

Romanos, J., Barisani, D., Trynka, G., Zhernakova, A., Bardella, M.T., and Wijmenga, C. (2009). Six new coeliac disease loci replicated in an Italian population confirm association with coeliac disease. J Med Genet *46*, 60-63.

Roohi, J., Montagna, C., Tegay, D.H., Palmer, L.E., DeVincent, C., Pomeroy, J.C., Christian, S.L., Nowak, N., and Hatchwell, E. (2009). Disruption of contactin 4 in three subjects with autism spectrum disorder. J Med Genet *46*, 176-182.

Sabin, F.R. (1917). Preliminary note on the differentiation of angioblasts and the method by which they produce blood-vessels, blood-plasma and red blood-cells as seen in the living chick. 1917. J Hematother Stem Cell Res *11*, 5-7.

Sakakibara, T., Nemoto, Y., Nukiwa, T., and Takeshima, H. (2004). Identification and characterization of a novel Rho GTPase activating protein implicated in receptor-mediated endocytosis. FEBS Lett *566*, 294-300.

Salanga, M.C., Meadows, S.M., Myers, C.T., and Krieg, P.A. (2010). ETS family protein ETV2 is required for initiation of the endothelial lineage but not the hematopoietic lineage in the Xenopus embryo. Dev Dyn *239*, 1178-1187.

Schoenebeck, J.J., Keegan, B.R., and Yelon, D. (2007). Vessel and blood specification override cardiac potential in anterior mesoderm. Dev Cell *13*, 254-267.

Schultz, J., Ponting, C.P., Hofmann, K., and Bork, P. (1997). SAM as a protein interaction domain involved in developmental regulation. Protein Sci *6*, 249-253.

Shalaby, F. (1997). A Requirement for Flk1 in Primitive and Definitive Hematopoiesis and Vasculogenesis. Cell *89*, 981-990.

Shalaby, F., Rossant, J., Yamaguchi, T.P., Gertsenstein, M., Wu, X.F., Breitman, M.L., and Schuh, A.C. (1995). Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. Nature *376*, 62-66.

Shimoda, Y., and Watanabe, K. (2009). Contactins. Cell Adhesion & Migration *3*, 64-70.

Sittaramane, V., and Chandrasekhar, A. (2008). Expression of unconventional myosin genes during neuronal development in zebrafish. Gene Expr Patterns *8*, 161-170.

Smyth, D.J., Plagnol, V., Walker, N.M., Cooper, J.D., Downes, K., Yang, J.H.M., Howson, J.M.M., Stevens, H., McManus, R., Wijmenga, C*., et al.* (2008). Shared and distinct genetic variants in type 1 diabetes and celiac disease. N Engl J Med *359*, 2767-2777.

Soleimani, Vahab D., Punch, Vincent G., Kawabe, Y.-i., Jones, Andrew E., Palidwor, Gareth A., Porter, Christopher J., Cross, Joe W., Carvajal, Jaime J., Kockx, Christel E.M., Van Ijcken, Wilfred F.J*., et al.* (2012). Transcriptional Dominance of Pax7 in Adult Myogenesis Is Due to High-Affinity Recognition of Homeodomain Motifs. Dev Cell *22*, 1208-1220.

Sommer, C.A., Pavarino-Bertelli, E.C., Goloni-Bertollo, E.M., and Henrique-Silva, F. (2008). Identification of dysregulated genes in lymphocytes from children with Down syndrome. Genome *51*, 19-29.

Spierings, E. (2008). Minor histocompatibility antigens: targets for tumour therapy and transplant tolerance. Int J Immunogenet *35*, 363-366.

Stainier, D.Y., Weinstein, B.M., Detrich, H.W., Zon, L.I., and Fishman, M.C. (1995). Cloche, an early acting zebrafish gene, is required by both the endothelial and hematopoietic lineages. Development *121*, 3141-3150.

Sumanas, S., Gomez, G., Zhao, Y., Park, C., Choi, K., and Lin, S. (2008). Interplay among Etsrp/ER71, Scl, and Alk8 signaling controls endothelial and myeloid cell formation. Blood *111*, 4500-4510.

Sumanas, S., Jorniak, T., and Lin, S. (2005). Identification of novel vascular endothelial-specific genes by the microarray analysis of the zebrafish cloche mutants. Blood *106*, 534-541.

Sumanas, S., and Lin, S. (2006). Ets1-Related Protein Is a Key Regulator of Vasculogenesis in Zebrafish. Plos Biol *4*, e10.

Suzuki, G., Harper, K.M., Hiramoto, T., Sawamura, T., Lee, M., Kang, G., Tanigaki, K., Buell, M., Geyer, M.A., Trimble, W.S.*, et al.* (2009). Sept5 deficiency exerts pleiotropic influence on affective behaviors and cognitive functions in mice. Hum Mol Genet *18*, 1652-1660.

Takahashi, K., Matsumoto, C., and Ra, C. (2005). FHL3 negatively regulates human high-affinity IgE receptor beta-chain gene expression by acting as a transcriptional co-repressor of MZF-1. Biochem J *386*, 191-200.

Takaya, A., Kamio, T., Masuda, M., Mochizuki, N., Sawa, H., Sato, M., Nagashima, K., Mizutani, A., Matsuno, A., Kiyokawa, E.*, et al.* (2007). R-Ras regulates exocytosis by Rgl2/Rlf-mediated activation of RalA on endosomes. Mol Biol Cell *18*, 1850-1860.

Thalappilly, S., Suliman, M., Gayet, O., Soubeyran, P., Hermant, A., Lecine, P., Iovanna, J.L., and Dusetti, N.J. (2008). Identification of multi-SH3 domain-containing protein interactome in pancreatic cancer: a yeast two-hybrid approach. Proteomics *8*, 3071-3081.

Thisse, B., Thisse, C (2004). Fast Release Clones: A High Throughput Expression Analysis. ZFIN Direct Data Submission (http://zfin.org).

To, M.D., Faseruk, S.A., Gokgoz, N., Pinnaduwage, D., Done, S.J., and Andrulis, I.L. (2005). LAF-4 is aberrantly expressed in human breast cancer. Int J Cancer *115*, 568-574.

Tosoni, D., Puri, C., Confalonieri, S., Salcini, A.E., De Camilli, P., Tacchetti, C., and Di Fiore, P.P. (2005). TTP specifically regulates the internalization of the transferrin receptor. Cell *123*, 875-888.

Veldman, M.B., and Lin, S. (2012). Etsrp/Etv2 is directly regulated by Foxc1a/b in the zebrafish angioblast. Circulation Research *110*, 220-229.

Veldman, M.B., Zhao, C., Gomez, G.A., Lindgren, A.G., Huang, H., Martin, B.L., D, K., and Lin, S. (2013). Transdifferentiation of Fast Skeletal Muscle into Functional Endothelium in Vivo by Transcription Factor Etv2. Plos Biol *in press*.

Vigil, D., Martin, T.D., Williams, F., Yeh, J.J., Campbell, S.L., and Der, C.J. (2010). Aberrant overexpression of the Rgl2 Ral small GTPase-specific guanine nucleotide exchange factor promotes pancreatic cancer growth through Ral-dependent and Ral-independent mechanisms. J Biol Chem *285*, 34729-34740.

Visvader, J.E., Fujiwara, Y., and Orkin, S.H. (1998). Unsuspected role for the T-cell leukemia protein SCL/tal-1 in vascular development. Genes Dev *12*, 473-479.

Vogeli, K.M., Jin, S.-W., Martin, G.R., and Stainier, D.Y.R. (2006). A common progenitor for haematopoietic and endothelial lineages in the zebrafish gastrula. Nature *443*, 337-339.

Wang, C., Faloon, P.W., Tan, Z., Lv, Y., Zhang, P., Ge, Y., Deng, H., and Xiong, J.-W. (2007). Mouse lysocardiolipin acyltransferase controls the development of hematopoietic and endothelial lineages during in vitro embryonic stem-cell differentiation. Blood *110*, 3601-3609.

Wang, L., Li, L., Shojaei, F., Levac, K., Cerdan, C., Menendez, P., Martin, T., Rouleau, A., and Bhatia, M. (2004). Endothelial and hematopoietic cell fate of human embryonic stem cells originates from primitive endothelium with hemangioblastic properties. Immunity *21*, 31-41.

Wardle, F.C., Odom, D.T., Bell, G.W., Yuan, B., Danford, T.W., Wiellette, E.L., Herbolsheimer, E., Sive, H.L., Young, R.A., and Smith, J.C. (2006). Zebrafish promoter microarrays identify actively transcribed embryonic genes. Genome Biol *7*, R71.

Wareing, S., Eliades, A., Lacaud, G., and Kouskoff, V. (2012a). ETV2 expression marks blood and endothelium precursors, including hemogenic endothelium, at the onset of blood development. Dev Dyn *241*, 1454-1464.

Wareing, S., Mazan, A., Pearson, S., Göttgens, B., Lacaud, G., and Kouskoff, V. (2012b). The Flk1-Cre-mediated deletion of ETV2 defines its narrow temporal requirement during embryonic hematopoietic development. Stem Cells *30*, 1521-1531.

Weber, C.H., and Vincenz, C. (2001). The death domain superfamily: a tale of two interfaces? Trends Biochem Sci *26*, 475-481.

Weber, G.J., Choe, S.E., Dooley, K.A., Paffett-Lugassy, N.N., Zhou, Y., and Zon, L.I. (2005). Mutant-specific gene programs in the zebrafish. Blood *106*, 521-530.

Wittwer, F., van der Straten, A., Keleman, K., Dickson, B.J., and Hafen, E. (2001). Lilliputian: an AF4/FMR2-related protein that controls cell identity and cell growth. Development *128*, 791-800.

Wong, K.S., Proulx, K., Rost, M.S., and Sumanas, S. (2009a). Identification of vasculature-specific genes by microarray analysis of etsrp/etv2 overexpressing zebrafish embryos. Dev Dyn *238*, 1836-1850.

Wong, K.S., Proulx, K., Rost, M.S., and Sumanas, S. (2009b). Identification of vasculature-specific genes by microarray analysis of Etsrp/Etv2 overexpressing zebrafish embryos. Dev Dyn *238*, 1836-1850.

Xiong, J.-W., Yu, Q., Zhang, J., and Mably, J.D. (2008). An Acyltransferase Controls the Generation of Hematopoietic and Endothelial Lineages in Zebrafish. Circulation Research *102*, 1057-1064.

Yamamizu, K., Matsunaga, T., Katayama, S., Kataoka, H., Takayama, N., Eto, K., Nishikawa, S.-I., and Yamashita, J.K. (2012). PKA/CREB signaling triggers initiation of endothelial and hematopoietic cell differentiation via Etv2 induction. Stem Cells *30*, 687-696.

Yang, Y., Cochran, D.A., Gargano, M.D., King, I., Samhat, N.K., Burger, B.P., Sabourin, K.R., Hou, Y., Awata, J., Parry, D.A.D.*, et al.* (2011). Regulation of flagellar motility by the conserved flagellar protein CG34110/Ccdc135/FAP50. Mol Biol Cell *22*, 976-987.

Yoder, J.A., Haire, R.N., and Litman, G.W. (1999). Cloning of two zebrafish cDNAs that share domains with the MHC class II-associated invariant chain. Immunogenetics *50*, 84-88.

Yoshio, T., Morita, T., Kimura, Y., Tsujii, M., Hayashi, N., and Sobue, K. (2007). Caldesmon suppresses cancer cell invasion by regulating podosome/invadopodium formation. FEBS Lett *581*, 3777-3782.

Zadro, C., Alemanno, M.S., Bellacchio, E., Ficarella, R., Donaudy, F., Melchionda, S., Zelante, L., Rabionet, R., Hilgert, N., Estivill, X.*, et al.* (2009). Are MYO1C and MYO1F associated with hearing loss? Biochim Biophys Acta *1792*, 27-32.

Zeng, L., Zhang, C., Xu, J., Ye, X., Wu, Q., Dai, J., Ji, C., Gu, S., Xie, Y., and Mao, Y. (2002). A novel splice variant of the cell adhesion molecule contactin 4 (CNTN4) is mainly expressed in human brain. J Hum Genet *47*, 497-499.

Zhang, J., Guo, J., Dzhagalov, I., and He, Y.-W. (2005). An essential function for the calcium-promoted Ras inactivator in Fcgamma receptor-mediated phagocytosis. Nat Immunol *6*, 911-919.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zheng, P.-P., Severijnen, L.-A., Willemsen, R., and Kros, J.M. (2009). Caldesmon is essential for cardiac morphogenesis and function: in vivo study using a zebrafish model. Biochem Biophys Res Commun *378*, 37-40.

Zheng, P.-P., van der Weiden, M., and Kros, J.M. (2005a). Differential expression of Hela-type caldesmon in tumour neovascularization: a new marker of angiogenic endothelial cells. J Pathol *205*, 408-414.

Zheng, Y., Zhou, Z.-M., Min, X., Li, J.-M., and Sha, J.-H. (2005b). Identification and characterization of the BGR-like gene with a potential role in human testicular development/spermatogenesis. Asian J Androl *7*, 21-32.

Zhu, L.J., Gazin, C., Lawson, N.D., Pagès, H., Lin, S.M., Lapointe, D.S., and Green, M.R. (2010a). ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics *11*, 237.

Zhu, L.J., Gazin, C., Lawson, N.D., Pagès, H., Lin, S.M., Lapointe, D.S., and Green, M.R. (2010b). ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. In BMC Bioinformatics, pp. 237.

Zieger, B., Hashimoto, Y., and Ware, J. (1997). Alternative expression of platelet glycoprotein Ib(beta) mRNA from an adjacent 5' gene with an imperfect polyadenylation signal sequence. J Clin Invest *99*, 520-525.

Znosko, W.A., Yu, S., Thomas, K., Molina, G.A., Li, C., Tsang, W., Dawid, I.B., Moon, A.M., and Tsang, M. (2010). Overlapping functions of Pea3 ETS transcription factors in FGF signaling during zebrafish development. Dev Biol *342*, 11-25.