Available at www.sciencedirect.com
SciVerse ScienceDirect
journal homepage: www.elsevier.com/locate/modo

Global analysis of the haematopoietic and endothelial transcriptome during zebrafish development

J.E. Cannon [a,b,c], E.S. Place [d], A.M.J. Eve [d], C.R. Bradshaw [b], A. Sesay [d], N.W. Morrell [a], J.C. Smith [b,c,d,*]

[a] Department of Medicine, University of Cambridge, Cambridge CB2 0QQ, UK
[b] Wellcome Trust/Cancer Research UK Gurdon Institute, The Henry Wellcome Building of Cancer and Developmental Biology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, UK
[c] Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK
[d] Division of Systems Biology, MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

ARTICLE INFO

Article history:
Received 18 October 2011
Received in revised form
3 October 2012
Accepted 4 October 2012
Available online 13 October 2012

Keywords:
Zebrafish
Haematopoiesis
Endothelial cell
High-throughput sequencing

ABSTRACT

In this paper, we use zebrafish embryos to characterise the transcriptome of the developing blood and endothelium, two cell types that are closely associated during development. High-throughput sequencing identified 754 genes whose transcripts are enriched threefold or more in blood and/or vascular endothelial cells compared with the rest of the embryo at 26–28 h post fertilisation. Of these genes, 388 were classified as novel to these cell types after cross-reference with PubMed and the zebrafish information network (ZFIN). Analysis by quantitative PCR and *in situ* hybridisation showed that 83% ($n = 41$) of these novel genes are expressed in blood or vascular endothelium. Of 10 novel genes selected for knockdown by antisense morpholino oligonucleotides, we confirmed that two, *tmem88a* and *trim2a*, are required for primitive erythropoiesis and myelopoiesis. Our results provide a catalogue of genes whose expression is enriched in the developing blood and endothelium in zebrafish, many of which will be required for the development of those cell types, both in fish and in mammals.

1. Introduction

Zebrafish are widely used in studies investigating haematopoietic and vascular development. They have several advantages over other vertebrate model systems, including access to hundreds of externally fertilised, transparent embryos that allow the visualisation of developmental processes *in vivo*. There are also a number of haematopoietic and vascular mutants previously found in large scale ENU mutagenesis screens (reviewed by Baldessari and Mione, 2008), and transgenic lines are available including the Tg(*fli1a:egfp*)^{y1} line used in this study (Baldessari and Mione, 2008; Lawson and Weinstein, 2002). For genes where mutants are not available, antisense morpholino oligonucleotides (morpholinos) can be used to knock down genes of interest. Finally, and importantly, there is a high degree of conservation of genes known to be important for vascular and haematopoietic development between zebrafish and higher organisms (Jing and Zon, 2011).

During early vertebrate embryo development blood and endothelial cells are found closely associated. In mammals they are initially found in the blood islands of the extra-embryonic yolk sac (Park et al., 2005), while during segmentation in zebrafish they are found intra-embryonically in the intermediate cell mass (ICM) of the ventral mesoderm (Detrich et al., 1995). In view of this close relationship it has been suggested that blood and endothelial cells have a common

* Corresponding author at: Wellcome Trust/Cancer Research UK Gurdon Institute, The Henry Wellcome Building of Cancer and Developmental Biology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, UK. Tel.: +44 20 8816 2048; fax: +44 20 8816 2041.
E-mail address: director@nimr.mrc.ac.uk (J.C. Smith).
0925-4773/$ - see front matter
http://dx.doi.org/10.1016/j.mod.2012.10.002

precursor cell, the haemangioblast (Sabin, 1920). Although there has been evidence to support this hypothesis from *in vitro* studies, it has only recently been shown that the haemangioblast exists *in vivo* (Park et al., 2005; Vogeli et al., 2006).

The factors controlling haemangioblast formation and the development of angioblasts (vascular endothelial cell precursors) and haematopoietic stem cells are incompletely understood. Several transcription factors are important for the formation of the haemangioblast. Stem cell leukaemia (*scl*, also known as *tal1*) null mice die *in utero* due to the complete absence of blood (Shivdasani et al., 1995). In zebrafish, morpholino knockdown of *scl* phenocopies the null mouse, but these embryos also have impaired vascular gene expression in the dorsal aorta and loss of intersegmental vessel (ISV) formation (Patterson et al., 2005). The Ets-1 related protein (*etsrp*, also known as *etv2*) was identified in a screen for novel genes affected in the *cloche* mutant, that lacks both blood and endothelial cells (Sumanas et al., 2005). Morpholino knockdown of *etsrp* leads to impaired vasculogenesis and myelopoiesis (Sumanas et al., 2008; Sumanas and Lin, 2006). *Fli1*, like *etsrp*, is an ETS transcription factor that is also important for haemangioblast formation. It has been suggested to act at the top of a transcriptional network driving blood and endothelial development by regulating other genes required for haemangioblast formation including *scl* and *etsrp* (Liu et al., 2008). The VEGF signalling pathway is also critical for vascular development. Loss of *Vegf* or its receptor *Flk1* in mice leads to death *in utero* due to failure to form the vasculature (Carmeliet et al., 1996; Shalaby et al., 1995). For erythrocyte development *Gata1* is a master regulator. *Gata1*−/− mice die *in utero* due to the failure of pro-erythrocytes to differentiate into mature erythrocytes (Fujiwara et al., 1996).

The identification of genes involved in blood and endothelial development is of significant therapeutic interest. We therefore sought to determine the transcriptome of developing haematopoietic and vascular endothelial cells *in vivo*. Previous studies have attempted to answer this question using microarrays (Covassin et al., 2006; Kalén et al., 2009; Sumanas et al., 2005; Wallgard et al., 2008; Weber et al., 2005; Wong et al., 2009). Here, because *fli1* is one of the earliest factors involved in haemangioblast formation, we have used a fluorescence-activated cell sorting (FACS) technique (Covassin et al., 2006) to isolate *gfp* positive (*gfp*+) and negative (*gfp*−) cells from transgenic Tg(*fli1a:egfp*)^{y1} embryos prior to high-throughput sequencing. This transgenic line utilises the *fli1a* promoter to drive *gfp* expression in blood and vascular endothelial cells, pharyngeal arch and neural crest derivatives (Lawson and Weinstein, 2002). Using this technique we have identified 388 novel genes expressed in the enriched population of blood and endothelial cells. Using morpholino knockdown we confirm that two of the genes identified, *tmem88a* and *trim2a*, are novel genes required for erythropoiesis and myelopoiesis in zebrafish.

2. Results

2.1. Isolation of vascular and haematopoietic cells from whole embryos

To identify genes involved in the development of endothelial and blood cells, we isolated *gfp*+ cells from dissociated 26–28 hpf Tg(*fli1a:egfp*)^{y1} transgenic zebrafish, where the *fli1a* promoter drives *gfp* expression in endothelial and haematopoietic cells and pharyngeal arch tissue (Lawson and Weinstein, 2002). This time-point was chosen because the intersegmental vessels are forming by angiogenesis and the haematopoietic stem cells are starting to arise from the ventral floor of the aorta (Bertrand et al., 2010; Isogai et al., 2003).

Approximately 6.6% of cells in 26–28 hpf Tg(*fli1a:egfp*)^{y1} embryos were *gfp*+ by FACS (Supplementary Fig. 1A). A small proportion of cells from each sorted group were re-sorted to determine the purity of the cell populations. The *gfp*+ population was always greater than 95% pure (Supplementary Fig. 1B) and the *gfp*− population greater than 99% (Supplementary Fig. 1C). As further validation for the purity of each population, qRT-PCR was performed on cDNA made from RNA isolated from the sorted cells. Genes expected to be enriched in the *gfp*+ cells were indeed highly enriched. These included *fli1a* (77.2 ± 20.0-fold), *kdrl* (46.9 ± 23.8-fold), *flt4* (24.9 ± 6.2-fold) and *gata1a* (27.9 ± 14.0) (Supplementary Fig. 1D). Conversely, several genes not known to be highly expressed in vascular or haematopoietic cells (ZFIN, www.zfin.org) were more highly expressed in *gfp*- cells. These included *myoD* (5.1 ± 0.8-fold), *ptprn* (8.5 ± 2.5-fold), and *dlb* (7.4 ± 1.2-fold) (Supplementary Fig. 1E). Together these results indicate that we can isolate highly purified populations of *gfp*+ and *gfp*− cells from Tg(*fli1a:egfp*)^{y1} zebrafish embryos.

2.2. Global analysis of genes enriched in gfp positive cells by massively parallel sequencing

The transcriptome of developing zebrafish blood and vascular endothelial cells was defined by undertaking high-throughput sequencing of cDNA made from sorted *gfp*+ and *gfp*− cell populations derived from about 3000 Tg(*fli1a:egfp*)^{y1} embryos. We found that 754 protein-coding genes were enriched threefold or greater in the *gfp*+ compared to the *gfp*− population of cells in both biological replicates (Fig. 1 and Supplementary Table 1). This group includes genes expected to be enriched such as *scl*, *etsrp*, *fli1a*, *gata1a*, haemoglobins and vegf receptors (Supplementary Fig. 2 and Table 1). Some genes known to be important for vascular development and/or haematopoiesis (like *ephrinb2*, *ephB4*, *jag2*, *notch1*, *notch3* and *unc5b*) were not enriched in the *gfp*+ libraries (Adams et al., 1999; Hadland et al., 2004; Krebs et al., 2000; Lawson et al., 2002; Lu et al., 2004; Van de Walle et al., 2011; Wang et al., 1998). This is because these genes are also strongly expressed in other tissues including neural tissues (ZFIN).

To identify the biological functions of the 754 genes we used the Panther classification system (Thomas et al., 2003) Genes involved in blood circulation and gas exchange (2.77-fold, $p = 5.14E^{-04}$), immunity and defence (1.45-fold, $p = 5.27E{-}05$), transport (1.42-fold, $p = 1.91E{-}04$) and intracellular protein traffic (1.4-fold, $p = 1.82E{-}03$) were found to be the most significantly enriched compared to the whole zebrafish genome whereas neuronal activities (−2.28-fold, $p = 8.57E{-}06$), nucleoside, nucleotide and nucleic acid metabolism (−1.31-fold, $p = 2.42E{-}04$) and sensory perception (−1.59-fold, $p = 0.024$) were significantly under-represented (Table 1). One third of the genes, however, had an unclassified biological function (Supplementary Fig. 3). Use of ZFIN and PubMed revealed that