A zebrafish reporter line reveals immune and neuronal expression of endogenous retrovirus.

Holly A. Rutherford[1Ψ], Amy Clarke[1Ψ], Emily V. Chambers[3], Jessica J. Petts[1], Euan G. Carson[1], Hannah M. Isles[1], Alejandra Duque-Jaramillo[4], Stephen A Renshaw[1], Jean-Pierre Levraud[5], Noémie Hamilton[1,2]*.

1. The Bateson Centre, Department of Infection, Immunity and Cardiovascular Disease, University of Sheffield, Sheffield, UK
2. The Institute of Neuroscience, University of Sheffield, Sheffield, UK
3. The Bioinformatics Core, Faculty of Medicine and Dentistry, University of Sheffield, Sheffield S10 2TN.
4. Institute of Microbiology (IMUL), Lausanne University Hospital and University of Lausanne.
5. Macrophages et Développement de l'Immunité, Institut Pasteur, CNRS UMR3738, 25 rue du docteur Roux, 75015 Paris

*Corresponding author: n.m.hamilton@sheffield.ac.uk
[Ψ] Authors contributed equally

Abstract

Endogenous retroviruses (ERVs) are fossils left in our genome from retrovirus infections of the past. Their sequences are part of every vertebrate genome and their random integrations are thought to have contributed to evolution. Although ERVs are mainly kept silenced by the host genome, they are found activated in multiple disease states such as auto-inflammatory disorders and neurological diseases. What makes defining their role in health and diseases challenging is the numerous copies in mammalian genomes and the lack of tools to study them. In this study, we identified 8 copies of the zebrafish endogenous retrovirus (*zferv*). We created and characterised the first *in vivo* ERV reporter line in any species. Using a combination of live imaging, flow cytometry and single cell RNA sequencing, we mapped *zferv* expression to early T cells and neurons. Thus, this new tool identified tissues expressing ERV in zebrafish, highlighting a potential role of ERV during brain development and strengthening the hypothesis that ERV play a role in immunity and neurological diseases. This transgenic line is therefore a suitable tool to study the function of ERV in health and diseases.

Keywords
Retroelement, zebrafish, endogenous retrovirus, reporter line, *zferv*, LTR

Funding
This work has been supported by a European Leukodystrophy Association fellowship (ELA 2016-012F4) to NH, an MRC Programme Grant (MR/M004864/1) to SAR and JPL is supported by Agence Nationale de la Recherche (grant ANR-16-CE20-0002-03). Imaging was carried out in the Wolfson Light Microscopy Facility, supported by an MRC grant (G0700091) and a Wellcome Trust grant (GR077544AIA).

Conflict of Interest Statement
The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Introduction

Over 40% of the human genome comprises endogenous transposable elements capable of recombination and disruption of genes and modification of their expression (Lander et al. 2001). Endogenous retroviruses (ERVs) are transposable elements originating from old integrations of retroviruses so successful that they have become part of most vertebrate genomes studied. ERVs replicate autonomously using a copy-and-paste mechanism and although they form a smaller percentage of all retroelements, they still represent 8% of the human genome (Bourque et al. 2018; Lander et al. 2001). The majority of known ERVs have lost their ability to replicate, but those most recently acquired still have intact genomes with the ability to produce viral RNA and particles. However, many of competent ERVs are under strict transcriptional suppression by epigenetic mechanisms (Maksakova, Mager, and Reiss 2008; Rowe et al. 2010; Turelli et al. 2014), protecting the host organisms against potential retroviral insertions and viral activities.

Although immobilised by mutations or transcriptionally repressed, ERVs have a complex relationship with the human genome, which they can regulate by providing cis-regulatory elements to surrounding genes and by lifting their transcriptionally repressed state. Through these mechanisms, it is believed that transposable elements have fuelled some of the necessary genetic changes for evolution (Feschotte 2008; Kunarso et al. 2010). *Syncytin-1*, an ERV envelope gene essential for the vascularisation of the placenta, is at the origin of evolutionary diversification of the placenta (Chuong 2018; Mi et al. 2000). During human embryogenesis, expression of specific ERV families have been associated with cell identity and cell potency in early stem cells (Göke et al. 2015). Additionally, ERV expression has been reported in healthy human tissues such as ovary and testis for ERV-9 (Pi et al. 2004), pancreas (Shiroma et al. 2001), breast (Tavakolian, Goudarzi, and Faghihloo 2019), stomach and small intestine (Okahara et al. 2004). Mainly based on transcriptional studies, the expression of different families of ERV is likely to be extended to more tissues, however their role in tissue development and function remains largely unknown.

ERVs have been linked directly and indirectly to the evolution and functioning of the immune system. Enhancer regions of interferon stimulated genes key to the interferon pathway, such as IRF1 and STAT1, were found introduced and amplified by ERV elements, with the human inflammasome failing to form upon the deletion of a subset of ERVs (Chuong, Elde, and Feschotte 2016). Adaptive immunity also benefits from the presence of ERVs. Indeed, ERV peptide recognition is used in T cell selection to optimise antigen recognition and T cells have a higher sensitivity to exogenous virus infection when presented with ERV peptides during their initial thymic selection (Mandl et al. 2013; Young et al. 2012). The human ERV (HERV) envelope gene contains immunosuppressive domains that reduce the Th1 response during pregnancy and therefore promote foetal development (Knerr et al. 2004; Lokossou et al. 2020). The role of ERVs in our immune system, particularly in the training of our adaptive immunity, can be a double-edge sword as ERVs are linked to a range of different disease states, including autoimmunity.

Aberrant expression of ERVs contributes to multiple pathologies. ERVs are found in abundance in multiple forms of cancers and are considered tumour-promoting factors (extensively reviewed in (Bermejo et al. 2020). The pathology of auto-inflammatory diseases has also been strongly associated with ERVs. Systemic lupus erythematosus (SLE) is an autoimmune disorder with increased level of autoantigen for an ERV

(Jorgensen et al. 2014). Recently, one of the murine SLE susceptibility locus was identified as a key suppressor of ERV expression, consolidating the role of ERVs in the pathogenesis of SLE (Treger et al. 2019). A similar disorder is Acardi-Goutieres Syndrome (AGS), a type 1 interferonopathy which resembles congenital cytomegalovirus infection and is caused by mutation in several genes encoding enzymes responsible for nucleic acid metabolism (Crow et al. 2015). Mutations in some of these genes, such as TREX1, MDA5 and ADAR1 trigger aberrant presence of various retroelements and the upregulation of an antiviral immune response (Ahmad et al. 2018; Benitez - Guijarro et al. 2018; Chung et al. 2018; Li et al. 2017; Thomas et al. 2017). Interestingly, anti-reverse transcriptase therapy in AGS patients can decrease the IFN response, highlighting the role of aberrant presence of ERVs in triggering an immune response (G. Rice et al. 2018). Increased expression of ERVs has been found in brains of patient suffering from neurodegenerative diseases such as Motor Neuron Disease (Li et al. 2015) and multiple sclerosis (Johnston et al. 2001; Mameli et al. 2007). Overexpression of a human ERV in neurons was shown to be neurotoxic, suggesting a potential role of ERVs in triggering neuronal toxicity (Li et al. 2015). A direct link to the pathology of these disorders has yet to be made, but nonetheless ERVs appear as strong causal factors for autoimmune and neurological disorders.

Although ERV enrichment has been detected in neurological pathologies, little is known about the function of ERV in healthy tissues. The exact role of ERV in our immune system and brain pathologies has yet to be understood and there is no model system specifically looking at ERV function *in vivo*. Zebrafish is already established as a model to study the immune system, with significant homology with mammals in innate and adaptive immunity (Renshaw and Trede 2012; Trede et al. 2004). The genetic tractability and transparency of the zebrafish embryos have allowed the creation of transgenic reporter lines, some of which have elucidated the role of immune cell behaviour *in vivo* (Renshaw et al. 2006). The tractability of the zebrafish has already been exploited to visualise the expression of the human ERV-9 in oocytes, similarly to human expression (Pi et al. 2004).

In this study, we used the zebrafish as a tractable *in vivo* model to characterise the zebrafish endogenous retrovirus (*zferv*). We identified multiple *zferv* family members, including 2 complete genomes *zferv1a* and *zferv1b*. Using the transparency of the zebrafish larvae we developed a reporter line for *zferv1a* and imaged for the first time ERV activation in healthy tissues *in vivo*. We showed that *zferv1a* is expressed in the thymus and in the brain. Colocalisation analysis and single cell RNA sequencing revealed expression of *zferv1a* specifically in T-cells, suggesting a potential role for ERV in lymphocyte function or development. Brain expression of ZFERV appears to include neuronal expression. This transgenic line can be used as a tool to further investigate the role of ERVs in immunity and in neurological disorders.

Results

The zebrafish genome contains multiple endogenous retrovirus integrations.

The presence of an ERV in zebrafish, named *zferv*, has been reported by Shen and Steiner while screening a thymic cDNA library (Shen and Steiner 2004) (**Figure 1A**). We searched for related sequences in the most recent reference zebrafish genome (GRCz11, Tü strain) using BLASTN searches, with the original *zferv* sequence as a query. Limiting ourselves to sequences flanked by long terminal repeats (LTRs) on both sides, we retrieved 8 hits scattered in the zebrafish genome, but not the exact *zferv* sequence, possibly because of strain difference (**Figure 1+ Sup. Table 1**). We identified 2 sequences encoding apparently fully functional ERVs, with 91% and 90% identity to *zferv*, that we respectively named *zferv1a* and *zferv1b* (**Figure 1B**). These two ERVs encode almost identical proteins (95 to 97% identity at the amino-acid level) and have LTR promoter region highly homologous (95% identity at the nucleotide level). An additional 6 pseudo *zferv* genes (here called *zferv2* - *zferv7*) were identified (**Figure 1C**), *zferv2* containing a frameshift, *zferv3* and *zferv4* with large deletions and *zferv5*, *zferv6* and *zferv7* containing a large insertion, all resulting in a degenerated ERV genome.

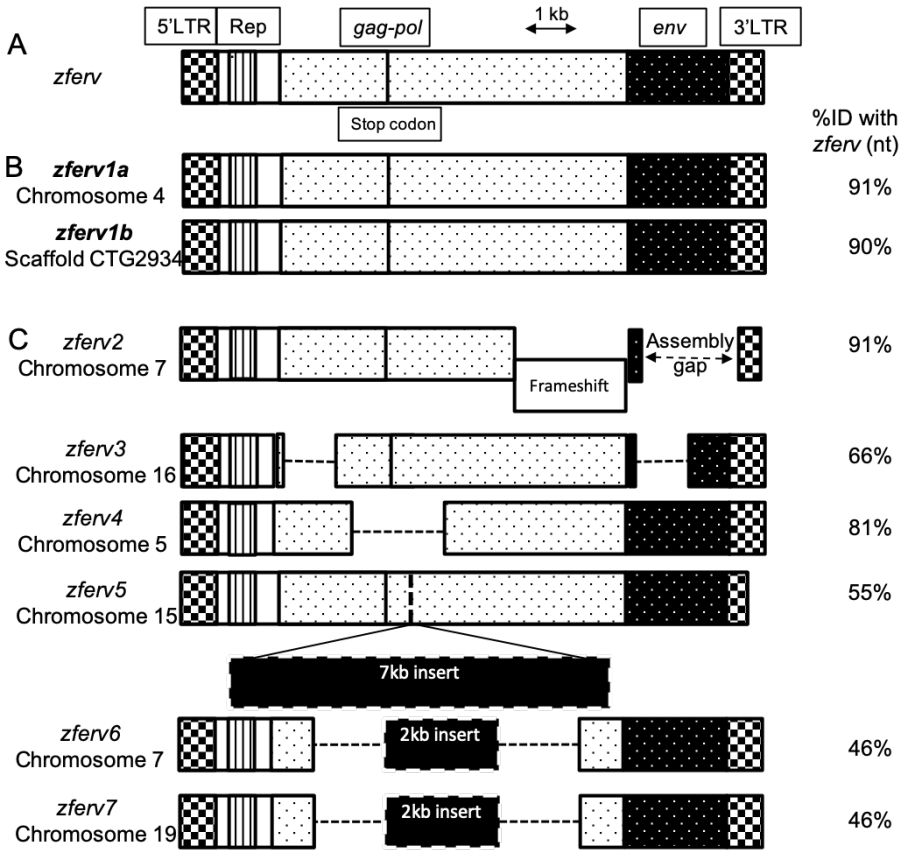


Figure 1. **Multiple copies of zferv are present in the zebrafish genome**.
A. Diagram of the original zferv genome found by Steiner et al. used as a reference for nucleotide identity (ID% nt)
B. Diagram representing the two closest related *zferv* genome found in most recent zebrafish genome GRCz11, named zferv1a and *zferv1b*.
C. Diagram of 6 pseudo zferv genes with degenerated genome (dotted line represents insertions). 5'-LTR: 5' –Long terminal repeat, Rep.: Repetitive element, gag-pol: genes encoding the polyprotein and reverse transcriptase, env: envelope gene.

***zferv1a* is actively expressed in the brain and in the thymus.**

zferv expression was initially found in the thymus at 5 days post fertilisation (dpf) using an RNA probe against the envelope (*env*) gene of the original *zferv* (Shen and Steiner 2004). To verify this observation, the entire *zferv1a* genome was subcloned from a fosmid provided by the Sanger Center and used to create an *in situ* hybridisation (ISH) RNA probe against the envelope gene of *zferv1a,* here called *env1a*. ISH performed on a time course starting from 2 cell-stage embryos until 5days post fertilisation (dpf) confirmed a strong signal in the thymus at 5dpf (**Figure 2A**), similarly to what was previously reported (Shen and Steiner 2004). Individually labelled cells were visible around the thymus following the branchial arches and around the ear (**Figure 2A, black box**). Additionally, we identified a signal at earlier timepoint during gastrulation and a clear signal in the brain, particularly strong at 3dpf (**Figure 2A**). Neuromasts were labelled after 24h exposure to the probe (**Sup. Figure 1**). There was some signal in the thymus using the sense probe signal, suggesting some bidirectional expression of *zferv1a*, as reported for some human ERV (Chiappinelli et al. 2015).

To ensure the specificity of this probe to the *zferv1a* gene, we created a probe against the reverse transcriptase gene of *zferv1a* (*pol1a*) and showed similar labelling of the brain and thymus (**Sup. Figure 2)**. We additionally created a probe against another *zferv* isoform, using specific primers (**Sup. Figure 3**) to the envelope gene of *zferv1b (*called here *env1b)*. Although we still identified a signal in the thymus at 5dpf, we observed a different expression pattern from *env1a* with no brain expression and a strong labelling of the liver in some embryos at 3dpf (**Figure 2B**).

***zferv1a* reporter line recapitulates endogenous expression.**

To follow the expression of the zebrafish endogenous retrovirus we identified as *zferv1a*, we took a transgenic approach taking advantage of the promoter activity of retroviral LTR. We used the 5' viral promoter *ltr5* from *zferv1a* to drive GFP expression by Gateway recombination (**Figure 3A**). Injected embryos showing expression in their thymus were raised and screened in adulthood for germline transmission; one founder was selected to generate the transgenic line bearing the sh533 allele, which was analysed at the F2 generation. Comparable to what was observed by ISH, transgenic embryos (obtained from transgenic mothers) displayed GFP signal at the blastula stage developed a strong GFP signal in the brain by 3dpf (**Figure 3C,D**). GFP signal observed in the eye was caused by the reflective nature of the retina and not by expression of the transgene, as shown using illumination from different laser also highlighting the eye and in situ hybridisation labelling with the env1a probe in a 5dpf embryo with pigment partially removed in the eye using a tyrosinase crispant (**Sup. Figure 4**). A weak signal in the thymus can be observed at 3dpf (**Figure 3C** white arrowhead) as the thymus starts to develop. At 5dpf, a strong signal was observed in the thymus (**Figure 3E,F**), comparable to that observed by ISH (**Figure 1A,B**). Consistent with ISH, our reporter line also labelled the optic tectum region of the brain and the spinal cord at 5dpf, with axons labelled in individual neurons (**Figure 3G,H**). GFP signal was also found in neuromasts (**Figure 3E, white arrows**), which we also observed in a small proportion of 5dpf embryos by ISH with longer development time (**Sup. Figure 1**). Transmission rate of 50% to F3 embryos of the brain signal was variable depending which F2 animal was used, including complete and partial brain labelling, whereas the thymic expression remaining constant throughout the generations.

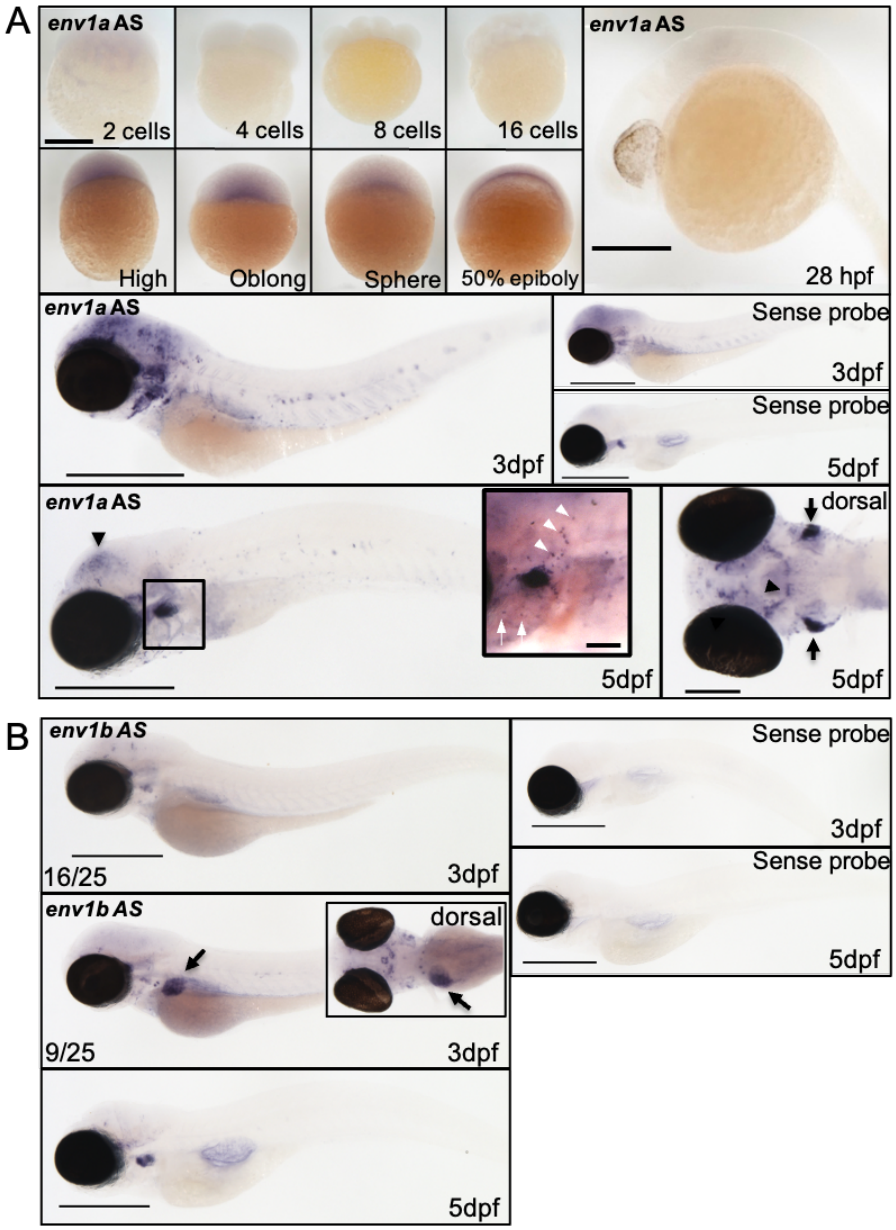


Figure 2: Reporter line for *zferv1a* recapitulates endogenous expression.
A. Expression of the envelope gene (*env*) of *zferv1a* (called *env1a* here) using an antisense (AS) and sense *in situ* hybridisation RNA probe from 2-cell stage until 5dpf (black arrowheads pointing to brain expression). Scale bar 500µm. Dorsal view of *env1aAS* expression in the brain (black arrowheads) and thymus (Black arrows). Scale bar 200µm. Black box is a zoomed image on the thymus area, highlighting strong expression around the thymus with single positive cells visible in the vicinity of the thymus around the ear (white arrowheads) and alongside the branchial arches (white arrows). Scale bar 70µm.
B. Expression of the envelope gene (*env*) of *zferv1b* (called *env1b AS* here) by *in situ* hybridization at 3dpf and 5dpf. Liver expression highlighted by black arrows. Scale bar 500µm.

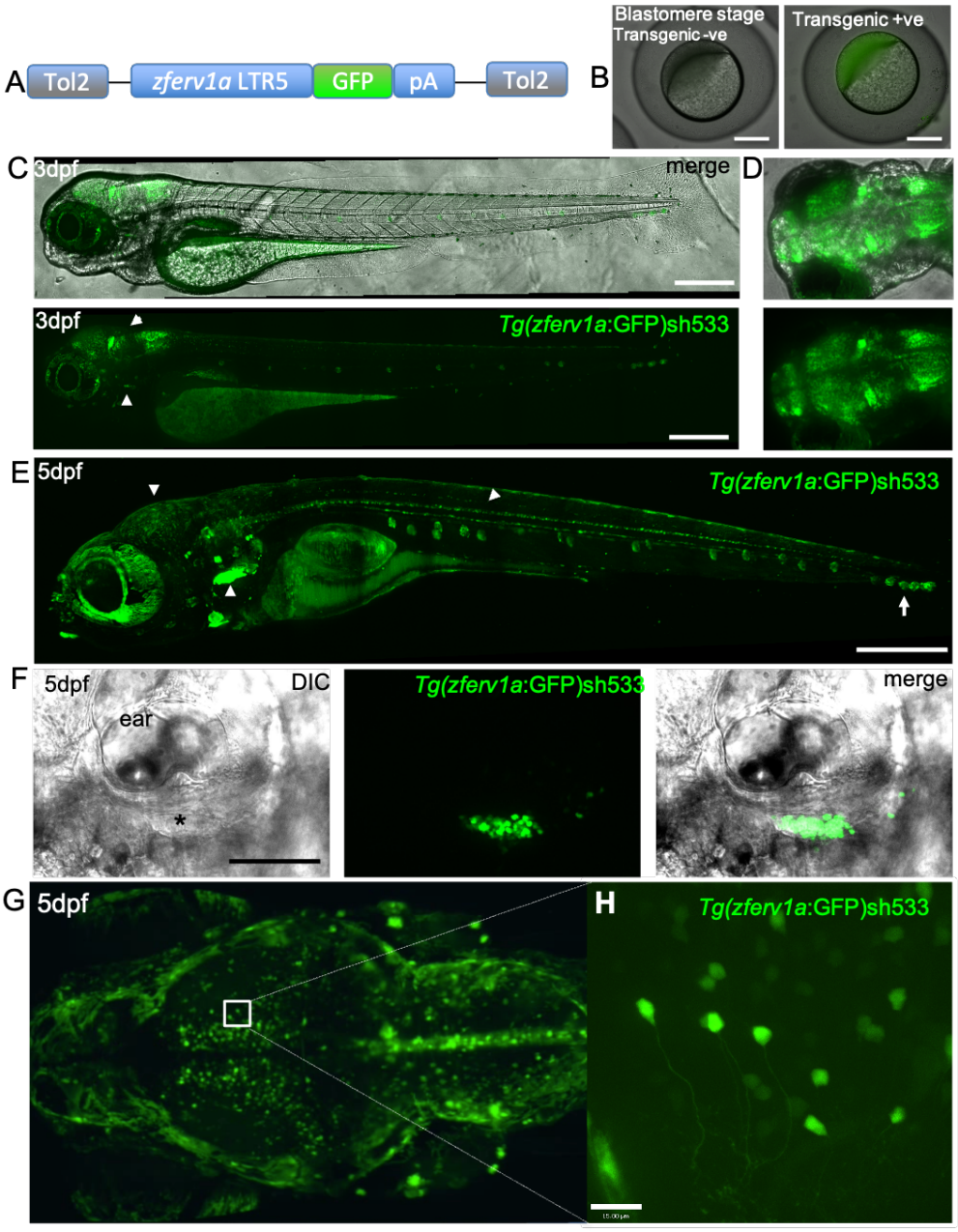


Figure 3: Reporter line for *zferv1a* recapitulates endogenous expression.

A. Diagram of gateway construct used to create the *Tg(zferv1a:GFP)* reporter line using the pDestCryCFP tol2 backbone.

B. Representative image at blastula stage in non-transgenic and *Tg(zferv1a:GFP:pA)sh533* transgenic embryos.

C. Lateral view of DIC and GFP high resolution images of a 3dpf *Tg(zferv1a:GFP:pA)sh533* embryos, displaying high expression of the transgene in the brain (white arrow) and the start of a signal in the thymus (white arrowheads). Scale bar 400μm.

D. Dorsal view of the brain at 3dpf, merged with with brightfield image and GFP only.

E. High resolution image of a 5dpf *Tg(zferv1a:GFP:pA)sh533* embryos, displaying high expression of the transgene in the thymus (white arrowheads) , brain and spinal cord (white arrowheads). Note signal in neuromast alongside the trunk (white arrow). Scale bar 500μm.

F. Single slice DIC and confocal GFP fluorescence merged image of the thymus (black star) situated underneath the ear in 5dpf embryos and the signal from the *Tg(zferv1a:GFP)sh533* reporter. Scale bar 100μm.

G. Dorsal view of a maximum projection of the brain from 5 dpf *Tg(zferv1a:GFP)* embryo acquired with a the lightsheet microscope.

H. High resolution single slice image from the optic tectum of *Tg(zferv1a:GFP)*. Scale bar 15 μm.

This transgenic line was not the first attempt to create a fluorescent reporter for *zferv1a*. Our initial reporter line was made using both 5' (*ltr5*) and 3' (*ltr3*) promoter regions in order to resemble the genome structure of an ERV. This line, called Tg(*ltr5:*GFP*:ltr3*)*sh500*, contained the CryCFP eye marker to help with selection of positive embryos. Although outcrossed F1 showed normal transmission rate (50% of positive embryos in F2, with CFP in lens and GFP in thymus), F2 transmission rate became unstable. Selected F2 positive embryos grown to adulthood failed to transmit the same signal pattern, with often reduced to complete loss of thymus labelling and random brain expression pattern despite transmission of the CFP lens phenotype at expected mendelian ratios. We also observed a dampening of the GFP signal in F3 and complete loss in F4. This suggested that the transgene was behaving as a transposable element and was ultimately silenced by the host or lost in the germline. This line was abandoned, and we continued our study with the *sh533* line.

Zebrafish endogenous retrovirus *zferv1a* expression is restricted to lymphoid cells. To further study the nature of the cells labelled in the thymus in the *sh533* line, we took advantage of the transparency and ease of access of the thymic tissue to perform live imaging. Cell tracking of timelapse imaging of the *sh533* line revealed a dynamic behaviour of GFP labelled cells, approaching and exiting the thymus following the ear and the branchial arches (**Figure 4A**, **Movie 1**), as previously described for thymocytes (Dee et al. 2016; Kissa et al. 2008). This thymic signal persisted beyond embryonic stages and was still visible in juveniles at 25dpf (**Figure 4B**). To analyse the origin of the *zferv1a* cell population, we dissociated 3 head kidneys from Tg(*zferv1a*:GFP)*sh533* and non-transgenic control adults and used size and granularity scatter by flow cytometry to distinguish between the various haematopoietic lineages (**Figure 4C**). According to these criteria and using a back-gating method for quantification (**Figure 5D,E**, **Sup. Figure 5**), we found that GFP positive cells have a typical lymphocyte profile (**Figure 4F**).

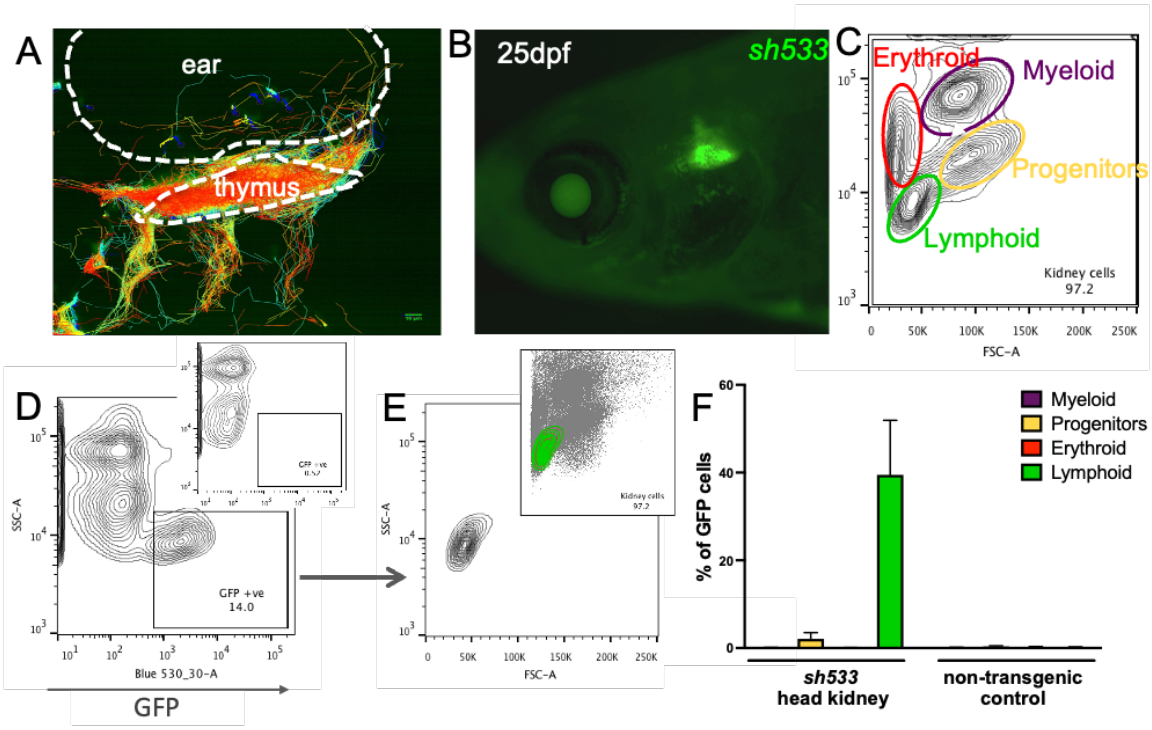


Figure 4: *zferv1a* is expressed in cells from the lymphoid lineage in adult haematopoietic tissues.

A. Cell tracking reveals dynamic behaviour of entry and exit of thymus by GFP positive cells in Tg(zferv1a:GFP)*sh533*. Whole stack analysis of cell movement by using TrackMate on FiJi. Scale bar 10µm.
B. Expression of *Tg(zferv1a:GFP)sh533* in a 25dpf zebrafish highlighting the triangular shaped thymus (white arrowhead).
C. FSC/SSC plot of a whole kidney marrow separating different haematopoietic lineages by size and granularity.
D. Side scatter (SSC) and GFP plot from single cell suspension of a whole kidney marrow from an adult *Tg(zferv1a:GFP)* with gate selecting GFP positive cells. Inset plot showing the distribution of cells from a non-transgenic adult whole kidney marrow.
E. Forward and side scatter (FSC/SSC) plot of selected GFP positive cells from panel D. Inset plot showing the location of the GFP positive cells in the haematopoietic lineage FSC/SSC plot.
F. Quantification of GFP positive cells in myeloid, erythroid, lymphoid lineages and progenitors, as defined in the dot plot D, from dissected sh533 and non-transgenic control adult head kidney (n=3).

To further corroborate the lymphoid identity of *zferv1a* expressing cells, we performed *in vivo* colocalisation analysis in 5dpf embryos by crossing the Tg(*zferv1*a:GFP)*sh533* line to different reporter lines labelling a wide spectrum of immune cell populations. We crossed F1 Tg(*zferv1a*:GFP)*sh533* to reporter lines labelling lymphoid cells, such as Tg(*lck*:mCherry), labelling T cells (Langenau et al. 2004) but also NK and innate lymphoid cells (Hernández et al. 2018), or Tg(*CD4*:mCherry), labelling CD4+ T cells and some macrophage/dendritic cells (Dee et al. 2016) and showed significant colocalisation between GFP and mCherry cells by colocalisation analysis (**Figure 5B,D,E)**. By contrast, crossing of the Tg(*zferv1a*:GFP)*sh533* to macrophage Tg(*mpeg1*:mCherryCAAX)*sh378* (**Figure 5A**) or neutrophil *Tg(lyz:nfsb*-mCherry)*sh260* (**Figure 5C**) (Buchan et al. 2019) reporter lines did not reveal significant colocalisation **(Figure 5E)**. To validate these results, we used single cell RNA sequencing datasets available on zebrafish blood cells (MTAB-5530 and E-MTAB-4617) as a robust assay to analyse gene expression in specific cell type and tissues. Using the extensive gene expression dataset available, we mapped the expression of the entire genome of *zferv1a* (*ENSDARG00000110878)* in erythrocytes, haematopoietic stem and progenitor cells (HSPCs), macrophages, neutrophils, CD4, T cells and thrombocytes. Only sporadic expression was observed in all cell lineages except for T cell subsets, CD4+and T cells, where significant expression of *env* was detected (**Figure 5F**). Altogether, these data are consistent with immune expression of *zferv1a* being restricted to T cells.

Modulation of *zferv1a* expression

To investigate whether *zferv1a* can affect the development of the tissues where it is expressed, we performed an overexpression assay. We injected the vector containing *zferv1a* alongside the same vector containing the eGFP cassette to use as a positive control. Each plasmid was injected in equimolar amount or GFP alone and the expression of zferv1a and GFP was measured by qPCR. Overexpression of GFP resulted in a high level of GFP transcript detected by qPCR, however we failed to observe an increase in zferv1a expression when injected (**Sup. Figure 6**). We then checked for induction of *zferv1a* (*ENSDARG00000110878*) in our datasets of larvae stimulated with type I interferon or with chikungunya (CHIKV) Virus ((Levraud et al. 2019); bioproject accession# PRJNA531581). No expression was detected in the dataset of type I IFN injected larvae (3dpf). Expression was detected in the dataset of CHIKV-infected fish (4dpf). However, while CHIKV infection triggers a very strong interferon response, with 1000-fold induction of some ISGs, there was only a statistically non-significant 1.7-fold increase for *zferv1a*.

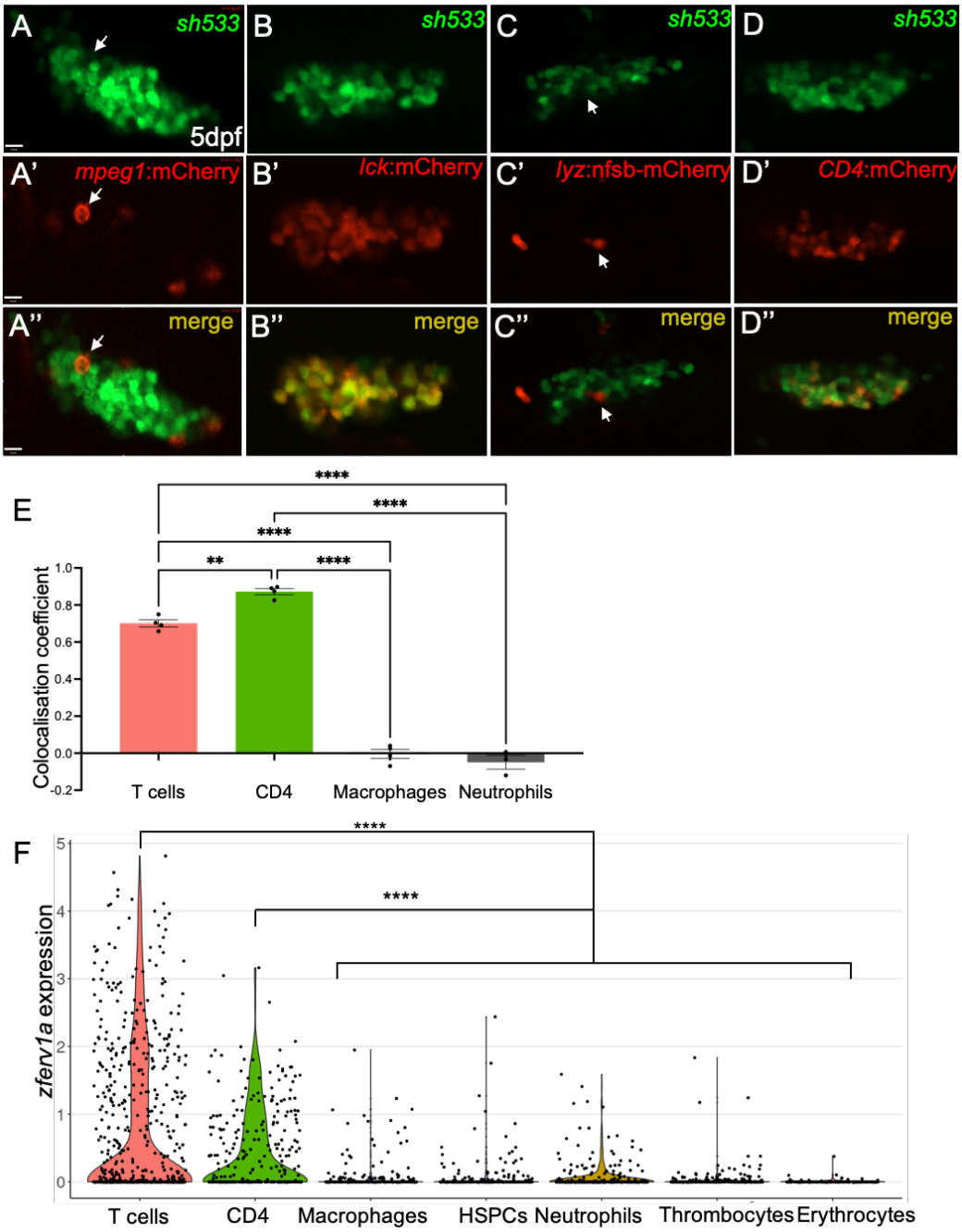


Figure 5: Immune expression of *zferv1a* is restricted to CD4 and T-cells.

A-D. Single slice confocal image of the thymus from double transgenic *Tg(zferv1a:*GFP*)sh533 x Tg(mpeg1:*mCherryCAAX)*sh378* labelling macrophage membrane (A-A''- white arrow pointing to a single macrophage), x *Tg(lck:*mCherry) labelling T cells (B-B''), *x Tg(lyz:nfsb*-mCherry)*sh260* labelling neutrophils (C-C''- white arrow pointing to a single neutrophil), and x *Tg(CD4:*mCherryCAAX) labelling CD4 cells (D-D'') with merge image used for colocalisation analysis.

E. Histogram plot of ImageJ-generated colocalisation coefficient of *Tg(zferv1a:*GFP*)sh533* crossed to *Tg(mpeg1:*mCherryCAAX)*sh378* for macrophages*, Tg(lyz:nfsb*-mCherry)*sh260* for neutrophils, *Tg(CD4:*mCherry*)* for CD4+ cells and *Tg(lck:*mCherry*)* for T cells (one-way ANOVA, n=4).

F. Violin plot of single-cell RNA sequencing datasets MTAB-5530 and E-MTAB-4617 from zebrafish adult whole kidney marrow and spleen showing expression of *zferv1a* (*ENSDARG00000110878*) in multiple haematopoietic lineages. Mann-Whitney with paired Wilcox Test. $p = 1.95 \times 10^{-11}$ for CD4 versus the other clusters and $p = 1.97 \times 10^{-45}$ for T cells versus the other clusters.

DISCUSSION

We describe here a novel reporter line in zebrafish to follow expression of an endogenous retrovirus in real time *in vivo*. Our results validate and refine previously published expression data of *zferv* in the thymus and specifically in the lymphoid lineage. We identified a close relative to *zferv*, here named *zferv1a*, and showed transcriptional activity mostly restricted to T lymphocytes as previously reported (Frazer et al. 2012; Shen and Steiner 2004), but also observed in neurons.

Stability of engineered-ERV transgene

Using a transgenic approach, we imaged *in vivo* in real time the behaviour of an ERV for the first time. Transgenesis entails random insertion of the transgene in the genome, which in some cases can disrupt neighbouring gene expression or induce silencing of the transgene (Akitake et al. 2011). Our initial ERV transgene Tg(*ltr5*:GFP:*ltr3*)*sh500* was engineered to resemble the ERV structure and keep both LTR sequences. However, this line started to show signs of genetic instability and silencing from the F2 generation. It is possible that the location of the insertion triggered silencing of the transgene, however, we suspect that due to the similarity to a retrovirus genome, the transgene was targeted by epigenetic mechanisms specific to ERV. Indeed, transcriptional repression of ERV has been extensively studied and key histone and DNA methylation modifications identified (Maksakova et al. 2008; Rowe et al. 2010; Turelli et al. 2014), which could be targeting the LTR promoter we used. Previous work on *zferv* identified amplification of the *zferv* locus initially in T cells from an acute lymphoblastic leukemia zebrafish model and, although at a lesser rate, also in normal T cells (Frazer et al. 2012). We suspect that genomic amplification of the initial transgene *ltr5*:GFP:*ltr3* would be facilitated in the initial sh500 transgenic line as the transgene would be subjected to the transposase activity of other functional ERVs. This in turn might have activated targeting and silencing of the transgene within a generation. Either or both of those mechanisms might explain why we still observed some mild (neuron-specific) silencing phenotypes in the second transgenic line using a less recognisable ERV structure, with the 3'LTR sequence replaced by a normal polyA tail in Tg(*ltr5*:GFP:*pA*)*sh533*.

Role of ERVs in T lymphocytes

Despite stability issues of the transgene, we consistently observed the signature expression of *zferv* in lymphocytes, already identified in zebrafish transcriptional studies by others (Frazer et al. 2012; Shen and Steiner 2004). The role of ERV expression in T cells is still under investigation. Each step of T cell differentiation is accompanied by a sequence of wide transcriptional changes that allows 1) commitment to the T cell lineage, 2) T cell receptor (TCR) rearrangement and 3) positive or negative selection (Mingueneau et al. 2013). It is possible therefore that transcriptional de-repression of neighbouring genes allows ERV to be expressed during those stages, which would explain such a strong expression during those early phases of T cell development.

Whether ERV expression is merely a side effect of neighbouring gene activation or plays a crucial role during T cell development is an important question to answer. ERV do confer advantages to our adaptive immune system, with important implications for infectious and cancer therapeutic approaches. T cells that have been presented with ERV peptides during their initial thymic selection have a higher sensitivity to exogenous retrovirus infection, conferring an advantage in fighting HIV (Young et al. 2012). And very recently, T cell populations were found with the ability to recognise the antigen of multiple HERV in patients with haematological cancers, with HERV-T cell less present in healthy donors (Saini et al. 2020). This finding might lead to selective therapies using HERV as a target to kill tumour cells.

Relevance of neuronal expression of ERV for human studies

We are providing evidence for a member of the ERV family to be expressed in the healthy developing brain of a vertebrate. Reporter expression suggests that this neuronal expression might be more easily silenced than in T cells; the relevance of this finding for endogenous *zferv* expression, and its physiological significance, remain to be investigated. The role of endogenous retrovirus expression in developing neurons deserves further investigation as the potential role of ERVs in neurodegenerative and neurological diseases is a subject under intense study. The implications of aberrant presence of ERVs in the brain in human neurological diseases are actively being investigated, with the human HERV-K found upregulated in cortical neurons of ALS and MS patients (ref), with currently active clinical trials targeting ERV with anti-retroviral therapies for ALS, MS and the autoinflammatory disease AGS (Gold et al. 2018, 2019; G. I. Rice et al. 2018).

Our finding in embryos and adults suggests a role of ERV in young neurons, potentially highlighting a role for ERV in neuron development and maturation. There is evidence that other types of retroelements are expressed in neurons where they might function as regulators of progenitor differentiation. Indeed, brain expression of other retroelements has previously been reported, mainly focused on the family of non-LTR element, the Long Interspersed Nuclear Element L1, comprising the largest percentage of the human genome of all RE accounting for 17% of all genomic DNA (Scott et al. 1987). In the brain, L1 elements are de-repressed to allow neuronal precursor cells differentiation and contribute to individual somatic mosaicism in human (Coufal et al. 2009) and mouse (Muotri et al. 2005). The latest study also used an engineered reporter approach by tagging GFP to the retroelement L1 (Muotri et al. 2005). Although L1 elements are not related to ERV which is the subject of our study, it was recently discovered that additional groups of RE are enriched in specific regions of the human brain, LTR elements in the cerebellum in particular (Bogu et al. 2019). With the advances in single cell sequencing, it is possible that expression of different groups of ERV, specifically human ERVs, will also be identified in the brain to begin a complete understanding of the presence and role of ERV in brain development and diseases (Evans and Ann 2020).

Conclusions and future work

The zebrafish has been successfully used to model human neurodegenerative and neurodevelopmental disorders. Early timepoint of brain development, particularly difficult to study in mammals, are studied with ease in the zebrafish due to its transparency. As such, zebrafish models of neurodevelopmental autoinflammatory disorders (Hamilton et al. 2020; Haud et al. 2011; Kasher et al. 2015) or neurodegenerative disease such as Multiple Sclerosis (Kulkarni et al. 2017) or ALS (Da Costa et al. 2014) can be combined to our *zferv1a* reporter line to elucidate the role of ERV in the development of the disease.

The first zebrafish reporter line for an ERV described in our study provides a tool to understand how certain pathological states activates ERV expression in specific cell lineages, using a real time *in vivo* approach.

METHODS

Zebrafish husbandry

We used the following transgenic lines: Tg(*mpeg1*:mCherryCAAX)*sh378* labelling the membrane of macrophages and microglia (Bojarczuk et al. 2016), Tg(*lck*:mCherry), labelling T cells (Langenau et al. 2004) but also NK (Hernández et al. 2018), Tg(*CD4*:mCherry) labelling CD4+ T cells and some macrophage/dendritic cells (Dee et al. 2016), the *Tg(lyz:nfsb*-mCherry)*sh260* for neutrophils (Buchan et al. 2019).
All zebrafish were raised in the Bateson Centre at the University of Sheffield in UK Home Office approved aquaria and maintained following standard protocols(Nüsslein-Volhard and Dham 2002). Tanks were maintained at 28°C with a continuous re-circulating water supply and a daily light/dark cycle of 14/10 hours. All procedures were performed on embryos less than 5.2 dpf which were therefore outside of the Animals (Scientific Procedures) Act, to standards set under AWERB (Animal Welfare and Ethical Review
Bodies) and UK Home Office–approved protocols.

Identification of *zferv* genomes

We used the published full *zferv* nucleotide sequence (AF503912) as a query for a nucleotide BLAST search on zebrafish genome assembly on the Ensembl website (http://www.ensembl.org/Danio_rerio/Tools/Blast?db=core), using the most recent reference genome (GRCz11, as of July 2017), choosing the "BLAT" option favouring highly similar sequences, and disabling the filter for species-specific repeats. We selected only hits with homology found in both the *gag/pol* and the *env* regions. Subsequent sequence alignments of LTRs, repetitive element, and open reading frames, and annotations were performed using DNA strider 2.0.

zferv1a full genome cloning

The full zferv1a sequence was subcloned from fosmid 1930h03 (Sanger Center, Cambridge, UK) predicted to contain a ~37kb of chromosome 4 encompassing the sequence, according the annotations of the zv9 version of the zebrafish genome (Ensembl.org). The sequence, plus ~100 bases of flanking genomic sequence on each side, was amplified using primers 5'-ccctgctcattcaacaccatac-3' and 5'-cccgtctgtgaattaccaagc-3', and cloned into PCR4-TOPO (Thermofisher).

RNA in situ hybridisation

The *env1a* probe was made from the plasmid containing the full *zferv1a* genome with the following primers: TGGGGATCCATGAATAAAATAAACAAATTGG, BamH1_env1a_fwd and CCACCCGGGCACCATATCCAATAGTTCCTCC, SmaI_env1a_rev generating a 1935bp (env1a) fragment. The *pol1a* probe was made from the same plasmid using TGGGGATCCGGCAGCAGACGCCGCTGCTA, BamH1_pol1a_fwd, TGGTCTAGACCTCAGGCTCCTCAGTGTCT, Xba1_pol1a_rev generating a 1295bp fragment. The *env1b* probe was made from cDNA using TGGGGATCCatgaATATAAACAAATTGGTGG, BamH1_env1b_fwd, CCACCCGGGCAGAACGCTATAGTCAGTGCTC,SmaI_env1b_rev generating a 635bp (env1a) fragment. Fragments were subsequently cloned into Zero Blunt™ TOPO™ vector (Invitrogen) and reverse transcribed into RNA using the DIG labelling kit (Roche) (SP6 enzyme for antisense probe and T7 for sense probe). Whole mount in-situ hybridisation on wild type (WT) nacre zebrafish was performed to identify the spatial and temporal pattern of *zferv1a* expression using the envelope (*env*) probe env1a and the reverse transcriptase probe *pol1a* and from *zferv1b* using the envelope probe *env1b*. We fixed WT embryos from the *nacre* strain at different stages and followed in situ hybridisation protocol as previously described (Thisse and Thisse 2008).

***zferv1a* and *lck:mCherry* transgenic lines generation**

LTR5 and LTR3 regions of *zferv1a* were cloned into p5E and p3E Gateway entry vectors respectively using the following primers: TGTGGATCCTCTCTTTCGAGATCAAGAGAGGG BamH1_LTR3'_fwd, ACACTCGAGGCTGCACCTTGTTGGAAAATGG Xho1_LTR3'_rev, TGTGGTACCCCTCTCTCTTGTAAAGGTTGAGGG kpn1_LTR5'_fwd, ACACCGCGGTTTAATAATGGTTTTGTCTCCC sacII_LTR5'_rev.

LR reaction were performed into the pDestCRY-CFP vector and injected into 1 cell stage zebrafish embryos to create the following transgenic lines: *Tg(LTR5:GFP:LTR3)sh500*, and *Tg(LTR5:GFP:pA)sh533*, subsequently called *Tg(zferv1a:*GFP*)sh533.*

The already described *lck* promoter (Langenau et al. 2004) was used to create the transgene *lck*:mCherry in the pDestCRYmCherry vector by Gateway recombination. Stable lines were created by injecting these constructs with Tol2 mRNA in one-cell stage embryos and selecting founders using the red eye marker.

Tyrosinase CRISPR/Cas9 crispant generation

Synthetic SygRNA (crRNA and tracrRNA) (Merck) in combination with cas9 nuclease protein (Merck) was used for gene editing. Transactivating RNAs (tracrRNA) and gene-specific CRISPR RNAs (crRNA) were resuspended to a concentration of 20 µM in nuclease-free water containing 10 mM Tris–HCl pH 8. SygRNA complexes were assembled on ice immediately before use using a 1:1:1 ratio of crRNA:tracrRNA:Cas9 protein. Gene-specific crRNAs were designed using the online tool CHOPCHOP (http://chopchop.cbu.uib.no/). Tyrosinase crRNA tyr: GGACTGGAGGACTTCTGGGG(AGG)(Isles et al. 2021).

Flow cytometry analysis

The head kidney of 3 adult *Tg(zferv1a:*GFP*)sh533* and 3 wild type *nacre* fish were dissected and collected into cold Leibovitz 15 media supplemented with 20% FBS and 0.5mM EDTA. The tissue was pipetted up and down to create a single cell suspension and kept on ice. Negative gate was set up using cells from *nacre* animals with the Blue530 detector on an FACSARIA IIu flow cytometry machine and all analysis done using the FlowJo software. Gating was performed as described in supplementary figure 5 using the GFP positive population back gated onto the different cell type population: monocytes, progenitors, T cells and erythrocytes.

Colocalisation imaging and analysis

*Tg(zferv1a:*GFP*)sh533* was crossed to Tg(*lck*:mCherry), Tg(*CD4*:mCherry), *Tg(lyz:nfsb-*mCherry)*sh260*, Tg(*mpeg1*:mCherry:CAAX)*sh378* and double positive animals were selected and imaged at 5dpf using the Perkin Elmer Spinning Disk system for 3 hours. Whole Z-stacks were imported into FiJi to perform colocalisation analysis using the Colocalisation Threshold Pluggin. Whole Z-ztack at 3 different time points during the timelapse in 4 different larvae were used with the same threshold values to determine the colocalisation coefficient by Pearson's correlation.

Single Cell RNA sequencing dataset analysis

Expression of *zferv1a (ENSDARG00000110878*) in different haematopoietic lineages was obtained from single cell RNA sequencing databases of 7 different transgenic lines each labelling a different lineage (Athanasiadis et al. 2017). Single-cell RNA-Sequencing raw counts files of zebrafish blood cells were obtained from the Single Cell Expression Atlas (https://www.ebi.ac.uk/gxa/sc/home, accession codes: E-MTAB-5530 and E-MTAB-4617). The data was then processed and analysed using the Seurat software package for R (Hao et al. 2021; Satija et al. 2015). The data was selected to have a minimum of 200 genes in each cell and then merged and normalised. The Seurat wrapper fastMNN was used to correct for batch effects and UMAP analysis was performed on the resulting combined dataset which resulted in 9 clusters. The function FindAllMarkers was used to

find the differentially expressed genes in each cluster and these were compared to marker genes published in the studies to assign cluster identity. Expression distributions across clusters were obtained for ENSDARG00000110878. Differential expression of CD4 and T cells versus all cells (Average log2 fold change = 1.98, adjusted p.value - 5.31e-81). Statistical differential expression was performed in Seurat between each cluster with Mann-Whitney test followed by a Kruskal-Wallis multiple testing correction.

zferv1a overexpression

The plasmid containing the full *zferv1a* gene was injected in equimolar concentration alongside a plasmid containing GFP into 1-cell stage zebrafish embryos, using GFP only injected embryos as control. RNA was extracted from 30hpf GFP positive embryos from both groups using Trizol and cDNA was synthesised as previously described (Hamilton et al. 2020). qPCR was performed to measure level of expression of GFP and *zferv1a* normalised to the *ef1a* reference gene using the following primers: egfpforward: CCATCTTCTTCAAGGACGAC, egfpreverse: CGTTGTGGCTGTTGTAGTTG, ZFERVpol-1S: GCTAGGACATCCCATTGTGT, ZFERVpol-2A: GGGAATGTGTTCTGGTGTCT, Ef1a-1S :GCTGATCGTTGGAGTCAACA, Ef1a-2A: ACAGACTTGACCTCAGTGGT.

Statistical analysis

All statistical analysis were performed in GraphPad Prism where data was entered using a grouped table for ANOVA analysis (more than 2 samples or 2 variables). Sample distribution was assessed using frequency of distribution analysis. n (experimental unit) number stated for each experiment in figure legends. p values are indicated and a star system is used instead for graph with multiple comparisons: *=p<0.05, **=p<0.01, ***=p<0.001, ****=p<0.0001. Following the recommendation of the American Statistical Association we do not associate a specific p value with significance (Wasserstein, Schirm, and Lazar 2019).

Acknowledgements

We thank the flow cytometry facility at the University of Sheffield, The Light Microscopy Facility and the Zebrafish facility staff at The Bateson Centre for their technical help. We thank Adam Hurlstone (The University of Manchester) for supplying the Tg(*lck*:mCherry) transgenic line, Ana Cvejic for the immune single cell sequencing datset and Angela Ciuffi from Institute of Microbiology (IMUL), Lausanne University Hospital and University of Lausanne for supervising ADJ.

Author contributions

NH and SAR acquired funding and designed the study; NH, AC, HAR, JJP, HMI, EGC, JPL, AC collected and analysed data. EVC performed all bioinformatic work. ADJ created crucial reagent. NH wrote the manuscript; NH, SAR, JPL revised the manuscript.

References

Ahmad, Sadeem, Xin Mu, Fei Yang, Emily Greenwald, Ji Woo Park, Etai Jacob, Cheng-Zhong Zhang, and Sun Hur. 2018. "Breaching Self-Tolerance to Alu Duplex RNA Underlies MDA5-Mediated Inflammation." *Cell* 172(4):797-802.e13. doi: 10.1016/j.cell.2017.12.016.

Akitake, Courtney M., Michelle Macurak, Marnie E. Halpern, and Mary G. Goll. 2011. "Transgenerational Analysis of Transcriptional Silencing in Zebrafish." *Developmental Biology* 352(2):191–201. doi: 10.1016/j.ydbio.2011.01.002.

Athanasiadis, Emmanouil I., Jan G. Botthof, Helena Andres, Lauren Ferreira, Pietro Lio, and Ana Cvejic. 2017. "Single-Cell RNA-Sequencing Uncovers Transcriptional States and Fate Decisions in Haematopoiesis." *Nature Communications* 8(1):2045. doi: 10.1038/s41467-017-02305-6.

Benitez-Guijarro, Maria, Cesar Lopez-Ruiz, Žygimantė Tarnauskaitė, Olga Murina, Mahwish Mian Mohammad, Thomas C. Williams, Adeline Fluteau, Laura Sanchez, Raquel Vilar-Astasio, Marta Garcia-Canadas, David Cano, Marie-Jeanne HC Kempen, Antonio Sanchez-Pozo, Sara R. Heras, Andrew P. Jackson, Martin AM Reijns, and Jose L. Garcia-Perez. 2018. "RNase H2, Mutated in Aicardi-Goutières Syndrome, Promotes LINE-1 Retrotransposition." *The EMBO Journal* 37(15). doi: 10.15252/embj.201798506.

Bermejo, Amaia Vergara, Emeline Ragonnaud, Joana Daradoumis, and Peter Holst. 2020. "Cancer Associated Endogenous Retroviruses: Ideal Immune Targets for Adenovirus-Based Immunotherapy." *International Journal of Molecular Sciences* 21(14):1–21. doi: 10.3390/ijms21144843.

Bogu, Gireesh K., Ferran Reverter, Marc A. Marti-Renom, Michael P. Snyder, and Roderic Guigó. 2019. "Atlas of Transcriptionally Active Transposable Elements in Human Adult Tissues." *BioRxiv* : https://. doi: 10.1101/714212.

Bojarczuk, Aleksandra, Katie A. Miller, Richard Hotham, Amy Lewis, Nikolay V. Ogryzko, Alfred A. Kamuyango, Helen Frost, Rory H. Gibson, Eleanor Stillman, Robin C. May, Stephen A. Renshaw, and Simon A. Johnston. 2016. "Cryptococcus Neoformans Intracellular Proliferation and Capsule Size Determines Early Macrophage Control of Infection." *Scientific Reports* 6(21489). doi: 10.1038/srep21489.

Bourque, Guillaume, Kathleen H. Burns, Mary Gehring, Vera Gorbunova, Andrei Seluanov, Molly Hammell, Michaël Imbeault, Zsuzsanna Izsvák, Henry L. Levin, Todd S. Macfarlan, Dixie L. Mager, and Cédric Feschotte. 2018. "Ten Things You Should Know about Transposable Elements." *Genome Biology* 19(1):1–12. doi: 10.1186/s13059-018-1577-z.

Buchan, Kyle D., Tomasz K. Prajsnar, Nikolay V. Ogryzko, Nienke W. M. de Jong, Michiel van Gent, Julia Kolata, Simon J. Foster, Jos A. G. van Strijp, and Stephen A. Renshaw. 2019. "A Transgenic Zebrafish Line for in Vivo Visualisation of Neutrophil Myeloperoxidase." *PLoS ONE* 14(4):e0215592. doi: 10.1101/456541.

Chiappinelli, Katherine B., Pamela L. Strissel, Alexis Desrichard, Huili Li, Christine Henke, Benjamin Akman, Alexander Hein, Neal S. Rote, Leslie M. Cope, Alexandra Snyder, Vladimir Makarov, Sadna Buhu, Dennis J. Slamon, Jedd D. Wolchok, Drew M. Pardoll, Matthias W. Beckmann, Cynthia A. Zahnow, Taha Mergoub, Timothy A. Chan, Stephen B. Baylin, and Reiner Strick. 2015. "Inhibiting DNA Methylation Causes an Interferon Response in Cancer via DsRNA Including Endogenous Retroviruses." *Cell* 162(5):974–86. doi: 10.1016/j.cell.2015.07.011.

Chung, Hachung, Jorg J. A. Calis, Xianfang Wu, Tony Sun, Yingpu Yu, Stephanie L. Sarbanes, Viet Loan Dao Thi, Abigail R. Shilvock, H. Heinrich Hoffmann, Brad R. Rosenberg, and Charles M. Rice. 2018. "Human ADAR1 Prevents Endogenous RNA from Triggering Translational Shutdown." *Cell* 172(4):811-824.e14. doi: 10.1016/j.cell.2017.12.038.

Chuong, Edward B. 2018. "The Placenta Goes Viral: Retroviruses Control Gene Expression in Pregnancy." *PLoS Biology* 16(10):e3000028. doi: 10.1371/journal.pbio.3000028.

Chuong, Edward B., Nels C. Elde, and Cédric Feschotte. 2016. "Regulatory Evolution of Innate Immunity through Co-Option of Endogenous Retroviruses." *Science* 351(6277):1083.

Da Costa, Marc M. J., Claire E. Allen, Adrian Higginbottom, Tennore Ramesh, Pamela J. Shaw, and Christopher J. McDermott. 2014. "A New Zebrafish Model Produced by TILLING of SOD1-Related Amyotrophic Lateral Sclerosis Replicates Key Features of the Disease and Represents a Tool for in Vivo Therapeutic Screening." *DMM Disease Models and Mechanisms* 7(1):73–81. doi: 10.1242/dmm.012013.

Coufal, Nicole G., José L. Garcia-Perez, Grace E. Peng, Gene W. Yeo, Yangling Mu, Michael T. Lovci, Maria Morell, K. Sue O'Shea, John V. Moran, and Fred H. Gage. 2009. "L1 Retrotransposition in Human Neural Progenitor Cells." *Nature* 460(7259):1127–31. doi: 10.1038/nature08248.

Crow, Yanick J., Diana S. Chase, Johanna Lowenstein Schmidt, Marcin Szynkiewicz, Gabriella M. a Forte, Hannah L. Gornall, Anthony Oojageer, Beverley Anderson, Amy Pizzino, Guy Helman, Mohamed S. Abdel-hamid, Ghada M. Abdel-salam, Sam Ackroyd, Alec Aeby, Guillermo Agosta, Catherine Albin, Stavit Allon-shalev, Montse Arellano, Giada Ariaudo, Vijay Aswani, Riyana Babul-hirji, Eileen M. Baildam, Nadia Bahi-buisson, Kathryn M. Bailey, Christine Barnerias, Magalie Barth, Roberta Battini, Marika Bianchi, Thierry Billette De, Michael W. Beresford, Edward M. Blair, Miriam Bloom, Alberto B. Burlina, Maria Luisa Carpanelli, Daniel R. Carvalho, Manuel Castro-gago, Anna Cavallini, Cristina Cereda, Kate E. Chandler, David a Chitayat, Abigail E. Collins, Concepcion Sierra Corcoles, Nuno J. V Cordeiro, Giovanni Crichiutti, Lyvia Dabydeen, Russell C. Dale, Stefano D. Arrigo, Christian G. E. L. De Goede, Corinne De Laet, Liesbeth M. H. De Waele, Ines Denzler, Isabelle Desguerre, Koenraad Devriendt, Maja Di Rocco, Michael C. Fahey, Blanca Gener, Cyril Goizet, Nirmala R. Gowrinathan, Kalpana Gowrishankar, Donncha Hanrahan, Bertrand Isidor, Nasaim Khan, Mary D. King, Edwin P. Kirk, Ram Kumar, Lieven Lagae, Jean-pierre S. Lin, Tarja Linnankivi, Mark T. Mackay, Daphna R. Marom, Charles Marques Lourenc, Shane a Mckee, Isabella Moroni, Jenny E. V Morton, Mohnish Suri, Uta Tacke, Tiong Y. Tan, Water Naude, Keng Wee Teik, Maya Mary Thomas, Marianne Till, Davide Tonduti, Enza Maria Valente, Rudy Noel Van Coster, Marjo S. Van Der Knaap, Grace Vassallo, Raymon Vijzelaar, Julie Vogt, Geoffrey B. Wallace, Evangeline Wassmer, Hannah J. Webb, William P. Whitehouse, Robyn N. Whitney, Maha S. Zaki, Sameer M. Zuberi, John H. Livingston, Flore Rozenberg, Pierre Lebon, Adeline Vanderver, Simona Orcesi, and Gillian I. Rice. 2015. "Characterization of Human Disease Phenotypes Associated with Mutations in TREX1, RNASEH2A, RNASEH2B, RNASEH2C, SAMHD1, ADAR, and IFIH1." *American Journal of Medical Genetics Part A* 296–312.

Dee, Christopher T., Raghavendar T. Nagaraju, Emmanouil I. Athanasiadis, Caroline Gray, Laura Fernandez del Ama, Simon A. Johnston, Christopher J. Secombes, Ana Cvejic, and Adam F. L. Hurlstone. 2016. "CD4-Transgenic Zebrafish Reveal Tissue-Resident Th2- and Regulatory T Cell–like

Populations and Diverse Mononuclear Phagocytes." *The Journal of Immunology* 197(9):3520–30. doi: 10.4049/jimmunol.1600959.
Evans, Taylor A., and Jennifer Ann. 2020. "Retroelement-Derived RNA and Its Role in the Brain." *Seminars in Cell and Developmental Biology* doi.org/10.1016/j.semcdb.2020.11.001. doi: 10.1016/j.semcdb.2020.11.001.
Feschotte, Cédric. 2008. "Transposable Elements and the Evolution of Regulatory Networks." *Nature Reviews Genetics* 9(5):397–405. doi: 10.1038/nrg2337.
Frazer, J. Kimble, Lance A. Batchelor, Diana F. Bradley, Kim H. Brown, Kimberly P. Dobrinski, Charles Lee, and Nikolaus S. Trede. 2012. "Genomic Amplification of an Endogenous Retrovirus in Zebrafish T-Cell Malignancies." *Adv Hematol* doi: 10.1155/2012/627920. doi: 10.1155/2012/627920.
Göke, Jonathan, Xinyi Lu, Yun Shen Chan, Huck Hui Ng, Lam Ha Ly, Friedrich Sachs, and Iwona Szczerbinska. 2015. "Dynamic Transcription of Distinct Classes of Endogenous Retroviral Elements Marks Specific Populations of Early Human Embryonic Cells." *Cell Stem Cell* 16(2):135–41. doi: 10.1016/j.stem.2015.01.005.
Gold, Julian, Monica Marta, Ute C. Meier, Tove Christensen, David Miller, Daniel Altmann, David Holden, Lucia Bianchi, Rocco Adiutori, David MacManus, Tarek Yousry, Klaus Schmierer, Benjamin Turner, and Gavin Giovannoni. 2018. "A Phase II Baseline versus Treatment Study to Determine the Efficacy of Raltegravir (Isentress) in Preventing Progression of Relapsing Remitting Multiple Sclerosis as Determined by Gadolinium-Enhanced MRI: The INSPIRE Study." *Multiple Sclerosis and Related Disorders* 24:123–28. doi: 10.1016/j.msard.2018.06.002.
Gold, Julian, Dominic B. Rowe, Matthew C. Kiernan, Steve Vucic, Susan Mathers, Ruben P. A. van Eijk, Avindra Nath, Marta Garcia Montojo, Gina Norato, Ulisses A. Santamaria, Mary Louise Rogers, Andrea Malaspina, Vittoria Lombardi, Puja R. Mehta, Henk Jan Westeneng, Leonard H. van den Berg, and Ammar Al-Chalabi. 2019. "Safety and Tolerability of Triumeq in Amyotrophic Lateral Sclerosis: The Lighthouse Trial." *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* 20(7–8):595–604. doi: 10.1080/21678421.2019.1632899.
Hamilton, Noémie, Holly A. Rutherford, Jessica J. Petts, Thomas Weber, Marco Henneke, Stephen A. Renshaw, and Mark J. Dunning. 2020. "The Failure of Microglia to Digest Developmental Apoptotic Cells Contributes to the Pathology of RNASET2-Deficient Leukoencephalopathy." *GLIA* 68(7):1531–45. doi: 10.1002/glia.23829.
Hao, Yuhan, Stephanie Hao, Erica Andersen-Nissen, William M. Mauck, Shiwei Zheng, Andrew Butler, Maddie J. Lee, Aaron J. Wilk, Charlotte Darby, Michael Zager, Paul Hoffman, Marlon Stoeckius, Efthymia Papalexi, Eleni P. Mimitou, Jaison Jain, Avi Srivastava, Tim Stuart, Lamar M. Fleming, Bertrand Yeung, Angela J. Rogers, Juliana M. McElrath, Catherine A. Blish, Raphael Gottardo, Peter Smibert, and Rahul Satija. 2021. "Integrated Analysis of Multimodal Single-Cell Data." *Cell* 184(13):3573-3587.e29. doi: 10.1016/j.cell.2021.04.048.
Haud, Noémie, Firat Kara, Simone Diekmann, Marco Henneke, Jason R. Willer, Melissa S. Hillwig, Ronald G. Gregg, Gustavo C. Macintosh, Jutta Gärtner, A. Alia, and Adam F. L. Hurlstone. 2011. "Rnaset2 Mutant Zebrafish Model

Familial Cystic Leukoencephalopathy and Reveal a Role for RNase T2 in Degrading Ribosomal RNA." *Proceedings of the National Academy of Sciences of the United States of America* 108(3):1099–1103. doi: 10.1073/pnas.1009811107.

Hernández, Pedro P., Paulina M. Strzelecka, Emmanouil I. Athanasiadis, Ana F. Robalo, Catherine M. Collins, Pierre Boudinot, Jean-Pierre Levraud, and Ana Cvejic. 2018. "Single-Cell Transcriptional Analysis Reveals Naïve Helper ILC-like Cells in Zebrafish." *SCIENCE IMMUNOLOGY* 342477. doi: DOI: 10.1126/sciimmunol.aau5265.

Isles, Hannah M., Catherine A. Loynes, Sultan Alasmari, Fu Chuen Kon, Katherine M. Henry, Anastasia Kadochnikova, Jack Hales, Clare F. Muir, Maria Cristina Keightley, Visakan Kadirkamanathan, Noémie Hamilton, Graham J. Lieschke, Stephen A. Renshaw, and Philip M. Elks. 2021. "Pioneer Neutrophils Release Chromatin within in Vivo Swarms." *ELife* 10. doi: 10.7554/eLife.68755.

Johnston, James B., Claudia Silva, Janet Holden, Kenneth G. Warren, Arthur W. Clark, and Christopher Power. 2001. "Monocyte Activation and Differentiation Augment Human Endogenous Retrovirus Expression: Implications for Inflammatory Brain Diseases." *Annals of Neurology* 50(4):434–42. doi: 10.1002/ana.1131.

Jorgensen, T. N., B. L. Kotzin, T. J. Vyse, B. J. Morley, C. L. Roark, L. H. Evans, L. Fossati-Jimack, S. Kikuchi, S. Hirose, S. Izui, L. Baudino, K. Yoshinobu, R. M. Tucker, and N. Morito. 2014. "Dissection of Genetic Mechanisms Governing the Expression of Serum Retroviral Gp70 Implicated in Murine Lupus Nephritis." *The Journal of Immunology* 181(4):2846–54. doi: 10.4049/jimmunol.181.4.2846.

Kasher, Paul R., Emma M. Jenkinson, Valérie Briolat, David Gent, Catherine Morrissey, Leo A. H. Zeef, Gillian I. Rice, Jean-Pierre Levraud, and Yanick J. Crow. 2015. "Characterization of *Samhd1* Morphant Zebrafish Recapitulates Features of the Human Type I Interferonopathy Aicardi-Goutières Syndrome." *The Journal of Immunology* 194(6):2819–25. doi: 10.4049/jimmunol.1403157.

Kissa, Karima, Emi Murayama, Agustin Zapata, Alfonso Cortés, Emmanuelle Perret, Christophe Machu, and Philippe Herbomel. 2008. "Live Imaging of Emerging Hematopoietic Stem Cells and Early Thymus Colonization." *Blood* 111(3):1147–56. doi: 10.1182/blood-2007-07-099499.

Knerr, Ina, B. Huppertz, C. Weigel, J. Dötsch, C. Wich, R. L. Schild, M. W. Beckmann, and W. Rascher. 2004. "Endogenous Retroviral Syncytin: Compilation of Experimental Research on Syncytin and Its Possible Role in Normal and Disturbed Human Placentogenesis." *Molecular Human Reproduction*. doi: 10.1093/molehr/gah070.

Kulkarni, Pushkar, Swapna Yellanki, Raghavender Medishetti, Dharmarajan Sriram, Uday Saxena, and Perumal Yogeeswari. 2017. "Novel Zebrafish EAE Model: A Quick in Vivo Screen for Multiple Sclerosis." *Multiple Sclerosis and Related Disorders* 11(August 2016):32–39. doi: 10.1016/j.msard.2016.11.010.

Kunarso, Galih, Na Yu Chia, Justin Jeyakani, Catalina Hwang, Xinyi Lu, Yun Shen Chan, Huck Hui Ng, and Guillaume Bourque. 2010. "Transposable Elements Have Rewired the Core Regulatory Network of Human Embryonic Stem Cells." *Nature Genetics* 42(7):631–34. doi: 10.1038/ng.600.

Lander, Eric S., Lauren M. Linton, Bruce Birren, Chad Nusbaum, Michael C. Zody, Jennifer Baldwin, Keri Devon, Ken Dewar, Michael Doyle, William Fitzhugh, Roel Funke, Diane Gage, Katrina Harris, Andrew Heaford, John Howland, Lisa Kann, Jessica Lehoczky, Rosie Levine, Paul McEwan, Kevin McKernan, James Meldrim, Jill P. Mesirov, Cher Miranda, William Morris, Jerome Naylor, Christina Raymond, Mark Rosetti, Ralph Santos, Andrew Sheridan, Carrie Sougnez, Nicole Stange-Thomann, Nikola Stojanovic, Aravind Subramanian, Dudley Wyman, Jane Rogers, John Sulston, Rachael Ainscough, Stephan Beck, David Bentley, John Burton, Christopher Clee, Nigel Carter, Alan Coulson, Rebecca Deadman, Panos Deloukas, Andrew Dunham, Ian Dunham, Richard Durbin, Lisa French, Darren Grafham, Simon Gregory, Tim Hubbard, Sean Humphray, Adrienne Hunt, Matthew Jones, Christine Lloyd, Amanda McMurray, Lucy Matthews, Simon Mercer, Sarah Milne, James C. Mullikin, Andrew Mungall, Robert Plumb, Mark Ross, Ratna Shownkeen, Sarah Sims, Robert H. Waterston, Richard K. Wilson, Ladeana W. Hillier, John D. McPherson, Marco A. Marra, Elaine R. Mardis, Lucinda A. Fulton, Asif T. Chinwalla, Kymberlie H. Pepin, Warren R. Gish, Stephanie L. Chissoe, Michael C. Wendl, Kim D. Delehaunty, Tracie L. Miner, Andrew Delehaunty, Jason B. Kramer, Lisa L. Cook, Robert S. Fulton, Douglas L. Johnson, Patrick J. Minx, Sandra W. Clifton, Trevor Hawkins, Elbert Branscomb, Paul Predki, Paul Richardson, Sarah Wenning, Tom Slezak, Norman Doggett, Jan Fang Cheng, Anne Olsen, Susan Lucas, Christopher Elkin, Edward Uberbacher, Marvin Frazier, Richard A. Gibbs, Donna M. Muzny, Steven E. Scherer, John B. Bouck, Erica J. Sodergren, Kim C. Worley, Catherine M. Rives, James H. Gorrell, Michael L. Metzker, Susan L. Naylor, Raju S. Kucherlapati, David L. Nelson, George M. Weinstock, Yoshiyuki Sakaki, Asao Fujiyama, Masahira Hattori, Tetsushi Yada, Atsushi Toyoda, Takehiko Itoh, Chiharu Kawagoe, Hidemi Watanabe, Yasushi Totoki, Todd Taylor, Jean Weissenbach, Roland Heilig, William Saurin, Francois Artiguenave, Philippe Brottier, Thomas Bruls, Eric Pelletier, Catherine Robert, Patrick Wincker, André Rosenthal, Matthias Platzer, Gerald Nyakatura, Stefan Taudien, Andreas Rump, Douglas R. Smith, Lynn Doucette-Stamm, Marc Rubenfield, Keith Weinstock, Mei Lee Hong, Joann Dubois, Huanming Yang, Jun Yu, Jian Wang, Guyang Huang, Jun Gu, Leroy Hood, Lee Rowen, Anup Madan, Shizen Qin, Ronald W. Davis, Nancy A. Federspiel, A. Pia Abola, Michael J. Proctor, Bruce A. Roe, Feng Chen, Huaqin Pan, Juliane Ramser, Hans Lehrach, Richard Reinhardt, W. Richard McCombie, Melissa De La Bastide, Neilay Dedhia, Helmut Blöcker, Klaus Hornischer, Gabriele Nordsiek, Richa Agarwala, L. Aravind, Jeffrey A. Bailey, Alex Bateman, Serafim Batzoglou, Ewan Birney, Peer Bork, Daniel G. Brown, Christopher B. Burge, Lorenzo Cerutti, Hsiu Chuan Chen, Deanna Church, Michele Clamp, Richard R. Copley, Tobias Doerks, Sean R. Eddy, Evan E. Eichler, Terrence S. Furey, James Galagan, James G. R. Gilbert, Cyrus Harmon, Yoshihide Hayashizaki, David Haussler, Henning Hermjakob, Karsten Hokamp, Wonhee Jang, L. Steven Johnson, Thomas A. Jones, Simon Kasif, Arek Kaspryzk, Scot Kennedy, W. James Kent, Paul Kitts, Eugene V. Koonin, Ian Korf, David Kulp, Doron Lancet, Todd M. Lowe, Aoife McLysaght, Tarjei Mikkelsen, John V. Moran, Nicola Mulder, Victor J. Pollara, Chris P. Ponting, Greg Schuler, Jörg Schultz, Guy Slater, Arian F. A. Smit, Elia Stupka, Joseph Szustakowki, Danielle Thierry-Mieg, Jean Thierry-Mieg, Lukas

Wagner, John Wallis, Raymond Wheeler, Alan Williams, Yuri I. Wolf, Kenneth H. Wolfe, Shiaw Pyng Yang, Ru Fang Yeh, Francis Collins, Mark S. Guyer, Jane Peterson, Adam Felsenfeld, Kris A. Wetterstrand, Richard M. Myers, Jeremy Schmutz, Mark Dickson, Jane Grimwood, David R. Cox, Maynard V. Olson, Rajinder Kaul, Christopher Raymond, Nobuyoshi Shimizu, Kazuhiko Kawasaki, Shinsei Minoshima, Glen A. Evans, Maria Athanasiou, Roger Schultz, Aristides Patrinos, Michael J. Morgan, Pieter de Jong, Joseph J. Catanese, Kazutoyo Osoegawa, Hiroaki Shizuya, Sangdun Choi, and Yu Juin Chen. 2001. "Erratum: Initial Sequencing and Analysis of the Human Genome: International Human Genome Sequencing Consortium (Nature (2001) 409 (860-921))." *Nature* 412(6846):565–66. doi: 10.1038/35087627.

Langenau, D. M., A. A. Ferrando, D. Traver, J. L. Kutok, J. P. D. Hezel, J. P. Kanki, L. I. Zon, A. T. Look, and N. S. Trede. 2004. "In Vivo Tracking of T Cell Development, Ablation, and Engraftment in Transgenic Zebrafish." *Proceedings of the National Academy of Sciences* 101(19):7369–74. doi: 10.1073/pnas.0402248101.

Levraud, Jean-Pierre, Luc Jouneau, Valérie Briolat, Valerio Laghi, and Pierre Boudinot. 2019. "IFN-Stimulated Genes in Zebrafish and Humans Define an Ancient Arsenal of Antiviral Immunity." *The Journal of Immunology* 203(12):3361–73. doi: 10.4049/jimmunol.1900804.

Li, Peng, Juan Du, John L. Goodier, Jingwei Hou, Jian Kang, Haig H. Kazazian, Ke Zhao, and Xiao Fang Yu. 2017. "Aicardi-Goutières Syndrome Protein TREX1 Suppresses L1 and Maintains Genome Integrity through Exonuclease-Independent ORF1p Depletion." *Nucleic Acids Research* 45(8):4619–31. doi: 10.1093/nar/gkx178.

Li, Wenxue, Myoung Hwa Lee, Lisa Henderson, Richa Tyagi, Muzna Bachani, Joseph Steiner, Emilie Campanac, Dax A. Hoffman, Gloria Von Geldern, Kory Johnson, Dragan Maric, H. Douglas Morris, Margaret Lentz, Katherine Pak, Andrew Mammen, Lyle Ostrow, Jeffrey Rothstein, and Avindra Nath. 2015. "Human Endogenous Retrovirus-K Contributes to Motor Neuron Disease." *Science Translational Medicine*. doi: 10.1126/scitranslmed.aac8201.

Lokossou, Adjimon G., Caroline Toudic, Phuong Trang Nguyen, Xavier Elisseeff, Amandine Vargas, Éric Rassart, Julie Lafond, Line Leduc, Steve Bourgault, Caroline Gilbert, Tatiana Scorza, Jorge Tolosa, and Benoit Barbeau. 2020. "Endogenous Retrovirus-Encoded Syncytin-2 Contributes to Exosome-Mediated Immunosuppression of T Cells." *Biology of Reproduction* 102(1):185–98. doi: 10.1093/biolre/ioz124.

Maksakova, I. A., D. L. Mager, and D. Reiss. 2008. "Endogenous Retroviruses - Keeping Active Endogenous Retroviral-like Elements in Check: The Epigenetic Perspective." *Cellular and Molecular Life Sciences* 65(21):3329–47.

Mameli, Giuseppe, Vito Astone, Giannina Arru, Silvia Marconi, Laura Lovato, Caterina Serra, Stefano Sotgiu, Bruno Bonetti, and Antonina Dolei. 2007. "Brains and Peripheral Blood Mononuclear Cells of Multiple Sclerosis (MS) Patients Hyperexpress MS-Associated Retrovirus/HERV-W Endogenous Retrovirus, but Not Human Herpesvirus 6." *Journal of General Virology* 88(1):264–74. doi: 10.1099/vir.0.81890-0.

Mandl, Judith N., João P. Monteiro, Nienke Vrisekoop, and Ronald N. Germain. 2013. "T Cell-Positive Selection Uses Self-Ligand Binding Strength to Optimize Repertoire Recognition of Foreign Antigens." *Cell* 38(2):263–74. doi: 10.1016/j.immuni.2012.09.011.

Mi, Sha, Xinhua Lee, Xiang-ping Li, Geertruida M. Veldman, Heather Finnerty, Lisa Racie, Edward LaVallie, Xiang-Yang Tang, Philippe Edouard, Steve Howes, James C. Keith, and John M. McCoy. 2000. "Syncytin Is a Captive Retroviral Envelope Protein Involved in Human Placental Morphogenesis." *Nature* 403(6771):785–89. doi: 10.1038/35001608.

Mingueneau, Michael, Taras Kreslavsky, Daniel Gray, Tracy Heng, Richard Cruse, Jeffrey Ericson, Sean Bendall, Matthew H. Spitzer, Garry P. Nolan, Koichi Kobayashi, Harald Von Boehmer, Diane Mathis, Christophe Benoist, Adam J. Best, Jamie Knell, Ananda Goldrath, Daphne Koller, Tal Shay, Aviv Regev, Nadia Cohen, Patrick Brennan, Michael Brenner, Francis Kim, Tata Nageswara Rao, Amy Wagers, Katherine Rothamel, Adriana Ortiz-Lopez, Natalie A. Bezman, Joseph C. Sun, Gundula Min-Oo, Charlie C. Kim, Lewis L. Lanier, Jennifer Miller, Brian Brown, Miriam Merad, Emmanuel L. Gautier, Claudia Jakubzick, Gwendalyn J. Randolph, Paul Monach, David A. Blair, Michael L. Dustin, Susan A. Shinton, Richard R. Hardy, David Laidlaw, Jim Collins, Roi Gazit, Derrick J. Rossi, Nidhi Malhotra, Katelyn Sylvia, Joonsoo Kang, Anne Fletcher, Kutlu Elpek, Angelique Bellemare-Pelletier, Deepali Malhotra, and Shannon Turley. 2013. "The Transcriptional Landscape of Aβ T Cell Differentiation." *Nature Immunology* 14(6):619–32. doi: 10.1038/ni.2590.

Muotri, Alysson R., Vi T. Chu, Maria C. N. Marchetto, Wei Deng, John V. Moran, and Fred H. Gage. 2005. "Somatic Mosaicism in Neuronal Precursor Cells Mediated by L1 Retrotransposition." *Nature* 435(7044):903–10. doi: 10.1038/nature03663.

Nüsslein-Volhard, C., and R. Dham. 2002. "Zebrafish: A Practical Approach." *New York: Oxford University Press* 2002. doi: 10.1017/S0016672303216384.

Okahara, Goichi, Shinobu Matsubara, Takaya Oda, Jun Sugimoto, Yoshihiro Jinno, and Fuminori Kanaya. 2004. "Expression Analyses of Human Endogenous Retroviruses (HERVs): Tissue-Specific and Developmental Stage-Dependent Expression of HERVs." *Genomics* 84(6):982–90. doi: 10.1016/j.ygeno.2004.09.004.

Pi, W., Z. Yang, J. Wang, L. Ruan, X. Yu, J. Ling, S. Krantz, C. Isales, S. J. Conway, S. Lin, and D. Tuan. 2004. "The LTR Enhancer of ERV-9 Human Endogenous Retrovirus Is Active in Oocytes and Progenitor Cells in Transgenic Zebrafish and Humans." *Proceedings of the National Academy of Sciences* 101(3):805–10. doi: 10.1073/pnas.0307698100.

Renshaw, S. A., and N. S. Trede. 2012. "A Model 450 Million Years in the Making: Zebrafish and Vertebrate Immunity." *Disease Models & Mechanisms* 5(1):38–47. doi: 10.1242/dmm.007138.

Renshaw, SA, CA Loynes, DM Trushell, S. Elworthy, PW Ingham, and KBM Whyte. 2006. "A Transgenic Zebrafish Model of Neutrophilic Inflammation." *Blood* 108(13):3976–78. doi: 10.1182/blood-2006-05-024075.The.

Rice, GI, C. Meyzer, N. Bouazza, M. Hully, N. Boddaert, M. Semeraro, LAH Zeef, F. Rozenberg, V. Bondet, D. Duffy, and YJ Crow. 2018. "Reverse-Transcriptase

Inhibitors in the Aicardi – Goutières Syndrome." *N Engl J Med.* 6(379(23)):2275–77.

Rice, Gillian I., Candice Meyzer, Naïm Bouazza, Marie Hully, Nathalie Boddaert, Michaela Semeraro, Leo A.H. Zeef, Flore Rozenberg, Bondet Vincent, Darragh Duffy, Alba Llibre, Jinmi Baek, Mame N. Sambe, Elodie Henry, Valerie Jolaine, Christine Barnerias, Magalie Barth, Alexandre Belot, Claude Cances, and Yanick J. Crow. 2018. *Reverse-Transcriptase Inhibitors in the Aicardi–Goutières Syndrome*. Vol. 379.

Rowe, Helen M., Johan Jakobsson, Daniel Mesnard, Jacques Rougemont, Séverine Reynard, Tugce Aktas, Pierre V. Maillard, Hillary Layard-Liesching, Sonia Verp, Julien Marquis, François Spitz, Daniel B. Constam, and Didier Trono. 2010. "KAP1 Controls Endogenous Retroviruses in Embryonic Stem Cells." *Nature* 463(7278):237–40. doi: 10.1038/nature08674.

Saini, Sunil Kumar, Andreas Due Ørskov, Anne Mette Bjerregaard, Ashwin Unnikrishnan, Staffan Holmberg-Thydén, Annie Borch, Kathrine Valentini Jensen, Govardhan Anande, Amalie Kai Bentzen, Andrea Marion Marquard, Tripti Tamhane, Marianne Bach Treppendahl, Anne Ortved Gang, Inge Høgh Dufva, Zoltan Szallasi, Nicola Ternette, Anders Gorm Pedersen, Aron Charles Eklund, John Pimanda, Kirsten Grønbæk, and Sine Reker Hadrup. 2020. "Human Endogenous Retroviruses Form a Reservoir of T Cell Targets in Hematological Cancers." *Nature Communications* 11(1):1–14. doi: 10.1038/s41467-020-19464-8.

Satija, Rahul, Jeffrey A. Farrell, David Gennert, Alexander F. Schier, and Aviv Regev. 2015. "Spatial Reconstruction of Single-Cell Gene Expression Data." *Nature Biotechnology* 33(5):495–502. doi: 10.1038/nbt.3192.

Scott, Alan F., Barbara J. Schmeckpeper, Mona Abdelrazik, Catherine Theisen Comey, Bruce O'Hara, Judith Pratt Rossiter, Tim Cooley, Peter Heath, Kirby D. Smith, and Louise Margolet. 1987. "Origin of the Human L1 Elements: Proposed Progenitor Genes Deduced from a Consensus DNA Sequence." *Genomics* 1(2):113–25. doi: 10.1016/0888-7543(87)90003-6.

Shen, Ching-hung, and Lisa A. Steiner. 2004. "Genome Structure and Thymic Expression of an Endogenous Retrovirus in Zebrafish." *Journal of Virology* 78(2):899–911. doi: 10.1128/JVI.78.2.899.

Shiroma, T., J. Sugimoto, T. Oda, Y. Jinno, and F. Kanaya. 2001. "Search for Active Endogenous Retroviruses: Identification and Characterization of a HERV-E Gene That Is Expressed in the Pancreas and Thyroid." *Journal of Human Genetics* 46(11):619–25. doi: 10.1007/s100380170012.

Tavakolian, Shaian, Hossein Goudarzi, and Ebrahim Faghihloo. 2019. "Evaluating the Expression Level of HERV-K Env, Np9, Rec and Gag in Breast Tissue." *Infectious Agents and Cancer* 14(1):1–5. doi: 10.1186/s13027-019-0260-7.

Thisse, Christine, and Bernard Thisse. 2008. "High-Resolution in Situ Hybridization to Whole-Mount Zebrafish Embryos." *Nature Protocols* 3:59–69. doi: 10.1038/nprot.2007.514.

Thomas, Charles A., Leon Tejwani, Cleber A. Trujillo, Priscilla D. Negraes, Roberto H. Herai, Pinar Mesci, Angela Macia, Yanick J. Crow, and Alysson R. Muotri. 2017. "Modeling of TREX1-Dependent Autoimmune Disease Using Human Stem Cells Highlights L1 Accumulation as a Source of Neuroinflammation." *Cell Stem Cell* 21(3):319-331.e8. doi: 10.1016/j.stem.2017.07.009.

Trede, Nikolaus S., David M. Langenau, David Traver, A. Thomas Look, and Leonard I. Zon. 2004. "The Use of Zebrafish to Understand Immunity." *Immunity* 20(4):367–79. doi: 10.1016/S1074-7613(04)00084-6.

Treger, Rebecca S., Scott D. Pope, Yong Kong, Maria Tokuyama, Manabu Taura, and Akiko Iwasaki. 2019. "The Lupus Susceptibility Locus Sgp3 Encodes the Suppressor of Endogenous Retrovirus Expression SNERV." *Immunity* 50(2):334-347.e9. doi: 10.1016/j.immuni.2018.12.022.

Turelli, Priscilla, Nathaly Castro-Diaz, Flavia Marzetta, Adamandia Kapopoulou, Charlène Raclot, Julien Duc, Vannary Tieng, Simon Quenneville, and Didier Trono. 2014. "Interplay of TRIM28 and DNA Methylation in Controlling Human Endogenous Retroelements." *Genome Research* 24(8):1260–70. doi: 10.1101/gr.172833.114.

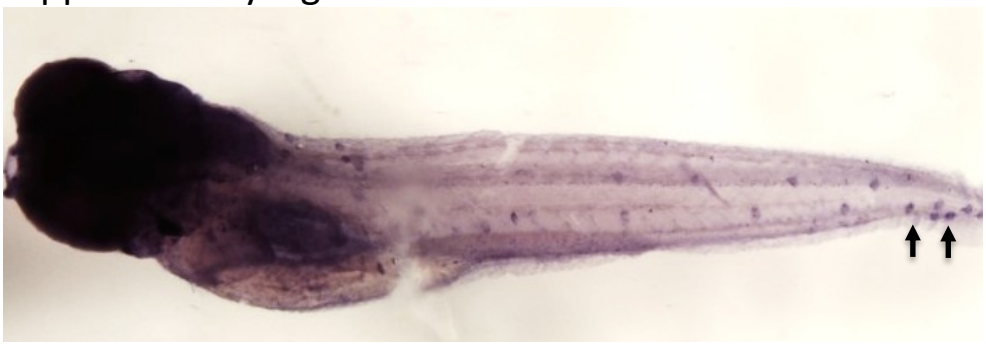
Wasserstein, Ronald L., Allen L. Schirm, and Nicole A. Lazar. 2019. "Moving to a World Beyond 'p < 0.05.'" *American Statistician* 73(sup1):1–19. doi: 10.1080/00031305.2019.1583913.

Young, George R., Mickaël J. Y. Ploquin, Urszula Eksmond, Munisch Wadwa, Jonathan P. Stoye, and George Kassiotis. 2012. "Negative Selection by an Endogenous Retrovirus Promotes a Higher-Avidity CD4+ T Cell Response to Retroviral Infection." *PLoS Pathogens* 8(5). doi: 10.1371/journal.ppat.1002709.

Supplementary Table 1: Coordinates of the *zferv1-7* genomic segments

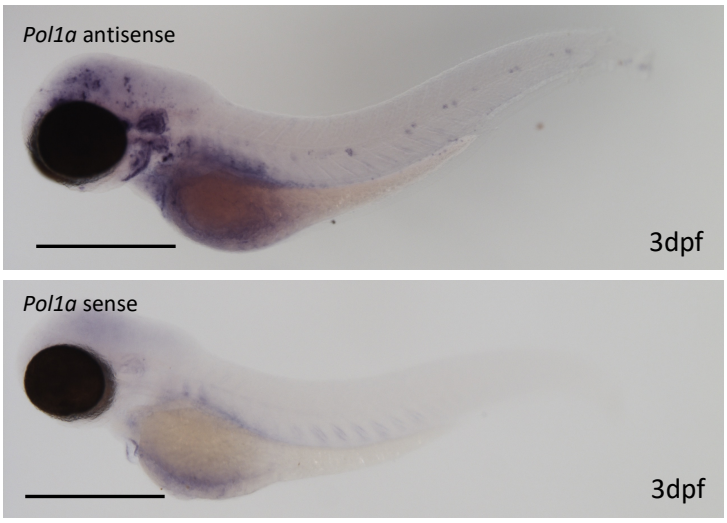
ERV name	chromosome (scaffold)	orientation	5' limit	3' limit
zferv1a	4	reverse	62,546,461	62,557,677
zferv1b	KN149691.1 - NW_001884427.4	forward	211.988	223.087
zferv2	7	reverse	19,379,508	19,389,176
zferv3	16	reverse	9,986,975	9,996,169
zferv4	5 - NW_018394575	reverse	168.525	178.11
zferv5	15	forward	44,144,056	44,162,279
zferv6	7 - NW_018394672	reverse	116.062	124.356
zferv7	19	forward	7,960,461	7,968,734

Supplementary Figure 1



Supplementary Figure 1: Long development of the *env* probe from *zferv1a* by ISH revealed labelling of neuromasts (Black arrows).

Supplementary Figure 2

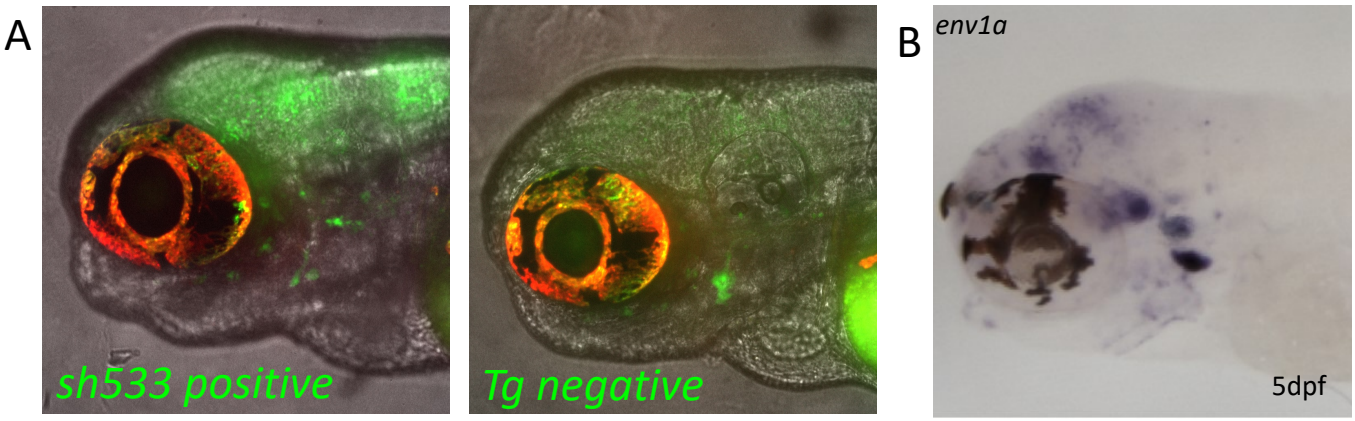


Supplemental Figure 2: Expression of the reverse transcriptase gene (*pol*) of *zferv1a* (called *pol1a* here) by *in situ* hybridisation at 3dpf. Scale bar 500µm.

Supplementary Figure 3

Supplementary Figure 3: Alignment of the envelope gene from zferv1a and zferv1b to highlight to regions with significant differences between the two genes. Primers used to design the probes are underlined in red.

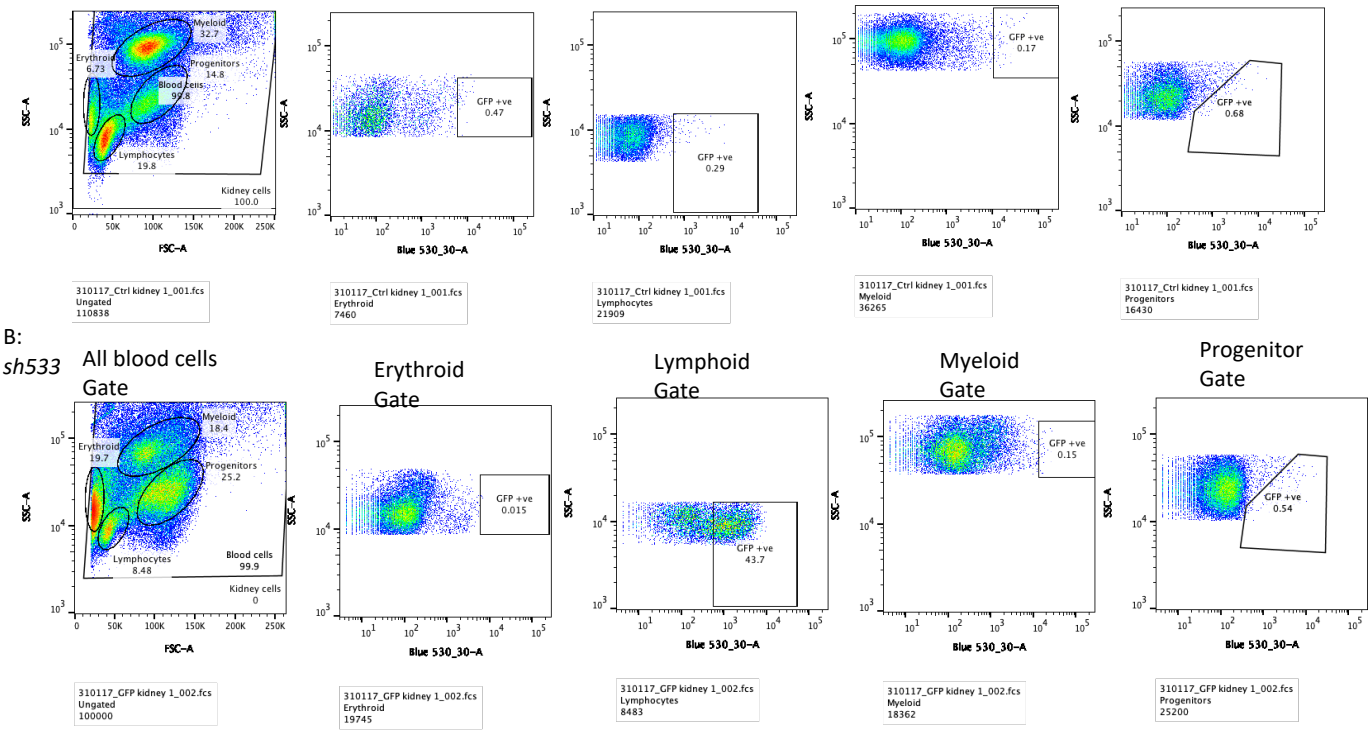
Supplementary Figure 4



Supplementary Figure 4: A. *sh533* positive and negative embryo illuminated with the GFP and mCherry laser to highlight the similar light reflection in the lenses that act as mirrors. **B. ISH with RNA probe env1a in a 5dpf embryos injected with a crRNA against the tyrosinase gene generating a mosaic pattern of**

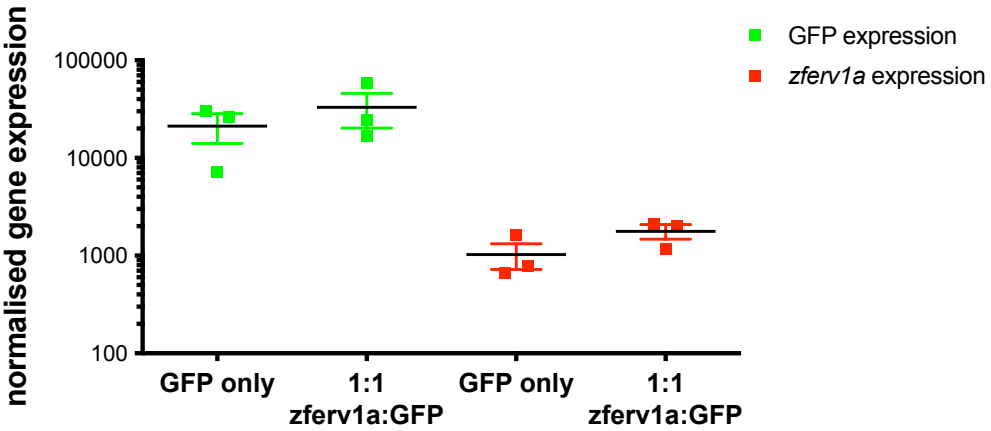
Supplementary Figure 5

A: Non transgenic control



Supplementary Figure 5: Gating strategy for each heamatopoietic lineages (erythroid, lymphoid, myeloid and progenitor) drawn first in non transgenic control (A) and applied to the sh533 transgenic kidney samples.

Supplementary Figure 6



Supplementary Figure 6: Expression of GFP (green) and zferv1a (red) by qPCR after 1-cell stage injection of a GFP expressing vector alone (GFP only) or in equimolar concentration with zferv1a expression vector (1:1 zferv1a:GFP).