Tampere University

MERI UUSI-MÄKELÄ

Zebrafish as a Model for Human Genetic and Infectious Diseases

Tampere University Dissertations 587

Tampere University Dissertations 587

MERI UUSI-MÄKELÄ

Zebrafish as a Model for Human Genetic and Infectious Diseases

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty of Medicine and Health Technology
of Tampere University,
for public discussion in the auditorium F114
of the Arvo building, Arvo Ylpön katu 34, Tampere,
on 8th June 2022, at 12 o'clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Medicine and Health Technology
Finland

Responsible supervisor and Custos	Professor Mika Rämet Tampere University Finland	
Pre-examiners	PhD Stefan Oehlers Agency for Science, Technology and Research Singapore	Docent Ilkka Paatero University of Turku Finland
Opponent	Professor Marko Salmi University of Turku Finland	

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

Copyright ©2022 author

Cover design: Roihu Inc.

ISBN 978-952-03-2367-7 (print)
ISBN 978-952-03-2368-4 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-2368-4

PunaMusta Oy – Yliopistopaino
Joensuu 2022

"Science is not about building a body of known 'facts'. It is a method for asking awkward questions and subjecting them to a reality-check, thus avoiding the human tendency to believe whatever makes us feel good."

-Terry Pratchett

ACKNOWLEDGEMENTS

This doctoral thesis was conducted in the Experimental Immunology research group in the Faculty of Medicine and Health Technology, Tampere University. I would like to thank everyone who helped me and participated in making this thesis. I sincerely thank the foundations; Väinö and Laina Kivi Foundation, Maud Kuistila Memorial Fund, City of Tampere Science Fund, Tampere Tuberculosis Foundation, Finnish Cultural Foundation and Tampere University Doctoral school for their financial support which facilitated my thesis work through the years.

I thank Professor Mika Rämet, for giving me an opportunity to work in his research group on these projects. I am grateful for all the help you could give me. I would like to thank the pre-examiners of the thesis, doctor Stefan Oehlers and docent Ilkka Paatero for their helpful comments on the manuscript of the book. Likewise, I would like to thank Professor Marko Salmi for acting as the opponent in the public thesis defence. In addition, I would like to acknowledge the valuable input from the members of the follow up group, Professors Olli Silvennoinen and Seppo Parkkila and Olli Laitinen. Thank you for your opinions and discussions. I would especially like to thank Professor David Tobin in Duke University (NC, USA), for your kindness, interest towards my work and most importantly showing me that hard core science can be thoroughly enjoyable. The co-authors and collaborators of the studies have my most sincere appreciation; without you, the science would not have happened.

I would like to express my gratitude to the experts who are always available for my never ending questions; the bioinformatician Harlan Barker, Juha Määttä from the Protein Core facility, biostatistician Heini Huhtala, Sari Toivola from histology core and Laura Kummola from FACS core facility. I am not afraid to ask stupid questions when I know they are always met with your kindness, never ending optimism and an attitude for getting things done. I also wish to thank the Zebrafish core facility for providing the help and support with the animal experiments. Through the years I have had the opportunity to supervise and teach a number of students. These students deserve my most sincere appreciation for participating in

this research during their thesis projects and through summer months, just to make my work easier. It is your motivation that helps me find mine every time.

My heartfelt thanks to the current and former group members of Experimental Immunology. The fly people Susanna Valanne, Mirva Järvelä-Stölting, Matthew Maasdorp, Laura Vesala, and Tiina Salminen,; thank you for your help and advice. Thanks to Carina Bäuerlein, Markus Ojanen, Anni Saralahti, Mirja Niskanen, and Sanna Harjula, for setting an example and being there during times when science is hard and nothing seems to work. I also wish to thank the former technicians Leena Mäkinen and Hannaleena Piippo who kept the lab running and my samples in order.

Thanks to my comrade since the first year at the Uni, Niila Saarinen for thoughtful plan B life/career planning sessions and to Tanja Kuusela who I can always rely on for emotional support over our *feminism and sympathy* sessions. Thanks to fellow climbers at TK and the pole dancers at RTP who help me keep my feet off the ground and the monkey off my back.

I have a wonderful family as an endless source of support. You cannot be thanked enough. The deepest thanks go to Joonas and the cats. I could not have done any of this without you.

Lastly I would like to thank (and apologise to) the zebrafish (*Danio rerio*) who were sacrificed for the experiments.

Tampere, February 2022.
Meri Uusi-Mäkelä

ABSTRACT

Due to our shared evolution, humans share parts of their genetics with related animals. We share most of our genes with the chimpanzee (*Pan troglodytes*) and with mouse (*Mus musculus*), but perhaps surprisingly also with zebrafish (*Danio rerio*). Despite its aquatic lifestyle and fishy looks, zebrafish shares most of its genes and several parts of its physiology and micropathology with human. These biological similarities of related species are invaluable for medicine, as they allow us to study our own biology through evolutionary homology in even species which are considered very primitive in phylogenetic terms. The selection of a model species should be based on ethics as well as the question at hand. In some cases, we can use a very distant species, provided that we can show the conservation of the gene or pathway biology between us and the model organism. In this thesis we have used zebrafish as a model for human genetics in four projects with focus on developmental and disease genetics.

In each project, we employed gene silencing or mutagenesis to study zebrafish genes. In our first project, we set out to study the differences observed in the efficiency of the CRISPR-Cas9 (clustered regularly interspaces short palindromic repeats, CRISPR associated 9) mutagenesis system. We observed differences between *in vitro* and *in vivo* efficiencies of mutagenesis between genes in larval zebrafish. This led us to consider if the efficiency of mutagenesis was determined by DNA packaging *in vivo*. Our results did not reveal strong correlation between epigenetic factors and CRISRP-Cas9 efficiency. However, our results indicate mutagenesis functions differently between genes independent of *in vitro* functionality, suggesting an influence from DNA packaging.

In the second project, we generated a zebrafish model for a newly identified human genetic disease, FINCA (fibrosis, neurodegeneration and cerebral angiomatosis). FINCA manifests in early childhood and leads to death of affected children before the age of two and is a result of a mutation in a gene *NHLRC2*. Mutations in mouse *Nhlrc2* lead to early lethality, which significantly hampers its applicability in studying FINCA. Silencing of *nhlrc2* in zebrafish lead to alterations in integrity of brain cells in zebrafish larvae based on the analysis of TEM images in comparison to controls, suggesting it replicates some parts of the phenotype of the

patients. Our results support notion that the loss of NHLRC2 is the cause of the FINCA disease.

With the same gene silencing technology, we analysed the effect of *intelectin 1* gene on *intelectin 3* mutated zebrafish. Intelectins are a family of proteins involved in recognition of bacteria in cells. Primarily, *intelectin 3* was found to be upregulated in mycobacterial infection in our microarray data of adult zebrafish. We failed to see a phenotype in *intelectin 3* mutated zebrafish, which lead us to consider if *intelectin 1*, due to similar expression kinetics, could be compensating for its loss, masking the phenotype of the mutants. However, simultaneous silencing and knockout of *intelectin 1* and *intelectin 3*, respectively, did not result in a phenotype in larval zebrafish in mycobacterial infection.

In the last project we set out to determine the role of the gene *pycard* in mycobacterial infection. *pycard* codes an adaptor protein of the inflammasome complex, which has also been indicated to play a role in adaptive immune response. Results showed that CRISPR-Cas9 generated *pycard* knockout zebrafish present decreased survival in mycobacterial infection. In addition, they present alteration in bacterial burden and granuloma size. RNA-sequencing analysis of the kidney transcriptome of the fish suggest multiple changes in transcription, which makes the knockout fish susceptible to mycobacterial infection. Our model suggests inflammasome signalling has a central role in orchestrating host's defence and neutrophil function during mycobacterial infection.

In these projects we explored the versatile applications of zebrafish in modelling human diseases. With the study of intelectins and *nhlrc2*, gene silencing led to near total elimination of the transcript, even though with intelectins this did not cause detectable changes in the phenotype. With mutagenesis of inflammasome adaptor gene *pycard*, we were able to generate a model suitable for generating a detailed understanding on its function in mycobacterial infection. Our results demonstrate the utility of the zebrafish in modelling human genetics.

TIIVISTELMÄ

Yhteisen evoluutiohistoriamme vuoksi osa ihmisen geeneistä voidaan tunnistaa meille läheisistä sukua olevissa eläimissä. Jaamme suuren osan perimästämme simpanssin (*Pan troglodytes*) ja hiiren (*Mus musculus*), mutta kenties yllättäen myös seeprakalan (*Danio rerio*) kanssa. Huolimatta sen vedenalaiseen elämään sopivista piirteistä, osa seeprakalan geeneistä, fysiologiasta ja mikropatologiasta muistuttavat ihmisen vastaavia ominaisuuksia. Sukulaislajien biologinen vastaavuus mahdollistaa oman biologiamme mallintamisen jopa hyvin alkukantaisina pidetyissä lajeissa, ja on olennainen osa lääketieteellistä tutkimusta. Joissain tapauksissa voimme näet hyödyntää jopa hyvin kaukaista sukulaislajia, kunhan voimme osoittaa tutkimamme geenin tai signalointireitin vastaavuuden lajien välillä. Tässä väitöstutkimuksessa olemme käyttäneet seeprakalaa mallina ihmisen geneettisessä tutkimuksessa neljässä projektissa, jotka keskittyvät kehitysbiologiseen genetiikkaan ja tautigenetiikkaan.

Näissä projekteissa olemme hyödyntäneet geenihiljennystä ja mutageneesiä seeprakalan geenien tutkimisessa. Ensimmäisessä tutkimuksessa pyrimme selvittämään CRISPR-Cas9 (clustered regularly interspaces short palindromic repeats, CRISPR associated 9) mutageneesin tehokkuudessa havaitsemiemme erojen perimmäistä syytä. Havaitsimme tutkimuksessamme eroja mutageneesin *in vivo*- ja *in vitro* -tehokkuudessa geenien välillä seeprakalan poikasessa. Päättelimme, että tämä saattaisi johtua DNA:n erilaisesta pakkausmekanismeista *in vivo*. Emme kuitenkaan havainneet merkittävää korrelaatiota epigeneettisten tekijöiden ja CRISPR-Cas9 mutageneesin tehokkuuden välillä. Tuloksemme kuitenkin osoittivat, että mutageneesin tehokkuudessa *in vivo* ja *in vitro* havaitut erot ovat geenikohtaisia, joten havaitut erot voivat osin johtua DNA:n pakkausmekanismeista.

Toisessa projektissa mallinsimme seeprakalassa ihmisessä hiljattain tunnistettua geneettistä FINCA-sairautta (engl. fibrosis, neurodegeneration and cerebral angiomatosis). FINCA ilmenee varhaislapsuudessa ja johtaa potilaan kuolemaan ennen kahden vuoden ikää. Sairauden syyksi on tunnistettu mutaatio geenissä nimeltä *NHLRC2*. Hiiressä mutaatiot vastaavassa geenissä johtavat elinkyvyttömyyteen varhaisissa kehitysvaiheissa, minkä vuoksi hiiri ei sovellu sairauden mallintamiseen. Vastaavan *nhlrc2* geenin hiljennys seeprakalassa johti kalanpoikasissa muutoksiin keskiaivojen solukossa. Tuloksen perusteella

seeprakalamallissa havaitut muutokset vastaavat osin potilaissa havaittuja muutoksia. Tuloksemme osaltaan vahvistavat, että mutaatiot NHLRC2 geenissä aiheuttavat FINCA-sairauden.

Samaa hiljennysmenetelmää käyttäen tutkimme myös *intelectin 1*-geenin merkitystä kaloissa, joiden *intelectin 3*-geeni oli mutatoitu toimimattomaksi. Intelektiinit ovat proteiineja, jotka osallistuvat bakteerien tunnistukseen elimistössä. Havaitsimme matalalla annoksella infektoiduissa kaloissa *intelectin3*-geenin ilmentymistason olevan merkittävästi koholla. Emme kuitenkaan havainneet muutoksia *intelectin3* poistogeenisissä seeprakaloissa. Päättelimme, että samanlaisen ilmentymiskinetiikkansa vuoksi *intelectin1* voisi kompensoida *intelectin3* puuttumista mutatoiduissa kaloissa ja näin estää mutaation aiheuttamien muutoksen havaitsemisen. *intelectin1*-geenin hiljentäminen mutatoiduissa kaloissa ei kuitenkaan johtanut havaittaviin muutoksiin seeprakalan poikasissa tehdyssä infektiokokeessa.

Viimeisessä projektissa tutkimme *pycard*-geenin merkitystä mykobakteeri-infektiossa. *pycard* koodaa proteiinia, joka toimii osana inflammasomikompleksia, minkä on havaittu osallistuvan hankinnaisen immuunivasteen säätelyyn. CRISPR-Cas9 -menetelmällä tuotetuissa poistogeenisissä kaloissa havaittiin heikennyt vaste selviämiskokeissa. Tämän lisäksi poistogeenisillä kaloilla havaittiin suurempi bakteerimäärä elimistössä ja niissä ilmenneet granuloomat olivat suurempia. RNA-sekvensoinnin avulla osoitimme kalan puolustussolujen geenien ilmentymisessä useita muutoksia, joiden perusteella kaloilla on muuttunut vaste mykobakteeri-infektiolle.

Projekteissa hyödynsimme monipuolisesti seeprakalaa ihmisen tautimallinnuksessa. Intelektiinien ja *nhlrc2*-geenin tutkimuksessa geenihiljennyksellä saatiin aikaan transkriptin lähes täydellinen eliminointi, vaikka intelektiinien kohdalla tämä ei johtanutkaan fenotyypissä havaittaviin muutoksiin. Inflammasomin *pycard*-geenin mutatoinnilla saatiin aikaan malli, jonka avulla voimme hahmottaa yhä tarkemmin geenin merkitystä mykobakteerin aiheuttamassa infektiossa. Tuloksemme ovat osoitus seeprakalan soveltuvuudesta ihmisen genetiikan tutkimuksessa.

CONTENTS

List of Figures

List of Tables

ABBREVIATIONS

ADME	Absorption, distribution, metabolism, and excretion
BCG	Bacillus Calmette Guerin
BLAST	Basic local alignment search tool
CARD	Caspase activation and recruitment domain
Cas9	CRISPR associated 9
CD	Complementarity determining
cDNA	Complementary DNA
CFU	colony forming unit
cGAS	Cyclic GMP-AMP Synthase
CRISPR	Clustered regularly interspaces short palindromic repeats
dpf	days post fertilization
dpi	days post infection
DSB	double stranded break
ESX-1	ESAT-6 secretion system 1
FACS	Fluorescence activated cell sorting
FINCA	Fibrosis, neurodegeneration and angiomatosis
HDR	Homology directed repair
HIV	Human immunodeficiency virus
IFN	Interferon
Ig	Immunoglobulin
IL	Interleukin
itln	Intelectin
M1/M2	Macrophage subtype 1 or 2
MHC	Major histocompatibility
MO	Morpholino
mRNA	messenger RNA
NHEJ	non-homologous end joining
nhlrc	NHL repeat containing
NK-cell	Natural Killer cell
NLRP	NOD like receptor protein

OECD	Organisation for Economic Cooperation and Development
PAM	protospacer adjacent motif
PCR	polymerase chain reaction
PYCARD	PYD and CARD domain containing
PYD	Pyrin domain
qPCR	quantitative polymerase chain reaction
RAG	Recombination activated gene
RD-1	Region of difference
RFLP	Restriction fragment length polymorphism
RNAi	RNA interference
RNA-seq	RNA-sequencing
ROS	Reactive oxygen species
scRNA-seq	single cell RNA-sequencing
sgRNA / gRNA	short guide RNA / guide RNA
STING	Stimulator Of Interferon Response cGAMP Interactor 1
TALEN	Transcription activator-like effector nuclease
Th	T helper
Treg	T regulatory
TSWGD	teleost specific whole genome duplication
VDJ	Variable diversity joining
wpi	weeks post infection
ZFN	Zinc Finger Nuclease

ORIGINAL PUBLICATIONS

Publication I **Uusi-Mäkelä MIE**, Barker HR, Bäuerlein CA, Häkkinen T, Nykter M, Rämet M. Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (Danio rerio). *PLoS One*. 2018 Apr 23;13(4):e0196238. doi: 10.1371/journal.pone.0196238.

Publication II Uusimaa J, Kaarteenaho R, Paakkola T, Tuominen H, Karjalainen MK, Nadaf J, Varilo T, **Uusi-Mäkelä M**, Suo-Palosaari M, Pietilä I, Hiltunen AE, Ruddock L, Alanen H, Biterova E, Miinalainen I, Salminen A, Soininen R, Manninen A, Sormunen R, Kaakinen M, Vuolteenaho R, Herva R, Vieira P, Dunder T, Kokkonen H, Moilanen JS, Rantala H, Nogee LM, Majewski J, Rämet M, Hallman M, Hinttala R. NHLRC2 variants identified in patients with fibrosis, neurodegeneration, and cerebral angiomatosis (FINCA): characterisation of a novel cerebropulmonary disease. Acta Neuropathol. 2018 May;135(5):727-742. doi: 10.1007/s00401-018-1817-z.

Publication III Ojanen MJT, **Uusi-Mäkelä MIE**, Harjula SE, Saralahti AK, Oksanen KE, Kähkönen N, Määttä JAE, Hytönen VP, Pesu M, Rämet M. Intelectin 3 is dispensable for resistance against a mycobacterial infection in zebrafish (Danio rerio). Sci Rep. 2019 Jan 30;9(1):995. doi: 10.1038/s41598-018-37678-1.

Manuscript **Uusi-Mäkelä MIE**, Harjula SE, Sillanpää A, Nätkin R., Niskanen MT, Nykter, M., Rämet M. The inflammasome adaptor *pycard* is essential for immunity against *Mycobacterium marinum* infection in adult zebrafish. (Submitted manuscript).

AUTHOR'S CONTRIBUTIONS

Publication I Meri Uusi-Mäkelä participated in designing the zebrafish and *in silico* experiments. Meri Uusi-Mäkelä performed the experiments on mutagenesis timing and Cas9 *in vitro* cleavage efficiency. Meri Uusi-Mäkelä participated in writing and revision of the manuscript together with other authors.

Publication II Meri Uusi-Mäkelä designed the *nhlrc2* morpholino and verified its function and effect on the phenotype of the zebrafish larvae. Meri Uusi-Mäkelä prepared the larval zebrafish samples for transmission electron microscopy. Meri Uusi-Mäkelä participated in writing and revision of the manuscript together with other authors.

Publication III Meri Uusi-Mäkelä designed the *itln1* morpholino and verified its function and effect on the phenotype of the zebrafish larvae. Meri Uusi-Mäkelä participated in conducting the infection experiments on the morphant larvae, as well as the analysis of the results. Meri Uusi-Mäkelä participated in writing and revision of the manuscript together with other authors.

Manuscript Meri Uusi-Mäkelä designed the *pycard* targeting gRNA and generated the knockout zebrafish lines. Meri Uusi-Mäkelä performed the infection experiments on the larval and adult zebrafish together with others. Meri Uusi-Mäkelä supervised the histological analyses on adult zebrafish. Together with others, Meri Uusi-Mäkelä performed flow cytometry and cell sorting experiments, RNA-seq sample preparation and data analysis. Meri Uusi-Mäkelä participated in writing and revision of the manuscript together with other authors.

1 INTRODUCTION

Due to our shared evolution, humans share much of their physiology and genetics with related animals. We share over 99% of our genes with the chimpanzee (Pääbo 2001) and around 98% with mouse (Mural et al. 2002). Animals which replicate human genetics and physiology, perhaps surprisingly, also include zebrafish (*Danio rerio*). Zebrafish shares 70% of its genes and many levels of physiology and micropathology with human (Howe et al. 2013; Patton, Zon, and Langenau 2021). The similarities shared between related species are invaluable for medicine as they allow us to probe our similar structures in experimental settings, which would not be possible with humans. Thus, at the moment, studying human genetics still relies heavily on animal models. Despite that novel technologies including organ-on-chip and body-on-chip offer novel 3D tissue culture solutions for more human-specific research, their use is still limited by the complexity of tissue organization and interplay (Wu, 2020; Al-lamki 2017). This is because artificial biology systems such as organ on a chip are only as complex as we make them and fail to include parameters which still remain unknown to us. Therefore, research often turns back to its roots, in studying animals.

Zebrafish (*Danio rerio*) has been used for modelling human genetics since the 1980s (Streisinger et al. 1981). The first experiments took advantage of the transparency of the larval zebrafish in finding phenotypic changes in large scale forward genetic mutagenesis screens (Nüsslein-Volhard 2012). Due to the transparency of the zebrafish in the early days of development, fluorescently labelled cells, chemical compounds, and bacteria can be observed in real time in a non-invasive manner. This has led to the emergence of a number of zebrafish strains carrying fluorescent markers specific for a cell type. Today, zebrafish is applied in research where mouse is less applicable, in large scale drug screens, in studies of early development and in researching host-pathogen interactions in mycobacterial infection to model human tuberculosis. In addition, due to its small size and well documented development, zebrafish larvae have been used for single cell transcriptomic experiments, which can decipher developmental trajectories of cells

from early developmental stages (Carmona et al. 2017; Farnsworth, Saunders, and Miller 2020; Wagner et al. 2018).

Besides its role in developmental biology, the larval zebrafish model has been used for studying biomedicine and genetics. In immunology, it has helped reveal many novel genes essential in immunity against tuberculosis (Cronan and Tobin 2014). Furthermore, the adult zebrafish has proved to be useful in modelling reactivation of latent tuberculosis infection, and screening for novel vaccine antigens (Myllymäki et al. 2018; Parikka et al. 2012). With the emergence of sequence specific genome modifying tools, including CRISPR-Cas9, generating mutant zebrafish has become straightforward (Hwang et al. 2013). Reverse genetic studies can help determine the function of individual genes in host resistance against mycobacterial infection. Despite its simplicity, recent research has revealed genetic modification actually results in transcriptome level changes which can compensate for gene loss (El-Brolosy et al. 2019; Rossi et al. 2015).

Our research investigates the applicability of the zebrafish in modelling human genetic and infectious diseases through gene silencing and gene knockout technologies. The review of literature will highlight the similarities and differences of human and zebrafish to evaluate the suitability of zebrafish as a genetic model for human diseases. It will also explore the use of gene silencing and gene modification technologies in relation to modelling human genetic diseases.

2 REVIEW OF THE LITERATURE

2.1 Studying genetics using model organisms

Animal models remain essential for deciphering the basic biology of organisms as well as for modelling human diseases. When trying to determine the function of a single gene, it is often too narrow an approach to look into a single tissue, or a single cell line, or use organ-on-a-chip models. Instead, gaining a broader understanding of biological and physiological phenomena requires studies on the level of the whole organism. This is important especially when examining phenomena which involve the migrating cells of the immune system or totipotent stem cells, which differentiate to ultimately give all the structures of a mature organism. Many animals have been used as *models* for human biology, and each of them has their own advantages and drawbacks for this purpose, as each species is specialised to live in its own preferred niche and thus has developed specialised means to succeed. In research, ethics often dictate the selection of the organism, and the selection must be based on the importance of the question at hand – are you studying a gene that has the potential to aid in overcoming a pandemic, or a gene which is essential in the response to an agent used in cosmetics? Can the organism even answer the question at hand, that is, if I want to study a disease transmitted by aerosols, does this organism have a set of lungs? Is the model relevant – can testing this novel drug in mice (*Mus musculus)* get us any closer to the drug's function in humans? If a patient's life is at risk, how fast can we find a cure?

Primates are the closest living relatives to humans, but their use in research presents ethical problems and thus they are not the preferred choice in initial studies on novel drugs and treatments. Human physiology, genetics and developmental biology resemble those of the mouse, which is much smaller than primates and thus easier – and cheaper - to use. The mouse is also the most studied model to date, and therefore transgenic and mutant lines and tools are readily available, lessening the need to develop completely new methodological solutions. However, even more *primitive* organisms can be used for modelling basic genetics. Primitive, in phylogeny, simply describes the occurrence of traits earlier in evolution – thus, primitive species are more like their ancestors than species which have emerged later. Primitive model

species include the fruit fly (*Drosophila melanogaster*) and the nematode (*Caenorhabditis elegans*), which present genetic similarity, but more divergence in the physiology and organization of their organs and tissues to mammals. Chickens (*Gallus gallus*) and frogs (*Xenopus laevis*) are physiologically similar to humans, but due to their limited use in applied biological research, they lack the genetic tools that have been developed for mouse. When looking into tissue biology and genetic programs controlling development and tissue regeneration, even more distant relatives can be studied. Ascidians, including the tunicate *Ciona intestinalis* are thought to present the simplest shared body plan of all chordates (Liu et al. 2006). Echinoderms, including sea urchins can regenerate most of their organs and even reproduce asexually by shedding limbs (Dupont and Thorndyke 2007), but regeneration can even be modelled in the lophotrochozoan flat worm (Mouton et al. 2018). However, when looking into more distant relatives, most physiological relevance is lost. A phylogenetic relationship of commonly used model organisms is presented in Figure 1.

Human, primates, and mice are all vertebrate tetrapod mammals and share similarities at the level of tissues and organs. The fruit fly and the nematode are both invertebrates, but still share their core cell biology and genetics with all metazoan – multicellular animal - organisms. However, due to the differences in the physiological organization of tissues the fly and the worm offer a limited amount of information on the development of vertebrate organisms.

The zebrafish emerged as a vertebrate model for human biology in 1980s. It is genetically similar to humans, with a shared microanatomy and physiology and several advantageous genetic tools are available for manipulating zebrafish, compared to invertebrates. The zebrafish is not a mammal, amniote or tetrapod, but shares many conserved aspects of developmental biology, physiology and immunology with other vertebrates. However, as no organism presents a spectrum of genetic and physiological disease identical to that of humans, one must be cautious when interpreting the applicability of the results to humans. When considering the suitability of a model organism for an application, it is the researcher's responsibility to consider the advantages and the limitations of the model, to justify the experimental setup and to evaluate the impact of the results accordingly.

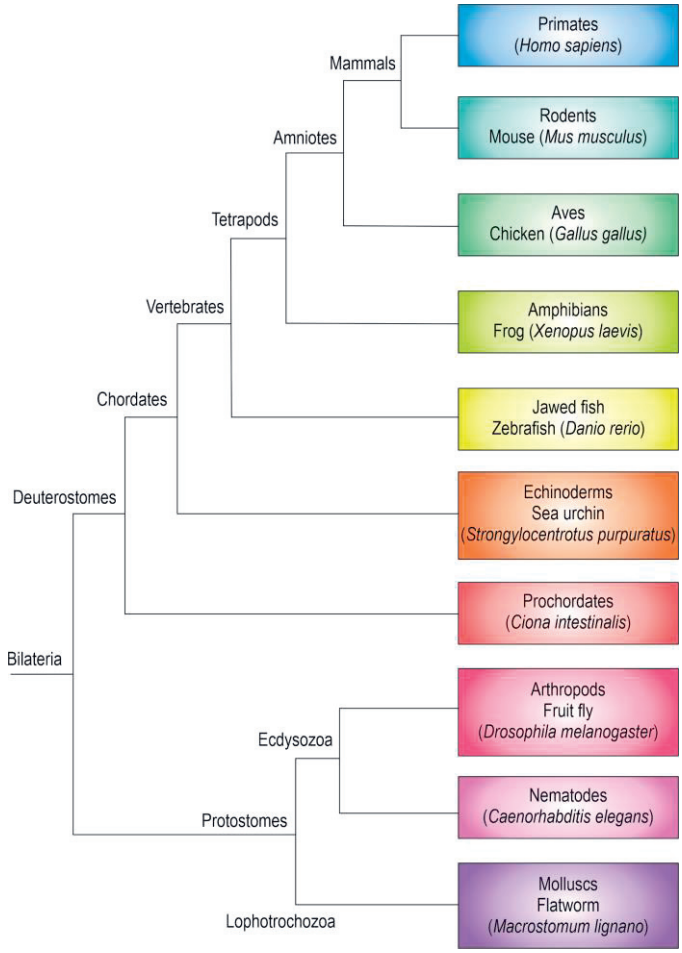


Figure 1. The morphology based phylogenetic relationship of selected model organisms in relation to humans. In coloured boxes, the group of animals is indicated followed by an exemplar representative. According to morphology based phylogenetic division, bilateria – animals with bilateral axis symmetry- are divided as follows: Deuterostomes develop their body plan with an anus forming before a mouth. Chordates include all deuterostomes which have a notochord at any point during development. Vertebrates have a backbone. Tetrapods include all four-legged vertebrates. Amniotes includes animals which develop inside protective extra embryonic membranes, such as an amnion, chorion or allantois. Mammals includes amniotes which feed their young from mammary glands of the mother. Protostomes develop a mouth first, but the developmental order is actually more varied than this. Ecdysozoa have either an insect-like or worm-like body plan. Lophotrochozoans present a spiral cleavage pattern and trochophora larvae with a distinct feeding tube. Compiled from Campbell and Reece (2005).

2.2 Zebrafish biology

2.2.1 Zebrafish as a model organism

Zebrafish are freshwater fish and belong to the cypriniform teleost, and the minnow (Cyprinidae) family. A zebrafish reaches adulthood in 3 months and measures approximately 3 cm in length. An adult female can lay several hundred eggs at a time, which are fertilized and develop *ex utero* – the two main reasons why the zebrafish has become an important model organism for studying vertebrate development and genetics (Driever et al. 1996; Haffter et al. 1996). In addition to an external fertilization and high fecundity, zebrafish larvae are optically translucent, which makes them attractive for studies on developmental biology. Moreover, housing the fish is relatively simple and inexpensive. In addition to basic genetic research, toxicological and pharmacokinetic drug screens have shown that biomedical compounds often display similar pharmacokinetics in zebrafish as in the human and mouse (Cassar et al. 2020; MacRae and Peterson 2015; Patton et al. 2021). Biotechnological tools for the fish are many, and the development of novel methods, such as CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats – CRISPR associated 9) have made the genetic manipulation of model organisms, including zebrafish, easier (Hruscha et al. 2013). While the larval zebrafish is useful in developmental genetics, the adult zebrafish has applications for example in immunology, regenerative biology, haematology and behavioural genetics (González-Rosa, Burns, and Burns 2017; Harjula et al. 2020; Norton and Bally-Cuif 2010; Traver et al. 2003). Zebrafish offers a tool for limiting the number of mammalian model organisms required for research by serving as an intermediate model species between invertebrates and mouse. Its suitability for this purpose will be reviewed in the upcoming chapters.

2.2.2 Ethics of animal experiments

In an ideal world, research would utilize human tissues and organ systems for modelling human genetics and diseases. However, we are still unable to model many phenomena in cell cultures and in organ-on-a-chip models. In order to provide an accurate understanding of the complex molecular, cellular and tissue level events, animal experiments are still sometimes required despite ethical conflicts. Model

organisms overcome the variation caused by diet, lifestyle and the environment, and thus allow the straightforward follow-up of genetic alterations on the phenotype and genotype.

In order to minimise animal suffering, researchers should follow the three Rs of animal experiments: Replace, Reduce, Refine. Accordingly, animal experiments should be replaced with *in vitro* assays whenever possible (replace). Replacing is not always possible, especially in the field of immunology, disease biology, and developmental biology, where interactions between multiple cell types, tissue types and organs, make it difficult to study the phenomena *in vitro*. Therefore, ethical conduct also dictates that the least developed model should be used, whenever possible. In toxicological studies, zebrafish could be used to replace experiments in rodents, provided the genes, tissues and organs relevant for the experimental question, are present in fish (Cassar et al. 2020). The size of the experiment should be limited to a minimal number of animals required for reaching reliable conclusions (reduce). Still, the number of animals required for an experiment should be determined using appropriate power calculations. Using zebrafish as a primary platform for experiments could reduce the number of mammals needed in preclinical testing (Cassar et al. 2020). For refinement, researchers should ensure minimal stress and suffering to the animals both during regular housing and maintenance as well as during experiments; this includes the appropriate use of analgesics and anaesthetics whenever possible to minimize the stress and suffering caused to the animals. In addition, the possibility to analyse drug action *in vivo* using the transparent zebrafish larvae provides new possibilities for improving preclinical trials by non-invasive experiments (Cassar et al. 2020; Geisler et al. 2017).

The three Rs also apply in zebrafish research. As with other non-primate model organisms, the debate on whether zebrafish larvae and adults are able to feel pain is ongoing, but the fish are at least known to respond to an unpleasant stimulus and thus they require analgesia and anaesthesia similarly to other animals (Ohnesorge, Heinl, and Lewejohann 2021; Strähle et al. 2012). Zebrafish embryos younger than 5 days post fertilization (dpf) are often exempted from animal experiments, as they are not able to feed themselves. Embryos which are old enough to hunt for their food (≥5 dpf) are considered to have a nervous system capable of orchestrating complex behaviour and perception of external stimuli, and thus to be developed enough to be classified as experimental animals (Geisler et al. 2017). However, from the point of view of research ethics and the three Rs, the principal aim is to replace all animals whenever possible (Hubrecht and Carter 2019).

2.2.3 Zebrafish biology

Zebrafish are fertilized *ex utero* and remain unpigmented for the first day of development. The embryos are large enough to examine with a basic stereo microscopy equipment, and the first cell division of the fertilized ovum takes place about 45 minutes after fertilization (Kimmel et al. 1995). The development of the zebrafish is well characterized and many aspects of organogenesis are conserved across vertebrate species (Irie and Kuratani 2011). A heartbeat can first be detected at around 24 hours post fertilization (hpf), and the blood flow begins at 25 hpf (Herbomel, Thisse, and Thisse 1999). The primary organ systems are formed, and the larvae begin to hatch from the chorion at 48 hpf (Kimmel et al. 1995). Larval zebrafish start to produce pigmentation at 24 hpf (Kimmel et al. 1995), but with 1-phenyl 2-thiourea, the pigmentation can be delayed in order to preserve full translucency, however with potential adverse effects on several biochemical processes (Bohnsack, Gallina, and Kahana 2011; X.-K. Chen et al. 2021; Elsalini and Rohr 2003). In experiments with zebrafish larvae, the larvae are typically followed until the self-feeding stage, 5 to 7 dpf depending on local animal ethics regulations. Zebrafish reach adulthood and reproductive age at approximately 10 - 12 weeks, when they also present a fully developed adaptive immune system (Lam et al. 2004; Westerfield 2000). Experiments requiring a fully developed immune system are typically conducted after 3 - 4 months of age, when the fish have achieved a roughly uniform size. However, zebrafish present a number of variations in body size measurements, and a single measurement offers limited information on developmental stage (Parichy et al. 2009). In future, including a more specific criteria could thus be used to improve experiments.

2.2.4 Zebrafish physiology

Zebrafish organs resemble the organs of mammals at both the functional and histological level but are often simpler in their structural organization (Lieschke and Currie 2007). The main differences between mammals and zebrafish are seen in the development and reproduction, sex determination, breathing (gills vs. lungs), external barriers (mucosal tissues, scales and skin) and the digestive system. In comparison with humans, the zebrafish intestine is structurally simpler. The zebrafish digestive system consists of a single tube without sphincters, and the fish do not have a stomach with gastric glands or an acidic pH (Wang et al. 2010). Instead, fish have an enlarged section called the intestinal bulb which functions as a stomach

or a storage for food prior to digestion. At the tissue level, the gastrointestinal tract is formed of a villous epithelium with enterocytes, goblet cells and endocrine cells like in humans, but does not have intestinal crypts, Peyer's patches, Paneth cells or a submucosal layer which connects muscle to mucosa (Cassar et al. 2020; Wallace et al. 2005). When segmented according to gene expression profiles, a small and a large intestine can be distinguished from the intestine (Wang et al. 2010). With relation to intestinal function, studies have interrogated the zebrafish gut microbiota and its colonization and detailed the conserved microbial colonization and its immunological significance (Murdoch and Rawls 2019). The pancreas functions similarly in zebrafish and in humans, secreting digestive enzymes to the gut from the exocrine pancreas and blood sugar maintaining hormones from the endocrine pancreas, however these functions are separated in ventral and dorsal buds (Tehrani & Lin, 2011). Similarly to pancreatic function, the endocrine system is conserved despite several structural differences (absence of a hypothalamic-pituitary portal system in fish) and has been reviewed in detail by others (Löhr and Hammerschmidt 2011).

The zebrafish liver is functionally similar to the human liver except that it is not structurally organized in bilayer plates or lobules but tubules – still, all functional cell types are present (Cassar et al. 2020; Katoch and Patial 2021). The cellular organization of the liver is simpler, and the tissue can regenerate (Katoch and Patial 2021). Many of the genes responsible for liver development and disease are conserved in human and zebrafish but differences in tissue organization place constraints on how far the physiology is conserved and applicable to humans (Wilkins and Pack 2013).

The zebrafish kidney functions similarly to the mammalian kidney, except that it is less complex and also acts as the main haematopoietic organ. The larval zebrafish pronephros is a very basic kidney, which develops into the mesonephros in adulthood and despite lesser complexity, is similar to the mammalian kidney based on its cellular composition and function (Outtandy et al. 2019; Swanhart et al. 2011). The zebrafish kidney functions to maintain osmoregulation, but processes such as glomerular filtration and renal tubular clearance are conserved between fish and mammals (Outtandy et al. 2019).

The zebrafish heart has one atrium and one ventricle with a conserved histology and structure, and like the other organs, it is simpler compared to its mammalian counterpart (Beffagna 2019). Like the other organs, the zebrafish heart can fully regenerate, even after a 20% amputation (Beffagna 2019).

The zebrafish recapitulates the main functions and the organization of the human central nervous system, including brain subdivisions and key neurotransmitters (dopamine, GABA, glutamate, noradrenaline, serotonin, histamine and acetylcholine) (Cassar et al. 2020; Kalueff, Stewart, and Gerlai 2014). The anatomy of the zebrafish eyes and ears also resemble those in humans, except for the absence of fovea, a higher rod to cone ratio and the ability to detect UV light, and in the case of the ear, the absence of a cochlea (Richardson et al. 2017; Whitfield 2002).

In relation to behaviour, responses to stimuli, social interactions, learning, memory and hunting are also displayed by fish (Cassar et al. 2020; Kalueff et al. 2014). Finally, a circadian clock function is conserved in fish, and the diurnal circadian rhythm is more similar to humans than nocturnal rodents (Vatine et al. 2011).

In conclusion, many zebrafish and human tissues are fundamentally similar. This has enabled the use of zebrafish for studies on congenital heart diseases and cardiomyopathies (Bakkers 2011), kidney and glomerular diseases (Outtandy et al. 2019), liver (Katoch and Patial 2021; Pham, Zhang, and Yin 2017), and haematopoietic diseases (Gore et al. 2018) and primary immunodeficiencies (e.g. Iwanami 2014; Rissone and Burgess 2018). The greatest difference between zebrafish and humans is perhaps that the zebrafish tissues can regenerate in both larvae and adult zebrafish. Natural regeneration is largely absent in humans, but inducing it is a major area of interest for threating many human diseases (Beffagna 2019).

2.2.5 Zebrafish genetics

Zebrafish originate from South-East Asia and domesticated lineages have been bred in laboratories since the 1980s. The typically used wild type (WT) zebrafish lineages include AB and Tüpfel long fin (TL). AB is the original laboratory line crossed from two shop bought fish lines A and B in the Streisinger lab in Albany, Oregon (Crim and Lawrence 2021). TL are homozygous for the *leo*^{t1} and *lof*^{dt2}alleles, which cause spotted skin and long fins, respectively. In the wild, there are many different species of the *Danio* genus, and they are able to form hybrids, indicating recent speciation event (Postlethwait et al. 2000). In comparison with mice, there is a lack of genetic congruency between laboratory zebrafish (Crim and Lawrence 2021). Zebrafish maintained in laboratories are known to harbour a level of genetic diversity similar to populations in the wild (Balik-Meisner et al. 2018). The diversity stems from

standard husbandry practices with zebrafish; in contrast to many other model organisms, zebrafish are regularly outcrossed to maintain a high level of population diversity as continuously incrossed zebrafish lose their fecundity and viability (Balik-Meisner et al. 2018). As a result, zebrafish populations are estimated to present a higher number of variation, measured in single nucleotide polymorphism per kilobase (kb) of unique sequence, than human populations (Butler et al. 2015). This genetic diversity has its pros and cons; especially in toxicological studies, genetic diversity comparable to that in human is essential for discerning the actual effects of toxic exposure (Balik-Meisner et al. 2018). However, compared to isogenic mouse lines with a recognized genetic background, zebrafish publications rarely state the detailed background of the wild type strains, which can lead to an inability to replicate results in different laboratories (Crim and Lawrence 2021; Suurväli et al. 2020). Therefore, it is important to note that zebrafish strains are not inbred strains as is the case with mice; the standard practice with mice requires that the mice are inbred for 20 generations. This divergence in the standard practice should be considered when comparing results obtained from mice and zebrafish (Nasiadka and Clark 2012).

2.2.5.1 Zebrafish genome

When the zebrafish genome was published in 2013, it became the largest vertebrate genome sequenced so far. The zebrafish genome contains more than 26 000 protein coding genes in over 1.412 Gb of DNA (Howe et al. 2013). Zebrafish have 25 chromosomes (Howe et al. 2013) with no apparent sex determining chromosomes. Humans have 23 pairs of chromosomes with 22 pairs of autosomes, with the whole genome covering approximately 20 000 protein coding genes (Glasauer and Neuhauss 2014). The genome of a zebrafish has one or more orthologs for about 70% of the human genes, and *vice versa* (Howe et al. 2013). For disease associated genes, this number is slightly greater (82%). The result of the teleost specific genome duplication is that many human genes have multiple homologues, and only 47% have a single homologue. The abundance of the genes in the zebrafish compared to other vertebrates is a result of a whole genome duplication event.

2.2.5.2 Zebrafish and the teleost specific genome duplication

Zebrafish belong to the order of cypriniforms, infraclass of teleost fishes under ray-finned fishes (Actinopterygii). A characteristic of the Teleostei is a common, teleost specific whole genome duplication (TSWGD) event, which occurred early at its emergence 350 to 320 million years ago (Glasauer and Neuhauss 2014). As duplication events generate a vast amount of raw material for genetic diversification, they have likely contributed to the incredible species diversity of the teleost fish, which account for approximately half of the 64 000 vertebrate species (Glasauer and Neuhauss 2014). The TSWGD creates challenges for using the zebrafish as a model for human diseases, as it generated a number of *ohnologues*, which are paralogues resulting from a genome duplication (according to the first person to suggest it (Ohno, 1970)). Ohnologues are subject to a different kind of selective pressure than paralogs, and have undergone either non-functionalization (inactivation), sub-functionalisation (specialization) or neofunctionalization (acquire novel functions). Non-functionalization is the most likely outcome as mutations are more likely to be deleterious, generating negative selection pressure, since there is no need to have more than one functional gene copy (Glasauer and Neuhauss 2014). Still, it has been estimated that for an unknown reason the zebrafish retained more than 20% of the duplicated genes resulting from TSWGD (Postlethwait et al. 2000). The retainment of the gene copies most likely resulted from a requirement to adapt to new surroundings, and the TSWGD facilitated adaptation by generating a number of gene copies able to take up new functions and create divergence. Unfortunately, the TSWGD generated complexity for genome analyses, as ohnologues and paralogues generate similar but different variants which complicate deducing homology between species. Differentiating between the two is important for assigning correct nomenclature and in assigning homology (Gasanov et al. 2021).

Finding homologous genes in two organisms requires thorough analysis. Similarities in proteins or amino acids offer a robust way to determine similarity, but sequence similarity does not always confer homology. In the case of zebrafish, many studies have aimed at showing how similar it is with other organisms. Research, however, cannot neglect the limitations of the model and simply assume similarity based on limited evidence. Differentiating paralogs from ohnologues, and differentiating homologs from diverged copies requires complex analyses, and can often be done only with the help of detailed transcriptomic and synteny analyses (Parey et al. 2020). The emergence of single cell level transcriptomic studies has

facilitated this greatly, which partly overcomes the difficulties in discerning genes at the protein level.

One drawback of zebrafish as a model is the limited availability of specific antibodies, which is a direct result of the frequency of gene duplications generated by the TSWGD as well as the lack of commercial incentive. Antibodies from mouse often cross-react and can be used for genes where only a single copy is present. In case of two very similar proteins, it is often impossible to tell them apart using polyclonal antibodies. Detailed expression specific information can be used to analyse if both (or all) copies of the protein are actually expressed – and thus next generation sequencing tools are substituting, in part, the need for antibodies in zebrafish studies. Transcriptomic analysis will be further covered later in this review of literature.

Synteny conservation refers to the arrangement of genetic loci in clusters with shared genes and gene organization and a shared transcriptional orientation within the cluster. Recently diverged organisms often share a common organization of genes in sets of chromosomes, which enables position-based discovery of orthologous genes despite divergence in sequence homology (Gasanov et al. 2021). A strong syntenic relationship between genetic loci can indicate positive selection of allelic combinations or shared regulatory regions of genes. For example, identification of the pattern recognition receptors like NLRPs (NOD like receptor protein) NLRP3 and NLRP1 proved challenging in zebrafish, and were thought to be absent (Laing et al. 2008). Recently, perhaps due to updates in reference genomes, Li et al (2019) published articles on both genes, where they present (among structural and functional evidence) positional synteny based identification of the zebrafish Nlrp3 (J.-Y. Li et al. 2018, 2019). In a similar study, 20 of the solute carrier 22 family proteins were analysed by synteny and gene expression profiles, which helped in assigning and orthologous relationship with their human counterparts (Mihaljevic et al. 2016).

2.3 Immune system

The immune system protects us from a myriad of pathogens. The immune system is considered to have evolved early and all invertebrates and vertebrates share relatively similar core immune functions composed of cells capable of surveillance, detection, migration, phagocytosis, and elimination. In addition to this, vertebrates possess an adaptive immunity capable of eliciting an immunological memory. The classical view

of the immune system supported a dichotomous division into innate and adaptive immunities – the first presenting in *lower* invertebrate organisms and the second only appearing in vertebrate species. This was supported by an evolutionary view of the presence of memory possessing, antigen specific T- and B-cells only in vertebrates. In turn, invertebrate species, such as the fruit fly, were thought to present an unspecific recognition system sensing a number of common epitopes by a fixed number of receptors. The emergence of NKT-cells (Natural Killer T-cells) and γδT-cells has blurred the border between the two arms of immunity, as they have characteristics of the adaptive immunity as well as essential innate immune functions (Figure 2). Regarding immune memory, it is currently recognized that the innate immunity has a memory, which is unspecific but results in an enhanced response resulting from previous infectious stimulus (Boraschi and Italiani 2018). Interestingly, this enhanced response can be passed on to future generations, unlike adaptive immune response (Boraschi and Italiani 2018).

The innate immune system is the first responder to a pathogen. It recognizes common pathogenic structures such as foreign DNA and bacterial cell wall structures and is present in all animalia. Recognition is achieved through receptors present on the cell surface and in the cytoplasm. Due to this, innate immune response is also fast acting. Cells of the innate immune system include mast cells, macrophages, neutrophils, dendritic cells, basophils, eosinophils, and natural killer cells (NK-cells). Macrophages, neutrophils and dendritic cells are the main phagocytes. Neutrophils, mast cells, basophils and eosinophils are granulocytes which are triggered to release cytotoxic granules which kill infected target cells. NK-cells recognize pathological cells by the low levels or the lack of major histocompatibility complex I (MHC-I) expression on their surface and trigger cytolytic activity in target cells by releasing cytotoxic granules from their cytoplasm.

Regarded to reside in the midst of innate and adaptive responses are γδT-cells and NKT-Cells. γδT-cells are lymphocytes, which do not require antigen processing and MHC-presentation of peptide epitopes on their surface for activation, but recognize lipid antigens. NKT-cells recognize CD-1, which is similarly associated with the binding of bacterial lipids and glycolipids.

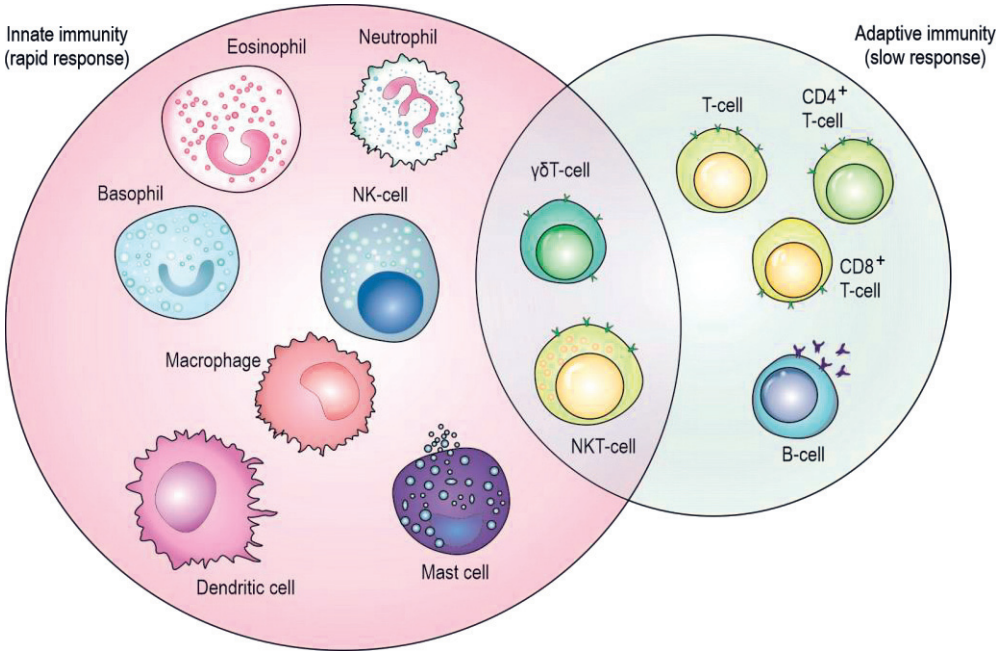


Figure 2. The main cell types of innate and adaptive immunities in vertebrates. Of the mentioned components, basophils have not been described in zebrafish. Adapted from Dranoff (2004).

The adaptive immunity is based on tolerance – the recognition of self and the ability to distinguish self from foreign components - and the generation of an incredible number of antigen-recognizing receptors. The generation of diversity is based on the recombination activating gene (RAG) transposable element which facilitates the variable-diversity-joining (V(D)J)-recombination. V(D)J recombination leads to a vast number of T-cell receptors and immense antibody diversity, mediated by T- and B-cells. The mammalian adaptive immune system consists of lymphocytic antigen receptors diversified by somatic rearrangements of variable elements. This adaptive immune system originated in jawed fish around 500 million years ago, and is found in all vertebrates, except for jawless fish, which possess different, variable lymphocyte receptors based on leucine-rich repeat structures (Flajnik and Kasahara 2010). The emergence of the adaptive immune system is a result of two evolutionary events: the appearance of the RAG-transposon which is essential for V(D)J-recombination, as well as the following genome duplication events which generated a vast amount of receptor diversity (Flajnik and Kasahara 2010). These events lead to the development of the B-cell receptor, the T-cell receptor and the MHC which are the essential components of the vertebrate adaptive immune system (Flajnik and

Kasahara 2010). Zebrafish and human have both innate and adaptive immune systems, however with several notable differences, which will next be explored. Since the cellular and molecular mechanisms have been conserved in evolution, zebrafish can be used to explore vertebrate haematopoiesis and immunogenetics.

2.3.1 Innate immunity in zebrafish

Innate immune cells are the first to emerge during the development of a zebrafish, with the first macrophages appearing around 1 dpf and the first neutrophils at 2 dpf (Meijer and Spaink 2011). The first emerging macrophages are functional and capable of eradicating an infection (Herbomel et al. 1999). Furthermore, the transcriptomes of neutrophils and macrophages have been characterized, and both M1 and M2 macrophages have been detected based on their transcriptomic signature (Kenyon et al. 2017; Rougeot et al. 2019). Neutrophils in both humans and zebrafish are the first responders to a damage or infection, as they are the most abundant leukocyte in the circulatory system (Harvie and Huttenlocher 2015). The role of eosinophils in immunity is not established as well as for other granulocytes. A population of granulocytic leukocytes resembling the eosinophils of mammals is found in zebrafish, and they express transcripts known to be characteristic of eosinophils (Balla et al. 2010). It is not known if basophils are present in zebrafish, but other innate immune cell types are, and their respective marker genes and sorting parameters are listed in Table 1.

2.3.1.1 Pattern recognition receptors

As zebrafish are fully aquatic organisms, they need a vastly different defence in the tissues exposed to their surroundings than land-dwelling mammalian species like humans and mice. Therefore, receptors associated with sensing invading pathogens at mucosal surfaces of skin and gills have evolved to fight fish specific pathogens, and the TSWGD has added to their diversity. Despite its divergence, the innate immune system of zebrafish contains the key pattern recognition receptors, such as Toll Like Receptors (H. Chen et al. 2021; Palti 2011), NLRs, C-type lectin like receptors, Retinoic acid inducible gene I (RIG-I-like) receptors and Cyclic GMP-AMP Synthase – Stimulator Of Interferon Response cGAMP Interactor 1 (cGAS-STING) (Ge et al. 2015; Li et al. 2017; Liu et al. 2020).

Table 1. Zebrafish blood cells, their fluorescence activated cell sorting (FACS) identification parameters and marker genes.

Cell/Process	FACS	Marker gene	Reference
Definitive haematopoiesis	-	*runx1*	(Kalev-Zylinska et al. 2002)
Erythrocyte	FSC^{low} SSC^{high}	*gata1*	(Brownlie et al. 2003; Detrich et al. 1995; Moore et al. 2016)
Thrombocyte	FSC^{low} SSC^{low}	*cd41*	(Jagadeeswaran et al. 1999; Lin et al. 2005)
Innate immunity			
Macrophage	FSC^{high} SSC^{high}	*mpeg1, mfap4*	(Ellett et al. 2011; Herbomel et al. 1999; Nguyen-Chi et al. 2015; Walton et al. 2015)
Dendritic cell	FSC^{high} SSC^{high}	*il12, iclp1 (MHCII), csfr1, mpx*	(Lugo-Villarino et al. 2010; Shao et al. 2015)
Neutrophil	FSC^{high} SSC^{high}	*mpx, lyz*	(Hall et al. 2007; Henry et al. 2013; Lieschke et al. 2001; Renshaw et al. 2006)
Eosinophil	FSC^{high} SSC^{high}	*gata2*	(Balla et al. 2010; Lieschke et al. 2001)
Mast cell	-	*cpa5*	(Dobson et al. 2008)
NK-cell	FSC^{low} SSC^{low}	*lck, rag1(-/-)*	(Carmona et al. 2017; Moore et al. 2016; Tang et al. 2017)
ILCs	FSC^{low} SSC^{low} (gut)	multiple markers	(Hernández et al. 2018)
Adaptive immunity			
T-Cell	FSC^{low} SSC^{low}	*lck*	(Moore et al. 2016; Willett et al. 2001; Willett, Cherry, and Steiner 1997)
B-cell	FSC^{low} SSC^{low}	*rag2, IgM*	(Danilova and Steiner 2002; Moore et al. 2016; Page et al. 2013)

HSC=Haematopoietic stem cells, NK-cell=natural killer cell, ILC=Innate lymphoid like cell

In addition to the abovementioned pattern recognition receptors, the zebrafish complement system has been described in detail. The complement system of the zebrafish contains the components of the classical, alternate and lectin pathway (Boshra, Li, and Sunyer 2006; Zhang and Cui 2014). However, zebrafish like other fish species have many copies of the c3 complement component, whereas humans have only a single copy, perhaps indicating an increased complement reactivity or functional divergence in zebrafish compared to humans (Najafpour et al. 2020; Zhang and Cui 2014). The zebrafish complement is especially important in maternal immunity, as many of the complement components are transferred to the zebrafish larvae by the mother fish in the form of mRNA or protein (Yang et al. 2014). The presence of maternal protection arsenal, including complement components, indicates an improved humoral protection of the embryo, which is perhaps necessary for the development outside of the mother's body (Yang et al. 2014). One of the mRNAs transferred to the embryos is *itln3* (Harvey et al. 2013). Intelectins are a scarcely known family of proteins, whose function will be reviewed in a later chapter.

NLR-receptor proteins are one of the most ancient pattern recognition domains, and have been suggested to have first originated already in archaea as nucleotide binding and leucine rich repeat domains (Meunier and Broz 2017). In multicellular organisms, the domains have fused and form the NLR-receptor protein family, the most diverse family of pattern recognition receptors known (Meunier and Broz 2017). Mammals have four subfamilies of NLRs, including acidic transactivation domain containing NLRAs, baculoviral inhibitor of apoptosis repeat containing NLRBs, caspase activation and recruitment domain (CARD) containing NLRCs and pyrin domain (PYD) containing NLRPs. PYDs and CARDs interact homotypically through their death domain fold (Meunier and Broz 2017). The functionality of the NLRs is conserved in zebrafish, which has over 200 NLRs, which include a unique Fisna-domain. Some NLRs activate inflammasomes, which are essential for bacterial clearance and amplification of the immune response. Inflammasome activation, has also been characterized in zebrafish (Forn-Cuní, Meijer, and Varela 2019; Kuri et al. 2017; Y. Li et al. 2018). In addition, the inflammasome activation associated NLRP1 and NLRP3 have recently been described in zebrafish, increasing the complementarity of the inflammasome activation between mammals and zebrafish (Forn-Cuní et al. 2019; J.-Y. Li et al. 2018, 2019).

2.3.2 Adaptive immunity in zebrafish

The adaptive immunity in zebrafish has been characterized for its main cell types, T- and B-cells, and it has been shown that they possess an adaptive immune memory like human (Lugo-Villarino et al. 2010). Zebrafish larvae do not have an adaptive immunity as functional T-cells are detected only at 4 - 6 weeks post fertilization whereas B-cells are detected at 20 dpf (Lam et al. 2004; Page et al. 2013). The diversity of adaptive immune recognition is mediated by RAG recombination in the zebrafish (Willett et al. 1997). As for now, it is not known where antigen presentation takes place, as zebrafish lack lymph nodes, which in mammals are the main tissue for antigen presentation and lymphocyte maturation (Renshaw and Trede 2012). On the other hand, the development of the thymus, where T-cells home to mature and undergo thymic selection to maintain tolerance is conserved between zebrafish and humans (Bajoghli et al. 2019).

B-cells are essential for the humoral adaptive immune response and immune memory. Zebrafish have been shown to express CD79a and CD79b, the B-cell receptor subunits, in the adult zebrafish kidney and spleen (Liu 2017). In zebrafish,

immunoglobulin diversity is generated through somatic rearrangements, somatic mutations and gene conversion (Page et al. 2013). However, B-cells do not undergo somatic hypermutations to the same extent as the mouse and human (Marianes and Zimmerman 2011). Zebrafish have only three immunoglobulin (Ig) classes, IgM, IgD and IgZ, the last of which is specific only to some teleosts. Zebrafish lack IgG, IgA and IgE class Igs (Fillatreau et al. 2013). Zebrafish B-cells were shown to mature without the pre-B-cell ($Rag^{hi}CD79^{+}IgH\text{-}\mu^{+}$) stage which has been described in humans and mice (Liu et al. 2017).

The zebrafish T-cell specific transcription factors (*tbet*, *stat6* and *foxp3*) have all been identified in fish (Mitra et al. 2010). T-cells have also been shown to be activated by dendritic cells in an antigen specific manner (Lugo-Villarino et al. 2010). Their gene expression has been characterized on the single cell level (Moore et al. 2016), confirming the presence of *cd4* (T-helper) and *cd8* (killer T cell) expressing sub-populations among others. *cd4*$^{+}$ T-cells present a typical cytokine and transcription factor profile upon antigen stimulation (Yoon et al. 2015). The transgenic (Tg) *Tg*(*cd4-1:mCherry*) zebrafish have been shown to harbour tissue resident Th2 and Treg-like T-cell sub-populations (Dee et al. 2016). Zebrafish T-regs promote immunological tolerance and present a typical T-reg cytokine profile, including *foxp3* (Kikuchi 2020; Quintana et al. 2010). When transfected into mice, the zebrafish *foxp3* is able to suppress cell proliferation and cytokine expression upon immune activation (Quintana et al. 2010). Zebrafish γδT-cells present typical gene expression profiles, are capable of phagocytosis and antigen presentation and are critical for the antigen specific *IgZ* antibody production in the intestinal mucosa (Wan et al. 2017). Moreover, the zebrafish has also been successfully used for modelling T-cell derived diseases such as T-cell acute lymphoblastic leukaemia and severe combined immunodeficiency (Baeten and de Jong 2018; Punwani et al. 2016). However, functional studies on T- and B-cells in zebrafish are still few in numbers.

2.3.3 Zebrafish model for pathogen infections

Infection is a battle between the invading pathogen and its host. Many pathogens have been infecting their hosts for thousands of years and so the pathogen and the host immune system have co-evolved. Thus, model organisms, which harbour the same immune system as humans can be used to study the same or related pathogens that infect humans. Evolution of pathogens leads into them developing new ways to evade the immune system, and the evolution of the host immune system, in turn,

finds new ways to eliminate the pathogen. In case of contagious diseases, a lethal pathogen kills the host, which often complicates the spread of the pathogen to a new host. Therefore, not all pathogens evolve into more lethal forms, but rather seek to take advantage of their host for as long as possible, as is the case with mycobacterial species which cause tuberculosis, as will be explained in the upcoming chapters.

Due to its similarity to human, the zebrafish offers a valuable model to study the human immune defence against pathogens. The zebrafish has been used as a model for many bacterial infections due to the convenience of the transparent larvae, and the availability of fluorescent markers. When combined, these features enable non-invasive high-resolution time-lapse experiments to probe the interaction of the bacteria and its host (Tobin, May, and Wheeler 2012). In comparison with zebrafish, both *in vitro* models for phagocytes and mice models have several drawbacks which can be overcome with zebrafish. With the *in vitro* models, the phagocytes need to be extracted from their natural surroundings, which prevents dissemination studies (Tobin et al. 2012). *In vitro* growth media does not often mimic the *in vivo* communication between cells through contact induced activation by other cells, through cytokines and chemo-attractants or through contacts with the extracellular matrix (Tobin et al. 2012). With the mouse simultaneous analysis of the whole animal is not possible, and without invasive methods, imaging is limited to the skin (Tobin et al. 2012).

Zebrafish has been successfully used to study many pathogens, including *Mycobacterium marinum*, *Mycobacterium leprae*, *Listeria monocytogenes*, *Staphylococcus aureus*, *Burkholderia enocepacia*, *Salmonella typhimurium*, *Shigella flexneri* and *Streptococcus pneumoniae* species – these have been reviewed e.g. in Yoshida et al. (2017) and Meijer and Spaink (2011). In addition to bacteria, zebrafish can also be infected by fungal species (e.g. *Aspergillus fumigatus* and *Candida albicans*) and viruses (e.g. spring viremia carp virus, rhabdovirus and viral haemorrhagic septicemia virus) as reviewed by Rosowski et al. (2018) and Varela et al. (2017). A limitation to the use of zebrafish is the rearing temperature, which limits the utility of the model at temperatures close to 28°C (22-33°C) (Tobin et al. 2012). However, some pathogens have aquatic relatives which have molecular biology similar to their cousins, as is the case with mycobacteria.

2.4 Tuberculosis and the zebrafish *M. marinum* infection model

2.4.1 Tuberculosis

Tuberculosis is a bacterial infection caused by an acid-fast bacterium, *Mycobacterium tuberculosis*. Tuberculosis causes roughly 1.5 million deaths yearly and was the leading cause of death from a single infectious agent prior to Severe acute respiratory syndrome corona virus 2019 (SARS-CoV-19) (World Health Organization 2020). It is especially prevalent in areas with a high HIV (human immunodeficiency virus) incidence. Over 10 million new cases are diagnosed yearly, and an increasing number of strains are multi drug resistant (World Health Organization 2020). The only available tuberculosis vaccine, the bacillus Calmette-Guérin (BCG) does not provide a lasting protection besides against the most severe forms of the disease in children (Andersen and Doherty 2005). The treatment of tuberculosis requires a prolonged regimen of several antibiotics, which has led to the development of multiple drug resistant strains of tuberculosis (World Health Organization 2020). The eradication of tuberculosis is further complicated by the fact that the disease most commonly develops into a latent form, without symptoms, and can lay dormant for decades with about a 10% risk of reactivation for each infected person (World Health Organization 2020).

The hallmark symptom of tuberculosis is the presence of bacteria-harbouring granulomas in the infected tissue – most commonly in the pulmonary area. As the bacteria are contracted through the airways, alveolar macrophages engulf the bacteria and travel deeper into the tissue for antigen presentation. Mycobacteria are able to survive inside the phagosome of the macrophages by manipulating host immune pathways (Chai et al. 2020). Infected macrophages become shielded by other innate and adaptive immune cells, with epithelioid macrophages surrounding the core of this granuloma structure, which contains the live bacteria (Pagán and Ramakrishnan 2015). The bacteria can reside inside the granuloma without causing symptoms, waiting for the opportune moment, until the host's immune system is compromised and the bacteria are able to thrive.

2.4.2 Animal models for tuberculosis

Developing novel vaccines and treatments for tuberculosis necessitates the use of animal models. As the immune system involves multiple cell types residing and

circulating between multiple tissues, organ model systems are not yet sufficient for modelling interaction with pathogens. Suitable animal models can recapitulate the full spectrum of the disease, however with obvious ethical problems. Although non-human primates recapitulate most features of human tuberculosis, their use is limited by ethics as well as cost (Myllymäki et al. 2015; Saralahti et al. 2020). Mice do not recapitulate the spontaneous latency of the disease, but offer many advantages such as available genetic and biomedical analysis tools as well as well characterized strains (Myllymäki et al. 2015; Saralahti et al. 2020). As explained in previous chapters, the zebrafish offers many advantages for preclinical immunological research. In addition, the zebrafish is naturally susceptible for mycobacterial infection by *M. marinum*, which is a close relative of the human tuberculosis causing bacterium, *M. tuberculosis*. The two bacteria share common infection mechanisms, including the ESX-1 system which is essential for virulence (Tobin and Ramakrishnan 2008). *M. marinum* grows at a lower temperature (*M. tuberculosis* +37°C and *M. marinum* +25 - 35°C) and therefore infects mainly ectotherms like fish, even though it can also infect humans (Ramakrishnan 2012).

2.4.2.1 Larval zebrafish model

The transparency of zebrafish larvae allows *in vivo* visualization of fluorescently labelled immune cells and bacteria, and it has been used for the characterization of early granuloma formation (Davis et al. 2002). Rather than being solely a passive host protective structure, the current understanding emphasizes the role of the granuloma in shielding the bacteria from eradication by the host immunity (Ramakrishnan 2012). Studies on zebrafish have revealed the dual role of macrophages in both limiting the infection by *M. marinum* as well as transporting the phagocytosed bacteria into host tissues (Clay et al. 2007). Observations on super-infected zebrafish provided the first indication on the dynamic nature of the granuloma by exposing the migration of superinfecting mycobacteria into already established granulomas (Cosma, Humbert, and Ramakrishnan 2004). Later studies have investigated the role of mycobacterial lipids in mediating the formation of vasculature into the granuloma (Walton et al. 2018). Studies in zebrafish have also strengthened the understanding of the function of cytokines and chemokines in a tuberculosis infection (Meijer 2016). Other important findings in zebrafish include e.g. the role of eicosanoid production in the host defence (Tobin et al. 2010), the role of the ESX-1/RD1 virulence locus in recruiting uninfected macrophages to facilitate mycobacterial spread from the granulomas (Davis and Ramakrishnan 2009), the role of host

macrophages in the development of drug tolerance (Adams et al. 2011), and the role of matrix metalloproteinase 9 in granuloma formation (Volkman et al. 2010). In addition, multiple studies have explored genes essential for the early mycobacterial infection in larval zebrafish (Meijer 2016; Renshaw and Trede 2012). The zebrafish model has been expertly reviewed on many occasions (Cronan and Tobin 2014; Meijer 2016; Meijer and Spaink 2011; Myllymäki et al. 2015).

When conducting infection experiments, the zebrafish larvae are typically infected intravenously to cause a systemic infection. Intravenous injection can be conducted at 1-2 dpf, as primary macrophages are present in the tissue and are able to interact with the bacteria, but other strategies can also be adopted depending on the desired outcome (Benard et al. 2012). In a study by Carvalho et al. (2011), a high throughput screen for yolk sack infected zebrafish was used for assessing disease progression and the effect of anti-tuberculosis compounds on fish survival and bacterial growth. They first showed that a low dose (20 - 40 colony forming units (CFU)) yolk sac infection of <1024 cell stage embryos using polyvinylpyrrolidone as a carrier offers a consistent infection dose, easy validation of a successful infection as well as limits the physical injury caused by the caudal vein injection (Carvalho et al. 2011). Primary granulomas were observed throughout the larvae and no adverse effects due to the injection were observed, indicating that the setup can be used for the high throughput screening of anti-tuberculosis drugs in an *in vivo* setting (Carvalho et al. 2011). A similar semi-high throughput screen based on image analysis was published soon after by others, but using a manual infection method (Takaki et al. 2012).

2.4.2.2 Adult zebrafish model

The adult zebrafish presents many advantages for modelling tuberculosis, as it has an adaptive immune system and fully developed physiology. Due to the developed immune system of an adult, it can present with both the latent and the chronic forms of the disease depending on the infection dose (Parikka et al. 2012). Moreover, the latent disease can be reactivated with irradiation or chemical treatment (Myllymäki et al. 2018; Parikka et al. 2012). Adult zebrafish infected with *M. marinum* present capsulated, caseating, necrotic and hypoxic granulomas as seen in humans (Myllymaki et al. 2018; Oehlers et al. 2015; Parikka et al. 2012; Swaim et al. 2007). The granulomas of zebrafish have fewer lymphocytes, but these lymphocytes are nevertheless essential, as the lymphocyte deficit *rag1*$^{-/-}$ mutants are hypersusceptible to the infection (Swaim et al. 2007). An important application of the adult zebrafish

is modelling vaccine responses. Adult zebrafish have been used for studying the use of mycobacterial antigens as DNA-vaccines (Myllymäki et al. 2018; Oksanen et al. 2013, 2016). A review by us on zebrafish as a vaccine model was recently published, so the model will not be further reviewed here (Saralahti et al. 2020).

2.4.3 Intelectins in tuberculosis

Intelectins are a family of carbohydrate binding lectins which are found in a number of vertebrate species as well as cephalochordates, urochordates and placozoan species (Chen, Li, and Yang 2020). The bacterial cell wall contains carbohydrate bearing proteins, which distinguishes them from mammalian cells. Many of these surface proteins are recognized by pattern recognition receptors or lectins such as intelectins. Binding of bacterial sugar moieties marks the bacteria for opsonization, which in turn activates cells capable of phagocytosis. Humans have two intelectins, intelectin 1 (*ITLN1*) and 2 (*ITLN2*). While human *ITLN1* is expressed in multiple tissues (heart, thymus, ovaries, testis, small intestine and colon) *ITLN2* has been shown to be expressed in the intestine (Lee et al. 2001).

Intelectins have been studied in multiple species regarding bacterial binding and their effect on infection. *ITLN1* has been shown to be able to bind BCG, which originates from *Mycobacterium bovis* (Tsuji et al. 2009). However, the over-expression of intelectins in mouse lungs did not protect the mice against a *M. tuberculosis* infection (Voehringer et al. 2007). Despite the lack of evidence on its effect against tuberculosis, human *ITLN1* has been indicated as a biomarker in cancer, and a polymorphism in its sequence has been linked to several diseases, including Chron's disease (Chen et al. 2020; Nonnecke et al. 2021). *ITLN1* has been shown to bind beta-D-galactofuranose present on the surface of bacteria (Tsuji et al. 2009; Wesener et al. 2015). In mouse, both intelectins were found to be upregulated early after a *Trichinella spiralis* infection (Chen et al. 2020; Pemberton et al. 2004). In the same study, ITLN2 was shown to be deleted in the genome of a mouse strain susceptible for *T. spiralis*.

Species specific duplication events (such as TSWGD) may have led to diversification in relation to the function of intelectins, as evidence of bacterial recognition and elimination has been obtained in fish. Upregulation during infection has been observed in both rainbow trout and grass carp intelectin genes (Chen et al. 2020). While agglutination was not observed with recombinant intelectin 1, two species of catfish were found to present agglutination in relation to intelectin binding

(Chen et al. 2020). Evidence of bactericidal activity has been observed in blunt snout bream where intelectin has been shown to increase the macrophage mediated killing of *Aeromonas hydrophila* (Ding et al. 2019). Yan et al. showed with a mutagenesis study in Chinese Amphioxus that intelectin can bind lipopolysaccharide through its non-canonical carbohydrate recognition domain, and furthermore, that this domain is conserved in Itln1 in zebrafish (J. Yan et al. 2018). As in humans, zebrafish *iltn1* is expressed in a number of different tissues, with the highest levels of expression found in the intestine, spleen and kidney, whereas *itln2* is expressed mainly in the intestine (L. Chen et al. 2016; Chen et al. 2018). Both *itln1* and *itln2* expression is also upregulated after a bacterial challenge with *Escherichia coli* or *S. aureus*. In addition, Zebrafish *itln1* and *itln2* were able to agglutinate $Gram^+$ and $Gram^-$ bacteria (L. Chen et al. 2016).

2.4.4 Inflammasomes in tuberculosis

Inflammasomes are multimeric protein complexes assembled in response to cellular damage or pathogens (Martinon, Burns, and Rg Tschopp 2002). Inflammasome signalling leads to the release of the proinflammatory cytokines interleukin 1β (IL-1β) and IL-18 in their mature form from infected macrophages. IL-1β is an essential proinflammatory cytokine which activates the T-cell mediated adaptive immune response, and essential for survival in a mycobacterial infection in mouse (Mayer-Barber et al. 2010). Inflammasomes contain a receptor protein which is able to detect Pathogen Associated Molecular Patterns (PAMPs) or endogenous cell Damage Associated Molecular Patterns (DAMPs). In the case of the activation of either NLRPs or AIM2 receptor, inflammasome pathway is induced. The receptor protein becomes linked to an interleukin activating caspase by an adaptor protein PYCARD (PYD and CARD domain containing, also known as ASC, Apoptosis Speck-like protein containing a CARD-domain) (Martinon et al. 2002). The adaptor protein is not essential for all inflammasomes, at least not for NLPR1 and NLRC4. Incorporation of the receptor, PYCARD and the caspase into the inflammasome leads to cleavage of the pro-interleukin protein into its active form. In addition, inflammasome activation can also activate pore formation by the Gasdermin D protein, which causes the release of the cellular contents of the infected cell into the extracellular space (Beckwith et al. 2020). Pore mediated cell death is referred to as pyroptosis, as it is mediated by infection (from lat. *pyros*, fire). Besides releasing cellular components, pyroptosis has been suggested to trap intracellular bacteria into

the cell debris in pore-induced intracellular traps, analogous to neutrophil extracellular traps (Jorgensen et al. 2016). In addition, inflammasome signalling also activates the STING/cGAS pathway which induces the interferon response (Liu et al. 2016).

The best characterized inflammasomes are the NLRP3 and AIM2 inflammasomes, both of which are known to be activated upon a *M. tuberculosis* infection (Koo et al. 2008; Saiga et al. 2012). Upon infection with *M. tuberculo*sis, the inflammasome has been described to be activated in macrophages and dendritic cells in human and mouse (Ma et al. 2021) (Figure 3). Similarly, other mycobacteria, including *M. marinum* have been shown to activate the NLRP3 inflammasome, this, however is not true for *M. bovis* BCG (Dorhoi et al. 2012). Upon infection with mycobacteria, Toll-like receptors activate NF-κB signalling, which leads to the production of pro-IL-1β, which becomes activated by the inflammasome (Ma et al. 2021). A mycobacterial infection thus leads to activation of the NLRP3 inflammasome through NLRP3, Toll-like receptors and NF-κB, but it also leads to activation of the AIM2 inflammasome through an ESX-mediated pathway (Carlsson et al. 2010; Subbarao et al. 2020). The importance of the AIM2 inflammasome is highlighted by the finding that the loss of AIM2 leads to impaired IFN-γ (Interferon γ) signalling in response to BCG as well as an increase in the bacterial burden (Liu et al. 2016). In addition to NLRP3 and AIM2, polymorphism in NLRC4 associates with a high risk of lung damage in tuberculosis patients with human immunodeficiency virus (HIV) infection, though the direct mechanism is not known (Ravimohan et al. 2020).

Pro-IL-1β activation independent of inflammasomes is possible through serine proteases (Netea et al. 2015), and consequently deletion of the NLRP3 receptor protein has been shown to be dispensable in a mycobacterial infection in mice (McElvania Tekippe et al. 2010). Only the adaptor protein PYCARD has been shown to be essential for survival from late phase *M. tuberculosis* infection independent of inflammasome activation (McElvania Tekippe et al. 2010), though there is some debate about the results (Mayer-Barber et al. 2010). In contrast to the results by McElvania TeKippe et al. (2010), Mayer-Barber et al. (2010) found that the survival of Casp-1 deficient mice was impaired, suggesting a less important role for Pycard. Research has indicated that Pycard could also present inflammasome independent immunological functions including intrinsic mediation of $CD4^+$ T-cell homeostasis (Javanmard Khameneh et al. 2019) and cytotoxic lymphocytes independent of inflammasome activation (Cheong et al. 2020), which could explain the dispensable role of the other inflammasome components. Other evidence has

pointed to inflammasome independent DUSP10, MAPK and cytokine activation through PYCARD (Taxman et al. 2011). In immortalised mouse bone marrow derived macrophages, Pycard was found to be essential for IL-1β and pro- IL-1β production independent of inflammasome activation (Subbarao et al. 2020). Activated PYCARD regulates IFN-β signalling, which suppresses the host protective IFN-γ signalling, potentially by interacting with STING, which prevents downstream signalling to TANK binding kinase 1 (TBK1) (S. Yan et al. 2018). In addition, samples from tuberculosis patients displayed a negative correlation between IFN-β and PYCARD mRNA levels. (S. Yan et al. 2018). As PYCARD is conserved in evolution (Meunier and Broz 2017), it is not surprising that it has so many inflammasome independent roles.

Many bacterial species are known to inhibit or suppress inflammasome activation (Lamkanfi and Dixit 2011; Shin and Brodsky 2015). It has also been suggested that *M. tuberculosis* is able to inhibit the AIM2-inflammasome activation in the host cell in an ESX-I mediated manner, and NLRP3 via its serine/threonine kinase PknF (Rastogi et al. 2021; Shah et al. 2013). Deletion of either the ESX-I or PknF leads to increased inflammasome activity, which can also be detrimental to the host. Inflammasome inhibition has been suggested as a host-directed therapy, which in combination with pathogen directed therapies could limit the damage to the host resulting from the infection and immune response (Ma, 2021). An advantage of the host-directed therapy is that pathogens are less likely to develop resistance to it, unlike to antibiotics, which target the pathogen.

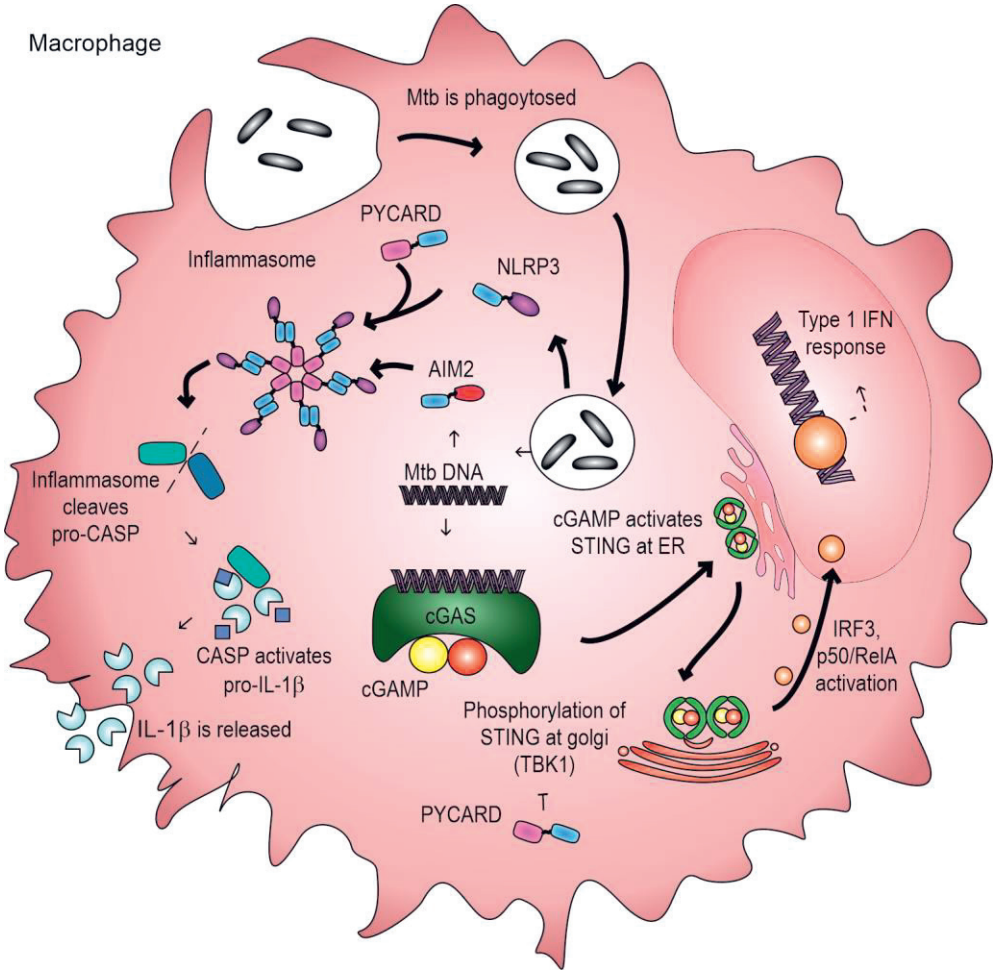


Figure 3. Schematic presentation of inflammasome signalling in a human macrophage upon a *Mycobacterium tuberculosis* (Mtb) infection. Once Mtb is phagocytosed by a macrophage, it activates the NLRP3, which recruits PYCARD to form an inflammasome. Inflammasome activation leads to cleavage of pro-caspase (pro-casp) into an active caspase (casp). Caspase cleaves pro-IL-1β into active IL-1β, which is released. The degrading Mtb releases its DNA, which can activate the AIM2 inflammasome (PYCARD is not required). Also, Mtb DNA can activate cGAS, which triggers cGAMP (Cyclic guanosine monophosphate–adenosine monophosphate). cGAMP activates STING at the endoplasmic reticulum (ER). STING/cGAMP migrates to the Golgi, where it is phosphorylated by TBK1. Activation of STING leads to activation of transcription factors IRF3 or p50/RelA. Transcription factors translocate to the nucleus to drive the expression of type I interferon responsive genes. PYCARD has been suggested to block the TBK1 mediated cGAS-STING signalling. Adapted from Wassermann et al. (2015).

2.5 Modelling developmental genetics and diseases in zebrafish

2.5.1 Zebrafish model for development

All deuterostomes share many genetic features of development, and the developmental processes in all vertebrates have further shared features. As described in previous chapters, zebrafish share a high number of their genes with humans (Howe et al. 2013). In addition to the >80% of disease associated genes, many developmentally important genes are also shared between fish and man. With regard to developmental diseases, deleterious alleles of genes important in development can cause diseases which manifest during childhood. Based on the paediatric disease annotation database, 75% of genes associated with human paediatric diseases have a clear homolog in zebrafish (Varga et al. 2018). Recently, transcriptional profiles of human and zebrafish oocytes showed that highly expressed genes are orthologous between the two species, suggesting that the zebrafish could also serve as a valid model for human oocyte studies (Can et al. 2020). Recent single cell studies facilitate the probing of genes and assessing their significance during vertebrate development and disease (Farnsworth et al. 2020; Farrell et al. 2018; Raj et al. 2018; Tambalo, Mitter, and Wilkinson 2020; Wagner et al. 2018; Xue et al. 2019) (Table 2).

2.5.2 Zebrafish model for rare diseases

Rare genetic diseases have an occurrence of less than 1:2000 people. Of these, 80% have a genetic aetiology, and 75% of them manifest in childhood (Adamson, Sheridan, and Grierson 2018; Wangler et al. 2017). Especially for genes whose deletion leads to an embryonic lethal phenotype in the mouse, the zebrafish could be used to study their significance due to *ex utero* fertilization and development (van Rooijen et al. 2009; Uusimaa et al. 2018). Of course, like limitations caused by the absence of a placenta, studies on diseases of tissues with known differences or absences, such as lungs, can't be studied in zebrafish. Similarly, studies using zebrafish are often limited to diseases caused by loss of function mutations, as gene regulatory elements have evolved and differentiated to the extent that diseases caused by aberrant gene expression patterns can't always be recapitulated.

Besides enabling the discovery of gene functions by reverse genetics, the zebrafish is also suitable for large scale drug screens. Screening for the molecular function of a disease gene and generating personalized zebrafish models can also

help in screening for effective drugs to treat a rare disease phenotype (Cully 2019). Large scale chemical screening using zebrafish embryos is fast and cost effective (Cassar et al. 2020). Several clinical trials based on zebrafish drug screens are ongoing and listed in Patton et al. 2021 (Patton et al. 2021). The zebrafish has been used for studying rare diseases as well as screening for drugs in several occasions.

Dravet syndrome is a rare condition (incidence 1:15700 in the Unites States (Wu et al. 2015)), which includes a severe form of epilepsy and intellectual disability, and can be caused, among other reasons, by a mutation in the gene SCN1A, which encodes a voltage gated sodium channel protein. By using a mutant zebrafish replicating the mutation observed in patients, Baraban, Dinday, and Hortopan (2013) found that a compound called clemizole inhibited seizures and convulsive behaviour. By September 2021, the drug had reached phase II clinical trials (ClinicalTrials.gov Identifier: NCT04462770).

In a similar case, Diamond Blackfan anaemia (incidence 5-7:1000000, National Organization for Rare Disorders, rarediseases.org/, Accessed: 01.10.2021), which causes red cell aplasia and craniofacial abnormalities was replicated in a zebrafish model. A zebrafish mutant for the rps29 protein recapitulated the disease phenotype and was used for a chemical compound screen (Taylor et al. 2012). The screen revealed that a calmodulin inhibitor trifluoperazine, was able to rescue the phenotype in zebrafish (Taylor et al. 2012). This compound is also in an ongoing phase I clinical trial (ClinicalTrials.gov Identifier: NCT03966053).

More recently, zebrafish were used for screening treatments for a patient suffering from a lymphatic anomaly resulting from an unknown genetic cause. A whole exome screen identified a gain of function mutation in the gene ARAF which is conserved between human and zebrafish (D. Li et al. 2019). A zebrafish overexpression model for the disease was generated using a transgene with a promoter for lymphatic expression, and this model was used for testing potential treatments (D. Li et al. 2019). Finally, an existing MEK inhibitor was found to be able to rescue the zebrafish model phenotype, and was administered to the patient as a compassionate treatment, which saved the patient's life (D. Li et al. 2019).

In addition to the diseases mentioned here, the zebrafish has been used for studying for example the disease pathology of von Hippel-Lindau disease (van Rooijen et al. 2009), the disease genetics of Hirschsprung disease (Gui et al. 2017) and osteoporosis causing ATP6V1H deficiency (L. Li et al. 2018).

2.5.3 Personalised medicine

Personalized medicine aims at generating patient-specific treatments for diseases. Next generation sequencing data directly from the patient helps in understanding the tissue level pathology of the tumour microenvironment, instead of selecting treatments based on trial and error. Especially when treating multigenic malignancies of complex origin, such as cancer, or studying monogenic diseases affecting multiple tissues, such as in immunology, personalized medicine will help stratify patients based on genomic data. To facilitate personalised drug selection, zebrafish avatars carrying injected human cancer tissue cells have been created (Costa et al. 2020). This is similar to developing a humanised zebrafish or generating mutated zebrafish with patient specific genetic alterations. The cancer-avatars can then be tested for phenotypic responses to drugs. As zebrafish produce hundreds of offspring within one spawning, semi-large scale (96/384 well plate) screens for drugs are feasible. Zebrafish embryos develop all essential organs within a few days from spawning, and drugs can be administered directly into the water. However, a few pitfalls remain, including non-mammalian physiology, differences in drug absorption, distribution, metabolism, and excretion (ADME) in relation to toxicology, as well as requirement for a higher than normal rearing temperature as per required by the transplant (Costa et al. 2020). Nevertheless, zebrafish have been successfully used for pharmacological studies which will be covered next.

2.5.4 Pharmacological studies with zebrafish

The pharmacological assessment of new drugs requires rigorous evaluation, as the safety of human patients is the priority. Drugs and chemicals are assessed for their safety and toxicity with well validated tests *in vitro* and *in vivo* using mammalian models. However, it has been suggested that improving preclinical testing could aid in minimizing unpredicted risks later on (Cassar et al. 2020). In addition, the later the unpredicted hazards are identified, the more costly the testing becomes, so information on drug toxicity before selecting candidates for preclinical testing could spare expenses. The zebrafish embryo has been suggested as a middle man between *in vitro* and mammalian tests, in order to assist in developing the lead compound for a candidate for mammalian studies in drug development (Cassar et al. 2020). In comparison with mammalian assays, pharmacological testing with zebrafish larvae is faster, and offers an inexpensive technology for safety testing before selecting a drug candidate for time consuming and costly mammalian *in vivo* studies (Barros et al.

2008). The parameters which can be analysed from zebrafish include changes in heart rate and atrial-ventricular dissociation, changes in the activity of the nervous system (controlled movement, convulsant and proconvulsant activity, cognitive impairment, drug dependence potential, changes in vision and hearing), renal function, gastrointestinal function and bone density (Barros et al. 2008; Cassar et al. 2020). However, zebrafish toxicity assays have not been validated for testing human drug toxicity so far. In turn, the European commission has accepted the zebrafish embryo model as a validated alternative (not default) assay to assess acute toxicity in fish (Organisation for Economic Cooperation and Development, OECD, No. 236) (Cassar 2020). For testing drugs aimed for human use, the simplicity of immersing the larva into a droplet containing the substance of interest is both the advantage and the pitfall of the model. The drugs are readily absorbed through the skin and gills making administration easy for soluble drugs, but makes the evaluation of the ADME untranslatable to mammalian organisms and finally the human patient (Cassar et al. 2020). An option to overcome the limitations presented and to increase zebrafish applicability would be to develop organ specific assays and micro delivery systems for drugs. One hindrance is that zebrafish are not as inbred as mice, and thus present more variation in results. Therefore, ideally, zebrafish strains should be harmonized and inbred to achieve congruent results. However, this is actually paradoxical as patients are also not expected to be a homogeneous population with a standardised lifestyle, diet and age.

2.6 Genetic and transcriptomic tools for zebrafish

2.6.1 Transgenic zebrafish

Much of the popularity of the zebrafish model derives from the transparency of the larvae, which allows non-invasive observations of internal structures and fluorescent compounds. This has led to the development of a number of transgenic zebrafish lines carrying fluorescent markers, generated using insertional mutagenesis. These transgene carrying zebrafish have permitted the non-invasive analysis of cell and protein functions *in vivo* in the fields of oncology, development, and immunology, which has partly overcome the disadvantage of having only a few monoclonal antibodies available for studying the changes in the protein levels. Fluorescently labelled blood cells have been used for studying zebrafish haematology, immunology

and host pathogen interactions (Carradice and Lieschke 2008; Cronan and Tobin 2014; Rounioja et al. 2012). Especially useful in the field of immunology, this system has enabled combined analyses with fluorescently labelled bacteria, and has shed light on previously unexplored interactions between the host and its pathogen (Gomes and Mostowy 2020; Meijer and Spaink 2011; Phelps and Neely 2005). Similar to larval zebrafish, transgenic adult zebrafish carrying fluorescent markers can be used for studying individual cell types. Although the adult zebrafish are not by default transparent, mutant strains such as Casper (roy$^{-/-}$;nacre$^{-/-}$) are fairly unpigmented and allow the detection of fluorescent labelling *in vivo* (White et al. 2008). This is especially useful in studies involving transplants and oncology (Costa et al. 2020; White et al. 2008). Along with the emergence of next generation sequencing technologies, this has enabled the straightforward separation of fluorescently labelled cells using fluorescence activated cell sorting (FACS) and their subsequent *omics* (i.e., genomic, transcriptomic, proteomic, epigenomic, etc.) analyses.

2.6.2 Transcriptomic profiling

Transcriptomic assays – first based on microarrays, now RNA sequencing (RNA-seq) – and single cell RNA-sequencing (scRNA-seq) have become a standard for profiling changes in gene expression. RNA-seq experiments can cover transcriptional events during embryogenesis and early development, but also uncover changes in gene expression caused by different treatments, effects of mutations or gene silencing. With the development of single cell transcriptome analyses (Tang et al. 2009) the emergence of droplet based barcoding, the scRNA-seq analysis has plunged forward (Klein et al. 2015; Macosko et al. 2015). ScRNA-seq covers the transcriptome of thousands of individual cells simultaneously, and with the help of marker genes, data can be grouped for a comprehensive analysis across tissues. Especially with zebrafish embryos, scRNA-seq has provided unique glimpses into developmental dynamics, lineage trancing of individual cells during development and the formation of the brain (Farrell et al. 2018; McKenna et al. 2016; Raj et al. 2018; Satija et al. 2015; Wagner et al. 2018). Clustering of the profiled cells from zebrafish larvae by tissue specific marker genes across multiple time points can aid in deducing the effect of a single gene in, for example, organogenesis. In adult zebrafish, transcriptome profiling has largely concentrated on haematopoietic analyses of kidney marrow (Tang et al. 2009) and spleen (Hernández et al. 2018). In

turn, analyses of infected larval and adult zebrafish have shed light on the genetic response to pathogen invasion (Benard et al. 2012; Harjula et al. 2020; Hernández et al. 2018).

Traditionally, the separation of immune cell types for analysis has relied on the existence of antibodies, and this has hindered the research on zebrafish immune cells. Cell sorting based on size (Forward Scatter, FSC) and granularity (Side Scatter, SSC) can be used in separating the blood cells of zebrafish to erythrocytes, lymphocytes, granulocytes and precursor cells, and Wright Giemsa-staining combined with a gene expression analysis has been used for confirming blood cell identity (Langenau et al. 2004; Traver et al. 2003). New methods, such as scRNA-seq analysis partly circumvent the lack of reliable antibodies as they allow the grouping and analysis of cells expressing cell type specific marker genes. Many single cell studies have recently explored haematopoietic lineages in zebrafish to further characterize the cell types (Athanasiadis et al. 2017; Carmona et al. 2017; Hernández et al. 2018; Moore et al. 2016; Tang et al. 2017).

Mining of open access data can provide researchers with essential cues on the selection of a gene, tissue and timepoint for an experiment. However, several problems can arise when determining, for example, the transcriptional response to a pathogen. An infection can develop differently between individuals due to genetic as well as environmental factors, which are altered as the infection progresses. Some of this is well understood, as are the role of early and late response genes in an infection, but some are less well understood. Selecting only a single timepoint during a course of an infection can thus give limited information of the progression of the infection in a flock of zebrafish. Furthermore, a large enough group size is required to overcome the variation caused by environmental factors, which often presents problems with novel, more expensive technologies. Ultimately, the sensitivity of the assay caused by variation between cells and the possibility to only sample around several thousand cells can lead to the desired signal becoming lost in the noise in the data. This is also a problem at the level of data analysis (Lähnemann et al. 2020)

Table 2. Zebrafish single cell experiments available for data visualization

Topic	Treatment	Age	Tissue	Ref
Visualization of Adult Zebrafish Whole Kidney Marrow	Untreated	Adult	Kidney marrow	1.
Transcriptional states and fate decisions in haematopoiesis	Transplantation of fluorescently labelled cells into rag2 mutants	Adult	Kidney marrow	2.
Lymphocytes and ILC-like cells in zebrafish	Untreated, infected	Adult	Kidney marrow, thymus, spleen, gut, gills	3.
Profiling of lineages and cell types in the vertebrate brain	CRISPR barcode labelling	23-25 dpf	Brain	4.
Reconstruction of developmental trajectories during zebrafish embryogenesis	Untreated	3.3 - 12 hpf	Whole embryo	5.
Gene expression landscapes and lineage in the zebrafish embryo	Untreated	4 - 24 hpf	Whole embryo	6.
A 3D Atlas of Hematopoietic Stem and Progenitor Cell Expansion	Morpholino silencing, CRISPR KO, overexpression of transcription factors	55 hpf	Caudal haematopoietic tissue	7.
Transcriptome atlas for zebrafish development	Untreated	1 - 5 dpf	Whole embryo	8.
Transcriptome atlas of the developing zebrafish hindbrain	Heat shock activation of transgene	16 - 44 hpf	Hindbrain	9.

Reference, data availability:

1. (Tang et al. 2017) https://molpath.shinyapps.io/zebrafishblood/
2. (Athanasiadis et al. 2017) https://www.sanger.ac.uk/tool/basicz/
3. (Hernández et al. 2018) https://www.sanger.ac.uk/science/tools/lymphocytes/lymphocytes/
4. (Raj et al. 2018) https://krishna.gs.washington.edu/content/members/aaron/fate_map/harvard_temp_trees/
5. (Farrell et al. 2018) https://singlecell.broadinstitute.org/single_cell/study/SCP162
6. (Wagner et al. 2018) www.tinyurl.com/scZfish2018
7. (Xue et al. 2019) https://www.picb.ac.cn/hanlab/ichtatlas/Home/
8. (Farnsworth et al. 2020) http://cells.ucsc.edu/?ds=zebrafish-dev
9. (Tambalo et al. 2020) https://singlecell.broadinstitute.org/single_cell/study/SCP667

2.6.3 Genetic engineering to study gene function

Prior to target specific mutagenesis, the technologies for whole genome sequencing and the establishment of reference genomes, studies on gene function had to rely on phenotype-specific analyses of random mutagenesis events. For the detailed analysis of developmental pathways, researchers relied on the use of model organisms, which allow the visual inspection of development, rather than mammalian species, which develop *in utero*. The first screens on the fruit fly and *C. elegans* shed light on many important genes essential for the first stages of development (Brenner 1974; Nüsslein-Volhard and Wieschaus 1980). As zebrafish is a vertebrate and thus a closer relative to humans than fruit fly or *C. elegans*, it quickly gained popularity among developmental scientists.

Prior to targeted genome modifying technologies, morpholino silencing was the standard for analysing the functions of embryonic genes in zebrafish. Still, morpholinos are essential for gene silencing especially in cases where knockout of the gene is lethal to the host, as the effects of morpholinos can be titrated. Since the early 2010s, the development of targeted mutagenesis technologies has flooded the field. In comparison with forward genetics, which determines the genetic basis responsible for a phenotype, reverse genetics mutates a gene to observe its effect on a phenotype. In combination with the availability of the complete genome sequence of the organism and the explosion in the omics studies across species, targeted mutagenesis has enabled genotype-based discoveries. After the emergence of three methods, Zinc Finger Nucleases (ZFN), Transcription-Activator Like Effector Nucleases (TALEN) and CRISPR-Cas9, the beginning of 2020s has clearly underscored CRISPR-Cas9 as the method of choice for targeted mutagenesis (Figure 4). However, each technology has its advantages, limitations, and pitfalls.

2.6.4 Forward genetic screens

Zebrafish has been used for developmental forward genetic screens since the 1990s mainly due to the very first onerous screens completed in the labs of Christiane Nüsslein-Volhard and George Streisinger (Nüsslein-Volhard 2012). Zebrafish, once again, offers a simple vertebrate platform to genetically manipulate and to detect morphological changes in embryos based on its *ex utero* development. Prior to chemical mutagenesis, genetic interference was accomplished with X-ray irradiation (Chakrabarti et al. 1983). The first genome wide ENU-mutagenic (N-ethyl-N-nitrosourea) screen in zebrafish was done solely based on phenotype (Haffter et al.

1996b). As random mutagenesis generates heterozygous alterations into the DNA, the mutated progeny is screened for marker positive founders to generate future genetically altered lines. The selected founders are then crossed to obtain homozygous offspring. Early studies relied on generating haploid organisms by producing impotent sperm that trigger egg division which leads to the development of haploid homozygous individuals (Streisinger et al. 1981).

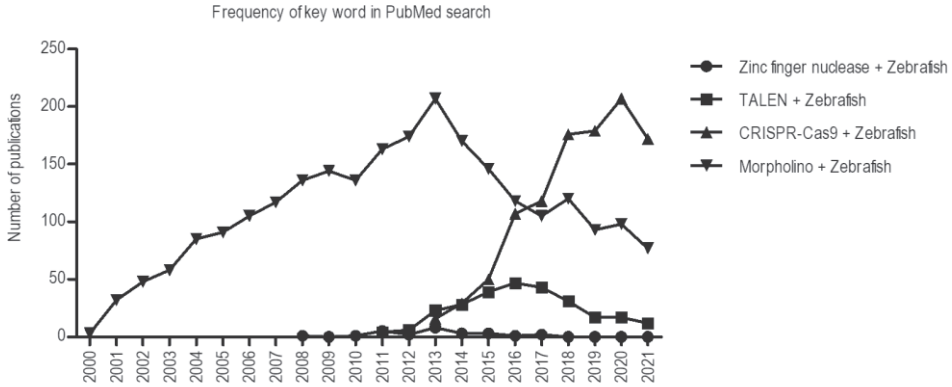


Figure 4. The number of search results for a keyword in PubMed per year. The searched words are indicated in the legend.

The disadvantage of random mutagenesis is the generation of a number of mutations simultaneously across the genome (Lawson and Wolfe 2011). Furthermore, irradiation can cause chromosomal rearrangements as well as large deletions, which impair the selection of the phenotype causing mutations (Lawson and Wolfe 2011). ENU in turn causes mainly point mutations, however with the ensuing difficulty of identifying the correct, phenotype causing lesion, as mutations can also be linked and co-selected (Lawson and Wolfe 2011). Lastly, mutations on the coding regions can cause dysregulation of a number of genes simultaneously, and data of gene regulatory regions in zebrafish has only recently started to emerge (Yang et al. 2020).

Later screens have applied a number of insertional techniques, including retrovirus- and transposon-based methods (Lawson and Wolfe 2011). The resulting mutations can be identified with an inverse polymerase chain reaction (PCR) or next generation sequencing based methods. Insertional mutagenesis lessens the burden of screening for successful mutagenesis with the aid of fluorescent tags (termed protein traps) carried within the insert. However, similarly to other random mutagenesis methods, identification of the genes involved in the phenotype can be challenging when the insertion is located outside protein coding regions of the

genome, or when multiple insertions are found. In addition, mutagenesis is less efficient than with ENU, so larger screens are required for the identification of the same number of phenotypes (Lawson and Wolfe 2011).

Along with decrease in costs, emergence of targeted mutagenesis technologies and reference sequences made it possible to effectively conduct reverse genetics. This has decreased the requirement to conduct laborious random genetic screens. However, forward genetics still serves a purpose of identification of genes with unknown functions.

2.6.5 Gene knockdown with Morpholinos

At their simplest, reverse genetic experiments aim to knock down or knock out the expression of a gene to observe its function by preventing its translation into a functional protein. While transient silencing is achieved with the use of RNA interfering compounds, a gene knockout can be generated with mutagenesis. Early studies in zebrafish utilising double stranded RNA for gene silencing lead to unspecific effects, later interpreted as antiviral effects, which lead to the adoption of the newly introduced morpholino technology instead of RNAi in zebrafish (Blum et al. 2015; Oates, Bruce, and Ho 2000). Zebrafish gene silencing experiments most commonly utilise morpholino mediated gene silencing due to its simplicity and well-established protocols. The efficacy and ease of morpholinos explain their popularity and give a clear indication as to why other RNA-interfering technologies (e.g., short interfering RNAs) have not become popular in zebrafish research.

Morpholinos are nucleotide analogues formed by DNA base pairing attached to methylenemorpholine rings, linked together by phosphorodiamidate bonds (Summerton and Weller 1997). Due to their structure, morpholinos are stable and resistant to enzymatic degradation, as well as unlikely to bind to unspecific target sites due to the lack of charge on the backbone of the structure (Eisen and Smith 2008; Summerton and Weller 1997). Morpholinos are usually designed over a 25-basepair complementary target site to bind to the mRNA molecule prior to mRNA processing and translation (Eisen and Smith 2008). A 40-60% GC-content (guanine and cytosine content) on the target sequence is optimal, as is avoiding four or more continuous guanines (Eisen and Smith 2008). Morpholinos can be designed to bind to or close to the translation start site of the mRNA to prevent attachment of the translation machinery, or to bind to a splice site, which will either result in exon skipping or intron inclusion. This will ultimately alter the mature mRNA-sequence,

leading to either mRNA degradation or the production of a non-functional protein product (Eisen and Smith 2008). In the case of non-coding RNAs, morpholinos can also be designed to prevent the formation of the RNA secondary structure (Kloosterman et al. 2007). Lastly, photoactivated morpholinos can be targeted to a limited number of cells and tissues (Tallafuss et al. 2012). Morpholinos are typically injected into the fertilized zebrafish egg, and their effect starts to wane after the first four days of development, which is also why they are almost exclusively used in larval zebrafish. The functionality of morpholinos can be assayed by western blot or by quantification of the altered mRNA with a quantitative polymerase chain reaction (qPCR).

Since the initial success with morpholinos in the early 2000s, it became evident that some morpholinos display off-target effects and fail to recapitulate the phenotype of the knockouts of the target gene (Kok et al. 2015; Robu et al. 2007). Since the publication of their paper proposing the abandonment of morpholino technology, Kok et al. have faced criticism for their conclusions, and especially for their methods which did not replicate the protocols of the primary studies they criticised, e.g. in the case of the morphant for the *megamind* gene (Blum et al. 2015; Morcos, Vincent, and Moulton 2015). Moreover, it has been largely accepted that off-targeting is a recognized down side of all antisense (and mutagenesis) technologies and the off-target effects need to be taken into account when interpreting the results of the experiments (Eisen and Smith 2008). However, criticism is still essential for pointing out flaws which can be overcome. In another study, Joris et al. (2017) reported that only 11 bp at the splice site were needed for a morpholino binding its target site, and blocking the splicing, increasing the number of potential off-targets. They suggested the number of secondary binding targets could be reduced by using a higher rearing temperature for the embryos (Joris et al. 2017). In addition, morpholinos can cause dose related toxicity, which triggers p53 expression and apoptosis (Robu et al. 2007). Appropriate controls, such as more than one morpholino, a mismatch or a random sequence control, dose-response titration, examining the effect of the mutation on a null-background quantifiable knockdown, simultaneous knockdown of *tp53* and mRNA rescue are encouraged, as well as comparison with a mutant phenotype, when available (Eisen and Smith 2008; Stainier, Kontarakis, and Rossi 2015). It has also been suggested that the morpholino phenotype should be replicated by generating mutations in the 5'-untranslated region of the gene to silence its expression and testing if the phenotype persists (Cunningham et al. 2020).

2.6.6 Targeted mutagenesis

Sequence specific targeting technologies rely on nucleases but differ on the mechanism of sequence specific targeting. While ZFN and TALENs rely on modified protein sequences which are laborious to produce and have to be obtained commercially in most labs, CRISPR-Cas9 is directed to its target site by a guide RNA molecule (referred to as short guide RNA (sgRNA) or simply gRNA) that homes to the targeted site instructed by complementary base pairing (Jinek et al. 2012). Generation of the gRNA molecule from a DNA template is straightforward using *in vitro* transcription, and at its simplest the nuclease molecule can be purchased ready made without the need for customization. This simplicity explains the enormous popularity CRISPR-Cas9 has gained since its commercialization (Figure 4). All three major technologies available for targeted mutagenesis utilise the generation of double stranded breaks (DSB) into the genome, taking advantage of the error prone repair machinery of the cell (Gaj, Gersbach, and Barbas 2013). The DSB are repaired either with homology directed repair (HDR) or non-homologous end joining (NHEJ), the former leading to homologous template mediated repair, and the latter often leading to incorporation of mutations into the original DNA sequence (Gaj et al. 2013). HDR uses the other intact allele as a template, but it can be manipulated to use instead a partly homologous template to repair the site of the double stranded break to introduce targeted modifications to the nuclease binding site. In comparison to HDR, NHEJ repairs the DSB without a complementary template, and is considerably faster than HDR and thus also more common (Miyaoka et al. 2016). NHEJ often leads to small indels, as the repair complex trims the ends of the strands to match each other before paired-end ligation.

ZFN are artificial restriction enzymes, which are attached to the FokI endonucleic restriction enzyme cleavage domain. Engineering the zinc finger domain can modify the DNA sequence binding specificity of the nuclease, which can be exploited to generate targeted DSB at a desired location (Beumer et al. 2008; Bibikova et al. 2002). Originally, ZFNs used a dimer of three fingers recognizing three base pairs (bp) each on the opposite strands of the target DNA, generating an 18 bp binding site. Later, however, the binding site has been extended to improve specificity (H. Li et al. 2020). The DSB is repaired with NHEJ/HDR, leading to potential mutations at the cut site (Beumer et al. 2008; Bibikova et al. 2002). The downside of the ZFN technology is that it requires protein level modifications for target design (H. Li et al. 2020; Urnov et al. 2010). Also, in comparison to ZFN, TALENs have greater specificity and efficiency (H. Li et al. 2020). TALENs function

similarly to ZFN, but instead of zinc fingers, it employs the transcription activator-like (TAL) domains for sequence recognition. The TAL DNA-binding domains are fused to the restriction endonuclease FokI which cleaves both DNA strands (Christian et al. 2010). Their binding to the target site is directed by repeat variable di-residues, which bind to the target site with a one-to-one correspondence (Christian et al. 2010).

CRISPR-Cas9 technology stems from the bacterial acquired immune system, which recognizes and memorises viral DNA by cleaving and pasting it into the bacterial genome. Many types of CRISPR-Cas have been described in different bacteria (Makarova and Koonin 2015), however perhaps the most applied of these is the type II CRISPR-Cas9 systems found from *Streptococcus pyogenes*. From now on in the text, CRISPR-Cas9 will refer to this system for the sake of clarity. In principle, the CRISPR fragments are transcribed from the bacterial genome and processed, and they anneal to form short, trans activating CRISPR RNAs (tracrRNA) which guide Cas9 to silence foreign viral genetic material. The type II CRISPR-Cas9 system has become the most commonly applied CRISPR tool to date. In this system, synthetic, CRISPR guide RNA (gRNA) molecules combine both the structure and specificity of the tracrRNA and the CRISPR-RNA (crRNA) to modify eukaryotic genomes, and when injected into zebrafish embryos, to modify the genome of an organism (Hwang et al. 2013) (Figure 5). The requirements for a guide RNA molecule are the uniqueness of the target site, a 40–60% GC-content as well as the presence of an NGG protospacer adjacent motif at the end of the target site. Similar to ZFN and TALENs, Cas9 generates a DSB and leads to NHEJ or HDR, facilitating the generation of mutations or the incorporation of foreign DNA material. The ease of CRISPR-Cas9 stems from the base pair complementarity of the guide RNA with the edited target, as well as the simple requirements for the target site. Multiple online tools facilitate this process and ordering oligos for the tabletop synthesis of gRNAs is an everyday task for any biotech researcher. Several hindrances have emerged, namely in the targeting of CRISPR to open chromatin and the relationship of the chromatin of the target area and its effect on off-targeting, as well as the emergence of genetic compensation.

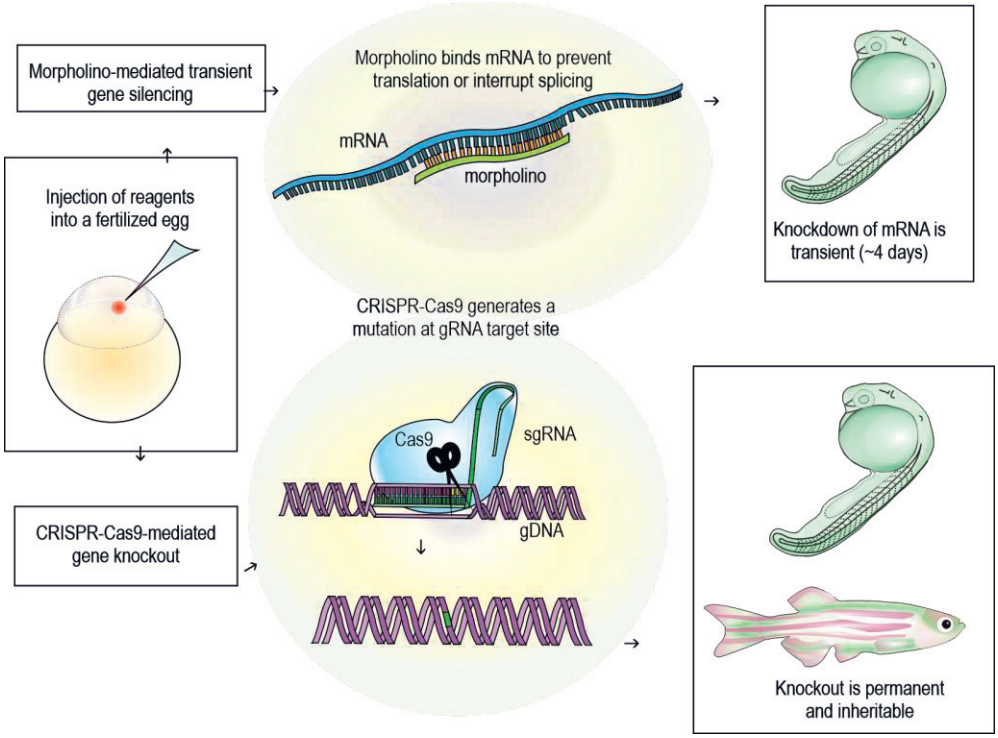


Figure 5. Schematic explanation of zebrafish gene silencing and genome editing. While the knockdown of an mRNA using morpholinos is transient, gene editing with CRISPR-Cas9 generates permanent changes in the genomic DNA (gDNA). Cas9 is directed to its target site using the complementary sequence from the short guide RNA (sgRNA). Mutations will also be inherited if the mutation occurs in the germ line cells. Adapted from Elmonem et al. (2018).

2.6.6.1 Chromatin hindrance and genetic compensation

All targeted genome engineering technologies are subject to errors which are largely dictated by their specificity and efficiency. Especially in a clinical setting, the technology must be precise and free of side and off-target effects. The simplest way to avoid off-target modifications is to choose a target site that is unique. Still, all technologies are prone to misguided binding, especially if targets appear in open euchromatin and the true target is located in the closed heterochromatin (Kuscu et al. 2014; Wu et al. 2014).

Studies (as reviewed by Verkuijl and Rots (2019) have found CpG-methylation and nucleosome positioning have the greatest effect on Cas9 binding, whereas

DNAse hypersensitivity and DNA methylation cause less hindrance. Direct evidence pointing towards the influence of heterochromatin on CRISRP-Cas9 efficiency has been provided by multiple studies, which have measured Cas9 binding and cleavage on sites either located on open euchromatin or closed heterochromatin controlled by small drug availability, or in or outside of nucleosomes (X. Chen et al. 2016; Daer et al. 2017; Hinz, Laughery, and Wyrick 2015). Hinz et al. tested the cleavage efficiency of Cas9 using *in vitro* assays and concluded that targets buried in nucleosomes are inaccessible and that the location of the PAM sequence outside the nucleosome is especially important to achieve binding and cleavage. Daer et al. used a drug inducible switch which regulates chromatin state in a cell line and showed that switching to an open chromatin status was able to restore the cleavage efficiency of Cas9. In a similar assay, X. Chen et al. (2016) used cell-based reporter systems to demonstrate the effect of 3D chromatin structure on the efficacy of CRISPR-Cas9 and TALEN cleavage activity. These results consistently point to the inaccessibility of the sites positioned in closed chromatin. A logical explanation to why this is, is that in a dividing cell and the chromatin landscape, Cas9 more actively scans areas located on open chromatin, which also increases the probability of binding and cleavage, as shown by Knight et al. (2015). Using chromatin modifiers and altering the chromatin state towards euchromatin can also improve Cas9 binding (Horlbeck et al. 2016; Isaac et al. 2016; Yarrington et al. 2018). Moreover, at the level of genetically modifying a whole organism, information on the chromatin landscape can be used to improve guide-selection and experimental design (Chen et al. 2017; Singh et al. 2015), as the gRNAs with the highest efficiency are found in chromatin areas with low nucleosome occupancy (Horlbeck et al. 2016).

Even though studies on cell lines give an easier read out of the phenomenon, they fail to recapitulate the changing chromatin environment of a developing embryo, as the totipotent cells adopt a lineage specific chromatin signature. To conclude, maximising the specificity and efficiency of mutagenesis is important for both ensuring the safety of patients and minimising off-target effect mediated bias in animal models. Taking chromatin status into account can also serve the experimental design: recent results indicate that heterochromatin is more prone to be repaired via HDR, increasing the probability of insertional mutagenesis (Liscovitch-Brauer et al. 2021). However, studies also imply that the chromatin-mediated hindrance is transient and also dependent on other experimental factors (Isaac et al. 2016; Kallimasioti-Pazi et al. 2018). This is likely, as chromatin is not a static structure but loosens and remodifies as cells undergo mitosis. In addition, most available gRNA design tools are based on the uniqueness of the target site and the secondary

structures of the gRNA, and do not incorporate *in vivo* evidence or chromatin status in embryonic context. Two tools have been developed, which take into account the chromatin status of the target cell. A gRNA design tool by Singh et al. (2015) employs data from human cell lines, and similarly, Chen et al. (2017) have constructed a tool with improvements based on the accessibility of the zebrafish chromatin.

2.6.6.2 Genetic compensation

As mentioned, during the 2010s, morpholino studies have been disregarded due to the detected off-target effects, but their use has also diminished due to the ease of generating zebrafish mutants. Already in 2008, Robu et al. showed that morpholino injections can lead to the induction of *tp53* transcription, and tp53 interference in the phenotype. Around 15-20% of morpholinos have been estimated to present off-targeting (Robu et al. 2007). However, this effect is not limited to morpholinos but is also present in short interfering RNA technology (Robu et al. 2007). Mutants are not either, however, without experimental drawbacks. Whole transcriptome analyses have since shown that instead of off-target effects, the discrepancy between morphants and mutants was caused by genetic compensation in the mutants (Rossi et al. 2015). Genetic compensation was caused by deleterious mutations, which trigger a number of related genes to become upregulated as a result of a loss of a gene, masking the phenotype observed with the knockdown (Rossi et al. 2015). Knockdowns are not subject to the same compensatory effects, justifying the use of morphants in comparison to mutants (Peng 2019). The presence of compensation has become evident through studies which have described similar results (Eve, Place, and Smith 2017; Wei et al. 2017; Zhu et al. 2017). The current viewpoint is that the genetic robustness of an organism instructs protein feedback loops to maintain fitness in the case of deleterious mutations leading to premature mRNA degradation, and in addition, independent of the protein feedback loops, transcriptional adaptation leads to upregulation of genes in a sequence dependent manner (El-Brolosy et al. 2019; Ma et al. 2019). This *robustness* is a result of the redundancy of the genome and transcriptome to overcome the presence of loss of function mutations, and it seems that compensatory genes are often members of the same pathway, or a paralogue generated by a gene duplication. As the zebrafish often has multiple paralogues of the same gene due to the TSWGD, it should be asked whether the zebrafish is more prone to or able to compensate for the loss of a single gene than humans, for example.

Indeed, transcriptional adaptation is not triggered by the loss of the protein product but results from the presence of the mutated mRNA. It can result in both the upregulation of adapting genes as well as opening of the chromatin structure due to the increase in mutated mRNA decay (El-Brolosy et al. 2019). Genes sharing sequence similarity to the mutated gene became upregulated due to transcriptional adaptation (El-Brolosy et al. 2019; Ma et al. 2019). This effect has been suggested to be due to the mRNA decay products which translocate to the nucleus to interact with histone modifiers and chromatin remodelling complexes, as well as due to decayed transcripts interfering with the expression of the antisense transcripts which should downregulate the compensatory genes through nonsense mediated decay (El-Brolosy et al. 2019). To work around this, silencing a gene by mutating it so that an mRNA is not produced at all (e.g. by mutating promoters) was shown to lead to a desired phenotype without compensation (El-Brolosy et al. 2019). However, evidence for this phenomenon is still scarce. Still, research on genetic compensation is important for progressing in understanding disease genetics behind deleterious and less deleterious mutations leading to different phenotypes observed in patients with genetic diseases (Peng 2019).

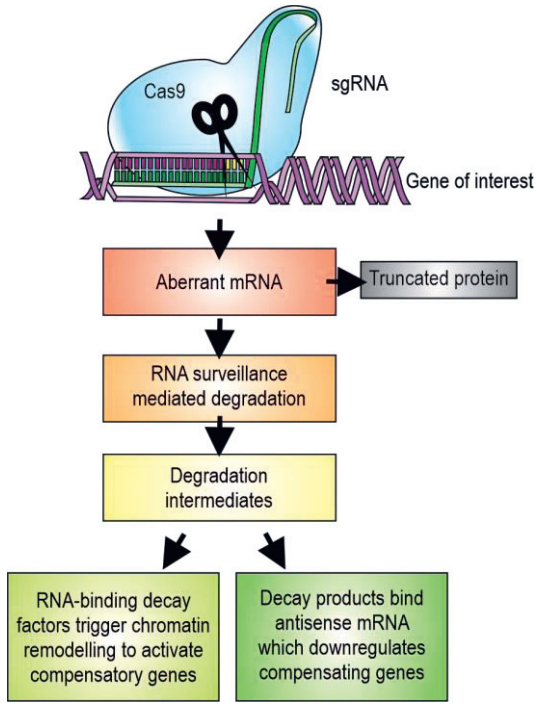


Figure 6. Mechanisms of genetic compensation mediated by aberrant mRNA. Adapted from Bunton-Stasyshyn, Wells, and Teboul (2019).

3 AIMS OF THE STUDY

Human disease genetics can be studied using animal models. The selection of model organism is dictated by its similarity to human, as well as its suitability to answer the question at hand. Human and zebrafish share common immunological cell types and pathogen recognition receptors. Moreover, in the case of tuberculosis, they also share a susceptibility to mycobacterial infections. The *M. marinum* / zebrafish model has proven useful in modelling human tuberculosis on many occasions. Furthermore, zebrafish and humans also have common vertebrate development. Due to fast reproduction and development, the zebrafish offers a valuable model to study human developmental biology. In cases where a mouse model presents an embryonic lethal phenotype, the zebrafish can offer options to study gene function due to its extrauterine development. The study of zebrafish genetics has greatly advanced since the emergence of genetic engineering technologies. In this thesis, the zebrafish model was applied to studying human disease associated genes of immunological and developmental significance.

The aims of the thesis were

1) to analyse the effect of chromatin on CRISPR-Cas9 during the early development of zebrafish (Study I);

2) to generate a zebrafish model for the human developmental disease FINCA (Study II);

3) to examine the role of the *itln1* in *itln3*-knockout zebrafish model for a human tuberculosis infection (Study III);

4) to establish the role of the inflammasome adaptor gene *pycard* in a mycobacterial infection in zebrafish (Study IV).

.

4 MATERIALS AND METHODS

4.1 Zebrafish lines and maintenance

In this study, wild type zebrafish strains, AB and TL, were used as indicated in later sections. Both the fish lines were initially obtained from the Wellcome trust Sanger Institute (Hinxton, Cambridgeshire, UK). The fish were maintained at 28°C in the Tampere Zebrafish Core facility in Aquatic habitats flow through systems (Apopka, Florida, USA) at a density of 7 fish per litre, 14h/10h light/dark cycle. Embryos were maintained in standard embryo medium (5mM NaCl, 0.17 mM KCL, 0.33 mM CaCl2, 0.33 mM MgSO4, 0.00001% Methylene Blue) in a 28.5°C incubator with a similar artificial daytime setting. Larvae and adult fish were fed with either Gemma Micra (Skretting, Stavanger, Norway) or SDS (Special Diet Services, Witham, Essex, UK) feed of a granularity appropriate to their age, as well as artemia and spirulina. For infection experiments, adult fish were transferred into an Aqua Schwarz GmbH (Göttingen, Germany) flow through system in a separate laboratory for infected fish.

4.2 Ethics statement

All experiments were conducted in accordance with local and European Union (EU) animal regulations (Act on the Protection of Animals Used for Scientific or Educational Purposes 497/2013 and the EU directive on the protection of Animals used for Scientific purposes 2010/63/EU)), and for each experiment, the animal experiment permissions were obtained from the Animal Experiment board of the Regional State Administrative Agency of Southern Finland. Animal experiment permissions (ESAVI/2235/04.10.07/2015, ESAVI/10823/04.10.07/2016, ESAVI/2464/04.10.07/2017, ESAVI/11133/04.10.07/2017, and ESAVI/2776 /2019).

The wellbeing of each animal was assessed daily. Animals which showed signs of discomfort or symptoms of an infection were euthanised with an overdose of a tricaine anaesthetic (Ethyl 3-aminobenzoate methanesulfonate, MS-222) (Merck, Kenilworth, New Jersey, USA). Typically, for untreated fish these signs include unresponsiveness, redness of skin or signs of a parasitic infection (lighter, swollen areas on scales). For infection experiments and experimental *M. marinum* infections, signs of discomfort include abnormal swimming behaviour, lesions in the skin, protruding scales, and gasping.

4.3 Mutagenesis with CRISPR-Cas9

4.3.1 gRNA design and synthesis

Mutated zebrafish were generated with a type II *S. pyogenes* CRISPR-Cas9 system (Cong et al. 2013; Jinek et al. 2012). The gRNA sequences were selected based on combined results of multiple online tools, such as CHOPCHOP (https://chopchop.cbu.uib.no/) and CRISPR-Scan (https://www.crisprscan.org/) (Labun et al. 2016; Moreno-Mateos et al. 2015). The target site uniqueness was analysed using basic local alignment search tool BLAST (NCBI) (Altschul et al. 1990) against the zebrafish reference genome (GRCz10). The oligo was checked for secondary structures using online oligo analyser tools. Synthetic oligos were ordered from Thermo Scientific. The selected sequences for gRNAs are listed in Table 3. The gRNA synthesis protocol was modified from (Hruscha and Schmid 2015). Briefly, the T7 promoter oligo and the oligo containing the gRNA sequence were annealed and used as a template for the MEGAShortScript T7 *In vitro* Transcription kit (Ambion, Thermo Fisher scientific, Waltham, Massachusetts, USA). The quality of the *in vitro* transcription was analysed with agarose gel separation or with a Fragment Analyzer system (Agilent Technologies, Santa Clara, California, USA), and the concentration was measured using NanoDrop (Thermo Fisher) or a Qubit RNA Broad Range kit (Thermo Fisher).

Table 3. gRNA sequences used

Gene ID	Gene name	Target sequence 5'-3' (PAM)
ENSDARG00000024211	*ifng1-2*	TGTTTGCTGTTTTCGGGA(TGG)
ENSDARG00000054975	*cxcr2*	TTCTTCATGGACAACCGC(AGG)
ENSDARG00000041041	*cxcr3.2*	TAGTGGGCACGAGATGTT(AGG)
ENSDARG00000056499	*ca6*	GGGGAGGCTGGGACTTGG(AGG)
ENSDARG00000088143	*sema4gb*	GATCCCACTAAGGGGTACAC(TGG)
ENSDARG00000089581	*nhlrc2*	GTTCCTTCCTGAGTGACAG(CGG)
ENSDARG00000040076	*pycard*	CGTGTTCACATCAAAAGACG(CGG)

4.3.2 *In vitro* digestion assay

To confirm gRNA functionality, the guide was tested *in vitro*. For this, an amplicon of around 1kb (kilobase) around the gRNA binding site was selected for amplification with a PCR using the Maxima Hot Start polymerase (Thermo Fisher). The primers used in the *in vitro* digestion assay are listed in Table 4. The amplicon was purified using a GeneJET PCR purification kit to remove unbound nucleotides (Thermo Fisher). The Cas9 protein (ToolGen Inc. Seoul, South-Korea) and the synthetised gRNAs were incubated at 37°C in NEB3 buffer (Ipswich, Massachusetts, USA) before adding the template (final ratio of Cas9:gRNA:template 10:10:1). The digestion reaction was incubated for 3 h at 28°C. The sample was then incubated with 300U of proteinase K (Thermo Fisher) at 37°C to release the Cas9. Finally, Proteinase K was inactivated at 65°C for 10min. The cutting efficiency was analysed by separating the fragments on an agarose gel.

4.3.3 Timing of mutagenesis and heteroduplex mobility assay

In order to determine when mutagenesis can be detected for each gRNA, embryos were collected immediately after spawning, injected with CRISPR-Cas9 components, and collected for samples between 0 - 8hpf. The primers used in genotyping are listed in Table 4. DNA was extracted with ethanol precipitation method, and afterwards threated with 15 units of RNase A (Thermo Scientific) per 100µl of sample, for 1h at 37°C. 100 µl of lysis buffer (10 mM Tris (pH 8.2), 10 mM ethylenediaminetetraacetic acid (EDTA), 200 mM NaCl, 0.5% SDS) with Proteinase K (0.2 mg/ml) was added to the sample and the sample was incubated for a minimum 2h at 55°C with occasional mixing. The debris was pelleted with a 20-minute centrifugation at 16 000 g at room temperature. The supernatant was

transferred into a fresh microcentrifuge tube, and 2 volumes of ice cold 100% ethanol was added. The DNA was precipitated for minimum 30 minutes in -20°C. The sample was separated by centrifugation for 10 minutes at 16 000 g at room temperature. Supernatant was discarded and the DNA containing pellet was washed with two volumes of 70% ethanol. The DNA was pelleted with a 5-minute centrifugation at 16 000 g at room temperature. The supernatant was removed, and the pellet was allowed to dry for several minutes. The dried DNA pellet was resuspended in nuclease free water.

Table 4. Primers used for genotyping and *in vitro* digestion assay

Genotyping			
Gene name	Gene ID	Direction	Sequence (5'-3')
ca6	ENSDARG00000056499	Forward	AGCATGCAACACCTTCGGTC
		Reverse	ATTTCAGGCATAAGTCCAGC
ca10a	ENSDARG00000052644	Forward	CTGCAAATCATCCCTTTGTG
		Reverse	GTTCCTCGCATCAAAACACC
cxcr2	ENSDARG00000054975	Forward	AGTGTTTCTCCCTCCACAGC
		Reverse	TAATGCGAAGGCTCATCCTC
pycard	ENSDARG00000040076	Forward	GACCCAACTGTGAGGAACCATG
		Reverse	GCTTTCTTCAGACTTAAACGCCTTC
sema4gb	ENSDARG00000088143	Forward	GGACTCACGCCTTCAGAC
		Reverse	GCCTTATATCAGCGATGTTAC
IVDA			
Gene	Accession number		Primer sequence (5'-3')
ca6	ENSDARG00000056499	Forward	TAGTCCACGAATGCAACAGC
		Reverse	GGCATGTCTGGCACAAATAG
cxcr2	ENSDARG00000054975	Forward	AGTGTTTCTCCCTCCACAGC
		Reverse	GAGAAATCAGCAACTGGTTACG
cxcr3.2	ENSDARG00000041041	Forward	GTACTCTACTCTTCCCAGGTTTACAC
		Reverse	CTGTGCGTTTTAATCTGGCA
ifng1-2	ENSDARG00000024211	Forward	CAAAAAGATCCCCGAAGACA
		Reverse	GTGCCAGCCTCTCCTTTGTA
pycard	ENSDARG00000040076	Forward	CAGCATTTGTGAGCAGAAGC
		Reverse	AAAGGTAGATTGGGGTGTTTG
sema4gb	ENSDARG00000088143	Forward	ACCCCGCTGTGCTTACATAG
		Reverse	TCACTTTCATTCTGCCCAATC

The results were analysed with heteroduplex mobility assay. Briefly, FastAP and ExoI (Thermo Scientific) were used for the purification of an amplified PCR product. In the presence of NEB3 buffer (New England Biolabs), the PCR product was heated to 95°C and slowly cooled to room temperature. This re-annealed

product was then run on a 10% polyacrylamide gel to separate the amplicons with heteroduplexes.

4.3.4 Generating stable mutant lines with CRISPR-Cas9

After the synthesis of gRNA molecules and confirming their functionality, CRISPR-Cas9 mutagenesis was conducted. For the mutagenesis, either *Cas9* mRNA or the Cas9 protein was used. *Cas9* mRNA was ordered from ToolGen (Seoul, South-Korea). The Cas9 protein was produced in house in the Tampere facility for protein services. The gRNA and the Cas9 protein were first incubated at 37°C for 15 minutes to allow the formation of the CRISPR-Cas9 complex. The complex was then injected into a fertilized zebrafish egg with phenol red as a tracer dye in a volume of 1 - 2 nl. The detailed use of each gRNA is described in the respective studies. Borosilicate capillary needles for the injections were generated with a Flaming/Brown Micropipette puller. Injections were performed with a micro injector (PV830 Pneumatic PicoPump, World Precision Instruments) and a micromanipulator (Narishige International, London UK). For stable mutant lines, the CRISPR-Cas9 injected AB founder fish (F0) were outcrossed to TL. The F1 (AB x TL) generation are carriers of a single mutation and were screened for carriers of a frameshift mutation. Mutagenesis efficiency was confirmed with a heteroduplex mobility assay or T7 endonuclease assay in the founder generation. A typical crossing scheme is presented in Figure 7.

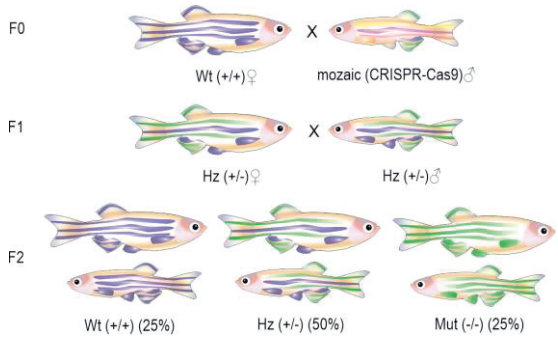


Figure 7. Schematic crossing scheme for obtaining homozygous mutant zebrafish after generating mosaic founders with CRISPR-Cas9. Chubbier fish (♀) are female and slimmer fish are male (♂). F0 is the founder generation where the mutations are generated to create a mosaic founder. F1 is the progeny of the F0 crossed to wild type (WT). F2 is the progeny of an incross of two F1 heterozygotes (Hz). This cross will result in assumed mendelian ratios of WT, Hz and homozygous mutant (Mut) zebrafish.

4.4 Genotyping of mutated zebrafish

4.4.1 Genotyping methods

A tissue sample was collected from the tail fin of adult zebrafish under anaesthesia. Zebrafish larvae were collected in tubes and kept on ice until adding the lysis buffer. An area around the gRNA binding site was amplified with PCR using either the Maxima Hot Start or the Dream Taq Hot Start DNA polymerase (Thermo Scientific). Genotyping from the PCR produced amplicon (250 – 700 bp) was done with either a heteroduplex mobility assay or a T7 endonuclease I (NEB) assay to identify mutation carriers from the mosaic F0 generation. The primers used in genotyping are listed in Table 5. The amplified sequence was purified using exonuclease I (ExoI, Thermo Scientific) and a fast alkaline phosphatase (FastAP, Thermo Scientific) incubation to remove single stranded primers and unbound nucleotides. The product was heated to 95°C and slowly cooled to 4°C to re-anneal the amplified sequence. The kinks caused by heteroduplex formation were analysed by separation with polyacrylamide gel electrophoresis on a 10% gel, or by digestion with the T7 endonuclease and subsequent agarose gel electrophoresis. Subsequent generations were genotyped either with Sanger sequencing or restriction fragment length polymorphism (RFLP). *Itln3* results were analysed by mutation specific primers.

Sanger sequencing was done with the BigDye™ Terminator v3.1 Cycle Sequencing Kit, and the run was completed by the Tampere Genomics facility. For RFLP, the mutation site was scanned for altered restriction enzyme binding sites. The amplicon was digested, and results were analysed by running the samples on agarose gel. The restriction enzyme for *pycard* was CseI/HgaI. The results were confirmed by sequencing.

Table 5. Genotyping primers for the zebrafish mutant lines

Gene (mutant)	Genotype	Direction	Sequence (5'-3')
itln3 (uta145)	WT	Forward	ATGCTAGGTTGAGGAGCATC
	Mutant	Forward	ATGCTAGGTTGAGGAGCTCG
itln3 (uta148)	WT	Forward	CTAGGTTGAGGAGCATCGCT
	Mutant	Reverse	CCGAGCTGATACTTACCTAGC
itln3	WT	Forward	GGAGCTGTCACTCCAAAGCC
	WT	Reverse	GTGGTTGATCAACCATTCAGCAC
pycard	WT & mutant	Forward	GACCCAACTGTGAGGAACCATG
		Reverse	GCTTTCTTCAGACTTAAACGCCTTC

4.5 Morpholino silencing

Morpholinos for *itln1* and *nhlrc2* were designed by the design service of the manufacturer GeneTools (Oregon, USA) upon detailed request. The morpholinos were designed to block the splice site of the transcript, to generate a non-functional protein. Morpholino sequences are listed in Table 6. Prior to experiments, the morpholino was tested at three different doses and the dose which did not result in toxicity related morphological changes was selected (Study II, Supplementary figure 14; Study III, Supplementary figure 4). The final concentrations used in each experiment are stated in the original communications II and III. Morpholinos were heated to 65°C for 10 minutes prior to injection to ensure dissolving. The *itln1* morpholino was administered with the bacterial injection solution during infection experiments. For *nhlrc2* the morpholino was injected with PBS and rhodamine dextran as a tracer dye.

Table 6. Morpholinos used

Gene ID	Gene name	Morpholino*	Sequence 5'-3'
ENSDARG00000089581	*nhlrc2*	SB	ACCCGATTCTGCTGATTTACCTTTC
ENSDARG00000089581	*nhlrc2*	TB	CTGTCAGCTTACTGTAAGACGCCAT
ENSDARG00000007534	*itln1*	SB	CTAATTCTGTACTTACTCGATTCAC
ENSDARG00000035559	*tp53*	TB	standard (Robu et al. 2007)
-	-	Random control	standard

*SB=Splice site blocking, TB=Translation blocking

4.5.1 Quantitating morpholino functionality

The effect of the morpholino on the target transcript was assayed with qPCR primers designed around the morpholino binding site, with one of the primers spanning an exon-intron boundary to prevent binding to contaminating DNA. The resulting amplicons were run on agarose gel to determine the fraction of the altered transcript in relation to the wild type transcript. The efficacy of the silencing was measured at multiple time points. The primers designed for morpholino quantitation are listed in Table 7.

Table 7. Primers for quantitation of morpholino efficiency

Gene name	Direction	Sequence (5'-3')
nhlrc2	Forward	GTCTCCAATACTGGGCAGGTG
	Reverse	GCCTGCTAATGTTGAGACTTTTCC
itln1	Forward	ATGATGCAGTCAGCTGGTTTTCTTCTG
	Reverse	GCAGTGACCGACTCTGGAAATTCTCC

4.6 Gene expression analyses

4.6.1 RNA extraction

RNA was extracted from whole larvae for gene expression analyses using RNeasy mini Plus kit (Qiagen, Hilden, Germany). 2-mercaptoethanol was used in the lysis buffer to denature RNases. The extraction was completed according to the manufacturer's instructions. RNA was resuspended in nuclease free water.

Adult fish kidneys were extracted for an RNA-seq analysis with the same RNeasy mini plus kit, however the kidneys were homogenized prior to extraction. Kidneys were homogenized in screw cap tubes with an RNA preserving buffer with 6 ceramic beads (2.8mm ∅, Omni International, GA, USA). Homogenization was done with 6.5m /s for 2 x 30 seconds on dry ice with a FastPrep-24™ 5G bead beating grinder and lysis system (MP Biomedicals LCC, CA, USA). For the analysis of adult fish organs, DNA removal columns were not used, instead contaminating genomic DNA was removed with a Rapid Out DNA removal kit (Thermo Scientific).

4.6.2 Gene expression analysis with qPCR

For quantitative PCR (qPCR), complementary DNA (cDNA) was synthetised with a SensiFast cDNA synthesis kit (Bioline Meridian bioscience, OH, USA). A gene expression analysis was performed with the PowerUp Sybr Green master mix (ThermoFisher Scientific) with a CFX96 Real-time PCR detection system (BioRad). Cycle threshold (Ct) values were normalized to the Ct values of the reference gene eukaryotic translation elongation factor 1α, *eef1a*, transcript, and amplicon sizes were confirmed by separating them on an agarose gel. The cycling parameters of the qPCR

programme were 50°C 02:00 (mm:ss), 95°C 02:00, 40 cycles of 95°C 00:03 and 60°C 00:30, followed by a melt curve from 55°C to 95°C at a 0.5°C increment. All samples were measured at least in duplicate. qPCR primers are listed in Table 8.

Table 8. qPCR primers for gene expression in zebrafish

Gene	ENSEMBL gene ID	Direction	Sequence (5'-3')
pycard	ENSDARG00000040076	Forward	AAGTCTGCAATCGAAAAGCTG
		Reverse	CTTTGCTTTCTGATTGCCCT
eef1a	ENSDARG00000020850	Forward	CTGGAGGCCAGCTCAAACA (1)
		Reverse	ATCAAGAAGAGTAGTACCGCTAGCATTAC (1)

(1) (Tang et al. 2007)

4.6.3 Transcriptome analysis with RNA-seq

Zebrafish were infected with a low dose of *M. marinum*. At 4 wpi, the fish were euthanised with an overdose of anaesthetic. The kidneys were extracted and homogenized for RNA-extraction. RNA-extraction was performed with the RNeasy Mini Plus kit as detailed in the previous section. The RNA quality was assessed with a Fragment Analyzer instrument (Agilent Technologies, Santa Clara, California, USA). Samples were sent for analysis to the Finnish Functional Genomics Centre at Turku Bioscience. The run was performed with NovaSeq 6000 S4 v1.5. Paired end sequencing of 100 bp was completed at 20 million reads per sample depth.

4.6.4 RNA-seq data analysis

The quality of the RNA-sequencing data was analysed with FastQC version 0.11.7 (Andrews 2010). The alignment was done with STAR aligner version 2.5.3a in the ENSEMBL reference genome GRCz11.104 (Dobin et al. 2013; Hubbard et al. 2002). Differentially expressed genes were determined with DESeq version 1.24.0 (Love, Huber, and Anders 2014) using R (v.3.6.1). Differential expression was set at $P<0.05$ after adjustment for multiple testing, with a log2 fold change >1 and with the absolute median difference of library-size normalized read counts >13 between the two conditions.

4.7 Infections with *M. marinum*

M. marinum (ATCC 972, Aronson strain) was cultured on BD Difco Middlebrook 7H10 plates (BD Biosciences, NJ, USA) and grown at 29°C protected from light. For infection experiments, *M. marinum* was inoculated in 7H9 media (BD Biosciences) with 0.2% Tween80 (Merck). After 72 hours, the culture was passaged to OD600~0.07 and allowed to reach the logarithmic growth phase.

For infections, 1 ml of culture at the logarithmic growth phase was pelleted. The bacterial pellet was then resuspended in a volume corresponding to approximately 30 CFU / 5 μl in PBS containing phenol red as tracer dye. For embryonic zebrafish infections, 2% polyvinylpyrrolidone was added to the solution. With adult fish, bacteria were injected intraperitoneally into the organ cavity using a 30G needle, under Tricaine anaesthesia. Infected adult fish were followed for 24 weeks for survival experiments, for 4 weeks for bacterial burden analyses, for 4 weeks for the RNA-seq analysis of gene expression and 8 weeks for histopathology. The wellbeing of the fish was assessed daily. Embryos were injected into the yolk sack using a microinjector system as described before. The injection dose was estimated by plating the injection mix onto 7H10 plates. Embryos were dechorionated at 1 dpf and transferred into 24-well plates with embryonic medium. Embryos and larvae were followed until their death, with a daily assessment of their wellbeing.

4.7.1 Survival experiment

The adult offspring of heterozygous fish, born at approximately mendelian ratios of each genotype (25% of WT, 50% heterozygous, 25% mutant) and equal numbers of both sexes were infected with a low dose of *M. marinum* as described. The wellbeing of the fish was assessed daily, and fish presenting any sign of infection were euthanised according to the humane end point criteria. Euthanised fish were genotyped after the experiment. The adult fish survival experiment was terminated at 24 weeks, or when 40% of the fish had died or been euthanised.

4.7.2 Bacterial burden assay

Bacteria in the organ block of the fish were analysed at 4 wpi using a qPCR-based analysis according to a previously described protocol (Parikka et al. 2012). Briefly, organ block samples were homogenized with ceramic beads as described in section

4.8. DNA was extracted using the Tri-reagent (MRC Inc. OH, USA). 1 μg of DNA was used for analysis with a Sensifast Sybr No-ROX kit (Bioline) against a dilution series of a sample with a known bacterial concentration. Primers used in the quantitation were Mmits F: 5'-CACCACGAGAAACACTCCAA-3', Mmits R: 5'-ACATCCCGAAACCAACAGAG-3' (Parikka et al. 2012). The detection limit was set to 100 CFU as samples with fewer bacteria do not give consistent results, and samples measuring below 100 CFU were given an arbitrary value of 0 (mutant) or 100 CFU (wild type) to minimise the bias in the statistical analysis.

4.7.3 Flow cytometry and cell sorting

Adult fish were euthanised with an overdose of anaesthetic and decapitated. The kidneys were released with tweezers and resuspended in 100 μl PBS with 0.5% foetal bovine serum (FBS) (Merck) and kept on ice. The kidney was homogenized by pipetting and passed through a 35μM cell strainer. Flow cytometry experiments were completed with FACS Arya Fusion (BD biosciences). The viability stain FVS510 (BD Biosciences) was used for excluding dead cells. Cells were gated according to their size (side scatter SSC) and granularity (forward scatter FSC) according to the data by Langenau et al. (2004) and Traver et al. (2003). 20 000 events were recorded per sample. Sorted cells were kept on ice and pelleted after the flow cytometry was completed. RNA was extracted with an RNeasy Micro plus kit (Qiagen). The flow cytometry results were analysed with FlowJo 10.7.1.

4.7.4 Histology

Adult zebrafish were infected with a low dose of *M. marinum* and euthanised 4 wpi with an overdose of anaesthetic. Fish were injected with 60μg of pimonidazole hydrochloride during terminal anaesthesia and were decapitated 10 minutes after injection. The fins, tail and the head above gills were removed. The fish were fixed in 10% phosphate buffered formalin at room temperature for 7 days. The fish were then decalcified with 0.5M EDTA pH 8.0 for 7 days. After, the torsos were rinsed under running tap water for 2h and the samples were incubated in 70% ethanol overnight at room temperature. The samples were cast into paraffin blocks and cut with a Leica Microtome SM2010R. 4 x 5μm sections were collected every 200μm on StarFrost advanced adhesive glass slides (Waldemar Knittel Glasbearbeitungs

GmbH, Braunschweig, Germany). The samples were deparaffinised and stained for analysis. The glasses were stained either with Ziehl-Neelsen or Mallory's trichrome staining according to standard protocols. For the hypoxia stain, a Hypoxyprobe-1 kit was used (Hypoxyprobe, Burlington, MA, USA). After staining the slides were dehydrated and embedded with DPX mountant for histology (Merck). The slides were analysed with a Hamamatsu NanoZoomer S60 digital slide scanner and analysed with NDPview v.2.8.24. The granuloma size, location and structure of granulomas, and the result of the hypoxic stain were recorded.

4.8 Statistical analysis

The number of animals required for the survival and bacterial burden experiment was estimated based on our previous data or pilot experiments. Power calculations were done with the online tool (https://clincalc.com/stats/samplesize.aspx). For example, the required group size for the *pycard* mutant survival experiment was estimated to be 19 fish (α=0.05, β =0.8, difference between groups for outcome 40%). As the experiment was done blinded, a moderately higher group size of 25 was selected so that it contained a sufficient number of each genotype and allowed for the loss of several fish during the infection procedure. Statistical analyses were done with Graph Pad Prism V5.02. For survival curves, the Log-Rank test was used. For other analyses including gene expression, the U-test was used. A linear mixed model was used for the analysis of granuloma size. The model included a fixed effect for genotype and a random effect for individual fish. The statistical analysis was done using a linear mixed model R-package (R package: lme4).

5 RESULTS

5.1 CRISPR-Cas9 efficiency is affected by closed chromatin

5.1.1 gRNAs which function *in vitro* do not always function *in vivo*

The functionality of a gRNA is determined by its folding and sequence specificity. Testing the functionality of gRNAs *in vitro* can tell if the gRNA is able to cleave its PCR amplified genomic target region when there is no interference from off-target binding sites or chromatin. In turn, when the target is located in the genomic DNA, the gRNA-Cas9 complex has to search for its highest affinity binding site among many available candidates. As shown by Knight et al. (2015), targets buried in closed heterochromatin take more time to be found than the targets located in open euchromatin. To experimentally investigate whether chromatin organization affects CRISPR-Cas9 mutagenesis in zebrafish larvae *in vivo*, a 1kb sequence around the gRNA binding site was amplified. The cutting efficiency was evaluated by separating the digested fragments by an agarose gel electrophoresis. This was then compared with the activity of the same gRNA cleavage efficiency *in vivo* in zebrafish larvae. The results show that for three selected gRNAs, the activity detected *in vivo* can differ from that observed *in vitro*, where chromatin is missing (Study I, Figure 1).

5.1.2 The onset of gRNA function is different between gRNAs

Several gRNA molecules that target genes, which are differentially expressed during development were analysed for the onset of mutagenesis. Prior to injection, Cas9 was incubated with the gRNA to promote complex formation. The results show that gRNAs can differ in their onset of mutagenesis. This can, however, be affected by target sequence specificity and affinity, which is different for each gRNA. The stability of the R-loop formed by the genomic DNA and the gRNA bound to Cas9

has been found to strongly correlate with the cleavage efficiency of Cas9 (Xu, Duan, and Chen 2017) Thus, the value of the result is merely that it demonstrates that not all gRNAs are equally efficient or act at a same rate (Study I, Figure 2).

5.2 The number of degrading cells is increased in the brains of *nhlrc2* morphants

5.2.1 Successful knockdown of *nhlrc2* with morpholinos

Morpholinos were designed against the translation start site and the splicing site at the beginning of exon 4 of the *nhlrc2* sequence. As the translation blocking morpholino does not cause changes in the transcript sequence, its effect can only be measured at the protein level and it cannot be quantitated with qPCR-based methods. The splice site blocking morpholino alters the transcript, so its effect can be quantitated using qPCR. Primers for the quantitation were placed in exon 3 and exon 5 (Figure 8). The qPCR product was separated using agarose gel electrophoresis, which clearly shows that only exon exclusion seems to take place (Figure 8). Based on the results, knockdown of *nhlrc2* using the splice site blocking morpholino was successful.

5.2.2 *nhlrc2* morphants display alterations in the midbrain region

Morphant fish were imaged through the first 7 days of development but marked changes in their appearance were not seen (Study II, Supplementary figure 14). However, the fish did appear lethargic and unresponsive, despite not displaying differences in their ability to break through the chorion (Study II, Supplementary figure 15). As FINCA patients displayed changes in their brain structure, the morphant brain was also analysed using transmission electron microscopy (TEM). As the effect of the translation blocking morpholino could not be quantitated, a splice site blocking morpholino was used for the experiment. Morphants were analysed with TEM at 2 dpf, and the analysis of the midbrain region revealed an increased number of disintegrating cells or cells with an accumulation of vacuoles in comparison to controls injected with random control morpholino (Figure 9).

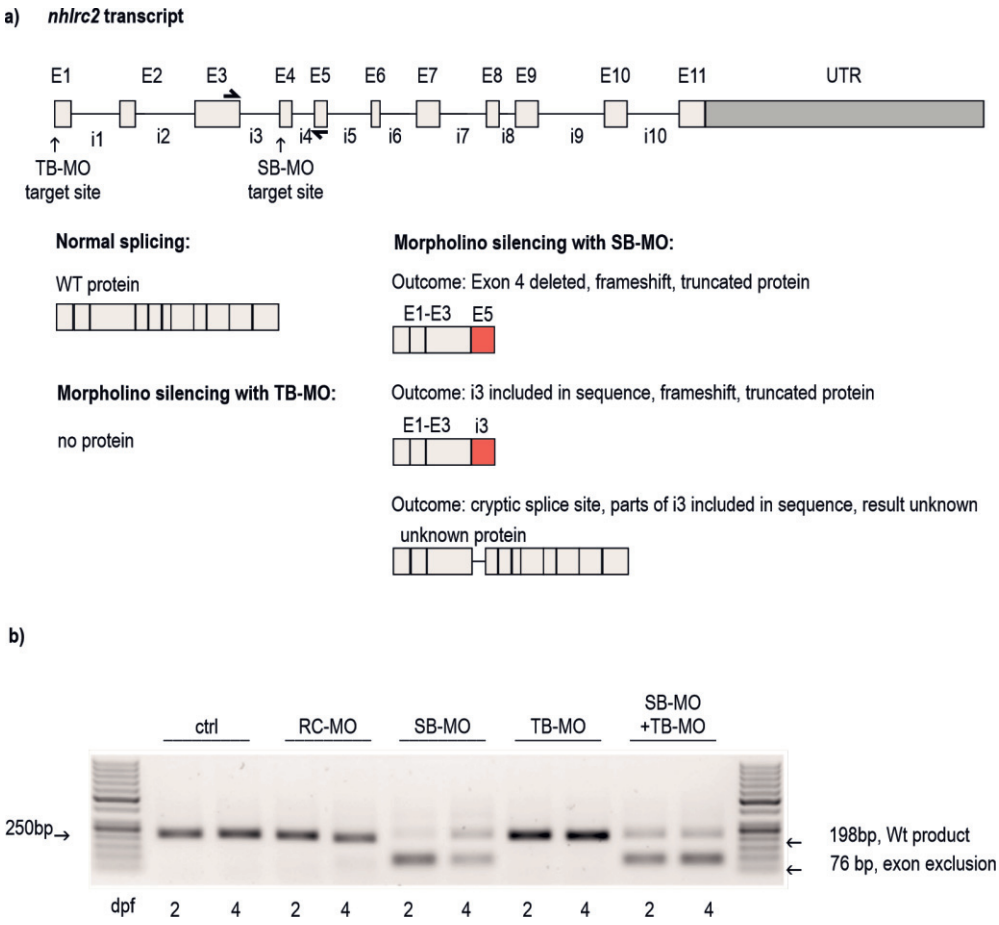


Figure 8. Effects of morpholino silencing on *nhlrc2* mRNA. a) The *nhlrc2* transcript is depicted with light grey boxes for exons (E) and lines for introns (i) numbered from 5' to 3', as well as the untranslated region (UTR) in dark grey. Two morpholinos (MO) were designed to prevent normal translation, a translation blocking MO (TB-MO) and a splice site blocking (SB-MO). Half-arrows indicate the target sites for the primers for quantitation. TB-MO binds to the translation start site. SB-MO binds to the i3-E4 boundary, where it prevents normal splicing. SB-MO can cause either exclusion of e4, or inclusion of i3, which both lead to frameshift and a truncated protein. b) Quantitation of the SB-MO effect with qPCR indicates that exon exclusion takes place, leading to a shorter 76bp product. ctrl = untreated control. Part a) has been drawn from nhlrc2 Ensembl ID: ENSDARG00000089581. Part b) of the figure has been adapted from Study II Supplementary figure 14.

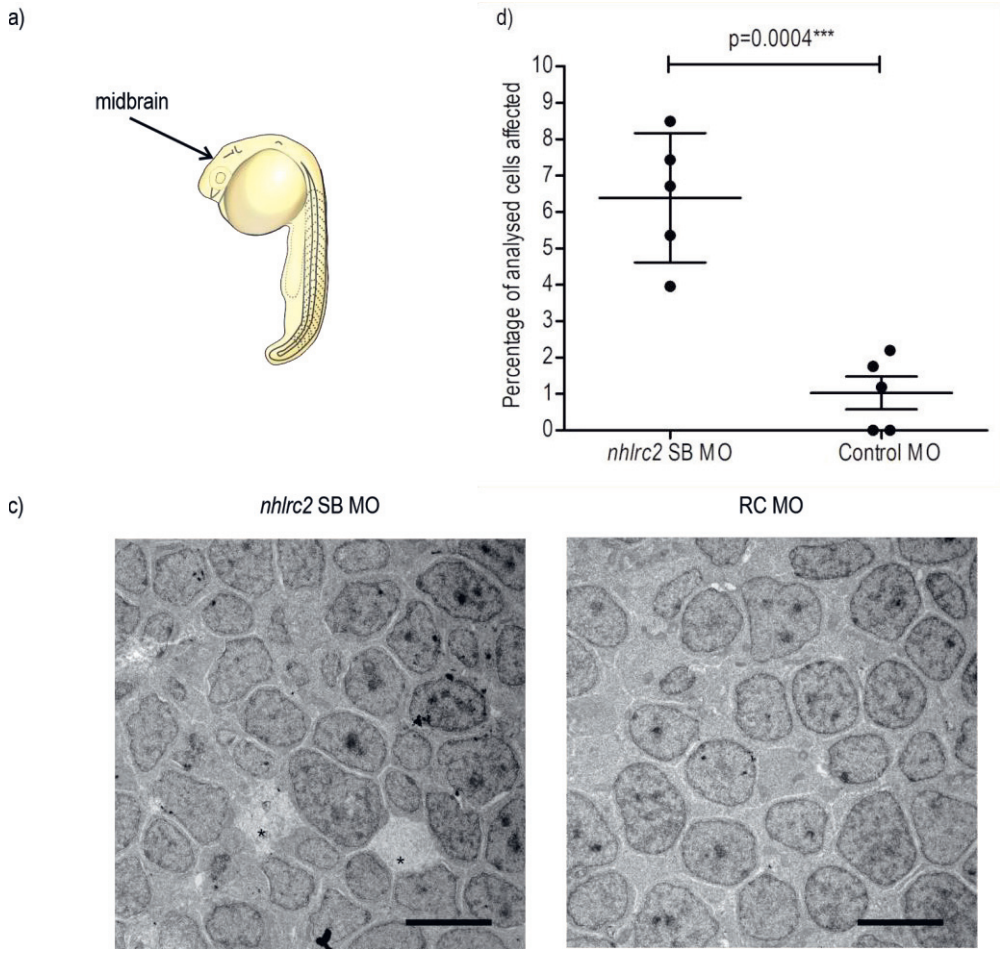


Figure 9. The knockdown of *nhlrc2* increases the number of degrading cells and cells with increased vacuole formation. a) The mid brain region of morphants (indicated) was analysed with TEM. b) TEM images from five *nhlrc2* morphant and five random control injected morphants showed that *nhlrc2* morphants have a greater number of cells that are affected (5,8%, 47 of 810 total cells) in comparison with (1%, 9 of 850 cells). The central line indicates the mean and whiskers the standard deviation. The p-value has been determined with the χ^2 test. c) Exemplar images of the results, with stars indicating degrading or vacuolarised cells from the same data presented in original publication II. The black bar depicts a 5µm scale bar.

5.3 Knockdown of *itln1* does not impair the survival of *itln3*-mutants

5.3.1 *itln1* morpholino silencing reduced transcript levels

A morpholino injection causes a reduction in the levels of the *itln1* transcript (Figure 10), and the effect is lost with time, as the morpholino gets diluted between dividing cells. Transcripts were measured with qPCR and the results were verified on agarose gel, which also indicates that only the exon skipping product is observed on the gel (Figure 10). In comparison with the controls injected with a random control morpholino, the reduced amount of transcript was visible both in the Ct values as well as in the gel run (Figure 10).

5.3.2 *itln1* morpholino knockdown does not affect the survival of *itln3* knockout zebrafish

As the *itln3* mutants did not show a phenotype in our survival experiments, it could be that another gene of the same family of genes compensates for the loss of *itln3* in the CRISPR-Cas9 mutated fish as suggested by El-Brolosy et al. 2019 and Rossi et al. 2015. As *itln1* is induced at the early points of a mycobacterial infection in a similar manner as *itln3* in larval zebrafish, it is a potential candidate to compensate for the loss of *itln3* in the knockout fish. To test this, *itln1* was silenced with morpholinos. The offspring of heterozygous parents were infected without prior knowledge of their genotypes, and the genotypes were analysed *post-mortem*. A significant change was observed in the survival of the *itln3* mutated fish upon silencing *itln1* expression with morpholinos. However, this was seen in only one of the two mutant lines, and only in WT and knockout larvae (Figure 11). A later analysis also revealed that knockout of *itln3* does not alter the transcript levels of the other *itln*-family members (Study III, Supplementary Figure 2).

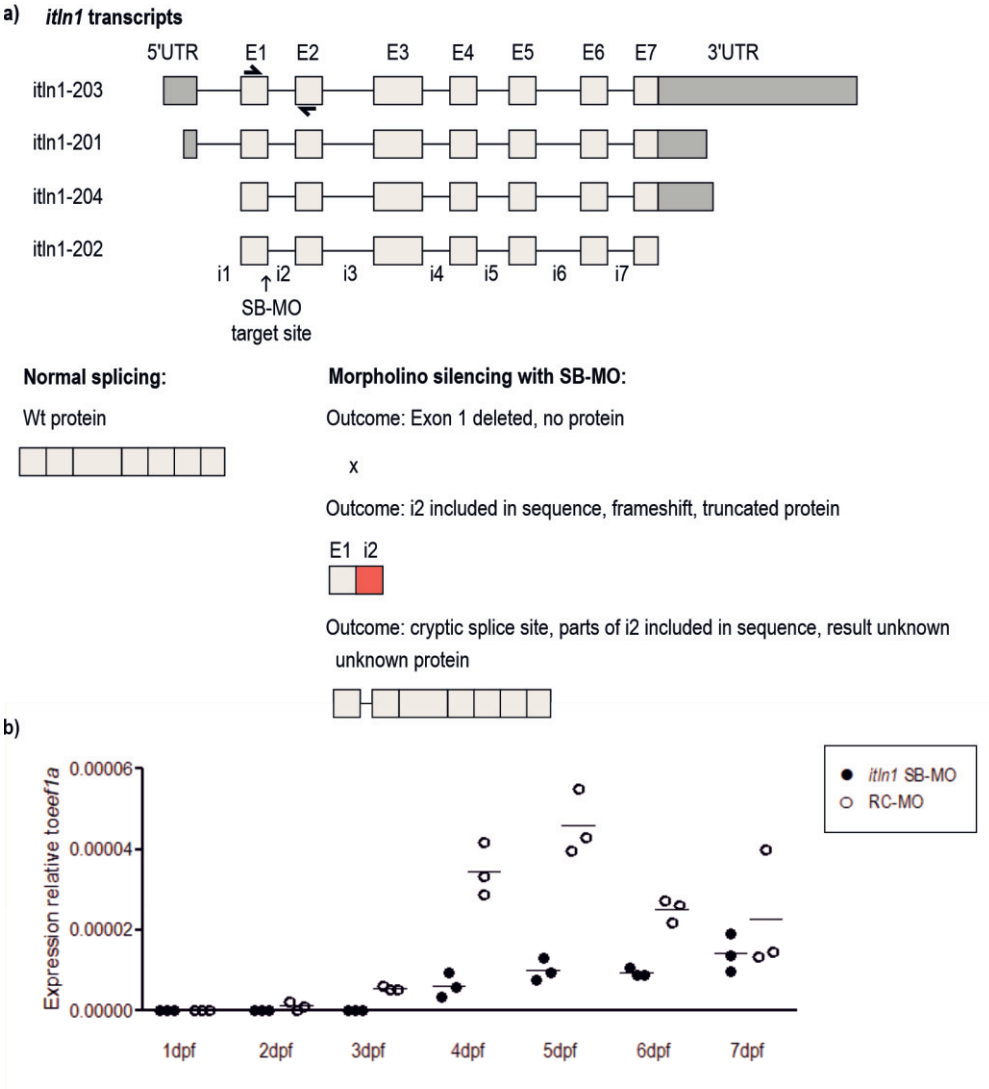


Figure 10. The effect of morpholino silencing on itln1 transcript levels. a) A schematic presentation of the 4 itln1 transcripts. Dark grey boxes indicate the untranslated regions (UTR), light grey boxes exons (E1-E7) and lines introns (i1-i7, not drawn to scale). The target site for the splice site blocking morpholino is indicated, with half-arrows indicating primer binding sites for measuring morpholino efficacy. Normal splicing and the three predicted outcomes of morpholino binding are presented. b) qPCR measurements of itln1 transcript in comparison to random control morpholino injected controls (normalized to *eef1a*). The figure has been modified from Study III. Supplementary Figure 4 and Figure 5)

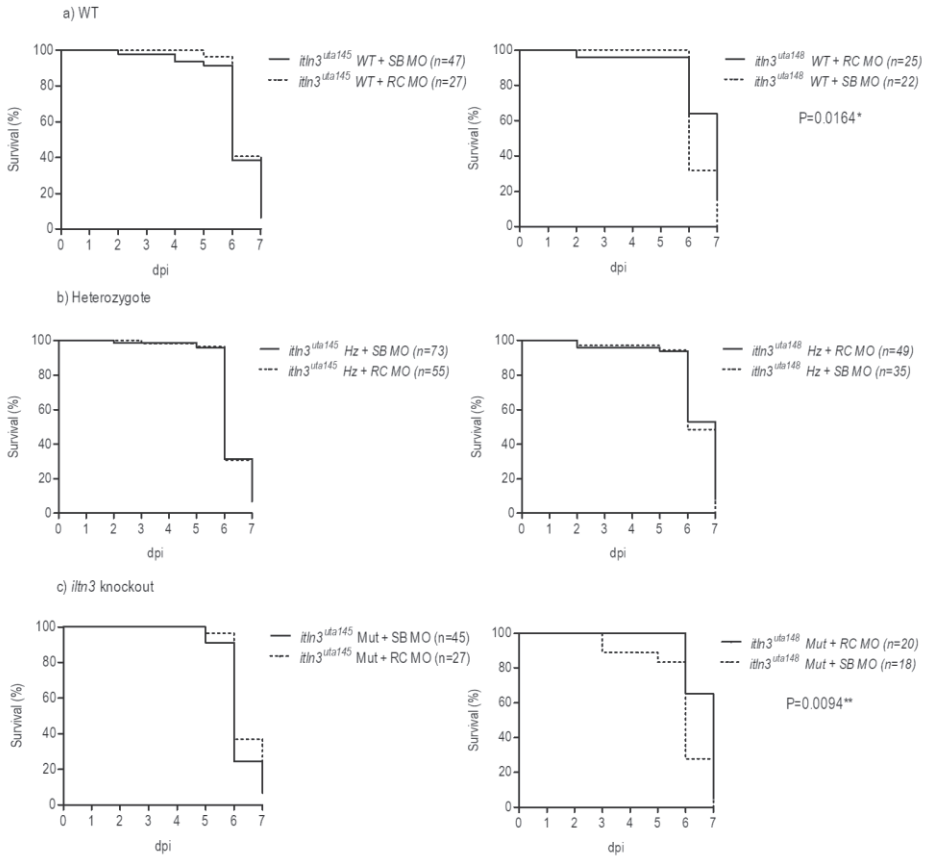


Figure 11. Silencing *itln1* does not alter *itln3*-mutant survival in a *M. marinum* infection. Data is represented for each genotype a) WT, b) heterozygous (Hz) and c) knockout (Mut) with left graphs presenting the *itln3*^{uta145} and the right side *itln*^{uta148} line. When infected with a low dose of *M. marinum* and the morpholino (either random control, RC, or *iltn1* SB MO) *iltn3*$^{uta145/uta145}$ survival is not altered for any genotype a-c. The survival of both *iltn3*$^{+/+}$ and *iltn3*$^{uta148/uta148}$ is significantly improved in comparison to RC MO injected controls, whereas the survival of *iltn3*$^{uta148/+}$ remains unaltered a-c. As the result was not repeated with both strains, it was concluded not to be significant. The data has been modified from Study III, Figure 5. Study III. dpi = days post infection

5.4 Pycard is essential for the defence against an *M. marinum* infection in zebrafish

5.4.1 *pycard*$^{-/-}$ knockout mutants show a reduction in transcript levels

To determine, if *pycard* is essential for the immune defence in zebrafish, two mutant zebrafish lines were generated using CRISPR-Cas9 mutagenesis. The success of the knockout was determined by measuring the transcript levels in zebrafish larvae at multiple timepoints. *pycard*$^{-/-}$ fish display reduced levels of transcripts at each point measured. As small amount of RNA was present at each timepoint, it can be concluded that the aberrant mRNA product produced as a result of the frameshift mutation is a potential trigger for compensatory effects. If a functional antibody was available, the presence or absence of the protein could be analysed.

Sanger sequencing revealed that the selected mutants carry two types of frameshift causing mutations, a (Study IV, Supplementary figure 1). When translated, the mutated sequence results in a truncated protein product, as displayed in Figure 12 b.

5.4.2 *pycard*$^{-/-}$ larvae do not show altered susceptibility to an *M. marinum* infection

The embryos from both knockout strains from two heterozygous parents were infected with a low dose of *M. marinum* to determine if their survival is impaired upon a mycobacterial infection. As the parent fish were heterozygotes, the offspring was 25% WT (*pycard*$^{+/+}$), 50% heterozygous (*pycard*$^{tpu4/+}$ and *pycard*$^{tpu5/+}$) and 25% homozygous knockout (*pycard*$^{tpu4/tpu4}$ and *pycard*$^{tpu5/tpu5}$). Larvae were genotyped postmortem using RFLP. As no difference was observed between genotypes, the maternal expression of *pycard* either as an mRNA or a protein could be masking the phenotype by protecting the embryos from infection. Harvey et al. results suggest that the *pycard* mRNA is not present in the developing embryo before 9 hpf, however they did not measure protein levels. The same experiment was done with the F3 generation, which were either offspring of homozygous or WT parents. As the parent fish (*pycard* knockout homozygous) do not have a functional copy of the gene, the offspring should not have a maternal transcript or protein, which could protect them from the infection. However, there was no change in the results. Therefore,

the survival of the larvae from an *M. marinum* infection is not affected by the lack of a functional *pycard* (Study IV, Figure 1).

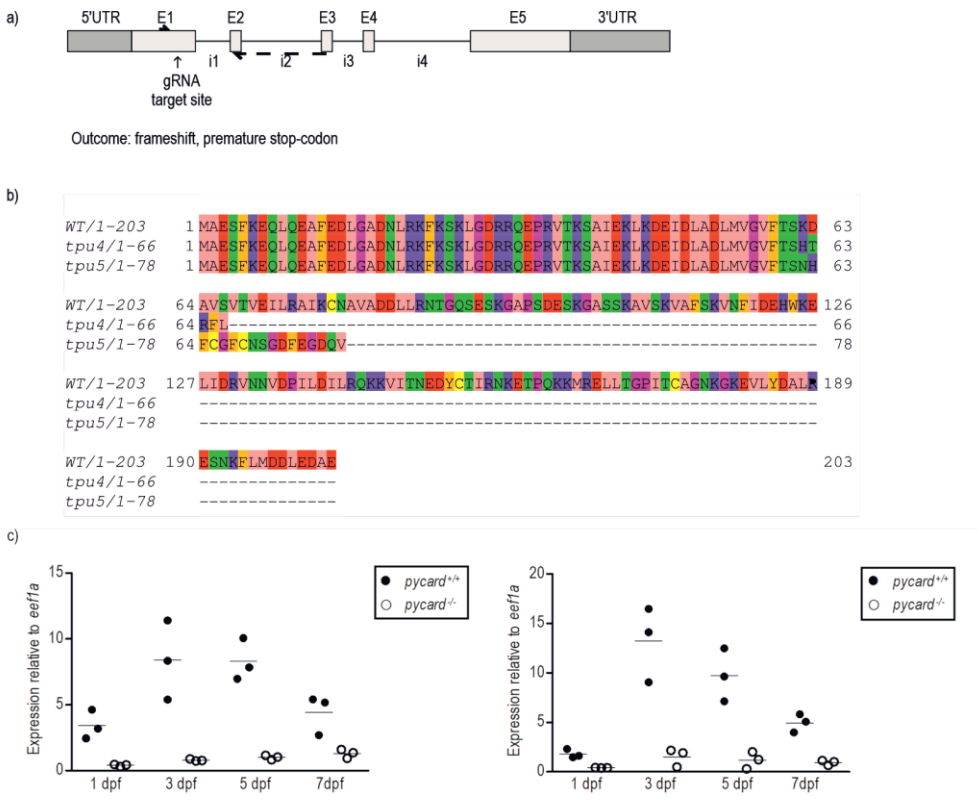


Figure 12. Knockout of *pycard* using CRISPR-Cas9 results in reduced transcript levels. a) The structure of the gene *pycard* where untranslated regions (UTR) are indicated with dark grey and exons (E1-E5) in lighter grey. Introns are indicated by lines joining exons and have not been drawn to scale. Gene structure has been adapted from Ensembl ID: ENSDARG00000040076. The gRNA binding site is indicated with an arrow and a label. Horizontal half arrows above and below exons indicate primer binding sites for transcript quantitation, with the dashed arrow line indicating the intron spanning primer. The resulting premature stop codon and the amino acid sequences for both selected mutants are shown in b) the alignment of the proteins generated with Clustal Omega (Sievers et al. 2011). Colours indicate amino acid properties (Zappo theme (Salmon is aliphatic or hydrophobic, orange is aromatic, blue is positive charge, red is negative charge, green is hydrophilic, purple is conformationally special and yellow is cysteine)). c) The result of the transcript quantitation in both mutant lines *pycard*^{tpu5} *and pycard*^{tpu4} measured in larval zebrafish at 1, 3, 5, and 7 days post fertilization (dpf). Part c) has been adapted from Manuscript IV, Supplementary figure 1.

5.4.3 *pycard*$^{-/-}$ adult zebrafish display an increased susceptibility to *M. marinum*

Adult zebrafish were infected with *M. marinum* to determine if *pycard* plays a role in the adaptive immune response during an infection. In the first experiment, 66 adult fish (offspring of heterozygous parents *pycard*$^{tpu5/+}$) were infected without knowing the genotypes of the infected fish. This was done to minimize the experimental bias caused by the researcher. As with the larval experiments, the fish were genotyped *post-mortem* using RFLP with CseI. A clear impaired survival was observed in the knockout *pycard*$^{tpu5/tpu5}$ fish, but the group size was too small (14 *pycard*$^{+/+}$ to 22 *pycard*$^{tpu5/tpu5}$) to reach statistical significance. Heterozygous and WT fish did not show difference in survival. When comparing the heterozygous group and WT group together against the knockout mutant group, a significant difference was observed (Study IV, Figure 2). In the repeated experiment with a larger group size, a significant difference between WT and knockout *pycard*$^{tpu5/tpu5}$ fish zebrafish was observed, indicating that the lack of a functional *pycard* impairs the survival of the adult zebrafish. As this effect was not seen in larvae, the results suggests that *pycard* could be essential in mediating the adaptive immune response against *M. marinum* instead of the innate response.

5.4.4 *pycard*$^{-/-}$ adult zebrafish present increased bacterial burden

In order to determine whether the resistance or tolerance to *M. marinum* was decreased in *pycard*$^{-/-}$ mutants, it was necessary to determine if the number of bacteria was higher in the mutants compared to WT fish. Fish from both strains were infected with a low dose of bacteria and the bacterial burden in the whole organ block at 4 weeks post infection (wpi) was determined with a qPCR-based method. The survival and the bacterial burden of the fish is dependent on the infection dose (Parikka et al. 2012). Thus, the load of bacteria in the organs reflects the ability of the fish to contain the progression of the infection, as a higher number of bacteria causes the fish to succumb faster to the infection. Infected *pycard*$^{-/-}$ knockout fish display a sight, however, significant increase in the number of bacteria detected in the whole organ block compared to WT fish (Study IV, Figure 2). This indicates that the immune system of the knockout fish is unable to contain the bacterial growth.

When divided by sex, a slight increase in the bacterial burden was observed in the knockout (*pycard*$^{tpu5/tpu5}$) female fish in comparison to WT male fish. However, a two-way ANOVA showed that sex did not have a statistically significant effect on

bacterial burden (p = 0.6284 after log transformation). (Figure 13). The slight increase of burden in females could be a result of the larger abdomen of female fish in general. If significant sex-specific differences between groups were observed, it would be sensible to investigate as to why this is and how it affects results in experiments where fish of both sexes were included. Unfortunately, sex of the fish was not recorded for all experiments, so no conclusive results can be drawn from this limited data.

Mycobacteria reside in granulomas, and thus it could be that the increase in the bacterial burden represents a change in the structure of the granulomas of the knockout fish.

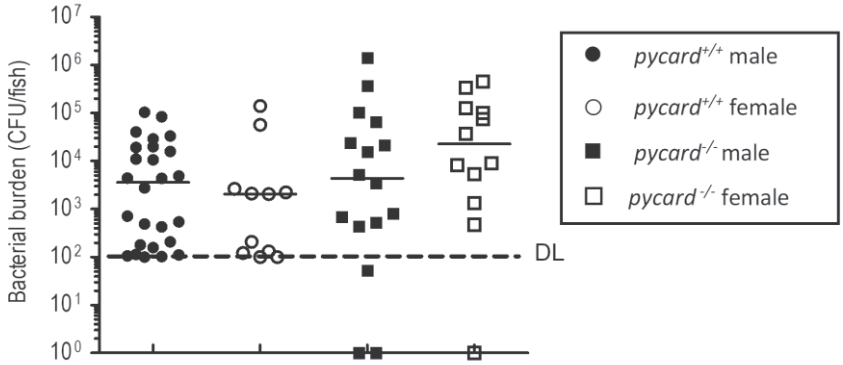


Figure 13. Example of a burden experiment data divided by sex. The line indicates median. Samples below detection limit (DL) were assigned a value of 100 CFU for WT and 0 CFU for *pycard*$^{-/-}$ for statistical analyses. The difference between sexes was not significant (p=0.6284) according to a two-way ANOVA analysis after log transformation. The data has been modified from Study IV, Supplementary figure 3.

5.4.5 Histological analysis reveals larger granulomas in *pycard*$^{-/-}$ fish in comparison to WT siblings

Pycard-knockout mice cannot contain mycobacteria in granulomas, and a higher number of bacteria are found outside granulomas in the lungs of the knockout mice than in WT mice (McElvania Tekippe et al. 2010). A likely explanation for the compromised survival and higher bacterial burden of the mutated *pycard* is that the mutant fish are unable to contain bacteria in granulomas. For a histological analysis, adult zebrafish (*pycard*$^{tpu4/tpu4}$ and WT siblings) were infected with a low dose of *M. marinum*. At 8 wpi, the fish were prepared for histological sectioning. Three different

histological stains were used: the Ziehl-Neelsen stain allows the detection of acid-fast bacteria in tissue sections; Mallory's trichrome stain colours collagenous structures and allows the visualization of the cuff around capsulated granulomas; lastly, the hypoxia staining shows if the granulomas have a necrotic, hypoxic core. Granuloma structure, bacterial content and hypoxicity were evaluated from the tissue sections stained every 200 μm through the fish. Granulomas which were not round were measured in two directions and the average of the two was recorded (Figure 14). Using the linear mixed models R-package (lme4), the granulomas in the KO fish were determined larger than those observed in the WT siblings ($p = 0.0217^*$) (Study IV, Figure 5). Differences in hypoxicity, the bacterial content, progression or number of granulomas per fish were not observed (Study IV, Supplementary figure 6).

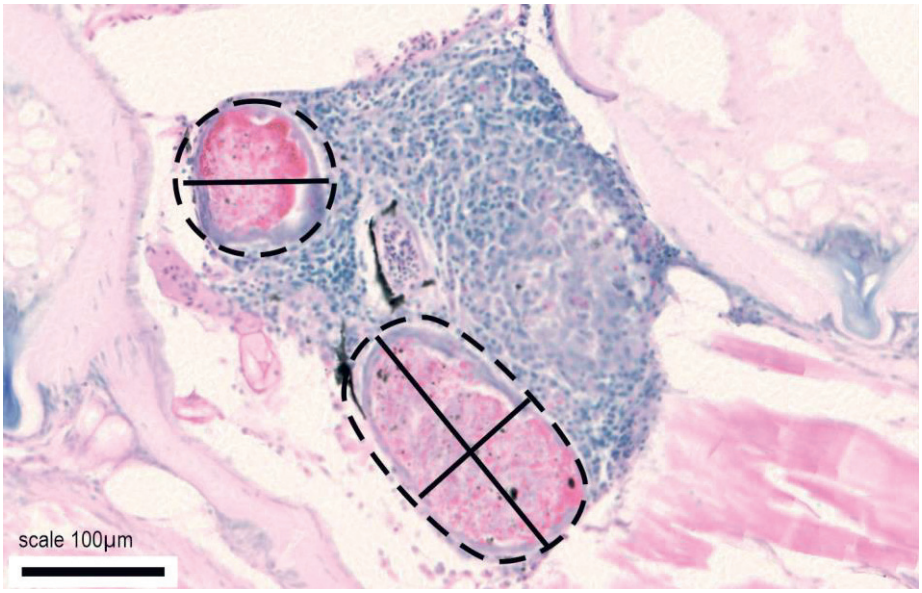


Figure 14. Granuloma diameter measurements. For round granulomas (upper left) the diameter was measured as the circle radius (solid black line crossing the dashed line). For oblong granulomas two antiparallel measurements were taken and their average was recorded. Granuloma boundaries are shown by dashed lines, and the diameter measurement by a solid line. A scale bar of 100μm is shown in the bottom left corner.

5.4.6 RNA-seq reveals a number of differentially regulated genes in the *pycard* $^{tpu5/tpu5}$ fish

In order to determine which changes in the mutants were responsible for the phenotype, RNA-seq was applied to analyse the differences between *pycard*$^{-/-}$ and WT zebrafish at the 4-week timepoint during an *M. marinum* infection. Control fish were injected with PBS.

Analysis of WT *pycard*$^{+/+}$ and knockout *pycard*$^{tpu5/tpu5}$ PBS injected fish revealed that the expression of only 4 genes was significantly altered. The expression of *pycard, akr1a1a* and *tmem176l.3a* was downregulated, whereas the expression of *nocta* was upregulated in the mutant fish. Between infected WT and *pycard*$^{tpu5/tpu5}$ fish, the expression of 124 genes were significantly altered (Study IV, Supplementary table 1). Of these, 22 were upregulated in *pycard*$^{tpu5/tpu5}$ fish and 102 downregulated, when compared to WT fish. Many immunologically significant genes were on this list, including the mutated gene *pycard* being the second most downregulated gene. 12 unnamed genes with no known function were also on this list (Study IV, Supplementary table 1).

In humans, the NLRP3 inflammasome is thought to mainly act in macrophages and other leukocytes, but also in epithelial cells (Martinon, Mayor, and Tschopp 2009). As open access single cell data are available for zebrafish, the RNA-seq data for the differentially expressed genes for *pycard*$^{tpu5/tpu5}$ was analysed with reference to the cell type. One of the significantly downregulated genes in our mutants was lysozyme ortholog *lyz*. We did not see a significant change in the myeloperoxidase homologue (*mpx*), which is another neutrophil cell marker gene (Figure 16). In line with the reduction of *lyz*, most of the differentially expressed, downregulated genes in our mutated fish are expressed in neutrophils, according to the single cell data by Tang et al. (2017) (Figure 15). In addition, we observed dysregulated expression of a number of myelopoiesis or haematopoiesis related transcription factors (*runx3*, *csrnp1a*, *rfn145b*, *cbx7* and *klf9* (Figure 16). Moreover, Tyrkalska et al. (2019) suggested that the inflammasome is essential in regulating the balance of *gata1a* and *spi1b* in relation to haematopoiesis. In our data we observed the downregulation of *spi1b* in the *pycard*$^{-/-}$ fish.

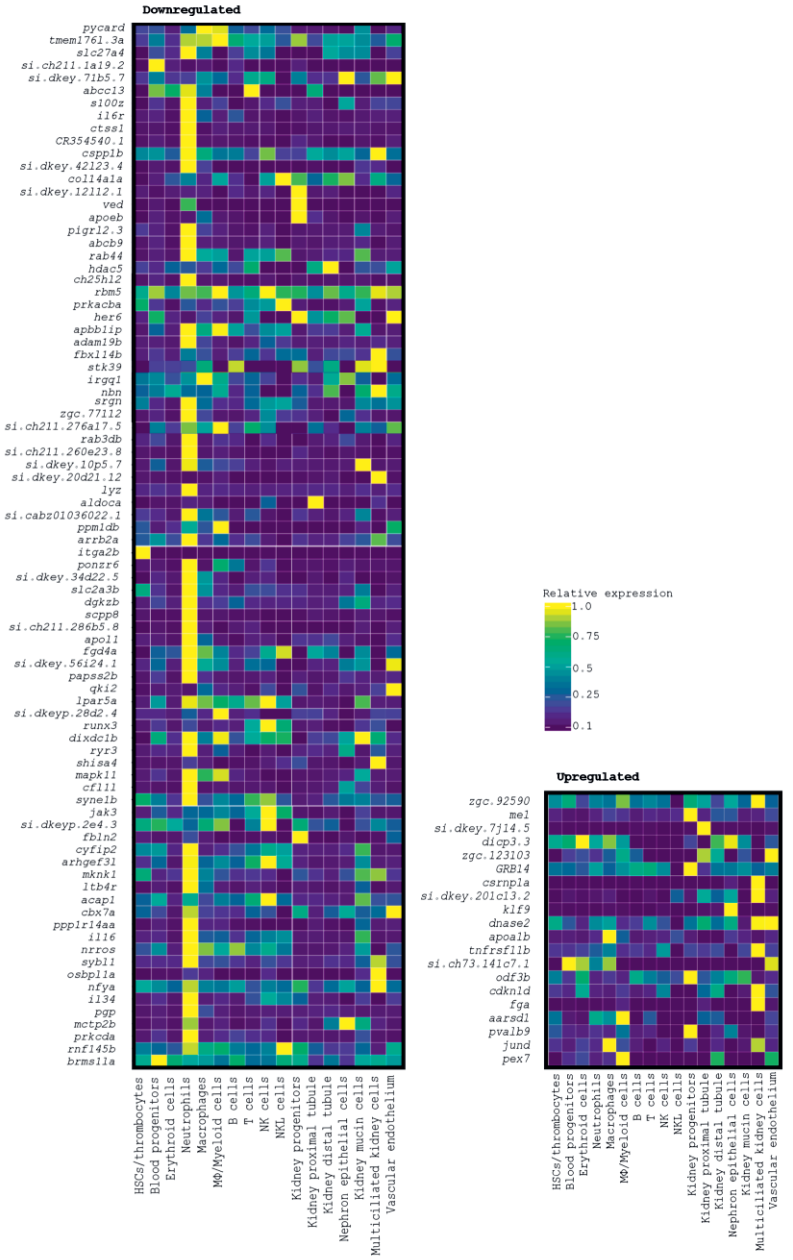


Figure 15. Expression of differentially expressed genes in blood cells. Differently expressed genes ($p<0.05$, fold change ≥ 2) in infected WT and infected knockout pycard$^{tpu5/tpu5}$ fish based on RNA-seq results and their expression in different blood cell populations according to the scRNA-seq data by Tang et al. 2017, derived from Lareau et al. (2017) (https://molpath.shinyapps.io/zebrafishblood/). Relative expression has been determined as the fragment of cells of the parent population (determined by transgenic fluorescent signatures) expressing the named transcript.

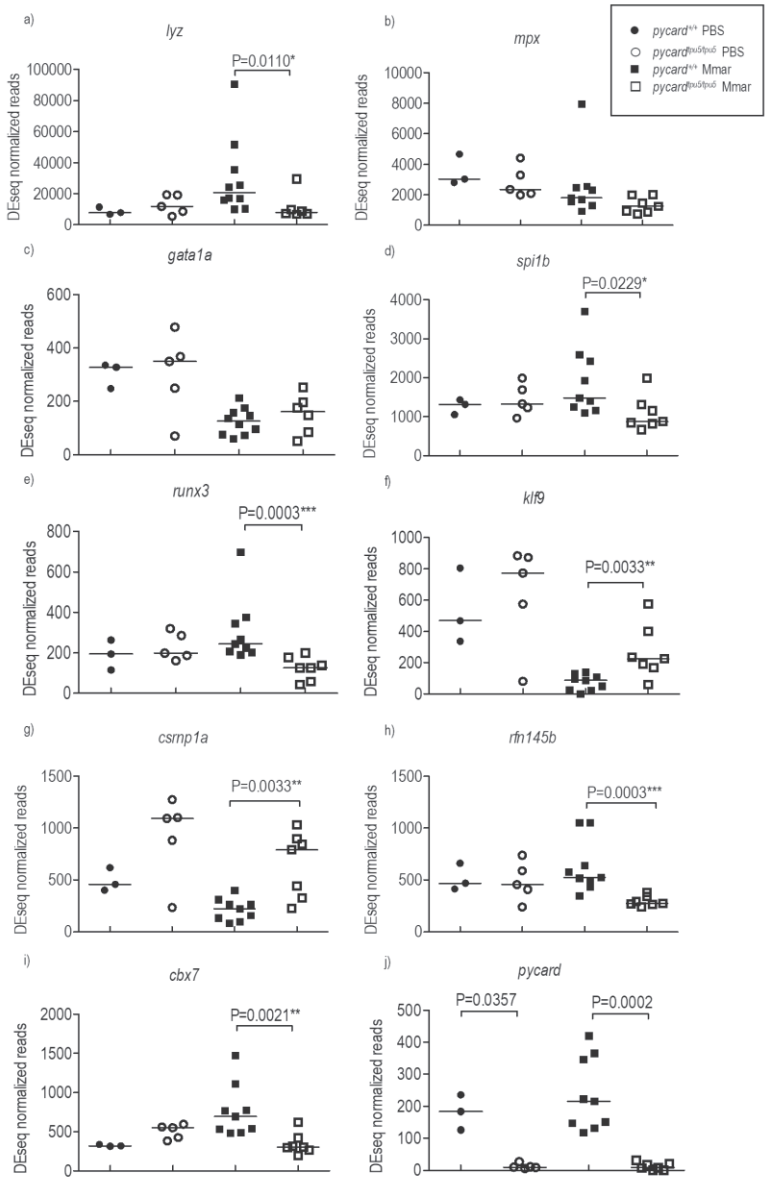


Figure 16. Findings from the RNA-seq of knockout *pycard* fish infected with low dose of *M. marinum*. Gene expression values are indicated as DEseq normalizes reads, normalised to gene length. Indicated are neutrophil marker genes a) *lyz*, b) *mpx*, haematopoietic transcription factors c) *gata1*, d) *spi1b*, e) polycomb group protein *cbx7a,* transcription factors f) *klf9* and g) runx3, and lastly our gene of interest h) *pycard*. A significant difference is indicated with the corresponding p-value, significance has been determined with a two-tailed U-test. The legend for all charts is displayed on the top right.

6 DISCUSSION

6.1 The chromatin landscape influences mutagenesis efficiency in zebrafish

In our study, we determined whether CRISPR-Cas9-based mutagenesis is equally efficient when targeting genes differentially expressed in early development. As mutating transcriptionally active genes was more often successful than was targeting genes which were only expressed late in embryogenesis, a correlation with chromatin seemed evident (Study I, Figure 3). However, using the available open access data, we were merely able to probe the chromatin status in a very superficial manner at irrelevant timepoints, and subsequently failed to find a significant correlation (Study I, Table 1). This is not surprising, as CRISPR-Cas9 acts in a complex developmental chromatin landscape that undergoes changes constantly in the zebrafish embryo.

During development, the totipotent fertilized egg will divide and give rise to all the cells of the mature organism. The dividing cells will lose their totipotency and become pluripotent stem cells and after further divisions will only carry the transcriptomic potency to express the genes of a specialised cell type. Thus, when injecting genome engineering reagents into a fertilized egg cell, the editing machinery will act in a chromatin landscape typical for the cell type at a given time. As the rate of cell division is different between organisms, the editing machinery will also have varying amounts of time to act before becoming diluted. Therefore, it can be reasoned that CRISPR-Cas9 could be subject to differences in the timing of the onset of the mutagenesis, due to the changes in chromatin landscape. Typically, chromatin is opened when genes are more active and is closed when genes are inactive (Tsompana and Buck 2014). However, the detection of transcripts is not an absolute readout for initiation of transcription, as maternal mRNAs are present in cells at

early timepoints (Harvey et al. 2013). Also, open chromatin does not always correlate with active transcription because developmentally important genes are epigenetically poised with permissive histone signals even if they are not actively transcribed (Lesch et al. 2013). Histones are the proteins around which DNA binds, and which bundle together to give nucleosomes. Histones are modified with acetylation and methylation which alters their activity. Typical signatures for actively transcribed genes are for example H3K27 (Histone 3 Lysine 27) acetylation, H3K4 methylation and H3K9 acetylation at the enhancer and promoter sequences of a gene (Perino and Veenstra 2016). As there is a number of different histones which implicate the transcriptional state of a gene, it would be interesting to test if genes wrapped in different kinds of histones are more readily accessible for CRISPR-Cas9 than others (Perino and Veenstra 2016). We attempted to see if this could be shown experimentally by testing the onset of the mutagenesis for functional gRNAs in zebrafish embryos, but with only three different genes (Study I, Figure 2). It may be, however, that the transcription machinery is also inhibitory to the cleavage function of Cas9, although there is no direct evidence to support this notion.

Different experimental readouts can be used as a measure of chromatin openness (Tsompana and Buck 2014). Nucleosome occupancy at CRISPR-Cas9 target sites has also been indicated by several studies (Hinz et al. 2015; Horlbeck et al. 2016; Isaac et al. 2016). Nucleosome breathing during cell division, however, partly lifts this barrier and allows CRISPR to act on its target (Isaac et al. 2016). Even if the assays probing the effect of the chromatin landscape on genome editing tools were done using cell lines, the hindrance caused by chromatin seems to be an established fact and likely applies similarly during early development.

In addition to the hindrance presented by the changing chromatin of the embryo, merely introducing a prokaryotic enzyme to a eukaryotic system can present challenges. Prokaryotic and eukaryotic cells are fundamentally different in their methods for arranging and controlling genetic material. Firstly, prokaryotic cells are defined by a lack of nuclear structures. Secondly, prokaryotic genomes are often smaller than those in eukaryotes, and they lack introns and hence all of the splicing machinery. The genetic material of bacteria is supercoiled, but devoid of highly organized packaging resembling that of eukaryotic chromatin. These are some of the many examples demonstrating how prokaryotes differ from eukaryotes. Despite the differences between prokaryotic and eukaryotic cells, the CRISPR-Cas9 immune system of bacteria is in fact functionally analogous to eukaryotic RNA-interference (RNAi), as RNAi similarly protects the eukaryotic cell from foreign RNAs (Makarova et al. 2006).

Cleaving genomic DNA using a prokaryotic DNA-cleaving enzyme is vastly different from cleaving extracted DNA *in vitro*, which is the basic procedure in, for example, molecular cloning. In applications like sequencing, DNA is not in its native state as DNA extracted from cells is often treated with proteinases, which also removes DNA packing histones. Also, editing genomic DNA in cell lines is different from editing the genomic DNA of a developing organism due to the dynamic chromatin landscape.

Furthermore, the efficiency of CRISPR-Cas9 is also dictated by the gRNA sequence specificity and stability, which was not taken into account in our research (Xu et al. 2017). The compatibility of the gRNA secondary structure with Cas9 binding is a likely parameter for the formation of a functional editing complex. The only readout without technological interference is whether a gRNA-Cas9 complex is able to cleave DNA *in vitro* but is unable to cleave it *in vivo* (Study I, Figure 1). In this setting, the only interfering component in the system is the number of binding sites available for cleavage, as well as the chromatin.

For the above reasons, simple assays which measure mutagenesis efficiency *in vitro* and compare this to the situation *in vivo* with only a single gRNA do not offer a conclusive description of the complex phenomenon. Moreover, this approach inaccurately measures the efficacy of gRNA design tools, and gRNA synthesis efficiency. Much like in Chen et al. (2017), it would be more conclusive to design multiple gRNAs targeted to open or closed chromatin and test their efficacy first *in vitro* and then *in vivo* (to minimize false negatives resulting from failures in *in vitro* gRNA synthesis). Large scale studies could reveal correlation and help minimize the influence of the target sequence.

6.2 *nhlrc2* knockdown in zebrafish as a model for the FINCA disease

A novel congenital disease FINCA (fibrosis, neurodegeneration, and cerebral angiomatosis) was described in three non-consanguineous patients. Despite being asymptomatic when born, the patients started manifesting with a multiorgan disease at around two months after birth and died before two years of age (Study II, Table 1). The patients were found to be compound heterozygous for two mutations (Asp148Tyr and Arg201GlyfsTer6) in a previously unidentified gene NHLRC2 (Study II, Supplementary figure 10a). An Nhlrc2 knockout mouse proved to be

homozygous lethal at the E8.5 stage in early development (Study II, Table 2). In addition to humans, a deleterious mutation in Nhlrc2 has been described in Angus cattle (Denholm 2017). The morpholino knockdown of *nhlrc2* in zebrafish resulted in cerebral vacuolisation phenotype according to the TEM analysis (Study II, Figure 6). This result resembles the cerebral alterations observed in the human patients (Study II Figure 1). The same amount of control morpholino did not evoke similar changes, which suggests that morpholino toxicity is not the cause of the changes observed (Study II, Figure 14). Tissue cultures from FINCA-patients suggested that the mutations cause severe tissue fibrosis and increased conversion of fibroblasts to myofibroblasts (Paakkola et al. 2018). Patients with FINCA had axial hypotonia and dystonia, which also indicates a role for NHLRC2 in maintaining muscle functionality (Study II, Table 1). Interestingly, a recent study connected *nhlrc2* to a group of genes associated with skeletal muscle atrophy induced by excessive exercise in zebrafish (Sun et al. 2021), which would also indicate a role for *nhlrc2* in maintaining normal muscle function in fish.

The NHLRC2 gene encodes a protein, which contains a thioredoxin domain. However, a reduction assay did not reveal any thioredoxin activity the NHLRC2 produced in bacteria (Study II, Figure 4c). Thus, the true role of NHLRC2 remains elusive. X-ray crystallography combined with a bioinformatic analysis revealed that NHLRC2 possesses a negative surface charge cleft which is a likely ligand binding site, and that the residues at this site are conserved across species (Biterova et al. 2018) (Figure 17). Similarly, a synteny analysis indicates the general organization of genes around NHLRC2 is conserved in human, mouse and zebrafish (Figure 17). The residue mutated in patients (Asp148) is conserved across species (Study II, Figure 4d). The other disease allele (Arg201GlyfsTer6) is processed through mRNA mediated decay (Study II, Supplementary figure 10a).

Paakkola et al. (2018) found that NHLRC2 localizes to the cytosol. They also investigated the potential interaction partners for NHLRC2 using proximity-mass spectrometry, which revealed that NHLRC2 binding partners are involved in cell-cell adhesion, cell division and protein transport (Paakkola et al. 2018). Supporting this, patient derived cell cultures were subjected to overexpression, knockdown and complementation assays, which suggested that NHLRC2 is involved in the organization of subcellular organelles, cellular morphology and the formation of intermediate filaments (Paakkola et al. 2018). In a knockout screen completed in human macrophages, NHLRC2 was found to be involved in actin dynamics and salmonella resistance (Yeung et al. 2019). Similar evidence from Haney et al. (2018) revealed the involvement of Nhlrc2 in RhoA-Rac1 signalling, which controls actin

polymerization and filopodia formation and ultimately phagocytosis. Using two-dimensional difference gel electrophoresis on neuronal precursor cells from heterozygous knockout mice, Hiltunen et al. (2020) showed how NHLRC2 affects especially vesicle associated proteins.

As the function of NHLRC2 remains elusive, deciding the phenotype which to look for proved to be challenging. The Harvey et al. (2013) dataset measuring maternal mRNA from 2-cell stage embryos showed that the *nhlrc2* coding mRNA is already present in the embryos and is likely to be maternally provided. Besides speaking for the importance of *nhlrc2* supply to the developing embryo, this also explains why *nhlrc2* knockdown zebrafish are not embryonically lethal like *Nhlrc2*-knockout mice, which die before the morula stage (Study II, Table 2).

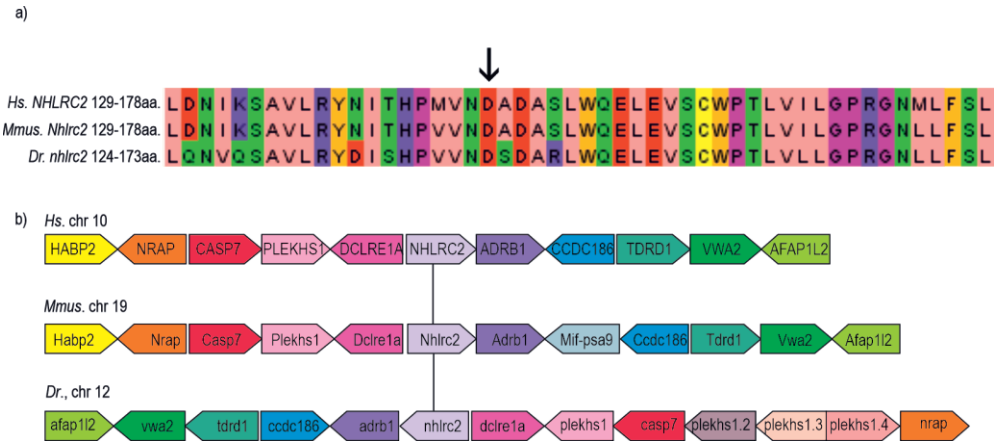


Figure 17. NHLRC2 synteny in human, mouse and zebrafish. a) The Clustal Omega alignment of the thioredoxin domain region of NHLRC2 homologues from human (*Homo sapiens, Hs.*), mouse (*M. musculus, Mmus.*), and zebrafish (*D. rerio, Dr.*). The patient mutation site Asp148 is indicated with an arrow. Colouring is according to Zappo. b) Gene synteny for human, mouse and zebrafish. Genes are displayed by their transcriptional direction in the respective chromosomes (chr.) for each species. NHLRC2 is aligned at the centre and connected with a line. Corresponding homologous genes are indicated with the same colour. Ensembl ID:s for NHLCR2 homologues: Hs. ENSG00000196865, Mmus. ENSMUSG00000025078, Dr. ENSDARG00000089581.

Despite the potential off-target effects of morpholinos, their use offers a fast and easy readout of a gene knockdown. Due to the lack of a functional antibody, the functionality of the translation blocking RNA could not be verified, and thus we decided to proceed solely with the splice site blocking morpholino. As the simultaneous silencing of tp53 did not seem to affect the phenotype, but it limited the amount of morpholino that could be administered, tp53 was left out to achieve

a sufficient level of silencing. A high knockdown efficiency was essential as both heterozygous humans and mice were healthy.

In order to increase the reliability of the zebrafish model, further characterisation would be essential. This would include an analysis of the tissue distribution, gene expression kinetics and also functional analysis of the protein. *In situ* hybridisation could be used for analysis of the tissue distribution of *nhlrc2* in the larval tissues to determine if it resembles that observed in human tissues. Some open access scRNA-seq data already exists which can be used for the analysis of the expression of *nhlrc2* in different tissues, and for example according to the data by Tang et al. (2017), the highest expression of *nhlrc2* is in macrophages, which would support its role in phagocytosis detected by Haney et al. (2018) and Yeung et al. (2019). However, whole embryo scRNA-seq analyses on mouse or human samples cannot be completed, so mining this open access zebrafish data could produce a completely novel understanding of the function of *nhlrc2* in vertebrate development.

Once verified, the zebrafish knockdown model could be used for further analysis of the phenotype caused by the loss of *nhlrc2*. A histological analysis of the larval zebrafish could reveal if fibrosis can already be detected at the early days of development. It seems unlikely that a drug treatment could abolish the symptoms of patients carrying the mutations. However, if more patients, which survive past childhood are found, screening for compounds which could alleviate the symptoms could also be done using the zebrafish model.

6.3 *itln1* does not alter the response to an infection in *itln3* knockout zebrafish

Intelectins, especially *itln3* were upregulated in our microarray data upon mycobacterial infection (Study III, Figure 1). To further study the function of *itln3* in a mycobacterial infection, *iltn3* was mutated to generate a knockout using CRISPR-Cas9 mutagenesis (Study III, Figure 3). *itln3* displayed the highest expression in the liver (Study III, Figure 2) and was upregulated upon a low dose *M. marinum* infection in both adult and larval zebrafish (Study III, Figure 1). The loss of *itln3* did not affect the survival of zebrafish from a mycobacterial infection in neither larvae (Study III, Figure 4) nor adults (Study III, Figure 6). In summary, the knockout of the gene did not result in an immunological phenotype. Rossi et al. (2015) suggest, this could be a result of genetic compensation mechanisms. Therefore, silencing compensating genes could reveal the underlying phenotype. When identifying genes

which could compensate for the loss of *itln3* in zebrafish, other members of the intelectin gene family are the most potential answer. Of these, *itln2* and *itln2-like* are both expressed mostly in the intestine (Study III, Figure 2). *iltn1* in turn, presents similar kinetics in infection as *iltn3* (Study III, Figure 1) and it is also expressed in the spleen and kidney, which are the main immunological organs of the fish (Study III, Figure 2). Thus, *iltn1* seemed like a potential candidate for substituting *itln3*. Morpholino silencing was selected as a strategy firstly because generating double mutants with CRISPR-Cas9 is time consuming and because it could result in further compensatory effects. Morpholino silencing is straightforward and less time consuming.

Morpholino silencing using splice site blocking morpholinos can be quantitated using a qPCR-based method. Quantitation showed that at early timepoints morpholino silencing lead to a near total loss of the transcript (Figure 10). A phenotype was observed in one of the mutants when comparing it to a group injected with a control morpholino (Figure 11). However, this cannot be attributed to the loss of both *iltn1* and *itln3*, as also the WT siblings of the mutants displayed a phenotype when *itln1* was silenced. Moreover, as the heterozygous siblings or the other mutant line did not recapitulate this result, it is most likely that the significant result is merely an artefact (Figure 11). However, a recent study has indicated that Itln1 is able to agglutinate bacteria and could potentially play a role in innate immunity (Chen et al. 2018:1). In conclusion, more studies to determine the role of itln1 are required.

The lack of a survival phenotype in both larvae and adult fish can be explained by the observation that the zebrafish Itln3 was not able to bind bacteria *in vitro* (Study III, Figure S7). However, this could also be a result of the protein purification tags, which could prevent the formation of a multimeric protein.

6.4 The response to an *M. marinum* infection is altered in *pycard* knockout zebrafish

Inflammasomes are important mediators of the inflammatory response, as they mediate the release of mature IL-1β (Martinon et al. 2002). Inflammasomes have a renowned role in the defence against tuberculosis (Master et al. 2008; Mayer-Barber et al. 2010; McElvania Tekippe et al. 2010) and their therapeutic potential has been recognised (Saiga et al. 2015). In addition, PYCARD has been indicated to work

independently of the inflammasome in several studies (e.g. Ellebedy et al. 2011; Fang et al. 2019; Taxman et al. 2011).

The knockout of *pycard* resulted in an increased susceptibility to a mycobacterial infection in adult zebrafish (Study IV, Figure 2). A reason for this seems to be the inability to control the bacterial growth (Study IV, figure 2). However, the reason for the inability to contain bacterial growth remains elusive. As the lack of *pycard* does not seem to affect the number of leukocytes (Study IV, Figure 3), it could be that it alters a more distinct gene expression pattern.

An RNA-seq analysis revealed that only four genes were significantly altered between steady state PBS injected *pycard*$^{-/-}$ and WT control fish (*nocta*, *akr1a1a* and *tmem176l.3a*) (Study IV, Supplementary table 1). Of these, nocturnin (*nocta*) is a gene regulated by the circadian clock. The NOCT homologue of Nocta, has been implicated in metabolic regulation, development and differentiation, but its specific role remains unclear (Hughes, Abshire, and Goldstrohm 2018). *akr1a1a* is a homologue of *Akr1a1* which is an aldehyde reductase which has also been indicated in the vitamin C synthesis pathway in knockout mice (Lai et al. 2017). *akr1a1a* has recently been found to play a role in maintaining normal insulin signalling in zebrafish (Qi et al. 2021). *tmem176l.3a* has no known function.

Of cell type specific genes, significant changes were observed in *lyz* (neutrophil), *cd7al* (T-cell), *jak3* (NK-cell), *itga2b* (thrombocyte) and *apoeb* (Mazzolini et al. 2020; Tang et al. 2017). As not all cell type markers are down, it can be these are specific for a subtype of cells (Study IV, Supplementary table 3).

According to the expression heat map, generated from the scRNA-seq data from Tang et al. (2017), it is evident that most of the genes downregulated by the lack of *pycard* are expressed in neutrophils (Figure 15). This could be explained by a decrease in the number of neutrophils, which unfortunately cannot be seen in the FACS data (Study IV, Figure 3) as all granulocytes and monocytes pool into the same group in the plot. Nevertheless, of the neutrophil marker genes, only the lysozyme homologue (*lyz*) was significantly downregulated in the knockout *pycard*$^{tpu5/tpu5}$., though a modest decrease was also seen in myeloperoxidase (*mpx*) (Figure 16). Previously, it has been proposed that inflammasome activation regulates neutrophil and macrophage differentiation by finetuning the balance of *spi1b* and *gata1a* by post transcriptionally cleaving *gata1* (Tyrkalska et al. 2019). This could indeed be the case, though it cannot be verified from our transcriptional data.

In turn, evidence from (Kalev-Zylinska et al. 2002) indicate a role for *runx3* in haematopoiesis. In their paper, morpholino knockdown of *runx3* leads to the depletion of *runx1*, which leads to a decrease in Lyz and Spi1b presenting cells in

zebrafish larvae (Kalev-Zylinska et al. 2002). Similarly, in human patient samples, the decline of RUNX3 has been indicated to transcriptionally control the expression of GATA1 (Balogh et al. 2020). In our data, both *runx3* and *lyz* are downregulated in *pycard*$^{tpu5/tpu5}$ knockout mutants, indicating that *pycard* is required for *runx3* expression, hence causing a decrease in *lyz* expression. Even so, *runx3* was not the only gene in our data associated with haematopoiesis or myelopoiesis. In addition to haematopoiesis, zebrafish *runx3* has been implicated in T-cell development (Iwanami et al. 2020) and in adult fish its expression is concentrated in T-cells (Tang et al. 2017).

Tuberculosis and other infectious bacterial species have been shown to modulate host epigenetics to control the expression of proinflammatory genes (reviewed by e.g. Fatima et al. (2021) and Gomez-Diaz et al. (2012)). The role of histone deacetylases has been shown to affect the outcome of a mycobacterial infection in both zebrafish and human macrophages (Moreira et al. 2020). Our data also indicated a differential expression of epigenetic, or transcription factor related modifiers: the nuclear transcription factor yα (*nfya*), histone deacetylase 5 (*hdac5*) and polycomb group protein chromobox 7 (*cbx7*), ring finger protein 145b (*rnf145b*) and cysteine-serine-rich nuclear protein 1a (*csrnp1a*) in response to a mycobacterial infection, in the absence of the inflammasome (Study IV, Supplementary table 2 and 3). Each of these, or their orthologue, has been suggested to control either haematopoietic cells or myelopoiesis. Mouse NF-Ya is a regulatory subunit of NF-Y transcription factor that interacts with DNA (Zhu et al. 2005). It is expressed in haematopoietic stem cells in the bone marrow and its overexpression has been shown to control renewal of the haematopoietic cells in mouse (Zhu et al. 2005). HDAC5 has been shown to suppress transcriptional potential of GATA1 in murine erythroleukemia cells (Watamoto et al. 2003). Mouse C*bx7* is has been indicated in controlling haematopoietic cell renewal and myelopoiesis (Jung et al. 2019). In this study, over expression of *Cbx7* enhanced mice myelopoiesis (Jung et al. 2019). CBX7 has also been indicated to correlate with increased number of blood cells including neutrophils in cancer patients (D. Li et al. 2020). Ring finger protein 145b (*rnf145b*) silencing has been shown to lead to ablation of primitive erythropoiesis and thrombocyte formation (Gieger et al. 2011). Morpholino knockdown of *csrnp1a* has been shown to result in a decrease in blood cells and in expression of *gata1* (Espina et al. 2013). Curiously, only *csrnp1a* is upregulated in our data, implying it could be compensating for the downregulation of the other genes. To our knowledge, of these transcription factors, *runx3* and *hdac5* has been investigated in relation to

mycobacterial infection (Lavalett, Ortega, and Barrera 2020; Moreira et al. 2020; Subbian et al. 2013).

The zinc finger transcription factor kruppel-like factor 9 (*klf9*) was upregulated in our mutants (Figure 16). Knockdown of klf9 has been shown to result in decreased haematopoiesis, erythroid maturation and T lymphopoiesis (Zhang et al. 2017). Curiously, *klf9* has recently been indicated in the regulation of proinflammatory genes in response to cortisol treatment in RNA-seq data from whole zebrafish larvae (Gans et al. 2020). A previous paper from the same group, Hartig et al. (2016), showed that cortisol treatment leads to a proinflammatory phenotype, which explains why genes regulated by *klf9* in response to cortisol are inflammatory genes. Interestingly, 11 genes upregulated in the Gans et al. (2020) study carried out in cortisol treated zebrafish larvae in a *klf9* mediated manner, include hit genes observed also in our data on infected fish, including lysozyme (*lyz*), si:dkey-8k3.2 (*si:dkey-8k3.2*), transmembrane protein 176l.3a (*tmem176l.3a*), chromobox homolog 7a (*cbx7a*), si:ch211-147j13.3 (*si:ch211-147j13.3*), interleukin 34 (*il34*), serglycin (*srgn*), si: ch211-223l2.4 (*si: ch211-223l2.4*), diverse immunoglobulin domain-containing protein 3.3 (*dicp3.3*), interleukin 6 receptor (*il6r*), and Kruppel-like factor 9(*klf9*) itself. As *klf9* is upregulated in *pycard*$^{tpu5/tpu5}$ knockout fish, our data indicates it either compensates for the loss of *pycard* or is upregulated in its absence. In other models, mice *Klf9* has been indicated to be a feedforward regulator of oxidative stress (Zucker et al. 2014). It has been shown to independently cause increased levels of reactive oxygen species (ROS) in cultured cells and in mouse tissues (Zucker et al. 2014). ROS are known to play a role in defence against tuberculosis (Shastri et al. 2018). The results of Zucker et al. (2014) also point towards the role of Klf9 in bleomycin induced pulmonary fibrosis in mice. Previous research also indicates that KLF9 is upregulated in the neutrophils of patients suffering from acute respiratory distress syndrome and in patient derived neutrophils treated with septic plasma (Juss et al. 2016; Khaenam et al. 2014). Considering that genes regulated by *klf9* in cortisol treated embryos are down in our data, *pycard* could be a downstream factor from *klf9*. *klf9* has not, to our knowledge, been previously indicated in tuberculosis.

Whether the indicated chromatin modifying proteins and transcription factors implicated in our data (*nfya*, *hdac5*, *runx3, cxb7, rnf145b, csrnp1a* and *klf9*) are directly or indirectly associated with *pycard* cannot be determined from the data. However, as all of these have been associated with haematopoiesis or myelopoiesis, our data does suggest another potential pathways for the impairment in neutrophils in *pycard* mutants besides the caspase mediated cleavage of Gata1 (Tyrkalska et al. 2019). It

can also be theorized that the results observed by Tyrkalska et al. (2019) are perhaps not recapitulated in our mutant, as some effects observed by us can be the result of genetic compensation mechanisms, which are not observed in the morpholino knockdown. Besides the differences in morphants and mutants, Tyrkalska et al. studied larval zebrafish in a salmonella infection. Still, our results do not rule out the effect of post translational effects on Gata1.

Of interest is also the downregulation of *nkl.1* (Study IV, Supplementary table 1). Zebrafish *nkl.1* is the first of the four tandem duplicates, and its expression is highest in the kidney, liver, head, and muscle (Pereiro et al. 2015). In the kidney, *nkl.1* is most expressed in myeloid cells (Pereiro et al. 2015). It is also elevated at 24 hours post a *Spring viremia carp virus* challenge (Pereiro et al. 2015). Human NK-lysin or granulysin is known to participate in killing intracellular mycobacteria (Stenger et al. 1999), but evidence for the homology of the zebrafish *nkl* genes remains to be shown.

In contrast to the clear indication of the impact on neutrophils in the downregulated genes, the genes upregulated in knockout *pycard*$^{tpu5/tpu5}$ do not indicate a specific expression pattern (Figure 15). They are of interest as they are potential compensating genes, and genes potentially downregulated by *pycard* in WT fish. Of these genes negative regulator of ROS (*nrros*) and the malic enzyme 1 (*me1*) are associated with mitigating cytosolic ROS (Noubade et al. 2014; Shao et al. 2020). *dicp3.3* in turn, is associated with T-cell activation, pointing towards a role in adaptive immunity.

Out of the unknown genes in our dataset, *tmem176l.3a* is the only gene implicated to be differentially regulated in both infection and in the basal state, and has recently been shown to be a negative regulator of the inflammasome in mice (Segovia et al. 2019) (Study IV, Supplementary table 1). Moreover, in the scRNA-seq data by Farnsworth et al. (2020), *tmem176l.3a* strongly colocalizes with *pycard* expression in zebrafish larvae. To our understanding, no connection has previously been made with these two genes. The TMEM176 protein family is conserved in vertebrates, and members of the family are known to be expressed at least in macrophages (Zuccolo et al. 2010), whereas in the Tang et al. (2017) data the highest expression was in all myeloid lineage cells (Figure 15).

Most of the genes implicated in our RNA-seq data are known to be involved in the immune response or haematopoiesis. The dysregulated genes also present pathways not previously connected to *pycard* or inflammasome activation, which can be of value also outside zebrafish tuberculosis research. Whether these results can be translated to a human setting remains to be determined.

6.5 Mimicking a disease phenotype with knockouts and knockdowns

The applicability of a model to answer the research question needs to be considered carefully. In some cases, as with the FINCA disease, the patients were compound heterozygous, presenting one allele which causes a frameshift and a truncated protein and another allele, which results in the production of a protein with an abolished ligand binding capacity. This can lead to a situation, where the production of a full-length mRNA fails to trigger compensatory effects, and also compensates for the loss of the other allele (Study II, Figure 4). Still, the production of the NHLRC2 protein is abolished (Study II, Figure 4). This is effectively recapitulated by the knockdown of *nhlrc2* in zebrafish, as mRNA-knockdown does not result in compensatory effects. However, the concept of compensation introduces an interesting paradigm which should be carefully considered when generating models or mimicking human diseases using morphants or mutants.

A notion which at least I have not seen put forward at any point of the debate is the tissue specificity of the gene silencing and mutagenesis. Morpholinos are generated by injecting the reagents into a 1 - 16 cell stage embryo, which causes the reagents to diffuse into the dividing cells in supposedly equal manner. This leads to the transcript interfering in all of the tissues, also supposedly in tissues where the target transcript is not expressed at all. Logic would state that in tissues where the target transcript is not present, off-targeting is more prevalent, as targeting compounds have more time to scan alternative binding sites, similarly to the effect presented by the chromatin landscape on genome editing tools like CRISPR-Cas9. Thus, it cannot be completely verified if our morphants or mutants are entirely free of off-target effects. However, if a knockout mutant line was available, injecting it with morpholinos could reveal potential off-target effects.

With mutants, mutagenesis can be targeted to a tissue with tissue specific vector systems. However, if trying to mimic a disease phenotype, tissue specific loss of function is more often not the intended aim. It would be of interest to determine, if morpholinos targeting genes expressed in a specific tissue cause more off-target effects than those targeting genes expressed ubiquitously. When considering the potential off-targeting, toxicity related effects and genetic compensation, it turns out no tool is perfect on its own. In addition, when modelling human genetic diseases, it should be analysed if the mutation in humans leads to a production of a non-functional protein, or if it triggers compensatory responses.

In relation to the compensatory effects of the mutagenesis, it has been estimated that on average humans bear around 100 heterozygous and 20 homozygous loss of function alleles which do not affect our general survival (Bunton-Stasyshyn et al. 2019). Humans are known to have higher number of transcriptome complexity, with an estimated double the number of transcript variants per gene compared to the mouse (Bunton-Stasyshyn et al. 2019). It can be that mammals have developed other mechanisms to compensate for potentially deleterious alleles than zebrafish, which seem to combat loss of function mutations with the sheer number of paralogues. When considering the selection of a suitable model for human genetic diseases resulting from a loss of function, compensatory mechanisms should be taken into account. Ideally, the compensatory mechanisms should be the same in both the patient and the model. In summary, both morphants and mutants present challenges to research. Therefore, an ideal experiment should always include both the mutant and the morphant and, to be more reliable, at least two of each with appropriate controls (Stainier et al. 2017).

7 SUMMARY AND CONCLUSIONS

The research on human genetic and infectious diseases largely relies on animal models. When a new drug compound is introduced for testing, it first undergoes a number of preclinical trials, which are mostly done using mice and primates, before entering clinical trials on human patients. The success rate for a drugs entering clinical trials is only 10%, when measured as a fraction of FDA (U.S. Food and Drug Administration) approved drugs from the total number of drugs that entered the clinical phase (Hay et al. 2014). One potential reason for this is that preclinical testing is not able to provide conclusive results on the *in vivo* function of drugs, due to an incomplete understanding of genetic events related to the disease. Besides serving as a model for genetic interactions, the zebrafish model can offer new viewpoints in assessing drug efficacy and administration. The knowledge from zebrafish can be used to complement the knowledge gained from mouse studies to assist in selecting more potent lead candidates for clinical testing, with the aim in improving the success rate of novel drug candidates. Besides improving the success rate, implementing zebrafish on preclinical testing could reduce the need for experimenting on mammalian species.

The zebrafish has been shown to be suitable for modelling developmental and infectious diseases. Moreover, it provides options for experiments, which are less applicable with other models due to biological differences. In this thesis we apply the zebrafish model in research of human genetic disease FINCA. Mutations observed in FINCA-patients result in pre-natal lethality in mouse, making it difficult to study FINCA in mouse. The gene silencing experiments in zebrafish larvae provided the necessary complementarity to the phenotype observed in human patients to confirm that the mutations in NHLRC2 are responsible for the symptoms observed in the FINCA patients.

Zebrafish was also applied for studying the significance of the intelectin gene family in mycobacterial infection. Despite being upregulated in infection, our experiments show that *intelectin 3* is dispensable for survival in mycobacterial infection. In addition, silencing the expression of *intelectin 1* did not alter this phenotype, suggesting that genetic compensation is not responsible for the lack of

phenotype in *intelectin 3* knockout zebrafish. Our results suggest that despite being induced during infection, intelectins 1 or 3 are not required for zebrafish survival from mycobacterial infection.

The ease of generating genetically modified organisms can be limited by the accessibility of the gene in early development. Despite not being able to pinpoint the role of epigenetic factors limiting the efficiency of CRISPR-Cas9 *in vivo* in zebrafish larvae, our results do support that target gene accessibility can influence CRISPR-efficiency. Our study indicates that chromatin accessibility should be taken into account in experimental design and thus emphasises the need for solutions which could improve chromatin accessibility to also limit off-target activity of CRISPR-Cas9.

Using the CRISPR-Cas9 technology, we generated a knockout zebrafish model for *pycard*. The *pycard* gene encodes an adaptor protein required for inflammasome complex and signalling. Our results indicate that *pycard* knockout adult zebrafish are more susceptible for mycobacterial infection than WT, and, that this is a result of a defect in haematopoiesis and neutrophil differentiation. Our results support the recent observations regarding the role of inflammasome in demand-driven haematopoiesis and extrapolate this to infection (Gans et al. 2020; Tyrkalska et al. 2019). Moreover, our results indicate immune cell transcriptome changes which can be responsible for these observations (Frame et al. 2020; Tyrkalska et al. 2019). The number of transcription factors among the genes affected by the loss of *pycard* indicated in our data present interesting directions for future studies on epigenetic factors in mycobacterial defence. Furthermore, these results also strengthen the use of zebrafish in modelling human tuberculosis and inflammasome signalling.

In conclusion, our results display, that zebrafish is an applicable model for studying genetics behind human diseases. Using zebrafish as a model for genetics brings forth a number of experimental options, which would be inapplicable or less suitable for mammalian studies. In addition, in situations where mouse is less optimal, zebrafish can offer a valuable secondary option. Our studies successfully employed zebrafish in modelling human genetic and infectious diseases in four studies which will each pave way for future research.

8 REFERENCES

Adams, K. N., K. Takaki, L. E. Connolly, H. Wiedenhoft, Kathryn Winglee, O. Humbert, P. H. Edelstein, C. L. Cosma, and L. Ramakrishnan. 2011. "Drug Tolerance in Replicating Mycobacteria Mediated by a Macrophage-Induced Efflux Mechanism." *Cell* 145(1):39–53. doi: 10.1016/j.cell.2011.02.022.

Adamson, Kathryn Isabel, Eamonn Sheridan, and Andrew James Grierson. 2018. "Use of Zebrafish Models to Investigate Rare Human Disease." *Journal of Medical Genetics* 55(10):641–49. doi: 10.1136/jmedgenet-2018-105358.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. "Basic Local Alignment Search Tool." *Journal of Molecular Biology* 215(3):403–10. doi: 10.1016/S0022-2836(05)80360-2.

Andersen, P., and T. M. Doherty. 2005. "The Success and Failure of BCG - Implications for a Novel Tuberculosis Vaccine." *Nature Reviews Microbiology* 3(8):656–62. doi: 10.1038/nrmicro1211.

Andrews, S. 2010. "Fastqc. a Quality Control Tool for High Throughput Sequence Data." 2017(04.12.).

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* 8:2045. doi: 10.1038/s41467-017-02305-6.

Baeten, Jeremy T., and Jill L. O. de Jong. 2018. "Genetic Models of Leukemia in Zebrafish." *Frontiers in Cell and Developmental Biology* 6:115. doi: 10.3389/fcell.2018.00115.

Bajoghli, Baubak, Advaita M. Dick, Annisa Claasen, Larissa Doll, and Narges Aghaallaei. 2019. "Zebrafish and Medaka: Two Teleost Models of T-Cell and Thymic Development." *International Journal of Molecular Sciences* 20(17):4179. doi: 10.3390/ijms20174179.

Bakkers, Jeroen. 2011. "Zebrafish as a Model to Study Cardiac Development and Human Cardiac Disease." *Cardiovascular Research* 91(2):279–88. doi: 10.1093/cvr/cvr098.

Balik-Meisner, Michele, Lisa Truong, Elizabeth H. Scholl, Robert L. Tanguay, and David M. Reif. 2018. "Population Genetic Diversity in Zebrafish Lines." *Mammalian Genome: Official Journal of the International Mammalian Genome Society* 29(1–2):90–100. doi: 10.1007/s00335-018-9735-x.

Balla, Keir M., Geanncarlo Lugo-Villarino, Jan M. Spitsbergen, David L. Stachura, Yan Hu, Karina Bañuelos, Octavio Romo-Fewell, Raffi V. Aroian, and David Traver. 2010. "Eosinophils in the Zebrafish: Prospective Isolation, Characterization, and Eosinophilia Induction by Helminth Determinants." *Blood* 116(19):3944–54. doi: 10.1182/blood-2010-03-267419.

Balogh, Peter, Emmalee R. Adelman, John V. Pluvinage, Brian J. Capaldo, Katie C. Freeman, Sandeep Singh, Kamaleldin E. Elagib, Yukio Nakamura, Ryo Kurita, Goro Sashida, Eli R. Zunder, Hui Li, Alejandro A. Gru, Elizabeth A. Price, Stanley L. Schrier, Irving L. Weissman, Maria E. Figueroa, Wendy W. Pang, and Adam N. Goldfarb. 2020.

"RUNX3 Levels in Human Hematopoietic Progenitors Are Regulated by Aging and Dictate Erythroid-Myeloid Balance." *Haematologica* 105(4):905–13. doi: 10.3324/haematol.2018.208918.

Baraban, Scott C., Matthew T. Dinday, and Gabriela A. Hortopan. 2013. "Drug Screening in Scn1a Zebrafish Mutant Identifies Clemizole as a Potential Dravet Syndrome Treatment." *Nature Communications* 4(1):2410. doi: 10.1038/ncomms3410.

Barros, T. P., W. K. Alderton, H. M. Reynolds, A. G. Roach, and S. Berghmans. 2008. "Zebrafish: An Emerging Technology for in Vivo Pharmacological Assessment to Identify Potential Safety Liabilities in Early Drug Discovery." *British Journal of Pharmacology* 154(7):1400–1413. doi: 10.1038/bjp.2008.249.

Beckwith, Kai S., Marianne S. Beckwith, Sindre Ullmann, Ragnhild S. Sætra, Haelin Kim, Anne Marstad, Signe E. Åsberg, Trine A. Strand, Markus Haug, Michael Niederweis, Harald A. Stenmark, and Trude H. Flo. 2020. "Plasma Membrane Damage Causes NLRP3 Activation and Pyroptosis during Mycobacterium Tuberculosis Infection." *Nature Communications* 11(1):2270. doi: 10.1038/s41467-020-16143-6.

Beffagna, Giorgia. 2019. "Zebrafish as a Smart Model to Understand Regeneration After Heart Injury: How Fish Could Help Humans." *Frontiers in Cardiovascular Medicine* 6:107. doi: 10.3389/fcvm.2019.00107.

Benard, Erica L., Astrid M. van der Sar, Felix Ellett, Graham J. Lieschke, Herman P. Spaink, and Annemarie H. Meijer. 2012. "Infection of Zebrafish Embryos with Intracellular Bacterial Pathogens." *Jove-Journal of Visualized Experiments* (61):e3781. doi: 10.3791/3781.

Beumer, Kelly J., Jonathan K. Trautman, Ana Bozas, Ji-Long Liu, Jared Rutter, Joseph G. Gall, and Dana Carroll. 2008. "Efficient Gene Targeting in Drosophila by Direct Embryo Injection with Zinc-Finger Nucleases." *Proceedings of the National Academy of Sciences* 105(50):19821–26. doi: 10.1073/pnas.0810475105.

Bibikova, Marina, Mary Golic, Kent G. Golic, and Dana Carroll. 2002. "Targeted Chromosomal Cleavage and Mutagenesis in Drosophila Using Zinc-Finger Nucleases." *Genetics* 161(3):1169–75.

Biterova, Ekaterina, Alexander Ignatyev, Johanna Uusimaa, Reetta Hinttala, and Lloyd W. Ruddock. 2018. "Structural Analysis of Human NHLRC2, Mutations of Which Are Associated with FINCA Disease." *PLoS ONE* 13(8):e0202391. doi: 10.1371/journal.pone.0202391.

Blum, Martin, Edward M. De Robertis, John B. Wallingford, and Christof Niehrs. 2015. "Morpholinos: Antisense and Sensibility." *Developmental Cell* 35(2):145–49. doi: 10.1016/j.devcel.2015.09.017.

Bohnsack, Brenda L., Donika Gallina, and Alon Kahana. 2011. "Phenothiourea Sensitizes Zebrafish Cranial Neural Crest and Extraocular Muscle Development to Changes in Retinoic Acid and IGF Signaling." *PLOS ONE* 6(8):e22991. doi: 10.1371/journal.pone.0022991.

Boraschi, Diana, and Paola Italiani. 2018. "Innate Immune Memory: Time for Adopting a Correct Terminology." *Frontiers in Immunology* 9:799. doi: 10.3389/fimmu.2018.00799.

Boshra, H., J. Li, and J. O. Sunyer. 2006. "Recent Advances on the Complement System of Teleost Fish." *Fish & Shellfish Immunology* 20(2):239–62. doi: 10.1016/j.fsi.2005.04.004.

Brenner, S. 1974. "The Genetics of Caenorhabditis Elegans." *Genetics* 77(1):71–94.

Brownlie, Alison, Candace Hersey, Andrew C. Oates, Barry H. Paw, Arnold M. Falick, H. Ewa Witkowska, Jonathan Flint, Doug Higgs, Jason Jessen, Nathan Bahary, Hao Zhu, Shuo Lin, and Leonard Zon. 2003. "Characterization of Embryonic Globin Genes of

the Zebrafish." *Developmental Biology* 255(1):48–61. doi: 10.1016/s0012-1606(02)00041-6.

Bunton-Stasyshyn, Rosie K. A., Sara Wells, and Lydia Teboul. 2019. "When All Is Not Lost: Considering Genetic Compensation in Laboratory Animals." *Lab Animal* 48(10):282–84. doi: 10.1038/s41684-019-0397-4.

Butler, Matthew G., James R. Iben, Kurt C. Marsden, Jonathan A. Epstein, Michael Granato, and Brant M. Weinstein. 2015. "SNPfisher: Tools for Probing Genetic Variation in Laboratory-Reared Zebrafish." *Development* 142(8):1542–52. doi: 10.1242/dev.118786.

Campbell, N. A., and J. B. Reece. 2005. *Biology*. 7th ed. Boston: Benjamin Cummings /Pearson.

Can, Handan, Sree K. Chanumolu, Elena Gonzalez-Muñoz, Sukumal Prukudom, Hasan H. Otu, and Jose B. Cibelli. 2020. "Comparative Analysis of Single-Cell Transcriptomics in Human and Zebrafish Oocytes." *BMC Genomics* 21(1):471. doi: 10.1186/s12864-020-06860-z.

Carlsson, Fredric, Janice Kim, Calin Dumitru, Kai H. Barck, Richard A. D. Carano, Mei Sun, Lauri Diehl, and Eric J. Brown. 2010. "Host-Detrimental Role of Esx-1-Mediated Inflammasome Activation in Mycobacterial Infection." *PLOS Pathogens* 6(5):e1000895. doi: 10.1371/journal.ppat.1000895.

Carmona, Santiago J., Sarah A. Teichmann, Lauren Ferreira, Iain C. Macaulay, Michael J. T. Stubbington, Ana Cvejic, and David Gfeller. 2017. "Single-Cell Transcriptome Analysis of Fish Immune Cells Provides Insight into the Evolution of Vertebrate Immune Cell Types." *Genome Research* 27(3):451–61. doi: 10.1101/gr.207704.116.

Carradice, Duncan, and Graham J. Lieschke. 2008. "Zebrafish in Hematology: Sushi or Science?" *Blood* 111(7):3331–42. doi: 10.1182/blood-2007-10-052761.

Carvalho, Ralph, Jan de Sonneville, Oliver W. Stockhammer, Nigel D. L. Savage, Wouter J. Veneman, Tom H. M. Ottenhoff, Ron P. Dirks, Annemarie H. Meijer, and Herman P. Spaink. 2011. "A High-Throughput Screen for Tuberculosis Progression." *PLOS ONE* 6(2):e16779. doi: 10.1371/journal.pone.0016779.

Cassar, Steven, Isaac Adatto, Jennifer L. Freeman, Joshua T. Gamse, Iñaki Iturria, Christian Lawrence, Arantza Muriana, Randall T. Peterson, Steven Van Cruchten, and Leonard I. Zon. 2020. "Use of Zebrafish in Drug Discovery Toxicology." *Chemical Research in Toxicology* 33(1):95–118. doi: 10.1021/acs.chemrestox.9b00335.

Chai, Qiyao, Lin Wang, Cui Hua Liu, and Baoxue Ge. 2020. "New Insights into the Evasion of Host Innate Immunity by Mycobacterium Tuberculosis." *Cellular & Molecular Immunology* 17(9):901–13. doi: 10.1038/s41423-020-0502-z.

Chakrabarti, Sankar, George Streisinger, Fred Singer, and Charline Walker. 1983. "Frequency of γ-Ray Induced Specific Locus and Recessive Lethal Mutations in Mature Germ Cells of the Zebrafish, BRACHYDANIO RERIO." *Genetics* 103(1):109–23.

Chen, Hong, Yue Liang, Yawen Han, Tengfei Liu, and Shulin Chen. 2021. "Genome-Wide Analysis of Toll-like Receptors in Zebrafish and the Effect of Rearing Temperature on the Receptors in Response to Stimulated Pathogen Infection." *Journal of Fish Diseases* 44(3):337–49. doi: 10.1111/jfd.13287.

Chen, Lei, Jinyi Li, and Guiwen Yang. 2020. "A Comparative Review of Intelectins." *Scandinavian Journal of Immunology* 92(1):e12882. doi: 10.1111/sji.12882.

Chen, Lei, Jie Yan, Jing Shi, Wenbo Sun, Zhi Chen, Jiang Yu, Jing Qi, Yijun Du, Haiqing Zhang, and Lijun Feng. 2018. "Zebrafish Intelectin 1 (ZITLN1) Plays a Role in the

Innate Immune Response." *Fish & Shellfish Immunology* 83:96–103. doi: 10.1016/j.fsi.2018.09.004.

Chen, Lei, Jie Yan, Weiping Sun, Yan Zhang, Chao Sui, Jing Qi, Yijun Du, and Lijun Feng. 2016. "A Zebrafish Intelectin Ortholog Agglutinates Both Gram-Negative and Gram-Positive Bacteria with Binding Capacity to Bacterial Polysaccharide." *Fish & Shellfish Immunology* 55:729–36. doi: 10.1016/j.fsi.2016.06.023.

Chen, Xiang-Ke, Joseph Shiu-Kwong Kwan, Raymond Chuen-Chung Chang, and Alvin Chun-Hang Ma. 2021. "1-Phenyl 2-Thiourea (PTU) Activates Autophagy in Zebrafish Embryos." *Autophagy* 17(5):1222–31. doi: 10.1080/15548627.2020.1755119.

Chen, Xiaoyu, Marrit Rinsma, Josephine M. Janssen, Jin Liu, Ignazio Maggio, and Manuel A. F. V. Goncalves. 2016. "Probing the Impact of Chromatin Conformation on Genome Editing Tools." *Nucleic Acids Research* 44(13):6482–92. doi: 10.1093/nar/gkw524.

Chen, Yunru, Shiyang Zeng, Ruikun Hu, Xiangxiu Wang, Weilai Huang, Jiangfang Liu, Luying Wang, Guifen Liu, Ying Cao, and Yong Zhang. 2017. "Using Local Chromatin Structure to Improve CRISPR/Cas9 Efficiency in Zebrafish." *PLOS ONE* 12(8):e0182528. doi: 10.1371/journal.pone.0182528.

Cheong, Melody, Kate H. Gartlan, Jason S. Lee, Siok-Keen Tey, Ping Zhang, Rachel D. Kuns, Christopher E. Andoniou, Jose Paulo Martins, Karshing Chang, Vivien R. Sutton, Greg Kelly, Antiopi Varelias, Slavica Vuckovic, Kate A. Markey, Glen M. Boyle, Mark J. Smyth, Christian R. Engwerda, Kelli P. A. MacDonald, Joseph A. Trapani, Mariapia A. Degli-Esposti, Motoko Koyama, and Geoffrey R. Hill. 2020. "ASC Modulates CTL Cytotoxicity and Transplant Outcome Independent of the Inflammasome." *Cancer Immunology Research* 8(8):1085–98. doi: 10.1158/2326-6066.CIR-19-0653.

Christian, Michelle, Tomas Cermak, Erin L. Doyle, Clarice Schmidt, Feng Zhang, Aaron Hummel, Adam J. Bogdanove, and Daniel F. Voytas. 2010. "Targeting DNA Double-Strand Breaks with TAL Effector Nucleases." *Genetics* 186(2):757–61. doi: 10.1534/genetics.110.120717.

Clay, Hilary, J. Muse Davis, Dana Beery, Anna Huttenlocher, Susan E. Lyons, and Lalita Ramakrishnan. 2007. "Dichotomous Role of the Macrophage in Early Mycobacterium Marinum Infection of the Zebrafish." *Cell Host & Microbe* 2(1):29–39. doi: 10.1016/j.chom.2007.06.004.

Cong, Le, F. Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D. Hsu, Xuebing Wu, Wenyan Jiang, Luciano A. Marraffini, and Feng Zhang. 2013. "Multiplex Genome Engineering Using CRISPR/Cas Systems." *Science (New York, N.Y.)* 339(6121):819–23. doi: 10.1126/science.1231143.

Cosma, Christine L., Olivier Humbert, and Lalita Ramakrishnan. 2004. "Superinfecting Mycobacteria Home to Established Tuberculous Granulomas." *Nature Immunology* 5(8):828–35. doi: 10.1038/ni1091.

Costa, Bruna, Marta F. Estrada, Raquel V. Mendes, and Rita Fior. 2020. "Zebrafish Avatars towards Personalized Medicine—A Comparative Review between Avatar Models." *Cells* 9(2):293. doi: 10.3390/cells9020293.

Crim, Marcus J., and Christian Lawrence. 2021. "A Fish Is Not a Mouse: Understanding Differences in Background Genetics Is Critical for Reproducibility." *Lab Animal* 50(1):19–25. doi: 10.1038/s41684-020-00683-x.

Cronan, Mark R., and David M. Tobin. 2014. "Fit for Consumption: Zebrafish as a Model for Tuberculosis." *Disease Models & Mechanisms* 7(7):777–84. doi: 10.1242/dmm.016089.

Cully, Megan. 2019. "Zebrafish Earn Their Drug Discovery Stripes." *Nature Reviews. Drug Discovery* 18(11):811–13. doi: 10.1038/d41573-019-00165-x.

Cunningham, Carlee MacPherson, Gianfranco Bellipanni, Raymond Habas, and Darius Balciunas. 2020. "Deletion of Morpholino Binding Sites (DeMOBS) to Assess Specificity of Morphant Phenotypes." *Scientific Reports* 10(1):15366. doi: 10.1038/s41598-020-71708-1.

Daer, René M., Josh P. Cutts, David A. Brafman, and Karmella A. Haynes. 2017. "The Impact of Chromatin Dynamics on Cas9-Mediated Genome Editing in Human Cells." *ACS Synthetic Biology* 6(3):428–38. doi: 10.1021/acssynbio.5b00299.

Danilova, Nadia, and Lisa A. Steiner. 2002. "B Cells Develop in the Zebrafish Pancreas." *Proceedings of the National Academy of Sciences of the United States of America* 99(21):13711–16. doi: 10.1073/pnas.212515999.

Davis, J. Muse, Hilary Clay, Jessica L. Lewis, Nafisa Ghori, Philippe Herbomel, and Lalita Ramakrishnan. 2002. "Real-Time Visualization of Mycobacterium-Macrophage Interactions Leading to Initiation of Granuloma Formation in Zebrafish Embryos." *Immunity* 17(6):693–702. doi: 10.1016/S1074-7613(02)00475-2.

Davis, J. Muse, and Lalita Ramakrishnan. 2009. "The Role of the Granuloma in Expansion and Dissemination of Early Tuberculous Infection." *Cell* 136(1):37–49. doi: 10.1016/j.cell.2008.11.014.

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.

Denholm, L. 2017. "Genotype Disclosure in the Genomics Era: Roles and Responsibilities." *Australian Veterinary Journal* 95(9):308–16. doi: 10.1111/avj.12606.

Detrich, H. W., M. W. Kieran, F. Y. Chan, L. M. Barone, K. Yee, J. A. Rundstadler, S. Pratt, D. Ransom, and L. I. Zon. 1995. "Intraembryonic Hematopoietic Cell Migration during Vertebrate Development." *Proceedings of the National Academy of Sciences of the United States of America* 92(23):10713–17. doi: 10.1073/pnas.92.23.10713.

Ding, Zhujin, Xiaoheng Zhao, Jixiu Wang, Feng Zhang, Weimin Wang, and Hong Liu. 2019. "Intelectin Mediated Phagocytosis and Killing Activity of Macrophages in Blunt Snout Bream (Megalobrama Amblycephala)." *Fish & Shellfish Immunology* 87:129–35. doi: 10.1016/j.fsi.2019.01.001.

Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. "STAR: Ultrafast Universal RNA-Seq Aligner." *Bioinformatics* 29(1):15–21. doi: 10.1093/bioinformatics/bts635.

Dobson, J. Tristan, Jake Seibert, Evelyn M. Teh, Sahar Da'as, Robert B. Fraser, Barry H. Paw, Tong-Jun Lin, and Jason N. Berman. 2008. "Carboxypeptidase A5 Identifies a Novel Mast Cell Lineage in the Zebrafish Providing New Insight into Mast Cell Fate Determination." *Blood* 112(7):2969–72. doi: 10.1182/blood-2008-03-145011.

Dorhoi, Anca, Geraldine Nouailles, Sabine Jörg, Kristine Hagens, Ellen Heinemann, Lydia Pradl, Dagmar Oberbeck-Müller, Maria Adelaida Duque-Correa, Stephen T. Reece, Jürgen Ruland, Roland Brosch, Jürg Tschopp, Olaf Gross, and Stefan H. E.

Kaufmann. 2012. "Activation of the NLRP3 Inflammasome by Mycobacterium Tuberculosis Is Uncoupled from Susceptibility to Active Tuberculosis." *European Journal of Immunology* 42(2):374–84. doi: 10.1002/eji.201141548.

Dranoff, Glenn. 2004. "Cytokines in Cancer Pathogenesis and Cancer Therapy." *Nature Reviews Cancer* 4(1):11–22. doi: 10.1038/nrc1252.

Driever, W., L. Solnica-Krezel, A. F. Schier, S. C. Neuhauss, J. Malicki, D. L. Stemple, D. Y. Stainier, F. Zwartkruis, S. Abdelilah, Z. Rangini, J. Belak, and C. Boggs. 1996. "A Genetic Screen for Mutations Affecting Embryogenesis in Zebrafish." *Development (Cambridge, England)* 123(Journal Article):37–46.

Dupont, S., and M. Thorndyke. 2007. "Bridging the Regeneration Gap: Insights from Echinoderm Models." *Nature Reviews Genetics* 8(4):320–320. doi: 10.1038/nrg1923-c1.

Eisen, Judith S., and James C. Smith. 2008. "Controlling Morpholino Experiments: Don't Stop Making Antisense." *Development (Cambridge, England)* 135(10):1735–43. doi: 10.1242/dev.001115.

El-Brolosy, Mohamed A., Zacharias Kontarakis, Andrea Rossi, Carsten Kuenne, Stefan Günther, Nana Fukuda, Khrievono Kikhi, Giulia L. M. Boezio, Carter Takacs, Shih-Lei Lai, Ryuichi Fukuda, Claudia Gerri, Antonio J. Giraldez, and Didier Y. R. Stainier. 2019. "Genetic Compensation Triggered by Mutant MRNA Degradation." *Nature* 568(7751):193–97. doi: 10.1038/s41586-019-1064-z.

Ellebedy, Ali H., Christopher Lupfer, Hazem E. Ghoneim, Jennifer DeBeauchamp, Thirumala-Devi Kanneganti, and Richard J. Webby. 2011. "Inflammasome-Independent Role of the Apoptosis-Associated Speck-like Protein Containing CARD (ASC) in the Adjuvant Effect of MF59." *Proceedings of the National Academy of Sciences* 108(7):2927–32.

Ellett, Felix, Luke Pase, John W. Hayman, Alex Andrianopoulos, and Graham J. Lieschke. 2011. "Mpeg1 Promoter Transgenes Direct Macrophage-Lineage Expression in Zebrafish." *Blood* 117(4):e49-56. doi: 10.1182/blood-2010-10-314120.

Elmonem, Mohamed A., Sante Princiero Berlingerio, Lambertus P. van den Heuvel, Peter A. de Witte, Martin Lowe, and Elena N. Levtchenko. 2018. "Genetic Renal Diseases: The Emerging Role of Zebrafish Models." *Cells* 7(9):E130. doi: 10.3390/cells7090130.

Elsalini, Osama A., and Klaus B. Rohr. 2003. "Phenylthiourea Disrupts Thyroid Function in Developing Zebrafish." *Development Genes and Evolution* 212(12):593–98. doi: 10.1007/s00427-002-0279-3.

Espina, Jaime, Carmen G. Feijóo, Camila Solís, and Alvaro Glavic. 2013. "Csrnp1a Is Necessary for the Development of Primitive Hematopoiesis Progenitors in Zebrafish." *PLOS ONE* 8(1):e53858. doi: 10.1371/journal.pone.0053858.

Eve, Alexander M. J., Elsie S. Place, and James C. Smith. 2017. "Comparison of Zebrafish Tmem88a Mutant and Morpholino Knockdown Phenotypes." *PLOS ONE* 12(2):e0172227. doi: 10.1371/journal.pone.0172227.

Fang, Rendong, Ryosuke Uchiyama, Shunsuke Sakai, Hideki Hara, Hiroko Tsutsui, Takashi Suda, Masao Mitsuyama, Ikuo Kawamura, and Kohsuke Tsuchiya. 2019. "ASC and NLRP3 Maintain Innate Immune Homeostasis in the Airway through an Inflammasome-Independent Mechanism." *Mucosal Immunology* 12(5):1092–1103. doi: 10.1038/s41385-019-0181-1.

Farnsworth, Dylan R., Lauren M. Saunders, and Adam C. Miller. 2020. "A Single-Cell Transcriptome Atlas for Zebrafish Development." *Developmental Biology* 459(2):100–108. doi: 10.1016/j.ydbio.2019.11.008.

Farrell, Jeffrey A., Yiqun Wang, Samantha J. Riesenfeld, Karthik Shekhar, Aviv Regev, and Alexander F. Schier. 2018. "Single-Cell Reconstruction of Developmental Trajectories during Zebrafish Embryogenesis." *Science* 360(6392). doi: 10.1126/science.aar3131.

Fatima, Samreen, Anjna Kumari, Meetu Agarwal, Isha Pahuja, Vinod Yadav, Ved Prakash Dwivedi, and Ashima Bhaskar. 2021. "Epigenetic Code during Mycobacterial Infections: Therapeutic Implications for Tuberculosis." *The FEBS Journal*. doi: 10.1111/febs.16170.

Fillatreau, Simon, Adrien Six, Susanna Magadan, Rosario Castro, J. Oriol Sunyer, and Pierre Boudinot. 2013. "The Astonishing Diversity of Ig Classes and B Cell Repertoires in Teleost Fish." *Frontiers in Immunology* 4:28. doi: 10.3389/fimmu.2013.00028.

Flajnik, Martin F., and Masanori Kasahara. 2010. "Origin and Evolution of the Adaptive Immune System: Genetic Events and Selective Pressures." *Nature Reviews Genetics* 11(1):47–59. doi: 10.1038/nrg2703.

Forn-Cuní, Gabriel, Annemarie H. Meijer, and Monica Varela. 2019. "Zebrafish in Inflammasome Research." *Cells* 8(8):901. doi: 10.3390/cells8080901.

Frame, Jenna M., Caroline Kubaczka, Timothy L. Long, Virginie Esain, Rebecca A. Soto, Mariam Hachimi, Ran Jing, Arkadi Shwartz, Wolfram Goessling, George Q. Daley, and Trista E. North. 2020. "Metabolic Regulation of Inflammasome Activity Controls Embryonic Hematopoietic Stem and Progenitor Cell Production." *Developmental Cell* 55(2):133-149.e6. doi: 10.1016/j.devcel.2020.07.015.

Gaj, Thomas, Charles A. Gersbach, and Carlos F. Barbas. 2013. "ZFN, TALEN, and CRISPR/Cas-Based Methods for Genome Engineering." *Trends in Biotechnology* 31(7):397–405. doi: 10.1016/j.tibtech.2013.04.004.

Gans, Ian, Ellen I. Hartig, Shusen Zhu, Andrea R. Tilden, Lucie N. Hutchins, Nathaniel J. Maki, Joel H. Graber, and James A. Coffman. 2020. "Klf9 Is a Key Feedforward Regulator of the Transcriptomic Response to Glucocorticoid Receptor Activity." *Scientific Reports* 10(1):11415. doi: 10.1038/s41598-020-68040-z.

Gasanov, Eugene V., Justyna Jędrychowska, Jacek Kuźnicki, and Vladimir Korzh. 2021. "Evolutionary Context Can Clarify Gene Names: Teleosts as a Case Study." *BioEssays* 43(6):2000258. doi: 10.1002/bies.202000258.

Ge, Rui, Yi Zhou, Rui Peng, Rui Wang, Mi Li, Yunbin Zhang, Chunfu Zheng, and Chen Wang. 2015. "Conservation of the STING-Mediated Cytosolic DNA Sensing Pathway in Zebrafish." *Journal of Virology* 89(15):7696–7706. doi: 10.1128/JVI.01049-15.

Geisler, Robert, Almut Köhler, Thomas Dickmeis, and Uwe Strähle. 2017. "Archiving of Zebrafish Lines Can Reduce Animal Experiments in Biomedical Research." *EMBO Reports* 18(1):1–2. doi: 10.15252/embr.201643561.

Gieger, Christian, Aparna Radhakrishnan, Ana Cvejic, Weihong Tang, Eleonora Porcu, Giorgio Pistis, Jovana Serbanovic-Canic, Ulrich Elling, Alison H. Goodall, Yann Labrune, Lorna M. Lopez, Reedik Mägi, Stuart Meacham, Yukinori Okada, Nicola Pirastu, Rossella Sorice, Alexander Teumer, Katrin Voss, Weihua Zhang, Ramiro Ramirez-Solis, Joshua C. Bis, David Ellinghaus, Martin Gögele, Jouke-Jan Hottenga, Claudia Langenberg, Peter Kovacs, Paul F. O'Reilly, So-Youn Shin, Tõnu Esko, Jaana Hartiala, Stavroula Kanoni, Federico Murgia, Afshin Parsa, Jonathan Stephens, Pim van der Harst, C. Ellen van der Schoot, Hooman Allayee, Antony Attwood, Beverley Balkau, François Bastardot, Saonli Basu, Sebastian E. Baumeister, Ginevra Biino, Lorenzo Bomba, Amélie Bonnefond, François Cambien, John C. Chambers,

Francesco Cucca, Pio D'Adamo, Gail Davies, Rudolf A. de Boer, Eco J. C. de Geus, Angela Döring, Paul Elliott, Jeanette Erdmann, David M. Evans, Mario Falchi, Wei Feng, Aaron R. Folsom, Ian H. Frazer, Quince D. Gibson, Nicole L. Glazer, Chris Hammond, Anna-Liisa Hartikainen, Susan R. Heckbert, Christian Hengstenberg, Micha Hersch, Thomas Illig, Ruth J. F. Loos, Jennifer Jolley, Kay-Tee Khaw, Brigitte Kühnel, Marie-Christine Kyrtsonis, Vasiliki Lagou, Heather Lloyd-Jones, Thomas Lumley, Massimo Mangino, Andrea Maschio, Irene Mateo Leach, Barbara McKnight, Yasin Memari, Braxton D. Mitchell, Grant W. Montgomery, Yusuke Nakamura, Matthias Nauck, Gerjan Navis, Ute Nöthlings, Ilja M. Nolte, David J. Porteous, Anneli Pouta, Peter P. Pramstaller, Janne Pullat, Susan M. Ring, Jerome I. Rotter, Daniela Ruggiero, Aimo Ruokonen, Cinzia Sala, Nilesh J. Samani, Jennifer Sambrook, David Schlessinger, Stefan Schreiber, Heribert Schunkert, James Scott, Nicholas L. Smith, Harold Snieder, John M. Starr, Michael Stumvoll, Atsushi Takahashi, W. H. Wilson Tang, Kent Taylor, Albert Tenesa, Swee Lay Thein, Anke Tönjes, Manuela Uda, Sheila Ulivi, Dirk J. van Veldhuisen, Peter M. Visscher, Uwe Völker, H.-Erich Wichmann, Kerri L. Wiggins, Gonneke Willemsen, Tsun-Po Yang, Jing Hua Zhao, Paavo Zitting, John R. Bradley, George V. Dedoussis, Paolo Gasparini, Stanley L. Hazen, Andres Metspalu, Mario Pirastu, Alan R. Shuldiner, L. Joost van Pelt, Jaap-Jan Zwaginga, Dorret I. Boomsma, Ian J. Deary, Andre Franke, Philippe Froguel, Santhi K. Ganesh, Marjo-Riitta Jarvelin, Nicholas G. Martin, Christa Meisinger, Bruce M. Psaty, Timothy D. Spector, Nicholas J. Wareham, Jan-Willem N. Akkerman, Marina Ciullo, Panos Deloukas, Andreas Greinacher, Steve Jupe, Naoyuki Kamatani, Jyoti Khadake, Jaspal S. Kooner, Josef Penninger, Inga Prokopenko, Derek Stemple, Daniela Toniolo, Lorenz Wernisch, Serena Sanna, Andrew A. Hicks, Augusto Rendon, Manuel A. Ferreira, Willem H. Ouwehand, and Nicole Soranzo. 2011. "New Gene Functions in Megakaryopoiesis and Platelet Formation." *Nature* 480(7376):201–8. doi: 10.1038/nature10659.

Glasauer, Stella M. K., and Stephan C. F. Neuhauss. 2014. "Whole-Genome Duplication in Teleost Fishes and Its Evolutionary Consequences." *Molecular Genetics and Genomics* 289(6):1045–60. doi: 10.1007/s00438-014-0889-2.

Gomes, Margarida C., and Serge Mostowy. 2020. "The Case for Modeling Human Infection in Zebrafish." *Trends in Microbiology* 28(1):10–18. doi: 10.1016/j.tim.2019.08.005.

Gomez-Diaz, Elena, Mireia Jorda, Miguel Angel Peinado, and Ana Rivero. 2012. "Epigenetics of Host-Pathogen Interactions: The Road Ahead and the Road Behind." *Plos Pathogens* 8(11):e1003007. doi: 10.1371/journal.ppat.1003007.

González-Rosa, Juan Manuel, Caroline E. Burns, and C. Geoffrey Burns. 2017. "Zebrafish Heart Regeneration: 15 Years of Discoveries." *Regeneration* 4(3):105–23. doi: 10.1002/reg2.83.

Gore, Aniket V., Laura M. Pillay, Marina Venero Galanternik, and Brant M. Weinstein. 2018. "The Zebrafish: A Fintastic Model for Hematopoietic Development and Disease." *Wiley Interdisciplinary Reviews. Developmental Biology* 7(3):e312. doi: 10.1002/wdev.312.

Gui, Hongsheng, Duco Schriemer, William W. Cheng, Rajendra K. Chauhan, Guillermo Antiňolo, Courtney Berrios, Marta Bleda, Alice S. Brooks, Rutger W. W. Brouwer, Alan J. Burns, Stacey S. Cherny, Joaquin Dopazo, Bart J. L. Eggen, Paola Griseri, Binta Jalloh, Thuy-Linh Le, Vincent C. H. Lui, Berta Luzón-Toro, Ivana Matera, Elly S. W. Ngan, Anna Pelet, Macarena Ruiz-Ferrer, Pak C. Sham, Iain T. Shepherd, Man-Ting So, Yunia Sribudiani, Clara S. M. Tang, Mirjam C. G. N. van den Hout, Herma C. van der Linde, Tjakko J. van Ham, Wilfred F. J. van IJcken, Joke B. G. M. Verheij,

Jeanne Amiel, Salud Borrego, Isabella Ceccherini, Aravinda Chakravarti, Stanislas Lyonnet, Paul K. H. Tam, Maria-Mercè Garcia-Barceló, and Robert M. W. Hofstra. 2017. "Whole Exome Sequencing Coupled with Unbiased Functional Analysis Reveals New Hirschsprung Disease Genes." *Genome Biology* 18(1):48. doi: 10.1186/s13059-017-1174-6.

Haffter, P., M. Granato, M. Brand, M. C. Mullins, M. Hammerschmidt, D. A. Kane, J. Odenthal, F. J. van Eeden, Y. J. Jiang, C. P. Heisenberg, R. N. Kelsh, M. Furutani-Seiki, E. Vogelsang, D. Beuchle, U. Schach, C. Fabian, and C. Nüsslein-Volhard. 1996a. "The Identification of Genes with Unique and Essential Functions in the Development of the Zebrafish, Danio Rerio." *Development (Cambridge, England)* 123(Journal Article):1–36.

Haffter, P., M. Granato, M. Brand, M. C. Mullins, M. Hammerschmidt, D. A. Kane, J. Odenthal, F. J. van Eeden, Y. J. Jiang, C. P. Heisenberg, R. N. Kelsh, M. Furutani-Seiki, E. Vogelsang, D. Beuchle, U. Schach, C. Fabian, and C. Nüsslein-Volhard. 1996b. "The Identification of Genes with Unique and Essential Functions in the Development of the Zebrafish, Danio Rerio." *Development (Cambridge, England)* 123:1–36.

Hall, Chris, Maria Vega Flores, Thilo Storm, Kathy Crosier, and Phil Crosier. 2007. "The Zebrafish Lysozyme C Promoter Drives Myeloid-Specific Expression in Transgenic Fish." *BMC Developmental Biology* 7:42. doi: 10.1186/1471-213X-7-42.

Haney, Michael S., Christopher J. Bohlen, David W. Morgens, James A. Ousey, Amira A. Barkal, C. Kimberly Tsui, Braeden K. Ego, Roni Levin, Roarke A. Kamber, Hannah Collins, Andrew Tucker, Amy Li, Daan Vorselen, Lorenzo Labitigan, Emily Crane, Evan Boyle, Lihua Jiang, Joanne Chan, Esther Rincón, William J. Greenleaf, Billy Li, Michael P. Snyder, Irving L. Weissman, Julie A. Theriot, Sean R. Collins, Ben A. Barres, and Michael C. Bassik. 2018. "Identification of Phagocytosis Regulators Using Magnetic Genome-Wide CRISPR Screens." *Nature Genetics* 50(12):1716–27. doi: 10.1038/s41588-018-0254-1.

Harjula, Sanna-Kaisa E., Anni K. Saralahti, Markus J. T. Ojanen, Tommi Rantapero, Meri I. E. Uusi-Mäkelä, Matti Nykter, Olli Lohi, Mataleena Parikka, and Mika Rämet. 2020. "Characterization of Immune Response against Mycobacterium Marinum Infection in the Main Hematopoietic Organ of Adult Zebrafish (Danio Rerio)." *Developmental and Comparative Immunology* 103(Journal Article). doi: 10.1016/j.dci.2019.103523.

Hartig, Ellen I., Shusen Zhu, Benjamin L. King, and James A. Coffman. 2016. "Cortisol-Treated Zebrafish Embryos Develop into pro-Inflammatory Adults with Aberrant Immune Gene Regulation." *Biology Open* 5(8):1134–41. doi: 10.1242/bio.020065.

Harvey, Steven A., Ian Sealy, Ross Kettleborough, Fruzsina Fenyes, Richard White, Derek Stemple, and James C. Smith. 2013. "Identification of the Zebrafish Maternal and Paternal Transcriptomes." *Development (Cambridge, England)* 140(13):2703–10. doi: 10.1242/dev.095091.

Harvie, Elizabeth A., and Anna Huttenlocher. 2015. "Neutrophils in Host Defense: New Insights from Zebrafish." *Journal of Leukocyte Biology* 98(4):523–37. doi: 10.1189/jlb.4MR1114-524R.

Hay, Michael, David W. Thomas, John L. Craighead, Celia Economides, and Jesse Rosenthal. 2014. "Clinical Development Success Rates for Investigational Drugs." *Nature Biotechnology* 32(1):40–51. doi: 10.1038/nbt.2786.

Henry, Katherine M., Catherine A. Loynes, Moira K. B. Whyte, and Stephen A. Renshaw. 2013. "Zebrafish as a Model for the Study of Neutrophil Biology." *Journal of Leukocyte Biology* 94(4):633–42. doi: 10.1189/jlb.1112594.

Herbomel, P., B. Thisse, and C. Thisse. 1999. "Ontogeny and Behaviour of Early Macrophages in the Zebrafish Embryo." *Development (Cambridge, England)* 126(17):3735–45.

Hernández, Pedro P., Paulina M. Strzelecka, Emmanouil I. Athanasiadis, Dominic Hall, Ana F. Robalo, Catherine M. Collins, Pierre Boudinot, Jean-Pierre Levraud, and Ana Cvejic. 2018. "Single-Cell Transcriptional Analysis Reveals ILC-like Cells in Zebrafish." *Science Immunology* 3(29):eaau5265. doi: 10.1126/sciimmunol.aau5265.

Hiltunen, Anniina E., Salla M. Kangas, Steffen Ohlmeier, Ilkka Pietilä, Jori Hiltunen, Heikki Tanila, Colin McKerlie, Subashika Govindan, Hannu Tuominen, Riitta Kaarteenaho, Mikko Hallman, Johanna Uusimaa, and Reetta Hinttala. 2020. "Variant in NHLRC2 Leads to Increased HnRNP C2 in Developing Neurons and the Hippocampus of a Mouse Model of FINCA Disease." *Molecular Medicine* 26:123. doi: 10.1186/s10020-020-00245-4.

Hinz, John M., Marian F. Laughery, and John J. Wyrick. 2015. "Nucleosomes Inhibit Cas9 Endonuclease Activity in Vitro." *Biochemistry* 54(48):7063–66. doi: 10.1021/acs.biochem.5b01108.

Horlbeck, Max A., Lea B. Witkowsky, Benjamin Guglielmi, Joseph M. Replogle, Luke A. Gilbert, Jacqueline E. Villalta, Sharon E. Torigoe, Robert Tjian, and Jonathan S. Weissman. 2016. "Nucleosomes Impede Cas9 Access to DNA in Vivo and in Vitro." *Elife* 5(Journal Article):e12677. doi: 10.7554/eLife.12677.

Howe, K., C. F. Torroja, J. Torrance, J. E. Collins, S. Humphray, K. McLaren, S. McLaren, I. Sealy, M. Caccamo, C. Churcher, C. Scott, J. C. Barrett, G. J. Rauch, W. Chow, B. Kilian, L. T. Quintais, J. A. Guerra-Assuncao, Y. Zhou, T. Eyre, S. Redmond, R. Banerjee, J. Chi, B. Fu, D. Lloyd, E. Kenyon, J. Almeida-King, J. Loveland, M. Jones, M. Quail, A. Hunt, B. Plumb, C. Clee, K. Oliver, C. Riddle, D. Elliot, G. Threadgold, G. Harden, D. Ware, B. Mortimore, P. Heath, C. Johnson, J. Wood, S. Clark, S. Pelan, M. Smith, R. Glithero, P. Howden, N. Barker, G. Panagiotidis, J. Lovell, H. Beasley, C. Henderson, D. Wright, J. Collins, K. Leung, K. Ambridge, D. Leongamornlert, S. McGuire, G. Barker, M. Kay, E. Gray, M. Humphries, D. Saunders, J. Wallis, S. Hammond, P. Wray, A. Ellington, M. Ellwood, R. Woodmansey, A. Tromans, R. Pandian, R. Anews, A. Kimberley, J. Garnett, R. Hall, C. Bird, I. Gehring, A. Berger, C. M. Dooley, Z. Ersan-Urun, C. Eser, L. Karotki, A. Kirn, J. Konantz, M. Oberlander, S. Rudolph-Geiger, M. Teucke, K. Osoegawa, B. Zhu, F. Yang, N. P. Carter, R. H. A. Plasterk, C. Lee, M. Westerfield, P. J. de Jong, J. H. Postlethwait, D. L. Stemple, C. Lanz, G. Raddatz, and S. C. Schuster. 2013. "The Zebrafish Reference Genome Sequence and Its Relationship to the Human Genome." *Nature (London)* 496(7446):498–503. doi: 10.1038/nature12111.

Hruscha, Alexander, Peter Krawitz, Alexandra Rechenberg, Verena Heinrich, Jochen Hecht, Christian Haass, and Bettina Schmid. 2013. "Efficient CRISPR/Cas9 Genome Editing with Low off-Target Effects in Zebrafish." *Development (Cambridge, England)* 140(24):4982–87. doi: 10.1242/dev.099085.

Hruscha, Alexander, and Bettina Schmid. 2015. "Generation of Zebrafish Models by CRISPR /Cas9 Genome Editing." *Methods in Molecular Biology (Clifton, N.J.)* 1254:341–50. doi: 10.1007/978-1-4939-2152-2_24.

Hubbard, T., D. Barker, E. Birney, G. Cameron, Y. Chen, L. Clark, T. Cox, J. Cuff, V. Curwen, T. Down, R. Durbin, E. Eyras, J. Gilbert, M. Hammond, L. Huminiecki, A. Kasprzyk, H. Lehvaslaiho, P. Lijnzaad, C. Melsopp, E. Mongin, R. Pettett, M. Pocock, S. Potter, A. Rust, E. Schmidt, S. Searle, G. Slater, J. Smith, W. Spooner, A. Stabenau, J. Stalker, E. Stupka, A. Ureta-Vidal, I. Vastrik, and M. Clamp. 2002. "The Ensembl Genome Database Project." *Nucleic Acids Research* 30(1):38–41. doi: 10.1093/nar/30.1.38.

Hubrecht, Robert C., and Elizabeth Carter. 2019. "The 3Rs and Humane Experimental Technique: Implementing Change." *Animals : An Open Access Journal from MDPI* 9(10):754. doi: 10.3390/ani9100754.

Hughes, Kelsey L., Elizabeth T. Abshire, and Aaron C. Goldstrohm. 2018. "Regulatory Roles of Vertebrate Nocturnin: Insights and Remaining Mysteries." *RNA Biology* 15(10):1255–67. doi: 10.1080/15476286.2018.1526541.

Hwang, Woong Y., Yanfang Fu, Deepak Reyon, Morgan L. Maeder, Shengdar Q. Tsai, Jeffry D. Sander, Randall T. Peterson, J. R. Joanna Yeh, and J. Keith Joung. 2013. "Efficient Genome Editing in Zebrafish Using a CRISPR-Cas System." *Nature Biotechnology* 31(3):227–29. doi: 10.1038/nbt.2501.

Irie, Naoki, and Shigeru Kuratani. 2011. "Comparative Transcriptome Analysis Reveals Vertebrate Phylotypic Period during Organogenesis." *Nature Communications* 2(1):248. doi: 10.1038/ncomms1248.

Isaac, R. Stefan, Fuguo Jiang, Jennifer A. Doudna, Wendell A. Lim, Geeta J. Narlikar, and Ricardo Almeida. 2016. "Nucleosome Breathing and Remodeling Constrain CRISPR-Cas9 Function." *Elife* 5(Journal Article):e13450. doi: 10.7554/eLife.13450.

Iwanami, Norimasa. 2014. "Zebrafish as a Model for Understanding the Evolution of the Vertebrate Immune System and Human Primary Immunodeficiency." *Experimental Hematology* 42(8):697–706. doi: 10.1016/j.exphem.2014.05.001.

Iwanami, Norimasa, Divine-Fondzenyuy Lawir, Katarzyna Sikora, Connor O´Meara, Kohei Takeshita, Michael Schorpp, and Thomas Boehm. 2020. "Transgenerational Inheritance of Impaired Larval T Cell Development in Zebrafish." *Nature Communications* 11(1):4505. doi: 10.1038/s41467-020-18289-9.

Jagadeeswaran, P., J. P. Sheehan, F. E. Craig, and D. Troyer. 1999. "Identification and Characterization of Zebrafish Thrombocytes." *British Journal of Haematology* 107(4):731–38. doi: 10.1046/j.1365-2141.1999.01763.x.

Javanmard Khameneh, Hanif, Keith Weng Kit Leong, Andrea Mencarelli, Maurizio Vacca, Bezaleel Mambwe, Kurt Neo, Alice Tay, Francesca Zolezzi, Bernett Lee, and Alessandra Mortellaro. 2019. "The Inflammasome Adaptor ASC Intrinsically Limits CD4+ T-Cell Proliferation to Help Maintain Intestinal Homeostasis." *Frontiers in Immunology* 10(Journal Article). doi: 10.3389/fimmu.2019.01566.

Jinek, Martin, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. 2012. "A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity." *Science (New York, N.Y.)* 337(6096):816–21. doi: 10.1126/science.1225829.

Jorgensen, Ine, Yue Zhang, Bryan A. Krantz, and Edward A. Miao. 2016. "Pyroptosis Triggers Pore-Induced Intracellular Traps (PITs) That Capture Bacteria and Lead to Their Clearance by Efferocytosis." *The Journal of Experimental Medicine* 213(10):2113–28. doi: 10.1084/jem.20151613.

Joris, Marine, Marie Schloesser, Denis Baurain, Marc Hanikenne, Marc Muller, and Patrick Motte. 2017. "Number of Inadvertent RNA Targets for Morpholino Knockdown in Danio Rerio Is Largely Underestimated: Evidence from the Study of Ser/Arg-Rich Splicing Factors." *Nucleic Acids Research* 45(16):9547–57. doi: 10.1093/nar/gkx638.

Jung, Johannes, Sonja C. Buisman, Ellen Weersing, Albertina Dethmers-Ausema, Erik Zwart, Hein Schepers, Mike R. Dekker, Seka S. Lazare, Franziska Hammerl, Yulia Skokova, Susanne M. Kooistra, Karin Klauke, Raymond A. Poot, Leonid V. Bystrykh, and Gerald de Haan. 2019. "CBX7 Induces Self-Renewal of Human Normal and Malignant Hematopoietic Stem and Progenitor Cells by Canonical and Non-Canonical Interactions." *Cell Reports* 26(7):1906-1918.e8. doi: 10.1016/j.celrep.2019.01.050.

Juss, Jatinder K., David House, Augustin Amour, Malcolm Begg, Jurgen Herre, Daniel M. L. Storisteanu, Kim Hoenderdos, Glyn Bradley, Mark Lennon, Charlotte Summers, Edith M. Hessel, Alison Condliffe, and Edwin R. Chilvers. 2016. "Acute Respiratory Distress Syndrome Neutrophils Have a Distinct Phenotype and Are Resistant to Phosphoinositide 3-Kinase Inhibition." *American Journal of Respiratory and Critical Care Medicine* 194(8):961–73. doi: 10.1164/rccm.201509-1818OC.

Kalev-Zylinska, Maggie L., Julia A. Horsfield, Maria Vega C. Flores, John H. Postlethwait, Maria R. Vitas, Andrea M. Baas, Philip S. Crosier, and Kathryn E. Crosier. 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 (Cambridge, England)* 129(8):2015–30.

Kallimasioti-Pazi, Eirini M., Keerthi Thelakkad Chathoth, Gillian C. Taylor, Alison Meynert, Tracy Ballinger, Martijn J. E. Kelder, Sébastien Lalevée, Ildem Sanli, Robert Feil, and Andrew J. Wood. 2018. "Heterochromatin Delays CRISPR-Cas9 Mutagenesis but Does Not Influence the Outcome of Mutagenic DNA Repair." *PLOS Biology* 16(12):e2005595. doi: 10.1371/journal.pbio.2005595.

Kalueff, Allan V., Adam Michael Stewart, and Robert Gerlai. 2014. "Zebrafish as an Emerging Model for Studying Complex Brain Disorders." *Trends in Pharmacological Sciences* 35(2):63–75. doi: 10.1016/j.tips.2013.12.002.

Katoch, Swati, and Vikram Patial. 2021. "Zebrafish: An Emerging Model System to Study Liver Diseases and Related Drug Discovery." *Journal of Applied Toxicology* 41(1):33–51. doi: 10.1002/jat.4031.

Kenyon, Amy, Daria Gavriouchkina, Jernej Zorman, Giorgio Napolitani, Vincenzo Cerundolo, and Tatjana Sauka-Spengler. 2017. "Active Nuclear Transcriptome Analysis Reveals Inflammasome-Dependent Mechanism for Early Neutrophil Response to Mycobacterium Marinum." *Scientific Reports* 7(1):6505. doi: 10.1038/s41598-017-06099-x.

Khaenam, Prasong, Darawan Rinchai, Matthew C. Altman, Laurent Chiche, Surachat Buddhisa, Chidchamai Kewcharoenwong, Duangchan Suwannasaen, Michael Mason, Elizabeth Whalen, Scott Presnell, Wattanachai Susaengrat, Kimberly O'Brien, Quynh-Ahn Nguyen, Vivian Gersuk, Peter S. Linsley, Ganjana Lertmemongkolchai, and Damien Chaussabel. 2014. "A Transcriptomic Reporter Assay Employing Neutrophils to Measure Immunogenic Activity of Septic Patients' Plasma." *Journal of Translational Medicine* 12(1):65. doi: 10.1186/1479-5876-12-65.

Kikuchi, Kazu. 2020. "New Function of Zebrafish Regulatory T Cells in Organ Regeneration." *Current Opinion in Immunology* 63:7–13. doi: 10.1016/j.coi.2019.10.001.

Kimmel, Charles B., William W. Ballard, Seth R. Kimmel, Bonnie Ullmann, and Thomas F. Schilling. 1995. "Stages of Embryonic Development of the Zebrafish." *Developmental Dynamics* 203(3):253–310. doi: 10.1002/aja.1002030302.

Klein, Allon M., Linas Mazutis, Ilke Akartuna, Naren Tallapragada, Adrian Veres, Victor Li, Leonid Peshkin, David A. Weitz, and Marc W. Kirschner. 2015. "Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells." *Cell* 161(5):1187–1201. doi: 10.1016/j.cell.2015.04.044.

Kloosterman, Wigard P., Anne K. Lagendijk, René F. Ketting, Jon D. Moulton, and Ronald H. A. Plasterk. 2007. "Targeted Inhibition of MiRNA Maturation with Morpholinos Reveals a Role for MiR-375 in Pancreatic Islet Development." *PLoS Biology* 5(8):e203. doi: 10.1371/journal.pbio.0050203.

Knight, Spencer C., Liangqi Xie, Wulan Deng, Benjamin Guglielmi, Lea B. Witkowsky, Lana Bosanac, Elisa T. Zhang, Mohamed El Beheiry, Jean-Baptiste Masson, Maxime Dahan, Zhe Liu, Jennifer A. Doudna, and Robert Tjian. 2015. "Dynamics of CRISPR-Cas9 Genome Interrogation in Living Cells." *Science (New York, N.Y.)* 350(6262):823–26. doi: 10.1126/science.aac6572.

Kok, F. O., M. Shin, C. W. Ni, A. Gupta, A. S. Grosse, A. van Impel, B. C. Kirchmaier, J. Peterson-Maduro, G. Kourkoulis, I. Male, D. F. DeSantis, S. Sheppard-Tindell, L. Ebarasi, C. Betsholtz, S. Schulte-Merker, S. A. Wolfe, and N. D. Lawson. 2015. "Reverse Genetic Screening Reveals Poor Correlation between Morpholino-Induced and Mutant Phenotypes in Zebrafish." *Developmental Cell* 32(1):97–108. doi: 10.1016/j.devcel.2014.11.018.

Koo, Ingrid C., Chen Wang, Sridharan Raghavan, J. Hiroshi Morisaki, Jeffery S. Cox, and Eric J. Brown. 2008. "ESX-1-Dependent Cytolysis in Lysosome Secretion and Inflammasome Activation during Mycobacterial Infection." *Cellular Microbiology* 10(9):1866–78. doi: 10.1111/j.1462-5822.2008.01177.x.

Kuri, Paola, Nicole L. Schieber, Thomas Thumberger, Joachim Wittbrodt, Yannick Schwab, and Maria Leptin. 2017. "Dynamics of in Vivo ASC Speck Formation." *The Journal of Cell Biology* 216(9):2891–2909. doi: 10.1083/jcb.201703103.

Kuscu, Cem, Sevki Arslan, Ritambhara Singh, Jeremy Thorpe, and Mazhar Adli. 2014. "Genome-Wide Analysis Reveals Characteristics of off-Target Sites Bound by the Cas9 Endonuclease." *Nature Biotechnology* 32(7):677–83. doi: 10.1038/nbt.2916.

Labun, Kornel, Tessa G. Montague, James A. Gagnon, Summer B. Thyme, and Eivind Valen. 2016. "CHOPCHOP v2: A Web Tool for the next Generation of CRISPR Genome Engineering." *Nucleic Acids Research* 44(W1):272. doi: 10.1093/nar/gkw398.

Lähnemann, David, Johannes Köster, Ewa Szczurek, Davis J. McCarthy, Stephanie C. Hicks, Mark D. Robinson, Catalina A. Vallejos, Kieran R. Campbell, Niko Beerenwinkel, Ahmed Mahfouz, Luca Pinello, Pavel Skums, Alexandros Stamatakis, Camille Stephan-Otto Attolini, Samuel Aparicio, Jasmijn Baaijens, Marleen Balvert, Buys de Barbanson, Antonio Cappuccio, Giacomo Corleone, Bas E. Dutilh, Maria Florescu, Victor Guryev, Rens Holmer, Katharina Jahn, Thamar Jessurun Lobo, Emma M. Keizer, Indu Khatri, Szymon M. Kielbasa, Jan O. Korbel, Alexey M. Kozlov, Tzu-Hao Kuo, Boudewijn P. F. Lelieveldt, Ion I. Mandoiu, John C. Marioni, Tobias Marschall, Felix Mölder, Amir Niknejad, Lukasz Raczkowski, Marcel Reinders, Jeroen de Ridder, Antoine-Emmanuel Saliba, Antonios Somarakis, Oliver Stegle, Fabian J. Theis, Huan Yang, Alex Zelikovsky, Alice C. McHardy, Benjamin J. Raphael, Sohrab P. Shah, and Alexander Schönhuth. 2020. "Eleven Grand Challenges in Single-Cell Data Science." *Genome Biology* 21(1):31. doi: 10.1186/s13059-020-1926-6.

Lai, Cheng-Wei, Hsiao-Ling Chen, Min-Yu Tu, Wei-Yu Lin, Theresa Röhrig, Shang-Hsun Yang, Ying-Wei Lan, Kowit-Yu Chong, and Chuan-Mu Chen. 2017. "A Novel Osteoporosis Model with Ascorbic Acid Deficiency in Akr1A1 Gene Knockout Mice." *Oncotarget* 8(5):7357–69. doi: 10.18632/oncotarget.14458.

Laing, Kerry J., Maureen K. Purcell, James R. Winton, and John D. Hansen. 2008. "A Genomic View of the NOD-like Receptor Family in Teleost Fish: Identification of a Novel NLR Subfamily in Zebrafish." *BMC Evolutionary Biology* 8(1):42. doi: 10.1186/1471-2148-8-42.

Lam, S. H., H. L. Chua, Z. Gong, T. J. Lam, and Y. M. Sin. 2004. "Development and Maturation of the Immune System in Zebrafish, Danio Rerio: A Gene Expression Profiling, in Situ Hybridization and Immunological Study." *Developmental & Comparative Immunology* 28(1):9–28. doi: 10.1016/S0145-305X(03)00103-4.

Lamkanfi, Mohamed, and Vishva M. Dixit. 2011. "Modulation of Inflammasome Pathways by Bacterial and Viral Pathogens." *The Journal of Immunology (1950)* 187(2):597–602. doi: 10.4049/jimmunol.1100229.

Langenau, David M., Adolfo A. Ferrando, David Traver, Jeffery L. Kutok, John-Paul D. Hezel, John P. Kanki, Leonard I. Zon, A. Thomas Look, and Nikolaus 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.

Lareau, Caleb, S. Iyer, D. M. Langenau, and Martin Aryee. 2017. *Single Cell InDrops RNA-Seq Visualization of Adult Zebrafish Whole Kidney Marrow*. Harvard University.

Lavalett, Lelia, Hector Ortega, and Luis F. Barrera. 2020. "Human Alveolar and Splenic Macrophage Populations Display a Distinct Transcriptomic Response to Infection With Mycobacterium Tuberculosis." *Frontiers in Immunology* 11:630. doi: 10.3389/fimmu.2020.00630.

Lawson, Nathan D., and Scot A. Wolfe. 2011. "Forward and Reverse Genetic Approaches for the Analysis of Vertebrate Development in the Zebrafish." *Developmental Cell* 21(1):48–64. doi: 10.1016/j.devcel.2011.06.007.

Lee, Jin-Kyu, Janet Schnee, Mabel Pang, Margreet Wolfert, Linda G. Baum, Kelley W. Moremen, and Michael Pierce. 2001. "Human Homologs of the Xenopus Oocyte Cortical Granule Lectin XL35." *Glycobiology* 11(1):65–73. doi: 10.1093/glycob/11.1.65.

Lesch, Bluma J., Gregoriy A. Dokshin, Richard A. Young, John R. McCarrey, and David C. Page. 2013. "A Set of Genes Critical to Development Is Epigenetically Poised in Mouse Germ Cells from Fetal Stages through Completion of Meiosis." *Proceedings of the National Academy of Sciences of the United States of America* 110(40):16061–66. doi: 10.1073/pnas.1315204110.

Li, Ding, YiRan Liu, Shuai Hao, Bo Chen, and AnHai Li. 2020. "Mining Database for the Clinical Significance and Prognostic Value of CBX Family in Skin Cutaneous Melanoma." *Journal of Clinical Laboratory Analysis* 34(12):e23537. doi: 10.1002/jcla.23537.

Li, Dong, Michael E. March, Alvaro Gutierrez-Uzquiza, Charlly Kao, Christoph Seiler, Erin Pinto, Leticia S. Matsuoka, Mark R. Battig, Elizabeth J. Bhoj, Tara L. Wenger, Lifeng Tian, Nora Robinson, Tiancheng Wang, Yichuan Liu, Brant M. Weinstein, Matthew Swift, Hyun Min Jung, Courtney N. Kaminski, Rosetta Chiavacci, Jonathan A. Perkins, Michael A. Levine, Patrick M. A. Sleiman, Patricia J. Hicks, Janet T. Strausbaugh, Jean B. Belasco, Yoav Dori, and Hakon Hakonarson. 2019. "ARAF

Recurrent Mutation Causes Central Conducting Lymphatic Anomaly Treatable with a MEK Inhibitor." *Nature Medicine* 25(7):1116–22. doi: 10.1038/s41591-019-0479-2.

Li, Hongyi, Yang Yang, Weiqi Hong, Mengyuan Huang, Min Wu, and Xia Zhao. 2020. "Applications of Genome Editing Technology in the Targeted Therapy of Human Diseases: Mechanisms, Advances and Prospects." *Signal Transduction and Targeted Therapy* 5(1):1–23. doi: 10.1038/s41392-019-0089-y.

Li, Jiang-Yuan, Ke Gao, Tong Shao, Dong-Dong Fan, Chong-Bin Hu, Cen-Cen Sun, Wei-Ren Dong, Ai-Fu Lin, Li-Xin Xiang, and Jian-Zhong Shao. 2018. "Characterization of an NLRP1 Inflammasome from Zebrafish Reveals a Unique Sequential Activation Mechanism Underlying Inflammatory Caspases in Ancient Vertebrates." *Journal of Immunology (Baltimore, Md.: 1950)* 201(7):1946–66. doi: 10.4049/jimmunol.1800498.

Li, Jiang-Yuan, Yue-Yi Wang, Tong Shao, Dong-Dong Fan, Ai-Fu Lin, Li-Xin Xiang, and Jian-Zhong Shao. 2019. "The Zebrafish NLRP3 Inflammasome Has Functional Roles in ASC-Dependent Interleukin-1β Maturation and Gasdermin E–Mediated Pyroptosis." *Journal of Biological Chemistry* 295(4):1120. doi: 10.1074/jbc.ra119.011751.

Li, Lin, Shaoqing Yang, Yanli Zhang, Dongrui Ji, Zuolin Jin, and Xiaohong Duan. 2018. "ATP6V1H Regulates the Growth and Differentiation of Bone Marrow Stromal Cells." *Biochemical and Biophysical Research Communications* 502(1):84–90. doi: 10.1016/j.bbrc.2018.05.124.

Li, Yajuan, Yi Huang, Xiaocong Cao, Xueying Yin, Xiangyu Jin, Sheng Liu, Jiansheng Jiang, Wei Jiang, Tsan Sam Xiao, Rongbin Zhou, Gang Cai, Bing Hu, and Tengchuan Jin. 2018. "Functional and Structural Characterization of Zebrafish ASC." *The FEBS Journal* 285(14):2691–2707. doi: 10.1111/febs.14514.

Li, Yajuan, Yuelong Li, Xiaocong Cao, Xiangyu Jin, and Tengchuan Jin. 2017. "Pattern Recognition Receptors in Zebrafish Provide Functional and Evolutionary Insight into Innate Immune Signaling Pathways." *Cellular & Molecular Immunology* 14(1):80–89. doi: 10.1038/cmi.2016.50.

Lieschke, G. J., A. C. Oates, M. O. Crowhurst, A. C. Ward, and J. E. Layton. 2001. "Morphologic and Functional Characterization of Granulocytes and Macrophages in Embryonic and Adult Zebrafish." *Blood* 98(10):3087–96. doi: 10.1182/blood.v98.10.3087.

Lieschke, Graham J., and Peter D. Currie. 2007. "Animal Models of Human Disease: Zebrafish Swim into View." *Nature Reviews. Genetics* 8(5):353–67. doi: 10.1038/nrg2091.

Lin, Hui-Feng, David Traver, Hao Zhu, Kimberly Dooley, Barry H. Paw, Leonard I. Zon, and Robert I. Handin. 2005. "Analysis of Thrombocyte Development in CD41-GFP Transgenic Zebrafish." *Blood* 106(12):3803–10. doi: 10.1182/blood-2005-01-0179.

Liscovitch-Brauer, Noa, Antonino Montalbano, Jiale Deng, Alejandro Méndez-Mancilla, Hans-Hermann Wessels, Nicholas G. Moss, Chia-Yu Kung, Akash Sookdeo, Xinyi Guo, Evan Geller, Suma Jaini, Peter Smibert, and Neville E. Sanjana. 2021. "Profiling the Genetic Determinants of Chromatin Accessibility with Scalable Single-Cell CRISPR Screens." *Nature Biotechnology* 1–8. doi: 10.1038/s41587-021-00902-x.

Liu, Chunfa, Ruichao Yue, Yang Yang, Yongyong Cui, Lifeng Yang, Deming Zhao, and Xiangmei Zhou. 2016. "AIM2 Inhibits Autophagy and IFN-β Production during M. Bovis Infection." *Oncotarget* 7(30):46972–87. doi: 10.18632/oncotarget.10503.

Liu, Li-ping, Jian-hai Xiang, Bo Dong, Pavanasam Natarajan, Kui-jie Yu, and Nan-er Cai. 2006. "Ciona Intestinalis as an Emerging Model Organism: Its Regeneration under

Controlled Conditions and Methodology for Egg Dechorionation." *Journal of Zhejiang University. Science. B* 7(6):467–74. doi: 10.1631/jzus.2006.B0467.

Liu, Xingjun, Yue-Sheng Li, Susan A. Shinton, Jennifer Rhodes, Lingjuan Tang, Hui Feng, Cicely A. Jette, A. Thomas Look, Kyoko Hayakawa, and Richard R. Hardy. 2017. "Zebrafish B Cell Development without a Pre–B Cell Stage, Revealed by CD79 Fluorescence Reporter Transgenes." *The Journal of Immunology Author Choice* 199(5):1706–15. doi: 10.4049/jimmunol.1700552.

Liu, Zhi-Fei, Jian-Fei Ji, Xiao-Feng Jiang, Tong Shao, Dong-Dong Fan, Xin-Hang Jiang, Ai-Fu Lin, Li-Xin Xiang, and Jian-Zhong Shao. 2020. "Characterization of CGAS Homologs in Innate and Adaptive Mucosal Immunities in Zebrafish Gives Evolutionary Insights into CGAS-STING Pathway." *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 34(6):7786–7809. doi: 10.1096/fj.201902833R.

Löhr, Heiko, and Matthias Hammerschmidt. 2011. "Zebrafish in Endocrine Systems: Recent Advances and Implications for Human Disease." *Annual Review of Physiology* 73:183–211. doi: 10.1146/annurev-physiol-012110-142320.

Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. "Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2." *Genome Biology* 15(12):550. doi: 10.1186/s13059-014-0550-8.

Lugo-Villarino, Geanncarlo, Keir M. Balla, David L. Stachura, Karina Bañuelos, Miriam B. F. Werneck, and David Traver. 2010. "Identification of Dendritic Antigen-Presenting Cells in the Zebrafish." *Proceedings of the National Academy of Sciences of the United States of America* 107(36):15850–55. doi: 10.1073/pnas.1000494107.

Ma, Jialu, Shasha Zhao, Xiao Gao, Rui Wang, Juan Liu, Xiangmei Zhou, and Yang Zhou. 2021. "The Roles of Inflammasomes in Host Defense against Mycobacterium Tuberculosis." *Pathogens* 10(2). doi: 10.3390/pathogens10020120.

Ma, Zhipeng, Peipei Zhu, Hui Shi, Liwei Guo, Qinghe Zhang, Yanan Chen, Shuming Chen, Zhe Zhang, Jinrong Peng, and Jun Chen. 2019. "PTC-Bearing MRNA Elicits a Genetic Compensation Response via Upf3a and COMPASS Components." *Nature* 568(7751):259–63. doi: 10.1038/s41586-019-1057-y.

Macosko, Evan Z., Anindita Basu, Rahul Satija, James Nemesh, Karthik Shekhar, Melissa Goldman, Itay Tirosh, Allison R. Bialas, Nolan Kamitaki, Emily M. Martersteck, John J. Trombetta, David A. Weitz, Joshua R. Sanes, Alex K. Shalek, Aviv Regev, and Steven A. McCarroll. 2015. "Highly Parallel Genome-Wide Expression Profiling of Individual Cells Using Nanoliter Droplets." *Cell* 161(5):1202–14. doi: 10.1016/j.cell.2015.05.002.

MacRae, Calum A., and Randall T. Peterson. 2015. "Zebrafish as Tools for Drug Discovery." *Nature Reviews Drug Discovery* 14(10):721–31. doi: 10.1038/nrd4627.

Makarova, Kira S., Nick V. Grishin, Svetlana A. Shabalina, Yuri I. Wolf, and Eugene V. Koonin. 2006. "A Putative RNA-Interference-Based Immune System in Prokaryotes: Computational Analysis of the Predicted Enzymatic Machinery, Functional Analogies with Eukaryotic RNAi, and Hypothetical Mechanisms of Action." *Biology Direct* 1:7. doi: 10.1186/1745-6150-1-7.

Makarova, Kira S., and Eugene V. Koonin. 2015. "Annotation and Classification of CRISPR-Cas Systems." *Methods in Molecular Biology (Clifton, N.J.)* 1311:47–75. doi: 10.1007/978-1-4939-2687-9_4.

Marianes, Alexis E., and Anastasia M. Zimmerman. 2011. "Targets of Somatic Hypermutation within Immunoglobulin Light Chain Genes in Zebrafish." *Immunology* 132(2):240–55. doi: 10.1111/j.1365-2567.2010.03358.x.

Martinon, Fabio, Kimberly Burns, and Jü Rg Tschopp. 2002. "The Inflammasome: A Molecular Platform Triggering Activation of Inflammatory Caspases and Processing of ProIL- That They Possess Several Distinct Protein/Protein Inter- Action Domains Which Are Used to Assemble Large Multi- Component Complexes. Apaf-1, for Example, Contains an N-Terminal CARD Followed by a NBS/Self-Oligomer-Ization Domain and a C-Terminal WD-40 Repeat (Jaro." *Molecular Cell* 10(Generic):417.

Martinon, Fabio, Annick Mayor, and Jürg Tschopp. 2009. "The Inflammasomes: Guardians of the Body." *Annual Review of Immunology* 27:229–65. doi: 10.1146/annurev.immunol.021908.132715.

Master, Sharon S., Silvana K. Rampini, Alexander S. Davis, Christine Keller, Stefan Ehlers, Burkhard Springer, Graham S. Timmins, Peter Sander, and Vojo Deretic. 2008. "Mycobacterium Tuberculosis Prevents Inflammasome Activation." *Cell Host & Microbe* 3(4):224–32. doi: 10.1016/j.chom.2008.03.003.

Mayer-Barber, Katrin D., Daniel L. Barber, Kevin Shenderov, Sandra D. White, Mark S. Wilson, Allen Cheever, David Kugler, Sara Hieny, Patricia Caspar, Gabriel Núñez, Dirk Schlueter, Richard A. Flavell, Fayyaz S. Sutterwala, and Alan Sher. 2010. "Caspase-1 Independent IL-1beta Production Is Critical for Host Resistance to Mycobacterium Tuberculosis and Does Not Require TLR Signaling in Vivo." *Journal of Immunology (Baltimore, Md.: 1950)* 184(7):3326–30. doi: 10.4049/jimmunol.0904189.

Mazzolini, Julie, Sigrid Le Clerc, Gregoire Morisse, Cédric Coulonges, Laura E. Kuil, Tjakko J. van Ham, Jean-François Zagury, and Dirk Sieger. 2020. "Gene Expression Profiling Reveals a Conserved Microglia Signature in Larval Zebrafish." *Glia* 68(2):298–315. doi: 10.1002/glia.23717.

McElvania Tekippe, Erin, Irving C. Allen, Paul D. Hulseberg, Jonathan T. Sullivan, Jessica R. McCann, Matyas Sandor, Miriam Braunstein, and Jenny P. Y. Ting. 2010. "Granuloma Formation and Host Defense in Chronic Mycobacterium Tuberculosis Infection Requires PYCARD/ASC but Not NLRP3 or Caspase-1." *PloS One* 5(8):e12320. doi: 10.1371/journal.pone.0012320.

McKenna, Aaron, Gregory M. Findlay, James A. Gagnon, Marshall S. Horwitz, Alexander F. Schier, and Jay Shendure. 2016. "Whole-Organism Lineage Tracing by Combinatorial and Cumulative Genome Editing." *Science* 353(6298). doi: 10.1126/science.aaf7907.

Meijer, Annemarie H. 2016. "Protection and Pathology in TB: Learning from the Zebrafish Model." *Seminars in Immunopathology* 38(2):261–73. doi: 10.1007/s00281-015-0522-4.

Meijer, Annemarie H., and Herman P. Spaink. 2011. "Host-Pathogen Interactions Made Transparent with the Zebrafish Model." *Current Drug Targets* 12(7):1000–1017. doi: 10.2174/138945011795677809.

Meunier, Etienne, and Petr Broz. 2017. "Evolutionary Convergence and Divergence in NLR Function and Structure." *Trends in Immunology* 38(10):744–57. doi: 10.1016/j.it.2017.04.005.

Mihaljevic, Ivan, Marta Popovic, Roko Zaja, and Tvrtko Smital. 2016. "Phylogenetic, Syntenic, and Tissue Expression Analysis of Slc22 Genes in Zebrafish (Danio Rerio)." *BMC Genomics* 17(1):626. doi: 10.1186/s12864-016-2981-y.

Mitra, Suman, Ayham Alnabulsi, Chris J. Secombes, and Steve Bird. 2010. "Identification and Characterization of the Transcription Factors Involved in T-Cell Development, t-Bet, Stat6 and Foxp3, within the Zebrafish, Danio Rerio." *The FEBS Journal* 277(1):128–47. doi: 10.1111/j.1742-4658.2009.07460.x.

Miyaoka, Yuichiro, Jennifer R. Berman, Samantha B. Cooper, Steven J. Mayerl, Amanda H. Chan, Bin Zhang, George A. Karlin-Neumann, and Bruce R. Conklin. 2016. "Systematic Quantification of HDR and NHEJ Reveals Effects of Locus, Nuclease, and Cell Type on Genome-Editing." *Scientific Reports* 6(1):23549. doi: 10.1038/srep23549.

Moore, Finola E., Elaine G. Garcia, Riadh Lobbardi, Esha Jain, Qin Tang, John C. Moore, Mauricio Cortes, Aleksey Molodtsov, Melissa Kasheta, Christina C. Luo, Amaris J. Garcia, Ravi Mylvaganam, Jeffrey A. Yoder, Jessica S. Blackburn, Ruslan I. Sadreyev, Craig J. Ceol, Trista E. North, and David M. Langenau. 2016. "Single-Cell Transcriptional Analysis of Normal, Aberrant, and Malignant Hematopoiesis in Zebrafish." *The Journal of Experimental Medicine* 213(6):979–92. doi: 10.1084/jem.20152013.

Morcos, Paul A., Alexandra C. Vincent, and Jon D. Moulton. 2015. "Gene Editing Versus Morphants." *Zebrafish* 12(5):319. doi: 10.1089/zeb.2015.1114.

Moreira, Jôsimar D., Bjørn E. V. Koch, Suzanne van Veen, Kimberley V. Walburg, Frank Vrieling, Tânia Mara Pinto Dabés Guimarães, Annemarie H. Meijer, Herman P. Spaink, Tom H. M. Ottenhoff, Mariëlle C. Haks, and Matthias T. Heemskerk. 2020. "Functional Inhibition of Host Histone Deacetylases (HDACs) Enhances in Vitro and in Vivo Anti-Mycobacterial Activity in Human Macrophages and in Zebrafish." *Frontiers in Immunology* 11:36. doi: 10.3389/fimmu.2020.00036.

Moreno-Mateos, Miguel A., Charles E. Vejnar, Jean-Denis Beaudoin, Juan P. Fernandez, Emily K. Mis, Mustafa K. Khokha, and Antonio J. Giraldez. 2015. "CRISPRscan: Designing Highly Efficient SgRNAs for CRISPR-Cas9 Targeting in Vivo." *Nature Methods* 12(10):982–88. doi: 10.1038/nmeth.3543.

Mouton, Stijn, Jakub Wudarski, Magda Grudniewska, and Eugene Berezikov. 2018. "The Regenerative Flatworm Macrostomum Lignano, a Model Organism with High Experimental Potential." *The International Journal of Developmental Biology* 62(6-7–8):551–58. doi: 10.1387/ijdb.180077eb.

Mural, Richard J., Mark D. Adams, Eugene W. Myers, Hamilton O. Smith, George L. Gabor Miklos, Ron Wides, Aaron Halpern, Peter W. Li, Granger G. Sutton, Joe Nadeau, Steven L. Salzberg, Robert A. Holt, Chinnappa D. Kodira, Fu Lu, Lin Chen, Zuoming Deng, Carlos C. Evangelista, Weiniu Gan, Thomas J. Heiman, Jiayin Li, Zhenya Li, Gennady V. Merkulov, Natalia V. Milshina, Ashwinikumar K. Naik, Rong Qi, Bixiong Chris Shue, Aihui Wang, Jian Wang, Xin Wang, Xianghe Yan, Jane Ye, Shibu Yooseph, Qi Zhao, Liansheng Zheng, Shiaoping C. Zhu, Kendra Biddick, Randall Bolanos, Arthur L. Delcher, Ian M. Dew, Daniel Fasulo, Michael J. Flanigan, Daniel H. Huson, Saul A. Kravitz, Jason R. Miller, Clark M. Mobarry, Knut Reinert, Karin A. Remington, Qing Zhang, Xiangqun H. Zheng, Deborah R. Nusskern, Zhongwu Lai, Yiding Lei, Wenyan Zhong, Alison Yao, Ping Guan, Rui-Ru Ji, Zhiping Gu, Zhen-Yuan Wang, Fei Zhong, Chunlin Xiao, Chia-Chien Chiang, Mark Yandell, Jennifer R. Wortman, Peter G. Amanatides, Suzanne L. Hladun, Eric C. Pratts, Jeffery E. Johnson, Kristina L. Dodson, Kerry J. Woodford, Cheryl A. Evans, Barry Gropman, Douglas B. Rusch, Eli Venter, Mei Wang, Thomas J. Smith, Jarrett T. Houck, Donald E. Tompkins, Charles Haynes, Debbie Jacob, Soo H. Chin, David R.

Allen, Carl E. Dahlke, Robert Sanders, Kelvin Li, Xiangjun Liu, Alexander A. Levitsky, William H. Majoros, Quan Chen, Ashley C. Xia, John R. Lopez, Michael T. Donnelly, Matthew H. Newman, Anna Glodek, Cheryl L. Kraft, Marc Nodell, Feroze Ali, Hui-Jin An, Danita Baldwin-Pitts, Karen Y. Beeson, Shuang Cai, Mark Carnes, Amy Carver, Parris M. Caulk, Angela Center, Yen-Hui Chen, Ming-Lai Cheng, My D. Coyne, Michelle Crowder, Steven Danaher, Lionel B. Davenport, Raymond Desilets, Susanne M. Dietz, Lisa Doup, Patrick Dullaghan, Steven Ferriera, Carl R. Fosler, Harold C. Gire, Andres Gluecksmann, Jeannine D. Gocayne, Jonathan Gray, Brit Hart, Jason Haynes, Jeffery Hoover, Tim Howland, Chinyere Ibegwam, Mena Jalali, David Johns, Leslie Kline, Daniel S. Ma, Steven MacCawley, Anand Magoon, Felecia Mann, David May, Tina C. McIntosh, Somil Mehta, Linda Moy, Mee C. Moy, Brian J. Murphy, Sean D. Murphy, Keith A. Nelson, Zubeda Nuri, Kimberly A. Parker, Alexandre C. Prudhomme, Vinita N. Puri, Hina Qureshi, John C. Raley, Matthew S. Reardon, Megan A. Regier, Yu-Hui C. Rogers, Deanna L. Romblad, Jakob Schutz, John L. Scott, Richard Scott, Cynthia D. Sitter, Michella Smallwood, Arlan C. Sprague, Erin Stewart, Renee V. Strong, Ellen Suh, Karena Sylvester, Reginald Thomas, Ni Ni Tint, Christopher Tsonis, Gary Wang, George Wang, Monica S. Williams, Sherita M. Williams, Sandra M. Windsor, Keriellen Wolfe, Mitchell M. Wu, Jayshree Zaveri, Kabir Chaturvedi, Andrei E. Gabrielian, Zhaoxi Ke, Jingtao Sun, Gangadharan Subramanian, J. Craig Venter, Cynthia M. Pfannkoch, Mary Barnstead, and Lisa D. Stephenson. 2002. "A Comparison of Whole-Genome Shotgun-Derived Mouse Chromosome 16 and the Human Genome." *Science (New York, N.Y.)* 296(5573):1661–71. doi: 10.1126/science.1069193.

Murdoch, Caitlin C., and John F. Rawls. 2019. "Commensal Microbiota Regulate Vertebrate Innate Immunity-Insights From the Zebrafish." *Frontiers in Immunology* 10:2100. doi: 10.3389/fimmu.2019.02100.

Myllymaki, Henna, Mirja Niskanen, Hanna Luukinen, Mataleena Parikka, and Mika Rämet. 2018. "Identification of Protective Postexposure Mycobacterial Vaccine Antigens Using an Immunosuppression-Based Reactivation Model in the Zebrafish." *Disease Models & Mechanisms* 11(3):UNSP dmm033175. doi: 10.1242/dmm.033175.

Myllymaki, Henna, Mirja Niskanen, Kaisa E. Oksanen, and Mika Rämet. 2015. "Animal Models in Tuberculosis Research - Where Is the Beef?" *Expert Opinion on Drug Discovery* 10(8):871–83. doi: 10.1517/17460441.2015.1049529.

Najafpour, Babak, João C. R. Cardoso, Adelino V. M. Canário, and Deborah M. Power. 2020. "Specific Evolution and Gene Family Expansion of Complement 3 and Regulatory Factor H in Fish." *Frontiers in Immunology* 11:2945. doi: 10.3389/fimmu.2020.568631.

Nasiadka, Andrzej, and Matthew D. Clark. 2012. "Zebrafish Breeding in the Laboratory Environment." *ILAR Journal* 53(2):161–68. doi: 10.1093/ilar.53.2.161.

Netea, Mihai G., Frank L. van de Veerdonk, Jos W. M. van der Meer, Charles A. Dinarello, and Leo A. B. Joosten. 2015. "Inflammasome-Independent Regulation of IL-1-Family Cytokines." *Annual Review of Immunology* 33(1):49–77. doi: 10.1146/annurev-immunol-032414-112306.

Nguyen-Chi, Mai, Béryl Laplace-Builhe, Jana Travnickova, Patricia Luz-Crawford, Gautier Tejedor, Quang Tien Phan, Isabelle Duroux-Richard, Jean-Pierre Levraud, Karima Kissa, Georges Lutfalla, Christian Jorgensen, and Farida Djouad. 2015. "Identification of Polarized Macrophage Subsets in Zebrafish." *ELife* 4:e07288. doi: 10.7554/eLife.07288.

Nonnecke, Eric B., Patricia A. Castillo, Amanda E. Dugan, Faisal Almalki, Mark A. Underwood, Carol A. De La Motte, Weirong Yuan, Wuyuan Lu, Bo Shen, Malin E. V. Johansson, Laura L. Kiessling, Edward J. Hollox, Bo Lönnerdal, and Charles L. Bevins. 2021. "Human Intelectin-1 (ITLN1) Genetic Variation and Intestinal Expression." *Scientific Reports* 11(1):12889. doi: 10.1038/s41598-021-92198-9.

Norton, William, and Laure Bally-Cuif. 2010. "Adult Zebrafish as a Model Organism for Behavioural Genetics." *BMC Neuroscience* 11(1):90. doi: 10.1186/1471-2202-11-90.

Noubade, Rajkumar, Kit Wong, Naruhisa Ota, Sascha Rutz, Celine Eidenschenk, Patricia A. Valdez, Jiabing Ding, Ivan Peng, Andrew Sebrell, Patrick Caplazi, Jason DeVoss, Robert H. Soriano, Tao Sai, Rongze Lu, Zora Modrusan, Jason Hackney, and Wenjun Ouyang. 2014. "NRROS Negatively Regulates Reactive Oxygen Species during Host Defence and Autoimmunity." *Nature* 509(7499):235–39. doi: 10.1038/nature13152.

Nüsslein-Volhard, C., and E. Wieschaus. 1980. "Mutations Affecting Segment Number and Polarity in Drosophila." *Nature* 287(5785):795–801. doi: 10.1038/287795a0.

Nüsslein-Volhard, Christiane. 2012. "The Zebrafish Issue of Development." *Development (Cambridge, England)* 139(22):4099–4103. doi: 10.1242/dev.085217.

Oates, A. C., A. E. Bruce, and R. K. Ho. 2000. "Too Much Interference: Injection of Double-Stranded RNA Has Nonspecific Effects in the Zebrafish Embryo." *Developmental Biology* 224(1):20–28. doi: 10.1006/dbio.2000.9761.

Oehlers, S. H., M. R. Cronan, N. R. Scott, M. I. Thomas, K. S. Okuda, E. M. Walton, R. W. Beerman, P. S. Crosier, and D. M. Tobin. 2015. "Interception of Host Angiogenic Signalling Limits Mycobacterial Growth." (Journal Article).

Ohnesorge, Nils, Céline Heinl, and Lars Lewejohann. 2021. "Current Methods to Investigate Nociception and Pain in Zebrafish." *Frontiers in Neuroscience* 15:378. doi: 10.3389/fnins.2021.632634.

Ohno, Susumu. 1970. *Evolution by Gene Duplication*. 1st ed. Springer, Berlin, Heidelberg.

Oksanen, Kaisa E., Nicholas J. A. Halfpenny, Eleanor Sherwood, Sanna-Kaisa E. Harjula, Milka M. Hammaren, Maarit J. Ahava, Elina T. Pajula, Marika J. Lahtinen, Mataleena Parikka, and Mika Rämet. 2013. "An Adult Zebrafish Model for Preclinical Tuberculosis Vaccine Development." *Vaccine* 31(45):5202–9. doi: 10.1016/j.vaccine.2013.08.093.

Oksanen, Kaisa E., Henna Myllymaki, Maarit J. Ahava, Leena Makinen, Mataleena Parikka, and Mika Rämet. 2016. "DNA Vaccination Boosts Bacillus Calmette-Guerin Protection against Mycobacterial Infection in Zebrafish." *Developmental and Comparative Immunology* 54(1):89–96. doi: 10.1016/j.dci.2015.09.001.

Outtandy, Priya, Claire Russell, Robert Kleta, and Detlef Bockenhauer. 2019. "Zebrafish as a Model for Kidney Function and Disease." *Pediatric Nephrology* 34(5):751–62. doi: 10.1007/s00467-018-3921-7.

Pääbo, Svante. 2001. "The Human Genome and Our View of Ourselves." *Science* 291(5507):1219–20. doi: 10.1126/science.1056972.

Paakkola, Teija, Kari Salokas, Ilkka Miinalainen, Siri Lehtonen, Aki Manninen, Mika Kaakinen, Lloyd W. Ruddock, Markku Varjosalo, Riitta Kaarteenaho, Johanna Uusimaa, and Reetta Hinttala. 2018. "Biallelic Mutations in Human NHLRC2 Enhance Myofibroblast Differentiation in FINCA Disease." *Human Molecular Genetics* 27(24):4288–4302. doi: 10.1093/hmg/ddy298.

Pagán, Antonio J., and Lalita Ramakrishnan. 2015. "Immunity and Immunopathology in the Tuberculous Granuloma." *Cold Spring Harbor Perspectives in Medicine* 5(9):a018499. doi: 10.1101/cshperspect.a018499.

Page, Dawne M., Valerie Wittamer, Julien Y. Bertrand, Kanako L. Lewis, David N. Pratt, Noemi Delgado, Sarah E. Schale, Caitlyn McGue, Bradley H. Jacobsen, Alyssa Doty, Yvonne Pao, Hongbo Yang, Neil C. Chi, Brad G. Magor, and David Traver. 2013. "An Evolutionarily Conserved Program of B-Cell Development and Activation in Zebrafish." *Blood* 122(8):e1–11. doi: 10.1182/blood-2012-12-471029.

Palti, Yniv. 2011. "Toll-like Receptors in Bony Fish: From Genomics to Function." *Developmental & Comparative Immunology* 35(12):1263–72. doi: 10.1016/j.dci.2011.03.006.

Parey, Elise, Alexandra Louis, Cédric Cabau, Yann Guiguen, Hugues Roest Crollius, and Camille Berthelot. 2020. "Synteny-Guided Resolution of Gene Trees Clarifies the Functional Impact of Whole-Genome Duplications." *Molecular Biology and Evolution* 37(11):3324–37. doi: 10.1093/molbev/msaa149.

Parichy, David M., Michael R. Elizondo, Margaret G. Mills, Tiffany N. Gordon, and Raymond E. Engeszer. 2009. "Normal Table of Post-Embryonic Zebrafish Development: Staging by Externally Visible Anatomy of the Living Fish." *Developmental Dynamics : An Official Publication of the American Association of Anatomists* 238(12):2975–3015. doi: 10.1002/dvdy.22113.

Parikka, Mataleena, Milka M. Hammaren, Sanna-Kaisa E. Harjula, Nicholas J. A. Halfpenny, Kaisa E. Oksanen, Marika J. Lahtinen, Elina T. Pajula, Antti Iivanainen, Marko Pesu, and Mika Rämet. 2012. "Mycobacterium Marinum Causes a Latent Infection That Can Be Reactivated by Gamma Irradiation in Adult Zebrafish." *Plos Pathogens* 8(9). doi: 10.1371/journal.ppat.1002944.

Patton, E. Elizabeth, Leonard I. Zon, and David M. Langenau. 2021. "Zebrafish Disease Models in Drug Discovery: From Preclinical Modelling to Clinical Trials." *Nature Reviews Drug Discovery* 20(8):611–28. doi: 10.1038/s41573-021-00210-8.

Pemberton, Alan D., Pamela A. Knight, John Gamble, William H. Colledge, Jin-Kyu Lee, Michael Pierce, and Hugh R. P. Miller. 2004. "Innate BALB/c Enteric Epithelial Responses to Trichinella Spiralis: Inducible Expression of a Novel Goblet Cell Lectin, Intelectin-2, and Its Natural Deletion in C57BL/10 Mice." *The Journal of Immunology* 173(3):1894–1901. doi: 10.4049/jimmunol.173.3.1894.

Peng, Jinrong. 2019. "Gene Redundancy and Gene Compensation: An Updated View." *Journal of Genetics and Genomics* 46(7):329–33. doi: 10.1016/j.jgg.2019.07.001.

Pereiro, P., M. Varela, P. Diaz-Rosales, A. Romero, S. Dios, A. Figueras, and B. Novoa. 2015. "Zebrafish Nk-Lysins: First Insights about Their Cellular and Functional Diversification." *Developmental & Comparative Immunology* 51(1):148–59. doi: 10.1016/j.dci.2015.03.009.

Perino, Matteo, and Gert Jan C. Veenstra. 2016. "Chromatin Control of Developmental Dynamics and Plasticity." *Developmental Cell* 38(6):610–20. doi: 10.1016/j.devcel.2016.08.004.

Pham, Duc-Hung, Changwen Zhang, and Chunyue Yin. 2017. "Using Zebrafish to Model Liver Diseases-Where Do We Stand?" *Current Pathobiology Reports* 5(2):207–21. doi: 10.1007/s40139-017-0141-y.

Phelps, Hilary A., and Melody N. Neely. 2005. "Evolution of the Zebrafish Model: From Development to Immunity and Infectious Disease." *ZEBRAFISH* 2(2):87–103.

Postlethwait, John H., Ian G. Woods, Phuong Ngo-Hazelett, Yi-Lin Yan, Peter D. Kelly, Felicia Chu, Hui Huang, Alicia Hill-Force, and William S. Talbot. 2000. "Zebrafish

Comparative Genomics and the Origins of Vertebrate Chromosomes." *Genome Research* 10(12):1890–1902. doi: 10.1101/gr.164800.

Punwani, Divya, Yong Zhang, Jason Yu, Morton J. Cowan, Sadhna Rana, Antonia Kwan, Aashish N. Adhikari, Carlos O. Lizama, Bryce A. Mendelsohn, Shawn P. Fahl, Ajithavalli Chellappan, Rajgopal Srinivasan, Steven E. Brenner, David L. Wiest, and Jennifer M. Puck. 2016. "Multisystem Anomalies in Severe Combined Immunodeficiency with Mutant BCL11B." *The New England Journal of Medicine* 375(22):2165–76. doi: 10.1056/NEJMoa1509164.

Qi, Haozhe, Felix Schmöhl, Xiaogang Li, Xin Qian, Christoph T. Tabler, Katrin Bennewitz, Carsten Sticht, Jakob Morgenstern, Thomas Fleming, Nadine Volk, Ingrid Hausser, Elena Heidenreich, Rüdiger Hell, Peter Paul Nawroth, and Jens Kroll. 2021. "Reduced Acrolein Detoxification in Akr1a1a Zebrafish Mutants Causes Impaired Insulin Receptor Signaling and Microvascular Alterations." *Advanced Science* 8(18):2101281. doi: 10.1002/advs.202101281.

Quintana, Francisco J., Antonio H. Iglesias, Mauricio F. Farez, Mario Caccamo, Evan J. Burns, Nasim Kassam, Mohamed Oukka, and Howard L. Weiner. 2010. "Adaptive Autoimmunity and Foxp3-Based Immunoregulation in Zebrafish." *PLOS ONE* 5(3):e9478. doi: 10.1371/journal.pone.0009478.

Raj, Bushra, Daniel E. Wagner, Aaron McKenna, Shristi Pandey, Allon M. Klein, Jay Shendure, James A. Gagnon, and Alexander F. Schier. 2018. "Simultaneous Single-Cell Profiling of Lineages and Cell Types in the Vertebrate Brain." *Nature Biotechnology* 36(5):442–50. doi: 10.1038/nbt.4103.

Ramakrishnan, Lalita. 2012. "Revisiting the Role of the Granuloma in Tuberculosis." *Nature Reviews Immunology* 12(5):352–66. doi: 10.1038/nri3211.

Rastogi, Shivangi, Sarah Ellinwood, Jacques Augenstreich, Katrin D. Mayer-Barber, and Volker Briken. 2021. "Mycobacterium Tuberculosis Inhibits the NLRP3 Inflammasome Activation via Its Phosphokinase PknF." *PLoS Pathogens* 17(7):e1009712. doi: 10.1371/journal.ppat.1009712.

Ravimohan, Shruthi, Pholo Maenetje, Sara C. Auld, Itai Ncube, Mandla Mlotshwa, William Chase, Caroline T. Tiemessen, Mboyo-Di-Tamba Vangu, Robert S. Wallis, Gavin Churchyard, Drew Weissman, Hardy Kornfeld, and Gregory P. Bisson. 2020. "A Common NLRC4 Gene Variant Associates With Inflammation and Pulmonary Function in Human Immunodeficiency Virus and Tuberculosis." *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 71(4):924–32. doi: 10.1093/cid/ciz898.

Renshaw, Stephen A., Catherine A. Loynes, Daniel M. I. Trushell, Stone Elworthy, Philip W. Ingham, and Moira K. B. Whyte. 2006. "A Transgenic Zebrafish Model of Neutrophilic Inflammation." *Blood* 108(13):3976–78. doi: 10.1182/blood-2006-05-024075.

Renshaw, Stephen A., and Nikolaus 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.

Richardson, R., D. Tracey-White, A. Webster, and M. Moosajee. 2017. "The Zebrafish Eye—a Paradigm for Investigating Human Ocular Genetics." *Eye* 31(1):68–86. doi: 10.1038/eye.2016.198.

Rissone, Alberto, and Shawn M. Burgess. 2018. "Rare Genetic Blood Disease Modeling in Zebrafish." *Frontiers in Genetics* 9:348. doi: 10.3389/fgene.2018.00348.

Robu, Mara E., Jon D. Larson, Aidas Nasevicius, Soraya Beiraghi, Charles Brenner, Steven A. Farber, and Stephen C. Ekker. 2007. "P53 Activation by Knockdown Technologies." *PLoS Genetics* 3(5):e78. doi: 10.1371/journal.pgen.0030078.

van Rooijen, Ellen, Emile E. Voest, Ive Logister, Jeroen Korving, Thorsten Schwerte, Stefan Schulte-Merker, Rachel H. Giles, and Fredericus J. van Eeden. 2009. "Zebrafish Mutants in the von Hippel-Lindau Tumor Suppressor Display a Hypoxic Response and Recapitulate Key Aspects of Chuvash Polycythemia." *Blood* 113(25):6449–60. doi: 10.1182/blood-2008-07-167890.

Rosowski, Emily E., Benjamin P. Knox, Linda S. Archambault, Anna Huttenlocher, Nancy P. Keller, Robert T. Wheeler, and J. Muse Davis. 2018. "The Zebrafish as a Model Host for Invasive Fungal Infections." *Journal of Fungi* 4(4):136. doi: 10.3390/jof4040136.

Rossi, Andrea, Zacharias Kontarakis, Claudia Gerri, Hendrik Nolte, Soraya Hölper, Marcus Krüger, and Didier Y. R. Stainier. 2015. "Genetic Compensation Induced by Deleterious Mutations but Not Gene Knockdowns." *Nature* 524(7564):230–33. doi: 10.1038/nature14580.

Rougeot, Julien, Vincenzo Torraca, Ania Zakrzewska, Zakia Kanwal, Hans J. Jansen, Frida Sommer, Herman P. Spaink, and Annemarie H. Meijer. 2019. "RNAseq Profiling of Leukocyte Populations in Zebrafish Larvae Reveals a Cxcl11 Chemokine Gene as a Marker of Macrophage Polarization During Mycobacterial Infection." *Frontiers in Immunology* 10:832. doi: 10.3389/fimmu.2019.00832.

Rounioja, Samuli, Anni Saralahti, Lilli Rantala, Mataleena Parikka, Birgitta Henriques-Normark, Olli Silvennoinen, and Mika Rämet. 2012. "Defense of Zebrafish Embryos against Streptococcus Pneumoniae Infection Is Dependent on the Phagocytic Activity of Leukocytes." *Developmental and Comparative Immunology* 36(2):342–48. doi: 10.1016/j.dci.2011.05.008.

Saiga, Hiroyuki, Shoko Kitada, Yosuke Shimada, Naganori Kamiyama, Megumi Okuyama, Masahiko Makino, Masahiro Yamamoto, and Kiyoshi Takeda. 2012. "Critical Role of AIM2 in Mycobacterium Tuberculosis Infection." *International Immunology* 24(10):637–44. doi: 10.1093/intimm/dxs062.

Saiga, Hiroyuki, Natalie Nieuwenhuizen, Martin Gengenbacher, Anne-Britta Koehler, Stefanie Schuerer, Pedro Moura-Alves, Ina Wagner, Hans-Joachim Mollenkopf, Anca Dorhoi, and Stefan H. E. Kaufmann. 2015. "The Recombinant BCG ΔureC::Hly Vaccine Targets the AIM2 Inflammasome to Induce Autophagy and Inflammation." *The Journal of Infectious Diseases* 211(11):1831–41. doi: 10.1093/infdis/jiu675.

Saralahti, Anni K., Meri I. E. Uusi-Mäkelä, Mirja T. Niskanen, and Mika Rämet. 2020. "Integrating Fish Models in Tuberculosis Vaccine Development." *Disease Models & Mechanisms* 13(8):dmm045716. doi: 10.1242/dmm.045716.

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.

Segovia, Mercedes, Sofia Russo, Mathias Jeldres, Yamil D. Mahmoud, Valentina Perez, Maite Duhalde, Pierre Charnet, Matthieu Rousset, Sabina Victoria, Florencia Veigas, Cédric Louvet, Bernard Vanhove, R. Andrés Floto, Ignacio Anegon, Maria Cristina Cuturi, M. Romina Girotti, Gabriel A. Rabinovich, and Marcelo Hill. 2019. "Targeting TMEM176B Enhances Antitumor Immunity and Augments the Efficacy of Immune Checkpoint Blockers by Unleashing Inflammasome Activation." *Cancer Cell* 35(5):767-781.e6. doi: 10.1016/j.ccell.2019.04.003.

Shah, Swati, Amro Bohsali, Sarah E. Ahlbrand, Lalitha Srinivasan, Vijay A. K. Rathinam, Stefanie N. Vogel, Katherine A. Fitzgerald, Fayyaz S. Sutterwala, and Volker Briken. 2013. "Cutting Edge: Mycobacterium Tuberculosis but Not Nonvirulent Mycobacteria Inhibits IFN-Beta and AIM2 Inflammasome-Dependent IL-1 Beta Production via Its ESX-1 Secretion System." *Journal of Immunology* 191(7):3514–18. doi: 10.4049/jimmunol.1301331.

Shao, Chang, Wenjie Lu, Ye Du, Wenchao Yan, Qiuyu Bao, Yang Tian, Guangji Wang, Hui Ye, and Haiping Hao. 2020. "Cytosolic ME1 Integrated with Mitochondrial IDH2 Supports Tumor Growth and Metastasis." *Redox Biology* 36:101685. doi: 10.1016/j.redox.2020.101685.

Shao, Tong, Lv-Yun Zhu, Li Nie, Wei Shi, Wei-Ren Dong, Li-Xin Xiang, and Jian-Zhong Shao. 2015. "Characterization of Surface Phenotypic Molecules of Teleost Dendritic Cells." *Developmental & Comparative Immunology* 49(1):38–43. doi: 10.1016/j.dci.2014.11.010.

Shastri, Madhur D., Shakti Dhar Shukla, Wai Chin Chong, Kamal Dua, Gregory M. Peterson, Rahul P. Patel, Philip M. Hansbro, Rajaraman Eri, and Ronan F. O'Toole. 2018. "Role of Oxidative Stress in the Pathology and Management of Human Tuberculosis." *Oxidative Medicine and Cellular Longevity* 2018:e7695364. doi: 10.1155/2018/7695364.

Shin, Sunny, and Igor E. Brodsky. 2015. "The Inflammasome: Learning from Bacterial Evasion Strategies." *Seminars in Immunology* 27(2):102–10. doi: 10.1016/j.smim.2015.03.006.

Sievers, Fabian, Andreas Wilm, David Dineen, Toby J. Gibson, Kevin Karplus, Weizhong Li, Rodrigo Lopez, Hamish McWilliam, Michael Remmert, Johannes Söding, Julie D. Thompson, and Desmond G. Higgins. 2011. "Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega." *Molecular Systems Biology* 7:539. doi: 10.1038/msb.2011.75.

Singh, Ritambhara, Cem Kuscu, Aaron Quinlan, Yanjun Qi, and Mazhar Adli. 2015. "Cas9-Chromatin Binding Information Enables More Accurate CRISPR off-Target Prediction." *Nucleic Acids Research* 43(18):e118. doi: 10.1093/nar/gkv575.

Stainier, Didier Y. R., Zacharias Kontarakis, and Andrea Rossi. 2015. "Making Sense of Anti-Sense Data." *Developmental Cell* 32(1):7–8. doi: 10.1016/j.devcel.2014.12.012.

Stainier, Didier Y. R., Erez Raz, Nathan D. Lawson, Stephen C. Ekker, Rebecca D. Burdine, Judith S. Eisen, Philip W. Ingham, Stefan Schulte-Merker, Deborah Yelon, Brant M. Weinstein, Mary C. Mullins, Stephen W. Wilson, Lalita Ramakrishnan, Sharon L. Amacher, Stephan C. F. Neuhauss, Anming Meng, Naoki Mochizuki, Pertti Panula, and Cecilia B. Moens. 2017. "Guidelines for Morpholino Use in Zebrafish." *PLOS Genetics* 13(10):e1007000. doi: 10.1371/journal.pgen.1007000.

Stenger, Steffen, Jean-Pierre Rosat, Barry R. Bloom, Alan M. Krensky, and Robert L. Modlin. 1999. "Granulysin: A Lethal Weapon of Cytolytic T Cells." *Immunology Today* 20(9):390–94. doi: 10.1016/S0167-5699(99)01449-8.

Strähle, Uwe, Stefan Scholz, Robert Geisler, Petra Greiner, Henner Hollert, Sepand Rastegar, Axel Schumacher, Ingrid Selderslaghs, Carsten Weiss, Hilda Witters, and Thomas Braunbeck. 2012. "Zebrafish Embryos as an Alternative to Animal Experiments—A Commentary on the Definition of the Onset of Protected Life Stages in Animal Welfare Regulations." *Reproductive Toxicology* 33(2):128–32. doi: 10.1016/j.reprotox.2011.06.121.

Streisinger, G., C. Walker, N. Dower, D. Knauber, and F. Singer. 1981. "Production of Clones of Homozygous Diploid Zebra Fish (Brachydanio Rerio)." *Nature* 291(5813):293–96. doi: 10.1038/291293a0.

Subbarao, Sathyavani, Julia Sanchez-Garrido, Nitya Krishnan, Avinash R. Shenoy, and Brian D. Robertson. 2020. "Genetic and Pharmacological Inhibition of Inflammasomes Reduces the Survival of Mycobacterium Tuberculosis Strains in Macrophages." *Scientific Reports* 10(1):3709. doi: 10.1038/s41598-020-60560-y.

Subbian, Selvakumar, Paul O'Brien, Nicole L. Kushner, Guibin Yang, Liana Tsenova, Blas Peixoto, Nirmalya Bandyopadhyay, Joel S. Bader, Petros C. Karakousis, Dorothy Fallows, and Gilla Kaplan. 2013. "Molecular Immunologic Correlates of Spontaneous Latency in a Rabbit Model of Pulmonary Tuberculosis." *Cell Communication and Signaling* 11(1):16. doi: 10.1186/1478-811X-11-16.

Summerton, J., and D. Weller. 1997. "Morpholino Antisense Oligomers: Design, Preparation, and Properties." *Antisense & Nucleic Acid Drug Development* 7(3):187–95. doi: 10.1089/oli.1.1997.7.187.

Sun, Chen-Chen, Zuo-Qiong Zhou, Zhang-Lin Chen, Run-Kang Zhu, Dong Yang, Xi-Yang Peng, Lan Zheng, and Chang-Fa Tang. 2021. "Identification of Potentially Related Genes and Mechanisms Involved in Skeletal Muscle Atrophy Induced by Excessive Exercise in Zebrafish." *Biology* 10(8):761. doi: 10.3390/biology10080761.

Suurväli, Jaanus, Andrew R. Whiteley, Yichen Zheng, Karim Gharbi, Maria Leptin, and Thomas Wiehe. 2020. "The Laboratory Domestication of Zebrafish: From Diverse Populations to Inbred Substrains." *Molecular Biology and Evolution* 37(4):1056–69. doi: 10.1093/molbev/msz289.

Swaim, Laura E., Lynn E. Connolly, Hannah E. Volkman, Olivier Humbert, Donald E. Born, and Lalita Ramakrishnan. 2007. "Mycobacterium Marinum Infection of Adult Zebrafish Causes Caseating Granulomatous Tuberculosis and Is Moderated by Adaptive Immunity (Vol 74, Pg 6108, 2006)." *Infection and Immunity* 75(3):1540. doi: 10.1128/IAI.00069-07.

Swanhart, Lisa M., Chiara Cianciolo Cosentino, Cuong Q. Diep, Alan J. Davidson, Mark de Caestecker, and Neil A. Hukriede. 2011. "Zebrafish Kidney Development: Basic Science to Translational Research." *Birth Defects Research. Part C, Embryo Today : Reviews* 93(2):141–56. doi: 10.1002/bdrc.20209.

Takaki, Kevin, Christine L. Cosma, Mark A. Troll, and Lalita Ramakrishnan. 2012. "An In Vivo Platform for Rapid High-Throughput Antitubercular Drug Discovery." *Cell Reports* 2(1):175–84. doi: 10.1016/j.celrep.2012.06.008.

Tallafuss, Alexandra, Dan Gibson, Paul Morcos, Yongfu Li, Steve Seredick, Judith Eisen, and Philip Washbourne. 2012. "Turning Gene Function ON and OFF Using Sense and Antisense Photo-Morpholinos in Zebrafish." *Development (Cambridge, England)* 139(9):1691–99. doi: 10.1242/dev.072702.

Tambalo, Monica, Richard Mitter, and David G. Wilkinson. 2020. "A Single Cell Transcriptome Atlas of the Developing Zebrafish Hindbrain." *Development* 147(6). doi: 10.1242/dev.184143.

Tang, Fuchou, Catalin Barbacioru, Yangzhou Wang, Ellen Nordman, Clarence Lee, Nanlan Xu, Xiaohui Wang, John Bodeau, Brian B. Tuch, Asim Siddiqui, Kaiqin Lao, and M. Azim Surani. 2009. "MRNA-Seq Whole-Transcriptome Analysis of a Single Cell." *Nature Methods* 6(5):377–82. doi: 10.1038/nmeth.1315.

Tang, Qin, Sowmya Iyer, Riadh Lobbardi, John C. Moore, Huidong Chen, Caleb Lareau, Christine Hebert, McKenzie L. Shaw, Cyril Neftel, Mario L. Suva, Craig J. Ceol, Andre Bernards, Martin Aryee, Luca Pinello, Iain A. Drummond, and David M. Langenau. 2017. "Dissecting Hematopoietic and Renal Cell Heterogeneity in Adult Zebrafish at Single-Cell Resolution Using RNA Sequencing." *The Journal of Experimental Medicine* 214(10):2875–87. doi: 10.1084/jem.20170976.

Tang, Rongying, Andrew Dodd, Daniel Lai, Warren C. McNabb, and Donald R. Love. 2007. "Validation of Zebrafish (Danio Rerio) Reference Genes for Quantitative Real-Time RT-PCR Normalization." *Acta Biochimica Et Biophysica Sinica* 39(5):384–90.

Taxman, Debra J., Elizabeth A. Holley-Guthrie, Max Tze-Han Huang, Chris B. Moore, Daniel T. Bergstralh, Irving C. Allen, Yu Lei, Denis Gris, and Jenny Pan-Yun Ting. 2011. "The NLR Adaptor ASC/PYCARD Regulates DUSP10, Mitogen-Activated Protein Kinase (MAPK), and Chemokine Induction Independent of the Inflammasome." *The Journal of Biological Chemistry* 286(22):19605–16. doi: 10.1074/jbc.M111.221077.

Taylor, Alison M., David M. Raiser, Jessica M. Humphries, Benjamin L. Ebert, and Leonard I. Zon. 2012. "Calmodulin Inhibition Rescues the Effects of Ribosomal Protein Deficiency by Modulating P53 Activity in Models of Diamond Blackfan Anemia." *Blood* 120(21):512–512. doi: 10.1182/blood.V120.21.512.512.

Tobin, David M., Robin C. May, and Robert T. Wheeler. 2012. "Zebrafish: A See-Through Host and a Fluorescent Toolbox to Probe Host–Pathogen Interaction." *PLOS Pathogens* 8(1):e1002349. doi: 10.1371/journal.ppat.1002349.

Tobin, David M., and Lalita Ramakrishnan. 2008. "Comparative Pathogenesis of Mycobacterium Marinum and Mycobacterium Tuberculosis." *Cellular Microbiology* 10(5):1027–39. doi: 10.1111/j.1462-5822.2008.01133.x.

Tobin, David M., Jay C. Vary, John P. Ray, Gregory S. Walsh, Sarah J. Dunstan, Nguyen D. Bang, Deanna A. Hagge, Saraswoti Khadge, Mary-Claire King, Thomas R. Hawn, Cecilia B. Moens, and Lalita Ramakrishnan. 2010. "The Lta4h Locus Modulates Susceptibility to Mycobacterial Infection in Zebrafish and Humans." *Cell* 140(5):717–30. doi: 10.1016/j.cell.2010.02.013.

Traver, David, Barry H. Paw, Kenneth D. Poss, W. Todd Penberthy, Shuo Lin, and Leonard I. Zon. 2003. "Transplantation and in Vivo Imaging of Multilineage Engraftment in Zebrafish Bloodless Mutants." *Nature Immunology* 4(12):1238–46. doi: 10.1038/ni1007.

Tsompana, Maria, and Michael J. Buck. 2014. "Chromatin Accessibility: A Window into the Genome." *Epigenetics & Chromatin* 7(1):33. doi: 10.1186/1756-8935-7-33.

Tsuji, Shoutaro, Makiko Yamashita, Donald R. Hoffman, Akihito Nishiyama, Tsutomu Shinohara, Takashi Ohtsu, and Yoshimi Shibata. 2009. "Capture of Heat-Killed Mycobacterium Bovis Bacillus Calmette-Guerin by Intelectin-1 Deposited on Cell Surfaces." *Glycobiology* 19(5):518–26. doi: 10.1093/glycob/cwp013.

Tyrkalska, Sylwia D., Ana B. Pérez-Oliva, Lola Rodríguez-Ruiz, Francisco J. Martínez-Morcillo, Francisca Alcaraz-Pérez, Francisco J. Martínez-Navarro, Christophe Lachaud, Nouraiz Ahmed, Timm Schroeder, Irene Pardo-Sánchez, Sergio Candel, Azucena López-Muñoz, Avik Choudhuri, Marlies P. Rossmann, Leonard I. Zon, María L. Cayuela, Diana García-Moreno, and Victoriano Mulero. 2019. "Inflammasome Regulates Hematopoiesis through Cleavage of the Master Erythroid Transcription Factor GATA1." *Immunity* 51(1):50-63.e5. doi: 10.1016/j.immuni.2019.05.005.

Urnov, Fyodor D., Edward J. Rebar, Michael C. Holmes, H. Steve Zhang, and Philip D. Gregory. 2010. "Genome Editing with Engineered Zinc Finger Nucleases." *Nature Reviews Genetics* 11(9):636–46. doi: 10.1038/nrg2842.

Uusimaa, Johanna, Riitta Kaarteenaho, Teija Paakkola, Hannu Tuominen, Minna K. Karjalainen, Javad Nadaf, Teppo Varilo, Meri Uusi-Mäkela, Maria Suo-Palosaari, Ilkka Pietila, Anniina E. Hiltunen, Lloyd Ruddock, Heli Alanen, Ekaterina Biterova, Ilkka Miinalainen, Annamari Salminen, Raija Soininen, Aki Manninen, Raija Sormunen, Mika Kaakinen, Reetta Vuolteenaho, Riitta Herva, Paivi Vieira, Teija Dunder, Hannaleena Kokkonen, Jukka S. Moilanen, Heikki Rantala, Lawrence M. Nogee, Jacek Majewski, Mika Rämet, Mikko Hallman, and Reetta Hinttala. 2018. "NHLRC2 Variants Identified in Patients with Fibrosis, Neurodegeneration, and Cerebral Angiomatosis (FINCA): Characterisation of a Novel Cerebropulmonary Disease." *Acta Neuropathologica* 135(5):727–42. doi: 10.1007/s00401-018-1817-z.

Varela, Mónica, Antonio Figueras, and Beatriz Novoa. 2017. "Modelling Viral Infections Using Zebrafish: Innate Immune Response and Antiviral Research." *Antiviral Research* 139:59–68. doi: 10.1016/j.antiviral.2016.12.013.

Varga, Máté, Dorottya Ralbovszki, Eszter Balogh, Renáta Hamar, Magdolna Keszthelyi, and Kálmán Tory. 2018. "Zebrafish Models of Rare Hereditary Pediatric Diseases." *Diseases* 6(2):43. doi: 10.3390/diseases6020043.

Vatine, Gad, Daniela Vallone, Yoav Gothilf, and Nicholas S. Foulkes. 2011. "It's Time to Swim! Zebrafish and the Circadian Clock." *FEBS Letters* 585(10):1485–94. doi: 10.1016/j.febslet.2011.04.007.

Verkuijl, Sebald AN, and Marianne G. Rots. 2019. "The Influence of Eukaryotic Chromatin State on CRISPR–Cas9 Editing Efficiencies." *Current Opinion in Biotechnology* 55:68–73. doi: 10.1016/j.copbio.2018.07.005.

Voehringer, David, Sarah A. Stanley, Jeffery S. Cox, Gladys C. Completo, Todd L. Lowary, and Richard M. Locksley. 2007. "Nippostrongylus Brasiliensis: Identification of Intelectin-1 and -2 as Stat6-Dependent Genes Expressed in Lung and Intestine during Infection." *Experimental Parasitology* 116(4):458–66. doi: 10.1016/j.exppara.2007.02.015.

Volkman, Hannah E., Tamara C. Pozos, John Zheng, J. Muse Davis, John F. Rawls, and Lalita Ramakrishnan. 2010. "Tuberculous Granuloma Induction via Interaction of a Bacterial Secreted Protein with Host Epithelium." *Science (New York, N.Y.)* 327(5964):466–69. doi: 10.1126/science.1179663.

Wagner, Daniel E., Caleb Weinreb, Zach M. Collins, James A. Briggs, Sean G. Megason, and Allon M. Klein. 2018. "Single-Cell Mapping of Gene Expression Landscapes and Lineage in the Zebrafish Embryo." *Science* 360(6392):981–87. doi: 10.1126/science.aar4362.

Wallace, Kenneth N., Shafinaz Akhter, Erin M. Smith, Kristin Lorent, and Michael Pack. 2005. "Intestinal Growth and Differentiation in Zebrafish." *Mechanisms of Development* 122(2):157–73. doi: 10.1016/j.mod.2004.10.009.

Walton, Eric M., Mark R. Cronan, Rebecca W. Beerman, and David M. Tobin. 2015. "The Macrophage-Specific Promoter Mfap4 Allows Live, Long-Term Analysis of Macrophage Behavior during Mycobacterial Infection in Zebrafish." *PLOS ONE* 10(10):e0138949. doi: 10.1371/journal.pone.0138949.

Walton, Eric M., Mark R. Cronan, C. J. Cambier, Andrea Rossi, Michele Marass, Matthew D. Foglia, W. Jared Brewer, Kenneth D. Poss, Didier Y. R. Stainier, Carolyn R. Bertozzi, and David M. Tobin. 2018. "Cyclopropane Modification of Trehalose

Dimycolate Drives Granuloma Angiogenesis and Mycobacterial Growth through Vegf Signaling." *Cell Host & Microbe* 24(4):514-525.e6. doi: 10.1016/j.chom.2018.09.004.

Wan, Feng, Chong-bin Hu, Jun-xia Ma, Ke Gao, Li-xin Xiang, and Jian-zhong Shao. 2017. "Characterization of Γδ T Cells from Zebrafish Provides Insights into Their Important Role in Adaptive Humoral Immunity." *Frontiers in Immunology* 7:675. doi: 10.3389/fimmu.2016.00675.

Wang, Zhengyuan, Jianguo Du, Siew Hong Lam, Sinnakarupan Mathavan, Paul Matsudaira, and Zhiyuan Gong. 2010. "Morphological and Molecular Evidence for Functional Organization along the Rostrocaudal Axis of the Adult Zebrafish Intestine." *BMC Genomics* 11(1):392. doi: 10.1186/1471-2164-11-392.

Wangler, Michael F., Shinya Yamamoto, Hsiao-Tuan Chao, Jennifer E. Posey, Monte Westerfield, John Postlethwait, Members of the Undiagnosed Diseases Network (UDN), Philip Hieter, Kym M. Boycott, Philippe M. Campeau, and Hugo J. Bellen. 2017. "Model Organisms Facilitate Rare Disease Diagnosis and Therapeutic Research." *Genetics* 207(1):9–27. doi: 10.1534/genetics.117.203067.

Wassermann, Ruth, Muhammet F. Gulen, Claudia Sala, Sonia Garcia Perin, Ye Lou, Jan Rybniker, Jonathan L. Schmid-Burgk, Tobias Schmidt, Veit Hornung, Stewart T. Cole, and Andrea Ablasser. 2015. "Mycobacterium Tuberculosis Differentially Activates CGAS- and Inflammasome-Dependent Intracellular Immune Responses through ESX-1." *Cell Host & Microbe* 17(6):799–810. doi: 10.1016/j.chom.2015.05.003.

Watamoto, Kouichi, Masayuki Towatari, Yukiyasu Ozawa, Yasuhiko Miyata, Mitsunori Okamoto, Akihiro Abe, Tomoki Naoe, and Hidehiko Saito. 2003. "Altered Interaction of HDAC5 with GATA-1 during MEL Cell Differentiation." *Oncogene* 22(57):9176–84. doi: 10.1038/sj.onc.1206902.

Wei, Shi, Miaomiao Dai, Zhaoting Liu, Yuanqing Ma, Hanqiao Shang, Yu Cao, and Qiang Wang. 2017. "The Guanine Nucleotide Exchange Factor Net1 Facilitates the Specification of Dorsal Cell Fates in Zebrafish Embryos by Promoting Maternal β-Catenin Activation." *Cell Research* 27(2):202–25. doi: 10.1038/cr.2016.141.

Wesener, Darryl A., Kittikhun Wangkanont, Ryan McBride, Xuezheng Song, Matthew B. Kraft, Heather L. Hodges, Lucas C. Zarling, Rebecca A. Splain, David F. Smith, Richard D. Cummings, James C. Paulson, Katrina T. Forest, and Laura L. Kiessling. 2015. "Recognition of Microbial Glycans by Human Intelectin." *Nature Structural & Molecular Biology* 22(8):603–10. doi: 10.1038/nsmb.3053.

Westerfield, Monte. 2000. *The Zebrafish Book, 4th Edition; A Guide For The Laboratory Use Of Zebrafish (Danio Rerio)*. 4th ed. Eugene: University of Oregon Press.

White, Richard Mark, Anna Sessa, Christopher Burke, Teresa Bowman, Jocelyn LeBlanc, Craig Ceol, Caitlin Bourque, Michael Dovey, Wolfram Goessling, Caroline Erter Burns, and Leonard I. Zon. 2008. "Transparent Adult Zebrafish as a Tool for In Vivo Transplantation Analysis." *Cell Stem Cell* 2(2):183–89. doi: 10.1016/j.stem.2007.11.002.

Whitfield, Tanya T. 2002. "Zebrafish as a Model for Hearing and Deafness." *Journal of Neurobiology* 53(2):157–71. doi: 10.1002/neu.10123.

Wilkins, Benjamin J., and Michael Pack. 2013. "Zebrafish Models of Human Liver Development and Disease." *Comprehensive Physiology* 3(3):1213–30. doi: 10.1002/cphy.c120021.

Willett, C. E., J. J. Cherry, and L. A. Steiner. 1997. "Characterization and Expression of the Recombination Activating Genes (Rag1 and Rag2) of Zebrafish." *Immunogenetics* 45(6):394–404. doi: 10.1007/s002510050221.

Willett, C. E., H. Kawasaki, C. T. Amemiya, S. Lin, and L. A. Steiner. 2001. "Ikaros Expression as a Marker for Lymphoid Progenitors during Zebrafish Development." *Developmental Dynamics: An Official Publication of the American Association of Anatomists* 222(4):694–98. doi: 10.1002/dvdy.1223.

World Health Organization. 2020. *Global Tuberculosis Report 2020*. Geneva.

Wu, Xuebing, David A. Scott, Andrea J. Kriz, Anthony C. Chiu, Patrick D. Hsu, Daniel B. Dadon, Albert W. Cheng, Alexandro E. Trevino, Silvana Konermann, Sidi Chen, Rudolf Jaenisch, Feng Zhang, and Phillip A. Sharp. 2014. "Genome-Wide Binding of the CRISPR Endonuclease Cas9 in Mammalian Cells." *Nature Biotechnology* 32(7):670–76. doi: 10.1038/nbt.2889.

Wu, Yvonne W., Joseph Sullivan, Sharon S. McDaniel, Miriam H. Meisler, Eileen M. Walsh, Sherian Xu Li, and Michael W. Kuzniewicz. 2015. "Incidence of Dravet Syndrome in a US Population." *Pediatrics* 136(5):e1310-1315. doi: 10.1542/peds.2015-1807.

Xu, Xiaojun, Dongsheng Duan, and Shi-Jie Chen. 2017. "CRISPR-Cas9 Cleavage Efficiency Correlates Strongly with Target-SgRNA Folding Stability: From Physical Mechanism to off-Target Assessment." *Scientific Reports* 7(1):143. doi: 10.1038/s41598-017-00180-1.

Xue, Yuanyuan, Denghui Liu, Guizhong Cui, Yanyan Ding, Daosheng Ai, Suwei Gao, Yifan Zhang, Shengbao Suo, Xiaohan Wang, Peng Lv, Chunyu Zhou, Yizhou Li, Xingwei Chen, Guangdun Peng, Naihe Jing, Jing-Dong J. Han, and Feng Liu. 2019. "A 3D Atlas of Hematopoietic Stem and Progenitor Cell Expansion by Multi-Dimensional RNA-Seq Analysis." *Cell Reports* 27(5):1567-1578.e5. doi: 10.1016/j.celrep.2019.04.030.

Yan, Jie, Lei Chen, Zhuang Liu, Yonglin Chen, Ying Sun, Jia Han, and Lijun Feng. 2018. "The D5 Region of the Intelectin Domain Is a New Type of Carbohydrate Recognition Domain in the Intelectin Gene Family." *Developmental and Comparative Immunology* 85:150–60. doi: 10.1016/j.dci.2018.02.021.

Yan, Shanshan, Hongbo Shen, Qiaoshi Lian, Wenlong Jin, Ronghua Zhang, Xuan Lin, Wangpeng Gu, Xiaoyu Sun, Guangxun Meng, Zhigang Tian, Zheng W. Chen, and Bing Sun. 2018. "Deficiency of the AIM2-ASC Signal Uncovers the STING-Driven Overreactive Response of Type I IFN and Reciprocal Depression of Protective IFN-γ Immunity in Mycobacterial Infection." *Journal of Immunology (Baltimore, Md.: 1950)* 200(3):1016–26. doi: 10.4049/jimmunol.1701177.

Yang, Hongbo, Yu Luan, Tingting Liu, Hyung Joo Lee, Li Fang, Yanli Wang, Xiaotao Wang, Bo Zhang, Qiushi Jin, Khai Chung Ang, Xiaoyun Xing, Juan Wang, Jie Xu, Fan Song, Iyyanki Sriranga, Chachrit Khunsriraksakul, Tarik Salameh, Daofeng Li, Mayank N. K. Choudhary, Jacek Topczewski, Kai Wang, Glenn S. Gerhard, Ross C. Hardison, Ting Wang, Keith C. Cheng, and Feng Yue. 2020. "A Map of Cis-Regulatory Elements and 3D Genome Structures in Zebrafish." *Nature* 588(7837):337–43. doi: 10.1038/s41586-020-2962-9.

Yang, Lili, Lingzhen Bu, Weiwei Sun, Lili Hu, and Shicui Zhang. 2014. "Functional Characterization of Mannose-Binding Lectin in Zebrafish: Implication for a Lectin-Dependent Complement System in Early Embryos." *Developmental & Comparative Immunology* 46(2):314–22. doi: 10.1016/j.dci.2014.05.003.

Yarrington, Robert M., Surbhi Verma, Shaina Schwartz, Jonathan K. Trautman, and Dana Carroll. 2018. "Nucleosomes Inhibit Target Cleavage by CRISPR-Cas9 in Vivo." *Proceedings of the National Academy of Sciences* 115(38):9351–58. doi: 10.1073/pnas.1810062115.

Yeung, Amy T. Y., Yoon Ha Choi, Amy H. Y. Lee, Christine Hale, Hannes Ponstingl, Derek Pickard, David Goulding, Mark Thomas, Erin Gill, Jong Kyoung Kim, Allan Bradley, Robert E. W. Hancock, and Gordon Dougan. 2019. "A Genome-Wide Knockout Screen in Human Macrophages Identified Host Factors Modulating Salmonella Infection." *MBio* 10(5):e02169-19. doi: 10.1128/mBio.02169-19.

Yoon, Sohye, Suman Mitra, Cathy Wyse, Ayham Alnabulsi, Jun Zou, Eveline M. Weerdenburg, Astrid M. van der Sar, Difei Wang, Christopher J. Secombes, and Steve Bird. 2015. "First Demonstration of Antigen Induced Cytokine Expression by CD4-1+ Lymphocytes in a Poikilotherm: Studies in Zebrafish (Danio Rerio)." *PloS One* 10(6):e0126378. doi: 10.1371/journal.pone.0126378.

Yoshida, Nagisa, Eva-Maria Frickel, and Serge Mostowy. 2017. "Macrophage–Microbe Interactions: Lessons from the Zebrafish Model." *Frontiers in Immunology* 8:1703. doi: 10.3389/fimmu.2017.01703.

Zhang, Shicui, and Pengfei Cui. 2014. "Complement System in Zebrafish." *Developmental and Comparative Immunology* 46(1):3–10. doi: 10.1016/j.dci.2014.01.010.

Zhang, Ying, Yuanyuan Xue, Chunwei Cao, Jiaojiao Huang, Qianlong Hong, Tang Hai, Qitao Jia, Xianlong Wang, Guosong Qin, Jing Yao, Xiao Wang, Qiantao Zheng, Rui Zhang, Yongshun Li, Ailing Luo, Nan Zhang, Guizhi Shi, Yanfang Wang, Hao Ying, Zhonghua Liu, Hongmei Wang, Anming Meng, Qi Zhou, Hong Wei, Feng Liu, and Jianguo Zhao. 2017. "Thyroid Hormone Regulates Hematopoiesis via the TR-KLF9 Axis." *Blood* 130(20):2161–70. doi: 10.1182/blood-2017-05-783043.

Zhu, Jiang, Yi Zhang, Gerard J. Joe, Richard Pompetti, and Stephen G. Emerson. 2005. "NF-Ya Activates Multiple Hematopoietic Stem Cell (HSC) Regulatory Genes and Promotes HSC Self-Renewal." *Proceedings of the National Academy of Sciences of the United States of America* 102(33):11728–33. doi: 10.1073/pnas.0503405102.

Zhu, Peipei, Zhipeng Ma, Liwei Guo, Wei Zhang, Qinghe Zhang, Ting Zhao, Kunpeng Jiang, Jinrong Peng, and Jun Chen. 2017. "Short Body Length Phenotype Is Compensated by the Upregulation of Nidogen Family Members in a Deleterious Nid1a Mutation of Zebrafish." *Journal of Genetics and Genomics = Yi Chuan Xue Bao* 44(11):553–56. doi: 10.1016/j.jgg.2017.09.011.

Zuccolo, Jonathan, Jeremy Bau, Sarah J. Childs, Greg G. Goss, Christoph W. Sensen, and Julie P. Deans. 2010. "Phylogenetic Analysis of the MS4A and TMEM176 Gene Families." *PLOS ONE* 5(2):e9369. doi: 10.1371/journal.pone.0009369.

Zucker, Shoshanna N., Emily E. Fink, Archis Bagati, Sudha Mannava, Anna Bianchi-Smiraglia, Paul N. Bogner, Joseph A. Wawrzyniak, Colleen Foley, Katerina I. Leonova, Melissa J. Grimm, Kalyana Moparthy, Yurij Ionov, Jianmin Wang, Song Liu, Sandra Sexton, Eugene S. Kandel, Andrei V. Bakin, Yuesheng Zhang, Naftali Kaminski, Brahm H. Segal, and Mikhail A. Nikiforov. 2014. "Nrf2 Amplifies Oxidative Stress via Induction of Klf9." *Molecular Cell* 53(6):916–28. doi: 10.1016/j.molcel.2014.01.033.

PUBLICATIONS

PUBLICATION I

Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (*Danio rerio*)

Uusi-Mäkelä MIE, Barker HR, Bäuerlein CA, Häkkinen T, Nykter M, Rämet M

PLoS One. 2018 Apr 23;13(4):e0196238.
doi: 10.1371/journal.pone.0196238.

Publication reprinted with the permission of the copyright holders. This publication is licensed under a Creative Commons Attribution 4.0 International licence (http://creativecommons.org/licenses/by/4.0/).

PLOS ONE

RESEARCH ARTICLE

Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (*Danio rerio*)

Meri I. E. Uusi-Mäkelä[1]*, Harlan R. Barker[2], Carina A. Bäuerlein[1], Tomi Häkkinen[3], Matti Nykter[3], Mika Rämet[1,4,5]*

1 Laboratory of Experimental Immunology, BioMediTech Institute and Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland, **2** Laboratory of Anatomy, Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland, **3** Laboratory of Computational Biology, BioMediTech Institute and Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland, **4** PEDEGO Research Unit, Medical Research Center Oulu, and Department of Children and Adolescents, Oulu University Hospital, Oulu, Finland, **5** Department of Pediatrics, Tampere University Hospital, Tampere, Finland

* mika.ramet@uta.fi (MR); meri.uusi-makela@staff.uta.fi (MIEUM)

Check for updates

OPEN ACCESS

Citation: Uusi-Mäkelä MIE, Barker HR, Bäuerlein CA, Häkkinen T, Nykter M, Rämet M (2018) Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (*Danio rerio*). PLoS ONE 13(4): e0196238. https://doi.org/10.1371/journal.pone.0196238

Editor: Bruce B. Riley, Texas A&M University, UNITED STATES

Received: February 2, 2018

Accepted: March 9, 2018

Published: April 23, 2018

Copyright: © 2018 Uusi-Mäkelä et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant experimental data are within the paper and its Supporting Information files. All files are available from the CRISPRz and the ArrayExpress databases (accession codes E-GEOD-45706; E-GEOD-4863; E-GEOD-52110; E-GEOD-74231).

Funding: This study was supported with the following grants: Tampere Tuberculosis Foundation (http://www.tuberkuloosisaatio.fi/) (MU;MR); Finnish Concordia Fund (http://www.konkordia-liitto.com/) (MU); Sigrid Juselius Foundation

Abstract

CRISPR-Cas9 technology is routinely applied for targeted mutagenesis in model organisms and cell lines. Recent studies indicate that the prokaryotic CRISPR-Cas9 system is affected by eukaryotic chromatin structures. Here, we show that the likelihood of successful mutagenesis correlates with transcript levels during early development in zebrafish (*Danio rerio*) embryos. In an experimental setting, we found that guide RNAs differ in their onset of mutagenesis activity *in vivo*. Furthermore, some guide RNAs with high *in vitro* activity possessed poor mutagenesis activity *in vivo*, suggesting the presence of factors that limit the mutagenesis *in vivo*. Using open access datasets generated from early developmental stages of the zebrafish, and guide RNAs selected from the CRISPRz database, we provide further evidence for an association between gene expression during early development and the success of CRISPR-Cas9 mutagenesis in zebrafish embryos. In order to further inspect the effect of chromatin on CRISPR-Cas9 mutagenesis, we analysed the relationship of selected chromatin features on CRISPR-Cas9 mutagenesis efficiency using publicly available data from zebrafish embryos. We found a correlation between chromatin openness and the efficiency of CRISPR-Cas9 mutagenesis. These results indicate that CRISPR-Cas9 mutagenesis is influenced by chromatin accessibility in zebrafish embryos.

Introduction

Since its discovery in *Streptococcus pyogenes*, the CRISPR-Cas9 (Clustered regularly interspaced short palindromic repeats–CRISPR associated 9) system has been extensively applied to modify the eukaryotic genome in a targeted manner [1,2]. CRISPR-Cas9 technology takes advantage of the bacterial Cas9 endonuclease, which generates a double stranded break in its DNA target [1]. The repair of the break by the error prone repair machinery of non-homologous end joining often leads to the incorporation of mutations and permanent modifications to the genome [2].

(http://sigridjuselius.fi/apurahat/) (MR); University of Tampere Doctoral School (http://www.uta.fi/english/doctoralschool/index.html) (MU); Finnish Cultural Foundation – Maili Autio Fund (https://skr.fi/) (HB). 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.

Cas9 is directed to bind its target sequence by a single chimeric guide RNA molecule (sgRNA), which recognizes an approximately 20 nucleotide target site, followed by the three nucleotide protospacer adjacent motif (PAM)-sequence (5'-NGG-3') [1–3]. The sgRNA sequence is considered the limiting step in mutagenesis design, as the genomic target site needs to be unique. An optimal GC-content and specific nucleotides at key positions in the target sequence can also alter the efficiency and the specificity of mutagenesis [4–8]. The efficiency and unspecific, off-target binding of the nuclease are not easy to predict. As a result, multiple algorithms and online tools have been created for the identification of guide RNA targets with optimal Cas9 loading scores and the least amount of off-targets [9–17]. However, the *in silico* predictions do not always correlate with the observed mutagenesis efficiency and specificity [11,18–20].

Eukaryotic gene expression is regulated at the epigenetic level by packing of DNA into nucleosomes, which are formed by wrapping 146bp of DNA around a histone octamer [21]. These eukaryotic chromatin structures fundamentally differ from bacterial DNA packing, and being a prokaryotic enzyme, it is plausible that Cas9 cannot fully operate around all chromatin structures. Indeed, recent evidence indicates that chromatin influences Cas9 binding by limiting the accessibility of the target site [10,18,22–25]. Cas9 takes longer to scan for the target sites buried in heterochromatin, whereas targets located in euchromatin are more accessible, and thus easier to locate [24]. However, heterochromatin does not entirely prevent Cas9 from binding to potential target sites and despite binding, cleavage does not necessarily occur [22,24]. Target site accessibility is reflected in the tendency of Cas9 to act on secondary targets, so it plays an important role when designing effective sgRNAs with maximum efficiency and a minimal number of off-targets [10,17,18]. If the intended target is buried in heterochromatin, it is more probable that Cas9 binds to secondary targets and is more likely to find those in the exon regions in euchromatin [18]. Evidence supporting the involvement of chromatin accessibility in Cas9 binding has emerged in *in vitro* models, cell lines and in the zebrafish (*Danio rerio*) [10,17,23–26]. However, detailed understanding on which chromatin features contribute to chromatin accessibility this is still lacking.

Compared to cell lines, zebrafish can present additional challenges for genome editing. Compared to other vertebrates, the teleost specific genome duplication has resulted in multiple similar genes or pseudogenes and this can, in some instances, complicate the identification of unique targets for sgRNA. Secondly, to generate mutant zebrafish, the sgRNA and Cas9 are microinjected into the fertilized embryo, and mutagenesis occurs during the first hours of development [27]. Compared to cell lines, the fertilized, CRISPR-injected zygote presents a challenge for all mutagenesis techniques as it undergoes developmental and differentiation processes that require global changes in chromatin. Lastly, the first cell division in zebrafish takes place very rapidly (40 minutes after fertilization), when compared to the cell divisions for example in mice (reaching E1.5 at 24 hours post fertilization, hpf). Mutagenesis occuring after this first cell division may more likely lead to mosaicism.

During development, the chromatin landscape is under constant change in order to enable coordinated growth and differentiation [28–30]. The zygote is supported by the available maternal transcripts and the zygotic genome remains transcriptionally inactive until the maternal to zygotic genome activation (MZT) at the mid blastula transition (MBT) [31]. Our current understanding of zygotic chromatin is limited, but it has been shown that a specific histone modification pre-patterning marks developmentally active and inactive genes during development [32]. The nuclease accessibility of the developing, chromatin-packed genome of embryos remains poorly understood. Previously, it was observed that chromatin does not influence CRISPR-Cas9 targeting in zebrafish embryos in an MNase assay (Micrococcal nuclease assay), but later ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) results suggested that CRISPR-Cas9 is more likely to be successful when targeting open chromatin [9,17]. More

information on the influence of chromatin on CRISPR-Cas9 mutagenesis in model organisms is needed in order to improve the efficiency of genome engineering methodologies.

In this study, we observed discrepancies between the *in vitro* and *in vivo* activities of sgRNAs, and that selected sgRNAs differ for their onset of mutagenesis. We saw an association between successful mutagenesis and the transcript levels during early development. We looked further into the involvement of gene activation and chromatin in explaining the CRISPR-Cas9 mutagenesis efficiency in zebrafish embryos. Our results indicate that gene expression and chromatin openness are associated with the efficiency of CRISPR-Cas9 mutagenesis. However, we saw no association of mutagenesis efficiency with either exon methylation or histone H3 Lysine 4 trimethylation (H3K4me3) at promoters.

Results

Good *in vitro* activity of sgRNA does not assure *in vivo* efficacy

Analyzing the efficacy of different sgRNAs *in vivo* is laborious. To improve the screening for efficient sgRNAs, *in vitro* digestion of the target sequence can be used. We analyzed the mutagenesis activity of six sgRNAs first *in vitro* and then selected three for analysis *in vivo*. As shown in Fig 1, some sgRNAs with good *in vitro* efficiency presented low or no *in vivo* activity. This suggests that factors present *in vivo* prevent Cas9 from acting on its target site.

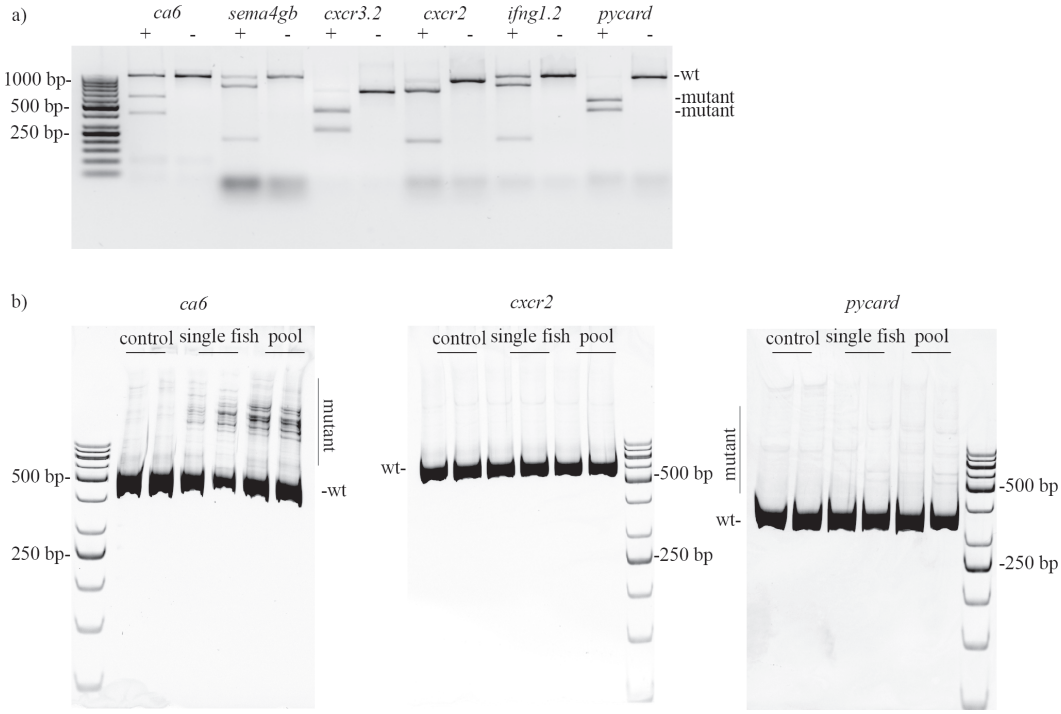


Fig 1. In vitro and in vivo CRISPR-Cas9 mutagenesis efficiencies do not correlate for all genes. a) An in vitro digestion assay shows that sgRNAs differ in their efficiencies. Below the gene name, + and - indicate the presence or absence of Cas9 protein in the reaction. On the right the wild type (wt) and the mutant products are indicated. b) The in vivo CRISPR-Cas9 mutagenesis visualized for ca6, cxcr2 and pycard with a heteroduplex mobility assay, with the wild type (wt) and the mutant products indicated. 5 embryos were collected per sample at 8hpf.

https://doi.org/10.1371/journal.pone.0196238.g001

Importantly, *cxcr2* had neither detectable gene expression nor mutagenesis efficiency, whereas the genes permissive for mutagenesis (*pycard*, *ca6*) showed early expression (Fig 1, S1 Fig). This led us to hypothesize that the onset and the level of gene expression could influence the CRISPR-Cas9 mutagenesis. The corresponding results using the T7 Endonuclease I assay are displayed in S2 Fig. In our hands the T7 Endonuclease I assay has a lower resolution compared to the heteroduplex mobility assay, especially with sgRNAs of lower efficiency. On the other hand, the T7 Endonuclease I assay can be readily used for quantitation of mutagenesis efficiency, especially with sgRNAs of higher efficiency.

The onset of mutagenesis differs between sgRNAs *in vivo*

As we saw a discrepancy between *in vivo* and *in vitro* mutagenesis efficiencies for some sgRNAs, we next analyzed whether the onset of mutagenesis correlates with the onset of gene expression. To avoid the delay of mRNA transcription for Cas9 activity, we used a ready Cas9 protein in our experiments with appropriate preincubation step to allow the sgRNA to complex with Cas9. Three of our functional sgRNAs were chosen for the analysis. The sgRNAs targeting *ca10a*, *sema4gb*, or *ca6* were co-injected with the Cas9 protein into the 1-cell stage embryo and the onset of mutagenesis was analyzed using both a heteroduplex mobility assay and a T7 Endonuclease I mutation detection assays. As shown in Fig 2 using the heteroduplex mobility assay, the first mutations become detectable as soon as 1hpf for *ca10a* and *sema4gb*, whereas the first mutations for *ca6* appeared at 3hpf (Fig 2). These results indicate that the onset of mutagenesis differs depending on the sgRNAs in zebrafish embryos. Based on these results, we analyzed the relationship of early gene expression and mutagenesis efficiency in more detail with all our sgRNAs. We were able to detect mutagenesis activity at 1hpf (roughly corresponding to 4-cell stage).

Likelihood of successful mutagenesis in relation to the expression level of the target gene in zebrafish embryos

Altogether, we have designed 86 sgRNAs using the crispr.mit.edu, ChopChop (V1 and V2) and CRISPRscan softwares [9,14,15]. Of these sgRNAs, 30% showed detectable *in vivo* activity (S1

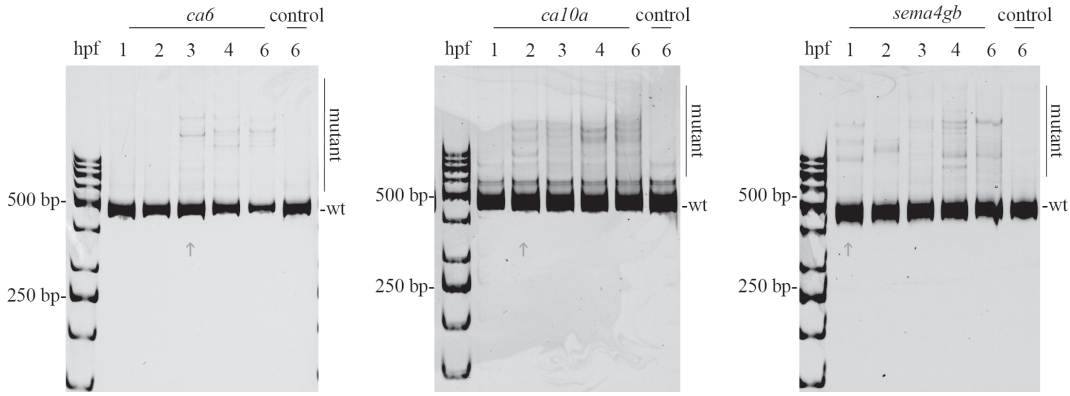


Fig 2. Onset of mutagenesis differs between sgRNAs. Heteroduplex mobility assay to demonstrate the onset of mutagenesis using high efficiency guide RNAs targeting three different genes with different gene expression patterns in early development. Embryos were collected at timepoints 1, 2, 3, 4, 6hpf (15–20 embryos per group). The gene name above the gel image indicates CRISPR-Cas9 injected embryos and control indicates uninjected controls. The legend on the side indicates the positions of wt (wild type) and mutant bands in the gel. Red arrows indicate the point at which first mutations can be detected.

https://doi.org/10.1371/journal.pone.0196238.g002

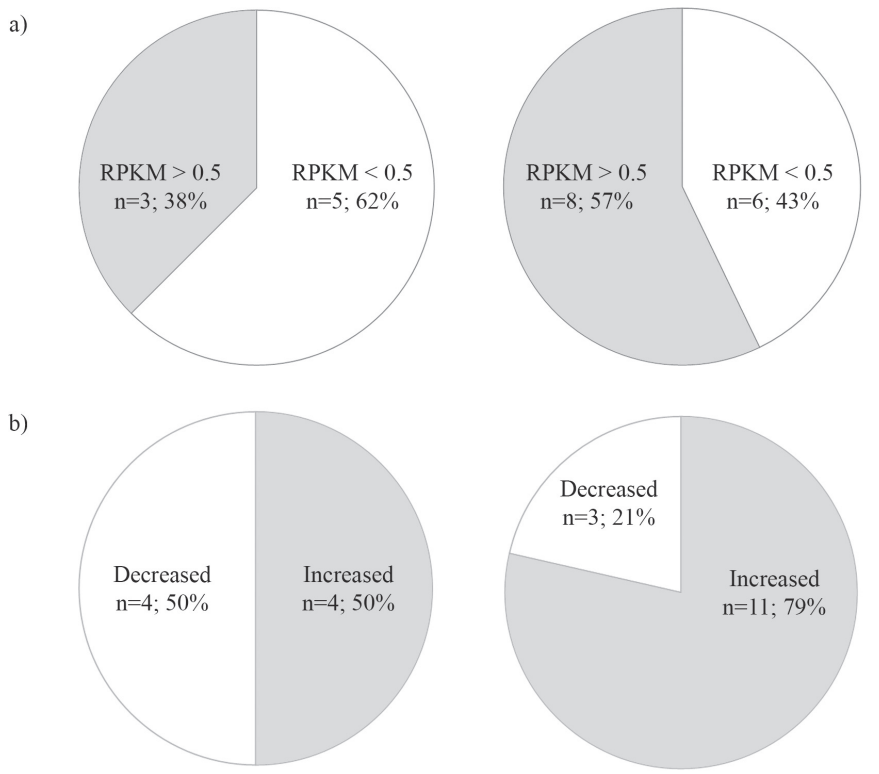


Fig 3. Relationship of transcript levels in early development and low mutagenesis efficiency. Pie charts of the RNA-seq data corresponding to graphs in S1 Fig. a) Number of transcripts for the genes resistant to (left) or permissive (right) for mutagenesis between the oblong sphere and the 15-somite stage (Fischer's exact test; not significant). 0.5 RPKM (Reads per Kilobase of transcript per Million mapped reads) was used as a limit for low expression. b) The number of genes resistant (left) or permissive (right) for CRISPR-Cas9 mutagenesis in which the number of transcripts is increased or decreased between the oblong sphere-stage and 50% epiboly (around the MZT) (Fischer's exact test; not significant).

https://doi.org/10.1371/journal.pone.0196238.g003

Table). As GC-content (%) has been suggested to influence the effectiveness of CRISPR-Cas9 mutagenesis, we analyzed the GC-content of our sgRNAs (S1 Table) [18]. The GC-content of our functional sgRNAs was found to be similar (Mann-Whitney U-test; p-value 0.452) to that of the non-functional sgRNAs.

When we compared the expression of the genes that we were able to mutate to those we were not, the genes resistant for mutagenesis more often had a very low expression level (Fig 3). However, the difference did not reach statistical significance (Fischer's exact test; not significant). Moreover, a majority of genes (79%) permissive for mutagenesis underwent an increase in the number of transcripts around the MZT (identified here as a positive change in the number of transcripts between the oblong sphere stage and 50% epiboly). This occurred more often than in the genes resistant to mutagenesis (50%). However, this observation was not statistically significant (Fischer's exact test) (Fig 3). To examine whether the lack of statistical significance was due to a type two error, we decided to determine whether there is a correlation between target gene expression and mutagenesis efficiency using larger datasets.

Table 1. Correlation between mutagenesis efficiency and gene expression at different developmental stages during early development. n = 209.

Developmental stage	Spearman correlation	p-value
64-cell	0.190	0.006*
oblong-sphere	0.227	0.001*
50%-epiboly	0.187	0.007*
15-somite	0.210	0.002*
36hpf	0.230	0.001*
48hpf	0.182	0.008*
60hpf	0.188	0.006*
72hpf	0.131	0.058

https://doi.org/10.1371/journal.pone.0196238.t001

Mutagenesis efficiency correlates with gene expression and chromatin accessibility in zebrafish embryos

In searching for factors that would explain the poor *in vivo* activity of some sgRNAs, we investigated available open access datasets. As genes with low expression values tended to be more difficult to mutate in our setting (Fig 3), we analyzed the association between expression levels and mutagenesis efficiency in greater depth, using large datasets in order to avoid type 2 error. We obtained CRISPR-Cas9 sgRNA efficiency data from CRISPRz database for all analyses [33]. We used open access RNA-seq data (E-GEOD-45706) for our primary analysis of the correlation between CRISPR-Cas9 mutagenesis and gene expression [33,34]. We found significant correlations in early development (between 64-cell stage and 36hpf), at the oblong sphere stage (3.66hpf, Spearman correlation 0.227; p-value 0.001) and at 36hpf (Spearman correlation 0.230; p-value 0.001). A strong correlation was observed at the oblong sphere stage which occurs shortly after MBT, around the time of zygotic genome activation. These results suggest that transcriptional activity influences CRISPR-Cas9 mutagenesis at early development (Table 1).

As methylation is known to correlate with transcriptional repression, we used zebrafish exon methylation data to analyze whether there is any correlation between exon methylation and the success of CRISPR-Cas9 mutagenesis [33,35]. As is shown in Table 2, there was no significant correlation between exon methylation and CRISPR-Cas9 mutagenesis efficiency at the 1-cell stage or at MBT (Table 2). Similarly, using open access data on embryonic histone methylation, we analyzed whether there is a correlation of H3K4me3 at promoters with CRISPR--Cas9 mutagenesis efficiency. As shown in Table 2, there seemed to be a correlation but this did not reach statistical significance (Spearman correlation 0.263; p-value = 0.074) [33,36].

ATAC-sequencing is a recent next generation sequencing method, which can be used to directly analyze chromatin accessibility. Open access ATAC-seq data for the zebrafish embryo is available at the 4hpf timepoint [37]. We compared mutagenesis efficiency data with ATAC-seq data at transcription start sites for a total of 263 genes. We discovered a significant, albeit rather weak correlation, indicating that chromatin accessibility appears to be one of the factors that explain the efficiency of CRISPR-Cas9 mutagenesis in zebrafish embryos (Table 2).

Table 2. Correlation between mutagenesis efficiency and chromatin features at different developmental stages and timepoints.

Chromatin feature	n	Developmental stage/Timepoint	Spearman correlation	p-value
Exon methylation	263	1-cell	0.115	0.063
		Mid blastula transition	0.107	0.084
H3K4me3	47	75–80% epiboly	0.263	0.074
Chromatin accessibility	263	4hpf	0.182	0.003*

https://doi.org/10.1371/journal.pone.0196238.t002

Discussion

In this study, we found discrepancies between the *in vitro* and *in vivo* efficiencies of some sgRNAs. These discrepancies suggested the presence of cellular factors which limit mutagenesis, and encouraged us to analyze chromatin involvement in more detail at the transcriptomic and epigenomic levels. Because the transcript counts of the early embryo can be masked by the presence of maternal transcripts, it is difficult to establish the exact relationship between gene expression and mutagenesis efficiency [38]. However, we found weak but significant correlations of gene expression with mutagenesis efficiency during early development, with the strongest correlation at the oblong sphere stage (3.66hpf, Spearman correlation 0.227; p-value 0.001) and later at 36hpf (Spearman correlation 0.230; p-value 0.001). The correlation at the oblong sphere stage suggests that genes which become active at the MZT are more accessible for Cas9 and hence undergo more efficient mutagenesis.

As chromatin structure is complex, its effect on target site accessibility has to be determined for each structural level, starting with direct modifications to DNA bases, continuing with analysis of histone modifications signaling for open chromatin, and ending with analysis of chromatin accessibility. A detailed analysis is required in order to understand how CRISPR--Cas9 mutagenesis activity could be manipulated at molecular level using for example chemical inhibitors of histone deacetylase activity. DNA methylation is known to mark transcriptional inactivity and recruit modified histones at the exons [39]. In our study, exon methylation was not found to significantly influence the activity of mutagenesis in zebrafish embryos. In confirmation, it has previously been suggested that Cas9 can act independently from DNA methylation in cell lines, and that, in general, most protein-DNA interactions are independent of DNA methylation [7,40]. If DNA methylation is not a limiting factor, we hypothesized that mutagenesis might correlate with higher order structures, specifically histone modifications. Various histone modifications mediate transcriptional activation and repression, and form nucleosome structures, which bind chromatin into an inactive heterochromatin state. H3K4me3 is a well known modification occurring in early development [41]. The most strongly suggestive, albeit not significant, correlation between experimental data and CRISPR-Cas9 mutagenesis efficiency was found with H3K4me3 data (Spearman correlation 0.263, p-value 0.07) [32]. This was expected, given the association with transcriptional activity at early developmental stages. As mutagenesis can already be detected at 1hpf it is possible that we fail to see a stronger correlation because the inspected timepoint is late and the histone landscape at 75–80% epiboly is dissimilar to that which is present before the MBT. In addition, if data from multiple timepoints would be available it would provide a more comprehensive view to opening of local chromatin structures. Also, observing only H3K4me3 signals might not accurately reflect the chromatin state in early embryos, as there are also other histone marks for open and closed chromatin, including H3K9me3 and H3K27me3 as well as H3K27ac at promoters [28,32,42]. A wider scale analysis of histone modifications could provide more insight into the association of CRISPR-Cas9 efficiency with histone landscape.

A higher order structure above the histone landscape is shaped by modified histones organizing into nucleosomes. Nucleosome occupancy, breathing and remodeling have previously been found to affect the cleavage activity of Cas9 and consequently, CRISPR-Cas9 mutagenesis is more successful when targeting the sequences depleted in nucleosomes [25,26,43]. The position of the PAM-sequence relative to nucleosomes has been found to be a key determinant of the Cas9 endonuclease activity *in vitro* but not in zebrafish [17,23]. Nucleosomes affect chromatin accessibility, which can be measured using ATAC-seq [44]. This state-of-the-art method has been used for identification of accessible chromatin regions during early development [37]. Using the publicly available data, we found a weak but significant correlation between

chromatin accessibility and mutagenesis efficiency at the MBT, indicating that chromatin influences the efficiency of CRISPR-Cas9 mutagenesis in zebrafish embryos, even though it is not the sole defining factor (Table 2) [37]. Our results are in line with those by others [17] with different analysis method and dataset. Moreover, our results suggest CRISPR-Cas9 mutagenesis efficiency to be independent of exon methylation and H3K4me3 at promoters.

Deciphering the effect of developmental chromatin on the activity of CRISPR-Cas9 mutagenesis model organisms ultimately leads us to an unanswered question about the regulation of zygotic genome activation and the signals that regulate this event at early stages before the MTZ [45]. The genome remains in a transcriptionally inactive state before the MZT, and it is likely that this inactive chromatin also limits the access of mutagenesis reagents such as Cas9. It is also likely that Cas9 can gain access during replication, and at sites that contain more permissive histone modifications or are depleted in nucleosomes, but only with limited efficacy. With further cell divisions, chromatin repressive signals then become diluted, leading to chromatin opening at the MZT and initiation of transcription [45]. Despite the biological significance of the MBT and MZT, we were already able to see mutagenesis taking place at 1hpf for some genes, so we propose that (when designing CRISPR-Cas9 mutagenesis strategies) chromatin structure should be taken into account at a very early timepoint (Fig 2).

Several studies have looked into the correlation of *in silico* predictions and *in vivo* activity of sgRNAs and found that CRISPR-sgRNA design tools often fail to accurately predict sgRNA activity [11,20,25]. Moreover, it has been observed, that the *in silico* predictions which are efficient for model organisms are not efficient for cell line based assays and *vice versa* [11]. As Haeussler *et al*. (2016) observed, CRISPR-Cas9 efficiency in mice is better predicted by the algorithms that have been trained on zebrafish experimental data, than by cell line based algorithms. It is logical to assume this is at least in part due to the fact that mice and zebrafish undergo similar, conserved developmental dynamics at the transcriptomic and epigenomic level (at the time when CRISPR-mutagenesis is taking place), and target site accessibility is largely defined by early chromatin. Thankfully, design tools, which also take into account target site accessibility, have recently become available [11,14,16,17]. Detailed analysis is required to pinpoint which are the most important chromatin structures impacting CRISPR-Cas9 activity. With a better understanding of these, we will hopefully achieve improvements in predictions for experimental design especially in the *in vivo* models. Eventually, it might be possible to modify local chromatin to increase target site accessibility and simultaneously decrease the likelihood of off-target binding. Our results confirm the involvement of chromatin in defining CRISPR-Cas9 mutagenesis efficiency in a vertebrate model *in vivo*.

Materials and methods

Zebrafish maintenance

Wild type AB fish were maintained in a flow-through system with a light/dark cycle of 14h/10h according to the standard procedure. Embryos and larvae were grown in an incubator (28.5˚C) in embryonic medium/E3 water (5mM NaCl, 0.17mM KCl, 0.33mM $CaCl_2$, 0.33mM $MgSO_4$, and 10–15% Methylene Blue).

Ethics statement and data availability

All experiments were carried out in accordance with the EU-directive 2010/ 63/EU on the protection of animals used for scientific purposes, and with the Finnish Act on the Protection of Animals Used for Scientific or Educational Purposes (497/2013) and the Government Decree on the Protection of Animals Used for Scientific or Educational Purposes (564/2013). We have only used zebrafish prior to their independently feeding larval stages in this study, which thus

do not require animal permits. Permit for the zebrafish housing and maintenance for the facility at the University of Tampere is ESAVI/10079/04.10.06/2015.

The computational data analysed in this study were collected from open access sources, as detailed in the appropriate sections.

Design and production of sgRNAs for CRISPR/Cas9 mediated genome editing

Target sequences (S1 Table) for sgRNA design were chosen using the online based CRISPR design tool (http://crispr.mit.edu/), ChopChop.V1 or V2 [14,15] or CRISPRscan [9]. Target site uniqueness was verified with the NCBI BLAST analysis against the zebrafish genome (GRCz10). sgRNAs were produced as described previously [46]. Briefly, the sgRNA oligo (Sigma-Aldrich) and the T7 promoter site oligo (S1 and S2 Tables) (Sigma-Aldrich) were annealed and *in vitro* transcribed using the MEGAshortscript T7 Transcription Kit (Ambion Life Technologies, CA, USA). The integrity and size of the produced sgRNAs were analyzed with gel electrophoresis (1% agarose in Tris-acetate-EDTA, TAE). The concentration of the sgRNAs was measured with the Qubit® RNA BR Assay kit (Thermo Fisher Scientific, MA USA 02451) and Nanodrop 2000 (Thermo Fischer Scientific).

sgRNA and Cas9 microinjection and genomic DNA extraction

The sgRNAs and the Cas9 protein (ToolGen Inc., Seoul, South Korea) were co-injected into one-cell stage zebrafish embryos with a micro injector (PV830 Pneumatic PicoPump, World Precision Instruments) under a Nikon microscope (SMZ645), using borosilicate needles prepared with a Flaming/Brown micropipette puller. Needles were calibrated by injecting solution into a halocarbon oil droplet to achieve a diameter of 12μm (approximately 1nl). The embryos were aligned on 1.2% agarose E3 water plates prior to the injection. An injection solution containing 130ng/μl sgRNA and 250ng/μl of the Cas9 protein in nuclease-free water was incubated 37˚C 15min. Rhodamine dextran was added to the solution for the visualization of the injections under a Zeiss Lumar V12 fluorescence microscope. To analyze the onset of the mutagenesis 10–20 CRISPR-Cas9 injected embryos were collected and frozen in liquid nitrogen for DNA extractions at 1, 2, 3, 4, 6hpf (hours post fertilization). To analyze the *in vivo* mutagenesis efficiency, 5 embryos were collectedat 8hpf and immediately frozen in liquid nitrogen. For DNA extraction, the embryos were lysed 4h 55˚C in lysis buffer (10mM Tris pH 8,2, 10mM EDTA, 200mM NaCl, 0.5% SDS, 200μg/ml Proteinase K). DNA was precipitated 1h -20˚C using two volumes of ethanol. DNA was then pelleted by centrifuging 16,000g 10min. The pellet was washed with 200μl of 70% ethanol before resuspending in 200μl of water. A purification step with phenol-chloroform was performed after treatment with 15u of RNase A (Thermo Fischer Scientific) per 100μl of sample, 1h 37˚C.

Heteroduplex mobility assay

Targeted loci were amplified from the genomic DNA by PCR using the Maxima Hot Start DNA polymerase (Thermo Fischer Scientific) according to the manufacturer's instructions. The PCR primers (S3 Table) were designed to anneal upstream and downstream of the expected cutting site. The PCR product was purified using Exo I and FastAP (Thermo Fischer Scientific) treatment 15min 37˚C, then 15min 85˚C. 10μl of the purified PCR product was annealed in a reaction containing 1x NEBuffer 2 (New England Biolabs, MA, USA) and was run on a 10% polyacrylamide gel. The gel was stained with GelRed (Bitium Inc., Fremont, CA).

T7 Endonuclease I mutation detection assay

After purifying and annealing the PCR amplified locus, 10μl of this product was incubated 30min 37˚C with 6 units of T7 Endonuclease I (New England Biolabs). The obtained products were separated on a 2.0% agarose TAE gel. The gel was stained with GelRed. The band sizes were compared to control samples.

In vitro digestion of DNA with the Cas9-gRNA complex

To test the *in vitro* cutting potential, equimolar amounts of the Cas9 protein (ToolGen Inc.) and sgRNA were pre-incubated 15min 37˚C in NEB 3 Buffer (New England Biolabs) and 1% Bovine serum albumin (Sigma Aldrich). For the template, a 850–1,200bp site around the target was amplified using Maxima Hot Start DNA polymerase according to the manufacturer's instructions. The template was then purified (GeneJET PCR Purification kit, Thermo Fischer Scientific). The template was then added to a final 10:10:1 ratio (Cas9:sgRNA:template PCR product). The reaction mix was incubated 3h 28˚C as this is the temperature at which zebrafish embryos are maintained. After this, we incubated the sample with 300U of Proteinase K 37˚C 10min to release the Cas9. Proteinase K was inactivated by incubation 65˚C 10min. Samples were run on a 1% agarose TAE gel to analyze the cutting efficiency.

Gene expression analysis of CRISPR targeted genes

The CRISPRz database contains a list of 1,398 validated zebrafish sgRNAs collected from various published resources [33]. In addition to sgRNA sequences, the associated mutagenesis efficiencies have been recorded in 325 unique zebrafish genes. We compared these mutagenesis efficiencies, from somatic cells, with a publicly available RNA-seq expression dataset housed in the ArrayExpress database [47]. The dataset (ArrayExpress E-GEOD-45706: https://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-45706) consists of RNA-seq data performed for samples from multiple stages of zebrafish development: 64-cell, oblong-sphere, 50%-epiboly, 15-somite, 36hpf, 48hpf, 60hpf and 72hpf (and 1 week, excluded from this analysis). Using the Stats package of the SciPy library, we performed Spearman rank correlation analyses of expression data for each sample in each ArrayExpress RNA-seq dataset; using the expression values for genes with available mutagenesis data for somatic cells in CRISPRz [48].

Histone modification in zebrafish promoters

The ArrayExpress dataset E-GEOD-4863 (https://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-4863/) is based on custom microarrays for the identification of ChIP binding sites of antibodies against the H3K4me3 in the promoters of zebrafish genes [36]. From the microarray datasets, the log of the median values of the 60-mer probes were summed for each gene, averaged, and then paired with CRISPRz mutagenesis values. Subsequently, these paired values were used to perform the Spearman rank correlation analysis.

Exon methylation analysis of zebrafish genes

McGaughey *et al*. showed that exon methylation was a better indication of mRNA expression than promoter methylation [35]. Their genome-wide ChIP-seq analysis of whole embryo zebrafish DNA methylation is available as an ArrayExpress dataset E-GEOD-52110 (https://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-52110/) at the 1-cell stage and at MBT. We first translated all ChIP-seq peaks from Zv9 genome coordinates to GRCz10 coordinates and then mapped them to exons annotated in the GRCz10 genome. A summation of all ChIP-seq peaks which overlapped exons was calculated for each gene, this sum was divided by the total length

of the gene's exons to generate a methylation coefficient. The methylation coefficients were then combined with mutagenesis data to compute Spearman rank correlations for each timepoint.

ATAC-seq analysis of zebrafish transcriptional units

ATAC-seq is a powerful method for identifying regions of accessible chromatin and it can be used to generate nucleotide resolution mapping of the hyperactive Tn5 transposase binding sites in the genome. An ATAC-seq analysis of 4hpf zebrafish has been previously completed and is available as an ArrayExpress dataset E-GEOD-74231 (https://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-74231/) [37]. SRR2747531 was downloaded from the NCBI Sequence Read Archive [37]. Reads were inspected using Fastqc version 0.11.5 and deemed to be of good quality and no further quality filtering or trimming was performed [49]. Subsequently, reads were aligned with Bowtie2 version 2.3.2 [50] using the parameter—very-sensitive-local against the Ensembl Zebrafish reference genome GRCz10. Alignments were filtered and sorted using samtools version 1.4 with the parameter -q 20. Duplicates were removed using Picard Markduplicates version 2.6.0 with the parameters REMOVE_DUPLICATES = TRUE VALIDATION_STRINGENCY = LENIENT [51]. As the alignment was performed against recent reference it was necessary to perform peak- calling independently of the original paper [37]. Furthermore, due to advances in peak calling softwares, peak calling was performed with macs2 version 2.1.1 using the parameters—nomodel—shift -100—extsize 200 -q 0.05 –broad [52]. Transcription start sites (TSSs) for all transcripts annotated in the GRCz10 genome were pooled for each gene. TSSs within 500nt were clustered as a single transcriptional unit. Of the total 22,152 zebrafish genes, 18,687 had a single transcript. Subsequent clustering created single transcriptional units in 2,233 of the remaining 3,465 genes with more than one annotated transcription start site. For clustered TSSs, the midpoint was used as the representative TSS. For each TSS, a +/- 1,000nt region was used to associate ATAC-seq peaks from the E-GEOD-74231 dataset. For each of these regions, an ATAC-seq coefficient was generated by summation of the product of ATAC-seq signal value by total overlap with the TSS region, divided by the length of the region (2,000nt). In cases where after clustering a gene still had more than one TSS ATAC-seq peak, a correlation was performed for all TSS regions and then averaged. Subsequently, all zebrafish genes possessed a single ATAC-seq coefficient. These were then combined with the mutagenesis data from the CRISPRz database in order to compute the Spearman rank correlation.

Supporting information

S1 Table. sgRNA target site sequences for each genomic target. sgRNAs used in the experiments in this paper are indicated by a * after the gene name. Functional (Yes/No) indicates observed in vivo activity.
(DOCX)

S2 Table. sgRNA template sequence. The extra 3' guanines (G/GG) were used if target sequence has one or two 5' guanines. N- indicates the position of the target sequence.
(DOCX)

S3 Table. Primers used in T7 Endonuclease I assay (T7EI), Heteroduplex mobility assay (HMA) and *In vitro* Digestion Assay (IVDA).
(DOCX)

S1 Fig. The relationship of mutagenesis efficiency and transcript level. The graph a) presents the expression of the genes resistant to CRISPR-Cas9 mutagenesis, at the early stages of

development. The graph b) presents the genes that were successfully mutated with CRISPR--Cas9. 2–10 sgRNAs have been used for mutagenesis. RPKM, Reads per Kilobase of transcript per Million mapped reads. All sgRNA sequences have been given in **S1 Table**.
(TIF)

S2 Fig. T7 endonuclease I assay results corresponding to Fig 1. The *in vivo* CRISPR-Cas9 mutagenesis efficiencies for selected genes estimated with the T7EI assay for *ca6*, *cxcr2* and *pycard*. 5 embryos were collected per sample at 8hpf. Black arrows indicate the mutated cleavage products for *ca6*.
(TIF)

Acknowledgments

We would like to thank Leena Mäkinen for assistance on conducting the zebrafish experiments and Heini Huhtala for invaluable advice on statistical analyses. We would also like to thank Anni Saralahti and Markus Ojanen for their input in the CRISPR-data collection and sgRNA design. Lastly, we would like to thank Helen Cooper for help in language revision and proofreading.

Author Contributions

Conceptualization: Meri I. E. Uusi-Mäkelä, Harlan R. Barker, Mika Rämet.

Data curation: Meri I. E. Uusi-Mäkelä, Harlan R. Barker, Tomi Häkkinen.

Formal analysis: Meri I. E. Uusi-Mäkelä, Harlan R. Barker.

Investigation: Meri I. E. Uusi-Mäkelä, Carina A. Bäuerlein.

Methodology: Meri I. E. Uusi-Mäkelä, Harlan R. Barker, Carina A. Bäuerlein, Matti Nykter.

Project administration: Meri I. E. Uusi-Mäkelä.

Resources: Mika Rämet.

Software: Harlan R. Barker, Tomi Häkkinen.

Supervision: Carina A. Bäuerlein, Matti Nykter, Mika Rämet.

Visualization: Meri I. E. Uusi-Mäkelä.

Writing – original draft: Meri I. E. Uusi-Mäkelä, Harlan R. Barker, Mika Rämet.

Writing – review & editing: Meri I. E. Uusi-Mäkelä, Harlan R. Barker, Carina A. Bäuerlein, Tomi Häkkinen, Matti Nykter, Mika Rämet.

References

1. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 2012 Aug 17,; 337(6096):816–821. https://doi.org/10.1126/science.1225829 PMID: 22745249
2. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science 2013 Feb 15,; 339(6121):819–823. https://doi.org/10.1126/science.1231143 PMID: 23287718
3. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature 2014 Mar 06,; 507(7490):62–67. https://doi.org/10.1038/nature13011 PMID: 24476820
4. Farboud B, Meyer BJ. Dramatic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. Genetics 2015 Apr; 199(4):959–971. https://doi.org/10.1534/genetics.115.175166 PMID: 25695951

5. Xu H, Xiao T, Chen C, Li W, Meyer CA, Wu Q, et al. Sequence determinants of improved CRISPR sgRNA design. Genome Res 2015 Aug; 25(8):1147–1157. https://doi.org/10.1101/gr.191452.115 PMID: 26063738
6. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. Science (New York, N.Y.) 2014 Jan 3,; 343(6166):80–84.
7. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nature biotechnology 2013 Sep; 31(9):827. https://doi.org/10.1038/nbt.2647 PMID: 23873081
8. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol 2014 Dec; 32(12):1262–1267. https://doi.org/10.1038/nbt.3026 PMID: 25184501
9. Moreno-Mateos MA, Vejnar CE, Beaudoin J, Fernandez JP, Mis EK, Khokha MK, et al. CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. Nat Methods 2015 Oct; 12(10):982–988. https://doi.org/10.1038/nmeth.3543 PMID: 26322839
10. Singh R, Kuscu C, Quinlan A, Qi Y, Adli M. Cas9-chromatin binding information enables more accurate CRISPR off-target prediction. Nucleic acids research 2015 Oct 15,; 43(18):e118. https://doi.org/10.1093/nar/gkv575 PMID: 26032770
11. Haeussler M, Schönig K, Eckert H, Eschstruth A, Mianné J, Renaud J, et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol 2016 - 7-5;17.
12. Zhu LJ, Holmes BR, Aronin N, Brodsky MH. CRISPRseek: a bioconductor package to identify target-specific guide RNAs for CRISPR-Cas9 genome-editing systems. PLoS ONE 2014; 9(9):e108424. https://doi.org/10.1371/journal.pone.0108424 PMID: 25247697
13. Rahman MK, Rahman MS. CRISPRpred: A flexible and efficient tool for sgRNAs on-target activity prediction in CRISPR/Cas9 systems. PLoS ONE 2017; 12(8):e0181943. https://doi.org/10.1371/journal.pone.0181943 PMID: 28767689
14. Labun K, Montague TG, Gagnon JA, Thyme SB, Valen E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. Nucleic Acids Res 2016 Jul 08,; 44(W1):272.
15. Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res 2014 Jul; 42(Web Server issue):401.
16. Chari R, Mali P, Moosburner M, Church GM. Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. Nat Methods 2015 Sep; 12(9):823–826. https://doi.org/10.1038/nmeth.3473 PMID: 26167643
17. Chen Y, Zeng S, Hu R, Wang X, Huang W, Liu J, et al. Using local chromatin structure to improve CRISPR/Cas9 efficiency in zebrafish. PLoS ONE 2017; 12(8):e0182528. https://doi.org/10.1371/journal.pone.0182528 PMID: 28800611
18. Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB, et al. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat Biotechnol 2014 Jul; 32(7):670–676. https://doi.org/10.1038/nbt.2889 PMID: 24752079
19. Friedland AE, Tzur YB, Esvelt KM, Colaiácovo MP, Church GM, Calarco JA. Heritable genome editing in C. elegans via a CRISPR-Cas9 system. Nat Methods 2013 Aug; 10(8):741–743. https://doi.org/10.1038/nmeth.2532 PMID: 23817069
20. Lee CM, Davis TH, Bao G. Examination of CRISPR/Cas9 design tools and the effect of target site accessibility on Cas9 activity. Exp Physiol 2017 Mar 16,.
21. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 1997 Sep 18,; 389(6648):251–260. https://doi.org/10.1038/38444 PMID: 9305837
22. Chen X, Rinsma M, Janssen JM, Liu J, Maggio I, Gonçalves, Manuel A F V. Probing the impact of chromatin conformation on genome editing tools. Nucleic Acids Res 2016 Jul 27,; 44(13):6482–6492. https://doi.org/10.1093/nar/gkw524 PMID: 27280977
23. Hinz JM, Laughery MF, Wyrick JJ. Nucleosomes Inhibit Cas9 Endonuclease Activity in Vitro. Biochemistry (N Y) 2015 DEC 8; 54(48):7063–7066.
24. Knight SC, Xie L, Deng W, Guglielmi B, Witkowsky LB, Bosanac L, et al. Dynamics of CRISPR-Cas9 genome interrogation in living cells. Science 2015 Nov 13,; 350(6262):823–826. https://doi.org/10.1126/science.aac6572 PMID: 26564855
25. Smith JD, Suresh S, Schlecht U, Wu M, Wagih O, Peltz G, et al. Quantitative CRISPR interference screens in yeast identify chemical-genetic interactions and new rules for guide RNA design. Genome Biol 2016 Mar 08,; 17:45. https://doi.org/10.1186/s13059-016-0900-9 PMID: 26956608

26. Horlbeck MA, Witkowsky LB, Guglielmi B, Replogle JM, Gilbert LA, Villalta JE, et al. Nucleosomes impede Cas9 access to DNA in vivo and in vitro. Elife 2016 MAR 17; 5:e12677. https://doi.org/10.7554/eLife.12677 PMID: 26987018

27. Hruscha A, Krawitz P, Rechenberg A, Heinrich V, Hecht J, Haass C, et al. Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. Development 2013 Dec; 140(24):4982–4987. https://doi.org/10.1242/dev.099085 PMID: 24257628

28. Bogdanovic O, Fernandez-Miñán A, Tena JJ, de la Calle-Mustienes E, Hidalgo C, van Kruysbergen I, et al. Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis. Genome Res 2012 Oct; 22(10):2043–2053. https://doi.org/10.1101/gr.134833.111 PMID: 22593555

29. Ho L, Crabtree GR. Chromatin remodelling during development. Nature 2010 Jan 28,; 463(7280):474–484. https://doi.org/10.1038/nature08911 PMID: 20110991

30. Andersen IS, Østrup O, Lindeman LC, Aanes H, Reiner AH, Mathavan S, et al. Epigenetic complexity during the zebrafish mid-blastula transition. Biochemical and Biophysical Research Communications 2012 January 27,; 417(4):1139–1144. https://doi.org/10.1016/j.bbrc.2011.12.077 PMID: 22209792

31. Kane DA, Kimmel CB. The zebrafish midblastula transition. Development 1993 Oct; 119(2):447–456. PMID: 8287796

32. Lindeman LC, Andersen IS, Reiner AH, Li N, Aanes H, Østrup O, et al. Prepatterning of developmental gene expression by modified histones before zygotic genome activation. Dev Cell 2011 Dec 13,; 21(6):993–1004. https://doi.org/10.1016/j.devcel.2011.10.008 PMID: 22137762

33. Varshney GK, Zhang S, Pei W, Adomako-Ankomah A, Fohtung J, Schaffer K, et al. CRISPRz: a database of zebrafish validated sgRNAs. Nucleic acids research 2016 Jan 4,; 44(D1):D826.

34. Yang H, Zhou Y, Gu J, Xie S, Xu Y, Zhu G, et al. Deep mRNA sequencing analysis to capture the transcriptome landscape of zebrafish embryos and larvae. PLoS ONE 2013; 8(5):e64058. https://doi.org/10.1371/journal.pone.0064058 PMID: 23700457

35. McGaughey DM, Abaan HO, Miller RM, Kropp PA, Brody LC. Genomics of CpG methylation in developing and developed zebrafish. G3 (Bethesda) 2014 /03; 4(5):861–869.

36. Wardle FC, Odom DT, Bell GW, Yuan B, Danford TW, Wiellette EL, et al. Zebrafish promoter microarrays identify actively transcribed embryonic genes. Genome Biol 2006; 7(8):R71. https://doi.org/10.1186/gb-2006-7-8-r71 PMID: 16889661

37. Kaaij LJT, Mokry M, Zhou M, Musheev M, Geeven G, Melquiond ASJ, et al. Enhancers reside in a unique epigenetic environment during early zebrafish development. Genome Biol 2016 Jul 05,; 17(1):146. https://doi.org/10.1186/s13059-016-1013-1 PMID: 27381023

38. Lee MT, Bonneau AR, Giraldez AJ. Zygotic Genome Activation During the Maternal-to-Zygotic Transition. Annual Review of Cell and Developmental Biology 2014 Oct 6,; 30(1):581–613.

39. Chodavarapu RK, Feng S, Bernatavichute YV, Chen P, Stroud H, Yu Y, et al. Relationship between nucleosome positioning and DNA methylation. Nature 2010 JUL 15; 466(7304):388–392. https://doi.org/10.1038/nature09147 PMID: 20512117

40. Domcke S, Bardet AF, Adrian Ginno P, Hartl D, Burger L, Schübeler D. Competition between DNA methylation and transcription factors determines binding of NRF1. Nature 2015 Dec 24,; 528(7583):575–579. https://doi.org/10.1038/nature16462 PMID: 26675734

41. Lorch Y, LaPointe JW, Kornberg RD. Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell 1987 Apr 24,; 49(2):203–210. PMID: 3568125

42. Vastenhouw NL, Schier AF. Bivalent histone modifications in early embryogenesis. Curr Opin Cell Biol 2012 JUN; 24(3):374–386. https://doi.org/10.1016/j.ceb.2012.03.009 PMID: 22513113

43. Isaac RS, Jiang F, Doudna JA, Lim WA, Narlikar GJ, Almeida R. Nucleosome breathing and remodeling constrain CRISPR-Cas9 function. Elife 2016 APR 28; 5:e13450. https://doi.org/10.7554/eLife.13450 PMID: 27130520

44. Tsompana M, Buck MJ. Chromatin accessibility: a window into the genome. Epigenetics Chromatin 2014; 7(1):33. https://doi.org/10.1186/1756-8935-7-33 PMID: 25473421

45. Pálfy M, Joseph SR, Vastenhouw NL. The timing of zygotic genome activation. Curr Opin Genet Dev 2017 Apr; 43:53–60. https://doi.org/10.1016/j.gde.2016.12.001 PMID: 28088031

46. Aspatwar A, Tolvanen MEE, Ojanen MJT, Barker HR, Saralahti AK, Bauerlein CA, et al. Inactivation of ca10a and ca10b Genes Leads to Abnormal Embryonic Development and Alters Movement Pattern in Zebrafish. Plos One 2015 JUL 28; 10(7):e0134263. https://doi.org/10.1371/journal.pone.0134263 PMID: 26218428

47. Kolesnikov N, Hastings E, Keays M, Melnichuk O, Tang YA, Williams E, et al. ArrayExpress update—simplifying data submissions. Nucleic Acids Res 2015 /01; 43(Database issue):1113.

48. Jones E, Oliphant T, Peterson P. SciPy: open source scientific tools for Python. 2014.

49. Andrews S. Fastqc. a quality control tool for high throughput sequence data. 2010;. Accessed 04.12., 2017.

50. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 2009; 10(3):R25. https://doi.org/10.1186/gb-2009-10-3-r25 PMID: 19261174

51. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009 AUG 15; 25(16):2078–2079. https://doi.org/10.1093/bioinformatics/btp352 PMID: 19505943

52. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based Analysis of ChIP-Seq (MACS). Genome Biol 2008; 9(9):R137. https://doi.org/10.1186/gb-2008-9-9-r137 PMID: 18798982

PUBLICATION II

NHLRC2 variants identified in patients with fibrosis, neurodegeneration, and cerebral angiomatosis (FINCA): characterisation of a novel cerebropulmonary disease.

Uusimaa J, Kaarteenaho R, Paakkola T, Tuominen H, Karjalainen MK, Nadaf J, Varilo T, **Uusi-Mäkelä M**, Suo-Palosaari M, Pietilä I, Hiltunen AE, Ruddock L, Alanen H, Biterova E, Miinalainen I, Salminen A, Soininen R, Manninen A, Sormunen R, Kaakinen M, Vuolteenaho R, Herva R, Vieira P, Dunder T, Kokkonen H, Moilanen JS, Rantala H, Nogee LM, Majewski J, Rämet M, Hallman M, Hinttala R.

Acta Neuropathologica 2018 May;135(5):727-742
doi: 10.1007/s00401-018-1817-z

Reprinted with the permission of Springer Nature.

NHLRC2 variants identified in patients with Fibrosis, Neurodegeneration, and Cerebral Angiomatosis (FINCA) – Characterisation of a novel cerebropulmonary disease

Johanna Uusimaa,[1–3,*] Riitta Kaarteenaho,[4,17] Teija Paakkola,[1,3,17] Hannu Tuominen,[5,6,18] Minna K. Karjalainen,[1,2,18] Javad Nadaf,[7,8,18] Teppo Varilo,[9,19] Meri Uusi-Mäkelä,[10,19] Maria Suo-Palosaari,[11] Ilkka Pietilä,[1,3] Anniina E. Hiltunen,[1,3] Lloyd Ruddock,[3,12] Heli Alanen,[3,12] Ekaterina Biterova,[3,12] Ilkka Miinalainen,[3] Annamari Salminen,[1,2] Raija Soininen,[3,12] Aki Manninen,[3,12] Raija Sormunen,[3,5] Mika Kaakinen,[3] Reetta Vuolteenaho,[3] Riitta Herva,[5] Päivi Vieira,[1,2] Teija Dunder,[1,2] Hannaleena Kokkonen,[13,14] Jukka S. Moilanen,[1,15] Heikki Rantala,[1,2] Lawrence M. Nogee,[16] Jacek Majewski,[7] Mika Rämet,[1,2,10] Mikko Hallman,[1,2] and Reetta Hinttala[1–3]

[1]PEDEGO Research Unit and Medical Research Center Oulu, University of Oulu, and Oulu University Hospital, PO Box 5000, FI-90014 Oulu, Finland
[2]Department of Children and Adolescents, Oulu University Hospital, PO Box 23, FI-90029 Oulu, Finland
[3]Biocenter Oulu, University of Oulu, PO Box 5000, FI-90014 Oulu, Finland
[4]Respiratory Research Unit and Medical Research Center Oulu, University of Oulu and Oulu University Hospital; Internal Medicine, Pulmonary Division, University of Oulu and Oulu University Hospital, PO Box 5000, FI-90014 Oulu, Finland
[5]Department of Pathology, Cancer and Translational Medicine Research Unit, University of Oulu, PO Box 5000, FI-90014 Oulu, Finland
[6]Department of Pathology, Oulu University Hospital, PO Box 23, FI-90029 Oulu, Finland
[7]McGill University and Génome Québec Innovation Centre, Montreal, Quebec, QC H3A 0G1, Canada
[8]St Jude Children's Research Hospital (SJCRH), 262 Danny Thomas Place, Memphis, TN 38105, USA
[9]Department of Medical Genetics, University of Helsinki, Haartmaninkatu 8, FI-00251 Helsinki, Finland
[10]BioMediTech Institute and Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland

[11]Department of Diagnostic Radiology and Medical Research Center Oulu, Oulu University Hospital and University of Oulu, PO Box 50, FI-90029 Oulu, Finland
[12]Faculty of Biochemistry and Molecular Medicine, University of Oulu, PO Box 5000, FI-90014 Oulu, Finland
[13]Northern Finland Laboratory Centre NordLab, Oulu University Hospital, PO Box 500, FI-90029 Oulu, Finland
[14] Department of Clinical Chemistry, University Oulu, PO Box 5000, FI-90014 Oulu, Finland
[15]Department of Clinical Genetics, Oulu University Hospital, PO Box 23, FI-90029 Oulu, Finland
[16]Division of Neonatology, Johns Hopkins University School of Medicine, CMSC 6-104A, 600 N. Wolfe St., Baltimore, Maryland 21287, USA

[17]These authors contributed equally to this work.
[18]These authors contributed equally to this work.
[19]These authors contributed equally to this work.

***Corresponding author:**

Johanna Uusimaa, MD, PhD

Professor of Paediatric Neurology

Tel: +358-8-3155819

Fax: +358-8-3155559

E-mail: johanna.uusimaa@oulu.fi

Abstract

A novel multi-organ disease that is fatal in early childhood was identified in three patients from two non-consanguineous families. These children were born asymptomatic but at the age of two months they manifested progressive multi-organ symptoms resembling no previously known disease. The main clinical features included progressive cerebropulmonary symptoms, malabsorption, progressive growth failure, recurrent infections, chronic haemolytic anaemia and transient liver dysfunction. In the affected children, a neuropathology revealed increased angiomatosis-like leptomeningeal, cortical and superficial white matter vascularisation and congestion, vacuolar degeneration and myelin loss in white matter, as well as neuronal degeneration. Interstitial fibrosis and previously undescribed granuloma-like lesions were observed in the lungs. Hepatomegaly, steatosis and collagen accumulation were detected in the liver.

A whole-exome sequencing of the two unrelated families with the affected children revealed the transmission of two heterozygous variants in the NHL repeat-containing protein 2 (NHLRC2); an amino acid substitution p.Asp148Tyr and a frameshift 2-bp deletion p.Arg201GlyfsTer6. NHLRC2 is highly conserved and expressed in multiple organs and its function is unknown. It contains a thioredoxin-like domain, however, an insulin turbidity assay on human recombinant NHLRC2 showed no thioredoxin activity. In patient-derived fibroblasts, NHLRC2 levels were low, and only p.Asp148Tyr was expressed. Therefore, the allele with the frameshift deletion is likely non-functional. Development of the *Nhlrc2* null mouse strain stalled before the morula stage. Morpholino knockdown of *nhlrc2* in zebrafish embryos affected the integrity of cells in the midbrain region.

This is the first description of a fatal, early-onset disease; we have named it FINCA disease based on the combination of pathological features that include fibrosis, neurodegeneration, and cerebral angiomatosis.

Keywords: central nervous system, cerebropulmonary disease, multi-organ disease, interstitial fibrosis, neurodegeneration, brain angiogenesis

Introduction

Recent developments in network-based approaches have led to the discovery of disease-causing genes, the unexpected links between apparently unrelated diseases, diagnostic biomarkers and therapeutic targets for cancer, diabetes and several neurodegenerative diseases including Alzheimer's, Parkinson's, and Huntington's diseases [1,22]. Rare, inherited mutations causing familial neurodegenerative diseases provide the molecular basis for the cellular pathways underlying the pathogenesis of diseases, including the accumulation of aberrant or misfolded proteins, protofibril formation, ubiquitin-proteasome system dysfunction, excitotoxic insult, oxidative and nitrosative stress, mitochondrial injury, synaptic failure, altered metal homeostasis and failure of axonal and dendritic transport [2]. Chronic, progressive pulmonary disorders include interstitial fibrosis, granulomatous changes and diffuse alveolar damage caused by various pathogenetic and genetic mechanisms [25]. However, the current understanding of the molecular basis of degenerative disease–related mechanisms, such as tissue fibrosis, neurodegeneration and angiogenesis, remains incomplete.

The research field focusing on molecular genetic aetiologies and cellular mechanisms of inherited human diseases has undergone a major revolution in the last decade because of modern research technology that includes whole-exome sequencing (WES) and bioinformatic analyses, which provide better means to identify novel proteins, cellular pathways and disease mechanisms. Herein, we characterise a novel, early-onset and progressive multi-organ disease presenting mainly as severe cerebropulmonary manifestations that eventually led to death in infancy. We present the clinical, radiological and histological characterisations of this multi-organ disease, which we named FINCA disease based on the unique histopathological findings of the patients (fibrosis, neurodegeneration, and cerebral angiomatosis). According to our WES analysis, all three affected children with FINCA disease harboured an identical combination of variants in the NHL repeat containing 2 (*NHLRC2*) gene encoding NHLRC2, a protein with a thioredoxin-like domain. However, its function is currently uncharacterized, enabling no direct biochemical assay to address the pathological role of the variants identified. To acquire more information about the function of NHLRC2 *in vivo*, we studied the effects of knocking out or down the expression of *Nhlrc2* in mice and zebrafish models, respectively. Finally, our data suggests that NHLRC2 is an important factor in the maintenance of multi-organ homeostasis, indicating a wider role for NHLRC2 in the pathogenesis of degenerative human diseases.

Materials and methods

Study subjects

As part of the diagnostic protocol, skin and skeletal muscle biopsy samples were collected from three patients with severe, early-onset undefined progressive cerebropulmonary and multi-organ disease. The patients were examined at the Department of Paediatrics of Oulu University Hospital from 2001–2003 (Patients 1 and 2 from Family A) and 2010–2011 (Patient 3 from Family B). Investigations included clinical assessments and radiological, histological, biochemical and molecular genetic analyses. Based on the severe, multi-organ phenotype, a mitochondrial disease was suspected, and muscle biopsies were performed; however, there were no specific abnormal findings. Mitochondrial respiratory chain enzyme activities were normal, there were no deletions in muscle mitochondrial DNA and the sequencing of the coding region of mtDNA revealed no pathogenic variants or exons and exon–intron boundaries of the genes encoding mitochondrial twinkle helicase and polymerase gamma. Furthermore, *CLN1*, *NKX2-1, SFTPB, ABCA-3* and *SFTPC* genes were sequenced to exclude infantile neuronal ceroid lipofuscinosis, brain–lung–thyroid syndrome and gene defects related to surfactant proteins; no pathogenic variants were found in these genes. The disease course was progressive, leading to the early deaths of all the patients.

Histopathological studies of tissue samples (Electronic Supplementary Material)

Tissue biopsies (lung and skeletal muscle biopsies) and autopsy samples were obtained from all three affected patients. The tissue was fixed in buffered 4% formaldehyde, routinely processed into paraffin blocks and cut into 5.0 μm sections. Hematoxylin-eosin

and luxol fast blue stainings were performed as previously described [8]

Transmission electron microscopy on patient tissue samples and zebrafish *nhlrc2* morphants

Autopsy samples from the lungs and liver from Patient 2 and whole zebrafish embryos at 2 dpf injected with random sequence (5.1 ng) or splice site-blocking morpholino (5.0 ng) were fixed in 1% glutaraldehyde and 4% formaldehyde in 0.1 M phosphate buffer and were then post-fixed with 1% osmium tetroxide (Electron Microscopy Sciences, Hatfield, PA, USA), dehydrated in acetone and embedded in Epon LX112 (#21210; Ladd Research Industries Inc., Williston, VT, USA). Thin sections were cut with a Leica Ultracut UCT microtome (FC6; Leica, Wetzlar, Germany), stained with uranyl acetate (Structure Probe Inc., West Chester, PA, USA) and lead citrate (Laurylab, Brindas, France) and examined with a Tecnai Spirit transmission electron microscope (FEI Company, Eindhoven, Netherlands). Images were captured with a Quemesa CCD camera (Olympus Soft Imaging Solutions, Münster, Germany). Specimens for electron microscopy were processed and analysed in the EM core facility at Biocenter Oulu (Oulu, Finland).

Molecular genetic analyses and genealogy (Electronic Supplementary Material)

Whole exome sequencing (WES) [26] and bioinformatics analysis [31] were performed on samples from both families. Using Sanger sequencing, we confirmed variants and segregation within the families. To investigate the heredity of the mutations, we traced the ancestors of the patients in the Finnish Population Registries and National Archives of Finland [30].

Screening of *NHLRC2* variants in population cohorts

To determine whether the two *NHLRC2* variants were present in the general Finnish population, we used two strategies. First, we examined these variants in a Finnish population of infants born at term (gestational age ≥37 wk, total *n*=306) sampled prospectively at Oulu University Hospital from 2004 to 2007 (*n* = 199) [11] and in 2014 (*n* =107). The individuals born in this hospital mostly originate from northern Finland; therefore, these infants represent a valid population control. Next, we searched the Sequencing Initiative Suomi (SISu) [7] and ExAC data [14] (http://exac.broadinstitute.org/;01,2017) for the variants. The SISu data currently covers exonic variants of 10,490 individuals of Finnish origin.

***NHLRC2* expression and function (Electronic Supplementary Material)**

After reverse-transcription PCR (RT-PCR), we used Sanger sequencing to investigate the expression of the mutant alleles. Protein expression in the whole-cell extracts of fibroblasts from all three patients and healthy control subjects, together with the homogenates of control autopsy samples from several types of human tissues, were analysed by immunoblotting. Two constructs containing either full-length human *NHLRC2* or the thioredoxin-like domain alone were expressed in the *E. coli* strain BL21 (DE3). An insulin turbidity assay to test thioredoxin activity was performed, as previously described [10].

Nhlrc2 knockout mice (Electronic Supplementary Material)

Heterozygous C57BL/6N-A^{tm1Brd} Nhlrc2$^{tm1a(KOMP)Wtsi}$/Wtsi mice were obtained from the Infrafrontier-EMMA repository (strain number EM:10219). These mice carry a mutated allele where the function of the *Nhlrc2* gene is inactivated by the insertion of a targeting cassette that contains the expression reporter beta galactosidase gene [27]. Heterozygous mice were cross-bred, pups were genotyped or pregnant females dissected and embryos analysed. The expression of *Nhlrc2* in embryos was studied by staining for the expression of the lacZ reporter. On embryonic day 2.5 (E2.5), morulae were flushed from the uterus in accordance with a previously described protocol [19]. Embryos that were grown overnight in a microculture drop were transferred onto a gelatinised 24-well plate, each into its own well with 300 µl of embryonic stem cell medium, and left to grow for 10–12 days, after which DNA was extracted, as described [23].

Generation of zebrafish morphants (Electronic Supplementary Material)

Two morpholinos against zebrafish *nhlrc2* (ENSDARG00000089581) were designed; the first morpholino was against the ATG-site to prevent translation and the second against the exon-intron boundary in the 3’ end of exon 4 (ENSDARE00000919598) to prevent splicing, which would both lead to a frameshift and premature termination of translation. The effect of the splice-site blocking morpholino for *nhlrc2* levels was quantified using PCR. The knockdown efficiency was determined from an agarose gel using ImageJ v1.49 by calculating the ratio of the wild-type (WT, the unmodified band) to the sum of all bands in a single lane (WT and two morpholino modified bands corresponding to exon exclusion and intron inclusion caused by splice site blocking).

Quantification of the affected cells in the midbrain region of zebrafish morphants from transmission electron micrographs

From five zebrafish injected with random sequence control morpholino (5.1 ng) and five zebrafish injected with *nhlrc2* splice site-blocking morpholino (5.0 ng), five random electron micrographs were taken from sagittal thin sections of the midbrain area. All the cells with visible nuclei and that were fully within the imaged area were counted and analysed for their integrity using iTEM software (Olympus Software Imaging Solutions GMBH). Cells with excess vacuolisation or disintegration were classified as affected.

Statistics

Differences were considered statistically significant (*) if the *p* value was 0.01– 0.05, very significant (**) if the *p* value was 0.001–0.01, and highly significant (***) if the *p* value was <0.001. A paired two-tailed student t-test was used to analyse the statistical significance. Zebrafish dechorionation timing differences were calculated with the log rank Mantel-Cox method, and values below 0.05 were considered significant.

Results

Clinical findings

Family A. Two brothers were born after normal consecutive singleton pregnancies to healthy parents in this non-consanguineous Finnish family. Both siblings presented at 2 months of age with feeding problems, muscular hypotonia, shoulder hypertonia and irritability. During the following months, both had delayed development, recurrent

infections, growth failure, respiratory problems and malabsorption. They subsequently developed dystonic tetraplegia, poor visual contact, progressive respiratory difficulties, transient liver dysfunction, and chronic haemolytic anaemia. The elder brother additionally presented with epileptic seizures and kidney dysfunction related to metabolic crisis at 9 months. The disease course was progressive and, despite intensive care, both patients died (at 1 year 9 months and 1 year 1 month, respectively) of progressive respiratory failure.

Family B. In this non-consanguineous Finnish family, the affected son was the first infant born to healthy parents. The mother had gestational hepatosis and toxaemia. At 2 months, the infant presented with irritability, floppiness, haemolytic anaemia, feeding problems, frequent diarrhoea and poor weight gain. At 4 months, he was hospitalised because of a suspected seizure, poor general condition and respiratory problems. During the following months, he had delayed development, dystonia, failure to thrive, strabismus, recurrent infections, and respiratory difficulties. The disease course was progressive, and despite intensive care, the patient passed away at 1 year 2 months due to respiratory failure.

All three patients appeared healthy at birth, but manifested at 2 months with unique multi-organ symptoms with a progressive disease course that resembled no known disease phenotype. Table 1 summarises the major clinical findings, including progressive cerebropulmonary symptoms, transient liver dysfunction, progressive growth failure and chronic haemolytic anaemia. Radiological investigations revealed over-inflated lungs with perihilar interstitial opacities, diffuse infiltrations and atelectasis (Online Resource Fig. 7) and thin corpus callosum (Online Resource Fig. 8). Additional findings included increased echogenicity of the kidneys (Patient 1) and hepatomegaly during metabolic crisis (Patients 1 and 2), enlarged thymus (Patient 2) (Online Resource Fig. 9), dilated lateral ventricles,

frontal sulci resembling cerebral atrophy (Patients 2 and 3) and increased signal intensity of the globus pallidum (Patient 3) (Online Resource Fig. 8). Visual evoked potentials (VEPs) revealed increased latencies and giant responses, indicating dysfunction of the visual tract (Patient 1). Electroencephalographs (EEGs) showed slow background activity and frontal bilateral rhythmic high-amplitude sharp delta transients, indicating metabolic encephalopathy (Patient 1). The EEG of Patient 3 was normal at 5 months, but at 11 months it was monotonic during sleep and consisted of 4-Hz delta waves without normal sleep spindles and vertex waves. Detailed descriptions of the clinical manifestations are presented in the Electronic Supplementary Material.

Histopathological findings in the brain

Post mortem neuropathological examinations were performed for all three patients. The findings in Patient 1 included brain atrophy (942 g, reference 1050 g); thin corpus callosum; small hippocampal regions and increased angiomatosis-like leptomeningeal, cortical and superficial white matter venous and capillary vascularisation and congestion (Fig. 1a). Superficial vascular changes were observed macroscopically. Widespread cerebral cortical oedema, focal neuronal loss and gliosis were observed (Fig. 1b and c) Symmetrical vacuolar degeneration and myelin loss of the fibres at the level of the crus cerebri in the midbrain were detected, as well as vacuolar degeneration of the white matter of the cerebellum, cerebellar oedema, focal Purkinje cell depletion and gliosis (Fig. 1d). Neuronal depletion of the anterior horns of the spinal cord was also found.

Patient 2 also had brain atrophy (764 g, reference 900 g), a thin corpus callosum and prominent congested leptomeningeal and superficial brain parenchymal vasculature.

Furthermore, the following neurodegenerative findings were detected: vacuolar degeneration and myelin loss of the corpus callosum and central white matter; vacuolar degeneration of the optic tract and the internal capsule; neuronal depletion of the corpora mammillaria, hypothalamic area and amygdaloid nucleus, as well as subthalamic and thalamic nuclei; vacuolar degeneration and myelin loss of the fibres of the crus cerebri and superior cerebellar peduncle and reticular formation in the midbrain, tegmentum and medial lemniscus in the pons (Fig. 1e-h). Other findings included mild vacuolar degeneration, myelin loss of the fibres of the pyramids and reticular formation in the medulla and in the white matter of the cerebellum. In addition, there was patchy neuronal loss in the anterior horns of the spinal cord.

Examinations of samples from Patient 3 revealed brain atrophy (742 g, reference 860 g), but a mostly normal corpus callosum, cerebral cortex and hippocampus. However, vascular congestion and angiectasia were detected. The temporal white matter exhibited vacuolar degeneration. Furthermore, the following findings were detected: vacuolar degeneration and neuronal depletion in the hypothalamic and amygdaloid nuclei and the putamen; vacuolar degeneration and myelin loss in the spinothalamic tract, medial lemniscus, superior cerebellar peduncle and reticular formation; degeneration of the anterior pontine nuclei; and minimal vacuolar degeneration in the white matter and dentate nucleus of the cerebellum.

Histopathological findings in the lungs

The lung biopsy samples from Patients 2 and 3 revealed no normal lung tissue; the lung tissue had been replaced by interstitial fibrosis resembling what is normally associated with

non-specific interstitial pneumonia (NSIP) (Fig. 2a and b). The alveolar epithelial cells appeared hyperplastic, and there were focal loose fibrotic areas positive after Alcian blue/periodic acid–Schiff (AB-PAS) staining. No hyaline membranes or diffuse alveolar damage (DAD) were visible.

Autopsy samples of the lungs were available from all three patients, and all the samples showed varying amounts of interstitial changes and fibrosis (Fig. 2c-h). Interstitial fibrosis in the lungs of Patient 1 was milder than in the lungs of his younger brother (Patient 2) (Fig. 2c). The lungs of Patient 2 revealed advanced interstitial fibrosis as a prominent feature, which was consistent with the findings in the lung biopsy samples; however, the degree of fibrosis in the autopsied lungs was more severe. In the lungs of Patient 3, the fibrosis was mainly moderate or mild. The general histological appearance of the autopsied lungs resembled that of NSIP. Intra-alveolar hyaline, hyaline membranes and squamous epithelial metaplasia reflecting DAD were intermittently observed in the lungs of Patients 1 and 3 but not in Patient 2 (Fig. 2g). Novel histopathological structures that we have termed granuloma-like lesions were observed in all cases, but in varying numbers; multiple lesions in five lung tissue sections in Patient 1, hundreds of lesions in five lung tissue sections in Patient 2 and fewer than 10 lesions in eight lung tissue sections of Patient 3 (Fig. 2c, e and h). Hyaline-resembling fibrotic areas were observed around the arteries in all cases, but the walls of the vessels were otherwise normal (Fig. 2d and e). Some airways were dilated, but no true honeycombing was found.

Most of the granuloma-like lesions were surrounded by hyaline-resembling extracellular material that contained large spindle-shaped cells. These large spindle-shaped cells were positive for alpha-smooth muscle actin (α-SMA) (myofibroblast marker) and

negative for cytokeratin (epithelial cell marker), phosphoglucomutase 1 (PGM1) (macrophage marker), smooth muscle myosin heavy chain (SMMHC) (smooth muscle cell marker) and caldesmon (smooth muscle cell marker) (Fig. 2f). The immunohistochemical phenotype indicated that the cells in the granuloma-like lesions were myofibroblasts.

Histopathological findings in the other organs

All three patients had hepatomegaly (435–485 g, reference weight 330–370 g). An autopsy liver sample from Patient 1 showed widespread microvesicular steatosis and foci of hepatocellular necrosis. Patient 1 also had cardiomegaly (90 g, reference weight 56 g) with slightly increased connective tissue. Autopsy revealed mild muscular atrophy in Patients 1 and 2; previous muscle biopsy samples from these patients revealed no histopathological changes. Furthermore, upon histological examination, the autopsy revealed atrophic thymus and adrenal glands without any specific abnormalities.

Electron microscopic findings in the lungs and liver

We analysed lung and liver autopsy samples from Patient 2 using a transmission EM (TEM). In the lungs, the most abundant cell types were myofibroblasts and smooth muscle and alveolar epithelial cells (Fig. 3a-c). Alveolar epithelial cells were mostly type II cells with intracellular lamellar bodies; only a few type 1 alveolar cells were detected (Fig. 3a and b). In the liver tissue samples, we observed steatosis (Fig. 3d) and collagen accumulation both inside- and outside of the hepatocytes, and a basement membrane was evident in the space of Disse (Fig. 3e and f).

Variants of *NHLRC2* in the two families

WES revealed compound heterozygous variants in *NHLRC2*, which encodes for the poorly characterised NHL repeat containing protein 2 in all three of the affected children (Fig. 4a). The *NHLRC2* gene is located in chromosome 10: 113,854,661-113,917,194 and the protein contains an N-terminal thioredoxin-like domain together with six NHL amino acid sequence repeats (named after three original genes *NCL-1, HT2A* and *LIN-41*) (Fig. 4b) [24]. Previously, mutations in *NHLRC2* have not been associated with any other human disease. The FINCA patients carried the heterozygous variant NM_198514:c.442G>T, which led to amino acid substitution p.Asp148Tyr in exon 3 within the highly conserved thioredoxin-like domain (Fig. 4b and d, Online resource Fig. 10). However, in the classical insulin-reduction activity assay for thioredoxins, neither full-length NHLRC2 nor the isolated thioredoxin-like domain produced in *E. coli* showed significant thioredoxin activity (Fig. 4c).

In addition, all three patients had a heterozygous frameshift 2-bp deletion NM_198514:c.601_602delAG, which caused p.Arg201GlyfsTer6 in exon 3 (Fig. 4a and b, Online Resource Fig. 10). Both the variants were segregated from their healthy parents. Genealogies of the patients, traced back seven to nine generations to the 1750s, revealed that the two families did not have common ancestors. We were unable to confirm the consanguinity of these two core families; their ancestors were born in northern Finland at least 150 miles apart.

Allele frequency of NHLRC2 variants among populations

We did not detect p.Asp148Tyr or p.Arg201GlyfsTer6 variants in the northern Finnish cohort of infants (*n*=306). In the Finnish population represented by Sequencing Initiative Suomi (SISu) data (*n*=10,488), p.Asp148Tyr (rs201701259) was rare, with a frequency of 0.003. In SISu, a frequency of 0.0001 was reported for p.Arg201GlyfsTer6 (rs757267294); however, the variant had not passed through quality control. In the Exome Aggregation Consortium (ExAc) data, the p.Asp148Tyr variant was present in non-Finnish European (*n*=33,351), South Asian (*n*=8,252), and African (*n*=5,202) populations with frequencies of 0.0003, 0.0002 and 0.0001, respectively. The p.Arg201GlyfsTer6 variant had a frequency of 0.0001 in non-Finnish Europeans (*n*=33,239) and was absent from other populations. The two variants were not detected as homozygous occurrences in any of the populations.

According to the SISu data, at least one of the two NHLRC2 variants, p.Asp148Tyr (rs201701259), was enriched in the Finnish population (n=10,488; natural logarithm of odds ratio 1.97), indicating that occurrence of the two variants in a compound heterozygous state in the two core families may be, at least partly, explained by their enrichment in the Finnish population.

Expression of Human *NHLRC2*

We investigated mRNA expression of variants NM_198514:c. 442G>T and c.601_602delAG in the patients' fibroblasts. *NHLRC2* mRNA expressed only the c.442G>T missense variant and not c.601_602delAG, confirming the origin of the variants from separate alleles and the degradation of mRNA with the frameshift 2-bp deletion

(Online Resource Fig. 10). The relative quantity of *NHLRC2* expression was studied by qPCR and showed approximately a two-fold decrease in patient-derived, immortalised fibroblasts compared to the controls (Fig. 4e).

The protein expression of NHLRC2 with the p.Asp148Tyr variant was investigated in the immortalised fibroblasts derived from patients and healthy controls. Immunoblotting revealed approximately a ten-fold decrease in the levels of NHLRC2 in whole-cell extracts from patients compared to control fibroblasts (Fig. 4f and g), further indicating the non-neutral role of the NHLRC2 variants.

Immunoblotting was also performed to investigate the expression of NHLRC2 in tissue homogenates from control autopsy samples (kidney, heart, muscle, liver, lung and brain tissue). NHLRC2 was detected in all human tissues included in the study (Online Resource Fig. 11). According to three public RNA expression databases (mouse) Brain RNA-Seq (https://web.stanford.edu/group/ barres_lab/brain_rnaseq.html) [32], Human Brain Transcriptome (http://hbatlas.org/) and the Atlas of the Developing Human Brain (www.brainspan.org)[13] the expression of mouse *Nhlrc2* is highest in astrocytes, oligodendrocyte progenitor cells, newly formed oligodendrocytes and endothelial cells. Mouse *Nhlrc2* is also expressed in neurons, myelinating oligodendrocytes and microglia. According to the Atlas of the Developing Human Brain (www.brainspan.org)[17], expression of human *NHLRC2* was highest in foetal brains during early pregnancy. The average expression levels of human *NHLRC2* are similar among the cerebellar cortex, mediodorsal nucleus of the thalamus, striatum, amygdala, hippocampus and 11 areas of the neocortex (Online Resource Fig. 12).

Nhlrc2^{tm1a} knockout mice

Nhlrc2 knockout (KO) mice (21 litters, *n*=154) were genotyped from heterozygous breeding pairs (Table 2). The ratio of born mice between WT and heterozygous mice was 1:2.125, which is in line with Mendelian segregation. The heterozygous mice, both female and male, were vital and bred well, indicating normal function of Nhlrc2^{tm1a} KO gametes. We also genotyped 53 embryos from six heterozygous breeding pairs on E10.5 and isolated and genotyped 45 morulae (E2.5) from six heterozygous breeding pairs (Online Resource Fig. 13). We did not detect any homozygous *Nhlrc2* KO mice, embryos or morulae, indicating that the lethality of Nhlrc2^{tm1a} KO homozygotes occurs between the fertilisation and morula (E2.5) stage. *Nhlrc2* expression analysis with LacZ-reporter mice indicated widespread expression throughout the body while WT littermates remained negative (Fig. 5). RNA *in situ* data of the Eurexpress atlas (www.eurexpress.org/ee/) showed strongest *Nhlrc2* expression in the brain, the central nervous system, the alimentary system and in the cardiovascular system of mouse embryos (E14.5) [4].

Zebrafish knockdown and dechorionation

We then decided to study the role of nhlrc2 in early embryogenesis using zebrafish, which is a convenient model organism for developmental studies because of the transparency of the embryos. The morphogenesis of the primary organ systems of a zebrafish is complete in 48 hours and the larvae display food seeking and avoidance behaviour 72 hours after conception [12]. First, we assessed the efficacy of *nhlrc2* splice site-blocking morpholino with PCR. As predicted, morphants had altered mRNA splicing compared to the untreated controls. The effect gradually faded toward 4 dpf (Online Resource Fig. 14). Besides minor

swelling in the abdomen, *nhlrc2* morphants developed normally (Online Resource Fig. 14) without any apparent developmental defects. Both splice site-blocking morpholino (3.6 ng) and translation-blocking morpholino (2.3 ng) were used, with both co-injected with an equal amount of p53 morpholino [21]. A slight delay was observed in the timing of the dechorionation (Online Resource Fig. 15a), but a similar effect was also observed in the random controls (Online Resource Fig. 15b), suggesting both were non-specific effects.

Because there were no apparent changes in the gross morphology of developing *nhlrc2* morphants, we next investigated the central nervous system in more detail using TEM. Because we were not able to ascertain the effect of the translation-blocking morpholinos without a validated antibody, we chose to use the splice site-blocking morpholinos for the TEM analysis. In this experiment, we achieved approximately 95.2% knock-down efficiency at the time point of 2 dpf by using the splice site-blocking *nhlrc2* morpholino (5.0 ng) (Online Resource Fig. 16). Using TEM, the quantitative analysis of the midbrain region from the 2 dpf morphants showed statistically significant increments in the number of affected cells in *nhlrc2* knock-downs compared to the controls injected with a random sequence morpholino (Fig. 6). The data indicate that *nhlrc2* has a key role in maintaining the cellular integrity of the central nervous system in developing zebrafish embryos.

Discussion

Based on clinical and molecular genetic data we have identified a novel cerebropulmonary and multi-organ disease, which we have named FINCA. This disease is characterised by a unique combination of tissue fibrosis, neurodegeneration and cerebral angiomatosis. Using

WES we identified identical compound heterozygous *NHLRC2* variants in three affected children from two unrelated families. In addition, we used a KO mouse model and zebrafish morphants to investigate the consequences of absent or decreased *Nhlrc2* expression; finally, we performed biochemical studies on human recombinant NHLRC2.

NHLRC2 is conserved in eukaryotes, having 84% sequence identity between human and mouse orthologues. The function of NHLRC2 is currently unknown. There are two NHLRC2 transcript variants in humans, both containing six NHL repeats that may form part of a TolB-like six-bladed beta-propeller. Six-bladed beta-propellers occur in a wide range of proteins, including 165 human proteins listed in InterPro [18]. They are often involved in protein–protein interactions. The longer isoform of NHLRC2 also has an N-terminal thioredoxin-like domain with an unusual motif (CCINC), where the active site of thioredoxin (WCGPC for human thioredoxins 1 and 2) is usually located. Interestingly, the missense variant identified from the patients in this study located in the thioredoxin-like domain and led, together with the frameshift variant, to significantly reduced expression of *NHLRC2* at the mRNA and protein level in patient-derived fibroblasts. A classical insulin-reduction activity assay for thioredoxins revealed that neither full-length NHLRC2 nor the isolated thioredoxin-like domain showed significant thioredoxin activity. This may be because of the substrate specificity of NHLRC2, which could be mediated by protein–protein interactions involving the six-bladed beta-propeller. Thus, insulin may be excluded from the active site; however, the thioredoxin fold motif is versatile, with both redox [16] and non-redox functions. An example of a non-redox function occurs in SasA, a circadian clock-associated histidine kinase [29]. This versatility plus the unusual motif in place of the active site indicates that the N-terminal domain has a thioredoxin fold, but no

thioredoxin activity. Conservation of the active-site cysteines indicates that the protein either has or regulates a redox-related function.

Based on large population cohort datasets, the rare variants identified in our patients are not restricted to the Finnish population; they are also present as simple heterozygotes in other populations. Our genotyping analysis of *Nhlrc2* KO mice, foetuses and morulae revealed that homozygosity was lethal whereas heterozygous mice were viable and non-symptomatic. Interestingly, the *nhlrc2* morphant zebrafish embryos had a significant number of affected cells in the midbrain, indicating that *nhlrc2* has an important role in the maintenance of cell survival in the developing brain. In contrast to the postnatal onset of the symptoms in humans, the timing of the effects of the zebrafish knockdown suggests the morphant zebrafish model does not fully reflect the features of the human disease. We are currently working on generating a knockout zebrafish model to better observe the potential symptoms at later stages of development.

Autopsy studies revealed that in FINCA patients, there was increased angiomatosis-like leptomeningeal, cortical and superficial white matter vascularisation and congestion, in addition to white matter degeneration and variable neuronal degeneration. According to currently available RNA expression data in Human Brain Transcriptome (http://hbatlas.org/), the average expression level of human *NHLRC2* is similar among different regions of the human brain. The expression of mouse *Nhlrc2* is the highest in astrocytes, oligodendrocyte progenitor cells, newly formed oligodendrocytes and in the endothelial cells [32]. Altogether, this expression data indicate that *NHLRC2* is widely expressed in the different regions and cell types of the brain. However, based on the current knowledge on the function of NHLRC2 it is difficult to conclude the primary mechanism

behind FINCA disease and to exclude the role of secondary processes such as ischemia as a cause of certain neuropathological findings described in our patients such as depletion of Purkinje cells and pyramidal neurons in Ammon's horn. Lung biopsy and autopsy studies showed severe interstitial fibrosis and previously undescribed granuloma-like lesions. The granuloma-like lesions were enriched with myofibroblasts that accumulate in areas surrounding granulomas in several granulomatous lung diseases such as sarcoidosis, atypical mycobacteriosis and tuberculosis [9]. The predominance of type II alveolar cells and paucity of type I cells were evident. Lung autopsy specimens additionally revealed DAD. Furthermore, all the patients had hepatomegaly. Liver samples for TEM were available only from one patient (Patient 2) revealing steatosis and accumulation of collagen bundles.

From the previous literature, we have found only one other "FINCA-like" cerebropulmonary disease called brain–lung–thyroid syndrome (BLT syndrome, OMIM 118700) or choreoathetosis and congenital hypothyroidism with or without pulmonary dysfunction (CAHTP, MIM #610978) that begins with muscular hypotonia followed by choreoathetosis, dystonia, ataxia, and dysarthria. This is an autosomal dominant infantile-onset disorder caused by mutations in thyroid transcription factor 1 (NKX2-1/TITF1; 14q13.3) [3,28]. The phenotype is less severe than in FINCA and varies between and within families; some patients show neonatal respiratory distress syndrome (RDS), developmental delay, symptoms of possible hypothalamic dysfunction or congenital cardiac septal defects [28]. Neuroimaging in some NKX2-1 related cases has shown structural brain abnormalities including agenesis of the corpus callosum [20], and autopsies have revealed reduced numbers of striatal and neocortical interneurons consistent with a defect in

neuronal migration [13]. In addition to RDS in neonates, pulmonary disease may present with interstitial lung disease in young children, and pulmonary fibrosis in older persons [20]. Histologic studies on patients with NKX2-1 related pulmonary dysfunction have identified interstitial widening and pneumocyte hyperplasia, desquamative interstitial pneumonia, accumulation of foamy alveolar macrophages, and pulmonary alveolar proteinosis [6]. The risk for pulmonary carcinoma is increased in young adults with an NKX2-1 related disorder. In BLT syndrome, thyroid involvement may present as subclinical hypothyroidism with mildly elevated thyroid stimulating hormone (TSH) or hyperthyrotropinemia. Additional symptoms, including failure to thrive, malabsorption, and intellectual deficit, have been reported in patients with deletions on chromosome 14, including NKX2-1 (www.orpha.net). In FINCA disease, in addition to cerebropulmonary symptoms, we also observed failure to thrive, malabsorption, early-onset chronic haemolytic anaemia, and recurrent infections in all three patients. Furthermore, transient liver dysfunction (Patients 1 and 2), transient kidney dysfunction (Patient 1), subclinical hypothyroidism (Patient 3, Electronic Supplementary Material), a transient increase in serum tumour markers and a non-specific decrease in the oxidative activity of monocytes (Patient 2, Electronic Supplementary Material) were noted.

Our studies indicate that NHLRC2 has a significant role in the CNS. However, further research is required to define the complex role of NHLRC2 in the pathogenesis of FINCA disease and, more commonly, its role in tissue fibrosis, neurodegeneration and cerebral angiomatosis. The presentation of FINCA disease and potentially other diseases may vary in response to the genetic defect in *NHLRC2*, environmental factors and other genetic factors. Recently, NHLRC2 has been listed as one of six novel blood-based

biomarkers for Alzheimer´s disease, indicating its yet undefined role in neurodegeneration [15]. Our expression analysis with LacZ-reporter mice indicates widespread expression of *Nhlrc2* throughout the mouse embryo. This, in connection with the severe multi-organ phenotype of FINCA patients, indicates that NHLRC2 has a vital role in early embryogenesis and in maintenance of multi-organ homeostasis after birth. We propose that NHLRC2 variants should be considered in patients with phenotypes that present as a combination of neurological and respiratory symptoms or additional multi-organ manifestations.

Acknowledgements

The authors would like to thank Professors Eric Shoubridge, Kalervo Hiltunen and Christer Betsholtz, Assistant Professor Michael Vanlandewijck, Adjunct Professor Siri Lehtonen and Dr. Riikka Pietilä for their expert advice and support and also Ms Pirjo Keränen, Ms Riitta Vuento, Ms Maarit Haarala, Ms Hanna Seppälä, Ms Kirsi Säkkinen, the Transgenic Core Facility at Biocenter Oulu, and the Laboratory Animal Centre at the University of Oulu for their expert assistance. Biocenter Oulu Electron Microscopy core facility, a member of Biocenter Finland, is acknowledged for their help with EM analysis. The zebrafish work was carried out at University of Tampere core facility, supported by Biocenter Finland. The digital pathology scanner of Northern Finland Biobank Borealis was used in imaging the neuropathological findings. This work was conducted with support from the Research Council for Health of the Academy of Finland (J.U., decision number 138566; R.H., decision numbers 266498, 273790 and 303996; M.H., decision number 1126662; L.R., decision numbers 266457 and 272573); the Sigrid Juselius Foundation

(J.U., R.H. and M.H.); the Foundation for Paediatric Research, Finland (J.U. and M.K.K.); the Alma and KA Snellman Foundation (J.U. and M.K.K.); a Marie Curie International Outgoing Fellowship of the European Union's Seventh Framework Programme (grant agreement number 273669 [BioMit]) (R.H.); Foundation of the Finnish Anti-Tuberculosis Association (R.K.); the Jane and Aatos Erkko Foundation (M.R.); the Competitive State Research Financing of the Expert Responsibility Area of Tampere University Hospital (MR); Special State Grants for Health Research in the Department of Paediatrics and Adolescence at Oulu University Hospital, Finland (J.U.); the National Heart, Lung and Blood Institute of the U.S. National Institutes of Health under award number HL-54703 (L.M.N.) and the Eudowood Foundation (L.M.N.).

Compliance with ethical standards

Study and ethical approval

All procedures performed in the studies involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Prior to the study, the guardians of the patients gave written informed consent to participate in the studies, and this was approved by the Ethics Committee of Oulu University Hospital (EETTMK 51/2008). Furthermore, the guardians of the patients in this manuscript have given written informed consent for the publication of their case details.

All procedures performed in the studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted. The National Animal Experiment Board of Finland approved the study protocol

(ESAVI/5882/04.10.07/2014). Animal care and experimental procedures were conducted in accordance with the national legislation and EU Directive 2010/63/EU. Zebrafish housing and maintenance were done according to facility permission ESAVI/10079/04.10.06/2015. All applicable international, national and/or institutional guidelines for the care and use of animals were followed.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Barabasi AL, Gulbahce N, Loscalzo J (2011) Network medicine: a network-based approach to human disease. Nat Rev Genet 12:56-68. Doi:10.1038/nrg2918

2. Bossy-Wetzel E, Schwarzenbacher R, Lipton SA (2004) Molecular pathways to neurodegeneration. Nat Med 10 Suppl:S2-9. Doi:10.1038/nm1067

3. de Vries BB, Arts WF, Breedveld GJ, Hoogeboom JJ, Niermeijer MF, Heutink P (2000) Benign hereditary chorea of early onset maps to chromosome 14q. Am J Hum Genet 66:136-142. Doi:S0002-9297(07)62240-X

4. Diez-Roux G, Banfi S, Sultan M, Geffers L, Anand S, Rozado D, et al. (2011) A high-resolution anatomical atlas of the transcriptome in the mouse embryo. PLoS Biol 9:e1000582. Doi:10.1371/journal.pbio.1000582

5. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792-1797. Doi:10.1093/nar/gkh340

6. Hamvas A, Deterding RR, Wert SE, White FV, Dishop MK, Alfano DN, et al. (2013) Heterogeneous pulmonary phenotypes associated with mutations in the thyroid transcription factor gene NKX2-1. Chest 144:794-804. Doi:S0012-3692(13)60595-4

7. Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland (2016) Sequencing Initiative Suomi project (SISu).http://sisuproject.fi. Accessed 1/10 2017

8. J. D. Bancroft, A. Stevens (1991) Theory and practice of histological techniques. John Wiley & Sons, Churchill Livingstone, Edinburgh.

9. Kaarteenaho-Wiik R, Sademies O, Paakko P, Risteli J, Soini Y (2007) Extracellular matrix proteins and myofibroblasts in granulomas of sarcoidosis, atypical mycobacteriosis, and tuberculosis of the lung. Hum Pathol 38:147-153. Doi:S0046-8177(06)00417-5

10. Karala AR, Ruddock LW (2010) Bacitracin is not a specific inhibitor of protein disulfide isomerase. FEBS J 277:2454-2462. Doi:10.1111/j.1742-4658.2010.07660.x

11. Karjalainen MK, Huusko JM, Ulvila J, Sotkasiira J, Luukkonen A, Teramo K, et al. (2012) A potential novel spontaneous preterm birth gene, AR, identified by linkage and association analysis of X chromosomal markers. PLoS One 7:e51378. Doi:10.1371/journal.pone.0051378

12. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF (1995) Stages of embryonic development of the zebrafish. Dev Dyn 203:253-310. Doi:10.1002/aja.1002030302

13. Kleiner-Fisman G, Rogaeva E, Halliday W, Houle S, Kawarai T, Sato C, et al. (2003) Benign hereditary chorea: clinical, genetic, and pathological findings. Ann Neurol 54:244-247. Doi:10.1002/ana.10637

14. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. (2016) Analysis of protein-coding genetic variation in 60,706 humans. Nature 536:285-291. Doi:10.1038/nature19057

15. Long J, Pan G, Ifeachor E, Belshaw R, Li X (2016) Discovery of Novel Biomarkers for Alzheimer's Disease from Blood. Dis Markers 2016:4250480. Doi:10.1155/2016/4250480

16. Martin JL (1995) Thioredoxin--a fold for all reasons. Structure 3:245-250. Doi:S0969-2126(01)00154-X

17. Miller JA, Ding SL, Sunkin SM, Smith KA, Ng L, Szafer A, et al. (2014) Transcriptional landscape of the prenatal human brain. Nature 508:199-206. Doi:10.1038/nature13185

18. Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, et al. (2015) The InterPro protein families database: the classification resource after 15 years. Nucleic Acids Res 43:D213-21. Doi:10.1093/nar/gku1243

19. Nagy A, Gertsenstein M, Vintersten K, Behringer R (2003) Manipulating the Mouse Embryo, A Laboratory Manual, 3rd ed. p.198-200. Cold Spring Harbor Laboratory Press, New York

20. Patel NJ, Jankovic J (2014) NKX2-1-Related Disorders. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Mefford HC, et al. (eds), GeneReviews(R), University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved, Seattle (WA).

21. Robu ME, Larson JD, Nasevicius A, Beiraghi S, Brenner C, Farber SA, et al. (2007) P53 Activation by Knockdown Technologies. PLoS Genet 3:e78. Doi:06-PLGE-RA-0378R3

22. Santiago JA, Potashkin JA (2013) Integrative network analysis unveils convergent molecular pathways in Parkinson's disease and diabetes. PLoS One 8:e83940. Doi:10.1371/journal.pone.0083940

23. Scavizzi F, Ryder E, Newman S, Raspa M, Gleeson D, Wardle-Jones H, et al. (2015) Blastocyst genotyping for quality control of mouse mutant archives: an ethical and economical approach. Transgenic Res 24:921-927. Doi:10.1007/s11248-015-9897-1

24. Slack F, Ruvkun G (1998) Heterochronic genes in development and evolution. Biol Bull 195:375-376. Doi:10.2307/1543152

25. Spagnolo P, Grunewald J, du Bois RM (2014) Genetic determinants of pulmonary fibrosis: evolving concepts. Lancet Respir Med 2:416-428. Doi:10.1016/S2213-2600(14)70047-5

26. Sulonen AM, Ellonen P, Almusa H, Lepisto M, Eldfors S, Hannula S, et al. (2011) Comparison of solution-based exome capture methods for next generation sequencing. Genome Biol 12:R94-2011-12-9-r94. Doi:10.1186/gb-2011-12-9-r94

27. Testa G, Schaft J, van der Hoeven F, Glaser S, Anastassiadis K, Zhang Y, et al. (2004) A reliable lacZ expression reporter cassette for multipurpose, knockout-first alleles. Genesis 38:151-158. Doi:10.1002/gene.20012

28. Thorwarth A, Schnittert-Hubener S, Schrumpf P, Muller I, Jyrch S, Dame C, et al. (2014) Comprehensive genotyping and clinical characterisation reveal 27 novel NKX2-1 mutations and expand the phenotypic spectrum. J Med Genet 51:375-387. Doi:10.1136/jmedgenet-2013-102248

29. Vakonakis I, Klewer DA, Williams SB, Golden SS, LiWang AC (2004) Structure of the N-terminal domain of the circadian clock-associated histidine kinase SasA. J Mol Biol 342:9-17. Doi:10.1016/j.jmb.2004.07.010

30. Varilo T, Savukoski M, Norio R, Santavuori P, Peltonen L, Jarvela I (1996) The age of human mutation: genealogical and linkage disequilibrium analysis of the CLN5 mutation in the Finnish population. Am J Hum Genet 58:506-512

31. Witkowski L, Carrot-Zhang J, Albrecht S, Fahiminiya S, Hamel N, Tomiak E, et al. (2014) Germline and somatic SMARCA4 mutations characterize small cell carcinoma of the ovary, hypercalcemic type. Nat Genet 46:438-443. Doi:10.1038/ng.2931

32. Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keeffe S, et al. (2014) An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. J Neurosci 34:11929-11947. Doi:10.1523/JNEUROSCI.1860-14.2014

Figure captions

Fig. 1 Histopathological findings in brain autopsy samples from patients with NHLRC2 variants. **a** Autopsy findings from Patient 1 revealed increased angiomatosis-like leptomeningeal vascularisation. **b** Degeneration and depletion of the pyramidal cells of the hippocampus. **c** Higher magnification of the Ammon's horn showing gradual depletion of pyramidal neurons. **d** Purkinje cells of the cerebellum (between arrowheads). **e-h** Patient 2 had vacuolar degeneration and myelin depletion in the pons. Haematoxylin and eosin stain (**a-f**) and luxol fast blue stain (**g, h**). The images were taken with an Aperio AT2 digital pathology slide scanner (Leica Biosystems, Wetzlar, Germany). Scale bars: 200 µm

Fig. 2 Histological findings in lung samples from patients with NHLRC2 variants. **a, b** Interstitial fibrosis was present in the lung biopsy samples. **c, e, f, h** Granuloma-like lesions (arrows) within regions with variable amounts of necrosis (stars) were present in the autopsy samples from each patient. **d** Perivascular fibrosis was visible around the artery (arrowhead). **f** Spindle-shaped cells in a granuloma-like lesion were positive for alpha–smooth muscle actin (short arrows). **g** Hyaline membranes (short arrows) were present in the alveoli. **a, b** Lung samples collected by lung biopsy (**a, b**) and during autopsy (**c-h**). Samples from Patient 1 (**c**), Patient 2 (**a, d-f**) and from Patient 3 (**b, g, h**). Haematoxylin and eosin stain (**a-e, g**). Scale bars: 80 µm

Fig. 3 Electron microscopy analysis of lung (a-c) and liver (d-f) autopsy samples. **a** Lung alveoli contained type II pneumocytes (P) and the alveolar space was very limited (L, the

lumen of the alveolus). **b** In addition to type II pneumocytes (P), myofibroblasts (M) were present with diverging 'limb-like' projections. **c** Myofibroblasts with ultrastructural features such as abundant endoplasmic reticulum (E) and adherens junctions (A) had produced fibrotic extracellular collagen bundles (C). **d** The liver had many large fat (F) vacuoles in the hepatocytes. **e** Thin and fragmentary basement membrane (B) was detected in the space of Disse adjacent to the endothelial cells surrounding the sinusoidal lumen (L). **f** The accumulation of fat vacuoles (F) and collagen bundles (C) in hepatocytes. Sinusoidal lumen (L). Scale bars: 10 µm (**a, d**), 5 µm (**b**), 2 µm (**c, e, f**)

Fig. 4 Variants p.Asp148Tyr and p.Arg201GlyfsTer6 in NHLRC2 in patients with FINCA disease from two non-consanguineous families. **a** Pedigrees and the segregation of heterozygous variants p.Asp148Tyr and p.Arg201GlyfsTer6 in Families A and B. The probands of the families are indicated with an arrow and the healthy siblings of the patients are indicated with empty symbols. **b** The structural domains of NHLRC2 and location of the variants. **c** Neither full-length NHLRC2 (NHLRC2-FL) nor the isolated thioredoxin-like domain (NHLRC2-trx) produced in *E. coli* showed significant thioredoxin activity in the insulin reduction assay. **d** Conservation of the thioredoxin-like domain region containing p.Asp148Tyr (p.D148Y) in red [5]. **e** QPCR showed a highly significant decrease in the relative quantity of *NHLRC2* mRNA in patient-derived fibroblasts compared to controls according to a two-tailed student t-test. **f** Immunoblotting indicated a significant loss of the p.Asp148Tyr NHLRC2 protein in fibroblasts from Patients 1–3 compared to the levels of GAPDH. **g** Immunoblotting data presented as numerical values measured using Image J and statistical analysis by a two-tailed student t-test.

Abbreviations: D148Y, p.Asp148Tyr; R201GfsX6, p.Arg201GlyfsTer6; Trx, thioredoxin; C, control; P, patient; FL, full length. Differences were considered as statistically significant (*), with *p* value 0.01-0.05 and highly significant (***) with *p* -value < 0.001

Fig. 5 LacZ mouse tissue staining. Heterozygote Nhlrc2^{tm1a} mice had a strong β-galactosidase signal throughout the body. Heart and thymus (**a**); Lungs (**b**); Kidneys, adrenal glands and gonads (**c**); Liver (**d**) and Brain (**e, f**) (**e** WT; **f** heterozygote Nhlrc2^{tm1a} mice). WT littermates were negative for β-galactosidase signal

Fig. 6 Transmission electron microscopy image from the midbrain region of the 2-dpf-zebrafish morphant. **a** A cross section of a zebrafish embryo at 2 dpf with a circled area specifying the exact location from which the electron micrographs were taken. **b** Cells with excess vacuolisation or disintegration were classified as affected (marked with an asterisk in Fig. 6d). In the midbrain 5.8 % of the 810 cells analysed were affected by the *nhlrc2* splice site-blocking morpholinos, whereas only 1% of the 850 cells were affected in MO controls. **c** Midbrain of the random sequence morpholino injected control. **d** Midbrain of the *nhlrc2* splice site-blocking morphant. Abbreviations: SB, splice site-blocking; MO, morpholino. Differences were considered as statistically very significant (**) with *p* value 0.001-0.01

Electronic supplementary material

Fig. 7 Radiological findings of chest X-rays and high-resolution computed tomography of Patients 1–3. **a** Patient 1 at 7 months of age shows flattened hemidiaphragms and hyperlucent lungs demonstrating over-inflated lungs, small air-filled bullae of the right mediobasal lung (short arrows) and perihilar small atelectasis (long arrow). Performed during an acute metabolic crisis at the age of 9 months, the chest X-ray demonstrates perihilar and basal airspace opacities and axial HRCT confirms bilateral consolidations (long arrows) and reticular opacities (short arrows). **b** Patient 2 at 4 months of age demonstrates mediastinal and hilar prominence (short arrows), over-inflated lungs, and diffuse reticular interstitial opacities. At 4 months of age, HRCT reveals widespread bilateral ground-glass opacities (short arrow), interstitial infiltrations (long arrows) and atelectasis (black arrowhead), mildly enlarged thymus (white arrowhead) and left hilar adenopathy (asterisk). At 1 year of age, chest X-rays show the progression of interstitial infiltrations and bilateral dense consolidations (long arrows). **c** Patient 3 at 4 months of age reveals over-inflated lungs and bilateral perihilar interstitial opacities (short arrows). At 13 months of age, the chest X-ray shows diffuse infiltrations with reticulation. HRCT confirms diffuse bilateral ground-glass opacification, peripheral interstitial septal thickening and lobular pleural thickening (long arrows)

Fig. 8 Findings from brain magnetic resonance imaging of Patients 1–3. **a** Sagittal T1-weighted magnetic resonance imaging (MRI) shows the thin corpus callosum of Patient 1 at 7 months of age. **b** Patient 2 at 10 months of age. **c** Patient 3 at 10 months of age (short arrows). The axial T2-weighted and coronal T1-weighted MRIs of Patients 2 (**b**) and 3 (**c**)

demonstrate slightly dilated lateral ventricles and cortical sulci. **c** Increased signal intensity of the globus pallidus on the axial T2-weighted MRI of Patient 3 (arrowheads)

Fig. 9 Findings from abdominal ultrasound and magnetic resonance imaging of internal organs and thymus of Patients 1 and 2. **a** The abdominal ultrasound findings (Patient 1 at 10 months of age) show increased cortical echogenity with the strand-like hypoechogenic outerzone of both kidneys (short arrows). **b** The abdominal T2-weighted axial MRI of Patient 2 at 4 months of age shows hepatomegaly and a non-expansive hypointense lesion of the right liver lobe (long arrows). **c** The lesion is also hypointense on the T1-weighted coronal MRI (long arrows). **d** The axial fat saturation T1-weighted MRI of Patient 2 shows that the lesion (arrowheads) in the liver is less enhanced by the gadolinium contrast agent than the rest of the liver. **e** Normal echogenity of the right liver lobe of Patient 2 at the age of 4 months. **f** A slightly enlarged thymus of Patient 2, which was seen in the coronal T1-weighted MRI (asterisk)

Fig. 10 a Electropherograms showing both heterozygote NM_198514:c.442G>T and c.601_602delAG variants in genomic DNA from the patient. **b** The RNA extracted from the patients' fibroblasts has only c.442G>T. Thus, Sanger sequencing of mRNA confirms that the variants reside in separate haplotypes and have been inherited as compound heterozygotes in all patients. Furthermore, the absence of the c.601_602delAG variant indicates that the transcript with the frameshift deletion is processed through nonsense mediated RNA decay

Fig. 11 Immunoblotting to detect NHLRC2 protein expression in human tissue homogenates from control autopsy samples. NHLRC2 was detected in all studied human organs including the heart (He), kidney (Ki), muscle (Mu), liver (Li), lung (Lu) and brain (Br). GAPDH was used as a loading control

Fig. 12 *Nhlrc2* is expressed in various cell types of the mouse brain and *NHLRC2* expression is highest in the early human life. **a** An RNA-Seq transcriptome and splicing database of the glia, neurons and vascular cells of the cerebral cortex show the highest *Nhlrc2* expression in astrocytes, oligodendrocyte progenitor cells, newly formed oligodendrocytes and endothelial cells. *Nhlrc2* was also expressed in neurons, myelinating oligodendrocytes and microglia. (OPG) oligodendrocyte progenitor cells [32]. Abbreviations: FPKM, fragments per kilobase per million. **b** The average expression levels of *NHLRC2* are similar between the cerebellar cortex (CBC), mediodorsal nucleus of the thalamus (MD), striatum (STR), amygdala (AMY), hippocampus (HIP) and 11 areas of the neocortex (NCX) in the human brain (http://hbatlas.org/). **c** The expression is highest in the early development of the human brain. The number of samples is indicated on the top of the x-axis. Data from www.brainspan.org [17]. Abbreviations: RPKM, reads per kilobase per million; pcw, post conception week; mos, months; yrs, years

Fig. 13 Wild-type (WT) and heterozygous (het) *Nhlrc2* KO morulae and their growth *in vitro*. The morulae were isolated at embryonic day (E) 2.5 and co-cultured in a micro drop culture overnight. On the next day, the E3.5 compacted morulae and early blastocysts were transferred to separate gelatinised wells. At E4.5, the embryos had matured to the blastocyst

stage and at E5.5 they hatched and attached to the bottom of the plate. Cells were grown for several days to detect possible phenotypes and to obtain enough material for genotyping (E7.5, E10.5 and E13.5). No homozygous morulae were detected

Fig. 14 *nhlrc2* splice site-blocking morpholino alters the mRNA spicing in zebrafish embryos but not the gross morphology of the embryos. Indicated amount of splice site targeting morpholino (together with an equal amount of p53 morpholino) was injected into the yolk sacs of fertilised eggs at the 1-2 cell stage. **a** Arrows on the left show the expected product sizes after the splice site-blocking events (334 bp intron inclusion, 198 bp WT product, 76 bp exon exclusion), and below the image are indicated the times of RNA extraction as days post fertilisation (dpf). **b** Un-injected controls on the left, *nhlrc2* splice site-blocking morpholino (3.6 ng) treated morphants in the middle and *nhlrc2* translation-blocking morpholino (2.3 ng) treated morphants on the right. The embryos appear to have WT morphology; only minor swelling was observed in the abdominal area in some individuals in both morphant groups at 3 dpf. However, this swelling diminished by 4 dpf. The images were taken with Axiovision software version 4.8 using a Zeiss SteREO Lumar V12 fluorescence microscope, equipped with 1.5X camera (Carl Zeiss MicroImaging GmbH, Göttingen, Germany)

Fig. 15 The timing of dechorionation of the *nhlrc2* morphants. The fraction of dechorionated embryos was recorded for selected hours post fertilisation (hpf). The significances were calculated with a log rank Mantel-Cox test. PBS control N=91,

untreated control N=47; p53 MO control N=23; untreated control N=36; random control MO 3.6 ng N=63; untreated control N=34 and splice site-blocking MO 3.6 ng N= 76, untreated control N=34. Graphs were generated with GraphPadPrism 5

Fig. 16 Knockdown of *nhlrc2* using splice site-blocking morpholino (nhlrc2 SB MO), translation-blocking morpholino (nhlrc2 TB MO), a combination (nhlrc2 SB+TB MO) of the two morpholinos compared to un-injected controls (Control) and random control morpholino (RC MO). **a** *nhlrc2* splice site blocking morpholino alters the mRNA splicing in zebrafish embryos. The indicated amount of splice site-targeting morpholino (together with an equal amount of p53 morpholino) was injected into the yolk sacs of fertilised eggs at the 1-2 cell stage. After the indicated time, RNA was isolated, PCR was performed and PCR product was run on 2% agarose gel. Arrows on the left annotate the expected product sizes after the splice site-blocking events (334 bp intron inclusion, 198 bp WT product, 76 bp exon exclusion), and below the image are the times of RNA extraction as days post fertilisation (dpf). **b** Knockdown efficiencies were calculated using ImageJ 1.49v. The levels of WT transcript are displayed above the bars for samples containing splice site-blocking morpholinos

Figure 1

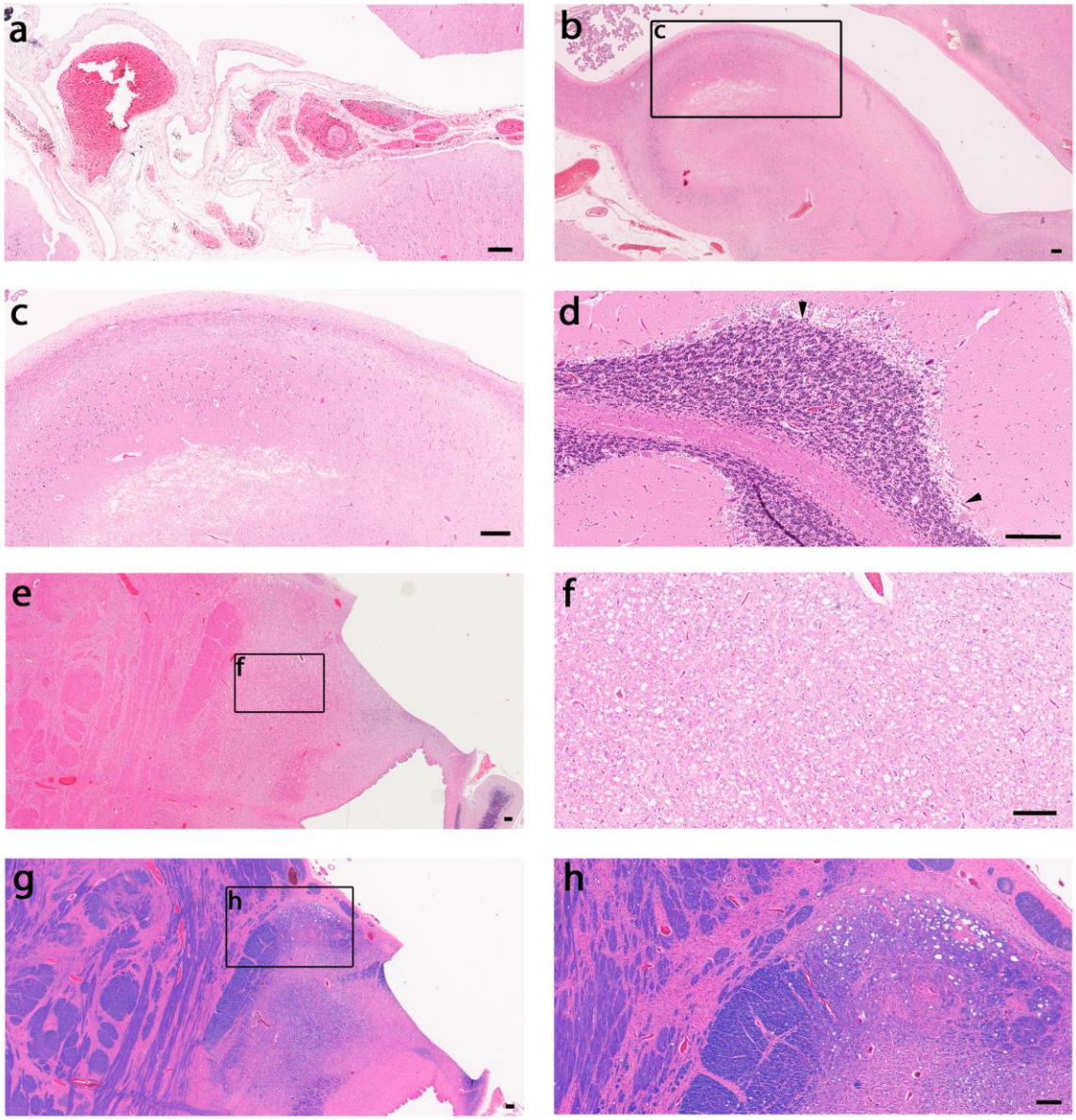


Figure 2

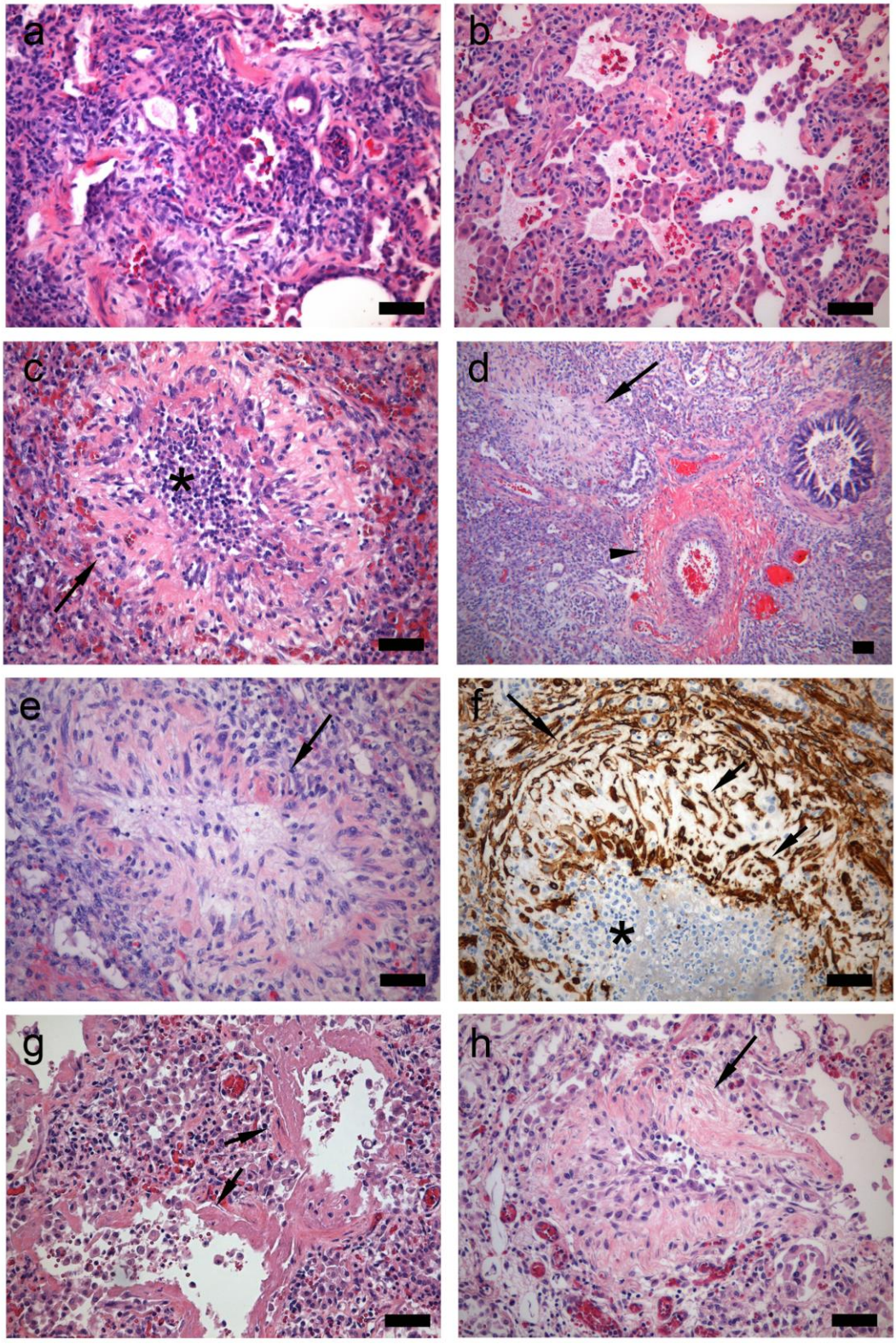


Figure 3

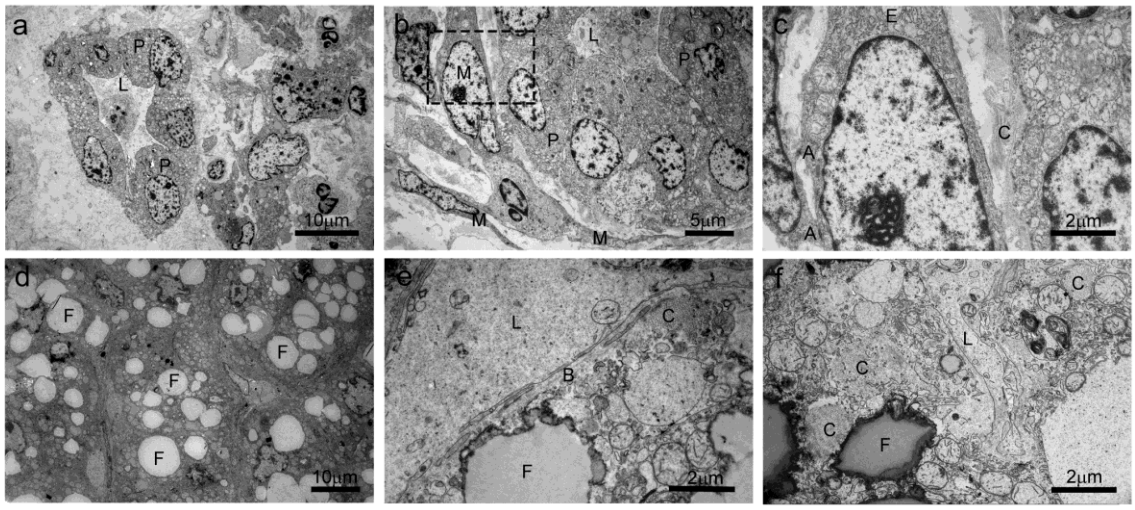

a
P
L
P
C
C
10μm
b
L
P
M
P
M
M
5μm
c
E
A
C
A
2μm
d
F
F
F
F
10μm
e
L
C
B
F
2μm
f
C
L
C
C
F
2μm

Figure 4

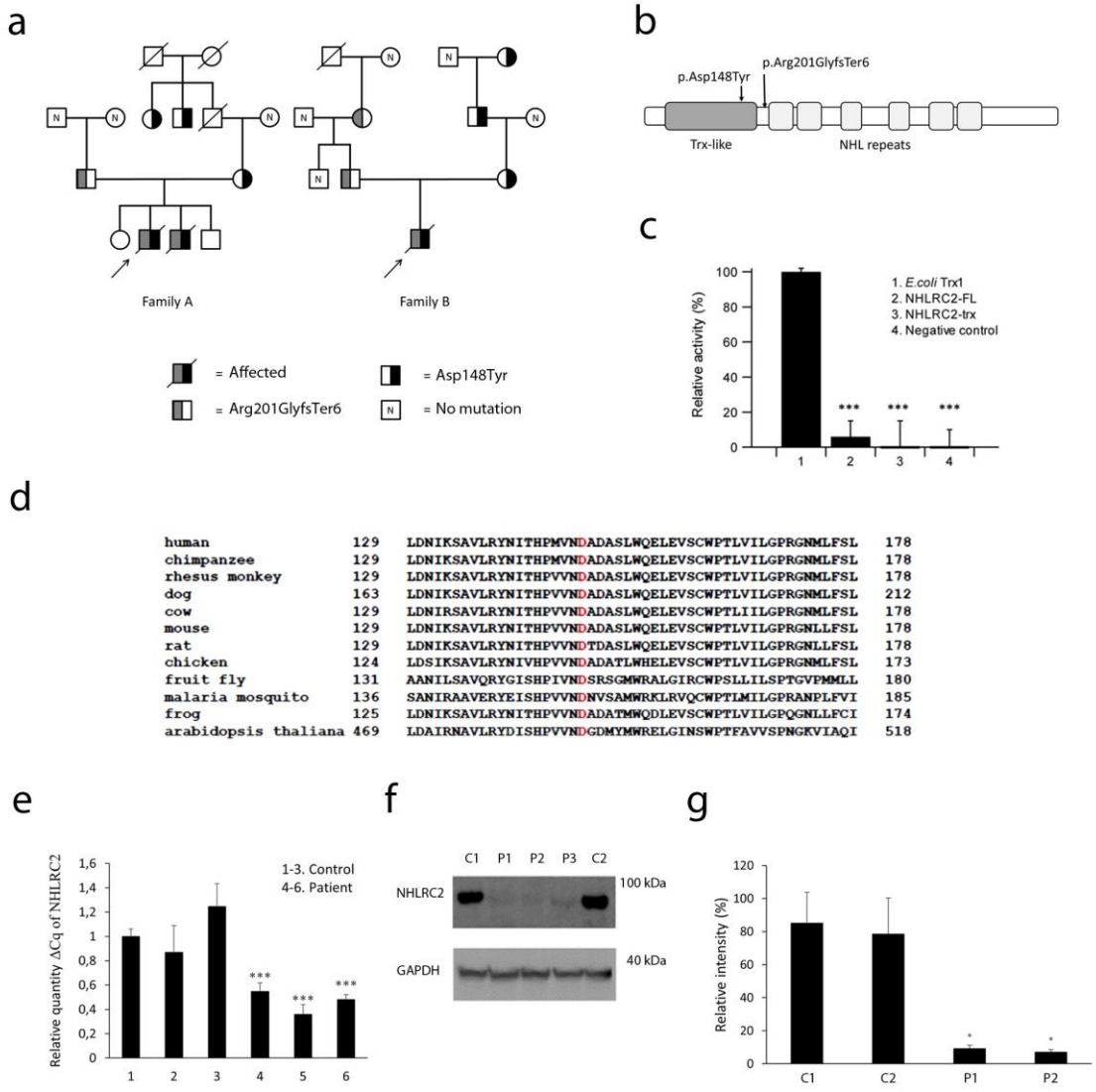

a
Family A
Family B
= Affected
= Arg201GlyfsTer6
= Asp148Tyr
= No mutation
b
p.Asp148Tyr
p.Arg201GlyfsTer6
Trx-like
NHL repeats
c
Relative activity (%)
1. E.coli Trx1
2. NHLRC2-FL
3. NHLRC2-trx
4. Negative control
d
human 129 LDNIKSAVLRYNITHPMVNDADASLWQELEVSCWPTLVILGPRGNMLFSL 178
chimpanzee 129 LDNIKSAVLRYNITHPMVNDADASLWQELEVSCWPTLVILGPRGNMLFSL 178
rhesus monkey 129 LDNIKSAVLRYNITHPVVNDADASLWQELEVSCWPTLVILGPRGNMLFSL 178
dog 163 LDNIKSAVLRYNITHPVVNDADASLWQELEVSCWPTLVILGPRGNMLFSL 212
cow 129 LDNIRSAVLRYNITHPVVNDADASLWQELEVSCWPTLIILGPRGNMLFSL 178
mouse 129 LDNIKSAVLRYNITHPVVNDADASLWQELEVSCWPTLVILGPRGNLLFSL 178
rat 129 LDNIKSAVLRYNITHPVVNDTDASLWQELEVSCWPTLVILGPRGNLLFSL 178
chicken 124 LDSIKSAVLRYNIVHPVVNDADATLWHELEVSCWPTLVILGPRGNMLFSL 173
fruit fly 131 AANILSAVQRYGISHPIVNDSRSGMWRALGIRCWPSLLILSPTGVPMMLL 180
malaria mosquito 136 SANIRAAVERYEISHPVVNDNVSAMWRKLRVQCWPTLMILGPRANPLFVI 185
frog 125 LDNIKSAVLRYNITHPVVNDADATMWQDLEVSCWPTLVILGPQGNLLFCI 174
arabidopsis thaliana 469 LDAIRNAVLRYDISHPVVNDGDMYMWRELGINSWPTFAVVSPNGKVIAQI 518
e
Relative quantity ΔCq of NHLRC2
1-3. Control
4-6. Patient
f
C1 P1 P2 P3 C2
NHLRC2
100 kDa
GAPDH
40 kDa
g
Relative intensity (%)
C1
C2
P1
P2

Figure 5

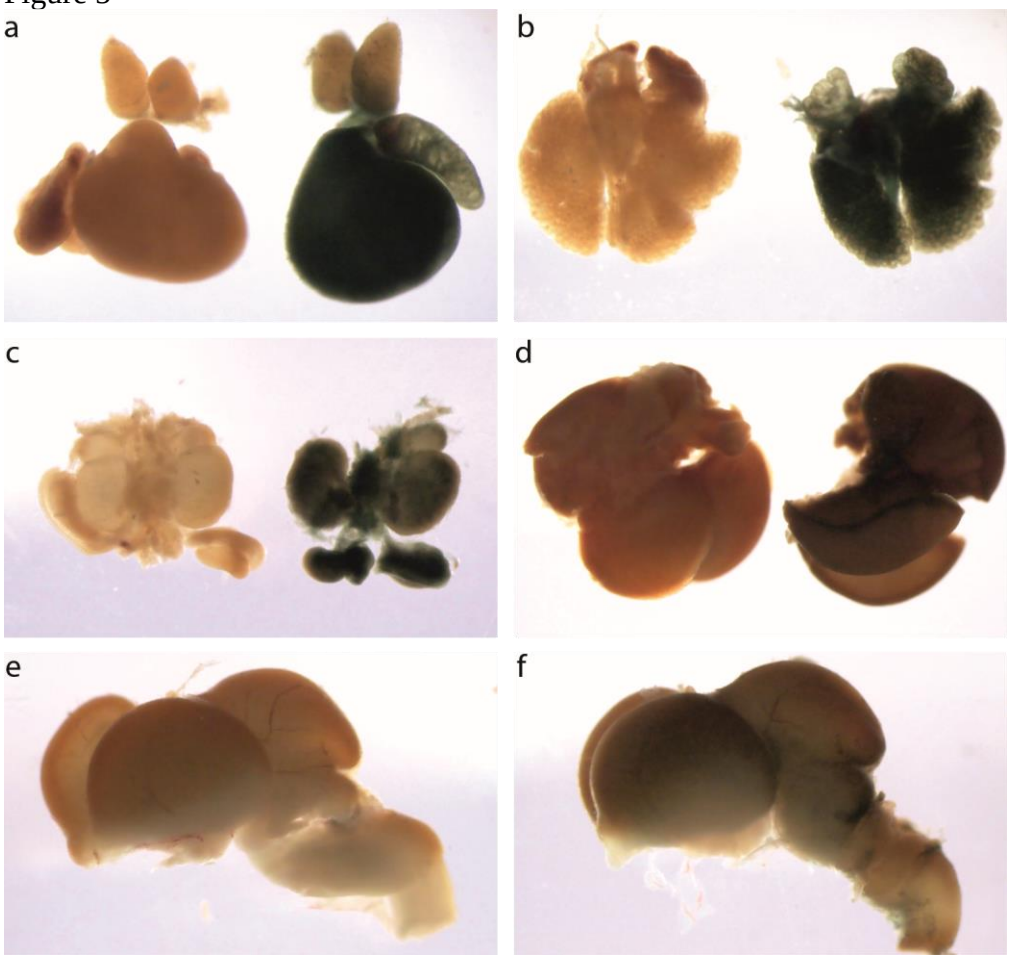
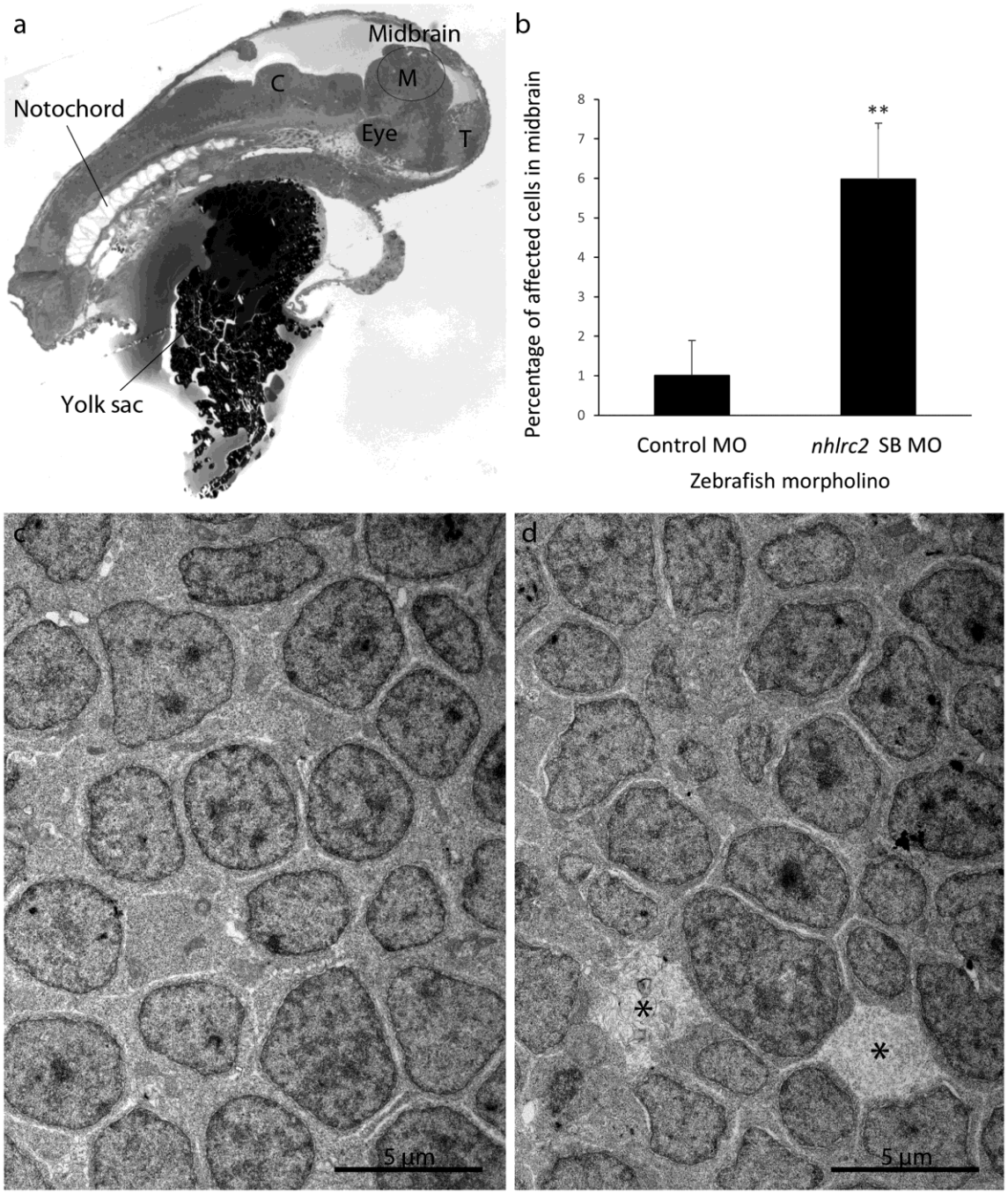


Figure 6


a
Midbrain
M
C
Notochord
Eye
T
Yolk sac
b
Percentage of affected cells in midbrain
0
1
2
3
4
5
6
7
8
**
Control MO
nhlrc2 SB MO
Zebrafish morpholino
c
5 µm
d
*
*
5 µm

Electronic supplementary material

NHLRC2 variants in patients with Fibrosis, Neurodegeneration, and Cerebral Angiomatosis (FINCA) – Characterisation of a novel cerebropulmonary disease

Johanna Uusimaa,[1–3,*] Riitta Kaarteenaho,[4,17] Teija Paakkola,[1,3,17] Hannu Tuominen,[5,6,18] Minna K. Karjalainen,[1,2,18] Javad Nadaf,[7,8,18] Teppo Varilo,[9,19] Meri Uusi-Mäkelä,[10,19] Maria Suo-Palosaari,[11] Ilkka Pietilä,[1,3] Anniina E. Hiltunen,[1,3] Lloyd Ruddock,[3,12] Heli Alanen,[3,12] Ekaterina Biterova,[3,12] Ilkka Miinalainen,[3] Annamari Salminen,[1,2] Raija Soininen,[3,12] Aki Manninen,[3,12] Raija Sormunen,[3,5] Mika Kaakinen,[3] Reetta Vuolteenaho,[3] Riitta Herva,[5] Päivi Vieira,[1,2] Teija Dunder,[1,2] Hannaleena Kokkonen,[13,14] Jukka S. Moilanen,[1,15] Heikki Rantala,[1,2] Lawrence M. Nogee,[16] Jacek Majewski,[7] Mika Rämet,[1,2,10] Mikko Hallman,[1,2] and Reetta Hinttala[1–3]

***Corresponding author:**
Johanna Uusimaa, MD, PhD
Professor of Paediatric Neurology
Tel: +358-8-3155819
Fax: +358-8-3155559
E-mail: johanna.uusimaa@oulu.fi

Case reports

Family A

Patient 1. Patient 1 was born at 38 weeks after a normal pregnancy and delivery to healthy, unrelated parents. The apgar score was 10/10, birth weight 3690 g, birth height 49 cm and head circumference 35 cm. He manifested with muscular hypotonia, dystonia and irritability at 2 months of age. He had delayed development and increased respiratory rate at the age of 3 months and alternating strabismus and loose faeces with transient diarrhoea at the age of 4 months.

When the patient was clinically examined at the age of 7 months, he was found to have dystonic tetraplegia, increased deep tendon reflexes, increased respiratory rate (48 breaths/min) with a bell-shaped chest but normal breath sounds and normal oxygen saturation values (SaO2) and poor visual contact. He also presented with unexpected unconsciousness and epileptic seizures. Laboratory investigations revealed haemolytic anaemia with decreased blood haemoglobin (75–92 g/l, reference 100–136 g/l) and blood haematocrit (0.21–0.26, reference 0.31–0.42), reticulocytosis (3.90–4.93%, reference 0.8–2.0%), increased plasma lactate dehydrogenase (1048–1698 U/l, reference 135–375 U/l) and plasma haemoglobin (129–414 mg/l, reference 0.29–2.00 g/l). The Coombs test was negative. Blood cell morphology analysis revealed anisocytosis, macrocytosis, poikilocytosis, polychromasia, hypochromia, microcytosis, ovalocytes and red cell fragments. Ophthalmological investigation revealed decreased vision in both eyes (Teller visus 20/800/1200 and 20/600/800, respectively).

From the age of 6 months, the child had breathing difficulties and recurrent respiratory and gastrointestinal symptoms but had negative screening results for common

viruses. Chest X-rays and high-resolution computed tomography (HRCT) at 7 months of age showed flattened hemidiaphragms and hyperlucent lungs, demonstrating over-inflated lungs, small air-filled bullae in the right mediobasal lung and perihilar small atelectasis (Fig. 7a).

During an acute metabolic crisis at the age of 9 months, a chest X-ray showed perihilar and basal airspace opacities and axial HRCT confirmed bilateral consolidations and reticular opacities (Fig. 7a). The patient's metabolic crisis included acute liver dysfunction with icterus and renal failure at the age of 9 months, which was triggered by viral gastroenteritis. During the subsequent 3-month period, he presented with hepatomegaly, increased liver transaminases and bilirubin, as well as icteric sclerae and skin. Extensive metabolic screening was negative, except for the following abnormal liver and kidney function tests taken during the acute metabolic crisis: increased plasma aspartyl aminotransferase (152–218 U/l, reference 10–50 U/l) and alanine aminotransferase (84–86 U/l, reference 10–50 U/l), slightly increased serum bilirubin (51–92 μmol/l, reference ≤25 μmol/l), conjugated bilirubin (62 μmol/l, reference ≤5 μmol/l), increased plasma gamma-glutamyltransferase (124–147 U/l, reference 5–80 U/l) and plasma ammonia (NH_4) (90–110 μmol/l, reference <80 μmol/l), decreased serum prealbumin (0.09–0.15 g/l, reference 0.24–0.42 g/l) and plasma albumin (27 g/l, reference 36–45 g/l), and increased plasma creatine (111–139 μmol/l, reference 15–35 μmol/l) and plasma urea (14.9–27.5 mmol/l, reference 1.7–8.3 mmol/l). Beta-hydroxybutyric acid and octandioic acid were present in urine during the metabolic crisis.

An abdominal ultrasound revealed hepatomegaly and increased cortical echogenity with a strand-like hypoechogenic outer zone in the kidneys (Fig. 9a). A nasogastric tube

was implanted when the boy was 1 year old due to feeding problems, poor weight gain and progressive growth retardation (retardation of growth from 0 standard deviations [SD] to -2 SD by the age of 7 months, decrease in relative weight from +12% to -20% and decrease in head circumference from -0.5 SD to -1.2 SD).

Brain magnetic resonance imaging (MRI) at the age of 7 months revealed a thin corpus callosum (Fig. 8a). Brain auditory evoked potentials and electroretinogram were normal, but visual evoked potentials showed increased latencies and giant responses that resembled the findings of a muscle–eye–brain (MEB) disease and late-infantile neuronal ceroid lipofuscinosis (Jansky–Bielschowsky). Therefore, a rectal biopsy was also performed but did not reveal any vacuolar cells. Electroencephalogram (EEG) revealed slow background activity and frontal bilateral rhythmic high-amplitude sharp delta transients, indicating metabolic encephalopathy.

A second episode of acute liver dysfunction at the age of 1 year 8 months was triggered by viral gastroenteritis and fever. The disease course was progressive, and the patient died at the age of 1 year 9 months due to respiratory failure.

Patient 2. Patient 2 was the younger brother of Patient 1. He was born at 38 + 6 weeks with an Apgar score of 9/9, birth weight 3280 g, height 51 cm and head circumference 34.5 cm. Similar to his older brother, feeding problems and tachypnoea were noticed at the age of 2 months.

At the age of 4 months, the infant manifested with right-sided chest prominence, a relative weight of -32% and muscular hypotonia. In contrast to his older brother, the infant was able to make eye contact and responded with a smile. However, during the follow-up,

he presented with poor weight gain, developmental delay and gastrointestinal symptoms (diarrhoea) by the age of 5 months. At that time, breath sounds and SaO2 were normal.

A chest X-ray at 4 months demonstrated mediastinal and hilar prominence, over-inflated lungs and diffuse reticular interstitial opacities (Fig. 7b). Also at 4 months, HRCT revealed widespread bilateral ground-glass opacities, interstitial infiltrations and atelectasis, mildly enlarged thymus, and left hilar adenopathy (Fig. 7b and 9f). A brain MRI at 10 months of age showed a thin corpus callosum and slightly dilated lateral ventricles and cortical sulci (Fig. 8b). At 1 year of age, chest X-rays showed the progression of interstitial infiltrations and bilateral dense consolidations (Fig. 7b). The child required oxygen support during respiratory failure triggered by a viral infection. An MRI of the liver revealed an abnormal signal (Fig. 9b-d), but the liver ultrasound was normal (Fig. 9e). A nasogastric tube was implanted at the age of 1 year due to feeding problems and progressive growth retardation that led to growth arrest. This patient did not present with seizures, and electrophysiological examinations were not performed.

Laboratory investigations revealed chronic haemolytic anaemia: decreased plasma haemoglobin (83–95 g/l, reference 100–136 g/l) and blood haematocrit (0.24–0.26, reference 0.31–0.42), reticulocytosis (3.46–3.99%, reference 0.8–2.0%), increased plasma lactate dehydrogenase (404–707 U/l, reference 150–450 U/l) and a negative Coombs test. The blood cell morphology analysis revealed anisocytosis, microcytosis, macrocytosis and polychromasia. Plasma alanine aminotransferase was elevated (184–257 U/l, reference 10–50 U/l), plasma aspartyl aminotransferase was within normal limits (48 U/l, reference ≤ 45 U/l) and both serum bilirubin (31–36 μmol/l, reference 5–25 μmol/l) and conjugated bilirubin (9 μmol/l, reference ≤5μmol/l) were slightly increased. Serum prealbumin (0.24–

0.27 g/l, reference 0.24–0.42 g/l) and plasma haemoglobin (63 mg/l, reference <100 mg/l) were within normal limits while serum haptoglobin (<0.08 g/l, reference 0.29–2.00 g/l) was at a lower level. Plasma ammonia (NH_4) was normal (37 μmol/l, reference <80 μmol/l). Furthermore, transient slightly elevated blood lactate (1.61 mmol/l, reference 0.33–1.33 mmol/l) and slightly elevated blood pyruvate (105 μmol/l, reference 30–80 μmol/l) were noted, but the lactate/pyruvate ratio was normal (15, reference <25). Serum carcinoembryonic antigen (S-CEA) was elevated at 5 months (13.2–22.1 μg/l, reference <3 μg/l), and serum neuron-specific enolase (S-NSE) levels were slightly elevated at the age of 6 months (20.1–20.4 μg/l, reference <12.5 μg/l). Because of the recurrent infections and granulomatosis that occurred in the patient's older brother, a nitroblue tetrazolium (NBT) assay was performed along with other immunological tests. NBT showed normal oxidative activity in neutrophils (95.5% versus reference 99.1%) but decreased activity in the monocytes (57% versus reference 90.9%). Urinary organic acid analysis revealed that ethylmalonic acid levels were slightly elevated.

At the age of 1 year, the boy's subcutis was thin, muscle mass was decreased, and muscle tone varied from hypotonia to hypertonia and dystonia. He presented with visual problems, alternating strabismus, and variable eye contact. The disease course was progressive, and the patient died due to respiratory failure at the age of 1 year 1 month after treatment at the intensive care unit (ICU) for 7 weeks.

Family B

Patient 3 was born at 36 + 1 gestational weeks by emergency caesarean section because of a failing heart rate. He was the first child. His mother had undergone ursodeoxycholic acid treatment because of gestational hepatosis and toxaemia. The infant's birth weight was

2910 g, and the Apgar score was 10/10. During the first week of life, he was treated twice for newborn jaundice with blue light therapy for a total duration of 4 days. ABO immunisation was suspected but not proven (serum bilirubin had increased up to 349 μmol/l at 8 days).

At 2 months of age, the patient presented with irritability, floppiness and feeding problems. Laboratory tests revealed haemolytic anaemia, with decreased plasma haemoglobin (69–84 g/l, reference 100–136 g/l) and plasma haematocrit (0.19–0.23, reference 0.31–0.42), reticulocytosis (4.0–5.1%, reference 0.8–2.0%), increased plasma lactate dehydrogenase (404–643 U/l, reference 150–450 U/l), undetectable plasma haptoglobin levels and a negative Coombs test. Blood cell morphology analysis revealed polychromasia, anisocytosis, macrocytosis, poikilocytosis and red cell fragments. The patient had frequent diarrhoea and poor weight gain. A milk allergy was suspected because of the constant diarrhoea, but changes in the formula only slightly improved his diarrhoea.

At 4 months, the infant presented with a suspected seizure during an influenza B infection, and he was hospitalised because of poor general condition, breathing problems (gasping for air), and low Sa02 values (70%). Chest X-ray at 4 months of age revealed over-inflated lungs and bilateral perihilar interstitial opacities (Fig. 7c).

At 13 months of age, chest X-rays showed diffuse infiltrates with reticulation (Fig. 7c). Thoracic HRCT confirmed diffuse bilateral ground-glass opacification, peripheral interstitial septal thickening and lobular pleural thickening (Fig. 7c). EEG was normal during wakefulness but unusually monotonic during sleep, during which it consisted of 4-Hertz delta waves without normal sleep spindles and vertex waves.

The child continued to be irritable with failure to thrive and, at age 7.5 months, he was referred to a neurologist because of slow development. Physiotherapy was started, and he was referred to an ophthalmologist because strabismus was noticed. His development did not progress.

At the age of 8 months, the child suffered almost daily from tachypnoea and cough. However, by auscultation, lung sounds were normal. At 10 months of age, he followed with his eyes but did not actively grasp toys or turn onto his stomach. He had brisk deep tendon reflexes and an indifferent plantar reflex. His breathing was tenuous, and tachypnoea was noted. The size of the child's liver was at the upper limit. Brain MRI showed slightly enlarged lateral ventricles and cortical sulci, as well as a thin corpus callosum. T2-weighted axial MRI revealed a mildly increased signal intensity of the globus pallidi (Fig. 8c).

The child underwent gastrostomy because of weight loss. Blood lactate was intermittently elevated (1.685–3.93 mmol/l, reference 0.33–1.33 mmol/l), and the plasma lactate/pyruvate ratio was at an increased level (35, reference <25). Serum thyroid-stimulating hormone (TSH) was slightly decreased (0.5–0.87 kU/l, reference 1.3–8.5 kU/l), but serum thyroid hormone levels (free T4 and T3) were normal. Immunological tests revealed that serum immunoglobulin G was slightly diminished (1.1–1.6 g/l, reference 2.7–10.1 g/l). In one instance, urinary organic acid analysis indicated small amounts of dicarboxylic acids without ketone bodies. Otherwise the extensive metabolic screening was normal. The child's breathing difficulties became progressively worse, and he died due to respiratory failure at 1 year 2 months of age after being treated at the ICU for 5 weeks.

Copy number variants and large regions of homozygosity were studied by HumanCytoSNP-12 (v2.1) (Illumina, San Diego, California, USA). The results were normal and no pathogenic variation was detected in Patient 3.

Methods

Histopathological studies of tissue samples

Tissue biopsies (lung and skeletal muscle biopsies) and autopsy samples were obtained from all three patients. The tissue was fixed in buffered 4% formaldehyde, routinely processed into paraffin blocks and cut into 5.0 μm sections. For immunohistochemistry, the sections were deparaffinised in xylene, rehydrated through graded ethanol, and stained with the following antibodies from Dako/Agilent Technologies (Santa Clara, CA, USA): CD68 (clone PG-M1, 1:100 dilution, Dako code M0876), Ck-PAN (Clone MNF-116, 1:700 dilution, Dako code M0821), caldesmon (clone h-CD, dilution 1:1000, Dako code M3557), alpha-smooth muscle actin (clone 1A4, dilution 1:500, Dako code M0851), and smooth muscle myosin heavy chain (SMMHC) (clone SMMS-1, dilution 1:100, Dako code M3558).

DNA extraction

Total genomic DNA from the blood samples of the patients and their family members, as well as from the patients' fibroblasts, was extracted with the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). Saliva samples were collected from the family members of the affected children using Oragene DNA Collection kits (DNA Genotek, Ottawa, Canada) and the prepIT L2P kit (DNA Genotek, Ottawa, Canada) was used for DNA extraction. For

the Northern Finnish population controls, DNA was extracted from umbilical cord blood with the UltraClean DNA Blood Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA). For a subset of the samples, whole-genome amplification was performed, as previously described [4].

Whole-exome sequencing and bioinformatics

Next-generation sequencing library preparation, enrichment, and sequencing were performed using the Nimblegen V2 Exome (http://www.nimblegen.com/products/seqcap/ez/v2/) at the Technology Centre of the Institute for Molecular Medicine Finland (FIMM), University of Helsinki, as previously described [13]. DNA from the affected patients and their parents was also analysed by exome sequencing. Bioinformatics analysis was performed, as previously described [16]. Briefly, the Burrows-Wheeler Aligner (BWA v. 0.5.9) was used to map all reads to UCSC hg19. PCR duplicates were removed from alignments using Picard version 1.87 (http://broadinstitute.github.io/picard/). Indels were realigned with the Genome Analysis Toolkit (GATK 1.0.5083). SAMtools version 0.1.18 was used to call the variants. All variants were annotated with ANNOVAR and in-house scripts, and the most likely protein damage variants (nonsense, splice-site, frame shift, indel and missense) were considered for further analysis. The variants were also annotated for allele frequency in public databases such as Exome Aggregation Consortium (ExAC) (http://exac.broadinstitute.org/); the National Heart, Lung, and Blood Institute (NHLBI) exomes (V.0.0.14, http://evs.gs.washington.edu/EVS/); the 1000 Genomes Project database (http://www.internationalgenome.org) and our in-house database of >1,000

exomes. To remove common variants and sequencing artefacts, variants with minor allele frequencies of >0.01 in any of the aforementioned databases were removed from further analysis. The only candidate gene that was common to both families was *NHLRC2*.

Variant analysis and genealogy of the two families

Using Sanger sequencing, we confirmed variants and segregation within the families. The region encompassing the two *NHLRC2* variants was amplified by PCR with the primers 5′-TGG TGT TCA CTC GGC TAA GT-3′ and 5′-GTC CAT TCT TCC AAA CGA CCA-3′ and then sequenced. To investigate the descent of the variants, we performed a genealogical study in accordance with published criteria [14]. We traced the ancestors of the patients from the Finnish Population Registries and scrutinised the microfiche copies available in the National Archives of Finland.

Fibroblast cultures

Fibroblasts from healthy individuals and patients were grown at 37°C in an atmosphere of 5% CO_2 in high-glucose Dulbecco's Modified Eagle's Medium (DMEM) (Corning, Manassas, VA, USA) supplemented with 10% foetal bovine serum. The patient and control cell lines were transduced with a retroviral vector expressing the E7 gene of type-16 human papilloma virus and a retroviral vector expressing the protein component of human telomerase to immortalise the cells [17].

RNA extraction from fibroblasts

We used trypsin (Lonza, Verviers, Belgium) to harvest 1×10^7 cultured fibroblasts from

healthy individuals and patients. Total RNA was extracted with the RNeasy Plus Mini kit (Qiagen, Hilden, Germany). After reverse-transcription PCR (RT-PCR), we used Sanger sequencing to investigate the expression of the mutant alleles. The OneStep RT-PCR kit (Qiagen, Hilden, Germany) was used for the patients' cDNA analysis by qPCR.

qPCR to measure the expression of NHLRC2 mRNA

qPCR was performed according to the manufacturer's instructions (IQ™ SYBR Green Supermix, Bio-Rad, Hercules, CA, USA) using a CFX Connect™ Real-Time System (Bio-Rad, Hercules, CA, USA). Melting temperature (T_m) of the analysis was adjusted to 61°C, and TATA-Box Binding Protein (*TBP*) and Glucuronidase Beta (*GUSB*) were chosen as endogenous control genes. Primers 5´-CAGACCTCGCTAGAGTACGC-3´ and 5´-GTTCAGCCATTCTAATCCTTCCG-3´ were used for *NHLRC2*, 5´-GTCTGCGGCATTTTGTCGG-3´ and 5´-CACACGATGGCATAGGAATGG-3´ for *GUSB* and 5´-CCACTCACAGACTCTCACAAC-3´ and 5´-CTGCGGTACAATCCCAGAACT-3´ for *TBP* (Primer bank IDs 304307770c1, 268834191c1 and 285026518c1) [11,12,15].

Immunoblotting

Protein expression in the whole-cell extracts of the fibroblasts from all three patients and healthy control subjects, together with the homogenates of control autopsy samples from several types of human tissues, were analysed by immunoblotting. Whole-protein extracts from the fibroblasts were prepared in 1.5% n-dodecyl maltoside (Sigma-Aldrich, St. Louis, MO, USA) in phosphate-buffered saline (PBS). Proteins from the whole-cell extracts and

tissue homogenates were precipitated with ethanol and a total of 20–100 μg of protein per sample was used for Tris-glycine SDS-PAGE (4–20% Precise Protein Gel; ThermoFisher Scientific, Waltham, MA, USA). Proteins were transferred to a nitrocellulose membrane (Amersham Hybond ECL; GE Healthcare, Buckinghamshire, UK). Mouse polyclonal antibodies against full-length human NHLRC2 (SAB1400871, Sigma-Aldrich, St. Louis, MO, USA; NBP1-85019, Minneapolis, MN, USA; ab88725, Abcam, Cambridge, UK) and against GAPDH (GeneTex, Irvine, CA, USA) were used, together with secondary antibodies (ThermoFisher Scientific, Waltham, MA, USA) to visualize the proteins by enhanced chemiluminescence (ECL) (GE Healthcare, Buckinghamshire, UK). Signal intensities were quantified (ImageJ) and the level of NHLRC2 was compared to the level of GAPDH [10].

Expression of human NHLRC2 in tissue

Autopsy tissue specimens from human heart, kidney, skeletal muscle, liver, lung and brain were homogenised to examine the overall expression of NHLRC2 in tissues. Tissue (5 mg) was dissected on ice, 'snap frozen' in liquid nitrogen and homogenised in 300 μl of ice-cold 20 mM Tris-HCl (pH 7.5) with Pierce protease inhibitors (ThermoFisher Scientific, Waltham, MA, USA). Homogenates were maintained with constant agitation for 2 hours at 4°C and centrifuged for 20 minutes at 13,500 × *g* at 4°C in a microcentrifuge (Eppendorf, Hamburg, Germany). Protein concentrations were measured from supernatants with the Coomassie Protein Assay kit (ThermoFisher Scientific, Waltham, MA, USA). Due to rapid autolysis of the lung and brain samples, the amount of total protein obtained from these autopsy specimens was low and required precipitation with ethanol prior to

immunoblotting.

Cloning of human *NHLRC2* in *E. coli*

A gene that encodes full-length human *NHLRC2* and is a codon optimised for *E. coli* expression was purchased from GenScript (Piscataway, NJ, USA). The domain organisation of human *NHLRC2* (M1-F726) is defined in InterPro [6] as having an N-terminal thioredoxin-like domain (Q43-D200) and a six bladed beta-propeller (G210-P592). Based on this information plus secondary structure prediction [1], structural homology searches and sequence analysis, we predicted that the protein contained both domains, with the thioredoxin-like domain having an α-helical N-terminal extension and ending around G202 and the beta-propeller domain starting around P222. The region between the two could belong to either or neither domain. Accordingly, we made a series of domain constructs to express the full-length protein and fragments in *E.coli*. The DNA fragments encoding these constructs were amplified by PCR with primers that were designed to include NdeI and BamHI restriction sites on the 5′ and 3′ ends, respectively. The PCR fragments were digested with the corresponding enzymes and ligated into a modified variant of pET23 with the T7 promoter replaced by Ptac [2]. All plasmids generated were sequenced to ensure there were no errors in the cloned genes.

Expression and purification of full-length human NHLRC2 protein and thioredoxin-like domain of NHLRC2

Sixteen *NHLRC2* constructs were expressed in *E. coli* strain BL21 (DE3) with EnPresso®B medium (BioSilta, Cambridgeshire, UK) in accordance with the manufacturer's

instructions. From the analysis of the constructs made, the predicted domain boundaries were correct, and the linker region was better behaved when part of the thioredoxin domain than when only part of the beta-propeller domain. Accordingly, two constructs were used for larger scale protein production and analysis. The first construct contained the gene encoding the full-length human NHLRC2 (M1-F726) protein while the second contained DNA corresponding to the NHLRC2 thioredoxin-like domain (M1-S221). Briefly, BL21 (DE3) cells were transformed with the appropriate construct and plated onto Luria-Bertani (LB) agar plates containing 100 μg/ml ampicillin. Plates were incubated overnight at 37°C. The next morning, several colonies were transferred into 5 ml of LB medium supplemented with 100 μg/ml ampicillin. Cultures were grown at 30°C with shaking at 200 rounds per minute (rpm) for 6 hours. Five millilitre aliquots of preculture were used to seed 500 ml of EnPresso®B with 100 μg/ml ampicillin in 5 L Erlenmeyer flasks. Cultures were incubated at 30°C with shaking at 250 rpm for approximately 16 hours.

Protein production was induced by the addition of 250 μM Isopropyl β-D-1-thiogalactopyranoside (IPTG), and booster media was added. Cells were harvested by centrifugation (3,220 x g, 20 min) after 24 hours of incubation at 30°C. Cells were resuspended in a lysis buffer (50 mM sodium phosphate [pH 7.4], 500 mM NaCl, 5% glycerol, and 10 mM imidazole) supplemented with PierceTM Protease Inhibitors EDTA-free (ThermoFisher Scientific, Waltham, MA, USA), 0.1 mg/ml lysozyme and 20 μg/ml DNAse. Cells were lysed using three cycles of freeze-thawing. Cell debris was removed by centrifugation, and the soluble fraction was loaded onto a 5 ml HiTrap HP column (GE Healthcare, Buckinghamshire, UK). Specifically, bound proteins were eluted from the column with increasing concentrations of imidazole. All the collected fractions were

analysed by SDS-PAGE. Fractions containing the protein of interest were pooled together and dialysed against 20 mM Tris (pH 8.0), 25 mM NaCl and 2 mM Dithiothreitol (DTT). Dialysed protein solution was loaded into a Q HiTrap (GE Healthcare, Buckinghamshire, UK), and the protein of interest was eluted with increasing concentrations of NaCl. Full-length NHLRC2 was additionally purified by size-exclusion chromatography on a Superdex 200 column (GE Healthcare, Buckinghamshire, UK). Samples containing full-length NHLRC2 or the thioredoxin-like domain were concentrated to 9.2 and 7.5 mg/ml, respectively, flash-frozen in liquid nitrogen, and stored at -70°C.

Insulin assay

The insulin turbidity assay to test for thioredoxin activity was performed, as previously described [3]. Briefly, the assay was done in 200 µl of 0.1 M sodium phosphate buffer (pH 7.0) with 1 mM EDTA and 10 mM DTT, as well as 1µM enzymes when appropriate. The reaction was started by the addition of 1 mg/ml insulin and monitored at 540 nm for 30 minutes. Measurements were taken every 30 seconds, and the lag phase for precipitation was monitored. *E. coli* thioredoxin was used as a positive control in the assay and was produced as previously described [8].

Culturing of *Nhlrc2*^{tm1a} mouse morulae

The mating of heterozygous female and male mice was confirmed by the presence of a vaginal plug. On embryonic day 2.5 (E2.5), morulae were flushed from the uterus in accordance with a previously described protocol [7]. The uterus was cut near the cervix, and a 27 G needle attached to a 3 ml syringe was inserted into the oviduct. Both uterine

horns were flushed with PrimeQ^{TM} M2 Mouse Embryo Culture Medium (Amsbio, Abingdon, UK). The morulae were transferred into a micro-culture drop of PrimeQ^{TM} KSOM Mouse Embryo Medium (Amsbio, Abingdon, UK), covered with embryo-tested light mineral oil (Merck-Sigma, Temecula, CA, USA) and equilibrated in an organ culture incubator (37°C, 5% CO_2). Embryos grown overnight in the microculture drop were transferred onto a gelatinized 24-well plate, each into its own well with 300 µl of embryonic stem cell medium, and left to grow for 10–12 days, after which DNA was extracted, as described [9].

Genotyping of *Nhlrc2^{tm1a}* knockout mouse offspring and embryos

Primers CAS_R1_Term, Nhlrc2_119354_F, Nhlrc2_119354_R, LacZ_probe_F, and LacZ_probe_R (https://www.infrafrontier.eu/search?keyword= Nhlrc2) were used with Phire Hot Start II Polymerase (ThermoFisher Scientific, Waltham, MA, USA) for genotyping.

X-gal staining

To demonstrate LacZ expression in the embryos, E14.5 *Nhlrc2^{tm1a}* heterozygote mutant embryos were dissected in ice-cold PBS and fixed at room temperature for 30 min in a solution containing 0.2% glutaraldehyde, 2 mM $MgCl_2$, 5 mM EGTA and 0.1 M kanolinite phenylphosphonate (KPP) (pH 7.3) in water. They were then washed three times for 10 min in 0.02% Nonidet P-40, 0.01% Na-deoxycholate, 5 mM EGTA, 2 mM $MgCl_2$ and 0.1 M KPP pH 7.3 in water and stained by incubating in a 5-bromo-4-chloro- 3-indolyl β-D-galactopyranoside staining solution (X-gal, #R0402; ThermoFisher, Waltham, MA, USA)

composed of 10 mM $K_3Fe(CN)_6$, 10 mM $K_4Fe(CN)_6$ and 1 mg/ml X-gal in DMSO at room temperature overnight. The embryos were thereafter fixed with 4% paraformaldehyde.

Generation of zebrafish morphants

Two morpholinos against the zebrafish *nhlrc2* (ENSDARG00000089581) were designed; the first morpholino was against the ATG-site to prevent translation, and the second one was against the exon-intron boundary at the 3' end of exon 4 (ENSDARE00000919598) to prevent splicing, which both would lead to a frameshift and premature termination of translation. The regions were sequenced for polymorphism before morpholino design. Morpholino sequences were 5'-ACCCGATTCTGCTGATTTACCTTTC-3' (splice site blocking) and 5'-CTGTCAGCTTACTGTAAGACGCCAT-3' (translation blocking) (GeneTools, LLC, Philomath, OR, USA).

The morpholinos were heated to 65ºC before use to achieve complete suspension. Then, 1 nl of each morpholino (1.0– 5.0 ng) in a solution containing PBS and 0.2% rhodamine B (Sigma-Aldrich, St. Louis, MO, USA) was injected into a 1-2 cell stage embryo using a microinjector (PV830 Pneumatic PicoPump, World Precision Instruments, Sarasota, FL, USA). The morphant embryos were screened for fluorescence (570/590 nm excitation/emission) 4 hours post fertilization (hpf), and images were taken at 24, 48, 72 and 96 hpf with a fluorescence microscope (Zeiss Lumar V12, Carl Zeiss, Oberkochen, Germany). A standard random control morpholino and a p53 morpholino (GeneTools, LLC, Philomath, OR, USA) were used to control the non-specific effects of the morpholinos. WT AB zebrafish embryos were maintained according to standard protocols.

Quantitation of knockdown

The effect of the splice site-blocking morpholino for *nhlrc2* levels was quantified using PCR. Primers were designed on exon 3 and exon 5 to detect the loss of exon 4 (forward primer 5'-GTCTCCAATACTGGGCAGGTG-3' and reverse primer 5'-GCCTGCTAATGTTGAGACTTTTCC-3'). Because an effective antibody against zebrafish Nhlrc2 was not available, the effect of the translation-blocking morpholino could not be evaluated.

RNA extractions were done from four to seven pooled embryos using Tri reagent (Molecular research centre, OH, USA), and 70-300 ng of the RNA was used for cDNA synthesis using a SensiFAST cDNA synthesis kit (Bioline, Taunton, MA, USA) according to the manufacturer's instructions. The obtained RNA was treated with DNase I (Thermo Scientific, Waltham, MA, USA) according to the manufacturer's instructions. Then, 70-300 ng DNase-free RNA was used for cDNA synthesis using SensiFAST cDNA synthesis kit (Bioline, Taunton, MA, USA) according to the manufacturer's instructions. The cycling conditions for PCR (SsoFast Evagreen Supermix, Bio-Rad, Hercules, CA, USA) were 95ºC 3:00, 95ºC 0:02, 59ºC 0:05 and 59ºC for 0:10, and a melt curve with an increment of 0.5ºC to 95ºC. 5 µl of the product was run on 2% agarose gel with 1:10 000 SybrSafe DNA gel stain (Thermo Fischer Scientific, Waltham, MA, USA). The knockdown efficiency was determined from an agarose gel using ImageJ v1.49 by calculating the ratio of the WT, unmodified band to the sum of all the bands on a single lane (WT and two morpholino modified bands corresponding to exon exclusion and intron inclusion caused by splice site blocking).

Timing of Dechorionation

Dechorionation of the embryos was recorded every 3 hours between 36 hpf and 53 hpf and then at 60 hpf and 70 hpf.

Supplementary figures

Fig. 7 Radiological findings of chest X-rays and high-resolution computed tomography of Patients 1–3. **a** Patient 1 at 7 months of age shows flattened hemidiaphragms and hyperlucent lungs demonstrating over-inflated lungs, small air-filled bullae of the right mediobasal lung (short arrows) and perihilar small atelectasis (long arrow). Performed during an acute metabolic crisis at the age of 9 months, the chest X-ray demonstrates perihilar and basal airspace opacities and axial HRCT confirms bilateral consolidations (long arrows) and reticular opacities (short arrows). **b** Patient 2 at 4 months of age demonstrates mediastinal and hilar prominence (short arrows), over-inflated lungs, and diffuse reticular interstitial opacities. At 4 months of age, HRCT reveals widespread bilateral ground-glass opacities (short arrow), interstitial infiltrations (long arrows) and atelectasis (black arrowhead), mildly enlarged thymus (white arrowhead) and left hilar adenopathy (asterisk). At 1 year of age, chest X-rays show the progression of interstitial infiltrations and bilateral dense consolidations (long arrows). **c** Patient 3 at 4 months of age reveals over-inflated lungs and bilateral perihilar interstitial opacities (short arrows). At 13 months of age, the chest X-ray shows diffuse infiltrations with reticulation. HRCT confirms diffuse bilateral ground-glass opacification, peripheral interstitial septal thickening and lobular pleural thickening (long arrows)

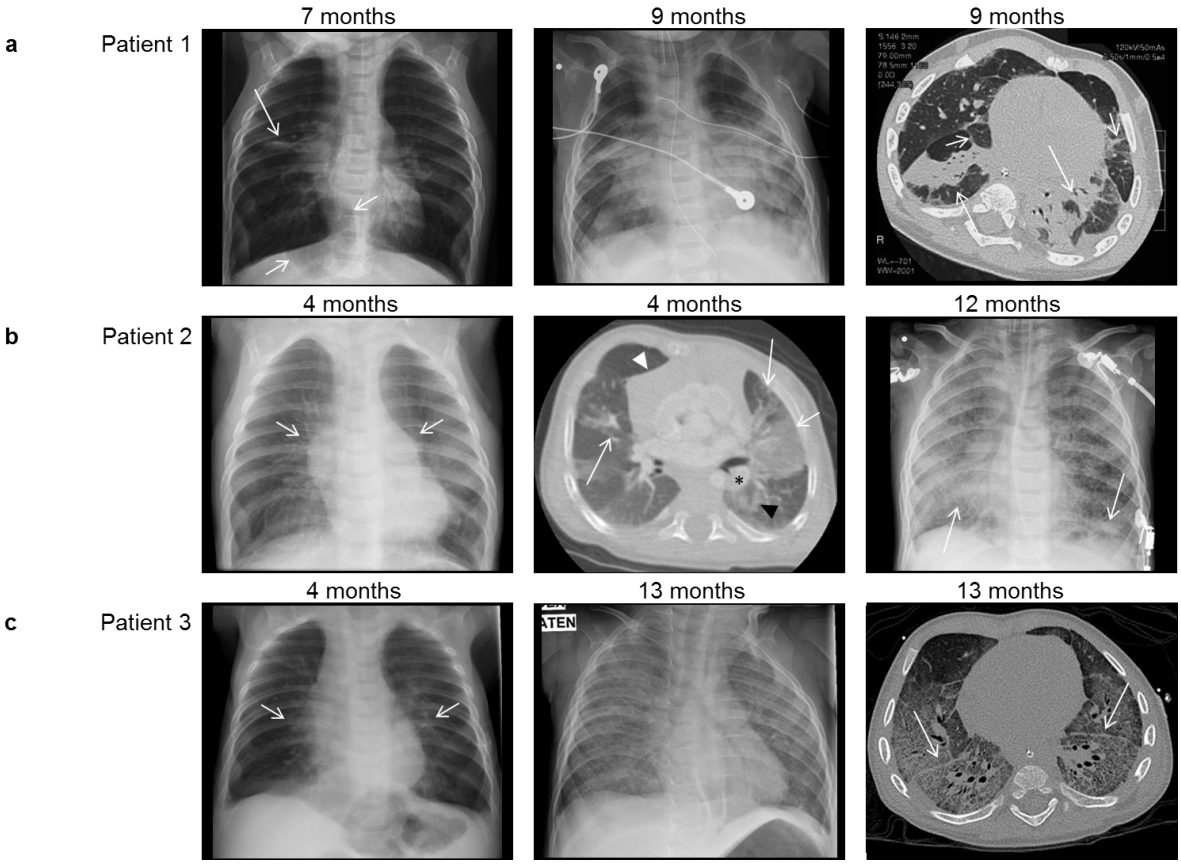

a
Patient 1
7 months
9 months
9 months
b
Patient 2
4 months
4 months
12 months
c
Patient 3
4 months
13 months
13 months

Fig. 8 Findings from brain magnetic resonance imaging of Patients 1–3. **a** Sagittal T1-weighted magnetic resonance imaging (MRI) shows the thin corpus callosum of Patient 1 at 7 months of age. **b** Patient 2 at 10 months of age. **c** Patient 3 at 10 months of age (short arrows). The axial T2-weighted and coronal T1-weighted MRIs of Patients 2 (**b**) and 3 (**c**) demonstrate slightly dilated lateral ventricles and cortical sulci. **c** Increased signal intensity of the globus pallidus on the axial T2-weighted MRI of Patient 3 (arrowheads)

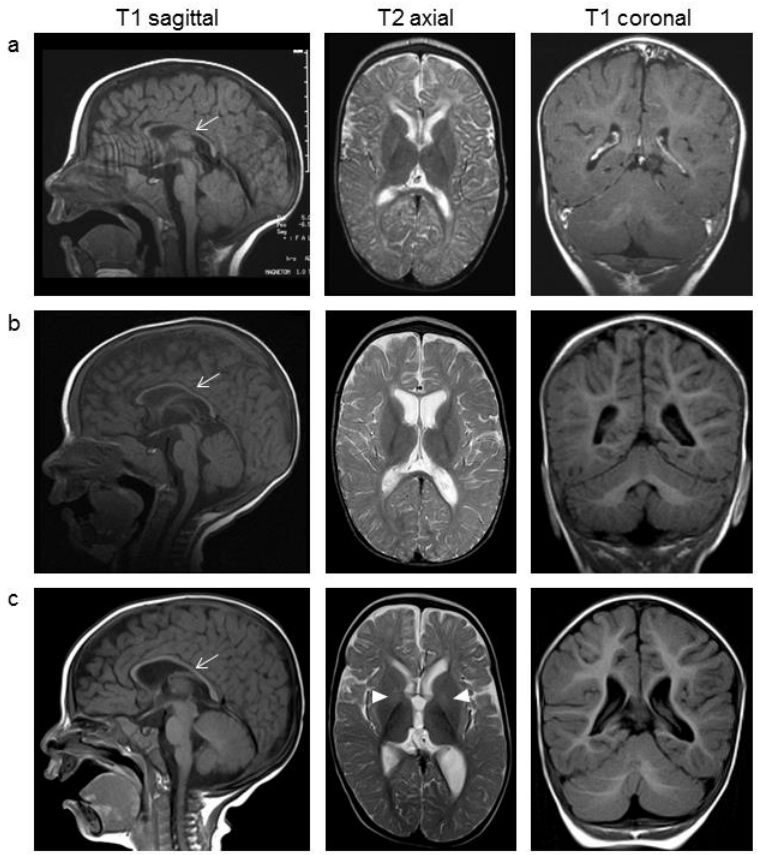


Fig. 9 Findings from abdominal ultrasound and magnetic resonance imaging of internal organs and thymus of Patients 1 and 2. **a** The abdominal ultrasound findings (Patient 1 at 10 months of age) show increased cortical echogenity with the strand-like hypoechogenic outerzone of both kidneys (short arrows). **b** The abdominal T2-weighted axial MRI of Patient 2 at 4 months of age shows hepatomegaly and a non-expansive hypointense lesion of the right liver lobe (long arrows). **c** The lesion is also hypointense on the T1-weighted coronal MRI (long arrows). **d** The axial fat saturation T1-weighted MRI of Patient 2 shows that the lesion (arrowheads) in the liver is less enhanced by the gadolinium contrast agent than the rest of the liver. **e** Normal echogenity of the right liver lobe of Patient 2 at the age of 4 months. **f** A slightly enlarged thymus of Patient 2, which was seen in the coronal T1-weighted MRI (asterisk)

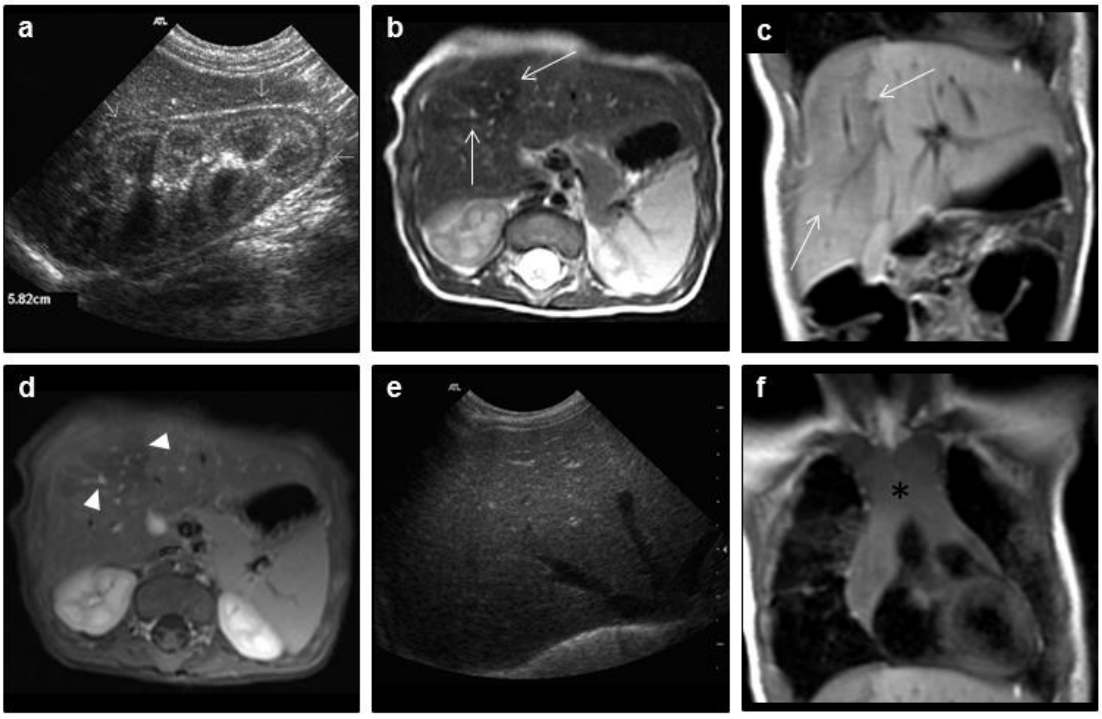


Fig. 10 a Electropherograms showing both heterozygote NM_198514:c.442G>T and c.601_602delAG variants in genomic DNA from the patient. **b** The RNA extracted from the patients' fibroblasts has only c.442G>T. Thus, Sanger sequencing of mRNA confirms that the variants reside in separate haplotypes and have been inherited as compound heterozygotes in all patients. Furthermore, the absence of the c.601_602delAG variant indicates that the transcript with the frameshift deletion is processed through nonsense mediated RNA decay

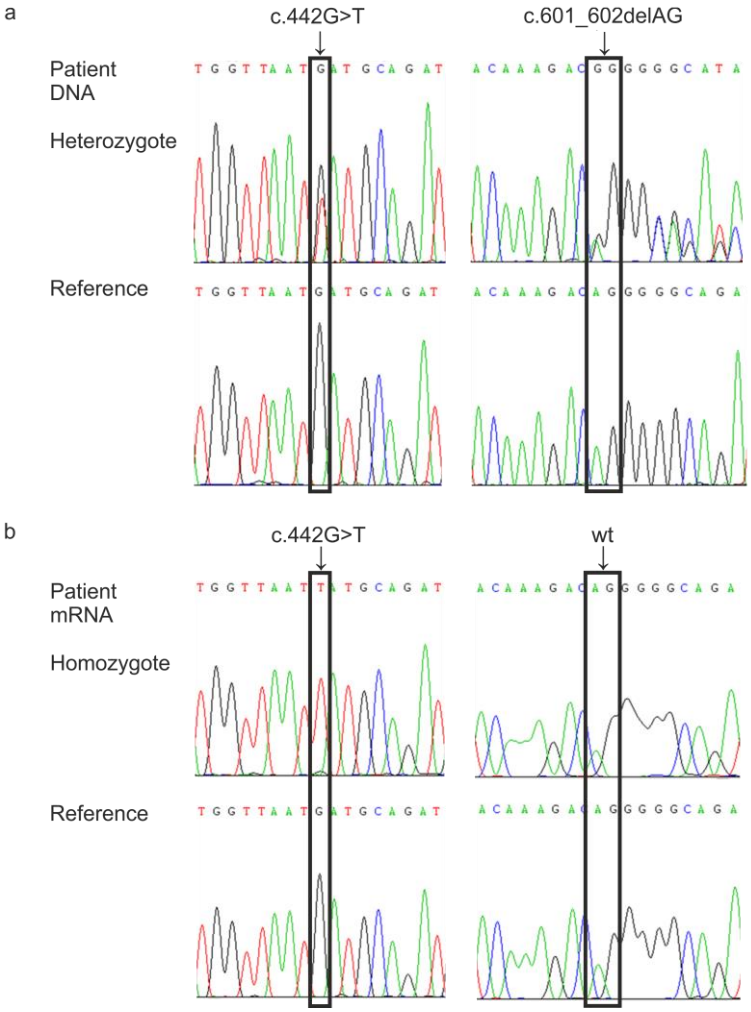


Fig. 11 Immunoblotting to detect NHLRC2 protein expression in human tissue homogenates from control autopsy samples. NHLRC2 was detected in all studied human organs including the heart (He), kidney (Ki), muscle (Mu), liver (Li), lung (Lu) and brain (Br). GAPDH was used as a loading control

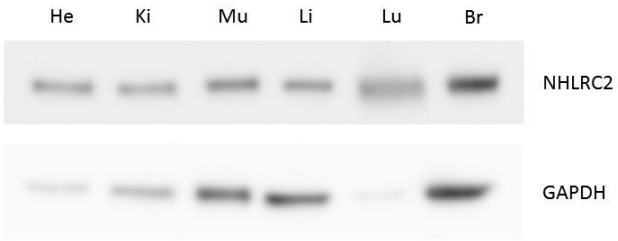
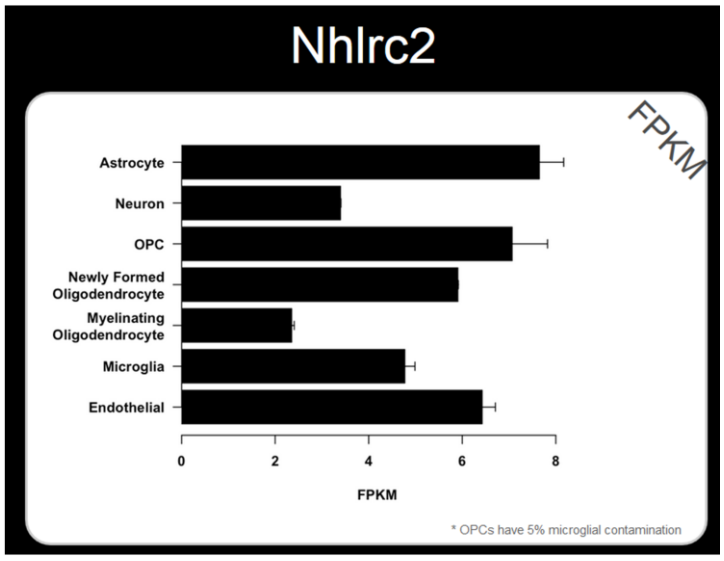
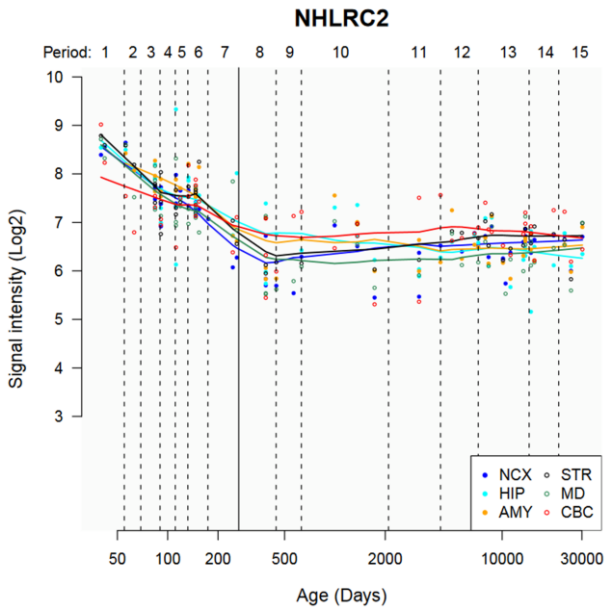


Fig. 12 *Nhlrc2* is expressed in various cell types of the mouse brain and *NHLRC2* expression is highest in the early human life. **a** An RNA-Seq transcriptome and splicing database of the glia, neurons and vascular cells of the cerebral cortex show the highest *Nhlrc2* expression in astrocytes, oligodendrocyte progenitor cells, newly formed oligodendrocytes and endothelial cells. *Nhlrc2* was also expressed in neurons, myelinating oligodendrocytes and microglia. (OPG) oligodendrocyte progenitor cells [18]. Abbreviations: FPKM, fragments per kilobase per million. **b** The average expression levels of *NHLRC2* are similar between the cerebellar cortex (CBC), mediodorsal nucleus of the thalamus (MD), striatum (STR), amygdala (AMY), hippocampus (HIP) and 11 areas of the neocortex (NCX) in the human brain (http://hbatlas.org/). **c** The expression is highest in the early development of the human brain. The number of samples is indicated on the top of the x-axis. Data from www.brainspan.org [5]. Abbreviations: RPKM, reads per kilobase per million; pcw, post conception week; mos, months; yrs, years

a



b


NHLRC2
Period: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Signal intensity (Log2)
10
9
8
7
6
5
4
3
NCX
HIP
AMY
STR
MD
CBC
50 100 200 500 2000 10000 30000
Age (Days)

c

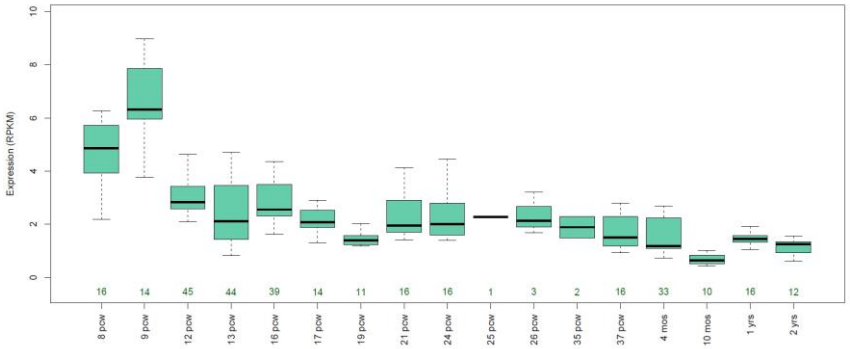

Expression (RPKM)
0 2 4 6 8 10
16 14 45 44 39 14 11 16 16 1 3 2 16 33 10 16 12
8 pcw
9 pcw
12 pcw
13 pcw
16 pcw
17 pcw
19 pcw
21 pcw
24 pcw
25 pcw
26 pcw
35 pcw
37 pcw
4 mos
10 mos
1 yrs
2 yrs

Fig. 13 Wild-type (WT) and heterozygous (het) *Nhlrc2* KO morulae and their growth *in vitro*. The morulae were isolated at embryonic day (E) 2.5 and co-cultured in a micro drop culture overnight. On the next day, the E3.5 compacted morulae and early blastocysts were transferred to separate gelatinised wells. At E4.5, the embryos had matured to the blastocyst stage and at E5.5 they hatched and attached to the bottom of the plate. Cells were grown for several days to detect possible phenotypes and to obtain enough material for genotyping (E7.5, E10.5 and E13.5). No homozygous morulae were detected

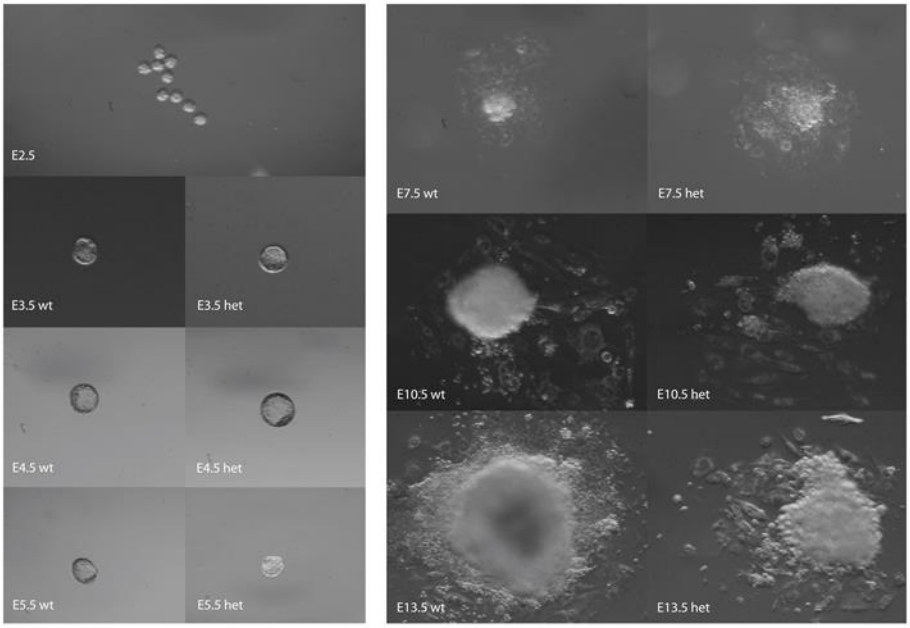
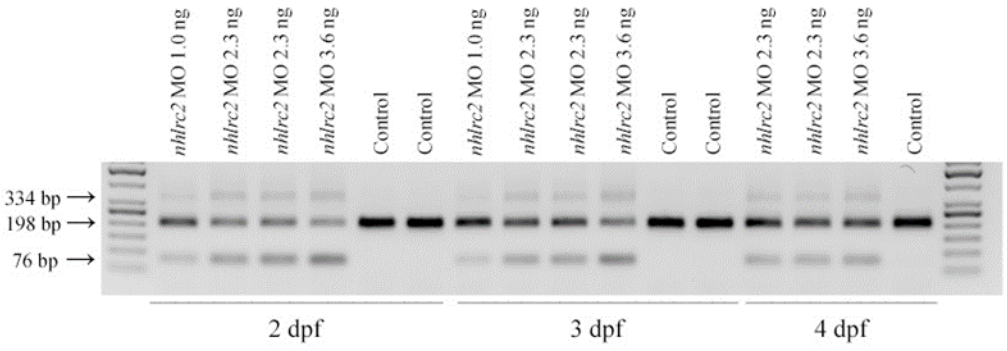


Fig. 14 *nhlrc2* splice site-blocking morpholino alters the mRNA spicing in zebrafish embryos but not the gross morphology of the embryos. Indicated amount of splice site targeting morpholino (together with an equal amount of p53 morpholino) was injected into the yolk sacs of fertilised eggs at the 1-2 cell stage. **a** Arrows on the left show the expected product sizes after the splice site-blocking events (334 bp intron inclusion, 198 bp WT product, 76 bp exon exclusion), and below the image are indicated the times of RNA extraction as days post fertilisation (dpf). **b** Un-injected controls on the left, *nhlrc2* splice site-blocking morpholino (3.6 ng) treated morphants in the middle and *nhlrc2* translation-blocking morpholino (2.3 ng) treated morphants on the right. The embryos appear to have WT morphology; only minor swelling was observed in the abdominal area in some individuals in both morphant groups at 3 dpf. However, this swelling diminished by 4 dpf. The images were taken with Axiovision software version 4.8 using a Zeiss SteREO Lumar V12 fluorescence microscope, equipped with 1.5X camera (Carl Zeiss MicroImaging GmbH, Göttingen, Germany)

a



b

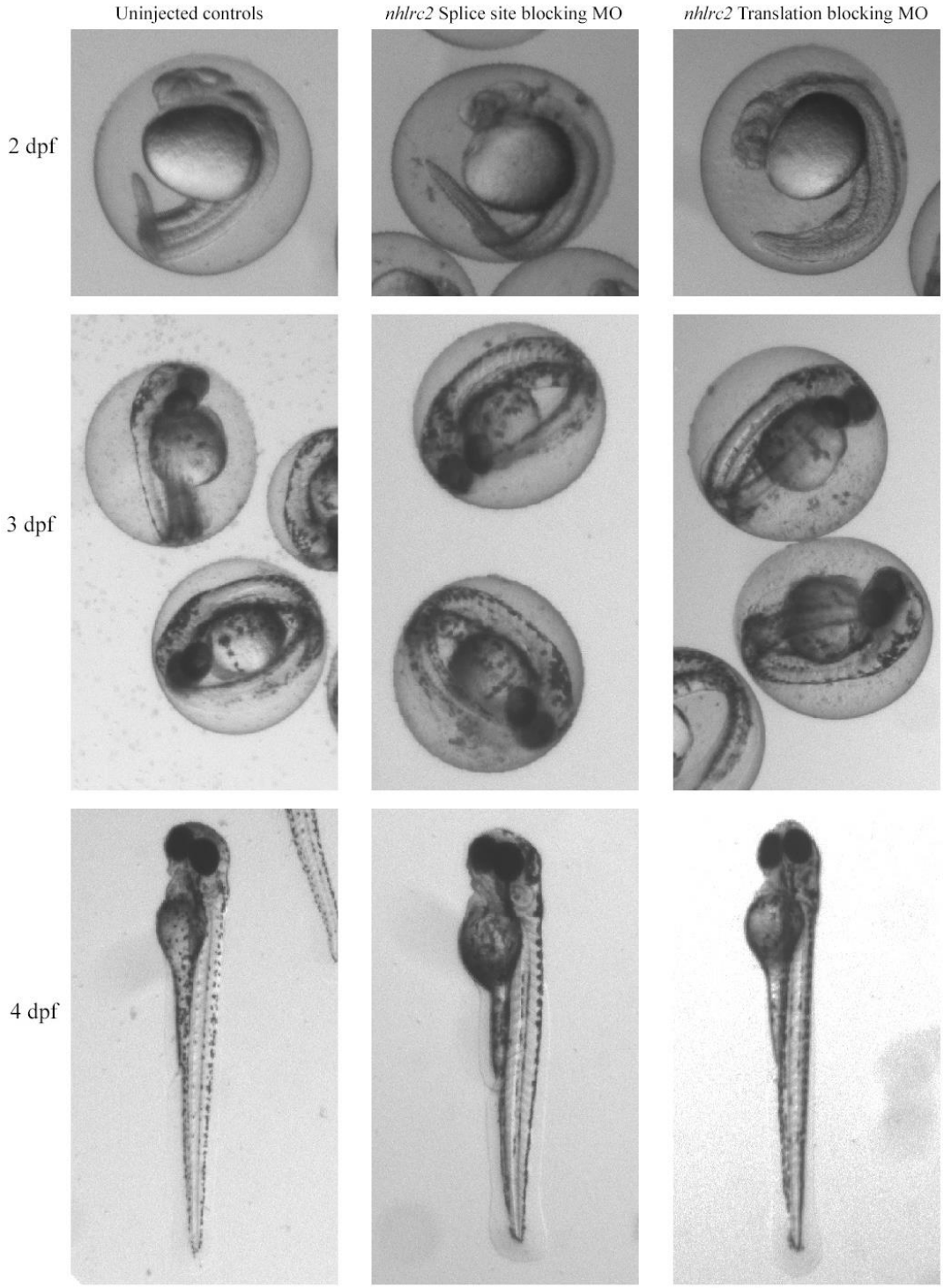

Uninjected controls
nhlrc2 Splice site blocking MO
nhlrc2 Translation blocking MO
2 dpf
3 dpf
4 dpf

Fig. 15 The timing of dechorionation of the *nhlrc2* morphants. The fraction of dechorionated embryos was recorded for selected hours post fertilisation (hpf). The significances were calculated with a log rank Mantel-Cox test. PBS control N=91, untreated control N=47; p53 MO control N=23; untreated control N=36; random control MO 3.6 ng N=63; untreated control N=34 and splice site-blocking MO 3.6 ng N= 76, untreated control N=34. Graphs were generated with GraphPadPrism 5

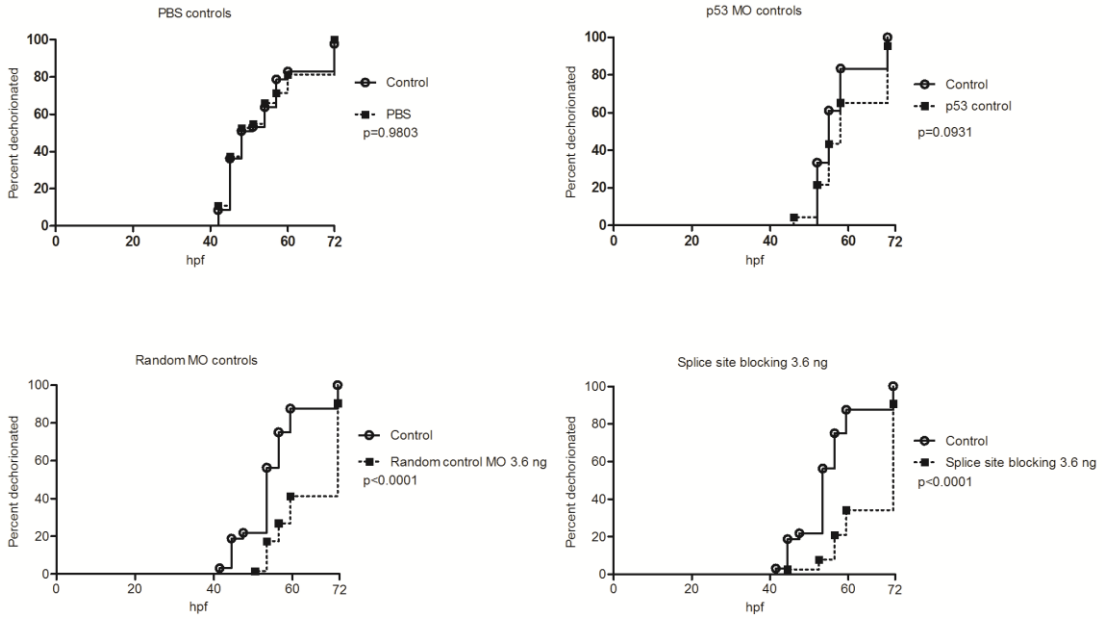
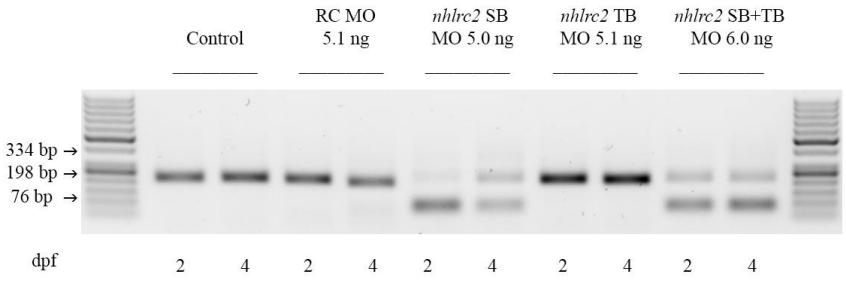
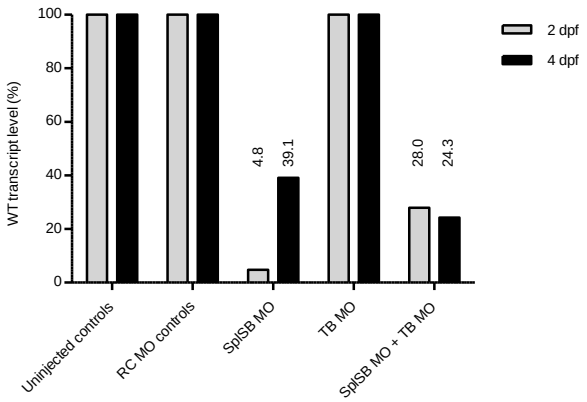


Fig. 16 Knockdown of *nhlrc2* using splice site-blocking morpholino (nhlrc2 SB MO), translation-blocking morpholino (nhlrc2 TB MO), a combination (nhlrc2 SB+TB MO) of the two morpholinos compared to un-injected controls (Control) and random control morpholino (RC MO). **a** *nhlrc2* splice site blocking morpholino alters the mRNA splicing in zebrafish embryos. The indicated amount of splice site-targeting morpholino (together with an equal amount of p53 morpholino) was injected into the yolk sacs of fertilised eggs at the 1-2 cell stage. After the indicated time, RNA was isolated, PCR was performed and PCR product was run on 2% agarose gel. Arrows on the left annotate the expected product sizes after the splice site-blocking events (334 bp intron inclusion, 198 bp WT product, 76 bp exon exclusion), and below the image are the times of RNA extraction as days post fertilisation (dpf). **b** Knockdown efficiencies were calculated using ImageJ 1.49v. The levels of WT transcript are displayed above the bars for samples containing splice site-blocking morpholinos

a



b


100
80
60
40
20
0
WT transcript level (%)
4.8
39.1
28.0
24.3
2 dpf
4 dpf
Uninjected controls
RC MO controls
SplSB MO
TB MO
SplSB MO + TB MO

References

1. Drozdetskiy A, Cole C, Procter J, Barton GJ (2015) JPred4: a protein secondary structure prediction server. Nucleic Acids Res 43:W389-94. Doi:10.1093/nar/gkv332

2. Gaciarz A, Veijola J, Uchida Y, Saaranen MJ, Wang C, Horkko S, et al. (2016) Systematic screening of soluble expression of antibody fragments in the cytoplasm of E. coli. Microb Cell Fact 15:22-016-0419-5. Doi:10.1186/s12934-016-0419-5

3. Karala AR, Ruddock LW (2010) Bacitracin is not a specific inhibitor of protein disulfide isomerase. FEBS J 277:2454-2462. Doi:10.1111/j.1742-4658.2010.07660.x

4. Karjalainen MK, Huusko JM, Ulvila J, Sotkasiira J, Luukkonen A, Teramo K, et al. (2012) A potential novel spontaneous preterm birth gene, AR, identified by linkage and association analysis of X chromosomal markers. PLoS One 7:e51378. Doi:10.1371/journal.pone.0051378

5. Miller JA, Ding SL, Sunkin SM, Smith KA, Ng L, Szafer A, et al. (2014) Transcriptional landscape of the prenatal human brain. Nature 508:199-206. Doi:10.1038/nature13185

6. Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, et al. (2015) The InterPro protein families database: the classification resource after 15 years. Nucleic Acids Res 43:D213-21. Doi:10.1093/nar/gku1243

7. Nagy A, Gertsenstein M, Vintersten K, Behringer R (2003) Manipulating the Mouse Embryo, A Laboratory Manual, 3rd ed. p.198-200. Cold Spring Harbor Laboratory Press, New york

8. Saaranen MJ, Karala AR, Lappi AK, Ruddock LW (2010) The role of dehydroascorbate in disulfide bond formation. Antioxid Redox Signal 12:15-25. Doi:10.1089/ars.2009.2674

9. Scavizzi F, Ryder E, Newman S, Raspa M, Gleeson D, Wardle-Jones H, et al. (2015) Blastocyst genotyping for quality control of mouse mutant archives: an ethical and economical approach. Transgenic Res 24:921-927. Doi:10.1007/s11248-015-9897-1

10. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9:671-675

11. Spandidos A, Wang X, Wang H, Seed B (2010) PrimerBank: a resource of human and mouse PCR primer pairs for gene expression detection and quantification. Nucleic Acids Res 38:D792-9. Doi:10.1093/nar/gkp1005

12. Spandidos A, Wang X, Wang H, Dragnev S, Thurber T, Seed B (2008) A comprehensive collection of experimentally validated primers for Polymerase Chain Reaction quantitation of murine transcript abundance. BMC Genomics 9:633-2164-9-633. Doi:10.1186/1471-2164-9-633

13. Sulonen AM, Ellonen P, Almusa H, Lepisto M, Eldfors S, Hannula S, et al. (2011) Comparison of solution-based exome capture methods for next generation sequencing. Genome Biol 12:R94-2011-12-9-r94. Doi:10.1186/gb-2011-12-9-r94

14. Varilo T, Savukoski M, Norio R, Santavuori P, Peltonen L, Jarvela I (1996) The age of human mutation: genealogical and linkage disequilibrium analysis of the CLN5 mutation in the Finnish population. Am J Hum Genet 58:506-512

15. Wang X, Seed B (2003) A PCR primer bank for quantitative gene expression analysis. Nucleic Acids Res 31:e154

16. Witkowski L, Carrot-Zhang J, Albrecht S, Fahiminiya S, Hamel N, Tomiak E, et al. (2014) Germline and somatic SMARCA4 mutations characterize small cell carcinoma of the ovary, hypercalcemic type. Nat Genet 46:438-443. Doi:10.1038/ng.2931

17. Yao J, Shoubridge EA (1999) Expression and functional analysis of SURF1 in Leigh syndrome patients with cytochrome c oxidase deficiency. Hum Mol Genet 8:2541-2549. Doi:ddc275

18. Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keeffe S, et al. (2014) An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. J Neurosci 34:11929-11947. Doi:10.1523/JNEUROSCI.1860-14.2014

PUBLICATION III

Intelectin 3 is dispensable for resistance against a mycobacterial infection in zebrafish (*Danio rerio*)

Ojanen MJT, **Uusi-Mäkelä MIE**, Harjula SE, Saralahti AK, Oksanen KE, Kähkönen N, Määttä JAE, Hytönen VP, Pesu M, Rämet M.

Scientific Reports. 2019 Jan 30;9(1):995.
doi: 10.1038/s41598-018-37678-1.

Publication reprinted with the permission of the copyright holders. This publication is licensed under a Creative Commons Attribution 4.0 International licence (http://creativecommons.org/licenses/by/4.0/).

SCIENTIFIC REPORTS

OPEN

Intelectin 3 is dispensable for resistance against a mycobacterial infection in zebrafish (*Danio rerio*)

Received: 9 October 2018
Accepted: 7 December 2018
Published online: 30 January 2019

Markus J. T. Ojanen[1,2], Meri I. E. Uusi-Mäkelä[1], Sanna-Kaisa E. Harjula[1], Anni K. Saralahti[1], Kaisa E. Oksanen[1], Niklas Kähkönen[3], Juha A. E. Määttä[3], Vesa P. Hytönen[3], Marko Pesu[2,4] & Mika Rämet[1,5,6,7]

Tuberculosis is a multifactorial bacterial disease, which can be modeled in the zebrafish (*Danio rerio*). Abdominal cavity infection with *Mycobacterium marinum*, a close relative of *Mycobacterium tuberculosis*, leads to a granulomatous disease in adult zebrafish, which replicates the different phases of human tuberculosis, including primary infection, latency and spontaneous reactivation. Here, we have carried out a transcriptional analysis of zebrafish challenged with low-dose of *M. marinum*, and identified *intelectin 3* (*itln3*) among the highly up-regulated genes. In order to clarify the *in vivo* significance of Itln3 in immunity, we created nonsense *itln3* mutant zebrafish by CRISPR/Cas9 mutagenesis and analyzed the outcome of *M. marinum* infection in both zebrafish embryos and adult fish. The lack of functional *itln3* did not affect survival or the mycobacterial burden in the zebrafish. Furthermore, embryonic survival was not affected when another mycobacterial challenge responsive *intelectin*, *itln1*, was silenced using morpholinos either in the WT or *itln3* mutant fish. In addition, *M. marinum* infection in dexamethasone-treated adult zebrafish, which have lowered lymphocyte counts, resulted in similar bacterial burden in both WT fish and homozygous *itln3* mutants. Collectively, although *itln3* expression is induced upon *M. marinum* infection in zebrafish, it is dispensable for protective mycobacterial immune response.

Tuberculosis is an epidemic multifactorial disease caused by *Mycobacterium tuberculosis*[1]. The susceptibility to tuberculosis depends on the *M. tuberculosis* strain and on a number of host-related factors such as environmental conditions, other underlying diseases as well as genetic variation[2–4]. Critical genes of the adaptive immunity required for the mycobacterial immune response such as *interferon gamma* (*IFNG*)[5,6] and *interleukin 12* (*IL12*)[7,8] were identified already in the 1980's and 1990's, respectively. The importance of these genes has later been verified in human tuberculosis patients[9] and by using experimental gene knockout mouse models of tuberculosis[10–13]. More recently, pattern recognition receptor (PRR) gene polymorphisms of Toll-like receptors (TLRs)[14–16] and C-type lectins[17,18], have been associated with *M. tuberculosis* susceptibility, delineating also the central role of the innate immunity in controlling the mycobacterial infection.

Lectins are carbohydrate-binding proteins important for numerous biological processes such as intracellular glycoprotein secretion, leukocyte trafficking and microbial recognition[19,20]. Consequently, lectins act as recognition molecules inside cells, on the cell surface and in extracellular fluids[20]. Intelectins (ITLNs) are a distinct family of lectins, which were first identified in *Xenopus laevis*[21] and were later found in a number of chordates including human, mouse and zebrafish (*Danio rerio*)[22–25]. Although ITLN function has been linked to a number of processes such as iron absorption[26], metabolic disorders[27] as well as cancer development[28,29], their exact biological functions are elusive. Suggesting a role for ITLNs in the immune response, *itln* gene expression is highly up-regulated upon a bacterial infection in fish[25,30–32]. Moreover, human ITLN1 (also known as Omentin) has been

[1]Laboratory of Experimental Immunology, BioMediTech Institute and Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland. [2]Laboratory of Immunoregulation, BioMediTech Institute and Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland. [3]Laboratory of Protein Dynamics, BioMediTech Institute and Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland. [4]Department of Dermatology, Tampere University Hospital, Tampere, Finland. [5]Department of Pediatrics, Tampere University Hospital, Tampere, Finland. [6]Department of Children and Adolescents, Oulu University Hospital, Oulu, Finland. [7]PEDEGO Research Unit and Medical Research Center Oulu, University of Oulu, Oulu, Finland. Correspondence and requests for materials should be addressed to M.R. (email: mika.ramet@uta.fi)

shown to bind to the *Mycobacterium bovis* Bacillus Calmette-Guérin (BCG)[33], and more specifically to exocyclic 1,2-diol glycan epitopes that are expressed selectively on microbial surfaces[34].

The importance of ITLNs for immunity *in vivo*, however, is less clear. Previously, Voehringer *et al*., used transgenic mice with lung-specific ITLN1 and ITLN2 over-expression to study the effects of these proteins in the mouse infection models of the parasite *Nippostrongylus brasiliensis* and the *M. tuberculosis* bacterium[35]. In these settings, the authors could not detect enhanced pathogen clearance in the *Itln* transgenic mice. In contrast, a so called "natural deletion" of the *Itln2* gene has been previously associated with a higher susceptibility against the parasite *Trichinella spiralis* in a C57BL/10 mouse strain[36]. Recently, an *Itln1* knockout mouse strain was created to study inflammatory bone diseases[37]. In the aforementioned study, the lack of Itln1 was associated with a proinflammatory phenotype characterized by elevated TNF and IL6 levels in bone tissue and in serum, and was shown to result in osteoporosis[37].

The genome of the zebrafish was assembled for the first time in 2002 and the prevalent 11th assembly (GRCz11) is an invaluable tool for research using zebrafish as a disease model[38]. Over 70% of human genes have at least one zebrafish orthologue and for this reason, the zebrafish immune system is highly similar compared to humans[38]. In fact, most of the human immune cell populations such as T- and B-cells[39–41], neutrophils and macrophages[42], dendritic cells[43] as well as the complement system[44] and immunoglobulins[45,46], are found in the zebrafish. Importantly, zebrafish can be modified genetically with the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated 9 (Cas9) mutagenesis[47,48], which allows disease modeling using reverse genetics, although some genes appear difficult to target successfully[49].

A *Mycobacterium marinum* infection of zebrafish is nowadays a commonly used model for studying tuberculosis in both larvae and adult fish[50,51]. Compared to several other tuberculosis models, the mycobacterial model of zebrafish is considered safe, cost-effective and ethical[52,53]. More importantly, *M. marinum* is closely related to *M. tuberculosis*, and the two bacterial species have comparable pathogenic characteristics in the natural hosts; macrophage mediated intracellular multiplication as well as the formation of granuloma structures[54–56]. The larval model enables studying specifically the innate immunity[57,58], whereas the adult zebrafish model allows studying also components of the adaptive immune system in both an acute mycobacterial infection[59] as well as during mycobacterial latency[56,60].

In order to identify candidate genes associated with the host response against mycobacteria, we conducted a gene expression microarray in *M. marinum* infected adult zebrafish. Here, we identified a zebrafish *ITLN* orthologue *itln3* among the genes that were most induced upon infection. In order to gain more insights into the function of ITLNs, we used CRISPR/Cas9 mutagenesis to create knockout *itln3* mutant zebrafish lines, and used the zebrafish *M. marinum* infection model to determine the *in vivo* significance of Itln3 in a mycobacterial infection.

Results

Genome-wide gene expression microarray analysis of *M. marinum* infected adult zebrafish. In order to identify genes involved in the host immune response against mycobacteria, we used the zebrafish *M. marinum* infection model and conducted a genome-wide gene expression analysis using the microarray platform. To this end, we infected wild-type (WT) AB zebrafish with *M. marinum* (20 CFU; SD 6 CFU) and isolated their organ blocks (includes all the organs of the abdominal cavity) for a transcriptomic analysis at 14 days post infection (dpi). From a total of 43603 probes used in the analysis, we found 93 probes, corresponding to 70 genes, that were up-regulated and 26 probes, corresponding to 21 genes, that were down-regulated (log2 fold change $>|3|$) compared to the mock-treated (PBS) controls (Supplementary Table 1). Further evaluation of the up-regulated probes with a GOrilla gene ontology (GO) enrichment analysis[61,62] revealed 22 enriched ($p < 0.001$) processes including response to carbohydrates (GO:0009743), cholesterol homeostasis (GO:0042632) and antigen processing and presentation (GO:0019882) (Supplementary Table 2). Among the up-regulated genes we found five genes; *si:busm1-194e12.11* (*mhc2* family gene), *arachidonate 5-lipoxygenase b, tandem duplicate 3* (*alox5b.3*), *zgc:113912* (*mhc2* family gene), *CD59 molecule* (*cd59*) and *si:busm1-194e12.12* (*mhc2* family gene) with well-known immunological functions in antigen processing, inflammation and in the regulation of the complement system (Fig. 1A, Supplementary Table 1). Of the 21 down-regulated genes, five were associated with the immune response; *CD58 molecule* (*cd58*), *myeloid-specific peroxidase* (*mpx*), *complement factor b-like* (*cfbl*), *immunoresponsive gene 1, like* (*irg1l*) and *si:busm1-266f07.1* (*mhc2* family gene) (Fig. 1A, Supplementary Table 1). Interestingly, approximately 38% of the up-regulated probes i.e. *parvalbumin 1* (*pvalb1*), *alpha-tropomyosin* (*tpma*), *troponin I, skeletal, fast 2b, tandem duplicate 2* (*tnni2b.2*) and *myosin, heavy polypeptide 1.1* (*myhz1.1*) were related to muscle associated biological processes including muscle contraction (GO:0006936), muscle system process (GO:0003012) and myofibril assembly (GO:0030239) (Supplementary Tables 1 and 2). The GO-analysis of the down-regulated probes also showed a significant enrichment of another 22 processes including response to external biotic stimulus (GO:0043207) and cholesterol biosynthetic process (GO:0006695) and immunological processes, such as response to other organism (GO:0051707), response to bacterium (GO:0009617) and the induction of bacterial agglutination (GO:0043152) (Supplementary Table 2).

Mycobacterial infection up-regulates *itln3* expression in both zebrafish embryos and adult fish. Previous studies in several animal models have shown the expression of the *Intelectin* (*ITLN*) gene to be induced upon a bacterial infection[25,30,32]. Accordingly, the expression of the zebrafish *itln3* (ENSDARG00000003523) was increased on average 3.3-fold (log2 change) upon a *M. marinum* infection in our microarray analysis (Fig. 1A, Supplementary Table 1). In contrast, two other *itln* genes; *itln2* (ENSDARG00000036084) and *itln2-like* (ENSDARG00000093796) were down-regulated compared to the PBS controls (-3.5 and -3.2 log2 fold change, respectively) (Fig. 1A, Supplementary Table 1), suggesting a diverse regulation of *itln* genes in the *M. marinum* infected zebrafish. Since both ENSDARG00000036084 and

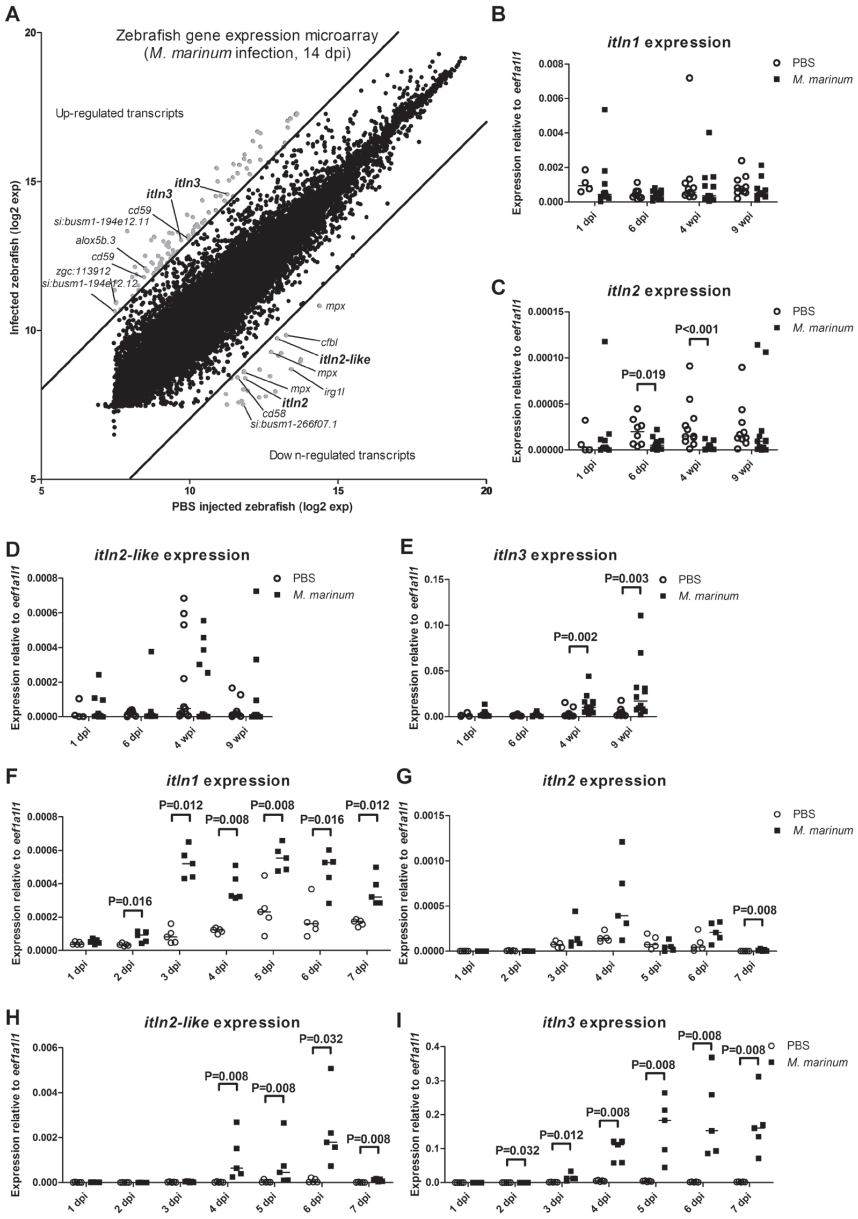


Figure 1. Zebrafish *intelectin* genes are differentially expressed upon *M. marinum* infection. (**A**) A genome-wide gene expression microarray was conducted in adult WT AB zebrafish injected with *M. marinum* (20 CFU; SD 6 CFU) (n = 2) or PBS (n = 3). Average numerical results (log2) for each probe in both infected fish (y-axis) and PBS controls (x-axis) are shown. Up- and down-regulated transcripts (log2 fold change|3|) in the organ blocks are shown in grey, and the common immunological genes are annotated. Two *itln3* probes as well as *itln2* and *itln2-like* probes are highlighted. (**B**–**E**) The expression of zebrafish *itln* genes (*itln1*, *itln2*, *itln2-like* and *itln3*) was measured with qPCR in the organ blocks of the *M. marinum* infected (6 CFU; SD 3 CFU) and PBS injected adult WT e46 zebrafish at 1 (n = 12 and n = 4, respectively) and 6 dpi (n = 12 and n = 8, respectively) as well as 4 (n = 12 and n = 11, respectively) and 9 wpi (n = 12 and n = 10, respectively). (**F**–**I**) The expression of *itln1*, *itln2*, *itln2-like* and *itln3* was determined with qPCR in the *M. marinum* (39 CFU; SD 47 CFU) infected WT AB embryos (n = 5 at all timepoints) and in PBS injected controls (n = 5 at all timepoints) at 1–7 dpi. Note the different scales of the y axes in B-I. Gene expressions were normalized to *eef1a1l1* expression and target genes were run once in the qPCR analyses. A two-tailed Mann-Whitney test was used in the statistical comparison of differences in B–I.

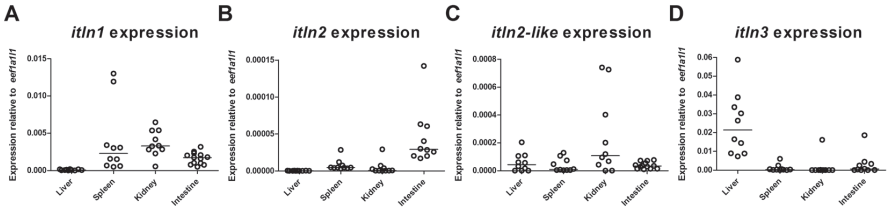


Figure 2. Expression of zebrafish *itln* genes in adult zebrafish tissues. Relative expression of (**A**) *itln1*, (**B**) *itln2*, (**C**) *itln2-like* and (**D**) *itln3* was measured with qPCR in the uninfected adult WT e46 zebrafish liver (n = 10), spleen (n = 10), kidney (n = 10) and the intestine (n = 10). Note the different scales of the y axes. Gene expressions were normalized to *eef1a1l1* expression and target genes were run once in the qPCR analyses.

ENSDARG00000093796 share the same gene name, *itln2*, in Ensembl genome browser, ENSDARG00000093796 is referred to as *itln2-like* throughout the text.

To confirm the differential expression pattern of the *itln* family members in the zebrafish mycobacterial infection and to study the kinetics of the host response more carefully, we analyzed *itln1* (ENSDARG00000007534), *itln2*, *itln2-like* and *itln3* gene expression from the abdominal cavity organ blocks of *M. marinum* infected (6 CFU; SD 3 CFU) WT e46 background adult zebrafish with qPCR at 1 and 6 dpi, as well as at 4 and 9 weeks post infection (wpi) (Fig. 1B–E). In line with our microarray data, *itln3* was significantly induced at 4 wpi (3.8-fold, $P = 0.002$) and 9 wpi (5.9-fold, $P = 0.003$) (Fig. 1E), whereas *itln2* was down-regulated compared to the PBS controls both at 6 dpi (0.3-fold, $P = 0.019$) and 4 wpi (0.2-fold, $P < 0.001$) (Fig. 1C). No significant differences in the relative mRNA expression levels of the *itln1* (Fig. 1B) or *itln2-like* (Fig. 1D) genes were observed between infected and the PBS injected adult fish at any of the measured time points.

Next, we infected WT AB zebrafish embryos with mycobacteria and performed an expression analysis of the *itln* genes by qPCR. Here, *M. marinum* (39 CFU; SD 47 CFU) was microinjected into the yolk sac of the embryos and the gene expression was quantified daily between 1 and 7 dpi (Fig. 1F–I). In the mycobacteria infected embryos we detected the up-regulation of both *itln1* (1.8 to 6.4-fold, $P = 0.008–0.016$) (Fig. 1F) and *itln3* (1.8 to 111.4-fold, $P = 0.008–0.032$) (Fig. 1I) starting at 2 dpi and continuing until 7 dpi, as well as the induction of *itln2* at 7 dpi (21.6-fold, $P = 0.008$) (Fig. 1G) and *itln2-like* (Fig. 1H) between 4 and 7 dpi (20.7 to 76.1-fold, $P = 0.008–0.032$) compared to the PBS controls. Also, in line with previous reports suggesting that other infectious diseases up-regulate *ITLN* expression, a significant induction of *itln3* expression (8.4-fold at 7hpi; 11.8-fold at 18hpi; 5.5-fold at 24hpi; 4.4-fold at 48hpi, $P = 0.002$ in all comparisons) was observed in *Streptococcus pneumoniae* (T4 serotype) infected (296 CFU; SD 32 CFU) embryos (Supplementary Figure 1).

In order to understand the infection-inducible nature of the zebrafish *itln* genes at steady state, we quantified the relative mRNA levels of *itln1*, *itln2*, *itln2-like* and *itln3* in the liver, spleen, kidney and intestine of unchallenged WT e46 zebrafish by qPCR (Fig. 2A–D). Here, we found that *itln2* expression was restricted to the intestine (Fig. 2B), whereas *itln3* showed the highest relative expression in the liver and the highest overall expression compared to the housekeeping gene (*eukaryotic translation elongation factor 1 alpha 1, like 1; eef1a1l1*) (Fig. 2D). Conversely, *itln1* was expressed in all of the studied tissues with the second highest overall expression levels (Fig. 2A), while *itln2-like* was primarily expressed in the zebrafish kidney and the intestine (Fig. 2C). These results are in line with a previous qPCR analysis of the *itln* gene family members in unchallenged adult zebrafish[25].

Creating *itln3* mutant zebrafish using CRISPR/Cas9 mutagenesis. The type II CRISPR/Cas system is an invaluable technology for targeted genome editing[63,64], and to date it has been utilized in a number of model organisms. We and others have used the CRISPR/Cas9 mutagenesis method successfully in the zebrafish[47,49,65,66]. Here, we used the CRISPR/Cas9 method to create zebrafish carrying nonsense *itln3* mutations for our *in vivo* studies (Fig. 3). To this end, we identified a functional gRNA targeting the second exon of the *itln3* gene with an average mutagenesis efficiency of 39.5% (Fig. 3A,B). After an outcross of parental mutation carriers (F0-generation) with WT TL zebrafish, we observed two germ-line transmitted frameshift mutations in the F1-progeny corresponding to a total loss of five base pairs (−5 bp; loss of GCATC) and to a total gain of eight base pairs (+8 bp; loss of GGAGCATC and gain of TGCTAGGTAAGTATCA) at the target loci (Fig. 3C). Analyses with the Translate tool (Expasy; SIB, Swiss Institute of Bioinformatics)[67] of both the −5 bp and +8 bp mutations confirmed the disrupted reading frames from amino acids 47 and 45 onwards resulting in premature stop-codons after 79 and 71 amino acids, respectively (Fig. 3C). These two different *itln3* null mutant zebrafish lines were named *itln3*^{uta145} (−5 bp mutation) and *itln3*^{uta148} (+8 bp mutation). qPCR analysis of uninjected and *M. marinum* infected (422 CFU; SD 221 CFU, 2 wpi) adult zebrafish revealed diminished *itln3* transcript levels in the homozygous *itln3*$^{uta145/uta145}$ (residual expression less than 1%, $P < 0.001$) and *itln3*$^{uta148/uta148}$ mutants (residual expression less than 0.1%, $P < 0.001$) compared to the WT controls (Supplementary Figure 2), suggesting that the indel-mutations lead to the nonsense-mediated RNA decay of the mutant mRNAs[68]. Furthermore, the inheritance of the mutations followed Mendelian ratios for both of the mutant lines, and the homozygous *itln3*$^{uta145/uta145}$ and *itln3*$^{uta148/uta148}$ mutants did not show any developmental defects nor phenotypical differences compared to their WT siblings (Supplementary Figure 3).

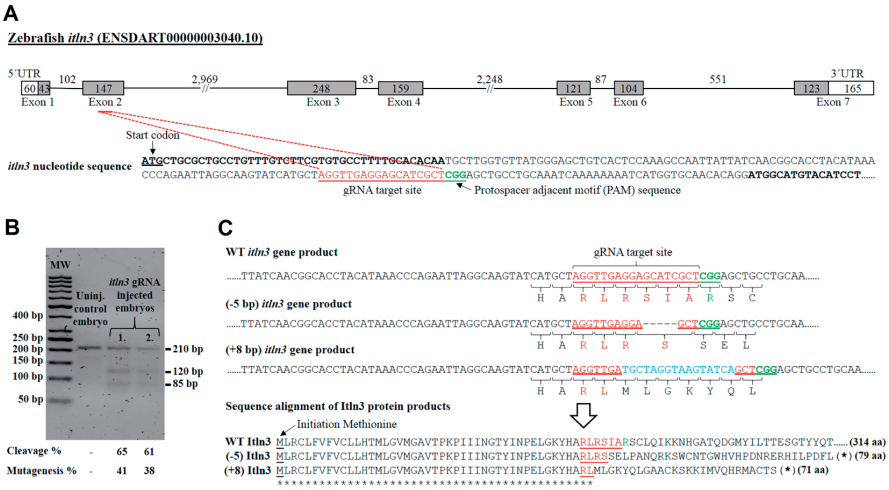


Figure 3. Generation of *itln3*^{uta145} and *itln3*^{uta148} mutant zebrafish lines using CRISPR-Cas9 mutagenesis. (**A**) An appropriate guide RNA (gRNA) target site was identified in the second exon of *itln3*. (**B**) 2.5% agarose TAE gel electrophoresis was performed to evaluate occurrence of target site mutations in zebrafish. The *in vivo* CRISPR/Cas9 mutagenesis efficiency was estimated with the T7EI assay in the gRNA and *Cas9* mRNA injected embryos. The size of the uninjected WT control PCR product is 210 bp, whereas the PCR products of the mutated embryos are partially cleaved at the target site. The cleavage efficiency was calculated from the band intensities and the mutagenesis efficiency calculated according to the following formula: % mutagenesis = 100 × (1 − (1- fraction of cleavage)$^{1/2}$)[64]. GeneRuler 50 bp DNA Ladder (#SM0373, Thermo Fischer Scientific) was used as a molecular weight marker (MW). Gel image is cropped to exclude portions that do not contain experimental samples. (**C**) gRNA target sites were sequenced from F1-generation mutant zebrafish and two frameshift mutations (-5 bp deletion, *itln3*^{uta145} and +8 bp insertion, *itln3*^{uta148}) detected, leading to truncated protein products of 79 and 71 amino acids, respectively.

Nonsense mutation in *itln3* does not affect host resistance against *M. marinum* in zebrafish embryos. The up-regulation of the expression of the *itln3* gene in a *M. marinum* infection suggests a possible role for Itln3 in the host immunity against mycobacterial infections. To test if the resistance towards a mycobacterial infection is altered in homozygous *itln3* mutant embryos, we first infected *M. marinum* (40 CFU; SD 30 CFU) into the yolk sac of the ungenotyped F2-progeny of heterozygous *itln3*$^{uta145/+}$ and *itln3*$^{uta148/+}$ zebrafish and followed their survival until 7 days post fertilization (dpf) (Fig. 4A,B). Post-experiment genotyping revealed an average survival of 47% in the *itln3*^{uta145} background embryos and 48% in the embryos with the *itln3*^{uta148} background. However, any significant differences in the survival between the homozygous and heterozygous *itln3* mutants or WT fish could not be observed in either *itln3*^{uta145} (Fig. 4A) or *itln3*^{uta148} fish lines (Fig. 4B) before 7 dpi (7 dpf). Next, we quantified the mycobacterial burden in the embryos that had survived by qPCR using primers for *M. marinum internal transcribed spacer* (*MMITS*)[56] (Fig. 4C). The *M. marinum* quantification revealed bacterial copy number medians (log10) of 4.18 and 4.15 in 100 ng of zebrafish DNA in the *itln3*^{uta145} and *itln3*^{uta145} WT groups, respectively. Comparably, heterozygous *itln3*$^{uta145/+}$ and *itln3*$^{uta148/+}$ fish had copy number medians of 4.02 and 4.22 (in 100 ng of zebrafish DNA, log10), respectively, and the homozygous *itln3*$^{uta145/uta145}$ and *itln3*$^{uta148/uta148}$ mutants 3.53 and 4.25 (in 100 ng of zebrafish DNA, log10). Thus, there were no statistically significant differences in the mycobacterial burdens between the different genotypes in neither *itln3*^{uta145} nor *itln3*^{uta148} zebrafish.

The site of the bacterial injection can affect the immune response in the embryos[50]. Therefore, we next treated the ungenotyped F2-progeny of *itln3*$^{uta145/+}$ and *itln3*$^{uta148/+}$ zebrafish by injecting *M. marinum* into the blood circulation valley of 2-day-old embryos. In these fish, the *M. marinum* infection (46 CFU; SD 31 CFU) was not able to cause any mortality prior to the experimental end-point of 5 dpi (7 dpf). However, this allowed us to quantify the *M. marinum* burden in all of the infected embryos at the end-point (Fig. 4D). Here, the bacterial copy number medians (log10) in 100 ng of zebrafish DNA were 3.61 (WT *itln3*^{uta145}), 3.83 (WT *itln3*^{uta148}), 3.63 (*itln3*$^{uta145/+}$), 3.77 (*itln3*$^{uta148/+}$), 3.75 (*itln3*$^{uta145/uta145}$) and 3.79 (*itln3*$^{uta148/uta148}$). Similarly to the yolk sac infection, mycobacterial quantification did not reveal any differences between the individuals of the different genotypes in either the *itln3*^{uta145} or the *itln3*^{uta148} zebrafish background. Noteworthy, we also infected the ungenotyped F2-progeny of *itln3*$^{uta145/+}$ and *itln3*$^{uta148/+}$ zebrafish with *S. pneumoniae* (serotypes 1 and T4, blood circulation valley infection at 2 dpf) and followed the survival of the fish to 5 dpi[69]. There was no difference between WT embryos and the *itln3* mutants (Supplementary Figure 1).

Deleterious mutations may lead to genetic compensation, which in turn can affect the observed phenotype in gene knockout models[70]. To address this, we used a morpholino-oligonucleotide to silence *itln1* in our *itln3* mutant zebrafish together with the yolk sac mycobacterial infection of zebrafish embryos. In order to ensure efficient termination of translation in all of the four zebrafish *itln1* transcripts, we targeted the second exon (E2)

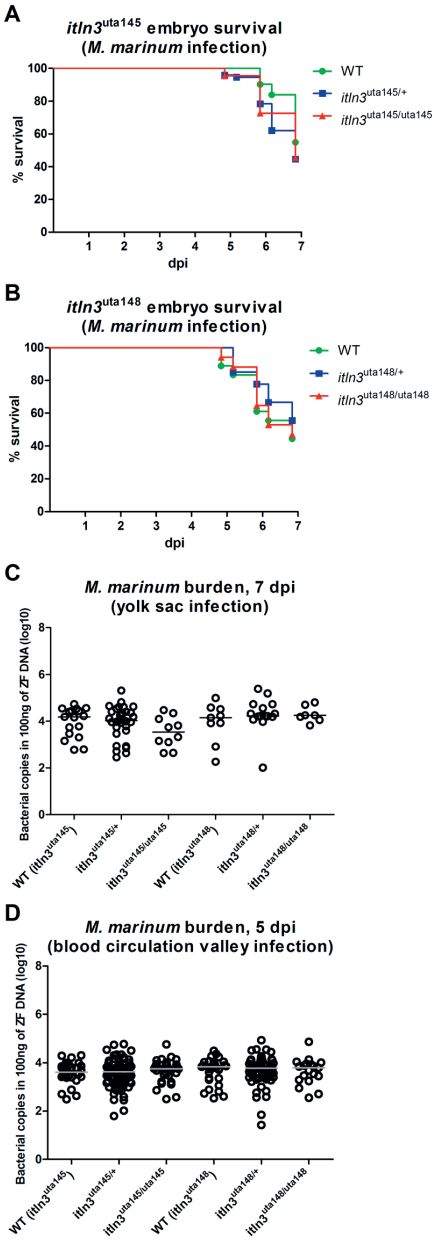


Figure 4. The lack of *itln3* does not affect the survival or the mycobacterial burden of *M. marinum* infected zebrafish embryos. (**A,B**) *M. marinum* (40 CFU; SD 30 CFU) was injected into the yolk sac of the WT (*itln3*^{uta145}) (n = 31), *itln3*$^{uta145/+}$ (n = 74), *itln3*$^{uta145/uta145}$ (n = 22), WT (*itln3*^{uta148}) (n = 19), *itln3*$^{uta148/+}$ (n = 27) and *itln3*$^{uta148/uta148}$ (n = 16) zebrafish embryos at 0 dpf and the survival recorded until 7 dpi. A log-rank (Mantel-Cox) test was used for the statistical comparison of differences. The data was collected from a single experiment. (**C**) Mycobacterial burden was measured by qPCR from the yolk sac infected WT (*itln3*^{uta145}) (n = 17), *itln3*$^{uta145/+}$ (n = 32), *itln3*$^{uta145/uta145}$ (n = 10), WT (*itln3*^{uta148}) (n = 9), *itln3*$^{uta148/+}$ (n = 15) and *itln3*$^{uta148/uta148}$ (n = 7) embryos that were alive at 7 dpi. (**D**) *M. marinum* (46 CFU; SD 31 CFU) was injected into the blood circulation valley of the WT (*itln3*^{uta145}) (n = 29), *itln3*$^{uta145/+}$ (n = 77), *itln3*$^{uta145/uta145}$ (n = 36), WT (*itln3*^{uta148}) (n = 31), *itln3*$^{uta148/+}$ (n = 57) and *itln3*$^{uta148/uta148}$ (n = 19) zebrafish embryos at 2 dpf and the *M. marinum* burden quantified at 5 dpi. Bacterial load is represented in panels C and D as bacterial copies (log10) in 100 ng of zebrafish DNA. A two-tailed Mann-Whitney test was used in the statistical comparison of differences in C and D.

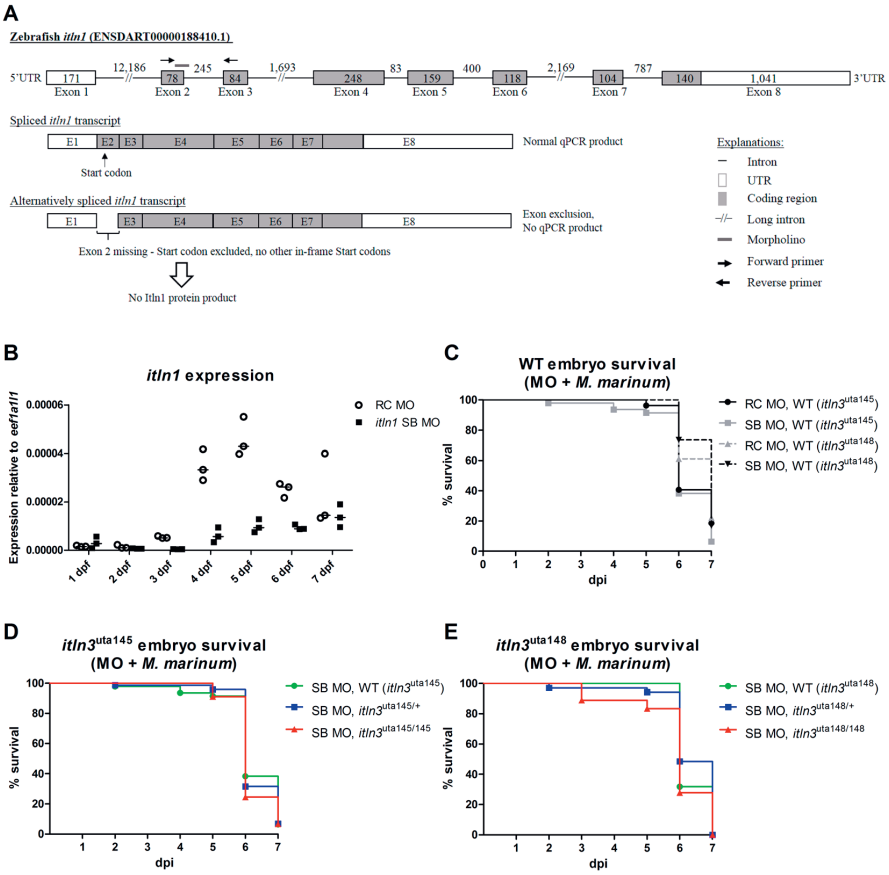


Figure 5. Morpholino mediated silencing of *itln1* expression does not alter the survival of the WT or *itln3* knockout zebrafish in a *M. marinum* infection. (**A**) A schematic representation of the effects of the morpholino mediated silencing of *itln1*. A splice site blocking morpholino (SB) was used to prevent the normal splicing event between exon 2 and exon 3 in *itln1*. Morpholino binding to its target site leads to an alternative splicing event that deletes the start codon containing exon 2 from the transcript. Consequently, this prevents translation of the Itln1 protein. In order to quantify the relative amount of the WT *itln1* transcript, qPCR primers were designed to specifically amplify only the WT *itln1* mRNA. (**B**) WT *itln1* expression was quantified with qPCR from the *itln1* SB morpholino (n = 3) and random control morpholino (RC) injected zebrafish (n = 3) at 1–7 dpf. Gene expression was normalized to *eef1a1l1* expression. All samples were run once as technical duplicates. (**C**–**E**) Survival of the morpholino and *M. marinum* (20 CFU; SD 19 or 13 CFU; SD 10 CFU) co-injected embryos were followed until 7 dpi. In panel C, WT (*itln3*^{uta145} and *itln3*^{uta148}) embryos injected with either SB (n = 47 and n = 53, respectively) or RC morpholino (n = 29 and n = 54) are shown, whereas in panels D and E the *itln3*^{uta145} background (n = 45–73) and *itln3*^{uta148} background embryos (n = 18–35) injected with SB morpholino are depicted, respectively. Note that the SB morpholino injected WT (*itln3*^{uta145}) embryo group is shown in both C and D panels in order to simplify data representation. The data in panel C was collected from two individual experiments, whereas other data is from a single experiment. A log-rank (Mantel-Cox) test was used for the statistical comparison of differences. MO = morpholino.

of the gene with a splice-blocking (SB) morpholino (Fig. 5A). In our initial SB morpholino titration experiments, 2.8 ng of *itln1*-blocking morpholino did not reveal any adverse effects on the survival or on the phenotype of unchallenged zebrafish embryos within the first 7 dpf. However, lower WT *itln1* mRNA levels were observed in the SB morphants with residual expression of 17.1% at 4 dpi, 21.8% at 5 dpi and 33.9% at 6 dpi compared to the random control (RC) injected embryos (Fig. 5B), demonstrating that this amount of the SB morpholino silences the expression of *itln1* efficiently during embryonic development. In addition, detectable *itln1* expression levels were observed in the RC morphants already at 1 dpi, whereas in the SB morpholino injected embryos *itln1* expression was evident later starting at 2 dpi based on qPCR (Fig. 5B, Supplementary Figure 4). Next, we performed morpholino-*M. marinum* co-injections (20 CFU; SD 19 CFU) into the yolk sac of the un-genotyped F2-progeny

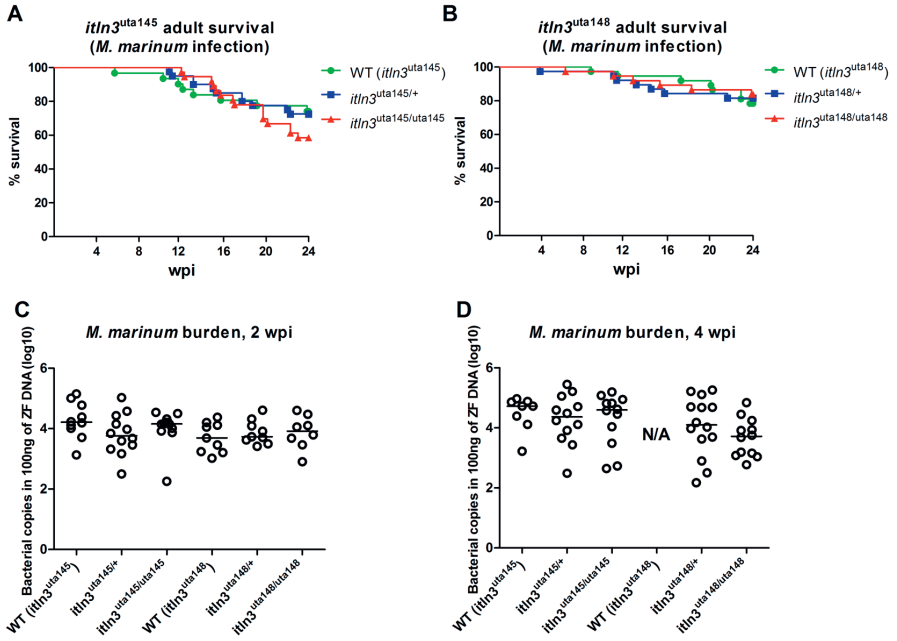


Figure 6. Adult *itln3* mutant zebrafish have comparable survival and mycobacterial burden compared to WT fish upon *M. marinum* infection. (**A**) The WT (*itln3*^{uta145}) (n = 38), *iltn3*$^{uta145/+}$ (n = 40), *iltn3*$^{uta145/uta145}$ (n = 38) and (**B**) WT (*itln3*^{uta148}) (n = 38), *iltn3*$^{uta148/+}$ (n = 38), *iltn3*$^{uta148/uta148}$ (n = 38) zebrafish were infected with *M. marinum* (48 CFU; SD 5), and their survival followed for 24 weeks. A log-rank (Mantel-Cox) test was used for the statistical comparison of differences. The data was collected from a single experiment. (**C**,**D**) The *itln3*^{uta145} and *itln3*^{uta148} background zebrafish were infected with *M. marinum* (422 CFU; SD 221 CFU) and bacterial burden (log10) in 100 ng of zebrafish DNA determined at 2 and 4 wpi from the organ blocks (without the kidney). Group sizes at 2 and 4 wpi, respectively, were as follows: WT (*itln3*^{uta145}) n = 10, n = 8; *iltn3*$^{uta145/+}$ n = 12, n = 12; *iltn3*$^{uta145/uta145}$ n = 12, n = 12; WT (*itln3*^{uta148}) n = 9, n = N/A; *iltn3*$^{uta148/+}$ n = 9, n = 14 and *iltn3*$^{uta148/uta148}$ n = 8, n = 12. All samples were run once. A two-tailed Mann-Whitney test was used in the statistical comparison of differences. N/A = no fish available for analysis.

of *itln3*$^{uta145/+}$ and *itln3*$^{uta148/+}$ zebrafish (Fig. 5C–E) and the F3-progeny of WT *itln3*^{uta148} (13 CFU; SD 10 CFU) (Fig. 5C) and followed their survival up to 7 dpi. There were few dying embryos among uninfected embryos upon RC or *itln1* morpholino injection (Supplementary Figure 4), whereas the mortality reached 77.8–100% in the morpholino-*M. marinum* co-injected embryos. Noteworthy, the comparison between the infected RC and SB morpholino injected WT *itln3*^{uta145} and *itln3*^{uta148} embryos did not show any differences in survival (Fig. 5C). Moreover, inhibiting *itln1* expression in homozygous *itln3*$^{uta145/uta145}$ and *itln3*$^{uta148/uta148}$ mutants lead to a similar mortality compared to the corresponding heterozygous and WT siblings of the same genetic background (Fig. 5D,E), indicating that the simultaneous lack of *itln1* and *itln3* functionality does not affect mycobacterial resistance in the zebrafish embryo. Consistently, we did not detect any differences in the mRNA levels of *itln1*, *itln2* and *itln2-like* between the homozygous *itln3* mutants and the WT controls either in uninjected (4 dpf) or *M. marinum* (25 CFU; SD 23 CFU, 4 dpf/4 dpi) infected embryos (Supplementary Figure 5), suggesting that there is no transcriptional compensation by the other studied *intelectin* gene members in the *itln3*$^{uta145/uta145}$ and *itln3*$^{uta148/uta148}$ mutant fish. Similarly, no transcriptional compensation by *itln1*, *itln2* or *itln2-like* was observed in the adult *itln3* mutant zebrafish either in steady state or upon *M. marinum* infection (Supplementary Figure 2).

Adult *itln3* mutant zebrafish have a normal immune response towards a *M. marinum* infection. In order to test the mycobacterial susceptibility of the *itln3* mutants in adult zebrafish, we performed a low-dose (48 CFU; SD 5 CFU) mycobacterial inoculation into the abdominal cavity of the fish and followed their survival for up to 24 wpi (Fig. 6A,B). After the follow-up, an average of 67% of the *itln3*^{uta145} background zebrafish had survived, corresponding to 74% of the WT, 73% of the *itln3*$^{uta145/+}$ and 59% of the *itln3*$^{uta145/uta145}$ fish. In the *itln3*^{uta148} background fish, a combined survival percentage of 81% was observed (78% in the WT, 82% in the *itln3*$^{uta148/+}$ and 84% in the *itln3*$^{uta148/uta148}$ fish). Similarly to the embryonic survival experiments, no statistically significant differences in the survival between the genotypes were observed.

We and others have previously shown that the outcome of a mycobacterial infection in adult zebrafish depends not only on the host genotype but also on the infection dose. While, a so called low-dose inoculate can result in latency and a chronic disease[56], a higher dose leads to a fast progressing acute infection[59,71]. We hypothesized

that the effects caused by the lack of Itln3 could be more prominent in an infection with a higher mycobacterial dose. Consequently, we infected WT fish as well as heterozygous and homozygous *itln3* mutants from both the *itln3*^{uta145} and *itln3*^{uta148} backgrounds with a higher *M. marinum* dose (422 CFU; SD 221 CFU) and quantified the bacterial burden at 2 and 4 wpi (Fig. 6C,D). In these fish, we detected *M. marinum* copy number medians (log10) of 4.19 (WT *itln3*^{uta145}), 3.76 (*itln3*$^{uta145/+}$), 4.16 (*itln3*$^{uta145/uta145}$), 3.69 (WT *itln3*^{uta148}), 3.73 (*itln3*$^{uta148/+}$) and 3.92 (*itln3*$^{uta148/uta148}$) in 100 ng of zebrafish DNA at 2 wpi and 4.72 (WT *itln3*^{uta145}), 4.36 (*itln3*$^{uta145/+}$), 4.60 (*itln3*$^{uta145/uta145}$), 4.10 (*itln3*$^{uta148/+}$) and 3.72 (*itln3*$^{uta148/uta148}$) at 4 wpi. Noteworthy, no WT *itln3*^{uta148} fish were available at 4wpi for a bacterial quantification. Altogether, these data indicate that the loss of Itln3 function is dispensable for the host resistance against abdominal cavity *M. marinum* infection in adult zebrafish.

Dexamethasone mediated lymphocyte depletion in *itln3* knockout zebrafish does not affect the survival or mycobacterial burden in a *M. marinum* infection. We have recently published a zebrafish immune-suppression model for mycobacterial reactivation using orally administered dexamethasone[60]. The dexamethasone treatment decreases the total amount of lymphocytes by an average of 36% (from a relative proportion of 19.3% to 12.4%), and consequently leads to reactivation of the *M. marinum* infection. In turn, a number of studies have suggested that Itln3 functions in microbial surveillance and therefore in the innate immunity[23,24,34]. In order to highlight the importance of innate immune mechanisms in the mycobacterial defense, we used the dexamethasone treatment to specifically deplete the lymphocyte population in the adult *itln3* mutation carrying zebrafish lines *itln3*^{uta145} and *itln3*^{uta148}, and subsequently infected both WT and homozygous *itln3* mutants with *M. marinum* (47 CFU; SD 4 CFU) (Fig. 7A). Expectedly, our flow cytometric analysis demonstrated a significant decrease in the lymphocyte counts of both WT *itln3*^{uta145} and *itln3*^{uta148} fish (31.5%, $P = 0.002$ and 23.7%, $P = 0.010$, respectively) as well as the *itln3*$^{uta145/uta145}$ and *itln3*$^{uta148/uta148}$ mutants (31.5% and 40.5%, $P < 0.001$ in both comparisons) three weeks after initiating the dexamethasone administration at 2 wpi (Fig. 7B–D). In addition, neither the total cell count nor the amount of myeloid cells and blood cell precursors were affected by dexamethasone (Supplementary Figure 6). We did not detect any substantial mortality of either the *itln3*^{uta145} or the *itln3*^{uta148} mutants or WT fish during the five-week follow-up period. As is shown in the Fig. 7D,E, the bacterial amounts did not differ between the groups; in 100 ng of zebrafish DNA, mycobacterial copy number medians (log10) of 2.60 and 2.65 in WT *itln3*^{uta145}, 2.55 and 3.10 in *itln3*$^{uta145/uta145}$, 2.87 and 2.43 in WT *itln3*^{uta148} and 2.25 and 2.91 in *itln3*$^{uta148/uta148}$ zebrafish were observed at 2 and 4 wpi, respectively.

In conclusion, our data are in accordance with previous literature on the possible role for *itln*s in immunity, as the zebrafish *itln3* is highly induced in a mycobacterial infection. However, *M. marinum* infection experiments using both zebrafish embryos and adult fish suggest that *itln3* is dispensable for a protective mycobacterial host response. Moreover, *itln1* does not seem to compensate for the lack of functional *itln3* in the embryonic infection model. Of note, unlike has been reported for human ITLN1, we were unable to demonstrate direct binding of recombinant Itln3 to mycobacteria (or *S. pneumoniae* or *Escherichia coli*) *in vitro* (Supplementary Figure 7), which may explain the nonessential role of Itln3 for zebrafish immunity in our models.

Discussion

The genetics of the host affect the outcome of a *M. tuberculosis* infection, i.e. the development of active tuberculosis[4]. Genome-wide expression analyses using microarray and RNA sequencing platforms are important for understanding complicated biological processes such as the host immune defense against pathogens. To date, a handful of transcriptome studies have been done in the zebrafish *M. marinum* infection model using microarray technology[59,72,73], the digital gene expression (DGE) method[74] and RNA sequencing[75–77]. Collectively, by using both zebrafish embryos and adult fish, these studies have provided important insights into the innate and adaptive host response against mycobacterial infections.

We used the adult zebrafish *M. marinum* (ATCC 927) infection model together with a zebrafish gene expression microarray to identify novel candidate genes in a mycobacterial infection. From this data, we identified a total of 91 differentially expressed genes (log2 fold change $>|3|$) that were linked to 44 enriched processes, including genes associated with the immune response. Previous studies have shown several genes of the complement system (e.g. *complement component c3b*, *c3b*; *complement component 6*, *c6*) to be up-regulated in an infection[59,72–74,76], whereas the expression of some complement associated genes (e.g. *complement factor b*, *cfb*; *mannose binding lectin*, *mbl*) has been shown to be reduced[72,74]. In line with previous results, we also saw an induction of *cd59* (regulation of membrane attack complex formation) as well as reduced expression of *cfbl* (component of the C3 convertase). Conversely, although previous transcriptomic studies have shown the induction of genes that are involved in neutrophil and macrophage related functions (e.g. *mpx* and *irg1l*)[73,76], our data indicated down-regulation of these transcripts in an infection. In summary, the aforementioned similarities and differences between these transcriptomic studies can result from a number of factors including the developmental stage of the host (embryos vs. adult fish), the time points chosen for sample collection, the different outcomes of an infection (chronic vs. acute), the use of different bacterial strains (E11, Mma20 or ATCC 927) and doses, and they can be due to differences in the technical execution of sample preparation and analyses.

Interestingly, circa 38% of the up-regulated probes were related to muscle associated biological processes including muscle contraction (GO:0006936), muscle system process (GO:0003012) and myofibril assembly (GO:0030239). Supporting the relevance of this finding, a genome-wide expression analysis in the fruit fly *Drosophila melanogaster* identified several muscle specific genes such as *actin88F* (A*ct88F*) and *tropomycin 2* (*Tm2*) to be induced after a *Pseudomonas aeruginosa* infection[78]. Consistently, the down-regulation of muscle expressed genes (*troponin C41C, TpnC41C*; *glutathione S-Transferase 2, Gst2*) was later connected to an increased susceptibility to infection, suggesting an immunological role for muscle tissue[79,80]. Although the differential expression of muscle specific genes can be indirectly linked to the immune response through the regulation of other physiological functions, as has been also suggested by Chatterjee *et al*.,[81], both mouse and zebrafish muscle

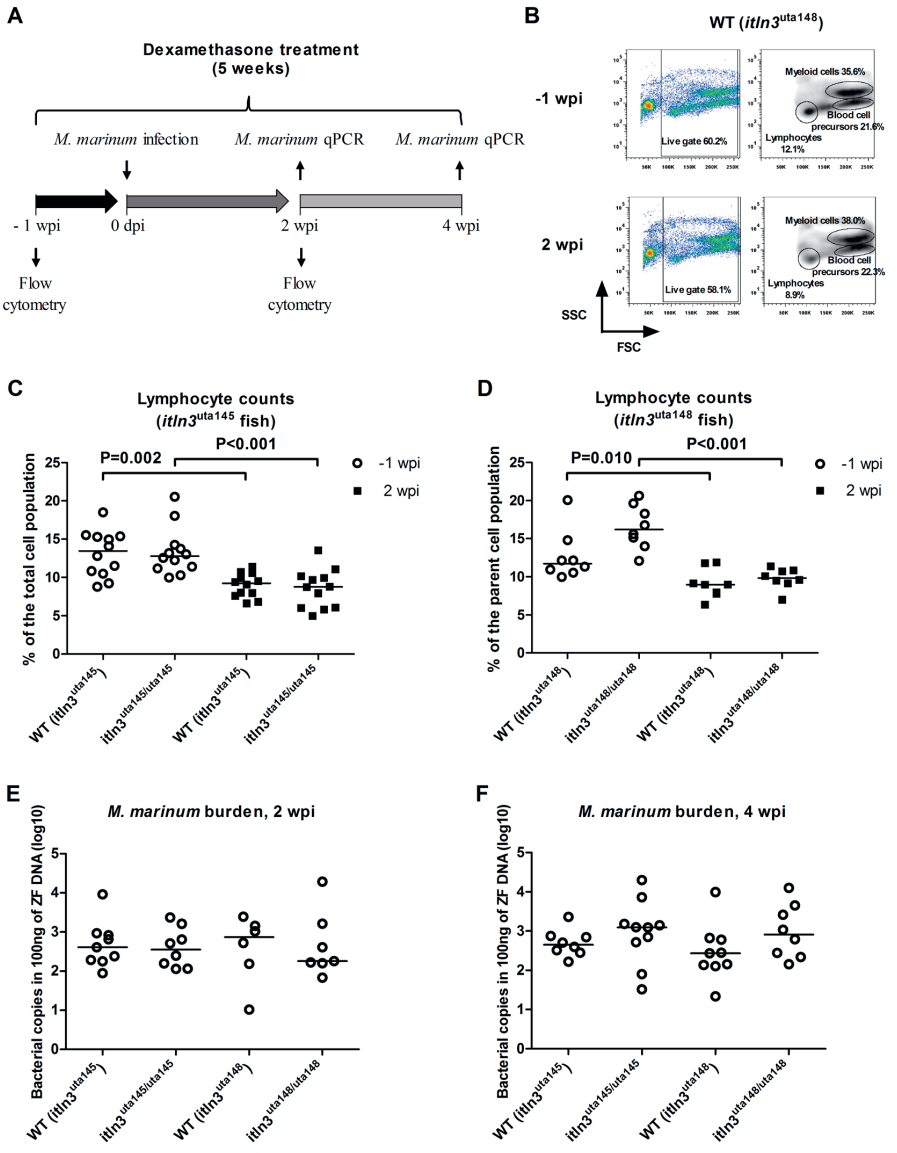


Figure 7. Dexamethasone mediated immunosuppression does not alter the survival of *itln3* deficient zebrafish in a mycobacterial infection. (**A**) A schematic representation of the performed experiment. *M. marinum* inoculate used in the infections were 47 CFU (SD 4 CFU). (**B**) Representative flow cytometry plots in WT (*itln3*^{uta148}) zebrafish at -1 wpi and 2 wpi used for quantifying lymphocyte, myeloid cell and precursor cell populations. FSC = forward scatter, SSC = side scatter. (**C,D**) Lymphocyte fractions of the total cell populations for both WT and *itln3* knockout zebrafish at -1 wpi and 2 wpi. Both the *itln3*^{uta145} (n = 12 in all groups) and *itln3*^{uta148} background fish (n = 8 in all groups) are shown. Blood cell samples were run as technical duplicates. (**E**,**F**) *M. marinum* burden (log10) in 100 ng of zebrafish DNA were measured at 2 wpi and 4 wpi by qPCR in the infected dexamethasone treated zebrafish organ blocks (without the kidney). All bacterial quantification samples were run once. Group sizes at 2 and 4 wpi, respectively, were as follows: WT (*itln3*^{uta145}) n = 9, n = 8; *iltn3*$^{uta145/uta145}$ n = 9, n = 10; WT (*itln3*^{uta148}) n = 6, n = 9 and *iltn3*$^{uta148/uta148}$ n = 7, n = 8. A two-tailed Mann-Whitney test was used in the statistical comparison of differences.

tissues have also been reported to control the expression of inflammatory cytokine genes (*Tnfa* and *Il6*) upon activation of the immune response[82,83]. In addition, relatively recent studies in the fruit fly and zebrafish have confirmed the importance of immunological signaling pathways in the muscle in both the humoral[81] and the

cellular[80] immune responses against pathogens. Further studies are required to understand the significance of muscle expressed genes also in the host response against mycobacterial infections.

Lee *et al*.,[21] described a new carbohydrate-binding lectin family, the Intelectins, with concomitantly proposed function in the innate immunity through microbial recognition[23,34,83]. Two *ITLN* genes (*ITLN1* and *ITLN2*) have been identified in humans, whereas the exact number of protein coding *itln* genes in zebrafish is elusive varying between six (The Zebrafish Information Network, ZFIN) and nine annotated members (Ensembl genome browser). However, not all are expressed in significant amounts. Our genome-wide gene expression analysis showed the highest expression levels for *itln1*, *itln2, itln2-like* and *itln3* in the PBS injected fish with an average log2 expression of 9.6, 11.9, 13.0 and 10.5, respectively (the lowest average log2 expression was 6.9 for a probe A_15_P113269; *ankyrin repeat and kinase domain containing 1*, *ankk1*). This was consistent with Lin *et al*., who reported the highest expression levels for *itln1*, *itln2* and *itln3*[25]. Since their discovery, a large number of studies have reported the induction of the expression of *ITLN* genes in different species in bacterial[25,30–32], parasite[84–86] and viral[87] infections. In addition, although the publicly available transcriptomic data of *M. marinum* infected zebrafish embryos by Benard *et al*., (Gene Expression Omnibus identifier: GSE76499) shows differential regulation of *itln* transcripts in this model[58], to our knowledge *itln* up-regulation in a mycobacterial infection has not previously been extensively reported in the literature. In our microarray analysis of *M. marinum* infected adult zebrafish, three of the differentially expressed genes were members of the *intelectin* family (*itln2, itln2-like* and *itln3*). The observed decrease in *itln2* expression as well as the induction of *itln3* in these fish was also later confirmed by a qPCR-based quantification of samples from a separate mycobacterial infection experiment. In line with this, a qPCR analysis of *M. marinum* infected zebrafish embryos revealed a significant induction of *itln3* expression post mycobacterial infection, a result which corresponds well with the data by Benard *et al*., (GSE76499)[58]. In embryos, we also observed up-regulation of *itln1*, *itln2* as well as an *itln2-like* in response to a mycobacterial infection. While *itln1* had identical induction kinetics to *itln3* with significant up-regulation starting at 2 dpi, the *itln2-like* and *itln2* genes were induced later in the infection at 4 and 7 dpi, respectively. Of note, the expression of *itln2* at 1 dpi (Fig. 1G) was below the limit of detection in the qPCR. Furthermore, up-regulation of *itln3* was observed after an *S. pneumoniae* infection in zebrafish embryos, suggesting that this gene can be induced also in an immune response against pneumococcus. All in all, in this study we demonstrated the inducibility of *intelectin* genes in both a mycobacterial and a *S. pneumoniae* infection as well as the down-regulation of *itln2* and *itln2-like* transcripts after mycobacterial inoculation in adult zebrafish.

Transcriptomic analyzes have identified several so called classical liver-expressed acute phase protein (APP) genes such as *c-reactive protein* (*crp*) and *serum amyloid a* (*saa*) also in fish species[88–90]. In addition, bacterial infections in rainbow trout (*Oncorhynchus mykiss*)[91], channel catfish (*Ictalurus punctatus*)[31] as well as in zebrafish[25] have resulted in liver-specific induction of certain *itln* gene members. Our mRNA expression analysis of unchallenged zebrafish confirmed the previously published tissue-restricted expression pattern of zebrafish *itln* genes[25]. While *itln2* is expressed almost exclusively in the intestine, the highest relative expression compared to the housekeeping gene was in the liver for *itln3*. In the current study, we produced Strep-tagged® zebrafish Itln3 in a mammalian expression system to study whether Itln3 could act as a potential APP in microbial recognition. Similarly to human ITLN1[34,83,92], zebrafish Itln3 was secreted into the culture media. However, although Tsuji *et al*., (2009) has reported the ability of human ITLN1 to bind to galactofuranosyl (Galf) residues on the mycobacterial cell membrane[83], recombinant Itln3 did not bind readily to *M. marinum in vitro*. Similarly, Itln3 did not bind to *E. coli* or *S. pneumoniae* in our hands.

To our knowledge, only two *in vivo* studies on the significance of Intelectins in the host response against pathogens have been conducted[35,36]. While the lack of ITLN2 was associated with an increased susceptibility toward the parasite *T. spiralis* in C57BL/10 mice[36], the over-expression of ITLN1 and ITLN2 in the lungs of transgenic mice could not restrict either a *Nippostrongylus brasiliensis* or a *Mycobacterium tuberculosis* infection differently from the littermate controls[35]. While genes of the innate immune system can be studied autonomously in zebrafish embryos, which lack a functional adaptive immunity[50], adult zebrafish have a highly similar immune system compared to humans[93]. Correspondingly, the embryonic *M. marinum* infection model has revealed important insights into the mechanisms of the innate immunity in mycobacterial host resistance (reviewed in[50,51]) and the adult model has proven its usefulness e.g. in modelling a latent infection[56] and disease reactivation[56,60]. In order to obtain a more comprehensive view about the *in vivo* significance of ITLNs in a mycobacterial infection, we used the *itln3* deficient zebrafish together with both the embryonic as well as the adult *M. marinum* models. In these experiments, the comparison of survival and mycobacterial burden between *itln3* mutant fish and their WT siblings did not reveal any differences in either of the mutant lines (*itln3*^{uta145} and *itln3*^{uta148}). Also, it was demonstrated that this was independent of the site or timing of the microinjection in the embryos (yolk sac at 0 dpf vs. blood circulation valley at 2 dpf). Collectively, we conclude that zebrafish *itln3* is not required for the resistance against a mycobacterial infection.

Genetic compensation is a well-known phenomenon in model organisms with experimental gene knockouts[70]. In this process, the specific function of a knockout gene can be restored by additional naturally occurring mutations or transcriptional changes in other genes[70]. Here, we report the up-regulation of *itln3* as well as another member of the *intelectin* family, *itln1*, in a *M. marinum* infection of zebrafish embryos with analogous induction kinetics. To overcome potential compensatory effects of *itln1* in our *itln3* mutants, we knocked down *itln1* by morpholinos and performed simultaneous *M. marinum* infections in the *itln3* mutant embryos. Silencing *itln1* in *itln3* mutants during a *M. marinum* infection did not reveal any differences compared to controls, demonstrating that *itln1* expression does not compensate for the lack of a functional *itln3* in a *M. marinum* infection. Moreover, our qPCR quantification of *itln1*, *itln2* and *itln2-like* in the uninjected and *M. marinum* infected zebrafish embryos as well as adult zebrafish did not reveal transcriptional compensation for *itln3* in the homozygous mutant background. Similarly, the depletion of lymphocytes in the adult zebrafish did not reveal the importance for Itln3 in the immunity against a *M. marinum* infection. All in all, our data indicate that despite being strongly

induced by a mycobacterial infection, *itln3* is dispensable for the immune response against *M. marinum* both in embryonic and adult zebrafish.

Methods

Zebrafish lines and maintenance. The zebrafish maintenance and all of the experiments were in accordance with the Finnish Act on the Protection of Animals Used for Scientific or Educational Purposes (497/2013) as well as the EU Directive on the Protection of Animals Used for Scientific Purposes (2010/63/EU). Experiments were approved by the Animal Experiment Board of Finland (permit for zebrafish maintenance: ESAVI/10079/04.10.06/2015; permits for the experiments: ESAVI/2464/04.10.07/2017, ESAVI/10823/04.10.07/2016, ESAVI/2235/04.10.07/2015 and ESAVI/11133/04.10.07/2017). WT AB fish as well as in-house CRISPR/Cas9 produced F2-generation *itln3*^{uta145} and F2- or F3-generation *itln3*^{uta148} mutant zebrafish were used in the embryonic experiments, whereas three- to seven-month-old AB, *il10*^{e46}, *itln3*^{uta145} and *itln3*^{uta148} zebrafish were used in the experiments with the adult fish. Zebrafish embryos were maintained according to standard protocols in embryonic medium E3 (5 mM NaCl, 0.17 mM KCl, 0.33 mM $CaCl_2$, 0.33 mM $MgSO_4$, 0.0003 g/l methylene blue) at 28.5 °C until 7 dpf. Maintenance of the adult zebrafish was as follows; unchallenged fish were kept in a conventional flow through system (Aquatic Habitats, Florida, USA) with an automated light/dark cycle of 14 h/10 h and fed once a day with Gemma Micro 500 (Skretting, Stavanger, Norway) or twice with SDS 400 (Special Diets Services, Essex, UK) feed. *M. marinum* infected adults were kept in a separate flow through system (Aqua Schwarz GMbH, Göttingen, Germany) with the above-mentioned light/dark cycle and fed once a day with Gemma Micro 500 (Skretting) or SDS 400 (Special Diets Services). Infected fish were monitored daily. Humane endpoint criteria pre-defined in the animal experiment permits were applied throughout the follow-up.

Experimental *M. marinum* infections. *M. marinum* (ATCC 927 -strain) culture and the adult zebrafish inoculations were performed as described previously[56]. In the *M. marinum* infections of the zebrafish embryos, a total volume of 1–2 nl was microinjected either into the yolk sac (0 dpf) or into the blood circulation valley (2 dpf) by using a borosilicate capillary needle (Sutter instrument Co., California, USA), a micromanipulator (Narishige International, London UK) and a PV830 Pneumatic PicoPump (World Precision Instruments, Florida, USA). 10 mM phosphate buffered saline (PBS) with 2% polyvinylpyrrolidone-40 (PVP) (Sigma-Aldrich) and 0.3 mg/ml phenol red (Sigma-Aldrich) was used as a mycobacterial carrier solution. Prior to circulation valley injections, the 2 dpf zebrafish were anesthetized with 0.02% 3-amino benzoic acid ethyl ester (Sigma-Aldrich). Embryonic infections were visualized with a Stemi 2000 microscope (Carl Zeiss MicroImaging GmbH, Göttingen, Germany) and the survival of the embryos followed daily. Adult zebrafish were first anesthetized with 0.02% 3-amino benzoic acid ethyl ester (Sigma-Aldrich, Missouri, USA), and then injected with 5 µl of *M. marinum* in a suspension of 10 mM PBS and 0.3 mg/ml phenol red (Sigma-Aldrich, Missouri, USA) into the abdominal cavity using a 30 gauge Omnican 100 insulin needle (Braun, Melsungen, Germany). The *M. marinum* amounts from both the embryonic and adult infections were verified by plating bacterial inoculates on 7H10 agar (Becton Dickinson, New Jersey, USA) plates and counting the colony forming units (CFU) 5-days after plating.

Gene expression microarray. RNA was extracted from the zebrafish organ blocks (includes all the organs of the abdominal cavity) with TRIreagent (Molecular Research Center, Ohio, USA) following the manufacturer's protocol. Microarray procedures were carried out by the Turku Centre for Biotechnology at the Finnish Microarray and Sequencing Centre by using a Zebrafish (V3) Gene Expression Microarray, 4 × 44 K (Agilent Technologies, California, USA). In short, 100 ng of total RNA was amplified and Cy3-labeled with Low Input Quick Amp Labeling kit, one-color (Agilent), processed using the RNA Spike-In Kit, one-color (Agilent) and quality controlled with 2100 bioanalyzer RNA 6000 Nano kit (Agilent). Labelling and hybridization of the transcripts were done onto Agilent's 4 × 44 K Zebrafish v3 array (Design ID 026437) using GE Hybridization Kit (Agilent). Microarrays were scanned with an Agilent scanner G2565CA using a profile AgilentHD_GX_1Color. Numerical results were obtained with Feature Extraction Software v. 10.7.3 (Agilent) with the protocol GE1_107_Sep09 and the signal intensities normalized prior further analysis. Cut-off value (log2 fold change $>|3|$) for the up- and down-regulated genes was chosen in order to obtain approximately 100 differentially expressed candidate genes for further evaluation. Gene ontology enrichment analysis was performed using GOrilla[61,62] with two unranked lists of genes (Target list: log2 fold change $>|3|$, Background list: log2 fold change $<|3|$) using *Danio rerio* genome assembly.

qPCR. For gene expression analysis of the zebrafish embryo samples, genomic DNA (gDNA) removal and RNA isolation were performed using the RNeasy Plus Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's guidelines. Adult zebrafish RNA was extracted from the organ blocks, liver, spleen, kidney and intestine with TRIreagent (Molecular Research Center) following the associated protocol. The genomic DNA (gDNA) from the RNA samples of the adult fish was removed with RapidOut DNA Removal Kit (Thermo Fischer Scientific, Waltham, USA). RNA quality was controlled with either 1.5% agarose Tris-acetate-EDTA (TAE) gel electrophoresis or by using Fragment Analyzer system (Advanced Analytical, Inc., Ankeny, USA) and the Standard Sensitivity RNA Analysis Kit (15 nt) (Advanced Analytical). All reverse transcriptions were done by using the SensiFAST™ cDNA synthesis kit (BioLine, London, UK), and the gene expression levels of the target genes were determined from the cDNA with quantitative PCR (qPCR) using the PowerUp™ SYBR® master mix (Thermo Fischer Scientific) and CFX96™ detection system (Bio-Rad Laboratories, California, USA). CFX Manager software (v. 3.1; Bio-Rad Laboratories) was used for data analysis. Target gene expression was normalized to the *eukaryotic translation elongation factor 1 alpha 1, like 1* (*eef1a1l1 or ef1a*)[94] expression using the $2^{-\Delta Ct}$ method. *M. marinum* burden from the zebrafish was determined from the total DNA by qPCR with *MMITS*-specific primers[56]. Embryo DNA for mycobacterial quantification was isolated with standard ethanol

precipitation procedure utilizing the following lysis buffer: 10 mM Tris (pH 8.2), 10 mM EDTA, 200 mM NaCl, 0.5% SDS, 200 µg/ml Proteinase K (Thermo Fischer Scientific), whereas TRIreagent (Molecular Research Center) was used for the adult fish DNA isolations. No reverse transcriptase control samples were added to the gene expression analyses, and no template control (H_2O) samples were included in all of the qPCR experiments to preclude contamination. Specificity and the correct size of the qPCR products were verified by melt curve analysis and 1.5% agarose TAE gel electrophoresis. Undetectable qPCR products with incorrect melt curves were given a Ct-value of 40 for the gene expression analyses, and the expression was considered to be below detection. qPCR primers used in the study are listed in Supplementary Table 3.

CRISPR/Cas9 mutagenesis. We have previously set-up our in-house zebrafish CRISPR/Cas9 mutagenesis method based on the protocol published by Hruscha and Schmid (2015)[65,95]. First, guide RNA (gRNA) target sequences for *itln3* were designed with the CRISPR design tool (http://crispr.mit.edu/), and validated with the Casellas laboratory sgRNA tool[96] and the standard nucleotide BLAST analysis[97]. *itln3* exon 2 gRNA was produced by *in vitro* transcription using the MEGAshortscript T7 Transcription Kit (Ambion Life Technologies, CA, USA). 2000 pg of gRNA, 330 pg of *cas9* mRNA (Sigma-Aldrich and Invitrogen, California, USA) and 1.5 ng of phenol red (Sigma-Aldrich) tracer were injected into one-cell-stage AB zebrafish embryos, and the success of mutagenesis was evaluated with the T7 endonuclease I (T7EI)- and the heteroduplex mobility assay (HMA) from isolated DNA of 2 dpf embryos[49]. Gel images were obtained with ChemiDoc™ XRS+ system (Bio-Rad Laboratories) and analyzed with Image Lab software (v. 5.2; Bio-Rad Laboratories). To establish the *itln3* knockout fish line, gRNA was microinjected into zebrafish embryos and the F0-generation fish grown to adulthood. Individual outcrosses of the F0-zebrafish with the Tupfel long fin (TL) fish allowed us to screen for germline transmitted mutations and to identify nonsense mutations of interest in the F1-progeny. The F1-progeny screen was done from the tailfin DNA of the adult zebrafish with HMA followed by Sanger sequencing in our institutes core-facility (MED, University of Tampere). The F1-zebrafish carrying individual mutations of interest were spawn together to obtain F2-generation progeny for the experiments. In the end, a total of two different *itln3* mutation (*itln3*^{uta145} and *itln3*^{uta148}) bearing zebrafish lines were used in the study.

PCR based genotyping. F2-generation *itln3*^{uta145} and *itln3*^{uta148} zebrafish lines were mainly genotyped using PCR. To this end, template DNA was either isolated using a standard ethanol precipitation protocol, or with a rapid tissue lysis protocol[98]. Primers were designed for both the WT and the mutated sequences at the gRNA target region; WT *itln3*^{uta145} F: 5′-ATGCTAGGTTGAGGAGCATC-3′, mutant *itln3*^{uta145} F: 5′-ATGCTAGGTTGAGGAGCTCG-3′, WT *itln3*^{uta148} F: 5′-CTAGGTTGAGGAGCATCGCT-3′, mutant *itln3*^{uta148} R: 5′-CCGAGCTGATACTTACCTAGC-3′, and amplified with the appropriate flanking primer: F: 5′-GGAGCTGTCACTCCAAAGCC-3′ or R: 5′-GTGGTTGATCAACCATTCAGCAC-3′. To determine the genotypes of the *itln3*^{uta145} and *itln3*^{uta148} zebrafish, individual PCR reactions with both WT and mutant primer pairs were prepared for each fish and 1.5% agarose TAE gel electrophoresis was performed to analyze the PCR products.

Morpholino injections. Splice-blocking morpholino (SB) for *itln1* (5′-CTAATTCTGTACTTACTC GATTCAC-3′) was designed by and ordered from GeneTools, LLC (Philomath, Oregon, USA). The targeted genomic sequence was verified from our AB and *itln3* knockout zebrafish lines by sequencing[99]. In order to ensure no adverse effects on survival or the phenotype of the morpholino injected embryos in later experiments, the oligonucleotide dosage was first titrated by using three different quantities (7.1 ng, 2.8 ng and 1.1 ng), and the survival of the embryos was observed daily until 7 dpf. The embryos were imaged using Zeiss Lumar V12 fluorescence microscope. The selected microinjection volume was set to 2 nl containing 2.8 ng of SB or random control (RC) morpholino as well as 7 mg/ml of tetramethylrhodamine dextran (Thermo Fisher Scientific) or 0.3 mg/ml phenol red tracer suspended in PBS. In the morpholino-*M. marinum* co-injections, the previously described suspension with 2% PVP was used as a mycobacterial carrier solution. All of the morpholino injections as well as morpholino and *M. marinum* co-injections were done before the 16-cell-stage of development into the yolk sac of the embryos. Similarly than in the other *M. marinum* infection experiments, the mycobacterial counts in the injections were verified by plating. Primers used for the morpholino target site sequencing were F: 5′-TGCACAGGTATTCACCATTTTATGATG-3′ and R: 5′-AAGTTCTCTGCAGCTTCTTGC-3′ and for the verification of the morpholino functionality as well as quantification of the WT *itln1* expression by qPCR: F: 5′-ATGATGCAGTCAGCTGGTTTTCTTCTG-3′ and R: 5′-GCAGTGACCGACTCTGGAAATTCTCC-3′.

Flow cytometry. Flow cytometry for the adult zebrafish kidney cells was performed as described previously[71]. Briefly, *itln3*^{uta145} and *itln3*^{uta148} fish were euthanized with 0.04% 3-amino benzoic acid ethyl ester and their kidneys isolated and suspended in PBS supplemented with 0.5% fetal bovine serum (Sigma-Aldrich). Prior analysis, the kidney cells were filtered through a cell strainer cap with a 35 µm mesh (Corning/Thermo Fisher Scientific). Relative amounts of lymphocytes, myeloid cells and blood cell precursors were determined with a FACSCanto II instrument (Becton, Dickinson, New Jersey, USA) and the data was analyzed with the FlowJo program (v. 7.5; Tree Star, Inc, Oregon, USA). Gating of the blood cell populations is based on previous publications[60,71,100,101].

Dexamethasone mediated immunosuppression. Similarly as described previously[60], 25 mg of dexamethasone (Sigma-Aldrich) was mixed with gelatin (Sigma-Aldrich) and used to coat 10 g of SDS400 food (Special Diets Services). During the experiment, a daily dose of 10 µg of dexamethasone (4 mg of food) was given per fish for a total of 5 weeks. A new batch of dexamethasone food was prepared for usage every second week. Dexamethasone was administered for a total of five weeks and the well-being of the fish monitored daily.

Statistical analysis. Sample size calculations have been described in our previous publication[100]. Statistical analyses were done with the Prism v. 5.02 (GraphPad Software, California, USA). In the survival experiments a log-rank (Mantel-Cox) test was used, whereas in the flow cytometry and qPCR analyses a nonparametric two-tailed Mann-Whitney analysis was performed. P values of < 0.05 were considered significant.

Data Availability

Gene expression microarray data has been submitted to Gene Expression Omnibus (GEO) repository and can be found with the identifier code: GSE120552. Other generated and analyzed data is available on reasonable request from the corresponding author.

References

1. World Health Organization. Global Tuberculosis Report. http://www.who.int/tb/publications/global_report/en/ (2017).
2. Lerner, T. R., Borel, S. & Gutierrez, M. G. The innate immune response in human tuberculosis. *Cellular Microbiology* **17**, 1277–1285 (2015).
3. Möller, M., de Wit, E. & Hoal, E. G. Past, present and future directions in human genetic susceptibility to tuberculosis. *FEMS Immunol. Med. Microbiol.* **58**, 3–26 (2010).
4. Meyer, C. G. & Thye, T. Host genetic studies in adult pulmonary tuberculosis. *Semin. Immunol.* **26**, 445–453 (2014).
5. Shimokata, K., Kawachi, H., Kishimoto, H., Maeda, F. & Ito, Y. Local cellular immunity in tuberculous pleurisy. *Am. Rev. Respir. Dis.* **126**, 822–824 (1982).
6. Onwubalili, J. K., Scott, G. M. & Robinson, J. A. Deficient immune interferon production in tuberculosis. *Clin. Exp. Immunol.* **59**, 405–413 (1985).
7. Denis, M. Interleukin-12 (IL-12) augments cytolytic activity of natural killer cells toward Mycobacterium tuberculosis-infected human monocytes. *Cell. Immunol.* **156**, 529–536 (1994).
8. Zhang, M. *et al*. Interleukin 12 at the site of disease in tuberculosis. *J. Clin. Invest.* **93**, 1733–1739 (1994).
9. Alcaïs, A., Fieschi, C., Abel, L. & Casanova, J. Tuberculosis in children and adults: two distinct genetic diseases. *J. Exp. Med.* **202**, 1617–1621 (2005).
10. Flynn, J. L. *et al*. An essential role for interferon gamma in resistance to Mycobacterium tuberculosis infection. *J. Exp. Med.* **178**, 2249–2254 (1993).
11. Cooper, A. M. *et al*. Disseminated tuberculosis in interferon gamma gene-disrupted mice. *J. Exp. Med.* **178**, 2243–2247 (1993).
12. Flynn, J. L. *et al*. IL-12 increases resistance of BALB/c mice to Mycobacterium tuberculosis infection. *J. Immunol.* **155**, 2515–2524 (1995).
13. Cooper, A. M., Magram, J., Ferrante, J. & Orme, I. M. Interleukin 12 (IL-12) is crucial to the development of protective immunity in mice intravenously infected with mycobacterium tuberculosis. *J. Exp. Med.* **186**, 39–45 (1997).
14. Dittrich, N. *et al*. Toll-like receptor 1 variations influence susceptibility and immune response to Mycobacterium tuberculosis. *Tuberculosis (Edinb)* **95**, 328–335 (2015).
15. Stappers, M. H. T. *et al*. TLR1, TLR2, and TLR6 gene polymorphisms are associated with increased susceptibility to complicated skin and skin structure infections. *J. Infect. Dis.* **210**, 311–318 (2014).
16. Qi, H. *et al*. Toll-like receptor 1(TLR1) Gene SNP rs5743618 is associated with increased risk for tuberculosis in Han Chinese children. *Tuberculosis (Edinb)* **95**, 197–203 (2015).
17. Tanne, A. *et al*. A murine DC-SIGN homologue contributes to early host defense against Mycobacterium tuberculosis. *J. Exp. Med.* **206**, 2205–2220 (2009).
18. Guo, Y., Liu, Y., Ban, W., Sun, Q. & Shi, G. Association of mannose-binding lectin gene polymorphisms with the development of pulmonary tuberculosis in China. *BMC Infect. Dis.* **17**, 210 (2017).
19. Wesener, D. A., Dugan, A. & Kiessling, L. L. Recognition of microbial glycans by soluble human lectins. *Curr. Opin. Struct. Biol.* **44**, 168–178 (2017).
20. Sharon, N. & Lis, H. History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology* **14**, 53R–62R (2004).
21. Lee, J. K. *et al*. Cloning and expression of a Xenopus laevis oocyte lectin and characterization of its mRNA levels during early development. *Glycobiology* **7**, 367–372 (1997).
22. Yan, J. *et al*. Comparative genomic and phylogenetic analyses of the intelectin gene family: implications for their origin and evolution. *Dev. Comp. Immunol.* **41**, 189–199 (2013).
23. Komiya, T., Tanigawa, Y. & Hirohashi, S. Cloning of the novel gene intelectin, which is expressed in intestinal paneth cells in mice. *Biochem. Biophys. Res. Commun.* **251**, 759–762 (1998).
24. Tsuji, S. *et al*. Human intelectin is a novel soluble lectin that recognizes galactofuranose in carbohydrate chains of bacterial cell wall. *J. Biol. Chem.* **276**, 23456–23463 (2001).
25. Lin, B. *et al*. Characterization and comparative analyses of zebrafish intelectins: highly conserved sequences, diversified structures and functions. *Fish Shellfish Immunol.* **26**, 396–405 (2009).
26. Suzuki, Y. A., Shin, K. & Lönnerdal, B. Molecular cloning and functional expression of a human intestinal lactoferrin receptor. *Biochemistry* **40**, 15771–15779 (2001).
27. Watanabe, T., Watanabe-Kominato, K., Takahashi, Y., Kojima, M. & Watanabe, R. Adipose Tissue-Derived Omentin-1 Function and Regulation. *Compr Physiol* **7**, 765–781 (2017).
28. Li, D. *et al*. Intelectin 1 suppresses the growth, invasion and metastasis of neuroblastoma cells through up-regulation of N-myc downstream regulated gene 2. *Mol. Cancer* **14**, 47 (2015).
29. Li, D. *et al*. Intelectin 1 suppresses tumor progression and is associated with improved survival in gastric cancer. *Oncotarget* **6**, 16168–16182 (2015).
30. Ding, Z. *et al*. Characterization and expression analysis of an intelectin gene from Megalobrama amblycephala with excellent bacterial binding and agglutination activity. *Fish Shellfish Immunol.* **61**, 100–110 (2017).
31. Peatman, E. *et al*. Expression analysis of the acute phase response in channel catfish (Ictalurus punctatus) after infection with a Gram-negative bacterium. *Dev. Comp. Immunol.* **31**, 1183–1196 (2007).
32. Takano, T. *et al*. The two channel catfish intelectin genes exhibit highly differential patterns of tissue expression and regulation after infection with Edwardsiella ictaluri. *Dev. Comp. Immunol.* **32**, 693–705 (2008).
33. Tsuji, S. *et al*. Capture of heat-killed Mycobacterium bovis bacillus Calmette-Guérin by intelectin-1 deposited on cell surfaces. *Glycobiology* **19**, 518–526 (2009).
34. Wesener, D. A. *et al*. Recognition of microbial glycans by human intelectin-1. *Nat. Struct. Mol. Biol.* **22**, 603–610 (2015).
35. Voehringer, D. *et al*. Nippostrongylus brasiliensis: identification of intelectin-1 and -2 as Stat6-dependent genes expressed in lung and intestine during infection. *Exp. Parasitol.* **116**, 458–466 (2007).
36. Pemberton, A. D. *et al*. Innate BALB/c enteric epithelial responses to Trichinella spiralis: inducible expression of a novel goblet cell lectin, intelectin-2, and its natural deletion in C57BL/10 mice. *J. Immunol.* **173**, 1894–1901 (2004).

37. Rao, S. *et al*. Omentin-1 prevents inflammation-induced osteoporosis by downregulating the pro-inflammatory cytokines. *Bone Research* **6**, 1–12 (2018).
38. Howe, K. *et al*. The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**, 498–503 (2013).
39. Willett, C. E., Zapata, A. G., Hopkins, N. & Steiner, L. A. Expression of zebrafish rag genes during early development identifies the thymus. *Dev. Biol.* **182**, 331–341 (1997).
40. Willett, C. E., Kawasaki, H., Amemiya, C. T., Lin, S. & Steiner, L. A. Ikaros expression as a marker for lymphoid progenitors during zebrafish development. *Dev. Dyn.* **222**, 694–698 (2001).
41. Danilova, N. & Steiner, L. A. B cells develop in the zebrafish pancreas. *Proc. Natl. Acad. Sci. USA* **99**, 13711–13716 (2002).
42. Bennett, C. M. *et al*. Myelopoiesis in the zebrafish, Danio rerio. *Blood* **98**, 643–651 (2001).
43. Lin, A. F. *et al*. The DC-SIGN of zebrafish: insights into the existence of a CD209 homologue in a lower vertebrate and its involvement in adaptive immunity. *J. Immunol.* **183**, 7398–7410 (2009).
44. Seeger, A., Mayer, W. E. & Klein, J. A complement factor B-like cDNA clone from the zebrafish (Brachydanio rerio). *Mol. Immunol.* **33**, 511–520 (1996).
45. Lam, S. H., Chua, H. L., Gong, Z., Lam, T. J. & Sin, Y. M. Development and maturation of the immune system in zebrafish, Danio rerio: a gene expression profiling, *in situ* hybridization and immunological study. *Dev. Comp. Immunol.* **28**, 9–28 (2004).
46. Danilova, N., Bussmann, J., Jekosch, K. & Steiner, L. A. The immunoglobulin heavy-chain locus in zebrafish: identification and expression of a previously unknown isotype, immunoglobulin Z. *Nat. Immunol.* **6**, 295–302 (2005).
47. Hwang, W. Y. *et al*. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–229 (2013).
48. Chang, N. *et al*. Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res.* **23**, 465–472 (2013).
49. Uusi-Mäkelä, M. I. E. *et al*. Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (Danio rerio). *PLoS ONE* **13**, e0196238 (2018).
50. Meijer, A. H. Protection and pathology in TB: learning from the zebrafish model. *Semin Immunopathol* **38**, 261–273 (2016).
51. Myllymäki, H., Bäuerlein, C. A. & Rämet, M. The Zebrafish Breathes New Life into the Study of Tuberculosis. *Front Immunol* **7**, 196 (2016).
52. van Leeuwen, L. M., van der Sar, Astrid, M. & Bitter, W. Animal models of tuberculosis: zebrafish. *Cold Spring Harbor perspectives in medicine* **5**, a018580 (2015).
53. Myllymäki, H., Niskanen, M., Oksanen, K. E. & Rämet, M. Animal models in tuberculosis research - where is the beef? *Expert Opin Drug Discov* **10**, 871–883 (2015).
54. Davis, J. M. *et al*. Real-Time Visualization of Mycobacterium-Macrophage Interactions Leading to Initiation of Granuloma Formation in Zebrafish Embryos. *Immunity* **17**, 693–702 (2002).
55. Swaim, L. E. *et al*. Mycobacterium marinum infection of adult zebrafish causes caseating granulomatous tuberculosis and is moderated by adaptive immunity. *Infect. Immun.* **74**, 6108–6117 (2006).
56. Parikka, M. *et al*. Mycobacterium marinum causes a latent infection that can be reactivated by gamma irradiation in adult zebrafish. *PLoS Pathog.* **8**, e1002944 (2012).
57. Meijer, A. H. & Spaink, H. P. Host-pathogen interactions made transparent with the zebrafish model. *Curr Drug Targets* **12**, 1000–1017 (2011).
58. Benard, E. L., Rougeot, J., Racz, P. I., Spaink, H. P. & Meijer, A. H. Transcriptomic Approaches in the Zebrafish Model for Tuberculosis-Insights Into Host- and Pathogen-specific Determinants of the Innate Immune Response. *Adv. Genet.* **95**, 217–251 (2016).
59. van der Sar *et al*. Specificity of the zebrafish host transcriptome response to acute and chronic mycobacterial infection and the role of innate and adaptive immune components. *Mol. Immunol* **46**, 2317–2332 (2009).
60. Myllymäki, H., Niskanen, M., Luukinen, H., Parikka, M. & Rämet, M. Identification of protective postexposure mycobacterial vaccine antigens using an immunosuppression-based reactivation model in the zebrafish. *Dis Model Mech* **11**, (2018).
61. Eden, E., Lipson, D., Yogev, S. & Yakhini, Z. Discovering motifs in ranked lists of DNA sequences. *PLoS Comput. Biol.* **3**, e39 (2007).
62. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10**, 48 (2009).
63. Jinek, M. *et al*. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
64. Ran, F. A. *et al*. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281–2308 (2013).
65. Aspatwar, A. *et al*. Inactivation of ca10a and ca10b Genes Leads to Abnormal Embryonic Development and Alters Movement Pattern in Zebrafish. *PLoS ONE* **10**, e0134263 (2015).
66. Hruscha, A. *et al*. Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development* **140**, 4982–4987 (2013).
67. Artimo, P. *et al*. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* **40**, 597 (2012).
68. Nickless, A., Bailis, J. M. & You, Z. Control of gene expression through the nonsense-mediated RNA decay pathway. *Cell Biosci* **7**, 26 (2017).
69. Rounioja, S. *et al*. Defense of zebrafish embryos against Streptococcus pneumoniae infection is dependent on the phagocytic activity of leukocytes. *Dev. Comp. Immunol.* **36**, 342–348 (2012).
70. Tautz, D. Problems and paradigms: Redundancies, development and the flow of information. *BioEssays* **14**, 263–266 (1992).
71. Ojanen, M. J. T. *et al*. The proprotein convertase subtilisin/kexin furinA regulates zebrafish host response against Mycobacterium marinum. *Infect. Immun.* **83**, 1431–1442 (2015).
72. Meijer, A. H. *et al*. Transcriptome profiling of adult zebrafish at the late stage of chronic tuberculosis due to Mycobacterium marinum infection. *Mol. Immunol.* **42**, 1185–1203 (2005).
73. van der Vaart, M., Spaink, H. P. & Meijer, A. H. Pathogen recognition and activation of the innate immune response in zebrafish. *Adv Hematol* **2012**, 159807 (2012).
74. Hegedus, Z. *et al*. Deep sequencing of the zebrafish transcriptome response to mycobacterium infection. *Mol. Immunol.* **46**, 2918–2930 (2009).
75. Rougeot, J. *et al*. RNA sequencing of FACS-sorted immune cell populations from zebrafish infection models to identify cell specific responses to intracellular pathogens. *Methods Mol. Biol.* **1197**, 261–274 (2014).
76. Benard, E. L., Rougeot, J., Racz, P. I., Spaink, H. P. & Meijer, A. H. Transcriptomic Approaches in the Zebrafish Model for Tuberculosis-Insights Into Host- and Pathogen-specific Determinants of the Innate Immune Response. *Adv. Genet.* **95**, 217–251 (2016).
77. Kenyon, A. *et al*. Active nuclear transcriptome analysis reveals inflammasome-dependent mechanism for early neutrophil response to Mycobacterium marinum. *Sci Rep* **7**, 6505 (2017).
78. Apidianakis, Y. *et al*. Profiling early infection responses: Pseudomonas aeruginosa eludes host defenses by suppressing antimicrobial peptide gene expression. *Proc. Natl. Acad. Sci. USA* **102**, 2573–2578 (2005).
79. Apidianakis, Y. *et al*. Involvement of skeletal muscle gene regulatory network in susceptibility to wound infection following trauma. *PLoS ONE* **2**, e1356 (2007).
80. Yang, H., Kronhamn, J., Ekström, J., Korkut, G. G. & Hultmark, D. JAK/STAT signaling in Drosophila muscles controls the cellular immune response against parasitoid infection. *EMBO Rep.* **16**, 1664–1672 (2015).

81. Chatterjee, A., Roy, D., Patnaik, E. & Nongthomba, U. Muscles provide protection during microbial infection by activating innate immune response pathways in Drosophila and zebrafish. *Dis Model Mech* **9**, 697–705 (2016).
82. Lin, S., Fan, T., Wu, J., Hui, C. & Chen, J. Immune response and inhibition of bacterial growth by electrotransfer of plasmid DNA containing the antimicrobial peptide, epinecidin-1, into zebrafish muscle. *Fish Shellfish Immunol.* **26**, 451–458 (2009).
83. Frost, R. A., Nystrom, G. J. & Lang, C. H. Lipopolysaccharide regulates proinflammatory cytokine expression in mouse myoblasts and skeletal muscle. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **283**, 698 (2002).
84. French, A. T. *et al*. Up-regulation of intelectin in sheep after infection with Teladorsagia circumcincta. *Int. J. Parasitol.* **38**, 467–475 (2008).
85. Pemberton, A. D., Knight, P. A., Wright, S. H. & Miller, H. R. P. Proteomic analysis of mouse jejunal epithelium and its response to infection with the intestinal nematode, Trichinella spiralis. *Proteomics* **4**, 1101–1108 (2004).
86. Datta, R. *et al*. Identification of novel genes in intestinal tissue that are regulated after infection with an intestinal nematode parasite. *Infect. Immun.* **73**, 4025–4033 (2005).
87. Podok, P., Xu, L., Xu, D. & Lu, L. Different expression profiles of Interleukin 11 (IL-11), Intelectin (ITLN) and Purine nucleoside phosphorylase 5a (PNP 5a) in crucian carp (Carassius auratus gibelio) in response to Cyprinid herpesvirus 2 and Aeromonas hydrophila. *Fish Shellfish Immunol.* **38**, 65–73 (2014).
88. Winkelhake, J. L., Vodicnik, M. J. & Taylor, J. L. Induction in rainbow trout of an acute phase (C-reactive) protein by chemicals of environmental concern. *Comp. Biochem. Physiol. C, Comp. Pharmacol. Toxicol.* **74**, 55–58 (1983).
89. Kovacevic, N., Hagen, M. O., Xie, J. & Belosevic, M. The analysis of the acute phase response during the course of Trypanosoma carassii infection in the goldfish (Carassius auratus L.). *Dev. Comp. Immunol.* **53**, 112–122 (2015).
90. Talbot, A. T., Pottinger, T. G., Smith, T. J. & Cairns, M. T. Acute phase gene expression in rainbow trout (Oncorhynchus mykiss) after exposure to a confinement stressor: A comparison of pooled and individual data. *Fish Shellfish Immunol.* **27**, 309–317 (2009).
91. Gerwick, L., Corley-Smith, G. & Bayne, C. J. Gene transcript changes in individual rainbow trout livers following an inflammatory stimulus. *Fish Shellfish Immunol.* **22**, 157–171 (2007).
92. Kerr, S. C. *et al*. Intelectin-1 is a prominent protein constituent of pathologic mucus associated with eosinophilic airway inflammation in asthma. *Am. J. Respir. Crit. Care Med.* **189**, 1005–1007 (2014).
93. Renshaw, S. A. & Trede, N. S. A model 450 million years in the making: zebrafish and vertebrate immunity. *Dis. Model. Mech.* **5**, 38–47 (2012).
94. Tang, R., Dodd, A., Lai, D., McNabb, W. C. & Love, D. R. Validation of zebrafish (Danio rerio) reference genes for quantitative real-time RT-PCR normalization. *Acta Biochim. Biophys. Sin. (Shanghai)* **39**, 384–390 (2007).
95. Hruscha, A. & Schmid, B. Generation of zebrafish models by CRISPR /Cas9 genome editing. *Methods Mol. Biol.* **1254**, 341–350 (2015).
96. Krebs, J. CRISPR design tool and protocol (2015). Data retrieved: 07:16, May 19, GMT. https://figshare.com/articles/CRISPR_Design_Tool/1117899 (2015).
97. Altschul, S. F. *et al*. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* **25**, 3389–3402 (1997).
98. Meeker, N. D., Hutchinson, S. A., Ho, L. & Trede, N. S. Method for isolation of PCR-ready genomic DNA from zebrafish tissues. *BioTechniques* **43**, 614 (2007).
99. Turpeinen, H. *et al*. Proprotein Convertase Subtilisin/Kexin Type 7 (PCSK7) Is Essential for the Zebrafish Development and Bioavailability of Transforming Growth Factor beta1a (TGFbeta1a). *J. Biol. Chem.* **288**, 36610–36623 (2013).
100. Harjula, S. E., Ojanen, M. J. T., Taavitsainen, S., Nykter, M. & Rämet, M. Interleukin 10 mutant zebrafish have an enhanced interferon gamma response and improved survival against a Mycobacterium marinum infection. *Sci Rep* **8**, 10360 (2018).
101. Traver, D. *et al*. Transplantation and *in vivo* imaging of multilineage engraftment in zebrafish bloodless mutants. *Nat. Immunol.* **4**, 1238–1246 (2003).

Acknowledgements

This study was financially supported by the Academy of Finland (M.R., 277495, M.P. 295814 and 286477), the Sigrid Juselius Foundation (M.R.), the Jane and Aatos Erkko Foundation (M.R.), the Competitive State Research Financing of the Expert Responsibility Area of Tampere University Hospital (M.R., M.P. 9U047 and 9V049), the Competitive State Research Financing of the Expert Responsibility area of Oulu University Hospital (M.R.), the Tampere Tuberculosis Foundation (M.O., M.R., S.-K.H., M.P.), the City of Tampere Science Foundation (S.-K.H), the Väinö and Laina Kivi Foundation (M.O., S.-K.H.), the Finnish Cultural Foundation, the Central Fund (S.-K.H.), the Finnish Concordia Fund (S.-K.H.), the Orion Research Foundation sr (S.-K.H), the Maud Kuistila Memorial Foundation (M.O.), the University of Tampere Doctoral Programme in Biomedicine and Biotechnology (M.O.), the Cancer Society of Finland (M.P.) and Tays tukisäätiö (Tays Support Foundation) (M.P.). We thank the Tampere Zebrafish Core Facility, partly funded by Biocenter Finland, for maintaining and providing the zebrafish. The use of the facilities and expertise of the Protein Technologies core facility of the University of Tampere, a member of Biocenter Finland, is also gratefully acknowledged. We also greatly acknowledge Hannaleena Piippo, Jenna Ilomäki, Leena Mäkinen, Carina Bäuerlein, Mirja Niskanen, Juha Saarikettu, Janey Barron, Christopher Gault, Janne Kärnä, Marianne Karlsberg, Ine Herman, Anna Grönholm and Latifeh Azizi for technical assistance and Jukka Lehtiniemi for taking the adult zebrafish images. In addition we thank Henna Myllymäki, Hannu Turpeinen, Mataleena Parikka and Tero Järvinen for scientific advice and support, as well as Hannah Pratt and Helen Cooper for proof-reading this manuscript.

Author Contributions

M.P., V.H. and M.R. provided materials and facilities for the research. M.O., M.U., J.M., V.H., M.P. and M.R. designed and/or directed the experiments. M.O., A.S., N.K., S.-K.H., K.O. and M.U. performed the experiments. M.O. and M.U. analyzed the data. M.O., M.U. and M.R. wrote the paper. All authors reviewed and approved the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-37678-1.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019

PUBLICATION IV

The inflammasome adaptor *pycard* is essential for immunity against *Mycobacterium marinum* infection in adult zebrafish.

Uusi-Mäkelä MIE, Harjula SE, Sillanpää A, Nätkin R., Niskanen M, Nykter, M., Rämet M.

Submitted manuscript