Key innate immune components controlling intracellular infection

Benard, E.L.

Chapter 3

A morpholino knockdown screen for genes involved in mycobacterium infection

Erica L. Benard [1], Peter I. Racz [1], Alexander E. Nezhinsky [2], Fons J. Verbeek [2], Herman P. Spaink [1], Annemarie H. Meijer [1]

[1] Molecular Cell Biology, Leiden University
[2] Section Imaging and Bioinformatics, Leiden Institute for Advanced Computer Science, Leiden University

Abstract

Zebrafish infection with *Mycobacterium marinum* is a well characterised animal model for human tuberculosis. Based on previous transcriptome data of zebrafish embryonic macrophages and infection studies, we selected genes for a knockdown screen to identify host genes involved in controlling mycobacterial growth during early stages of infection when granulomas are being formed. The level of infection in zebrafish embryos could be assessed by using specifically designed software that calculates the *M. marinum* fluorescent pixel count per embryo. This screen resulted in 7 hits of the 17 candidate genes tested that repeatedly led to an increased or decreased *M. marinum* infection load after a morpholino-based knockdown and 5 of these results could be reproduced by using a second morpholino with different target sites indicating their high specificity. Of these candidate hits, the transcription factor gene *atf3*, knockdown led to reduced infection, while knockdown of the other hits led to increased infection, including the perforin gene *mpeg1*, the scavenger receptor genes *cd36* and *marco*, and the signalling mediator gene *creld2*. RNA deep sequencing analysis of uninfected and infected samples of both morphants of *creld2* showed that although no general immune-deficiency was observed, a specific down-regulation of a group of apolipoprotein genes was detected. Further in-depth analysis of these genes will contribute to a better understanding of the innate immune response to mycobacterial infection and the mechanisms involved in granuloma formation.

Introduction

Mycobacterium tuberculosis is the cause of tuberculosis in humans. One third of individuals of the world population is infected with this pathogen and there is an increasing occurrence of multidrug resistant strains of *M. tuberculosis* [1], which highlights that understanding the mechanistic basis of host-pathogen interactions underlying tuberculosis has important medical implications. In order to address this topic, a combination of functional genomics approaches and modern cell biology are required in animal models that mimic the human disease. *Mycobacterium marinum* is the closest genetic relative of the *M. tuberculosis* complex and a natural pathogen of fish. Similar to human tuberculosis, *M. marinum* infected fish develop latency with survival of the pathogen in caseating granulomas, which are tight aggregates of infected and non-infected macrophages and other immune cells [2, 3]. Infecting the optically transparent zebrafish embryo with fluorescent bacteria allows us to monitor the infection *in vivo* and follow granuloma formation in real time [4].

A major benefit of using zebrafish embryos as a host for studying gene function is the use of research tools for reverse genetics by knocking down a target gene with, for instance, morpholinos. Morpholinos consist of antisense oligonucleotides, usually 25 nucleotides long, with bases connected to a morpholine ring structure as the backbone,

making them very stable. They can be designed to either block translation initiation in the cytosol (by targeting the 5' UTR or the first 25 bases of coding sequence) or they can modify pre-mRNA splicing in the nucleus (by targeting splice junctions or splice regulatory sites). Unfortunately, the use of morpholinos can also produce off-target effects (non-antisense effects due to interactions with structures other than gene transcripts) which can induce an aspecific phenotype. The exact underlying mechanisms of off-target effects are not completely understood, however, the phenotypes are often the result of p53-dependent neural toxicity [5] or a more extreme toxicity with severe overall deformity [6, 7]. The advantage of using a morpholino approach over a knockout approach, such as ENU mutagenesis, is that knockdown of the selected gene can be rapidly achieved in any transgenic or wildtype zebrafish line, whereas generating a stable mutant zebrafish line requires extensive labour and time. Furthermore, in the case of studies of essential genes, knockdown of which leads to lethality, translation-blocking morpholino knockdown can give complementary information because it can also lead to knockdown of maternally inherited mRNAs. Therefore, if applied with appropriate controls, morpholinos remain a powerful tool.

To investigate which genes play a role during *M. marinum* infection, we set up a morpholino knockdown screen in zebrafish to identify genes involved in susceptibility or resistance to mycobacterial infection. The target genes studied in this research were selected for their specific expression in macrophage cells, for their infection-inducible or -repressible expression during *M. marinum* infection, or for their suggested function based on literature. This selection process resulted in seventeen candidate genes to screen, in addition to *spi1b* (also called *pu.1*) which is known to be required for leukocyte development, as a positive control [8]. The selection included several macrophage-specific genes (such as *cd53*, *coro1a*, *mpeg1*, *marco*, and *mfap4*) previously shown to be expressed downstream of the highly conserved Spi1b transcription factor [9] or expressed in macrophages of *mpeg1* transgenic reporter lines (unpublished). These genes were considered to be promising candidates because two other Spi1-regulated genes (*cxcr3*.2 and *ptpn6*) were shown to be involved in infection with *M. marinum* [9, 10]. The selection also included infection-responsive genes, such as the transcription factor gene *atf3*, a gene shown to be required for optimal killing of bacteria by phagocytes in mice, *plac8* [11], and an immune-responsive gene with bactericidal activity of macrophage-lineage cells, *irg1*. The final selection of candidate genes for which morpholinos were designed is shown in Table 1.

Based on the initial screening results, a subset of candidate genes was selected for more detailed analyses and their role in mycobacterial infection was confirmed with additional morpholinos. The endoplasmic reticulum (ER) stress-inducible gene *cysteine-rich with EGF-like domains 2* (*creld2*) was selected as an example gene from the screen to confirm the specificity of the morpholino knockdown approach. Therefore, RNA sequencing was applied to further investigate the knockdown phenotype which was associated with a strongly increased mycobacterial burden. Notably, a group of genes

encoding apolipoproteins was down-regulated in absence or presence of infection in both *creld2* morphant groups.

Materials and methods

Zebrafish lines and handling of embryos

Zebrafish were handled in compliance with the local animal welfare regulations and were maintained according to standard protocols (zfin.org). The culture was approved by the local animal welfare committee (DEC) of the University of Leiden and all protocols adhered to the international guidelines specified by the EU Animal Protection Directive 2010/63/EU. The zebrafish line AB/TL was used in this study. Embryos were grown at 28.5-30°C in egg water (60 µg/ml Instant Ocean sea salts). For the duration of bacterial injections and imaging, embryos were kept under anaesthesia in egg water containing 200 µg/ml tricaine (Sigma-Aldrich). Embryos used for stereo fluorescence imaging were kept in egg water containing 0.003% 1-phenyl-2-thiourea (Sigma-Aldrich) to prevent melanisation.

Morpholino knockdown

Morpholino oligonucleotides (Gene Tools) were diluted to the desired concentration in 1x Danieau buffer (58 mM NaCl, 0.7 mM KCl, 0.4 mM $MgSO_4$, 0.6 mM $Ca(NO_3)_2$, 5.0 mM HEPES; pH 7.6) containing 1% phenol red (Sigma-Aldrich) and 1 nl was injected into the yolk at the 1-2 cell stage using a Femtojet injector (Eppendorf). An overview of the used morpholino sequences, exon-intron boundary sites, and concentrations can be found in Table 1. As a control, the standard control morpholino (Gene Tools) was used at the same concentrations as the other morpholinos.

Infection experiments

Embryos were staged at 24 hours post fertilisation (hpf) by morphological criteria and manually dechorionated. *M. marinum* infections were performed using the Mma20 strain expressing mCherry in a pSMT3 vector [12]. *M. marinum* was prepared as described previously, injected into the blood circulation at 28 hpf, and plated during injections to confirm the number of colony forming units (cfu's) injected [13].

RNA isolation, DNase treatment and RT-PCR

Whole embryo RNA isolation and removal of residual genomic DNA was performed as described in [14]. Morpholino knockdown was verified by RT-PCR with the SuperScript® One-Step RT-PCR System (Invitrogen, #10928-034) using 50 ng DNAse treated RNA template. RT-PCR primers for knockdown verification can be found in Table 2.

RNA for deep sequencing analysis was isolated using QIAzol lysis reagent and purified using the Rneasy MinElute Cleanup kit (QIAGEN Benelux B.V., Venlo, Netherlands). RNA sequencing was performed as previously described [15].

Chemically induced inflammation assay

Copper sulphate was administered via the egg water ($CuSO_4$ (Merck, #1027910250), 10 uM, 2 hours treatment at 28°C) [16]. Embryos were washed in egg water three times and fixed in 4% paraformaldehyde in phosphate buffered saline (PBS) overnight at 4°C.

Immunohistochemistry

Identification of leukocytes was performed by immuno-labelling with the leukocyte specific anti-L-plastin antibody and Alexa568-conjugated secondary antibody [17].

Image analysis

Fluorescence images were taken with a Leica MZ16FA stereo fluorescence microscope equipped with a DFC420C digital colour camera. In *M. marinum* infection experiments, bacterial pixel counts were obtained with stereo fluorescence images and analysed using dedicated pixel quantification software [18]. An image can be made of one, two or three embryos and contains two channels: bright field and fluorescence. The software allows the user to allocate the number of embryos per image and it selects the fluorescence channel for analysis. Images of non-infected embryos are used to set the threshold level for background fluorescence and the number of fluorescent pixels per embryo is determined for each group of morphants or control embryos. This is performed for all groups of embryos and the results are written to a comma-separated file for statistical analysis.

Statistical analysis

All data (mean ± SEM) were analysed using unpaired, two-tailed t-tests for comparisons between two groups (ns, no significant difference; *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$). For RNA deep sequencing, significant differences were calculated by DESeq analysis of transcript count data (*=fold change>2, $p<0.05$).

Gene synteny, protein alignment and homology

Gene synteny of *creld2* was analysed with Genomicus, version 75.01 [19]. Multiple protein alignment performed with ClustalW2 [20]. Amino acid identity and similarity percentages were determined using UniProt CLUSTAL O(1.2.0) multiple sequence alignment.

Results

Morpholino screen workflow

To study the effect of our selected genes on the development of *M. marinum* infection we chose to use a morpholino knockdown approach as a rapid screen while zebrafish from an ENU mutagenesis screen were being sequenced (Hubrecht and Sanger Institutes). We primarily used morpholinos targeting exon-intron boundaries of the pre-mRNA to block correct splicing. Splice-blocking sites were preferentially chosen to delete an exon encoding a functional domain of the protein or creating an early stop codon. Similarly, when an intron was included this introduced an early stop codon. When no splice blocking sites could be determined, a translational-blocking morpholino was designed. The specificity of the morpholinos was checked via BLAST (Ensembl) and the optimal concentration was determined by titrating the dose. The ideal dose of morpholino was set at the highest concentration that elicited no detectable off-target effects [21, 22] yet still produced an effective knockdown of the target gene. Efficient splice modification was determined by reverse-transcriptase polymerase chain reaction (RT-PCR) showing a band shift after gel electrophoresis of RT-PCR products.

After injecting the morpholinos into 1-2 cell staged embryos (fig. 1A), the morphants and their control embryos were treated with PTU at 12 hpf to prevent melanisation, then at 24 hpf they were dechorionated, staged according to morphological criteria [23] and subsequently infected via the blood island with 150-200 cfu of mCherry-expressing Mma20 bacteria (fig. 1B). Fluorescence images of the infected embryos at 4 days post infection (dpi) were created (fig. 1C) to assess the bacterial burden with dedicated pixel quantification software [18, 24]. A subset of these candidate genes yielding a significantly higher or lower bacterial pixel count compared to the standard control group were selected for more detailed analysis and their role in mycobacterial granuloma formation was confirmed with additional morpholinos targeting different intron-exon boundaries (fig. 1D).

During the first phase of the morpholino screen we observed slight differences between the pre-calculated cfu value and the actual injected cfu value that was determined by plating aliquots of the injection dose. Therefore, to enhance consistency within different experiments of the morpholino screen, we tested whether using aliquots of glycerol stocks of *M. marinum* stored at -80°C would enable a more consistent cfu number for injection. To test this, we used the *creld2* morphants that we showed during the first phase of the morpholino screen to have a reliable significant effect on the levels of infection generated by liquid cultured *M. marinum* and injected the morphants with two sources of *M. marinum*: a 7H9 liquid culture stock grown at 28.5°C overnight to logarithmic phase and a glycerol stock [13]. Image analysis at 4 dpi showed that the *creld2* morphants only showed a significant difference when injected with the liquid culture (fig. 2A) and not with the glycerol stock (fig. 2B) of *M. marinum*. In addition,

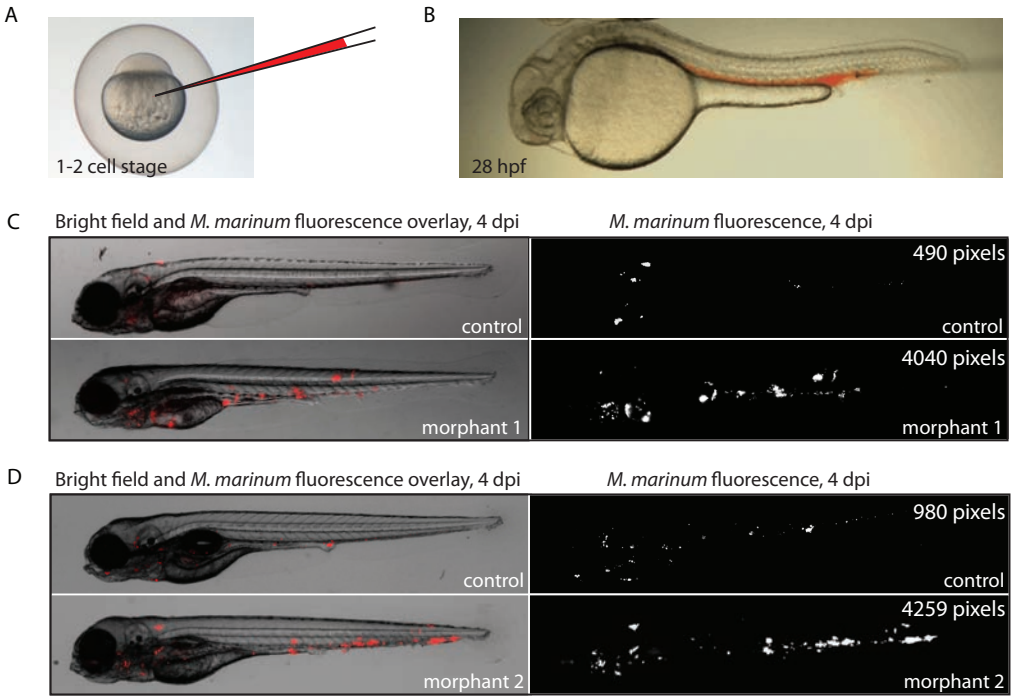


Figure 1. Overview of screen workflow. (A) Morpholinos are micro-injected into 1-2 cell stage AB/TL embryos. (B) Morphant and control embryos are injected with mCherry-expressing *M. marinum* bacteria (150-200 cfu) in 2% PVP. The injection substance contains 1% phenol red for visual support of correct injection into the blood circulation at 28 hpf. Image is still from [13]. (C and D) Images of an infected control (top) and morphant (bottom) embryo injected with the (C) first morpholino and the (D) second morpholino of the same gene (example used is from knockdown of *creld2*). Left hand images are overlays of bright field and mCherry-fluorescence channel of *M. marinum* and right images are the mCherry-fluorescence channel with fluorescence depicted as white pixels. *M. marinum* fluorescence levels as determined using dedicated pixel quantification software [18, 39] are indicated in the top right corner of the images.

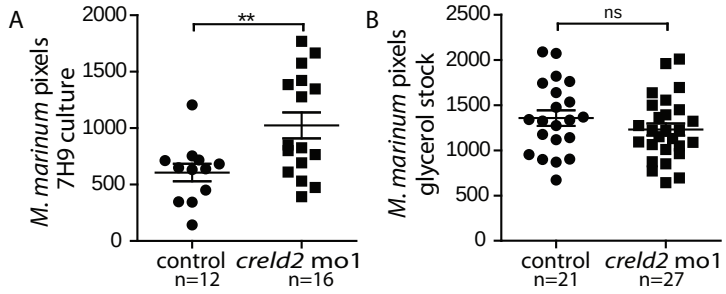


Figure 2. Infection with 7H9 liquid cultured *M. marinum* leads to increased bacterial burden in *creld2* morphants. Morphants of *creld2* and their controls only show a significant increase in *M. marinum* pixel count after injection with (A) 7H9 liquid cultured *M. marinum* and not after injection with (B) glycerol stock *M. marinum*. Bacterial burden was quantified by determining the number of fluorescent bacterial pixels with dedicated software.

when determining cfu numbers the colonies of the liquid culture *M. marinum* grew sufficiently for counting at 4 dpi, while the colonies of the glycerol stock *M. marinum* grew with a delay and could only be counted at 7 dpi. This experiment was repeated with morphants of *myeloid differentiation factor 88* (*myd88*), encoding an adaptor protein in Toll-like receptor (TLR) and interleukin-1 receptor signalling. Mutation or knockdown of this gene is known to lead to a severely increased bacterial load [25, 26]. Similar to knockdown of *creld2*, significant differences were only observed in the morphants injected with bacteria from liquid cultured *M. marinum* (data not shown) and not in morphants injected with bacteria from glycerol stock. Considering that morpholino effects are attenuated over time, we concluded that the more rapid growth of bacteria from liquid culture is better compatible with the assessment of morpholino effects. This impelled us to perform the remaining part of the morpholino screen with the liquid cultured *M. marinum*.

From the selected candidate genes, *spi1b* (*pu.1*) was used as a positive control to validate the designed workflow of the screen. Knockdown of this transcription factor gene causes a lack of macrophages in embryos and these morphants have previously been shown to be highly susceptible to *M. marinum* infection [27]. This result was confirmed during our morpholino knockdown screen (table 1 and supplementary fig. 1A).

Morpholino screen results

After initial screening of the first seventeen genes, seven positive hits were found that showed a significant difference in bacterial infection level and that was repeatable at least twice: one reducing *M. marinum* infection (*atf3)* and six increasing *M. marinum* infection (*cd36*, *marco*, *mpeg1*, *creld2*, *cd53*, *coro1a*) (table 1 and supplementary fig. 1). These seven positive hits were then re-tested with a second splice-blocking morpholino targeting a different exon-intron boundary or an ATG translational blocking morpholino (table 1 and supplementary table 2). This resulted in five hits that showed an infection phenocopy: *creld2*, *cd36*, *mpeg1*, *marco*, and *atf3*. Of these hits, *creld2* is further described below and *atf3, marco*, and *mpeg1* are further investigated in detail in the thesis Chapters 4, 5, and 6. Combined bacterial pixel count data for other morpholinos that showed consistent repeated effects on *M. marinum* infection are shown in supplementary fig. 1.

Detailed analysis of Creld2

The *creld2* gene shows high conservation between zebrafish (ENSDARG00000029071), humans (ENSG00000184164), mice (ENSMUSG00000023272), and rats (ENSRNOG00000004659) in terms of gene synteny and protein homology (supplementary fig. 3). SMART analysis of Creld2 protein reveals that it contains a signal peptide domain, a DUF3456 domain (TLR4 regulator and MIR-interacting saposin-like protein (MSAP)), an epidermal growth factor-like calcium binding domain, and a furin-like cysteine rich repeats domain (supplementary fig. 3B). When *creld2* was knocked down in zebrafish

Table 1. Results of morpholino screen

Gene symbol	Selection criteria	MO #	Screened # morphant/control	# Repeats	Significance	% *M. marinum* Pixel Count	Confirmed knockdown
spi1b	Positive control. Morphants have depleted macrophage population and have high susceptibility to *M. marinum* infection in zebrafish embryo model.	1	24/24	1	Yes, P<0.0001	**2181%**	Yes
atf3	Infection responsive; literature links with human TB and TLR signalling	1	321/392	3	Yes, P<0.0001	**46%**	Yes
		2	133/151	3	Yes, P<0.0001	**54%**	No
		3	95/112	2	Yes, P<0.0001	**31%**	N.A.
cd36	Infection responsive; literature link with human TB	1	174/135	4	Yes, P<0.0001	**220%**	Yes
		2	35/60	1	Yes, P=0.0014	**189%**	Yes
cd53	Macrophage specific; literature link with TNF signaling	1	128/117	3	Yes, P<0.0001	**469%**	No
		2	53/36	1	No	**233%**	N.A.
coro1a	Macrophage specific; literature link with human TB	1	30/30	2	Yes, P=0.0062	**617%**	No
		2	6/15	1	No	**390%**	No
creld2	Infection responsive; function in ER stress (important in infection)	1	288/215	5	Yes, P<0.0001	**236%**	Yes
		2	147/129	2	Yes, P<0.0001	**213%**	Yes
marco	Macrophage-specific; literature link with human TB	1	60/102	2	Yes, P<0.0001	**238%**	Yes
		2	109/73	2	Yes, P<0.0001	**213%**	Yes
mpeg1	Infection responsive; macrophage-specific	1	239/211	5	Yes, P<0.0001	**224%**	Yes
		2	158/141	3	Yes, P<0.0001	**536%**	Yes
cd180	Infection responsive; literature link with TLR signaling	1	79/167	3	No	**146%**	No
crel	Infection responsive; major transcription factor in innate immunity	1	16/31	1	No	**178,8%**	No
irg1l	Infection responsive; macrophage-specific	1	Phenotype	1			N.A.
irg1	Infection responsive; macrophage-specific	1	12/18	1	No	**72%**	No
lgals9l1	Macrophage-specific; literature link with human TB	1	Phenotype				N.A.
		2	Phenotype				N.A.
mfap4	Macrophage-specific; putative role in inflammation	1	51/51	2	No	**101%**	Yes
plac8.1	Infection responsive; literature on mouse infection models	1	26/12	1	No	**68%**	No
stat1a	Infection responsive; major transcription factor in innate immunity	1	41/15 and Phenotype	2	Yes, P<0.0001 and N.A.	**28%**	No
		2	Phenotype	1			N.A.
stat1b	Infection responsive; major transcription factor in innate immunity	1	Phenotype	2			N.A.

% Legend
<40
40-80
80-120
120-200
200-300
300-400
>500

* TB = tuberculosis, TLR = Toll-like receptor

Table 2. Description of used morpholinos.

Gene symbol	Ensembl ID	Target boundary site	Morpholino sequence 5' - 3'	Concentration [mM]
spi1b	ENSDARG00000000767	MO1 ATG	GATATACTGATACTCCATTGGTGGT	1
atf3	ENSDARG00000007823	MO1 intron2-exon3	GGACAGCCTGAAATAACAACATCTA	0.2
		MO2 intron3-exon4	CTTCTCAGACTCCTAAAGTCAAAAC	0.2
		MO3 ATG	GTGCTGAAGCATCATGCCGAAATCT	0.1
cd36	ENSDARG00000032639	MO1 exon2-intron2	CTATGAGGCCACAAATATTACCTGT	0.2
		MO2 exon3-intron3	TTTGAAATACATACTGCGACAGCAA	0.6
cd53	ENSDARG00000068233	MO1 intron5-exon6	AGCAACCAAACTGGTGAAAGAAGGA	0.1
		MO2 exon6-intron6	ATGTGAGAGCAGTATTACCTTATGC	0.1
coro1a	ENSDARG00000054610	MO1 exon3-intron3	ATGAAGTCCTTGTCACTCACCATGA	0.25
		MO2 exon4-intron4	TGGTTTCTGATGTGGTTTACCTGCA	0.25
creld2	ENSDARG00000029071	MO1 exon5-intron5	TGACAGTATTTTTGAATACCTTTGC	0.6
		MO2 exon2-intron2	CATTTTAATTTTCTTCTCACCTCGT	0.45
marco	ENSDARG00000059294	MO1 intron15-exon16	ACGGACCACTGATGAAGAAAAACAA	0.02
		MO2 exon3-intron3	TAACTAAAAGTACCTTGGCTTCCAC	0.3
mpeg1	ENSDARG00000055290	MO1 exon1-intron1	ATTTTGTACTTACTTGAACCCGTGC	0.2
		MO2 intron1-exon2	GGTTACGGACCTGAGAAACAAATTT	0.1
***cd180(2of2)* and *cd180(1of2)* ***	ENSDARG00000093750 and ENSDARG00000067536	MO1 exon1-intron1 and exon2-intron2	GGTTATTTAGCTTTGCTCACCTTGT	0.5
crel	ENSDARG00000055276	MO1 exon4-intron4	AAAAGATTGAACTTACGCAATGGCC	0.6
irg1	ENSDARG00000069844	MO1 intron3-exon4	GAGTGCACCTATAATATATATCCCA	Phenotype
irg1l	ENSDARG00000062788	MO1 intron2-exon3	TGCCACTGAAAACAGACAAGAGACA	N.A.
lgals9l1	ENSDARG00000025903	MO1 exon3-intron3	CGGAGTGAATGTACTTGTACCTTGT	Phenotype
		MO2 intron8-exon9	ATTATGCTGATGGAAGCAGCAAAAT	Phenotype
mfap4	ENSDARG00000090783	MO1 ATG	CGCCAAGAACAGCACGATTGCCATC	0.5
		MO2 exon3-intron3	TAAAGTCACACTCGTACCGTCCATC	0.2
plac8.1	ENSDARG00000043729	MO1 exon3-intron3	GCTATAAAGATACTAACCTCTATGT	0.5
plac8.2 *	ENSDARG00000094438	MO1 exon3-intron3	ACAGACAGACAGACATTCACCTGTA	0.25
		MO2 intron2-exon3	CCTAAGAAAAATACATGGTTGACGT	0.1
stat1a	ENSDARG00000006266	MO1 exon3-intron3	ATGTGGTCAACAGGCACCTGCAAGT	0.25
stat1b	ENSDARG00000076182	MO1 intron2-exon3	ATTCCCTGCAAACAACACTGACTAC	Phenotype
tnfrsf1a *	ENSDARG00000018569	MO1 intron2-exon2	TCCCTTCAAACTGTCCGTAAAGGGC	Inconclusive
		MO2 exon6-intron7	CTGCATTGTGACTTACTTATCGCAC	Inconclusive
tnfrsf1b *	ENSDARG00000070165	MO1 exon2-intron2	ATTCTTGTTTTATAACGTACCTGGT	Inconclusive
		MO2 exon3-intron3	CCTTTTTGTTTAAAACTCACCTTCT	Inconclusive

* *plac8.2*, *tnfrsf1a* and *tnfrsf1b* data is inconclusive (not consistent between repeated experiments) and is not shown in Table 1. Morpholino for *cd180* targets two paralogous genes: *cd180 (1of2)* and *cd180 (2of2).*

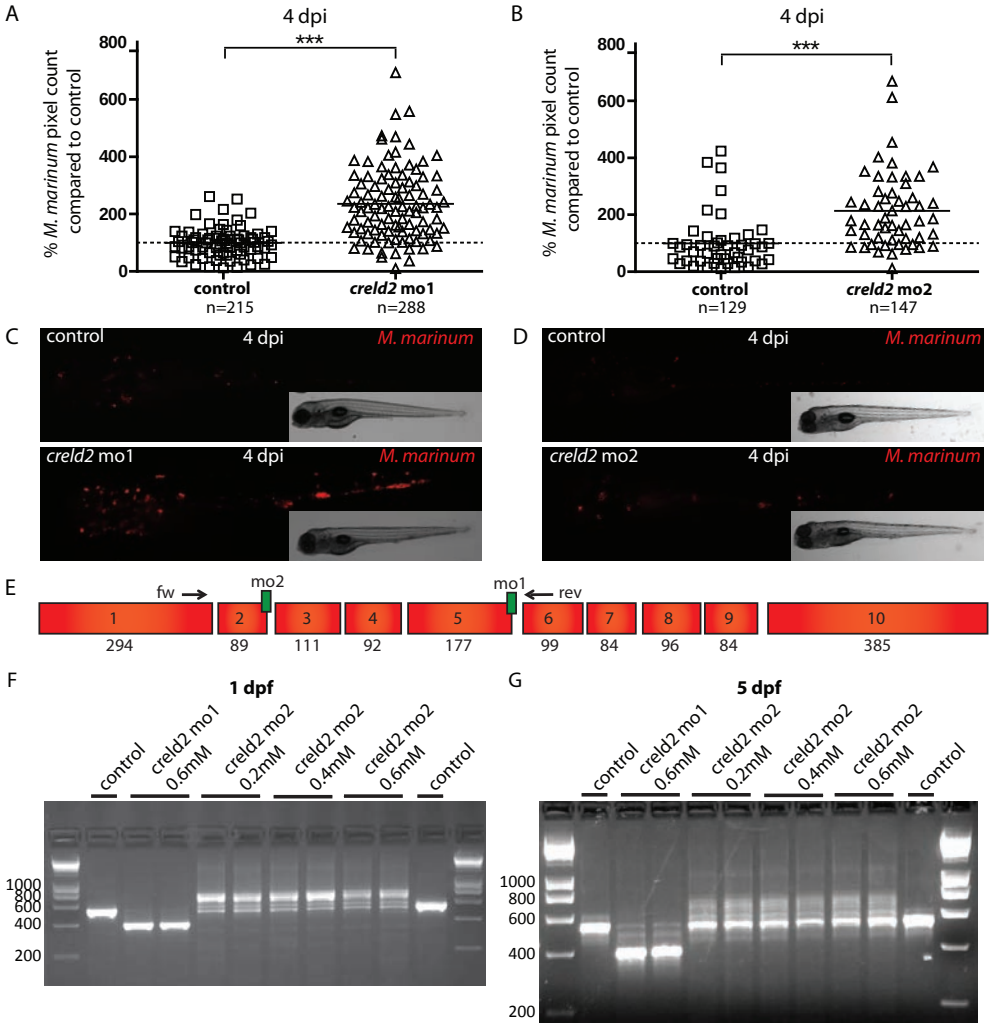


Figure 3. Knockdown of *creld2* with two separate morpholinos impairs control of *M. marinum* infection. AB/TL embryos were injected with different splice blocking morpholinos against *creld2* (mo1 or mo2) or with control morpholinos, subsequently injected with mCherry-expressing *M. marinum*, and infected embryos were imaged at 4 dpi. Graphs show combined representative results of *creld2* knockdown of 5 repeated independent experiments with (A) mo1 and 2 repeated independent experiments with (B) mo2. All pixel values of data points are compared against the average pixel value of the infected control embryos and presented as percentages. The mean ± SEM is indicated (n = number of embryos). (C and D) Representative stereo fluorescent images are shown below the graph of each experiment with an inlay of the bright field image of the same embryo. (E) Overview of the exon/intron structure of zebrafish *creld2* (ENSDARG00000029071) with base pair values given below the respective exon. Intron sizes are not depicted truly. Targeting sites of mo1, mo2, and forward and reverse primers of the RT-PCR reaction are also depicted. (F and G) Injection with both morpholinos against *creld2* generates sufficient knockdown as tested by RT-PCR. Gel electrophoresis shows a shift in fragment size of the PCR product of both morphants compared to controls (530 bp) at (F) 1 dpf and (G) 5 dpf.

embryos, it consistently led to an increased level of *M. marinum* fluorescence pixels at 4 dpi in over 200 screened embryos (Table 1 and fig. 3A+C). This infection phenotype could be phenocopied with a second morpholino (Table 1 and fig. 3B+D) and neither of the morphants showed any toxicity phenotypes (fig. 3C-D). By performing an RT-PCR with a forward primer located in exon 1 and a reverse primer located in exon 6, the knockdown of morpholino 1 (targeting exon2-intron2 boundary) and morpholino 2 (targeting exon5-intron5 boundary) (fig. 3E) was confirmed at 1 dpf (time point of *M. marinum* injection) and 5 dpf (time point of *M. marinum* pixel analysis) (fig. 3F-G).

To verify whether knocking down *creld2* has any effect on the capability of leukocytes to migrate towards a local inflammation site, an event that frequently occurs during granuloma formation, a chemical inflammation (ChIN) assay [16] was used. By treating embryos with copper sulphate, the hair cells of the neuromast along the lateral line of the zebrafish embryo are damaged which causes a local inflammation site, eventually leading to attraction of macrophages and neutrophils. The ChIN assay on *creld2* morphants and their control revealed no differences in the attraction of leukocytes to the inflammation sites (fig. 4).

To gain insight in whether the higher *M. marinum* infection could be caused by an immune-deficiency we performed RNA-Seq analysis of pools of uninfected and infected *creld2* morphants for both morpholinos and their controls at 4 dpi, the time point when granuloma-like structures have formed and *creld2* is still partially knocked down in both morphants. First, we focussed on a set of genes that are consistently induced during *M. marinum* infection (fig. 5). Although some differences with the control were noted, a clear immune-deficiency was not observed. A number of genes were similarly up- or down-regulated in both morphants, although not always significant in both cases. The genes that were slightly up-regulated in the morphants during infection included the general pro-inflammatory genes *il1b* and *cxcl-c1c,* the matrix metalloproteinases *mmp9* and *mmp13a* and the tissue inhibitor of matrix metalloproteinases *timp2b,* and the non-coding RNA *si:ch211-243g18.3*. The genes that were slightly down-regulated included *cyp1a*, *cyp24a1*, *hmox*, and *mfap4 (6 of 13)* (fig. 5).

When studying the total number of genes that were up- and down-regulated, both in uninfected and infected embryos, a substantial overlap could be observed between both morphants (fig. 6). Approximately 30% of all up-regulated genes in the uninfected samples (49 genes) and in the infected samples (52 genes) overlap between both morphants (fig. 6A). Furthermore, approximately 40% of all genes down-regulated in the uninfected samples (117 genes) and in the infected samples (114 genes) overlap between the two morphants (fig. 6B). When comparing these overlapping groups of the uninfected morphants with the infected morphants, 25% of the genes are identical in the up-regulated group (16 genes, listed in fig. 6A) and 30% of the genes are identical in the down-regulated group (55 genes, listed in fig. 6B).

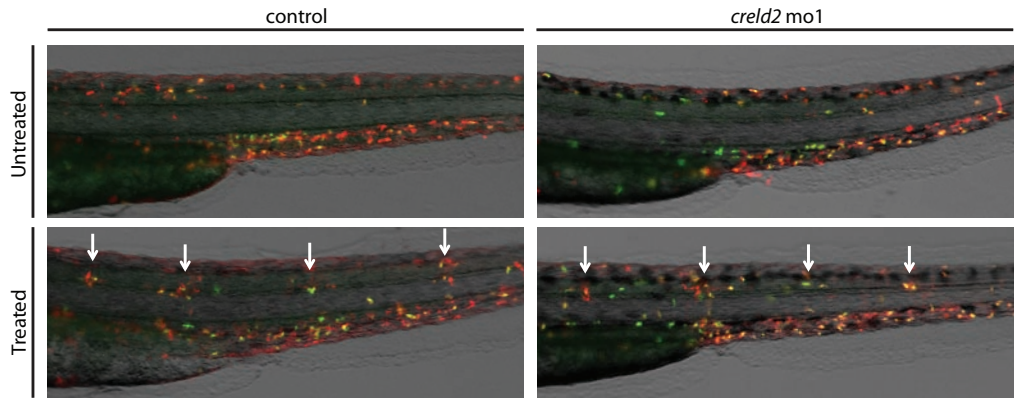


Figure 4. *creld2* does not play a role in leukocyte migration towards local inflammation. Representative images of untreated and copper sulphate treated control and *creld2* morphant (mo1) 3 dpf embryos. Embryos were immuno-labelled with Ab against the general leukocyte marker L-plastin (red signal) in combination with a neutrophil-specific Mpx TSA-staining (green signal). White arrows indicate accumulation of leukocytes at the local inflammation sites of the neuromasts due to chemically induced hair cell damage [16].

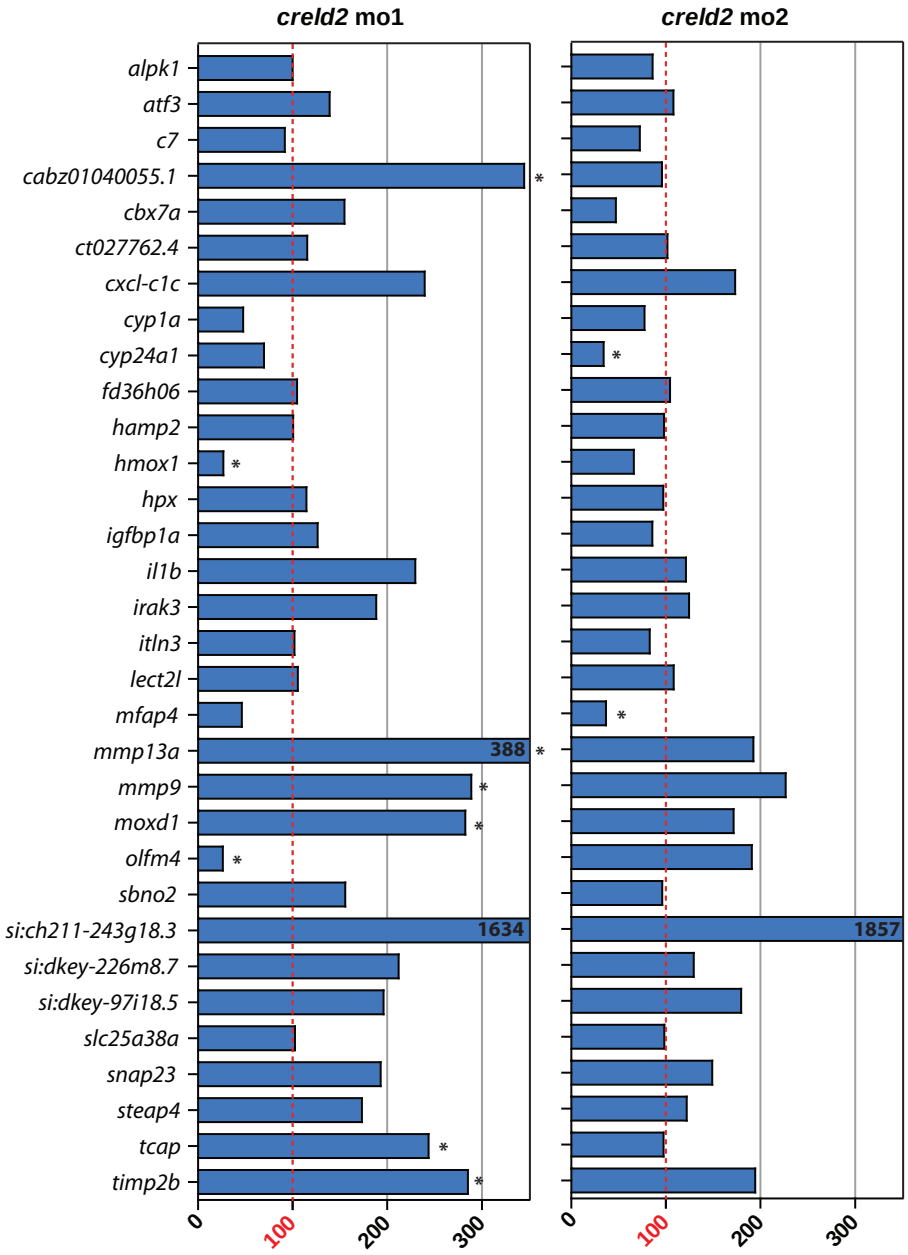
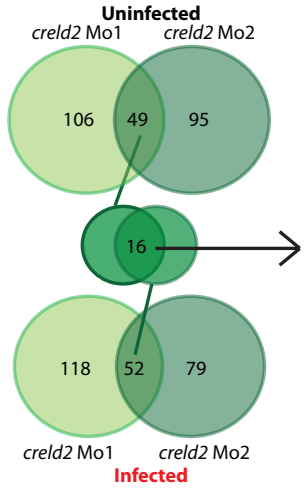


Figure 5. Effect of *creld2* knockdown on the innate immune response during *M. marinum* infection. Graphs show the effects of *creld2* mo1 and mo2 knockdown on a set of genes that showed reproducible induction by *M. marinum* infection in control embryos. The expression level of these genes under infection conditions of *creld2* deficiency is expressed as the percentage of the expression level in the corresponding infection conditions of the control embryos. Results are based on RNA-Seq analysis of pools of 30 infected and 30 uninfected embryos for each group. Embryos were injected with *M. marinum* Mma20 (250 cfu) or mock injected with 2% PVP, and RNA-Seq analysis was performed at 4 dpi. (*=fold change>2, $p<0.05$).

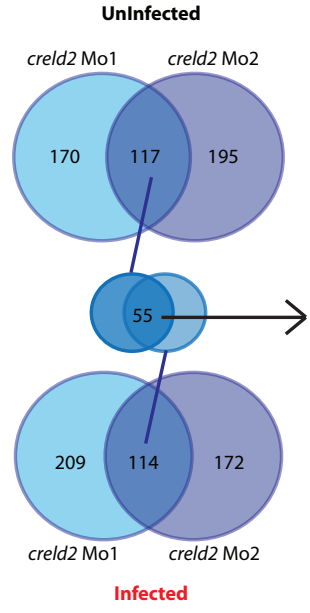
A



Up-regulated genes

Ensembl Gene ID	Associated Gene Name
ENSDARG00000087755	*afap1L1 (3 of 4)*
ENSDARG00000077389	*bx119910.6*
ENSDARG00000087080	*cabz01044356.1*
ENSDARG00000090805	*cr385054.2*
ENSDARG00000074879	*cr774186.2*
ENSDARG00000022614	*ddx43*
ENSDARG00000016713	*dhrs13l1*
ENSDARG00000070242	*enod1 (6 of 15)*
ENSDARG00000090425	*lrrtm2*
ENSDARG00000070801	*opa1*
ENSDARG00000078069	*rrm2*
ENSDARG00000009017	*scyl1 (1 of 2)*
ENSDARG00000094890	*si:ch211-163b2.4*
ENSDARG00000071644	*si:dkey-19a16.7*
ENSDARG00000094058	*si:dkeyp-118h3.1*
ENSDARG00000093888	*si:dkeyp-87e3.1*

B



Down-regulated genes

Ensembl Gene ID	Associated Gene Name	Ensembl Gene ID	Associated Gene Name
ENSDARG00000016412	*agt*	ENSDARG00000052734	*hmgcra*
ENSDARG00000026764	*ahsg*	ENSDARG00000037533	*mep1b*
ENSDARG00000013856	*amy2a*	ENSDARG00000087851	*mrpl53 (2 of 2)*
ENSDARG00000020866	*apoa4 (1 of 4)*	ENSDARG00000055876	*msmo1*
ENSDARG00000094929	*apoa4 (4 of 4)*	ENSDARG00000009153	*pla2g1b*
ENSDARG00000042780	*apoba*	ENSDARG00000079544	*si:ch1073-464p5.5*
ENSDARG00000060350	*apod (3 of 3)*	ENSDARG00000075270	*si:ch211-233g6.8*
ENSDARG00000086370	*apoea*	ENSDARG00000053448	*si:ch211-251f6.6*
ENSDARG00000024503	*c6ast3*	ENSDARG00000087390	*si:ch211-5k11.12*
ENSDARG00000052578	*c6ast4*	ENSDARG00000045835	*si:dkey-14d8.6*
ENSDARG00000030254	*cbln9*	ENSDARG00000086849	*si:dkey-187k19.2*
ENSDARG00000043168	*cela1 (2 of 7)*	ENSDARG00000094215	*si:dkey-240n22.2*
ENSDARG00000041569	*ces2*	ENSDARG00000069189	*si:dkey-242h9.3*
ENSDARG00000041595	*ces3*	ENSDARG00000028784	*si:dkey-24p1.6*
ENSDARG00000073898	*ckap5*	ENSDARG00000077192	*slc15a1a*
ENSDARG00000043722	*cpa4*	ENSDARG00000090469	*sycn (1 of 2)*
ENSDARG00000021339	*cpa5*	ENSDARG00000023176	*tdo2b*
ENSDARG00000090899	*cr450711.2*	ENSDARG00000023151	*ucp1*
ENSDARG00000045089	*CRP*	ENSDARG00000036833	*upp2*
ENSDARG00000021172	*cyp2ad2*	ENSDARG00000093844	*zgc:136461*
ENSDARG00000042953	*cyp2n13*	ENSDARG00000030893	*zgc:154142*
ENSDARG00000042641	*cyp51*	ENSDARG00000095863	*zgc:161979*
ENSDARG00000056744	*ela2*	ENSDARG00000086569	*zgc:172051*
ENSDARG00000007276	*ela3l*	ENSDARG00000034403	*zgc:173443*
ENSDARG00000053526	*enpp7*	ENSDARG00000076221	*zgc:198419*
ENSDARG00000019532	*fads2*	ENSDARG00000009443	*zgc:92137*
ENSDARG00000090689	*hbbe1.2*	ENSDARG00000040282	*zgc:92590*
ENSDARG00000023656	*he1a*		

Figure 6. Overlap of genes up- and down-regulated in uninfected and infected *creld2* morphants. Significantly (A) up-regulated genes and (B) down-regulated genes in *creld2* mo1 samples (left circles in light colour) and mo2 samples (right circles in dark colour) of uninfected (top circles) and infected (bottom circles) embryos. The overlap of each group is shown as an overlap between each circle (Venn-diagram). Total overlap between overlapping genes of uninfected and infected embryos are indicated as small Venn-diagrams in the middle (16 up-regulated and 55 down-regulated genes). Lists of these genes are provided on the right with their Ensembl gene IDs and associated gene names.

Discussion

This study focussed on a knockdown screen to identify host genes that are involved in control of mycobacterial infection. We used a morpholino-based approach to knockdown a list of candidate genes and studied their effect on *M. marinum* infection, the natural pathogen of fish and a close relative of *M. tuberculosis*. Approximately 30% of the screened genes resulted in hits that when knocked down gave rise to a reproducible increase or decrease of *M. marinum* infection load which could be phenocopied with a second morpholino. The high number of hits is not surprising considering that all genes selected are either infection-responsive, expressed in macrophages, the immune cell that plays a major part in the infection process of *M. marinum*, or have been published to have a link with mycobacterial infection. In follow up work, functional analyses of three of the most interesting genes from this knockdown screen (*atf3, marco*, and *mpeg1*) have been performed to identify which step of the infection process is affected and which genes and pathways are influenced (Chapters 4, 5, and 6). Here, we investigated one of the hits from the screen, *creld2*, in more detail.

CRELD2 in mice is a glycoprotein of approximately 60-kDa that is localised not only in the ER and Golgi apparatus [28], but is also spontaneously secreted [29]. CRELD2 has been suggested to function as a signalling mediator and to be involved in processing and trafficking of proteins through the ER–Golgi apparatus and is also associated with ER stress-related diseases [28]. During *M. tuberculosis* infection, mycobacteria are capable of secreting early secreted Antigen of 6 kDa (ESAT-6) which in turn can induce ER stress [30], a response that occurs in macrophages of tuberculosis granulomas in areas where apoptotic cells accumulate in both mice and humans [31]. The *Creld2* gene is conserved between humans, mice, and zebrafish; however, the molecular features of Creld2 in either species are unknown. In zebrafish embryos, *in situ* hybridisation revealed that *creld2* (ENSDARG00000029071) is expressed at 1 dpf in the yolk syncytial layer (YSL), liver primordium and intestinal bulb, at 2 dpf in the intestinal bulb and liver, and at 5 dpf in the gut and liver [32]. We have also detected *creld2* expression in macrophage, neutrophil, and lymphocyte populations isolated from transgenic zebrafish larvae carrying fluorescent markers for these cell types (Rougeot et al., unpublished). To gain information about how gene expression is affected in *creld2* morphants during the *M. marinum* infection process, we used RNA sequencing technology (RNA-Seq) which revealed that knockdown of *creld2* pre-mRNA does not lead to a general immune deficiency: however, we observed consistent down-regulation of a group of apolipoprotein genes in both uninfected and infected embryos of both morphants. This is of interest because RNA-Seq analysis of a time course of *M. marinum* infection shows that these genes are collectively regulated in a dynamic pattern of alternating up- and down-regulation (chapter 4). Here we find that this same cluster of genes is down-regulated in both *creld2* morphant groups, either in the absence or in the presence of infection. Further study will be required to dissect the role of Creld2 in mycobacterial pathogenesis and the possible functional relationship with apolipoprotein-mediated lipid transport.

Great care was taken during our screen to avoid false-positive results caused by any off-target effects of the injected morpholinos. By first injecting a dose curve of a morpholino, we could optimise the concentration to prevent development of morphological phenotypes. Once a morpholino produced reproducible differences in *M. marinum* infection levels compared to the control embryos, a second morpholino was used to validate the observed infection phenotype. It is well documented that 15-20% of morpholinos produce non-specific toxic off-target effects [5, 6, 33] and we also encountered morpholinos that produced non-specific morphological phenotypes, even when injected at a very low dose, such as *stat1b*, *irg1l*, and *lgals9l1* (Table 1 and table 2). An additional disadvantage of morpholinos is their temporary knockdown effect which generally lasts for the first few days post fertilisation. The exact time period of knockdown can be studied by performing an RT-PCR (in the case of splice blocking morpholinos used in this study). Despite this temporary knockdown effect, different levels of *M. marinum* infections could be observed, both increased and decreased, in our morpholino screen which is most likely the result of an accumulated effect on mycobacterial growth during the first days of infection.

Despite the two previously mentioned disadvantages, morpholinos remain a rapid and convenient resource for knockdown of a gene of interest and can be used in any wild type or transgenic line of zebrafish, which is ideal for screening purposes. Once an interesting gene has been identified, more time can be invested into creating a mutant zebrafish line. While previously the identification of mutations in zebrafish relied mostly on random ENU mutagenesis, new techniques for targeted mutagenesis purpose have been developed rapidly over the last few years, which use custom nuclease-targeted genome engineering tools like zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and the clustered regularly interspaced short palindromic repeats (CRISPR/Cas) system (reviewed in [34]). These techniques are less time consuming compared to traditional ENU mutagenesis because they do not need multiple outcrosses to reduce background mutations [35]. ZFNs, TALENs and CRISPR/Cas approaches can induce double-strand breaks in DNA that stimulate error-prone non-homologous end joining or homology-directed repair at specific genomic locations [36]. On the one hand, ZFNs proved to be difficult in production, have high costs, and modest efficacy in many applications. On the other hand, TALENs proved to be more easy to produce (open source), have low costs, and a reliably high activity. Recently, the CRISPR/Cas system has proven to also be applicable to zebrafish [37, 38] and although still in further development this system is very promising. It only requires as few as two plasmids, one containing a Cas9 sequence (encoding the Cas9 enzyme that generates a specific double-stranded DNA break) and the second containing the single guide template sequence which enables target recognition with its inserted 20 base pair spacer sequence. It is possible to synthesise this spacer commercially, integrate it into the single guide template plasmid, and synthesise the RNA. Alternatively to the use of the second plasmid, a synthetic template can be used to synthesise the guide RNA, thereby avoiding any cloning steps. Although all three of these targeted approaches

to create mutant zebrafish still can produce off-target effects, they can be an efficient tool to use during the next phase of detailed analysis of target genes identified in initial screens such as our morpholino screen for genes involved in mycobacterium infection.

References

1. Chiang, C.Y., Centis, R. & Migliori, G.B. Drug-resistant tuberculosis: past, present, future. *Respirology* **15**, 413-432 (2010).

2. Swaim, L.E. *et al.* Mycobacterium marinum Infection of Adult Zebrafish Causes Caseating Granulomatous Tuberculosis and Is Moderated by Adaptive Immunity. *Infection and Immunity* **74**, 6108-6117 (2006).

3. Parikka, M. *et al.* Mycobacterium marinum causes a latent infection that can be reactivated by gamma irradiation in adult zebrafish. *PLoS pathogens* **8**, e1002944 (2012).

4. 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).

5. Nasevicius, A. & Ekker, S.C. Effective targeted gene 'knockdown' in zebrafish. *Nat Genet* **26**, 216-220 (2000).

6. Ekker, S.C. & Larson, J.D. Morphant technology in model developmental systems. *Genesis* **30**, 89-93 (2001).

7. Woods, I.G. & Schier, A.F. Targeted mutagenesis in zebrafish. *Nat Biotechnol* **26**, 650-651 (2008).

8. Rhodes, J. *et al.* Interplay of Pu.1 and Gata1 Determines Myelo-Erythroid Progenitor Cell Fate in Zebrafish. *Developmental Cell* **8**, 97-108 (2005).

9. Zakrzewska, A. *et al.* Macrophage-specific gene functions in Spi1-directed innate immunity. *Blood* **116**, e1-11 (2010).

10. Kanwal, Z. *et al.* Deficiency in hematopoietic phosphatase ptpn6/Shp1 hyperactivates the innate immune system and impairs control of bacterial infections in zebrafish embryos. *J Immunol* **190**, 1631-1645 (2013).

11. Ledford, J.G., Kovarova, M. & Koller, B.H. Impaired host defense in mice lacking ONZIN. *Journal of immunology* **178**, 5132-5143 (2007).

12. van der Sar, A.M. *et al.* Mycobacterium marinum Strains Can Be Divided into Two Distinct Types Based on Genetic Diversity and Virulence. *Infection and Immunity* **72**, 6306-6312 (2004).

13. Benard, E.L. *et al.* Infection of zebrafish embryos with intracellular bacterial pathogens. *J Vis Exp* **pii: 3781. doi: 10.3791/3781** (2012).

14. Stockhammer, O.W., Zakrzewska, A., Hegedus, Z., Spaink, H.P. & Meijer, A.H. Transcriptome profiling and functional analyses of the zebrafish embryonic innate immune response to Salmonella infection. *J Immunol* **182**, 5641-5653 (2009).

15. Veneman, W.J. *et al.* A zebrafish high throughput screening system used for Staphylococcus epidermidis infection marker discovery. *BMC Genomics* **14**, 255 (2013).

16. d'Alençon, C.A. *et al.* A high-throughput chemically induced inflammation assay in zebrafish. *BMC Biology* **8**, 151 (2010).

17. Mathias, J.R. *et al.* Live imaging of chronic inflammation caused by mutation of zebrafish Hai1. *Journal of Cell Science* **120**, 3372-3383 (2007).

18. Nezhinsky A.E., V.F.J. Pattern recognition for high throughput zebrafish imaging using genetic algorithm optimization. *Lecture Notes in BioInformatics* **6282**, 301-312 (2010).

19. Louis, A., Muffato, M. & Roest Crollius, H. Genomicus: five genome browsers for comparative genomics in eukaryota. *Nucleic Acids Res* **41**, D700-705 (2013).

20. Larkin, M.A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948 (2007).

21. Ekker, S.C. Morphants: a new systematic vertebrate functional genomics approach. *Yeast* **17**, 302-306 (2000).

22. Eisen, J.S. & Smith, J.C. Controlling morpholino experiments: don't stop making antisense. *Development* **135**, 1735-1743 (2008).

23. Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B. & Schilling, T.F. Stages of embryonic development of the zebrafish. *Dev Dyn* **203**, 253-310 (1995).

24. Stoop, E.J.M. *et al.* Zebrafish embryo screen for mycobacterial genes involved in the initiation of granuloma formation reveals a newly identified ESX-1 component. *Disease Models & Mechanisms* **4**, 526-536 (2011).

25. van der Vaart, M., van Soest, J.J., Spaink, H.P. & Meijer, A.H. Functional analysis of a zebrafish myd88 mutant identifies key transcriptional components of the innate immune system. *Dis Model Mech* **6**, 841-854 (2013).

26. Cambier, C.J. *et al.* Mycobacteria manipulate macrophage recruitment through coordinated use of membrane lipids. *Nature* **505**, 218-222 (2014).

27. Clay, H. *et al.* Dichotomous Role of the Macrophage in Early Mycobacterium marinum Infection of the Zebrafish. *Cell Host & Microbe* **2**, 29-39 (2007).

28. Oh-hashi, K. *et al.* CRELD2 is a novel endoplasmic reticulum stress-inducible gene. *Biochemical and biophysical research communications* **387**, 504-510 (2009).

29. Oh-hashi, K., Kunieda, R., Hirata, Y. & Kiuchi, K. Biosynthesis and secretion of mouse cysteine-rich with EGF-like domains 2. *FEBS letters* **585**, 2481-2487 (2011).

30. Choi, H.H. *et al.* Endoplasmic reticulum stress response is involved in Mycobacterium tuberculosis protein ESAT-6-mediated apoptosis. *FEBS letters* **584**, 2445-2454 (2010).

31. Seimon, T.A. *et al.* Induction of ER stress in macrophages of tuberculosis granulomas. *PloS one* **5**, e12772 (2010).

32. Thisse, B. & Thisse, C. Fast Release Clones: A High Throughput Expression Analysis. *ZFIN Direct Data Submission* (2004).

33. Heasman, J. Morpholino oligos: making sense of antisense? *Developmental biology* **243**, 209-214 (2002).

34. Blackburn, P.R., Campbell, J.M., Clark, K.J. & Ekker, S.C. The CRISPR system--keeping zebrafish gene targeting fresh. *Zebrafish* **10**, 116-118 (2013).

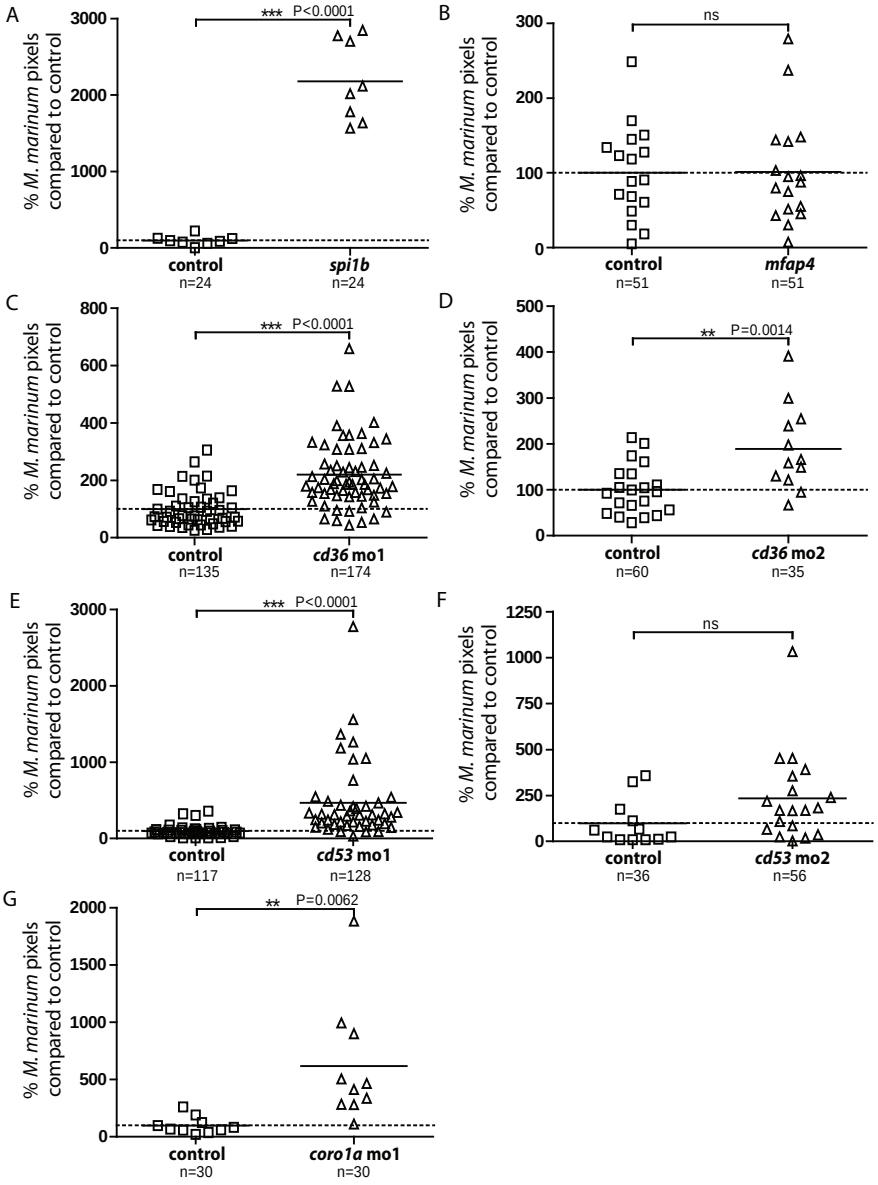
35. de Bruijn, E., Cuppen, E. & Feitsma, H. Highly Efficient ENU Mutagenesis in Zebrafish. *Methods in molecular biology* **546**, 3-12 (2009).

36. Gaj, T., Gersbach, C.A. & Barbas, C.F., 3rd ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* **31**, 397-405 (2013).

37. Hwang, W.Y. *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature biotechnology* **31**, 227-229 (2013).

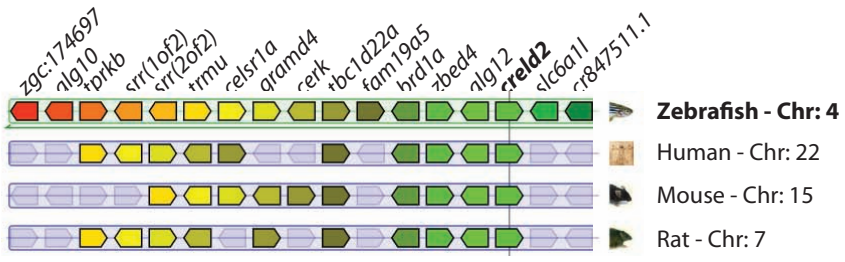
38. Jao, L.E., Wente, S.R. & Chen, W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc Natl Acad Sci U S A* **110**, 13904-13909 (2013).

39. Nezhinsky A.E., V.F.J. Efficient and Robust Shape Retrieval from Deformable Templates. *International Symposium on Leveraging Applications of Formal Methods* Part II, LNCS 7610 (T. Margaria, B. Steffen, and M. Merten (Eds.)), 42-55 (2012).



Supplementary figure 1. *M. marinum* pixel count results for morpholinos showing repeated consistent results. AB/TL embryos were injected with morpholinos against a target gene or with standard control morpholino, subsequently injected with mCherry-expressing *M. marinum*, and infected embryos were imaged at 4 dpi. Graphs show combined representative results of all experiments for knockdown of (A) *spi1b*, (B) *mfap4*, (C) *cd36* with morpholino 1, (D) *cd36* with morpholino 2, (E) *cd53* with morpholino 1, (F) *cd53* with morpholino 2, and (G) *coro1a* with morpholino 1. All pixel values of data points are compared against the average pixel value of the infected control embryos and presented as percentages. Results for knockdown of *atf3, marco*, and *mpeg1* can be found in thesis chapters 4, 5, and 6, respectively. The mean ± SEM is indicated (n = number of embryos).

A



B

```
                                  Signal peptide                    DUF3456 (TLR4 regulator
mCRELD2          MHLLLAAAFGLL--LLLPPPGAVASRK-PTMCQRCRTLVDKFNQGMANTARKNFGGGNTA 57
rCRELD2          MHLLLAAGFGLL--LLLLPP-PAASKK-PTQCQRCRTLVDKFNQGMANTARKNFGGGNTA 56
hCRELD2          MRLPRRAALGLLPLLLLLPPAPEAAKK-PTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTA 59
zCreld2          -MLLSCSIFRLFCIILLLQLGSIYTKDFTALCSTCRQLVDDFDKGLEKTAKQNFGGGNTA 59
                   *   : : *:  :**     .  ::. .: *  ** ***.*::*: .**::********
and MSAP)
mCRELD2          WEEKTLSKYEFSEIRLLEIMEGLCDSSDFECNQLLEQQEEQLEAWWQTLKKEHPNLFEWF 117
rCRELD2          WEEKTLSKYEFSEIRLLEIMEGLCDSSDFECNQLLEQQEEQLEAWWQSLKKDYPNLFEWF 116
hCRELD2          WEEKTLSKYESSEIRLLEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWF 119
zCreld2          WEERKLSKYETSEIRLTEILEGLCQSSNFECSHMLEENEEHLEAWWFKRKTKHPDLFKWF 119
                 ***:.***** ***** **:****:**:***.::** :**:*****   *..:*:**:**
                                  Epidermal growth factor-like domain
mCRELD2          CVHTLKACCLPGTYGPDCQECQGGSERPCSGNGYCSGDGSRQGDGSCQCHTGYKGPLCID 177
rCRELD2          CVRTLKACCLPGTYGPDCKECQGGSERPCSGNGYCSGDGSRQGDGSCQCHAGYKGPLCID 176
hCRELD2          CVKTLKVCCSPGTYGPDCLACQGGSQRPCSGNGHCSGDGSRQGDGSCRCHMGYQGPLCTD 179
zCreld2          CIETIKVCCPKGSFGPDCNTCIGGADRPCHGNGKCDGDGTRAGNGKCSCDEGYDGEFCLD 179
                 *:.*:*.**  *::****  * **::*** *** * .***:* *:*.* *. **.* :* *
                                                  Furin-like repeat
mCRELD2          CTDGFFSLQRNETHSICSACDESCKTCSGPSNKDCIQCEVGWAR-VEDACVDVDECAAET 236
rCRELD2          CMDGYFSLQRNETHSICLACDESCKTCSGPSNKDCVQCEVGWAR-VEDACVDVDECAAET 235
hCRELD2          CMDGYFSSLRNETHSICTACDESCKTCSGLTNRDCGECEVGWVL-DEGACVDVDECAAEP 238
zCreld2          CSDGYFNSLRNDTFFLCKECHESCVGCSGGTNQHCKECRNGWVKDQEGSCIDINECIKDP 239
                 * **:*.  **:*. :*  *.***  *** :*:.* :*. **.   *.:*:*::**   :.
                                                  Furin-like repeat
mCRELD2          SPCSDGQYCENVNGSYTCEDCDSTCVGCTGKGPANCKECIAGYTKESGQCTDIDECSLEE 296
rCRELD2          PPCSEAQYCENVNGSYICEECDSTCVGCTGKGPANCKECIAGYTKQSGQCADIDECSLEE 295
hCRELD2          PPCSAAQFCKNANGSYTCEECDSSCVGCTGEGPGNCKECISGYAREHGQCADVDECSLAE 298
zCreld2          APCSDDQYCLNTDGSFSCKACDIRCTGCKGDGASSCLNCADGYKDEEGTCTDIDECTEDP 299
                 .***  *:* *.:**: *: **  *.**.*.*...* :*  **   : * *:*:***:

mCRELD2          KACKRKNENCYNVPGSFVCVCPEGFEETEDACVQTAEGKVTE-ENPTQPPSREDL 350
rCRELD2          KACKRRNENCYNVPGSFVCVCPDGFEETEDACVQTAQSEVTE-ENPTQPLSHEDL 349
hCRELD2          KTCVRKNENCYNTPGSYVCVCPDGFEETEDACVPPAEAEATEGESPTQLPSREDL 353
zCreld2          ASCS-DNQHCLNTDGSFSC--------EEKVPAFNSEGAKTG-DSPEK---HEDL 341
                  :*   *::* *. **: *         *.. .  ::.  *  :.* :   :***
```

C

Identical / Similar	**z-Creld2**	m-CRELD2	r-CRELD2	h-CRELD2
z-Creld2	**100**	47	47	50
m-CRELD2	78	**100**	91	76
r-CRELD2	77	98	**100**	76
h-CRELD2	77	90	90	**100**

Supplementary figure 2. Conservation of vertebrate *creld2*. Genomic structure and sequence analysis of the zebrafish *creld2* gene. (A) Physical synteny of the region containing the *creld2* gene in zebrafish chromosome 4 (Chr: 4) with human chromosome 22 (Chr: 22), mouse chromosome 15 (Chr: 15), and rat chromosome 7 (Chr: 7). Seventeen genes are included in this figure, each represented by a unique coloured box. Surrounding *creld2* the genes *alg12*, *zbed4*, *brd1a*, *fam19a5*, *tbc1d22a*, *cerk*, *gramd4*, *celsr1a*, and *trmu* are arranged in the same

Supplementary figure 2 continued: order on zebrafish Chr: 4 as on the vertebrate species analysed (human, mouse, and rat). Coding direction of the genes is indicated by the pointed end. Pale blue symbols indicate non-syntenic genes. (B) Multiple alignment of the vertebrate Creld2 proteins. Grey boxes indicate the protein domains signal peptide, DUF3456 (TLR4 regulator and MIR-interacting saposin-like protein (MSAP)), Epidermal growth factor-like calcium binding domain, and furin-like cysteine rich repeats regions. *(asterisk) = single, fully conserved residue, : (colon) = conservation between groups of strongly similar properties - scoring > 0.5 in the Gonnet PAM 250 matrix, . (period) = conservation between groups of weakly similar properties - scoring =< 0.5 in the Gonnet PAM 250 matrix. (C) Homology of protein sequences was determined between zebrafish (zCreld2): ENSDARP00000048382; rat (rCRELD2): ENSRNOP00000006357; human (hCRELD2): ENSP00000332223; mouse (mCRELD2): ENSMUSP00000024042. Amino acid identity and similarity percentages were determined using UniProt CLUSTAL O(1.2.0) multiple sequence alignment.