5

Summary & discussion

In a world where bacterial infections are becoming harder and harder to fight due to rising antibiotic resistance, disease models that allow high throughput screening are needed more than ever. Such models will allow the discovery and study of novel and existing virulence factors of infectious agents by mutant screens, as well as discovery of novel antimicrobial compounds and compounds that enhance the immune system of the host. Model organisms used in disease studies range from invertebrates to mammals, for example the nematode *Caenorhabditis elegans* and the mouse. Of these two examples the mouse allows a fairly high extrapolation of the results to human, however it is not suitable for high throughput screening. In contrast *C. elegans* is suitable for high throughput screening. For example, it has been used to study *Pseudomonas aeruginosa* virulence, even to the extent of comparing the genomes of *P. aeruginosa* strains PAO1 and the more virulent PA14. However, for *C. elegans* extrapolation of the results to humans is difficult. Several other model organisms have similar shortcomings, thus, the need for a model organism that allows both extrapolation of the results and high throughput screening arose. This gap has now been successfully filled by the zebrafish.

The embryos and larvae of the zebrafish are highly suited for high throughput screening. A female can lay up to two hundred eggs, which are fertilized *ex utero*. Before the larvae start to be able to eat they are not considered an experimental animal under European guidelines for animal experimentation. At this stage the brain and neuronal system is still not fully developed and therefore large throughput screening for medical purposes at this stage can be considered broadly ethically acceptable. From a scientific perspective, the developing embryos are transparent, enabling real-time analysis of infection and the immune response. Several transgenic zebrafish lines with fluorescently labelled immune related cells are available. Like all jawed vertebrates, the zebrafish possesses the innate and adaptive immune system, however, in the developing zebrafish, these are temporally separated. The innate immune system is already functional at 1 day post fertilization (dpf), while the adaptive immune system does not reach full maturity until 2 - 4 weeks post fertilization. This allows the innate immune system to be studied separate from the adaptive immune system.

At the start of the research described in this thesis, several bacterial infection systems in the zebrafish had already been set up, both with Gram-

positive bacteria (e.g. *Streptococcus* sp.) and Gram-negative bacteria (e.g. *Salmonella typhimurium*). Additionally, extensive research was being done with *Mycobacterium marinum*, which forms granulomas in zebrafish just as its close relative *Mycobacterium tuberculosis* does in humans. However, almost all of the zebrafish infection models required intravenous injection of the infecting agent, laborious work that does not allow for high throughput screening. For high throughput screening it would be easier if infection could be achieved by immersion only within the first five days after infection and one bacterium capable of just that had been just described, *Edwardsiella tarda*.

E. tarda is a naturally occurring fish pathogen, capable of infecting many different hosts, not restricted to fish alone. This bacterium was found to be able to cause a lethal infection in zebrafish embryos after 5 hour immersion of these embryos in a bacterial suspension. Not much was known yet about this bacterium and we chose it to conduct our studies.

Genome sequence of *E. tarda*

With new sequencing techniques greatly reducing the cost and the time to sequence a whole genome, and with no available genome of *E. tarda*, we set out to be the first to sequence it and determine the virulence factors of the strain FL6-60 that was chosen for our study. Just after our first *de novo* construction of the genome however, the genome of another *E. tarda*, strain EIB202, was published. Therefore in **chapter 2** we used *E. tarda* EIB202 as a reference genome for the assembly of our *E. tarda* FL6-60 genome. This allowed us to almost completely assemble the genome, leaving only a small amount of *de novo* assembled contigs to be added manually. By matching flanking sequences of these contigs with the gaps in the reference assembly, we completed the genome. However, several contigs were left unassembled. *E. tarda* EIB202 contains a plasmid, but none of the reads generated in the sequencing could be matched to this plasmid, showing its absence in *E. tarda* FL6-60. The unassembled contigs however could be *de novo* assembled into a circular element with several overlapping reads. The coverage of this element, when corrected for the bias that we found towards the AT-rich regions, equalled that of the completed genome, showing that it is present in one copy per cell. When annotating the genes, they were primarily found to be phage related.

Moreover, the length of the element is similar to that of the P22-podoviridae, leading us to hypothesize that it is a novel phage-like element.

The genomes of the two *E. tarda* strains were compared and were found to be very similar. The gaps in the *E. tarda* FL6-60 genome that needed to be filled manually with *de novo* assembled contigs were all regions containing phage related genes. Other regions with high variation between the two strains were concentrated around predicted genomic islands, also containing phage-related genes and additionally transposons. These observations show that these two strains have rarely diverged in the none-phage-related genes. Comparison with *E. trade*'s close relative *Edwardsiella ictaluri* also showed a remarkable structural similarity. Only two regions had a different orientation and the identity between the two species was 35.4 %.

To look in more detail for differences between different *E. tarda* strains and *E. ictaluri*, we compared the T3SS and the T6SS of the different strains. This showed a high identity between the *E. tarda* strains EIB202 and FL6-60. The T3SS of *E. tarda* strain FK1051 showed the highest divergence from the other *E. tarda* strains, but this might be due to sequencing errors. This in depth comparison between the different strains of *E. tarda* shows that the deep sequencing technology used is highly suited to quickly identify differences between genomes.

Where *E. tarda* has quite a broad host range, its close relative *E. ictaluri* is a more specialized pathogen, mainly infecting catfish. Differences in hosts or host-range could be attributed to differences in the virulence factors. Therefore we took *E. ictaluri* along in the analysis of the secretion systems. We found that the order of the genes of both the T3SS and the T6SS is the same between *E. tarda* and *E. ictaluri*, but that there is a significant amount of divergence within the genes. It would be very interesting to study the effect of the differences between different *Edwardsiella* strains in a single model organism, as well as the contribution of other virulence-related genes with the help of mutants.

Exposure of zebrafish embryos to *E. tarda*

Having determined the presence of the virulence factors of *E. tarda* in our strain, we set out to set up the infection model by immersion in our lab, as described in **chapter 3**. In this chapter, we started out by attempting to confirm the ability of *E. tarda* FL6-60 to cause lethal infection in zebrafish embryos by

immersion. In addition to *E. tarda*, we also tested the pathogen *Pseudomonas aeruginosa* strains PAO1 and PA14, known to have a broad host range. Where *E. tarda* indeed caused mortality in zebrafish embryos, both strains of *P. aeruginosa* were unable to. The mortality of zebrafish embryos upon immersion in an *E. tarda* suspension was however very variable, ranging from 25 to 75 % after 4 days between different experiments. This kind of variation is too large when screening for attenuated mutants or antimicrobial compounds, so we wanted a different readout in stead of mortality.

In order to find a different readout of infection, microarray analysis was performed on zebrafish embryos immersed in an *E. tarda* suspension. Additionally, we wanted to see if *P. aeruginosa* PAO1 and PA14 did elicit a transcriptional response, since they did not cause mortality. As a non-pathogenic control *Escherichia coli* was taken along in the analysis. The microarray yielded some surprising results. *E. tarda*, the only bacterium among the four tested able to cause mortality after immersion, caused the smallest number of up- and down-regulated genes in the zebrafish, smaller even than the non-pathogenic *E. coli*. Interestingly, *P. aeruginosa* PAO1 and PA14 caused a very high response of the zebrafish. Further analysis of the different responses to the bacteria, revealed that barely any immune response gene was up-regulated. The highest up-regulated gene after exposure to *E. tarda* was *cyp1a*, a gene of the cytochrome p450 family, known to be induced by toxic chemicals in vascular endothelium and the epithelium of the gills. This gene was up-regulated by all four bacteria, and most strongly by *P. aeruginosa*, which is known to excrete large amounts of toxins and other virulence factors. Other genes that were found to be up-regulated upon exposure to *E. tarda* were *zgc:154020*, a gene with high identity to the mouse *immunoresponsive gene 1* (*irg1*) which we therefore named *irg1-like* (*irg1l*), and stanniocalcin 1 (*stc1*), which in humans has been implicated in the inflammatory response. Gene ontology analysis confirmed the lack of an immune response, showing only an enrichment of the genes with the GO-term “response to stimulus” and more specifically “response to stress”.

The lack of up-regulated immune-related genes at 5 hours post exposure (hpe) indicated that no tissue infection had been established yet, and that the regulated genes possibly reflected an epithelial response. Therefore we wanted to look at later time points after immersion in *E. tarda*. For this analysis, six

genes were chosen for qPCR analysis, *cyp1a*, *irg1l* and *stc1* from the microarray results and known immune markers *il1b*, *mmp9* and *tnfa*. The immune markers *il1b* and *mmp9* showed a response at 24 hpe which became much stronger at 48 hpe. *Irg1l* was induced at all time points. *Stc1* and *tnfa* had such variable results, that we excluded them from analysis, but these variations, in combination with the variation in mortality rate, led us to hypothesize that not all embryos were infected. Experiments fractionating the *E. tarda* suspension by centrifugation into a wash fluid and washed bacteria showed that the early response of *cyp1a* and *irg1l* could be explained as a response to shed cell membrane components or other released bacterial compounds. The response to these bacterial compounds could be epithelial, where that of *cyp1a* is relatively short lived and *irg1l* might be continued as an internal tissue response towards systemic infection or as a remaining response to the bacterial components. While it is possible that this is a purely epithelial response, one can imagine that shed bacterial compounds might diffuse into the embryo causing a general response. The response of *il1b* and *mmp9* showed to be primarily due to the bacterial fraction, suggesting that these are markers for systemic infection.

Using a single embryo RNA-extraction protocol, we tested the regulation of the genes *cyp1a*, *irg1l*, *il1b* and *mmp9* in individual embryos at 48 hpe with qPCR. The results confirmed our hypothesis stronger than expected. Where *cyp1a* was, as expected, not induced and *irg1l* was induced in all embryos as a possible result of a response to shed bacterial compounds, *mmp9* and *il1b* were only up-regulated in one out of five embryos, suggesting that only this embryo was systemically infected.

To check whether the found variation was due to individual differences between the embryos, we decided to inject zebrafish with *E. tarda*. The response of the same four genes as before was determined by qPCR at two time points, 4 hours post infection (hpi) and 8 hpi. *Cyp1a* showed little to no induction, consistent with our suggestion that it may be involved in a response of the skin epithelium upon external exposure to shed bacterial compounds. Of the other three genes *irg1l* and *mmp9* were not or only weakly induced at 4 hpi, but highly induced at 8 hpi in all embryos, while *il1b* was highly induced at both time points, also in all embryos. Not only the transcriptional response proved

reproducible in all embryos upon *E. tarda* injection, but also the mortality, since after 48 hpi all embryos had died.

Seeing that injection of embryos gave more reproducible results on the level of single embryos, we did a microarray analysis on *E. tarda* injected embryos. In contrast to the results of the microarray from the immersion experiment, this microarray showed up-regulation of many immune response genes other than *irg1l*, *mmp9* and *il1b*. A gene ontology analysis reflected this by a significant enrichment of the GO-term “immune system process” in addition to the “ response to stimulus”, the last one of which was the only one found after immersion.

The results described in this chapter raise some interesting points for discussion and further research. One of the most striking conclusions that can be drawn is that the zebrafish embryo is remarkably resistant to external infection. Of the many bacteria tested by different research groups, only two, *E. tarda* and *Flavobacterium columnare*, were reported to be able to infect 1 dpf zebrafish embryos by static immersion. However, as shown in this thesis, the ability of *E. tarda* to cause infection in the immersion assay remains doubtful. If infection occurs at all, this is with low efficiency, and it cannot be excluded that the embryos die as the result of a toxic response instead. A pilot with *E. tarda* immersion of 3 dpf embryos, embryos with open mouths, did not provide a suitable alternative, as all embryos survived the following 2 days and after this time frame the zebrafish embryo becomes less interesting for high throughput screening. These results would suggest a very strong immune system being present very early on, but the mortality rate of 100% after intravenous injection shows that this is not the case. Instead, the bacteria do not seem to be able to penetrate the zebrafish skin or gut epithelium, thereby unable to cause internal infection. While it might be possible that *E. tarda* and *F. columnare* are able to penetrate the skin, it is more likely that small unintended damage to the skin, caused during handling of the embryos, provides points of entry. In fact, at the beginning of this research, the *P. aeruginosa* strains PAO1 and PA14 seemed to cause mortality upon immersion, but when more experience was gained in working with zebrafish embryos, this mortality was reduced to none. Still, *E. tarda* and *F. columnare* are special, since they seem to be the only tested bacteria needing very little skin damage to be able to cause infection.

The resistance of zebrafish embryos to external exposure is likely to be by preventing entry. The question that remains is how this is achieved. Maybe the epithelium produces so far unidentified antimicrobial compounds. Another explanation is that the zebrafish epithelium is surrounded by mucus that traps the bacteria. More likely it is a combination of both, mucus invested with antimicrobials, but this should be studied in more detail. Maybe the ability of *E. tarda* and *F. columnare* to infect embryos arises from a resistance to possible antimicrobials.

The results of the microarray after immersion show a response of the zebrafish towards external bacteria and, as shown by qPCR, especially towards shed or excreted bacterial compounds, a response we think might be epithelial. The most interesting result was that *P. aeruginosa* strains PAO1 and especially PA14 elicited a much stronger response than *E. tarda*. Since *P aeruginosa* is known to excrete many compounds, it is interesting to use this bacterium in addition to *E. tarda* to further study this response and see whether it is epithelial. Differences between *E. tarda* and *P. aeruginosa* in the compounds involved in epithelial responses can be identified. To test if secretion systems are involved, bacterial mutants and additionally a trans-well assay could be used. In the trans-well assay, the bacteria and the embryos in a well are separated by a 0.2 µm filter mesh, barring the passage of bacteria, but allowing secreted compounds to pass and interact with the embryos. To test if the response was caused by LPS components (lipopolysaccharides) sheared of the bacterial membrane during the washing process, which is equally possible, studies with purified LPS and LPS-mutants can be done.

In conclusion to chapter 3 it can be said that the *E. tarda* immersion assay is not suitable for high throughput screening. Recently, a robotic injection system for high throughput screening has been described. In this system bacteria are automatically injected into the yolk of fertilized eggs, which slowly proliferate to cause an infection in a fully grown embryo. It is very useful for the study of *Mycobacterium marinum*. However, due to the pathogenicity of *E. tarda* it is not suitable for this bacterium, since it would kill the developing zygote before it becomes a full grown embryo. This may change if attenuated *E. tarda* mutants are found.

Analysis of a zebrafish mutant in Myd88

During this project the first knockout mutant in an essential immune-related gene was identified in a TILLING-screen. Therefore, in **chapter 4**, we used *E. tarda* to characterize this zebrafish with a knockout mutation in the gene encoding Myd88. Myd88 is an important adaptor protein in the Toll-like receptor (TLR) pathway. TLRs are one family of receptors of the innate immune system that recognize pathogen-associated molecular patterns. Several studies with *myd88*-morpholinos have already been conducted, but morpholinos have several downsides. They are only applicable during the first days of embryo development, since they will be diluted out or destroyed during further growth. Moreover, they can cause non-specific effects. The point mutation in *myd88* creates an early stop codon, leading to a truncated Myd88 protein, consisting only of a partial death domain, that has been cut off just before a residue essential for communication with downstream components, and completely lacking the TIR-domain, responsible for communication with the TLR. Comparison of transcriptional profiles between wild type and *myd88*$^{-/-}$ embryos at three different time points showed that only a small amount of genes was consistently changed, the changed genes including *myd88* itself. This latter observation suggests either a feedback loop of *myd88* to itself or a strongly reduced stability of the mutated *myd88* mRNA.

Raising the *myd88*$^{-/-}$ mutants to adulthood has proven difficult. When we started this investigation adults died as soon as they were fin clipped for genotyping. Therefore we used the wild type and mutant siblings from the offspring of heterozygous adults. Since the immersion assay was shown not to be suitable for analysis on the level of single embryos and did not produce a clear immune response, we used intravenous injection for characterization of the immune response of this mutant. qPCR analysis was performed on the genes *mmp9*, *irg1l*, *il1b* and *ifn*. This showed that *mmp9* is clearly dependent on Myd88 and *irg1l* and *il1b* are at least partially dependent. For *ifn* it was not possible to draw any conclusions, since this gene was up-regulated in only one of the wild type embryos.

The embryos used for the qPCR analysis were then subjected to microarray analysis. As expected, expression of many immune-related genes was shown to be dependent on Myd88. Making use of the presence of multiple probes for several immune-related genes on the array, we could confidently identify

several of them as being (partially) m*yd88*-dependent. Confirming previous results obtained with m*yd88* morpholino knockdown and *Salmonella* infection, we found *mmp9*, *il1b* and *irak3* expression to be *myd88*-dependent. Several other genes such as transcription factor genes (e.g. *junb*), signal transduction genes (e.g. *rip2k*), cytokine genes (e.g. *tnfb*), chemokine genes (e.g. *cxcl-c1c*) and other immune response genes (e.g. *ncf1*) were shown to be m*yd88*-dependent. These results are in line with findings in *MyD88*-deficient mice. Several genes that were shown to be (partially) m*yd88*-dependent, were previously shown to be *traf6*-dependent in a morpholino knockdown study. These included *mmp9*, *il1b*, *tnfb*, *ncf1*, zgc:*114032* (complement factor B), zgc:*112143* (TNFα-induced protein 9), and *tlr5b*. These genes can now be placed as targets of a *myd88-traf6* signalling pathway.

The most surprising observation was the difference in expression between *tnfa* and *tnfb* in the *myd88*$^{-/-}$ mutant. In contrast to *tnfb*, which was shown to be completely dependent on Myd88, *tnfa* was completely independent on Myd88. *Tnfa* and *tnfb* are two copies in the zebrafish genome of the gene encoding TNF in the mouse and human. While *tnfb* reflects the same dependence on Myd88 as human TNF, *tnfa* does not, possibly reflecting the functional divergence of these homologous genes.

The observation that several genes (such as *il1b*) are partially dependent on Myd88 may be due to tissue specific differences. Since in these experiments RNA was isolated from whole embryos, differences between different tissues are lost. In septic wild type and *Myd88*-deficient mice for example, striking differences were found between the transcriptional response of spleen and liver. The response of the liver showed a strong dependence on Myd88, while the response of the spleen was largely independent on Myd88.

In this chapter, we have shown that the *myd88*$^{-/-}$ mutant zebrafish is severely impaired in its innate immune response towards *E. tarda*. Studies using this zebrafish mutant with other infecting agents are already on its way. Initial results show that *myd88* plays an important role in the control of a *M. marinum* infection. Since the problems with raising the *myd88*$^{-/-}$ mutant zebrafish to adulthood are now resolved, the full scale possibilities of the zebrafish mutant, as opposed to a morpholino knockdown, can be taken advantage of. Furthermore, increased efficiency of the TILLING-screens due to high throughput sequencing techniques will lead to the identification of more

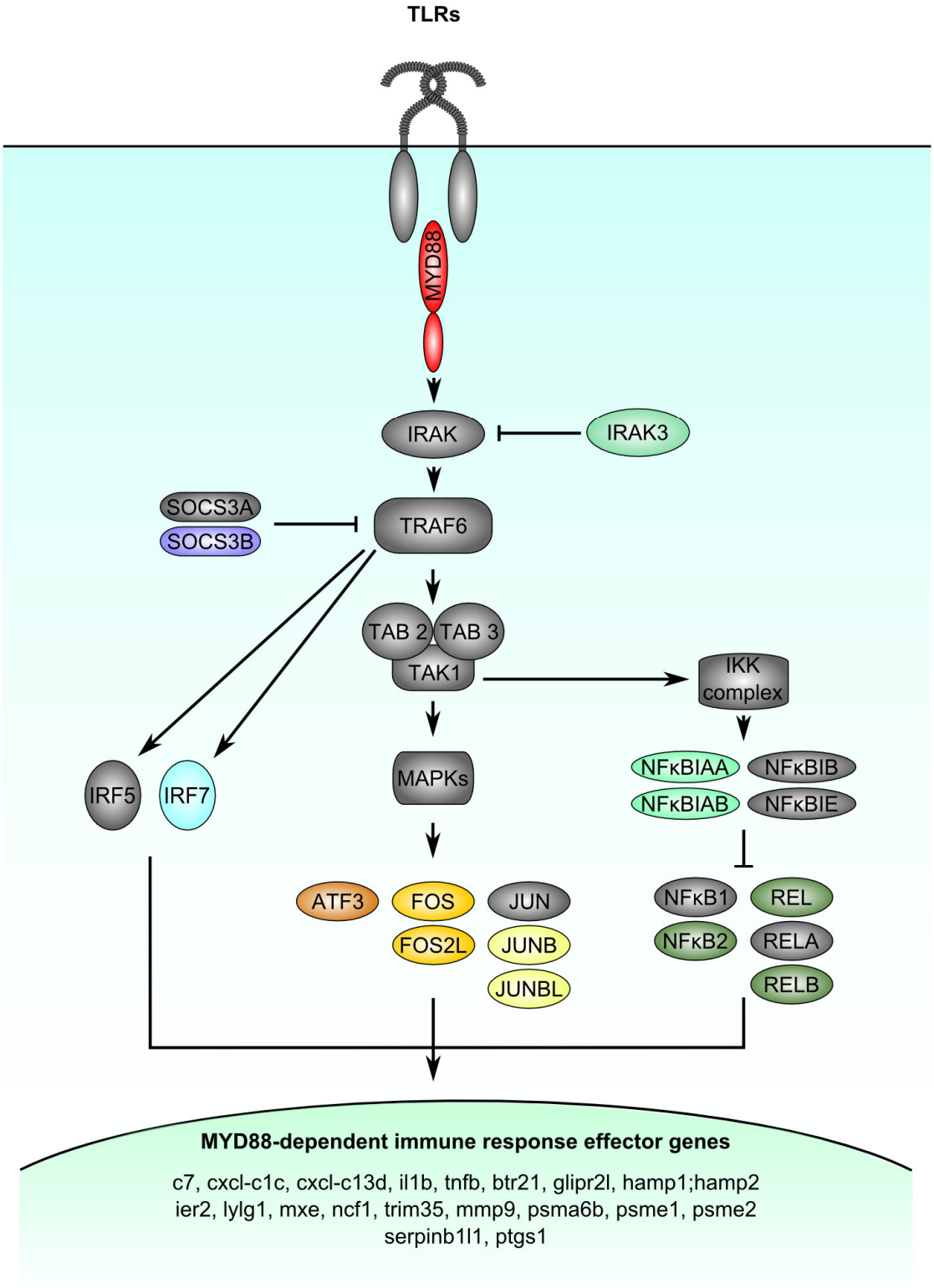

Figure 1. Schematic representation of the Tlr-Myd88 dependent pathway. Transcriptional regulation of components pictured in color was shown to be Myd88-dependent in this thesis, as well as all of the immune response effector genes shown. The pathway shown is based on knowledge of TLR signalling in mammalian species and it should be noted that most interactions remain to be experimentally confirmed in zebrafish.

immune-related mutants soon. In fact, a mutant in *tirap*, encoding another adaptor protein of the TLRs that cooperates with Myd88, has already been identified. The ambition of the Sanger Institute is that within a couple of years zebrafish mutants for each gene will be available. In addition to the TILLING-screens, zinc finger technology for the zebrafish is fast emerging, making targeted gene knock-out and gene knock-in possible. These developments will allow the zebrafish to quickly catch up with the mouse.

Conclusion

In conclusion, the research described in this thesis provides a strong basis for future research. The genome of *E. tarda* can provide insight in differences between pathogenic and non-pathogenic strains, as well as in broad host-range and more host-specific *Edwardsiella*-strains. The zebrafish embryo model will be especially suitable to study the interaction of *E. tarda* virulence factors with the host innate immune system, where comparison with the response to *E. ictaluri* infection will also be of interest. Unfortunately, we found that *E. tarda* is not completely suitable for high throughput screening, since the immersion assay is not reproducible enough, while on the other hand it is too pathogenic to use in automated yolk injection. Still, our studies into the immersion system gave many interesting leads for studying the epithelial response of the zebrafish, not only in response to *E. tarda*, but also with regard to the *P. aeruginosa* strains. Not only the epithelial response is of interest, but also the bacterial components that cause this response.

However, the characterization of the *myd88*$^{-/-}$ mutant will probably be the most important legacy of this thesis. It is the first zebrafish knock-out mutant in a gene central in the innate immune system. In the future, studies with many different pathogens and with different cancer models can be done, revealing the complete role of *myd88* in the zebrafish, with possible novel insights in the function of the innate immune system.