Cover Page

The handle http://hdl.handle.net/1887/32962 holds various files of this Leiden University dissertation.

Author: Veneman, Wouter Jurjen
Title: Developing systems for high-throughput screening of infectious diseases using zebrafish
Issue Date: 2015-05-12

Developing systems for high-throughput screening of infectious diseases using zebrafish

Wouter J. Veneman

Developing systems for high-throughput screening

of infectious diseases using zebrafish

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 12 mei 2015
klokke 13.45 uur

door

Wouter Jurjen Veneman
geboren te Almelo
in 1987

Promotiecommissie:

Promotor: Prof. dr. H.P. Spaink

Co-promotor: Dr. A.H. Meijer

Co-promotor: Dr. S.A.J. Zaat   (Academic Medical Center/UVA)

Overige leden: Prof. dr. S.A. Renshaw   (The University of Sheffield)

                        Prof. dr. J. den Hertog

                        Prof. dr. G.P. van Wezel

                        Dr. ir. F.J. Verbeek

Table of contents

Chapter 1

Outline of the thesis

Public health and infectious diseases

Biomaterial associated infections

Between the year 2000 and 2050 the percentage of people over the age of 60 worldwide is expected to rise from 11% to 22%. The total number of this population is estimated to grow to 2 billion in this same period. Eighty percent of these older people will live in low and middle income countries. This particular group of in low and middle income countries will have a 3 times higher chance of lower quality of life due to noncommunicable diseases such as cancer, diabetes, osteoporosis and heart diseases [1-5]. This group of rather vulnerable people will probably receive medical treatment once or more in their old age. With many procedures patients will get in contact with biomaterials, ranging from infusion needles to implants, such as pacemakers or artificial joints. These biomaterials as shown in Figure 1 are often associated with infections by the skin bacterium *Staphylococcus epidermidis* [6-12]. Additionally, infections with (multi-drug resistant) *Staphylococcus aureus* strains are a common problem in hospitals. The first publications about biomaterial-associated infections (BAI) go back to the early 1950's by Elek et al. [13]. They discovered higher infection rates of *S. aureus* inoculum on a soaked suture wire compared to standard inoculum in human volunteers. In more recent studies, the percentages of infection related to implantation are approximately between 1.2 and 53% [14-18]. Since materials are being implanted daily in many patients worldwide as shown in Table 1, the total number of resulting infections is very large. There would not be a problem if these infections were easy to treat, however
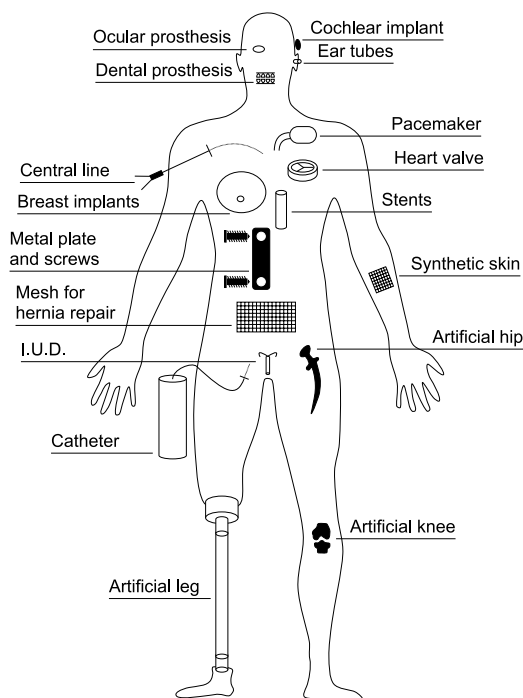

Figure 1: Overview of commonly implanted biomaterials and medical devices. This figure illustrated the variety of medical implants being used these days.

Table 1: List with 11 most implanted medical devices in the U.S.A. (data adapted from 24/7 Wallst 2011-07-11) [23].

#	Description	Number of procedures	Total cost annual	Average cost per procedure
1	Artificial eye lenses	2.582 million	$8-10 billion	$3.200-4.500
2	Ear tubes	715.000	$1-2 billion	$1.000-4.500
3	Coronary stents	560.000	$7,5 billion	$13.000
4	Artificial knees	543.000	$12 billion	$22.000
5	Metal pins, screws, plates and rods	453.000	$4,5 billion	$2.000-20.000
6	Intra-Uterine Devices	425.000	$340 million	$800
7	Spine screws, rods and artificial discs	413.000	$10 million	$25.000
8	Breast implants	366.000	$992 million	$3.351
9	Pacemakers	235.567	$4,5 billion	$20.000
10	Artificial hips	230.000	$10,5 billion	$45.000
11	Implantable Cardioverter Defibrillators	133.262	$5,5 billion	$40.000

S. epidermidis will adhere to the implanted material where it forms a biofilm. From there it will also infect the surrounding tissue and host cells, even after the material has been removed as has been shown in mice [10, 19]. This makes treatment of BAI more difficult, since some antibiotics like vancomycin cannot penetrate the infected tissue efficiently [20]. This can result in many complications with in the worst case scenario the necessity of the chirurgical removal of the implant. To prevent BAI we hope to find new molecular factors that could help understand how bacteria growing at or near implanted materials interact with the host. Although most research regarding BAI is performed with mammalian models such as mice or goats [21, 22], we used the zebrafish larvae as model as described in more detail below.

Tuberculosis and the link with metabolic syndromes

An infectious disease not related with biomaterials but forming another major societal problem is tuberculosis (TB), which is caused by *Mycobacterium tuberculosis* (Mtb). Mtb is one of the most successful human pathogens and is capable of establishing infection despite the host innate and adaptive immune responses. Each year 9 million people develop TB and over 1,5 million die from it. A third of the global human population is estimated to be latently infected, and has a 5-10% lifetime risk of developing TB reactivation disease [24]. Therefore finding better treatments and diagnostic tools for this disease is essential. With the upcoming antibiotic resistance for the first line antibiotics including Isoniazid (1952), Rifampicin (1966), Pyrazinamide (1952) and Ethambutol (1961), there is a great need for new antibiotics [25]. The World

Health Organization (WHO) reported that already over 300.000 people were reported with multi drug-resistant tuberculosis (MDR-TB) in 2013. Extensively drug-resistant tuberculosis (XDR-TB), bacteria are resistant also to second line drugs, and has been reported by 100 countries in 2013. It is estimated that 9% of the people diagnosed with MDR-TB have in fact XDR-TB [26].

Although *in vitro* models can provide detailed information of how TB bacteria are taken up and survive in macrophages, it does not reflect the infection *in vivo*. Unfortunately mice have limitations as a model to study the pathogenesis of TB, since they do not form the characteristic granulomas with necrotic centres after infection. Guinea pigs and rabbits do show the granuloma formation, however they are relatively expensive and do not allow advanced genetic approaches, real time imaging or high-throughput screening. The zebrafish larvae infection system (this thesis) can with the use of the natural pathogen *M. marinum*, a close relative of Mtb, model the early steps in human TB disease progression very well. After infection of zebrafish embryos with mycobacteria the innate immune cells will form granulomas, with in later stages of development also the formation of necrotic centres [27]. As outlined below zebrafish larvae allow the use of advanced genetic and real time imaging technologies and are excellently suited for high-throughput screening.

Although TB is already a major health issue and difficult to treat, it becomes more problematic when people are also infected with the human immunodeficiency virus (HIV) [28]. The WHO estimated that people have a 26 to 31 times higher change to develop TB when already diagnosed with HIV [29].

Unfortunately not only HIV is a great risk in combination with TB, also obesity is linked to TB [30-32]. Obesity and overweight are defined as abnormal fat accumulation which is a great health risk. It is also a large risk factor for example for diabetes, cardiovascular diseases and cancer [33, 34]. Although commonly perceived as a problem in high income countries, its incidence is now also rising in low- and middle-income countries. In 2008 more than 1.4 billion adults globally were overweight and more than half a billion were obese. Obesity among children is ever so problematic since already 42 million children worldwide are overweight who will likely become obese adults, and are more susceptible to diabetes and cardiovascular diseases [35]. Many treatments are known, however, due to ethnic backgrounds and different lifestyles globally, it is a challenge to develop new strategies to prevent obesity [36].

A particular gene that regulates appetite, called leptin, has been well described in the context of metabolic processes [37] and also functions in immunity and inflammation [38, 39]. Overnutrition and obesity are correlated with high concentrations of leptin. Having high levels of leptin can lead to leptin resistance. This in turn can lead to increased TNF-α production, altered T cell subset ratios and repressed T cell response [40], with as a result, higher incidence of infectious diseases. This suggests that leptin resistance would be disadvantageous in the context of TB treatment, and there are scientific papers that could confirm this statement [41, 42]. However there are also few papers indicating the opposite, that there is a positive effect on TB progression due

to obesity [43]. The role of leptin has been described in the context of TB and obesity [30] but clear evidence that obesity as a result of leptin resistance could lower the progression of TB has not been found yet.

Zebrafish as a model for innate immunity research of human diseases

The zebrafish finds it origin in the Himalayan region and it is a popular aquarium fish. The first scientific experiments with this organism date from the early 1970's by Dr. George Streisinger [29], who used it for studies of vertebrate development and genetics. Since then it showed to be a versatile model for studies of human diseases such as cancer, infectious diseases, cardiovascular disease, diabetes, haematopoiesis, and neural disorders [44-46]. Although zebrafish do not have lungs and are cold blooded they have high similarity with mammalians from a genetic and developmental perspective. Since the larvae are very small and transparent, they are easy to follow in time, also due their fast development. The first innate immune cells are present at 1 day post fertilization (dpf). One adult couple can produce up to 300 eggs per week and the larvae are therefore very well suited for high-throughput screens. The availability of many mutant and transgenic strains makes it possible for example to follow live fluorescently labelled macrophages or neutrophils interact with fluorescently labelled pathogens inside a larva. Because they are so small, housing costs per animal are low. Many tools are available to identify genes involved in specific biological processes, such as random mutagenesis or targeted mutagenesis using zinc-finger nucleases or the CRISPR-Cas system [47, 48]. Also, using a morpholino approach, which is a ~25 base artificial oligonucleotide, injected into the one cell stage, specific gene function can be knocked down for the first couple of days of embryonic development [49, 50]. Altogether, this has led to the common use of zebrafish larvae in biomedical research. In addition, the model is better suited for translational studies than invertebrate animals since over 80% of genes linked with human disease have at least one obvious zebrafish orthologue [51]. Therefore it holds a very strong position as a screening model with a high-throughput level in between cell and tissue culture systems and higher animal models.

Outline of the thesis

The work described in this thesis focusses on the establishment and improvement of automated and high-throughput techniques for immunological studies in zebrafish larvae. We studied the innate immune host response towards *S. epidermidis*, a normally innocent human skin bacterium that is one of the main causes of biomaterial-associated infections. To this end, we used transcriptome analysis tools such as micro-arrays and RNA deep sequencing (RNAseq). To compare the results of infection with *S.*

epidermidis with the response to a natural fish pathogen, we performed simultaneously experiments with *M. marinum*.

Chapter 2 describes a broad overview of all the applications of injecting automatically with a robotic micro-injector. These include the generation of transgenic animals by injecting DNA into the yolk of a one-cell stage fertilized egg, or the knockdown of particular genes by injection of morpholinos. Furthermore, injections of opportunistic or pathogenic bacteria as well human cancer cells are described in detail. Analyses of the large amounts of injected eggs are performed with either automated confocal laser scanning microscopy (CLSM) or high-throughput Complex Object Parametric Analysis and Sorting (COPAS XL).

Continuing on chapter 2 we optimized our high-throughput screening system using video recorded procedures as shown in **chapter 3**. Here we show in detail how multiple devices such as a large spawning tank, an automated micro-injector, a COPAS flow-cytometer, and a Vertebrate Automated Screening Technology (VAST BioImager) can yield a combination of a high-throughput low-resolution with a medium-throughput high-resolution screening technology. Furthermore we optimized the injection method by selecting the best developmental stage for early yolk injections with *S. epidermidis* and a highly and medium virulent *M. marinum* strain.

Chapter 4 is focused on the pathogenesis of *S. epidermidis* in zebrafish larvae. This was the first scientific report of injecting *S. epidermidis* into zebrafish larvae. We performed extensive imaging and automated the monitoring of bacterial burden progression in live larvae using the COPAS XL. It was found that injection of *S. epidermidis* into the caudal vein did not lead to an infection, however when injected into the yolk of a 16-128 cell stage embryo, we did manage to reproducibly create an infection spreading into the entire body of the larvae. Time course analysis of transcriptome host responses was performed using micro-arrays. This led to the discovery of a group of host genes that were differentially expressed after infection by *S. epidermidis* and not by *M. marinum*. The validation of the micro-array results was performed using a newer technique, the so called RNAseq. Since transcriptome analysis using micro-arrays and RNAseq can provide a lot of information on how the host reacts to an invasive agent such as a bacterium, it can reveal possible targets for new treatment strategies.

Although the RNAseq method is very promising, the analysis still relies on expensive 'black box' software packages or dedicated programming skills of the researcher. However the second option could be rather difficult for a biologist untrained in informatics. In **chapter 5** we describe the results of collaboration between bioinformaticians, statisticians and biologists. This led to a straight forward and easy to use software package called GeneTiles, for the analysis and visualization of large RNAseq datasets.

One of the main advantages is that is runs on a server with a user-friendly interface. This means that is can be accessed from computers with an internet connection and different operations systems such as Windows and Linux can be used. The implantation of additional software packages such as DEXSeq has led to the discovery of differentially expressed genes under infectious conditions. Also a direct visualization of the differential expression in a variety of biological pathways obtained from Wikipathways leads to a fast interpretation of functional data.

In **chapter 6** we present a time course analysis of the distribution of nanometer and micrometer sized polystyrene particles after zebrafish yolk injection. This could help us to develop a screening model for biomaterial-associated infection using zebrafish larvae. We determined the parameters for the diameter and shape of the particles that determine the possibility to inject these particles using glass micro capillaries into zebrafish larvae. The size is also very important for the distribution inside the larvae, since we found that the smaller sized particles spread more compared to the bigger particles.

In **chapter 7** the results are summarized and discussed in the context of possible applications of the results for biomedical purposes.

References

1. Mathers CD, Stevens GA, Boerma T, White RA, Tobias MI: Causes of international increases in older age life expectancy. The Lancet 2014.

2. Prince MJ, Wu F, Guo Y, Gutierrez Robledo LM, O'Donnell M, Sullivan R, Yusuf S: The burden of disease in older people and implications for health policy and practice. The Lancet 2014.

3. Chatterji S, Byles J, Cutler D, Seeman T, Verdes E: Health, functioning, and disability in older adults—present status and future implications. The Lancet 2014.

4. Steptoe A, Deaton A, Stone AA: Subjective wellbeing, health, and ageing. The Lancet 2014.

5. Bloom DE, Chatterji S, Kowal P, Lloyd-Sherlock P, McKee M, Rechel B, Rosenberg L, Smith JP: Macroeconomic implications of population ageing and selected policy responses. The Lancet 2014.

6. Zimmerli W, Sendi P: Pathogenesis of implant-associated infection: the role of the host. Semin Immunopathol 2011, 33(3):295-306.

1

7.   Zimmerli W: Implanted devices: biocompatibility, infection and tissue engineering. Semin Immunopathol 2011, 33(3):219-220.

8.   Anderson JM, Rodriguez A, Chang DT: Foreign body reaction to biomaterials. Semin Immunol 2008, 20(2):86-100.

9.   Broekhuizen CA, Schultz MJ, van der Wal AC, Boszhard L, de Boer L, Vandenbroucke-Grauls CM, Zaat SA: Tissue around catheters is a niche for bacteria associated with medical device infection. Crit Care Med 2008, 36(8):2395-2402.

10.   Broekhuizen CA, de Boer L, Schipper K, Jones CD, Quadir S, Vandenbroucke-Grauls CM, Zaat SA: Staphylococcus epidermidis is cleared from biomaterial implants but persists in peri-implant tissue in mice despite rifampicin/vancomycin treatment. J Biomed Mater Res A 2008, 85(2):498-505.

11.   Broekhuizen CA, Sta M, Vandenbroucke-Grauls CM, Zaat SA: Microscopic detection of viable Staphylococcus epidermidis in peri-implant tissue in experimental biomaterial-associated infection, identified by bromodeoxyuridine incorporation. Infect Immun 2010, 78(3):954-962.

12.   Zaat S, Broekhuizen C, Riool M: Host tissue as a niche for biomaterial-associated infection. Future Microbiol 2010, 5(8):1149-1151.

13.   Elek SD, Conen PE: The Virulence of Staphylococcus-Pyogenes for Man - a Study of the Problems of Wound Infection. Brit J Exp Pathol 1957, 38(6):573-586.

14.   Cocke WM: A Critical-Review of Augmentation Mammapalsty with Saline-Filled Prostheses. Ann Plas Surg 1994, 32(3):266-269.

15.   Pittet B, Montandon D, Pittet D: Infection in breast implants. The Lancet Infectious Diseases 2005, 5(2):94-106.

16.   Wang JT, Wang AY, Psarros C, Da Cruz M: Rates of revision and device failure in cochlear implant surgery: a 30-year experience. Laryngoscope 2014, 124(10):2393-2399.

17.   Engle MP, Vinh BP, Harun N, Koyyalagunta D: Infectious complications related to intrathecal drug delivery system and spinal cord stimulator system implantations at a comprehensive cancer pain center. Pain Physician 2013, 16(3):251-257.

18.   Rodriguez DJ, Afzal A, Evonich R, Haines DE: The prevalence of methicillin resistant organisms among pacemaker and defibrillator implant recipients. Am J Cardiovasc Dis 2012, 2(2):116-122.

19. Broekhuizen CA, de Boer L, Schipper K, Jones CD, Quadir S, Feldman RG, Dankert J, Vandenbroucke-Grauls CM, Weening JJ, Zaat SA: Peri-implant tissue is an important niche for Staphylococcus epidermidis in experimental biomaterial-associated infection in mice. Infect Immun 2007, 75(3):1129-1136.

20. Stevens DL: The role of vancomycin in the treatment paradigm. Clin Infect Dis 2006, 42 Suppl 1:S51-57.

21. Broekhuizen CA, de Boer L, Schipper K, Jones CD, Quadir S, Feldman RG, Vandenbroucke-Grauls CM, Zaat SA: The influence of antibodies on Staphylococcus epidermidis adherence to polyvinylpyrrolidone-coated silicone elastomer in experimental biomaterial-associated infection in mice. Biomaterials 2009, 30(32):6444-6450.

22. Custers RJ, Dhert WJ, Saris DB, Verbout AJ, van Rijen MH, Mastbergen SC, Lafeber FP, Creemers LB: Cartilage degeneration in the goat knee caused by treating localized cartilage defects with metal implants. Osteoarthritis Cartilage 2010, 18(3):377-388.

23. The Eleven Most Implanted Medical Devices In America. 24/7 Wallst 2011.

24. Sulis G, Roggi A, Matteelli A, Raviglione MC: Tuberculosis: Epidemiology and Control. Mediterr J Hematol Infect Dis 2014, 6(1):e2014070.

25. Gunther G: Multidrug-resistant and extensively drug-resistant tuberculosis: a review of current concepts and future challenges. Clin Med 2014, 14(3):279-285.

26. World-Health-Organization: GLOBAL TUBERCULOSIS REPORT 2014. 2014.

27. Hosseini R, Lamers GE, Hodzic Z, Meijer AH, Schaaf MJ, Spaink HP: Correlative light and electron microscopy imaging of autophagy in a zebrafish infection model. Autophagy 2014, 10(10):1844-1857.

28. Newman M, Foisy MM, Ahmed RA: The Use of Therapeutic Drug Monitoring in Complex Antituberculous and Antiretroviral Drug Dosing in HIV-Tuberculosis Coinfected Patients. J Int Assoc Provid AIDS Care 2014.

29. World-Health-Organization: Global update on the health sector response to HIV. 2014.

30. Zheng Y, Ma AG, Wang QZ, Han XX, Cai J, Schouten EG, Kok FJ, Li YC: Relation of Leptin, Ghrelin and Inflammatory Cytokines with Body Mass Index in Pulmonary Tuberculosis Patients with and without Type 2 Diabetes Mellitus. PLoS One 2013, 8(11).

31. Kumar NP, Banurekha VV, Nair D, Sridhar R, Kornfeld H, Nutman TB, Babu S: Coincident pre-diabetes is associated with dysregulated cytokine responses in pulmonary tuberculosis. PLoS One 2014, 9(11):e112108.

32.   Skowronski M, Zozulinska-Ziolkiewicz D, Barinow-Wojewodzki A: Tuberculosis and diabetes mellitus - an underappreciated association. Arch Med Sci 2014, 10(5):1019-1027.

33.   Gong J, Robbins LA, Lugea A, Waldron RT, Jeon CY, Pandol SJ: Diabetes, pancreatic cancer, and metformin therapy. Front Physiol 2014, 5:426.

34.   Triggle CR, Ding H: Cardiovascular impact of drugs used in the treatment of diabetes. Ther Adv Chronic Dis 2014, 5(6):245-268.

35.   10 Facts on Obesity [http://www.who.int/features/factfiles/obesity/en/]

36.   Caprio S, Daniels SR, Drewnowski A, Kaufman FR, Palinkas LA, Rosenbloom AL, Schwimmer JB: Influence of race, ethnicity, and culture on childhood obesity: implications for prevention and treatment: a consensus statement of Shaping America's Health and the Obesity Society. Diabetes Care 2008, 31(11):2211-2221.

37.   Lopez-Jaramillo P, Gomez-Arbelaez D, Lopez-Lopez J, Lopez-Lopez C, Martinez-Ortega J, Gomez-Rodriguez A, Triana-Cubillos S: The role of leptin/adiponectin ratio in metabolic syndrome and diabetes. Horm Mol Biol Clin Investig 2014, 18(1):37-45.

38.   Paz-Filho G, Mastronardi C, Franco CB, Wang KB, Wong ML, Licinio J: Leptin: molecular mechanisms, systemic pro-inflammatory effects, and clinical implications. Arq Bras Endocrinol Metabol 2012, 56(9):597-607.

39.   Matarese G, Moschos S, Mantzoros CS: Leptin in immunology. J Immunol 2005, 174(6):3137-3142.

40.   Schaible UE, Kaufmann SH: Malnutrition and infection: complex mechanisms and global impacts. PLoS Med 2007, 4(5):e115.

41.   Wang HT, Zhang J, Ji LC, You SH, Bai Y, Dai W, Wang ZY: Frequency of tuberculosis among diabetic patients in the People's Republic of China. Ther Clin Risk Manag 2014, 10:45-49.

42.   Karlsson EA, Beck MA: The burden of obesity on infectious disease. Exp Biol Med (Maywood) 2010, 235(12):1412-1424.

43.   Leung CC, Lam TH, Chan WM, Yew WW, Ho KS, Leung G, Law WS, Tam CM, Chan CK, Chang KC: Lower risk of tuberculosis in obesity. Arch Intern Med 2007, 167(12):1297-1304.

44.   Meijer AH, van der Vaart M, Spaink HP: Real-time imaging and genetic dissection of host-microbe interactions in zebrafish. Cell Microbiol 2014, 16(1):39-49.

45. Brittijn SA, Duivesteijn SJ, Belmamoune M, Bertens LF, Bitter W, de Bruijn JD, Champagne DL, Cuppen E, Flik G, Vandenbroucke-Grauls CM et al: Zebrafish development and regeneration: new tools for biomedical research. Int J Dev Biol 2009, 53(5-6):835-850.

46. Spaink HP, Cui C, Wiweger MI, Jansen HJ, Veneman WJ, Marin-Juez R, de Sonneville J, Ordas A, Torraca V, van der Ent W et al: Robotic injection of zebrafish embryos for high-throughput screening in disease models. Methods 2013, 62(3):246-254.

47. Foley JE, Maeder ML, Pearlberg J, Joung JK, Peterson RT, Yeh JRJ: Targeted mutagenesis in zebrafish using customized zinc-finger nucleases. Nat Protoc 2009, 4(12):1855-1868.

48. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JR, Joung JK: Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol 2013, 31(3):227-229.

49. Bedell VM, Ekker SC: Using engineered endonucleases to create knockout and knockin zebrafish models. Methods Mol Biol 2015, 1239:291-305.

50. Bedell VM, Westcot SE, Ekker SC: Lessons from morpholino-based screening in zebrafish. Brief Funct Genomics 2011, 10(4):181-188.

51. Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren K, Matthews L et al: The zebrafish reference genome sequence and its relationship to the human genome. Nature 2013, 496(7446):498-503.

1

Chapter 2

2

Robotic injection of zebrafish embryos for high-throughput screening in disease models

**Herman P. Spaink[1], Chao Cui[2], Malgorzata I. Wiweger[1,2],
Hans J. Jansen[2], Wouter J. Veneman[1], Rubén Marín-Juez[2],
Jan de Sonneville[3], Anita Ordas[1], Vincenzo Torraca[1],
Wietske van der Ent[1], William P. Leenders[4], Annemarie H. Meijer[1],
B. Ewa Snaar-Jagalska[1], Ron P. Dirks[2]**

[1] Institute of Biology Leiden, Leiden University
[2] ZF-screens BV
[3] Life Science Methods BV
[4] Radboud University Nijmegen Medical Center

Abstract

The increasing use of zebrafish larvae for biomedical research applications is resulting in versatile models for a variety of human diseases. These models exploit the optical transparency of zebrafish larvae and the availability of a large genetic tool box. Here we present detailed protocols for the robotic injection of zebrafish embryos at very high accuracy with a speed of up to 2000 embryos per hour. These protocols are benchmarked for several applications: (1) the injection of DNA for obtaining transgenic animals, (2) the injection of antisense morpholinos that can be used for gene knock-down, (3) the injection of microbes for studying infectious disease, and (4) the injection of human cancer cells as a model for tumor progression. We show examples of how the injected embryos can be screened at high-throughput level using fluorescence analysis. Our methods open up new avenues for the use of zebrafish larvae for large compound screens in the search for new medicines.

Introduction

The use of zebrafish as an animal model has an abundance of applications in fundamental research in vertebrate development, physiology and toxicology [1-5]. More recently, this model has also been shown to be highly applicable for studies of many types of disease [6-20]. The benefits of the relatively small and transparent larvae for optical imaging using transgenic fish lines expressing many colour varieties of the GFP protein have been widely exploited in these disease studies [6]. Currently the genetic tool box is comprised of a large variety of gene knock-down or knock-out systems [21-28]. Additionally many genomic-based techniques such as RNA deep sequencing, metabolomics and proteomics have been applied to zebrafish [29-39]. A comparison of parallel deep RNA sequencing and proteome analysis has been reported (Palmblad et al., submitted). The fact that the innate immune system of zebrafish is highly similar to that of mammals and is already fully functional as early as two days after fertilization makes zebrafish larvae extremely useful for studies of diseases related to the immune system [40, 41]. Examples given below are studies of cancer progression and infectious diseases caused by many bacteria, fungi or viruses.

Zebrafish micro-injection and screening tools

Micro-injection of zebrafish embryos is an essential technology for the following applications:

• The generation of transgenic zebrafish lines.

• The generation of gene knock-out lines using zinc fingers or TALEN technology.

• Gene knock-down using morpholinos, siRNA or antibodies.

• Overexpression of genes by injection of mRNA.

• The injection of tracer dyes or particles.

• Intraorganismal introduction of microbes in embryos or larvae for infection

 studies.

• Transplantation of cells between embryos.

• Xenograft implantation of cells for cancer studies.

Many general methods for these applications can be found in the book: Essential Zebrafish Methods: Cell and Developmental Biology [42]. More specifically for injection methods that can be used for these applications we can refer to detailed descriptions in four methodology compendia or books: (1) Zebrafish: Methods and Protocols [43], (2) The zebrafish book, A guide for the laboratory use of zebrafish (*Danio rerio*), 5th edition [44] and (3) Methods in Cell Biology issues 104 and 105 [45, 46], which provide comprehensive laboratory protocols and reviews for recent zebrafish methods related to disease models and chemical screens. Recently also video enhanced protocols for zebrafish micro-injection have been published, for instance by Benard et al. [47], describing in detail the application of micro-injection of bacterial pathogens. In brief, all methods use thin glass capillary needles to introduce compounds or biological materials inside various parts of embryos and larvae. In the embryonic and early larval stages the transparency and softness of the tissues warrants a high success rate of the injection protocol. In contrast, injection in the later larval stages is more difficult due to higher rigidity of tissue and can currently only be performed at relatively low-throughput. For the first four applications mentioned above, DNA, RNA or morpholinos can be injected into the yolk of early stage embryos. Due to the relatively large size of the yolk this offers a fast procedure for micro-injection that even can be completely automated. Robotic injection of zebrafish embryos using image recognition has been shown to accurately deliver morpholinos at a throughput level of 25 consecutive injections per run of 2 min [48]. Recently an alternative robotic injection method was shown to inject embryos at a speed of 2000 per hour with a success rate of 99% [49]. This method makes use of specially designed grids where embryos occupy the hemi-spherical wells of an agarose cast in a centred and completely reproducible manner, with the cell mass always resting to the side. In this paper we present further applications and detailed methods for the use of this robotic injection system. We will also provide examples of high-throughput screening of injected embryos. Screening of zebrafish embryos can make use of the rapid technical advances in high-throughput analysis methods for zebrafish embryos [6, 50, 51]. The COPAS XL (Complex Object Parametric Analyzer and Sorter) system (Union Biometrica) can be used for fluorescence imaging of zebrafish embryos at a throughput level of 200 embryos per minute. This system has been designed for the analysis, sorting and dispensing of objects up to 1.5 mm in diameter based on size,

optical density and fluorescence intensity. A profiler option simultaneously detects and analyzes up to 8000 data points per object for each of the channels of extinction and fluorescence, and includes advanced imaging options. The resulting profiles can be used to set parameters for zebrafish larvae to be sorted in 96-wells plates. In this paper we describe software to process the recorded data for further statistical analysis. Recently a Vertebrate Automated Screening Technology (VAST) with cellular-resolution and parallel animal processing has been reported in which the screening throughput is limited only by the image acquisition speed rather than by the fluidic or mechanical processes [52, 53]. In the near future we will also use this methodology for zebrafish high content image analysis aimed at disease screening at high-throughput such as discussed below.

Applications of micro-injection for studies of infectious disease

A growing list of pathogenic bacteria, filamentous fungi, yeasts, microsporidia, helminths, trypanosomes and viruses has been used for experimental infection studies in zebrafish, as detailed in several recent reviews [14, 18, 54-59]. Bacterial pathogens have been tested most frequently, as discussed by Meijer and Spaink [14] who present an overview of over 30 bacterial species for which disease studies in zebrafish have been published. In addition, zebrafish larvae have also been used to study the effects of a bacterial strain that normally do not cause a disease, *Staphylococcus epidermidis*, for the purpose of studying the effects of factors such as medical implant materials on defence mechanisms against commensal bacteria [39]. One of the most successful zebrafish disease models is the indirect study of human tuberculosis via the infection of zebrafish embryos with *Mycobacterium marinum*, as recently reviewed by Tobin, May and Wheeler [60]. The studies of *M. marinum* infection have already led to the clarification of many important processes in the life cycle of tuberculosis infection, in particular those underlying the mechanisms of granuloma formation in which the bacteria proliferate in macrophages [61, 62]. The context of the embryo's developing immune system makes it possible to study the contribution of different immune cell types to disease progression already at 1–2 days post fertilization, when functional macrophages and neutrophil develop. Furthermore, due to the clear temporal separation of innate immunity from adaptive responses, zebrafish larvae are particularly useful for dissecting the innate host factors involved in pathology. Recent studies have underscored the remarkable similarity of the zebrafish and human immune systems, which is important for biomedical applications [41]. Since conserved pathogen associated molecular pattern recognition systems are already functional at one day post fertilization, zebrafish embryos and larvae are highly suitable for rapid screening of disease progression up to the stage that feeding becomes necessary and ethical constraints become apparent [63]. For infection studies a common route of infecting zebrafish embryos is the injection of the pathogen into the caudal vein of 1 day old embryos [47]. This method is relatively labour-intensive and although it has been successfully used for drug screens [64] it compares to cellular screening technologies as a low-throughput technique, leading to major bottlenecks in

drug discovery. Since infection by immersion in most infection models is not an effective alternative [65], we sought to achieve a reliable high-throughput automatic injection system, drastically reducing the man hour requirement while vastly increasing the number of reproducibly infected embryos. As shown by Carvalho et al. [49] and Veneman et al. [39] we have successfully used robotic micro-injection technology for screening bacterial proliferation during the first 5 days of larval development. In these studies the COPAS technology mentioned above was used to monitor bacterial proliferation. Carvalho et al. [49] have also shown that this high-throughput injection method, by its versatile applicability in small high safety flow cabinets, can be used for the study of infection by dangerous human pathogens such as *Mycobacterium tuberculosis* owing to the fact that this bacterial species can survive within macrophages of zebrafish larvae.

Applications of micro-injection for studies of cancer progression

The zebrafish is increasingly used as a model for the study of cancer progression and metastatic potential of tumor cells [16, 66-70]. The optical transparency of zebrafish larvae has permitted novel insights into mechanisms underlying tumor cell migration and the role of host recognition factors that are highly useful for the study of human cancers [71-73]. High-throughput methodologies have been applied in oncologic small molecules screens in zebrafish [74-77], however, testing has mainly been performed by addition of potential therapeutic drugs to the swimming water of zebrafish as micro-injection of drugs has been hampered by technical limitations in the throughput level. However, the use of artificial carriers that can be used for slow release of drugs has been shown to be highly applicable for future screening for efficacy or toxicity of therapeutics in zebrafish [78]. Especially promising for disease studies in general is the use of photodegradable biogels that can be used for controlled release of compounds [79]. Micro-injection technology has been extensively applied for xenografting of cancer cells and non-transformed cells, including stem cells [68, 72, 80-86]. It was shown that many cancer cell types from mammalian origin can survive for extended periods during larval development and that proliferation and spreading can be monitored using automated methods [87]. Clonal zebrafish lines which allow serial transplantations of tumor cells from one fish to another without detrimental gamma-irradiation or usage of immunosupressants can be used for studying the behaviour of xenografted tumor cells in zebrafish after an adaptive immune system has developed [88]. Various studies have explored the possibilities for injecting cancer cells in different tissues of zebrafish embryos or larvae during different developmental stages. For instance injection of mammalian cancer cells into the yolk, cardinal vein, duct of Cuvier, or hindbrain can lead to efficient proliferation and spreading, depending on which tumor cell line is used [72, 80-83, 85, 87]. It has been indicated that automatic robotic injection of tumor cells can also be performed using robotic technology [49]. In this paper we provide methods for automated injection of tumor cells in early pre-gastrulation stage zebrafish embryos and we show how this method can be used for screening of cancer cell proliferation and spreading.

2

Methodology

General injection and screening protocol

Robotic micro-injection was performed using needles from the company Qvotek (Table 1). For comparisons we have used borosilicate glass capillary needles [47], made with a Harvard Apparatus needle puller. Although these needles are perfectly suitable for manual injection, the commercial needles were needed for robotic injection in order to obtain highly reproducible results. For DNA, morpholino and bacterial injections, needles with an inner tip diameter of 10 µm were used, whereas for cancer cell injections needles with an inner diameter of 20 µm were used. For injection we used a specially designed grid that has been described previously by Carvalho et al. [49]. Zebrafish embryos were placed on a 1% agarose covered grid with a featured honeycomb pattern consisting of 1024 hemi-spherical wells (1.3 mm diameter), except for DNA injection where a new grid with 9 blocks of 100 wells was used. The embryo grid was placed in a motorized stage coupled to a controller and a motorized micro-manipulator. The loaded injection needle was placed in the Injectman II (Eppendorf). The injection height was set using a new precise z-calibration unit (Life Science Methods BV), consisting of a prism fixed next to a hole in the multi-well plate holder which offers a side-view of the needle (Qvotech) when lowered into this hole. For calibration the hole is centered above the lens and a needle is lowered into the hole until a dot appears on the screen. The user then clicks on the dot to center the needle with a precision of smaller than 4 pixels (10 µm). Subsequently, the prism is centered above the lens and the needle is moved to the focal plane of the prism. On the screen a side view of the needle appears and the user clicks on the tip of the needle to calibrate the needle height with equal precision. This 3D datum position is then used to calculate the injection point in the

Table 1: Specifications of capillaries (Qvotek, Mississauga, Canada) for robotic intrayolk injection of zebrafish embryos.

Parameter	Value
Start capillary	Borosilicate 100 mm x 1 mm O.D. (0,75 mm I.D.)
Tip end	Beveled and heat-formed spike
Tip angle	No
Orientation of tip opening	Top
Flexibility of tip	Rigid
Taper	Short
Inner tip diameter, bacterial injections	10 µm
Inner tip diameter, tumor cell injections	20 µm

egg with an accuracy of 10 μm in 3D. All components were connected to the controlling computer (linutop, www.linutop.com), which was equipped with a software control program written in python (Life Science Methods BV). For subsequent high-throughput screening we made use of the COPAS XL (Union Biometrica) as described previously [39, 49]. For post sorting analysis of the data we have developed a new script, written in the software package Perl, which can show the distribution of fluorescence pixels in the body of zebrafish larvae as demonstrated with data shown in the bacterial infection section (Figures. 1 and 2).
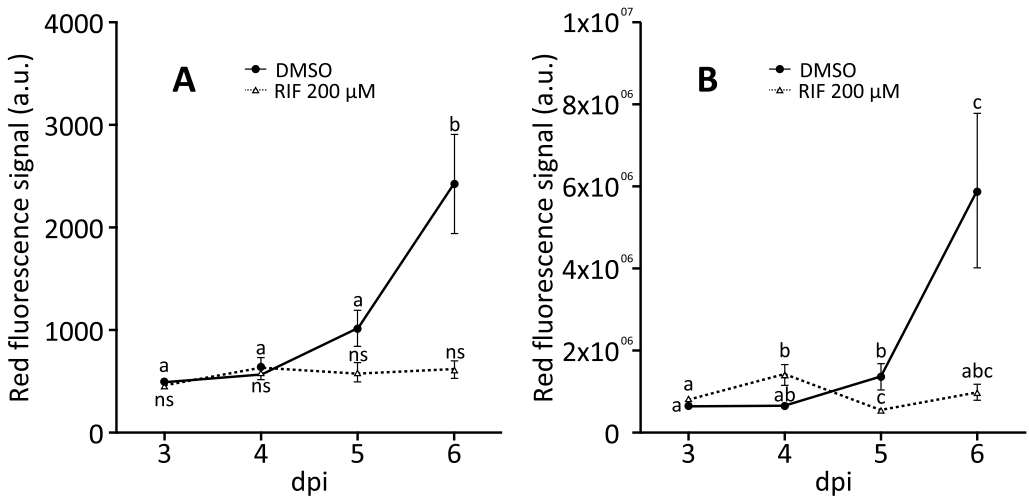
Figure 1: Progression of the bacterial burden in whole embryos (A) or the tail section of the same embryos (B). Groups of at least 50 embryos were analysed between 3 dpi to 6 dpi and red fluorescent signal was measured and shown in the graphs as mean ± S.E. Different letters indicate statistical significant differences between embryos of the same treatment ($P < 0.05$). ns, no significant differences.
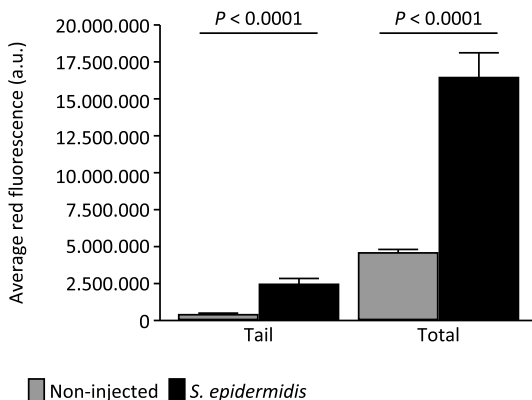

Figure 2: Quantification of fluorescence intensity of *S. epidermidis* injected embryos using the COPAS XL and PERL macro. The bar graph represents the average fluorescence intensity from 4 biological replica's (70 embryos per group) in the tail and the entire body at 5 dpi (error bars represent SEM).

Injection of DNA

We injected a DNA construct containing a GFP fusion gene (beta actin promoter-NLSmCherry-IRES-GFP) in the Gateway Tol2 vector together with Tol2 transposase RNA using standard protocols [43] in the middle of the yolk before the first cell-stage. This resulted in fluorescence outside of the yolk at 6 days post implantation (dpi) in 15% of the cases (results not shown). High-throughput injection in combination with subsequent COPAS sorting of fluorescent larvae could make up for this low yield. However, since it has been reported that DNA-constructs should be injected close to or inside the first cell of the embryo proper to improve integration efficiency, an alternative semi-automated solution was developed. A new macro was developed to move the stage to an egg, show an image of the egg, after which the user determines the best injection spot. The pointer is moved to this spot and a mouse-click initiates an injection. After the injection the macro moves the robot to the next egg. Injections were performed close to the cell and in the first cell, and yield was measured as number of embryos showing fluorescence outside the yolk divided by the total number of injected embryos at 6 dpi using a Leica MZ16FA stereo fluorescence microscope. To reduce the delay between different sets of injections, a new grid was used featuring nine blocks of 100 wells. Using a scalpel the agarose grid was cut after injections and the blocks were harvested separately. The yield of semi-automated DNA injection directed just under the first cell is shown in Table 2 demonstrating a yield (at 6 dpi) that is similar to manual injection when COPAS sorting is used (Table 2), however the throughput (<5 s per egg) as well as simplicity (one mouse click) of semi-automated injection is far greater. The semi-automated injections of DNA will be fully automated for future applications by recording a database of images together with their best injection spot. Automated image recognition can be used to find the best correlating image from this database. As such we expect to achieve a speed of 1500 eggs per hour for DNA constructs. In order to achieve this we still need to optimize the lighting as we currently use ambient light.

Table 2: Comparison of different DNA injection methods. Fully-automated and semi-automated methods show 2–3 times higher success rate per minute compared to manual methods (last column). COPAS sorting at 1 dpi can almost double (79% vs 43%) the chance of DNA expression at 6 dpi.

Method	Total inj.	% DNA-expressing embryos at 1 DPI	% DNA-expressing embryos at 6DPI	Inj. time (min)	# inj./min	Successful inj./min
Manual injection (in the cell) not sorted at 1dpi	63	79	38	25	2.5	1.0
Manual injection (yolk center) not sorted at 1dpi	120	83	29	15	8.0	2.3
Full-automated (yolk center) selected using COPAS at 1dpi	524	35	52	15	34.9	5.2
Semi-automated injection (close to the cell) not sorted at 1dpi	223	25	43	20	11.2	4.1
Semi-automated injection (close to the cell) selected using COPAS at 1dpi	223	25	79	20	11.2	2.1

Injection of bacteria and screening for bacterial infection

We previously published the development of an automatic injection system for yolk injection in zebrafish embryos [49]. In brief, at 2 h post-fertilization (hpf) or 4 hpf, zebrafish embryos were placed in the 1024 wells agarose grid and injected within half an hour. The system was successfully applied to *M. marinum* and *M. tuberculosis* injection in zebrafish embryos, which serve as a promising model for anti-tuberculosis drug screening [49]. In addition we have also applied this method for injection of *S. epidermidis* at 2 hpf. The results show that the number of injected bacteria is highly reproducible as estimated using plating of bacteria or COPAS analysis. We also have shown that COPAS analysis can reliably reproduce counting by bacterial plating thereby providing a versatile method for quantification of bacterial proliferation [39]. In addition to the capacity of the COPAS technology to sort embryos based on preset fluorescence or optical density properties, we show that it can also be useful for quantification of the distribution of injected particles using a specially designed Perl script. To demonstrate this we have analyzed the data obtained from a *M. marinum* infection experiment in which the injected larvae were analyzed for bacterial spreading in the larva in the presence or absence of the antibiotic rifampicin. Zebrafish embryos between 16 and 64 cell stage were injected with 40 cfu (colony forming units) of *M. marinum* strain E11 expressing the fluorescent protein mCherry. Three days post-injection larvae were analysed and presorted using the COPAS XL (Union Biometrica). Presorting settings were established in order to ensure a homogenous group of embryos presenting moderate infection levels, discarding non-infected and highly infected embryos. Sorted larvae were treated with either 0.2% DMSO (control) or 200 μM rifampicin. Embryos were treated and infection levels were analyzed with COPAS XL for three days. The resulting optical density and red fluorescence profiles were analysed with a custom made script. The script reads the optical density and the red fluorescence profiles from a larva. It then divides both profiles in two and sums the optical density values and red fluorescence values for each half. The anterior half of the larvae is optically more dense compared to the posterior half and will have a higher sum. The lower optical density sum will determine the posterior half of the profile. The sum of the red fluorescence values for this half is the tail fluorescence and the sum of both halves is the total fluorescence. This allows us to analyse infection levels either in the whole larvae or in the tail section. The results (Figure 1) show that bacteria can be detected with high sensitivity in the tail part of the larvae. We have also used this method for reanalysis of the previously published data set for *S. epidermidis* injection that was used for yielding a reference transcriptome database [97]. The results (Figure 2) show that injected staphyloccoci massively spread into the tail part starting two days after injection providing quantitative data that further support the images from confocal laser scanning microscopy [39].

Injection of bacteria in combination with antisense morpholinos

The conventional procedure of gene function analysis during bacterial infection in zebrafish embryos is to inject morpholino oligonucleotides at 1–2 cell stage into the yolk followed by bacterial injection into the caudal vein after the onset of blood circulation. This method is well-established but time consuming and low-throughput [89]. The labour intensity of such injections can be overcome by carrying out co-injection of morpholino and bacteria into the yolk of 1–2 cell stage zebrafish embryos using the automated injection system. This experimental setup was first tested using a morpholino oligonucleotide (GeneTools, Oregon, USA) designed against the transcription factor *Pu-1* (*Spi1*) that blocks myeloid development [49]. These results confirmed the methodology and prompted us to investigate the function of more specific immune regulatory factors such as a follow up study on myeloid differentiation factor 88 (*Myd88*), an adaptor protein in Toll-like receptor signalling that has been studied by gene knock-down and gene knock-out analysis in zebrafish embryos, and shown to be involved in proinflammatory innate immune response to microbial infection [89, 90]. For this, one nanoliter of *MyD88* morpholino at 1.6 mM concentration was co-injected with 40 cfu of *M. marinum* at the 1–2 cell stage into the yolk of zebrafish embryo. Analysis of the fluorescence profile of 5 days post-injected embryos using the COPAS technology showed an increased bacterial burden in the morphant group compared to the control group due to knock-down of this gene (Figure 3). These results mimic the already published infection studies with *Myd88* knock-down, confirming that automatic yolk co-injection of morpholino and bacteria is capable of creating reproducible results for morpholino screens in a high-throughput manner.
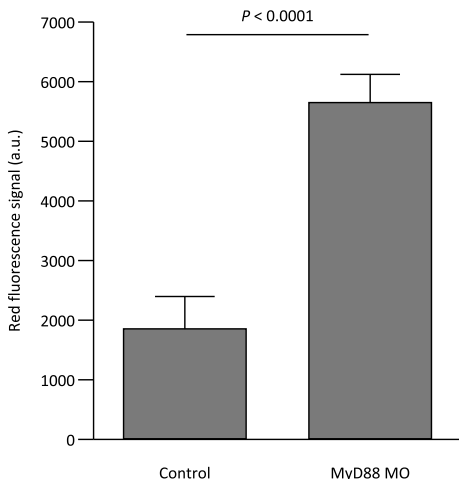

Figure 3: Automatic co-injection of antisense morpholino and bacteria into yolk. Increased bacterial burden was measured using the COPAS system in *MyD88* and control morphants at 5 dpi following co-injection of *M. marinum* with *MyD88* morpholino into 1–2 cell stage zebrafish embryos. Significant difference was tested by unpaired t-test.

Tumor cell implantation and screening for proliferation and spreading of cancer cells

1. Robotic injection of tumor cells

In manual injection experiments various tumor cell lines were shown to be highly different in their capacity to proliferate and spread into the entire zebrafish body after injection into the yolk of pre-gastrulation stage embryos. Pilot experiments using the injection robot with the highly aggressive osteosarcoma cell lines SJSA-1 (American Type Culture Collection, ATCC) and osteosarcoma line L2792 [91] showed a massive proliferation and strong spreading of fluorescent cells in the tail at 3 days post injection into the yolk at the 256 cell stage (Figure 4). For further implantation assays shown in this paper, human prostate cells (PC3 and LNCap) and a highly angiogenic melanoma line (Mel57-VEGF) were used. The latter line was chosen because a stably transfected line was available that yielded a bright fluorescent signal that could be measured reproducibly at low cell counts using the COPAS technology.
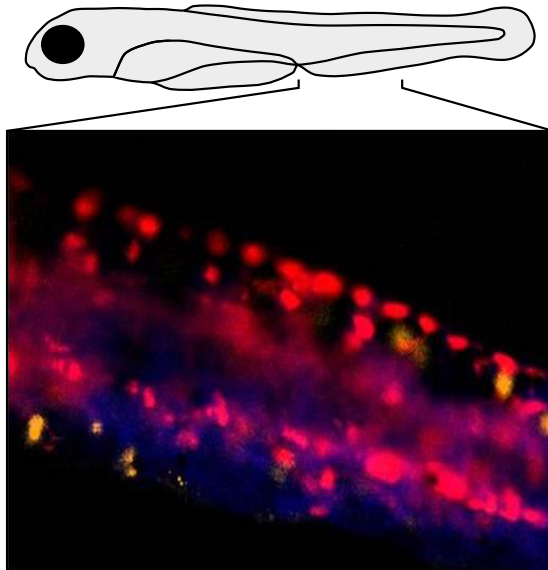
Figure 4: Confocal laser scanning analysis of 5 days old larvae injected with osteosarcoma cell line SJSA-1 at the 256 cell stage. For these studies a transgenic zebrafish *Tg(Fms:Gal4, UAS:mCherry)* was used [100]. The figure shows a z-series projection of a part of the tail area with the mCherry signal shown in orange and in red the proliferation and strong spreading of the CM-DiI labelled cancer cells. Blue color represents autofluorescence

In order to investigate the dissemination of tumor cells in zebrafish embryos, we developed an optimal setup for the establishment of a high-throughput xenograft model using the automated injection system. Cells were cultured as previously described i.e. in RPMI-1640 medium (Sigma Aldrich) supplemented with fetal bovine serum to a final concentration of 10%. Cells were subdivided 2–3 times a week at ratio of 1:5–1:10. Cultures were renewed every 8 passages. During subdividing care was taken that single cells were passed so no aggregates have formed. For transplantation, cells with 75–80% confluency should be used as they tend to clump less then more confluent cell cultures. For detection of cancer cells we mostly have used CM-DiI labeled cell cultures that were labeled using the following protocol: T75 flask with cells reaching 80% confluency were gently but quickly rinsed with 1 ml Trypsin–EDTA. For getting cells into suspension,

cells were covered with 0.5 ml of Trypsin–EDTA and incubated for 2–5 min at 37 ˚C. Trypsination was stopped by adding 1 ml of PBS with 10% FCS and cell lumps were broken by gentle pipetting up and down. From this step no visible cell aggregates should be observed. Single cells were pelleted by centrifugation at 100g for 4 min in a 2 ml Eppendorf tube. Pellets were resuspended in 0.5 ml PBS with CM-DiI (InVitrogen) incubated first at 37 ˚C for 5 min and later at 4 ˚C for an additional quarter of an hour. After this step, cells were centrifuged again at 100g for 4 min, washed once in 0.5 ml PBS and counted. Finally cells were resuspended in 14% PVP to a final concentration of minimum 200 cells per microliter. Prior to injection, the labelled cells were kept at 34 ˚C and implanted within 3 h. Transgenic cells were treated in a similar way but the step related to incubation with a dye was omitted. As a control, fish were injected with 14% PVP. Due to the size of tumor cells, the suspension was loaded into glass capillary needles with bigger opening (Qvotech, 20 micron, Table 1). Up to 5 nl volume with minimum 400 cells were implanted using the robotic injector. Successfully injected embryos (around 80–90% rate) were either sorted manually under a fluorescence stereoscope (Leica) or sorted by the COPAS XL. Zebrafish embryos were maintained in groups of 50 in a Petridish with 20–25 ml egg water and kept at 34 ˚C until 6 dpi. Using a *Tg(fli:GFP)* endothelial reporter transgenic zebrafish line with fluorescent vasculature [92], we found that PC3 and LNCap cells transplanted robotically behaved in the same way as manually transplanted cells [87]. The first clearly disseminated fluorescent tumor cells were observed at 5 dpi in the caudal vein, and also in other small vascular in the eye and intersegmental vessels (Figure 5). These experiments showed several important technical aspects of the micro-injection procedure. High concentration of PVP prevented cell clumping and needle clogging. However, as highly concentrated PVP is very viscous, it is difficult to work with. For cells that are kept in suspensions, lower concentration of PVP will be sufficient (e.g. 2%). It is also important to note that cell dissemination has only been observed when the number of implanted cells reached at least 400. When lower numbers of cells were implanted into an early embryo, even highly aggressive cells did not migrate out of the yolk.
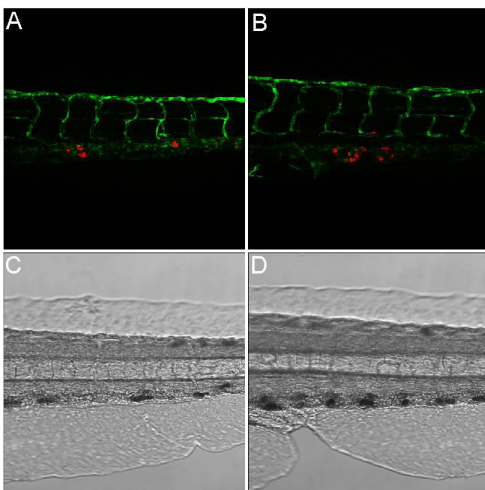

Figure 5: Automatic injection of tumor cells into zebrafish embryos resulted in tumor dissemination. (A & B) Confocal z-series projection and (**C** & **D**) bright-field of fluorescent cell tracker CM-DiI labeled LNCap cells disseminated to the tail vascular of a 7 day-old larva. Images were acquired using a Leica TCS SPE confocal microscope with an x20 dry objective.

2. Confocal screening and tumor cell dissemination

Taking advantage of the transparency of zebrafish larvae, we performed high-resolution confocal microscopy laser scanning imaging to characterize the dissemination of injected tumor cells in the case of the PC3 and LNCAP cell lines (500 cells injection). It is known that leukocytes contribute to different steps of tumor progression [72, 93, 94]. In order to investigate the role of embryonic myeloid cells in cancer cell spreading or killing of cancer cells, whole-mount L-plastin immunohistochemistry was carried out as described [95]. Antibodies and dilutions were used as follows: L-plastin rabbit anti-zebrafish primary antibody (provided by Dr. Huttenlocher), was diluted 1:500 and the secondary antibody (Alexa 568 anti-rabbit, Invitrogen) was used at a 1:200 dilution.

At 6 dpi, confocal imaging of fixed larvae revealed the presence of labeled cells in the vascular system, indicating that injected human tumor cells survived and disseminated outside the yolk (Figure 6). Noteworthy, we also detected that labeled tumor cells were accompanied by a leukocyte population at various positions in the tissue (Figure 6), suggesting that myeloid cells have a role in cancer cell recognition. In future studies we will study whether differences in dissemination are correlated with the aggressiveness of tumor cell type. For this we will use a combination of the automated injection system together with published high-throughput imaging technologies [87] to screen for the malignancy of many different mammalian cancer cell lines in the zebrafish embryo model at higher throughput levels than previously established.
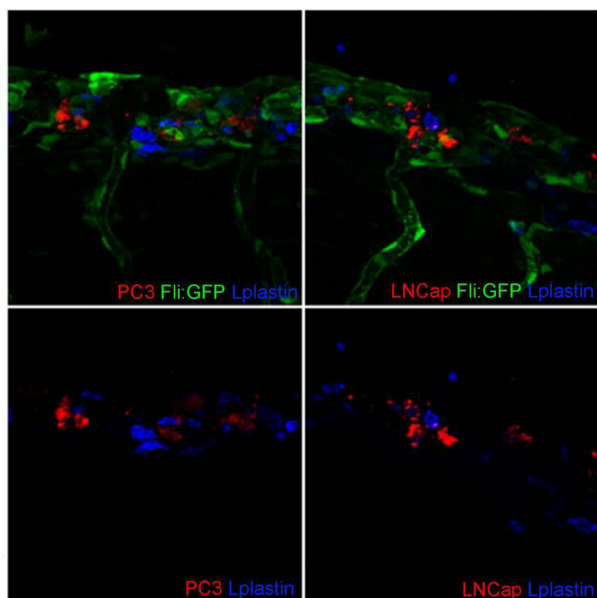

Figure 6: Confocal screening of tumor cell dissemination in zebrafish embryos. PC3 and LNCap (red color) injected in embryos showed clear dissemination from the original implantation site at 6 dpi. Leukocytes (in blue) associated with tumor cells are visible at the dissemination site with the vascular system in green (this signal is omitted in the in bottom panels). Two representative images are shown of embryonic leukocytes associated with fluorescent tumor cells at the caudal vein. Images were acquired using a Leica TCS SPE confocal microscope with an x20 dry objective.

3. COPAS screening for tumor cell proliferation

To demonstrate the survival rate and proliferation of injected tumor cells in zebrafish embryos, we employed a stably transfected melanoma cell line Mel57-VEGF expressing EGFP, [96] in our robotic injection system. A large set of embryos were injected with Mel57-EGFP cells (400 cell/embryo) at 2 and 4 hpf. The injected embryos were immediately run through the COPAS system to determine the total level of green fluorescence, which represents the injection load. The embryos were subsequently sorted with the following parameters: optical density threshold (extinction) = 390 mV (COPAS value: 20), minimum time of flight = 280 ms (COPAS value: 700), red photomultiplier tube (PMT) voltage = 0 V, green PMT voltage = 600 V, yellow PMT voltage = 0 V, fluorescent density threshold = 800. Successfully injected embryos were followed with the COPAS system until 6 dpi.

Combination of COPAS profiling and epifluorescence/brightfield imaging revealed significant tumor cell proliferation in the injected embryos (Figure 7). At the first day after injection, the fluorescent signal of tumor clusters slightly dropped, suggesting a decrease of live cells due to the implantation process. At 6 dpi, the average signal per larva had significantly increased, and was almost 2-fold higher than at 1 dpi (Figure 7). These results illustrated the ability of COPAS to accurately quantify the number of tumor cells in zebrafish embryos. This shows that the combination of the automatic injection and COPAS sorting system provides a powerful *in vivo* platform for the investigation of human cancer proliferation, which in turn highly facilitates the discovery and screening of anti-cancer compounds directed at targets from the implanted cells and the recipient host. Based on these results we have recently started with follow up studies using our methods of xenografting with many other cancer cell lines aimed at obtaining mechanistic insights in the spreading of cancer cells through the zebrafish body. These studies already have indicated that using this methodology new knowledge can be obtained in proliferation and metastatic behaviour of Ewing sarcoma cells (manuscript in preparation). This is of clinical relevance since at the time of diagnosis ""25% of
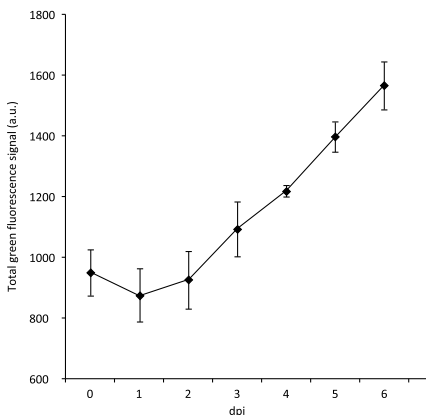

Figure 7: Proliferation of automatic injected Mel57 cells in zebrafish embryos. Total fluorescent signal intensity quantifies the increase of tumor cell number in the injected embryos over 6 days. The sample size for each experiment was >400 embryos. More than 150 successfully injected embryos were sorted out at 1 dpi. Final survival rate was >80% at 6 dpi.

Ewing sarcoma patients, predominantly children and young adults, present with metastatic disease and a poor survival prognosis. In case of recurrence of the disease, overall survival rates are as low as ''10–20%''. Although the causative underlying defect is known, there are currently no reliable treatment strategies for this malignant bone tumor. Therefore, high-throughput screening is greatly needed for rapid preclinical drug screens for this type of tumor.

Conclusions

The provided methods for robotic micro-injection of DNA, morpholinos, bacteria and cancer cells in zebrafish embryos enables the generation of an abundance of reproducibly treated living test models. In Table 3 we present a summary of variables that we have tested with our developed technology. We have shown that this technology benefits optimally from many new high-throughput screening methods. These methods are not only applicable to the disease models highlighted in this paper, i.e. cancer and infectious disease, but are widely applicable to studies that benefit from the use of small transparent zebrafish larvae. We anticipate that our technology is also applicable to other fish models. For instance the use of carp fish, which are easy to culture and have a clutch size of hundred thousands of eggs in one spawn, would be extremely useful for screens in larvae at ultra high-throughput level [6]. We have recently generated a draft genome of the European common carp (*Cyprinus carpio*) showing a striking gene similarity with the zebrafish, but, a much higher compactness of the genome, making it highly interesting for further comparative studies with zebrafish at the genetic level [97]. In our own research we will continue to use zebrafish and other fish species for applications in disease studies, but, we will also give more attention to the applicability of our methodology for toxicology research in zebrafish [98]. This is because we believe

Table 3: Overview of robotic injection applications. For details on the microbial robotic injection variables tested we also refer to Carvalho et al., 2011 [49] and Veneman et al., 2013 [39].

Robotic injection application	Variables tested	Limiting factors
DNA	• Position of injection • COPAS screening	• Higher accuracy of injection site will limit speed
Morpholinos	• Combination with microbial injection	• Morpholinos have to be injected as early as possible
Microbes	• Different types of microbes (e.g. Mycobacteria and Staphylococci) • Embryonic stage of injection • Microbial escape from yolk	• There is a delay period before bacteria from yolk will be exposed to cellular immune responses (e.g. phagocytosis) • For whole organism screening there will be background fluorescence from microbes in the yolk
Cancer cells	• Cancer cell types • PVP concentration • Spread to different body parts	• Embryo survival rates • Clumping of cells that will clog needles

that the combination of screens for potential new therapeutics will benefit greatly in speed by a combination with toxicology approaches. In the near future the use of our micro-injection technology will also be used for detailed drug RP/KD studies that should give indications on the translatability of drug treatment effects and might suggest protocols for optimized drug administration. For this the injection of synthetic slow or controlled release vehicles in combination with highly sensitive detection of compounds using mass spectrometry will be highly valuable [38, 78, 79, 99].

Acknowledgements

http://www.sciencedirect.com/science/article/pii/S104620231300203X

References

1.   Anderson KV, Ingham PW: The transformation of the model organism: a decade of developmental genetics. Nat Genet 2003, 33 Suppl:285-293.

2.   Briggs JP: The zebrafish: a new model organism for integrative physiology. Am J Physiol Regul Integr Comp Physiol 2002, 282(1):R3-9.

3.   Grunwald DJ, Eisen JS: Headwaters of the zebrafish -- emergence of a new model vertebrate. Nat Rev Genet 2002, 3(9):717-724.

4.   Spitsbergen JM, Kent ML: The state of the art of the zebrafish model for toxicology and toxicologic pathology research--advantages and current limitations. Toxicol Pathol 2003, 31 Suppl:62-87.

5.   Sprague J, Bayraktaroglu L, Clements D, Conlin T, Fashena D, Frazer K, Haendel M, Howe DG, Mani P, Ramachandran S et al: The Zebrafish Information Network: the zebrafish model organism database. Nucleic Acids Res 2006, 34(Database issue):D581-585.

6.   Ali S, Champagne DL, Spaink HP, Richardson MK: Zebrafish embryos and larvae: a new generation of disease models and drug screens. Birth Defects Res C Embryo Today 2011, 93(2):115-133.

7.   Allen JP, Neely MN: Trolling for the ideal model host: zebrafish take the bait. Future Microbiol 2010, 5(4):563-569.

8.   Bakkers J: Zebrafish as a model to study cardiac development and human cardiac disease. Cardiovasc Res 2011, 91(2):279-288.

9.   Berger J, Currie PD: Zebrafish models flex their muscles to shed light on muscular dystrophies. Dis Model Mech 2012, 5(6):726-732.

10.   Goessling W, North TE, Zon LI: New waves of discovery: modeling cancer in zebrafish. J Clin Oncol 2007, 25(17):2473-2479.

11.   Jing L, Zon LI: Zebrafish as a model for normal and malignant hematopoiesis. Dis Model Mech 2011, 4(4):433-438.

12.   Lieschke GJ, Currie PD: Animal models of human disease: zebrafish swim into view. Nat Rev Genet 2007, 8(5):353-367.

13.   Martin JS, Renshaw SA: Using in vivo zebrafish models to understand the biochemical basis of neutrophilic respiratory disease. Biochem Soc Trans 2009, 37(Pt 4):830-837.

14.   Meijer AH, Spaink HP: Host-pathogen interactions made transparent with the zebrafish model. Curr Drug Targets 2011, 12(7):1000-1017.

15.   Mione MC, Trede NS: The zebrafish as a model for cancer. Dis Model Mech 2010, 3(9-10):517-523.

16.   Mione M, Meijer AH, Snaar-Jagalska BE, Spaink HP, Trede NS: Disease modeling in zebrafish: cancer and immune responses--a report on a workshop held in Spoleto, Italy, July 20-22, 2009. Zebrafish 2009, 6(4):445-451.

17.   Mione M, Zon LI: Cancer and inflammation: an aspirin a day keeps the cancer at bay. Curr Biol 2012, 22(13):R522-525.

2

18.   Renshaw SA, Trede NS: A model 450 million years in the making: zebrafish and vertebrate immunity. Dis Model Mech 2012, 5(1):38-47.

19.   Richardson J, Zeng Z, Ceol C, Mione M, Jackson IJ, Patton EE: A zebrafish model for nevus regeneration. Pigment Cell Melanoma Res 2011, 24(2):378-381.

20.   Santana S, Rico EP, Burgos JS: Can zebrafish be used as animal model to study Alzheimer's disease? Am J Neurodegener Dis 2012, 1(1):32-48.

21.   Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug RG, 2nd, Tan W, Penheiter SG, Ma AC, Leung AY et al: In vivo genome editing using a high-efficiency TALEN system. Nature 2012, 491(7422):114-118.

22.   Bill BR, Petzold AM, Clark KJ, Schimmenti LA, Ekker SC: A primer for morpholino use in zebrafish. Zebrafish 2009, 6(1):69-77.

23.   Cade L, Reyon D, Hwang WY, Tsai SQ, Patel S, Khayter C, Joung JK, Sander JD, Peterson RT, Yeh JR: Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. Nucleic Acids Res 2012, 40(16):8001-8010.

24.   Chen S, Oikonomou G, Chiu CN, Niles BJ, Liu J, Lee DA, Antoshechkin I, Prober DA: A large-scale in vivo analysis reveals that TALENs are significantly more mutagenic than ZFNs generated using context-dependent assembly. Nucleic Acids Res 2013, 41(4):2769-2778.

25.   Clark KJ, Voytas DF, Ekker SC: A TALE of two nucleases: gene targeting for the masses? Zebrafish 2011, 8(3):147-149.

26.   Dahlem TJ, Hoshijima K, Jurynec MJ, Gunther D, Starker CG, Locke AS, Weis AM, Voytas DF, Grunwald DJ: Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. PLoS Genet 2012, 8(8):e1002861.

27.   Huang P, Zhu Z, Lin S, Zhang B: Reverse genetic approaches in zebrafish. J Genet Genomics 2012, 39(9):421-433.

28.   Moore FE, Reyon D, Sander JD, Martinez SA, Blackburn JS, Khayter C, Ramirez CL, Joung JK, Langenau DM: Improved somatic mutagenesis in zebrafish using transcription activator-like effector nucleases (TALENs). PLoS One 2012, 7(5):e37877.

29.   Chatterjee S, Lufkin T: Regulatory genomics: Insights from the zebrafish. Curr Top Genet 2012, 5:1-10.

30.   Chen E, Ekker SC: Zebrafish as a genomics research model. Curr Pharm Biotechnol 2004, 5(5):409-413.

31.   Chen K, Cole RB, Rees BB: Hypoxia-induced changes in the zebrafish (Danio rerio) skeletal muscle proteome. J Proteomics 2013, 78:477-485.

32.   Hegedus Z, Zakrzewska A, Agoston VC, Ordas A, Racz P, Mink M, Spaink HP, Meijer AH: Deep sequencing of the zebrafish transcriptome response to mycobacterium infection. Mol Immunol 2009, 46(15):2918-2930.

33.   Hogl S, van Bebber F, Dislich B, Kuhn PH, Haass C, Schmid B, Lichtenthaler SF: Label-free quantitative analysis of the membrane proteome of Bace1 protease knock-out zebrafish brains. Proteomics 2013, 13(9):1519-1527.

34.   Kabli S, Spaink HP, De Groot HJ, Alia A: In vivo metabolite profile of adult zebrafish brain obtained by high-resolution localized magnetic resonance spectroscopy. J Magn Reson Imaging 2009, 29(2):275-281.

35.   Lossner C, Wee S, Ler SG, Li RH, Carney T, Blackstock W, Gunaratne J: Expanding the zebrafish embryo proteome using multiple fractionation approaches and tandem mass spectrometry. Proteomics 2012, 12(11):1879-1882.

36.   Lucitt MB, Price TS, Pizarro A, Wu W, Yocum AK, Seiler C, Pack MA, Blair IA, Fitzgerald GA, Grosser T: Analysis of the zebrafish proteome during embryonic development. Mol Cell Proteomics 2008, 7(5):981-994.

37.   Ordas A, Hegedus Z, Henkel CV, Stockhammer OW, Butler D, Jansen HJ, Racz P, Mink M, Spaink HP, Meijer AH: Deep sequencing of the innate immune transcriptomic response of zebrafish embryos to Salmonella infection. Fish Shellfish Immunol 2011, 31(5):716-724.

38.   Raterink RJ, van der Kloet FM, Li JJ, Wattel NA, Schaaf MJM, Spaink HP, Berger R, Vreeken RJ, Hankemeier T: Rapid metabolic screening of early zebrafish embryogenesis based on direct infusion-nanoESI-FTMS. Metabolomics 2013, 9(4):864-873.

39.   Veneman WJ, Stockhammer OW, de Boer L, Zaat SA, Meijer AH, Spaink HP: A zebrafish high throughput screening system used for Staphylococcus epidermidis infection marker discovery. BMC Genomics 2013, 14:255.

40.   Renshaw SA, Ingham PW: Zebrafish models of the immune response: taking it on the ChIn. BMC Biol 2010, 8:148.

41.   van der Vaart M, Spaink HP, Meijer AH: Pathogen recognition and activation of the innate immune response in zebrafish. Adv Hematol 2012, 2012:159807.

42.   Detrich HW, Westerfield M, Zon LI: Essential Zebrafish Methods: Cell and Developmental Biology. Academic Press, Oxford 2009.

2

43.   Lieschke GJ: Methods and Protocols. Springer protocols, Heidelberg 2008.

44.   Westerfield M: The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio). Univ of Oregon Press, Eugene 2000.

45.   Detrich HW, Westerfield M, Zon L: The Zebrafish: Disease Models and Chemical Screens. Elsevier, Amsterdam 2011.

46.   Detrich HW, Westerfield M, Zon L: The zebrafish : genetics, genomics and informatics. Elsevier, Amsterdam 2011.

47.   Benard EL, van der Sar AM, Ellett F, Lieschke GJ, Spaink HP, Meijer AH: Infection of zebrafish embryos with intracellular bacterial pathogens. J Vis Exp 2012(61).

48.   Wang W, Liu X, Gelinas D, Ciruna B, Sun Y: A fully automated robotic system for microinjection of zebrafish embryos. PLoS One 2007, 2(9):e862.

49.   Carvalho R, de Sonneville J, Stockhammer OW, Savage ND, Veneman WJ, Ottenhoff TH, Dirks RP, Meijer AH, Spaink HP: A high-throughput screen for tuberculosis progression. PLoS One 2011, 6(2):e16779.

50.   Gehrig J, Reischl M, Kalmar E, Ferg M, Hadzhiev Y, Zaucker A, Song C, Schindler S, Liebel U, Muller F: Automated high-throughput mapping of promoter-enhancer interactions in zebrafish embryos. Nat Methods 2009, 6(12):911-916.

51.   Kokel D, Bryan J, Laggner C, White R, Cheung CY, Mateus R, Healey D, Kim S, Werdich AA, Haggarty SJ et al: Rapid behavior-based identification of neuroactive small molecules in the zebrafish. Nat Chem Biol 2010, 6(3):231-237.

52.   Chang TY, Pardo-Martin C, Allalou A, Wahlby C, Yanik MF: Fully automated cellular-resolution vertebrate screening platform with parallel animal processing. Lab Chip 2012, 12(4):711-716.

53.   Pardo-Martin C, Chang TY, Koo BK, Gilleland CL, Wasserman SC, Yanik MF: High-throughput in vivo vertebrate screening. Nat Methods 2010, 7(8):634-636.

54.   Brothers KM, Wheeler RT: Non-invasive imaging of disseminated candidiasis in zebrafish larvae. J Vis Exp 2012(65).

55.   Crim MJ, Riley LK: Viral diseases in zebrafish: what is known and unknown. ILAR J 2012, 53(2):135-143.

56.   Cui C, Benard EL, Kanwal Z, Stockhammer OW, van der Vaart M, Zakrzewska A, Spaink HP, Meijer AH: Infectious disease modeling and innate immune function in zebrafish embryos. Methods Cell Biol 2011, 105:273-308.

57.   Kanther M, Rawls JF: Host-microbe interactions in the developing zebrafish. Curr Opin Immunol 2010, 22(1):10-19.

58.   Milligan-Myhre K, Charette JR, Phennicie RT, Stephens WZ, Rawls JF, Guillemin K, Kim CH: Study of host-microbe interactions in zebrafish. Methods Cell Biol 2011, 105:87-116.

59.   Novoa B, Figueras A: Zebrafish: model for the study of inflammation and the innate immune response to infectious diseases. Adv Exp Med Biol 2012, 946:253-275.

60.   Tobin DM, May RC, Wheeler RT: Zebrafish: a see-through host and a fluorescent toolbox to probe host-pathogen interaction. PLoS Pathog 2012, 8(1):e1002349.

61.   Berg RD, Ramakrishnan L: Insights into tuberculosis from the zebrafish model. Trends Mol Med 2012, 18(12):689-690.

62.   Tobin DM, Ramakrishnan L: Comparative pathogenesis of Mycobacterium marinum and Mycobacterium tuberculosis. Cell Microbiol 2008, 10(5):1027-1039.

63.   Strahle U, Scholz S, Geisler R, Greiner P, Hollert H, Rastegar S, Schumacher A, Selderslaghs I, Weiss C, Witters H et al: Zebrafish embryos as an alternative to animal experiments--a commentary on the definition of the onset of protected life stages in animal welfare regulations. Reprod Toxicol 2012, 33(2):128-132.

64.   Takaki K, Cosma CL, Troll MA, Ramakrishnan L: An in vivo platform for rapid high-throughput antitubercular drug discovery. Cell Rep 2012, 2(1):175-184.

65.   van Soest JJ, Stockhammer OW, Ordas A, Bloemberg GV, Spaink HP, Meijer AH: Comparison of static immersion and intravenous injection systems for exposure of zebrafish embryos to the natural pathogen Edwardsiella tarda. BMC Immunol 2011, 12:58.

66.   Amatruda JF, Patton EE: Genetic models of cancer in zebrafish. Int Rev Cell Mol Biol 2008, 271:1-34.

67.   Etchin J, Kanki JP, Look AT: Zebrafish as a model for the study of human cancer. Methods Cell Biol 2011, 105:309-337.

68.   Konantz M, Balci TB, Hartwig UF, Dellaire G, Andre MC, Berman JN, Lengerke C: Zebrafish xenografts as a tool for in vivo studies on human cancer. Ann N Y Acad Sci 2012, 1266:124-137.

69.   Payne E, Look T: Zebrafish modelling of leukaemias. Br J Haematol 2009, 146(3):247-256.

2

70. White RM, Cech J, Ratanasirintrawoot S, Lin CY, Rahl PB, Burke CJ, Langdon E, Tomlinson ML, Mosher J, Kaufman C et al: DHODH modulates transcriptional elongation in the neural crest and melanoma. Nature 2011, 471(7339):518-522.

71. Feng Y, Santoriello C, Mione M, Hurlstone A, Martin P: Live imaging of innate immune cell sensing of transformed cells in zebrafish larvae: parallels between tumor initiation and wound inflammation. PLoS Biol 2010, 8(12):e1000562.

72. He S, Lamers GE, Beenakker JW, Cui C, Ghotra VP, Danen EH, Meijer AH, Spaink HP, Snaar-Jagalska BE: Neutrophil-mediated experimental metastasis is enhanced by VEGFR inhibition in a zebrafish xenograft model. J Pathol 2012, 227(4):431-445.

73. Tobia C, Gariano G, De Sena G, Presta M: Zebrafish embryo as a tool to study tumor/endothelial cell cross-talk. Biochim Biophys Acta 2013, 1832(9):1371-1377.

74. Stern HM, Zon LI: Cancer genetics and drug discovery in the zebrafish. Nat Rev Cancer 2003, 3(7):533-539.

75. Taylor KL, Grant NJ, Temperley ND, Patton EE: Small molecule screening in zebrafish: an in vivo approach to identifying new chemical tools and drug leads. Cell Commun Signal 2010, 8:11.

76. Terriente J, Pujades C: Use of zebrafish embryos for small molecule screening related to cancer. Dev Dyn 2013, 242(2):97-107.

77. Trede NS, Heaton W, Ridges S, Sofla H, Cusick M, Bearss D, Jones D, Fujinami RS: Discovery of biologically active oncologic and immunologic small molecule therapies using zebrafish: overview and example of modulation of T cell activation. Curr Protoc Pharmacol 2013, Chapter 14:Unit14 24.

78. Peng K, Cui C, Tomatsu I, Porta F, Meijer AH, Spaink HP, Kros A: Cyclodextrin/dextran based drug carriers for a controlled release of hydrophobic drugs in zebrafish embryos. Soft Matter 2010, 6(16):3778-3783.

79. Peng K, Tomatsu I, van den Broek B, Cui C, Korobko AV, van Noort J, Meijer AH, Spaink HP, Kros A: Dextran based photodegradable hydrogels formed via a Michael addition. Soft Matter 2011, 7(10):4881-4887.

80. Corkery DP, Dellaire G, Berman JN: Leukaemia xenotransplantation in zebrafish--chemotherapy response assay in vivo. Br J Haematol 2011, 153(6):786-789.

81. Dovey MC, Zon LI: Defining cancer stem cells by xenotransplantation in zebrafish. Methods Mol Biol 2009, 568:1-5.

82. Haldi M, Ton C, Seng WL, McGrath P: Human melanoma cells transplanted into zebrafish proliferate, migrate, produce melanin, form masses and stimulate angiogenesis in zebrafish. Angiogenesis 2006, 9(3):139-151.

83. Lee LM, Seftor EA, Bonde G, Cornell RA, Hendrix MJ: The fate of human malignant melanoma cells transplanted into zebrafish embryos: assessment of migration and cell division in the absence of tumor formation. Dev Dyn 2005, 233(4):1560-1570.

84. Mohseny AB, Xiao W, Carvalho R, Spaink HP, Hogendoorn PC, Cleton-Jansen AM: An osteosarcoma zebrafish model implicates Mmp-19 and Ets-1 as well as reduced host immune response in angiogenesis and migration. J Pathol 2012, 227(2):245-253.

85. Stoletov K, Klemke R: Catch of the day: zebrafish as a human cancer model. Oncogene 2008, 27(33):4509-4520.

86. Taylor AM, Zon LI: Zebrafish tumor assays: the state of transplantation. Zebrafish 2009, 6(4):339-346.

87. Ghotra VP, He S, de Bont H, van der Ent W, Spaink HP, van de Water B, Snaar-Jagalska BE, Danen EH: Automated whole animal bio-imaging assay for human cancer dissemination. PLoS One 2012, 7(2):e31281.

88. Mizgirev I, Revskoy S: Generation of clonal zebrafish lines and transplantable hepatic tumors. Nat Protoc 2010, 5(3):383-394.

89. Stockhammer OW, Zakrzewska A, Hegedus Z, Spaink HP, Meijer AH: Transcriptome profiling and functional analyses of the zebrafish embryonic innate immune response to Salmonella infection. J Immunol 2009, 182(9):5641-5653.

90. van der Vaart M, van Soest JJ, Spaink HP, Meijer AH: Functional analysis of a zebrafish myd88 mutant identifies key transcriptional components of the innate immune system. Dis Model Mech 2013, 6(3):841-854.

91. Buddingh EP, Schilham MW, Ruslan SE, Berghuis D, Szuhai K, Suurmond J, Taminiau AH, Gelderblom H, Egeler RM, Serra M et al: Chemotherapy-resistant osteosarcoma is highly susceptible to IL-15-activated allogeneic and autologous NK cells. Cancer Immunol Immunother 2011, 60(4):575-586.

92. Lawson ND, Weinstein BM: In vivo imaging of embryonic vascular development using transgenic zebrafish. Dev Biol 2002, 248(2):307-318.

93. Shojaei F, Ferrara N: Refractoriness to antivascular endothelial growth factor treatment: role of myeloid cells. Cancer Res 2008, 68(14):5501-5504.

2

94.   Stockmann C, Doedens A, Weidemann A, Zhang N, Takeda N, Greenberg JI, Cheresh DA, Johnson RS: Deletion of vascular endothelial growth factor in myeloid cells accelerates tumorigenesis. Nature 2008, 456(7223):814-818.

95.   Mathias JR, Dodd ME, Walters KB, Rhodes J, Kanki JP, Look AT, Huttenlocher A: Live imaging of chronic inflammation caused by mutation of zebrafish Hai1. J Cell Sci 2007, 120(Pt 19):3372-3383.

96.   Stollman TH, Scheer MG, Franssen GM, Verrijp KN, Oyen WJ, Ruers TJ, Leenders WP, Boerman OC: Tumor accumulation of radiolabeled bevacizumab due to targeting of cell- and matrix-associated VEGF-A isoforms. Cancer Biother Radiopharm 2009, 24(2):195-200.

97.   Henkel CV, Dirks RP, Jansen HJ, Forlenza M, Wiegertjes GF, Howe K, van den Thillart GE, Spaink HP: Comparison of the exomes of common carp (Cyprinus carpio) and zebrafish (Danio rerio). Zebrafish 2012, 9(2):59-67.

98.   Driessen M, Kienhuis AS, Pennings JL, Pronk TE, van de Brandhof EJ, Roodbergen M, Spaink HP, van de Water B, van der Ven LT: Exploring the zebrafish embryo as an alternative model for the evaluation of liver toxicity by histopathology and expression profiling. Arch Toxicol 2013, 87(5):807-823.

99.   Sharif F, Porta F, Meijer AH, Kros A, Richardson MK: Mesoporous silica nanoparticles as a compound delivery system in zebrafish embryos. Int J Nanomedicine 2012, 7:1875-1890.

100.   Gray C, Loynes CA, Whyte MK, Crossman DC, Renshaw SA, Chico TJ: Simultaneous intravital imaging of macrophage and neutrophil behaviour during inflammation using a novel transgenic zebrafish. Thromb Haemost 2011, 105(5):811-819.

2

Chapter 3

Establishment and optimization of a high-throughput setup to study *Staphylococcus epidermidis* **and** *Mycobacterium marinum* infection as a model for drug discovery

3

Wouter J. Veneman[1#], Rubén Marín-Juez[2#], Jan de Sonneville[3], Anita Ordas[1], Susanne Jong-Raadsen[2], Annemarie H. Meijer[1], Herman P. Spaink[1]

[1] Institute of Biology, Leiden University
[2] ZF-screens BV
[3] Life Science Methods BV

[#] Authors contributed equally

Abstract

Zebrafish are becoming a valuable tool in the preclinical phase of drug discovery screenings as a whole animal model with high-throughput screening possibilities. They can be used to bridge the gap between cell based assays at earlier stages and *in vivo* validation in mammalian models, reducing, in this way, the number of compounds passing through to testing on the much more expensive rodent models. In this light, in the present manuscript is described a new high-throughput pipeline using zebrafish as *in vivo* model system for the study of *Staphylococcus epidermidis* and *Mycobacterium marinum* infection. This setup allows the generation and analysis of large number of synchronous embryos homogenously infected. Moreover the flexibility of the pipeline allows the user to easily implement other platforms to improve the resolution of the analysis when needed. The combination of the zebrafish together with innovative high-throughput technologies opens the field of drug testing and discovery to new possibilities not only because of the strength of using a whole animal model but also because of the large number of transgenic lines available that can be used to decipher the mode of action of new compounds.

Introduction

To date the zebrafish (*Danio rerio*) has been successfully established as an efficient model to study a variety of infectious diseases [1]. The zebrafish embryo offers unique *in vivo* imaging possibilities due to their transparency and the large number of existing transgenic reporter lines expressing fluorescent proteins. This powerful combination makes it possible to track different immune cell types in time while interacting with pathogens such as *Mycobacterium marinum*, the closest relative of *M. tuberculosis* [2], or *Staphylococcus epidermidis*, the main causative of biomaterial-associated infection [3-5]. Different routes of infection can be used in zebrafish embryos depending on the purposes of the study [6].

One of these infection routes is yolk injection of the bacteria. The main advantage of this method compared to the others is that yolk infection can be performed automatically via robotic injection, significantly reducing the injection time and allowing high reproducibility of the infection [7, 8].

Previous work, using the zebrafish as a high-throughput *in vivo* model system for the study of *S. epidermidis* and *M. marinum* infection showed to be successful [7, 8]. This system is able to screen for disease progression via robotic yolk injection of early embryos and using fluorescence readout as a measure for the bacterial load. In agreement with this notion, this setup has been optimized and established a highly efficient high-throughput pipeline with the potential to generate large numbers of homogenously infected embryos and track the progression of the infection during the time after treatment with a number of compounds. With the established setup it is possible to generate up

to 8000 synchronous embryos to screen for disease progression, processing in this way up to 2500 embryos per hour. Embryos are sorted based on their bacterial load using an automated system, ensuring homogenous groups of infected larvae. Furthermore, to validate the setup, effects of reference known to prevent tuberculosis progression in mammals have been tested on embryos infected with *M. marinum* E11 strain or the more virulent M strain [9].

This study describes in detail the high-throughput pipeline that has been established to be able to generate large numbers of infected embryos and the subsequent analysis of the bacterial progression during development and after compound treatment.

Protocol

1) Bacterial strains and growth conditions

1.1) Prepare *S. epidermidis* inoculum

1.1.1) Take several individual colonies from *S. epidermidis* strain O-47, containing a pWVW189 derived mCherry expression vector (De Boer L. unpublished) from a Luria Bertani (LB) agar plate supplemented with 10 µg/ml chloramphenicol and culture overnight at 37°C in 25 ml LB medium supplemented with 10 µg/ml chloramphenicol to mid-log stage.

1.1.2) Centrifuge 1 ml of the culture at 12000 × g for 1 min and subsequently wash them 3 times with 1 ml sterile phosphate-buffered saline (PBS) with 0.3% V/V Tween$_{80}$.

1.1.3) Measure the optical density at 600 nm (OD$_{600}$), and dilute the bacterial suspension to an OD$_{600}$ of 0.3 in 2% W/V polyvinylpyrrolidone$_{40}$ (pvp$_{40}$) in PBS. Note: an OD$_{600}$ of 0.3 corresponds to 1.0×10^8 colony forming unit/ml (cfu/ml).

1.2) Prepare *M. marinum* inoculum

1.2.1) Take several individual colonies from *M. marinum* strain M or E11 containing the pSMT3-mCherry vector stably expressing mCherry [10] from a Middlebrook 7H10 agar plate with 10% V/V Middlebrook OADC enrichment supplemented with 50 µg/ml hygromycin and culture overnight at 28°C in 10 ml of Middlebrook 7H9 broth with 10% V/V Middlebrook ADC enrichment supplemented with 50 µg/ml hygromycin.

1.2.2) Centrifuge 1 ml of the culture at 12000 × g for 1 min and subsequently wash it 3 times with 1 ml sterile PBS with 0.3% V/V Tween$_{80}$.

1.2.3) Measure the OD$_{600}$, and dilute the bacterial suspension to an OD$_{600}$ of 0.3 in 2% W/V pvp$_{40}$ in PBS. Note: an OD$_{600}$ of 1 corresponds to 1.0×10^8 cfu/ml.

2) Prepare zebrafish eggs

2.1) Place maximum 70 male and 50 female wild type zebrafish separately into the large breeding vessel. Note: place the female fish in the lower part of the large breeding vessel.

2.2) Remove the separator the next day in the morning, to let the zebrafish start breeding.

2.3) Collect the eggs at the bottom of the large breeding vessel through the egg collector in a 50 ml tube filled with egg water (60 µg/ml Instant ocean sea salt).

3) Injection needles

3.1) Obtain commercially available custom made glass capillaries needles with an inner diameter of 10 µm.

4) Experimental outline of injection

4.1) Boil 100 ml of 1% W/V agarose in egg water, and cool until approximately 40˚C. Pour agarose into the automated micro-injectors plate and place the 1024 well stamp into the agarose. Note: the plate is ready to use when cooled down.

4.2) On the automated micro-injector operating software click on 'Calibrate stage', then click on '1024' well grid and place the agarose plate in the micro-injector and calibrate the plate by clicking on the screen at the centre position of the well.

4.3) Go to 'needle menu' and click on 'calibrate needle holder'.

4.4) Fill the injection needle using a microloader tip with either 10 µl pvp_{40} containing 100 cfu/nl *S. epidermidis* or 30 cfu/nl *M. marinum*, or use pvp_{40} as mock injection.

4.5) Place the needle in the automated micro-injector and calibrate the x, y position by lowering or moving the needle up and clicking on the screen at the position of the needle. Then calibrate the z position of the needle by clicking on the screen at the position of the tip of the needle.

4.6) Distribute the eggs over the agarose grid using a plastic transfer pipet, and remove excess egg water. Place the agarose grid in the automated micro-injector.

4.7) Go to the 'Injection menu' and adjust the 'Injection pressure' setting to 200 hPa, 'Injection time' 0.2 s and 'Compensation pressure' 15 hPa, which correlates with 1 nl, at

the Femtojet settings menu.

4.8) Click on 'Inject all' to inject the entire plate.

4.9) Collect the eggs after injection by washing them into a Petri dish (92 × 16 mm), with a maximum of 70 embryos per Petri dish, and incubate at 28°C.

5) Flow-cytometer analysis

5.1) Prepare the large particle flow cytometer according to the manufacturer's instructions and fill the sample cup and sheath fluid container with egg water.

5.2) At the operating software, go to the 'PMT' menu and use the following settings: 650 V for the 'Red' channel and 0V for the 'Green' and 'Yellow' channel. Then go to 'Thresholds' menu and use the following settings: 'Optical density' threshold signal: 975mV (COPAS XL value: 50) and the 'Time Of Flight' (TOF) minimum to 320 μs (COPAS XL value: 800) in order to reduce the influence of debris.

5.3 For analysis without sorting the embryos go to step 5.4, for analysis and sorting the embryos into a Petri dish go to step 5.5 or for analysis and sorting the embryos into a 96 well plate go to step 5.6.

5.4) Place the embryos in the sample cup and click on 'start' to start the analysis. When all embryos are analyzed stop the analysis by clicking on 'stop'. Save the data by clicking on 'Store'. Note: all data is stored as TXT, LMD, DAT, CSV and BSRT files. Follow the protocol at step 5.7.

5.5) Place the embryos in the sample cup and define the maximum of 70 embryos per plate to be sorted by entering 70 in the 'Sort' menu. Place an empty Petri dish under the sorter and click on 'Manual Sort. When the Petri dish is filled, save the data by clicking on 'Store'. Note: all data is stored as TXT, LMD, DAT, CSV and BSRT files. Follow the protocol at step 5.7.

5.6) Place the embryos in the sample cup and define the maximum of 1 embryo per well to be sorted by entering 1 in the 'Sort' menu. Place an empty 96 well plate into the left plate holder and click on 'Fill plate'. When the 96 well plate is filled, save the data by clicking on 'Store'. Note: all data is stored as TXT, LMD, DAT, CSV and BSRT files. Follow the protocol at step 5.7.

5.7) Get the TXT file to process the raw data, use the following data filter settings: 'Status select': 40, and if using the sort module 'Status sort': 6. Then use the numbers from the total fluorescence signal from the 'Red' channel to calculate the average and the standard error of the mean. Plot these data sets into bar or scatter graphs.

6) Drug treatment

6.1) Analyze and sort at 3 days post injection (dpi), *M. marinum* infected embryos in two equal groups using the large particle flow cytometer (step 5.5). Treat one group with a compound of interest in its carrier solvent and other with carrier solvent alone (control). Apply similar treatments to mock-injected control to test for antibiotic side effects.

6.2) Repeat at 4 and 5 dpi the analysis (step 5.5) and refresh the egg water or egg water containing the compound.

7) High-resolution imaging

7.1) Anesthetize the embryos with 0.02% W/V buffered 3-aminobenzoic acid ethyl ester (Tricaine) in egg water 10 min before analysis.

7.2) Prepare the Vertebrate Automated Screening Technology system and the Large Particle Sampler according to the manufacturer's instructions.

7.3) Select the reference images corresponding to the age of the embryos from the 'Imaging – Object – Detection Setup' menu.

7.4) Select the amount of pictures and orientation to be made by the Vertebrate Automated Screening Technology system from the 'Imaging – Auto store images' menu.

7.5) Place a 96 well plate filled with embryos (from step 5.6) into the left plate holder of the Large Particle Sampler, and click on 'Run plate'.

7.6) When an embryo is detected and correctly positioned; image the head and the tail separately with the CLSM using a 10X plain dry objective and stitch the images afterwards using image processing software.

Representative results

The present results show that the high-throughput pipeline to study *S. epidermidis* and *M. marinum* infection has been successfully established and that may be extended to other infection models. Firstly, the use of the large breeding vessel (Figure 1A), based on the published system by Adatto et al. 2011 [11], enables to generate large numbers of synchronous eggs in single events affording a high control of the spawning process. Next to be able to inject large number of embryos in a short period of time, an improved version of the previously developed automated micro-injection system [7] was used (Figure 1A). To asses which is the best developmental stage for yolk infection, injections
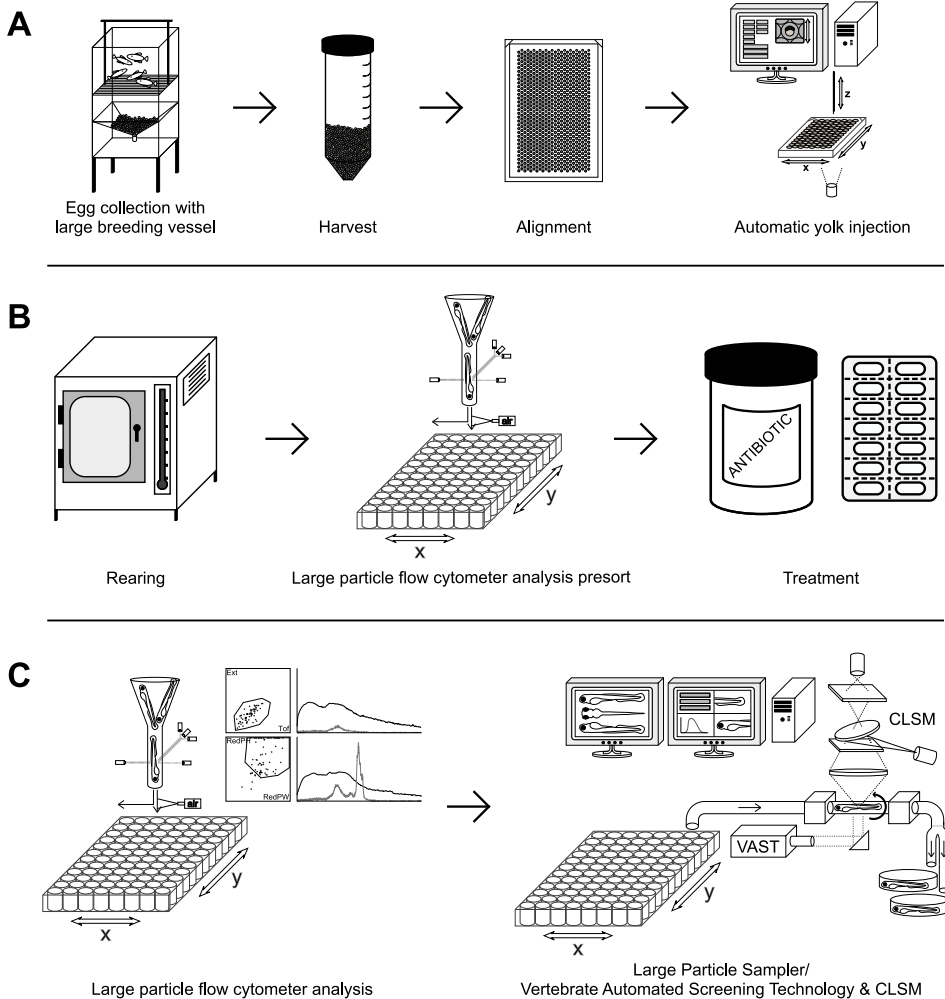
Figure 1: Mainstream experimental outline. (**A**) Adult fish are put together to mate, eggs are collected, aligned into an agarose plate and injected. (**B**) The injected eggs are incubated at 28°C and will be pre-sorted for possible drug treatment. (**C**) Subsequent analysis by large particle flow cytometer and/or Large Particle Sampler/Vertebrate Automated Screening Technology with CLSM.

with *S. epidermidis* and *M. marinum* were performed at all the different stages between 1 and 512 cell stage, according to the description made by Kimmel et al. 1995 [12]. Injections with 100 cfu *S. epidermidis* between the 16 and 128 cell stage provided the best infection pattern (Figure 2). The bacteria proliferated inside the yolk for 3 days and spread into the body from 3 dpi onwards. Performing injections before the 16 cell stage led to high mortality from 4 dpi, and injection after the 256 cell stage showed mainly bacterial growth inside the yolk with hardly any bacteria spreading inside the body of the embryo. Quantification of bacterial burden was performed by fluorescence intensity analysis using the large particle flow cytometer as described by Veneman et al. 2013 [8] (Figure 3).
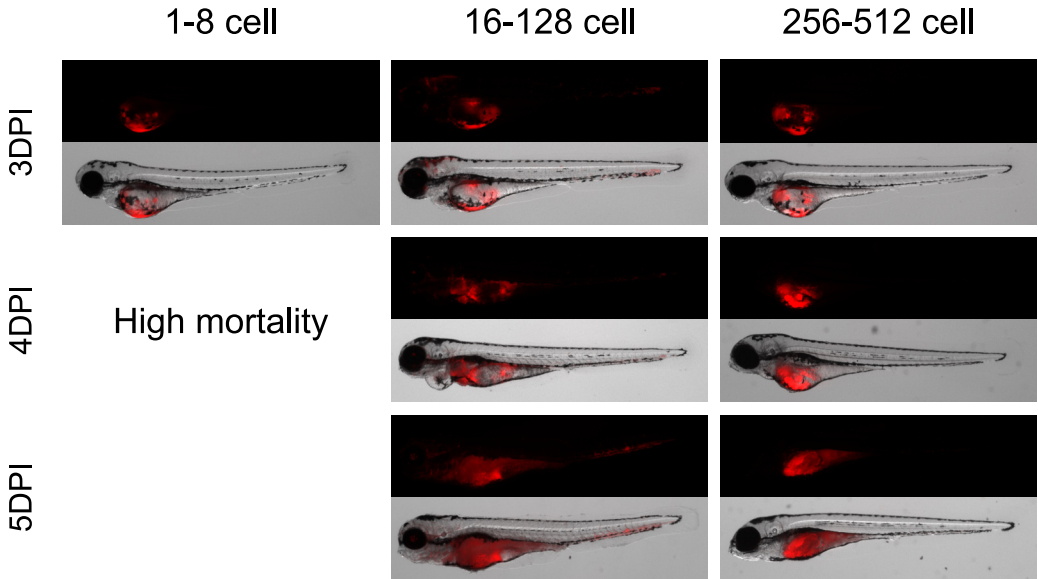
Figure 2: Establishment of best cell stage for *S. epidermidis* yolk injection. Zebrafish embryos were injected in the yolk at different developmental stages from 1 to 512 cell stage with 100 cfu of *S. epidermidis*. Embryos injected between 1 and 8 cell stage showed bacterial growth in the yolk and high mortality from 4 dpi. Embryos injected between 16 and 128 cell stage showed bacterial growth in the yolk and inside the body starting at 3 dpi. Embryos injected between 256 and 512 cell stage showed many bacterial growth inside the yolk.
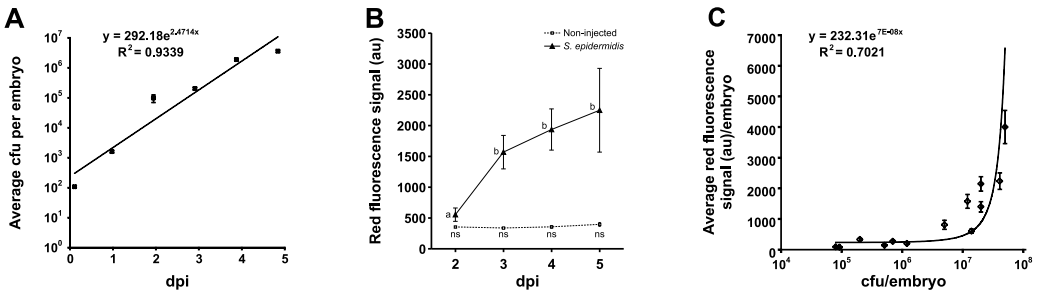
Figure 3: Quantification of bacterial burden using large particle flow cytometer. 100 cfu of *S. epidermidis* were injected into the yolk of zebrafish embryos. (**A**) Up to 5 dpi, each day, groups of 10 embryos were homogenized and plated directly, showing the average exponential growth based on two biological replicas (error bars = SEM). (**B**) Large particle flow cytometer analysis shows the average fluorescence signal from non-injected and *S. epidermidis* injected embryos. 30-160 embryos per condition were analyzed (error bars = SEM), different letters indicate statistical significant differences by one-way ANOVA followed by Tukey's post-hoc test ($P < 0.001$), ns: not significant differences. (**C**) Correlation between cfu and average fluorescence signal of groups of 10 *S. epidermidis* infected embryos (error bars = SEM). This figure has been modified from Veneman et al. 2013 [8].

Observations showed that the optimal developmental stage for injection of 30 cfu *M. marinum* injection is between 16 to 128 cell stage for the E11 strain (Figure 4A) and between 16-64 cell stage with the more virulent M strain (Figure 4B). Embryos injected at these stages showed bacterial growth inside the yolk and spreading of the bacteria through the embryo (Figure 7). Infection with both strains at earlier stages presented non-specific generalized bacterial growth leading the embryos to die after 4 dpi. On the other hand, in embryos injected at later stages bacterial burden was restricted to the yolk.
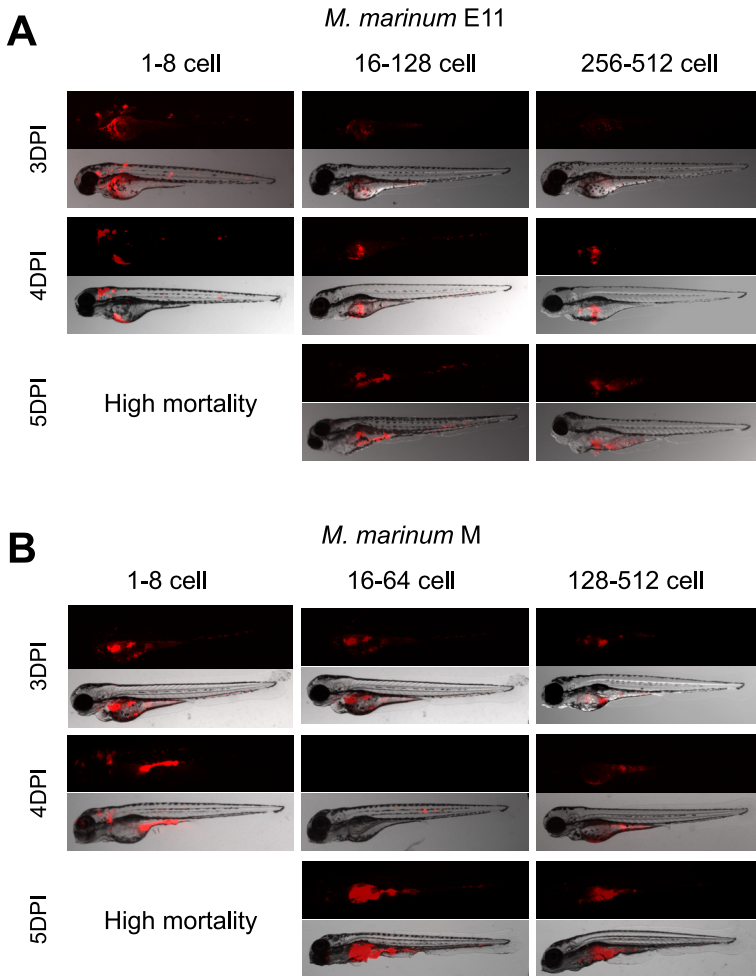
Figure 4: Establishment of the best cell stage for *M. marinum* yolk injection. Zebrafish embryos were injected at all the different developmental stages from 1 to 512 cell stage with 30 cfu of *M. marinum* E11 and M strain. (**A** & **B**) Embryos injected from 1-8 cell stage showed similar spreading and mortality with both strains. (**A**) Embryos injected between 16-128 cell stage with E11 strain showed formation of granulomas and systemic infection while those injected from 256 to 512 cell stage kept bacterial burden into the yolk. (**B**) Embryos injected between 16-64 cell stage with M strain showed formation of granuloma like structures and systemic infection while those injected from 128 to 512 cell stage kept bacterial burden into the yolk.

Next, pre-sorting with large particle flow cytometer (Figure 1B) generated large homogenous groups of infected fish excluding non- or highly infected embryos (Figure 5A and 6A). After pre-sorting, *M. marinum* infected embryos were treated with Rifampicin, a first-line anti-tuberculosis drug. Previous studies demonstrated that treatment with Rifampicin at a dose of 200 µM efficiently reduces *M. marinum* infection in zebrafish [7, 13]. Taking advantage of the large number of homogenously infected embryos generated with the high-throughput setup, treatment with different doses was performed. Embryos infected with *M. marinum* M strain and treated for 48 hours with 12, 24 and 200 µM Rifampicin showed to reduce efficiently mycobacterial infection in a dose dependent manner (Figure 5B). In view of the efficient reduction of the infection using Rifampicin at a dose of 200 µM this concentration was used for the future experiments. In line with the previous result, studying bacterial burden progression using *M. marinum* E11 strain a significant reduction 24 hours and onwards after treatment with 200 µM Rifampicin was observed (Figure 6B).
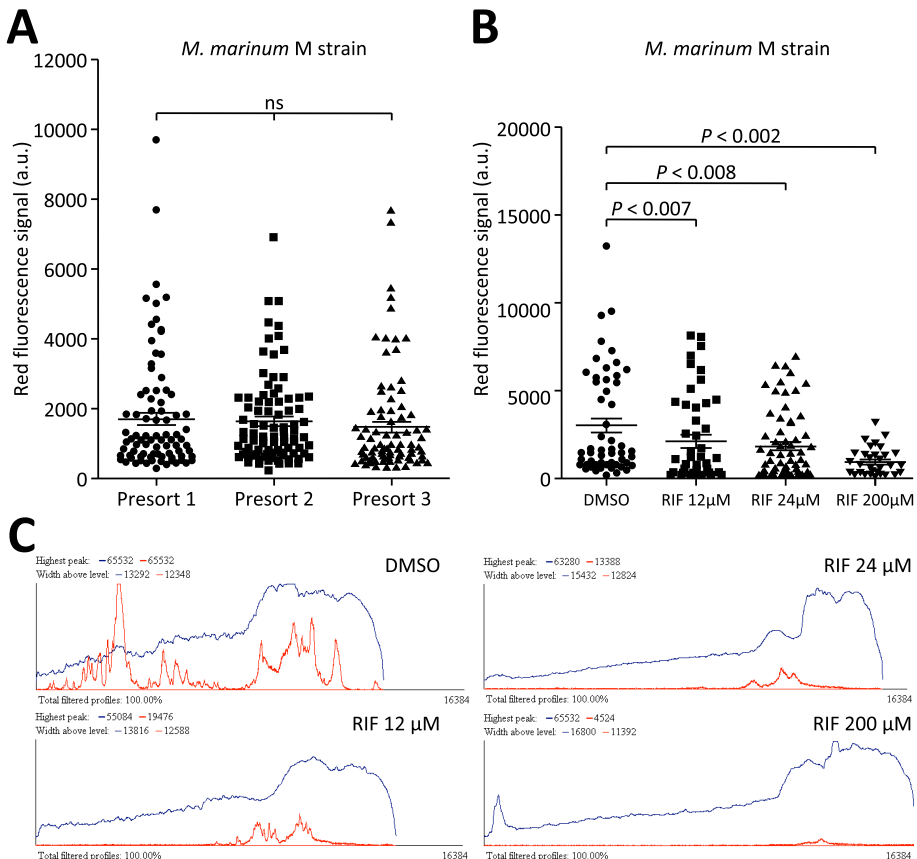

Figure 5: Treatment of *M. marinum* acute infection with a first-line anti-tuberculosis drug. Embryos injected between 16-64 cell stage with 30 cfu of *M. marinum* M strain were run through the large particle flow cytometer at 3 dpi to be sorted in two groups after discarding the non- and/or highly infected embryos. (**A**) Fluorescence of individual embryos in both groups. (**B**) Embryos treated with Rifampicin (RIF) for 48

hours at doses of 12, 24 and 200 μM were analyzed at 4 dpi; their bacterial load is significantly reduced. (**C**) Representative COPAS profiles of embryos treated with DMSO and Rifampicin at doses of 12, 24 and 200 μM for 24 hours. Bacterial load and distribution is indicated by the red peaks. Blue line represents the profile of the element sorted (4 dpf zebrafish embryo) by the COPAS. 60-90 embryos per condition were analyzed. Each data point represents an individual embryo. Values are indicated as mean ± SEM. ns: not significant differences. Analysis of statistical significance of differences was performed by one-way ANOVA followed by Tukey's post-hoc test.
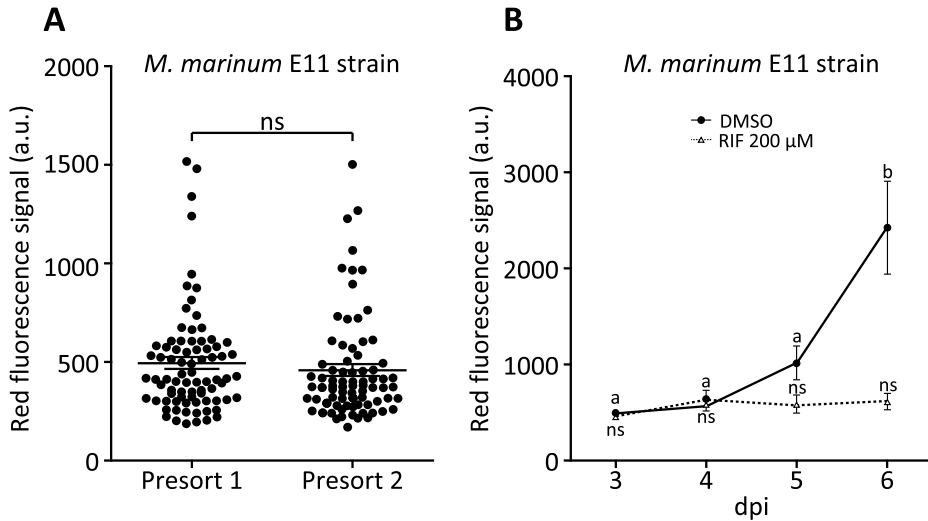

Figure 6: Treatment of *M. marinum* chronic infection with a first-line anti-tuberculosis drug. Embryos injected between 16-64 cell stage with 30 cfu of *M. marinum* E11 strain were run through the large particle flow cytometer at 3 dpi to be sorted in two groups after discarding the non- and/or highly infected embryos. (**A**) Fluorescence of individual embryos in both groups. (**B**) Embryos treated with Rifampicin (RIF) at 200 μM during 4 days were analyzed showing a significant reduction of the bacterial load after 1 day of treatment. 90 embryos per condition were analyzed. Values are indicated as mean ± SEM. Different letters indicate significant differences between time points of the same treatment. * indicates significant differences with control group. ns: not significant differences. Analysis of statistical significance of differences was performed by one-way ANOVA followed by Tukey's post-hoc test. (*P* < 0.05). Figure **B**) has been modified from Spaink et al. 2013 [13].

Furthermore, if high magnification imaging is required of these embryos, they can be displayed automatically in 96 well plates (Figure 1C), from where the samples can be analyzed using the Vertebrate Automated Screening Technology system with the Large Particle Sampler mounted onto a CLSM. The Vertebrate Automated Screening Technology system with the Large Particle Sampler is a system that can either be mounted onto a CLSM or stereo microscope. This device allows the loading of live or fixed embryos from a 96 well plate or bulk container automatically through a glass capillary, and orientates it in front of the camera at the desired angle (e.g. dorsal or lateral). Images of the embryo in all orientations can be made with the on board camera or with an external CLSM (Figure 7). Embryos will subsequently be transferred in the collection or waste container.
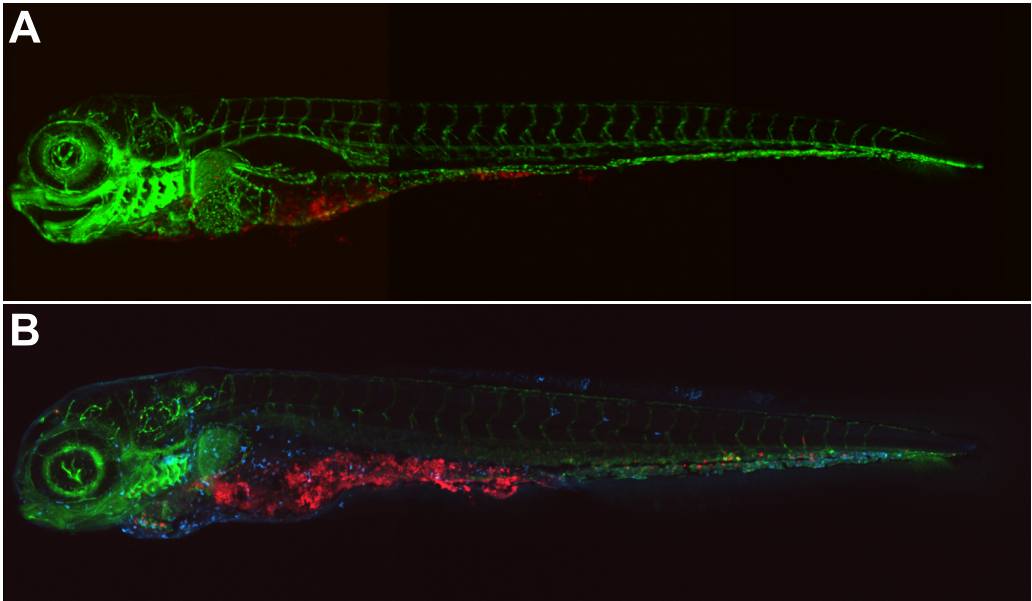
Figure 7: Result of *M. marinum* E11 yolk injection imaged using Vertebrate Automated Screening Technology and CLSM. Confocal Z stack (stitched 3 images) of a 5 dpi *Tg(fli1-egfp)* [17] embryo. (**A**) Live embryo showing proliferation of *M. marinum* E11 bacteria (red) throughout the body. (**B**) Fixed 5 dpi *Tg(fli-egfp)* embryo showing *M. marinum* E11 bacteria (red) throughout the body co-localizing with leucocytes (light blue) detected by L-plastin immunostaining [18].

Discussion

The high-throughput methodology described in this paper provides a fast and cost effective pipeline to screen high number of fish embryos and larvae with different types of infections. Using the large breeding vessel instead of traditional single or family breeding tanks facilitated control of the spawning process and generation of larger number of synchronous eggs. With an improved version of the automated micro-injection system [7], it is possible to inject up to 2500 eggs almost all in the same cell stage within 1 hour. With these updates and improved software it is feasible to inject more eggs than previously was possible which can be used to perform large drug screens with bacterial proliferation as a read out. However this method is still limited to yolk injection, other injection routes for example described by Benard et al. 2012. [6], will hopefully be incorporated in the automated micro-injection system in the near future.

Although these methods are benchmarked for screening zebrafish, it would be useful for applications with other fish species as well. For instance, the common carp has been indicated to have advantages for drug screens. Like zebrafish, eggs and early stage embryos from common carp are transparent but with the main advantage of its large spawn size of hundred thousands of eggs and the availability of inbred lines that offer a more constant genetic background [14].

The analysis of large amounts of infected embryos is done with the high-throughput large particle flow cytometer. This device can sort analyzed embryos into multi well plates or a Petri dish making it especially suitable for testing large numbers of compounds. If a higher imaging resolution is needed, than the setup is adapted in a way that the large particle flow cytometer technology can be used for pre-screening and subsequently analyze the samples at a medium throughput at a higher resolution. This can be done using the Vertebrate Automated Screening Technology [15, 16]. This device can automatically collect live or fixed embryos between 2 and 7 days post fertilization from a multi well plate or bulk container, image 360° through a capillary using CLSM or stereo microscopy and dispose again in 2 bulk containers allowing manual sorting of the embryos based on the microscopic images. Future improvements will allow the sorting of the embryo after imaging into the multi well plate, therefore making it possible to screen automatically large number of individual embryos over time with CLSM. Assuming that in future applications the Vertebrate Automated Screening Technology system can also be connected to the large particle flow cytometer technology without the need of prior dispensing larvae into multi well plates, will lead to a more advanced sorting.

This paper describes the establishment and optimization of a high-throughput setup to study *S. epidermidis* and *M. marinum* infection as a model for drug discovery. It demonstrates that the outcome of these bacteria injected into the yolk depends on the developmental stage of the eggs at the time of injection. Injecting *M. marinum* E11 at 16-128 cell stage or the M strain at 16-64 cell stage leads to the same infection pattern as caudal vein injection [2, 6]. However this setup is not limited to the proliferation of bacterial pathogens only. It was shown before that it is possible to robotically inject solutions containing DNA, RNA or morpholinos for transgenesis, over-expression and gene knock-down studies, respectively [13]. Furthermore, it was shown that this setup is also useful for the study of cancer cell proliferation and migration. Therefore this pipeline presents a versatile method for high-throughput screens of a variety of signal mechanisms in the context of innate immunity, applied to infectious disease and the development of cancer. These screens can be combined with others for medicine discovery but also with analysis of possible toxic effects of identified applicable drugs.

Acknowledgments

Supplementary material

Supplementary material can be found on:

http://www.jove.com/video/51649/establishment
-optimization-high-throughput-setup-to-study

References

1.   Meijer AH, Spaink HP: Host-pathogen interactions made transparent with the zebrafish model. Current drug targets 2011, 12(7):1000-1017.

2.   Tobin DM, Ramakrishnan L: Comparative pathogenesis of Mycobacterium marinum and Mycobacterium tuberculosis. Cell Microbiol 2008, 10(5):1027-1039.

3.   Boelens JJ, Dankert J, Murk JL, Weening JJ, van der Poll T, Dingemans KP, Koole L, Laman JD, Zaat SA: Biomaterial-associated persistence of Staphylococcus epidermidis in pericatheter macrophages. J Infect Dis 2000, 181(4):1337-1349.

4.   Broekhuizen CA, Schultz MJ, van der Wal AC, Boszhard L, de Boer L, Vandenbroucke-Grauls CM, Zaat SA: Tissue around catheters is a niche for bacteria associated with medical device infection. Crit Care Med 2008, 36(8):2395-2402.

5.   Busscher HJ, van der Mei HC, Subbiahdoss G, Jutte PC, van den Dungen JJ, Zaat SA, Schultz MJ, Grainger DW: Biomaterial-associated infection: locating the finish line in the race for the surface. Sci Transl Med 2012, 4(153):153rv110.

6.	Benard EL, van der Sar AM, Ellett F, Lieschke GJ, Spaink HP, Meijer AH: Infection of zebrafish embryos with intracellular bacterial pathogens. Journal of visualized experiments : JoVE 2012(61).

7.	Carvalho R, de Sonneville J, Stockhammer OW, Savage ND, Veneman WJ, Ottenhoff TH, Dirks RP, Meijer AH, Spaink HP: A high-throughput screen for tuberculosis progression. PloS one 2011, 6(2):e16779.

8.	Veneman WJ, Stockhammer OW, de Boer L, Zaat SA, Meijer AH, Spaink HP: A zebrafish high throughput screening system used for Staphylococcus epidermidis infection marker discovery. BMC genomics 2013, 14(1):255.

9.	van der Sar AM, Abdallah AM, Sparrius M, Reinders E, Vandenbroucke-Grauls CM, Bitter W: Mycobacterium marinum strains can be divided into two distinct types based on genetic diversity and virulence. Infect Immun 2004, 72(11):6306-6312.

10.	Stoop EJM, Schipper T, Rosendahl Huber SK, Nezhinsky AE, Verbeek FJ, Gurcha SS, Besra GS, Vandenbroucke-Grauls CMJE, Bitter W, van der Sar AM: Zebrafish embryo screen for mycobacterial genes involved in the initiation of granuloma formation reveals a newly identified ESX-1 component. Disease Models & Mechanisms 2011, 4(4):526-536.

11.	Adatto I, Lawrence C, Thompson M, Zon LI: A new system for the rapid collection of large numbers of developmentally staged zebrafish embryos. PloS one 2011, 6(6):e21715.

12.	Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF: Stages of embryonic development of the zebrafish. Dev Dyn 1995, 203(3):253-310.

13.	Spaink HP, Cui C, Wiweger MI, Jansen HJ, Veneman WJ, Marin-Juez R, de Sonneville J, Ordas A, Torraca V, van der Ent W et al: Robotic injection of zebrafish embryos for high-throughput screening in disease models. Methods 2013, 62(3):246-254.

14.	Henkel CV, Dirks RP, Jansen HJ, Forlenza M, Wiegertjes GF, Howe K, van den Thillart GE, Spaink HP: Comparison of the exomes of common carp (Cyprinus carpio) and zebrafish (Danio rerio). Zebrafish 2012, 9(2):59-67.

15.	Chang TY, Pardo-Martin C, Allalou A, Wahlby C, Yanik MF: Fully automated cellular-resolution vertebrate screening platform with parallel animal processing. Lab Chip 2012, 12(4):711-716.

16.	Pardo-Martin C, Chang TY, Koo BK, Gilleland CL, Wasserman SC, Yanik MF: High-throughput in vivo vertebrate screening. Nat Methods 2010, 7(8):634-636.

17.	Lawson ND, Weinstein BM: In vivo imaging of embryonic vascular development using transgenic zebrafish. Dev Biol 2002, 248(2):307-318.

3

18.   Mathias JR, Dodd ME, Walters KB, Rhodes J, Kanki JP, Look AT, Huttenlocher A: Live imaging of chronic inflammation caused by mutation of zebrafish Hai1. J Cell Sci 2007, 120(Pt 19):3372-3383.

3

Chapter 4

A zebrafish high-throughput screening system used for *Staphylococcus epidermidis* infection marker discovery

4

Wouter J. Veneman[1], Oliver W. Stockhammer[2], Leonie de Boer[2], Sebastian A.J. Zaat[2], Annemarie H. Meijer[1], Herman P. Spaink[1]

[1] Institute of Biology, Leiden University
[2] Department of Medical Microbiology, Academic Medical Center/UVA

Abstract

Staphylococcus epidermidis bacteria are a major cause of biomaterial-associated infections in modern medicine. Yet there is little known about the host responses against this normally innocent bacterium in the context of infection of biomaterials. In order to better understand the factors involved in this process, a whole animal model with high-throughput screening possibilities and markers for studying the host response to *S. epidermidis* infection are required.

We have used a zebrafish yolk injection system to study bacterial proliferation and the host response in a time course experiment of *S. epidermidis* infection. By combining an automated micro-injection system with Complex Object Parametric Analysis and Sorting (COPAS) technology we have quantified bacterial proliferation. This system was used together with transcriptome analysis at several time points during the infection period. We show that bacterial colony forming unit (cfu) counting can be replaced by high-throughput flow-based fluorescence analysis of embryos enabling high-throughput readout. Comparison of the host transcriptome response to *S. epidermidis* and *Mycobacterium marinum* infection in the same system showed that *M. marinum* has a far stronger effect on host gene regulation than *S. epidermidis*. However, multiple genes responded differently to *S. epidermidis* infection than to *M. marinum*, including a cell adhesion gene linked to specific infection by staphylococci in mammals.

Our zebrafish embryo infection model allowed (i) quantitative assessment of bacterial proliferation, (ii) identification of zebrafish genes serving as markers for infection with the opportunistic pathogen *S. epidermidis*, and (iii) comparison of the transcriptome response of infection with *S. epidermidis* and with the pathogen *M. marinum*. As a result we have identified markers that can be used to distinguish common and specific responses to *S. epidermidis*. These markers enable the future integration of our high-throughput screening technology with functional analyses of immune response genes and immune modulating factors.

Introduction

Infections with *Staphylococcus epidermidis* bacteria pose a serious problem associated with the use of biomaterials in modern medicine [1-5]. These bacteria can form biofilms on the surface of inserted biomaterials and persist in the surrounding tissues, where immune functions are disturbed due to the combined presence of a biomaterial and the bacteria [6, 7]. In order to better understand the cause of this phenomenon and to assess the propensity of different bacterial strains and biomaterials to alter and trigger the immune response in the host, a whole animal model with high-throughput screening possibilities is desired. This will help identifying which factors determine that innocent bacteria become less susceptible to host defence mechanisms or antibiotic treatments when associated with biomaterials. Mouse and rat models have been used

to investigate *S. epidermidis* infection and biomaterial-associated infection processes. However, histological examination of biopsies is time consuming and does not allow following the infection process over time [8-10]. Even with the use of bioluminescence and fluorescence imaging, high challenge doses are required to visualize bacterial colonization and high-throughput screening in rodents is not feasible. However, the zebrafish at the embryonal and larval stages is an excellent model for this purpose: it is translucent, fluorescently labelled immune cells and bacteria can be microscopically imaged in real time, and embryos can be obtained in high numbers [11-15]. The responses of many different pathogens such as *Escherichia coli*, *Mycobacterium marinum*, *Salmonella typhimurium*, *Edwardsiella tarda*, *Burkholderia cenocepacia*, and *Staphylococcus aureus* have already been assessed in zebrafish [16-21]. In previous work we successfully performed extensive transcriptome analyses with *M. marinum*, *S. typhimurium*, and *E. tarda* intravenous infection models using custom made Agilent micro-arrays and deep sequencing [19, 22-24]. The conventional infection method for zebrafish embryos is injection of pathogens into the caudal vein. However, this method is labour intensive and low throughput. For that reason we have recently developed and validated a high-throughput yolk infection model using *M. marinum* with an automated micro-injection system [12]. However, in this high-throughput model no transcriptome analysis has been performed until now. In the present study we have developed a high-throughput system for quantitating infection with *S. epidermidis* using the automated micro-injection system together with Complex Object Parametric Analysis and Sorting (COPAS) technology. This quantitative high-throughput technology has been used to study the transcriptome responses during non-lethal infection progression of *S. epidermidis* over time using micro-arrays and RNA deep sequencing. In order to understand which responses can be linked to defence mechanisms of the zebrafish towards fish pathogens, we have compared the host responses to *S. epidermidis* and to *M. marinum* at a time point when the initial yolk infection has further spread into the embryo's tissues. The obtained results allowed us to identify a number of genes as markers common for both infection models but also genes that can be used as markers to discriminate between pathogen specific responses.

Results and discussion

Pathogenesis of *S. epidermidis* and *S. aureus* in zebrafish embryos

We first set out to compare *S. epidermidis* infected zebrafish embryos with embryos infected with *S. aureus*. For this purpose we injected *S. epidermidis* O-47 and *S. aureus* RN4220 strains containing GFP or mCherry plasmids under the same conditions into the yolk of embryos at 2 hours post fertilization (hpf). Injections with 5 cfu of *S. aureus* already showed a high intensity of fluorescent bacteria inside the yolk at the first day after injection. At the second day after injection all embryos had died from infection with bacteria spread inside the entire body of the embryos (data not shown). Injection directly into the caudal vein at 28 hpf with approximately 2500 cfu resulted in 100%

mortality within several hours (data not shown). This high early mortality due to *S. aureus* is in accordance with earlier reports [20, 25, 26]. We subsequently tested *S. epidermidis* O-47 in yolk injections at doses of 5, 20, 50 or 100 cfu. At 1 day post injection (dpi) several small spots of fluorescent bacteria were observed inside the yolk with all cfu doses (Figure 1A), which were absent in mock-injected controls. From 2 dpi onwards, bacteria, indicated by their fluorescence signal, were visible inside the yolk in a dose-depended fashion. The fluorescence signal became detectable inside the body of the embryos starting at 3 dpi (Figure 1A). From this day onwards the bacteria were persisting in the vascular system and within various tissues. Confocal laser scanning microscopy was used to obtain a detailed image of bacteria spreading into the different tissues at 3, 4 and 5 dpi, the time points at which spreading of bacteria was observed (Figure 2). At 3 dpi bacteria were found intracellularly and extracellularly in the hematopoietic region (Figure 2, panel 1), and free existing bacteria were observed in the blood were taken up by mpeg1:KAEDE positive cells (Figure 2, panel 4). Free staphylococci in the blood have also been observed after intravascular catheter-related infections [27, 28]. At 4 dpi much more extracellular bacteria in the intersegmental vessels were seen (Figure 2, panel 2). No differences were found between the patterns observed at 4 and 5 dpi (Figure 2, panel 3). Although there was strong increase of the bacterial burden in tissues and blood at 4 and 5 dpi (Figure 2 panel 2 & 3 and Supplementary file 1), in most cases embryos survived the 5 days infection period similar as the mock-injected controls. Cfu counts of homogenized pooled embryos revealed that *S. epidermidis* proliferated exponentially inside the embryos during the 5 days of infection (Figure 1B). Comparing the yolk injection method with the traditional caudal vein injection method showed that embryos injected with as much as 5000 and 10000 cfu of *S. epidermidis* into the caudal vein at 28 hpf did not develop any signs of infection. Fluorescence microscopy
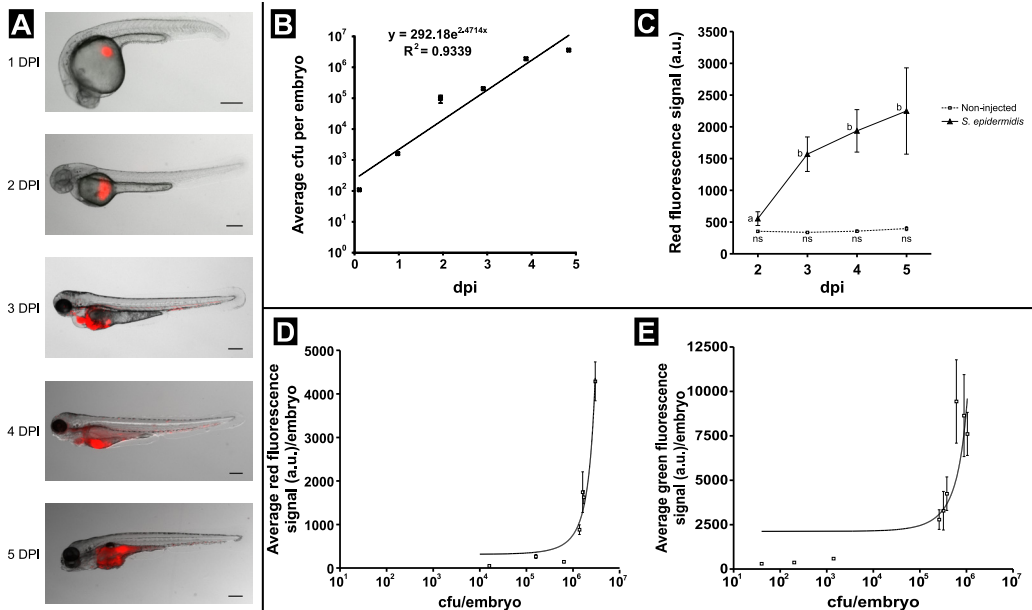
(See legend on next page)

Figure 1. Quantitation of fluorescence intensity in *S. epidermidis*-injected embryos using the COPAS system. (**A**) Bright field /fluorescence overlay images of mCherry-labelled *S. epidermidis*. Wild type zebrafish embryos injected with 100 cfu of *S. epidermidis* O-47 into the yolk at 2 hpf were imaged at 5 time points from 1 to 5 dpi, scale bar is 250 µm. (**B**) Cfu counts of *S. epidermidis*-infected embryos. Groups of 10 embryos were homogenized and plated directly after injection until 5 dpi. (**C**) The graphs represent the average fluorescence intensity from the entire group of non-injected and *S. epidermidis*-injected embryos, from 2 dpi until 5 dpi. An increase in fluorescence intensity is visible during this infection period. (Error bars = SEM). Different letters indicate statistical significant differences (*P* < 0.001). (**D** & **E**) Correlation between cfu counts and fluorescence intensity of embryos infected with mCherry-labelled (**D**) and GFP-labelled (**E**) bacteria. Pools of 10 infected embryos between 2 and 5 dpi were homogenized and plated. The average fluorescence intensity is plotted against the cfu count.
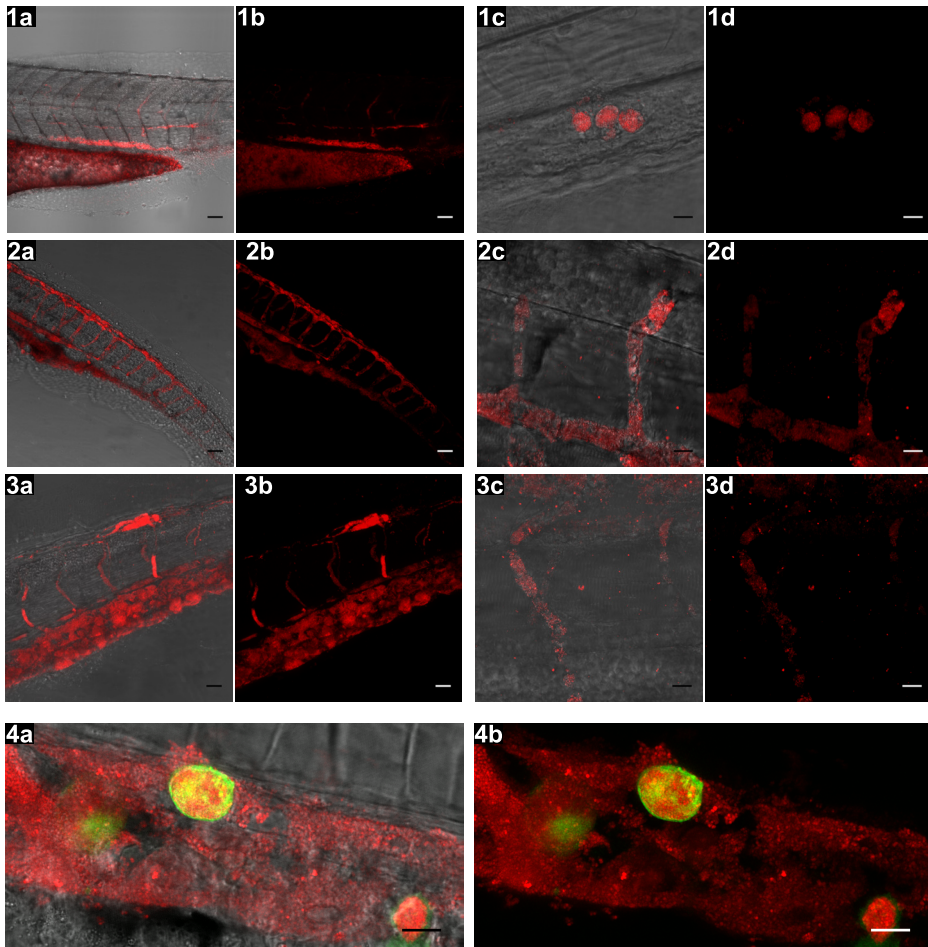

4

Figure 2: Invasion of *S. epidermidis* into the zebrafish embryo body. Confocal z-stacks are shown as transmission/fluorescence overlay (**A** & **C**) and fluorescence images (**B** & **D**). Panel 1: at 3 dpi mCherry labelled *S. epidermidis* is observed inside the body (**1A** &**1B**, scale bar: 50 µm), and intracellular in the hematopoietic region (**1C** & **1D**, scale bar: 10 µm). Panel 2: at 4 dpi bacteria are found inside the vasculature (**2A** & **2B**, scale bar: 50 µm), including the intersegmental vessels (**2C** & **2D**, scale bar: 10 µm). Panel 3: at 5 dpi bacteria are still persisting in the vasculature (**3A** & **3B**, scale bar: 25 µm) and in the intersegmental vessels (**3C** & **3D**, scale bar: 10 µm). (**4A** & **4B**) bacteria being taken up by mpeg1:KAEDE positive cells and extracellular in the hematopoietic region at 3 dpi (scale bar: 10 µm).

showed that all injected bacteria were cleared within several hours after injection (data not shown). In view of this, we conclude that the yolk infection system is therefore uniquely suitable to follow the proliferation of *S. epidermidis* and its effects on the host for at least 5 dpi. At the moment we can only speculate why the bacterial that were injected in the yolk had such better survival rates than bacteria injected in the caudal vein at later stages. Three possible explanations (or a combination of these factors) are that (1) there were repeated cycles of invasion from the yolk, (2) the bacteria in the yolk are primed to an infectious growth strategy for instance by using alternate sigma factors [29], or (3) the host immune system has been altered due to the prolonged exposure to high numbers of bacteria and possible associated anti-inflammatory compounds inside the embryos.

High-throughput infection quantification

The COPAS XL (Union Biometrica, USA) is a large cell flow cytometer designed for fluorescence screening of zebrafish embryos, Drosophila larvae and beads ranging from 1500 to 2000 microns in diameter [12, 30]. Samples are analysed for size, optical density and three fluorescence signals. Groups of up to 3000 embryos can simultaneously be analysed and sorted into multi well plates or Petri dishes within 15 minutes (Figure 3). The Profiler software package II detects and analyses up to 8000 data points per object for the extinction and fluorescence channels, and can be used to visualize every sample or to set sorting parameters. The Profiler shows the outline of a sample together with all fluorescence intensity traces, for each of the embryos in a sample. This is exemplified by a typical experiment of *S. epidermidis* infection of zebrafish embryos measured at 4 dpi (Figure 3, panel B, and detailed in Supplementary file 2). COPAS analysis was performed every day from 2 dpi until 5 dpi. The daily analysis did not cause noticeable damage to the embryos. We observed an increase in the fluorescence signal during the 5 days of infection with *S. epidermidis* (Figure 1C). Cfu count results showed good correlation with the increase of fluorescence signal in pools of embryos infected with mCherry-labelled (Figure 1D) or GFP-labelled (Figure 1E) bacteria. However, in the green channel some background signal produced by the embryonic yolk was detected, leading to less accurate quantification (Figure 1E). Since the red fluorescence channel did not show any background signal, the results with mCherry-labelled bacteria were quantitatively more reliable, showing a correlation with the cfu counts (Figure 1D). We did not find any influence of the orientation of the embryos in the flow chamber since we did not detect differences in embryos passing the laser dorsally or ventrally, or with the anterior or posterior side first. Therefore, our results show that combining the COPAS analysis with the automated micro-injection system provides a screening system of which the infection levels are statistically reliable.
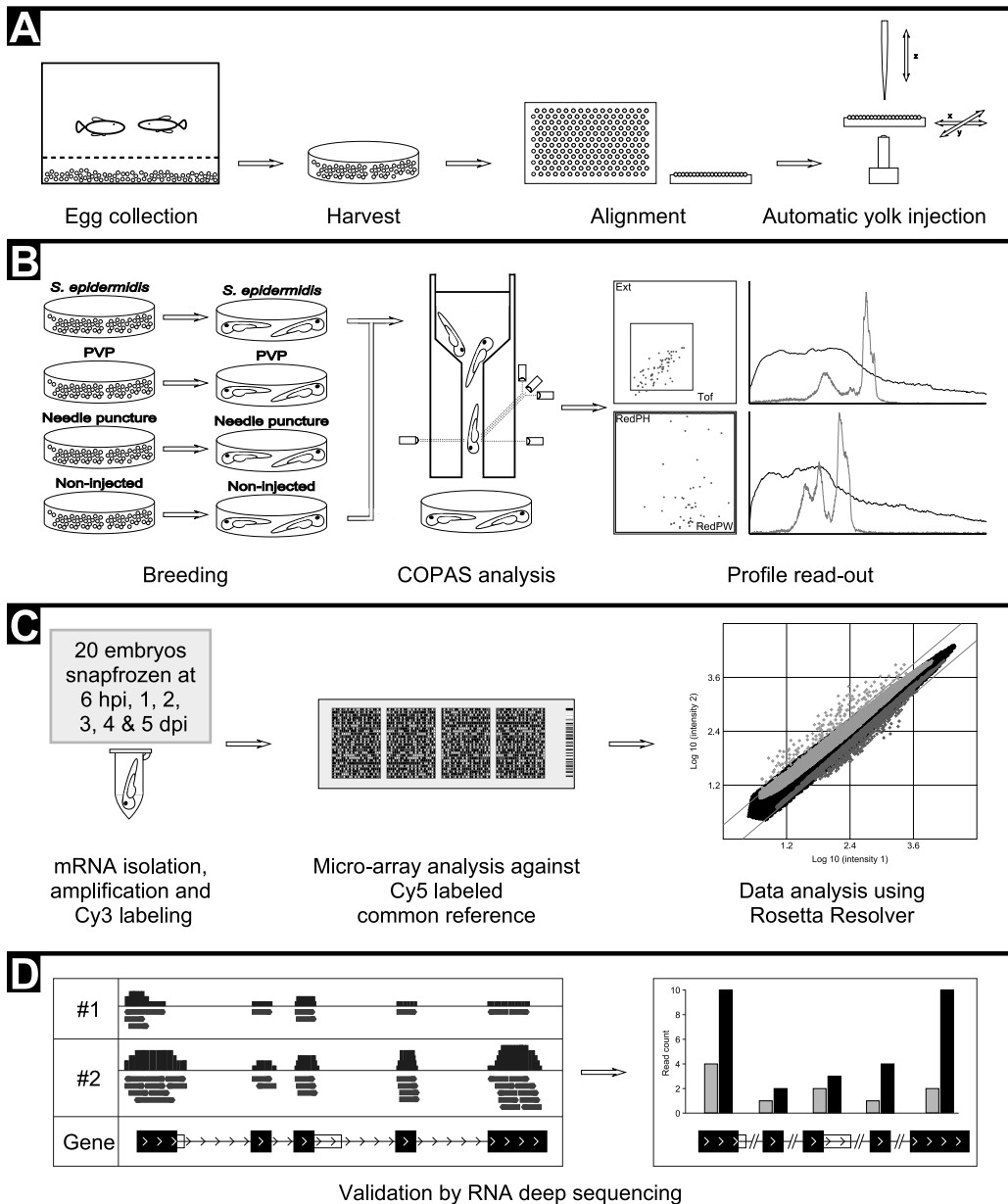
Figure 3. Workflow of high-throughput injection and subsequent analysis. (**A**) from left to right; a zebrafish pair are put together to mate, eggs are collected, eggs are distribute into a 1024 well agarose grid, eggs are injected into the yolk at 2 hpf using the automated micro-injection system. (**B**) from left to right; after injection, eggs are collected into Petri dishes and incubated at 28°C for a period of 5 days, COPAS analysis is performed on the *S. epidermidis* and non-injected embryos at 2, 3, 4 and 5 dpi. (**C**) from left to right; from all groups 20 embryos are snap frozen at 6 hpi, 1, 2, 3, 4 and 5 dpi for RNA isolation, amplification and Cy3 labelling, micro-array analysis against Cy5 labelled common reference and data analysis using Rosetta Resolver. (**D**) from left to right; validation of micro-array data was performed by RNAseq analysis of 4 biological replicas of *S. epidermidis* infected embryos at 5 dpi.

Specific marker genes for *S. epidermidis* infection

To characterize the transcriptome response of zebrafish embryos following *S. epidermidis* yolk injection, we performed a time resolved infection experiment using the high-throughput set up (Figure 3). Considering that all bacterial injections were carried out with polyvinylpyrrolidone$_{40}$ (PVP) as carrier, PVP-injected embryos were taken along to control for possible effects of the carrier. Furthermore, needle puncture treated and non-injected embryo groups were included as additional control groups. Injections were performed with groups of at least 150 embryos of the same parents of which sets of 20 embryos were sampled during 6 time points (Figure 3). In order to check for reproducibility of this experiment an independent experiment was performed with the same parents at 4 dpi. RNA from these samples was used for micro-array analysis using custom made Agilent 4x180k micro-arrays.

Principal component analysis showed a clear signature progression in time of all samples (Supplementary file 3). Results of statistical analyses are presented in the Venn diagrams of Figure 4. Comparing each time point with multiple control samples clearly shows that there is a false negative effect in the controls that can be corrected for by using the overlap of the ratios of the different controls. This led to a filtered dataset as used for Figure 5 as discussed below. However, we want to emphasize that the injection of PVP has a reproducible effect by itself (Figure 4). This could be of relevance, especially considering the effect of biomaterials on infection capacity of *S. epidermidis* in patients that make it worthwhile to further analyse the effect of PVP on infection in future experiments. We have analysed the effect of *S. epidermidis* infection over time on gene expression using Unigene clusters and ENSEMBL codes as specified in the raw data table of Supplementary file 4. Annotation of these probes by Gene Ontology (GO) shows that the most noticeable result is that *S. epidermidis* does induce many immune-related genes starting from 3 dpi, observing a maximum induction of their expression at 4 dpi. Filtering the results (Figure 5) we found an effect on the expression levels of many genes in the earliest measured time point (6 hpi) after exposure to injected bacteria. This effect is diminished to only a few genes whose expression is affected at the time point of 2 dpi, most of which cannot be assigned to a GO category (Figure 4). At 6 hpi, GO analysis indicated very broad classes of gene functions whose expression are affected by infection but did not reveal an obvious link to the immune response since the broad GO category "immune response" was not represented. At this stage the known innate immune responses to bacterial infection are not yet apparent. For instance neutrophils and macrophages have not yet developed and nothing is known about the function of pattern recognition receptors before this stage. We are currently studying the function of the expressed Toll-like receptors during early stages of embryogenesis [31]. We have manually annotated various functional categories of genes, of which the transcription levels were strongly affected by infection during time as shown in Figure 5 in a schematic representation and in Supplementary file 5 in a quantitative manner. Many of the immune genes indicated in Figure 5 have been previously linked to expression in cells
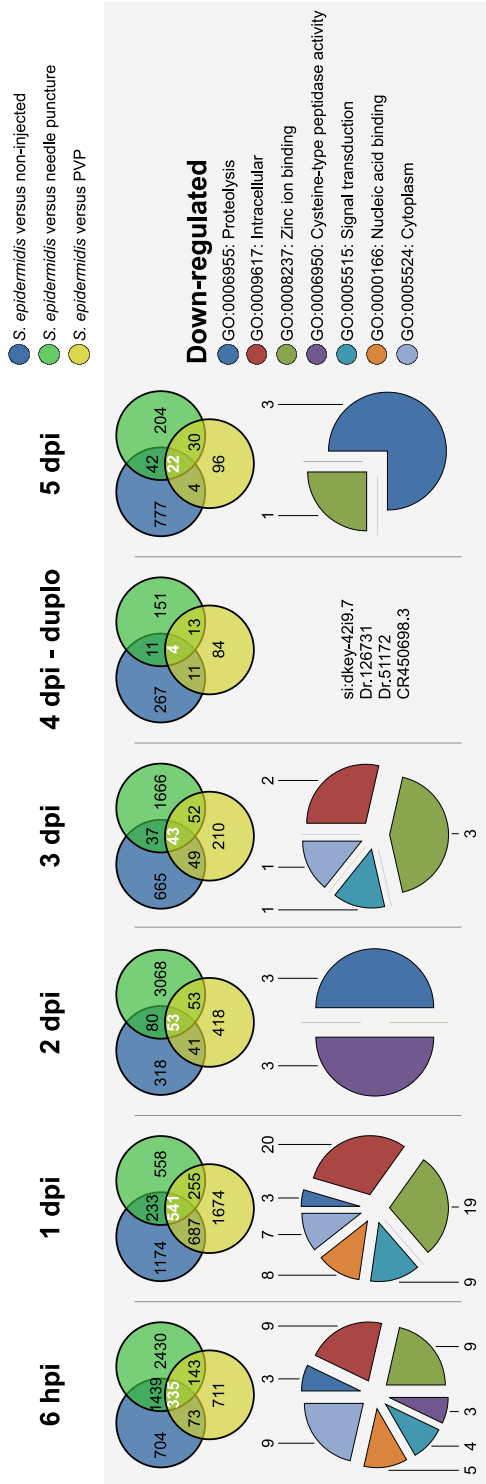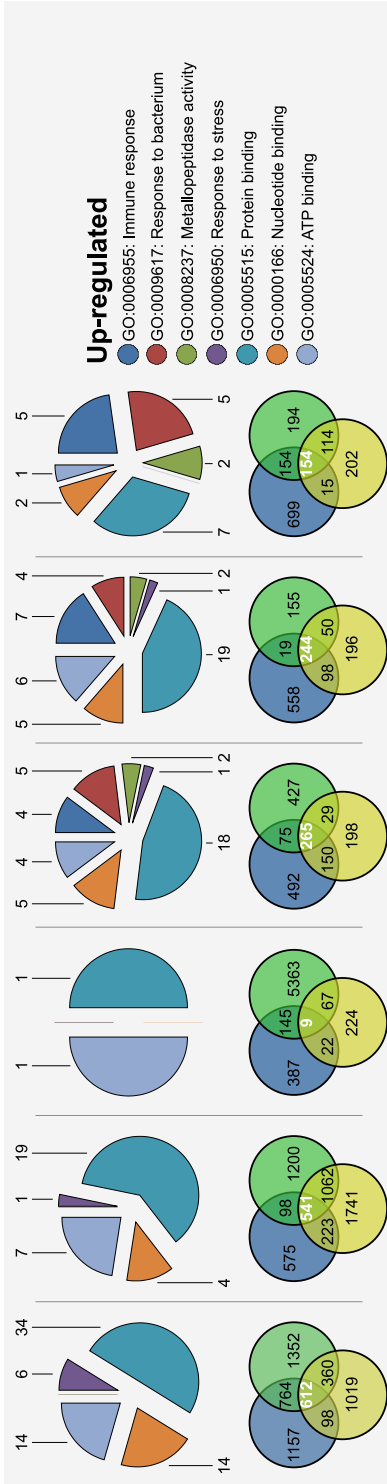
(See legend on next page)

Figure 4. Overlapping probes in time from micro-array analysis. The Venn diagrams show the number of significantly up-regulated probes (top row) or down-regulated probes (bottom row) (*P*-value smaller than 10^{-8} and fold changes larger than 2 or smaller than -2) between *S. epidermidis* injected versus non-injected, *S. epidermidis* injected versus needle puncture and *S. epidermidis* injected versus PVP at 6 hpi, 1, 2, 3, 4 and 5 dpi. Data at 4 dpi are based on a biological replica. Pie diagrams represent GO annotation using DAVID bioinformatics resources 6.7 [40, 41] from the overlapping genes (in white) at the representative time points.

of the myeloid lineage in zebrafish [32]. The immune transcriptome response correlates with the infection progression as described above. The first 3 days, the bacteria accumulate inside the embryonic yolk. This apparently does not lead to significantly induction or repression of many immune-related genes. From 3 dpi onwards many immune-related genes were significantly induced with a peak at 4 dpi (Figure 5). At 5 dpi there were slightly less immune-related genes significantly expressed than at 4 dpi. Expression levels of the 49 selected genes shown in Figure 4 at 5 dpi were also lower compared with 4 dpi (Supplementary file 5). Since the analysed larvae were from the same injected batch it seems that higher microbial burden is not strictly correlated with a stronger immune response. Validation of this micro-array experiment was performed by deep sequencing analysis of RNA samples derived of 4 batches of approximately 150 embryos at the 5 day time point of infection and 4 non-infected batches of embryos. The data confirms the micro-array data as exemplified for some of the most reliable probes (Table 1). Furthermore we have compared the normalized reads per kilobase per million mapped reads (RPKM) [33] for the genes shown in Figure 5. These comparisons show that only in a few cases there are discrepancies between the results of the two technologies. Since the RNA deep sequencing results are obtained with a pool of larger number of biological samples, this could indicate that in these cases the micro-arrays result are less trustworthy. However, it was noted that in cases of discrepancy there were extremely low levels of expression resulting in a very limited number of mapped reads, showing that even with a sequencing depth of at least 20 million reads per sample there is still a limitation of sensitivity of RNA sequencing. This is of note because in most publications currently a sequence depth of 20 million reads is standard for RNA deep sequencing [34, 35]. In several cases, such as *il8*, there was no ENSEMBL annotation of the gene that could be used for RPKM analysis. We have manually quantified the number of reads mapping to *il8* to show that there is also induction after infection as in the case of the micro-array analysis (Figure 6). RNA deep sequencing provides a much more detailed insight in gene regulation for instance showing also expression levels for every exon of the genes as shown for the representative genes *mmp9* and *il8* (Figure 6). With the expected progress in development of high-throughput bioinformatic pipelines for data visualization of RNA sequencing data sets, we and others will be able in the near future to further harvest information from our submitted expression datasets as to the effects of infection on differential splicing, transcription start sites or polyadenylation sites in the entire transcriptome.
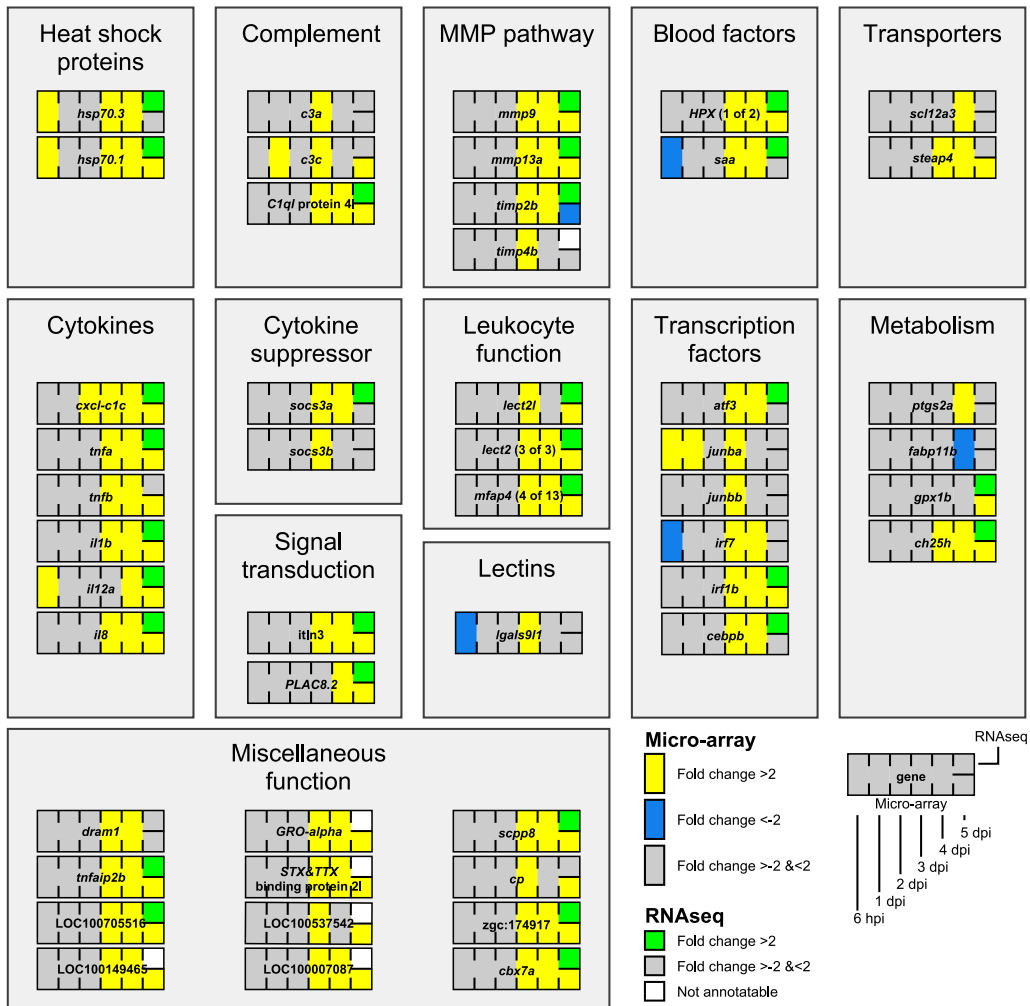
Figure 5. Gene expression during *S. epidermidis* infection. Micro-array data are shown of *S. epidermidis*-injected versus non-injected samples at 6 hpi, 1, 2, 3, 4 and 5 dpi time points. The yellow boxes represent up-regulation and the blue boxes down-regulation with a *P*-value smaller than 10^{-8} and fold changes larger than 2 or smaller than -2. The top right bar shows the RNA deep sequencing data, where the green boxes represent significant up-regulation. The white boxes could not be identified by RNA deep sequencing. Grey boxes mean that data did not meet the significant criteria. Genes were manually annotated and assigned to functional groups based on GO annotations of the zebrafish genes and their human homologues and on searching on PubMed abstracts. The *D. rerio* Uni-Gene Build # 124 or ENSEMBL Zv9 codes were used as shown with the raw data table in Supplementary file 4).

Table 1 Validation of micro-array data by RNA deep sequencing analysis. Shown are 8 representative immune-related genes that were significantly expressed in the micro-array and RNA deep sequencing experiment.

Gene	Sample	Micro-array		RNA deep sequencing	
		Fold change	*P*-value	Fold change	*P*-value
atf3	*S. epidermidis* versus non-injected (5 dpi)	+ 2.83	2.49×10^{-09}	+ 4.77	1.83×10^{-44}
cxcl-c1c	*S. epidermidis* versus non-injected (5 dpi)	+ 5.77	1.77×10^{-14}	+ 6.74	4.07×10^{-25}
HPX (1 of 2)	*S. epidermidis* versus non-injected (5 dpi)	+ 10.10	1.60×10^{-11}	+ 6.00	1.13×10^{-51}
il1b	*S. epidermidis* versus non-injected (5 dpi)	+ 11.94	1.22×10^{-10}	+ 8.54	9.97×10^{-38}
lect2l	*S. epidermidis* versus non-injected (5 dpi)	+ 6.99	1.80×10^{-20}	+ 3.12	1.17×10^{-21}
mfap4 (4 of 13)	*S. epidermidis* versus non-injected (5 dpi)	+ 21.27	4.15×10^{-41}	+ 3.67	4.30×10^{-09}
mmp9	*S. epidermidis* versus non-injected (5 dpi)	+ 5.18	9.73×10^{-20}	+ 10.71	1.54×10^{-95}
mmp13a	*S. epidermidis* versus non-injected (5 dpi)	+ 7.35	1.07×10^{-18}	+ 15.30	6.03×10^{-82}

Comparison of transcriptome responses to *S. epidermidis* and *M. marinum*

In order to compare the transcriptional response observed in zebrafish embryos infected with *S. epidermidis* with the response triggered by a pathogenic bacterium, we also performed an injection experiment with *M. marinum* using the same experimental protocol and sampling at 5 dpi. This time point was chosen in order to make it comparable to previous studies in which the caudal vein was used as the injection site [23]. We performed micro-array analysis, confirming the biological relevance of the yolk injection system since many markers that were previously identified to be differentially expressed in the caudal vein injection system [23] appeared regulated in a similar manner in the high-throughput yolk infection system (Figure 7 and Supplementary file 6). We observed a stronger transcriptional response of immune related genes than what we observed with caudal vein administration, which can be explained by the fact that bacteria accumulated more strongly after five days compared to the caudal vein injection method and have been present one more day inside the embryos. A number of the immune markers identified to be differentially expressed after infection with *S. epidermidis* appeared regulated in the same way after *M. marinum* yolk infection. These include the matrix metalloproteinases, complement factors, cytokines and heat shock proteins that were previously also identified in *M. marinum* infection in the caudal vein [23]. There are also distinct differences in genes responding to infection by these two different bacteria. Most obvious is a stronger transcriptional response of a number of relevant genes to *M. marinum* than to *S. epidermidis* infection. Interestingly, there is also a category of genes that are highly regulated by S. epidermis but not significantly by *M. marinum* in the yolk infection model. These genes include various immune related genes such as *il8*, *il12a*, *tnfb*, *lect2l*, and transcription factor *atf3*, *junba*, *junbb*,
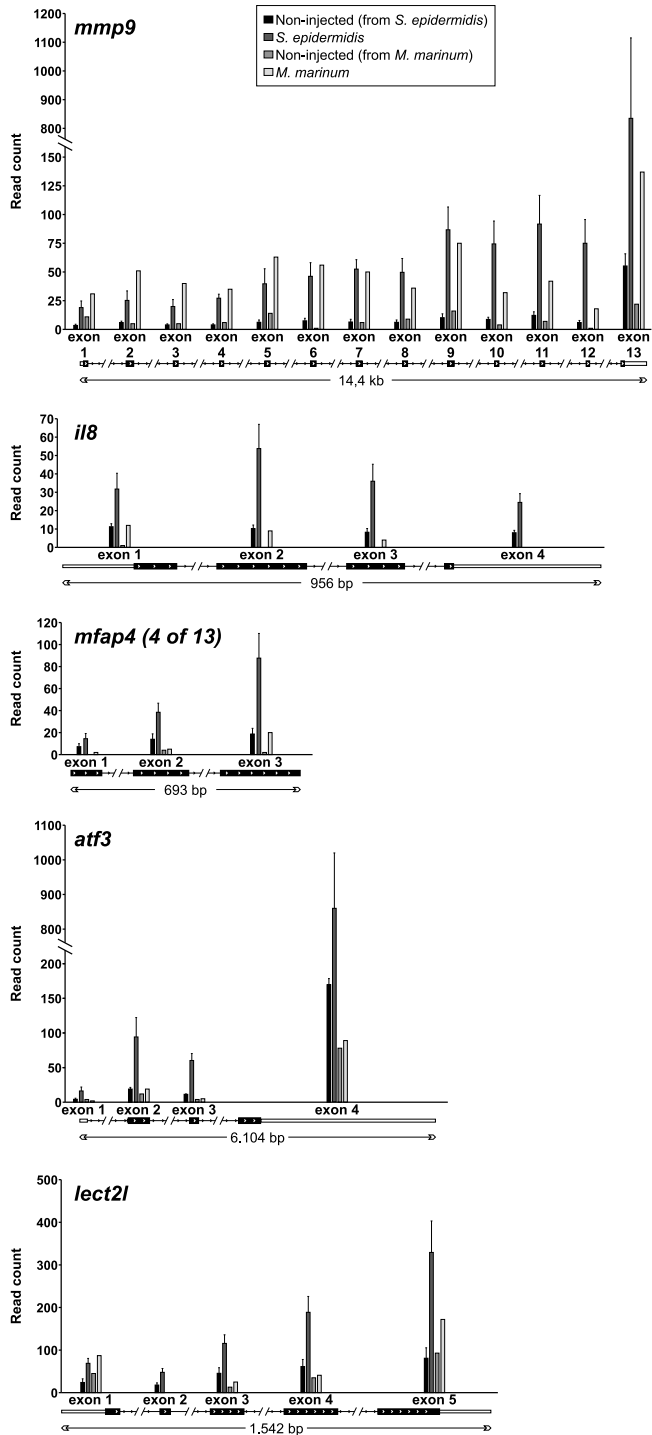
Figure 6. Expression levels of individual exons. All exons of *mmp9*, *il8*, *lect2l*, *mfap4* (4 of 13) and *atf3* were significant induced at 5 dpi following yolk infection with *S. epidermidis*. *M. marinum* yolk infection at the same time point only resulted in induction of all *mmp9* exons.

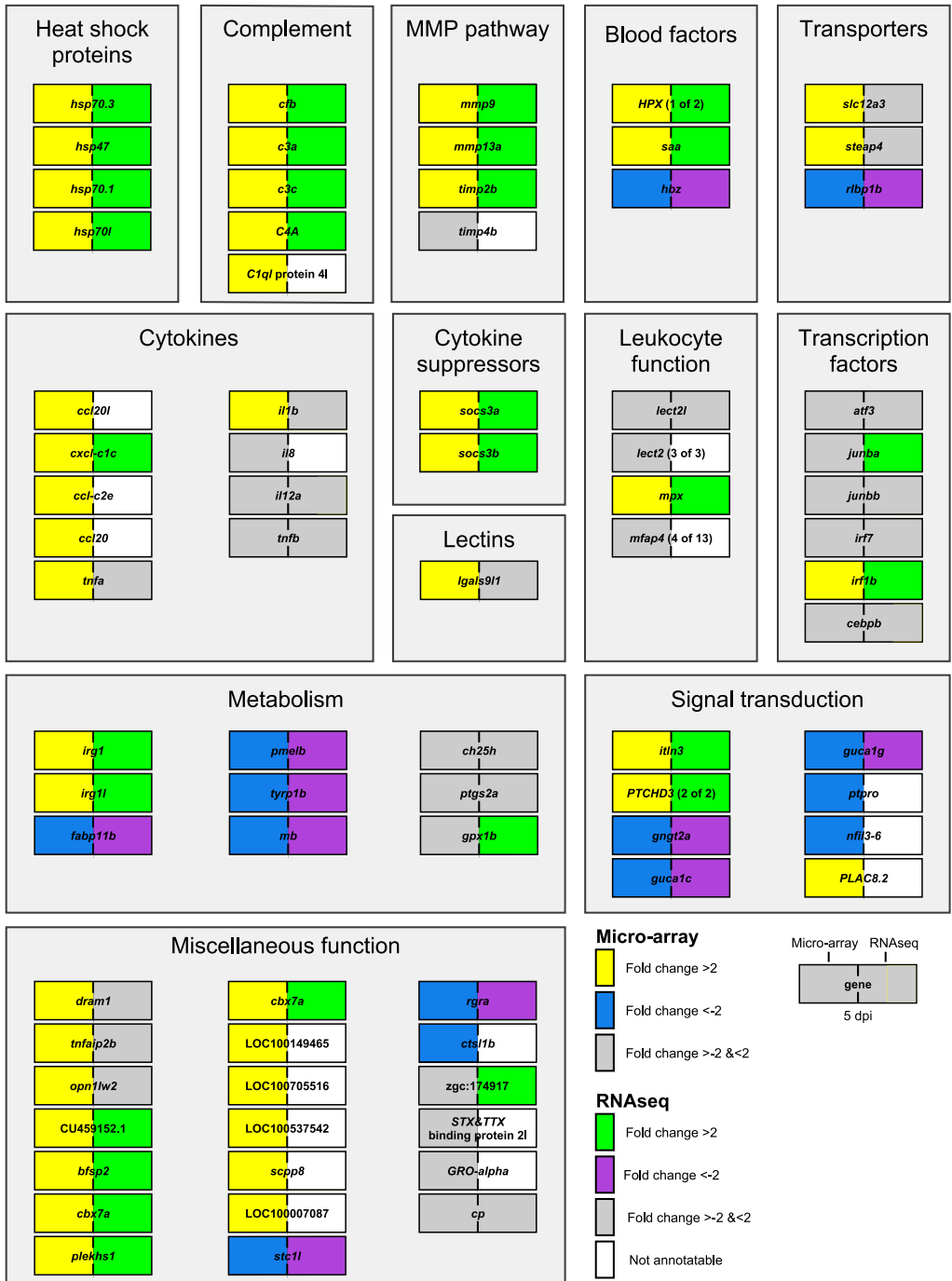**Figure 7. Gene expression during *M. marinum* E11 infection.** Micro-array data is shown for the infection of *M. marinum* at 5 dpi at the left side of each gene. The yellow boxes represent up-regulation and the blue boxes down-regulation with a *P*-value smaller than 10^{-8} and fold changes larger than 2 or smaller than -2. Probes with unchanged expression are indicated in grey. RNA deep sequencing data is shown for the same

sample at the right side of each gene. Green boxes represent up-regulation and the purple boxes down-regulation with a *P*-value smaller than 0.02 and fold changes larger than 2 or smaller than -2. Genes were manually annotated and assigned to functional groups based on GO annotations of the zebrafish genes and their human homologues and on searching on PubMed abstracts. The *D. rerio* Uni-Gene Build # 124 or ENSEMBL Zv9 codes were used as shown with the raw data table in Supplementary file 6).

irf7, *irf1b* and *cebpb*. However, in pilot micro-array studies where higher numbers of mycobacteria were injected, some of the markers were also induced or repressed with the exception of *il12a* and *cebpb* (data not shown). For *M. marinum* infection we observe some difference of gene regulation after yolk infection as compared to caudal vein injection. For instance, *atf3* is up-regulated in the latter system (unpublished results). One of the examples of genes induced specifically by *S. epidermidis* in the yolk infection system has high similarity to the mammalian microfibril-associated glycoprotein 4-like isoform 1 gene (*mfap4*-like isoform 1, Unigene accession number Dr.149043, also called *mfap4* (4 of 13) in ENSEMBL) which has been identified previously by Schlosser et al. [36] to bind to human Surfactant protein A (*SP-A*). Interestingly *SP-A* is a good marker for clearance of *S. aureus* since it is involved in binding to the staphylococcal adhesion extracellular adherence protein as well to the macrophage receptors *SP-A* receptor 210 and scavenger receptor class A, enhancing phagocytosis [37]. There are over 13 homologs of this gene clustered in a region on chromosome 1 that are extremely similar but not all inducible by *S. epidermidis* infection. Therefore we aim to further investigate specificity of induction of these genes by microbial infection in follow up studies. We also performed a RNA sequencing experiment of the *M. marinum* infection system at 5 dpi for verification of the micro-array data. These data show that all exons of *mfap4* (4 of 13), *atf3* and *lect2l* tested are significantly expressed at 5 dpi by *S. epidermidis*, whereas after *M. marinum* infection there is no significant expression of these exons. With *mmp9* as a positive control, the expression of all exons is significantly expressed after *M. marinum* infection (Figure 6). The expression levels for *mfap4* (4 of 13) are based on manual annotation of the unique reads in the known gene region since there are several repetitive DNA regions in common with the other 12 annotated *mfap4* gene family members (Figure 6). Here we clearly benefit from the power of RNA sequencing that can overcome the problems of micro-array probe annotation for complex gene families. Therefore using our unbiased approach we were able to confirm known immune genes as markers for staphylococci infection but also identify novel markers as good candidates for specific response markers of *S. epidermidis* in our infection model that we will study further in functional analysis in the near future.

Conclusions

Microscopic imaging showed that *S. epidermidis* when injected into the yolk or caudal vein proved to be far less virulent than *S. aureus*. Under the same conditions of yolk injection *S. aureus* immediately invades the body of embryo causing 100% mortality within 3 dpi. In contrast, during the five day time period analysed, *S. epidermidis* proliferates efficiently in the entire body of the infected embryos providing an excellent

system for analysis of factors that influence bacterial proliferation and virulence. Based on this advantage, we have developed a versatile high-throughput analysis system for bacterial proliferation that is much less time consuming than cfu count determinations, and which allows repeated measurements of the same embryos over time. COPAS analysis proved to be accurate to determine the bacterial burden inside embryos at high-throughput. With addition of sorting zebrafish embryos into multi well plates, for automated confocal laser scanning microscopy, a medium-throughput, high-resolution screenings system can be added. We therefore have extended the high-throughput infection methods developed by Carvalho et al. [12] to a quantitative level and showed the applicability for the analysis of proliferation of opportunistic pathogens such as *S. epidermidis*. Our over time transcriptome analysis results correlate very well with the infection pattern of *S. epidermidis*. The bacteria will grow for the first 2 to 3 days inside the yolk of the embryos, while from 3 and 4 dpi *S. epidermidis* invade the body of the embryo, at which stage a strong response of many immune related genes occurs. We have compared transcriptome response in the same system using *M. marinum*. These comparisons show that *M. marinum* has a far stronger effect on host gene regulation than *S. epidermidis*. However, some genes were identified that specifically responded to *S. epidermidis* and not to *M. marinum* infection including a cell adhesion gene (*mfap4*, ENSEMBL 4 of 13) that can be linked to specific infection by staphylococci in mammals. Vuong et al. [38] and Otto et al. [26] already reported that *S. epidermidis* itself does not seem to have particular specific virulence factors. All known putative virulence factors have origins in the commensal lifestyle of this species. However, the large difference between the outcome of injection of bacteria into the yolk or caudal vein could have been caused by an effect of prolonged growth of the bacteria in the host organism resulting in a higher virulence when the bacteria are release in other tissues. In our future research we aim to use our identified host marker genes to identify new bacterial traits involved in proliferation in host tissues and the factors that determine their expression during time with emphasis on the time points when bacteria get in contact with immune cells. We are particularly interested in the effect of biomaterials on possible virulence factors that make virulence deviate from the commensal life style. This can help to understand which host mechanisms and genes are involved during biomaterial-associated infections.

Material and methods

Bacterial strains and growth conditions

S. epidermidis strain O-47, containing the GFP expression vector pWVW189 or a derived mCherry expression vector (De Boer L. unpublished) and *S. aureus* strain RN4220 pWVW189 (De Boer L. unpublished) from LB (Luria Bertani) agar plates were cultured overnight at 37°C in 25 ml LB medium supplemented with 10 μg/ml chloramphenicol to mid-log stage. *M. marinum* strain E11 was grown as described in Carvalho et al. [12]. Two reaction vials with 1 ml of the culture were centrifuged at 14680 rpm for 1 min.

The pellets were combined and washed three times with 1 ml PBS. Suspensions were prepared based on the OD_{600} and by plating and cfu determination. The inocula were suspended in 2% polyvinylpyrrolidone$_{40}$ (PVP$_{40}$, CalBiochem) to $5.0×10^6$, $1.0×10^8$ cfu/ml.

Zebrafish husbandry

Zebrafish were handled in compliance with animal welfare regulations and maintained according to standard protocols (http://ZFIN.org). Embryos were grown at 28°C in egg water (60 µg/ml Instant ocean sea salt, Sera Marin). The egg water was refreshed every day.

Experimental design of infection study

Infection experiments were performed using mixed egg clutches from wild type AB×TL or *Tg(UAS:KAEDE/MPEG1:GAL4)* strain zebrafish [15]. Embryos were staged at 2 hpf by morphological criteria, and 20 cfu of mCherry or GFP expressing *S. epidermidis* O-47 bacteria suspended in 2% PVP$_{40}$ were injected into the yolk. As a control an equal volume of 2% PVP$_{40}$ was likewise injected. Manual injections were controlled using a Leica M50 stereomicroscope together with a FemtoJet micro-injector (Eppendorf) and a micromanipulator with pulled micro capillary needles. Automated micro-injections were performed as described in Carvalho et al. [12].

4

Microscopy

A Leica fluorescence (MZ 16 FA) stereo microscope and Leica TCS SPE confocal microscope were used to take images of zebrafish embryos. Embryos were kept under anaesthesia (0.02% buffered 3- aminobenzoic acid ethyl ester (Tricaine, Sigma) in egg water) during imaging.

COPAS analysis

Zebrafish embryos were measured alive every 24 hours until 5 dpi with the COPAS XL using the setting as described below. Photo multiplier tubes (PMT) voltage: 650 V for green/red and 0 V for yellow. Optical density threshold signal was set to 975 mV (COPAS value: 50) and the time of flight (TOF) minimum to 320 µs (COPAS value: 800) in order to reduce the influence of debris.

Cfu count

Injected embryos were collected into a 2 ml reaction vial with a sterile 5 mm stainless steel bead and PBS. The reaction vials were vigorously shaken for 30 seconds at 30 revolutions per second in a shaker (Retsch MM301). All suspensions were diluted, plated in duplicate on LB agar supplemented with 10 µg/ml chloramphenicol, and incubated overnight at 37°C. The following day, colonies were counted using a fluorescence stereo microscope (Leica MZ125).

Micro-array

Seven parent zebrafish couples kept separately from one another for mating the following week, to perform an identical experiment for a biological replicate. Injections were performed from the 16 cell stage onwards at approximately 2 hpf: the first group was injected with 1 nl 2% PVP_{40} solution containing 20 cfu/nl *S. epidermidis* O-47 pWVW189, the second group with 1 nl 2% PVP_{40} solution without bacteria, the third group only received a needle puncture in the yolk, and the last group was a non-injected control group. Groups consisted of approximately 150 embryos. The *S. epidermidis* O-47 pWVW189-injected group and the non-treated embryos were measured at 2, 3, 4 and 5 dpi with the COPAS XL just before snap freezing. At 6 hpi, 1, 2, 3, 4 and 5 dpi, 20 embryos were collected randomly from each group, snap frozen in liquid nitrogen, and stored at −80°C. Embryos were homogenized in 0.5 ml of TRIzol reagent (Invitrogen), and total RNA was extracted according to the manufacturer's instructions. RNA samples were treated with DNaseI, (Ambion) to remove residual genomic DNA. RNA integrity was analysed by Lab-on-a-chip analysis (Agilent). The average RIN value of the RNA samples was 8.1 with a minimum of 6.7. Per sample, 500 ng total RNA was combined with Spike A and amplified according to the Agilent Two-Color Micro-array-Based Gene Expression Analysis guide version 5.5 (G4140-90050, Agilent technologies). For the common reference an equimolar pool of all test samples was made and 500 ng samples were amplified similarly as the test samples with the exception that Spike B was used. Amino-allyl modified nucleotides were incorporated during the aRNA synthesis (2.5 mM rGAC (GE Healthcare), 0.75 mM rUTP (GE Healthcare), 0.75 mM AA-rUTP (TriLink Biotechnologies). Synthesized aRNA was purified with the E.Z.N.A. MicroElute RNA Clean Up Kit (Omega Bio-Tek). The quality was inspected on the BioAnalyzer (Agilent Technologies) with the Agilent RNA 6000 kit (5067–1511, Agilent Technologies). Test samples were labelled with Cy3 and the Reference sample was labelled with Cy5. For mycobacterium infected embryos a dye swap technical duplicate was performed in which the control was either labelled with Cy3 or Cy5. The overlap of the technical duplicates was used for the output files. Five µg of aRNA was dried down and dissolved in 50 mM carbonate buffer pH 8.5. Individual vials of Cy3/Cy5 from the monoreactive dye packs (GE Healthcare) were dissolved in 200 µl DMSO. To each sample, 10 µl of the appropriate CyDye dissolved in DMSO was added and the mixture was incubated for 1 h. Reactions were quenched with the addition of 5 µl 4 M hydroxylamine (Sigma-Aldrich). The labelled aRNA was purified with the E.Z.N.A. MicroElute RNA Clean Up Kit. Yields of aRNA and CyDye incorporation were measured on the NanoDrop ND-1000. Each hybridization mixture was made up from 825 ng Test (Cy3) and 825 ng Reference (Cy5) material. Hybridization mixtures were made as described in the Agilent Two-Color Micro-array-Based Gene Expression Analysis guide version 5.5 (G4140-90050, Agilent technologies). The samples were loaded onto 4x180k *D. rerio* micro-arrays (Design ID:028233, Agilent Technologies) and hybridized for 17 hours at 65°C. Afterwards the slides were washed and scanned (20 bit, 3 µm resolution) in an ozone-free room with the Agilent G2505C scanner as described in the Agilent Two-Color Micro-array-Based

Gene Expression Analysis guide version 5.5 (G4140-90050, Agilent technologies). Data was extracted with Feature Extraction (v10.7.3.1, Agilent Technologies) with the GE2_107_Sep09 protocol for two-color Agilent micro-arrays. Micro-array data was processed using Rosetta Resolver 7.2 (Rosetta Biosoftware). *S. epidermidis* infection groups were compared to the PVP, needle puncture and non-injected control groups using the Rosetta common reference re-ratio experiment pipeline. Significance cut off for the ratios of *S. epidermidis* versus PVP, *S. epidermidis* versus needle puncture and *S. epidermidis* versus non-injected were set at 2 fold change at *P*-value smaller than 10^{-8}. Pathway analysis was performed using the Pathvisio software package (www.pathvisio.org) [39] with the same significance cut off. The raw micro-array data have been deposited in the NCBI GEO database under accession number GSE42847 and GSE44352. DAVID bioinformatics resources 6.7 [40, 41] was used for gene ontology analysis.

RNA deep sequencing

Validation of micro-array data was performed by RNAseq analysis. Ten parent zebrafish couples were kept separately from one another for mating the following week, to perform an identical experiment for 4 biological replicates. Injections were performed at approximately 2 hpf using the automated micro-injection system. At 5 dpi, embryos were collected from the 2 hpf injected and non-injected group, snap frozen in liquid nitrogen, and stored at −80°C for RNA isolation. Twenty cfu of *S. epidermidis* O-47 pWVW189 were injected to obtain 150 embryos per sample. For *M. marinum* infected embryos approximately 1000 embryos were used with 30 cfu injected per embryo. Embryos were homogenized in 1 ml of TRIzol reagent (Invitrogen), and total RNA was extracted according to the manufacturer's instructions. RNA samples were treated with DNaseI, (Ambion) to remove residual genomic DNA. RNA integrity was analysed by Lab-on-a-chip analysis (Agilent). The average RIN value of the RNA samples was 9.7 with a minimum of 9.5. A total of 3 μg of RNA was used to make RNA-Seq libraries using the Illumina TruSeq RNA Sample Preparation Kit v2 (Illumina Inc., San Diego, USA). In the manufacturer's instructions two modifications were made. In the adapter ligation step 1 μl instead of 2.5 μl adaptor was used. In the library size selection step the library fragments were isolated with a double Ampure XP purification with a 0.7x beads to library ration. The resulting mRNA-Seq library was sequenced using an Illumina HiSeq2000 instrument according to the manufacturer's description with a read length of 2 × 50 nucleotides. Image analysis and base calling was done by the Illumina HCS version 1.15.1. Sequence reads were quality trimmed using the quality_trim module in the CLCbio Assembly Cell v4.0.6. Filtered reads were mapped to ENSEMBL transcripts (Zv9_63) using the ref_ assemle_short module in the CLCbio Assembly Cell v4.0.6. Accumulation of transcripts to ENSEMBL genes was done by first converting the mapping files to a table with the assembly_table module in the CLCbio Assembly Cell v4.0.6. Secondly, a custom script was used that sums all reads belonging to the same gene. Non-uniquely mapped reads were divided between genes according to their ratio of uniquely mapped reads. Finally,

read counts of transcripts belonging to the same gene were summed to obtain count data at ENSEMBL gene level. Fold-change and differential expression significance values were calculated from gene level read counts using the DESeq package version 1.8.3) available in Bioconductor (version 2.10). DESeq utilizes a negative binomial distribution for modelling read counts [42]. Secondly reads were counted per exon with a python script (Lodder R. unpublished). Sorted sam files were obtained from the raw fastq files through Bowtie2 [43] and samtools [44]. The raw RNAseq data have been deposited in the NCBI GEO database under accession number GSE42847 and GSE44352.

Acknowledgements

Supplementary material

Supplementary material can be found on:

http://www.biomedcentral.com/1471-2164/14/255

References

1.   Boelens JJ, Dankert J, Murk JL, Weening JJ, Poll Tvd, Dingemans KP, Koole L, Laman JD, Zaat SAJ: Biomaterial-Associated Persistence of Staphylococcus epidermidis in Pericatheter Macrophages. The Journal of Infectious Diseases 2000, 181(4):1337-1349.

2.   Broekhuizen CA, Schultz MJ, van der Wal AC, Boszhard L, de Boer L, Vandenbroucke-Grauls CM, Zaat SA: Tissue around catheters is a niche for bacteria associated with medical device infection. Crit Care Med 2008, 36(8):2395-2402.

3.   Busscher HJ, van der Mei HC, Subbiahdoss G, Jutte PC, van den Dungen JJ, Zaat SA, Schultz MJ, Grainger DW: Biomaterial-associated infection: locating the finish line in the race for the surface. Sci Transl Med 2012, 4(153):153rv110.

4.   Zaat S, Broekhuizen C, Riool M: Host tissue as a niche for biomaterial-associated infection. Future Microbiology 2010, 5(8):1149-1151.

5.   Zaat SAJ, Broekhuizen CAN, Boer Ld, Schultz MJ, Boszhard L, Wal ACvd, C..M.J.E.Vandenbroucke-Grauls: Biomaterial-associated infection breaking out of the biofilm. European Cells and Materials 2008, 16:10.

6.   Zimmerli W, Sendi P: Pathogenesis of implant-associated infection: the role of the host. Seminars in Immunopathology 2011, 33(3):295-306.

7.   Anderson JM, Rodriguez A, Chang DT: Foreign body reaction to biomaterials. Seminars in Immunology 2008, 20(2):86-100.

8.   Rupp ME, Ulphani JS, Fey PD, Bartscht K, Mack D: Characterization of the importance of polysaccharide intercellular adhesin hemagglutinin of Staphylococcus epidermidis in the pathogenesis of biomaterial-based infection in a mouse foreign body infection model. Infection and Immunity 1999, 67(5):2627-2632.

9.   Rupp ME, Fey PD, Heilmann C, Gotz F: Characterization of the Importance of Staphylococcus epidermidis Autolysin and Polysaccharide Intercellular Adhesin in the Pathogenesis of Intravascular Catheter-Associated Infection in a Rat Mode. The Journal of Infectious Diseases 2001, 183(7):1038-1042.

10.   Broekhuizen CAN, de Boer L, Schipper K, Jones CD, Quadir S, Feldman RG, Vandenbroucke-Grauls CMJE, Zaat SAJ: The influence of antibodies on Staphylococcus epidermidis adherence to polyvinylpyrrolidone-coated silicone elastomer in experimental biomaterial-associated infection in mice. Biomaterials 2009, 30(32):6444-6450.

11.   Brittijn SA, Duivesteijn SJ, Belmamoune M, Bertens LFM, Bitter W, Debruijn JD, Champagne DL, Cuppen E, Flik G, Vandenbroucke-Grauls CM et al: Zebrafish development and regeneration: new tools for biomedical research. The International Journal of Developmental Biology 2009, 53(5-6):835-850.

12.   Carvalho R, de Sonneville J, Stockhammer OW, Savage ND, Veneman WJ, Ottenhoff TH, Dirks RP, Meijer AH, Spaink HP: A high-throughput screen for tuberculosis progression. PLoS One 2011, 6(2):e16779.

13.   Cui C, Benard EL, Kanwal Z, Stockhammer OW, Vaart Mvd, Zakrzewska A, Spaink HP, Meijer AH: Infectious disease modeling and innate immune function in zebrafish embryos. Methods in cell biology 2011, 105:273-308.

4

14. Renshaw SA, Loynes CA, Trushell DMI, Elworthy S, Ingham PW, Whyte MKB: A transgenic zebrafish model of neutrophilic inflammation. Blood 2006, 108(13):3976-3978.

15. Ellett F, Pase L, Hayman JW, Andrianopoulos A, Lieschke GJ: mpeg1 promoter transgenes direct macrophage-lineage expression in zebrafish. Blood 2010, 117(4):e49-e56.

16. Meijer AH, Spaink HP: Host-pathogen interactions made transparent with the zebrafish model. Curr Drug Targets 2011, 12(7):1000-1017.

17. Davis JM, Clay H, Lewis JL, Ghori N, Herbomel P, Ramakrishnan L: Real-time visualization of mycobacterium-macrophage interactions leading to initiation of granuloma formation in zebrafish embryos. Immunity 2002, 17(6):693-702.

18. van der Sar AM, Musters RJP, van Eeden FJM, Appelmelk BJ, Vandenbroucke-Grauls CMJE, Bitter W: Zebrafish embryos as a model host for the real time analysis of Salmonella typhimurium infections. Cellular Microbiology 2003, 5(9):601-611.

19. van Soest JJ, Stockhammer OW, Ordas A, Bloemberg GV, Spaink HP, Meijer AH: Comparison of static immersion and intravenous injection systems for exposure of zebrafish embryos to the natural pathogen Edwardsiella tarda. BMC Immunol 2011, 12:58.

20. Prajsnar TK, Cunliffe VT, Foster SJ, Renshaw SA: A novel vertebrate model of Staphylococcus aureus infection reveals phagocyte-dependent resistance of zebrafish to non-host specialized pathogens. Cellular Microbiology 2008, 10(11):2312-2325.

21. Vergunst AC, Meijer AH, Renshaw SA, O'Callaghan D: Burkholderia cenocepacia Creates an Intramacrophage Replication Niche in Zebrafish Embryos, Followed by Bacterial Dissemination and Establishment of Systemic Infection. Infection and Immunity 2010, 78(4):1495-1508.

22. Stockhammer OW, Rauwerda H, Wittink FR, Breit TM, Meijer AH, Spaink HP: Transcriptome analysis of Traf6 function in the innate immune response of zebrafish embryos. Molecular Immunology 2010, 48(1-3):179-190.

23. van der Sar AM, Spaink HP, Zakrzewska A, Bitter W, Meijer AH: Specificity of the zebrafish host transcriptome response to acute and chronic mycobacterial infection and the role of innate and adaptive immune components. Molecular Immunology 2009, 46(11-12):2317-2332.

24. Ordas A, Hegedus Z, Henkel CV, Stockhammer OW, Butler D, Jansen HJ, Racz P, Mink M, Spaink HP, Meijer AH: Deep sequencing of the innate immune transcriptomic response of zebrafish embryos to Salmonella infection. Fish Shellfish Immunol 2011, 31(5):716-724.

4

25.   Tong SYC, Chen LF, Fowler VG: Colonization, pathogenicity, host susceptibility, and therapeutics for Staphylococcus aureus: what is the clinical relevance? Seminars in Immunopathology 2011, 34:185-200.

26.   Otto M: Molecular basis of Staphylococcus epidermidis infections. Seminars in Immunopathology 2011, 34(2):201-214.

27.   Mermel LA, Farr BM, Sherertz RJ, Raad II, O'Grady N, Harris JS, Crave DE: Guidelines for the Management of Intravascular Catheter-Related Infections. Clinical Infectious Diseases 2000, 32(9):1249-1272.

28.   Al-Solaiman Y, Estrada E, Allon M: The Spectrum of Infections in Catheter-Dependent Hemodialysis Patients. Clinical Journal of the American Society of Nephrology 2011, 6(9):2247-2252.

29.   Kazmierczak MJ, Wiedmann M, Boor KJ: Alternative Sigma Factors and Their Roles in Bacterial Virulence. Microbiology and Molecular Biology Reviews 2005, 69(4):527-543.

30.   Pulak R: Techniques for analysis sorting and dispensing of C. elegans on the COPAS flow-sorting system. Methods in molecular biology 2006, 351:275-286.

31.   van der Sar AM, Stockhammer OW, van der Laan C, Spaink HP, Bitter W, Meijer AH: MyD88 Innate Immune Function in a Zebrafish Embryo Infection Model. Infection and Immunity 2006, 74(4):2436-2441.

32.   Zakrzewska A, Cui C, Stockhammer OW, Benard EL, Spaink HP, Meijer AH: Macrophage-specific gene functions in Spi1-directed innate immunity. Blood 2010, 116(3):e1-e11.

33.   Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nature Methods 2008, 5(7):621-628.

34.   Haas BJ, Chin M, Nusbaum C, Birren BW, Livny J: How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes? BMC Genomics 2012, 13(1):734.

35.   Kogenaru S, Qing Y, Guo Y, Wang N: RNA-seq and microarray complement each other in transcriptome profiling. BMC Genomics 2012, 13:629.

36.   Schlosser A, Thomsen T, Shipley JM, Hein PW, Brasch F, Tornøe I, Nielsen O, Skjødt K, Palaniyar N, Steinhilber W et al: Microfibril-associated Protein 4 Binds to Surfactant Protein A (SP-A) and Colocalizes with SP-A in the Extracellular Matrix of the Lung. Scandinavian Journal of Immunology 2006, 64(2):104-116.

4

37. Sever-Chroneos Z, Krupa A, Davis J, Hasan M, Yang CH, Szeliga J, Herrmann M, Hussain M, Geisbrecht BV, Kobzik L et al: Surfactant Protein A (SP-A)-mediated Clearance of Staphylococcus aureus Involves Binding of SP-A to the Staphylococcal Adhesin Eap and the Macrophage Receptors SP-A Receptor 210 and Scavenger Receptor Class A. Journal of Biological Chemistry 2010, 286(6):4854-4870.

38. Vuong C, Otto M: Staphylococcus epidermidis infections. Microbes and Infection 2002, 4(4):481-489.

39. Pico AR, Kelder T, Iersel MPv, Hanspers K, Conklin BR, Evelo C: WikiPathways: Pathway Editing for the People. PLoS Biology 2008, 6(7):e184.

40. Huang DW, Sherman BT, Lempicki RA: Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009, 37(1):1-13.

41. Huang DW, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009, 4(1):44-57.

42. Anders S, Huber W: Differential expression analysis for sequence count data. Genome Biology 2010, 11(10):R106.

43. Langmead B, Salzberg SL: Fast gapped-read alignment with Bowtie 2. Nature Methods 2012, 9(4):357-359.

44. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R: The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009, 25(16):2078-2079.

4

Chapter 5

Analysis of RNAseq datasets from a comparative infectious disease zebrafish model using GeneTiles bioinformatics

Wouter J. Veneman[1#], Jan de Sonneville[2#], Kees-Jan van der Kolk[2,3#], Anita Ordas[1], Zaid Al-Ars[3], Robert-Jan Raterink[4], Koen Egberts[4], Thomas Hankemeijer[4], Annemarie H. Meijer[1], Herman P. Spaink[1]

[1] Institute of Biology, Leiden University
[2] Life Science Methods BV
[3] Computer Engineering, Delft University
[4] Leiden Academic Centre for Drug Research, Leiden University

[#]Authors contributed equally

5

Abstract

We present a RNA deep sequencing (RNAseq) analysis of a comparison of the transcriptome responses to infection of zebrafish larvae with *Staphylococcus epidermidis* and *Mycobacterium marinum* bacteria. We show how our developed GeneTiles software can improve RNAseq analysis approaches by more confidently identifying a large set of markers upon infection with these bacteria. For analysis of RNAseq data currently software programs such as Bowtie2 and Samtools are indispensable. However, these programs that are designed for a LINUX environment require some dedicated programming skills and have no options for visualization of the resulting mapped sequence reads. Especially with large data sets this makes the analysis time consuming and difficult for non-expert users. We have applied the GeneTiles software to the analysis of previously published and newly obtained RNAseq datasets of our zebrafish infection model and we have shown the applicability of this approach also to published RNAseq datasets of other organisms by comparing our data with a published mammalian infection study. In addition we have implemented the DEXSeq module in the GeneTiles software to identify genes, such as glucagon A, that are differentially spliced under infection conditions. In the analysis of our RNAseq data this has led to the possibility to improve the size of data sets that could be efficiently compared without using problem-dedicated programs, leading to a quick identification of marker sets. This approach will therefore also be highly useful for transcriptome analyses of other organisms for which well-characterized genomes are available. Based on the transcriptome analysis we have analysed the function of the leptin b gene that was identified as the most highly induced gene after infection. Using metabolomics analyses we have shown that infection with *M. marinum* results in a rapid wasting syndrome. Results of gene knock-down analysis suggest that the leptin b gene is essential for progression of the wasting syndrome after bacterial infection.

Introduction

In our previous research we have used zebrafish larval infection models to study the transcriptome response to infection by several pathogens [1-5]. In addition we have tested the response of zebrafish larvae to infection by the opportunistic bacterium *Staphylococcus epidermidis* as a model for biomaterial-associated infections that are often caused by this species in clinical practice [6-9]. These studies have led to a high-throughput model that resulted in a large set of RNAseq data sets highlighting a new bottleneck in our research: the fast and user-friendly analysis of large datasets that can be easily visualized for comparative purposes.

In the analysis of our former transcriptome data sets there was a need for specialized scripting languages to quickly find good marker genes for disease. We used an existing visualization program, Integrative Genomics Viewer (IGV) [10] that shows the data

solely along a line representing the genome, thereby requiring zooming in to view the aligned reads. IGV, and many other open source visualization programs such as MapView [11], Tablet [12], GenoViewer [13] and BamView [14] also require the user to scroll or manually search for other genes to bring these into focus and an overview of a selection of genes based on alignment results is not available. Finally, most of these data visualization programs do not allow for export of presented visual results, other than taking screenshots. In our previous analysis of RNAseq data we were reliant on manual counting of reads as guided by the Integrative GenomeViewer [1].

In this paper we used the transcriptome data set obtained from the zebrafish high-throughput screening system for *S. epidermidis* infection [1] as a case study to optimize and automate the data analysis pipeline. Using the resulting software package we also went further and added a larger RNAseq data set from *Mycobacterium marinum* infection data for comparisons of specificity of the transcriptome responses. We also integrated the DEXSeq algorithm that can be used to give an estimate of probability of the occurrence of differential splicing. This has led to the identification of genes that are differentially spliced after microbial infection in zebrafish larvae. Finally, we wanted to include a comparison with whole organism infection data in other vertebrate species. Unfortunately, there is still few RNAseq data for this available that can be mapped on ENSEMBL genome data and as a result we have only been able to compare our zebrafish infection data with RNAseq data from a bovine digital dermatitis (BDD) model as published by Scholey et al. 2013 [15]. However, the results are sufficient to show that our approach makes also such interspecies comparisons of RNAseq datasets very easy and can quickly lead to conclusions on conserved immune responses, even in comparisons between very different fish and mammalian infection models. One of the genes conserved between fish and mammals and highly induced by infection encodes the hormone leptin b and we selected this gene for functional analysis in this study.

Material and methods

Bacterial strains and growth conditions

The *S. epidermidis* strain O-47 containing a pWVW189 derived mCherry expression vector (De Boer L. unpublished) was grown as described in Veneman et al. (2013). The *M. marinum* strain E11 was grown as described in Carvalho et al. 2011 [16]. Two reaction vials with 1 ml of the culture were centrifuged at 14680 rpm for 1 min. The pellets were combined and washed three times with 1 ml phosphate-buffered saline (PBS). Suspensions were prepared based on the optical density at 600 nm and by plating and cfu determination. The inoculates were suspended in 2% polyvinylpyrrolidone$_{40}$ (PVP$_{40}$, CalBiochem) to 2.0×10^7 or 3.0×10^7 cfu/ml.

Morpholino knock-down experiments

For morpholino knock-down experiments, morpholino oligonucleotides (Gene Tools) were diluted to 0.5mM in nuclease free water. To block leptin b we used an ATG morpholino (5'- TTTTTTGCTTGTTAATATCATCCCT-3'). Embryos injected with leptin b morphlino were staged between the 16 and 128 cell stage and injected with 50 cfu of mCherry expressing *M. marinum* E11 bacteria suspended in 2% PVP_{40}. Embryos were screened at 3, 4 and 5 dpi for bacterial proliferation using the COPAS XL (Union BioMetrica) flow cytometer as described in Veneman et al. (2013).

Metabolic profiling

Samples for mass spectrometry were taken at 4 and 5 dpi. Eight embryos per group were collected into a 1.5 ml reaction vial. All egg water was removed and the embryos were washed 2 times with 200 µl Millipore water. Per sample 80 µl of stock C (table 1) was added and were snap frozen in liquid nitrogen for 30 minutes. Samples were

Table 1		MeOH	I.S. Carnithine stock (table 2)	I.S. Amine stock (table 3)	Glucose-d7 5.5mM
	Stock C	126.4 µl	16 µl	16 µl	1.6 µl

Table 2 (Carnithine stock)				Step 1: stock 7.3.1.2		Step 2: Spike solution 7.3.1.3		
		Internal STD's (1mg/ml)	µl	µg/ml	µl	µg/ml	µM
	1	Carnitine-d3 HCL			20	2.00	9.996
	2	Deoxycarnitine-d9 HCL			10	2.00	10.487
	3	Choline-d4 HCL			10	2.00	13.923
	4	Betaine-d3 HCL			20	2.00	12.769
	5	Acetyl-L-carnitine-d3 HCL			10	2.00	8.240
	6	Butyryl-L-carni	10	12.5		0.025	0.092
	7	Octanoyl-L-carnitine-d3 HCL	10	12.5		0.025	0.076
	8	Octadecanoyl-L-carnitine-d3 HCL	10	12.5		0.025	0.054
		Stock solution 7.3.1.2			20		
		MeOH	770		9910		
		Total volume	800		10000		

Table 3 (Amine stock)	Internal Standard	Conc. (mg/ml)	Solvent
	L-Ornithine D6	1	H_2O
	Beta-Alanine D4	1	H_2O
	L-2-aminobutyric acid d6	1	H_2O
	2-(4-hydroxy-3methoxyphenyl)ethyl-1,1,2,2-D4-amine	1	H_2O
	L-NT-Methyl-D3-L-Histidine	1	H_2O
	L-3-(4-Hydroxy-3methoxy-D3-phenyl)alanine	1	H_2O
	U-C13n15 cell free AA mix	0.875	H_2O
	Ethanolamine-d7	1	H_2O
	DL-Homocysteine-d4	1	H_2O
	1.4-Butane-d8-diamine (Putrescine)	1	H_2O

Ratio: 1 ml labeled amino acid mix: 40 µl all other amine standards, in a total volume of 10 ml in MeOH.

sonicated for 2 minutes at 40 kHz continuously changing positions in order for local sweet spots in the sonicator. The freeze-sonicate cycle was repeated 3 times. The lysate was vortexed for 5 seconds after which they were centrifuged for 5 minutes at 4 °C at 16.1 krcf. Per sample 75 µl supernatant was collected and 8 µl supernatant was collected for quality control. The samples were evaporated using speedvac for 1 hour until they were completely dry. All samples were stored at -80 °C. Samples were defrosted in 30 minutes and per sample 40 µl borate buffer was added and vortexed. Per sample 10 µl AccQ-Tag (3 mg/ml) (Waters BV) reagent was added for derivatization. The samples were incubated at 55 °C for 10 minutes and were quickly spun down for 2 seconds after which they were diluted 4 times using H_2O.

Mass spectrometry data processing

After Liquid Chromatography mass spectrometry (LC-MC) data acquisition, the data was processed in order to extract the proper target and internal standards (I.S.) peak area as follows. First, an Xcalibur (Thermo Fischer) Processing setup was generated in which all the targets and I.S. are identified (in terms of retention time and exact mass) from one of the QC samples and validated with the system suitability test (SST). Per batch the peak separation and retention time was checked and integrated into the processing setup. When every metabolite was correctly assigned and identified in the processing setup, the whole batch was analysed in the Xcalibur Quan browser. All peaks were integrated and manually checked in the Xcalibur Quan browser. When peaks were bad (i.e. bad peak shapes due to e.g. low concentration), the target was discarded from the study. From all the remaining targets the relative standard deviation (RSD) was calculated within all the QCs. Metabolites with RSDs above 20% were discarded from the study. On the remaining high-quality, normalized metabolite data set, univariate and multivariate analysis was performed. For the multivariate analysis, a principle component analysis (PCA) was used as a data exploratory tool [17]. For the univariate analysis, t-tests or ANOVAs were used.

Zebrafish husbandry

Zebrafish were handled in compliance with animal welfare regulations and maintained according to standard protocols (http://ZFIN.org). Embryos were grown at 28°C in egg water (60 µg/ml Instant ocean sea salt, Sera Marin). The egg water was refreshed every day.

Experimental outline

Infection experiments were performed with mixed egg clutches from wild type ABxTL strain zebrafish. Embryos were staged at 2 hours post fertilization by morphological criteria, and 20 cfu of mCherry expressing *S. epidermidis* O-47 or 30 cfu of mCherry expressing *M. marinum* E11 bacteria suspended in 2% PVP_{40} were injected into the yolk. Automated micro-injections were performed as described in Carvalho et al. [16].

At 5 days post fertilization, embryos (N ~100) were collected from the 2 hours post fertilization injected and non-injected group, snap frozen in liquid nitrogen, and stored at −80°C for RNA isolation.

RNA deep sequencing

RNA isolation was performed as described in Veneman et al. (2013). A total of 3 μg of RNA was used to make RNAseq libraries using the Illumina TruSeq RNA Sample Preparation Kit v2 (Illumina Inc., San Diego, USA). In the manufacturer's instructions two modifications were made. In the adapter ligation step 1 μl instead of 2.5 μl adaptor was used. In the library size selection step the library fragments were isolated with a double Ampure XP purification with a 0.7x beads to library ration. The resulting mRNAseq library was sequenced using an Illumina HiSeq2000 instrument according to the manufacturer's description with a read length of 2 × 50 nucleotides. Image analysis and base calling was done by the Illumina HCS version 1.15.1. The raw RNAseq data have been deposited in the NCBI GEO database with accession number GSE42846, GSE44351 and GSE57792.

Data analysis

GeneTiles was used for quantification and visualization of the RNAseq data. When using GeneTiles, the complete data processing pipeline, including the used parameters is available for download. This enables the user to perform the same analysis locally, or try small modifications (for bioinformaticians). Here we give a quick list of the used programs and their function within our current pipeline. A detailed explanation including the used parameters is available in Supplementary file 2.

In order of use in GeneTiles:

1. Bowtie2 [18] is used to align the reads in the fastq file to the genome (obtained from Ensembl). Bowtie2 generates SAM files that contain the reads together with the location on the genome. Upon multiple hits, the best quality hit is selected, or upon a tie of multiple best hits, the reads are randomly distributed (the manual of Bowtie2 is referred to for other default behaviour).

2. Samtools [19] is used to convert and compress the SAM files into a binary BAM file.

3. Samtools is furthermore used to sort the reads in the BAM files based on the aligned read location in the genome, resulting in a sorted BAM file.

4. The BAM files is indexed to be able to quickly find the aligned reads based on a location in the genome, i.e. to be able to quickly search the BAM file. The index is saved as a BAI file.

5. Using the available annotation from Ensembl we can search the BAM file for reads

within a gene. All reads that at least partially fall within the gene exon and intron regions are counted once. This is done with a python script which consists of the combination of HTseq and pysam [20]. The output of this script is a tab separated file (tsv) containing the read counts per gene.

6. We used DESeq, an R-script to perform statistical analysis. DESeq is used to normalise the reads using a DESeq scaling factor, computed as the median of the ratio, for each gene, of its read count over the geometric mean across samples. Then variance and average of the measurement compared to the control is expressed as a *P*-value, by calculating the dispersion per gene using DESeq. The size factors as well as the *P*-values are stored in 'tsv' files.

7. Using scripts, similarly as in step 5, also the input files for DEXSeq can be generated. DEXseq requires:

• A 'gtf' file containing the experiment design.

• For all samples a 'txt' file containing the counts obtained for the mapping data in the .sam files.

• Genome annotation (from Ensembl) in a 'gff' file.

Scripts to obtain these files are also available in the supplement.

8. Using DEXSeq, another R-script to perform more complex statistical analysis, we can look at the reads within exons, and compare the variance and average per exon between measurement and control groups of samples. DEXSeq uses binning, where exons are cut into bins, based on known exon boundaries. When a read overlaps multiple bins, it is counted in each bin. Per bin, based on the annotation, two comparisons can be made, a comparison between the same exon bins in different samples (groups), and a comparison between an exon-bin and its neighbour exon-bins within the same group of samples. Note that therefore DEXSeq requires at least two groups containing at least two samples. Based on both comparisons a likelihood test is performed resulting in a *P*-value. More details are available in Anders et al. 2012 [21]. The output of size factors and *P*-values are stored as 'tsv' files.

9. Using a script, all tsv files are combined into an excel file, available for download, e.g. per experiment, chromosome, per filtered results of most significant reads or highest ratio between measurement and control. In addition an index is built for fast visualization online (closed source).

Through the website of wikipathways [http://wikipathways.org/index.php/Download_Pathways] the SVG images were downloaded on the GeneTiles server. Using Javascript, on the client side, within the SVG images the gene-boxes are given a background colour

based on a user selection, e.g. *P*-value or ratio. For this, the genes and or proteins are matched to their Ensembl references on the selected genome. In addition also the human pathways are searched for find homologs of genes using Ensembl biomart. Using these homologs, predictions of homolog pathways can be accessed, this enables to search a larger set of pathways contained in the human section of wikipathways. It should be noted that using human pathways to find information about zebrafish biology should be treated with caution and can only lead to suggestions for further investigation.

Results and Discussion

Software design

RNAseq data, containing tens of millions of reads, is mostly processed using scripts. After processing, a selection of reads is analysed using RNAseq viewers. Directly browsing processed RNAseq data is difficult due to the large dynamic range of length scales of reads (50 bp), exons (~200 bp), introns (~3 kb), genes (~20 kb), and chromosomes (~65 Mb). Using a minimal size of one pixel per read, a computer screen allows only for ~50 kb to be visible. In addition, most RNAseq viewers show introns at the same scale as exons, which in most experiments means that 90 percent of the visible sequence data does not display aligned reads. We created an online viewer, GeneTiles (www. Genetiles.com), that does allow for browsing all the aligned reads, while eliminating almost completely the need for user intervention (such as zooming in). The genes in a chromosome are visible as tiles in a 2D array. The tile colour and intensity are a measure of the significance of the number of reads of experiment versus control, indicating changes in expression levels. When a tile is selected, the gene is loaded underneath, scaled to fit the width of the screen. In a schematic view all introns are shrunk to a fixed short length to visualize the aligned reads in a graph above the exons. To accomplish fast browsing, all reads are indexed on the server directly after data processing. This indexed data is also available for download to apply custom filtering in Excel or other programs. The export functions of the tiles and genes as scalable vector graphics makes it easy for the user to modify the final visualization for publication.

Workflow

Automated analysis of RNAseq data using GeneTiles does not need any programming steps anymore in a Linux environment by the user and performs directly a visualization of the differential expressed data, making it easier to interpret. To validate this new software package, RNAseq data from zebrafish bacterial infection experiments was obtained from Veneman et al. (2013) and used as the initial test model. All programs used by Veneman et al. (2013) are implemented and more visualization and export options are added in a server based environment (Figure 1, Supplementary file 1). Therefore the analysis pipeline of GeneTiles represents a combination of previously
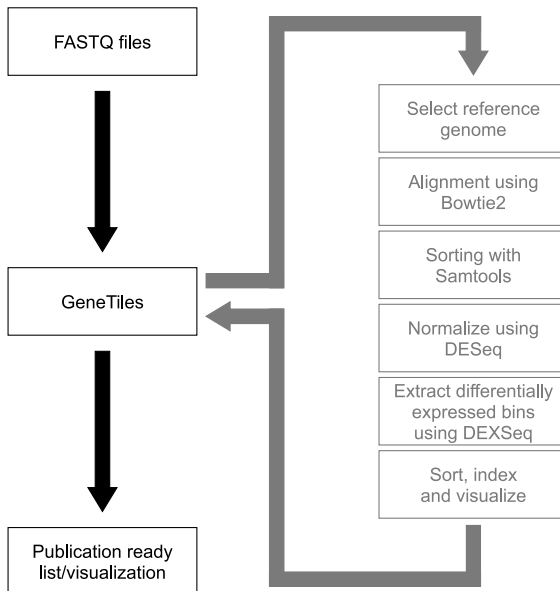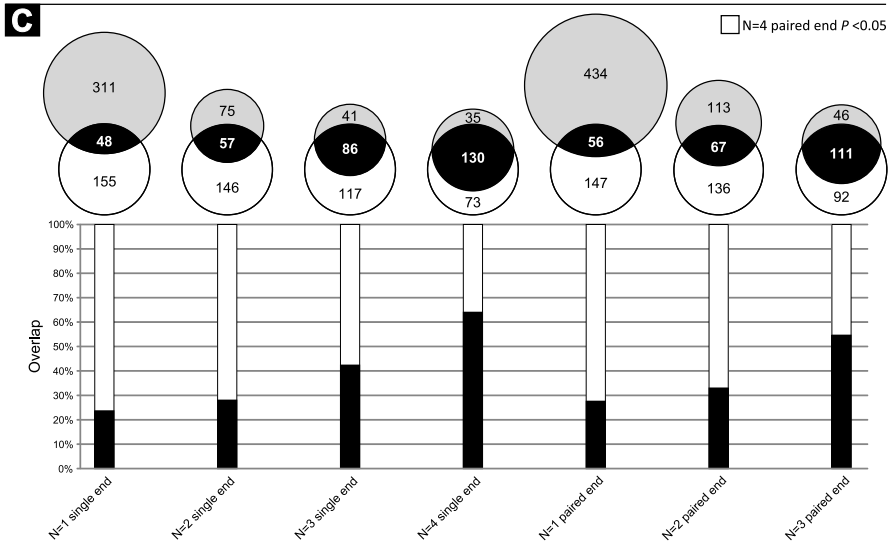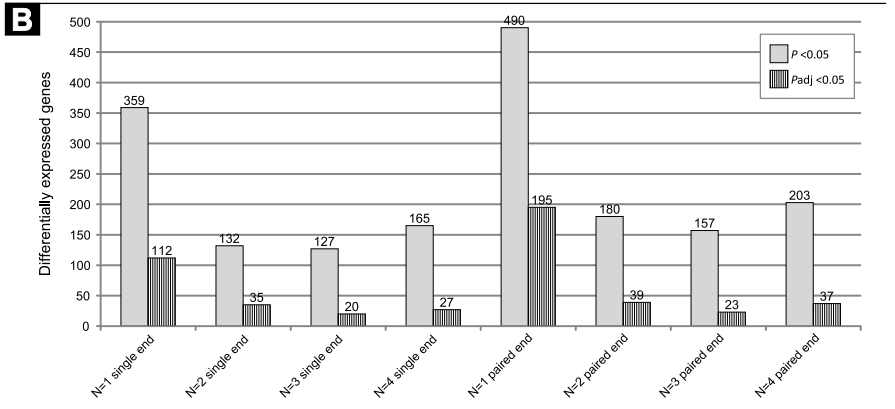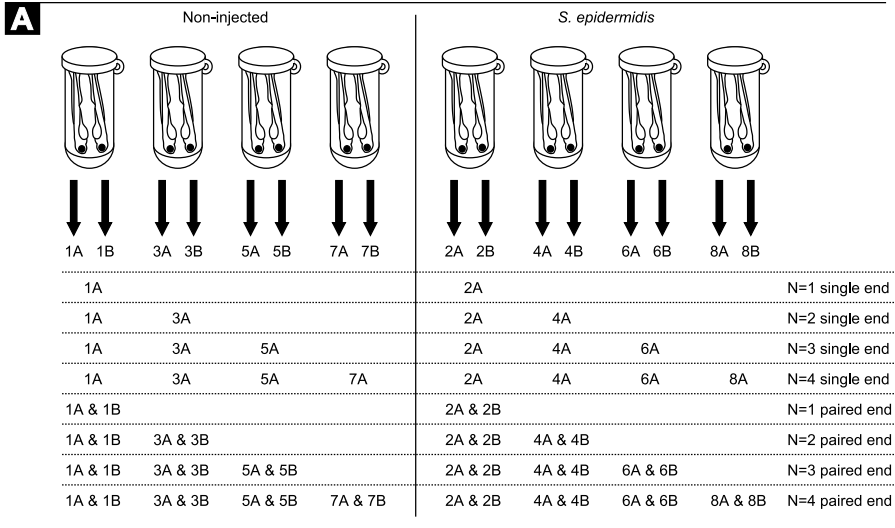
Figure 1: Pipeline of RNAseq data analysis. The diagram shows the workflow of the data analysis starting at the raw fastq files until the final visualization performed automatically by the GeneTiles server. The analysis pipeline is open source, and is available in the supplement.

described tools that have been previously shown to be useful for RNAseq analyses [22]. This makes it very manageable because it reduces the amount of high-end computers required in the research group for alignment and analysis, as all calculations are performed on the server. To start the analysis, the user can choose between various genomes that have been imported from ENSEMBL. Subsequently, the fastq files will be uploaded, followed by the option to analyse the data as single-end or paired-end. The files will be aligned automatically, after which the control or measurement treatment can be chosen. DESeq [23] will normalize the data, and subsequently DEXSeq [21] will extract differentially expressed bins that indicate differential splicing. A table containing the differentially expressed genes or visualizations containing tiles or individual genes can be exported at this point.

With respect to the use of Bowtie2 aligner, we want to point out that it will fail to map reads spanning exon-exon boundaries to the genome. This problem could be solved using a splice aware aligner based on Bowtie2, such as Tophat2 [24]. This option will be included in a future version of the GeneTiles package, However Tophat2 is more computationally intensive, and depends in the correct predictions of splice sites. Therefore the analysis without the splice aware aligner as used in this paper will remain present.

Different analysis methods

Considering that the RNA samples of the *S. epidermidis* infection experiments were paired-end sequenced we had the possibility to explore the added value of paired-end over single-end sequencing. We compared the outcome of the differential expression of these 2 methods as well as the difference in sample sizes as shown in Figure 2A. It can

A — Non-injected / S. epidermidis

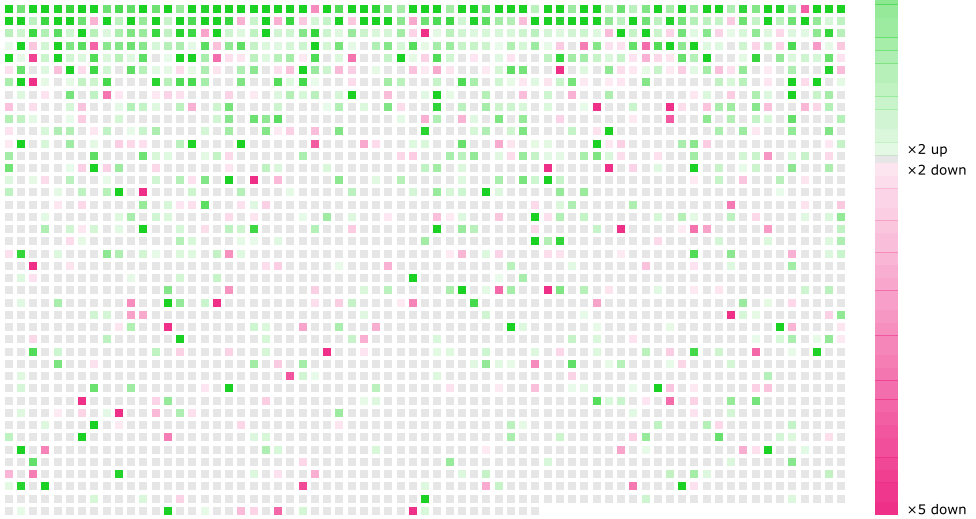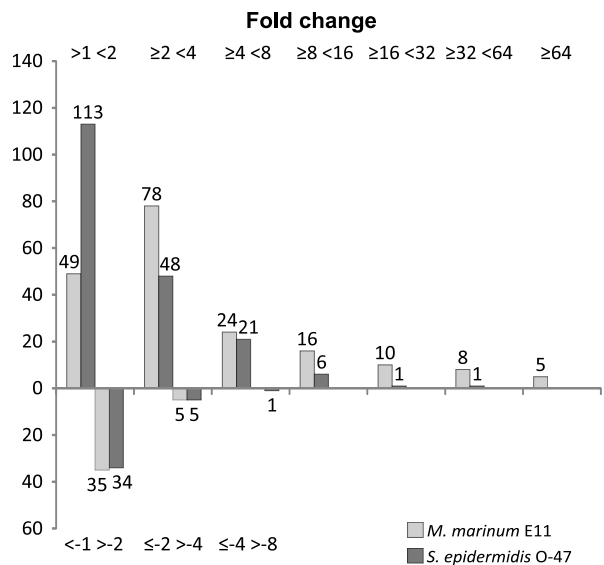| 1A | | | | 2A | | | | N=1 single end |
1A	3A			2A	4A			N=2 single end
1A	3A	5A		2A	4A	6A		N=3 single end
1A	3A	5A	7A	2A	4A	6A	8A	N=4 single end
1A & 1B				2A & 2B				N=1 paired end
1A & 1B	3A & 3B			2A & 2B	4A & 4B			N=2 paired end
1A & 1B	3A & 3B	5A & 5B		2A & 2B	4A & 4B	6A & 6B		N=3 paired end
1A & 1B	3A & 3B	5A & 5B	7A & 7B	2A & 2B	4A & 4B	6A & 6B	8A & 8B	N=4 paired end

Figure 2: Comparing single- and paired-end RNAseq analysis. (A) The different samples sizes (n=1-4) for analysis are visualized. (**B**) The total number of differentially expressed genes with a fold change larger than 2 or smaller than -2 and a *P*-value smaller than 0.05 (solid grey bars) or an adjusted *P*-value smaller than 0.05 (patterned dark bars). (**C**) The light grey circles of the Venn diagrams show the total number of differentially expressed genes for all the samples sizes with a fold change larger than 2 or smaller than -2 and a *P*-value smaller than 0.05 as shown in **B**, and the overlap of differentially expressed genes compared to the n=4 paired-end data set. The bar graphs show the overlap in percentage of these different sample sizes compared to the n=4 paired-end data set.

be noted that a the number of differentially expressed genes does not give an estimate of the reliability of the data, however, considering the high quality of RNAseq data it can be assumed that adding an extra biological sample provides more relevant information than analysing a smaller group of samples by paired-end sequencing. Our data support this assumption since we found only a slight increase of 23% in differentially expressed genes when performing paired-end analysis in the 4 *S. epidermidis* infected samples, using a *P*-value of 0.05 as cut-off filter.

Secondly we analysed the number of differentially expressed genes with all possible options regarding samples sizes (Figure 2B). We found a set of 359 differentially expressed genes when only analysing 1 sample with a *P*-value of 0.05 and compared this large set of genes to our reference set of 203 genes (resulting from the analysis of paired-end sequence data of all four samples) (Figure 2C. This comparison shows a rather small overlap of only 23.7% (single-end) and 27.6% (paired-end). As expected, this overlap increases and therefore the number of false-positives decreases when adding more samples (Figure 2C). In order to provide a statistically more stringent analysis of differentially expressed genes we have also included in the GeneTiles software a tool for minimizing false discovery based on the algorithm of Benjamini et al. 1995 [25], using the implementation of DESeq. The resulting adjusted values show far more stringent results (Figure 2B) but generally confirm the limited value of performing paired-end sequencing as compared to the added value of adding more biological controls.

5

Comparison *S. epidermidis* versus *M. marinum* infection in zebrafish embryos

We compared the different transcriptome host responses of zebrafish embryos upon *S. epidermidis* or *M. marinum* yolk injection. The previous comparison as shown by Veneman et al. (2013) was based on a single biological replicate of *M. marinum* infected zebrafish embryos. We used this single replicate and added 5 more independent biological replicas which led to a total of 6 replicas of *M. marinum*, 4 replicas of *S. epidermidis* infected zebrafish embryos and a total of 9 replicas of non-infected control samples. As found before [1], *S. epidermidis* infection elicits a much smaller transcriptional host response of immune related genes compared to *M. marinum* (Figure 3). However, the total number of differentially expressed genes was increased since several genes that were previously not found to be significantly regulated by *S. epidermidis* where they do show a response in this analysis. Another finding is the high induction of genes upon *M. marinum* infection compared to the *S. epidermidis*

A *S. epidermidis* O-47 genes with best *p*-value
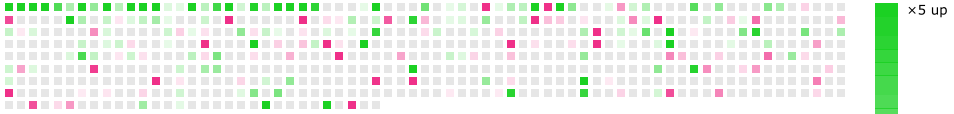
M. marinum E11 genes with best *p*-value

×5 up

×2 up
×2 down

×5 down

B

Up regulated

1722 | 190 | 189

Down regulated

1071 | 40 | 237

Fold change

>1 <2   ≥2 <4   ≥4 <8   ≥8 <16   ≥16 <32   ≥32 <64   ≥64

140
120
100
80
60
40
20
0
20
40
60

49 113
78 48
24 21
16 6
10 1
8 1
5
5 5
1
35 34

<-1 >-2   ≤-2 >-4   ≤-4 >-8

M. marinum E11
S. epidermidis O-47

(see legend on next page)

Figure 3: An overview of differentially expressed genes. (A) The GeneTiles output shows a much larger set of genes with a *P*-value smaller than 0.05 with the *M. marinum* infected samples compared to the *S. epidermidis* infected samples. Each tile visualizes one gene, sorted on *P*-value. The colour and intensity are a function of the ratio between measurement and control samples. **(B)** Comparing the data from **3A** is shown in the Venn diagrams on the left and the overlaps in white digits are shown in a quantitative manner in the bar graph on the right.

infected samples (Figure 3B). An explanation could be that the *M. marinum* bacterium is a natural fish pathogen, and therefore is better recognized as intruder. *S. epidermidis* is in large quantities also pathogenic for fish, however, since it is not a natural pathogen it could very well be that it does not get recognized as well as *M. marinum*. For instance, we now observe a high induction of the leptin b gene (*lepb*) upon infection with *S. epidermidis* (Figure 4). In this case the difference with the previous study is caused by errors in the automated annotation of probes in the micro-array used in Veneman et al. (2013) which was used as bench mark for the RNAseq analysis. The high expression of *lepb* found in *M. marinum* infected samples is in line with the results earlier described by Wieland et al. 2005 [26], where they found a higher mycobacterial load in the lungs of leptin-deficient ob/ob mice.
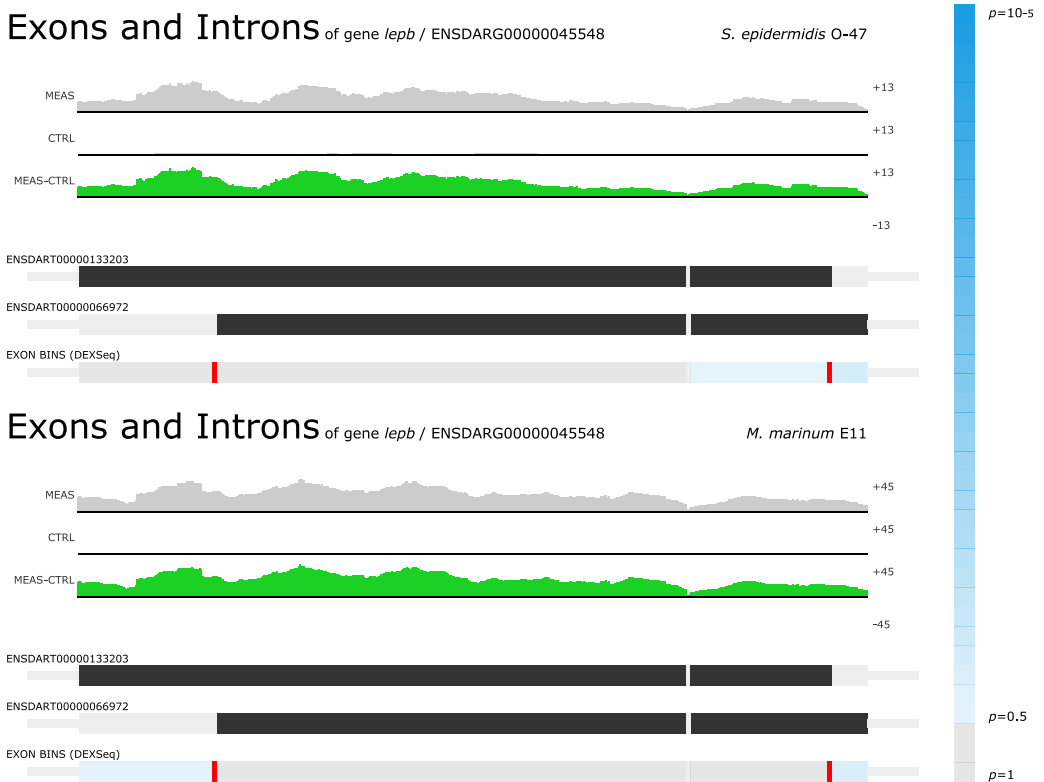

Figure 4: *Lepb* as highest induced gene. For both *S. epidermidis* O-47 (fc: 36, *P*-value: 4.99×10^{-6}) and *M. marinum* E11 (fc: 148, *P*-value: 1.54×10^{-10}) infected samples *lepb* was found to be the highest induced gene.
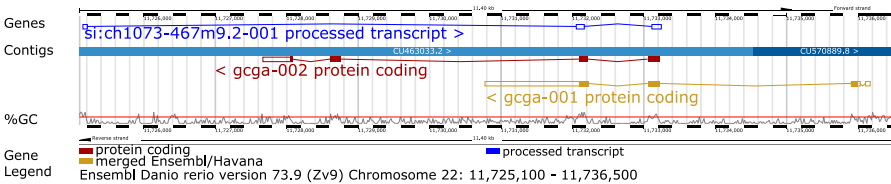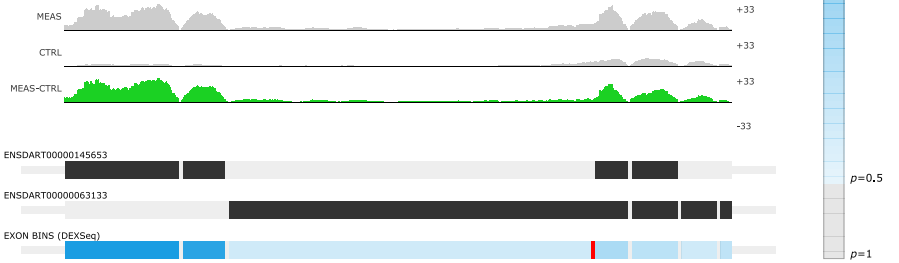
Differential splicing

Another feature in the GeneTiles software is the integration of the DEXSeq analysis tool [21], which allows searching for genes that are differentially spliced. The analysis strategy of differential splicing is schematically shown in Figure 5A, with an example that shows that 1 of 4 exons is spliced out from a pre-mRNA to form the mature mRNA. With both the *S. epidermidis* and *M. marinum* infection data set we found glucagon a (*gcga*) as top candidate to be differentially spliced with large enrichment of two 5' exons (Figure 5B), which are indicated by the dark blue bars underneath the representing exons. Supporting this finding, the pro-glucagon gene has been described before as being differentially spliced into multiple peptides in teleost fish [27].
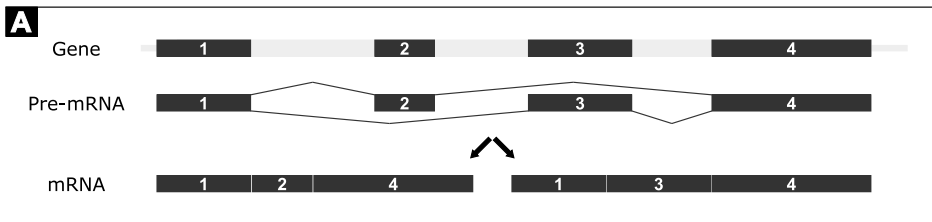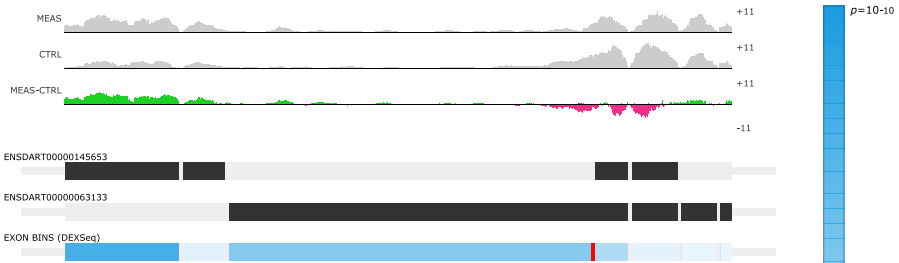
(see legend on next page)

(See Figure on previous page)

Figure 5: Finding differentially spliced genes. (**A**) Schematic view of the principal of differential splicing, where two different mRNAs are formed from one gene. (**B**) For both the *S. epidermidis* (fc: 1.17, *P*-value 4.62x10[-1]) and *M. marinum* (fc: 4.70, *P*-value 2.06x10[-10]) infected samples glucagon a (*gcga*) was found to be differentially spliced as shown by the dark blue bar under the left exon. The screenshot of the GeneTiles visualization demonstrates a schematic view, where introns are compressed to allow for more space for visualization of reads on exons. The gradient-blue bar on the right indicates the *P*-value predicting differential splicing. The bottom image is a gene representation from ENSEMBL.

We also demonstrate that GeneTiles can quickly point out false negative results based on DEXSeq as a result of ambiguous ENSEMBL annotations. For instance in our analysis of infection markers granulin antisense (*grnas*) appeared as a candidate for differential splicing. However, the actual differential expression found of grnas occurs from a fusion of two genes, granulin 1 (*grn1*) and granulin 2 (*grn2*), which are located at this same position as shown in Figure 6.
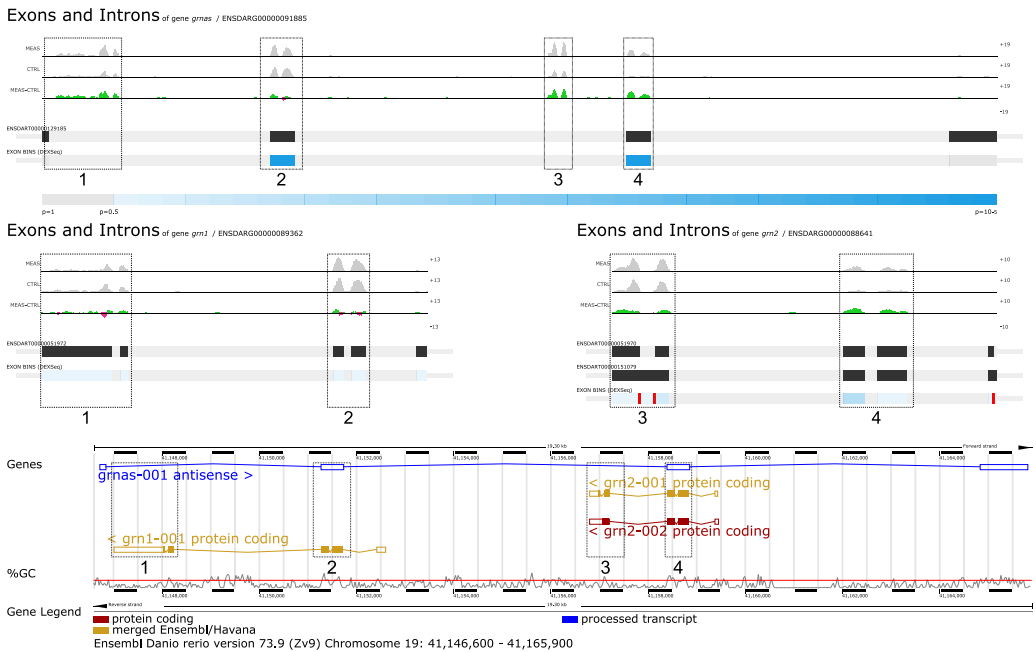

Figure 6: False discovery. The differential splicing indicated by the dark blue bars at exon 2 and 4 from grnas proved to be incorrect. The differential expression found indicated by the 4 boxes at *grnas* (fc: 1.87, *P*-value 3.05x10[-3]) derived from *grn1* (fc: 1.57, *P*-value 1.80x10[-2]) and *grn2* (fc: 3.33, *P*-value 2.23x10[-3]). The screenshot of the GeneTiles visualization demonstrates a non-schematic view, where introns are not compressed showing the actual length of the introns and exons. The gradient-blue bar indicates the *P*-value predicting differential splicing. The bottom image is a gene representation from ENSEMBL.

Comparing different host infection models

To date, we have only found 1 publication in PubMed describing RNAseq analysis of the host after a bacterial infection *in vivo* other than in zebrafish [15]. The data of Scholey et al. 2013 [15], describing bovine digital dermatitis (BDD, an infectious foot disease) was used to compare the host response in the cow and zebrafish. The raw data with accession number GSE41732 was obtained from the GEO database and analysed with the GeneTiles software package. Comparing the differential expression data from Scholey et al. 2013 [15], we could not validate all transcripts, since 23% of the transcripts are retired and are not available anymore, due to the updated version of the *Bos taurus* 4.0 ENSEMBL annotation to the COW UMD3.1 ENSEMBL annotation. All other transcripts could be validated.
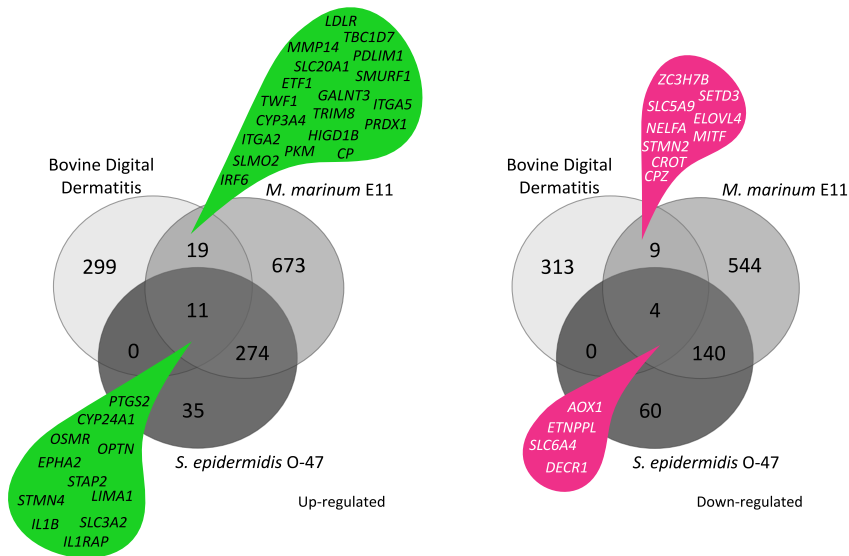
Cluster #	Enrichment score cluster	Ontology	GO #	Name	# Genes	*P*-value
1	3.2	Biological Process	GO:0007584	Response to nutrient	5	1.00E-04
1	3.2	Biological Process	GO:0031667	Response to nutrient levels	5	3.70E-04
1	3.2	Biological Process	GO:0009991	Response to extracellular stimulus	5	5.60E-04
2	2.82	Biological Process	GO:0050896	Response to stimulus	13	1.00E-02
4	2.5	Biological Process	GO:0006979	Response to oxidative stress	4	2.90E-03
6	2.41	Biological Process	GO:0009719	Response to endogenous stimulus	5	5.20E-03
6	2.41	Biological Process	GO:0010033	Response to organic substance	6	7.80E-03
14	1.7	Biological Process	GO:0050727	Regulation of inflammatory response	3	7.90E-03
19	1.62	Biological Process	GO:0080134	Regulation of response to stress	5	1.30E-03
21	1.59	Biological Process	GO:0001666	Response to hypoxia	3	2.30E-02
29	1.29	Biological Process	GO:0048522	Positive regulation of cellular process	8	3.70E-02
34	1.15	Biological Process	GO:0009611	Response to wounding	4	6.60E-02
55	0.7	Biological Process	GO:0050794	Regulation of cellular process	18	3.80E-02
1	2.13	Biological Process	GO:0019752	Carboxylic acid metabolic process	4	6.30E-03
1	2.13	Biological Process	GO:0043436	Oxoacid metabolic process	4	7.20E-03
2	1.94	Molecular Function	GO:0048037	Cofactor binding	3	1.10E-02
3	1.56	Biological Process	GO:0006631	Fatty acid metabolic process	3	7.20E-03
3	1.56	Biological Process	GO:0044255	Cellular lipid metabolic process	3	5.60E-02
4	0.59	Cellular Component	GO:0005622	Intracellular	11	2.30E-01

(Rows grouped under labels "Upregulated" for the first 13 and "Downregulated" for the last 6.)

(see legend on next page)

Figure 7: Overlap with other host pathogen RNAseq experiments. The Venn diagrams show the number of human orthologs of differentially expressed (fc: >2 or <-2 and *P*-value <0.05) genes from the Bovine digital dermatitis, *M. marinum* E11 and *S. epidermidis* O-47 infection data. The gene ontology analysis [28, 29] is based on the overlapping groups indicated by the green and magenta drops.

Both zebrafish and bovine gene identifiers were linked to the human orthologs using the ENSEMBL database. The results show that 61% of the zebrafish genes and 95% of the bovine genes could be translated to human orthologs; a comparison of the differentially expressed gene sets in both disease models (*P*-value < 0.05) is shown in the Venn diagrams in Figure 7. Gene ontology (GO) [28, 29] analysis on the overlapping set of differentially expressed genes between the cow and zebrafish showed that the differentially expressed genes are categorized in multiple response processes (Figure 7). The group of up-regulated overlapping genes between BDD, *M. marinum* and *S. epidermidis* infection includes the following genes; prostaglandin-endoperoxide synthase (*PTGS2*), cytochrome P450 family 24 subfamily A polypeptide 1 (*CYP24A1*), oncostatin M receptor (*OSMR*), Optineurin (*OPTN*), EPH receptor A2 (*EPHA2*), signal transducing adaptor family member 2 (*STAP2*), stathmin-like 4 (*STMN4*), LIM domain and actin binding 1 (*LIMA1*), interleukin 1, beta (*IL1b*), solute carrier family 3 (amino acid transporter heavy chain), member 2 (*SLC3A2*) and interleukin 1 receptor accessory protein (*IL1RAP*). As expected, most of these genes are related to the immune system such as *OPTN* which can activate Fas-ligand pathways to induce apoptosis or anti-inflammatory responses [30], and *IL1B* that is a well-known cytokine produced by activated macrophages, which then can indirectly activate *PTGS2* that also is significantly expressed [31]. The *IL1RAP* is essential for signal transduction of *IL1* in order to induce proinflammatory proteins upon infection [32].

Pathway analysis

The visualization is not only limited to the coloured tiles as described above, but can also be used for functional analysis using WikiPathways [33]. With 96 zebrafish and 267 human pathways at the moment implemented in the software package this allows the user a fast overview of differential expression in biological networks. An example is given in Figure 8, where the Toll-like receptor signalling pathway is showing the differential expression data of *M. marinum* infected embryos using the pathway we submitted to WikiPathways that has been accepted in the curated collection.
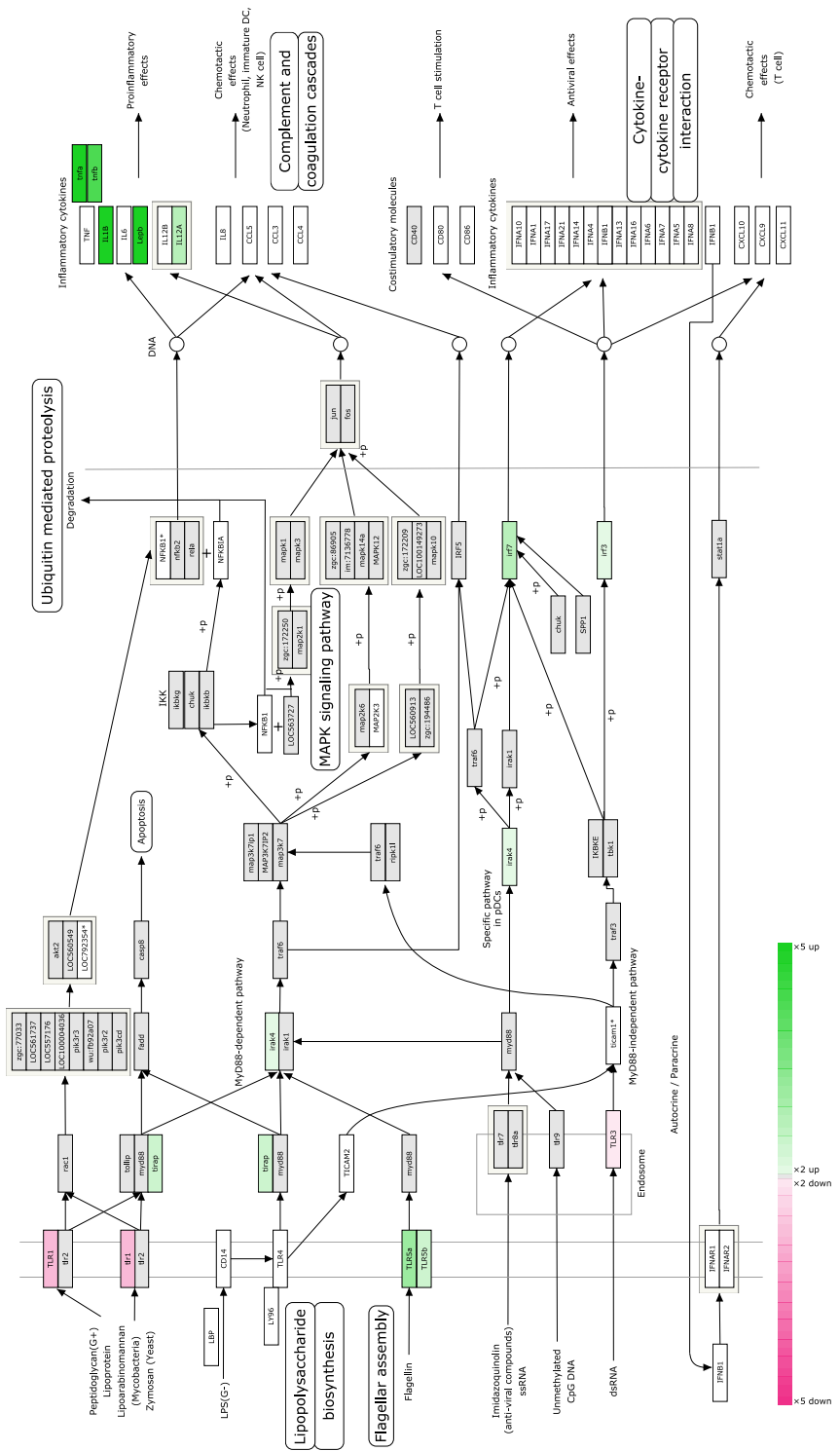
(see legend on next page)

Figure 8: Toll-like receptor pathway showing *M. marium* expression data. RNAseq expression data is shown of *M. marinum* E11 infected zebrafish at 5 dpi. The green and magenta boxes show differential expression respectively up- and down-regulation (fc >2 or <-2) with a *P*-value smaller than 0.05. The genes with an asterisk are discontinued in current databases, the white boxes could not be identified and the grey boxes did not meet the expression criteria.

Lepb as highest induced gene

As mentioned earlier, we found *lepb* as highest induced gene upon infection with *S. epidermidis* and *M. marinum*. Although leptin has been studied over decades for its function in controlling fat balance via brain signalling, its function in innate immune responses has only recently been discovered [34, 35]. For instance, it was shown to be an important factor in mucosal immunity; however, little is known about the underlying mechanisms. For that reason, we designed a morpholino to knock down the function of *lepb* in embryos. At a concentration of 0.5mM we did not found any observable phenotype in development. Combining the *lepb* knock down with the injection of *M. marinum* E11 suggested that loss of *lepb* leads to an increase of the bacterial burden (Figure 9), indicating that *lepb* could have a role during infection. Since leptin has been described to be involved in metabolic processes and the wasting syndrome [36], we performed liquid chromatography mass spectrometry (LC-MS)on 5 day old *lepb* morphants versus wild type control embryos, each with and without *M. marinum* infection. We used a reference set of wasting syndrome associated metabolites (R. Marín-Juez, unpublished results) and found 7 metabolites not influenced by the *M. marinum* infection in the *lepb* knockdown embryos (Figure 10). This indicates that there could be a cross regulation between the nutritional status and immune response against *M. marinum*. Based on these results it could very well be that weight loss due to tuberculosis in humans could be linked to the high levels of *lepb*, as a result of the infection as illustrated in Figure 11. This might be the result of the role of leptin in suppressing appetite or a hitherto undiscovered other function of leptin in metabolic control. However, there are some
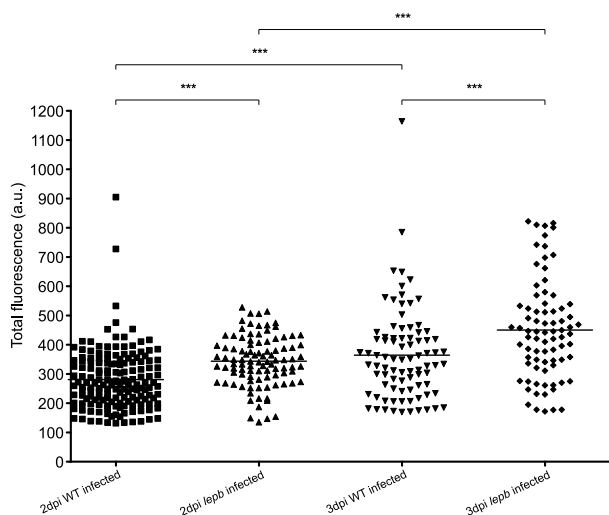
5

Figure 9: The *lepb* morphants show an increased bacterial burden compared to the wild type infected embryos at 2 and 3 dpi. Analysis was performed using the COPAS XL flow cytometer. The asterisk represent the *P*-value obtained by a one way anova; ***P<0.001.
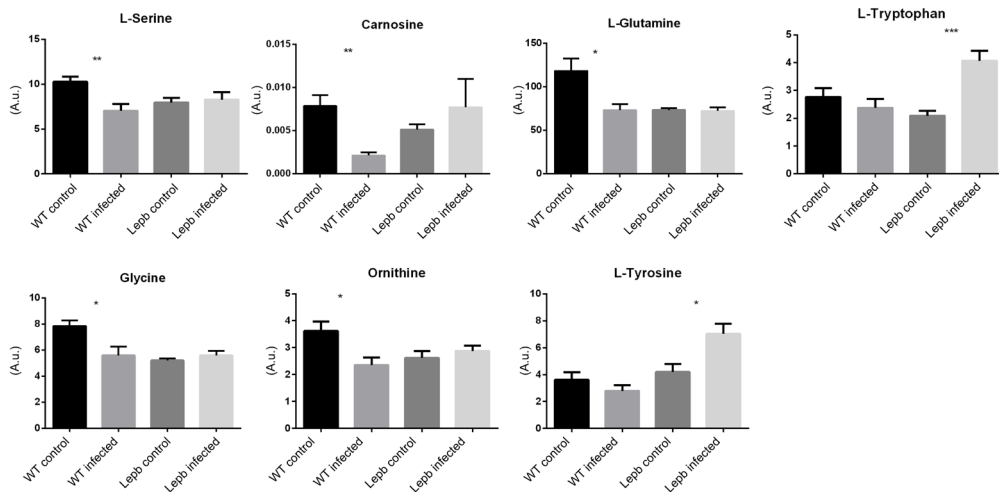
Figure 10: Bar graphs showing 7 wasting metabolites differentially changes in *lepb* morphants with or without *M. marinum* infection. The asterisk represent *P*-values obtained by a student T-test; *P<0.05, **P<0.01, ***P<0.001 and ****P<0.0001. n=8
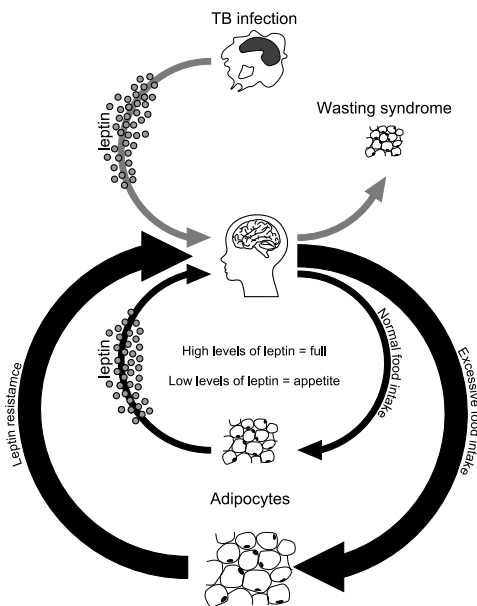

Figure 11: Model of the potential link between leptin production and TB progression. Under normal circumstances leptin regulates appetite and prevents overnutrition. Due to overnutrition leptin resistance can occur, which then can lead to uncontrolled appetite. However, leptin levels are also increased upon TB infection, which in turn can lead to wasting.

papers describing that leptin is not the missing link between the immune response against TB and weight loss [37-39]. For example Schwenk et al. (2003) used in their study heparinized blood samples from humans with pulmonary TB and did not find a correlation of high leptin levels and progression of TB. However, plasma concentrations may not always reflect the biological activity of compounds such as leptin and therefore do not represent the whole systemic situation. This means that our approach of using a whole model organism could still indicate new networks between the function of leptin in the context of a TB infection that are also relevant to the human situation.

Conclusions

The described toolbox for RNAseq data analysis offers two different levels of support in an integrative setting. First, the software combines several programs needed for open source RNAseq analysis such as Bowtie2, Samtools, the 'R' statistical package, DESeq, DEXSeq, HTSeq and pysam. These programs are placed in a pipeline (script) that runs these programs in the required order, with correct in- and output settings. Secondly, the processed data is visualized in a user friendly way and made available for export with a choice of quantitative settings.

The advantages and ease of use of this combined toolbox is demonstrated by analysis of previously published RNAseq datasets from zebrafish and cow infectious disease models, as well as new RNAseq data of a zebrafish mycobacterial infection experiments. This resulted in a highly confident innate marker set for systemic innate immune response to infection by pathogenic and non-pathogenic bacterial species in zebrafish. Furthermore, the data is viewable in a pathway view using the pathways stored at WikiPathways. In this way it was also possible to quickly determine the effect of the number of replicates and the evaluation of potential false positive results, as is the case for the analysis of differential splicing using the DEXSeq algorithm. Comparing our experiences with our previous analyses [1] and the re-analyses performed here we can estimate that we have saved several months of working time while obtaining far superior output files that could be rapidly compared to new RNAseq data sets also from other organisms. The data analysed in this study is available at the GeneTiles website for further analysis and as demonstration material. This makes it possible to rapidly evaluate new immune markers in the datasets described in this paper but also can be used to identify new markers based on other search criteria such as the discovery of the *lepb* induction after infection, which is currently under further investigation.

Availability and requirements

The analysis is open source and available for download, as well as offered as supplementary material. All visualization images are also available for download on the demo page of http://www.genetiles.com. The analysis pipeline, including the source and a complete script to run the same analysis locally is available for download. It is offered together with the open demo for everyone, and it is possible to apply small changes locally to change the analysis or to change the input files.

Acknowledgements

5

Supplementary material

Supplementary material can be found on:

http://link.springer.com/article/10.1007%2Fs00251-014-0820-3

References

1. Veneman WJ, Stockhammer OW, de Boer L, Zaat SA, Meijer AH, Spaink HP: A zebrafish high throughput screening system used for Staphylococcus epidermidis infection marker discovery. BMC Genomics 2013, 14(1):255.

2. van der Vaart M, van Soest JJ, Spaink HP, Meijer AH: Functional analysis of a zebrafish myd88 mutant identifies key transcriptional components of the innate immune system. Dis Model Mech 2013, 6(3):841-854.

3. van Soest JJ, Stockhammer OW, Ordas A, Bloemberg GV, Spaink HP, Meijer AH: Comparison of static immersion and intravenous injection systems for exposure of zebrafish embryos to the natural pathogen Edwardsiella tarda. BMC Immunol 2011, 12:58.

4. Ordas A, Hegedus Z, Henkel CV, Stockhammer OW, Butler D, Jansen HJ, Racz P, Mink M, Spaink HP, Meijer AH: Deep sequencing of the innate immune transcriptomic response of zebrafish embryos to Salmonella infection. Fish Shellfish Immunol 2011, 31(5):716-724.

5. Stockhammer OW, Rauwerda H, Wittink FR, Breit TM, Meijer AH, Spaink HP: Transcriptome analysis of Traf6 function in the innate immune response of zebrafish embryos. Mol Immunol 2010, 48(1-3):179-190.

6. Broekhuizen CA, Schultz MJ, van der Wal AC, Boszhard L, de Boer L, Vandenbroucke-Grauls CM, Zaat SA: Tissue around catheters is a niche for bacteria associated with medical device infection. Crit Care Med 2008, 36(8):2395-2402.

7. Boelens JJ, Dankert J, Murk JL, Weening JJ, van der Poll T, Dingemans KP, Koole L, Laman JD, Zaat SA: Biomaterial-associated persistence of Staphylococcus epidermidis in pericatheter macrophages. J Infect Dis 2000, 181(4):1337-1349.

8. Busscher HJ, van der Mei HC, Subbiahdoss G, Jutte PC, van den Dungen JJ, Zaat SA, Schultz MJ, Grainger DW: Biomaterial-associated infection: locating the finish line in the race for the surface. Sci Transl Med 2012, 4(153):153rv110.

9. Zaat S, Broekhuizen C, Riool M: Host tissue as a niche for biomaterial-associated infection. Future Microbiol 2010, 5(8):1149-1151.

10. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP: Integrative genomics viewer. Nat Biotechnol 2011, 29(1):24-26.

11. Bao H, Guo H, Wang J, Zhou R, Lu X, Shi S: MapView: visualization of short reads alignment on a desktop computer. Bioinformatics 2009, 25(12):1554-1555.

12. Milne I, Bayer M, Cardle L, Shaw P, Stephen G, Wright F, Marshall D: Tablet--next generation sequence assembly visualization. Bioinformatics 2010, 26(3):401-402.

13. Laczik M, Tukacs E, Uzonyi B, Domokos B, Doma Z, Kiss M, Horvath A, Batta Z, Maros-Szabo Z, Torok Z: Geno viewer, a SAM/BAM viewer tool. Bioinformation 2012, 8(2):107-109.

14. Carver T, Harris SR, Otto TD, Berriman M, Parkhill J, McQuillan JA: BamView: visualizing and interpretation of next-generation sequencing read alignments. Brief Bioinform 2013, 14(2):203-212.

15. Scholey RA, Evans NJ, Blowey RW, Massey JP, Murray RD, Smith RF, Ollier WE, Carter SD: Identifying host pathogenic pathways in bovine digital dermatitis by RNA-Seq analysis. Vet J 2013, 197(3):699-706.

16. Carvalho R, de Sonneville J, Stockhammer OW, Savage ND, Veneman WJ, Ottenhoff TH, Dirks RP, Meijer AH, Spaink HP: A high-throughput screen for tuberculosis progression. PLoS One 2011, 6(2):e16779.

17. Xia J, Psychogios N, Young N, Wishart DS: MetaboAnalyst: a web server for metabolomic data analysis and interpretation. Nucleic Acids Res 2009, 37(Web Server issue):W652-660.

18. Langmead B, Salzberg SL: Fast gapped-read alignment with Bowtie 2. Nat Methods 2012, 9(4):357-359.

5

19. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R: The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009, 25(16):2078-2079.

20. Anders S, Pyl PT, Huber W: HTSeq-a Python framework to work with high-throughput sequencing data. Bioinformatics 2014.

21. Anders S, Reyes A, Huber W: Detecting differential usage of exons from RNA-seq data. Genome Res 2012, 22(10):2008-2017.

22. Hatem A, Bozdag D, Toland AE, Catalyurek UV: Benchmarking short sequence mapping tools. BMC Bioinformatics 2013, 14:184.

23. Anders S, Huber W: Differential expression analysis for sequence count data. Genome Biol 2010, 11(10):R106.

24. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL: TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 2013, 14(4):R36.

25. Benjamini Y, Hochberg Y: Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc B Met 1995, 57(1):289-300.

26. Wieland CW, Florquin S, Chan ED, Leemans JC, Weijer S, Verbon A, Fantuzzi G, van der Poll T: Pulmonary Mycobacterium tuberculosis infection in leptin-deficient ob/ob mice. Int Immunol 2005, 17(11):1399-1408.

27. Holland LZ, Short S: Alternative splicing in development and function of chordate endocrine systems: a focus on Pax genes. Integr Comp Biol 2010, 50(1):22-34.

28. Huang da W, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009, 4(1):44-57.

29. Huang da W, Sherman BT, Lempicki RA: Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009, 37(1):1-13.

30. Wild P, Farhan H, McEwan DG, Wagner S, Rogov VV, Brady NR, Richter B, Korac J, Waidmann O, Choudhary C et al: Phosphorylation of the autophagy receptor optineurin restricts Salmonella growth. Science 2011, 333(6039):228-233.

31. Lappas M: NOD1 and NOD2 regulate proinflammatory and prolabor mediators in human fetal membranes and myometrium via nuclear factor-kappa B. Biol Reprod 2013, 89(1):14.

32. Subramaniam S, Stansberg C, Cunningham C: The interleukin 1 receptor family. Dev

5

Comp Immunol 2004, 28(5):415-428.

33. Kelder T, van Iersel MP, Hanspers K, Kutmon M, Conklin BR, Evelo CT, Pico AR: WikiPathways: building research communities on biological pathways. Nucleic Acids Res 2012, 40(Database issue):D1301-1307.

34. Paz-Filho G, Mastronardi C, Franco CB, Wang KB, Wong ML, Licinio J: Leptin: molecular mechanisms, systemic pro-inflammatory effects, and clinical implications. Arq Bras Endocrinol Metabol 2012, 56(9):597-607.

35. Matarese G, Moschos S, Mantzoros CS: Leptin in immunology. J Immunol 2005, 174(6):3137-3142.

36. Gaetke LM, Oz HS, de Villiers WJS, Varilek GW, Frederich RC: The leptin defense against wasting is abolished in the IL-2-deficient mouse model of inflammatory bowel disease. J Nutr 2002, 132(5):893-896.

37. Schwenk A, Hodgson L, Rayner CF, Griffin GE, Macallan DC: Leptin and energy metabolism in pulmonary tuberculosis. Am J Clin Nutr 2003, 77(2):392-398.

38. van Crevel R, Karyadi E, Netea MG, Verhoef H, Nelwan RH, West CE, van der Meer JW: Decreased plasma leptin concentrations in tuberculosis patients are associated with wasting and inflammation. J Clin Endocrinol Metab 2002, 87(2):758-763.

39. Santucci N, Diaz A, Bianchi E, Spinelli S, D'Attilio L, Bongiovanni B, Didoli G, Brandan N, Nannini L, Bay ML et al: Leptin does not enhance cell-mediated immune responses following mycobacterial antigen stimulation. Int J Tuberc Lung Dis 2014, 18(8):981-987.

40. Inferring deferential exon usage in RNA-Seq data with the DEXSeq package [http://www.bioconductor.org/packages/2.13/bioc/vignettes/DEXSeq/inst/doc/DEXSeq.pdf]

5

Chapter 6

Distribution of micro and nano range sized biomaterials after yolk injection into zebrafish larvae

Wouter J. Veneman[1,2], Martina G. Vijver[2],
Annemarie H. Meijer[1], Herman P. Spaink[1]

[1]Institute of Biology Leiden, Leiden University
[2]Institute of Environmental Sciences, Leiden University

6

Abstract

Biomaterial-associated infection is a major problem in modern medicine. The presence of foreign body materials such as artificial hips or intravenous needles can in combination with the normally innocent skin bacterium, *Staphylococcus epidermidis* cause infections that are difficult to treat. To test and develop new compositions that could be used as biomaterials we developed a screening method. Here we report the outcome of a variety of nanometer and micrometer sized biomaterials injected into the zebrafish embryos. We found a size depended outcome of the polystyrene biomaterial distribution throughout the embryo. These results could hopefully contribute to the early stage of development of biocompatible materials for modern medicine.

Introduction

Thanks to modern medicine, the past decades people have a longer life expectation and their physical condition can be improved up to old age, since it is easier to replace body parts such bones or cartilage. However, the risk of the accompanying surgical treatments lead to a great chance of infection due to the implantation of the used biomaterial/biomedical devices such as an artificial hips, intravenous needles or catheters. These biomaterial-associated infections (BAI) are mainly caused by the skin bacterium *Staphylococcus epidermidis* [1-4]. This means that there is a great demand for biomaterials and medical devices that are biocompatible in such a way that it does not get rejected by the human body but also does not induce infections.

In previous studies we already showed that the zebrafish embryo is a versatile model to study the pathogenesis of *S. epidermidis* [5, 6]. The use of zebrafish larvae as whole vertebrate animal model in modern research has been well established. The larvae have great *in vivo* imaging potential due to their transparency and the large amount of transgenic strains expressing fluorescent proteins [7-9]. Secondly they are very well suited for high-throughput screening since zebrafish are relatively cheap to house, the eggs can be obtained in large amounts in a single time, and solutions containing DNA, RNA, morpholinos and pathogens can be introduced robotically [10, 11]. Furthermore, the analysis of these injected embryos can be followed over time using automated screening methods [5]. All these advantages have led to the fact that many human diseases such as tuberculosis, cancer, and cardiovascular defects have been modelled using zebrafish larvae [12-20].

In order to develop a screening model for BAI using zebrafish we report in this paper the outcome of distribution of biomaterials after injection. For this we used a set of fluorescently labelled polystyrene particles with a diameter ranging from 15 micrometer to 70 nanometer. Our first strategy was to inject particles as large as possible, so that these would not get phagocytized, but still be small enough to be injected easily using glass micro capillaries. The second strategy was to inject nanometer sized particles,

118

since these will probably spread easier due to their phagocytosis by motile immune cells, and could therefore provide more details about the distribution pattern of how injected agents would behave. We mainly chose the yolk of eggs between the 16 and 256 cell stage as injection site, so results would be comparable with earlier reports of injection of *S. epidermidis* into the yolk. Secondly this method can easily be automated for future experiments or large high-throughput screens.

Material and methods

Zebrafish husbandry

Zebrafish were handled in compliance with animal welfare regulations and maintained according to standard protocols (http://ZFIN.org). Embryos were grown at 28°C in egg water (60 µg/ml Instant ocean sea salt, Sera Marin). The egg water was refreshed every day.

Experimental design

Injections were performed using mixed egg clutches from the *Tg(bactin:Hras-EGFP VU119)* [21] strain. For yolk injections: eggs were staged between 16 and 128 cell stage by morphological criteria, and injected into the yolk with 1 nl of 4% polyvinylpyrrolidone$_{40}$ (PVP$_{40}$) containing 5 mg/ml nanometer or micrometer particles (see table 1). For tail muscle injection: 2 and 3 day old larvae were staged by morphological criteria and anaesthetized (0.02% buffered 3-aminobenzoic acid ethyl ester (Tricaine, Sigma) in egg water). Larvae were injected into the tail muscle [22] with approximately 1-10 nl containing one particle. Injections were controlled using a Leica M50 stereomicroscope together with a FemtoJet micro-injector (Eppendorf) and a micromanipulator with pulled and bevelled micro capillary needles.

Table 1: commercial available nanometer and micrometer sized particles as used in the experiments. nm= nanometer, µm= micrometer

Size	Composition	Fluorescence dye	Manufacture	Reference number
70 nm	Polystyrene	Nile red	Corpuscular inc. (USA)	103125-05
250 nm	Polystyrene	Nile red	Corpuscular inc. (USA)	103127-05
2.2 µm	Polystyrene	Nile red	Corpuscular inc. (USA)	103235-05
4.2 µm	Polystyrene	Nile red	Corpuscular inc. (USA)	103129-05
9.9 µm	Polystyrene	Nile red	Corpuscular inc. (USA)	103245-05
15 µm	Polystyrene	Nile red	Corpuscular inc. (USA)	103251-05

Microscopy

Embryos were examined and imaged daily using a fluorescence stereo microscope (MZ 205 FA, Leica). Embryos were kept under anaesthesia (0.02% buffered 3-aminobenzoic acid ethyl ester (Tricaine, Sigma) in egg water) during imaging. The image processing package Fiji [23] was used for processing and analysis of the images.

Results and discussion

We used a set of polystyrene particles as proof of principle for studying the distribution of injected biomaterials into embryos. These polystyrene particles are chosen as they are commercially available in a monodispersed form in many different sizes. Earlier biomaterial injection experiments with materials that are more related to medical devices such as titanium, poly(lactic-co-glycolic acid) or polycaprolactone led to difficulties. These particles are irregular in shape as well size, with as result clogging of the glass micro capillaries and inconsistent outcomes. However, with the monodispersed polystyrene particles these problems did not occur. Using a 4% PVP_{40} solution as a carrier made it possible to inject polystyrene particles into the embryonic yolk because the viscosity of the solvent prevented the particles from flowing out after injection. The visualization of the distribution of particles was facilitated by the use of the *Tg(bactin:Hras-EGFP VU119)* strain which expresses membrane-targeted GFP as shown in Figure 1 and 2. Following injection into the yolk of 16-128 cell stage embryos, we observed a difference in the distribution of biomaterials of various sizes throughout the body of the embryo (Figure 3). The nanometer sized biomaterials spread more compared to the micrometer sized biomaterials. One possible explanation could be that the concentration of particles per nanoliter of the small particles was higher since all suspensions were diluted to a 5 mg/ml concentration and the same volume of materials was injected. Another explanation could be that there is indeed a size dependent factor, which makes that the smaller beads spread easier than larger sized beads. This is in line with earlier findings of Desai et al. (1996 & 1997) [24, 25], who also found that nanometer sized particles had a higher uptake in rat tissue and *in vitro* compared to the micrometer sized particles. We did find that the smaller sized particles (70 & 250 nm) accumulate around the heart region at 3 to 4 days post injection (dpi) (Figure 2) from where they probably distribute throughout the body. In a few experiments we found a very small number of embryos with larger sized beads (≥ 9.9 μm) throughout the body. Since the embryos did not show any signs of distributed particles yet at 1 day post fertilization (dpf), we can rule out the possibility that the yolk-injected particles were spread as a result of uptake by cells during the first hours of development, as in the case with early RNA or DNA injections before the 16 cell stage. It is also unlikely that these large particles were distributed at later developmental stages via the vasculature, since the maximum diameter of the vasculature is also around 10 μm in diameter. We performed a whole mount L-plastin immunohistochemistry [26] on larvae injected with the larger sized particles (≥ 9.9 μm) to check for possible uptake by leukocytes at later stages from 2 until 5 dpi, but we could not colocalize the particles with leukocytes throughout the body (data not shown).

Even though the percentage of embryos showing a distribution of particles through the body is low, it indicates that it is possible for such materials being taken up into the body of the embryo after injection into the yolk. This has implications for the results of earlier reports [5, 6, 10, 11] on spreading of bacteria and human cancer cells after
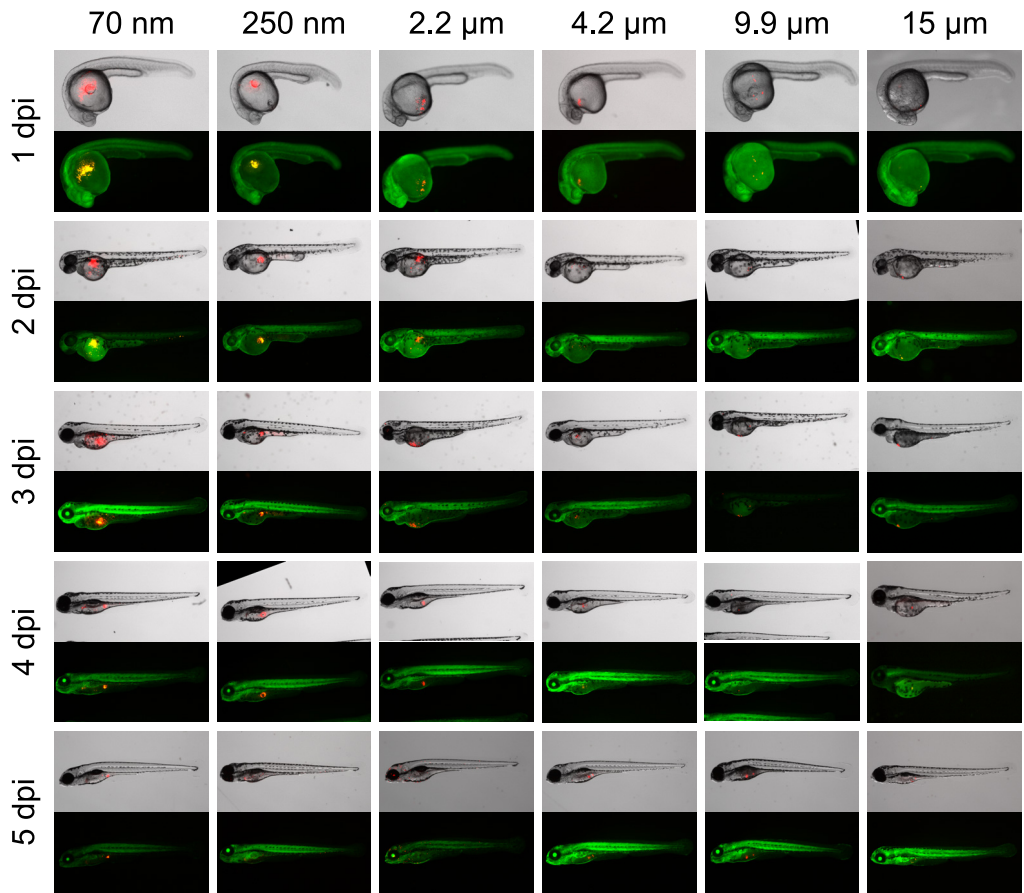
Figure 1: Representative bright field /fluorescence overlay images showing migration of Nile red fluorescence particles throughout the *Tg(bactin:Hras-EGFP VU119)* embryos over a 5 day period. The x axis shows the different sizes of particles injected and the y axis shows the time period from 1 until 5 dpi.

being injected into the early embryonic yolk since it indicates that distribution could be driven by unknown aspecific processes from the host. We are hoping to unravel the mechanism(s) of such aspecific transport from yolk to body in future research.

All previously described experiments were based on the yolk injection method. We also performed injections of beads in the tail muscle of 2 and 3 day old larvae, which mimics the implantation of an artificial implant in human chirurgical applications, since the implantation is performed directly into tissue. However, when using the 9.9 μm sized particles these injections proved to be very labour intensive, and did not result in the biomaterial to remain in the same place after injection (data not shown). In almost all cases the particles were pushed out of the tissue while injecting. Even when the particle would stay at the implantation side, a considerable amount of tissue damage occurred.

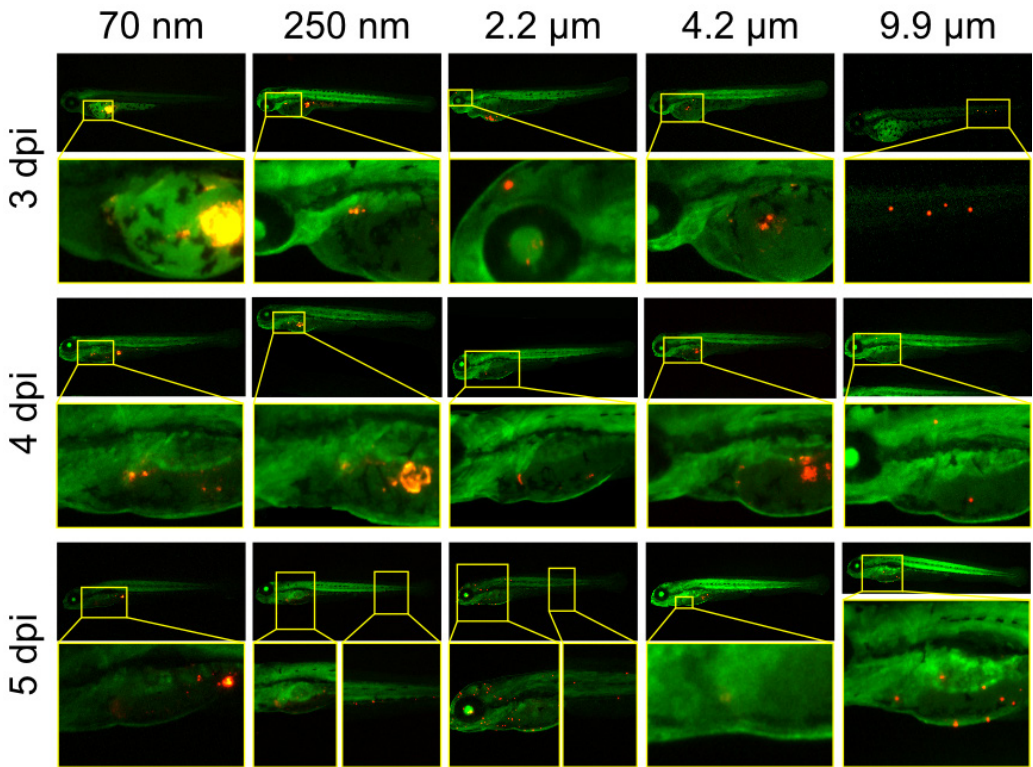**Figure 2: Detailed images of *Tg(bactin:Hras-EGFP VU119)* larvae injected with particles.** The yellow boxes show the particles close to the heart region or already migrated away from the yolk injection side.
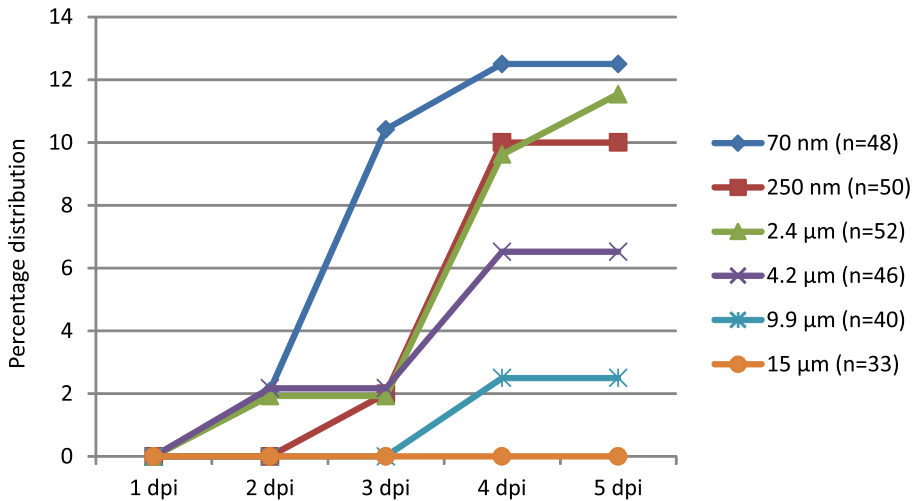

Figure 3: Quantification of distributed particles. The graph shows the percentage of embryos with nano or micrometer sized beads that were spread away from the yolk injection side at subsequent dpi.

Conclusions

We found that there was a difference based on size in distribution of biomaterials. The preferable larger sized biomaterials (≥ 9.9 μm) migrated less than the smaller sized biomaterials (≤ 2.2 μm). The location where they accumulate still remains unknown, however if they spread, such as the smaller particles (≥ 2.2 μm), it probably starts at the heart region from where they distribute out inside the body. The mechanism of how particles spread also remains unknown, however we think that the larger (≥ 9.9 μm) particles do not get taken up by leucocytes and there is no evidence for spreading via the vasculature. Nevertheless we hope to find the mechanisms/processes that are involved during the migration of these particles in the future.

Acknowledgements

References

1. Broekhuizen CA, Schultz MJ, van der Wal AC, Boszhard L, de Boer L, Vandenbroucke-Grauls CM, Zaat SA: Tissue around catheters is a niche for bacteria associated with medical device infection. Crit Care Med 2008, 36(8):2395-2402.

2. Busscher HJ, van der Mei HC, Subbiahdoss G, Jutte PC, van den Dungen JJ, Zaat SA, Schultz MJ, Grainger DW: Biomaterial-associated infection: locating the finish line in the race for the surface. Sci Transl Med 2012, 4(153):153rv110.

3. Zaat S, Broekhuizen C, Riool M: Host tissue as a niche for biomaterial-associated infection. Future Microbiol 2010, 5(8):1149-1151.

4. Boelens JJ, Dankert J, Murk JL, Weening JJ, van der Poll T, Dingemans KP, Koole L, Laman JD, Zaat SA: Biomaterial-associated persistence of Staphylococcus epidermidis in pericatheter macrophages. J Infect Dis 2000, 181(4):1337-1349.

5. Veneman WJ, Marin-Juez R, de Sonneville J, Ordas A, Jong-Raadsen S, Meijer AH, Spaink HP: Establishment and optimization of a high throughput setup to study Staphylococcus epidermidis and Mycobacterium marinum infection as a model for drug discovery. J Vis Exp 2014(88):e51649.

6

6. Veneman WJ, Stockhammer OW, de Boer L, Zaat SA, Meijer AH, Spaink HP: A zebrafish high throughput screening system used for Staphylococcus epidermidis infection marker discovery. BMC Genomics 2013, 14:255.

7. Brittijn SA, Duivesteijn SJ, Belmamoune M, Bertens LF, Bitter W, de Bruijn JD, Champagne DL, Cuppen E, Flik G, Vandenbroucke-Grauls CM et al: Zebrafish development and regeneration: new tools for biomedical research. Int J Dev Biol 2009, 53(5-6):835-850.

8. Meijer AH, Spaink HP: Host-pathogen interactions made transparent with the zebrafish model. Curr Drug Targets 2011, 12(7):1000-1017.

9. Meijer AH, van der Vaart M, Spaink HP: Real-time imaging and genetic dissection of host-microbe interactions in zebrafish. Cell Microbiol 2014, 16(1):39-49.

10. Carvalho R, de Sonneville J, Stockhammer OW, Savage ND, Veneman WJ, Ottenhoff TH, Dirks RP, Meijer AH, Spaink HP: A high-throughput screen for tuberculosis progression. PLoS One 2011, 6(2):e16779.

11. Spaink HP, Cui C, Wiweger MI, Jansen HJ, Veneman WJ, Marin-Juez R, de Sonneville J, Ordas A, Torraca V, van der Ent W et al: Robotic injection of zebrafish embryos for high-throughput screening in disease models. Methods 2013, 62(3):246-254.

12. Lieschke GJ, Currie PD: Animal models of human disease: zebrafish swim into view. Nat Rev Genet 2007, 8(5):353-367.

13. Cronan MR, Tobin DM: Fit for consumption: zebrafish as a model for tuberculosis. Dis Model Mech 2014, 7(7):777-784.

14. Tobin DM, Ramakrishnan L: Comparative pathogenesis of Mycobacterium marinum and Mycobacterium tuberculosis. Cell Microbiol 2008, 10(5):1027-1039.

15. van der Ent W, Jochemsen AG, Teunisse AF, Krens SF, Szuhai K, Spaink HP, Hogendoorn PC, Snaar-Jagalska BE: Ewing sarcoma inhibition by disruption of EWSR1-FLI1 transcriptional activity and reactivation of p53. J Pathol 2014, 233(4):415-424.

16. Amatruda JF, Patton EE: Genetic models of cancer in zebrafish. Int Rev Cell Mol Biol 2008, 271:1-34.

17. Etchin J, Kanki JP, Look AT: Zebrafish as a model for the study of human cancer. Methods Cell Biol 2011, 105:309-337.

18. Konantz M, Balci TB, Hartwig UF, Dellaire G, Andre MC, Berman JN, Lengerke C: Zebrafish xenografts as a tool for in vivo studies on human cancer. Ann N Y Acad Sci 2012, 1266:124-137.

6

19.   Payne E, Look T: Zebrafish modelling of leukaemias. Br J Haematol 2009, 146(3):247-256.

20.   Bakkers J: Zebrafish as a model to study cardiac development and human cardiac disease. Cardiovasc Res 2011, 91(2):279-288.

21.   Cooper MS, Szeto DP, Sommers-Herivel G, Topezewski J, Solnica-Krezel L, Kang HC, Johnson I, Kimelman D: Visualizing morphogenesis in transgenic zebrafish embryos using BODIPY TR methyl ester dye as a vital counterstain for GFP. Dev Dynam 2005, 232(2):359-368.

22.   Benard EL, van der Sar AM, Ellett F, Lieschke GJ, Spaink HP, Meijer AH: Infection of zebrafish embryos with intracellular bacterial pathogens. J Vis Exp 2012(61).

23.   Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B et al: Fiji: an open-source platform for biological-image analysis. Nat Methods 2012, 9(7):676-682.

24.   Desai MP, Labhasetwar V, Amidon GL, Levy RJ: Gastrointestinal uptake of biodegradable microparticles: Effect of particle size. Pharmaceut Res 1996, 13(12):1838-1845.

25.   Desai MP, Labhasetwar V, Walter E, Levy RJ, Amidon GL: The mechanism of uptake of biodegradable microparticles in Caco-2 cells is size dependent. Pharmaceut Res 1997, 14(11):1568-1573.

26.   Mathias JR, Dodd ME, Walters KB, Rhodes J, Kanki JP, Look AT, Huttenlocher A: Live imaging of chronic inflammation caused by mutation of zebrafish Hai1. J Cell Sci 2007, 120(Pt 19):3372-3383.

6

Chapter 7

General discussion and summary

Thanks to all the great medical inventions in the last couple of centuries the life expectancy of the human population is significantly increasing. By the year 2050 twenty-two percent of the human population will be older than the age of 60. This puts great pressure on the health care, since this group will most likely need medical care at a certain point due to cancer, diabetes, osteoporosis, heart diseases or any other kind of age-related disease. However, there can still be great improvements achieved in the field of medical instruments and implants. Although there are already many implants or medical tools available, there is still a risk of infection, related to these biomaterials. One of the main priorities during development is the functionality of the material, such as a titanium hip that lasts a long time. However, biocompatibility is also a factor of great importance. Testing biocompatibility of new materials in early stages of development could save time and money. However choosing the right model for the research question can be difficult. Does the data from cell and tissues culture provide sufficient information or does it lack the complex networks that you can find in animal models, and which animal represents the human situation the best? At the same time the requirements for the laboratories and facilities are also playing a very important role, since some pathogens or animals can be examined in biosafety level 1 laboratories where other are restricted to biosafety level 3 or 4 laboratories.

Development and optimization of high-throughput zebrafish techniques

The work in this thesis uses zebrafish larvae as a test organism for studying various infectious diseases. Most of the advantages of using the zebrafish larvae as a model are already discussed in detail in the previous chapters. One of these advantages and the focus of this thesis is that zebrafish eggs and larvae are ideal for high-throughput screening. We analysed which developmental stage is optimal for spreading of bacterial though the body after yolk injection with 2 types of bacteria, namely *Staphylococcus epidermidis* and *Mycobacterium marinum*. We also compared the outcome with other injection methods like injecting into the caudal vein at 1 day post fertilization (dpf). We determined that injection into the yolk with *S. epidermidis* and *M. marinum* between the 16-128 cell stage results in reproducible infections with spreading of the bacteria into the tissues of the embryo at later stages. In contrast, injection before the 16 cell stage leads to a high mortality and injecting after the 128 cell stage does not lead to a representative infection. The more virulent *M. marinum* M strain does not even give a representative infection after the 64 cell stage (**chapter 3**).

The early developing zebrafish egg is ideal to automatically inject into the yolk, and we showed that this could be performed with RNA, DNA, bacteria or cancer cells (**chapter 2**). The automated micro-injector allowed us to inject up to 2000 eggs per hour in a consistent manner into the yolk of a developing zebrafish egg. This kind of specifications cannot be reproduced with manual injection techniques. Using this kind of automated micro-injector yields therefore a strongly increased number of samples that can be used for compound screening or testing other biological questions, and with that reduces valuable research time.

128

However handling these amounts of samples also requires fast and efficient analysis techniques. We made use of a model of flow-cytometry that can handle extraordinary large particle sizes (called COPAS XL), which could analyse and sort large amounts of infected larvae (2000/30min), without harming the larvae. Therefore this analysis could be performed daily to monitor the fluorescence signal of the bacterial burden in large groups of infected larvae (**chapter 4**). The only drawback of performing such flow-cytometry is the low resolution. However, to overcome this problem we also implemented a medium throughput, high resolution method. Therefore, we used the Vertebrate Automated Screening Technique (VAST BioImager), that automatically loads and positions the larvae to predefined settings. The imaging was performed using a confocal laser scanning microscope (CLSM) with an additional colour camera (**chapter 3**).

Using this high-throughput pipeline we were able to understand the pathogenesis of *S. epidermidis* in zebrafish larvae better. We performed intensive microscopy, flow cytometry and transcriptome analysis, on the infection process of *S. epidermidis* in zebrafish larvae. The use of the COPAS XL proved to be accurate to analyse the bacterial burden without homogenizing and plating the embryos. Secondly the use of micro-array analysis led to detailed host transcriptome expression, which was replaced by RNA deep sequencing (RNAseq) within the project.

Transcriptome analysis

Although the micro-array approach provides useful expression patterns of the transcriptome, the method of detection is rather biased. Micro-arrays use labelled probes for the detection which means that if a probe is not present on the chip it will not be found. RNAseq on the other hand counts short reads based on a reference genome as illustrated in Figure 1. This means that also expression of non-annotated sequences can be discovered and the analysis can be repeated when updated genomes are released. We chose to use the RNAseq analysis method using Illumina sequencing which is based on the sequencing of cDNA that is de novo replicated on flow cells. However, there are also other options available for the analysis of differential expression of RNA such as Ion Torrent sequencing, which also measures cDNA instead of RNA but on a semiconductor chip. In this technology small fragments bind to specific beads from where hydrogen ions will be released when a nucleotide binds to the fragment. This leads to changes of the pH that can be measured with a voltage meter. These changes in voltage can therefore be used to read the sequence. Another option that might be used in the near future is direct sequencing of RNA using the MinION device. This is based on consumable flow cells with nanopores that can sense single molecules. This little device with the size of a USB-stick can be used for DNA sequencing but could, theoretically, also be used also for RNA and protein sequencing. However, the nanopore technology is still under development and although it offers good perspectives for further decreasing cost of sequencing its suitability for direct RNA sequencing has still not been demonstrated.
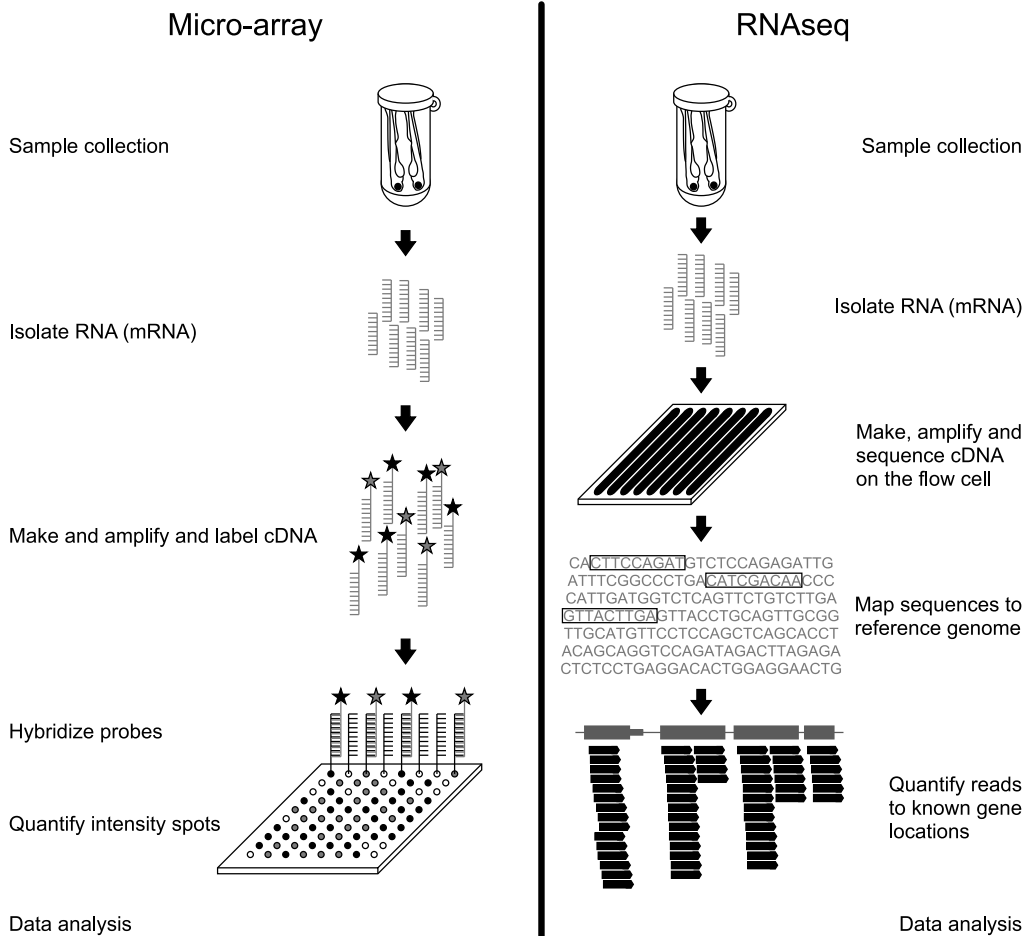
7

Figure 1: The differential expression analysis methods. The left side shows the micro-array technique with hybridization of the probes on the chip. Only targets present on the chip be found, and due to multiple probes of a target to improve sensitivity, this will reduce the total amount of total targets on the chip. The right side shows the RNAseq technique, which uses short reads to quantify the expression levels. Therefore unknown sequences can be found and the analysis can be repeated with newer and updated versions of the reference genome.

A possible draw-back of RNAseq technology is that the analysis of sequence data requires some dedicated bioinformatics skills that are not always available in biology institutes. A possible solution would be to outsource the bioinformatic analyses to companies but this is currently very costly. For that reason we collaborated with bioinformaticians and statisticians to design a platform for fast and user friendly analysis of large RNAseq data files. This resulted in the design of the GeneTiles software package (**chapter 5**), which combines all the programs needed for basic RNAseq analysis such as Bowtie2, Samtools, the 'R' statistical package, DESeq, DEXSeq, HTSeq and pysam. All these programs run in a pipeline in a server based environment, which can be accessed from every computer with operations systems such as Windows and Linux with an internet connection. One of the advantages of having the analysis run on a server is that there is no need any more

for powerful computers in the laboratory. To offer a complete package that includes the entire analysis and direct visualization, multiple options are implemented, such as differential expression sorting on ratio, *P*-value, adjusted *P*-value, differential splicing or just chromosome view. In addition to this we also included all available biological pathways from Wikipathways. This allows quicker and better interpretation of the expression data in a more complex manner. As result of this improved RNAseq analysis method we found a gene that was differentially spliced under infectious conditions, namely glucagon a (*gcga*).

Metabolic changes upon infection

Another finding was the high induction of the leptin b (*lepb*) gene in RNAseq data both *S. epidermidis* and *M. marinum* infected embryos. This was not discovered earlier since the *lepb* probe was not present on the micro-array. The leptin hormone is normally produced by fat cells, which regulate the fat balance in the human body. If a person consumes food, a signal is given to the hypothalamus, which then regulates the appetite, and prevents from overeating. However, if a person continuously overeats, this can lead to leptin resistance. This means that the leptin signal does not arrive at the hypothalamus and that there is no feedback signal that prevents overeating. Therefore people will eat more than they need and gain weight with obesity as a result (Figure 2). This can in turn also lead to continuous production of too high levels of insulin leading to insulin resistance. Insulin resistance is the hall mark of diabetes type 2 that results in chronic problems in diminishing blood glucose levels (leading to hyperglycaemia), and a lack of conversion of glucose to glycogen in tissues. Due to the high glucose levels in the circulation, people will also get more susceptible to heart diseases, stroke, blindness and kidney failure.

Since the leptin gene is a very interesting gene which is already linked to multiple metabolic and immune functions, and *lepb* was the highest induced gene upon infection with *M. marinum* and *S. epidermidis* we designed a morpholino to knock down the *lepb* gene functionality in the early development of the zebrafish embryo. Preliminary results suggested an induced bacterial burden in the *lepb* morphants, which indicates that *lepb* could have a role upon a *M. marinum* infection in zebrafish larvae. This prompted us to perform mass spectrometry on these *lepb* morphants with and without a *M. marinum* infection versus wild type controls. We found metabolites associated with the wasting syndrome (R. Marín-Juez, unpublished results) to be no longer influenced by infection in the leptin knockdown situation. The results therefore showed that leptin is an important player in the wasting syndrome caused by infection. This could indicate that there is cross regulation between nutritional status of the host and the immune response against *M. marinum*. This could also be a possible explanation of acute weight loss due to TB, since appetite is supressed by high leptin levels, due to infection. At later stages of infection by TB bacteria one could speculate that leptin resistance in the hypothalamus might occur, leading to a complex situation where a return of appetite in combination with wasting might lead to a new stage of tuberculosis disease progression.
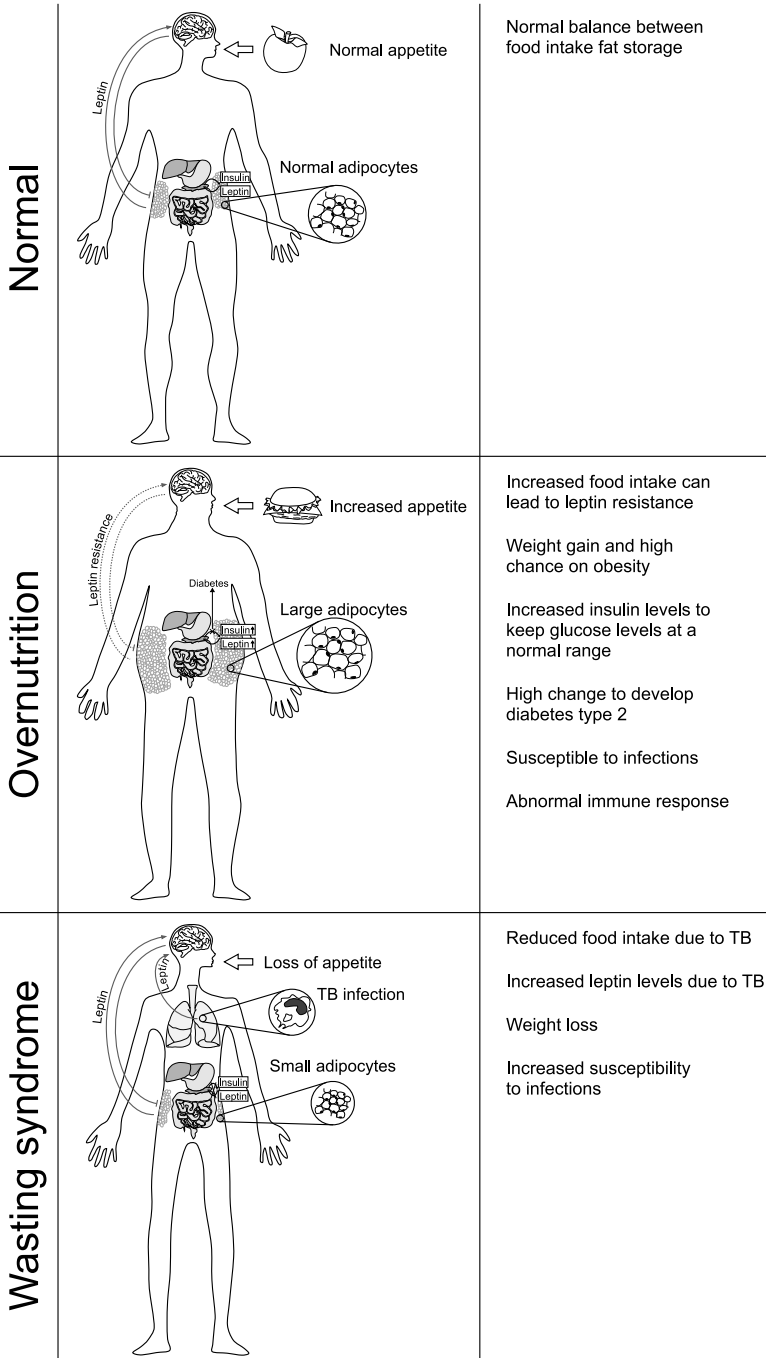
Figure 2: Functions of the leptin gene. Under normal conditions leptin regulates appetite and prevents from overeating. When a person overeats, the adipocytes will increase and the chance of leptin resistance occurs. This leads to uncontrolled appetite and chances of developing diabetes type 2. Leptin is also increased in TB infected people, which could suppress appetite with weight loss as result. This figure is based on Amitani et al. Front Neureosci. 2013 and Matarese et al. J Immunol. 2005.

Biomaterial-associated infection

As described above, we developed many different injection and implantation techniques for biomaterial-associated infection in zebrafish larvae (**chapter 6**). However, reproducibly implanting biomaterials into zebrafish embryos appeared to be more difficult than expected. Injection of polystyrene beads into the tail muscle proved to be rather labour intensive, whereas injection into the yolk did not lead to high frequencies of distributed biomaterials. We did find a size dependent distribution of beads ranging from 70 nm to 15 μm. The smaller sized beads did distribute more than the larger ones. We have not found the mechanism yet that explains the distribution of these particles, but hope to elucidate this in future work. The implantation of beads could in combination with the infection model of *S. epidermidis* in zebrafish as described in chapter 3, 4 and 5, be an important addition to the already known mammalian research models on biomaterial-associated infection.

Conclusion

Infectious diseases are everywhere around us, and we need to keep improving our knowledge about our defence mechanisms and the evasion strategies of the pathogens. The work in this thesis describes multiple techniques that can contribute to fast screening methods in order to come up with new strategies against infectious diseases. The use of automated micro-injectors in combination with large flow cytometers and automated microscopy has shown added value (**chapters 2, 3 & 4**) for research questions about (opportunistic) pathogens. The collaboration between scientists of different research areas has proved to be very successful in the development of an easy to use analysis platform for the analysis of RNAseq data (**chapter 5**). This has led to very detailed description of host molecular expression patterns following infection by these pathogens. This could be used to gain more insight in how biomaterials behave in a host environment in the presence or absence of infection (**chapter 6**). All together this has led this to a variety of research methods that can be used for studies of infections caused by many bacteria such as *S. epidermidis* and *M. marinum* described in this thesis, but, also by other microbes, such as fungi.

7

Chapter 8

Nederlandse sammenvatting

Met alle ontwikkelingen in de laatste paar decennia is de levensverwachting van de mens ook erg vooruit gegaan. De verwachting is dat in het jaar 2050 tweeëntwintig procent van de wereldbevolking ouder is dan 60 jaar. Dit heeft natuurlijk een groot effect op de gezondheidszorg, gezien deze bevolkingsgroep op een moment in het leven medische zorg nodig zal hebben als gevolg van kanker, diabetes, osteoporose, hartziekten en veel andere vormen van ouderdom gerelateerde ziekten. Om de medische zorg die men nodig heeft te waarborgen, moeten we zorgen voor een goede ontwikkeling van medische instrumenten, zowel wat betreft functionaliteit als als biocompatibiliteit. Op dit gebied kunnen nog altijd grote verbeteringen plaatsvinden. Het gebruik van het juiste onderzoeksmodel is hierbij erg belangrijk. Geeft de data van cel- of weefselkweek genoeg informatie of schiet het tekort aan diepte in biologische context, welke je wel kan vinden in diermodellen? En welk model geeft de humane situatie het beste weer? De laboratoria en faciliteiten spelen hierbij ook een belangrijke rol, aangezien sommige pathogene organismen onderzocht kunnen worden in een biologisch veiligheidslaboratorium van niveau 1, daar waar andere alleen in biologisch veiligheidslaboratoria van niveau 3 of 4 onderzocht kunnen worden.

Ontwikkeling en optimalisatie van high-throughput-technieken voor zebravisonderzoek

Het werk in dit proefschrift is in het bijzonder gericht op het gebruik van embryo's en larven van de zebravis als model voor verschillende infectieziekten. Een belangrijk voordeel en het onderwerp van dit proefschrift is het feit dat de embryo's en larven ideaal zijn voor high-throughput screening. Zo hebben we bepaald welk ontwikkelingsstadium het beste is voor injecties in de dooier van het embryo met twee verschillende soorten bacteriën, namelijk *Staphylococcus epidermidis* en *Mycobacterium marinum*. We hebben de resultaten vergeleken met andere injectiemethoden zoals injectie via de ader in de staart van een embryo van 1 dag oud. We hebben aangetoond dat injectie van *S. epidermidis* en *M. marinum* in de dooier van embryo's tussen het ontwikkelingsstadia van 16 tot 128 cellen optimaal is. Er vindt dan in latere ontwikkelingsstadia verspreiding van de bacteriën naar andere weefsels plaats. Injectie in een vroeger stadium leidt tot hoge mortaliteit en in een later stadium leidt ook niet tot een representatieve infectie. De meer virulente *M. marinum* M stam kon het beste voor het stadium van 64 cellen worden geïnjecteerd (**hoofdstuk 3**)

Aangezien de vroege ontwikkelingsstadia van zebraviseitjes ideaal zijn om automatisch te injecteren in de dooier, hebben we dit uitgevoerd met RNA, DNA, bacteriën of kankercellen (**hoofdstuk 2**). De automatische micro-injector is namelijk in staat om consistent 2000 eitjes per uur te injecteren. Deze aantallen injecties zijn niet meer te evenaren door middel van handmatige injecties. Door het gebruik van dit soort automatisering zijn we in staat om meer monsters te produceren voor bijvoorbeeld antibioticatesten of andere biologische vraagstellingen in minder tijd. Dit houdt in dat waardevolle onderzoekstijd bespaard kan worden. Echter, het onderzoeken van grote aantallen monsters vraagt ook om geavanceerde analysemethoden zoals flow-cytometrie. Wij hebben gebruik gemaakt van een extra groot model flow-cytometer

(COPAS XL) dat in staat is om zebravislarven te analyseren en sorteren (2000/30min), zonder ze daarbij te beschadigen. Daarom kan deze analyse dagelijks herhaald worden om het fluorescente signaal van de aanwezige bacteriën in deze larven te meten (**hoofdstuk 4**). Er is echter wel één probleem bij flow-cytometrie, en dat is de lage resolutie. Dit probleem hebben we opgelost door gebruik te maken van een geautomatiseerde screeningstechniek (VAST BioImager). Daardoor zijn we in staat om larven automatisch te laden in een glazen capillair en te oriënteren in de gewenste hoek, om deze vervolgens te analyseren met een confocale laser scanning microscoop met extra kleurencamera (**hoofdstuk 3**).

Door het gebruik van deze high-throughput methode hebben we een beter begrip gekregen van de pathogenese van *S. epidermidis* in zebravislarven. We hebben hiervoor microscopie, flow-cytometrie en transcriptoomanalyse uitgevoerd. Door het gebruik van de COPAS XL hebben een snelle analyse weten uit te voeren om de hoeveelheid bacteriën in de larven te bepalen, zonder de larven hiervoor te moeten opofferen.

Transcriptoomanalyse

We hebben door gebruik te maken van micro-arrays een basis gelegd voor analyse van transcriptoomveranderingen na een infectie in het zebravisembryo. Deze methode werd al snel vervangen door RNA deep sequencing. Micro-array maakt gebruik van gelabelde probes voor de detectie, wat betekent dat als een bepaalde probe niet aanwezig is ook de expressie van dit bepaalde gen niet gevonden zal kunnen worden. RNAseq daarentegen telt het aantal sequenties van verschillende genen gebaseerd op een referentiegenoom. Om die reden kunnen er ook onbekende genen gevonden worden, en de analyse kan altijd herhaald worden met betere en nieuwere versies van een referentiegenoom.

Wij hebben de keuze gemaakt voor Illumina sequencing, dat gebaseerd is op analyse van expressie van cDNA dat de novo gerepliceerd wordt op de flow cell. Echter, er zijn ook nog andere opties beschikbaar voor de expressieanalyse van RNA, zoals Ion Torrent sequencing, dat ook gebruik maakt van cDNA in plaats van RNA, maar dan wel op een semiconductorchip. Met deze methode worden kleine fragmenten gebonden op specifieke beads waarvan waterstofmoleculen worden los gelaten wanneer een nucleotide bindt. Dit leidt tot een verandering van de pH, wat door een voltagemeter gemeten kan worden. Deze veranderingen in voltage kunnen dus omgezet worden in een sequentie. Een andere optie in de toekomst zou de directe methode van RNA-sequencing met behulp van de MinION kunnen zijn. Deze methode is ook gebaseerd op flow cells met nanopores die individuele moleculen kunnen meten. Dit apparaat heeft slechts de grootte van een USB-stick en wordt al gebruikt voor de sequencing van DNA maar kan theoretisch ook gebruikt worden voor RNA of eiwitten. De nanoporetechnologie is nog altijd in ontwikkeling, maar biedt vooralsnog goede vooruitzichten voor goedkopere sequencingmethoden.

8

Een nadeel van RNAseq is de complexiteit van de analyse van de data. Het vereist namelijk nogal wat programmeerkunde, wat niet altijd beschikbaar is in biologieinstituten. Een oplossing zou kunnen zijn om het door bedrijven te laten uitvoeren, maar hier hangt dan wel een prijs aan. Om die reden hebben wij samen met bio-informatici en statistici een platform ontwikkeld voor het snel en gemakkelijk analyseren dan RNAseq-data. Het resultaat is een softwarepakket genaamd GeneTiles (**hoofdstuk 5**), dat alle programma's zoals Bowtie2, Samtools, 'R' statistiek programma, DESeq, DEXSeq, HTSeq en pysam combineert in één programma. Aangezien GeneTiles op een server draait, kan het bereikt worden van alle computers met verschillende besturingssoftware zoals Windows of Linux met een internetconnectie. Eén van de grote voordelen van het feit dat GeneTiles op een server draait is dat daardoor geen snelle en dure computers in het laboratorium meer nodig zijn. Om een zo compleet mogelijk pakket te hebben voor de analyse zijn er extra opties toegevoegd voor directe visualisatie van differentiële expressie gebaseerd op ratio, *P*-waarde, gecorrigeerde *P*-waarde, differentiële splitsing of gewone chromosomen. Tevens hebben we alle beschikbare biologische pathways van Wikipathways toegevoegd. Daardoor kan expressiedata direct gevisualiseerd worden in complexe netwerken en data sneller en beter geïnterpreteerd worden. Het resultaat van deze nieuwe methode is de vondst van een gen dat onder infectieuze condities differentieel gesplitst kan worden, namelijk glucagon a (*gcga*).

Metabole veranderingen als gevolg van infectie

Een andere opmerkelijke vondst is de hoge expressie van het gen voor leptine b (*lepb*) in larven geïnfecteerd met *S. epidermidis* en *M. marinum*. Dit was niet eerder ontdekt omdat er geen probe van *lepb* aanwezig was op de micro-array. Het leptine gen wordt onder normale condities geproduceerd door vetcellen, die de vethuishouding reguleren in het lichaam. Wanneer een persoon eet, wordt een signaal afgegeven naar de hersenen om aan te geven wanneer je vol zit. Echter, wanneer een persoon continu te veel eet, kan dit tot leptineresistentie leiden. Dit heeft tot gevolg dat er geen signaal meer is dat je waarschuwt om te stoppen met eten. Hierdoor eten men meer dan nodig en zal daarbij in gewicht aankomen met obesitas als resultaat. Dit kan dan leiden tot een continue productie van te hoge waarden van insuline, met insulineresistentie als gevolg. Insulineresistentie is 1 van de kenmerken van diabetes type 2 dat kan leiden tot chronische problemen van het verlagen van het bloedsuikerspiegel, en een tekort aan glucose omzetting tot glycogeen. Als gevolg van deze continue hoge bloedsuikerspiegel is men meer vatbaar voor hartziekten, beroertes, blindheid en nieruitval.

Aangezien leptine al erg interessant was vanwege verschillende functies in het metabolisme en het immuunsysteem en omdat *lepb* het meest geïnduceerde gen was na infectie met *M. marinum* en *S. epidermidis*, hebben we een morpholino ontworpen om de functie van *lepb* te onderdrukken (knockdown) gedurende de vroege ontwikkeling van een zebravisembryo. De resultaten wezen op een verhoogde hoeveelheid *M. marinum*-bacteriën in de knockdown-embryo's. Dit zou kunnen betekenen dat *lepb* een belangrijke rol heeft in het infectieproces van *M. marinum* in zebravislarven.

Daarom hebben we massaspectrometrie uitgevoerd op knockdown-embryo's en wild type embryos's, beide met en zonder infectie. We vonden metabolieten geassocieerd met het wasting syndroom (R. Marín-Juez, ongepubliceerde data) niet langer meer beïnvloed na infectie in de leptine knockdown-embryo's. De resultaten geven aan dat leptine een belangrijke factor is gedurende het wasting syndroom als gevolg van een infectie. Dit kan een indicatie zijn dat een verband is tussen de voedingsbalans en de immuunrespons tegen *M. marinum*. Tevens zou dit een mogelijke verklaring zijn voor gewichtsverlies als gevolg van tuberculose, omdat de eetlust is onderdrukt door hoge concentraties leptine als gevolg van de infectie. In latere stadia van een tuberculose infectie, zou men kunnen speculeren dat leptineresistentie in de hypothalamus kan optreden, met als gevolg een complexe situatie waarbij de terugkeer van de eetlust in combinatie met het wasting syndroom, een nieuw stadium van de tuberculose infectie kan vormen.

Biomateriaal-geassocieerde infecties

Zoals eerder beschreven hebben we veel verschillende injectie- en implantatietechnieken ontwikkeld voor biomateriaal-geassocieerde infecties. Echter het reproduceerbaar implanteren van biomaterialen in zebravislarven bleek moeilijker dan gedacht (**hoofdstuk 6**). Injectie van polystyreenpartikels in de staartspier bleek erg arbeidsintensief en injectie in de dooier resulteerde slechts in een geringe verspreiding door het lichaam. Desondanks hebben we aangetoond dat de diameter van de biomaterialen een effect heeft op de verspreiding. De kleinere partikels verspreiden namelijk meer dan de grotere partikels. Het mechanisme dat de verspreiding kan verklaren hebben we nog niet ontdekt, maar hopen we op korte termijn te kunnen ontrafelen. De implantatie van biomaterialen in combinatie met de gedetailleerde pathogenese van *S. epidermidis* zoals beschreven in hoofdstuk 3, 4 en 5, zou een perfecte toevoeging kunnen zijn voor reeds bekende zoogdiermodellen voor biomateriaal-geassocieerde infecties.

Conclusie

Infectieziekten zijn overal om ons heen en daarom moeten we onze kennis over virulentiefactoren en afweermechanismen blijven ontwikkelen. Het werk beschreven in dit proefschrift gaat over verschillende technieken die bij kunnen dragen aan snelle screeningsmethoden met het doel om nieuwe strategieën te ontwikkelen tegen infectieziekten. Geautomatiseerde micro-injectoren in combinatie met grote flow-cytometers hebben hun toegevoegde waarde laten zien (**hoofdstuk 2, 3 & 4**) voor onderzoeksvragen met betrekking tot (opportunistische) pathogenen. Samenwerkingsverbanden tussen verschillende onderzoeksgroepen bleken erg succesvol met als resultaat de ontwikkeling van een gemakkelijk te gebruiken analyseplatform voor RNAseq-data (**hoofdstuk 5**). Dit had veel gedetailleerde expressieprofielen van deze pathogenen in de gastheer tot resultaat. Deze aanpak kan ook meer inzicht geven in hoe biomaterialen zich gedragen in een gastheeromgeving

8

in aanwezigheid of afwezigheid van een infectie (**hoofdstuk 6**). Alles bij elkaar heeft dit onderzoek geresulteerd in verschillende onderzoekmodellen voor de studie van infecties veroorzaakt door bacteriën zoals *S. epidermidis* en *M. marinum* beschreven in dit proefschrift, maar ook voor andere microben zoals schimmels of gisten.

8

List of Publications

Carvalho R, de Sonneville J, Stockhammer OW, Savage ND, Veneman WJ, Ottenhoff TH, Dirks RP, Meijer AH, Spaink HP: A high-throughput screen for tuberculosis progression. PLoS One, 2011

Veneman WJ, Stockhammer OW, de Boer L, Zaat SA, Meijer AH, Spaink HP: A zebrafish high-throughput screening system used for *Staphylococcus epidermidis* infection marker discovery. BMC Genomics, 2013 (chapter 4)

Spaink HP, Cui C, Wiweger MI, Jansen HJ, Veneman WJ, Marin-Juez R, de Sonneville J, Ordas A, Torraca V, van der Ent W, Leenders WP, Meijer AH, Snaar-Jagalska BE, Dirks RP: Robotic injection of zebrafish embryos for high-throughput screening in disease models. Methods, 2013 (chapter 2)

Kanwal Z, Wiegertjes GF, Veneman WJ, Meijer AH, Spaink HP: Comparative studies of Toll-like receptor signalling using zebrafish. Developmental and Comparative Immunology, 2014

Veneman WJ, Marín-Juez R, de Sonneville J, Ordas A, Jong-Raadsen S, Meijer AH, Spaink HP: Establishment and optimization of a high-throughput setup to study Staphylococcus epidermis and *Mycobacterium marinum* infection as a model for drug discovery. Journal of Visualized Experiments, 2014 (chapter 3)

Veneman WJ, de Sonneville J, van der Kolk KJ, Ordas A, Al-Ars Z, Meijer AH, Spaink HP: Analysis of RNAseq datasets from a comparative infectious disease zebrafish model using GeneTiles bioinformatics. Immunogenetics, 2014 (chapter 5)

Veneman WJ, Vijver MG, Meijer AH, Spaink HP: Distribution of micro and nano range sized biomaterials after yolk injection into zebrafish larvae. Manuscript in preparation (chapter 6)

van der Ent W, Veneman WJ, Groenewoud A, Chen L, Tulotta C, Hogendoorn PCW, Spaink HP, Snaar-Jagalska BE: Automation and manipulations for cancer discovery. Submitted for publication

Wiweger MI, Veneman WJ, Dirks RP, Spaink HP: "No big fish" - zebrafish as model organism for studies on skeletal tuberculosis. Manuscript in preparation

Curriculum Vitae

Wouter Jurjen Veneman was born on February 11th 1987 in Almelo, the Netherlands. He finished high school at the S.G. St. Canisius in Almelo. In 2003 he started the Medical Microbiology laboratory study at the ROC Oost Nederland in Hengelo, the Netherlands, during which he performed an internship at the Medical Microbiology department of the Medical Center Haaglanden hospital in The Hague under the supervision of Mr. Johan Mutsears and Drs. Casper Jansen. In 2007 he started his bachelor of applied science in Biology and Medical laboratory research in Leiden, the Netherlands during which he performed two internships. One at the pathology department of the Leiden University Medical Center under the supervision of Dr. Ronald van Eijk and Prof. Dr. Hans Morreau and another shared internship between the Medical Microbiology department of the Academic Medical Center Amsterdam, the Netherland and the Molecular Cell biology department of the Institute of Biology of the University of Leiden, the Netherlands under the supervision of Dr. Bas Zaat and Prof. Dr. Herman Spaink. Directly after he started working at ZF-screens B.V. in Leiden, the Netherlands together with Dr. Ralph Carvalho and Dr. Oliver Stockhammer. In September 2010 he started his PhD research in the group of Prof. Dr. Herman Spaink and Dr. Annemarie Meijer. During his PhD he focused on the development of a biomaterial-associated infection model using zebrafish larvae. He established a model to study the pathogenesis of *Staphylococcus epidermidis* in zebrafish larvae and the molecular host response towards this kind of infections. In 2014 while finishing his PhD project, he also started working for the Institute of Environmental Sciences (CML) of the University of Leiden, the Netherlands in the ecotoxicology group of Dr. Martina Vijver to develop and optimize a fast screening method of nanoparticles toxicity using zebrafish larvae.