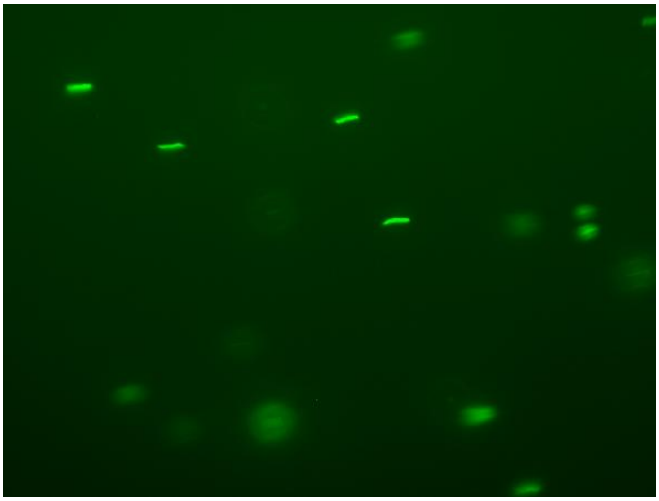
SLU

The role and genetic control of non-coding RNAs in bacterial infection



Jonas Åkerblom

Degree project 30 hp

Uppsala 2010

Department of Molecular Biology

SLU, Sveriges Lantbruksuniversitet
Department of Molecular Biology (MOLBIO)

Author:
Jonas Åkerblom

Title:
The role and genetic control of small non-coding RNAs in bacterial infection
Den genetiska kontrollen av små icke-kodande RNA-molekyler och deras roll inom bakteriell infektion

Keywords:
Dictyostelium discoideum, non-coding RNAs, RNAi, Legionella pneumophila, bacterial infection, miRNAs, siRNAs, promoter

Supervisor:
Fredrik Söderbom (SLU, MOLBIO)

Examinator:
Jerry Ståhlberg (SLU, MOLBIO)

Course description:
Independent project in biology E
Självständigt arbete i biologi E, 30hp (EX0565)

Publication:
Uppsala, January 2010

Abstract

Promoters are genetic elements that facilitate the transcription of a gene and they have been found in front of non-coding RNA (ncRNA) genes in different organisms, e.g. the model plant *Arabidopsis thaliana*. A similar element, DUSE, has been found in front of ncRNAs in the social amoeba *Dictyostelium discoideum* and a part of this project has been to analyze the function of this putative promoter element through cloning and expression studies. A construct to analyze the function of DUSE was successfully designed and introduced into *D. discoideum* but full expression studies were not finished because of shortage of time. However, the first preliminary tests with northern blot showed a distinct loss in the expression of a model gene when the putative promoter was rendered non-functional by site-directed mutagenesis, indicating that DUSE truly functions as a promoter.

The main part of this project has been to set up an infection system for *Legionella pneumophila* utilizing the social amoeba *Dictyostelium discoideum* as a host organism. A big focus of the infection studies has been to study the RNA interference response to infection by utilizing two *D. discoideum* knock out strains and SOLiD sequencing. Deep sequencing using the SOLiD technique has been carried out on infected- and uninfected cells to compare the difference in the small RNA population and to analyze the response to infection.

The infection system was successfully set up with good confirmation of intracellular bacteria by two different screening methods. The infection efficiency showed to differ substantially between the tested strains AX4 and AX2, 10,0% to 3,4% respectively. The protein Dicer B has shown to play a big role in the biogenesis of miRNA and a *D. discoideum* strain with this gene knocked out showed to be more resistant to infection, having an infection efficiency of only 1.2%. At the same time, the growth of this knock out strain was inhibited when grown in proximity to *L. pneumophila*. This could indicate that there are miRNA(s) that are involved in the amoeba's response to bacterial infection.

SOLiD sequencing showed that the non-coding Class I RNA gene, *DdR-31*,was the gene most affected in infected cells compared to uninfected cells also indicating the role for a non-coding RNA in infection.

Populärvetenskaplig sammanfattning

Det här projektet har eftersträvat att studera hur icke-kodande RNA-molekyler regleras i den encelliga amöban *Dictyostelium discoideum* samt att utforska dessa molekylers roll vid bakteriell infektion. RNA-molekyler delas idag in i två olika grupper, mRNA (budbärar-RNA) som avläses till ett funktionellt protein och icke-kodande RNA som har specifika funktioner på egen hand, så som tRNA (transport-RNA) och rRNA (ribosomalt-RNA). Man har hittills funnit att icke-kodande RNA är inblandade i nedbrytandet av mitokondriellt rRNA i *D. discoideum* när dessa infekteras av *Legionella pneumophila.* Tidigare har man även funnit att icke-kodande RNA har en roll i kontrollen av gener vid infektion i andra organismer, så som modellväxten *Arabidopsis thaliana*. Vad som studerats i detta projekt är en klass icke-kodande RNA som är unika för *D. discoideum*, Class I RNA, med ännu okänd funktion och deras möjliga roll i infektionsresponsen. Inom denna klass finns genen *DdR-21* som använts som modellgen i detta projekt.

En del av infektionsanalysen har varit att studera hur RNA-interferensmaskineriet, inaktivering av gener via dubbelsträngat RNA, påverkas av infektionen. Detta har gjorts via användandet av *D. discoideum* med specifikt avstängda gener samt en väldigt omfattande sekvenseringsteknik, SOLiD. För att studera genregleringen av icke-kodande RNA har funktionen av en sekvens som är en tänkbar promotor, DUSE, analyserats. Dessa element har sen tidigare bara bekräftats i andra organismer och här för första gången visas preliminära bevis för en faktisk promotor framför icke-kodande RNA i *D. discoideum*.

D. discoideum används som modellorganism för att det finns många molekylära och genetiska verktyg tillgängliga och för att det patogena gensvaret är väldigt likt det i makrofager hos djur. Därmed kan infektionsprocessen studeras på ett mycket enklare sätt. Anledningen till att *L. pneumophila*, som orsakar lungsjukdomen legionärssjukan, och att andra likartade bakterier är patogena för människor är just denna likhet mellan bakteriernas naturliga värdorganismer och makrofagerna. Detta innebär att bakterien kan använda sig av samma mekanismer vid infektion.

För att studera DUSE så konstruerades en utomgenomisk plasmid innehållande två versioner av modellgenen *DdR-21* med omkringliggande sekvens. Den ena versionen gjordes med en normalt fungerande DUSE medan den andra versionen gjordes med en muterad och förhoppningsvis ickefungerande DUSE. För att

skilja dessa två modellgener från det normala genuttrycket av *DdR-21* så ändrades även en loop-sekvens i genen, och därmed RNAt, på två olika sätt i de olika versionerna. För att studera skillnaden mellan de båda versionerna gjordes en Northern blot där man bestämmer mängden av ett visst RNA. En första preliminär Northern blot visade att mängden RNA var lägre då den troliga promotorn, DUSE, var söndermuterad jämfört med när den var hel och fungerade normalt. Detta pekar på att detta element faktiskt har en stor påverkan på uttrycket av modellgenen och därmed fungerar som en promotor.

Att få till en hög reproducerbar infektionseffektivitet visade sig vara komplicerat då *L. pneumophila* utvecklar sin rörlighet, som gör att den kan infektera, vid en viss celldensitet. Då denna celldensitet visade sig variera kraftigt från experiment till experiment kunde inget samband påvisas för att standardisera infektionen. Genom att öka koncentrationen bakterier relativt koncentrationen av *D. discoideum* vid infektionen gick det dock att få till en kvantifierbar infektionseffektivitet och ett system som kan användas för framtida tester med *L. pneumophila* och andra intressanta bakterier. Vid test av RNA-interferensmaskineriet upptäcktes att när man slog ut genen som kodar för proteinet Dicer B, som aktiverar det dubbelsträngade RNAt via klyvning, i *D. discoideum* så blev infektionseffektiviteten lägre. Det kunde även fastställas att *D. discoideum* utan Dicer B hade sämre tillväxt än normala D. discoideum när den växte i närhet till *L. pneumophila*.

SOLiD sekvenseringen visade att det sker en omfattande upp- och nedreglering av icke-kodande RNA när *D. discoideum* blir infekterade av *L. pneumophila*. Den gen där störst förändring sker är *DdR-31,* ett Class I RNA precis som modellgenen *DdR-21*. Detta kan vara en indikation på att denna gen och kanske klassen i sig har en funktion i gensvaret på infektion. En annan sak som kunde observeras var att storleksfördelningen på RNA-molekylerna i infekterade gentemot ickeinfekterade celler var förändrad. En större mängd av kortare RNA kunde hittas i infekterade celler, vilket tyder på att det kan ha skett en klyvning och trolig inaktivering av en hel del RNA vid infektion.

Det verkar som att icke-kodande RNA kan ha en funktion inom gensvaret vid infektion av bakterier och en fortsatt analys av den stora mängd data som SOLiD analysen gav kan förhoppningsvis styrka detta faktum. En studie där man stänger av *DdR-31,* genen som påverkades mest av infektion, i *D. discoideum* och sen gör infektionsstudier på den kan också vara ett bra framtida test. När det gäller genregleringen och analysen av funktionen av DUSE så kommer ett framtida test med Northern blot med den färdiga utomgenomiska plasmiden ge definitiva besked om det verkligen är en promotor eller inte.

Abbreviations

A	Adenine
Amp (r)	Ampicillin resistence
APS	Ammonium persulfate
ATP	Adenosinetriphosphate
AYE	ACES buffered yeast extract
bp	Basepair(s)
BSA	Bovine serum albumine
C	Cytosine
cAMP	cyclic adenosine monophosphate
CYE	Charcoal yeast extract
dil	Dilution
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DUSE	Dictyostelium upstream sequence element
e.g.	Exempli gratia
EDTA	ethylenediaminetetraacetic acid
et al.	And others
G	Guanine
GFP	Green fluorescent protein
i.e.	Id est
IPTG	Isopropyl-D-thiogalactopyranoside
kb	Kilobase(s)
LB	Luria Bertani (medium)
Mb	Megabase(s)
MCS	Multiple cloning site
miRNA	micro RNA
MOI	Multiples of infection
mRNA	messenger RNA
ncRNA	Non-coding RNA

nt	Nucleotide(s)
OD	Optical density
oligo	Oligonucleotide
Ori	Origin of replication
P	Phosphate(s)
PCR	Polymerase chain reaction
PNK	Polynucleotide kinase
RdRp	RNA-dependent RNA polymerase
REMI	Random enzyme mediated mutagenesis
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNase	Ribonucleas
rpm	Rotations per minute
rRNA	ribosomal RNA
SDS	Sodium dodecryl sulfate
siRNA	small interfering RNA
SM	Standard medium
snoRNA	Small nucleolar RNA
snRNP	Small nuclear RNP
SOLiD	Sequencing by Oligonucleotide Ligation and Detection
SRP	Signal recognition particle
SSC	Sodium chloride-sodium citrate buffer
T	Thymine
TBE	tris-borate-EDTA
TEMED	N,N,N',N'-Tetramethylethylenediamine
Tm	Melting temperature
tRNA	transfer RNA
UV	Ultraviolet light

Table of contents

1 Introduction

The aim of this project has been divided into two parts where one has been to set up a system for infection of *Dictyostelium discoideum* with *Legionella pneumophila* and to analyze how the RNA interference (RNAi) machinery is influenced by infection. We have also started to analyze how the small RNA population is affected by infection by the use of deep sequencing, i.e. SOLiD sequencing where we can compare the difference in the levels of primarily small RNAs in infected and uninfected cells. The other part has been to study a ncRNA named *DdR-21*, with unknown function, to see if it might be involved in the infectious process. It has previously been shown that small ncRNAs are involved in infection in other organisms and in this part of the project we wanted to see if longer ncRNAs might have a role in the response to infection in *D. discoideum*. The primary focus of this second part has been to analyze a putative promoter element, DUSE, for *DdR-21*, which if confirmed would be the first promoter element found that control ncRNA gene expression in *D. discoideum*. This promoter element will then be used to tinker with the expression of the *DdR-21* gene which could give information about its function.

1.1 Dictyostelium discoideum

D. discoideum is a unicellular eukaryotic social amoea, commonly known as a slime mold, and a member of the Mycetozoa phylum [1]. The mycetozoa is a branch of the eukarotyic evolutionary tree placed before both the fungi and the metazoan branches. *D. discoideum* as a species was discovered in 1935 as a soil-living organism living among decayed forest material feeding on different bacteria prominent in those habitats [2,3].

The haploid genome was fully sequenced in 2005 and it is 34 Mb in size and built up of six chromsomes containing approximately 10300 protein coding genes

[4,5]. An important aspect of the *D. discoideum* genome is the high A/T content at 78 % and that it has a high composition of simple sequence repeats and transposable elements [3,5]. Another interesting and useful characteristic of the genome is that it contains several extra chromosomal plasmids which among others contain the ribosomal genes. Some of these plasmids have been modified and are used in *D. discoideum* research as transformation vectors [5].

There are many features of *D. discoideum* that make it a useful and interesting model organism for many research areas. A number of cell biological, molecular genetic and biochemical techniques are available for studies of *D. discoideum* and the haploidity of the genome together with the possibility of RNA interference (RNAi) make the creation of mutants and gene manipulation easy [1]. Much that is possible to do in the famous yeast model organism *Saccharomyces cerevisiae* can also be performed in *D. discoideum* but because of its intriguing life cycle, *D. discoideum* can also be used to study the development of multicellularity and cell differentiation. Other areas often studied are phagocytosis, cell signaling, cell migration and morphogenesis [6].

1.1.1 Life cycle

D. discoidium has a very special life cycle. Apart from its normal unicellular stadium, where it goes through normal vegetative cell division, *D. discoidium* also has a multicellular phase which it enters upon starvation. When there is a lack of nutrition and starvation commences there is a switch in the life style of *D. discoideum* and a set of genes involved in chemotaxis for cAMP are turned on. Cells under starvation develop the ability to produce, excrete and recognize cAMP in the vecinity and respond to it [3]. Under normal conditions this process leads to entrance into the asexual multicellular life cycle, the formation of spores within fruiting bodies, figure 1.

Around 12 hours after initiation of multicellularity a structure known as a mound is formed and from there there are two ways to go; the aggregate can either go into the slug structure or the finger structure. The slug is the only development phase when the multicellular *D. discoideum* is motile. This moving structure is essential if the environmental conditions are not favourable at the site the cells currently are in. The slug is able to move to a better environment, by the use of light and heat, and there it continues development through the finger structure [3,7]. The cells in the finger structure continue to differentiate to become stalk- and spore cells which make up the mature fruiting body. After approximately 24 hours the fruiting body is fully developed, containing dead stalk cells surrounded

by a shell of cellulose and a ball of spores that will stay dormant until conditions are better and germination is induced [3].

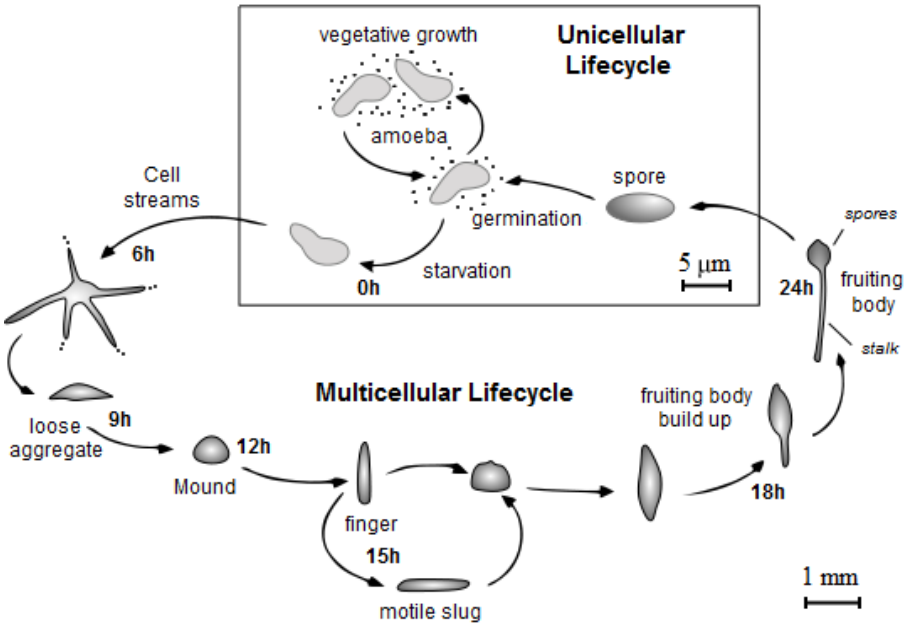


Figure 1. Schematic view of the life cycle of *Dictyostelium discoideum* where the different multicellular stages and the time it takes for the organism to reach them are visualized. Under normal conditions it stays in the unicellular stages of life.

Under certain conditions, submersion and darkness, an alternative sexual development pathway is available. This is initiated by the fusion of two cells of different mating types. The fused cells start to excrete cAMP which attracts other *D. discoideum* cells. The big fusion cell, the zygote, engulfs the attracted amoebas and uses them as nutrients. This structure known as a macrocyst, can subsequently produce meiotic offspring [3,8].

1.1.2 As a host for infection

The use of more simple organisms as hosts for bacterial infection is of great importance because the basic cell biology is evolutionary conserved between the non-vertebrates and the vertebrates. By studying extensively used model organisms, e.g. *Drosophila melanogaster*, *Caenorhabditis elegans* and *D. discoideum*, where many molecular and genomic tools are available, the pathogenic response of higher eukaroytes can be established by much simpler methods [9,10]. *D. discoideum* has several advantages in phagocyte-pathogen interaction studies, e.g. the

cells are easily infected and there are many genetic tools available, such as random mutagenesis by restriction enzyme mediated integration (REMI) and RNAi, making mutant screening and host mutant analysis easy to perform [10]. Other features that make *D. discoideum* a useful host is that it can easily be grown on a bacterial carpet in petri dishes or without bacteria in solution and it have uncomplicated demands for nutrients and environmental conditions [11].

A very intriguing and important trait of *D.discoideum* is that it shares many similarities with human macrophages, such as phagocytosis and the way of killing bacteria. Because of the fact that *D. discoideum* shares environment with many pathogenic bacteria it is believed that because these bacteria have developed ways to infect the amoeba it has at the same time by coincedence gained the ability to infect the similar human macrophages and thereby becoming opportunistic human pathogens [12]. The protozoa can be seen as environmental incubators that propel the bacteria to adapt to avoid being killed by similar functioning macrophages. The similarities between human macrophages and amoeba are plentyful and besides having similar killing mechanisms, bacteria like *L. pneumophila* are internalized and grow in a very analogous ways in both cell types [13].

There seems to be a vast number of bacteria and other microorganisms that are equally pathogenic for humans and amoeba like *D. discoideum*. The first infection studies made with *D. discoideum* was with *L. pneumophila* [14] and up until today many more bacteria have been found suitable to use in similar studies, table 1 [13].

Table 1. *Bacteria that have been used in infection studies with D. discoideum.*

Bacteria	Reference
Cryptococcus neoformas	Steenbergen *et al*., 2001 [15]
Klebsiella pneumoniae	Benghezal et al., 2006 [16]
Legionella pneumophilia	Hägele et al., 2000 [14]
Mycobacterium avium	Skriwan *et al*., 2002 [17]
Mycobacterium marinum	Solomon et al., 2003 [18]
Mycobacterium tuberculosis	Hagedorn *et al*., 2009 [19]
Neisseria meningitidis	Colucci *et al*., 2008 [20]
Pseudomonas aeruginosa	Pukatzi *et al*., 2001 [21]
Salmonella typhimurium	Skriwan *et al*., 2002 [17]
Vibrio cholerae	Pukatzi *et al*., 2005 [22]
Yersinia pseudotuberculosis	Vlahou *et al*., 2009 [23]

1.2 Legionella pneumophila

L. pneumophila was first discovered in 1976 as the causative agent of a clinical outbreak of a respiratory disease named Legionnaires' disease [24]. *L. pneumophila* are gram-negative rod-structured obligate aerobic bacteria with a size variation between 0.3-0.9 µm in width and 2-20 µm in length [25]. They are found in both man-made and natural fresh water environments where they live as parasites within protozoan hosts. Even though they are almost exclusively found in aquatic environments they have also been found in protozoa from moist soil samples [24].

L. pneumophila is only one out of 52 species and 72 sero groups within the *Legionellaceae* family, which fits in the evolutionary tree within the phylum of proteobacteria [11,26]. Some species of *Legionella* are able to infect human macrophages and are therefore regarded as human pathogens although most of the species are harmless saphrophytes, feeding on dead organic material [27]. The bacteria are slow-growing and the temperature interval for growth is 20°C to 45°C but survival is possible at temperatures close to 70°C for a short period of time [25,26].

The genome of *L. pneumophila* consist of one single circular choromosmome and up to today three different isolates, Philadelphia, Paris and Lens, have been fully sequenced. The sequencing showed that the average G/C content is 38% and that the total genome size varies between 3.3-3.5 Mbp and contains approximately 3000 genes [28,29]. Around 60% of the genes in *L.pneumophila* have homologues in other known bacteria with the closest known relative being *Coxiella burnetii*, an intracellular pathogen. Hence, approximately 40% of the genes are genus specific [28].

1.2.1 Life cycle and host interaction

L. pneumophila goes through two distinct phases in life, a transmission phase and a replication phase, where much of the morphology of the bacteria changes which reflects its life as a parasite. *L. pneumophila* that are not living within a host protozoa are believed to be part of biofilms, an aggregation of cooperating microogranisms, which they use as protection against predators and to obtain necessary nutrients [30]. It has been shown that these biofilms are only used for survival and no evidence has been found that growth can occur in them [31,32]. *L. pneumophila* are not forming biofilms on their own but instead they attach to already existing biofilms created by other bacteria which they then colonize and take over.

These biofilms are a perfect food source for protozoa that engulf the bacteria by normal phagocytosis through an actin-mediated process [32,33]. *L. pneumophila* that are in its transmissive phase are able to survive within the protozoa while the ones in the replicative phase are rapidly killed and degraded. The transmissive bacteria that survive and resist degradation by the protozoan lysosome do so by forming a protective membrane-bound vacuole inside the host-cell [27]. There are different secretion systems encoded by the bacteria that are responsible for its ability to survive and replicate within the vacuole of protozoa and macrophages. The most important ones are the Dot/Icm type IVB secretion system and the Lvh type IVA secretion system [27]. When the protective vacuole is fully established the bacteria go into a replicative phase during which they repress the genes involved in transmission [30,34].

The internalized bacteria affect the host *D. discoidium* in several ways by just staying within their replication vacuole. It has been shown that several genes involved in metabolism are upregulated in the amoeba upon infection, which is probably done to supply the bacteria with the vital nutrients for survival. Genes that are downregulated are instead genes involved in bacterial degradation, fatty acid modification and protein biosynthesis [35].

When nutrition levels are low the bacteria yet again change between the replicative phase and the transmissive phase and environmental signals causes blockage of the RNA-binding repressor protein CsrA (carbon storage regulator A) by the signal transduction system, LetA/LetS [36]. This leads to the induction of several genes and the most important gene is the flaA, coding for the protein flagellin, which is the building stone of the flagellum. The flagellum gives the bacteria its motility making it highly infectious. Furthermore, the *flaA* gene is also involved in mediating the other important transmissive traits, lysosome avoidance and the ability to induce host cell death [37]. The resilient and infectious bacteria can now upon cell lysis start the life cycle all over again by either invading a new host or becoming a part of a biofilm.

An important difference between protozoa and macrophages as hosts for *L. pneumophila* is that the bacteria is able to continue its life cycle after infection and lysis of protozoa but so far this have not been seen in humans. This together with the fact that the temperature of 37°C within macrophages is not an optimal temperatur for *L. pneumophila* neither for growth nor for virulence which indicates that macrophages are not a natural host for the bacteria [38].

1.2.2 Pathogenesis

L. pneumophila is the causative agent of two serious human respiratory diseases, Legionnaires' disease and Pontiac fever, both affecting mostly adults but also to a lesser extent children and infants [26]. The legionella bacteria are transmitted to humans through inhalation of aerosols that contain bacteria which infect the alveolar macrophages in the patient's lungs. Common man-made sources for infection are water fountains, water misters, cooling towers etc. and increased prolifiration comes with elevated temperatures in these sources [25,26].

Legionnaires' disease is one of the most common causes of pneumonia and it is often fatal if not diagnosed and treated [25,27]. The disease is much more common for elderly people since it usually affects people with underlying disease or respiratory problems. Pontiac fever is the milder and less studied disease caused by the legionella bacteria. It is a non-fatal disease that has similar symptoms to influenza and in contrast to Legionnaires' disease it does not cause pneumonia [26,39].

1.3 Non-coding RNAs

RNAs have for a long time been divided into two major groups, mRNAs which are coding for proteins and non-coding RNAs (ncRNAs) like tRNAs, rRNAs and small RNAs, which have a distinct function on their own. The ncRNAs can operate on RNA stability, transcription and protein transport among others and they are involved in many different processes in the cells ranging from apoptosis to chromosome maintenance. ncRNAs can be found in a large size range with the smallest known being 20 nt while the biggest found can be several thousand nucleotides long [40].

1.3.1 ncRNAs and infection

Studies on the transcriptional response in the host cell *D. discoideum* upon infection with *L. pneumophila* have showed that many protein coding genes are being up- and downregulated but so far almost nothing is known about the role of the ncRNAs in the response to infection[35]. The one thing that has been shown to occur upon infection with *L.pneumophila* is that the mitochondrial rRNA of *D. discoideum* is specifically cleaved, indicating that the bacteria hinder the mitochondrial protein synthesis [41].

In addition it has been shown that ncRNAs play a big role in the pathogen *L. pneumophila* upon infection of the amoeba *Acanthamoeba castellanii*. Two

ncRNAs named RsmY and RsmZ are linking the two regulatory networks involving the two-component system LetA/LetS and the repressor CsrA together. The transcription of these ncRNAs is activated by the LetA/LetS system and they have the ability to bind to the repressor CsrA and thereby block its repressive activity. When CsrA no longer represses the LetA/LetS system the bacterium is able to enter its transmissive phase and infect its host [42]. From this it is obvious that ncRNAs are very important for the bacteria in its ability to infect host cells and the question is then if there are equally important ncRNAs in the host organism involved in the response to infection.

In a study on the plant model organism *Arabidopsis thaliana* it was found that a class of ncRNAs with a size of 30-40 nt are induced when the plant is infected with the bacteria *Pseudomonas syringae*. They showed that one of these ncRNAs is involved in disease resistance by silencing its target gene [43]. Previousely it had been shown that a miRNA in *A. thaliana* was induced upon stimulation with a bacterial peptide from the flagellin protein. This induction caused a repression of auxin signals needed for the internal growth of the bacteria *P. syringae* [44]. These reports both show that in plants ncRNAs can play very important roles in the fight against pathogens and it naturally leads to the question if this occurs in other organisms outside the plant kingdom, e.g. in *D. discoideum*.

1.3.2 ncRNAs and Class I RNAs in *D. discoideum*

Most of the ncRNAs that have been found in *D. discoideum* have homologues in other organisms, rRNAs, tRNAs, spliceosomal RNAs, small nucleolar (sno)RNAs, signal recognition particle (SRP) RNAs, RNase P RNAs, antisense RNAs and small interfering (si)RNAs. The spliceosomal RNAs are found within the nucleus and together with specific proteins they form snRNPs, ribonucleoproteins, which identify and cut out introns from pre-mRNAs. There are 17 expressed spliceosomal RNAs identified in *D. discoideum* and they fold into predicted conserved structures also found in many other organisms [40]. SnoRNAs are involved in modifying other RNAs either by cleaving or by adding a chemical modification to the RNA.

The two main groups of snoRNAs are also present in *D. discoideum*, the C/D and the box H/ACA snoRNAs. The box C/D snoRNAs are modifying RNAs by methylation and the box H/ACA modifies RNAs by pseudouridylation [40]. Two other large groups of ncRNAs found in *D. discoideum* that also exist in other organisms are SRP RNAs and RNase P RNA. SRP RNAs are involved in protein localization while the RNase P RNAs are involved in the processing of tRNAs.

The antisense RNAs and siRNAs are involved in the regulation of transcriptional and post-transcriptional gene expression [40].

In addition to all the common ncRNAs, some additional unique groups of ncRNAs have been found in *D. discoideum*, i.e. D1/Dd8, dutA, msRNA, Class I and Class II RNAs [45,46]. In this project *DdR-21* is used as a model gene and it is a member of the Class I RNA genes. So far 14 uniqe Class I RNAs and 2 unique Class II RNAs have been identified and because of the structure and sequence similarities between them, the Class II RNAs have been considered to be a subclass of the Class I RNAs. The Class I RNAs are long cytosolic ncRNAs with a size distribution between 55-65 nt forming a stem-loop structure with conserved 5´- and 3´sequences, figure 2 [45] (Avesson and Söderbom, unpublished). Until today no homologues to these ncRNAs have been found in any other organism besides *D. discoideum*. An interesting finding is that two of the Class I genes, *DdR-21* and *DdR-32* most probably are transcribed as precursors which later are cleaved into their correct active form [45].

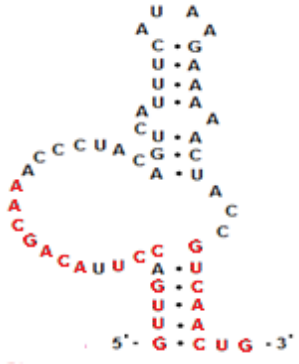


Figure 2. The structure of the representative Class I RNA *DdR-21* (adapeted from Hinas and Söderbom, 2007). As can be seen in the figure the RNAs folds into a stem-loop structure where the red nucleotides are conserved in all the known members of the class.

The function of the Class I RNAs is still not known. However, it has been shown that the expression levels are decreased during the course of the multicellular development. Furthermore, disrupting one of the Class I RNA genes, *DdR-21*, affects early development. Taken together, these results indicate a function in the developmental process. The stretch of 11 nt that is shared between almost all the members of the Class I RNAs and also by the Class II RNAs could be a possible interaction site with proteins but this have not yet been confirmed [45] (Avesson and Söderbom, unpublished).

1.3.3 Promotors of ncRNAs

Promoters are short stretches of nucleotides facilitating transciption of specific genes. The promoters are situated upstream of the target gene and the RNA polymerase utilize it to locate the gene and attaches to it and from there continue downstream to initiate transcription [47]. A putative promoter sequence of 8 nucleotides named DUSE, Dictyostelium upstream sequence element, has been found positioned approximately 63 nt upstream of the transcription start site of the majority of the different classes of ncRNAs ≥ 50 nt in *D. discoideum* [48].

The first upstream promoter element found in front of ncRNA genes was identified in front of U-RNA genes, involved in the splicing machinery, in *A. thaliana*. The element is situated approximately 70 nt upstream of the start site of transcription of the studied U5 RNA gene [49]. Promoters of ncRNA genes have also been confirmed in non-plants in both the nematode *C. elegans* and in the fruit fly *D. melanogaster*, table 2. In *C. elegans*, these promoter sequences are known as UM1-3 and they are built up of a number of different consensus sequences of 7-11 nt. UM1 is the most frequent and found in front of approximately 10% of the ncRNA so far identified in *C. elegans*. It consists of a stretch of 50 nt containing two conserved core sequences of 21 nt known as the PSEA and PSEB elements [50].

Table 2. *Comparison of the putative promoter sequence from D. discoideum to the confirmed promoter sequences in two other famous model organisms.*

Organism	Sequence	Distance between promoter and transcription start site	Reference
D. discoideum	ATCCCACTAA	63 bp	Hinas *et al*., 2006
A. thaliana	ATCCCACATCG	70 bp	Vankan *et al*., 1988
C. elegans	GCGGAACCCG – 5bp - TGTCGGCCGC	30 bp	Li *et al*., 2008

The fact that there are ncRNA promotors in other organisms and the high similarity between the putative elements found in *D. discoideum* and the confirmed promoter in *A. thaliana* strongly suggests that DUSE in fact is a promoter element.

1.3.4 The RNAi machinery

In this project two different *D. discoideum* strains were used, AX4 and AX2, and two different AX2 knock out strains were also used in the experiments, RdpC$^-$ (RNA-dependent RNA-polyemerase) and DicerB$^-$. Both the knocked out genes are important components of the RNAi machinery and play important roles in the defence of the organism.

The mechanism of RNAi was first reported in 1998 in the nematode *C. elegans* as a system where double stranded RNA cause silencing of specific genes, figure 3 [51]. The RNAi machinery has now been found in diverse eukaryotes, from plants and animals to fungi and it is believed that a simple RNAi system existed in the last common ancestor to all the eukaroytes. The RNAi machinery has been proposed to have evolved as a defense system against viruses and selfish genetic elements. Even though the RNAi machinery is a complex system there are four components that are of specific importance, a Dicer protein, a RNA-dependent RNA polymerase, a Piwi-like protein and an Argonaute-like protein [52]. The process of RNAi starts with dsRNAs in the cytoplasm where they are are recognized and cleaved by the Dicer protein (a RNA specific endonuclease) into small siRNAs, 21-25 nucleotides. The source of dsRNAs can either be endogenous pre-mRNAs or exogenous coming from viral infections [53].

After the generation of siRNAs by Dicer there are two routes to take. The first is the generation of even more siRNAs to get an even better silencing effect. This happens when the siRNAs binds to the complementary mRNAs and act as primers for a RNA-dependent RNA polymerase. The long dsRNAs created in this way are then cleaved by Dicer, leading to a bigger pool of double stranded siRNAs [52]. The other route is to directly bind to the RNA-induced silencing complex (RISC). Only one strand of the siRNA, the guide strand, binds to the RISC complex while the other strand, the anti-guide strand, is degraded during RISC activation. The RISC complex is built up of endonucleic proteins of the Argonaute-Piwi family and becomes active when the siRNA has bound. The active RISC binds to the complementary mRNA guided by the bound single stranded siRNA and the argonaute protein then cleaves the target mRNA, causing silencing of the gene [52,53].

RNAi has been confirmed to exist in *D. discoideum* where siRNAs in the size of 21-23 nt are thought to be responsible for the silencing effect. The RNAi machinery in *D.discoideum* seems to have a partial capacity because at some developmental stages no RNAi effect at all can be seen [54].

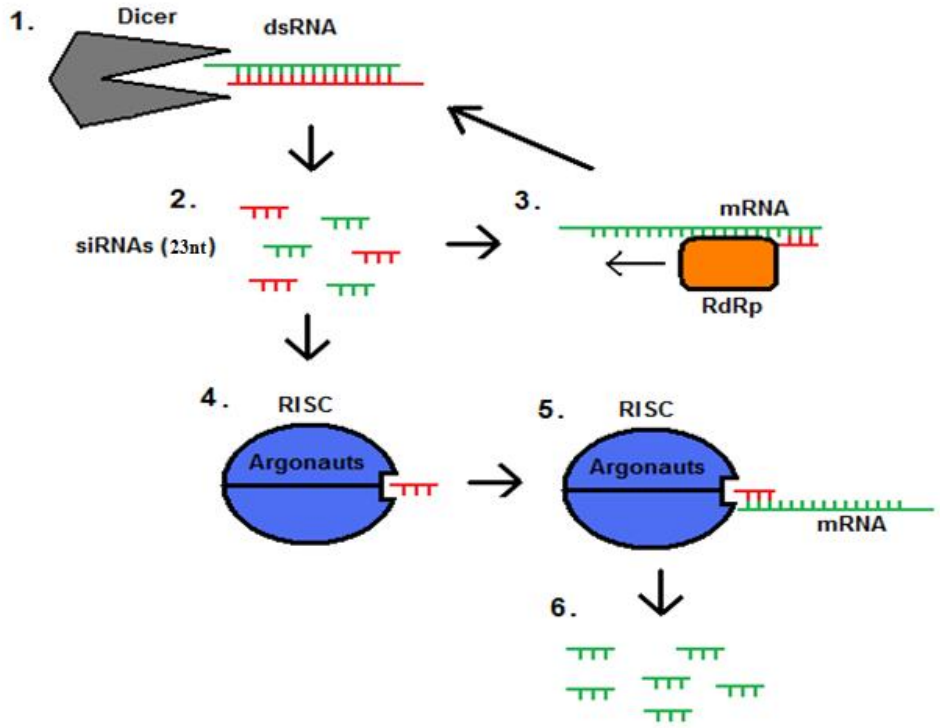


Figure 3. Scehmatic view of the RNAi mechanism. Dicer cleaves the dsRNA generating siRNAs which either are incorporated into RISC or works as primers on the mRNAs for RdRp to generate more dsRNA, amplifying the silencing signals. siRNAs bound to RISC then binds to the complementary mRNA so that the active RISC endonuclease components, Argonaute proteins, can cleave the mRNA and cause silencing of the gene.

2 Material and methods

2.1 Growing *D. discoideum*

Four different *D. discoideum* strains were used during the project, AX2, AX4 (Boston), AX2RdpC$^-$ and AX2DrnB$^-$ which all were available in -80°C storage at the department of moleular biology, SLU. To retrieve them the frozen *D. discoideum* were spread on SM-plates [55] together with *Klebsiella pneumoniae* and incubated at 22°C until plaques appeared, normally 4-6 days. To get *D. discoideum* growing in liquid culture, growing cells from the outer part of the plaques were scraped off and inoculated in 2 ml of a rich axenic medium, HL5 [55], together with PenStrep, dil 1:100, (GIBCO) in sterile round-bottom glass tubes.

For continousely growing *D. discoideum* the 2 ml culture were then transfered to a 100 ml E flask and diluted to a total volume of 25 ml HL5 with PenStrep and kept at 22°C on a shaking table at 155 rpm. The cells could be kept growing for approximately a month, if the growth medium was continusely changed so that the concentration of cells was kept below $4x10^6$ cells/ml, before new cells had to be restarted. To calculate the concentraton of cells in liquid culture a hemocytometer (Bürker chamber) was used to count the cells using a light microscope (Zeiss Axioskop).

2.2 Plasmids and oligonucleotides

All the oligonucleotides used through the project are listed in table 6 (appendix). They were all customly made and ordered from Invitrogen (Desalted, scale of synthesis 25 nmol). The T_m values of the primers were calculated from the following formula: Tm = 64.9°C + 41°C x (number of G's and C's in the primer – 16.4)/N.

Two different plasmids have been used in the project, the pCR 2.1 TOPO (Invitrogen) as a cloning vector and the pDXA-HC as a transformation vector, figure 19 (appendix). The TOPO cloning system utalizes the fact that a normal PCR with *Taq* polymerase adds a deoxyadenosin on the 3´end of the PCR products. The TOPO-vector is constructed as a linearized plasmid with a deoxythymidine at the 3' ends together with a covalently bond topoisomerase. This makes it very easy to ligate the PCR product with the A overhang into the vector with the complementary T overhang with the aid of the topoisomerase, which cleaves and rejoins complementary strands of DNA [56]. The pCR 2.1 TOPO vector contains two important selection inserts, kanamycin and ampicilin resistence, which are used to select for and isolate bacteria that contain the plasmid.

The second plasmid pDXA-HC is designed to be transformed into *D. discoideum*, containing a origin of replication (Ori) for a high copy number plasmid, a multiple cloning site (MCS) and selecion inserts Amp (R) and Tn5 neo (R) conveying resistence for Ampicilin and G418 respectively [57]. In the pDXA-HC vector used in this project the actin promoter insert has been removed to prevent the expression of the inserted *DdR-21* gene from this promoter. pDXA-HC will refer to this modified plasmid from here on.

2.3 Site-directed mutagenesis of *DdR-21*

The work with the promoter analysis have been a continuation of a project started before this thesis and not everything written in this section have been done during the time of this master thesis. For easier understanding of the project, all methods used to reach the final double construct are included in this report.

2.3.1 Extraction of genomic DNA and genomic PCR

To extract the genome 1-2·10^8 *D. discoideum* cells were harvested by centrifugation at 300 g for 5 min at 4°. Subsequently, supernatant was poured off and the pellet were resuspended in 1,5 ml nucleic buffer (40 mM Tris pH 7,8, 1,5% sucrose, 0,1 mM EDTA, 6 mM $MgCl_2$, 50 mM KCl, 5 mM DTT, 0.4% NP-40). The tube was then incubated on ice for 5 min which was followed by a centrifugation at 13000 rpm for 5 min at 4°C. The steps after the first centrifugation was then repeated twice before the pellet was resuspended in 20 ml 0,5 M EDTA and dH_2O to a total volume of 100 ml. 100 ml 10% sodium lauryl sulphate (sarcosine) was then added to the tube, which then was gently mixed and incubated at 55°C for 15

min. After that 250 ml 4M NH4Ac was added and was followed by centrifugation at 13000g for 15 min at 4°C. The supernatant was then transferred to a fresh 1,5 ml eppendorf tube and precipitated by adding 1 ml ice cold 99% EtOH. The tube was then centrifuged at 13000 g for 10 min which was followed by a wash with 1.5 ml cold 70% EtOH. The pellet was briefly air-dried and resuspended in a suitable volume of 10 mM Tris-Cl pH 8.5. Finally10 mg/ml of RNase A was added to the DNA solution, followed by incubation at 37°C for 1 hour.

The *DdR-21* gene with surrounding sequence was isolated from the genome by PCR, the Stratagene Robocycler® Gradient 96 were used for all PCRs. Two sets of primers (489/490 + 491/492) were used to introduce two different pairs of restriction sites, Sal I/Kpn I + Sac I/Xba I, flanking the *DdR-21* insert. PCR was carried out by using 2,5 µl of each primer together with, 1 µl genomic DNA, 5 µl Pfu reaction buffer (10x), 1 µl Pfu Turbo polymerase (2,5 units/ µl), 1 µl dNTP solution (10mM of each nucleotide) and dH_2O to a final volume of 50 µl. The PCR cycling conditions were 95°C for 2 min, 35 cycles of 95°C for 30 s, 55°C for 30 sec, 62°C for 2 min, followed by a final extension of 62°C for 10 min.

2.3.2 Agarose gel electrophoresis

Agarose gel electrophoresis was the method used during the project to analyze the size and purity of the DNA or RNA of interest. The agarose gels were made with concentrations variying between 0.8-2% depending on the size of the fragments to analyze. TBE (0,5x) was used to make the gels and as running buffer. To be able to visualize the DNA fragments ethidiumbromide (0.5 µg/ml) was added to the gels before they were casted. The gels were run at 100V and 60mA. A UV-table with a connected camera (GelPhotoSystem GFS1000) visualized the fragments in the gel. As size markers a, 1Kb ladder or a 100 bp ladder (Fermentas) were used depending on the size of the fragments.

2.3.3 TOPO TA cloning of PCR products and plasmid purification

To get the PCR products into the competent DH5α strain of Escherichia coli the TOPO TA Cloning® kit (Invitrogen) was used with the pCR®2.1 TOPO® vector. This would create two TOPO plasmids, named v001 (*DdR-21* amplified with oligos 489/490) and v002 (*DdR-21* amplified with oligos 491/492), containing the wild type *DdR-21* with surrounding nucleotides but flanked by different restriction sites. The cloning was performed according to the manufacturer´s recommendations with the difference that in the procedure of addition of A overhangs a 2x

Master Mix containing 0.05 units/µl Taq DNA Polymerase in reaction buffer, 4 mM $MgCl_2$, 0.4 mM dNTPs (Fermentas) was used instead of seperate solutions for H_2O, buffer, dNTPs and polymerase. The TOPO cloning mixture was incubated at 30 min to get efficient cloning before transformation into DH5α. Transformation was done by heat-shock where 40 µl of the competent cells were mixed with 2 µl of TOPO cloning mixture before incubation in a 42°C watherbath for 30 sec and then i put back on ice. After the heat-shock, the cells were imideatly put back on ice. 10 µl of the mixture were mixed with 200 µl of LB [58] and then spread on selective LB + ampicillin (50 ug/ml) plates and incubated at 37°C over night.

For isolation of more plasmides the transformed colonies were grown in 2 ml LB + ampicillin (50 ug/ml) in a shaking incubator (LAB-LINE) at 250 rpm and 37°C before plasmid purification were performed. The plasmid purifications in this project were performed with one of three different kits, QIAprep Spin Miniprep Kit or QIAGEN Plasmid Midi Kit both from Qiagen or Fermenta´s The GeneJET™ Plasmid Miniprep Kit. In all instances the protocols supplied with the kits from the manufacturer were followed when performing plasmid preparations. The only difference was the amount of starting material where 2 ml bacterial cultures were used for the miniprep kits and 100 ml cultures for the midikit. A NanoDrop1000 spectrophotometer (Thermo Scientific) was used to analyze the concentration of DNA or RNA.

2.3.4 Sequencing of DNA

Sequencing was performed to analyze the constructed plasmids in order to verify that they contained the correct inserts. Sequencing was done both at Uppsala Genome Center and at Macrogen Inc. The sequencing files were analyzed by the program Chromas Lite 2.01 and the online program ClustalW2 was used for sequence alignments.

2.3.5 PCR-based mutagenesis

To introduce the mutations in the promoter and loop regions oligonucleotides were designed for site directed mutagenesis by PCR. In order to optimize the mutagenic PCR many different annealing temperatures were tried before the mutations were succesfull incorporated. Included here are only the experimental conditions which worked succesfully. One of the inserts contains mutations both at the promoter and loop region so here the mutagenesis was done in two steps. First the v001 plasmid, described in 2.2.3, is used as template to insert the promoter mutation, creating the

v011 plasmid. The v011 plasmid is then used as a template for the second loop mutation, creating the v211 plasmid containing the first cassette. For the first loop mutation the v002 plasmid, described in 2.2.3, is used as a template, creating the v102 plasmid containing the second cassette.

To introduce the promoter mutation primers 496 and 497 was used and to get the two different loop mutations forward primers 517 and 528 was used together with reverse primer 419. The PCRs were performed by using 1 µl of each primer together with, 1 µl plasmid template (v001 for the promoter mutation, v002 for the loop 1 mutation and v011 for the loop 2 mutation), 5 µl Pfu reaction buffer (10x), 1 µl Pfu Turbo polymerase (2,5 units/ µl), 1 µl dNTP solution (10mM of each nucleotide) and dH_2O to a final volume of 50 µl. The PCR cycling conditions were 95°C for 2 min, 14 cycles of 95°C for 30 s, 55°C for 30 sec for the promoter mutation and 52°C for the loop mutations, 72°C for 5 min.

The PCR-based mutation reactions were Dpn I treated in order to destroy the original fully methylated template DNA and leave the PCR products intact. Agarose gel electrophoresis was performed to verify that the template DNA had been degraded. Transformation and plasmid preparation of the Dpn I treated PCR products were then done as in the TOPO cloning. Further analysis to confirm the mutated plasmids was performed by analytical PCR. For the analytical PCRs 1 µl template was mixed with 1 µl primer (forward primers 498 and 489 for the promoter mutation and loop mutations respectively and reverse primers 489 and 493/494 for the promoter and the loop mutations 1 and 2 respectively), 12,5 µl 2 x PCR Master Mix and 9,5 µl dH_2O. The PCR cycling conditions were 95°C for 5 min, 30 cycles of 95°C for 30 sec, 55°C for 30 sec, 60°C for 1 min and a final extension at 60°C for 10 min. The analytical PCR was followed by agarose gel electrophores to visualize the results. Finally sequencing was performed as a definitive confirmation.

2.4 Transformation into *D. discoideum*

2.4.1 Gel extraction and ligation into the pDXA-HC vector

FastDigest® restriction enzymes (Fermentas) matching the flanking restriction sites, Sal I + Kpn I and Sac I + Xba I, were used to cleave out the confirmed inserts from 10 µg of the TOPO-plasmids. The protocoll included from the manufacturer for double cleavage were followed. 10 µg of the pDXA-HC vector was also cleaved with the same restriction enzymes as the construct to be inserted in it

and 2 µl FastAP, Thermosensitive Alkaline Phosphatase, (Fermentas) was added to remove phosphates and prevent self-ligation of the vector. All the cut out inserts were loaded on a 1.2% agarose gel and the linearized pDXA-HC vector were loaded on a 0.8% agarose gel. Fragments of the correct sizes were excised by a sterile scalpel blade and used for gel extraction. Gel extraction was performed with the Fermentas GeneJET™ Gel Extraction Kit and it was done following the manufacturer´s instructions.

T4 DNA Ligase (Fermentas) was used for ligation of the extracted inserts into the vector; the molratio 3:1 of vector to insert was used. Because of low yields from the gel extraction only 50 ng of vector was used and the amount of insert was adjusted accordingly. The ligation reaction was otherwise performed after the protocol supplied from the manufacturer and the reactions were incubated over night at 22°C. To achieve the double construct with both the cassettes into one pDXA-HC plasmid the process of cleaveage and ligation was done a second time but this time with the other pair of restriction enzymes.

2.4.2 *D. discoideum* transformation by electroporation

The *D. discoideum* AX3 strain was used for transformation at a cell density of 1-4·10^6 cells/ml. 10^7 cells were aliqouted in 12 ml falcon tubes and centrifuged at 300 g at 4°C for 4 min. The supernatant was discarded and the pellets resuspended in cold Zap buffer [59] to 10^7 cells/ml. 800 ml of the resuspended cells were mixed with 10 mg of the transformation construct and added to a chilled electroporation cuvette (0,4 cm gap, Bio-Rad) and incubated on ice for a few minutes. The electroporator (Bio-Rad GenePulser) was set to 3 mF and 1KV (2,5 KV/cm) and the cells were zaped with a time constant of 0.9-1.4 being acceptable. The electroporated cells were then transfered to three petri dishes (15 cm) filled with 10 ml HL5 + PenStrep (10µl/ml) medium, i.e 100 µl, 200 µl and 500 µl respectively. As controls, plates with electroporated cells with no DNA were included.

The pDXA-HC vector contains the Tn5 neo (R) gene which confere G418 resistence which is used for selection. G418 selection was performed by first incubating the electroporated cells in HL5 + PenStrep (10µl/ml) for 24 hours before the medium is changed to include 5 µg/ml G418 and incubated for three days at 22°C. The medium was then changed to 10 µg/ml G418for yet another three days before adding medium without G418 for three days to improve the viability of cells. The last step of the selection is to again add medium with 10 µg/ml G418 each third day until colonies are claerly visible on the plates.

To analyze the visible colonies and get larger cultures the colonies were transferred to a 24-well plate where they were kept growing before single colonies were isolated. To isolate single colonies a 200 µl pipette was used to suck the liquid up and down in the wells to make the cells detach from the bottom. 100 µl were then transferred to a 1,5 ml eppendorf tube and the cell density was counted in a hemocytometer. Approximately 50 cells are spread together with 250 µl *K. pneumoniae*, which was grown in SM-medium, on a SM-plate containing 0,5 M MES, 40 ml/L (SMM-plate). The plates are then incubated at 22°C for 4-6 days until large plaques appear which can be analyzed by PCR.

To get a good template for the PCR cells are scraped of the plaques and resuspended in 10 µl dH_2O to which 25 ml LysB buffer (50 mM KCl, 10 mM Tris pH 8,3, 2,5 mM $MgCl_2$, 0,45% Nonidet P40 and 0,45% Tween 20) and 1µl proteinase K was added. The mixture is then incubatedt at 95°C for 5 min to inactivate the proteinase K before it is ready to be used as template for PCR. The analytical PCR is done as described in 2.3.5 and the primers used depend on which insert is in the plasmid that has been transformed into the cells. For the first insert primer 489 was used as forward primer and primers 499 and 528 as reverse primers for confirmation of the promoter mutation and the loop mutation two respectively. For the second insert primer 489 was used as a forward prime and primer 518 as a reverse primer to confirm the loop mutation one. 25 ml cultures are started from the plaques with cells that contain confirmed correct plasmids as stated earlier. Transformed cells with all the different plamids were succesfully made and stored in -80°C to be later used for RNA extraction and expression studies.

2.5 RNA extraction from *D. discoideum*

Approximately 10^8 *D. discoideum* cells are collected in a 50 ml falcon tube and centrifuged in a swing-out rotor at 300 g for 5 min at 4°C. The pellets are resuspended in an equal volume as the culture volume of cold PDF (20 mM KCl, 5 mM $MgCl_2$, 13 mM KH_2PO_4, 7 mM K_2HPO_4, pH 6.2) and spun down again. The pellets are resuspended in 1ml TRIzol reagent (Invitrogen) in a 1,5 eppendorf tube, vortexed and inbubated for 5 min at room temperature. 200µl of chloroform is added and the mixture is vortexed and incubated for 3 min at room temperature. A 15 min centrifugation at 12000 rpm in a microcentrifuge (BIOFUGE pico Heraeus) was followed by a separation of the phenol/chloroform and aqueous phases. The RNA is in the aqueous phase so that phase is transfered to a new 1,5 ml eppendorf tube to which 500 µl room-tempered isopropanol is added followed by a

10 min incubation at room temperature. Precipitation of RNA by a 10 min centrifugation at 13000 rpm was then followed by discarding the supernatant and washing the RNA pellet with 1 ml room-tempered 70% ethanol. The wash was followed by a 5 min centrifugation at 13000 rpm and all ethanol were then discarded and the pellets were allowed to air-dry for a short time. The dry pellets were then disolved in RNase free dH_2O and the final RNA concentration were determined by NanoDrop. To analyze the integrity of the RNA a 1,5% agarose gel electrophores weas done and two sharp bands representing the 16S and 26S subunits of the ribosome were expected if the RNA were intact.

2.6 Expression analysis by northern blot

For expression analysis by northern blot 20 µg total RNA were first mixed 1:1 with 2 x RNA loading dye and denatured at 95°C and chilled on ice before they were loaded on the gel. A 10% polyacrylamide gel containing1xTBE and 7M Urea were used, 10% APS (1:100) and TEMED (1:1000) were added to initiate polymerization of the gel. The gel was casted to an approximate size of 20x20x0.1cm and prerun at 21W in 1xTBE for 1 hour. The prepared RNA samples were then added to the washed wells and run for approximately 2 hours for separation of the small RNAs. As size markers a DNA ladder ([γ-^{32}P]-ATP end-labeled pUC197MspI, AMbion) and a RNA ladder ([γ-^{32}P]-ATP end-labeled Decade marker) were used.

Next, electroblotting of the gel was done where the RNA was transfered to a Hybond-N+ (Amersham Biosciences) nylon membrane by running it in TBE (1x) buffer at 20V in 4°C over night in a Bio-Rad Trans-Blot Cell. To fix the RNA to the membranes UV-crosslinking was done at 150mJ in a Bio-Rad GS Gene Linker. For the hybridization step, 15µl of the oligonucleotides (15 pmol) used were radioactively labeled using 10µl [γ-^{32}P]-ATP (100µCi), 2,5 µl 10 x T4 PNK reaction buffer, 0,5µl T4 PNK (10 U) and 25µl dH_2O. The probes were ready to use after 1 hour incubation at 37°C. QIAquick nt removal (QIAGEN) was used to remove unincorporated nucleotides.

Before hybridization the membranes were incubated in Church buffer (1 % BSA, 1 mM EDTA, 0,5 M NaH_2PO_4 (pH 7,2), 7 & SDS) at 42°C for 1 hour and the oligos were denatured at 95°C for 5 min then put on ice for 5 min. The membranes were then soaked in approimately 30 ml of fresh 42°C Church buffer together with the denatured oligos in a hybridization bottle (Amersham Biosciences). The bottles are then put in a hybridization oven (Hybaid Mini Hybridization Oven) at 42°C over night. The next day the membranes are washed in 42°C

washing buffers of equal volumes as the hybridization buffer in the subsequent order: rinse with 2xSSC/0.1%SDS, 2x5min 2xSSC/0.1%SDS, 2x10 min 1xSSC/0.1%SDS, 2x 5min 0.5xSSC/0.1% SDS. The membranes were then sealed in a plastic hybridization bag and exposed in a phosphorimager (Molecular Dynamics PhosphorImager) and the expression results could then be easily analyzed.

2.7 Set up of a bacterial infection system of *D. discoideum*

2.7.1 Growing *L. pneumophila*

The *L. pneumophila* Philadelphia 1 (NC_002942) strain used in the project was a kind gift from Professor R. Isberg (Tufts University, Boston, USA). The strain contains a gene for streptomycin resistence used for selection and also contains a plasmid that express GFP from an isopropyl-D-thiogalactopyranoside (IPTG)-inducible promoter, GFP is used when looking at the bacteria in a fluorescense microscope [60]. The bacteria were succesfully grown both on agar plates and in solution. Charcoal yeast extract (CYE) agar supplied with ACES (N-(2-acetamido)-2-aminoethanesulfonic acid, Sigma) was the medium used for growth on agar plates and for growth in liquid medium the same medium but without charcoal (AYE-broth) was used [61]. Both on plates and in liquid culture the bacteria were grown at 37°C and for adequate growth in liquid culture a constant shaking at approximately 200 rpm was needed. *L. pneumophila* growing on plates could be kept and used for appoximately 14 days in 4°C before new had to be retrieved from the -80°C stock.

2.7.2 Infection of *D. discoideum*

The infection system that was set up in this project was adapted from existing protocols from other groups that perform succesful infections of *D. discoideum* [60,62]. Fresh *L. pneumophila* from CYE-plates were picked with a sterila toothpick and spread in 1x1cm patches on a freshly made CYE-plate and incubated at 37°C for 2 days. 18ml AYE-broth containing 5 µg/ml streptamycin antibiotic and 1mM IPTG was prepared and aliqouted to five 12 ml falcon tubes, 6 ml in the first and 3 ml in the rest. All bacteria from a patch is collected and added to the first tube, which then is vortexed for 30 sec to resuspend the bacteria. From the first tube 3 ml is then added to the second tube, which then is vortexed for 30 sec and the procedure is repeated until all five tubes contain 3 ml with different dilutions

of *L. pneuomophila* (1:1, 1:2, 1:4, 1:8 and 1:16). The tubes were then incubated at 37°C over night shaking at 200 rpm.

The next morning the cell density of the cultures were analyzed by measuring the OD_{600} where OD = 1 corresponds to approximately 10^9 cells. An OD of over 3 was aimed for since that is approximately when the bacteria starts to develop motility. However, variation between cultures was very high. The apperance and motility of the bacteria was analyzed by light microscopy (40x magnification). High percentage of short and fat cells with over 10% motile were the requirements for the culture to be used in the infection. The *D. discoideum* culture that were to be used in the infection were also analyzed so that the cell density were at the required level of $2\text{-}3\cdot10^6$ cells/ml before infection could be commenced. Approximately $1\cdot10^7$ cells of the *D. discoideum* culture was then collected at 300g for 5 min for AX4 cells and 1000 rpm for 2 min for AX2 cells. The pelleted cells were then washed twice with rich 1xMB medium (5 ml 10xMB (35 g yeast extract, 70 g thiotone, 0,5 L dH_2O), 5ml 400 mM MES (pH 6.9), 40 ml sterile H_2O) before the pelled cells were resuspended in rich 1xMB medium and diluted to a concentration of $8\cdot10^5$ cells/ml.

The infection was performed in a 24 well-plate where different infection conditions could be used in different wells. In the wells where screening by microscopy was to be done, round cover slips were added. Before infection could be started, 500 µl of the diluted *D. discoideum* were added to all wells, except in the control lane where only *L. pneumophila* was present, and let to settle for 2 hours. The *L. pneumophila* culture chosen for infection was diluted in rich 1xMB so that the addition of 20 µl gave the right totalt concentration, depending on the wanted MOI. 20µl of the diluted bacteria was added to all wells were infection were to take place. To the control wells with only *L. pneumophila* the 20 µl of bacteria solution were added together with 500 µl rich 1xMB. Infection was then initiated by centrifugation at 1000 rpm for 5 min before the plate were incubated at 25.5°C.

For removal of non-internalized *L. pneumophila* two different procedures were done. 2 hours after initiation of infection Gentamicin to a final concentration of 100 ug/ml were added to the wells before incubation were continued at 25.5°C. After an additional 2 hours the wells were washed 3 times with 500µl 1xMB to remove the gentamicin and to remove any non-internalized bacteria not killed by the antibiotic. Finally 500µl of 1xMB were added to all wells. In addition IPTG to a final concentration of 0,1mM was added to the wells to be screened by microscopy. Subsequently, the plate were incubated for the final 2 hours at 25,5°C.

2.7.3 Screening of infection

To screen for infected *D. discoideum* two different analytical methods were used, one focusing on quantfication of infection and the other on accurate confirmation of infection.

2.7.3.1 Confirmation of infection by microscopy

For analyzis of infection by microscopy round cover slips are placed in the wells of interest before infection is started in the 24-well plate. To be certain that the cover slips are sterile and un-contaminated, the 24-well plate with cover slips is put in UV-light for approximately 20 min before start of infection. After infection was completed the wells were aspirated of all solution and 500µl formaldehyd solution (3,7% formaldehyd in 1xPBS (137 mM NaCl, 2,7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4) was added, followed by 20 min incubaion at room temperature for fixating the *D. discoideum* cells. The formaldehyd solution was then apirated and the wells were immideatly washed with 1xPBS, which was repeated three times.

Next the coverslip with the fixated cells was mounted on a microscope cover slide. A drop of Fluoro Guard antifade Reagent (BioRad 170-3140) was added to the slide and the coverslip from the 24-well infection plate was put on top of it. Nail polish was then put all around the cover slip for fixation before the slide was put to dry in a dark drawer for approximately 1 hour. The slides were then stored in -20°C and ready to be analyzed by microscopy. A Zeiss Axioskope flourescent microscope was used for the microscopy analysis and the program AxioVision Rel. 4.8 was used for taking pictures of the infected *D. discoideum*.

2.7.3.2 Quantitative screening of infection

To get a quantitative screening of the infection a plating assay was performed. To the wells of interest 0,02% saponin was added to break the *D. discoideum* cells. To mix a 200µl pipett was used to suck up and down 10-15 times and all the solution in the well were put in a 1,5 ml eppendorf tube. 500µl dH_2O was then added to the well and mixed before the solution was put in the same tube, which then vortexed for 30 sec. The previous intracellular *L. pneumophila* could be quantified by spreading it on CYE plates. To be able to get quantifible data a dilution series could be necessesary depending on the MOI used in the infection. For a MOI of 10:1 no dilution was needed to get a countable amount of colonies on the plates. Of the sample, 10µl were added to 90µl sterile 0,9 % NaCl, which were then

spread on CYE plates containing Cm (5 µg/ml) and incubated at 37°C for 3-4 days. The plates were then analyzed by counting the number of colonies which then gave approximate results of the infection efficiency.

2.8 Bacterial predation of *D. discoideum*

This method was modified from Shevchuk and Steinert 2009 [62] and is used for screening of virulence traits in *L. pneumophila* and as an analysis of the *D. discoideum* susceptibility to infection. Different strains of *D. discoideum*, AX2, AX2DrnB$^-$ and AX2RdpC$^-$, were tested for growth together with only *K. pneumoniae* or with *L. pneumophila* and *K. pneumoniae*.

The *D. discoideum* strain of interest were grown to a cell density of 1-3·10^6 cells/ml and centrifuged at 2000 rpm for 2 min and resuspended and washed in an equal volume of 1xSørensen buffer (2.0 g KH_2PO_4 0.29 g Na_2HPO_4, pH 6.0). The pellet is then resuspended in infection medium (1:1 HL5 and 1xSørensen buffer) to a final cell density of 1·10^4 cells/ml. Both *K. pneumoniae* and *L. pneumophila* are suspended in 100µl dH_2O and centrifuged at 300g for 5 min and the pellets are then resuspended in 2 ml dH_2O. The bacterial cultures are then diluted to a final concentration of 1·10^9 cells/ml by measuring OD 550nm (OD 1 ≈ 1,31·10^9 cells/ml).

100µl of the *K. pneumoniae* suspension is mixed with 100µl of the *D. discoideum* suspension and 100µl dH_2O and spread on a SMM-plate. In addition, 100 µl of *K. pneumoniae* suspension and 100µl *D. discoideum* and 100 µl *L. pneumophila* is spread on a second SMM-plate. These plates are then incubated at 22°C for 3-5 days until plaques appear.

2.9 SOLiD analysis of infected *D. discoideum*

The RNA used from uninfected and infected cells for the SOLiD sequencing was a generous gift from Professor R. Isberg in the US. The same methods were used as in Li et al., [63] to separate and extract the RNA. The RNA was then sent to Uppsala Genome Sequencing that performed the preparation of small RNAs and the SOLiD sequencing using an ABI SOLiD 3 system (Applied Biosystems), which resulted in approximately 3 million sequences for each sample. The SOLiD data was then carefully collated and filtrated to divide the sequences into different libraries of interest.

3 Results

3.1 Overview

The first goal of this project was to set up an infection system based on existing protocols where the social amoeba *D. discoideum* is used as a host for *L. pneumophila*. Adaptation and optimization of the infection system used by other resarch groups have given us good insights in what was necessary to succesfully achieve infection. However, it has shown to be harder then expected to get high infection efficiency and a good detection system. The bioinformatic analysis of RNAs in infected and unifected cells presented here give a good first indication of what can be found in the upcoming extensive analysis of the huge amount of data that was recieved by SOLiD sequencing.

As for the analysis of the putative promoter element the plasmid with the double construct (see below) has been sucesfully created and expression studies was started but not finished during the scope of this project.

3.2 Functional investigation of DUSE

The function of Class I RNA genes, so far only identified in *D. discoideum*, is still unknown but earlier findings have demonstrated that ncRNAs are involved in infection in plants and bacteria which lead us to the tought of them playing a role in the respons to bacterial infection [43,44]. To get a first insight into this we used SOLiD sequencing which gave us expression data of RNAs in the size range of ~10-40 nt (including fragmented longer RNAs). By comparing the RNA populations derived from infected and non-infected *D. discoideum* cells, RNAs that were up- or down-regulated during infection could be identified. To further study this class of long ncRNAs we decided to investigate a potential promoter element,

DUSE, which has been found in front of the Class I RNAs genes. This putative promoter element could then be used to tinker with the expression of the genes and thereby give us a better insight into their function. For this to be possible the element had to be confirmed as a promoter element and that was one big part of this thesis. This project was previousely started by me as part of a research training project a few months before the master thesis started.

To study the putative promoter element, DUSE, the Class I ncRNA gene *DdR-21* was used as a model gene since it had been well studied in our lab. In addition to the upstream promoter region additional sequences surrounding the transcriptional unit of the gene were also included with the model gene. This was done since we could not be sure if there are additional regulatory sequences besides the DUSE that could affect the experiment if they were left out. To study the function of the promoter element our strategy was to create an extrachromosomal plasmid containing two modified versions of the *DdR-21* gene. This plasmid double contruct was then transformed into *D. discoideum* from which RNA could be extracted and expression analysis made by northern blot.

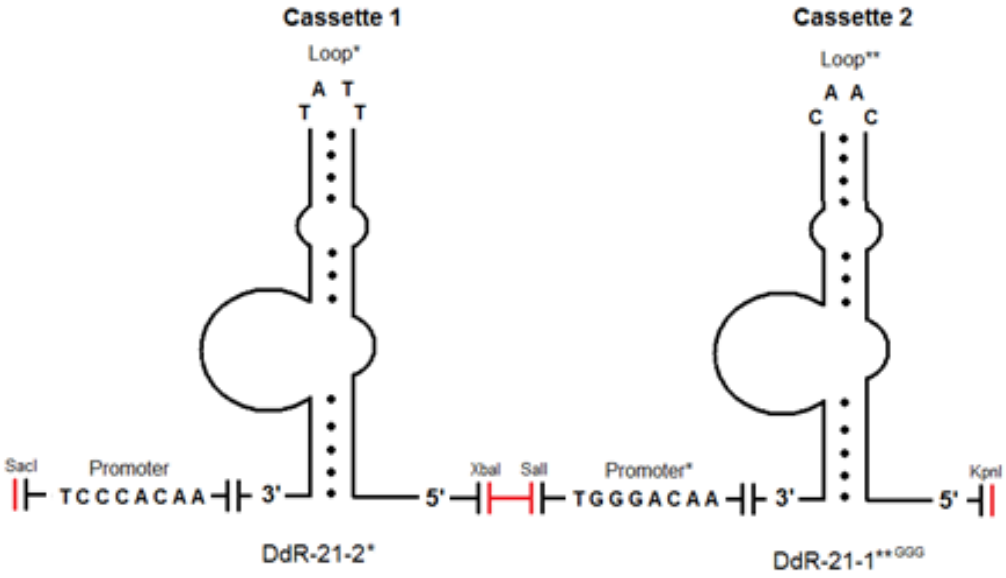


Figure 4. The constructed two cassettes of two differently modified *DdR-21* genes with surroundings inside a pDXA-HC plasmid. Cassette 1 contains the wild type promoter and loop mutation 1 while the second cassette contains a mutated promoter and loop mutation 2. The cassettes are depicted in black and the parts of the plasmid included are shown in red.

3.2.1 Construction of the extrachromosomal plasmid constructs

To analyse the putative promoter altered versions of the Class I gene *DdR-21* was used and inserted into an extra chromosomal plasmid, pDXA-HC. To be able to modify the gene it was first inserted into a TOPO-vector where mutations are easier to incorporate by site-directed mutagenesis. The modified genes were cut out of the TOPO-vector and subsequently ligated into the pDXA-vector and then succesfully transformed into *D. discoideum*.

For analysis of the promoter element (DUSE) three nucleotides in the promoter element, CCC, were changed to Gs, figure 4. In order to study the effect of this promoter mutation the expression of *DdR-21* variants preceeded by the wild type or mutated DUSE was monitored. To be able to distinguish between the endogenous *DdR-21* and the plasmid encoded constructs we also changed four nucleotides in the loop sequences of the genes. This made it possible to specifically detect the altered extrachromosomal *DdR-21* in the gene expression analysis.

In addition to the insert with the mutated promoter and loop sequence, DdR21-1**_GGG, another insert with a different loop mutation and a wild type promoter, DdR21-2*, was constructed and inserted after the first insert, creating two cassettes in the plasmid, figure 4. This was done for quantification purposes and to be able to use the other insert as a monitor of copy number effects and to see that the transformation was succesful. Hence, if no RNA can be detected from the promoter mutation construct we can determine that this is not due to a low plasmid copy number as long as the insert with wild type promoter is expressed.

A set of primers were designed for site directed mutagenesis to be able to make the different constructs, table 7 (appendix). The use of site directed mutagenesis to introduce the mutations was harder and more time consuming then expected because of long A and T-stretches in close proximity to the gene and because of the overall high AT-content. However, all the plasmids were successfully created and could be transformed into *D. discoideum*.

3.2.2 Verification of the double construct

To confirm that the created extra chromosomal plasmid contained the correct insert an analysis first by restriction digestion and then by sequencing was done. In order to see that both cassettes had been introduced correctly into the plasmid, cleavage by restriction enzymes were performed giving clear products of the expected size, data not shown. To confirm that the cassettes contained the correct modifications the plasmid was sent for sequencing, figure 5. The first cassette con-

tained the wild type promoter and loop mutation 1 while the second cassette contained the mutated promoter and loop mutation 2, confirming the intended construct.

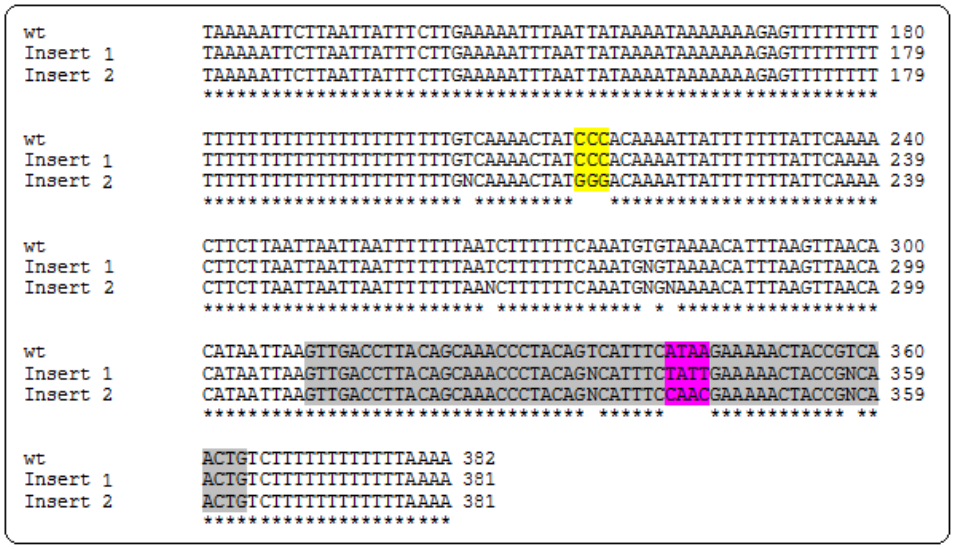


Figure 5. Alignment of the interesting parts of the sequences of the two modified constructs compared to the wild type. The *DdR-21* gene is marked in grey and the loop sequences are colored purple. The altered nucleotides of the putative promoter upstream of the gene are marked in yellow. N marks uncertain nucleotides in the sequences.

3.2.3 Tranformation into *D. discoideum* and screening of single colonies

The constructed plasmids, pDXA_DdR21-1, pDXA_DdR21-2, pDXA_DdR21-2**, pDXA_DdR21-1*_GGG and pDXA_DdR21_DC were subsequently transformed into a *D. discoideum* AX3 strain. The first four plasmids were succesfully created and stored in -80°C previous to the start of this master thesis. Transformation was carried out by electroporation and to identify the transformants G418 (Geneticin) selection was used. Primary transformants were harvested and spread on SM-agar plates together with *K. pneumoniae* to get single clones from which DNA could be extracted for screening.

PCR screening for the double construct was performed with a forward primer targeting the vector in front of the first cassette and reverse primers targeting the loop mutations within the cassettes, giving expected products of ~400 bp and ~1050 bp, respectively. As controls, cells were transformed with pDXA-HC vector without insert and these were screened for by a forward and reverse primer within

the vector, expected to generate a product of ~750 bp. Five clones of the control with only vector were analyzed and three positives were confirmed, figure 6. Only two single plaques of the double construct were obtained and screened and one was positive for both cassettes and the other only for the first, figure 6. All the positive clones were stored in -80°C and will be used in future expression studies.

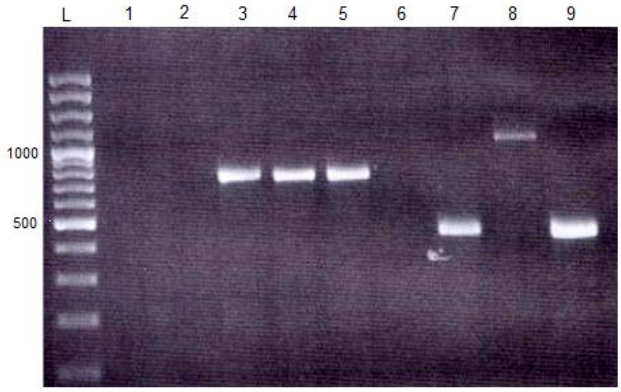


Figure 6. PCR identification of single colonies from *D. discoideum* containing only the pDXA-HC vector and colonies containing the pDXA-HC vector with the double construct. Lanes 1-5 contains five different clones, analyzed with primers 516+517, with just the vector and lanes 3-5 are clearly positive. Lanes 6-9 contains two different clones containing the double construct. Lanes 6-7 are for one clone, lane 6 is analyzed with primers 516 as forward and 493 as reverse primer and lane 7 is analyzed with 516 as forward and 495 as reverse primer. Lanes 8-9 are for the other clone with the same primers used as for lanes 6 and 7. The analysis was made so that both cassettes were searched for and it seems as both are at least there in the second clone.

3.2.4 Expression analysis of the putative promoter element

Even though all constructs were succefully created and transformed into *D. discoideum* there was not enough time during the span of this project to perform the expression analysis on the final extra chromosomal plasmid containing the double construct. To get a first insight into the possibility of the DUSE functioning as a promoter an expression analysis (northern blot) was performed on a construct with only the mutated promoter, pDXA_DdR21-2_GGG, comparing it to the wild type, figure 7.

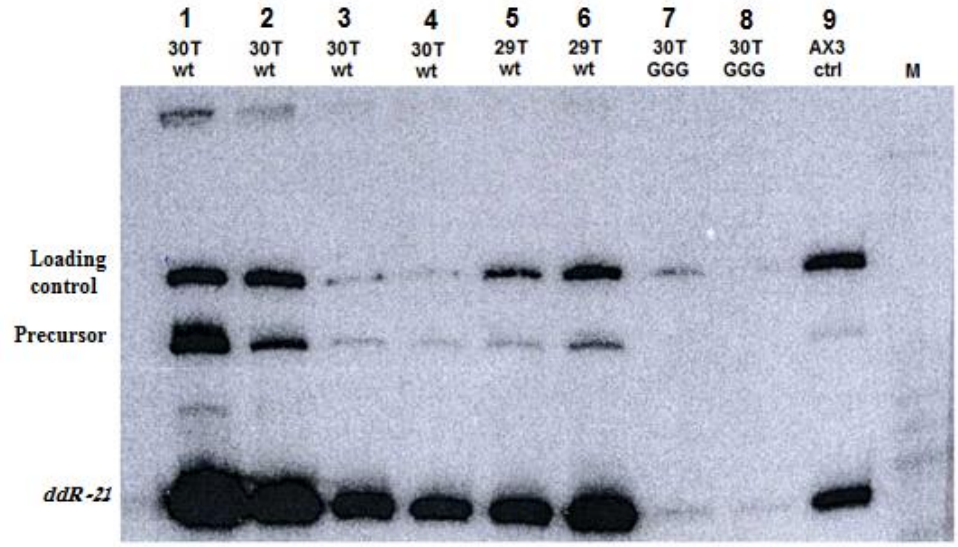


Figure 7. Expression of the extrachromosomal *DdR-21* in two different clones containing the same construct. Two different versions of the wild type were used because of variations in cloning. The AX3 control and size marker (M) are included for reference. Experiment performed by Lotta Avesson.

The probe used in this northern blot detects both the endogenous *DdR-21* and the plasmid expressed *DdR-21*. The loading control, U6 snRNA, shows that the amount of RNA loaded in each lane varies very much. The weak bands in lanes 7-8 with the promoter mutant can be compared with lanes 3-4 containing the wild type promoter. A big difference can be seen here with almost no expression of *DdR-21* when the promoter is mutated while the expression is still high in the wild type with the normal promoter. These data clearly points in the direction that the DUSE has an important role as a promoter element in front of ncRNA genes like *DdR-21*. Further confirmation will be gained when the plasmid containing the double construct can be analyzed in a similar fasion.

Looking at lanes 2 and 6 that have approximately the same amount of RNA loaded and comparing the expression of *DdR-21* it can be concluded that there seems to be no dfference between the 30T and 29T wild type. This was analyzed because the problems with the unstable T stretch located near the putative promoter sequence.

It can also be clearly seen that endogenous *DdR-21* is present in all lanes, which should give as strong bands as the loading control. The AX3 which was loaded in lane 9 is used as a control for the endogenous expression and as can be seen here the bands for the loading control and the *DdR-21* are approximately equally strong. The *DdR-21* gene is transcribed as a precursor before it is cleaved into its functional form and the probe also binds to this precursor explaining the third bands seen in figure 7.

3.3 *L.pneumophila* infection of *D. discoideum*

The aim of this part of the project was to to set up a good infection system in the lab with *D. discoideum* as a host for different bacteria. As stated above (table 1) many different bacteria have been succesfully used to infect *D. discoideum* but so far *L. pneumophila* has been most extensively studied, hence it was a logical choice to start with and the only one used in this work. A logical continuation of this project will be to try to develop the infection system for other interesting bacteria like *Mycobacterium marinum* and *Klebsiella pneumoniae*. In addition to setting up the infection system with *L. pneumophila* we also wanted to investigate the RNA interference response in the host upon infection with the use of different knock out strains. In connection to this we also analyzed the small RNA population in infected and uninfected cells by SOLiD sequencing.

3.3.1 Optimization of infection

Infection with *L. pneumophila* has been performed successfully in other labs with similar methods but with a high variation of infection percentage indicating that very small differences can lead to big variance in infection efficiency. Previousely to this project an infection system was unexisting in the lab meaning that we had to start everything up from the beginning and go through lots of adaptation and optimization steps to achieve successful and confirmed infection. The infection was done by growing *L. pneumophila* to a certain cell density at which it developed motility and then adding it at a specific MOI to wells containing *D. discoideum* at a concentration of $4 \cdot 10^5$ cells/ml, figure 8. After infection was started by a short centrifugation the 24-well plate were incubated at 25.5°C for 6 hours. During these 6 hours Gentamicin was added 2 hours after initiation of infection and was subsequently removed 2 hours later by a washing procedure. This was done to remove non-internalized bacteria. The infection was then either screened by microscopy or quantified by plating. The different parameters that were analyzed for optimization were the motility of the bacteria, the concentration of *D. discoideum* and *L. pneumophila*, the use and concentration of the antibiotic Gentamicin and the number of washes needed to kill and remove non-internalized bacteria.

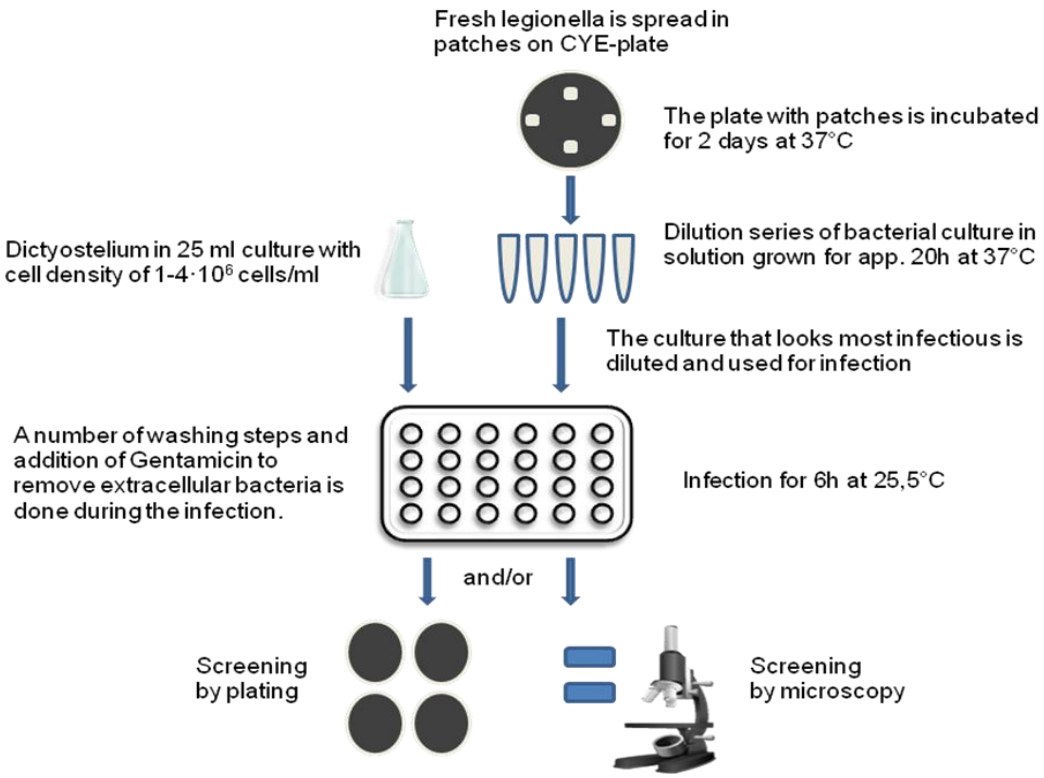


Figure 8. Flowchart of infection including a short description of the most important steps.

The motility of the bacteria is probably the most important parameter determining the infection efficiency and motility is dependent on the growth phase of *L. pneumophila*. By measuring the OD at 600nm of the liquid bacteria cultures and then checking the motility we hoped to see a connection between these parameters. What was found from this was that the variation was very big between cultures with motility varying between 0% and 30% cells for the same OD. What could be concluded was that an OD of over 3 was required to obtain more then 10% motility of the cells.

Another important parameter for efficient infection is to have a proper concentration of *D. discoideum* cells in the wells where the infection is performed. Too tightly packed cells would lead to the cells being protected by each other from being infected by the bacteria and therby affect the experiment negatively. Concentrations of *D.discoideum* between $1 \cdot 10^5$-$8 \cdot 10^5$cells/ml were tested and $4 \cdot 10^5$cells/ml were found to be the most suitable. The concentration of *L. pneumophila* compared to the host give the parameter multiplicites of infection (MOI) and we found that 10 times more bacteria to *D. discoideum* (MOI 10:1) gave the highest infection efficiency with the strains used.

To be able to quantify the infection efficiency all non-internalized bacteria had to be killed and removed. This was achieved by adding the antibiotic Gentamicin (100µg/ml) for 2 hours during the infection to kill bacteria outside the host cells, followed by at least three washes, adding fresh media and removing the killed non-internalized bacteria. It could be confirmed that no living non-internalized bacteria were left after this treatment when screening it by plating.

3.3.2 Screening of infection by microscopy

A bacterial strain containing a GFP-expressing plasmid with an IPTG inducable promoter was used in the experiments to be able to confirm that *L. pneumophila* had succesfully infected *D. discoideum*. This gave the opportunity to easily identify very scarce numbers of bacteria with a fluorescence microscope both as free living and more importantly post infection within the host, figure 9-11. Growth of *L. pneumophila* showed to be just as good both on plates and in liquid culture. The advantage of growing on liquid, AYE-broth, was that it was easy to supply ITPG at the desired concentration needed for good GFP induction. The bacteria also showed to be more motile, making it more infectious, when grown in liquid medium.

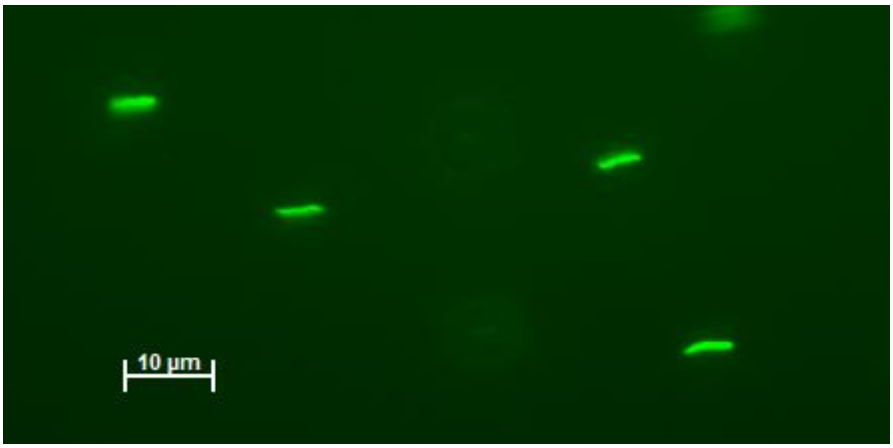


Figure 9. Microscopy photograph of free living *L. pneumophila* induced with IPTG, 1mM, at 63x magnification. The bacteria expressing GFP can be seen as rods of an approximate size of 5-8µm fluorescing GFP.

D. discoideum cells with internalized bacteria were not detectable with the microscopic settings used for visualizing GFP expressing *L. pneumophila*. This forced us to use a different microscope. Bifocal lighting was used to see the fixed *D. discoideum* cells, figure 10A. The irregular structures seen in the red ring are two *D. discoideum* cells and they are as expected between 10-20µm long. Visible in the

figures are also many grey dots which are contamination from the microscope. For visualizing the GFP fluorescing bacteria the settings of the micropsope were switched so that flouresence at 440nm could be detected, figure 10B. The image shows two rod-shaped bacteria at approximatly the same position as the *D. discoideum* cells indicating infection.

(a) (b)

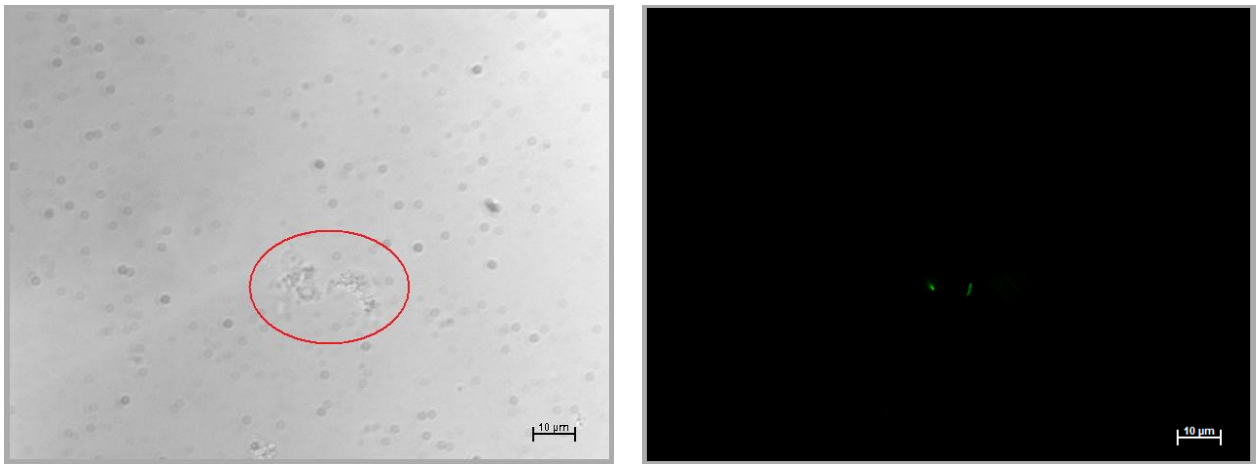


Figure 10. Microscopic photograph of *D. discoideum* infected with *L. pneumophila* at 40x magnification at normal bifocal light, (a), and at 440nm flouresence, (b). The *D. discoideum* cells are visible at bificoal light in the red circle and the bacteria are seen at the 440nm at the same location indicating that these cells are infected.

To get a better view of the probable infection figures 10A and 10B were superimposed, figure 11. The image clearly illustrates that the left most bacteria lies at the same position as the host, however it cannot be said for sure that the bacteria lies within the cell and not on top of it. It is even more difficult to say if the right most bacteria have infected the *D. discoideum* cell or if it is just outside the cell, since it is situated at the edge of the *D. discoideum* cell. Through the microscope screening it could alsod be seen that the frequency of infection was low. This could be said because most of the fixed *D. discoideum* did not seem to be at the same position as any bacteria. To be able to analyze the infection in a more quantitative way an additional plating screening method was performed.

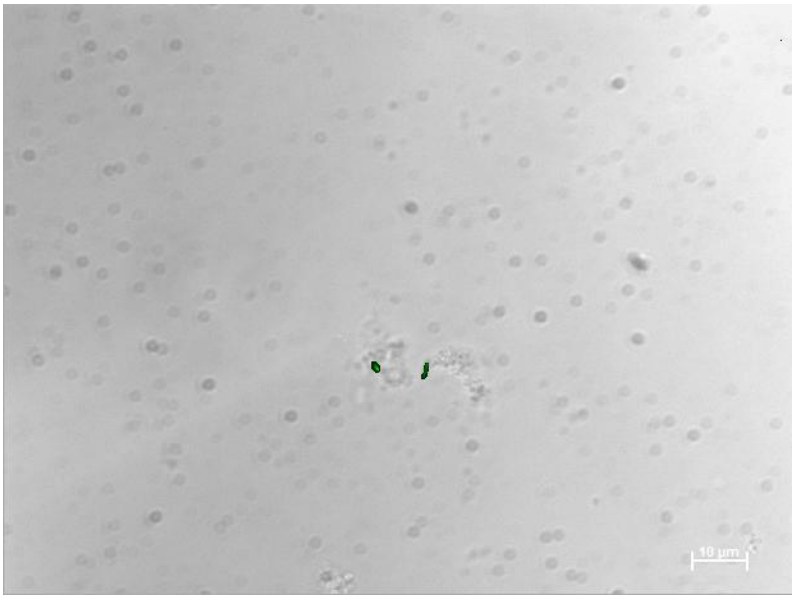


Figure 11. Illustratic picture of the pictures from figure 13 superimposed. It can clearly be seen that the bacterium on the left is at the same position as the host cell, indicating that it is within the host. The bacterium to the right is situated at the very edge of the host cell and could possibly also be within the host.

3.3.3 Screening of infection by plating

The method chosen to quantify the infection was the plating screening were *D. discoideum* cells are infected by *L. pneumophila*, treated with antibiotic to kill the free bacteria, washed and subsequently treated with saponin to release the intracellular bacteria. This mixture is then diluted and spread on CYE plates on which colonies are formed from the living bacteria. From the number of colonies on the plates the total number of internalized bacteria can be estimated and the number of infected *D. discoideum* can be determined. Two different wild type strains of *D. discoideum*, AX2 and AX4, and two AX2 knock out strains were used for the infection experiment, table 3.

Table 3. *Data from the plating screening assay where $4 \cdot 10^5$ D. discoideum cells were infected with $4 \cdot 10^6$ L. pneumophila cells, i.e. MOI 10:1.*

Dicty strain	# colonies	Infection efficiency
AX4	250	10,0%
AX2	85	3.4%
AX2 RdpC$^-$	80	3.2%
AX2 DrnB$^-$	30	1.2%

The experiments showed that the AX4 strain is clearly more prone to infection compared to the AX2 strain, table 3. Another interesting result is that the RdpC$^-$ strain shows approximately the same infectivity as the wild type but the DrnB$^-$ strain is almost three times as resistant to infection. This indicates that Dicer B somehow is involved in the hosts' response to bacterial infection while RdpC is not. It should be noted that these experiments were only performed once and they need to be repeated.

3.3.4 *L. pneumophila* effect on *D. discoideum* growth

To further analyze the effect *L. pneumophila* has on the different strains of *D. discoideum* a plaque assay was performed were the amoeba was grown on plates with a lawn of *K. pneumoniae* alone or together with *L. pneumophila*. *D. discoideum* normally grow on *K. pneumoniae* forming plaques as they consume the bacteria. By growing *D. discoideum* on *L. pneumophila* mixed with *K. pneumoniae* and comparing the plaque size and number with the ones formed on *K. pneumoniae* only it is easy to see the impact the pathogenic bacteria has on the amoeba. The data from the assay shows that all the strains have approximately the same number of plaques without *L. pneumophila*, table 4 and figure 12. A difference can be seen when looking at the number of plaques when the different *D. discoideum* strains were grown on a mix of *L. pneumophila* and *K. pneumoniae* where the DrnB$^-$ strain has more then 50% fewer plaques then the other strains. This data also points in the direction that Dicer B is involved in the response to *L. pneumophila*.

Table 4. *Results from the plaque assay with the different AX2 strains. These data are the average value from two different experiments.*

Dicty strain	With/Without *L. pneumophila*	# Plaques
AX2 DrnB$^-$	Without	435
AX2 DrnB$^-$	With	97,5
AX2 RdpC$^-$	Without	420
AX2 RdpC$^-$	With	217,5
AX2 wild type	Without	420
AX2 wild type	With	220

Another test that was made was to grow *D. discoideum* on a bacterial lawn consisting of only *L. pneumophila*. No plaques could be seen on these plates indicating that *D. discoideum* can not grow on only *L. pneumophila*.

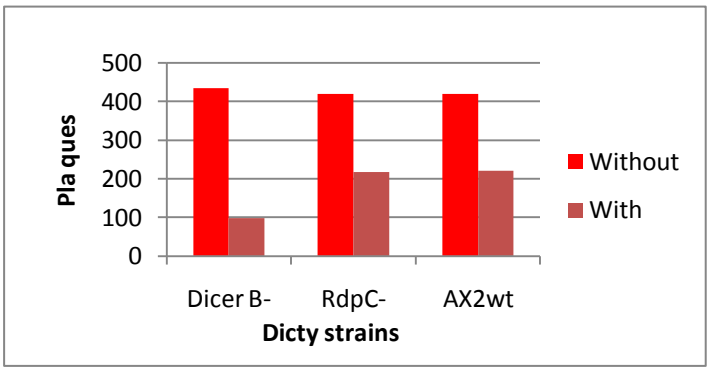


Figure 12. Plaque assay of the AX2 knock out strains with and without the presence of *L. pneumophila*.

An additional difference could be seen in the plate assay when comparing the size of the plaques in the assay, figure 13. The Dicer B knock out strains clearly shows smaller plaques then the other strains both in the presence and absence of *L. pneumophila*.

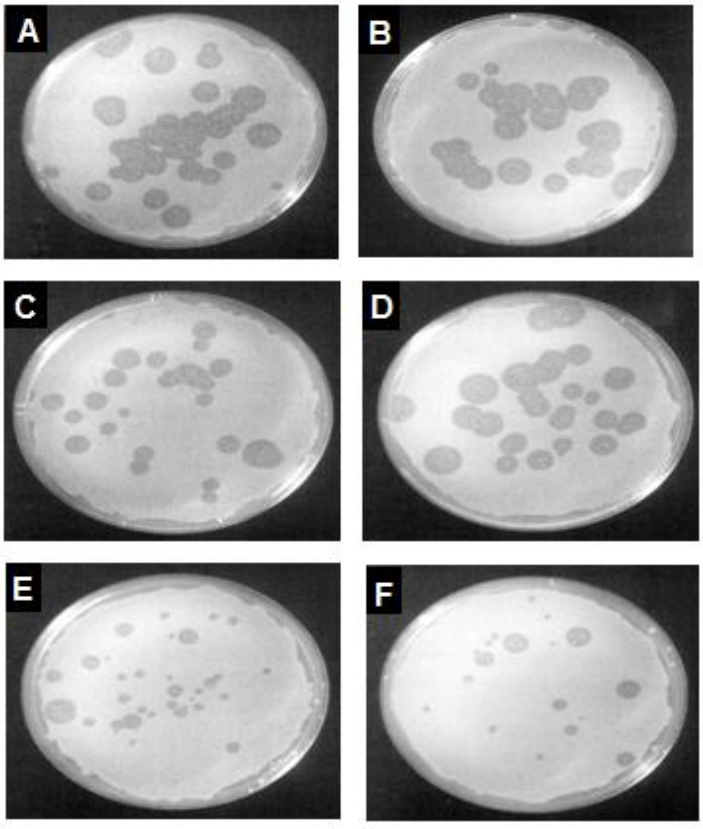


Figure 13. Photpgraphs of the plates from a plaque assay after 6 days. A) AX2 wild type without *L. pneumophila*. B) AX2 wild type with *L. pneumophila*. C) AX2 RdpC$^-$ without *L. pneumophila*. D) AX2 RdpC$^-$ with *L. pneumophila*. E) AX2 DrnB$^-$ without *L. pneumophila*. F) AX2 DrnB$^-$ with *L. pneumophila*.

3.3.5 The role of small ncRNAs in bacterial infection

The SOLiD sequencing gave massive amounts of data and making an extensive analysis of the SOLiD data was beyond this project. In the analysis in this thesis we decided to focus on the distribution of the different classes of RNAs, size distribution of the different libraries and comparative expression studies of the RNAs in infected and uninfected cells.

The distribution of the different classes of RNAs found in the SOLiD data is illustrated as pie charts in figure 14. In both infected and uninfected cells most of the hits belonged to complex repeats regions (60,9% and 40,4% respectively). The second largest fraction of hits lies within the coding genes (23,5% and 35,1% respectively). Another rather large portion was derived from possible degradation products of rRNA and tRNA (4.3% and 8.7% respectively). An unexpectedly small portion hits from the retrotransposon DIRS-1 was found (3,3% in both). In similar analyses of small RNAs from growing cells, DIRS derived RNA made up >50% of the total population of small RNAs [46] (Avesson, Reimegård and Söderbom, unpublished). Putative degradation products of other non-coding RNAs like snoRNAs and SRP RNAs constitute only a very small fraction in both libraries (1,1% and 0,44%).

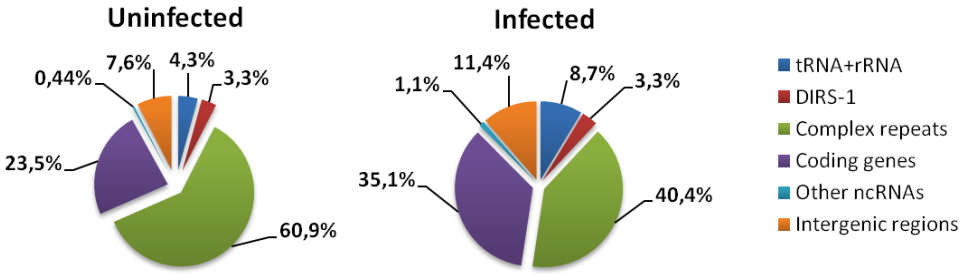


Figure 14. Schematic view of the relative differences between the classes of RNAs found in the SOLiD sequencing. To the left are data from uninfected cells and to the right are data from cells 6 hours after infection.

Looking at the expression patterns data from the SOLID sequencing there were in total 5171 specific gene hits with a threshold value of at least 10 hits per gene. The top ten genes, with the largest difference in expression in infected compared to uninfected cells are listed in table 5. Out of the ten genes only one was downregulated whereas the rest were upregulated in the infected *D. discoideum* cells. The most interesting finding is that one of the Class I RNAs, *DdR-31*, is the most affected gene, showing that this class of RNAs might play an important role in the

host upon infection. The downregulated gene *nad7* together with many other of the most downregulated genes are situated on the mitochondrial genome (data not shown).

Table 5. *The ten RNA genes with the highest difference in expression in infected cells compared to uninfected cells. The ratio is calculated from the natural logarith of the number of hits in uninfected cells divided by the number of hits in infected cells. Hence a ratio of 2,90 means that the gene is expressed approximately 18 times as much in infected compared to un-infected cells,*

Gene	Chromosome	Ratio	Gene description
r31	5	2,90	Class I RNA, unknown funtction
fcf1	1	2,24	FCF1 family protein, unknown function
DDB_G0291598	6	2,13	Unknown gene
abpC	1	2,10	Actin-binding protein C, gelation factor
DDB_G0286651	4	2,08	Saposin B domain-containing gene, lipid- and sphingolipid metabolism
tRNA-Ser-AGA-7	3	2,04	Serine transfer RNA
tRNA-Ile-AAU-14	6	2,03	Isoleucine transfer RNA
nad7	M	-2,01	NADH dehydrogenase subunit 7
kif11	5	1,97	Kinesin family member 11, microtubule-based movement
DDB_G0280619	3	1,97	Unknown gene

A previous study where DNA microarray was used to analyze the gene expression of *L. pneumophila* infected *D. discoideum* cells showed that there are lots of transciptional changes upon infection [63]. Comparing the significantly regulated genes 6 hours post infection in the microarray experiment to the SOLiD data analyzed in this project showed that only one of the top matches in the SOLiD experiment matched one of the affected in the microarray study, the DDB_G0286651. This gene which was highly overexpressed in infected cells compared to uninfected cells in our SOLiD data showed to be downregulated in the microarray analysis. This could indicate higher levels of expression but also an increased turnover of the gene.

The filtration steps involved in the bioinformatic analysis of the SOLiD data can be seen in figure 15. The analyzis of the size distribution of the sequences in the different libraries maping to different regions in the genome are shown in the diagrams in figure 16-18. Looking at what the different regions have in common it is clear that there is a size bias towards 21 nt, a normal size for siRNAs and miRNAs. In three of the diagrams there is a great number of small RNAs in the size range of 9-15 nt which may be degradation products.

Analysing the size range of all sequences (figure 16) a big difference is seen between the uninfected, and the infected cells. There are more then twice as many sequences at 21 nt in the uninfected compared to the infected which instead has a larger fraction of sequences in the 10-12 nt range. This difference is also prominent when looking at the intergenic regions (figure 17) even though the difference is not as substantial. As for the library of the ncRNA repeats (figure 18) almost all sequences have a size of approximately 21 nt.

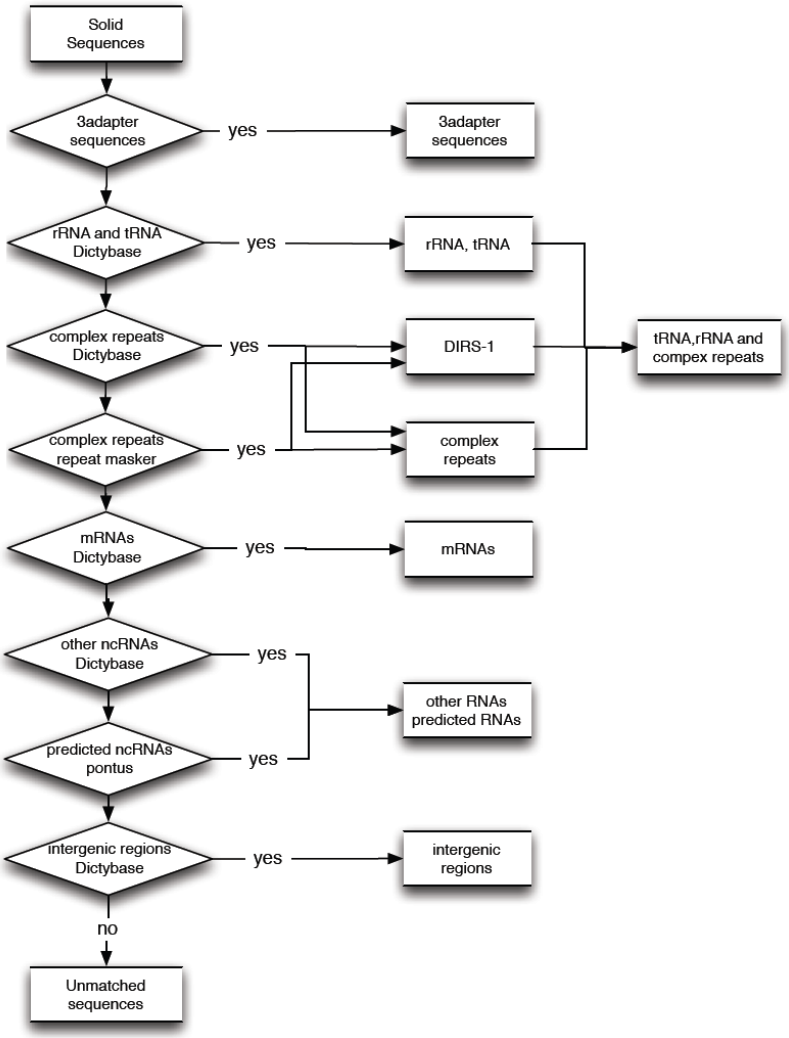


Figure 15. Schematic view of the bioinformatic filtration performed on the SOLiD sequencing data. The diamonds to the left show what the sequences were searches against and the boxes show the different libraries the sequences were put in.

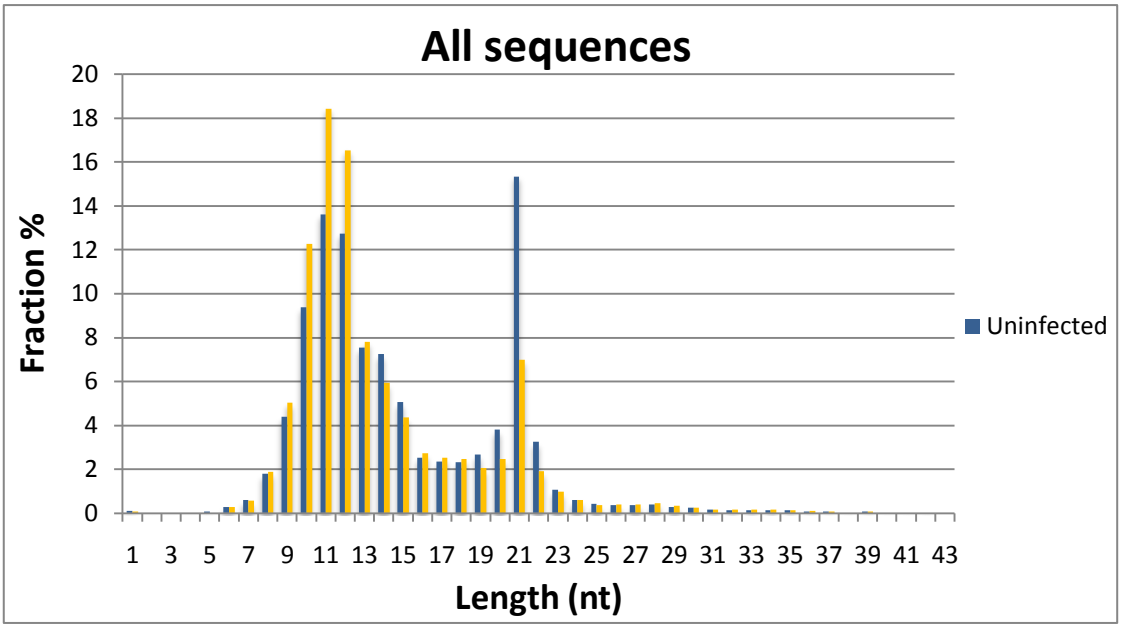


Figure 16. Size distribution of all the sequences from the SOLiD sequencing comparing uninfected and infected cells.

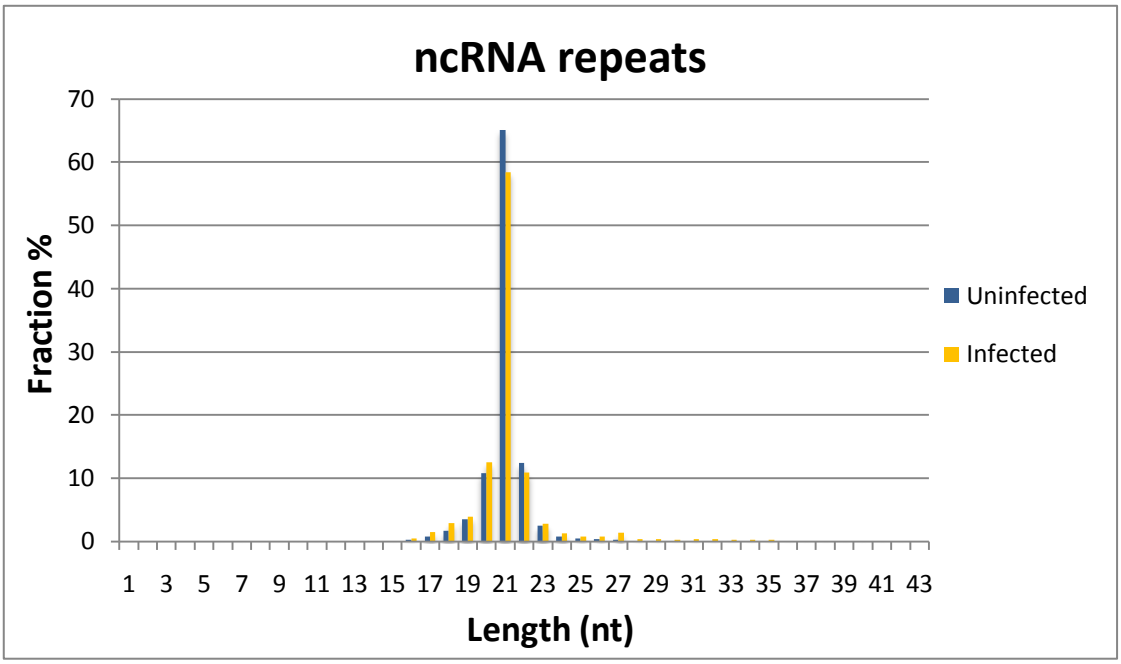


Figure 17. Size distribution of the sequences in the intergenic library by the SOLiD sequencing comparing uninfected and infecteed cells.The sequences in this diagram corresponds to the sequences in the rRNA and tRNA Dictybase and complex repeats Dictybase seen in figure 15.

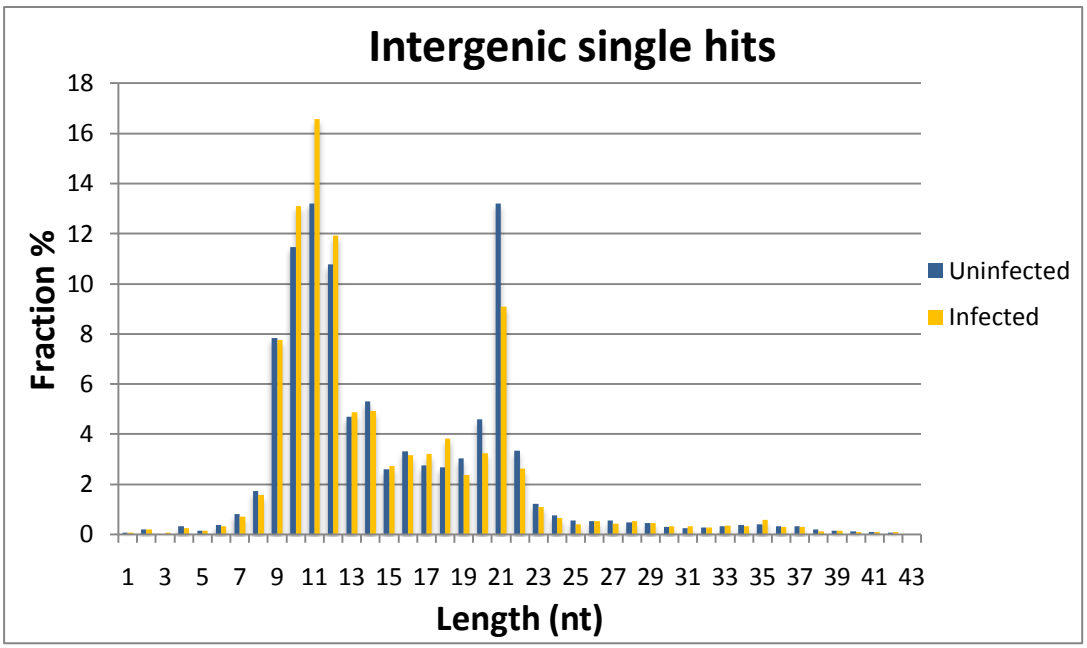


Figure 18. Size distribution of the sequences in the ncRNA repeats library from the SOLiD sequncing comparing uninfected and infected cells. These sequences correspond to the sequences in the intergenic regions seen in figure 15.

4 Discussion

4.1 Expression studies of a putative ncRNA promoter

To construct an extra chromosomal plasmid containing two modified versions of the *DdR-21* gene showed to be much harder then expected. Because of the extremly high A/T richness of the *D. discoideum* genome it was really difficult to achieve specific binding of primers for site-directed mutagenesis and incorporation of the wanted mutations. Another problem was that the region in close proximity to the putative promoter element contained a stretch of Ts which showed to be very unstable. During the mutational PCRs there was often a deletion of a T leaving 29 Ts instead of the normal 30 Ts preceding the putative promoter element. Since we could not be sure how this would affect the promoter element we had to analyze these modifications in case we could not get a construct with the wild type 30 Ts for a certain mutation. This led to a lot of extra work for this part of the project and in the end we could conclude that the stretch of Ts did not seem to affect the function of the promoter if it were 29Ts compared to 30Ts for a certain mutation, figure 7.

Even though there was a lot of trouble in making the double construct it proved to be succesfull in the end. This construct will be extremly useful for further detailed functional studies of the promoter and it can also be used for deletion/mutational studies of other interesting elements involved in gene regulation identified in the *D. discoideum* genome. Even though the northern blot from the preliminary expression study showed great variance in RNA loading it gave a strong first indication that DUSE is a true promoter element. The expression test showed a clear loss in expression of the model gene *DdR-21* when the promoter sequence was mutated. The conclusion that can be drawn from the test is that the DUSE does have an effect on the expression levels of the model gene, which a promoter element should. Even though no real controls were included in this pre-

liminary expression, as will be in the final construct with the two cassettes and double loop mutations, it still gave some initial results. The upcoming expression studies of the double construct will give more decisive evidence and hopefully strengtening these preliminary results and confirming the DUSE as a promoter element for ncRNAs in *D. discoideum*.

4.2 *L. pneumophila* infection studies

The infection system set up was based on infection studies performed in other research groups. Even though it had been succesfully performed before, it was much harder then expected to get the *L. pneumophila* infectious enough to get a quantifible infection of *D. discoideum*. A very important parameter for infection is the motility of *L. pneumophila* which is dependent on the current life phase of the bacteria. The bacteria develops flagella when the concentration of bacteria reach a certain cell density, giving *L. pneumophila* motility and the possiblity to infect a new host. To get bacteria with a high degree of motility was shown to be a big problem and no good correlation was found between motility and exact levels of cell density, measured by OD. This made it very hard to get a system that could be repeated over and over again with similar levels of motility of the bacteria in each experiment.

A possible explanation to our problems to reach a high motility could be that the strain has lost some of its infectious potency. This happens if the bacteria are kept running to long and since the *L. pneumophilia* sample we got were from a lab in the US it is hard to know how long that sample had been grown before it were sent. It is known that the same train works well in the infectious studies done on it in the US indicating that it is not with the strain the problem lies. A good idea would be to get an additional sample of the bacteria and compare it with the bacteria used in this study to see if a more motile sample of *L. pneumophila* can be found. Another potenital parameter that could be changed is possibly the growth medium. The bacteria seem to be very sensitive to the medium and a small change in the medium could lead to a big change in the infectivity of the bacteria.

Another problem that occured was that it was difficult to exactly determine if a bacterial cell was within the *D. discoideum* cells just by fluorescence microscopy. Even though the bacteria could be clearly seen at the same position as the host cell it was still a chance that the bacteria might be stuck on the outer membrane and had not infected the cell. To overcome this problem a confocal microscope could

be used as a further confirmation of infection. A confocal microscope gives an increased contrast and the possibility to see the cells in a three-dimensional way.

For the growth studies of *D. discoideum* on a lawn of *L. pneumophila* we found that no plaques were formed on these plates. The reason for this could be that *D. discoideum* takes up the bacteria but is kiled by it or that it can not consume the bacteria and therefore not grow.

4.2.1 The role of DrnB and RdpC in *L. pneumophila* infection

We also wanted to investigate the RNAi response in *D. discoideum* upon bacterial infection. One way this was done was by utilizing existing strains where important genes of the RNAi machinery have been knocked out. Comparative studies were then used to find differences with the knock out strains compared to the normal wild type strain when infected with *L. pneumophila*.

It has previousely been shown that *D. discoideum* contains two Dicer-like proteins, DrnA and DrnB. Analysis of their involvement in RNAi and miRNA biogenesis showed that only DrnB had an significant effect when knocked out. In the strain lacking DrnB the analyzed miRNA could no longer be found indicating that DrnB is responsible for the processing of the pre-miRNA [45]. *D. discoideum* also contains three RNA-dependent RNA polymerases, RdpA-C, where only RdpC had a clear effect when removed. The strain without active RdpC had upregulated levels of a tested miRNA [45]. From these findings we decided to focus on these proteins and work with the $DrnB^-$ and $RdpC^-$ strains in addition to the wild type strain in our infection studies. From the observations made on the RdpC- strain no effects could be found on the infection rate or the growth of *D. discoideum* when in proximity to *L. peumophilia*, figure 12 and 13. This was not unexpected from the fact that the loss of RdpC led to higher siRNA and miRNA levels [45]. Putative miRNAs with possible function in response to infection would then still be present in the cells and the response may function as in wild type cells.

Because of the problems to get the infection system to work properly and that the optimized experiments were not repeated as many times as desired, we have no solid statistical data. This means that the conclusions drawn here are only a first model which has to be further analyzed and more data is desired to get solid proof of the results.

4.2.2 Working models for DrnB involvment in response to *L. pneumophila*

In contrast to the RdpC$^-$ an effect could be seen both in the infection assay and the plaque assay with the DrnB$^-$ strain indicating a role for the protein in response to bacterial infection. The data from the plaque assay showed an approximate decrease of 50% in the number of plaques in the DrnB$^-$ strain compared to the other strains. This suggests that this strain has a lower survival rate e.g. that *L. pneumophila* were more pathogenic to this strain. Another clear difference between the DrnB$^-$ strain and the other strains was the size difference of the plaques both in the presence and absence of *L. pneumophila*, indicating that this strain of bacteria was more resistent to predation by *D.discoideum*.

Analyzing the results of the infection efficieny of the different strains shows a different story. Comparing the two different wild type strains AX2 and AX4 shows a big difference in how prone they are to infection where the AX4 strain is approximately three times easier to infect. This shows that the AX4 would be a better strain to work with in future infection studies since higher infection efficiency is strived for. When comparing the two knock out strains to the isogenic wild type AX2 strain, it was clear that the infection efficiency is approximately the same in the RdpC$^-$ strain and the wild type strain but more then 50% lower in the DrnB$^-$ strain. This indiated that the *D. discoideum* DrnB$^-$ strain is more resistent to bacterial infection by *L. pneumophila*. However, those experiments have to be repeated several times to get statistical solid data. The observations from the infection and the plaque assays seem contradictory since one suggests that the DrnB$^-$ strain is more resistent to infection at the same time as the other implies that the survival is lower when grown together with *L. pneumophila*.

A normal effect of miRNA(s) is that they cause a down-regulation of target genes. This would lead to an up-regulation of these target genes in the DrnB$^-$ strain. A possible explanatory model is that these target genes code for proteins that inhibit a gene coding for a receptor protein that is needed for the uptake of bacteria. The down-regulation of the receptor protein in this *D. discoideum* strain would then lead to an inhibitied growth and at the same time give a lower susceptibility to infection.

Another possible model for the observations made with the DrnB$^-$ strain is that miRNA(s) are involved in *D. discoideum* survival upon bacterial infection. In the DrnB$^-$ knock out strain the miRNAs cannot be processed to active miRNA and hence, cannot protect the host cell from being killed by *L. pneumophila*. This would mean that the DrnB$^-$ strain is not more resistent to infection as the infection

data suggests but that the ability of *D. discoideum* to survive after infection, with intracellular *L. pneumophila*, is decreased. This could be a possible explanation since one problem with the experimental set up is that the number of living *D. discoideum* is not acounted for which could be a big variable. To get around this problem a future experiment that should be included is to also quantify the number of live *D. discoideum* after infection and compare that to the number of *L. pneumophila*. By doing this a more definite and trustworthy infection efficiency of the different strains could be calculated.

4.2.3 SOLiD analysis of ncRNAs and infection

SOLiD sequencing gives an enormous amount of data so for this thesis project we limited the analysis to only include parts of the data that were handlable to analyze during the time of this project. Upcoming analyses will certainly give lots of more interesting data giving an even clearer view of expression differences of ncRNAs in infected and uninfected *D. disoideum* cells.

What could be seen in the libraries created from the SOLID data was that there is a clear size bias against 21 nt in both uninfected and infecte cells, which was expeced since this is the common size for miRNAs and many siRNAs. The largest portion of sequences was found in the 9-12 nt size range which most probably is degraded RNA. The finding that there is a much higher number of sequences in the 21 nt range in non-infected cells compared to infected is very interesting. It can also be clearly seen that the number of sequences in the 9-12 range is larger in infected cells indicating that there could be some kind of cleavage of the mi/siRNA when the cells are infected. This also coincides with the findings by Zhang and Kuspa that mitochondrial rRNA is cleaved upon infection and possibly this cleavage can happen to many other sorts of ncRNAs in the cells as well [41].

Another interesting finding was that most affected gene in the expression analysis (SOLiD) when comparing infected cells to uninfected was the Class I RNA, *DdR-31*. In this thesis the Class I *DdR-21* gene was used as a model gene for analyses of the putative ncRNA promoter (DUSE) and we also had planned to analyze the possible involvement of this class of ncRNAs in infection by knock out studies. From the SOLiD sequencing it seems that at least the *DdR-31* is involved in the response to infection and it would be a good model gene for the Class I RNAs to use for future infection studies.

The finding that the most downregulated genes were situated on the mitochondrial genome is not very unexpected. This also correlates with what Zhang and

Kuspa [41] showed, i.e. that the mitochondrial rRNA is cleaved upon infection which would lead to a down regulation of these genes in infected cells, corraboralating our findings.

One problem with the analysis of the SOLiD data is that not all RNA genes are found because they are too similar to other genes. Due to the method of analysis this will lead to that genes end up among the unmatched sequences and are not included in the expression analysis. An example of this is the *DdR-21* gene; hence no conclusions about its involvement in infection can be drawn. To be able to include *DdR-21* and other genes of interest in future expression analysis, a search against the exact sequence of the genes of interest that have not been found needs to be performed to fish them out.

Acknowledgments

First and foremost I want to show my gratitude towards my supervisor Dr. Fredrik Söderbom for giving me the opportunity to do this project. I want to thank him for always being there when needed and always being open to questions and advice. It has been both interesting and fun to work with the project and I have gained a vast amount of experience from doing it.

I also want to give big thanks to Lotta Avesson who has been my guide in the lab. I have appreciated that you have been open to even the stupidiest of questions and for being patient, kind and helpful.

In addition I want to show my appreciation towards the other members in our group Åsa Fransson and Anna Åsman for being very good people to work with and for kindly assisting when needed.

I would also like to thank Johan Reimegård and Andrea Hinas for the help with the SOLiD sequencing analysis and the fluorescence microscope.

Finally I want to show my gratitude towards all the people in the department of molecular biology, SLU, for being very nice and helpful during my time in the lab.

References

1. Eichinger L. and Noegel A.A. (2003). Crawling into a new era – the *Dictyostelium* genome project. *The EMBO journal* **22**, 1941-1946.

2. Raper B.K. (1935). Dictyostelium discoideum, a new species of slime mold from decaying forest leaves. *Journal of Agricultural Research* **50**, 135-147.

3. Kessin R. H. (2001). *Dictyostelium* – evolution, cell biology, and the development of multicellularity. Cambridge Univeristy Press, Cambridge, UK.

4. Kuspa A. and Loomis W.F. (2006). Dictyostelium discoideum protocols. Humana Press, 1 edition.

5. Eichinger L. *et al.*, (2005). The genome of the social amoeba *Dictyostelium discoideum*. *Nature* **435**, 43-57.

6. Chisholm L.R., Gaudet P., Just M.E., Pilcher E.K., Fey P., merchant N.S. and Kibbe A.W. (2006). dictyBase, the model organism database for Dictyostelium discoideum. *Nucleic Acids Research* **34**, D423-D427.

7. Bonner T.J. (1998). A way of following individual cells in the migrating slugs of Dictyostelium discoideum. *PNAS*, **95**, 9355-9359.

8. Urushihara H. (2009). The cellular slime mold: Eukaryotic model microorganism. *Experimental animals*. **58**, 97-104.

9. Fuchs B.B. and Mylonakis E. (2006). Using non-mammalian hosts to study fungal virulence and host defence. *Current Opinion in Microbiology*, **9**, 367-372.

10. Dorer S.M. and Isberg R.R. (2006). Non-vertebrate hosts in the analysis of host-pathogen interactions. *Microbes and Infection*, **8**, 1637-1646.

11. Steinert M., Heuner K., Buchrieser C., Albert-Weissenberger C. and Glöckner G. (2007). Legionella pathogenicit: Genome structure, regulatory networks and the host cell response. *International Journal of Medical Microbiology* **297**, 577-587.

12. Clarke M. (2009). Recent insights into host-pathogen interactions from Dictyostelium. *Cellular microbiology*.

13. Russel G.D. and Gordon S. (2009). Phagocyte-pathogen interactions: Macrophages and the host response to infection. American Society for Microbiology Press, Washington, D.C.

14. Hägele S., Köhler R., Merkert H., Schleicher M., Hacker J. and Steinert M. (2000). Dictyostelium discoideum: a new host model system for intracellular pathogens of the genus Legionella. *Cellular microbiology*. **2**, 165-171.

15. Steenbergen N.J., Shuman A.H. and Casadevall A. Cryptococcus neoformans interactions with amoebae suggest an explanation for its virulence and intracellular pathogenic strategy in macrophages. *PNAS* **98**, 15245-15250.

16. Benghezal M., Fauvarque M-O., Tournebize R., Froquet R., Marchetti A., Bergeret E., Lardy B., Klein G., Sansonetti P., Charette J. S. and Cosson P. (2006). Specific host genes required for killing of Klebsiella bacteria by phagocytes. *Cellular Microbiology* **8**, 139-148.

17. Skriwan C., Fajardo M., Hägele S., Horn M., Wagner M., Michel R., Krohne G., Schleicher M., Hacker J. and Steinert M. (2002). Varioud bacterial pathogens and symbionts infect the amoeba Dictyostelium discoideum. *International Journal of Medical Microbiology* **291**, 615-624.

18. Solomon M.J., Leung S.G. and Isberg R.R. (2003) Intracellular replication of Mycobacterium marinum within Dictyostelium discoideum: Efficient replication in the absence of host coronin. *Infection and Immunity* 3578-3586.

19. Hagedorn M., Rohde H.K., Russel G.D. and Soldati T. (2009). Infection by tubercular Mycobacteria is spread by nonlytic ejection from their amoeba hosts. *Science* **323**, 1729-1733.

20. Colucci R.M.A., Peracino B., Talá A., Bozzaro S., Allfano P. and Bucci C. (2008). Dictyostelium discoideum as a model host for meningococcal pathogenesis. *Medical Science Monitor* **14**, 134-140.

21. Pukatzki S., Kessin H.R. and Mekalanos J.J. (2001). The human pathogen Pseudomonas aeruginoa utilizes conserved virulence pathways to infect the social amoeba Dictyostelium discoideum. *PNAS* **99**, 3159-3164.

22. Pukatzki S., Ma T.A., Sturtevant D., Krastins B., Sarracino D., Nelson C.W., Heidelberg F.J. and Mekalanos J.J. (2005). Identification of a con-

served bacterial protein secretion system in Vibrio cholerae using the Dictyostelium host model system. *PNAS* **103**, 1528-1533.

23. Vlahou G., Schmidt O., Wagner B., Uenlue H., Dersch P., Rivero F. and Weissenmayer A.B. (2009). Yersinia outer protein YopE affects the actin cytoskeleton in Dictyostelium discoideum through targeting of multiple Rho family GTPases. *BMC Microbiology* **9**:138.

24. Fields B.S. (1996). The molecular ecology of legionellae. *Trends in microbiology* **4**, 286-290.

25. Diederen B.M.W. (2007). Legionella spp. and Legionnaires' disease. *Journal of Infection* **56**, 1-12.

26. Palusińska−Szysz M. and Cendrowska−Pinkosz M. (2009). Pathogenicity of the family Legionellaceae. *Archivum immunologiae et therapiae experimentalis*. **57**, 279-290.

27. Albert-Weissenberger C., Cazalet C. and Buchrieser C. (2006). Legionella Pneumophila – a human pathogen that co-evolved with fresh water protozoa. *Cellular and Molecular Life Sciences* **64**, 432-448.

28. Chien M., Morozova I., Shi S., Sheng H., Chen J., Gomez S.M., Asamani G., Hill K., Nuara J., Feder M., Rineer J., Greenberg J.J., Steshenko V., Park S.H., Zhao B., Teplitskaya E., Edwards J.R., Pampou S., Georghiou A., Chou I., Lannucilli W., Ulz M.E., Kim D.H., Geringer-Sameth A., Goldsberry C., Morozov P., Fischer S.G., Segal G., Qu X., Rzhetsky A., Zhang P., Cayanis E., De Jong P.J., Ju J., Kalachikov S., Shuman H.A. and Russo J.J. (2004). The genomic sequence of accidental pathogen *Legionella pneumophila*. *Science* **305**, 1966-1968.

29. Cazalet C., Rusniok C., Brüggermann H., Zidane N., Magnier A., Ma L., Tichit m., Jarraud S., Bouchier C., Vandenesch F., Kunst F., Etienne J., Glaser P. and Buchrieser C. (2004). Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity. *Nature Genetics* **36**, 1165-1173.

30. Molofsky B.A. and Swanson S.M. (2004). Differentiate to thrive: lessons from the *Legionella pneumophila* life cycle. *Molecular Microbiology* **53**, 29-40.

31. Murga R., Forster S.T., Brown E., pruckler M.J., Fields S.B. and Donlan M.R. (2001). Role of biofilms in the survival of *Legionella pneumophila* in a model portable-water system. *Microbiology* **147**, 3121-3126.

32. Declerck P. (2009). Biofilms: the environmental playground of Legionella pneumophila. *Environmental Microbiology*.

33. Lu H. and Clarke M. (2005). Dynamic properties of Legionella-containing phagosomes in Dictyostelium amoebae. *Cellular Microbiology* **7**, 995-1007.

34. Jules M. and Buchrieser C. (2007). Legionella pneumophila adaptation to intracellular life and the host response: Clues from genomics and transcriptomics. *FEBS Letters* **581**, 2829-2838.

35. Farbrother P., Wagner C., Na J., Tunggai B., Morio T., Urushihar H., Tanaka Y., Schleicher M., Steinert M. and Eichinger L. (2006). *Dictyostelium* transcriptional host cell response upon infection with *Legionella*. *Cellular Microbiology* **8**, 438-456.

36. Molofsky B.A. and Swanson S.M. (2003). *Legionella pneumophila* CsrA is a pivotal repressor of transmission traits and activator of replication. *Molecular Microbiology* **50**, 445-461.

37. Molofsky B.A. Shetron-Rama M.L. and Swanson S.M. (2005). Components of the Legionella pneumophila flagellar regulon contribute to multiple virulence traits, including lysosome avoidance and macrophage death. *Infection and Immunity* **73**, 5720- 5734.

38. Swanson S.M. and Hammer K.B. (2000). *Legionella pneumophila* pathogenesis: A fateful journey from amoebae to macrophages. *Annual review of Microbiology* **54**, 567-613.

39. Tossa P., Deloge-Abarkan M., Zmirou-Navier D., Hartemann P. and Mathieu L. (2006). Pontiac fever: an operational definition for epidemiological studies. *BMC public health* **6**:112.

40. Hinas A. and Söderbom F. (2007). Treasure hunt in an amoeba: non-coding RNAs in *Dictyostelium discoideum*. *Current Genetics* **51**, 141-159.

41. Zhang C. and Kuspa A. (2009). Transcriptional down-regulation and rRNA cleavage in *Dictyostelium discoideum* mitochondria during *Legionella pneumophila* infection. *PLoS ONE* **4**(5): e5706.

42. Sahr T., Brüggemann H., Jules M., Lomma M., Albert-Weissenberger C., Cazalet C., and Buchrieser C. (2009). Two small ncRNAs jointly govern virulence and transmission in *Legionella pneumophila*. *Molecular Microbiology* **72**, 741-762.

43. Katiyar-Agarwal S., Gao S., Vivian-Smith A. and Jin H. (2007). A novel class of bacteria-induced small RNAs in *Arabidopsis*. *Genes & Development* **21**, 3123-3134.

44. Navarro L., Dunoyer P., Jay F., Arnold B., Dharmasiri N., Estelle M., Voinnet O. and Jones D. G. J. (2006). A plant miRNA contributes to antibacterial resistence by repressing auxin signaling. *Science* **312**, 436-439.

45. Hinas A., Reimegård J., Wagner E.G.H., Nellen W., Ambros R.V. and Söderbom F. (2007). The small RNA repertoire of *Dictyostelium discoideum* and its regulation by components of the RNAi pathway. *Nucleic Acids Research* **35**, 6714-6726.

46. Aspegren A., Hinas A., Larsson P., Larsson A. and Söderbom F. (2004). Novel non-coding RNAs in *Dictyostelium discoideum* and their expression during development. *Nucleic Acids Research* **32**, 4646-4556.

47. Alberts B., Johnson A., Lewis J., Raff M., Roberts K. and Walter P. Molecular Biology of the Cell. Fourth edition. Garland science, New York, USA.

48. Hinas A., Larsson P., Avesson L., Kirsebom A. L., Virtanen A. and Söderbom F. (2006). Identification of the major spliceosomal RNAs in *Dictyostelium discoideum* reveals developmentally regulated U2 variants and polyadenylated snRNAs. Eukaryotic Cell, 924-934.

49. Vankan P., Edoh D. and Filipowicz W. (1988). Struture and function of the U5 snRNA gene of *Arabidopsis thaliana*. Conserved upstream sequence elements in plant U-RNA genes. *Nucleic Acids Research* **16**, 10425- 10440.

50. Li T., He H., Wang Y., Zheng H., Skogerbø G. and Chen R. (2008). *In vivo* analysis of *Caenorhabditis elegans* noncoding RNA promoter motifs. *BMC Molecular Biology* **9**:71.

51. Fire A., Xu S., Montgomery K. M., Kostas A.S., Driver E.S. and Mello C.C. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806-811.

52. Cerutti H. and Casas-Mollano J.A. (2006). On the origin and functions of RNA-mediated silencing: from protists to man. *Current Genetics* **50**, 81-99.

53. Siomi H. and Siomi C. M. (2009). On the road to reading the RNA-interference code. *Nature* **457**, 396-404.

54. Martens H., Novotny J., Oberstrass J., Steck L. T., Postlethwait P. and Nellen W. (2002). RNAi in *Dictyostelium*: The role of RNA-directed RNA polymerases and double-stranded RNase. *Molecular Biology of the Cell* **13**, 445-453.

55. Fey P., Kowal S. A., Gaudet P., Pilcher E. K., and Chisholm L. R. (2007). Protocols for growth and development of *Dictyostelium discoideum*. *Nature protocols* **2**, 1307-1316.

56. Schuman S. (1994). Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA Topoisomerase. The journal of biological chemistry **269**, 32678-32684.

57. Manstein J. D., Schuster HP, Morandini P. and Hunt M.D. Cloning vectors for the production of proteins in Dictyostelium discoideum. *Gene* **162**, 129-134.

58. Sambrook J. and Russel D. W. (2001). Molecular Cloning – A laboratory manual, 3 edition (New York, Cold Spring Harbor Laboratory Press.

59. Howard K.P., Ahern G.K and Firtel A.R. (1988). Establishment of a transient expression system for *Dictyostelium discoideum*. *Nucleic Acids Research* **16**, 2613-2623.

60. Solomon M.J., Rupper A., Cardell A. J. and Isberg R.R. (2000). Intracellular growth of *Legionella pneumophila* in *Dictyostelium discoideum*, a system for genetic analysis of host-pathogen interactions. *Infection and immunity*, 2939-2947.

61. Feeley J.C., Gibson J.R., GormanW.G., Langford C.N., Rasheed K.J., Mackel D.C. and Charcoal-yeast extract agar: primary isolation medium for Legionella pneumophila. Journal of Clinical Microbiology **10**, 437-441.

62. Shevchuk P. and Steinert M. (2009). Host-Pathogen Interactions: Methods and Protocols, chapter 4. Screening of Virulence Traits in Legionella pneumophila and Analysis of the Host Susceptibility to Infection by Using the Dictyostelium Host Model System. 47-56. Institute for Microbiology, Technische Universität Braunschweig, Braunschweig, Germany

63. Li Z., Dugan S.A., Bloomfield G., Skelton J., Ivens A., Losick V. and Isberg R.R. The amoebal MAP kinase response to *Legionella pneumophila* is regulated by *DupA*. *Cell Host & Microbe* **6**, 253-267.

Appendix

Table 6. *List of all the oligonucleotides used in the project*

Oligo name	Type	Sequence 5' to 3'	Size (nt)	G/C%	T_m value
(489) Upstream 21-1	PCR primer	GTCGACCAAAGAGATCCTTCACCAACT	27	48,2	59,7
(490) Downstream 21-1	PCR primer	GGTACCCAACATGATAGTTCAACAAAT	27	37,0	55,2
(491) Upstream 21-2	PCR primer	GAGCTCCAAAGAGATCCTTCACCAACT	27	48,2	59,7
(492) Downstream 21-2	PCR primer	TCTAGACAACATGATAGTTCAACAAAT	27	29,6	52,1
(496) DUSEmut1-1	PCR primer	TTTTTGTCAAAACTATGGGACAAAATTATTTTTTTATTCAAAACTTC	47	21,3	59,3
(497) DUSEmut1-2	PCR primer	GAAGTTTTGAATAAAAAAATAATTTTGTCCCATAGTTTTGACAAAAA	47	21,3	59,3
(498) DUSEmut1-3	Analytical primer	TTTTTTTTTGTCAAAACTATGGG	23	26,1	46,4
(499) DUSEmut1-4	Analytical primer	GAATAAAAAAAATAATTTTGTCCC	23	21,7	44,6
(518) 21-loop1-for	PCR primer	CAGCAAACCCTACAGTCATTTCTATTGAAAACTACCGTCAACTGTC	47	40,4	67,2
(528) 21-loop2-for	PCR primer	CAGCAAACCCTACAGTCATTTCCAACGAAAAACTACCGTCAACTGTC	47	44,7	68,9
(519)21-loop-rev	PCR primer	GACAGTTGACGGTAGTTTTCGTTGGAAATGACTGTAGGGTTTGCTG	47	44,7	68,9
(516) pDXAfor	Sequencing primer	GGTATTGAAATGACCTCCGTTT	22	40,9	51,1
(517) pDXArev	Sequencing primer	CATGTGGTGTATCAAACTTTGT	22	36,4	49,2
(573) pDXArev2	Sequencing primer	CTTGAATTGATCCTCTAGATTAA	23	30,4	48,1
21DoubleConstruct	Sequencing primer	ATGTTGGGTACCGAGCTCC	19	57,9	53,2
(493) 21-1northern	Northern probe + Analytical	CAGTTGACGGTAGTTTTTCAATA	23	34,8	49,9
(494) 21northern	Northern probe	CAGTTGACGGTAGTTTTTCTTAT	23	34,8	49,9
(495) 21-2northern	Northern probe + Analytical	CAGTTGAGGGTAGTTTTTCGTTG	23	43,5	53,5

Table 7. An overview of the different extrachromosomal plasmids constructed

Name	Working name	Restriction sites	Promoter sequence	Loop sequence
pDXA_DdR21-1	v001	SalI / KpnI	CCC	ATAA
pDXA_DdR21-2	v002	SacI / XbaI	CCC	ATAA
pDXA_DdR21-2_GGG	v011	SalI / KpnI	GGG	ATAA
pDXA_DdR21-2*	v102	SacI / XbaI	GGG	CAAC
pDXA_DdR21-1**_GGG	v211	SalI / KpnI	CCC	TATT
pDXA_DdR21_DC	DC	SalI / KpnI + SacI / XbaI	GGG + CCC	CCC & GGG

Figure 19. Schematic figures of the plasmids used. The pCR 2.1 TOPO vector was used for making the modified DdR-21 genes and the pDXA-HC was used to insert the finished constructs *into D. discoideum*. The actin promoter in the pDXA-HC vector is not present in the modified versions used all through the project.