31-06-2019

Aalborg University

Functional analysis of microRNA-target gene interactions related to Intestinal Bowel Disease

by Sofie Raunbak

AALBORG UNIVERSITY
STUDENT REPORT

Main Supervisors: Meg Duroux

External supervisor: Bo Thomsen & Knud Larsen

AALBORG UNIVERSITY
STUDENT REPORT

Translational Medicine

Title:	Functional analysis of microRNA-target gene interactions related to Intestinal Bowel Disease
Semester:	4th Semester
Project period:	1/9-2018 - 31/6-2019
ECTS:	60 ECTS
Main supervisor:	Meg Duroux
External supervisor:	Bo Thomsen & Knud Larsen
Project group:	9033
Pages:	84
Characters:	137,098
Appendix:	7
Pages of appendix:	11 pages

Abstract

Introduction: The burden of intestinal bowel diseases is an increasing problem worldwide, and at the movement, the treatments available have insufficient long term effects. The regulation of expressions in 60% of the human genome is regulated by microRNA, and additionally, the regulation by miRNA have been associated with the development of several diseases. Due to this association researchers are investigating the use of microRNA as a possible treatment, but prediction and validation of functional pairs of miRNA target gene interactions have proven difficult.

Aim: To identify and validate functional pairs of miRNA candidate target genes interactions

Methods: This has been investigated by using PCR to amplify the investigated candidate target genes sequences, which was fused into the reporter plasmid psiCHECK2. The reporter system takes advantage of the expression of a transcript expressing either luciferase or Renilla protein regulated by the presence of a recognition target sequence in the 3'UTR region.

Results: This study found statistically significant differences in the expression of the reporter gene, which indicate interactions between the miRNAs, *miR-21-5p, miR-150-5p and miR-155-5p,* and their predicted targets *IL12A, ZEB1,* and *SOCS1*, respectively.

Conclusion: In conclusion, the regulations of the reporter gene expression show interactions between the miRNA and target genes investigated. This may contribute to the investigation of a long-term therapeutic treatment against IBD. However, further research related to the regulation of these target genes and the effect of this regulation is needed.

Abbreviations

BCR	B-cell receptor
CD	Crohns Disease
CNS	Central nervous system
DMEM	Dulbecco's modified Eagles Medium
ENS	Enteric nervous system
GI	Gastrointestinal
IBD	Intestinal bowel disease
IL	Inter Leukin
INF-ɣ	Interferon-gamma
LARII	Luciferase Assay Reporter 2
LY	Landrace and Yorkshire race
MHC	Major Histocompatibility complex
Mill	million
miRISC	microRNA silencing complex
miRNA	microRNA
NC	negative control
NK	Natural killer
PBS	phosphate buffered saline
PC	positive control
PCR	polymerase chain reaction
Pen/strep	Penicillin/Streptomycin
PLB	passive lysis buffer
TCR	t-cell receptor
UC	Ulcerative colitis

Table of content

1. Introduction

Inflammatory Bowel Disease (IBD), such as Crohn's disease (CD) and Ulcerative Colitis (UC), are a growing problem worldwide (1,2), and 0.5% of the general world population is estimated to suffer from IBD (2). It is a modern disease, which began in the western world, as this became more industrialized (1–3). The disease has evolved into a global burden, due to the fact that the estimated amount of residents with an IBD is 1 million (mill) in the USA and 2.5 mill in Europe (2), and with no way to treat the affected population effectively (1,2,4). Now, the incidence of IBD has reached a plateau in the western world, but in other continents like Asia, the previously low incidence is increasing (1–3). However, the development of IBD, in both incidence and prevalence, have been observed sporadically over the years, whereby only an estimate of the global burden can be made (2), but awareness of these diseases is increasing and combined with increased registration of disease, the number of patients affected gets more precise (2). IBD's are autoimmune chronic inflammatory disorders in the gastrointestinal tract, which related to phenotype change in pathology and clinical manifestation (4–6). Caused by this and the influence of environmental factors, the pathogenesis of IBD is poorly understood (3–5,7). However, the fact that genetic and environmental factors are influencing the pathogenesis of IBD is widely accepted (4,7). Over the last decades, the treatment possibilities have developed and different therapeutic approaches, like early intervention with immunomodulators and biological agents, have been introduced (1,7). The treatments have reduced the mortality and the risk of surgery at the moment, but both short-term and long-term effects of the new treatments have not revealed yet. However, the disease does not only burden the countries, but also the patient, who experience not only physical symptoms but the disease also impact on their possible career aspirations, quality of life and the social stigma (2,5). Due to the increased awareness of IBD, the burden hereof and the fact that no usable treatment is developed, recent studies have investigated new treatment possibilities, hereunder microRNA(miRNA). It is widely accepted that miRNA regulates gene expression, and by this, investigations with a focus on beneficial use of miRNA both as a biomarker and a therapeutic treatment have been conducted (8). Investigating miRNAs and their relation to IBD have been difficult, but still, over 100 genes have been associated with IBD (8). However, the link between the functionality and regulation of expression of the genes and IBD become clearer as knowledge of the pathogenesis increases. For further development of a therapeutic treatment against IBD, based on miRNA, more research focussing on the impact on the pathogenesis by genes and miRNA

is required, and furthermore, the impact of gene expression and the regulation hereof need more research.

1.1 The Gastro-Intestinal Tract

To understand the full influence of an IBD on the patient, an introduction to the structures and functions of the gastrointestinal (GI) tract is needed. In general, the main functions of the GI tract is to digest and absorb nutrients and thereby supply the entire body with the necessary components for it to function. This is performed by using smooth muscle peristalsis to move the content of the lumen; meanwhile, the content is exposed to absorptive epithelial on the mucosal surface. Furthermore, the epithelial mucosal surface is developed to prevent the transmucosal movement of antigenic, toxic and infectious agents. Both the digestion and the health of the GI tracts is coordinated by a series of neural events, which include the central nervous system (CNS), the enteric nervous system (ENS) and a special GI endocrine peptide, which can target both cells and tissue in the GI tract. Furthermore, special epithelial cells can be found throughout the entire GI tract, and control secretion, motility and regulate food intake, postprandial glucose levels and metabolisms. Immune functions, protecting the GI tract, involve inflammatory mediators, which influence the recruitment of immune cells to the GI tract, and also influences the activity of the neural network in the GI tract. Furthermore, the presence of 10^{14} microorganisms is estimated in the GI tract, which is involved in the development of the ENS and in general the overall health in the body, due to mechanisms activating the immune system. (9,10).

1.1.1 Anatomy of the GI tracts

The GI tract is composed of several organs and can be divided into two groups. The upper tracts, which refers to mouth, esophagus, stomach, duodenum, jejunum, and ileum, while the lower GI tract, which refers to the colon, rectum, and anus. (9,10).

The ingested food is moved from the pharynx, by peristalsis through the esophagus and into the stomach, where digestion enzymes and gastric acid disintegrate the food bolus. The digested material, the chime, is transported into the duodenum and is in the small intestines, which include the duodenum, jejunum, and ileum, further disintegrated. In the small intestines the proteins, fats, and carbohydrates are disintegrated to smaller components, which allow absorption. Salivary glands, pancreas, liver, and gallbladder assist this process, which contributes, with the secretion of e.g. gastric acid and hormones. From the small intestines, the

content enters the colon, where all remaining fluids and salt are absorbed and the faeces are prepared for expulsion through the rectum and anal canal. (9,10).

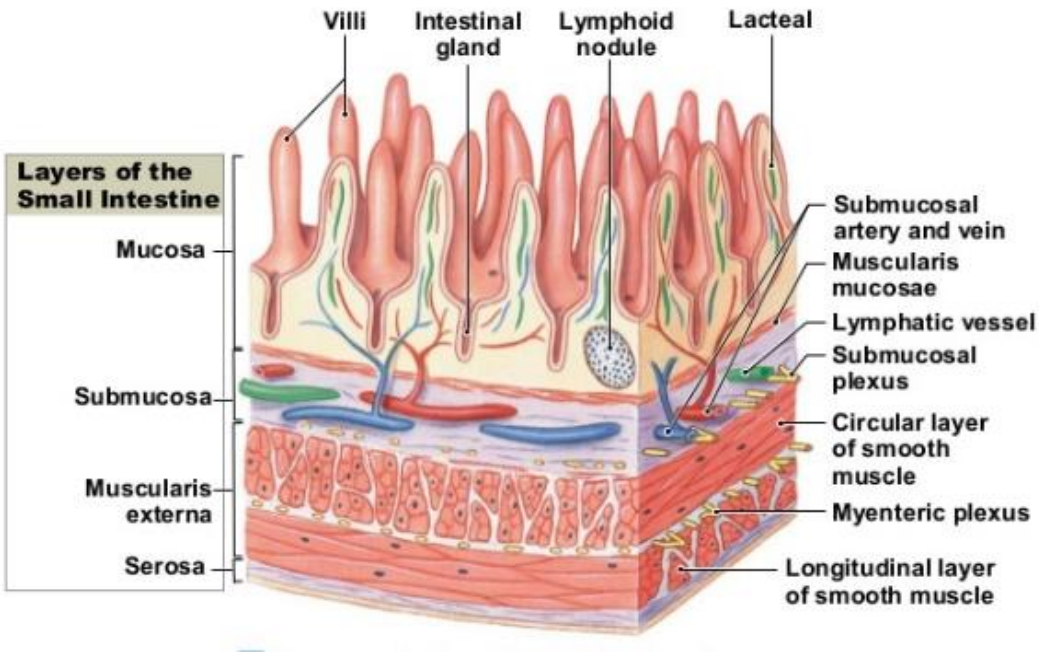


Figure 1.1, ***Structural layers of the GI tract.*** *Source: https://socratic.org/questions/from-the-lumen-outward-what-are-the-layers-of-the-gastrointestinal-tract*

To accomplish the digestive process of the intestines, the GI tract has several structure layers, as presented in *Figure 1.1, Structural layers of the GI tract,* which each have their function. This includes an inner mucosal layer, submucosal layer, smooth muscle layer, and the outer serosal layer. The inner mucosal layer consists of epithelial cells, and four types hereof is distributed in this layer. The cell types are paneth-, goblet- and enteroendocrine cells and enterocytes. The paneth cell, can be found in the bottom of the intestinal glands, and synthesize and secrete antimicrobial peptides and proteins. Furthermore, these antimicrobial molecules influence the balance of colonizing microbiota and the protection of the innate immune system. The goblet cells have the ability to secrete gel-forming mucus, which protects the inner wall of the intestines. The enteroendocrine cells produce hormones and peptides, as a response to stimuli, which are transported to the bloodstream and where they work systemically or stimulate the response of the ENS. Lastly, there are the enterocytes, which is an absorbing cell and amongst other absorbs water, sugar, and ions. The cells in the mucosal layer have the ability to sense the luminal content, which releases signal-molecules and through the circulation affect the adjacent cells and distinct neuronal pathways. The submucosal layer contains connective tissue, nerves and lymphatics, and the smooth muscle layer, consist of both longitudinal and

circular smooth muscles, and is responsible for the peristalsis of the intestines. The last layer is the serosa layer, which consists of connective tissue and is covert by mucus to prevent any damage to the surrounding tissue when the intestines are moving. (9,10).

1.1.2 The Gastro-intestinal barrier

The gastrointestinal tract is inhabited by 10^{14} microbes and additionally, it is exposed to dietary antigens daily (11). Beside digestion of nutrients, the GI tract participates in maintaining the immune homeostasis and keep a balance in the microbial environment (3).

By a balanced microbial environment, the GI tract is able to stay tolerant to different food groups, e.g. spicy food, and commensal bacteria and self-antigens. Furthermore, the GI tract can induce an immune response in the presence of harmful organisms, and protection of the GI tract is accomplished through the presence of an epithelial barrier and the mucosal immune system (3). *The mucosal immune system* can be divided into three parts, where the intestinal epithelial barrier is one of these. *The epithelial barrier* consists of a single layer of epithelial cells, which is organized into crypts and villi. The special epithelial cells, such as the goblet cell help establishing the barrier function by secretion of mucin and the secretion of antimicrobial proteins. The epithelial mucosal defense form the physical and chemical barrier, which keep gut content and cells apart. This limits the amount and the diversity of bacteria, which reach the epithelial surface and thereby met the underlying mucosa. To maintain the barrier the epithelial cells are in constant renewal. (3). However, by IBD being an autoimmune disease, the immune system in IBD is defective and react on self-antigens which cause the inflammatory process to launch and create tissue damage (3,12).

1.2 Clinical manifestation of IBD

IBD being an autoimmune chronic inflammatory process of the GI tract (4–7), manifests with a relapsing pattern and is often diagnosed in early childhood or adolescence (5,6). Studies have shown IBD to be one of the most common autoimmune inflammatory diseases just after rheumatoid arthritis (5), and since no usable treatment is developed, the disease has evolved into a burden worldwide (1,2,4). Normally IBD is characterized by the two phenotypes, CD and UC, which both lead to serious complications (5–7), but present similar symptoms like vomiting, diarrhea, and pain (5,6). CD can develop in the whole GI tract, and cause patches of inflammation in the deep muscle tissue(3,5–7), leading to severe complications like e.g. intestinal stricture formation, fistulization and the development of abscesses (6). Furthermore,

continuous abdominal pain, increased susceptibility to skin tags and perianal fistulas can be observed in CD patients (5). The manifestation of UC contrast, by developing in the superficial mucosal layer, and is only found in the colon (3,5–7). UC is presented by an ascending inflammation beginning from the rectum (5,6), by this UC patients experience rectal bleeding and sporadic pain (5). IBD is a very idiopathic disease, which is influenced by both environmental, genetic and immunological factors(3,5–7), and the state of the disease vary according to the mentioned factors and often according to specific food groups(5). Furthermore, IBD increases the risk of developing colon adenocarcinoma (5), which speculated to be caused by treatments with immunosuppressive over a longer period of time(5,6). However, as mentioned IBD is often diagnosed in early childhood or adolescent, and distinct differences have been observed in early-onset IBD and later-onset IBD (6). In children, UC is often represented by pancolitis and CD is presented by ileocolonic disease, but rarely just with ileum disease (6). Furthermore, the phenotype and genotype difference according to age. Evidence shows that children younger than the age of six years often have more severe infections, which raise concerns of immune-deficiency. (6). The changing behavior of the disease, along with the fact that the pathogenesis is not fully understood. This increases the difficulty of creating a treatment, but by the increasing knowledge of the pathogenesis occurrence of novel therapeutic approached (7).

1.2.1 Development of IBD

As mentioned, the pathogenesis is not fully understood but is influenced by both environmental, immunological and genetic factors which all contribute to the pathogenesis (3,7).

The changing epidemiology, which has been seen over time, have suggested environmental factors to influence the pathogenesis of IBD. Several studies suggest the westernized lifestyle and industrialization have great impact, due to the diet, hygiene status, microbial exposure, and pollutions, which follow the industrialization (1,2,4,13). As mentioned the microbiome of the gut influence the immune system by maintaining the balance in the GI tract. However, environmental exposure is altering the microbiome through time and thereby altering the risk of not only IBD but also diseases in general (13). The microbiome can in natural ways change in the light of what we eat, and changes during intestinal inflammation (13). From the food, we digest, macro- and micronutrients occur, which might influence the risk of IBD. The micronutrients have important roles in the GI tract, by maintaining the intestines and by contributing to the gut microbiome or the immune response in the GI tract (13). For instance,

zinc deficiency can be caused by a change in the diet, and reduce the integrity of the epithelial barrier and increase the permeability of the GI tract (13). In IBD, the inflammation is characterized by smaller diversities in the normal microbiota, which can cause colonization of either commensal or pathogenic bacteria (13), by which the increased colonization of bacteria in the GI tract might cause inflammation. In relation, studies investigating intestinal infection show, that resolution of the infection is linked with re-established microbiota and thereby the balance of the GI tract (13). In relation, the diversity of microbiota is enhanced by the use of antibiotics. A study investigating the microbiome in CD patients, exposed and non-exposed to antibiotics, show antibiotics to amplify the diversity (14), and additionally, another study suggests the fail of bacterial recovery after use of antibiotics (13,14). A shift in the diversity was seen only 3-4 days after the beginning of treatment with antibiotics, and antibiotics have been widely used in both developing and developed countries, whereby it is reasonable to consider antibiotics as an inducing factor for IBD (13). The environmental factor in the form of pollutions, related to the increased industrialization and urbanization has been suggested, due to IBD being more common in urban centers than developing unindustrialized countries(13). Many components have been associated with health conditions, where air pollution is one of them. A short-term exposure to the airborne particles in a have shown to increase the permeability of the GI tract and the innate immune response (13), while a long-term exposure has shown to increase the expression of pro-inflammatory cytokines and alter the microbiota, which as described previously, also influence the immune system in the GI tract (13). According to the development of IBD, an incidence study with a focus on air pollution have shown an increased diagnose of CD in regions with high NO_2 and higher development of UC in areas with high levels of SO_2 in the air (13,15). Another study indicates air pollution to influence the risk of IBD, but show no individual association for either CD or UC patients (13,16).

The immune-pathogenesis of IBD is divided in a three-stage model, which begins with luminal antigen penetration of the lamina propria of the inner tissue layer in the GI tract(4). The penetration by the antigen causes an impaired natural antigenic clearance, which gives a compensatory immune response (5).

The genetic distribution of IBD also contributes to part of the pathogenesis, even if not fully understood (5,7). In addition, the genetic explanation contributed to the investigation related to the risk of developing an IBD (5). Some studies reveal 12% of IBD patients have a family history with IBD (5) and now 20-25% of genetic factors contribute to disease heritability (6,7).

However, even if the inheritance patterns are not clear, a greater genetic influence is seen on the pathogenesis of CD, compared to UC, which relates to the age of onset and location of manifestation (7). Furthermore, it is suggested that the disease and phenotype manifested is heritable (7). The pathogenesis of IBD is very complex and involves genetic and environmental factors(4–7). The complexity of the pathogenesis and the factors influencing it is investigated and slowly being clarified(4,5). This has resulted in more than 100 genes being investigated and 50-100 genes suspected to influence the development of IBD (4,5,7). The first association between genes and IBD was made by Hugot et al. (5,17), who showed the development of CD being affected with an insertion mutation in the NOD2 gene (5). Later, GWAS (Genome Wide Association Studies) found associations of mutations in IL23R and the related signal mediators; JAK, STAT and IL12B, and TNF-α, the resulting cytokine (5,7). Furthermore, the genes IL-10, NLRP3 and CCR6 have been linked with IBD as inflammatory mediators, and the gene ATG16L1 has been linked with intestinal autophagy (5). Furthermore, GWAS found hereditary differences between ethnic groups, which the genetic-environmental interplay is more complex than first thought (5). The genetic predisposition of IBD is not easy to investigate since some genes are primarily detected in white individuals of European descent (2). An example hereof is the genes NOD2 and ATG16L1, which have been highly associated with the phenotype CD (2,5). Differences in the mutation of NOD2 is observed when comparing the European population with an Asian population, and the gene ATG16L1 is not associated with Crohn's disease in Asia (2).

The host genetic in children is known to have a predominant role in the diagnosis of IBD (6). Evidence suggests that both monogenic and multigenic disease involve genes, which are associated with immunodeficiencies (6). The genetic risk is found in MHCII (Major Histocompatibility Complex Class II) molecules, and NOD2 was a key discovery, which explains the link in the IBD pathogenesis, between the intestinal microbiota and the innate immune system (6). However, mutations in certain proteins, which control the mechanisms maintaining the anti-inflammatory state in the GI tract, make the bowels recessive to uncontrollable inflammation (5).

1.3 The Immune system

As mentioned, IBD is an autoimmune disease, which means the immune system reacts on self-antigens and thereby attacks the tissue and cells of the body, causing cell and tissue damage (9,10). However, to understand how the initiation of the chronic inflammatory state occurs, the understanding of the basic functions of the immune system is needed.

Initially, it is important to understand the immune system is divided into two parts, the innate and the adaptive immune system, which both are influenced by IBD (12). The innate immune system is the first response to a penetrating microorganism, and include the epithelial barrier, the mucosal layer of the epithelium and the innate immune cells. The innate immune system in contrast to the adaptive has no memory and by having a similar response every time activated, where the adaptive immune system is advancing and remember the microorganisms met before. Furthermore, the adaptive immune system, are advancing the response according to the time the microorganism invades the body, and for every turn, this particular microorganism enters the body, a quicker more powerful and specific response is created. However, establishing the adaptive immune system can take from days to weeks the first time it met a microorganism, where it only takes seconds to hours for the innate immune system to activate. Caused by this, the innate immune system protects the body until the adaptive immune system is ready. (12).

1.3.1 The Innate Immune system

As mentioned both the epithelial barrier, the mucosal layer and the innate immune cells are part of the initiating response to infection in the intestines (11,12). The epithelial barrier and mucosal layer function, as described previously, as a chemical and physical barrier against infectious agents (11,12). If this barrier should fail and harmful microorganisms should enter the tissue, the innate immune system is activated.

Normally, the innate immune system is activated by recognition of the infectious antigen, but in the case of IBD, it is a self-antigen that activates the immune system (11,12). However, the antigen is recognized by one of the five immune cells, which could be macrophages or neutrophil granulocytes, B1-cells or dendritic cells, or a Natural Killer (NK) cell (12). The innate immune system is activated by pattern recognition of molecules found on the surface of an infectious agent. When activated, secretion of chemokines, cytokines, and antimicrobial peptides occur (11,12). Caused by the secretion, recruitment of different innate immune cells and activation of the dendritic cell is initiated. In the infected tissue both macrophages and neutrophile granulocytes, NK-cells and B1-cells are present, along with the activated dendritic

cell. Both the macrophage and the neutrophil granulocyte is a phagocyte, and function by killing and devouring the infectious agent. They are assisted by the complement system, which increases the binding of the phagocyte to the infectious agent, with aid from 20 serum proteins. (11). The B1-cell produce natural antibodies against the antigens and additionally, function as an antigen-presenting cell. This is the same ability as a B-cell of the adaptive immune system, but in the adaptive, the B-cell has memory and is here called a B2-cell. The fourth cell of the innate immune system is the NK-cell, which primarily is used against virus infections, where the NK-cell kills the virus-infected cells. (11,12). The NK-cell is activated by the previously mentioned dendritic cell. Furthermore, have the ability to recognize patterns and activate both NK-cells and T-cells.

The T-cell is part of the adaptive immune system and is activated by the dendritic cell by the presentation of the antigen. First the dendritic cell bind to the antigen and evolve to a potent APC (antigen presenting cell), whereafter it migrates to the paracortex in the lymph nodes, which is the T-cells whereabouts (11,12). In the paracortex, the antigen is presented for a naïve T-cell, either by MHCI or MHCII, which activate the T-cell. By this, the innate immune system plays a major role in the adaptive immune response, which is activated after a couple of hours to days (11,12).

1.3.2 The Adaptive Immune system

The adaptive immune system can be activated after a couple of hours if an infection in the body is too tough for the innate immune system. The adaptive immune system includes two types of lymphocytes, T- and B2-cells, and the production of antibodies. The adaptive immune system is amongst others, activates by the dendritic cell by antigen presentation to the T-cell. The antigen is presented to either a $CD8^+$- or a $CD4^+$ T-cell, through their respective MHC classes, class one and two respectively, and a co-stimulation from the dendritic cell (11). The $CD4^+$ T-cell have many differential possibilities like differentiate to either Th1-, Th2-, Th17-, T-follicular-helper-cells or regulatory t-cells. The $CD8^+$ T-cell differentiate to a cytotoxic T-cell. The function of the evolved $CD4^+$ T-cell is to help the infection-fighting immune cells, by the production of e.g. IL-2, which makes the macrophages cell-killing mechanism more effective (12). The $CD8^+$ evolves to a fighting immune cell, and kill the cells that are mainly infected with viruses, or in other ways damaged cells (12). The B2-cells evolves by activation from the T-follicular-helper-cell, and mature to an antibody-producing plasma cell. Through the production of cytokines, the development of the B2-cell and the activity of the innate immune

cells, like macrophages, is controlled. Activation of the cytokine receptors increases the phagocyte activity, whereby it is easier to phagocyte and kill the microorganisms.

When activated, the T-cells, plasma cells and memory T- and B-cells leave the lymph nodes via the efferent lymphatics and migrate to the mucosal tissue where the immunological response was induced (11,12).

In relation, the adaptive immune system differs from the innate by recognizing antigens instead of molecular patterns, and remember the infectious agents according to their antigens. However, caused by this recognition method, the adaptive immune system needs to expand to be able to give the needed support to the innate immune system. This process takes time since only a few specific B- and T-cells perform the clonal expansion, which is the process where the B- and T-cell with the right antibody is multiplied (11,12). The specific B- and T-cells recognize the antigens with specific B-cell receptors (BCR) and T-cell receptors (TCR). The specific cells expand until enough cells have been made, and when the cells have reached the right size, attach on the infectious agent begins. The B-cells evolve to plasma cells and begin the production of antibodies against the infectious agent. Furthermore, the B-cells have the ability over time to produce better binding antibodies, which is called affinity maturation. This is a process caused by a genetic mutation in the genes coding for the antibodies, combined with a selection of the strongest antibodies produced. This mostly results in the response from the immune system being so effective, that the infectious agent is eliminated. (12).

However, the second time this infectious agent enters the body; the B- and T-cells recognize the antigens from the agent because of the memory. This cause a quicker reaction from the adaptive immune system, resulting in the infectious agent not expanding to a damaging level.

1.3.3 Autoimmunity

The immune system is a strong defense, but defects occur, which is the case for IBD (12). In the case of IBD, the immune responds on a self-antigen and attack the body tissue instead of an infectious agent. This can lead to chronic and life-threatening diseases, like IBD, and causing tissue damage and malfunction. Most types of autoimmune diseases are aiming at specific cell types in the blood or specific organs, like type-1 diabetes, which affects the β-cells in the pancreas, or in the case of IBD, where the GI tract is a target (12).

1.3.4 The autoimmune IBD

IBD is seen as an autoimmune disease, where CD is driven by TH1-cells, but no self-antigen is defined. However, due to the changing associations to different genes, scientists speculate in IBD being an immunodeficiency (6,12). By this, the interplay between microbiota and the immune system have an essential role (3,12). The strongest associated polymorphism involves the pattern recognizing receptor NOD2, and the proteins, ATG16L1 and IRGM, involved in the protection against intracellular bacteria (5,12). Genetic studies show changes in the genetic variations for each patient, by which the genetic mutations cannot be fully related to the disease development (7,12). Furthermore, the same mutations have been found in healthy individuals as well. (12)

However, studies show an insufficient differentiation of Th17-cells influence the immune-pathogenesis of CD (12). The gene for IL-23-receptor and five others, including IL-21 and TGF-β, are involved in the differentiation of Th17-cells, whereby these might be influenced in IBD's as well. Additionally, deficiencies in the IL-8 production in macrophages and accumulation of neutrophil granulocytes have been observed in areas of the GI tract infected by CD. This indicates, along with a high expression of NOD2 from paneth cells, in the small intestines, and the association between the NOD2-genotype and the expression of α-defensins from the paneth cells, that CD development is caused by insufficient bacterial suppression. (12).

1.4 Treatment of IBD

IBD is characterized from mild to severe IBD, and the treatment type for the patient is dependent on this characteristic (5). Some of the events, to which patients seek hospitalization, are symptoms in IBD like e.g. abdominal cramping and tenderness, bloody stools, substantial diarrhea and anemia (5). Typically, the patient is hospitalized, and kept under surveillance for diagnostic purposes (5).

1.4.1 Diagnostics of IBD

The diagnostic of IBD is a combination of several investigations, including radiological, histological and biochemical assays (5). With the results, the characteristic of the patients IBD is clarified (5). The characterization accumulated from mild to severe, and depend on the bowel movement per day, severity of symptoms, and how many times the patient has been and subsequent is in need of treatment (5).

The biochemical analysis is performed, as early investigations, and include blood count, C-reactive protein, liver function, and erythrocyte sedimentation rates (5). Furthermore, a diagnostic tool is the cytokine level in the blood (5), and both TNF-α, IL-12, IL-8, IL-4, and IL-18 have been found in high concentrations in patients with IBD (5). With the diagnostic measures, the two phenotypes UC and CD can be distinguished (5). CD is characterized by fissuring, cobblestoning, and generalized mucosal erythema, where UC differs by being characterized by ulceration, bleeding and luminal dilation (5). As a histological analysis, a biopsy of the rectal tissue is performed, which provides information on possible cellular changes that follow with IBD (5). The radiological assay is the least invasive method and provides imaging of the GI tract (5). This method can provide observations of areas in the small and the large intestines, which cannot be reached with an endoscopy (5).

1.4.2 Management of IBD

IBD has no treatment yet, and until developed, the clinicians manage to suppress the diseases. For the hospitalized patient, the hemodynamics of the body is monitored, and intravenous administration of both fluids, nutrients, and corticosteroids are provided (5). Furthermore, spicy, fried and low-fiber foods are restricted from the patient's diet to relieve both symptoms and diarrhea (5).

In the management of IBD, the focus lies on nutrition, anti-inflammatory and immunomodulatory agents including combinations thereof (5), which is the most common approach in the treatment against IBD (13). However, the treatments of IBD follows a “step-up” principle, where the characterization of the IBD is influencing the treatment method. The characterization is, as mentioned previously, mild, moderate and severe, and management of IBD is treated cumulatively related to the characteristics (5).

Mild IBD is often treated with anti-inflammatory compounds like 5-aminosalicylic acid, which is used for induction and maintenance of treatment of IBD (5). Sulfasalazine was the first 5-aminosalicylic treatment and is still in use today, but due to the side effects, about 30% of patients cannot tolerate the drug (5). However, the newer 5-aminosalicylic acid, mesalamine, is more tolerable for the patients and has better outcomes for patients with UC (5). The used treatment for the management of moderate or refractory mild IBD is immunosuppressives, like the corticosteroids Azathioprine, methotrexate, and budesonide (5). The corticosteroids have, like mesalamine against mild IBD, been found more effective in both UC and CD (5). However, long-term use of corticosteroids give some concerns, since they have been proven as a

contributor to infection in patients due to their immunosuppressive effects (5). The immunomodulatory drugs, or immunosuppressives, contrast to the anti-inflammatory drugs, by reducing the inflammatory response, and inhibit both proliferation and activation of lymphocytes (5). The main immunosuppressive drugs used against IBD is Azathioprine, 6-mercaptopurine, tacrolimus, and methotrexate (5). The use of thiopurines have been found to increase the effect of anti-TNF therapy, but it is unclear whether the treatments are synergistic, or if the thiopurine therapy works independently (5). However, the tolerance of thiopurine is very individual inpatient, and a study conducted by Chaparro et al. showed that about 17% of patients receiving this treatment interrupted the treatment due to adverse effects. Furthermore, the treatments at some point have been withdrawn, which caused high rates of IBD recurrence. Caused by this, if tolerated by the patient, continuing the treatment is preferred (5). Anti-Tnf therapy is proven effective and safe to use, but the agents are only effective in two-thirds of the patients (5). In addition, the treatment is more expensive, caused by the fact that the drugs are administered by injection, whereby the patient needs regular contact with a doctor (5). The drug methotrexate, which as mentioned is one of the main immunosuppressives used against IBD, is used against a brought spectrum of chronic inflammatory diseases (5). In CD patient, the drug has proved a 65% remission, when compared with a placebo, but more clinical trials are still needed with a focus on the effects of methotrexate against IBD (5). The severe or refractory moderate IBD is managed with biologic agents, like molecule inhibitors and immunoglobulins, and the last stage, severe refractory IBD, requires surgical resection of the infected bowel (5). However, due to the recent advances in the treatment of IBD, where new immunomodulators and biological agents have been introduced, the need for surgery has been reduced (1,5). A study by Lakatos et al, have investigated the surgical rate of IBD between 1977 and 2006, which were 14.6 %, 30.1% and 51.6% after 1, 5 and 10 years of disease in a patient diagnosed before 1999, respectively (5). A lower surgical rate was observed in patients diagnosed after 1999, and the lower surgical rates were associated with increased early use of the immunomodulatory Azathioprine (1,5). New understandings of the immune-pathogenesis caused changes in treatment methods, which including targeted therapy, are evident in decreasing the surgical rate (1,6). However, the side effects of the immunomodulatory and biologic agents are increasing the complication rates, and the impact hereof on cancer risk and the mortality is yet to be discovered (1). In general, the mortality and risk of cancer are increased in CD patients, wherein contrast the mortality is stable in UC patients (1).

1.4.3 Microbiota as IBD treatment

As mentioned, both anti-inflammatory drugs and immunomodulatory drugs are used against IBD, but clinical evidence shows that microbiota in genetically susceptible hosts induces chronic intestinal inflammation (5). Caused by this, therapeutic manipulation of the intestinal microbial imbalance have been investigated in the further research of usable treatments against IBD and might be a beneficial approach (5). The therapeutic manipulation of the microbiota involves both individual and combined use of antibiotics, probiotics, and prebiotics in different combinations (5). However, not all antibiotics have been found useful against IBD, but antibiotics like metronidazole, ciprofloxacin, and clarithromycin have (5). The probiotics are recommended as both a preventive and a therapeutic measure since different strains have proven to increase the integrity of the intestinal barrier, induce immune tolerance, reduce GI infections and prevent pathogenic bacterial translocation (5). By combining antibiotics and probiotics it might be possible to both knocks down the existing infection and prevent a new. Furthermore, the probiotics have the ability to regulate the cell response and the intestinal immunity in the GI tract (5). Studies have demonstrated probiotics to reduce inflammation by suppressing IL-17 production and inducing regulatory T-cell accumulation in the secondary lymphoid organs, spleen, and lymph nodes, and in another study, a mixture of probiotics have been demonstrated to relieve the symptoms of IBD (5). Restoring the healthy composition of the microbiome in the gut might result in a reduction of inflammation and might benefit IBD patients (5). However, even if the treatments are promising, more research in the area is needed.

The prebiotics is a way of altering the diet to assist the beneficial microbiome and suppress the pathogenic microbiome. However, no evidence supporting the theory is available, and more research necessary to prove the efficacy of this treatment (5).

1.4.4 Nutritional therapy

For clinical management of IBD, nutrition therapy has been studied, and evidence suggests that changing the diet affect the gut microbiome (5). Some microbe species are believed to influence the pathogenesis of IBD, whereby changing the diet might influence this (5). Studies have proven dietary formulations to be effective as pediatric CD treatment, but further investigation is needed of the effects in adults (5). However, nutritional deficiencies are observed in IBD patients, to which formulated diets might help against (5).

1.4.5 Mortality of IBD

The mortality of IBD is not very high, but the illness itself has, as mentioned previously, a great impact on the society and the patient self (1,2,4). Furthermore, with the new treatment, and the new use of immunomodulators and biologic agent, the impact on mortality is unknown (1). The situation now is that the mortality rates and rates of intestinal cancer have been decreasing, whereby the reduction of the inflammation burden and spread of the disease in the bowels, have benefitted from the new treatment (1). The mortality of IBD patients was investigated in a large study performed in Denmark, which compares data from the general population with data from patients with UC and CD (1). The study showed an increase in the mortality of approximately 10% in UC and 50% in CD patients when compared with the general population (1). Even if the mortality is still higher, a reduction in e.g. UC have been seen, due to the decrease from gastrointestinal disorders and colorectal cancer, which often is a complication of UC (1). However, the subgroups of patients, who experience immense disease in the first few years after diagnostics, may be at greater risk of dying from IBD (1). Due to the reduction in mortality, we might in the future see severe side effects including mortality, caused by the aggressive medical treatments used (1).

Furthermore, the cost of the treatment with biological agents is US$25,000 per patient (2), and when adding the increasing incidence, especially in countries like India and China, who have a population of over 1 billion each (2), the burden of the world become massive. Not only the health-care system will be overloaded, but also the economic costs of the treatment can become a problem, and maybe create diversity in the society with people who can afford the treatment, and people who cannot.

1.5 Current approach to IBD

The integration of bioinformatics and genome profiling analysis has contributed to the pathogenesis and in general the investigation of a usable treatment for IBD. The genome profiling analysis has helped illuminate the functions of miRNA, like e.g. miR-214, which activate inflammatory responses and upregulate the pathway of STAT3 in UC patients and thereby have revealed a possible target for the therapeutic treatment of IBD. Furthermore, it is acknowledged that both the genome, gut microbiome, immune system and the exposomes (definition of the total non-genetic environmental exposure factors), can all be used as a possible target for therapeutic treatment (13). However, not all possible targets are preferred

for targeted therapy, since e.g. genomic deletion or replacement in the mutation/variation causing of the disease have been technically feasible in animal studies, but since humans have a more complex genome the risk of replacing or deleting variations is unknown (13).

Due to the genetic dispositions present in IBD, the search for treatment against IBD has led to research of miRNAs. MiRNA is a non-coding RNA, at approximately 22 nucleotides of length, which have the ability to influence the regulation of gene expression (18). Primarily miRNA induces translation repression and mRNA degradation by interacting with the 3' untranslated region (3'UTR) of the target mRNA, but also activate translation and regulate transcription (7,18). Furthermore, interactions with 5' UTR, coding sequences and gene promoters have been reported (18), whereby this might be a beneficial applicant in the search for an IBD treatment.

1.5.1 Biogenesis of miRNA

miRNA is transcribed from DNA sequences, and undergo three stages of development, which are primary miRNA, precursor miRNA and mature miRNA (18). The development of miRNA follows one of two pathways, either the canonical dominant pathway or the non-canonical pathway, but both pathways lead to functional miRNA silencing complexes (miRISC) (18).

In the canonical pathway, the primary miRNA is transcribed and cleaved by a microprocessor complex to produce the precursor miRNA (13). The precursor miRNA is exported to the cytoplasm with aid from Exportin5 and produce the mature miRNA duplex. From the duplex, either the 5p or the 3p strand is loaded into AGO as a guide strand, which also determines the name of the miRNA, and form miRISC (13). In the non-canonical pathway, small hairpin RNA is cleaved by the microprocessor complex and exported with Exportin 5, which transport them to the cytoplasm where they by mature with aid from AGO2-dependent cleavage. (18).

1.5.2 Interaction between miRNA and target mRNA

The interaction between miRNA and their target gene depend on factors like the location of the miRNA subcellular, the amount of miRNA and target, and the affinity between the miRNA-mRNA interactions (18). Normally, the miRNA is secreted into the extracellular fluids, and from here, either by aid from a protein or exosome, transported to the target cell (13). However, recent studies suggest the occurrence of miRNA being transported from one subcellular location to another, and by this controlling both translation- and transcription rate (13). MiRNA was first discovered in 1993 by the research of the gene lin-4 (13). A mutation of this gene was proven to influence another gene, by post-transcriptional regulation, and lin-4 was discovered

as a small non-coding RNA and not a protein-coding RNA as first expected (13). By this, further investigation of miRNAs have been performed, leading to the detection of miRNA in animal models, with many miRNA conserved across species. (18). The role of miRNA influence on gene expression are well recognized and critical for the development but have also been associated with the development of many diseases (7,18).

1.5.3 miRNA gene regulation

As mentioned, miRNA bind to the 3' UTR of their target, but can also bind to other mRNA regions and by this, influence the expression of the gene (18). Binding to the 3'UTR cause translational repression and mRNA degradation, while binding to the 5' UTR and the coding region only have silencing effect on the gene expression, where in contrast, interactions with the promoter region induced the transcription (7,18).

Silencing effect of the mRNA is performed by the miRISC, which consist of a guide strand and AGO, and the target specificity depends on the interaction between the miRNA and the complementary sequence on the target (13). The complementary interaction determines whether an AGO-dependent slicing of the target or if miRISC translation inhibition and mRNA degradation happen. (18).

Translation activation has been observed by upregulation of the gene expression under the influence of miRNA. However, the association between the increased gene expression and miRNA have only been observed in the serum-starved cell, while the miRNA inhibits the expression in proliferating cells (13). Other examples of increased gene expression include the binding of 5'UTR on mRNA, but here during amino acid starvation (13). This indicates the upregulation of gene expression only happen under special circumstances. (18).

Transcription and post-transcriptional regulation were found to be influenced by the miRISC in the nucleus (18). However, the understanding of how this occurs and the mechanisms behind are unknown, and still, need to research. (18).

The dynamic of the miRNA mediated gene regulation influence the gene expression to find a steady state (18). However, as previously mentioned, different factors contribute to the regulation, hereunder the location of miRNAs, which can also be observed in the extracellular fluid (18). By their presence in the extracellular fluids, the miRNA can be used as a biomarker, due to their relation to diseases (3,18). In contrast, the search of candidate genes is difficult, whereby even if 100 susceptible genes have been found related to IBD, most of them have not yet been functionally linked to the disease (7). With miRNA regulation of gene expression

being a powerful regulator several studies, transfecting miRNA mimic into different cell lines, were conducted (18). However, more research focusing on target specificity, function and miReceptor is needed (18).

1.5.4 miRNA and IBD

As known, the GI tracts have the ability to absorb nutrients, and additionally, secrete waste products into the intestines (19). Furthermore, the intestinal barrier, consisting of the epithelial barrier and the mucosal layer, prevent infectious agents from entering the submucosal layer. The breakdown of this barrier is essential in IBD, but difficulties in the identification of signaling in the immune response and the inflammatory cascade, and additionally the regulation hereof are present (19).

However, miRNAs are involved in several autoimmune diseases, hereunder beside IBD, rheumatoid arthritis and SLE (19), and miRNA thereby have an essential role in the treatment of the immune-related diseases (19). According to Zang et al. (19), the present investigations are focussing on how miRNA relates to intestinal disease like IBD (19), and in 2008 Wu et al. reported the presence of miR-192 in healthy individuals, but not in UC patients (19)). In addition, Bian et al (19,20)demonstrated a significant increase of miR-150 in UC patients (19). Both studies demonstrate differences in the expression of miRNA when comparing healthy individuals with UC patient, where other studies additionally have demonstrated the role of miRNA related to the immune system, and hereunder their influence on the function of B- and T-cells (19). For instance, a study(19,21)found a deficiency in Dicer, which is one of the enzymes processing miRNA, which resulted in a decreased development of T-cells (15). Several miRNAs. E.g. miR-150, miR-155 and miR181a, are involved in the regulation of both B- and T-cells (19), and for instance overexpression of e.g. miR-150, have shown to cause blocking in the development of B-cells (19). Furthermore, the pattern recognition receptors in the innate immune system like Toll-like receptors, which are found on both macrophages and dendritic cells, are regulated by miRNA. Loss of miR-155 in dendritic cells has been showed to cause an impairment of the dendritic cells antigen-presentation and their ability to create a co-stimulation in the T-cell (19). By this, the t-cell will not activate and develop, since bot an antigen and a co-stimulation is needed for this (13,19). The target gene of miR-155 is the gene SOCS1, which regulate the antigen-presentation in the dendric cell, and by deregulation, the impairment in the dendritic cells function is caused (19). Furthermore, Wu et al. (19)found miR-21 increased in UC patients compared with the healthy individual (19). The findings stated from the different studies indicate miRNAs involvement in the pathogenesis of IBD, and by

targeting the right gene, miRNA might contribute to the improvement of IBD. However, even if some studies find miRNA to suppress removal of bacteria by targeting ATG16L1, or activation of NOD2 resulting in pro-inflammatory and anti-bacterial molecule production, the targeting of the miRNA prove troublesome.

1.5.5 miRNA targeting

In the case of IBD, the dysregulation of miRNAs and their target genes are vital to identify (19), and according to published data (19,22,23)), miR-21 is the only miRNA which is consistently upregulated in the tissue or serum of IBD patients, whereby target genes for miR-21 has good potential as a biomarker (19). In addition, approximately 5 and 6 expressed miRNA have been found in the mucosa or serum of CD patients with active disease, when comparing to patients with inactive disease (19). However, the influence of miRNA in IBD has not yet been confirmed, but miRNA seems to have a promising approach in the treatment against IBD (19).

As mentioned, miRNA is one of the non-coding RNAs, and is found to regulate both developmental features and are related to disease development (24). The miRNAs regulate protein-coding genes by post-transcriptional regulation of the mRNA complementary to the miRNA. Approximately 60% of the protein-coding genes are thought to be regulated by miRNA, and due to research the list with known miRNA is long and still increasing (24). However, even if the genome is estimated to contain more than 1000 miRNAs, not many regulatory targets are known (24). This cause difficulty in the prediction and validation of miRNA targets. Moreover, the biology of the miRNAs has poor understanding and both experimental and computational studies have been conducted to find a more beneficial approach (24). Due to miRNA:mRNA pairing, site location, conservation, site accessibility and the expression profile of the miRNA, the prediction and validation become difficult (24).

miRNA:mRNA pairing is one of the most important features in a target recognition, where the complementary site in the mRNA is called the seed site. Two types of seed sites have been found, the stringent seed, which have a 100% nucleotide match to the miRNA and the moderate-stringent seed which can have mismatches (24). The stringent seed site has the advantage of being very specific but might miss potential targets, where the moderate-stringent seed site increases the sensitivity and find more potential targets, but have an increased number of false-positives (24). As for the *site location*, most target sites are placed within the 3' UTR, but theoretically, the miRISC have the ability to bind to any segment of the mRNA. However,

seed sites in the coding sequence and the 5' UTR can cause downregulation, since the miRISC compete with both ribosome complexes and translation initiative complexes (24). Due to this, the 3' UTR might be more accessible for longer binding and thereby an optimal regulation by the miRNA. Therefore most studies are focusing on the 3' UTR target sites, and the data available might be biased hereof (24). As mentioned previously, the conservation of the miRNA including their targets can be seen across species, but some miRNA are species-specific, and a study (24,25)show that approximately 30% of the valid targets found at this moment might not be conserved, which can cause difficulties in the investigation of the function and biology of the miRNA (24). By filtering the predicted targets site according to conservation across species, the possibility of false-positive target sites might decrease, but this method only functions on conserved miRNA and thereby possible target sites might not be found (24).

Furthermore, the *site accessibility* of the mRNA is relevant, since the secondary structure of the mRNA affects the interaction between the miRNA and the mRNA (24). By this, the target site need an open structure, and additionally, the most effective target sites often have a prosperous amount of A:U nucleotides approximately 30 nucleotides both up- and downstream from the target site (24). To find the best seeding site an analysis of the folding pattern of the mRNA is preferred, but this requires a fair amount of time and computer power, and with this only 90-95% folding analysis is correct, and result in only 50-70% of the base pairs being correct, by which the analysis of the folding pattern cause difficulties as well. (24).

When predicting and validating target sites, the *expression profile* is important, since miRNA often regulate more than one gene, and the expression profile can change related to the gene which is regulated, and the expression of a certain gene may also change related to the different tissue it is found in (24). By this filtering based on the expression profile of the gene would be an effective approach and additionally also reduce the possibility of false-positive rates, but the filtering of the target sites according to the expression profiles also exclude targets since some targets only show an effect on protein level (24).

The tools used to predict and validate the miRNA expression are databases like TargetScan, Mirdb, mirTarbase, and others are poorly comparable since not much overlapping data is present(24). Furthermore, none of the prediction tools available have incorporated the optimal distance between the miRNA sites, which is 17-35 nucleotides between the sites, and supplementary sites in the 5' UTR have not been incorporated as well (24). However, Target

scan, which is used in this study, show the best performance in comparing between the databases, but only consider stringent seeds, whereby not all potential target are found (24).

To summarize, the functional miRNA targets are challenging to validate, since no prediction tool is able to incorporate all the known features relevant for the function of the miRNA. Furthermore, the prediction of target interactions between different miRNAs cause trouble, and the expression profile of the miRNA has a changing expression between both cell types and are changing according to the cellular conditions (24). Thereby, determining how miRNA expression affects the gene regulation in diseases like IBD cause major difficulties. (24). Even so, clinical trials regarding the therapeutic efficacy of miRNAs against IBD have been performed (19), but future studies are needed to provide a basis for the clinical trials and thereby elucidate miRNA-based therapies in IBD (19).

2. Problem delimitation

Intestinal Bowel diseases are, as accounted for in this study, an increasing, and soon to be, global disease, with approximately 0.5 % of the general population affected by IBD. Presently the treatments available for IBD are antibiotics, immunosuppressives, biological agents and in the worst case, surgery is needed. However, with these treatments, both complications and consequences follow since e.g. antibiotics contribute to decreased diversity of the microbiome. This causes colonization of both commensal and harmful bacteria in the intestines and worsens the condition. Furthermore, the use of immunosuppressives have been suspected of contributing to infection, after long-term treatment, and since IBD is a chronic relapsing disease, long-term treatment is needed. Additionally, immunosuppressives inhibit the immune response, not only in the GI tracts but also in the entire body. The adverse effect of this treatment also caused 17% of patients who received the treatment to stop treatment, by which an increase in relapse of the disease was seen. The last option of treatment is surgery, where the infected part of the intestines is removed, but with this complication with malabsorption, which influence the general body functions, and the microbiota in the intestines can be affected. The microbiota helps to maintain the balance in the GI tracts and contribute to the immune response as well, by which a worsening of the disease can be observed. In addition, the surgery might not always be effective, since a relapse of the disease can occur in another part of the intestines. The patients affected by IBD do not only face the physical challenges caused by the manifestation of the disease, like e.g. diarrhea and abdominal pain, but also their social stigma, possible carrier aspirations and in their general quality of life. Additionally, the challenges which follow the disease can also cause a psychological embarrassment.

Caused by the presented problems, the awareness and investigational focus on IBD has increased, and several studies have been conducted in the search of a new therapeutic approach against IBD. For this, the focus lies on miRNA, since associations have been found between miRNA and autoimmune diseases, like rheumatoid arthritis. miRNA has been proven to regulate gene expression, and theoretically, by inhibiting or increasing the regulation by miRNA the processes initiating IBD can be blocked. However, the prediction and validation of targets are challenging. Studies revealed several miRNAs and proven certain miRNA to influence IBD, like e.g. miR-21, which also prove the beneficial approach of this perspective.

However, further research is needed to reduce the challenges found in the prediction and validation of miRNA.

The investigations conducted in this study is based on the presented literature and data provided by a previously non-published study performed in 2017 at the Department of Molecular Biology and Genetics at Aarhus University by student Astrid Nielsen. In her previous study, by applying a newly developed porcine model, Astrid Nielsen was able to investigate the influence of miRNA's on the pathogenesis of Intestinal Bowel Disease, more specifically Ulcerative Colitis (UC). The porcine model, based on feeding piglets with Dextran Sodium Sulphate (DSS), resulted in the piglet's gastrointestinal system mimicking the gastrointestinal system of a patient with UC. Thus measuring the regulation of expressed miRNAs in the porcine model gave an indication of whether miRNA's had an influence on the development and effect of UC the UC affected tissue. In the study, 12 miRNA's were investigated, but only six miRNAs, *miR-125a-5p, miR-30b-5p, miR-27b-3p, miR-21-5p, miR-150-5p and miR-155-5p,* were assessed relevant by a significant regulation of their expression in the porcine model, which indicate these six miRNAs have an influence on the development and function of UC.

By comparison, of the results presented in the study by Astrid Nielsen, the literary finding and the data available by different prediction databases, three miRNA's were selected for further investigation. Due to the higher regulative differences in DSS treated piglets and the non-treated controls, and the literature suggesting an influenced expression in studies conducted related to IBD and the immune system, the miRNAs, *miR-155-5p, miR-150-5p and miR-21-5p*, chosen for further investigation.

3. Aim of study

The aim of this master thesis is to identify and validate functional pairs of miRNA target genes interactions in the transcriptional response to intestinal bowel disease, by combining available literary studies and experimental work. This has been investigated by using PCR to amplify the investigated candidate target sequences, which was then fused into the reporter plasmid psiCHECK2, by using restriction enzymes and ligation. This reporter system takes advantage of the expression of a transcript expressing wither luciferase or Renilla protein regulated by the presence of a recognition target sequence in the 3'UTR region. The plasmid DNA was co-transfected into cultured HEK293 cells with either miRNA Mimic or miRNA power inhibitor related to the tested miRNA.

4. Materials and Methods

4.1 Prediction of candidate target sequences for miRNA

The data relevant for the three miRNA's, *miR-155-5p, miR-150-5p, and miR-21-5p,* investigated in this study, were retrieved and further evaluated by using data available in different databases, such as miRWalk, TargetScan, GeneCards, and the NCBI database, and by comparing the data between the databases.

Initially, a search in the database miRWalk (Medical Research Center and Medical Faculty Mannheim, version 3.0) was conducted with each of the miRNA's, *miR-155-5p, miR-150-5p and miR-21-5p,* which revealed all the target genes related to each of the miRNA's. The search was conducted with human sequences since no porcine sequences were available in the database. Furthermore, cross-referencing with human sequences makes it possible to relate the investigation of each miRNA to the pathogenesis in humans. The validation and available information were sorted in miRWalk, by filtering the data according to other databases, namely TargetScan Human 7.2 (Whitehead Institute for Biomedical Research), Mirdb and Mirtarbase (OmicX). By cross-referencing data between several databases, the data retrieved is expected to be more reliable. An additional search in TargetScan was conducted to investigate the conservation of the target site in human and pig. The target sites not found in both human and pig were not investigated further. In *Figure 4.1, Alignment of 3'UTR sequences from different species containing a recognition sequence for the microRNA miR-150-5p, and appendix A*, TargetScan reveals a target sequence and compare the sequence cross-species, thereby stating the presence of the target site in both species and the similarities of the target sites by looking at the nucleotide match. To lower the possibility of changing the function of the target sites, cross-species, only the target sites with a 100% nucleotide match was considered for further investigation. Furthermore, the translation direction of each target sites was relevant for this study, whereby this was investigated with the database TargetScan, illustrated in *Figure 4.2, Example of pairing with target site,* and *Appendix B.* The target sites without the correct translation direction from 5' UTR to 3' UTR were, as the target sites without 100% nucleotide match, excluded from this study.

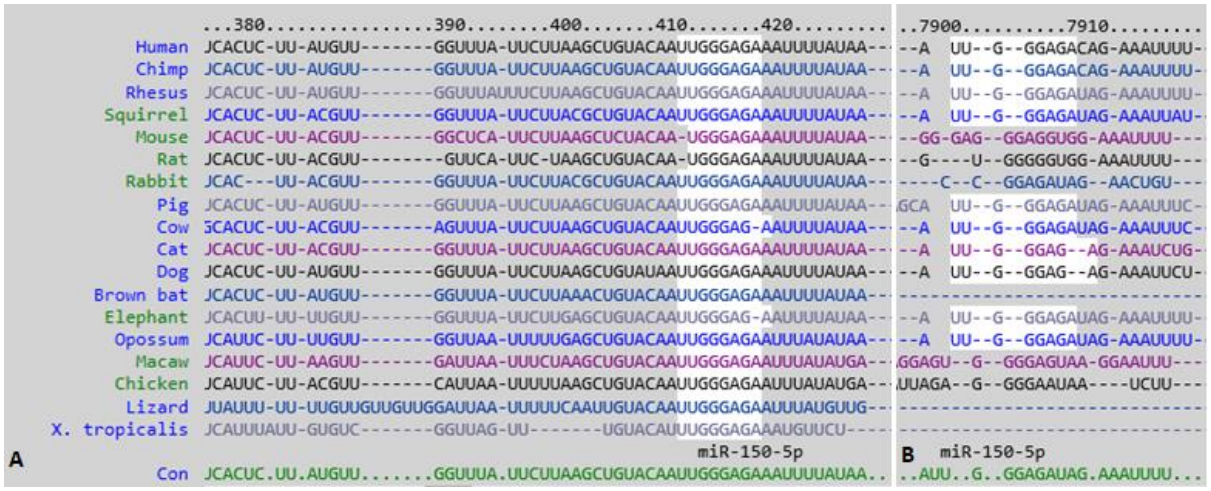


Figure 4.2, ***Alignment of 3'UTR sequences from different species containing a recognition sequence for the microRNA miR-150-5p.*** *In figure A, the target sequence for ZEB1 is presented, which is recognized by the miRNA, miR-150-5p. All species where the target sequence is present is illustrated by the white markings, including pig and human. In addition, the hyphens found in the nucleotide sequences, represent sites with nucleotide mismatches. Both human and pigs contain a 100% match of the target site, whereby this is used for further investigation. However, figure B displays the matchings of the target sequence for CBL, where multiple nucleotide mismatches are observed. This meaning, that even if the target site is present in both human and pig, this gene is excluded, due to the multiple nucleotide mismatches in the gene sequence.*

	Predicted consequential pairing of target region (top) and miRNA (bottom)	Site type	Context++ score	Context++ score percentile	Weighted context++ score	Conserved branch length	P_{CT}
Position 412-419 of ZEB1 3' UTR	5' ...UCUUAAGCUGUACAA---UUGGGAGA...	8mer	-0.22	98	-0.22	4.276	< 0.1
hsa-miR-150-5p	3' GUGACCAUGUUCCCAACCCUCU						

Figure 4.3, ***Example of pairing with target sites.*** *The target site is presented from 5' UTR to 3' UTR, and the pairing of the miRNA, miR-155-5p, is shown in the complementary sequence.*

The function of the remaining target sites was found by using Gene Cards (GeneCards.org, Weizmann Institute of Science, v4.8.2 Build 6). Evaluation of the functions, interacting pathways and interacting proteins for each target site, and their effect on the inflammatory processes or the immune system was conducted, where only the target sites influencing the inflammatory process or the immune system was included in this study, as a candidate sequence with recognition from a miRNA. This was executed to ensure the target sites investigated in this study might have a relation to Intestinal Bowel Diseases (IBD's).

Sequence information related to each candidate sequence was retrieved from the NCBI nucleotide database. This yielded information about translation and placement of the candidate sites in genomic porcine DNA. The cDNA sequence related to each candidate sequence was, retrieved from the NCBI nucleotide database, used to perform a BLAST search in the program Ensembl BLAST. By performing a BLAST search with cDNA from each candidate sequence,

several suggestions of a genomic sequence were found, as presented in *Figure 4.3, Example of Result hits from Ensembl BLAST*. The blast hit with the longest sequence match, and the highest percentage ID, was processed by configuring the data to show 500 bp in each direction from the candidate site, and by checking the presence of the candidate site for the respective miRNAs in this sequence.

Show All entries Show/hide columns (2 hidden) Filter

Genomic Location	Overlapping Gene(s)	Orientation	Query start	Query end	Length	Score	E-val	%ID
10:42213602-42216139 [Sequence]	ZEB1	Forward	2789	5324	2538 [Sequence]	4904.0	0.0e+00	99.92 [Alignment]
10:42209449-42211261 [Sequence]	ZEB1	Forward	794	2607	1814 [Sequence]	3546.0	0.0e+00	99.94 [Alignment]
10:42192452-42192654 [Sequence]	ZEB1	Forward	485	687	203 [Sequence]	395.0	2.7e-110	100.00 [Alignment]
10:42147410-42147610 [Sequence]	ZEB1	Forward	59	259	201 [Sequence]	392.0	2.4e-109	100.00 [Alignment]
10:42213280-42213460 [Sequence]	ZEB1	Forward	2608	2788	181 [Sequence]	353.0	2.2e-97	100.00 [Alignment]
10:42185715-42185876 [Sequence]	ZEB1	Forward	323	484	162 [Sequence]	317.0	1.3e-86	100.00 [Alignment]
10:42194808-42194913 [Sequence]	ZEB1	Forward	688	793	106 [Sequence]	206.0	2.8e-53	100.00 [Alignment]
15:7625862-7626006 [Sequence]	ZEB2	Forward	2830	2974	145 [Sequence]	196.0	4.1e-50	82.07 [Alignment]
15:7615978-7616106 [Sequence]	ZEB2	Forward	811	945	135 [Sequence]	157.0	1.4e-38	76.30 [Alignment]

*Figure 4.4, **Example of Result hits from Ensembl BLAST.** A list of result hits is presented, where all influencing factors for the genomic sequence selection are seen, both the length of the sequence (yellow arrow), percentage ID (green arrow) and overlapping Genes (black arrow) if any is present.*

4.2 Primer designing for each of the candidate target sequences

For DNA amplification, with Polymerase Chain Reaction (PCR), a specific primer set was designed manually for each of the four candidate sites relevant for further investigation. General guidelines, displayed in *Appendix C*, were used to design the specific primer sets for each of the candidate sites.

The primers were designed to amplify a 3' UTR sequence containing the target-binding site for each miRNA, respectively. The genomic DNA sequence, previously found with Ensembl BLAST, was used to design primer sites, which comply with the general guidelines, for each candidate sequence. The primer sites were contemplated to be placed at a distance of 200-250 bp, both upstream and downstream from the candidate site. The candidate target sites were observed at variable places in the genome, where some were localized 250 bp downstream of the stop codon in the 3'UTR, and others only 23 bp downstream from the stop codon. While designing the primer sets, the nucleotide grouping of the primers was compared with the candidate sequence, to make sure the primers could only bind to the primer sites in the candidate sequence.

For the later investigation, linkers containing restriction sequences for selected restriction enzymes were added to each primer. To exclude the presence of internal recognition sites for the chosen restriction enzymes, the DNA sequences for the respective candidate genes were analyzed with Webcutter 2.0 (Labtools, 2018). From this, the program indicated which restriction enzymes was possible to use, without cutting the DNA sequence. The information obtained for each candidate sequence was compared with the restriction enzymes placed in a multicloning sites in the vector, psiCHeck2 MCS2, which were used later in this study. Different linkers were attached to the forward and the reverse primers, to ensure correct directional cloning of the candidate sequence in the vector. The same restriction enzymes, *Xho* I, were used for all forward primers, and the restriction enzyme *Spe* I, was attached to the reverse primer. These two linkers were chosen according to their procession in the vector, and additionally, the linkers had protruding ends to ensure proper digest from the restriction enzymes.

4.3 Genomic DNA purification of porcine muscle tissue

Muscle tissue from a crossbreed Landrace and Yorkshire race (LY) sow, previously dissected and frozen at the Department of Molecular Biology and Genetics at Aarhus University, was applied for the DNA purification. Approximately 100 mg, was cut from the frozen muscle, and finely divided with a scalpel. Cut muscle was transferred into two Eppendorf centrifuge tubes, and Lysis buffer, 300 µL (containing 50 mM Tris, 100 mM EDTA, 100 mM NaCl, 1% SDS and H_2O) and 30 µL Proteinase K (10mg/mL Proteinase K, 50 mM Tris, 1 mM $CaCl_2$ and H_2O) was added to each tube. The content was mixed by inverting the tubes five times, and incubated in a heating block (Eppendorf Thermomixer Comfort) at 55°C for approximately 20 hours. After incubation, 200 µL 6M NaCl was added, and the tubes were vortexed vigorously for 20 seconds, then centrifuged for 10 minutes (min.) at 1500 rounds per minute (rpm). After centrifugation, the supernatant was transferred to new Eppendorf centrifuge tubes and isopropanol (Merck, 99.8%), 0.7 times the volume of the supernatant, was added to each tube, which was inverted for mixing until DNA threads were gathering in lumps. The precipitated DNA solution was centrifuged for 10 min. at 1,500 rpm, and the supernatant with isopropanol was discarded. The DNA was washed with 70% ethanol (VWR Chemicals), centrifuged at 1,500 rpm for 10 min. and the ethanol discarded. The edges of the tubes were dried with a paper towel to remove the remaining ethanol, whereafter the tubes were left open while drying for 30 min. at room temperature (RT).

The DNA pellet was dissolved in 100 µL 1xTE buffer and periodically vortexed while incubating in a heating block at 45°C for 30-45 min. The quality of the DNA was analyzed by using a spectrophotometer (DeNovix, DS-11 FX), and by gel electrophoresis in a 1% agarose gel (Invitrogen, UltraPure™ Agarose) supplied with 15 µl 1mg/ml Ethidium bromide (Sigma Aldrich, 1mg/mL, Cas. No. 1239-45-8).

4.4 Amplification of candidate target sequences

PCR was used to amplify sequences of the selected candidate gene including target sequences for the respective miRNA's. This was conducted by creating a PCR master mix for each primer set, by mixing 5xPhusion HF Buffer (Thermo Scientific, cat. no.:F-518), 10mM dNTPs, Phusion DNA Polymerase (Thermo Scientific, cat. no.:F-530L, 2 U/mL) and H_2O with each of the separate primer sets, as described in *table 4.1, Components for a 5xReaction*.

Components	**5 X Reaction (10µL/Well)**
5xPhusion HF Buffer	10 µL
10 mM dNTPs	1 µL
Forward Primer, 10µM	2,5 µL
Reverse Primer, 10µM	2,5 µL
Phusion DNA Polymerase, 2U/mL	0,5 µL
H_2O	33,5 µL
Total amount:	50 µL

Table 4.1, ***Components for a 5xReaction.*** *The amount of each component used to create the PCR product for each of the primer sets.*

A 5xReaction was mixed for each of the primer sets, to ensure PCR-product was available for further investigation of the candidate genes. The PCR master mix from each primer set was distributed in 96-well plates (SARSTEDT. Ref no. 72.1979.202) with 9µL in each well for every primer set, where to either 1µL genomic DNA or 1µL H_2O was added. Of the genomic porcine DNA previously purified was added to four of the five wells, and 1µL H_2O was added to the last of the five wells. The well containing H_2O was used as a negative control to ensure no contamination of the samples had happened.

Due to the differences in the primer Tm°, two PCR protocols, a 2-step, and a 3-step protocol were configured for the thermos cycler (Applied Biosystems, GeneAmp, PCR System 9700). A 2-step protocol with an annealing temperature of 72°C, displayed in *Figure 4.4, 2-step PCR protocol,* used with primer sets for candidate sequences containing target sites for TRIM32 and IL12A. The 3-step protocol with an annealing temperature of 68°C, displayed in *Figure 4.5, 3-step PCR protocol,* used with primer sets for the candidate sequences containing SOCS1 and ZEB1. The two protocols ran in different thermocyclers (Applied Biosystems, Gene amp PCR system 9700), while preparations for the further analysis was conducted.

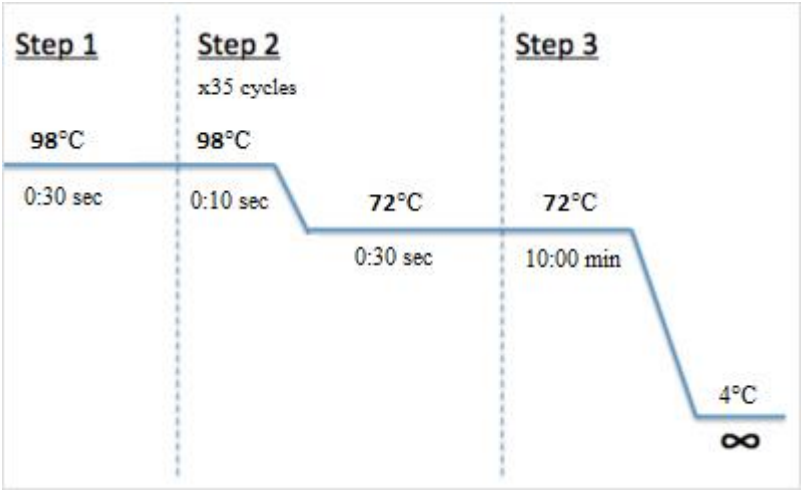


Figure 4.5, ***2-step PCR protocol.*** *The 2-step protocol had a denaturation period of 30+10 seconds in total at 98°C; where after an annealing phase with a temperature of 72°C lasted for 30 seconds. This cycle was run 35 times and continued in the elongation phase at 72°C for 10 minutes. Lastly, the PCR product was cooled down to 4°C and stored at this temperature until use.*

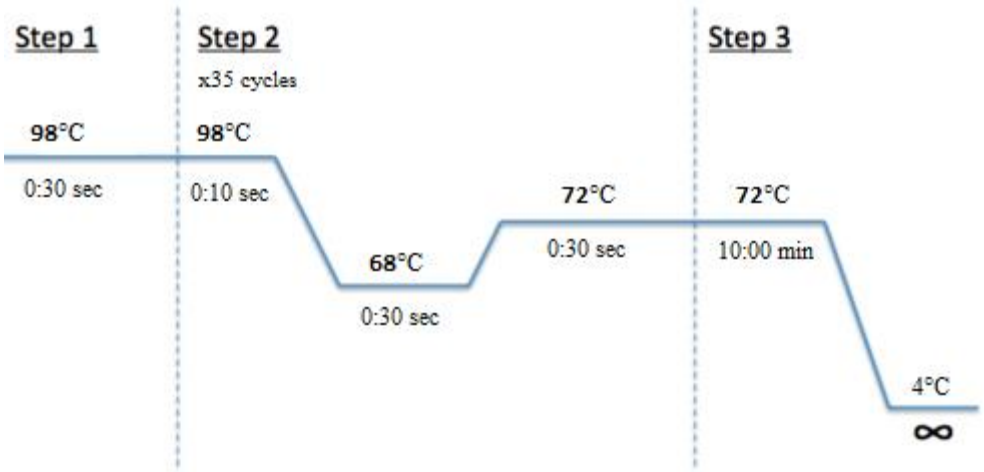


Figure 4.6, ***3-step PCR protocol.*** *The 3-step protocol had, like the 2-step protocol, a denaturation phase of 98°C. This phase lasted for 30 seconds, whereafter the annealing phase began. The annealing phase began with a temperature of 98°C lasting for 10 seconds, whereafter the temperature decreased to the annealing temperature at 68°C lasting 30 seconds. The cycle ends on an elongation temperature of 72°C lasting for 30 seconds, and the cycle was run 35 times. In the last elongation phase, the 72°C, held for 10 min, and the PCR-product was cooled down to 4°C, where it was held until use.*

4.5 Gel electrophoresis

Agarose gel electrophoresis was used to examine the results of the PCR reaction. A 1.8% agarose gel supplied with 15 µl 1mg/ml Ethidium bromide for the electrophoresis was prepared, by mixing agarose (Invitrogen, UltraPure^{TM} Agarose, ref. no. 16500-500) with 1xTBE buffer. The agarose was melted by heating, where after 15 µl 1mg/ml Ethidium bromide (Sigma Aldrich, 1mg/mL, Cas. No. 1239-45-8) was added, and the mixture was poured into a chamber with a chamber comb and left to solidify at room temperature (RT).

The gel was placed in an electrophoresis vessel (VWR Chemicals) containing 1xTBE buffer, and the chamber combs were carefully removed. One reaction of each respective PCR product and the negative control was mixed with 5µL loading dye, where after the PCR Product along with a homemade pUC 19 marker, containing pUC19 plasmid DNA and 6xloading buffer, was loaded onto the gel. The electrophoresis was run at 200V for 35 min., where after pictures were taken with a High Performance Ultraviolet Trans illuminator (UVP) at resolution 2592x1944.

4.6 Sequencing of the candidate target sequences

To investigate whether the candidate target sequences were correctly amplified, the gel product related to each primer set was purified, to collect all DNA fragments obtained. This was performed by using a MinElute Gel Extraction Kit (QIAGEN), containing Buffer QG, Buffer PE, Buffer EB, and MinElute Spin Column with belonging 2 mL collection tube.

To excise the DNA fragments from the gel, the gel was placed on a MacroVue UV 20 (Hoefer) to make the DNA bands visible, where after the bands were cut with a sharp clean scalpel. Each respective DNA band was transferred to a 2 mL Eppendorf tube (AXYGEN scientific) and weighted. A volume of Buffer QG, 3x the weight of the gel slices containing the DNA bands, was added to their respective tubes and incubated at 55°C for about 10 min. To help the gel dissolve in the QG buffer, the gel pieces were vortexed 2-3 times during the incubation time. After complete dissolvent of the gel pieces, 1xVolume of isopropanol was added to each sample and mixed by inverting five times. MinElute spin columns, provided by the gel extraction kit, was placed in 2 mL collection tubes and the samples were transferred to a MinElute spin column each. The samples were centrifuged for 1 min at 13000xG, and the flow-through discarded, where after the spin columns were placed back into their collection tubes. An amount of 500µL Buffer QG was added to each MinElute column, and again the samples were centrifuged for 1 min at 13000xG and the flow-through discarded.

The MinElute spin column was placed back into the same collection tube, and 750µL Buffer PE was added to each column. Since the DNA obtained from each sample were later being subjected to direct sequencing, whereby it is used for a salt-sensitive application, the columns were left to incubate for five min. at RT, after adding Buffer PE. After five min. the samples were centrifuged two times for 1 min, wherein between each centrifuge, the flow-through was discarded and the column again placed in the collection tube. After the second centrifugation, each MinElute spin column was placed in a clean 1.5 mL Eppendorf tube. To elute the DNA, 11µL Buffer EB was added to the center of the MinElute membrane without touching the membrane. The MinElute spin column was left to incubate for 1 min at RT and then centrifuged for 1 min.

The extracted DNA product was transferred to new tubes, 5µL in two tubes for each respective candidate fragment, and 5µL of either the forward or the reverse primer for the candidate gene. Each tube was marked and send to Macrogen Europe, Amsterdam for sequencing. After sequencing, the results from Macrogen Europe were analyzed by using Adobe Acrobat Pro

DC, version 2015 and the Trace view function in the program CodonCode Aligner 8.0.2 (CodonCode) to ensure the product correctly amplified for the further investigation in this study.

4.7 TOPO Cloning of each candidate target sequence

Before inserting the candidate target sequences into the psiCheck-2 MCS2 vector, the sequences were inserted into a TOPO vector, which functioned as a carrier plasmid. This was conducted by using a Zero Blunt TOPO PCR Cloning Kit (Invitrogen). Firstly, 50µL fresh PCR product of each candidate target sequence was amplified for insertion in the TOPO vector. The PCR product was mixed according to *table 4.1, Components for a 5xreaction,* described in section *4.4 Amplification of Candidate target sequences,* where after the PCR products from each gene was run through a 1.8% Agarose gel at 200V for 35 min. The DNA fragment in the gel was cut out with the use of a scalpel and extracted from the gel by using the MinElute Gel Extraction Kit (QIAGEN), as described in *section 4.6 sequencing of the selected candidate target sequences,* resulting in 20µL fresh PCR product.

Reagents	**Volume**
Fresh PCR Product	3 µL
Salt solution	1 µL
H_2O	1 µL
pCR-BluntII-TOPO vector	1 µL

Table 4.2, ***TOPO Cloning Reaction mix.***

The PCR product was used in the TOPO Cloning Reaction mixed with the pCR-BluntII-TOPO vector (Invitrogen), H_2O and a salt solution, containing 1.2 M NaCl and 0.06 M $MgCl_2$, as described in *table 4.2, TOPO Cloning Reaction mix*. One reaction was mixed for each of the candidate target sequences and incubates in 5 min at RT. The TOPO Cloning Reaction mixes were placed on ice until use. An amount of 2µL TOPO Cloning Reaction mix was used for transformation, by adding the TOPO cloning reaction to a vial of One Shot TOP10 chemically competent *E. coli* (Invitrogen, cat. no.: K280040), and mix by inverting the vial. Each vial was incubated on ice for 15 min., where after the cells were given a heat-shock at 42°C for 30 seconds, and quickly placed on ice again. Tempered S.O.C. medium (Invitrogen, cat. no. 15544-034) of 100µL was added to each vial, which was shaken horizontally at 200 rpm in

37°C for 1 hour. Meanwhile, five 2xYT kanamycin culture plates, containing 16g/L Tryptone, 10g/L yeast extract, 5 g/L NaCl, 20g/L agarose and 50mg/mL Kanamycin, was dried half open in an incubator at 37°C.

After incubation, each mixture of competent cells was plated, one by one to the 2xYT culture plate, and evenly distributed with a Drigalski spatula after transfer to each a plate. The plates were incubated at 37°C overnight, where four of the plates were containing colonies for the candidate target sequences, and the last plate, which was a control plate only with the TOPO Vector.

From the colonies grown overnight, six overnight cultures were composed for each of the candidate target sequences, by collecting a colony with a pipette tip and drop it into a tube containing 4-5 mL 2xYT Medium supplemented with 50 mg/mL Kanamycin. The overnight cultures were incubated at 37°C for about 16 hours in a shaker.

4.7.1 Plasmid DNA purification

After incubation of overnight cultures, the plasmid DNA was purified with the use of a Plasmid DNA Purification Kit, NucleoSpin Plasmid (Machery-Nagel, cat. no. 740588.250), whereby an amount of 1.5 mL overnight culture was transferred to individual tubes giving 24 tubes for the purification in total for all candidate target sequences. The tubes were centrifuged twice for 1 min at 13,000 x g, where after the supernatant was discarded and any excess fluid was sucked with a tissue paper from the edge of the tube.

The cells were lysed by adding 250µL of A1 buffer to each of the tubes; where after the pellet in the tubes were resuspended with a vortex-mixer (VELP SCIENTIFICA, Wizard X). Each pellet needed to be fully dissolved, where after 250µL A2 buffer, with a blue color, was added to each tube and mixed with plasmids by inverting 8 times. After the content of the tubes obtained a uniform color, 300µL A3 buffer was added to each tube, which is mixed again by inverting the tubes. The tubes are inverted 8 times, whereby the color of the content changed from blue to a clear liquid with white cloudy material in, whereby the cells were lysed. The tubes were then centrifuged at 13,000 x g for 15 min, while NucleoSpin plasmid columns are put into the supplied 2 mL collection tubes. The supernatant, about 700µL, is transferred to each a Spin column and centrifuged for 1 min at 13,000 x g.

The flow-through was discarded, and 500µL AW buffer was added to each Spin column, and centrifuged again at 13,000 x g for 1 min, to wash the cells. The flow-through from the wash was discarded, and 600µL A4 buffer was transferred to each Spin column. The Spin columns

were centrifuged at 13,000 x g for 1 min, whereafter the flow-through was discarded. Again, the Spin columns were centrifuged at 13,000 x g, now for 2 min., to dry the membrane in the column. After drying the membrane, the spin columns were placed in new collection tubes, and 75µL of AE buffer was added to each membrane, without touching the membrane with the pipette. Before centrifugation, the spin columns were incubated at RT for 1 min, where after they were centrifuged at 13,000 x g for 1 min.

After centrifugation, the collected DNA is transferred to new tubes, then marked and stored at -20°C until use.

4.7.2 Sequencing of the TOPO plasmid and targets.

To ensure the TOPO cloning was successful, the cloning was inspected by sequencing the purified plasmid DNA from each Candidate target sequence.

Initially, the TOPO plasmid DNA was analyzed with restriction enzyme digestion, by adding 10µL of each TOPO plasmid DNA to each a well in a 96-well plate. *Xho*I restriction enzyme (New England BioLabs inc., cat. No.: R0146S, 20000 U/mL) and *Spe*I restriction enzyme (New England Biolabs inc., cat. No.: R0133S, 10000 U/mL), previously inserted in the primer sites, H_2O and buffer 2 (New England BioLabs inc., cat. No.: B7002C, 10XConcentration) was mixed according to *table 4.3, Enzyme mix for TOPO plasmid DNA.*

Reagents	**Volume (1 x rxn)**	**Total Volume (25 x rxn)**
Buffer 2	1,5 µL	37,5 µL
XhoI enzyme, 20000 U/mL	0,5 µL	12,5 µL
SpeI enzyme, 10000 U/mL	0,5 µL	12,5 µL
H_2O	2,5 µL	62,5 µL

Table 4.3, ***Enzyme mix for TOPO plasmid DNA.***

Buffer 2 and H_2O was first mixed in a tube, whereafter the enzymes were added to the mix. An amount of 5µL of the enzyme mixture was added to each well, where after the wells were closed with a plastic membrane and incubated at 37°C for one hour, while making a 1% agarose gel (Invitrogen, UltraPure™ Agarose) containing 15 µl 1mg/ml Ethidium bromide (Sigma Aldrich, 1mg/mL, Cas. No. 1239-45-8).

After incubation, 5µL loading dye was added to each well, and the well was spun in a Mini Plate spinner (Labnet, MPS 1000). The gel was set up in the gel electrophoresis chamber, and six uncut and six cut plasmids from each target is transferred to each a well. Between the plasmids from each candidate target sequence, a pUC19 marker was loaded and the gel was run at 200V for 35 min.

A fragment of the expected size was seen in all TOPO clones, but only one of each sample was sequenced. The purified DNA, 5µL, was transferred to two tubes per candidate target sequence and 5µL of either a forward or the reverse primer, related to the respective candidate gene, was added. The tubes were closed, marked with labels, and centrifuged before being sent to Macrogen Europe, Amsterdam, for sequencing.

4.8 Subcloning of candidate target sequences

Before continuing to the subcloning, the content of DNA was measured by using a spectrophotometer (DeNovix, DS-11 FX), to calculate the amount needed to reach a concentration of 1500 ng of each TOPO plasmid DNA.

A total amount of 30 µL TOPO plasmid DNA, digested with restriction enzymes, was wanted of each inserted target sequence, whereby the amount of each plasmid DNA was mixed according to *table 4.4, Plasmid DNA and restriction enzyme composition*.

Content:	**Volume:**
Plasmid DNA (1500ng)	X µL
Buffer 2(x10)	3 µL
XhoI enzyme, 20000 U/mL	1 µL
SpeI enzyme, 10000 U/mL	1 µL
H_2O	x µL

Table 4.4, ***Plasmid DNA and restriction enzyme composition.***

The amount of TOPO plasmid DNA was individual for each candidate target and the psiCheck-2 MCS2 vector, which was cut like the TOPO plasmid DNA. The calculated amount of the DNA was added to each a well in a 96-well plate and supplied with the amount of H_2O needed.

A mixture of the restriction enzymes, *Xho* I and *Spe* I and buffer 2 was prepared for all the candidate target sequences. An amount of 4 µL of the enzyme mixture was added to each well, where after the plate was sealed, spun in a Mini Plate spinner (Labnet, MPS 1000) and incubated at 37°C for 1 hour in a thermocycler (Applied Biosystems, Gene amp PCR system 9700). Meanwhile, a 1% agarose gel, supplemented with 15 µl 1mg/ml Ethidium bromide, was prepared, as previously described in *section 4.7.2, Sequencing of TOPO plasmid and targets.* Before loading the gel, 5µL loading dye was added to each well with the digested TOPO plasmid DNA. The pUC19 marker was loaded to the first chamber in the gel, followed by the cut DNA loaded with an empty chamber in between. The gel was run at 200V for 35 min, where after pictures were taken with a High Performance Ultraviolet Trans illuminator (UVP) at resolution 2592x1944.

The DNA fragments of expected size were cut from the gel with scalpel, and extracted with a Gel Extraction Kit (QIAGEN), as described in *section 4.6, Sequencing of the selected Candidate target sequences*. Before transformation, the psiCheck-2 MCS2 vector and the candidate target sequences needed to be ligated, which was conducted by mixing T4 DNA Ligase buffer (New England BioLabs, cat. No.: B02025, 10x concentration), T4 DNA Ligase (New England BioLabs, cat. No.: M0318L, 3000000 U/mL), H_2O and both the digested vector and the digested candidate target sequences, as displayed in *table 4.5, Content in Ligation of Vector and Insert gene.*

Reagents	**Control**	**Insert gene**
psiCheck-2 MCS2 vector, 324 ng/µL	2 µL	2 µL
Candidate sequence	-	4 µL
T4 DNA Ligase buffer, 10x	1 µL	1 µL
T4 DNA Ligase, 3000000 U/mL	1 µL	1 µL
H_2O	6 µL	2 µL

Table 4.5, ***Content in Ligation of Vector and Insert gene***

The vector and candidate target sequence was ligated in a 96-well plate, where a control ligation without candidate target sequence and ligations containing each their individual target sequence was mixed in each a well. First, the vector and target sequence was added to each well, and since no insert gene was wanted in the control, 4 µL H_2O was added to the vector

instead of a target sequence. A ligase mixture containing the T4 DNA Ligase buffer, T4 DNA Ligase and the rest of the water was mixed, and 4 µL of the ligase mix was added to each well. The vector, *psiCheck2,* and target sequences were ligated in a thermos cycler, which switched between 30 min at 20°C and 45 min at 16°C for 10 cycles and then stored at 4°C.

For transformation, the ligated vector and target sequence were transferred to each a tube containing One Shot TOP10 chemically competent cells (Invitrogen, cat.no. K280040) and incubated on ice for 15 min. After the 15 min, the cells were exposed to a heat-shock at 42°C for 30 sec., where after they were placed on ice again and 125µL tempered S.O.C. medium was added to each tube. The tubes were shaken horizontally at 37°C for one hour. Meanwhile, 2xYT culture plates supplemented with 100µg/mL ampicillin were dried in an incubator, with half-open lid, at 37°C.

The content of the tubes was, after incubation, transferred to each a dried culture plate supplemented with ampicillin, and evenly distributed with a Drigalski spatula. The Drigalski spatula had, prior to use, been cleaned with ethanol and burned. This procedure was performed for all four target sequences and the control, whereafter it was placed in an incubator at 37°C overnight. From the colonies grown, overnight cultures were made by pouring 4-5 mL 2xYT media containing 100µg/mL ampicillin into a 15 mL tube, and inoculate the colonies from each candidate target sequence with a pipette tip. These tubes were placed in a shaker at 37°C for approximately 16 hours, where after a plasmid DNA purification was conducted by using the plasmid DNA purification Kit, NucleoSpin Plasmid, as described in *section 4.7.1*, *Plasmid DNA purification*.

The subcloning the candidate target sequences into the vector, were inspected by digesting the clones with the restriction enzymes *Xho* I and *Spe* I, as used when digesting the TOPO cloning, where after the clones were run through a 1,8% agarose gel (Invitrogen, UltraPure™ Agarose), supplemented with 15% ethidium bromide, for 35 min at 200V. The tubes containing the purified DNA from the subcloning of each candidate target sequence was stored at -20°C and two subclones of each candidate target sequence was send to EZ sequencing at Macrogen Europe, Amsterdam.

4.9 Retransforming plasmid constructs containing candidate target sequences

For investigating the orientation of the candidate target sequence in the psiCheck-2 MCS2 vector, specific primers designed for the vector was used to conduct a sequencing by transferring 5µL of each subclone to two individual tubes and adding 5µL of the forward to one and 5µL of the reverse primer to the other tube. The tubes were sent to Macrogen Europe, Amsterdam, for sequencing, and the results from each sequence were analyzed with Adobe Acrobat Pro DC, version 2015 and the Trace view function in the program CodonCode Aligner 8.0.2 (CodonCode).

When confirmed by sequencing, each plasmid DNA with the inserted candidate target sequence and the vector psiCheck2 was retransformed to amplify the available content of DNA. This was conducted by adding 1µL of each plasmid DNA to each a tube containing TOP10 competent E. coli cells, where after they were incubated on ice for 15 min. Meanwhile, 2xLB culture plates containing 100µg/mL ampicillin, one for each plasmid DNA, was dried in an incubator at 37°C. The TOP10 competent *E.coli* cells, with each plasmid DNA, were transferred to the 2xLB culture plate and evenly distributed on the plate. The plates were incubated at 37°C for approximately 16 hours, whereafter overnight culture with 10 mL 2xYT Media supplemented with 100mg/mL Ampicillin, was made for each of the genes. One colony, containing each a candidate target sequence, was transferred to each a tube and shaken at 37°C for 16 hours. The amount of 10µL overnight culture from each respective plasmid DNA was transferred to new 2xLB culture plates, again supplemented with ampicillin, and propagation with an inoculation needle was conducted after the plasmid DNA was absorbed into the plate.

New overnight cultures were made, by adding 500mL 2xYT media and 100mg/µL Ampicillin per mL to an autoclaved flask and then transferring 1 mL of the small overnight cultures to each a flask. One overnight culture of each subcloned target sequence and two overnight cultures with psiCheck2 were incubated on a shaker at 37°C for 16 hours.

The overnight cultures were purified with a QIAprep Spin Miniprep Kit (Qiagen, cat. No. 27106) one at a time, by distributing 40 mL of overnight culture in 50 mL tubes. The overnight cultures were centrifuged at 4000 rpm for 10 min, and the supernatant was discarded. The pellet dissolved in 1 mL A1 buffer by vortex until being homogenous. The dissolved pellet was transferred to 2 mL Eppendorf tubes, 500µL in each tube, whereafter it was mixed with 500µL A2 buffer by inverting 6 times. An amount of 700 µL A3 buffer was added to each tube, which again was inverted until the blue color from the A2 buffer was gone. The tubes were centrifuged

at 17000xG for 10 min, whereafter the supernatant was transferred to new 2 mL Eppendorf tubes. The supernatant was gradually transferred to SpinColumns and centrifuged at 13000xG for 1 min over two turns. After each centrifuge, the flow through was discarded. The columns were washed 2 times with 400 µL AW buffer and centrifuged at 13000xG for 1 min, wherein between the flow through was discarded, and again washed with 600 µL of A4 buffer and centrifuged at 13000xG for 1 min, and flow through discarded. The SpinColumns were centrifuged for 2 min. at 13,000 x g to dry the membrane, and then moved to new collection tubes, where 80µL AE buffer was added over two turns, and at each turn incubated at RT for 1 min and centrifuged at 13,000 x g for 1 min. The eluted DNA was transferred to two 1.5 mL Eppendorf tubes and the concentration of the DNA is measured with a spectrophotometer (DeNovix, DS-11 FX). Furthermore, the content was inspected by digesting each plasmid DNA and the vector with the restriction enzymes, *Xho* I and *Spe* I, as previously described in section 4.8, *Subcloning Candidate target sequences*, and compare these results with previous results.

4.9.1 Glycerol stock

In case more plasmid DNA from the subclones were needed, a glycerol stock of each subclone with the candidate target sequences and the vector was frozen. The glycerol stock was made by transferring 500 µL of a new overnight culture, made as previously described in this section, and 500 µL 50% Glycerol, where after they were mixed by inverting 6 times and frozen at -80°C.

4.10 Cell culture

For the later transfection, one vial of HEK293 cells, kindly provided by Thomas Birkballe, was thawed and seeded in a T75 culture flask (SARSTEDT AG & Co, REF: 83.3911.002) with preheated sterile filtered growth media containing Dulbecco's Modified Eagles Medium (1X) (DMEM) + Glutamax (Gibco, REF. 32430-027) supplemented with 10% FCS and 1% pen/strep (penicillin/Streptomycin). The cell culture was incubated at standard conditions at 37°C and 5% CO_2 for approximately 24 hours, whereafter the media was changed. The two following days, the cell confluence was checked by using a LEICA DMIL microscope with an x4 lens, and when confluence reached 80-90 %, the cells were subcultured in two new T75 flasks.

The subculture was performed by trypsinizing the cells, where first the media was poured from the cells. The cells were washed 2 times with 9 mL sterile 1x Dulbecco's Phosphate Buffered Saline (PBS) (GIBCO, REF. 14040-091), and 2 mL 0,25% Trypsin/EDTA (Gibco, 1x) was added. The cell culture was incubated at standard conditions for 3-5 min for detachment, which

was checked with a microscope. After detachment, 6 mL growth media was added to the cell culture, to neutralize the trypsin/EDTA. The bottom of the T75 culture flask was washed 3 times with the supplied growth media, where after the media, now cell suspension, was transferred to a 15 mL centrifuge tube and centrifuged at 500xG for 5 min. The supernatant was poured and the cell pellet was resuspended in 3 mL growth media. Preheated growth media, 14 mL, was transferred to each three new T75 culture flasks, and 1 mL cell suspension was added to each flask. The presence of cells was checked in the microscope, and the cell culture was incubated at standard conditions.

This procedure was repeated several times until enough cells were obtained for cell freezing.

4.10.1 Cryopreserving cells

During the project, cells were prepared for cryopreservation and were stored in a -140°C freezer. This was performed by first trypsinizing the cells, as mentioned in the previous section, where after the cell suspension was transferred to a 15 mL centrifuge tube and centrifuged for 5 min at 500xG. The supernatant was poured and the cell pellet was resuspended in 5 mL growth media. From the cell suspension, 10µL was transferred to a micro Eppendorf tube, and 10µL 0.4% Trypan blue (Sigma-Aldrich, CAS 72-57-1) was added and resuspended 5 times. A hemocytometer, kept in 96% ethanol, was dried and the cover glass was attached by breathing on the hemocytometer and sliding the glass on. The hemocytometer was loaded with 10µL of the cell suspension containing the Trypan blue, in both chambers and then counted. Calculations were performed by using the following equation:

$$\frac{cell\ count}{4} * 2 * 10^4 = cells/mL$$

The $2x10^6$ cells in 695µL media were distributed in each cryotube, and 305µL freezing agent, consisting of 75µL DMSO and 230µL FCS was added to each tube. The cell production resulted in 36 cryotubes with a cell concentration of $1{,}5x10^6$ cells, and 18 cryotubes with a cell concentration of $2x10^6$ cells.

4.11 Transfection and Co-transfection of HEK293 cells.

The process of transfecting the HEK293 cells, previously grown, was conducted by first seeding the cells. After 24 hours, the cells were transfected with and without plasmid DNA respectively as a positive and negative control (NC). Furthermore, cells were co-transfected with plasmid DNA and either miRCURY LNA miRNA Mimic (QIAGEN, 66,67 pmol/µL) or miRCURY LNA miRNA Power Inhibitor (QIAGEN, 50 pmol/µL), related to the miRNA, *miR-155-5p, miR-150-5p and miR-21-5p*, expected to regulate the expression of the respective candidate target site in the plasmid DNA. Prior to the trial, several optimization trials were conducted, and to determine the proper cell density to seed in 24-well plate to obtain a confluence of 70%, seedings with a cell density of $0.5x10^5$, $1x10^5$, $1.5x10^5$ and $2x10^5$ cells/well were conducted. To determine the concentration of the transfection reagent Lipofectamine 2000 Reagent (Invitrogen, 1mg/mL REF no. 11668-027), to avoid toxicity, transfection trials with 1µL/well, 1.5µL/well 2µL/well and 2.5µL/well of Lipofectamine 2000 was conducted. Furthermore, trials were conducted to find the most effective proportion of plasmid DNA and lipofectamine 2000 when transfecting. For this, a fixed concentration of 1µL/well Lipofectamine 2000, which was the one proven most effective and not toxic, and plasmid DNA concentrations of 0.5µg/well, 1.5µg/well and 2.5µg/well were used. Similar trials, with the miRCURY LNA miRNA Mimic and miRCURY LNA miRNA power Inhibitor, were performed to determine which concentration to use to avoid toxicity, but still, conduct the most effective co-transfection. For miRCURY LNA miRNA Mimic, concentrations of 1pmol/well, 2.5pmol/well and 5pmol/well was tried, and for miRCURY LNA miRNA power Inhibitor concentrations of 1pmol, 2.5pmol, 5pmol, 10pmol, and 25pmol per well were conducted.

The same procedure was used for both pilot trial and experiment and performed as described in the following sections.

4.11.1 Seeding of HEK293 cells

A similar procedure, as used when sub cultivating cells, was used when seeding the HEK293 cells for the transfection and co-transfection.

Initially, the growth medium was poured from the cells, and they were washed 2 times with 9 mL 1xPBS, which was poured after each wash. An amount of 2 mL 0.25% Trypsin/EDTA was added, and the cells were incubated for 3-5 min until detachment, which was checked with a microscope. When detached, 6 mL media was added to neutralize the Trypsin/EDTA. The cell suspension was used to wash the bottom of the T75 culture flask, where after the cell

suspension was transferred to a 15 mL centrifuge tube. The HEK293 cells were centrifuged at 500 x g for 5 min. and the supernatant was discarded. The remaining cells were resuspended in 3 mL media and counted with the use of 0.4% Trypan blue and a hemocytometer as described in the previous section 4.10, *Cell culture.*

Due to the previous optimization trials, a cell density of $1x10^5$ was estimated as the optimal cell density due to the confluence level obtained, whereby $1x10^5$ was seeded in 21 wells in three 24-well plates. Before the transfer to each plate, the cells were resuspended in the proper amount of media for all wells. While seeding, the cells were resuspended for every third transfer, to ensure the same cell density in all wells. After seeding, the plates were incubated for 24 hours at standard conditions. The transfection of HEK293 was made in triplicates, whereby three similar 24-well plates were prepared before the transfection.

4.11.2 Transfection of HEK293 cells

In all of the three plates prepared, a triplicate test of each control and the investigational parameter was conducted, giving three NC without plasmid DNA, and 3 wells with respectively PC, co-transfection with miRCURY LNA miRNA Mimic and miRCURY LNA miRNA power Inhibitor, for the investigated candidate sequence and the vector psiCheck2 as a control. Three miRCURY LNA miRNA Mimic and power Inhibitors were used, which each recognized the candidate sites in each respective candidate target sequence, and are supposed to regulate the expression thereof. The plate setup for each experiment can be seen in *table 4.6, Transfection trial*.

psiCheck2_SOCS1			**psiCheck2**		
NC	NC	NC			
PC	PC	PC	PC	PC	PC
miRNA Mimic – miR-155-5p	miRNA Mimic – miR-155-5p	miRNA Mimic – miR-155-5p	miRNA Mimic – miR-155-5p	miRNA Mimic – miR-155-5p	miRNA Mimic – miR-155-5p
miRNA Inhibitor – miR-155-5p	miRNA Inhibitor – miR-155-5p	miRNA Inhibitor – miR-155-5p	miRNA Inhibitor – miR-155-5p	miRNA Inhibitor – miR-155-5p	miRNA Inhibitor – miR-155-5p

*Table 4.6, **Transfection Trial.** Here the plate setting of each transfection trial illustrated, with triplicate of each test conducted. Both the plasmid DNA construction and the vector, psiCheck2, was transfected at the same time, and again repeated 2 times.*

In both the PC and the wells containing miRNA Mimic and Inhibitor, an amount of 0,5µg/well plasmid DNA and 1µL/well Lipofectamine 2000 was used, and a concentration of 1pmol/well of each respective miRNA Mimic and 2,5pmol/well of each respective miRNA Inhibitor was used.

The growth media in the wells were first replaced with 450µL growth media without pen/strep by aspiration, where after the cells were washed one time with 1xPBS and the growth media without pen/strep was added. While preparing the reagents for the transfection, the plates were incubated at standard condition until use.

Four different reagents were made for the transfection and co-transfection of the HEK293 cells. The negative controls did not contain plasmid DNA, which was replaced with DMEM + Glutamax, and the positive control did not contain either miRNA Mimic or Inhibitor, with which this was replaced with DMEM + Glutamax. Each reagent, both DNA, Lipofectamine 2000 and miRNA Mimic or Inhibitor, were firstly prepared in each a vial in the proper amount used for the trial, where after they were incubated at RT for 5 min. When incubated the DNA-Lipofectamine and DNA-Lipofectamine-miRNA Mimic/Inhibitor- complexes were prepared by mixing the reagents to each well and incubate at RT for 25 min. After incubation, each well was supplemented with 50µL of the respective complex, as described in table 4.10, *Plate setup for transfection in HEK293 cells.* After transferring the complexes to each well, the plates were incubated at standard conditions for 48 hours, where the results were analyzed by using Luciferase Assay.

4.12 Analysis of transfected HEK293 cells

The results from the transfection were analyzed with Dual-Luciferase Reporter Assay System (Promega, REF. E1910), which included Luciferase Assay Reagent (LARII), Passive Lysis Buffer (PLB) and a Stop & Glo reagent. The procedure was performed by first removing the growth media without pen/strep from the cells and then wash with 1xPBS. The cells were lysed with 100µL 1xPLB added to each well, where after they were placed on a shaker for 15 min. at RT. The lysed cells were transferred to each a tube, 20µL of the lysed cell samples were transferred to each well in an optic 96-well plate. The lysed cell samples were analyzed with a FLUOstar OPTIMA luminometer (BMG Labtech) and the program OPTIMA Control version 2.10 R2 (BMG Labtech). First, the Firefly Luciferase was measured by adding 100µL LARII, set the settings and measure the luminescence. Quickly after measuring firefly Luciferase,

100µL 1xStop & Glo reagent was added to each well, and the second measure of the lysed cell samples was conducted. The second measures were conducted 2-3 min. after the measurement of Firefly Luciferase, and is measuring the expression of Renilla Luciferase in the lysed cell samples.

4.12.1 Evaluation of outcome results from the Luciferase Assay

From the Luciferase Assay, various outcome measures were obtained. Results of the raw data were conditioned by the cells seeded in the 24-well plates, both survival, and growth pattern, during transfection. During trials, the growth pattern and the HEK293s ability to maintain attachment to the wells for unknown reasons changed. Caused by the differences in the growth patterns, normalizing the data were performed in relation to the growth patterns. The data were normalized according to background measures retrieved from negative controls in both optimization trials and experimental trials. Due to the growth pattern of the cells, background measures were calculated, one background measure for each growth situation of the cells, for firefly luciferase and Renilla luciferase separately. After calculation of each background measure, the measures were subtracted from the raw data, and the ratio between firefly luciferase and Renilla luciferase were calculated. Both subtractions of the background measures and calculation of the ratio were calculated in Microsoft Office Excel 2016 for all experimental samples related to the candidate target sequences under investigation, and the concurrent control trials. The ratio was calculated with the equation displayed in the following:

$$Ratio = Renilla/Firefly$$

The mean ratio for the positive control and the target sequence influenced by either miRCURY LNA miRNA Mimic or miRCURY LNA miRNA Power Inhibitor were calculated for both the candidate target sequences, and psiCHECK2. The mean ratio was used in the statistical analysis of the Luciferase Assay results.

4.12.1 Statistical evaluation

Statistical analysis of the raw data was performed in GraphPad Prism 8 (GraphPad Software, version 8.1.1), according to the ratios between Renilla luciferase and firefly luciferase. In the analysis, the normal distribution and statistical significance were investigated. Furthermore, GraphPad Prism 8 was used to create a boxplot of the mean ratio of each candidate target

sequence and the correlated control trial with psiCHECK2. Additionally, the standard deviation was calculated and inserted in each boxplot.

To find the normal distribution of the raw data, the Shapiro-Wilks normality test was performed for each candidate target sequence and their related control trials with psiCHECK2. Since the data was normal distributed, a parametric analysis was performed to test for statistical significance. In the experimental trial, more than two groups were analyzed, whereby a one-way ANOVA was performed for each candidate target sequence, and the related control trial with psiCHECK2. To investigate which of the groups were significantly different from the positive control of each candidate target sequence and the control trials with psiCHECK2, Dunnet's post hoc analysis was performed.

5. Results

5.1 Identification of the Candidate target sequence for each respective miRNA

As mentioned earlier in this study, three specific miRNA's were investigated, based on research in a previous study, performed at the Department of Molecular Biology and Genetics at Aarhus University by Astrid Nielsen in 2017.

The miRNA's, *miR-155-5p, miR-150-5p, and miR-21-5p,* were, as mentioned in the problem delimitation, selected from the results of the study by Astrid Nielsen, to which this study represents an extension. Potential target sequences related to *miR-155-5p, miR-150-5p and miR-21-5p* were identified, by using several target mining functions available in the databases miRWalk, TargetScan, GeneCards, and the NCBI database, as described in *Section 4.1, Data collection on candidate target sequences for miRNA.* The databases, used in this study, provides predicted and valid information on miRNA-target interactions, and by comparing and sorting the information by certain criteria, the most reliable candidate target sequences for each of the miRNA's were revealed. From the information found in the databases and the comparison hereof, the candidate target sequences found relevant for miR-21-5p was IL12A, for miR-155-5p the targets were TRIM32 and SOCS1, and ZEB1 for miR-150-5p. IL12A, SOCS1, ZEB1, and TRIM32 were found relevant as candidate target sequences since they all have a regulatory effect on the inflammatory processes or the immune system. They are all present in both human and pig and had a 100% nucleotide match within the target recognition sequence when comparing the two species.

IL12A is a protein-coding gene, which encodes a subunit of a cytokine. This cytokine can act as a growth factor for activated T- and NK-cells (natural killer cells), and enhance the lytic activity of NK and lymphokine-activated killer cells. Furthermore, this cytokine affects some biological activities in the body. The presence of the cytokine, encoded by the *IL12A* gene, is necessary for a T-cell independent induction of INF-ɣ (Interferon Gamma) and the differentiation of both Th1 and Th2 cells.

SOCS1 belongs to a group of proteins, which is part of a classical negative feedback system. *SOCS1* is encoding a member of a STAT-induced STAT-inhibitor, also known as suppressor of cytokine signaling (SOCS). The expression of SOCS1 can be induced by several cytokines, including IL-2 (Interleukin 2), IL-3 (Interleukin 3), EPO and INF-ɣ. The protein encoded by

this gene takes part in a negative feedback loop, which attenuates cytokine signaling. SOCS1 regulates cytokines, which use the STAT3- pathway, and is e.g. an important regulator for signaling by IL-6 (Interleukin 6) and INF-ɣ mediated sensory neuronal survival.

ZEB1 is a protein-coding protein, which acts as a transcriptional repressor. The *ZEB1* gene product induces, inhibits and enhances different promoters and transcriptions; including the gene expression of IL-2 and either enhances or represses the promoter activity related to the *ATP1A1* gene, depending on the amount of cDNA and the cell type. In addition, ZEB1 promotes tumorigenicity by repressing rigidness inhibiting miRNA's. ZEB1 have related pathways to cytokine signaling in the Immune system and ERK signaling, and mutations of this gene are associated with diseases as Fuchs endothelial and corneal dystrophy.

The ***TRIM32*** gene is encoding a protein, which is a member of the tripartite motif (TRIM) family. The protein has been located in both the cytoplasmic body and the nucleus, where it interacts with the activation of HIV-1 Tat protein, which activated the HIV-1 gene. The TRIM32 gene is associated with diseases such as muscular dystrophy, Bardet-Biedl syndrome 11 and Limb-Girdle. Furthermore, this gene is related to MHC class I mediated antigen processing and presentation.

The cDNA sequence of each candidate target sequences, found in the NCBI nucleotide database, was used in this study as a checklist. Furthermore, the cDNA sequences were used when designing primers for further analysis of each target sequence. The sequences are displayed below, where the target site for the corresponding miRNA is marked in yellow and gene-specific primer sequences with light blue. In addition, both the Linker (underlined sequence) and the four extra nucleotides have been inserted in the 5'-end of the sense (forward) primer and in the 3' end of the antisense (reverse) primer.

IL12A cDNA sequence

5' –

actgCTCGAGTGACTGGTCCTATCTAAGCTGGTATAGGAATTCCCATGCTTGTTTAC
GTTAGTCATCACCCCAAATTTGAAAGATGTGAGGCTCCCATCCACACAATTGATA
CTCTGATCAAGTGTATTTCACAGTGGATAAGTTATACTACGGATAAGTTATTTTTT
TAAGTTTTCATGAGCAAACTGCTAAGGAGGGAAAATGTCCTCACAGAACATGTT
GGGTTTTTTTTTCCCTTTAATAGAAGAGCAAGAATTTATAAGCTATTTCAGTACCA
AAGTGTTTGTAGAAACAAACACTTGAGCCTAATTTATTTTTTAAGTATTTATTTAT
ATAATTTTGTACTCATGAAAGCATATGAACTAACTTATATTTATTTATGTTATATT

TATTAAAATATTTATTCCCAAGTGGATTTGAAATATACCTTGTTATTCTAAAAATA
AAATGACTGAATTaaagtaatgctcatatttttctaatgtgactgaaatactctgagtgttaaaatacatatcgtgggaaaactg
aaaactgatgaaatttgaacaaaaccaccttggaaaaaaagccacaatattcttactcattttctaggtcctttaatctcatattttaaagaga
gccacaagactagatctgcACTAGTcagt – 3’
(DNA fragment of 634 bp – excluding the linker sequences and the extra four nucleotides)

SOCS1 cDNA sequence

5’–
actgCTCGAGCAAGCATCCGCGTGCACTTCCAGGCCGGCCGCTTCCACCTGGACGG
CAGCCGCGAGAGCTTCGACTGCCTCTTCGAGCTGCTGGAGCACTACGTGGCGGC
GCCGCGCCGCATGCTGGGGGCCCCGCTGCGCCAGCGCCGCGTGCGGCCGCTGCA
GGAGCTGTGCCGCCAGCGCATTGTGGCTACCGTGGGCCGCGAGAACCTGGCGCG
CATCCCCCTCAACCCCGTCCTCCGCGATTACTTGAGCTCCTTCCCCTTCCAGATAT
GATCGGCCGCGCCCGCCCTGCACGGAGCATTAACTGGGGCGCCTTATTATTTTCT
ATTATTAATTATTATTTCCTTGGAACCATGTGGGTTGGAGGGAGCGAGTGTGGGG
GCGAGGCGCCTCCCGCCCTCGGCTGGAGACCCCTCCCGTAGACCCCTCCCCACCT
CTTGTGGGGGTGCCCCCTCTTGGTGCTCCCTCTGGGTCCCCCTGGTTGTCCCCCTG
GTAGTAGCTTAACTTAACACTAGTcagt – 3’
(DNA fragment of 503 bp – excluding the linker sequences and the extra four nucleotides)

ZEB1 cDNA sequence

5’–
actgCTCGAGATCCGGGTGTGCCTGAACCTCAGACCTAGTAATTTTTCATGCATTTTTC
AAGTTAGGAACAAGTTTGTAACATGAAGCAGATTAGAAAACTTAATGACTCCGA
AAGCAAAGATTTAACAGGTGAAAGGAAACCAATATAATTAGATAAGCATCTGGT
GTCGTTTCACTTTATCAGTATTAATTATCACTCTTATGTTGGTTTATTCTTAAGCTG
TACAATTGGGAGAAATTTTATAATTTTTTATTGGTAAACATAGGCTAAATCCGCT
TCAGTATTTTATTATGTTTTTTAAAATGTGAGAACTTCTGCACTACAAAATTCCCT
TCACAGAGAACTATAATGCAGTTCCAAACCGTGCTAACTACCTTTTATAAATTCA
ACCTAGAAGGTAGTCATTTCTCATATTTAGATGTCATAGTAGAGTGTGTTCACTA
GTcagt – 3’
(DNA fragment of 431 bp – excluding the linker sequences and the extra four nucleotides)

TRIM32 cDNA sequence

5'–

actgCTCGAGCACTGACCTGTGCTAAGAGCTGTCTGTGGAATTGTACAACACAGGCTG
GGGCCAACACAGTGCAGAGGGGGACACATCTCATTTTCATCATTAAAACTGTGG
GGTCCTTATTCTGTATTTTCCTCAAGTGAGTGCATCATATTTGGTTTGGGGATTTA
TTTCCATTCCTCTATCAAGCATTAAATGATTGATAACTGTTTCTTCATCACTGGTG
TTCGTGTCTTTTCATTAAATATGGGATTTGGGAGGTGGGGGTGGTATCTGCTGCTT
TGTTTCCACACCTGACACTCCATCCTTTGAGGATGGTAAGAATCTACCTAGGTCA
CCATTCAGCCAAAAGTACCATTATGTCAAGTTAATCCAGATCCTACTCCCACATA
ACCTGTGTATGGAATCAGCCTCCAAGGTCTGCAGTCACCAGCACACTAGTcagt – 3'
(DNA fragment of 424 bp – excluding the linker sequences and the extra four nucleotides)

5.2 Purification of Genomic porcine DNA

In order to PCR amplify and investigate the four candidate target sequences, genomic DNA was isolated from muscle tissue sampled from a crossbreed Yorkshire and Landrace sow. The purification was made from a muscle piece divided into two tubes, which as presented in *Figure 5.1, Results from Genomic DNA Purification*, was successful. The two bands containing the DNA of high molecular weight is above the pUC19 marker, and results from measurement in a spectrophotometer showed a concentration of 180.8 ng/µL purified DNA, with a purity of 2.05 (OD_{260}/OD_{280}). The recommended purity of DNA samples has a value of 1.8 (OD_{260}/OD_{280}), which lies under the purity value of the DNA purified in this study. However, when purifying from tissue this value is hard to obtain, and from the figure, a smear is observed under each DNA bands, which displays degraded genomic DNA and possibly RNA not fully degraded in the purification process, which the higher purity values can be explained by.

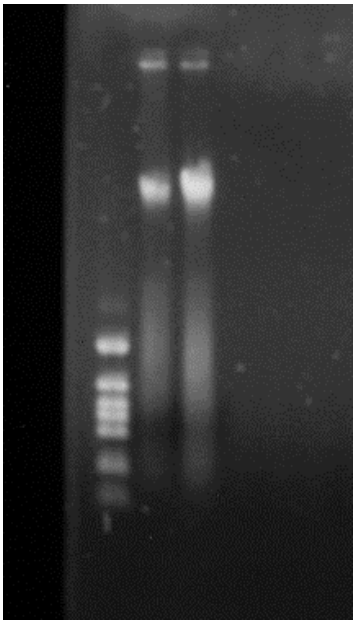


Figure 5.1 ***Results from Genomic DNA Purification.*** *Results from gel electrophoresis of the DNA purified is seen in this figure, where the content of both tubes with purified DNA has been tested. The content of both tubes is similar, with a small amount of remaining protein, which is displayed by the smear at the bottom of the figure.*

5.3 Candidate target sequence amplification

The target-specific primer sets for each candidate sequence, previously designed from the cDNA sequences available from the NCBI nucleotide database, was used to PCR amplify each 3'UTR region, containing a target sequence recognized by the respective miRNA's. Each PCR product was further inspected by gel electrophoresis, with results displayed in *Figure 5.3, Amplification of Candidate target sequences*. By comparing the bands position in the gel presented in figure 5.3, with the pUC19 marker in the figure and presented in *figure 5.2, pUC19 marker*, an estimate of fragment size in base pairs of each band is given. The comparison between the size of fragment amplified by each primer set and the number of bp available from the candidate target sequences, presented in *Section 5.1, Identification of candidate genes for each respective miRNA*, indicates the amplification of each candidate target sequences was successful.

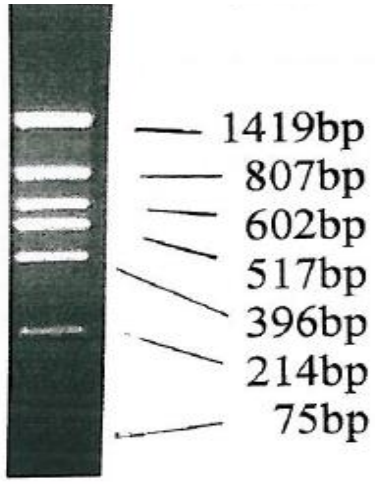


Figure 5.2, ***pUC19 Marker.*** *The pUC19 marker is giving the displayed ladder when performing gel electrophoresis. By this, the number of base pairs in investigated target sequences can be estimated*

This is assessed based on the pUC19 marker values, which is displayed in *Figure 5.2, pUC19 Marker*. By comparing with the results displayed in *Figure 5.3, Amplification of Candidate target sequences*, it can be assessed that the target sequence for the candidate gene *IL12A* (lane 1) is just above 602 bp band in the pUC19 marker. The results for *TRIM32* (lane 2) and *ZEB1* (lane 4) can be observed between the marker with 517 bp and the marker with 396 bp, and lastly, the gene *SOCS1* can be seen just under the marker of 517 bp, which all fits with the number of base pairs previously mentioned for each of the candidate genes. The PCR amplification was successful, and for IL2A, SOCS1 and ZEB1 also very specific.

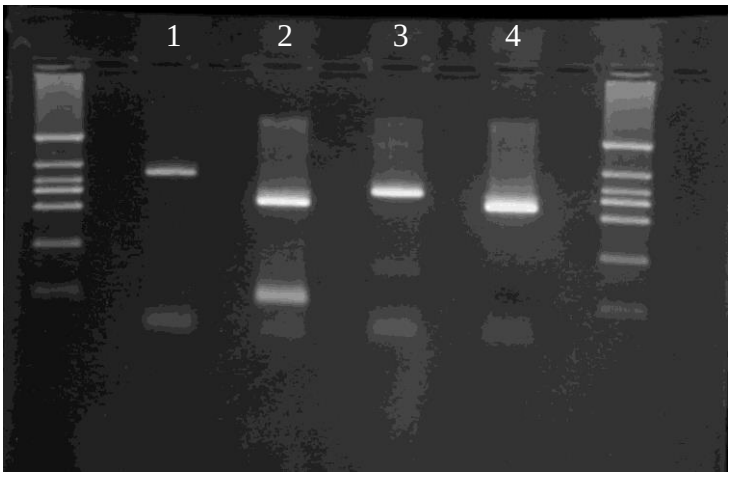


Figure 5.3, ***Amplification of Candidate target sequences.*** *Results from each primer set, designed from the target sequences available from the database, displayed with a pUC19 marker in each end. Each primer is numbered from 1-4, in following order 1) IL12A, 2) TRIM32, 3) SOCS1 and 4) ZEB1. By comparing each DNA band with the pUC19 marker the number of base pairs in each DNA strand amplified can be estimated*

DNA from each band from the results presented in *Figure 5.3, Amplification of Candidate target sequences*, was extracted from the agarose gel and sequenced by Macrogen to investigate if the target sequences had been amplified. The results sequence received from Macrogen was compared with the target sequence, previously presented in *section 5.1, Identification of Candidate genes for each respective miRNA*, available from the databases. An example of a sequence received from Macrogen is presented in *Figure 5.4, Sequencing result of IL12A*, where the target sequence from IL12A can be observed in the sequence from base pair 225-245. The full sequences from each of the candidate target sequences can be seen in *Appendix E*.

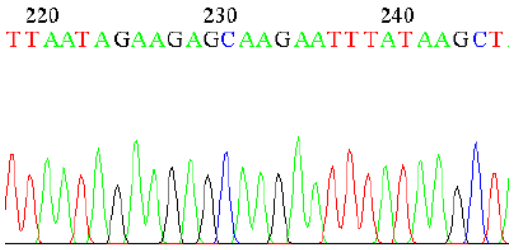


Figure 5.4, ***Sequencing result of IL12A.*** *A result of sequencing the DNA band from the candidate gene IL12A.*

The results from sequencing revealed the sequences being identical with the sequences available from the database, and in the further investigation of these genes, they were cloned firstly into TOPO Vector by using a TOPO cloning kit and then subcloned into the vector psiCheck-2 MCS2, which is presented in *Figure 5.5, Strategy of cloning*.

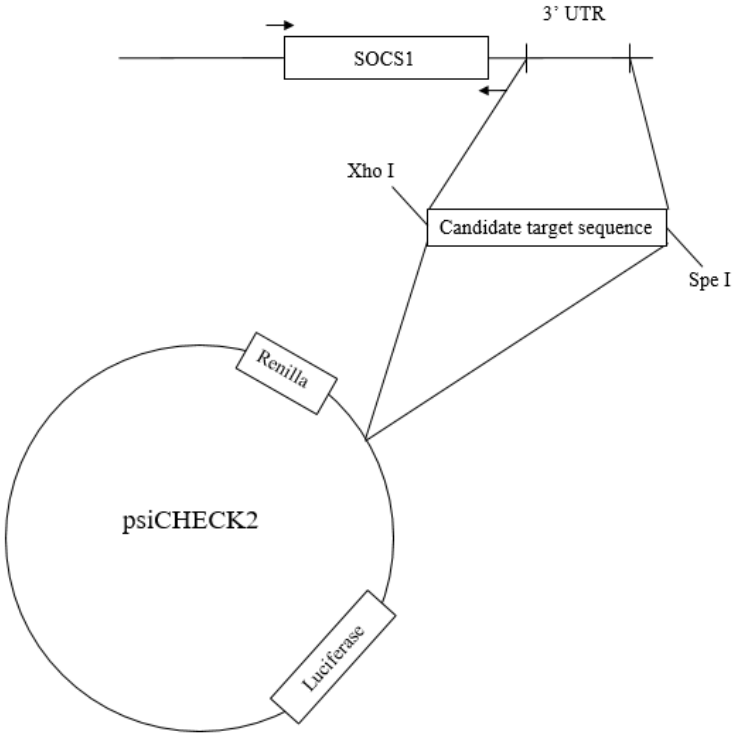


Figure 5.5, ***Strategy of cloning.*** *The 3'UTR sequence containing the candidate target sequence is inserted in a multicloning site in the psiCHECK2 vector, by digestion of the restriction enzymes Xho I and Spe I.*

5.4 Insertion of candidate target sequences in TOPO Vector

Before inserting each candidate target sequence into the psiCheck2 vector, the target sequences were inserted into the pCR-BluntII-TOPO vector, to make it easier to later insert them in the psiCheck2 vector. After TOPO cloning each candidate target sequence, the recombinant vector the target sequences were cloned into, was transformed into competent *E. coli*, an overnight culture was made and the DNA was purified with a plasmid DNA purification kit, resulting in six mini-preps from each candidate target sequence. The DNA was run through a 1% agarose gel to see if the clones were successful.

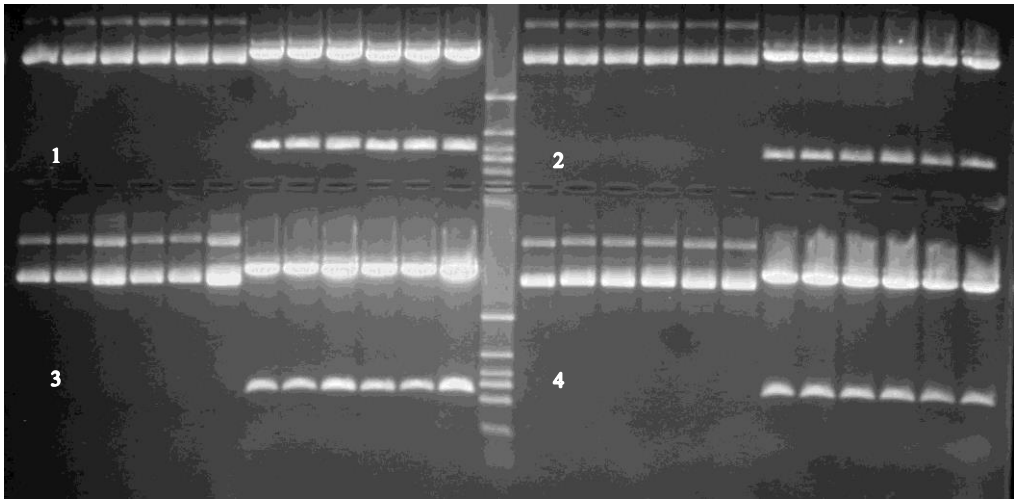


Figure 5.6, ***TOPO cloned Candidate target sequence.*** *Presence of the Candidate target 3'UTR sequences was tested by restriction enzyme digest and gel electrophoresis after cloning into the TOPO vector. In the first six wells, each tested candidate target sequence with digested mini-preps, and undigested mini-preps of the tested in the following six wells can be seen.*

The results are presented in *Figure 5.6, TOPO cloned Candidate target sequence,* where both digested and undigested mini-prep DNA from each candidate target sequence can be seen. The first six wells contain undigested mini-prep DNA, where after the corresponding digested mini-prep DNA can be seen, for each respective candidate target sequence. The candidate target sequence IL12A is presented as number 1 in the figure, where after SOCS1 comes as number 2, TRIM32 as 3 and lastly ZEB1 as number 4. Furthermore, the pUC19 marker can be seen in the middle of the wells, and by comparing these results with the results, seen in *Figure 5.3, Amplification of Candidate target sequences*, it can be observed the sequences has the same number of bp. To make sure each candidate target sequence was correctly cloned, two of each mini-preps were sent to Macrogen for sequencing.

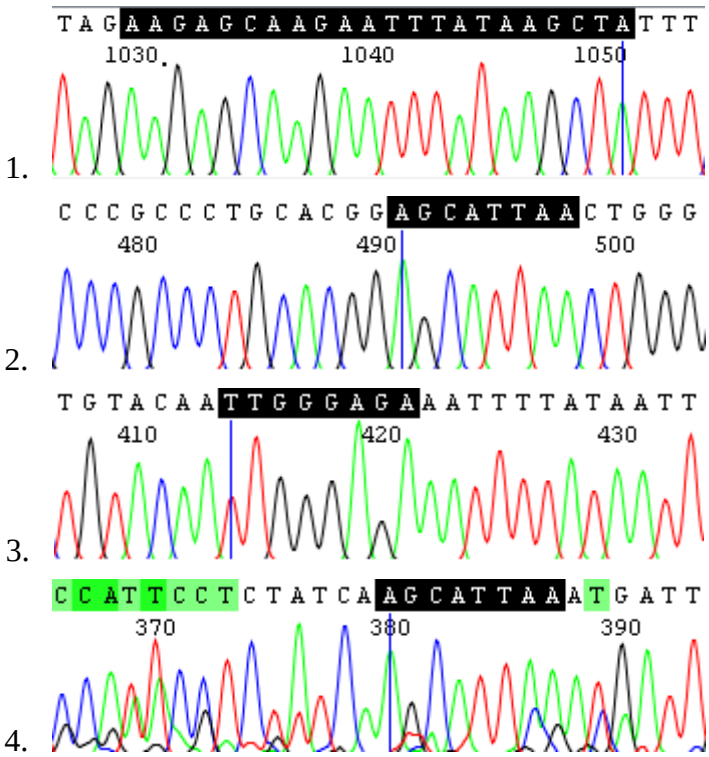


Figure 5.7, ***Result from sequencing Candidate target sequences in TOPO vector.*** *Black markings indicate the microRNA target sequences. Numbers from one to four is IL12A, SOCS1, ZEB1, and TRIM32*

The result from the sequencing is presented in *Figure 5.7, Results from sequencing Candidate target sequences in TOPO vector*, and by comparing these results with the target sequences found with aid from different databases, it can be determined the sequences are identical. In addition, also the linkers and the four extra nucleotides can be found in the results, in both the 3'UTR and 5'UTR end of the sequence, proving that the cloning was successful.

5.5 Subcloning candidate target 3'UTR sequences into psiCheck2 vector

As mentioned previously, the candidate 3'UTR target sequences were inserted into the TOPO vector, to make the transfer of the target sequences into the psiCheck2 vector easier. However, even if the candidate genes were inserted into the TOPO vector, difficulties with the insertion of the genes into the psiCheck2 vector was experienced, and several trials were conducted.

In the subcloning, each mini-prep potentially containing candidate target sequences were digested from the TOPO vector using the restriction enzymes *Xho* I and *Spe* I, and subsequently ligated with the psiCheck2 vector, and then transformed into competent *E.coli* cells, as described in *section 4.7 TOPO cloning of each candidate target sequence.* Transformations were plated onto 2xYT-AMP plates and colonies were picked for preparation of DNA mini-

preps. The DNA was purified, and as with the TOPO cloning, the DNA was analyzed by gel electrophoresis in an agarose gel, and the mini-prep DNA of the correct size sent to Macrogen for sequencing.

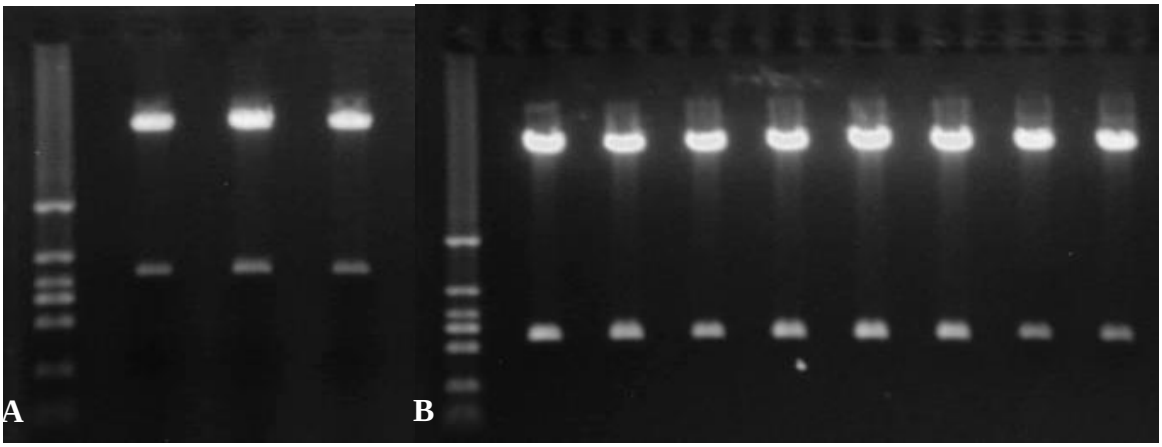


Figure 5.8, ***Insertion of IL12A (A) and ZEB1 (B) in psiCheck2.*** *In the figures presented here, the results from the gel electrophoresis of the Candidate gene IL12A marked A, and the Candidate gene ZEB1 marked B, is presented. From these, three IL12A was digested successfully and have the expected size (bp). Eight successfully digested mini-preps of ZEB1 can be seen in figure B, with the correct size of bp.*

From the transformation of the candidate genes into the psiCheck2 vector, three colonies of IL12A, eight colonies of SOCS1, eight colonies of ZEB1 and five colonies of TRIM32 appeared. By analyzing the DNA mini-prep digests with electrophoresis on a 1.8% agarose gel, comparison with the results presented in *section 5.3, Candidate target sequence amplification,* was possible. All colonies related to IL12A and ZEB1 had an insert, which is presented in *Figure 5.8, Insertion of IL12A (A) and ZEB1 (B) in psiCheck2.* The candidate target sequence SOCS1 had five of eight colonies with insert in the psiCHECK2 vector, as seen in *Figure 5.9, Insertion of the Candidate target sequence SOCS1 into psiCHECK2.* By comparing the size of these results with previous results, presented in *Figure 5.3, Amplification of Candidate target sequences*, it can be determined the bands of each colony had the right size, whereby two mini-preps from each candidate target sequence were selected for sequencing by Macrogen.

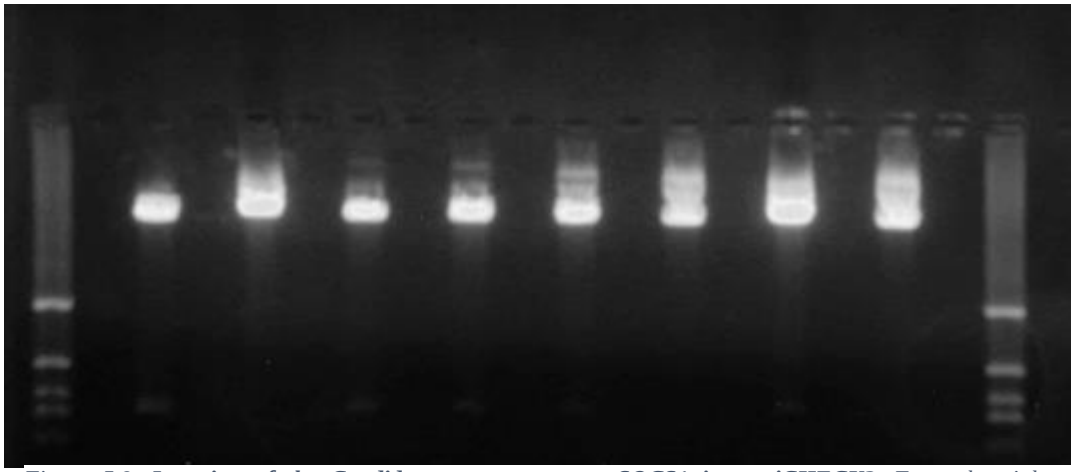


Figure 5.9, ***Insertion of the Candidate target sequence SOCS1 into psiCHECK2.*** *From the eight colonies tested in this gelelectrophresis, five inserts can be seen*

However, not all transformations were so successful, which can be seen with the candidate target sequence *TRIM32*, presented in *Figure 5.10, Insertion of TRIM32 in psiCHECK2*. The transformation with TRIM32 yielded resulted in five colonies from which DNA was propagated and purified. As judged from the digestion and gel electrophoresis, only one colony had a band of the correct size, whereby only this was sent for sequencing at Macrogen.

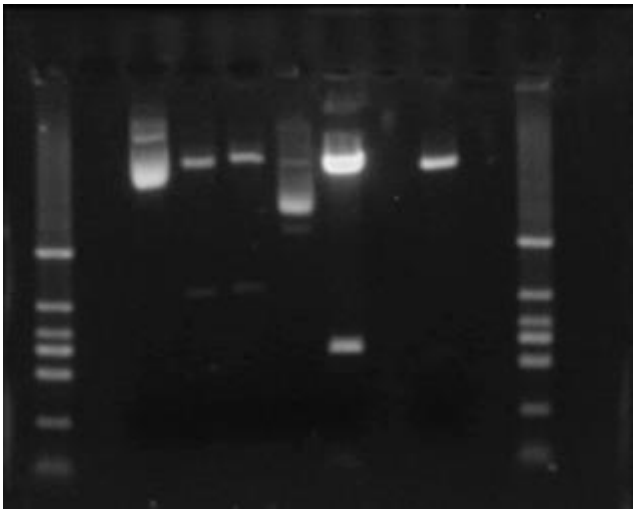


Figure 5.10, ***Insertion of TRIM32 in psiCHECK2.*** *The five colonies from the transformation of TRIM32 into the psiCHECK2 vector is here analyzed by gel electrophoresis. From this, only one of the five colonies had an insert the right size.*

The firm Macrogen conducted two sequencings, where one was conducted with the primer sets related to each candidate gene, and other was conducted with a new primer set made for the psiCHECK2 vector. In *Figure 5.11, Sequence result of the respective candidate 3'UTR target sequences in psiCHECK2,* the results from the sequencing with the primer sets related to each

candidate gene is presented, from which the presence of each Candidate gene in the vector can be determined.

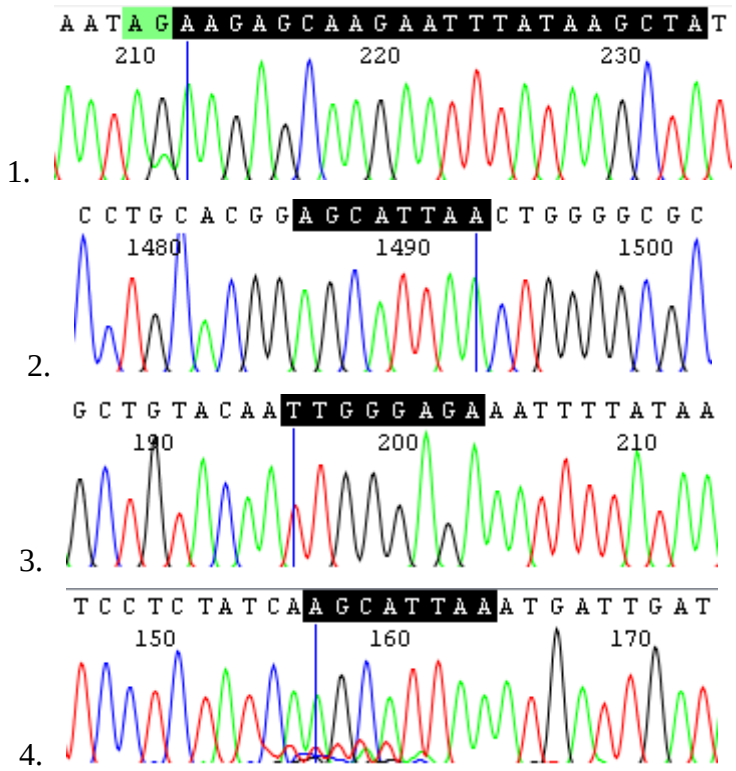


Figure 5.11, ***Sequence result of the respective candidate 3'UTR target sequences in psiCHECK2.*** *The results from the sequencing done by Macrogen shows the presence of all candidate target sequences in the vector*

The results from the vector-specific primer sets derived from the psiCHECK2 vector are presented in *Figure 5.12, Insert of candidate target sequences in psiCHECK2*. The miRNA target sites can be seen by the black markings, where just before the inserted four bp and the Linkers Xho1 can be found. When comparing the sequences received from Macrogen, with the sequence of the psiCHECK2 vector, which can be seen in *Appendix D*, only three of the four target sequences can be found in the multicloning site of the vector, which is IL12A (1), SOCS1 (2) and ZEB1 (3), by which a successful transformation can be determined.

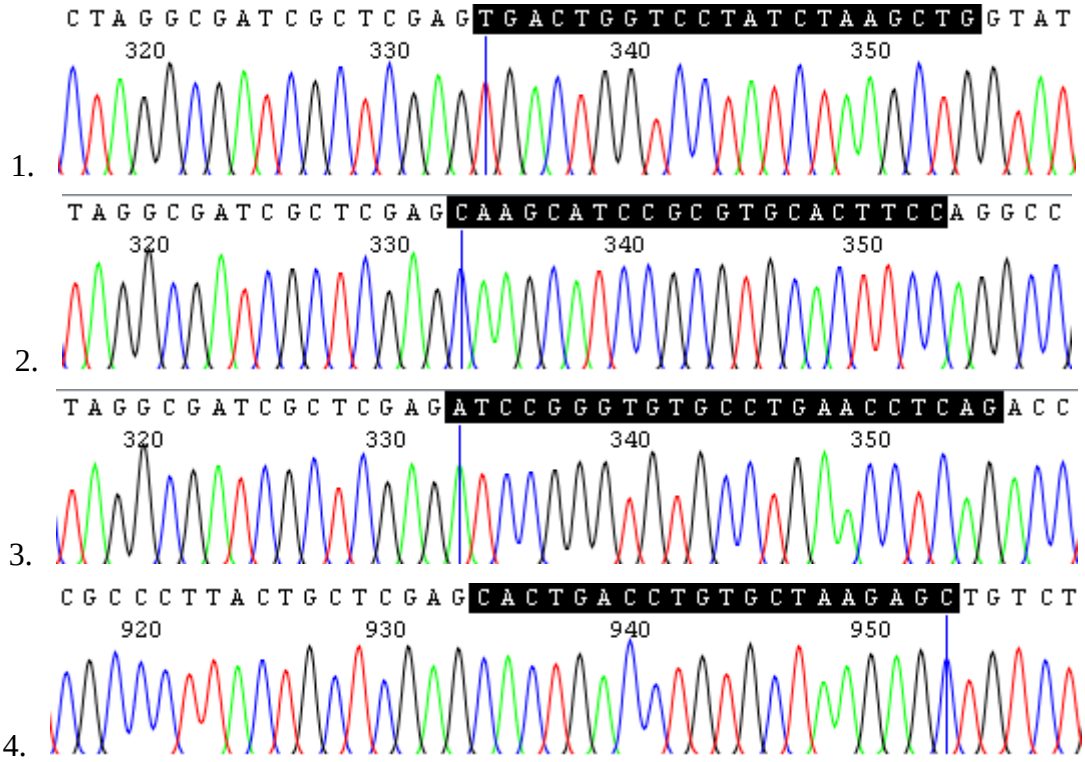


Figure 5.12, ***Insert of candidate target sequences in psiCHECK2****. Before all target sites related to all the candidate genes, extra inserted base pairs and the sequence for the linker Xho I can be observed by the sequence atcgCTCGAG*

For the last candidate target sequence, TRIM32 (4) the insert of the target sequence is present, however a small piece of the TOPO vector have been digested and inserted along with the candidate target sequence, and the candidate target sequence is not inserted in the same place at the 3' UTR, as the other target sequences. Due to the sequencing, it can be determined the candidate target sequence is present in the psiCHECK2 vector, which is seen in *Figure 5.12, Insert of candidate target sequences in psiCHECK2*, where also the end of the pCR-BLUNTII-TOPO vector can be observed. Even though the subclone was partly correct (in some degree), the candidate target sequence was excluded for subsequent trials conducted in this study, and only the clones, containing *IL12A, SOCS1, and ZEB1*, with a 100% success was used in the further investigation of the miRNA's *miR-155-5p, miR-150-5p, and miR-21-5p.*

5.6 Statistical Analysis of reporter gene expression regulated by microRNAs binding to candidate target sequences

From the data analysis of the expression of the reporter gene regulated by miRNA bindings to each candidate target sequence in the transfected and co-transfected HEK293 cells, effects of both miRCURY LNA miRNA Mimic and Power Inhibitor were evaluated.

The data outcome from the Luciferase Assay, was first corrected by subtraction of background signals related to the different growth pattern of the cells, where after the ratio between the Renilla Luciferase and Firefly Luciferase were calculated, which can be found in *Appendix F*. The ratios were used for further statistical analysis, in GraphPad Prism 8 (GraphPad Software), where boxplot with standard deviations over the expression of the reporter gene, regulated by miRNA bound to each candidate target sequence, and the corresponding control trial with psiCHECK2, a normality analysis, and a one-way ANOVA was performed.

5.6.1 One-way ANOVA

The ratio between Renilla luciferase and Firefly luciferase, previously calculated in Microsoft Office Excel 2016, were analyzed with the Shapiro-Wilks normality test in GraphPad Prism 8. The results obtained showed all data, except the psiCHECK2 positive control in the ZEB1 experiment, were normal distributed. This was supported by three scatterplots, displayed in *appendix G*, presenting the distribution of the ratio between Renilla luciferase and firefly luciferase, of each individual candidate target sequences and their related control trials with psiCHECK2.

Due to the normal distribution of the data, the parametric analysis One-way ANOVA was performed since more than two groups needed to be analyzed. The expression of each candidate target sequence and their correlated control trials were analyzed separately, since the mean expression influenced by the miRNA Mimic and Power Inhibitor needed to be compared with the positive control of reporter gene expression related to the candidate target sequence and the control trial psiCHECK2, respectively.

5.6.2 Expression of the Renilla reporter gene construct containing the candidate target sequence ZEB1

The mean expression of the reporter gene related to the candidate target sequence *ZEB1*, and the influence of miRCURY LNA miRNA Mimic and Power Inhibitor, HSA-miR-150-5p, is plotted in *Figure 5.13, Expression of the Renilla reporter gene construct containing the candidate target sequence for ZEB1 and psiCHECK2.* From this, as expected, a decrease in the expression of the reporter gene related to *ZEB1* can be observed, when influenced by the miRNA Mimic HSA-miR-150-5p. Additionally, an increase in the expression can be observed under the influence of miRNA Power Inhibitor. The control experiment, where psiCHECK2 was co-transfected with both miRCURY LNA miRNA Mimic and Power Inhibitor HSA-miR-150-5p, is displayed along with the expression of the reporter gene related to *ZEB1*. Here

almost no effect of the miRNA Mimic and Inhibitor can be observed, and it can be discussed whether this is caused by the miRNA Mimic or Inhibitor affecting the vector.

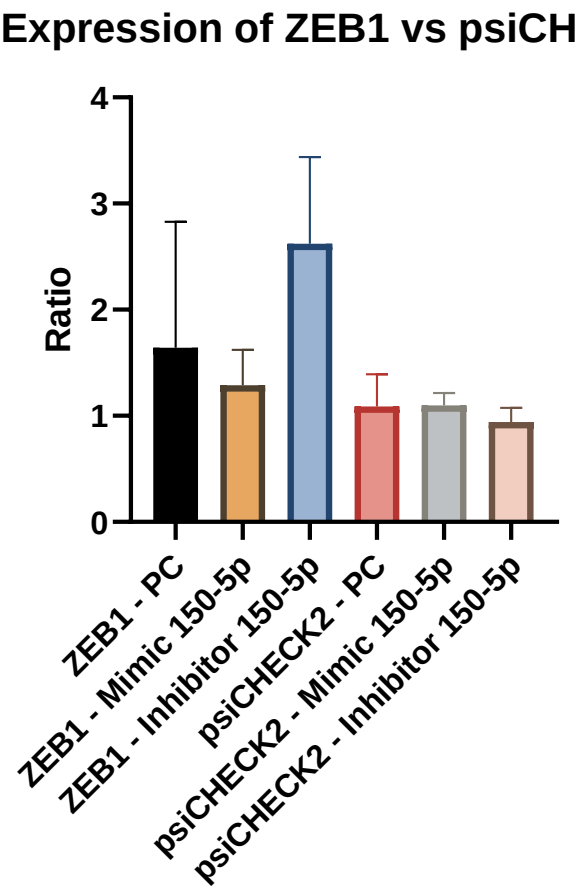


Figure 5.13, ***Expression of*** **the renilla reporter gene construct containing the** ***candidate target sequence*** **for** ***ZEB1 and psiCHECK2.*** *The mean expression of renilla is displayed in the first three columns, in the order of PC, under influence of Mimic and under influence of miRNA Inhibitor. From this, a decrease in the expression of renilla can be observed, while under influence of miRNA Mimic, and an increase in the expression, when influenced by miRNA inhibitor. The last three columns, display the mean expression of psiCHECK2, in the order of PC, influenced by miRNA Mimic, and lastly the expression of psiCHECK2 is displayed, while under influence of miRNA Inhibitor.*

Furthermore, the standard deviation, seen in *Figure 5.13*, is higher in ZEB1 - PC and the ZEB1 group influenced by Power Inhibitor HSA-miR-150-5p. This indicates there is a higher variation in these two groups when comparing the spread of all groups.

One-way ANOVA of reporter gene expression related to ZEB1

The one-way ANOVA analysis of the reporter gene expression related to ZEB1 resulted in a p-value of 0.008, which is under 0.05 and show a significant difference between the groups in terms of the means. The R-square value shows an amount of variance on 32.9% between the

three groups analyzed. In addition, the F-test, which is part of the one-way ANOVA, tell if the groups in the one-way ANOVA analysis have the same variation between them. In this analysis, Bartlett's test was used, since more than five data point were analyzed. The F-test displayed a p-value of 0.007, which lies under 0.05 and indicate the standard deviation is significantly different.

Post Hoc Analysis was performed to analyze which group was significantly different from the positive control. The Dunnet's post hoc test was performed and showed a p-value of 0.591 when comparing the mean ratio between the positive control and the expression influenced by miRNA Mimic. When comparing the mean ratio of the positive control and the expression influenced by miRNA Inhibitor, a p-value of 0.041 is displayed. This shows a statistically significant difference between the mean ratio of the pc and the expression influenced by miRNA Inhibitor.

The one-way ANOVA of the control group with psiCHECK2 had a p-value of 0.201, which is above the limit of statistical significance on 0.05. This indicates no significant difference between the groups was found. In addition, the R-square showed a variance between the groups of 12.5 %. Bartlett's F-test showed a p-value of 0.016, which indicate a significant difference between the standard deviations in the control group. No statistical significance was indicated for the One-way ANOVA, whereby no post hoc analysis was performed.

5.6.3 Expression of Renilla reporter gene construct containing the candidate gene IL12A

The expression of the Renilla reporter gene containing the candidate target gene IL12A is displayed in *Figure 5.14, Expression Renilla reporter gene containing the candidate target sequence of the IL12A and psiCHECK2*, respectively. From the first three columns in the figure, which display the expression related to the candidate target sequence IL12A, and the influence of either miRNA Mimic and Power Inhibitor HSA-miR-21-5p, variations in the expression can be observed. When comparing the columns depicting the expression under influence of miRNA Mimic with the positive control, a decrease can be observed and corresponding an increase when comparing the expression related to IL12A influenced by miRNA Inhibitor with the positive control. However, only small variations in the renilla reporter gene expression containing psiCHECK2 are seen in the last three columns in *Figure 5.13*. However, the variations observed does not have the expected pattern, since the expression influenced by both miRNA Mimic HSA-miR-21-5p and Power Inhibitor HSA-miR-21-5p has small increases.

The standard deviations presented in *Figure 5.14, Expression Renilla reporter gene containing the candidate target sequence of the IL12A and psiCHECK2,* are uniform with only a few elevated standard deviations. The elevations in the standard deviation can be observed for the reporter gene containing IL12A – PC, IL12A influenced by power inhibitor and in the expression of the reporter gene control for psiCHECK2 influenced by the power inhibitor. This indicates the data has a more even spread, when comparing to the spread of the data related to *ZEB1*, where higher elevations in the standard deviation were seen, for the ZEB1 – PC and ZEB1 influenced by the inhibitor.

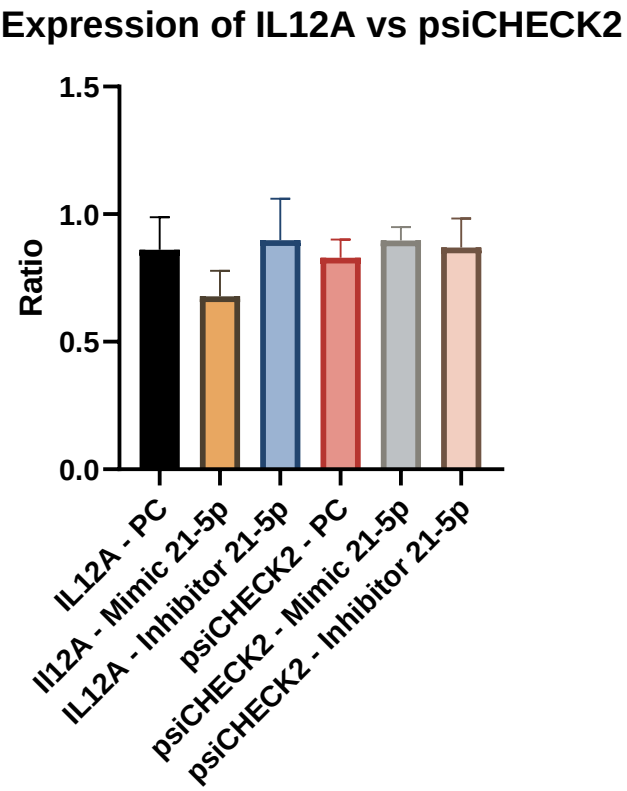


Figure 5.14, ***Expression*** **Renilla reporter gene containing the candidate target sequence** ***of the IL12A and psiCHECK2,*** **respectively.** *In the first three columns, the expression of the candidate target sequence IL12A is displayed, followed by the expression of psiCHECK2 in the next three columns. In the first column of the expression of IL12A can be observed, without influences from either miRNA Mimic or Inhibitor. In the second column, the expression of IL12A influenced by miRNA Mimic can be observed, where a small decrease in the expression is seen. In the third column, the expression of IL12A influenced by miRNA Inhibitor is displayed, and here an increase of the expression is observed. For the columns displaying the expression of psiCHECK2, both with and without influence from miRNA Mimic and Inhibitor, a small variation of the expression can be observed. However, both the expression of pdiCHECK2 influenced by miRNA Mimic and Inhibitor has a higher expression than the PC.*

5.6.3.1 Statistical Evaluation of the Renilla reporter gene construct IL12A and the control group psiCHECK2

The one-way ANOVA analysis of the reporter gene expression containing IL12A resulted in a p-value of 0.004, which indicate a statistically significant difference between the groups in terms of the mean. The R-square show a 37.3% difference in the amount of variance between the three groups. The Bartlett's F-test has a p-value of 0.416, which show the standard deviation between the three expression groups of IL12A not are significantly different.

Dunnet's post hoc Analysis showed a p-value of 0.014 when comparing the mean ratio of the PC and the expression related to *IL12A* influenced by miRNA Mimic. This indicates a significant difference between these two groups. For the comparison of the mean ratio of the PC and the expression related to *IL12A* influenced by miRNA Power Inhibitor, a p-value of 0.777 was calculated, thereby indicating the difference between these two groups were not statistically significant.

With the control group with psiCHECK2, the one-way ANOVA showed a p-value of 0.250, which indicate no statistical difference between the groups were present. This was backed up by the R-square, which showed an amount of variation between the three groups on 10 %. Furthermore, the Bartlett's F-test showed the standard deviation had no significant difference since the p-value was on 0.101 and thereby above 0.05. Since no significance was found, the post hoc test was not performed.

5.6.4 Expression of renilla reporter gene construct containing the candidate gene SOCS1

In *Figure 5.15, Expression of renilla reporter gene construct with SOCS1 and the control psiCHECK2,* the expression of the reporter gene related to SOCS1, both with and without influence of miRNA Mimic and Inhibitor HSA-miR-155-5p, is presented in the first three columns. From the mean ratio in *Figure 5.14*, a decrease in the expression is seen, when influenced by miRNA Mimic HSA-miR-155-5p, which is expected. In contrast, a small decrease in the reporter gene expression of SOCS1 is seen, when the expression is influenced by miRNA Power Inhibitor, which is not expected. The expression related to SOCS1 is higher when influenced by miRNA Inhibitor than when under influence of miRNA Mimic, but still not higher than the positive control, which was expected. The standard deviation presented in *Figure 5.14*, show a higher standard deviation in the PC, than in the reporter gene expression

related to SOCS1 influenced by miRNA Mimic and Power Inhibitor. This indicates the raw data of the PC have a higher spread when comparing the data.

For the psiCHECK2 control trial, a decrease in the expression is observed, when influenced by both miRNA Mimic HSA-miR-155-5p and Power Inhibitor HSA-miR-155-5p. From this, it can be speculated whether the psiCHECK2 vector contains a target site for the miRNA, *miR-155-5p*. Furthermore, it can be speculated whether the behavior of the HEK293 cells is influencing the miRNA expression. In addition, the results might indicate a higher toxicity level of the miRNA Mimic and Power Inhibitor, than the PC. This might be caused by the different growth pattern observed by the HEK293 cells.

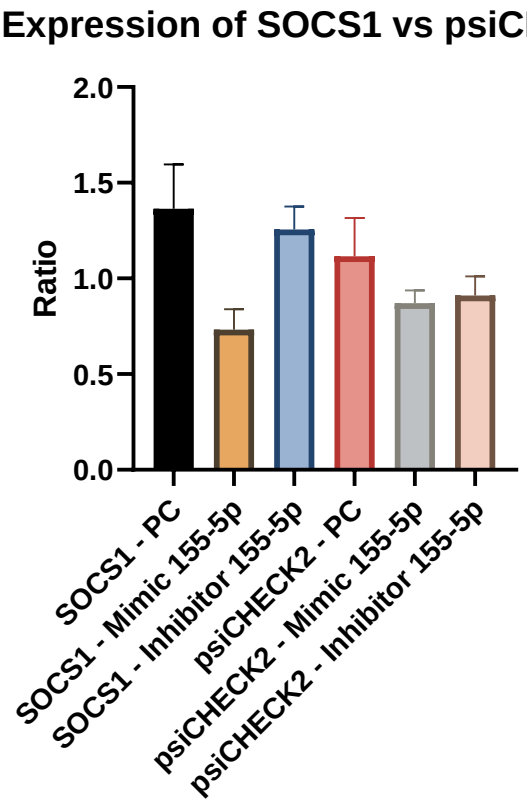


Figure 5.15, ***Expression of Renilla reporter gene construct with SOCS1 and the control psiCHECK2.*** *The expression of the candidate target sequence SOCS1 is depicted in the first three columns, in the order of PC, expression influenced by miRNA Mimic and the expression of SOCS1 influenced by miRNA Inhibitor. A decrease in the expression of SOCS1 influenced by miRNA Mimic can be observed when comparing with the expression PC, which has no influence of miRNA Mimic or miRNA inhibitor. In addition, a small decrease in the expression influenced by the miRNA Inhibitor can be observed. However, the expression of SOCS1 is still increased under influence of miRNA Inhibitor, when comparing with the expression of SOCS1 influenced by miRNA Mimic*

The normal distribution, presented in *Figure 5.15*, show the same tendency for the different groups. For both standard deviations in the groups influenced by miRNA Mimic or miRNA Power Inhibitor, show a lower standard deviation than what is seen in the PC of the trials. Furthermore, the miRNA Mimic standard deviations are lower than the miRNA Power Inhibitor standard deviations. However, the standard deviations in general, are a bit higher for the expression of the reporter gene containing the candidate target sequence SOCS1, than the control trial psiCHECK2. This indicates a higher spread in the expression, when the cells are not influenced by miRNA Mimic or miRNA Power Inhibitor, and it indicates a higher spread for the reporter gene expression related to the candidate target sequence in general when compared with the control trial.

In general, a lower spread of the data can be observed in the control trials with psiCHECK2, than in the trials with the candidate target sequence. Furthermore, higher variations in the expression of the candidate target sequences can be observed, when comparing to their correlated control trials with psiCHECK2.

5.6.4.1 Statistical Evaluation of the Renilla reporter gene construct SOCS1

The one-way ANOVA analysis of the Expression of SOCS1, gave a p-value under 0.001, which is indicated a statistically significant difference between the groups. In relation, the R-square showed a variance of 76.3% between the groups. Bartlett's F-test displayed a p-value of 0.059, which indicate the standard deviation of the expression of SOCS1 was not significantly different.

In addition, ***Dunnet's post hoc Analysis*** showed a p-value of 0.001, when comparing the mean expression of SOCS1 positive control and the expression of SOCS1 influenced by miRNA Mimic, which indicates a significant difference between the two groups. However, when comparing the positive control with the expression of SOCS1 influenced by the miRNA Power Inhibitor, a p-value of 0.284, show no significant difference between the positive control and the SOCS1 expression influenced by miRNA Power Inhibitor.

The control group with psiCHECK2 had a p-value of 0.001 in the one-way ANOVA, which indicate there is a significant difference between the groups. In relation, the R-square show an amount of variation of 42% between the three groups analyzed. The Bartlett's F-test, show a p-value of 0.010, which is below 0.05. This indicates the standard deviations of the control

trials with psiCHECK2 is significantly different from each other. Due to the significant difference showed with the one-way ANOVA, Dunnet's post hoc analysis was performed.

The ***Dunnet's post hoc Analysis*** compared the two groups influenced by either miRNA Mimic or miRNA Power Inhibitor with the PC. The comparison with the expression influenced by miRNA and the PC resulted in a p-value of 0.001, and the p-value when comparing the expression influenced by miRNA Power Inhibitor and the PC, resulted in a p-value of 0.006, which indicate that both groups were significantly different from the control group.

6. Discussion

The findings of this study suggest IL12A, ZEB1, and SOCS1 as target genes to the related miRNAs, *miR-21-2p*, *miR-150-5p* and *miR-155-5p.* This was confirmed by the statistical evaluation of the Renilla reporter gene expression regulated by the miRNAs binding to the candidate target sequences. The statistical evaluation showed statistically significant differences related to each miRNAs binding to their respective target gene. However, significant differences were also found in the control experiment related to the miRNA, *miR-155-5p.* A post hoc analysis was performed for all Renilla reporter gene expressions with significant differences and revealed significant differences of the reporter gene expression influenced by miRNA Mimic HSA-miR-21-5p, miRNA Mimic HSA-miR-155-5p, and the miRNA power inhibitor HSA-miR-150-5p. Furthermore, statistical differences were found for the Renilla reporter gene expression, in the control experiment, influenced by both the miRNA Mimic and power inhibitor related to the miRNA *miR-155-5p.* When comparing the results of the Renilla reporter gene expressions with the control expression, only the control experiment related to miR-155-5p had a statistical difference between the positive control and the related groups. The regulation of the control experiment, which was not seen in relation to the miR-21-5p and miR-150-5p, might indicate the presence of a miR-155-5p target site in the vector. However, no studies explaining this regulation was found, whereby further research in this area is needed.

Previous studies indicate that expression of 60% of the genes in the human genome is regulated by miRNA, which has been associated with both development and general function of the body (24). Additionally, associations with the development of autoimmune diseases like rheumatoid arthritis and SLE have been conducted(5). Due to the increased awareness of IBD, becoming a global disease influencing 0.5% of the general population, and with unsatisfying available treatments, researchers are focusing on the possibility of using miRNA as a therapeutic treatment against IBD (24). However, many miRNAs have been discovered, but the identification of target genes have been difficult, and challenges with associating target genes and diseases require more research of e.g. target genes and their function in diseases.

Despite this study discovered the association between three target genes and their related miRNAs, the prediction of the target genes, the cloning process, with the cell culturing of

HEK293 and additionally the transfection of the cells and the measurement on the expression of the Renilla reporter gene have been challenging. As mentioned previously the data in the prediction tool available are very hard to compare, since not much coherent data is available. The prediction tool TargetScan was primarily used in this study, which shows the best performance when comparing data available in prediction tool (24). However, with TargetScan, only targets with stringent seeds recognition sequences are found whereby possible target with a non-stringent seed was not considered as a target gene. However, a moderate-stringent seed site would increase the sensitivity and find more possible targets, but this seed site is not very specific, like the stringent seed site, and thereby increase the chances of a false-positive seed site (24). Furthermore, the prediction of the target sites was conducted according to the conservation of the miRNA in both human and pig, where some are species-specific and was not revealed in the search for relevant target genes to investigate. the conversation of the miRNA decrease the possibility of false-positive target sites but only work on conserved miRNAs by which not all target sites was revealed and taken into consideration. However, the conservation of the miRNA also increase the chances of investigating the function and biology related to the miRNA, which this project might lead to in further studies of the target genes.

Each target gene was inserted into first the TOPO cloning vector, and after this into the reporter gene vector psiCHECK2 containing structural genes for luciferase and renilla. The process with the insertion of target genes into a TOPO cloning vector was very time consuming, due to the initial lack of results. However, three out of four clonings were successful, whereby the successful cloning with the target genes IL12A, ZEB1 and SOCS1 were used for further analysis. The influence of the specific restriction enzymes and their influence on the cloning vector and the reporter gene vector psiCHECK2 was thoroughly investigated, but the target gene TRIM32 was only partly successful. Several precautions were taken by several sequencings between each step of the cloning process, and steps were optimized by e.g. changing buffers used during the transformation. Even so the digestion by the restriction enzymes of the TRIM32 gene in the TOPO cloning vector partly failed, and a piece of the TOPO cloning vector was inserted in the Renilla reporter gene vector psiCHECK2. By this, and the lack of knowledge related to the associations between the miRNA investigated and the TOPO cloning vector, the target gene TRIM32 was excluded for further investigation.

Before the transfection of the reporter plasmid with inserted target sequence into the HEK293 cells, cell culturing and cryopreservation of the HEK293 cells was conducted. During the cell culturing, adhesion problems of the cells were observed. Furthermore, during this process, the growth pattern of the cells varied, to which no explanation was found. The growth media used was composed containing DMEM + GlutaMax, 10% FCS and 1% pen/strep, and additionally, the same procedures of subcultivation were used during the cell culture. In addition, the HEK293 cell is normally straightforward to grow in cell culture and is used as a host for gene expression, including transfections. However, the cell culture was obtained from another group at Department of Molecular Biology and Genetics at Aarhus University, and the passage number and cell age were unknown, by which the influence of both age and passage number can be speculated.

Before the transfection of the plasmid with the inserted target sequence, several optimization studies were conducted. This process was very time consuming since many parameters were tested, and if just one of the parameters changed this affected all of them. The components Lipofectamine 2000, miRNA Mimic and miRNA Power Inhibitor were all toxic above certain concentrations, by which these relations were thoroughly investigated one parameter at the time. Testing was initiated with the number of cells seeded in a 24-well plate to obtain a confluence of 70%. This was followed by varying the added amount of Lipofectamine 2000 to achieve the highest measures of the Renilla reporter gene and continued by investigating the amount of plasmid DNA to transfect to achieve the best measure of the reporter gene. When these optimal parameters were set, the same evaluations were conducted related to the miRNA Mimic and the miRNA power inhibitor. The transfection of the HEK293 cells with the reporter vector constructs harboring the target gene sequences was conducted as technical triplicates, and during this time period, the HEK293 cells growth pattern decreased extensively, which was reflected in the results. At the end of the trial period, toxicity levels could be observed after transfection, which due to the extensive optimization trials should have been avoided.

After transfection, the expression of renilla protein from the HEK293 cells was assessed with a Luciferase reporter assay kit. Since a technical triplicate was conducted a lot of samples needed to be analyzed. The transfected cells were first lysed with passive lysis buffer, where after a small amount of each sample was transferred to a 96-well plate. During the optimization trials, the effect of the time used between the measure of Luciferase and Renilla was discovered to be very influential, and by this, the measure of the Renilla reporter gene

expression was conducted over several turns to avoid changing time intervals of the measures. Furthermore, the results were conditioned of the number of cells transfected, the toxicity level of the components used and the growth pattern of the cells. In the first trials conducted with SOCS1, the cells grew between the transfection and the measure with luciferase assay, where they in the last trial with IL12A did not grow between the transfection and the analysis with luciferase assay, which caused toxicity levels to increase and cell death was observed.

7. Conclusion

In conclusion, by regulation of the reporter gene expression, this thesis has shown interactions between the miRNAs miR-21-5p, miR-150-5p and miR-155-5p, and their predicted target genes. These findings provide a good contribution for the further investigation of a possible therapeutic treatment, but still, more research is needed on both the regulation of the target genes and the effect of regulations by either mimicking or inhibiting the miRNA. Furthermore, the findings of this study showed interactions between the miRNA and the psiCHECK2 vector, which might indicate the presence of a miRNA target site in psiCHECK2.

8. References

1. Burisch J, Munkholm P. Inflammatory bowel disease epidemiology. Curr Opin Gastroenterol. 2013;29(4):357–62.
2. Kaplan GG. The global burden of IBD: From 2015 to 2025. Nat Rev Gastroenterol Hepatol [Internet]. 2015;12(12):720–7. Available from: http://dx.doi.org/10.1038/nrgastro.2015.150
3. Chan S-N, Low EN Den, Raja Ali RA, Mokhtar NM. Delineating inflammatory bowel disease through transcriptomic studies: current review of progress and evidence. Intest Res. 2018;16(3):374.
4. Ye Y, Phang Z, Chen W, Ju S, Zhou C. Epidemiology and risk factors of inflammatory bowel diseases. Bull Acad Natl Med. 2007;191(6):1105–23.
5. Rohr M, Chandrakala Aluganti Narasimhulu, Dhara Sharma, Mitsushita Doomra, Aladdin Riad, Saleh Naser and SP. Inflammatory diseases of the Gut. Pediatr Crit Care Med Vol 2 Respir Cardiovasc Cent Nerv Syst. 2014;21(2):467–81.
6. Crowley E, Muise A. Inflammatory Bowel Disease: What Very Early Onset Disease Teaches Us. Gastroenterol Clin North Am [Internet]. 2018;47(4):755–72. Available from: https://doi.org/10.1016/j.gtc.2018.07.004
7. Meddens CA, van der List ACJ, Nieuwenhuis EES, Mokry M. Non-coding DNA in IBD: from sequence variation in DNA regulatory elements to novel therapeutic potential. Gut. 2019;68(5):928–41.
8. Schaefer JS. miRNAs: how many in IBD? Curr Opin Gastroenterol [Internet]. 2016;32(4):258–66. Available from:
9. Meerveld BG-V, Johnson AC, Grundy D. Gastrointestina Physiology and Function. In: Handbook of Experimental Pharmacology.
10. Hall JE. Guyton and Hall Texbook of Medical Physiology. 2012.
11. Ahluwalia B, Magnusson MK, Öhman L. Mucosal immune system of the gastrointestinal tract: maintaining balance between the good and the bad. Scand J Gastroenterol. 2017;52(11):1185–93.

12. Agger R, Nielsen CH, Leslie G, Aasted B. Immunologi. 1. udgave, 2012.

13. Ananthakrishnan AN, Bernstein CN, Iliopoulos D, Macpherson A, Neurath MF, Ali RAR, et al. Environmental triggers in IBD: A review of progress and evidence. Nat Rev Gastroenterol Hepatol [Internet]. 2018;15(1):39–49. Available from: http://dx.doi.org/10.1038/nrgastro.2017.136

14. Dethlefsen L, Relman DA. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. Proc Natl Acad Sci U S A [Internet]. 2011;108 Suppl:4554–61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20847294%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3063582

15. Kaplan GG, Hubbard J, Korzenik J, Sands BE, Panaccione R, Ghosh S, et al. The inflammatory bowel diseases and ambient air pollution: A novel association. Am J Gastroenterol [Internet]. 2010;105(11):2412–9. Available from: http://dx.doi.org/10.1038/ajg.2010.252

16. Opstelten JL, Beelen RMJ, Leenders M, Hoek G, Brunekreef B, van Schaik FDM, et al. Exposure to Ambient Air Pollution and the Risk of Inflammatory Bowel Disease: A European Nested Case–Control Study. Dig Dis Sci. 2016;61(10):2963–71.

17. Hugot JP, Laurent-Puig P, Gower-Rousseau C, Olson JM, Lee JC, Beaugerie L, et al. Mapping of a susceptibility locus for Crohn's disease on chromosome 16. Nature. 1996;379(6568):821–3.

18. O'Brien J, Hayder H, Zayed Y, Peng C. Overview of microRNA biogenesis, mechanisms of actions, and circulation. Front Endocrinol (Lausanne). 2018;9(AUG):1–12.

19. Zhang L, Cheng J, Fan XM. MicroRNAs: New therapeutic targets for intestinal barrier dysfunction. World J Gastroenterol. 2014;20(19):5818–25.

20. Bian Z, Li L, Cui J, Zhang H, Liu Y, Zhang CY, et al. Role of miR-150-targeting c-Myb in colonic epithelial disruption during dextran sulphate sodium-induced murine experimental colitis and human ulcerative colitis. J Pathol. 2011;225(4):544–53.

21. Lu LF, Liston A. MicroRNA in the immune system, microRNA as an immune system. Immunology. 2009;127(3):291–8.

22. Balzola F, Bernstein C, Ho GT, Lees C. Identification of restricted subsets of mature microRNA abnormally expressed in inactive colonic mucosa of patients with inflammatory bowel disease: Commentary. Inflamm Bowel Dis Monit. 2011;11(3):126–7.

23. Zahm AM, Thayu M, Hand NJ, Horner A, Leonard MB, Friedman JR. Circulating microRNA is a biomarker of pediatric crohn disease. J Pediatr Gastroenterol Nutr. 2011;53(1):26–33.

24. Saito T, Sætrom P. MicroRNAs - targeting and target prediction. N Biotechnol. 2010;27(3):243–9.

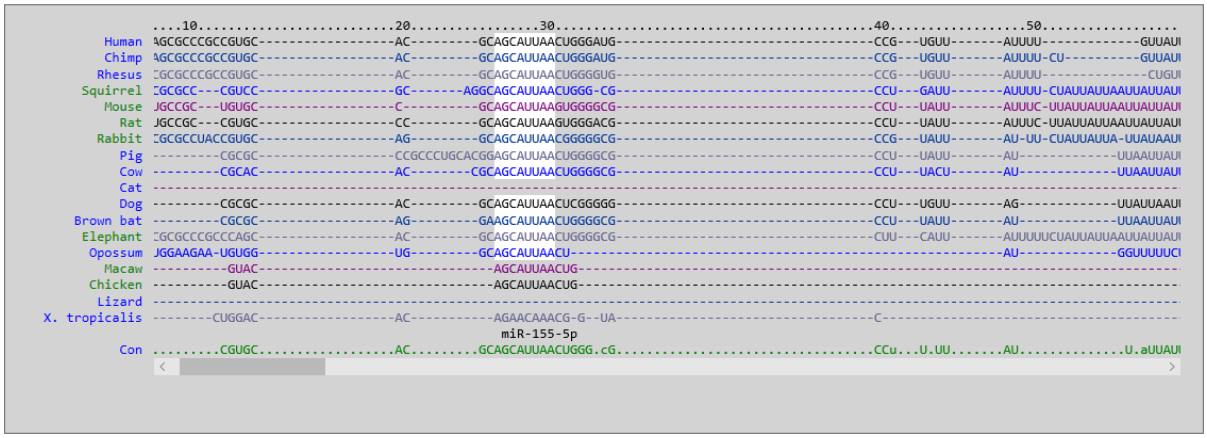
25. Sethupathy P, Megraw M, Hatzigeorgiou AG. A guide through present computational approaches for the identification of mammalian microRNA targets. Nat Methods. 2006;3(11):881–6.

Appendix

Appendix A

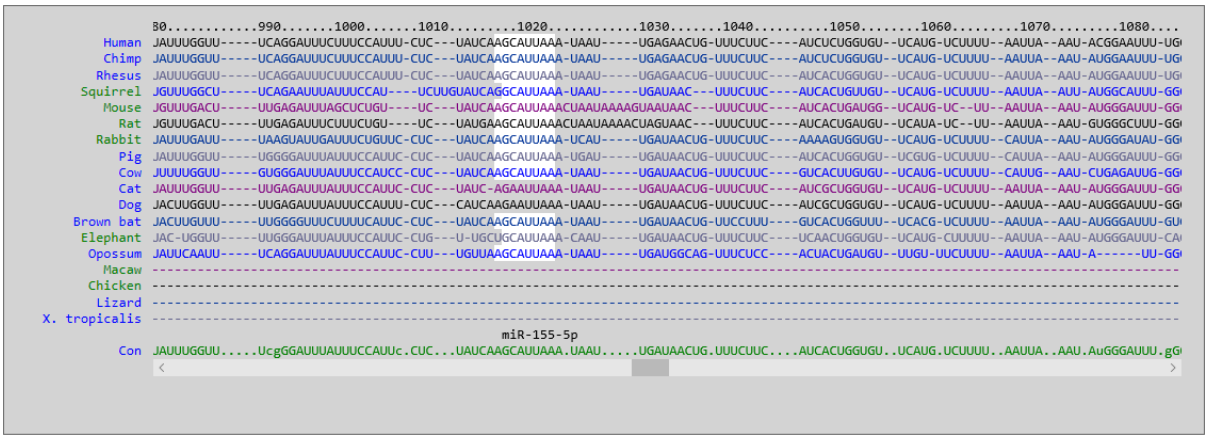
Target sites of the candidate genes identifying conservation cross species.

Target site - SOCS1



Target site – IL12A

Target site – TRIM32



Appendix B

Translation direction

SOCS1

	Predicted consequential pairing of target region (top) and miRNA (bottom)	Site type	Context++ score	Context++ score percentile	Weighted context++ score	Conserved branch length	P_{CT}
Position 24-31 of SOCS1 3' UTR hsa-miR-155-5p	5' ...GCCCGCCGUGCACGC--AGCAUUAA... 3' UGGGGAUAGUGCUAAUCGUAAUU	8mer	-0.33	97	-0.33	3.691	0.38

IL12A

	Predicted consequential pairing of target region (top) and miRNA (bottom)	Site type	Context++ score	Context++ score percentile	Weighted context++ score	Conserved branch length	P_{CT}
Position 256-263 of IL12A 3' UTR hsa-miR-21-5p	5' ...GAAGGGCAAAUAUUUAUAAGCUA... 3' AGUUGUAGUCAGACUAUUCGAU	8mer	-0.65	99	-0.65	4.829	0.73
Position 256-263 of IL12A 3' UTR hsa-miR-590-5p	5' ...GAAGGGCAAAUAUUUAUAAGCUA... 3' GACGUGAAAAUACUUAUUCGAG	8mer	-0.59	99	-0.59	4.829	0.73

TRIM32

	Predicted consequential pairing of target region (top) and miRNA (bottom)	Site type	Context++ score	Context++ score percentile	Weighted context++ score	Conserved branch length	P_{CT}
Position 1017-1024 of TRIM32 3' UTR hsa-miR-155-5p	5' ...UCCAUUUCUCUAUCA-----AGCAUUAA... 3' UGGGGAUAGUGCUAAUCGUAAUU	8mer	-0.40	98	-0.39	3.279	0.28

Appendix C

General guidelines for primer design.

General Guidelines for Primer Design
- Primers length from 18-25 nucleotides
- Primer T_m optimal at 60-70°C; T_m for P_1 and P_2 should be within 5°C
- GC>50%, tend to finish 3´with G or C
- Have a balanced proportion between GC and AT rich domains if possible
- Avoid four or more repeats of certain base or dinucleotids e.g. ACCCC or ATATATATATAT
- Avoid primer homology between the two used primers, since this can create primer-dimers

Table 3.7, ***Guidelines used for Manually Primer Design.***

Appendix D

Sequence and primer designed for psiCHECK2

Forward primer:

Antal bp: 28 bp

Gc indhold: 16 ≈ 57 %

Sequence: 5' – gttcgctgcctacctggagccattcaag – 3'

Reverse primer:

Antal bp: 24 bp

GC indhold: 11 ≈ 46 %

Sequence: 5' – gcttaagaggatcgcggcgctcttctgc – 3'

```
1 agatctgcgc agcaccatgg cctgaaataa cctctgaaag aggaacttgg ttaggtacct
       61 tctgaggcgg aaagaaccag ctgtggaatg tgtgtcagtt agggtgtgga aagtccccag
      121 gctccccagc aggcagaagt atgcaaagca tgcatctcaa ttagtcagca accaggtgtg
      181 gaaagtcccc aggctcccca gcaggcagaa gtatgcaaag catgcatctc aattagtcag
      241 caaccatagt cccgccccta actccgccca tcccgcccct aactccgccc agttccgccc
      301 attctccgcc ccatggctga ctaatttttt ttatttatgc agaggccgag gccgcctcgg
      361 cctctgagct attccagaag tagtgaggag gcttttttgg aggcctaggc ttttgcaaaa
      421 agcttgattc ttctgacaca acagtctcga acttaagctg cagaagttgg tcgtgaggca
      481 ctgggcaggt aagtatcaag gttacaagac aggtttaagg agaccaatag aaactgggct
      541 tgtcgagaca gagaagactc ttgcgtttct gataggcacc tattggtctt actgacatcc
      601 actttgcctt tctctccaca ggtgtccact cccagttcaa ttacagctct taaggctaga
      661 gtacttaata cgactcacta taggctagcc accatggctt ccaaggtgta cgaccccgag
      721 caacgcaaac gcatgatcac tgggcctcag tggtgggctc gctgcaagca aatgaacgtg
      781 ctggactcct tcatcaacta ctatgattcc gagaagcacg ccgagaacgc cgtgattttt
      841 ctgcatggta acgctgcctc cagctacctg tggaggcacg tcgtgcctca catcgagccc
      901 gtggctagat gcatcatccc tgatctgatc ggaatgggta agtccggcaa gagcgggaat
      961 ggctcatatc gcctcctgga tcactacaag tacctcaccg cttggttcga gctgctgaac
     1021 cttccaaaga aaatcatctt tgtgggccac gactgggggg cttgtctggc ctttcactac
     1081 tcctacgagc accaagacaa gatcaaggcc atcgtccatg ctgagagtgt cgtggacgtg
     1141 atcgagtcct gggacgagtg gcctgacatc gaggaggata tcgccctgat caagagcgaa
     1201 gagggcgaga aaatggtgct tgagaataac ttcttcgtcg agaccatgct cccaagcaag
     1261 atcatgcgga aactggagcc tgaggagttc gctgcctacc tggagccatt caaggagaag
     1321 ggcgaggtta gacggcctac cctctcctgg cctcgcgaga tccctctcgt taagggaggc
     1381 aagcccgacg tcgtccagat tgtccgcaac tacaacgcct accttcgggc cagcgacgat
     1441 ctgcctaaga tgttcatcga gtccgaccct gggttctttt ccaacgctat tgtcgaggga
     1501 gctaagaagt tccctaacac cgagttcgtg aaggtgaagg gcctccactt cagccaggag
     1561 gacgctccag atgaaatggg taagtacatc aagagcttcg tggagcgcgt gctgaagaac
     1621 gagcagtaat tctaggcgat cgctcgagcc cgggaattcg tttaaaccta gagcggccgc
     1681 tggccgcaat aaaatatctt tattttcatt acatctgtgt gttggttttt tgtgtgagga
     1741 tctaaatgag tcttcggacc tcgcgggggc cgcttaagcg gtggttaggg tttgtctgac
     1801 gcgggggggag ggggaaggaa cgaaacactc tcattcggag gcggctcggg gtttggtctt
     1861 ggtggccacg ggcacgcaga agagcgccgc gatcctctta agcaccccc cgccctccgt
     1921 ggaggcgggg gtttggtcgg cgggtggtaa ctggcgggcc gctgactcgg gcgggtcgcg
     1981 cgccccagag tgtgaccttt tcggtctgct cgcagacccc cgggcggcgc cgccgcggcg
```

Appendix E

This appendix displays the full sequences of each candidate target sequence. The primer sites for each sequence and the target site for the sequence redulated by the miRNA is marked.

>ssc-miR-155-5p MIMAT0022959 UUAAUGCUAAUUGUGAUAGGGG

SOCS1 Homo sapiens suppressor of cytokine signaling 1; chr3:31882192-31882520

5´-CGCGCCCGCCCUGCACGGAGCAUUAACUGGGGCGCCUUAUUAUUUAAUUAU-3´

3´-UGGGGAUAGUGCUAAUCGUAAUU-5´

```
31881641 CCACGCCCCCCTCTCTTAACCAAACTTTGAGCGACTCCGAGCCCCACCGAGCTCCCAGCC 31881700
31881701 GCGTGGAAGTACTCTGCGGACGGGGAAGCCCATTTTCTAGATGAGACCACTGAGGCTCAA 31881760
31881761 AGGGCCGCCCCGACCTGGTCTTGCGCCCTCGACCCATTAGAAACGGATCCCCACCCCTCC 31881820
31881821 TAAGTTGCCTCGCGGGTTCTCACCTGCCCAGCCCCACCCAGACTCCCCGCTCACTGACTG 31881880
31881881 TCTCCCCCATCAGCGCACCCCCGGACGCTATGGCCCACCCCTCCGGCCGACCCCGCGTGT 31881940
31881941 AGGATGGTAGCACACAACCAGGTGGCAGCCGACAATGCAATCTCCACGGCAGCAGAGCCC 31882000
31882001 CGACGGCGGCCAGAGCATTCCTCCTCCTCCTCCTCCTCTTCATCTTCATCCTCCTCCTCC 31882060
31882061 TCCTCACCCGGGGTCCCGGCGCGCCCGCGGCCCTGCCCGGCGGCCCCGGCTCCAGCCCCG 31882120
31882121 GGCGACACGCACTTCCGCACGTTCCGCTCGCACGCCGACTACCGGCGCATCACCCGCGCC 31882180
31882181 AGCGCGCTCCTCGACGCCTGCGGCTTCTACTGGGGGCCTCTGAGCGTGCACGGGGCGCAC 31882240
31882241 GAGCGGCTGCGCGCCGAGCCCGTGGGCACCTTCCTGGTGCGCGACAGCCGCCAGCGGAAC 31882300
31882301 TGCTTCTTCGCCCTCAGTGTGAAGATGGCTTCGGGCCCCACAAGCATCCGCGTGCACTTC 31882360
31882361 CAGGCCGGCCGCTTCCACCTGGACGGCAGCCGCGAGAGCTTCGACTGCCTCTTCGAGCTG 31882420
31882421 CTGGAGCACTACGTGGCGGCGCCGCGCCGCATGCTGGGGGCCCCGCTGCGCCAGCGCCGC 31882480
31882481 GTGCGGCCGCTGCAGGAGCTGTGCCGCCAGCGCATTGTGGCTACCGTGGGCCGCGAGAAC 31882540
31882541 CTGGCGCGCATCCCCCTCAACCCCGTCCTCCGCGATTACTTGAGCTCCTTCCCCTTCCAG 31882600
31882601 ATATGATCGGCCGCGCCCGCCCTGCACGGAGCATTAACTGGGGCGCCTTATTATTTTCTA 31882660
31882661 TTATTAATTATTATTTCCTTGGAACCATGTGGGTTGGAGGGAGCGAGTGTGGGGGCGAGG 31882720
31882721 CGCCTCCCGCCCTCGGCTGGAGACCCCTCCCGTAGACCCCTCCCCACCTCTTGTGGGGGT 31882780
31882781 GCCCCCTCTTGGTGCTCCCTCTGGGTCCCCCTGGTTGTCCCCCTGGTAGTAGCTTAACTT 31882840
31882841 AACTATCTGGGGCCAGGACCTGAACTCAGTACCTCCTACCTCTTCATGTTTACATATACC 31882900
31882901 CAGTATCTTTGCACAAACCAGGGGTTGGGGGAGGGTCTCTGGCTTTATTTTTCTGCTGTG 31882960
31882961 CAGAATCCTATT
```

TRIM32 Tripartite motif containing 32; chr1:257063847-257066911

5´-ACAACAUAACUUUAGAAAACAAAGCAUUAAAAAAAAAGUCCAUCAGUAUUAUGGGGAGUGUG-3´

3´-UGGGGAUAGUGCUAAUCGUAAUU-5´

257063597 GTTTGGGGAAGTTTAAGTTACTTCCTTAAGGTCACATAACTAGTATCAGAGACTCTAGAA 257063656
257063657 CCTGAAGTCTCACCCACTCATGCAGCACTTAGTTTCCCCTTGTAGTACCGTCAGCTAACA 257063716
257063717 CGGAATGGATTTGAGTAAATACTTATTAAATAAGTCAGCGACTGAATGAGTGATGGGAGC 257063776
257063777 CATTCTCTACACGTGTCTTGAGTGCGTTAACTCATGGAGGCGGGGAATAATGGCTTCTTT 257063836
257063837 TTCTCTTTAGCAGGACTTGGCCCTCTGGAGCATGTACACTAGGCTGTTCAGTCCTGGGCT 257063896
257063897 GTGCTAGCAGCCCCTTCACAGGAAGAGCAATGGCTGCAGCCGCAGCCTCTCACCTGAACC 257063956
257063957 TGGATGCCCTCCGGGAAGTGCTGGAATGCCCCATCTGCATGGAGTCCTTCACAGAGGAGC 257064016
257064017 AGCTGCGGCCCAAGCTCCTGCACTGTGGCCATACCATCTGCCGCCAGTGCCTGGAGAAGC 257064076
257064077 TGCTGGCCAGCAGCATCAACGGCGTGCGCTGTCCCTTTTGCAGTAAGATTACCCGCATAA 257064136
257064137 CCAGCCTGAGCCAGCTGACGGACAACCTGACCGTGCTGAAGATCATTGACACGGCTGGGC 257064196
257064197 TCAGCGAGGCCGTGGGGCTGCTCATGTGCCGGGCCTGCGGGCGGCGGCTGCCCCGGCAGT 257064256
257064257 TCTGCCGGAGCTGCAGCGTGGTGTTATGTGAGCCCTGCCGGGAGGCGGACCACCAGCCTC 257064316
257064317 CCGGCCACTGCACGCTCCCTGTCAAAGAGGCCGCCGAGGAGCGGCGCCGGGACTTTGGAG 257064376
257064377 AGAAGTTGGCCCGTCTACGGGAGCTCATGGGGGAGCTGCAGCGGCGGAAGGTGGCCTTGG 257064436
257064437 AAGGTGTCTCCAAGGACCTGCAGGCACGGTATAAAGCCGTTCTCCAGGAGTACGGGCACG 257064496
257064497 AGGAGCGCCGGGTCCAGGAGGAGCTGGCTCGCTCTCGGAAGTTCTTCACAGGCTCTTTGG 257064556
257064557 CTGAGGTTGAAAAGTCCAACAGTCAAGTGGTGGAGGAGCAGAGTTACCTGCTCAACATCG 257064616
257064617 CAGAGGTGCAGGCTGTGTCCCGCTGTGACTACTTCCTGGCCAAGATCAAGCAGGCGGATG 257064676
257064677 TGGCGCTCCTGGAAGAGACGGCGGAGGAGGAGGAGCCCGAGCTCACCGCCAGCCTGCCCC 257064736
257064737 ACGAGCTGACCCTGCAGGACGTGGAGCTCCTCAAGGTCGGCCACGTGGGCCCCCTCCAAA 257064796
257064797 TCGGGCAGGCGGTTAAGAAGCCCCGGACAGTCAACATGGAAGACTCCTGGGCCATGGAGG 257064856
257064857 CTGCGGCTTCTGCTGCCTCTACTTCTGTGACATTTAGAGAGATGGACATGAGCCCCGAGG 257064916
257064917 AAGTTGCCAGCCCGAGGGCCTCGCCTGCTAAGCAGCGGGGAGCCGAGACAGCCGCCAGCA 257064976
257064977 TCCAGCAGTGCCTCTTTCTCAAGAAGATGGGGGCCAAAGGCAGCACTCCAGGCATGTTCA 257065036
257065037 ACCTTCCGGTCAGCCTCTACGTGACTGGTCAAGGCGAGGTGCTGGTCGCTGACCGAGGCA 257065096
257065097 ACTACCGTATACAAGTCTTTACCCGCAAAGGCTTCTTGAAGGAGATCCGCCGCAGCCCCA 257065156
257065157 GCGGCATCGATAGCTTTGTGCTGAGCTTCCTTGGGGCTGATCTGCCCAATCTCACTCCTC 257065216
257065217 TCTCAGTGGCCATGAACTGCCACGGGCTGATCGGTGTGACTGACAGCTATGACAACTCCC 257065276
257065277 TCAAGGTGTACACCTTGGATGGCCACTGCGTGGCCTGTCACAGGAGCCAGTTGAGCAAGC 257065336
257065337 CCTGGGGCATCACAGCCCTTCCATCCGGCCAGTTCGTGGTAACTGATGTGGAGGGTGGAA 257065396
257065397 AGCTCTGGTGCTTCACTGTTGACCGAGGAGCAGGGGTGGTCAAGTACAGCTGCCTGTGCA 257065456
257065457 GTGCCGTTCGGCCCAAGTTTGTGACCTGCGACACCGAGGGCACCGTCTACTTCACCCAGG 257065516
257065517 GCTTGGGCCTTAATCTGGAGAATCGGCAGAATGAGCACCACCTGGAGGGCGGCTTCTCCA 257065576
257065577 TTGGCTCTGTGGGCCCCGACGGGCAGCTGGGTCGCCAGATCAGCCACTTCTTCTCAGAGA 257065636
257065637 ACGAGGATTTCCGCTGTATCGCCGGCATGTGTGTGGATGCGCGTGGGGATCTCATCGTGG 257065696
257065697 CTGACAGCAGTCGCAAGGAGATTCTCCACTTCCCCAAGGGGGGCGGCTATAGTGTCCTTA 257065756
257065757 TCCGAGAGGGGCTCACCTGTCCAGTGGGCATCGCCCTCACTCCTAAGGGGCAGCTGCTGG 257065816
257065817 TCTTGGACTGTTGGGATCATTGCATCAAGATTTACAGCTACCATCTGAGAAGATATTCTA 257065876
257065877 CCCCTTAGGGGGCTGAGTGATAGCAGTTTCTTGTGCTCCTCAGCCATCTTCCCTTCCCTC 257065936
257065937 ATTCCTTGGTGGTTGGTGGTACTGTAGCGTAGACTTGGCCTGTGTCCTGATTCCAGCTGG 257065996
257065997 CTAGTCCTAGAAGTTTAGAAGCTCTGTCTCTTTTAGTGTGTTTGTTTGTTTTAATTTGCA 257066056
257066057 TTATCTCCTCCCTAAGTTTAGAGCTTTAACAGATGCACTGCCCTAAATAGGACACGATAT 257066116
257066117 GAGGTTAGCCGATGTTTGAGTAGAACCTAGACACTTCCAGGGCTTCAGTAGCCTTTTGCC 257066176
257066177 CCTGTATTTGAAATTTGCCCTTGTATTGAATCCTTGTTGATTTTTCTCTCTTTTGGCTTT 257066236
257066237 GACCACTCTGGTCCATTATTTCCTCCCTTATTGATATACACTCCATCTGTGACTTTCTCT 257066296
257066297 TCAGCTCTGCTGCATTTAGCATGCATGCTCCAGTGGGACCTGCTGCACCTTTCAGTGACA 257066356
257066357 CTTCAGACATCCTATCCCTGTGGCATGATGTCTCGGGTCAGAGCACCTTTCACTTAGTGC 257066416
257066417 CACCAAGGGAGACATACAGTCCCTCTGGGAAGCAGGGCCTCTCAGCTCTGGGGGGGTCTG 257066476
257066477 TGCTATCTTGGATTGGTCTCCAGTTGTGGCAGTGACAGAGGATGGATTGACCCTTGTTGG 257066536
257066537 CTGGTATTGAAGGCCTTAGCCTGGAAATTGCTAGCCTTTCAAAGGAGCATATGGATTGGA 257066596
257066597 ATTAGGCCAAAGCATAGGAAAATGGGAATAAACAGAAATAAATGCTTTTATAGTGAAAAG 257066656
257066657 ACACTGTTGGGTAGTCTCTAGAGCTCATCGTTTTTCATACTACCTCTCTCCACTGACCTG 257066716
257066717 TGCTAAGAGCTGTCTGTGGAATTGTACAACACAGGCTGGGGCCAACACAGTGCAGAGGGG 257066776
257066777 GACACATCTCATTTTCATCATTAAAACTGTGGGGTCCTTATTCTGTATTTTCCTCAAGTG 257066836
257066837 AGTGCATCATATTTGGTTTGGGGATTTATTTCCATTCCTCTATCAAGCATTAAATGATTG 257066896
257066897 ATAACTGTTTCTTCATCACTGGTGTTCGTGTCTTTTCATTAAATATGGGATTTGGGAGGT 257066956
257066957 GGGGGTGGTATCTGCTGCTTTGTTTCCACACCTGACACTCCATCCTTTGAGGATGGTAAG 257067016
257067017 AATCTACCTAGGTCACCATTCAGCCAAAAGTACCATTATGTCAAGTTAATCCAGATCCTA 257067076
257067077 CTCCCACATAACCTGTGTATGGAATCAGCCTCCAAGGTCTGCAGTCACCAGCACCCTGTT 257067136
257067137 TTAAACTCAGCCTCCACCTGCTTTA

>ssc-miR-150-5p MIMAT0025365 UCUCCCAACCCUUGUACCAGUG

ZEB1 Zinc finger E-box binding homeobox 1; chr10:42213602-42216139

5´-UUAUGUUGGUUUAUUCUUAAGCUGUACAAUUGGGAGAAAUUUUAUAAUUUUUUAUUGGUAAACAUAG-3´

3´-GUGACCAUGUUCCCAACCCUCU-5´

```
      CCCAAAAAGAAGATGCGAAAGACAGAAAACGGAATGTATGCCTGCGATTTGTGTGACAAG 42213411
42213412 ATTTTCCAAAAGAGTAGCTCACTGTTGAGACATAAATATGAACACACAGGTATGTTAACA 42213471
42213472 AATCTGGCAGTATCTTAAAAATGAATCTCTAACCACTTCCATAAACCTGAAGTGCCCCAA 42213531
42213532 AACCTGATGAGGATTGAATTGGTTGGATGCCATTTCATTTCACAGCTTTTACCTGTGCTC 42213591
42213592 TTATTTTCAGGTAAAAGACCTCATGAGTGTGGCATCTGCAAAAAGGCCTTCAAACACAAG 42213651
42213652 CATCACTTGATCGAACACATGCGCTTGCATTCTGGAGAGAAGCCCTACCAATGTGACAAG 42213711
42213712 TGCGGGAAGCGCTTCTCGCACTCGGGCTCCTACTCCCAGCACATGAACCACCGCTACTCC 42213771
42213772 TACTGCAAGCGGGAGGCAGAGGAGCGGGACGCAGGCCCCGAGGACGCAGGCCCGGGGGCC 42213831
42213832 TGGCCGAGCACGAGCACGCGGGGGCCCGCGCATCCCCATCACAGGGCGACTCGGACGAGA 42213891
42213892 GGGAGAGCTCGACGCGGGAAGAGGATGAAGACAGCGACAAGGAGGAAGAGGAGGAGGAGA 42213951
42213952 GAGAGATGGAGGAACTGCAGGAGGAAAAGGGAAGTGGAGAACCCGGAGGAGAGGAGGAGG 42214011
42214012 AGGAGGAGGAAGAGGAAGAGGAGATGGAGGAGGAGGAGGAGCAGGAGACGGAGGAGCAGG 42214071
42214072 AGGAGGGAGGCAGAGCCAAGACCGCAGGTGCGACAGCGGAGAGGGAAGGCGCTGGAAGTC 42214131
42214132 CAGAGAGCAGCCTAGAACACAAAGTGAGCGAGGGTAGTGAGCAGGTGTCGGAAGAAAAGA 42214191
42214192 CAAATGAAGCCTAATAGTTTTTCTAGAAAGAAAAAAATTCTAATTGGTAATGAAGTTTGT 42214251
42214252 TCTATATTATACATGCTTTTCAAGGAAACGCAGTAACCTGTATACTGTGATTTCTGTTCA 42214311
42214312 CTACTGTGTAAAGTAAAAATTAAAAAAAATACAAAATACAAAAAAAAAAAAAAAAAAAAC 42214371
42214372 CACACAAAAAAAAAATCCGGGTGTGCCTGAACCTCAGACCTAGTAATTTTTCATGCATTT 42214431
42214432 TTCAAGTTAGGAACAAGTTTGTAACATGAAGCAGATTAGAAAACTTAATGACTCCGAAAG 42214491
42214492 CAAAGATTTAACAGGTGAAAGGAAACCAATATAATTAGATAAGCATCTGGTGTCGTTTCA 42214551
42214552 CTTTATCAGTATTAATTATCACTCTTATGTTGGTTTATTCTTAAGCTGTACAATTGGGAG 42214611
42214612 AAATTTTATAATTTTTTATTGGTAAACATAGGCTAAATCCGCTTCAGTATTTTATTATGT 42214671
42214672 TTTTTAAAATGTGAGAACTTCTGCACTACAAAATTCCCTTCACAGAGAACTATAATGCAG 42214731
42214732 TTCCAAACCGTGCTAACTACCTTTTATAAATTCAACCTAGAAGGTAGTCATTTCTCATAT 42214791
42214792 TTAGATGTCATAGTAGAGTGTGTTCTCATTGAAAGTGTATTGTTAGCCTTAAGAAAGCAG 42214851
42214852 CCGAGAGCAGAACTGAAGTTTCCCTCATGTGGTTTCAAGTGGGGTTGAAGAGACTGTCGT 42214911
42214912 TGAGTTCTGAGTGCAGGGACGAACAGTGTTAATACTGGGGAGGGCAGCAGAGTCGTCAGA 42214971
42214972 ATCAGTGTTTGTGATTGTGTCTGAGTATGTGGTAAACAAATATGAAGGATGTGATAGGAA 42215031
42215032 AGCTTTGTCTCTCCTTTGGCCTTAAGCAAGACCCGTGTGCTGTAAGTGCCATTTATCAGT 42215091
42215092 ATTTTCAAGGCTCTGACCAGCCGTCGCTAACGTGTGGCCTACAATAACTAGCATTTGTTG 42215151
42215152 ATTTGTTTGTTGTACCGAAATTCAAAAATAAAATGTAAAACCACTGACTCTGTCAGAGAA 42215211
42215212 GCCAAACACAGGGACATTTTCATCCTTTAATTCCTCAAGATTCATTTTGCGTTAATTGAC 42215271
42215272 TTCCAGAATTTCTCTACCGAAATGAAAGGGAAATTTTCTAATCTGCTTTATCCATGTACT 42215331
42215332 TGCATTTCAGACATGGACATGCCATTGTTATTTGGCTCATAACTGTTTCCAAATGTTAGT 42215391
42215392 TATTATGGACCCAGTTTATTAACAACATTAGCTGATTTTTACCTATCAGTATTATTTTAT 42215451
42215452 TTCTTTTAGTTTATAGATCTGTGCAACATTTTTGTACTGTATGTCTTCAAACCTGGCAGT 42215511
42215512 ATTAATACCCTTCTTACTGACATATGTACTTTTAGTTTTAGAAAACTTTTATATTTATGT 42215571
42215572 GTCTTATTTTTATATTTCTTTATTTATTACACAGTGTAGTGTATAATACTGTAGTTTGTA 42215631
42215632 TTAATACAATAATATATTTTAGTATGAAAATTTGGAAAGTTGATAAGATTTAAAGTAGAG 42215691
42215692 ATGCAATTGGTTTTCTTGCATTGAGATTTGATTTAACAGTGTTATGTTAACATTTATACT 42215751
42215752 TGCCTTGGACTGTAGAACAGAATTTAAATGGGAATGTATTAGTTTTACAACTACAATCAA 42215811
42215812 GTCATTTTACCTTTACCCAGTTTTTAATATAAAACTCTTAAATTTTGAAATTCACTGTGT 42215871
42215872 GACTAATAGCATGATGCTCTGTGGTTTTATTAAAAATTAGCCTAACCATAAAACTCTCAT 42215931
42215932 TTCCTTATTAGCAAGCCAAATTAGGATAACCTTATATAAACAGTGTTGGGAACAACGTTT 42215991
42215992 TAACATTTTGTGCCAATTTGTTCCTGTATTCATGTATGTAAGTTACAAATCTGACTCTTC 42216051
42216052 ATTTTTAAGTTCCTTGTTACACCATGGTCATTTTCTAGTTTTTTACCAGACTCCCCCATC 42216111
42216112 TCACAATAAAATGCATCAACAAGCCTGACTTGCCGTCATTCTTTTCATCGTTATCAAGAC 42216171
42216172 TTTTTTTGGAAAACTCTGTGAAATGGGGTGTTTACACCATCATCCTGCACTTGACTTATC 42216231
42216232 TTGGCCTGAGTTTGGGTAACTAATTATTTTAGACATTCACCACTCAATTATGGATTTTCT 42216291
42216292 TGTGGCCACGGCCACATTCTCTGGGAAGGTCAGAGTTGACAACCAAAGGCCTCCTCAGAG 42216351
42216352 CCTACTTCCTTTGCCTGCCTCCCTCCCCCCTCCACTCT
```

>ssc-miR-21-5p MIMAT0002165 UAGCUUAUCAGACUGAUGUUGA

IL12A (p35) interleukin 12A; chr13:99733394-99741078

5′-
CCCUUUAAUAGAAGAGCAAGAAUUUAUAAGCUAUUUCAGUACCAAAGUGUUUGUAGAAACAA
AC-3′
3′-AGUUGUAGUCAGACUAUUCGAU-5′

```
>chromosome:Sscrofa11.1:13:99739875:99741578:1
99739875 AGGTCTTACATTGTGTTATCAAATTAAAGGGTTCAAACTGGGAGACCCATTGCCAGCTTT 99739934
99739935 ATTTATCTCACACGAATGGGTGGTTTTAATCAATAGGTCTGATACTGACGGAGAGTTTCC 99739994
99739995 GATCAATAGCAATCCATCACCTCCCTGTTCTAAAAAGGGCACTGCTTCTTAATGTAACAT 99740054
99740055 ATAAAAATTAATTAGATCCCACAGCAGATGTCATCTTACCAGAAATGTTCCTATCTAAAA 99740114
99740115 TGCAACTCCCACTGCTGATTTTCTAAATAGCTTTCAGGGGCTTTTCAGAGCTCACTGAAG 99740174
99740175 ATCCATGTGCTTGGATAATAAGTCTCCTCTCAGGTTCTGCCTGGCCATCTGGCTGGAAGT 99740234
99740235 AGCTTTGCATCAGGTGTAGGAGATGGAATGAGTGGGTTGTAGGCTCTGAGACATGCATTG 99740294
99740295 GCTTCACTCGTTTTTACAGATGAACGTTTTAATTTTGGAATATCAGCCATGTCAAGCTCA 99740354
99740355 ACCAATATATATTCCCCTAGGCTCTGAATTTCAACAGTGAGACTGTGCCACAGAAGCCCT 99740414
99740415 CCCTGGAAGAACTGGATTTTTATAAGACTAAAATCAAGCTCTGCATACTTCTTCATGCCT 99740474
99740475 TCAGAATTCGTGCGGTGACCATCGACAGAATGATGAGCTATCTGAATTCTTCCTAAAAAG 99740534
99740535 CTGAAGTCTCTCCTAACCTTAAAGTCATTTTTATAAAAATTGGAACCAAAGAAATGTTAA 99740594
99740595 TAGGGTATGGGTTAGGAACTGGGGCGGGGGTGGCTTGACTGGTCCTATCTAAGCTGGTAT 99740654
99740655 AGGAATTCCCATGCTTGTTTACGTTAGTCATCACCCCAAATTTGAAAGATGTGAGGCTCC 99740714
99740715 CATCCACACAATTGATACTCTGATCAAGTGTATTTCACAGTGGATAAGTTATACTACGGA 99740774
99740775 TAAGTTATTTTTTTAAGTTTTCATGAGCAAACTGCTAAGGAGGGAAAATGTCCTCACAGA 99740834
99740835 ACATGTTGGGTTTTTTTTTCCCTTTAATAGAAGAGCAAGAATTTATAAGCTATTTCAGTA 99740894
99740895 CCAAAGTGTTTGTAGAAACAAACACTTGAGCCTAATTTATTTTTTAAGTATTTATTTATA 99740954
99740955 TAATTTTGTACTCATGAAAGCATATGAACTAACTTATATTTATTTATGTTATATTTATTA 99741014
99741015 AAATATTTATTCCCAAGTGGATTTGAAATATACCTTGTTATTCTAAAAATAAAATGACTG 99741074
99741075 AATTAAAGTAATGCTCATATTTTTCTAATGTGACTGAAATACTCTGAGTGTTAAAATACA 99741134
99741135 TATCGTGGGAAAACTGAAAACTGATGAAATTTGAACAAAACCACCTTGGAAAAAAAGCCA 99741194
99741195 CAATATTCTTACTCATTTTCTAGGTCCTTTAATCTCATATTTTAAAGAGAGCCACAAGAC 99741254
99741255 TAGATCTGCATTAATACTGAAACTGTTGAATGACTTAGTTTTTGAGAATATATTCTGTGA 99741314
99741315 TTTTTGTTTAACATCTGAATCTAAAAGTAAAATACATAAAAATGTTTTTCCCTATTAACC 99741374
99741375 ACAGCCCCATAACATACTGGGCATTAGACAGATACATAATAACATCCTCCAACTGGGAAA 99741434
99741435 AACATCGTTAAAAGGGCTTCTAGTTTCATTATTCTAATTTTTGTGACCCTTTAAGCTCTC 99741494
99741495 AAGGGGAAGTTGTTTCTAAATTACAACAAGAAAAACCTCTCCATTTGCTTTAAATTATAT 99741554
99741555 AGAACTTCTGCCTGCACCACATGG                                     99741578
```

Appendix F

In this section, the ratio between Renilla Luciferase and Firefly Luciferase can be observed. The ratio of each candidate target sequence, and their control trials with psiCHECK2, is displayed in individual tables in the following.

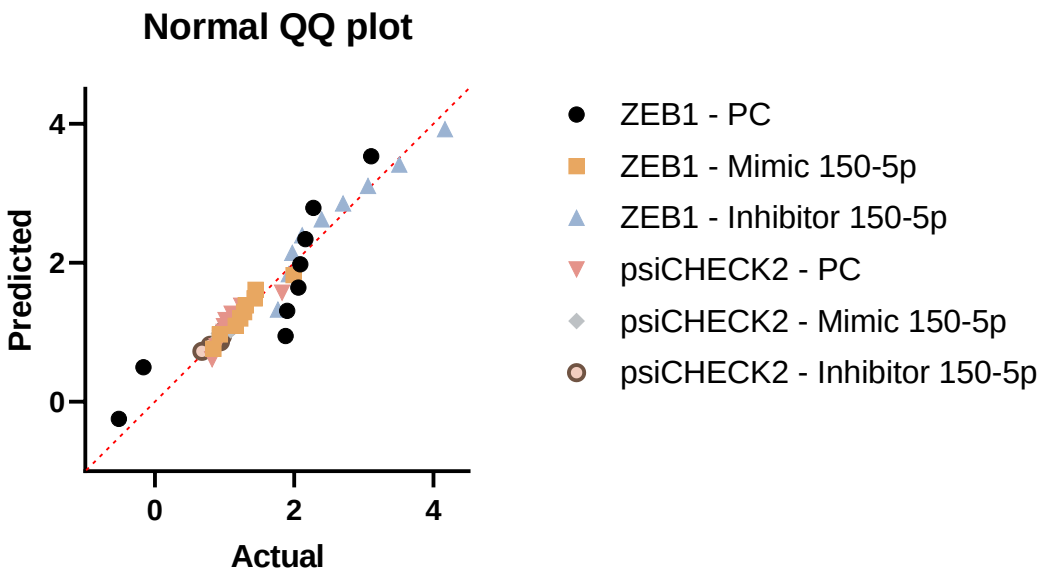
Renilla/Firefly Ratio	**Plate 1**	**Plate 1**	**Plate 1**	**Plate 2**	**Plate 2**	**Plate 2**	**Plate 3**	**Plate 3**	**Plate 3**
SOCS1 – PC	1.16	1.28	1.39	1.06	1.49	1.85	1.21	1.34	1.48
SOCS1 – Mimic HSA-miR-155-5p	0.53	0.70	0.79	0.64	0.71	0.83	0.69	0.85	0.85
SOCS1 – Inhibitor HSA-miR-155-5p	1.28	1.21	1.49	1.17	1.17	1.33	1.13	1.16	1.36
psiCHECK2 - PC	1.11	1.32	1.52	0.90	0.92	1.10	1.01	1.02	1.15
psiCHECK2 – Mimic HSA-miR-155-5p	0.85	0.96	0.81	0.73	0.90	0.90	0.88	0.91	0.91
psiCHECK2 – Inhibitor HSA-miR-155-5p	0.90	1.11	0.88	0.79	0.93	0.79	0.98	0.95	0.86

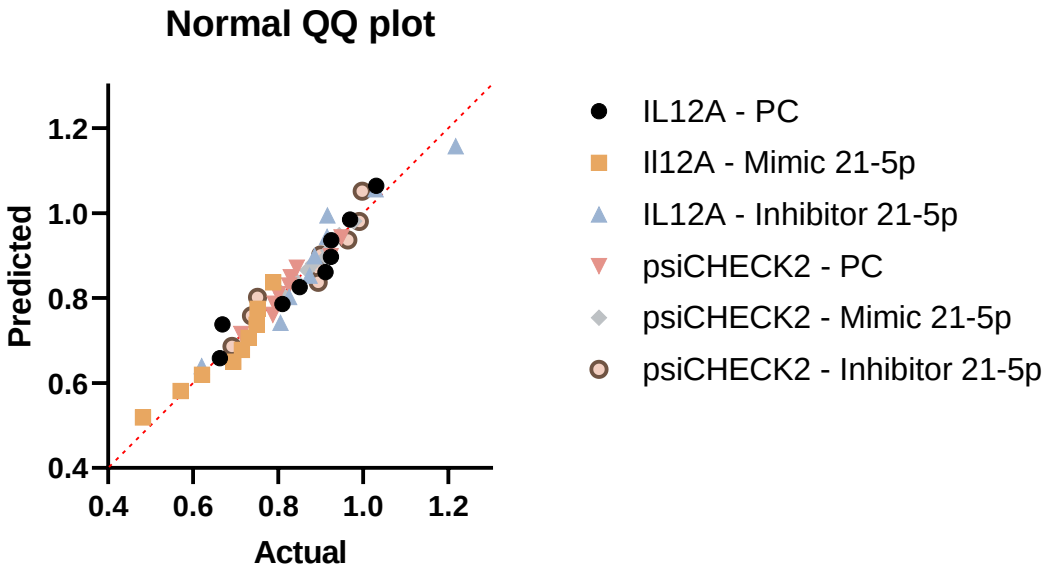
Renilla/Firefly Ratio	**Plate 1**	**Plate 1**	**Plate 1**	**Plate 2**	**Plate 2**	**Plate 2**	**Plate 3**	**Plate 3**	**Plate 3**
IL12A – PC	0.91	0.92	0.92	0.66	0.67	0.81	0.85	0.97	1.03
IL12A – Mimic HSA-miR-21-5p	0.79	0.75	0.69	0.48	0.57	0.62	0.71	0.73	0.75
IL12A – Inhibitor HSA-miR-21-5p	0.89	1.22	1.03	0.62	0.81	0.83	0.87	0.92	0.91
psiCHECK2 - PC	0.71	0.80	0.92	0.83	0.79	0.95	0.84	0.79	0.83
psiCHECK2 – Mimic HSA-miR-21-5p	0.81	0.94	0.92	0.84	0.91	0.92	0.88	0.87	0.98
psiCHECK2 – Inhibitor HSA-miR-21-5p	0.96	0.99	0.90	0.89	1.00	0.74	0.69	0.75	0.90

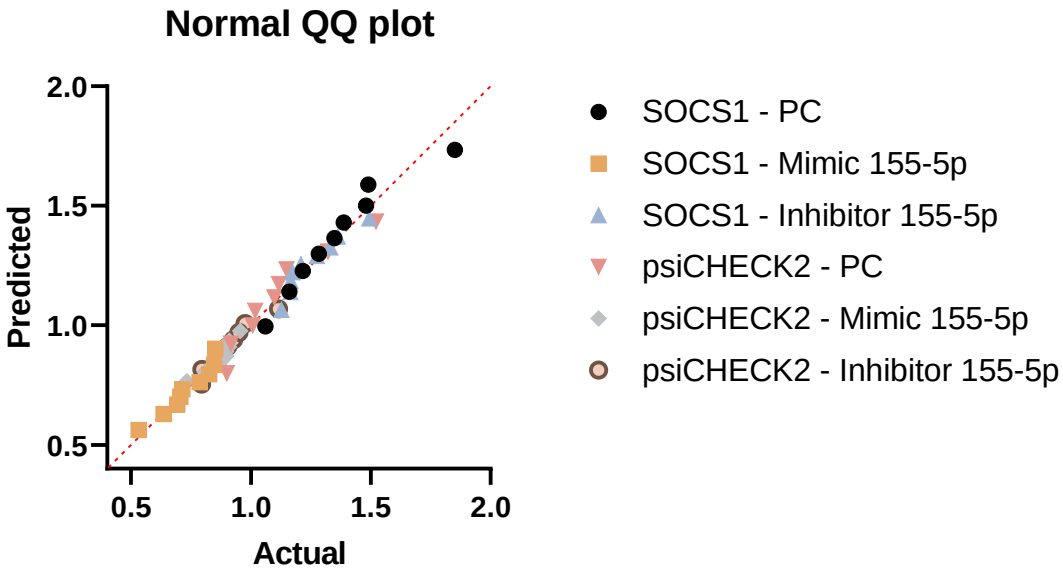
Renilla/Firefly Ratio	**Plate 1**	**Plate 1**	**Plate 1**	**Plate 2**	**Plate 2**	**Plate 2**	**Plate 3**	**Plate 3**	**Plate 3**
ZEB1 – PC	1.88	2.09	2.06	-0.16	3.11	1.90	-0.52	2.28	2.16
ZEB1 – Mimic HSA-miR-150-5p	0.93	1.28	1.45	1.30	1.99	1.43	0.84	1.16	1.23
ZEB1 – Inhibitor HSA-miR-150-5p	3.5	1.9	1.77	1.97	4.16	2.39	2.11	3.06	2.70
psiCHECK2 - PC	0.88	1.24	0.82	0.99	1.82	0.99	0.94	1.10	1.02
psiCHECK2 – Mimic HSA-miR-150-5p	1.13	1.16	1.16	1.09	0.90	1.27	0.92	1.14	1.10
psiCHECK2 – Inhibitor HSA-miR-150-5p	0.94	0.98	1.06	0.95	0.97	0.79	0.68	1.13	0.97

Appendix G

Scatterplots displaying the normal distribution of each individual candidate target sequence and their correlated control trials.






Normal QQ plot
Predicted
Actual
2.0
1.5
1.0
0.5
0.5
1.0
1.5
2.0
SOCS1 - PC
SOCS1 - Mimic 155-5p
SOCS1 - Inhibitor 155-5p
psiCHECK2 - PC
psiCHECK2 - Mimic 155-5p
psiCHECK2 - Inhibitor 155-5p