Specificity and mechanism of RNA trafficking from mouse to bacteria in the gut

Xiaochen Du

Thesis submitted for the degree of Doctor of Philosophy

The University of Edinburgh

2023

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where states otherwise by reference or acknowledgment, the work presented is entirely my own.

Parts of Chapter 1 are previously published as part of a review in Microlife which was co-authored with Ruth Ley and Amy Buck (Du et al., 2021).

Xiaochen Du

Date: 30/12/23

Contributions

Ms. Elaine Robertson and the author were responsible for collecting mice serum. Villin-cre transgenic mice (B6.Cg-Tg(Vil1-cre)997Gum/J) were provided by Dr Tom Youdale. All mice used in this study were bred within the in-house facilities at the University of Edinburgh.

The bacterial 16S rDNA libraries in Chapter 3 were prepared by Dr Rowan Bancroft and subsequently sequenced by Edinburgh Genomics at the University of Edinburgh.

The bioinformatic analysis of the sRNA sequencing data in Chapters 3 and 5, as well as the CLEAR-CLIP data, was performed by Dr Jose Roberto Bermudez-Barrientos and the author, with valuable input from Dr Cei Abreu-Goodger and Dr Sujai Kumar.

In Chapter 5, the methodology for measuring EVs using Zetaview and MemGlow™ 640 labelling was guided by Dr Ruby White. TEM imaging of EVs was carried out by Mr Stephen Mitchell. Dr Prerna Vohra provided the bacterial strains *Salmonella* typhimurium SL1344 and *E. coli* W3110, and contributed invaluable comments and suggestions regarding the bacterial experimental design and data analysis.

Acknowledgements

Starting my PhD journey in 2019 during the global COVID-19 pandemic brought unexpected challenges that significantly impacted my study and personal life. The support I received from my supervisor, colleagues, family, and partner was invaluable in helping me navigate through these tumultuous times!

First and foremost, I would like to express my deepest gratitude to my supervisor, Professor Amy Buck. Thank you for providing me with the opportunity to pursue my doctoral studies at the University of Edinburgh and for entrusting me to explore the gut microbiota research. Your initiative in arranging for me to conduct research at the renowned Ley lab in Germany, known for its ground-breaking work in gut microbiota research, has been invaluable. Regrettably, the onset of the COVID-19 pandemic significantly disrupted the trajectory of my project. Initially, I found myself disheartened by these unforeseen challenges. However, your unwavering support, continuous brainstorming sessions, and exploration of new project avenues alongside me were crucial in guiding me through this difficult period.

A sincere thank you to Dr Prerna Vohra, our esteemed collaborator, for your invaluable assistance. Your provision of bacterial strains, lab space, and techniques are crucial to my research. Your insightful comments and guidance on experimental design and data analysis have been immensely valuable.

Thank you to my thesis committee members, Dr Joanne Thompson, and Prof Amy Pedersen, for their support and feedback throughout my doctoral journey. Your guidance at every stage has been pivotal in shaping the direction of my research in meaningful ways.

Thank you to the members of the Buck lab, both past and present. Each one of you has contributed significantly to making my time in the lab an enriching experience. I would like to extend special thanks to Kat and Kyriaki for their invaluable guidance in Ago2 immunoprecipitation and western Blot techniques. Sarah, thank you for your guidance on CLEAR-CLIP. Ruby, your guidance on Zetaview and EVOS usage has been invaluable. Tom, your guidance on gut sectioning and immunofluorescence staining have been immensely helpful. Elaine, I am grateful for your assistance with

mouse sample collection. Beto, thank you for your invaluable contributions to the bioinformatic analyses.

I am also deeply thankful to Dr Cei, Dr Rowan, and Dr Yu Huo for their contributions to my project, each in their unique way.

Finally, many heartfelt thanks to my family in China! Despite being thousands of miles apart, your boundless understanding and unwavering support have meant everything to me. Special thanks to my wife, Yuwei, for accompanying me during my studies in Edinburgh, and sharing both my moments of joy and hardship!

感谢你们，助我成长!

Lay Summary

In recent years, the relationship between human health and the gut microbiota has gained significant recognition. The term “gut microbiota” refers to the diverse community of microorganisms inhabiting the host’s gastrointestinal tract. A balanced and healthy gut microbiota plays a pivotal role in maintaining the overall health of the host. An imbalance in gut microbiota has been linked to a wide spectrum of human diseases. Several approaches, such as probiotics, prebiotics, and faecal microbiota transplantation, are being actively explored as potential interventions to restore or maintain a healthy gut microbiota and address associated diseases. However, strategies are still lacking to precisely manipulate the gut microbiota. One emerging strategy is the use of microRNAs (miRNAs).

miRNAs are small non-coding RNAs, with an average length of 22 nucleotides that function by binding to messenger RNAs (mRNAs) within cells and regulating the amount of protein that gets produced from these mRNAs. miRNAs have also been found to be released outside of the cells and detected in a cell-free form in serum, saliva, urine, and faeces. Recent studies suggest that mammalian miRNAs can move from epithelial cells directly into gut bacteria and they can modulate the composition and activity of the microbiota by directly regulating bacterial genes. However, many of these studies are done in a test tube and questions persist about whether and how miRNAs enter bacteria under natural conditions. There is also little understanding of the specificity: which bacteria can internalize and respond to intestinal miRNAs and which cannot? In this thesis, we aim to study the specificity and mechanism of RNA trafficking from mouse to bacteria in the gut.

Firstly, we developed a purification method to isolate pure gut microbiota from mouse gut contents in order to sequence the miRNAs that were transferred to the microbiota naturally. Our results confirm that host miRNAs are naturally found inside bacterial cells. In particularly, we find that miR-21a-5p is the most abundant miRNA, and our data suggest it may be specifically internalized by *Lactobacillus.*

Our next aim was to build a model for how the miRNAs can reach bacteria. In mammals, miRNAs usually function with an Argonaute protein and it has been observed that miRNA-Ago2 (Argonaute 2) complexes are secreted extracellularly by mouse intestinal epithelial cells. We hypothesize that miRNAs move with the Ago2

protein and that the Ago2 protein is involved in coordinating the binding between miRNAs and the bacterial RNA targets. To test this hypothesis, we first investigated whether mouse Ago2 protein is detected inside gut microbiota using western blot analysis and immunofluorescence staining. Our results were inconclusive because the Ago2 signal was low and not consistent. We then employed the method called CLEAR-CLIP (covalent ligation of endogenous Argonaute-bound RNAs–Crosslinking and immunoprecipitation) to capture Ago2 directly from gut tissue in order to sequence bacterial genes that could be bound by the miRNA-Ago2 complexes. We identified high-confidence bacterial targets in the gut tissues. However, further bioinformatic analysis and experimental validation are necessary to confirm these.

We then examined extracellular vesicles (EVs) as another potential carrier that could transport miRNAs into bacteria. EVs are lipid-bound vesicles released by cells into the extracellular space that play various roles in communication between organisms. We isolated EVs from an intestinal epithelial cell line (Mode-K cells) and profiled the miRNA content of Mode-K EVs. Our results show that miRNAs from the let-7 family are abundant and enriched in Mode-K EVs. We then tested whether Mode-K EVs are internalized by bacteria, and probed specificity by comparing the uptake properties of the mouse EVs to the uptake of the parasite (*Heligmosomoides bakeri*) EVs. We found Mode-K EVs are taken up by *Salmonella enterica serovar* Typhimurium (*S.* Typhimurium) but not *E. coli*. Strikingly parasite EVs are not taken up by either bacterial species, suggesting specificity. Furthermore, we show that Mode-K EVs can promote the growth of *S.* Typhimurium and we detected mouse miRNAs in *S.* Typhimurium after co-culturing with Mode-K EVs. Our data suggest that Mode-K EVs have the capability to deliver RNA cargo to specific bacteria and affect bacterial growth. Whether this promotion effect of Mode-K EVs on *S.* Typhimurium is induced by mouse miRNAs requires further investigation.

Finally, we used a genetic reporter system to investigate whether EVs have the capability to transport functional cargo into bacteria. This technique involves the donor cells secreting EVs containing Cre mRNA, which are then internalized by reporter cells. The activity of Cre within the reporter cells leads to the production of fluorescent molecules, serving as an indicator of functional transmission. Our co-culture experiments of mouse Cre-expressing EVs with reporter *S.* Typhimurium demonstrated the functional transmission of cargo from mouse to bacteria via EVs.

In summary, this thesis demonstrates that host miRNAs are present within gut microbiota naturally and EVs derived from intestinal epithelial cells can act as a transport mechanism for host miRNAs into specific bacteria. These findings contribute to a deeper comprehension of the specificity and mechanism of RNA trafficking from mouse to bacteria within the gut environment which could hold promising implications for modulating the gut microbiota to treat associated diseases in the future.

Abstract

The human gastrointestinal tract hosts a complex and diverse population of microorganisms collectively known as the gut microbiota. The gut microbiota plays pivotal roles in maintaining human health. Research links an imbalance in gut microbiota to several human diseases like inflammatory bowel diseases, obesity, diabetes mellitus, and gastrointestinal cancers. Consequently, modulating the gut microbiota is increasingly recognized as vital for treating these conditions, and several approaches, including probiotics, prebiotics, and faecal microbiota transplantation, are being explored for this purpose. However, strategies are lacking for precisely manipulating the composition of the gut microbiota. One emerging strategy is the use of microRNAs (miRNAs) which may be involved in the specific regulation of bacterial genes.

miRNAs are a class of small, single-stranded RNA, typically 18–25 nucleotides in length. These small RNAs primarily regulate gene expression by either facilitating the degradation of messenger RNA (mRNA) or repressing mRNA translation in mammals. Over the last 15 years, extensive literature has demonstrated the presence of miRNAs in a cell-free form including serum, saliva, urine, and faeces. Recent reports suggest that mammalian miRNAs can directly influence the composition and activity of gut bacteria by entering bacteria and interacting with bacterial genes. However, many questions remain regarding how miRNA-bacteria interactions occur under physiological conditions. This thesis aims to investigate the specificity and mechanisms underlying small RNA trafficking from mouse intestinal cells to bacteria within the gut environment.

To explore the miRNA-bacteria interactions, a crucial first step is to identify the miRNAs that are naturally transferred to the gut microbiota. We developed a purification method to isolate pure gut microbiota from mouse gut contents and sequenced small RNAs in the purified gut microbiota (PGM). Our data suggest host miRNAs are naturally present within gut microbiota and show that the miRNA composition within PGM is distinct to that found in total gut contents, implying specificity. We then employed miRNA-FISH (Fluorescence In Situ Hybridization) to visualize the presence of host miRNAs within gut bacterial cells and to detect miR-21a-5p, which is the most abundant miRNA in PGM. Moreover, co-staining

experiments conducted in PGM using probes for miR-21a-5p and *Lactobacillus* demonstrated co-localization, suggesting the uptake of miR-21a-5p by *Lactobacillus*. Our work provides a promising foundation on which to discover additional interactions between miRNA and bacteria within the gut.

Next, we investigated the transport mechanism of host miRNAs into bacteria. In mammals, miRNAs usually function with an Argonaute protein and it has been observed that miRNA-Ago2 (Argonaute 2) complexes are secreted by mouse intestinal epithelial cells. We hypothesized that miRNAs move with Ago2 protein and the miRNA-Ago2 complexes regulate gene expression within the gut microbiota. To test this hypothesis, we investigated whether the mouse Ago2 protein is detectable within gut microbiota and adapted a method to identify bacterial genes directly targeted by miRNA-Ago2 complexes. The detection of Ago2 was performed in PGM using western blot analysis and immunofluorescence imaging and our results suggest that there is no consistent Ago2 signal in PGM. However, PGM does not encompass the entire spectrum of bacterial species within the mouse gut. To account for possibility that Ago2 signal could be lost in the PGM purification, we employed another method to capture Ago2 and isolate associated RNAs from the intact gut tissue: CLEAR-CLIP (covalent ligation of endogenous Argonaute-bound RNAs–Crosslinking and immunoprecipitation). We identified high-confidence bacterial targets co-purifying with Ago2 in samples obtained from different segments of the gut, suggesting the transfer of host miRNAs with Ago2 protein and their potential functional role in gut bacteria. Nevertheless, further bioinformatic analysis and experimental validation are required to confirm these as genuine targets.

In order to account for other transport methods beyond Ago2, we then investigated extracellular vesicles (EVs). These lipid-bound vesicles are secreted by cells into the extracellular space and play important roles in both intercellular and inter-organismal communication. We isolated EVs from a mouse intestinal epithelial cell line (Mode-K) and profiled the miRNAs content of Mode-K EVs. Our results show that miRNAs from the let-7 family are abundant and enriched in Mode-K EVs. To explore potential specificity in the uptake of EVs by bacteria, we conducted a comparison using fluorescently labelled EVs obtained from Mode-K cells and from the gastrointestinal parasite *Heligmosomoides bakeri*. We show that *Salmonella* Typhimurium SL1344 internalize Mode-K EVs but not parasite EVs when cultured in M9 medium and the

uptake of mouse or parasite EVs is not observed in *E. coli* W3110. These data demonstrate specificity in the interaction between mammalian EVs and SL1344. Additionally, our data indicated that Mode-K EVs exerted a stimulatory effect on the growth of SL1344 and this growth-promoting impact is dependent on the dosage of Mode-K EVs, as higher concentrations correlated with increased growth. Co-culture experiments using Triton X-100-treated Mode-K EVs confirmed the necessity of intact EVs for the observed growth promotion in SL1344. We also show that mouse miRNAs are detected in *S.* Typhimurium after co-culturing with Mode-K EVs. In conclusion, Mode-K EVs have the capability to deliver RNA cargo to bacteria and affect bacterial growth. However, whether this promotion effect of Mode-K EVs on SL1344 is induced by mouse miRNAs requires further investigation.

Finally, we used a genetic reporter system (Cre-loxP system) to investigate whether Mode-K EVs can deliver functional cargo into *S.* Typhimurium. This technique involves the donor cells secreting EVs containing Cre mRNA, which are then internalized by the reporter cells. If the Cre mRNA is translated after internalization, Cre recombinase can remove the loxP-flanked transcription terminator, leading to the production of fluorescent molecules in the reporter cells. Our co-culture experiments of mouse Cre-expressing EVs with reporter *S.* Typhimurium demonstrated the functional transmission of cargo from mouse to bacteria via EVs.

In summary, this thesis demonstrates that host miRNAs are present within gut microbiota under physiological conditions and provides an approach to identify interactions between miRNA and bacteria within the gut. We also demonstrate that EVs derived from intestinal epithelial cells can act as a transport mechanism for host miRNAs into specific bacteria and impact bacteria growth. These findings enhance our understanding of the specificity and mechanism of RNA trafficking from mouse to bacteria within the gut environment which could hold promising implications for modulating the gut microbiota to treat associated diseases in the future.

List of Abbreviations

Ago - Argonaute	**CFP** - cyan fluorescent protein
CLASH - Crosslinking ligation and sequencing of hybrids	**CLEAR** - Covalent ligation of endogenous Argonaute-bound RNAs
CLIP - Cross-linking immunoprecipitation	**CMV** - cytomegalovirus
Cre - Cre recombinase	**Cre+ cells** - Cre-expressing cells
Ct - cycle threshold	**Cy3** - Cyanine 3
Cy5 - Cyanine 5	**DIG** - digoxigenin
EAE - Experimental autoimmune encephalomyelitis	**EF1α** - elongation factor-1α
eGFP - enhanced green fluorescent protein	**ELISA** - Enzyme-Linked Immunosorbent Assay
EVs - extracellular vesicles	**FDR** - false discovery rate
FISH - fluorescence in situ hybridization	**FMT** - faecal microbiota transplantation
GI - gastrointestinal	**HRP** - horseradish peroxidase
LB - Lysogeny broth	**LNA** - locked nucleic acid
MDS - multidimensional scaling	**miRNAs** - microRNAs
OD600 - optical density at 600 nm	**OMVs** - outer membrane vesicles
PC – positive control	**PCoA -** Principal Coordinate Analysis
PCR - Polymerase chain reaction	**PGK** - phosphoglycerate kinase
PGM - purified gut microbiota	**RBPs** - RNA-binding proteins
rfp - red fluorescent protein	**RISC** - RNA-induced silencing complex
RT - reverse transcription	**sRNA** - small RNA

TSA - tyramide signal amplification	**UBC** - ubiquitin C
WT - wild type	

Table of Contents

List of Figures and Tables

Chapter 1: Introduction

Chapter 2: Material and Methods

Chapter 3: Identification of transferred host miRNAs in purified gut microbiota

Chapter 4: Identification of the role of mouse Ago2 protein in host miRNA-bacteria interaction

Chapter 5: Transport of RNA from mouse to bacteria via extracellular vesicles: probing specificity and function

Chapter 1: Introduction

1.1 The gut microbiota and its modulation

The human gastrointestinal (GI) tract hosts a complex and diverse population of microorganisms collectively known as the gut microbiota (Figure 1.1). It is estimated that the number of microorganisms inhabiting the GI tract exceeds 10^{14}, and the gut microbiome contains over 100 times as many genes as the human genome (Thursby and Juge, 2017). Despite the wide variety of microorganisms in the gut, the majority belong to just a few dominant gut microbial phyla. The most prevalent phyla in the gut microbiota are *Firmicutes*, *Bacteroidetes*, *Actinobacteria*, *Proteobacteria*, *Fusobacteria*, and *Verrucomicrobia* (Rinninella et al., 2019). Among these, *Firmicutes* and *Bacteroidetes* alone make up around 90% of the gut microbiota. The *Firmicutes* phylum is incredibly diverse, comprising more than 200 different genera. Some of the well-known genera within *Firmicutes* include *Clostridium*, *Lactobacillus*, *Bacillus*, *Enterococcus*, and *Ruminicoccus*. Notably, *Clostridium* genera account for 95% of the *Firmicutes* phylum. *Bacteroidetes* phylum mainly includes genera such as *Bacteroides* and *Prevotella*. The maintenance of the normal gut microbiota composition plays a vital role in supporting host health by contributing to various physiological functions, such as maintaining intestinal barrier (Natividad and Verdu, 2013), modulating the host's energy metabolism (Cani et al., 2019), protecting against infection by pathogenic bacteria (Buffie et al., 2015) and shaping host immunity (Wu et al., 2010; Gaudet et al., 2015).

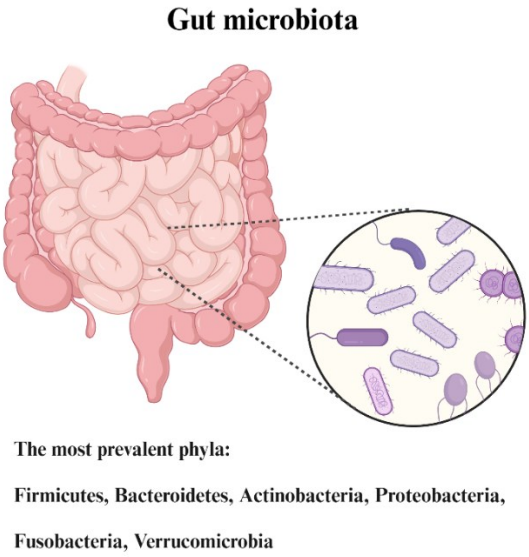


Figure 1.1 The gut microbiota and its most prevalent phyla. Figure created with BioRender.com.

Dysbiosis of the gut microbiota is associated with various human diseases, including inflammatory bowel diseases (Nishino et al., 2018), obesity (Walters et al., 2014), diabetes mellitus (Wen et al., 2008) and GI cancers (Serban, 2014) and modulating its composition is being considered a promising treatment for these diseases. Several strategies have been developed in order to restore or rebalance the gut microbiota, including the use of probiotics, prebiotics, and faecal microbiota transplantation (FMT) (Figure 1.2). Probiotics are living microorganisms that, when consumed in adequate quantities, confer health benefits to the human host (Kristensen et al., 2016). The most frequently utilized probiotic species belong to the *Lactobacillus* and *Bifidobacteria* genera, as well as some yeasts. These beneficial microorganisms are believed to enhance health through various mechanisms, including the promotion of beneficial gut microbial populations (Stratiki et al., 2007), competition with pathogenic species for adhesion sites (Servin, 2004), and the production of antimicrobial compounds (Cleusix et al., 2007). Prebiotics are substances characterized as selective fermentation components that induce specific alterations in the activity or composition of the gut microbiota, which are thought to confer benefits to the host (Rossen et al., 2015). These compounds primarily promote the growth and activity of beneficial microorganisms like *Lactobacilli* and *Bifidobacteria* (Tabibian et al., 2013). The effects of prebiotics include increased production of short-chain fatty acids and a reduction in pH levels (Fernández et al., 2016). FMT involves transferring faecal bacteria from

healthy donors to patients with intestinal diseases, dysbiosis, or alterations in their natural gut microbiota. The goal of FMT is to restore the community and functionality of the gut microbiota (Khoruts and Sadowsky, 2016). This approach has shown significant therapeutic potential in treating various diseases, including *Clostridium* difficile infection (Van et al., 2013), irritable bowel syndrome (Johnsen et al., 2018), and metabolic diseases (Vrieze et al., 2012).

However, various limitations associated with these modulation methods. For example, while many microorganisms used as probiotics are generally considered safe, their administration might raise safety concerns in certain vulnerable populations, notably immune-compromised individuals such as pregnant women, infants, the elderly, as well as critically ill or hospitalized patients (Ayichew et al., 2017). The drawbacks of FMT include the potential transfer of pathogens, temporary outcomes, stool toxicity, and challenges in replicating the procedure (Khan et al., 2022). Therefore, there is an ongoing demand for the development of novel, safe, and effective approaches for modulating gut microbiota.

Understanding the natural host mechanisms that regulate the gut microbiota may offer novel and safe strategies for precisely manipulating it. It is well established that host-derived products, such as mucus, antimicrobial peptides, and immunoglobulin A (IgA) can either promote the growth of certain microbial species or inhibit others (Suzuki et al., 2004; Carvalho et al., 2012) (Figure1.2). Over the past five years, another type of regulator of the microbiota from host has been proposed: Ribonucleic acids. In particularly, several studies have demonstrated that mammalian microRNAs (miRNAs) can have a specific impact on the composition and activity of gut bacteria, through direct RNA-RNA interactions with bacterial genes. These findings intersect with a growing body of research showing that small RNAs, including miRNAs, can function beyond the cell of origin in mammals, through transport in extracellular vesicles (EVs) or other carriers.

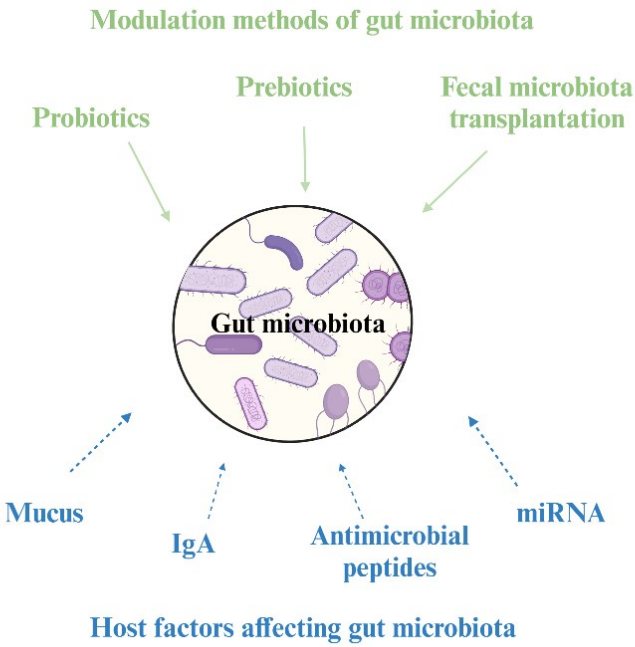


Figure 1.2 Current strategies for modulating the gut microbiota (top) and host factors influencing microbiota composition (bottom). Figure adapted from Hasan and Hongyi, 2019.

1.2 miRNAs: function and regulation in the gut

There are three predominant classes of small RNA in eukaryotes that mediate gene regulation and genome defence: small interfering RNAs, PIWI-associated RNAs and miRNAs (Kutter and Petr, 2008). Of these classes, miRNAs are the most well studied in mammals for their roles in development and disease (Kloosterman and Ronald, 2006). miRNAs derive from endogenous genes that are typically transcribed by RNA polymerase II into primary miRNA transcripts that undergo processing by Drosha and other cofactors in the nucleus (Pasquinelli, 2012) (Figure 1.3). This processing produces the pre-miRNAs (~60-70 nucleotides in length) that are transported to the cytoplasm by Exportin 5 where they are further processed by Dicer, which removes the loop region from pre-miRNAs. One strand of the resulting duplex is bound by an Argonaute (Ago) protein to form the RNA-induced silencing complex (RISC). Within RISC, the miRNA serves as the guide to bind to specific messenger RNA targets through sequence complementarity and this results in either cleavage of the mRNA or inhibition of translation (Bartel, 2018). miRNAs play important regulatory roles in various biological processes including early development, cellular differentiation, proliferation, apoptosis, developmental timing, and hematopoiesis (Dong et al., 2013). The aberrant expression levels of miRNAs have been linked to the initiation and development of human diseases, genetic disorders, and altered immune system function (Saal and Harvey, 2009; Tsitsiou and Lindsay, 2009).

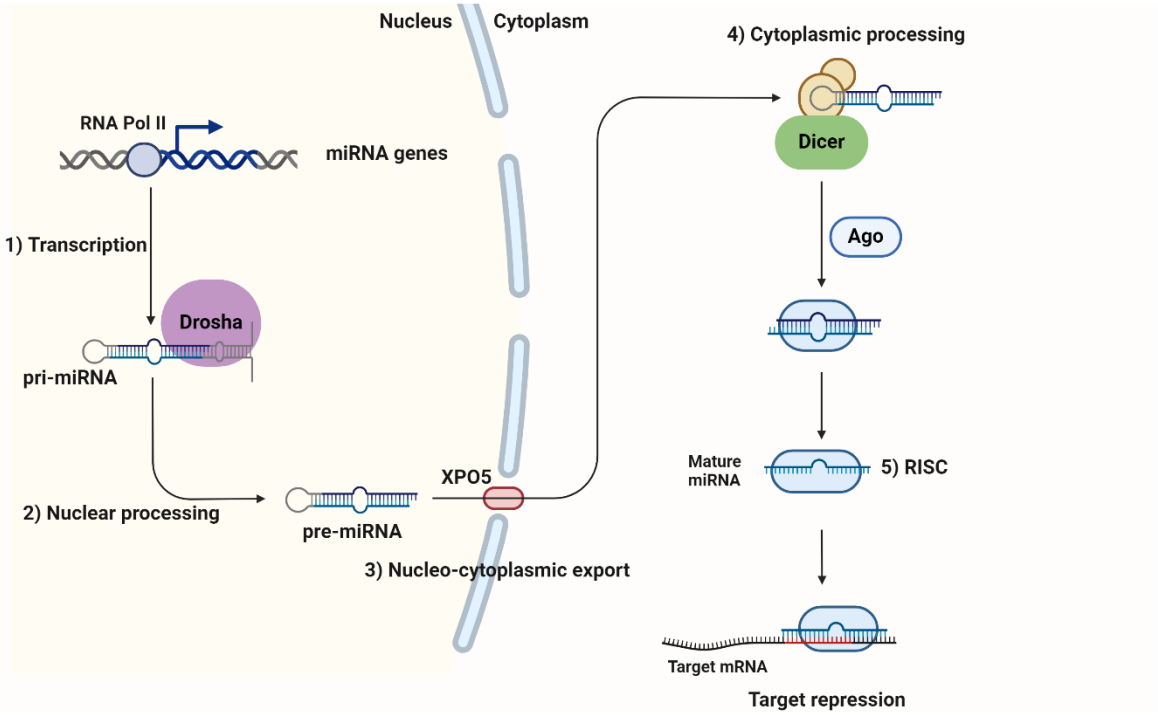


Figure 1.3 miRNAs biogenesis. miRNAs biogenesis involves several key steps including 1) the transcription of primary miRNA transcripts, 2) nuclear processing by Drosha, 3) nucleocytoplasmic export facilitated by XPO5, 4) cytoplasmic processing by Dicer, and 5) the formation of the RNA-induced silencing complex (RISC) in conjunction with Ago proteins. Figure created with BioRender.com.

Many reports have now demonstrated the importance of miRNAs in gut function. For example, in mice in which Dicer was ablated from intestinal epithelial cells there were fewer goblet cells, impaired nutrient absorption and barrier function (McKenna et al., 2010). The physiological characteristics of human intestinal epithelial cells were also shown to be directly regulated by transfection of synthetic miRNA mimics or inhibitors (Dalmasso et al., 2010). Moreover, miRNA expression profiles within intestinal cells have been shown to change during various intestinal diseases, and some of the miRNAs are likely to be causative agents in disease progression and development. For example, the level of miR-21 increased in colonic tissue from active ulcerative colitis and Crohn's disease, and its pathogenic role was confirmed in miR-21 knockout mice (Wu et al., 2008; Shi et al., 2013; Yang et al., 2013). miRNAs also play a role in the ability of intestinal cells to sense and respond to changes in the environment. For example, intestinal miRNA levels are altered when exposed to high levels of dietary lipids (Gil-Zamorano et al., 2020), and the presence or absence of the microbiome

directly affects miRNA expression in intestinal cells (Dalmasso et al., 2011; Singh et al., 2012). miRNA expression levels were further shown to be altered in a highly cell type-specific manner, with the strongest deregulation in stem cells, suggesting important roles of miRNAs in regulating intestinal homeostasis (Peck et al., 2017).

1.3 Extracellular miRNAs and miRNA-bacteria interactions in the gut

The above work suggests that the biogenesis and function of miRNAs may be closely tied to sensing changes in the intestinal environment. At the same time, miRNAs are also now recognized to be released into the environment, with scope to act outside of the cells from which they derive. Extensive literature in the last 15 years has shown that miRNAs can be detected in a cell-free form in different body fluids such as plasma, serum, saliva, and urine (Murillo et al., 2019). Furthermore, miRNAs are stable and can be reliably detected in faeces from human and mice (Ahmed et al., 2009; Wu et al., 2012), providing a non-invasive method to sample gut luminal miRNAs in different contexts. For example, the levels of miR-21 and miR-92a were significantly higher in the stool of patients with colorectal cancer compared with healthy controls (Wu et al., 2012). Accordingly, there is growing interest in using faecal miRNAs as non-invasive markers of intestinal malignancies (Yau et al., 2019).

miRNAs and other classes of extracellular small RNAs have been reported to participate in communication between different organisms and species (Claycomb et al., 2017). The relative abundance and stability of mammalian miRNAs in the gut lumen therefore leads to the question of whether these could modulate gene targets in other organisms in the gut. Several recent studies have reported direct RNA–RNA interactions between host miRNAs and bacterial genes in the gut (Table 1.1), based on a combination of *in vivo* genetic experiments in mice and *in vitro* validation experiments. Liu et al., 2016 used Dicer knockout mice in a first demonstration that the deficiency of miRNAs in intestinal epithelial cells caused an imbalanced gut microbiota, which could be restored by transplantation of wild-type faecal miRNAs (where the native RNA is extracted from faeces and administrated to mice by gavage). To explore whether specific miRNAs can directly interact with gut bacteria, the authors selected two important gut bacteria, *Fusobacterium nucleatum* and *Escherichia coli*, to predict their potential interacting miRNAs by sequence similarity. When culturing the strains *in vitro* with synthetic miRNAs, they found that miRNA-515-5p stimulated *F. nucleatum* growth and that miRNA-1226-5p stimulated *E. coli* growth. The expression levels of predicted miRNA targets in these two strains were also found to be increased in response to the synthetic RNA. However, further experiments, such as employing miRNA inhibitors (Robertson et al., 2010) to block miRNA regulation of target gene expression, are necessary to corroborate the direct RNA-RNA interactions

between host miRNA and bacterial mRNA. In a subsequent study, the same authors showed that miR-30d, which is enriched in the faeces of animals with experimental autoimmune encephalomyelitis (EAE), promoted the growth of *Akkermansia muciniphila* by increasing the expression of β-galactosidase (Liu et al. 2019). These examples suggest that host miRNA interactions with bacterial transcripts can lead to an increase in bacterial gene expression and bacterial growth. A more recent study suggests that miRNAs can also suppress bacterial growth. In particular, the composition of gut microbiota in miR-21a-5p knockout mice was characterized by an increase in *Lactobacillus* and incubation of synthetic miR-21 with *Lactobacillus reuteri* led to reduced growth (Santos et al., 2020).

Table 1.1 Reported miRNA-bacteria interactions in the gut

miRNA	Bacteria species	Target transcript	Effect on bacteria	Reference
miR-1226-5p	*Escherichia coli*	yegH mRNA	Promotion	Liu et al., 2016
miR-515-5p	*Fusobacterium nucleatum*	16S rRNA	Promotion	Liu et al., 2016
miR-30d-5p	*Akkermansia muciniphila*	β-Galactosidase mRNA	Promotion	Liu et al., 2019
miR-21a-5p	*Lactobacillus reuteri*	Unknown	Inhibition	Santos et al., 2020

1.4 Implications and questions about miRNA–bacteria interactions

Given the many diverse functions and mechanisms by which miRNAs can mediate cellular and organismal biology, it is an exciting prospect to consider their capacity to interact with other organisms. There are a growing number of examples of cross-species interactions of small RNAs in diverse living systems. In plants, mobile small RNAs have been shown to be part of a bidirectional arms race with fungal parasites (Weiberg et al., 2013; Cai et al., 2018). In honeybees, small RNAs are thought to mediate some of the caste-determining effects of royal jelly (Zhu et al., 2017) and are transmissible between organisms and across generations (Maori et al., 2019). As detailed in Section 1.3, a handful of studies also show specific mammalian miRNAs are naturally transported into some bacterial cells, where the miRNA forms a direct RNA-RNA interaction with a bacterial gene, resulting in changes in the microbiome with phenotypic effects in intestinal health. The potential impact of these findings is significant, since this suggests that intestinal miRNAs are an integral part of how the microbiome is regulated. Furthermore, since it is easy to synthesize different RNA sequences, one could envision this as a new, programmable method to dictate changes to the composition or activity of the microbiome.

These findings push forward the idea that RNA is central to cross-species interactions. Yet, many questions remain regarding how host miRNA-bacteria interactions occur under physiologically relevant conditions (Figure 1.4). Since there are hundreds of miRNAs as well as bacterial species/strains in the gut, it is a daunting prospect to systematically determine or predict which miRNA-bacteria interactions occur. Furthermore, the precise mechanisms governing the uptake of natural miRNAs by gut bacteria and how these transferred miRNAs integrate into functional pathways to regulate bacterial genes remain unclear.

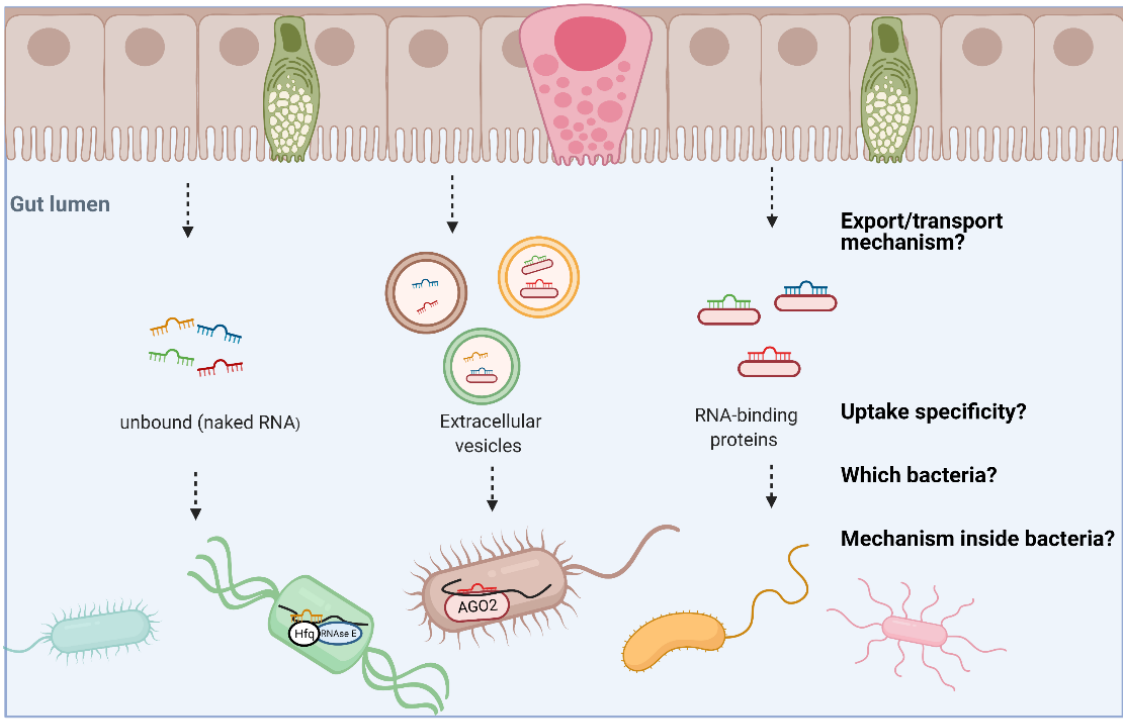


Figure 1.4 Knowledge gaps in miRNA-bacteria interactions. The mechanisms and specificity around RNA trafficking from host to bacteria in the gut are largely unknown. Figure from Du et al., 2021.

1.5 Strategies to identify specific miRNA-bacteria interactions|

The analysis of altered miRNAs in faecal samples from certain disease context and linking these changes to alterations in the microbiome can provide insights into miRNA-bacteria interactions. For example, *A. muciniphila* was found to be increased in the faeces from EAE mice compared with controls, and miR-30d was also enriched in faeces from EAE mice (Liu et al., 2019). However, it should be noted that the level of a miRNA in faeces is not a direct proxy for miRNAs internalized by bacteria, since dead host cells and extracellular luminal miRNAs will also be present. Nonetheless, the author showed that miR-30d could directly impact the growth of *A. muciniphila* by culturing the strain in the presence of synthetic miR-30d *in vitro* and the oral administration of synthetic miR-30d to mice *in vivo*. The miRNA-dependent increase in bacterial growth was not observed using scrambled miRNA sequence controls.

miRNA-bacteria interactions can also be inferred based on studies in miRNA knockout mice. In the miR-21 knockout mouse, *Lactobacillus* was found to be significantly increased in the small intestinal lumen when compared with wild-type animals (Santos et al., 2020). Although deletion of the miRNA could also indirectly affect the microbiome (e.g., through altered host cell populations and functions), the authors showed that synthetic miR-21a-5p directly suppressed the growth of *Lactobacillus in vitro*. Sequence specificity was demonstrated in this study too, since the scrambled miR-21a-5p sequence did not show effects.

Correlation analysis can be another method to identify potential miRNA-bacteria interactions. Horne et al., 2019 found fourteen significant correlations between miRNAs and bacteria taxa by assessing the association of the relative abundance of predicted taxa and level of miRNA in faecal samples. However, it is essential to note that while correlation analysis is informative, it lacks precision in the absence of direct experimental evidence.

1.6 Transport mechanisms for miRNA into bacteria cells

In the studies described above, the authors have cultured specific bacteria with synthetic miRNAs *in vitro* to show that naked miRNAs can enter bacteria, and that uptake or activity is sequence-specific. However, the natural miRNA uptake mechanism in the gut environment is not known, nor is it clear whether only specific bacteria internalize RNA. Here I introduce two potential mechanisms for miRNA transport: extracellular vesicles (EVs) and RNA-binding proteins.

1.6.1 Extracellular vesicles

In mammals, EVs are the best understood mechanism for miRNA stabilization and transfer between cells. There are three general classes of EVs, depending on their biogenesis: (i) exosomes (generally 50–100 nm in size), which are of endosomal origin and derive from multivesicular bodies; (ii) microvesicles (generally 100–1000 nm in size), which bud off the plasma membrane; and (iii) apoptotic bodies (generally 50–5000 nm in size), which are generated when cells undergo apoptosis (Colombo et al., 2014). Several key studies demonstrated that intestinal epithelial cells release exosomes, suggesting these are involved in antigen presentation (Van et al., 2001; Van et al., 2003). However, we now know that exosomes and other classes of EVs also contain RNAs, including miRNAs (Veziroglu and Mias, 2020). Seminal work by Valadi et al., 2007 showed that miRNAs can be transported from one mast cell to another via exosomes. Microvesicles derived from tumours were shown to deliver miRNAs to CD4+ T cells and thereby increase the number of regulatory T cells, promoting tumour growth (Yin et al. 2014). EVs can also mediate communication between different species in the gut. For example, the GI nematode *Heligmosomoides bakeri* releases miRNA-containing EVs that can be internalized by host mouse cells and modulate innate immunity during infection (Buck et al., 2014).

EVs therefore represent a potential transport mechanism for mammalian miRNAs in the gut, and there is some evidence that mammalian EVs interact with bacteria, with diverse functional outcomes. For example, EVs released from the respiratory epithelium during respiratory viral infection interact with *Pseudomonas aeruginosa* to provide nutrients and promote biofilms (Hendricks et al., 2021). EVs released from neutrophils or macrophages have been shown to have antimicrobial activities (Timár et al., 2013; García-Martínez et al., 2019). Interestingly, although Liu et al., 2016

demonstrated that synthetic miRNAs can enter bacteria *in vitro* (without cofactors), they also showed that EVs containing miRNAs were present in human and mouse faecal samples. Yet the question of specificity remains: Do all bacteria internalize EVs? Are there ligand/receptor interactions that dictate which bacteria internalize these? Intriguingly, a recent paper showed that plant-derived EVs from *Zingiber officinale* could also transport miRNAs into *Lactobacillus rhamnosus*, and proposed that the lipid content of the EVs was important for uptake and potentially specificity (Teng et al., 2018). Given the ubiquity of EVs derived from organisms in the gut, as well as potential dietary sources of EVs, it will be important to test the breadth of cross-species transfer of miRNA by EVs and understand the underlying properties that define specificity. It remains possible that other types of gut inhabitants (e.g., fungi) could also internalize host miRNAs. There is little information of this type of interaction in mammalian systems but a burgeoning body of data suggesting plant–fungi RNA–RNA interactions (Cai et al., 2018).

1.6.2 RNA-binding proteins

Another potential mechanism for stabilization and transport of mammalian miRNAs is RNA-binding proteins (RBPs). RBPs are involved in miRNA biogenesis as well as function in the RNA-induced silencing complex (Connerty et al., 2015). Recent studies have uncovered a role of RBPs also in the selection, transportation and stabilization of miRNAs that are exported into the extracellular environment. For example, miRNAs have been reported to be delivered by high-density lipoprotein to recipient cells and regulate the expression of target genes (Vickers et al., 2011). Several studies have demonstrated that the majority of extracellular miRNAs present in blood plasma and cell culture exist in a non-membrane-bound form, associating specifically with Ago proteins (Arroyo et al., 2011; Turchinovich et al., 2011). This association is proposed to play a critical role in stabilizing miRNAs in the extracellular environment. Within mouse muscle cells, the miRNA-Ago2 complex has been observed to translocate from the cytoplasm into mitochondria, which resemble bacteria in some aspects (Zhang et al., 2014). Recent findings from our lab demonstrate that Ago2 is primarily secreted by mouse intestinal epithelial cells in a soluble non-vesicular form, associating specifically with distinct populations of miRNAs. Therefore, transport via Ago2 provides a possible mechanism for how miRNAs could enter bacteria and elicit gene regulation in the bacteria.

1.7 Functional integration of miRNAs in bacteria

In mammals, it is assumed that miRNAs only function through association with an Ago protein, where their role is to guide this protein to a nucleic acid target with sequence complementarity (Figure 1.3). Human and mouse have four Ago subfamily members (Ago1 to Ago4) and all of them can bind miRNAs and trigger translational repression (Liu et al., 2004; Adiliaghdam et al., 2020). Previous data from our laboratory indicates that Ago2 protein is the most abundant Ago protein present in mouse gut tissue. Moreover, Ago2 is reported to be transported into mitochondria, which are evolutionarily related to bacteria, and miRNA-Ago2 complexes can enhance mitochondrial translation (Zhang et al., 2014). An important mechanistic question here is whether and how Ago2 would operate in gene silencing or activation in bacteria.

Bacteria can regulate gene expression with their own small RNAs (sRNAs) categorized as cis-encoded and trans-encoded sRNAs (Waters et al., 2009; Storz et al., 2011). Cis-encoded sRNAs are transcribed from the DNA strand opposite their mRNA target. They are primarily involved in post-transcriptional regulation employing various mechanisms such as transcription attenuation, translation inhibition, and mRNA decay modification. Transcription attenuation occurs when the sRNA binding affects the formation of a termination structure in the target mRNA. The most prominent mechanism is translation inhibition, where the sRNA binds perfectly to the ribosome binding site (RBS) of the mRNA, competing with ribosome loading. Additionally, binding of the sRNA to its target mRNA can also alter mRNA decay. Trans-encoded sRNAs, associated with intergenic regions, regulate multiple mRNAs distantly from their origin through partial complementarity. They can activate translation via the 5’ untranslated region (5’ UTR), relieving structural inhibitions on translation. Conversely, they can inhibit translation through transcription attenuation, translation repression, and mRNA degradation mechanisms, with regulatory signals frequently located in the 5' UTR of target mRNAs. In Gram-negative bacteria, sRNAs often function alongside RNA chaperone proteins like Hfq, which facilitates sRNA-mRNA interactions, translation modulation, and mRNA degradation (Felden and Yoann, 2021). Additionally, other bacteria use proteins like ProQ and ProQ-like proteins for RNA binding. Since several reports have shown that culturing bacteria with synthetic miRNA mimics *in vitro* directly impacts bacterial gene expression (Liu et al., 2019; Santos et al., 2020), it is also important to consider whether and how

miRNAs integrate into bacterial sRNA machinery and function with bacterial protein partners to regulate gene expression.

1.8 PhD Aims

Mammalian miRNAs can influence the composition of gut microbiota by entering bacteria and interacting with bacterial genes. However, many questions remain regarding how host miRNA-bacteria interactions occur and the detailed mechanisms by which these occur under physiological conditions. My PhD project focuses on profiling transferred mouse miRNAs in the gut microbiota, and investigating the specificity and mechanism of RNA trafficking from mouse to bacteria. There are four aims for my thesis:

Aim 1 (Chapter 3): Determine whether mouse miRNAs are naturally taken up by gut bacteria *in vivo* and identify transferred host miRNAs and bacteria that internalize them.

Aim 2 (Chapter 4): Investigate whether mouse Ago2 protein could transport miRNAs into gut bacteria and thereby mediate miRNA-bacterial RNA interactions.

Aim 3 (Chapter 5): Investigate whether host miRNAs are delivered by mammalian EVs into specific bacteria and whether they impact bacterial growth and transfer functional RNA cargo.

Aim 4 (Chapter 6): To complement Aim 3 I explored the capacity to use a Cre-loxP genetic reporter system to probe functional transmission of cargo from mouse cells to specific bacteria.

Chapter 2: Material and Methods

2.1 Cell lines, bacterial strains, and culture conditions

Mode-K cells (a mouse intestinal epithelial cell line) were cultured in complete Dulbecco's Modified Eagle Medium (DMEM) (Sigma-Aldrich) supplemented with 10% (v/v) fetal bovine serum, 1% (v/v) L-Glutamine, 1% (v/v) Penicillin/Streptomycin, 1% (v/v) Sodium Pyruvate and 1% (v/v) Non-Essential Amino Acids (Vidal et al., 1993) at 37°C with 5% CO2.

Salmonella typhimurium SL1344 and *Escherichia coli* W3110 were routinely grown on LB agar plates (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, and 15 g/L agar, pH at 7.0). A single colony was picked and incubated in LB liquid medium (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl, pH at 7.0) or M9 minimal medium (1 x M9 minimal salts, 1M MgSO4, 1M CaCl2, and 20% (w/v) glucose) supplemented with 40 µg/ml histidine at 37 °C for experiments.

2.2 Animals used

C57BL/6J mice and Villin-cre transgenic mice (B6.Cg-Tg(Vil1-cre)997Gum/J) were bred by the in-house facilities at The University of Edinburgh. Experimental procedures were executed under a UK Home Office licence as approved by qualified veterinarians and were carried out by qualified personnel in accordance with the UK Home Office guidelines.

2.3 DNA and RNA methods

2.3.1 DNA extraction

Genomic DNA from gut contents (Input), purified gut microbiota (PGM) and faeces was extracted using the QIAamp Fast DNA Stool Mini Kit (Qiagen, Cat#: 51604) according to the manufacturer's protocol with minor adjustments. Approximately 100 mg of material per sample was employed for DNA extraction. Samples were added in a 2 ml microcentrifuge tube and placed on ice. The remaining steps were as recommended with the exception that 50 µl of DNase-free H_2O was used to elute DNA. For the extraction of genomic DNA from Mode-K cells, the PureLink™ Genomic DNA Mini Kit (Invitrogen™, Cat#: K182001) was used according to the manufacturer's protocol. In the final step, 50 µl of DNase-free H_2O was used to elute DNA. Extraction

of plasmid DNA from bacteria was carried out using the GeneJET Plasmid Miniprep Kit (Thermo Scientific, Cat#: K0502) according to the manufacturer's protocol. In the final step, 50 µl of DNase-free H_2O was used to elute plasmid DNA. The DNA concentration was measured using the NanoDrop® ND-1000 UV-Vis Spectrophotometer. DNA was stored at -20°C.

2.3.2 RNA extraction

Total RNA from PGM, Input, and faeces was extracted using the Norgen Stool Total RNA Purification Kit (Norgen, Cat#: 49500) with 120 mg of materials processed per sample. The protocol was carried out according to manufacturer's recommendations. with the following modifications. In the final step, 50 µl of RNase-free H_2O was used to elute RNA. For the extraction of total RNA from EVs, EV-depleted supernatant, and mouse serum, the miRNeasy Serum/Plasma Kit (QIAGEN, Cat#: 217004) was used according to manufacturer's recommendations. In the final step, 14 µl of RNase-free H_2O was used to elute RNA. Total RNA from Mode-K cells was extracted using the miRNeasy Mini Kit (QIAGEN, Cat#: 217184). In the final step, 50 µl RNase-free H_2O was used to elute RNA. The RNA concentration was determined using the Qubit® RNA HS Assay Kits (Invitrogen™, Cat#: Q32852). The purified RNAs were stored at -80°C for long term storage.

2.3.3 Polymerase Chain Reaction (PCR)

PCR was performed using Q5 High-Fidelity Hot Start Polymerase (New England Biolabs, Cat#: M0493) under the following conditions (Table 2.1 and 2.2). PCR primers used in this study are listed in Table 2.3. Following PCR, samples were run on 1-3% (w/v) agarose gels in TAE buffer, and visualised using a Licor D-DiGit gel scanner.

Table 2.1 PCR reaction

Component (25µl reaction)	Volume/reaction
5x Q5 Reaction Buffer	5 µl
10 mM dNTPs	0.5 µl
10 µM Forward Primer	1.25 µl
10 µM Reverse Primer	1.25 µl
Q5 Hot Start High-Fidelity DNA Polymerase	0.25 µl
Template DNA	variable
Nuclease-Free Water	to 25 µl

Table 2.2 Thermocycling Conditions

Step	Temperature	Time
Initial denaturation	98°C	30 seconds
30 Cycles	98°C	10 seconds
	*Tm-5°C	30 seconds
	72°C	30 seconds/kb
Final extension	72°C	5 minutes
Hold	4–10°C	

* Tm: Melting Temperature

Table 2.3 Forward and reverse primer sequence and amplicon length

Gene	Forward primer	Reverse primer	Lengths	Tm	Reference
Mouse 18S rDNA	GGACCAGAGCGAAAGCATTTGCC	TCAATCTCGGGTGGCTGAACGC	495bp	61.1°C / 62°C	Jensen et al., 2004
Mouse eEF2	TGTCAGTCATCGCCCATGTG	CATCCTTGCGAGTGTCAGTGA	123bp	53.5°C / 53.9°C	Eissa et al., 2017
Bacteria 16S rDNA	AGAGTTTGATCATGGCTCAG	TACGGTTACCTTGTTACGACTT	1500bp	52.2°C / 53.5°C	Frank et al., 2008
Cre	GACATGTTCAGGGATCGCCA	CAGCGTTTTCGTTCTGCCAA	525bp	57.4°C / 56.6°C	This study
Cre (RT-PCR)	CCCTGTTTCACTATCCAGGT	GGGTAACTAAACTGGTCGAG	353bp	54°C / 52.4°C	Zhou et al., 2012

2.3.4 Reverse transcription (RT)

Reverse transcription of mRNA was carried out using the QuantiTect Reverse Transcription Kit (QIAGEN, Cat#: 205311). The genomic DNA elimination reaction was prepared on ice, following the instructions outlined in Table 2.4. This reaction was incubated for 2 minutes at 42°C, after which it was promptly placed on ice. The reverse-transcription master mix was prepared on ice, as specified in Table 2.5. Subsequently, 14 µl of template RNA (entire genomic DNA elimination reaction) was added to each tube containing the reverse-transcription master mix. The reaction was then incubated for 15 minutes at 42°C, followed by a 3-minute incubation at 95°C to deactivate the Quantiscript Reverse Transcriptase.

Table 2.4 Genomic DNA elimination reaction components

Component	Volume/reaction
gDNA Wipeout Buffer, 7x	5 µl
Template RNA	0.5 µl
RNase-free water	1.25 µl
Total volume	1.25 µl

Table 2.5 Reverse-transcription reaction components

Component	Volume/reaction
Quantiscript Reverse Transcriptase	1 µl
Quantiscript RT Buffer, 5x	4 µl
RT Primer Mix	1 µl
Entire genomic DNA elimination reaction	14 µl
Total volume	20 µl

Reverse transcription of miRNA was carried out using the miScript II RT kit (QIAGEN, Cat#: 205311). Template RNA was thawed on ice. 10x miScript Nucleics Mix and 5x miScript HiFlex Buffer were thawed at room temperature (15–25ºC). The reverse-transcription master mix was prepared on ice according to Table 2.6. The reverse-transcription master mix contains all components required for first-strand cDNA synthesis except template RNA. Template RNA was added to each tube containing reverse-transcription master mix and the reactions were incubated for 60 minutes at 37ºC, followed by a 5-minute incubation at 95°C to inactivate miScript Reverse Transcriptase.

Table 2.6 Reverse-transcription reaction components

Component	Volume/reaction
5x miScript HiFlex Buffer	4 µl
10x miScript Nucleics Mix	2 µl
miScript Reverse Transcriptase Mix	2 µl
Template RNA	Variable (200 ng)
RNase-free water	Variable
Total volume	20 µl

2.3.5 Quantitative real-time PCR (qPCR)

Real-Time PCR for Detection of Mature miRNA or noncoding RNA was carried out using the miScript SYBR Green PCR kit (QIAGEN, Cat#: 218073). The reaction mix was prepared according to Table 2.7 for a 10 µl per well reaction volume (used in 384-well plates) and mixed thoroughly but gently. The 384-well plates were tightly with film and centrifuged for 1 minutes at 1000 x g at room temperature (15–25°C) to remove bubbles. The real-time cycler was programmed according to Table 2.8 in the LightCycler® 480 Instrument I (Roche). Melting curve analysis was used to document amplicon specificity and cycle threshold (Ct) value was determined.

Table 2.7 Reaction setup for real-time PCR

Component	Volume/reaction (384-well)
2x QuantiTect SYBR Green PCR Master Mix	5 µl
10x miScript Universal Primer	1 µl
10x miScript Primer Assay	1 µl
Template cDNA	1 µl
RNase-free water	2 µl
Total volume	10 µl

Table 2.8 Cycling conditions for real-time PCR

Step		Temperature	Time
PCR Initial activation step		95°C	15 min
3-step cycling (40 cycles)	Denaturation	94°C	15 s
	Annealing	55°C	30 s
	Extension	70°C	30 s

2.4 Method of purifying gut microbiota from mouse gut content

Three C57BL/6J male mice were used for gut content collection. The whole intestines from duodenum to rectum were dissected from the sacrificed mice quickly. Faecal pellets were collected from rectum then stored in -80°C. The intestines were cut into several pieces and the luminal contents were squeezed out, homogenized, and suspended in 20ml PBS. The suspensions were filtered through 70 µm nylon mesh to

remove big particles and gut debris. After the filtrates were collected, materials retained on the filter were re-homogenized in 2 ml of PBS, and re-filtered. The filtrates were centrifuged at 2,300 x g, 4°C for 5 minutes and then the pellets collected were washed twice with 15 ml PBS. The pellets were re-suspended in 15 ml 1% (v/v) Triton X-100 (Sigma, Cat#: 9036-19-5) and incubated for 10 minutes at room temperature. The pellets were washed twice with PBS and re-suspended in 1 ml of 0.5 mg/ml proteinase K (Life Technologies, Cat#: AM2546) and incubated at 37°C for 10 minutes. The pellets were washed twice with PBS and re-suspended in 1ml 1x TURBO™ DNase Buffer with 5 µl TURBO™ DNase (Life technologies, Cat#: AM2238) and incubated at 37°C for 10 minutes. The pellets were washed twice with PBS and re-suspended in 1 ml of 0.8 mg/ml RNaseA (Thermo Scientific™, Cat#: EN0531) and incubated at 37°C for 5 minutes. The pellets were washed twice with PBS and re-suspended in 1 ml of 10 mg/ml lysozyme (Thermo Fisher Scientific, Cat#: J60701.06) and incubated at 37°C for 30 minutes. The pellets were collected by centrifugation at 2,300 x g for 5 minutes and stored in -80°C.

2.5 Bacterial 16S rDNA sequence-based survey

The amplicon libraries that span the fourth hypervariable domain of microbial 16S ribosomal RNA (rRNA) genes were prepared using NEXTflex™ 16S V4 Amplicon-Seq Kit 2.0 (Bioo scientific, Cat#: NOVA-4203-01) and sequenced on the Illumina Miseq sequencing platform using 250PE protocol by Edinburgh Genomics. Sequencing data was processed using QIIME 2 (Version 2022.8) following an established protocol (Bolyen et al., 2019). Briefly, reads were trimmed to remove low quality bases and forward and reverse reads were then merged. The quality settings for DADA2 (Callahan et al., 2016) were “- -p-trunc-len-f 230” and “- -p-trunc-len-r 150”. After demultiplexing sequences, operational taxonomic units (OTUs) were assigned using a closed-reference OUT picking protocol using the QIIME 2 filtered at 97% identity. The phylogenetic tree was constructed from the representative sequences and the sampling depth is 59400. The rarefied OTU frequency matrix and phylogenetic tree were used with beta diversity analysis to produce weighted UniFrac, and bray-curtis distance matrices. A Principal coordinates analysis (PCoA) plot was generated based on Unifrac distances matrices to test the similarity of the community members.

2.6 sRNA library preparation

Libraries for small RNA sequencing were prepared using the CleanTag small RNA library prep kit (Trilink, Cat#: L-3206) according to the manufacturer's instructions.

2.6.1 3' Adapter Ligation to RNA Template

The 3' adapter was diluted according to the RNA input, using a 1:2 ratio for 100 ng and 1:10 for samples with low concentrations. The RNA samples were heated to 70°C for 2 minutes before being added to the reaction mix (Table 2.9). The reaction was incubated in a thermocycler for 1 hour at 28°C, followed by an additional incubation for 20 minutes at 65°C.

Table 2.9 Reaction mix for 3' adapter ligation

Component	Volume/reaction
CleanTag 3' adapter	0.5 µl
Buffer 1	2.5 µl
Enzyme 1	0.5 µl
RNase inhibitor	0.5 µl
RNA template	1 µl
Total volume	5 µl

2.6.2 5' Adapter Ligation to 3' Tagged RNA Template

The 5' adapter was similarly diluted as the 3' adapter, heated to 70°C for 2 minutes, and then added to the reaction mix (Table 2.10). The reaction was incubated in a thermocycler for 1 hour at 28°C, followed by an additional incubation for 20 minutes at 65°C.

Table 2.10 Reaction mix for 5' adapter ligation

Component	Volume/reaction
3' tagged RNA	5 µl
Nuclease-free water	2 µl
Buffer 2	0.5 µl
Enzyme 2	1 µl
RNase inhibitor	0.5 µl
CleanTag 5'adapter	1 µl
Total volume	10 µl

2.6.3 Reverse Transcription (RT) Reaction of Tagged RNA Library

The tagged RNA was mixed with 1 µl RT primer (5 µM) and incubated at 70 °C for 2 minutes before adding the following reagents (Table 2.11) to the mix. The mixture was incubated at 50 °C for 1 hour. The generated cDNA was stored at -20 °C overnight or used immediately for PCR amplification.

Table 2.11 Reaction mix for RT

Component	Volume/reaction
5' tagged RNA + RT primer	11 µl
Nuclease-free water	0.96 µl
RT Buffer	2.88 µl
dNTPs (10 mM)	0.72 µl
DTT (1 M)	1.44 µl
RNase inhibitor	0.5 µl
RT enzyme	0.5 µl
Total volume	18 µl

2.6.4 PCR Amplification of RT Product

The PCR Mix was prepared according to Table 2.12. Specific index primers were incorporated into the reverse primers during PCR amplification to distinguish the various samples within the sequencing pool.

Table 2.12 Reaction mix for PCR amplification

Component	Volume/reaction
cDNA	18 µl
High Fidelity PCR Master Mix	20 µl
Forward primer (20 µM)	1 µl
Index primer	1 µl
Total volume	40 µl

The samples were subjected to amplification in a thermocycler following the conditions outlined in Table 2.13. The number of cycles used varied based on the quantity of input material: RNA derived from Input, PGM, and Faeces were amplified for 15 cycles, while EVs and EVs-depleted supernatants were amplified for 22 cycles.

Table 2.13 The cycling conditions for PCR

Step	Time	Temperature	Cycles
Initial denaturation	30 s	98 °C	1
Denaturation	10 s	98 °C	15-20
Annealing	30 s	60 °C	
Elongation	15 s	72 °C	
Final extension	10 min	72 °C	1

2.6.5 Library pooling and purification

The profiles of each sRNA library were evaluated using the High Sensitivity DNA Bioanalyser chip. In order to achieve uniform sequencing depth coverage across all libraries, I combined equal amount of each sRNA library based on the High Sensitivity DNA Bioanalyzer chip analysis. Pooled libraries were run on a 6% Novex™ TBE gel (Invitrogen™, Cat#: EC62655BOX). The quick load pBR322 DNA digested with MspI (NEB) was used as size indicator. Gels were run at 100 V for 80 minutes and stained in SYBR gold (1:10,000 dilution in TBE buffer) (Invitrogen™, Cat#: S11494) for 15 minutes on shaker covered with foil. Stained gels were then imaged on the D-Digit gel scanner (Li-cor). The region between 140-180 bp was then excised over a blue LED transilluminator (IO Rodeo). Gel pieces were fragmented by centrifugation (3 minutes, 5,900 x g) through a hole in a 0.5 ml tube into a 1.5 ml tube. Libraries were eluted from gel fragments by rotation overnight at 4°C in nuclease-free water and subsequently separated from the gel by spinning the sample through a Spin-X centrifuge tube filter (Corning Costar) for 2 minutes at 16,100 x g. The libraries were precipitated overnight at -80 °C using 2.5 volumes of 100% Ethanol, 1/10 volume Sodium Acetate (3 M, pH 5.5) and 1 µl GlycoBlue Coprecipitant (Invitrogen™, Cat#: AM9515). Samples were then pelleted at 16,100 x g, 4 °C for 30 min and precipitated RNA was washed with 500 µl of 70% ethanol, air dried on ice for 15 min and re-suspended in 20 µl nuclease-free water. Concentrations of the purified, pooled libraries were measured using Qubit dsDNA HS kit (Invitrogen™, Cat#: Q33230). The libraries were sequenced using a single-end 100 base pair sequencing run on the Illumina NextSeq 2000 platform by the Edinburgh Clinical Research Facility.

2.7 Fluorescence in situ hybridization for miRNA and bacterial rDNA detection

2.7.1 Cell fixation and permeabilization

The bacterial cells were fixed in 400 µl of 4% (w/v in PBS) paraformaldehyde (Sigma, Cat#: 30525-89-4) for 1 hour at room temperature, then pelleted from this mixture by centrifugation at 5,900 x g for 3 minutes and washed twice with 400 µl PBS. The fixed cells were suspended in 500 µl of 50% (v/v in PBS) ethanol and incubated at -20˚C for at least 30 minutes. After two time 400 µl PBS washes, the cells were treated with 800 µl of the 1 mg/ml lysozyme solution (Sigma, Cat#: 9001-63-2) for 15 minutes at 37˚C with vortexing and then pelleted from this mixture by centrifugation at 5,900 x g for 3 minutes and washed twice with 400 µl PBS. Next the cells were treated with 800 µl of the 20 µg/ml proteinase K solution (in Tris-EDTA buffer) (Thermo Scientific™, Cat#: EO0491) for 37 °C for 15 minutes with vortexing and then pelleted from this mixture by centrifugation at 5,900 x g for 3 minutes and washed twice with 400 µl PBS.

2.7.2 Hybridization and post-hybridization washes

The hybridization solution was made up of 50% (v/v) formamide (Thermo Scientific™, Cat#: 17899), 10% (w/v) dextran sulfate (Sigma-Aldrich, Cat#: D8906), 1x Denhardt's solution (Sigma-Aldrich, Cat#: D2532), 50 mM sodium phosphate pH 7.0, 2x sodium citrate (SSC) (Applied Biosystems, Cat#: AM9770), 20 µg of sheared salmon sperm DNA (SSSD) (Applied Biosystems, Cat#: AM9680) and 20 µg yeast tRNA (Life Technologies, 15401-011). For 16S rDNA detection, the cells were suspended in 98 µl hybridization solution and 2 µl of 100 µM 16S rDNA probe or scramble 16S probe and incubated at 46°C for 4 hours. For miRNA detection, the LNA miRNA probes were denatured by heating at 90°C for 4 min, then chilled on ice. The cells were suspended in 99.5 µl hybridization solution and 0.5 µl of 10 µM LNA miRNA probe (final con. 50 nM) and incubated at 54°C for 2 hours. The probes used are listed in Table 2.14.

Table 2.14 List of fluorescence in situ hybridization probes

Probes	Sequence 5'-3'	Modification	Target	Manufacturer
DNA probes				
Eub338	GCTGCCTCCCGTAGGAGT	5'-Cy3	Bacterial 16S rDNA	Integrated DNA Technologies
Scramble Eub338	ACTCCTACGGGAGGCAGC	5'-Cy3	Negative control	Integrated DNA Technologies
Lab158	GGTATTAGCACCTGTTTCCA	5'-Cy3	*Lactobacillus* 16S rDNA	Integrated DNA Technologies
LNA probes				
mmu-miR-21-5p miRCURY LNA miRNA Detection probe	UAGCUUAUCAGACUGAUGUUGA	5'-DIG and 3'-DIG	mmu-miR-21-5p	QIAGEN
mmu-let-7c-5p miRCURY LNA miRNA Detection probe	UGAGGUAGUAGGUUGUAUGGUU	5'-DIG and 3'-DIG	mmu-let-7c-5p	QIAGEN
Scramble-miR miRCURY LNA Detection probe	--	5'-DIG and 3'-DIG	Negative control	QIAGEN

Following hybridization, 1.0 mL 0.1x SSC with 0.1% (v/v) tween-20 (Sigma, Cat#: P9416) (SSCT) was added in the hybridization mixture then the cells were centrifuged at 5,000 rpm for 3 minutes and the supernatants were aspirated. The pellets were suspended in 200 µl of 50% formamide, 2x SSC, 0.1% tween-20 to the cell pellets, mixed thoroughly and incubated for 30 minutes at 55 °C in a heating block. 1 ml 0.1x SSCT was added to the mixture, then the cells were centrifuged at 2,300 x g for 3 minutes and the supernatants were aspirated. The pellets were suspended in 500 µl 0.1x SSCT and incubated for 40 minutes at 55 °C in a heat block then centrifuged at 2,300 x g for 3 minutes and the supernatants were aspirated.

2.7.3 Immunofluorescence

Cells were incubated in 0.3% (v/v) hydrogen peroxide H_2O_2 (Sigma, Cat#:7722-84-1) (diluted in distilled water) for 15 minutes at RT and then washed in 400 µl 1x PBS-T three times. The pellets were suspended in 300 µl of 1x DIG blocking buffer (Roche, Cat#: 11585762001) and incubated at RT for 1 hour. 1x DIG blocking buffer was prepared fresh by diluting the Blocking solution (10x) with 1x Maleic acid buffer (Roche, Cat#: 11585762001) to a 1x DIG blocking buffer. The cells were centrifuged at 2,300 x g for 3 minutes and the supernatants were aspirated. The cells were treated with 300 µl of HRP Anti-Digoxigenin antibody (Abcam, Cat#: ab51949) (1:200 dilution in 1x DIG blocking buffer) and incubated at RT for 1 hour. After three PBS-T washes, the cells were incubated with 200 µl Cy5-TSA working solution (Stratech, Cat#: K1052) at RT for 1 hour. Cy5-TSA working solution was prepared by diluting fluorophore tyramide stock solution 1:200 in 1x Amplification diluent to make fluorophore tyramide working solution with 0.3% H_2O_2. The cells were washed three times with PBS-T and PBS once. The cells were treated with 100 µl DAPI solution (300 nM) (Thermo

Scientific™, Cat#: 62248) and incubated for 5 minutes at RT. The cells were then washed three times with PBS.

2.7.4 Microscopy

5 µl bacterial suspension was spotted on microscope slides with 2% agarose pad and then was covered with 10 µl ProLong™ Gold Antifade Mountant (Invitrogen™, Cat#: P36930) and a coverslip. Samples were imaged with Zeiss LSM880 Airyscan with a 63x oil objective.

2.8 Ago2 Immunoprecipitation

2.8.1 Preparation of lysate

Ago2 was immunoprecipitated from several starting materials including: mouse gut tissue, gut contents, faeces, and purified gut microbiota (PGM). The gut tissue, specifically from the cecum, was promptly flash-frozen in liquid nitrogen upon euthanizing the mice and subsequently ground into a fine powder under liquid nitrogen. 100 mg of the tissue powders were lysed using 1 ml lysis buffer (50 mM Tris.HCl pH 7.5, 300 mM NaCl, 1% IGEPAL CA-630, 5 mM EDTA, 10% glycerol; filtered with 0.22 µm Millex-GP filter) containing protease inhibitors (1 tablet per 10 ml) (Roche, Cat#: 11873580001) for 30 minutes on ice. To generate lysates from mouse gut contents and faeces, the whole intestine was dissected from the sacrificed mouse and faecal pellets were collected from the rectum and the luminal contents were squeezed out from the remaining intestine by tweezer. PGM was prepared as the protocol in section 2.4. 100 mg of gut contents and faecal sample and 200 mg of PGM were homogenized and lysed in 1 ml lysis buffer with 0.1mm zirconium grinding beads using the Tissue Lyser II (Qiagen) in pre-cooled cartridges for 5 minutes at 30 Hz twice and then incubated on ice for 20 minutes. Cell/Tissue debris were removed by centrifugation at 16,100 x g, 4°C for 20 minutes.

2.8.2 Antibody conjugation

50 µl Dynabeads Protein G (Invitrogen™, Cat#:1004D) per immunoprecipitation (IP) were placed in a 1.5 ml tube on a magnet stand, and the beads' buffer was subsequently removed. The beads were washed five times with 1 ml of filtered binding wash buffer (BWB), consisting of 50 ml of PBS and 10 µl of Tween 20. The beads were re-suspended by gently tapping the tube while adding BWB. 10 µl of mouse anti-

mAGO2 antibody (1.15 mg/ml) per IP in 600 µl BWB was added to the beads and this mixture was placed on a spinning wheel at 4°C for at least 5 hours.

2.8.3 Sample addition, immunoprecipitation, and elution

The conjugated beads were subjected to three washes using 500 µl of lysis buffer. The beads tube was then positioned on a magnet, and the buffer was carefully aspirated. Following this step, the samples were added to each beads tube and placed on a spinning wheel at 4°C overnight. Post-incubation, the tubes were placed back on the magnet to remove the unbound fraction, which was subsequently stored at -80°C until further use. The beads underwent a series of washes to eliminate nonspecific binding. Firstly, a wash with Low Salt buffer (composed of 50 mM Tris.HCl pH 7.5, 300 mM NaCl, 5 mM MgCl2, 0.5% Triton X-100, and 2.5% glycerol; filtered with a 0.22 µm Millex-GP filter) was performed, followed by two washes with High Salt buffer (containing 50 mM Tris.HCl pH 7.5, 800 mM NaCl, 10 mM MgCl2, 0.5% Triton X-100, and 2.5% glycerol; filtered with a 0.22 µm Millex-GP filter) for 5 minutes each at 4°C while rotating. Another wash with Low Salt buffer and a subsequent wash with PNK buffer (consisting of 50 mM Tris.HCl pH 7.5, 50 mM NaCl, 10 mM MgCl2, 0.5% Triton X-100; filtered with a 0.22 µm Millex-GP filter) were conducted. Finally, the proteins bound to the beads were eluted using 4x NuPAGE LDS sample buffer (Thermo, Cat#: NP0008) with 0.1 M DTT, and incubated for 10 minutes at 70°C with 1,000 rpm rotation.

2.9 Western Blot analysis

Protein lysates were mixed in 4x NuPAGE LDS sample buffer with 0.1 M DTT then denatured for 10 minutes at 70°C (For Ago2 IP samples, disregard this step). The samples were run on 4-12% Bis-Tris NuPAGE SDS gels (Thermo) in 1x NuPAGE MOPS SDS running buffer (Thermo, Cat#: NP000102) at 120V for varying times depending on the size of the protein. The Precision Plus Protein All Blue Prestained Protein Standards ladder (Bio-Rad, Cat#: 1610373) was used as a size marker. The separated proteins were transferred on an Immobilon-FL PVDF membrane (Millipore, Cat#: IPFL00010) pre-activated in 100% methanol. Wet transfer was performed in 1x NuPAGE Transfer buffer (Thermo, Cat#: NP00061) with 10% (v/v) methanol for 105 minutes at 100V. The membrane was then blocked in 3% (w/v) milk in TBS/T (1xTBS with 0.1% Tween 20) or 5% (w/v) Bovine Serum Albumin (BSA) in TBS/T depending on the antibody used for 2 hours. The blocked membrane was incubated with the

primary antibody (Table 2.15) at 4°C overnight. Following four TBS/T washes (15 minutes for each wash), the membrane was incubated with a fluorescently-labelled secondary antibody (Goat anti-mouse IgG (H + L) AlexaFluor 680 or Goat anti-rabbit IgG (H + L) Dylight 800) for 1 hour at room temperature in dark. After four TBS/T washes (15 minutes for each wash) and a TBS wash (5 minutes), the membrane was imaged at 700 and/or 800 nm using the Odyssey CLx infrared imaging system (LI-COR). Western blot images were analysed using Image Studio Lite (Li-cor).

Table 2.15 Antibodies used for Western blot

Antibody information	Dilution & Buffer	Manufacturer	Catalogue no.
Mouse anti-mAGO2	1:4,000 in 3% milk/TBST	Provided by Dónal O'Carroll, unpublished	
Anti-CD9 antibody	1:2,000 in 5% BSA/TBST	Abcam	ab223052
Anti-β-Actin antibody	1:1,000 in 5% BSA/TBST	Cell signalling	4967S
Anti-Cre recombinase antibody	1:2,000 in 3% milk/TBST	Sigma	SAB2702246
Goat anti-mouse IgG (H + L) AlexaFluor 680	1:10,000 in 3% milk/TBST	Invitrogen	A-21058
Goat anti-rabbit IgG (H + L) Dylight 800	1:10,000 in 5% BSA/TBST	ThermoFisher Scientific	SA5-35571

2.10 Immunofluorescence imaging for Ago2 protein

The immunofluorescence staining was performed in Mode-K cells and purified gut microbiota using Alexa Fluor™ 488 Tyramide SuperBoost™ Kit (Invitrogen™, Cat#: B40912) with some modifications.

2.10.1 Prepare cells (fixation and permeabilization)

For Mode-K cells:

Mode-K cells were seeded in 100 µl complete DMEM per well in a 12-well chamber slide and incubated at 37°C. When cells reached ~80% confluency, the culture medium was aspirated and cells were washed 3 times using 200 µl of PBS. Cells were fixed using 100 µl 4% (w/v in PBS) paraformaldehyde for 10 minutes at 37°C then washed 3 times using 200 µl of PBS. Cells were treated with 100 µl 0.1% (v/v in PBS) Triton X-100 for 15 minutes at room temperature and then washed 3 times using 200 µl of PBS.

For purified gut microbiota:

Purified gut microbiota was fixed in 400 µl of 4% paraformaldehyde for 1 hour at room temperature, then washed twice using 500 µl of PBS. The fixed cells were suspended in 500 µl of 50% (v/v in PBS) ethanol and incubated at -20˚C for at least 30 minutes.

After two times PBS washes, the cells were treated with 800 µl of the 1 mg/ml lysozyme solution for 15 minutes at 37˚C with vortexing and then washed twice with 400 µl PBS.

2.10.2 Peroxidase labelling with primary and secondary antibody

Cells were suspended in 100 µl of Hydrogen Peroxide Solution (Component C2) and incubated for 60 minutes at room temperature to quench the endogenous peroxidase activity. After three times PBS washes, cells were incubated with Blocking buffer (Component A) for 60 minutes at room temperature. Cells were then treated with primary antibody (mouse anti-mAgo2 antibody or mouse IgG) overnight at 4˚C. After three times washes with PBS for 10 minutes, cells were incubated with poly-HRP-conjugated secondary antibody for 60 minutes at room temperature. Cells were washed three times with PBS for 10 minutes.

2.10.3 Tyramide labelling and microscopy

Cells were suspended in 100 µl of the tyramide working solution (1% 100x Tyramide stock solution and 1% 100x H_2O_2 solution in 1x Reaction buffer) and incubated for 10 minutes at room temperature. Reaction Stop Reagent was added and incubated for 5 minutes at room temperature. Cells were washed three times with PBS. The cells were treated with 100 µl DAPI solution (300 nM) and incubated for 5 minutes at RT. The cells were then washed three times with PBS. Mode-K cells were imaged with EVOS Cell Imaging Systems M7000 with 10x objective and purified gut microbiota were imaged with Zeiss LSM880 Airyscan with a 63x oil objective.

2.11 CLEAR-CLIP

2.11.1 Conjugate antibody to Protein G beads

50 µl Dynabeads Protein G (Invitrogen™, Cat#:1004D) per immunoprecipitation (IP) were placed in a 1.5 ml tube on a magnet stand, and the beads' buffer was subsequently removed. The beads were washed five times with 1 ml of filtered binding wash buffer (BWB), consisting of 50 ml of PBS and 10 µl of Tween 20. The beads were re-suspended by gently tapping the tube while adding BWB. 10 µl of either mouse anti-mAGO2 antibody (1.15 mg/ml) or mouse IgG (1 mg/µl) per IP in 600 µl BWB was added to the beads and this mixture was placed on a spinning wheel at 4°C for at least 5 hours.

2.11.2 Gut tissue sampling and crosslinking

For fresh tissue:

The intestines were quickly dissected from the sacrificed mice, and then cut into small pieces. Approximately 120 mg of tissue was weighed on ice and triturated using a pestle in a 1.5 ml tube. The samples were suspended in 3 ml of cold PBS in 60 mm cold dishes. The samples were cross-linked through three rounds of 254 nm UV irradiation (400 mJ each) while kept on ice. The samples were mixed between each irradiation step. After crosslinking, the samples were scraped off into a 15 ml tube and spun down at 2000 xg for 5 minutes at 4°C. The pellets were suspended in 1.2 ml of ice-cold Lysis Buffer (50mM Tris/HCl pH7.5, 100mM NaCl, 1mM $MgCl_2$, 0.1 mM $CaCl_2$, 1% Igepal, 0.5 % sodium deoxycholate, 0.1% SDS) in a 2 ml tube.

For frozen tissue:

The intestines were quickly dissected from the sacrificed mice and snap-frozen using liquid nitrogen (the intestines were cut into pieces before freezing). The intestines were ground separately under liquid nitrogen using a pestle and mortar placed on dry ice. Approximately 120 mg of powdered tissues were spread thinly and evenly onto 100 mm cold dishes. The samples were cross-linked through three rounds of 254 nm UV irradiation (400 mJ each) on dry ice. Cells were scraped off and suspended in 1.2 ml of ice-cold Lysis Buffer in a 2 ml tube.

2.11.3 Lysis and bind sample to beads

1 ml 0.1mm zirconium grinding beads were added to the sample and homogenized immediately using the TissueLyser II, 30 Hz for 5 minutes, twice (after the first disruption, shake in a reverse direction). The sample was incubated on ice for 10 minutes. 30 µl RQ1 DNase/tube (Promega, Cat#: M6101) was added and incubated at 37°C, 1000 x g for 5 minutes. RNase treatment was performed by adding 20 µl of 1:500 dilution of RNase IT (Agilent, Cat#:400720) per lysate (1:25,000 dilution total) and incubating at 37°C, 1000 x g for 5 minutes. After adding 5 µl RNase inhibitor (Promega, Cat#: N2115) per tube (~0.2 U/µl), the sample was spun down at 16,100 x g, 4°C for 15 minutes. 60 µl per sample was kept as input. The conjugated beads were washed 3 times with 500 µl lysis buffer. The beads tube was placed on the magnet and the buffer was aspirated. The samples were added into each beads tube and put

on a spinning wheel at 4°C overnight. The samples were placed on the magnet and the unbound fraction was saved. The beads were washed twice with 1 ml of Lysis Buffer (3 minutes rotating at 4°C each), three times with 1 ml of High Salt Buffer (50 mM Tris-HCl pH 7.4, 1 M NaCl, 1 mM EDTA, 1% Igepal, 0.5% sodium deoxycholate, 0.1% SDS) (3 minutes rotating at 4 °C each), and twice with 1 ml of PNK+Tween Buffer (20 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 0.2% Tween-20).

2.11.4 On beads 5'end phosphorylation

The beads were spun down after taking off PNK+Tween Buffer to remove most buffer and then re-suspended in 80 µl of the reaction mix and incubate at 20°C for 2.5 hours with intermittent shaking. The composition of the reaction mix is listed in Table 2.16. The sample was washed 3 times in PNK+Tween buffer (1x the beads were directly resuspended in buffer, 1x with 3 minutes rotation at 4°C, 1x the beads were directly resuspended in buffer).

Table 2.16 Reaction mix for 5' end phosphorylation

Component	Volume per sample (µl)
Nuclease free water	58
PNK (3' phosphatase, NEB, Cat#: M0236S)	8.0
10x PNK buffer (from kit)	4.0
10 mM ATP (NEB, Cat#: P0756S)	8
RNasin (Promega, Cat#: N2111)	2.0
Total volume:	80.0

2.11.5 Intermolecular RNA-RNA ligation

The beads were spun down after taking off PNK+Tween Buffer to remove most buffer and then re-suspended in 100 µl ligation mix per sample. The composition of the ligation mix was listed in Table 2.17 (Prepare mix without PEG-8000, add mix to tube and then add PEG-8000 separately to every tube). Reaction mixes were incubated overnight at 16°C with intermediate shaking. The next morning 2.5 µl T4 RNA ligase I and 10 µl of rATP (10 mM) were added to each sample and ligated for another 5 hours. The samples were washed two times with lysis buffer followed by one wash in PNK/EDTA/EGTA Buffer (50 mM Tris-HCl pH 7.4, 10 mM EDTA, 10 mM EGTA, 0.5% Igepal) for 3 minutes rotating at 4 °C and two PNK+Tween Buffer washes.

Table 2.17 Reaction mix for intermolecular ligation

Component	Volume per sample (µl)
Nuclease free water	41.25
10x ligase buffer (NEB, Cat#:	10
10 mM ATP (NEBy, Cat#: P0756S)	10
BSA 10mg/ml	1
RNasin (Promega, Cat#: N2111)	1.5
T4 RNA ligase I (NEB, Cat#:	6.25
PEG-8000 (50%) (NEB, Cat#:	30
Total volume:	100

2.11.6 Dephosphorylation of RNA

The beads were spun down after taking off PNK+Tween Buffer to remove the remaining buffer and then re-suspended in 80 µl TSAP reaction mix per sample. The composition of the reaction mix was listed in Table 2.18. The sample was incubated at 45 minutes, 20°C with intermediate shaking. The sample was washed with PNK+Tween Buffer twice (1x re-suspend, 1x 3 minutes with rotation at 4°C).

Table 2.18 Reaction mix for TSAP dephosphorylation

Component	Volume per sample (µl)
Nuclease free water	62
10x TSAP buffer	8
RNasin (Promega, Cat#: N2111)	2
TSAP (Promega, Cat#: M9910)	8
Total volume:	80

2.11.7 3’ end adaptor ligation on beads (in the dark)

The beads were spun down after taking off PNK+Tween Buffer to remove the remaining buffer and then re-suspended in 80 µl reaction mix per sample. The composition of the reaction mix was listed in Table 2.19. The sample was incubated overnight at 16°C with intermediate shaking. The sample was washed with PNK+Tween Buffer twice.

Table 2.19 Reaction mix for 3' adapter ligation

Component	Volume per sample (µl)
Nuclease free water	42
10x ligase buffer (TriLink, Cat#:	8
RNasin (Promega, Cat#: N2111)	2
3' Clean Tag-IR800 (10 µM)	8
T4 RNA Ligase 2, truncated	4
PEG 8000 (50%)	16
Total Volume:	80

2.11.8 Elution from beads and SDS-PAGE purification of the Ago2-RNA complex

The beads were spun down briefly to take off the residual wash buffer and re-suspended in 50 µl elution buffer (15 µl 4x NuPAGE LDS sample buffer (Invitrogen™, Cat#: NP0007), 45 µl PNK buffer, 3 µl 1 M DTT). The samples were incubated at 70 °C with constant shaking (1,000 rpm) for 10 minutes. The eluate containing Ago2-RNA complexes were separated from the beads on the magnet and run on a 12-well 4-12% NuPAGE Bis-Tris protein gel (20 µl per lane, 2 lanes per sample) at 120 V for 150 minutes in MOPS SDS running buffer (Invitrogen™, Cat#: NP0001). Gels were run on ice in the dark to prevent RNA degradation and protect the 3' adapter fluorophore. The gels were put onto the overhead-projector films and visualised using the Odyssey CLx scanner (Li-cor). The regions corresponding to 115-185 kDa were excised from the gel using a scalpel and transferred into 0.5 ml tubes with hole in it (pierce hole in with G20 needle). Gel pieces were flash frozen on dry ice for 5 minutes and spun through the hole into a 2 ml tube at 5,900 x g for 5 minutes.

2.11.9 RNA extraction and library preparation for sequencing

The sample were treated with 10 µl proteinase K (20 mg/ml) and 1 µl RNase inhibitor in 400 µl Proteinase K buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.1% TritonX-100, 10 mM Imidazole, 1% SDS, 5 mM EDTA) for 2 hours at 55°C with 1,000 rpm constant shaking. An equal volume of Acid-phenol:chloroform (Ambion) was added to the sample to extract the RNA and the tube was shaken for 15 seconds. After centrifugation at 16,100 x g, 4 °C for 15 minutes, the upper aqueous phase (~200-250 µl) was transferred into a new 1.5 ml tube. RNA was precipitated overnight at -80°C using 2.5 volumes of 100% Ethanol, 1/10 volume Sodium Acetate (3 M, pH 5.5) and

1 µl GlycoBlue Coprecipitant (Invitrogen™, Cat#: AM9515). Samples were then pelleted at 16,100 x g, 4 °C for 30 minutes and precipitated RNA was washed with 500 µl of 70% ethanol, air dried on ice for 15 minutes and re-suspended in 2.5 µl RNase-free H_2O. sRNA library was prepared using the TriLink CleanTag Small RNA library kit (TriLink, Cat#: L-3206) with the following modifications. As the 3' adapter was already ligated, 2.5 µl of buffer 1 from the kit was added to the RNA and library preparation was continued at the 5' adapter ligation step. The 5' adapter was diluted 1:10 in RNase-free H_2O, PCR amplification were performed using 20 PCR cycles and a size range of 160-240 bp was purified from the TBE gel (to avoid the miRNA only band). The libraries were sequenced using a single-end 100 base pair sequencing run on the Illumina NextSeq 2000 platform by the Edinburgh Clinical Research Facility.

2.12 Isolation and characterization of EVs

2.12.1 EVs isolation

Mode-K cells were grown in complete DMEM medium (described in 2.1) to 70-80% confluency in T-150 cm^2 flasks (~7 flasks, 30 ml medium per flask). Then the growth media was removed, cells were washed with PBS twice and shifted to serum-free advanced DMEM medium (Gibco™, Cat#: 12491015), 30 ml medium per flask. After 24 hours incubation, conditioned media (~200 ml) was collected, cells and cellular debris were removed by centrifugation at 200 x g for 10 minutes 4°C, supernatant decanted and centrifuged at 2000 x g for 20 minutes 4°C to remove dead cells/apopotic bodies. Supernatants were further filtered through 0.22 µm Millex GP filter units (Millipore, Cat#: SLGPR33RB) and subsequently concentrated using Vivaspin 20 (5 kDa MWCO Polyethersulfone, Cytiva) to ~12 ml. The concentrated medium was transferred into a Beckman polyallomer centrifuge tube (Beckman, Cat#: 331374) and ultracentrifuged at 100,000 x g, 90 minutes, 4°C using a Beckman ultracentrifuge with a SW40 swing out rotor. Pelleted material was then washed with 0.22 µm filtered PBS twice. After the second wash, pellets were resuspended in 200 µl PBS and transferred to the protein LoBind tube (Eppendorf, Cat#: 0030108094). EVs from *Heligmosmoides bakeri* were collected from the conditioned media of *H. bakeri* where they release their excretory/secretory products known as "HES" (Johnston et al, 2015) with the same ultracentrifugation method. EVs were stored at -80°C until required.

2.12.2 Characterization of EVs

The total protein content of EVs was measured using the Qubit Protein Assay kit (Thermo, Cat#: Q33211). Western Blot analysis of EV-associated marker CD9 and the ubiquitous eukaryotic protein β-actin was performed as 2.9. The primary antibodies are Anti-CD9 antibody (Abcam, Cat#: ab92726) and Anti-β-Actin antibody (Cell signalling, Cat#: 4967S).

The particle concentration and size of EVs was determined using ZetaView® according to the manufacturer's protocol. Briefly, the flow cell was rinsed with 0.2 µm filtered dH_2O after initialisation of the machine then loaded with 1 ml 100 nm polystyrene standard beads at a dilution of 1:250,000. The alignment was performed using the diluted beads using scatter mode on the 488 laser. Prior to the measurement of EVs samples, the flow cell was rinsed with 0.2 µm filtered PBS. EVs samples were diluted in 0.2 µm filtered PBS with appropriate dilutions (the number of visible particles on the Zetaview® was within the acceptable range) and measured using scatter mode on the 488 laser with the standardised settings of sensitivity -85, shutter -, and 3 repeated measurement of 11 different positions were taken ensuring >1000 particles were tracked. For visualization of the vesicles, EVs were fixed in 2% paraformaldehyde and TEM imaging was performed using Electron Microscope JEOL TEM-1400 Plus TEM.

2.13 EVs labelling and uptake assays

2.13.1 Memglow labelling and fluorescent particles measurement

Memglow640 stock solution (Universal Biologicals, Cat#: MG04-10) was thawed on ice before use. Memglow640 was suspended in 200 µl PBS and EVs were diluted in 300 µl PBS. Memglow640 and EVs were mixed immediately by pipetting up and down and incubated for 30 minutes at room temperature covered from light. The labelled EVs were suspended in 12 ml PBS and ultracentrifuged at 100,000 x g, 90 minutes, 4°C to remove the free dye. Finally labelled EVs were suspended in 200 µl PBS.

The concentration of fluorescent particles post-labelling was assessed with ZetaView®. Initially, a specific laser (488 nm or 640/660 nm) and the corresponding filter were chosen for measurements. To optimize focus, standardized DR660 fluorescent standard beads (as recommended in the Zetaview® user guide) were used,

and a measurement of the beads was taken. Subsequently, the fluorescently labelled EVs were loaded into the chamber, and measurements were conducted using established settings of sensitivity = 95, trace length = 7, and shutter = 100-150, with the low bleaching mode enabled. To determine labelling efficiency, the same EVs sample underwent measurement in both scatter and fluorescent modes. Depending on label efficiency, certain samples required different dilutions for scatter and fluorescence measurements due to the machine's measurement range.

2.13.2 EV uptake assays

SL1344 and *E. coli* W3110 were cultured in LB or M9 medium at 37°C to log phase (4~6 hours, OD600=0.5). Following this, 150 µl of bacterial cultures were treated with labelled Mode-K EVs, *H. bakeri* EVs, or PBS at a concentration of $1x10^9$ fluorescent particles/ml and incubated at 37°C for 6 hours. The bacterial cells were washed three times with 400 µl PBS to remove any free-floating EVs. Afterward, the cells were fixed in 4% (w/v in PBS) paraformaldehyde (400 µl) for 1 hour at room temperature, followed by two washes with 400 µl PBS. Finally, the bacteria were suspended in 20 µl PBS and microscopy was carried out as detailed in Section 2.7.4.

2.14 Growth measurement of bacteria

Bacteria was cultured in 3 ml liquid LB or M9 medium with bacteria colony from a fresh plate and grown aerobically to log phase (4~6 hours, OD600=0.5) at 37ºC with 250 rpm shaking. These cultures were used to inoculate aliquots of medium (1/50), with 3 µl inoculum volumes of 150 µl per well on a 96-well plate. EVs were added to the culture at the specified concentration, while an equal volume of PBS was included in the control group. Three replicates for each treatment. Growth was continuously monitored by measuring absorbance at 600 nm (OD 600) hourly for up to 16 hours using Tecan Infinite M200 Plate Reader.

2.15 Plasmid transformation for bacteria

2.15.1 Preparing competent cells

The recipient strains (SL1344, *E. coli* W3110) were grown overnight in 5 ml LB at 37°C with the appropriate antibiotic. 2 ml of the overnight cultures were sub-cultured into 25 ml LB broth (x2) and grown until OD600 reached 0.4-0.6 at 37°C with shaking (typically 4 hours). Once ready, the cultures were placed on ice for 5 minutes, then centrifuged

at 2,300 x g for 10 minutes at 4°C in 50 ml tubes. 5 ml ice-cold sterile distilled water was added to each pellet and mixed well by tapping the base of the tube. 25 ml with ice-cold sterile distilled water was used to top up the suspension and samples were centrifuged at 2,300 x g for 5 minutes at 4°C. The same washing step using ice-cold sterile distilled water were repeated three times. The pellet was resuspended in 3 ml of ice-cold sterile 10% glycerol, then centrifuged at 2,300 x g for 10 minutes at 4°C. Finally, the cells were resuspended in 1x volume of ice-cold sterile 10% glycerol (typically 250 µl) and kept on ice until electroporation.

2.15.2 Electroporation

70 µl volumes of electrocompetent cells and varying amounts of plasmid DNA (ranging from 0.1 to 2 µg) were added to a 1.5 ml Eppendorf tube and gently mixed by tapping. The cell-plasmid mixture was transferred to ice-cold 0.1 cm cuvettes and kept on ice. The outside of the cuvettes was wiped thoroughly and the electroporation step was performed at 1.8 kV, 25 µF, and 200 Ω. Immediately following electroporation, 450 µl pre-warmed SOC Broth was added to the cuvette and mixed well. Then the mixture was transferred into a 1.5 ml Eppendorf tube and incubated at 37°C for 2 hours with gentle shaking at 250 rpm. After the incubation, 100 µl of the culture was spread onto LB agar plates containing the appropriate antibiotics and incubated overnight at 37°C.

2.16 Plasmid transfection for Mode-K cells

Mode-K cells were seeded in a 6-well plate using 1 ml of DMEM medium. Transfection was carried out when cells reached 70-80% confluency using Lipofectamine 2000 (Invitrogen™, Cat#: 11668027). Three replicates were used for each transfection, and a control utilizing only Lipofectamine 2000 without a plasmid was included. pcDNA3.1-CMV-CFP;UBC-Cre25nt was a gift from Jacco van Rheenen (Addgene plasmid # 65727 ; http://n2t.net/addgene:65727 ; RRID:Addgene_65727). For each well, 4 µg of plasmid DNA was diluted in 250 µl of Opti-MEM (Gibco™, Cat#: 11058021). In control wells, an equal volume of sterile water was used instead of plasmid DNA. Additionally, 10 µl of Lipofectamine 2000 was diluted in 250 µl of Opti-MEM, followed by incubation of the mixture for 5 minutes at RT. The diluted DNA and Lipofectamine 2000 were combined and further incubated for 20 minutes at RT. Growth media were removed from the wells, and a PBS wash was performed. Subsequently, the mixture, totalling 1 ml Opti-MEM, was drop-wise applied to the cells. After a 6-hour incubation period,

the cells were visually inspected under a dissecting microscope, indicating good tolerance to the transfection and Opti-MEM. Then, 2 ml of complete DMEM was added to each well, and cells were incubated overnight. The following day, the transfected cells were reseeded into a T75 flask using 20 ml DMEM, and 150 µg/ml Zeocin (Thermo Scientific Chemicals, Cat#: J67140.8EQ) was introduced for selection. Untransfected cells and those that have not stably integrated the construct into their genome were expected to perish during Zeocin selection, allowing only cells containing the plasmid construct to survive and proliferate (this process may take up to 2-3 weeks).

2.17 Mouse serum and serum EVs collection

The blood collection from C57BL/6J mice and Villin-cre transgenic mice was carried out by Ms. Elaine Robertson. Upon collection, the whole blood was allowed to clot by keeping it undisturbed at room temperature for 2 hours. Afterward, it was centrifuged at 1,000 x g for 15 minutes. The resulting supernatant was transferred to a new 1.5 ml Eppendorf tube and centrifuged again at 2,300 x g for 15 minutes to obtain cell-free serum. Cell-free serum was diluted with PBS to a total volume of 12.5 ml and subsequently subjected to ultracentrifugation to pellet the EVs, following the procedure outlined in Section 2.12.1.

2.18 Bioinformatic analysis

Bioinformatic analysis was performed by Dr Jose Roberto Bermudez-Barrientos with valuable input from Dr Cei Abreu-Goodger and Dr Sujai Kumar.

2.18.1 Small RNA sequencing analysis

The raw sequencing data were checked using FastQC (v0.11.9) for quality issues (Andrews, 2010). Then the 3’ adapter sequence was trimmed using reaper (Davis et al, 2013), and reads smaller than 18 nt in size after trimming were removed. Reads between 18-60 nt were selected and collapsed into unique reads with their counts using tally (Davis et al, 2013). Count matrices were prepared across all samples, as well as plots showing the length distribution and first-nucleotide preference, using ad hoc R scripts (R Core Team, 2021).

For analysis of miRNA expression, miRNAs were re-identified from the raw read sequences using 16 sSeq (Zhao et al., 2017) with standard settings. The

QuickMIRSeq output file with miRNA counts was then uploaded on Degust web tool (Version 4.1.3) (Powell et al., 2019) for differential expression analysis and data visualisation. Differential expression analysis on Degust was executed using the Voom/Lima method, and data were analysed using a false discovery rate cut-off of 0.05.

2.18.2 CLEAR-CLIP analysis

The fastq files were adapter trimmed with cutadapt and then converted to a uniq fasta format. Bowtie2 version 2.4.4 (Langmead et al., 2012) were used to map reads to the mouse mature miRNA fasta file obtained from miRBase version 22.1 (Kozomara et al., 2019) and the reads where the first part matches a mature miRNA and the remaining part is longer than 18 nucleotides were selected. These second part of these reads was stored to a temporary fasta file called btmpXXXXXXX (where bXXXXX is randomly generated). In a second step, bowtie2 in local alignment mode was used to determine which section of the second part of the putative chimeric reads matched the target genome and whether it was low-complexity. These results were stored in two temporary files b2tmpXXXXXXX and b3tmpXXXXXXX respectively. In the following step we created a multifasta from the unique fasta and run ShortStack v 3.8.5 (Johnson et al., 2016). ShortStack helped us to deal with reads that have multiple mapping sites in the target genome. In a following step in the pipeline, we made a bam file whose read's name contains the id of the miRNA that the first part hit. Later we merged the information of all overlapping chimeric reads that shared the same miRNA with bedtools version 2.30.0 with its merge functionality (Quinlan et al., 2010) resulting in chimeric read clusters. Then we used this merged bed file and count the number of overlapping reads in each of the library bam files to get a sample counts tab separated file.

To be able to tell which genomic features overlap with our chimeric read clusters, we need annotations for our target genome. When assessing mouse targeting chimeras, we used Gencode mouse version M32 (Frankish et al., 2019) annotations (CDS, 3' UTR, 5' UTR, miRNA, tRNA, lncRNA and rRNA) and Ensembl mouse release 109 (Martin et al., 2023) for regulatory regions (CTCF binding site, enhancer, open chromatin, promoter, and transcription factor binding site). For the gut bacterial targeting chimeras, we used the mouse gut metagenome described by Xiao and

collaborators (Xiao et al., 2015) which is a catalogue of ~2.6 million non-redundant genes by sequencing DNA from faecal samples of 184 mice. We annotated RNA features (rRNA, tRNAs, etc.) in the mentioned bacterial gene catalog using Infernal tool suite v1.1.4 (Nawrocki et al., 2013) with the command cmsearch with the trusted threshold cutoff (--cut_tc) and Rfam database v 14.8 (Kalvari et al., 2021). We used bedtools intersect to annotate chimeric reads cluster. A column for each available annotation was added including the gene symbol of the overlapping gene or the feature id when a gene symbol was not available for a given genomic feature such as regulatory regions. This step was done separately for Gencode features and regulatory regions since the later ones did not have a strand and this was considered in the bedtools intersect command.

To find miRNA seeds in the target portion of the chimeric reads, we first extracted the actual miRNA sequence and the target region from the genome. This was achieved with bedtools getfasta. The seed finding search was made with a perl script developed in the lab by Dr Sujai Kumar. Finally, all the information gathered in the above-mentioned steps was combined in a single table within R 4.2.2, using the packages dplyr 1.1.2 (Wickham et al., 2023) and GenomicRanges (Lawrence et al., 2023).

Non-chimeric reads were defined as those reads that were not present in the btmpXXXXXXX file of putative chimeras mentioned above which means reads that do not have a mouse miRNA in the left portion and at least contiguous 18 nt extra nucleotides. Non-chimeric reads were aligned with STAR 2.7.10b (Dobin et al., 2013) using the parameters: outMultimapperOrder=Random, outFilterMultimapNmax=20, to both previously mentioned references Gencode mouse genome and to the mouse gut metagenome gene catalog. Then either Gencode or the mouse gut metagenome annotations were used to summarize the mapping information from the resulting bam files.

2.18.3 miRNA target prediction

The genome of *Salmonella enterica subsp. enterica serovar* typhimurium str. LT2 (GenBank: AE006468.2) was obtained from NCBI genomes. The query miRNA database was all mature mouse sequences, encompassing both 5p and 3p, retrieved from miRBase version 22.1 (https://mirbase.org/). Target prediction software TargetFinder was employed to predict the miRNA target, allowing a maximum penalty

score (-c) of 6, scanning for potential target sites in the sense DNA strand (-r), and producing the output in table format (-p table). After getting TargetFinder results, overlapping annotations were included such as CDS, minus_35_signal, minus_10_signal, ribosome_entry_site, protein_binding_site, tRNA, rRNA, ncRNA, and pseudogene from the associated general feature file (GFF) within R v4.2.2 using the packages rtracklayer v1.58.0 and GenomicRanges v1.50.2. The seed information was inferred from TargetFinder ASCII binding representation for 6mer, 7mer and 8mer.

Chapter 3: Identification of transferred host miRNAs in purified gut microbiota

3.1 Introduction

Increasing evidence supports the extracellular existence of miRNAs in the gut, and several reports suggest miRNAs can mediate host-bacteria communication. However, there remains a lack of systematic studies elucidating the specific miRNAs involved in host-bacteria interactions *in vivo*. Previous researches focus on identifying miRNAs within faecal samples to study miRNA-bacteria interaction. For example, Liu et al., (2016) for the first time performed a comprehensive profiling of mouse and human faecal miRNAs using the NanoString nCounter platform in the normal gut and revealed their essential role in maintaining normal gut microbiota using IEC-miRNA-deficient mice. In their study, they showed that two mouse miRNAs were functional in bacteria based on culturing bacteria with synthetic miRNA mimics. Specifically, they showed that miR-515-5p entered and promoted the growth of *Fusobacterium nucleatum* and miR-1226-5p entered and promoted the growth of *E. coli*. Another study showed that host faecal miRNA levels are differentially and specifically impacted by microbiota composition in colitic $IL10^{-/-}$ and colitic $TLR5^{-/-}$ mice and could indicate the colitogenic potential of the gut microbiota (Viennois et al., 2019).

It is important to note that faecal samples contain a complex mixture of components, including dead host cells, microorganisms, extracellular vesicles, undigested nutrients, and various other constituents. Therefore, the miRNAs detected in faecal samples represent a combination of miRNAs that could be present in different contexts (host cells, extracellular vesicles, and bacteria). Consequently, it is challenging to draw direct conclusions about miRNAs specifically present inside bacteria based solely on faecal miRNA analysis. To gain a more comprehensive understanding of the scope for RNA communication between host cells and gut microbiota under physiological conditions, it is crucial to systematically profile the total transferred host miRNAs in gut microbiota. This necessitates the purification of gut microbiota to remove any contaminants from mouse cells and extracellular vesicles that may interfere with the analysis.

The same type of strategy has been used to identify sRNAs that can traffic between fungal and host cells. A sequential protoplast purification method was developed to isolate pure fungal cells from infected *Arabidopsis* tissues based on the different cell wall compositions of plants and fungal cells (Cai et al., 2018). The authors identified 42 *Arabidopsis* sRNAs in purified *B. cinereal* protoplasts by sRNA sequencing and showed that some transferred host sRNAs were capable of silencing fungal virulence genes. By leveraging the differences in cell surface structures between eukaryotic cells and bacterial cells, some studies have been able to use selective lysis with nonionic surfactants like Triton X-100, saponin, and Tween 20 to obtain pure bacterial cells or enrich bacterial nucleic acids from various sample types, including blood, cerebrospinal fluid, nasopharyngeal aspirate, and arthroplasty fluids (Brennecke et al., 2017; Hasan et al., 2016; Thoendel et al., 2016). Triton X-100, in particular, has shown effectiveness as a detergent for the selective lysis of human cells (Brennecke et al., 2017; Hasan et al., 2016). This approach holds promise for my research to remove mouse cell contaminants and isolate pure gut microbiota from total gut contents, allowing for subsequent identification of host miRNAs within the bacterial community by sRNA sequencing.

In previous studies, imaging techniques have been employed to provide direct evidence of host miRNA uptake by gut bacteria *in vitro*. For instance, synthetic fluorescence-conjugated miRNAs were cultured with bacteria and found to co-localize with bacterial nucleic acids, implying the uptake of miRNAs by bacteria (Liu et al., 2016). However, this method relies on synthetically labelled RNAs and is not suitable for visualizing the natural transmission of miRNAs to bacteria in the mouse gut. Rather, fluorescence in situ hybridization (FISH) is a powerful molecular method to visualize target DNA or RNA molecules, where the probes are designed to be complementary to the sequence of interest (Young et al., 2020). However, considering the short length and diverse GC content of miRNAs, it is challenging to have sufficient sensitivity and specificity with conventional RNA or DNA probes (Yaylak and Bünyamin, 2022). To overcome these difficulties, locked nucleic acid (LNA)-modified oligonucleotide probes have been developed. LNA is a class of RNA analog characterized by a locked ribose ring structure, maintaining an ideal conformation for Watson-Crick binding (Kauppinen et al., 2005). This structural modification significantly enhances the thermal stability when they hybridize to complementary DNA or RNA strands, enabling higher affinity

and specificity in base pairing interactions. miRNA-FISH with LNA probes has become a visual detection method for imaging the spatial localization of miRNAs at the tissue, cellular and even subcellular level (Silahtaroglu et al., 2007; Lu and Andrew, 2009). To confirm the presence of host miRNAs within gut bacteria, I aim to conduct miRNA FISH with LNA probes tailored to detect specific miRNAs within bacterial cells. This methodology can offer direct evidence of miRNA uptake by gut bacteria under physiological conditions, which is a key first step in testing the hypothesis that miRNAs could naturally impact bacterial gene expression.

The primary goal of this chapter is to establish a method to purify gut microbiota from mouse gut contents and subsequently identify host miRNAs that are transferred to these microbial populations. Through this work, I aim to overcome the limitations of previous faecal studies that do not demonstrate association of miRNAs specifically with bacteria.

3.2 Results

3.2.1 Purifying gut microbiota from mouse gut content

One main aim of this chapter is to develop a purification method to isolate pure gut microbiota from mouse gut lumen contents such that sRNA sequencing could be carried out. The method involves several steps as described in Figure 3.1. The whole intestine was dissected from the sacrificed mouse, and the luminal contents were squeezed out with tweezers and suspended in PBS. The gut contents in PBS were filtered through a 70 μm nylon mesh to remove large particles and gut debris. The sample were then treated with 1% Triton X-100 to selectively lyse mouse cells and EVs. This step removes dead host cells and luminal miRNAs that could interfere with the analysis of host miRNAs in bacteria. Any remaining extracellular DNAs and RNAs were further removed using proteinase K, TURBO™ DNase, RNase A, and lysozyme treatment. These steps ensure the elimination of extracellular genetic material that may confound the analysis of host miRNAs in bacteria. Between each step, the bacterial pellet is washed three times with PBS. These washing steps help ensure removal of miRNAs in extracellular forms and further eliminate any remaining impurities.

To assess the extent to which mouse cells were removed during the purification process, the mouse 18S rDNA gene (which has hundreds of copies in the genome) and the housekeeping gene eEF2 were measured from DNA by PCR. The absence of amplification of these genes is used to indicate successful removal of mouse cells. By comparing the miRNA profiles between the Input (total gut contents isolated) and PGM (purified gut microbiota), host miRNAs that are abundant or enriched in the gut microbiota can be identified. Additionally, faecal samples from the same mice are used to compare differences in the Input, PGM and Faeces in terms of the bacterial populations present and miRNAs that can be detected.

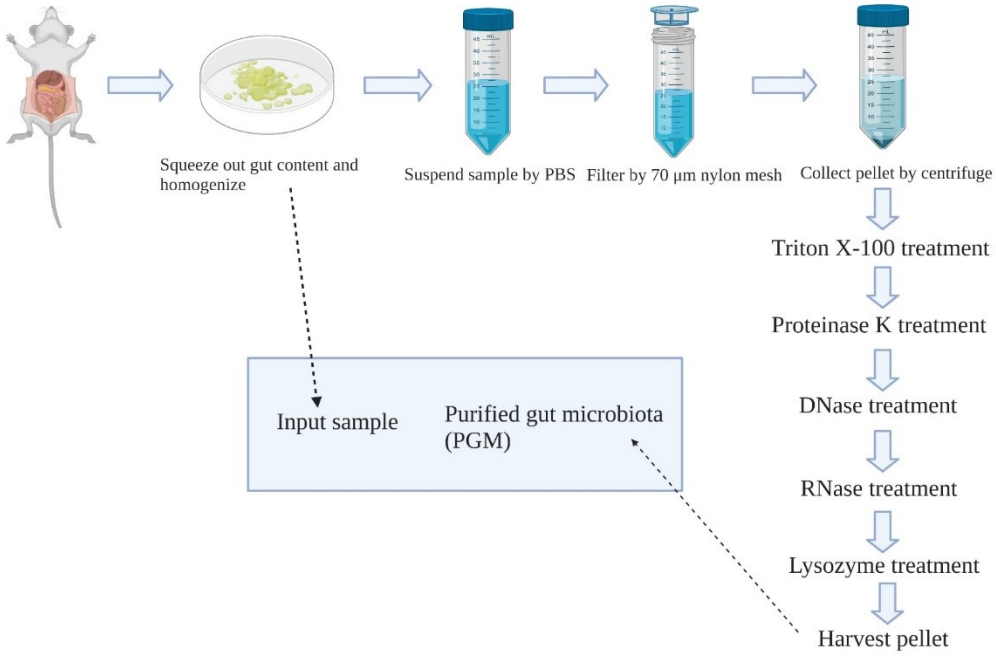


Figure 3.1 Schematic overview of the method for selective removal of mouse cells and other sources of extracellular miRNAs from bacteria. Triton X-100 is applied to selectively lyse mouse cells and extracellular vesicles.

Genomic DNA from Input, PGM, and Faeces was extracted using the QIAamp Fast DNA Stool Mini Kit. The DNA concentration of each sample was measured using the NanoDrop and showed in Table 3.1. PCR results (Figure 3.2) demonstrated that the mouse 18S rDNA and eEF2 genes were detected in the Input samples (Input_A, Input_B, and Input_C) and faecal samples (Faeces_A, Faeces_B, and Faeces_C). However, these genes were not detected in the PGM samples (PGM_A, PGM_B, and PGM_C). The bacterial 16S rDNA gene was detected in all samples, including the Input, Faecal, and PGM samples, with a strong signal. The absence of mouse-specific genes (18S rDNA and eEF2) in the PGM samples and the presence of bacterial 16S rDNA in all samples collectively suggest that the purification method effectively removed mouse cell contamination from the PGM samples and keep bacteria intact.

Table 3.1 DNA concentration of each sample measured by Nanodrop

Sample	DNA con. (ng/µl)	Sample	DNA con. (ng/µl)	Sample	DNA con. (ng/µl)
Input_A	244.8	Faeces_A	160.4	PGM_A	128.0
Input_B	185.9	Faeces_B	190.2	PGM_B	113.3
Input_C	200.1	Faeces_C	155.7	PGM_C	107.2

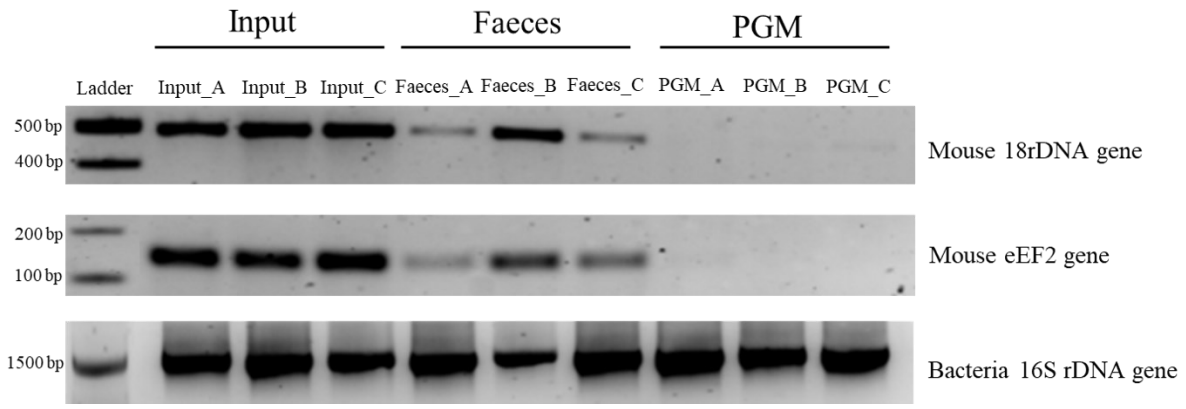


Figure 3.2 PCR results of mouse 18rDNA gene, eEF2 gene and bacteria 16S rDNA gene in Input, Faeces, and PGM. PCR products (20%, v/v) were run on 1-3% agarose gels in TAE buffer, and visualised using a Licor D-DiGit gel scanner.

3.2.2 Bacterial 16S rDNA sequence-based survey of PGM

One complexity of the above protocol is the possibility that not all members of the gut microbiota survive the purification process. For example, some conditions including exposure to cell wall-degrading enzymes, nutrient limitations, antibodies, and the influence of bacteriophages have been identified as potential factors leading to the generation of cell wall-deficient bacteria (Lazenby et al., 2022). To assess the changes in microbiome composition after the purification, a bacterial 16S rDNA sequence-based survey was conducted for the Input (Input_A, Input_B, and Input_C) and PGM (PGM_A, PGM_B, and PGM_C), with Faeces (Faeces_A, Faeces_B, and Faeces_C) included for comparison. The V4 region of the 16S rRNA gene was sequenced using an Illumina MiSeq platform, and the resulting sequencing data were analyzed using the QIIME2 software following an established protocol.

Over 72,000 sequences were generated for each sample (Table 3.2) and DADA2 (Callahan et al., 2016) on QIIME 2 was used to filter the reads. Table 3.2 shows that > 90% of reads passed the initial input filter and 83-87% of reads were merged. Unique sequences were classified and grouped into 413 OTUs based on 97% nucleotide sequence identity. The taxonomic classification of the unique sequences reveals the average number of bacterial taxa in the PGM is lower than that in the Input and Faeces (Figure 3.3A). This suggests that some bacteria did not survive the purification process and were depleted from the PGM. Principal Coordinate Analysis (PCoA) (Figure 3.3B) based on weighted UniFrac distances (A weighted UniFrac distance uses species abundance information and weights the branch length with abundance difference) was performed to assess the clustering patterns of bacterial communities. The results indicate that the bacterial communities of PGM form a separate cluster from those of the Input and Faeces. At the family level, it is observed that there is an increase in the abundance of *Lactobacillus* and *Lachnospiraceae* in the PGM, and a decrease in the abundance of *Muribaculaceae* compared to Input (Figure 3.3C). This suggests a shift in the relative abundance of specific bacterial taxa following the purification process.

Table 3.2 The number and percentage of reads filtered by DADA2

Samples	No. of input reads	No. of filtered reads	Percentage of input passed filter	Merged	Percentage of input merged
Input_A	204539	189202	92.5%	176405	86.25%
Input_B	102576	95342	92.95%	89649	87.4%
Input_C	121164	109326	90.23%	103194	85.17%
Faeces_A	150047	139347	92.87%	130293	86.83%
Faeces_B	141155	128427	90.98%	118391	83.87%
Faeces_C	72912	67611	92.73%	62970	86.36%
PGM_A	93484	86939	93%	79784	85.35%
PGM_B	117265	108306	92.36%	99885	85.18%
PGM_C	92554	85626	92.51%	80458	86.93%

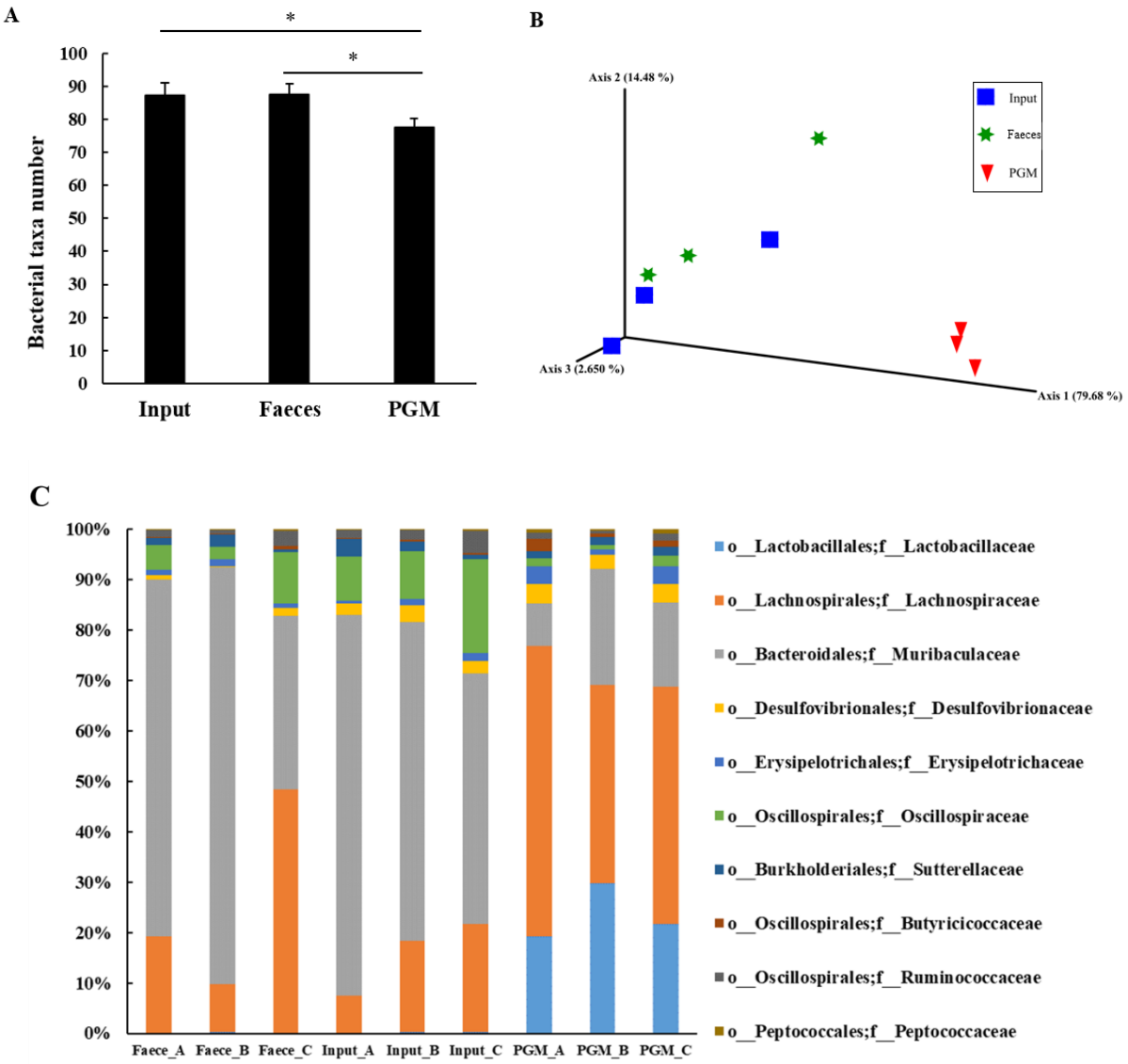


Figure 3.3 Microbiome composition of Input, Faeces and PGM. A) Bacterial taxa number detected in each sample. Three biological replicates for each type of sample. *p < 0.05 (two-tailed t test). B) Principal Coordinate Analysis of Weighted UniFrac Distances showing differences in bacterial community composition among samples. C) Relative abundance of bacteria classified at a family-level taxonomy. The top 10 abundant bacteria are shown and the order of taxa in the legend follows the order of taxa in the community bar (from bottom to top).

3.2.3 Identification of host miRNAs in PGM by sRNA sequencing

To determine whether host miRNAs could be identified in the PGM, I isolated total RNA from Input, Faeces and PGM using Norgen Stool Total RNA Purification Kit with 120mg of material per sample. The RNA concentration was measured with Qubit® RNA HS Assay Kits (Table 3.3). The same amount of RNAs were then used to generate sRNA libraries for sequencing using the CleanTag small RNA library prep kit. Prior to sequencing, the profiles of each sRNA library were evaluated using the High Sensitivity DNA Bioanalyser chip. The chip results (Figure 3.4) indicate that all the sRNA libraries have a main band around 150 bp. In order to obtain a similar coverage of sequencing depth in each library I pooled 30 ng of each sRNA library according to concentrations reported by the High Sensitivity DNA Bioanalyser chip analysis. The products between 140-180 bp (corresponding to the size of 20-60 bp sRNAs) were size selected and the libraries were sequenced using a single-end 100 base pair sequencing run on the Illumina NextSeq 2000 platform by the Edinburgh Clinical Research Facility. Sequencing yielded between 9 and 19 million high quality reads per library (Table 3.4).

Table 3.3 RNA concentration in each sample measured by Qubit

Sample	RNA con. ng/µl	Sample	RNA con. ng/µl	Sample	RNA con. ng/µl
Input_A	223	Faeces_A	122	PGM_A	142
Input_B	184	Faeces_B	144	PGM_B	146
Input_C	154	Faeces_C	150	PGM_C	113

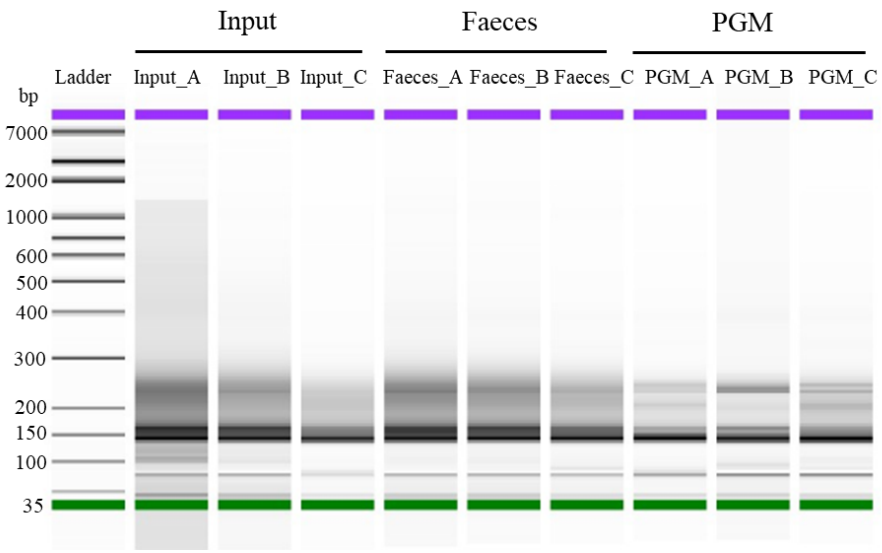


Figure 3.4 Bioanalyzer High Sensitivity DNA analysis of the sRNA libraries. Size in base pairs (bp) is indicated on the left of the gel image.

Table 3.4 The number of reads passing the high-quality score for each library.

Sample	No. of reads with the average quality score > 30*	Sample	No. of reads with the average quality score > 30*	Sample	No. of reads with the average quality score > 30*
Input_A	9,178,016	Faeces_A	12,864,373	PGM_A	14,317,408
Input_B	12,929,751	Faeces_B	12,651,484	PGM_B	9,570,553
Input_C	17,790,284	Faeces_C	18,679,859	PGM_C	11,052,389

* A quality score exceeding 30 implies that the likelihood of erroneously calling a base is less than 1 in 1000, ensuring an accuracy of over 99.9%.

The sequencing data was processed by Dr Jose R. Bermudez-Barrientos. The adapter was removed with cutadapt 3.4 and only sequences longer than 18 nt were kept. To compare the sRNA composition of these three types of samples (Input, Faeces and PGM), we conducted length distribution and first nucleotide analysis for these sRNA reads. Figure 3.5 shows that the sRNA patterns from Input and Faeces samples exhibit remarkable similarity, possibly indicating sRNAs originating from the mouse. However, the sRNA profile of PGM displays significant divergence from both Input and Faeces. There is a very clear peak at 29 nt with the majority beginning with uridine (U) in all PGM libraries. Table 3.5 lists the top 10 sRNA sequences identified in each PGM library, all of which have been annotated as bacterial ribosomal RNA or transfer RNA by blasting (megablast) against the NCBI Nucleotide database. Notably, a

significant proportion of the sRNAs in PGM are annotated as 5S ribosomal RNA, characterized by their 29-nucleotide length and initiation with the nucleotide U.

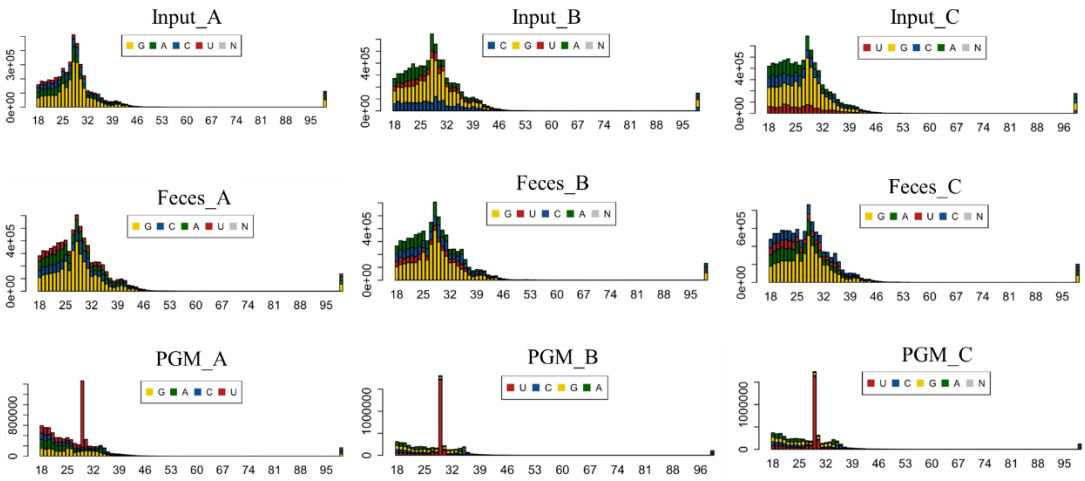


Figure 3.5 Length distribution and first nucleotide plots of the sRNAs in Input, Faeces, and PGM. The x-axis represents the read length in nucleotides and the y-axis shows the number of reads. U= uridine, A= adenine, C= cytosine, G=guanine.

Table 3.5 The top 10 most abundant sRNAs sequenced in the PGM libraries. Reads were annotated in the NCBI Nucleotide database.

sRNA reads	Length	Percentage of reads	Product description
TCAGAGACTGGTGGAGATATGAATCCAGA	29 bp	19.7%	Uncultured bacterium 5S ribosomal RNA
TCAGAGATTGGTAAAGACACGAATCCAGA	29 bp	18.3%	Uncultured bacterium 5S ribosomal RNA
TCAGGGATTGGTGGAGATATGAATCCAGA	29 bp	14.8%	Uncultured bacterium 5S ribosomal RNA
TGGAAGAGTGTCTTCCAGCGATCTTTGAAA	30 bp	4.1%	*Faecalibaculum rodentium* tRNA
ATCGGGCATAAGTCCGGTGCCGAAGACGTGGACT	34 bp	2.6%	*Duncaniella sp.* 23S ribosomal RNA
GAGCGGGAGAGTAGGTAATCGCCGCC	26 bp	2.2%	Uncultured bacterium 5S ribosomal RNA
GCAGATGTAGTTCAATGGTAGAACA	25 bp	2.1%	*Clostridium innocuum* tRNA-Gly
TGGAAGAGTGTCTTCCAGCGATCTTTG	27 bp	1.9%	*Faecalibaculum rodentium* tRNA
CGGGACGTAGCACAGCTTGGTAGTGCACTTGCATG	35 bp	1.7%	*Faecalibaculum rodentium* tRNA
TGGAAGAGTGTCTTCCAGCGATCTTTGA	28 bp	1.4%	*Faecalibaculum rodentium* tRNA

The number of reads mapping to miRNA loci is shown in Figure 3.6A. PGM libraries display the lowest counts mapping to miRNAs loci, followed by Faecal libraries, while Input libraries have the most counts mapping to miRNA loci. In total 352 distinct miRNAs were detected across all samples. A multidimensional scaling (MDS) plot was

generated to visualize the differences in miRNA composition between PGM, Input, and Faecal samples. The MDS plot (Figure 3.6B) illustrates distinct clustering of PGM samples towards the right bottom of the graph, while Input samples cluster towards the left bottom, and Faecal samples gather in the middle top region. This pattern suggests that PGM exhibits a unique miRNA profile compared to the Input and Faeces, implying an alteration in miRNA composition following the purification process. Differential expression analysis of miRNAs in PGM and Input was performed with Degust (Powell et al., 2019) using the Voom/Lima method. A false discovery rate (FDR) of less than 0.05 was used to identify significance in differential expression of miRNAs. A total of 57 miRNAs (FDR < 0.05) are differentially expressed in PGM compared to Input (Supplementary Table 3.1). Of these 24 are upregulated and 33 downregulated in PGM (Figure 3.6C). The top 20 most abundant miRNAs in the PGM samples are displayed in the Figure 3.7. Among the miRNAs analyzed here, mmu-miR-215-5p is found to be the most abundant miRNA in the Input samples, which could derive from mouse cells, extracellular miRNAs, and gut bacteria. mmu-miR-192-5p is identified as the most abundant miRNA in the Faecal samples. mmu-miR-21a-5p is the most abundant miRNA in PGM samples and seven miRNAs from the let-7 family are also among the top abundant miRNAs in PGM. Among the top 20 most abundant miRNAs, only mmu-miR-22-3p, mmu-miR-378a-3p and mmu-miR-10a-5p show enrichment in PGM with FDR < 0.05 (Figure 3.6 C).

These results suggest that host miRNAs can be detected within gut bacterial communities directly after isolating the bacteria from the gut environment. The identification of abundant or enriched host miRNAs in PGM samples provides high confidence miRNA candidates for further investigation of their transfer mechanisms and potential functions within the bacterial species.

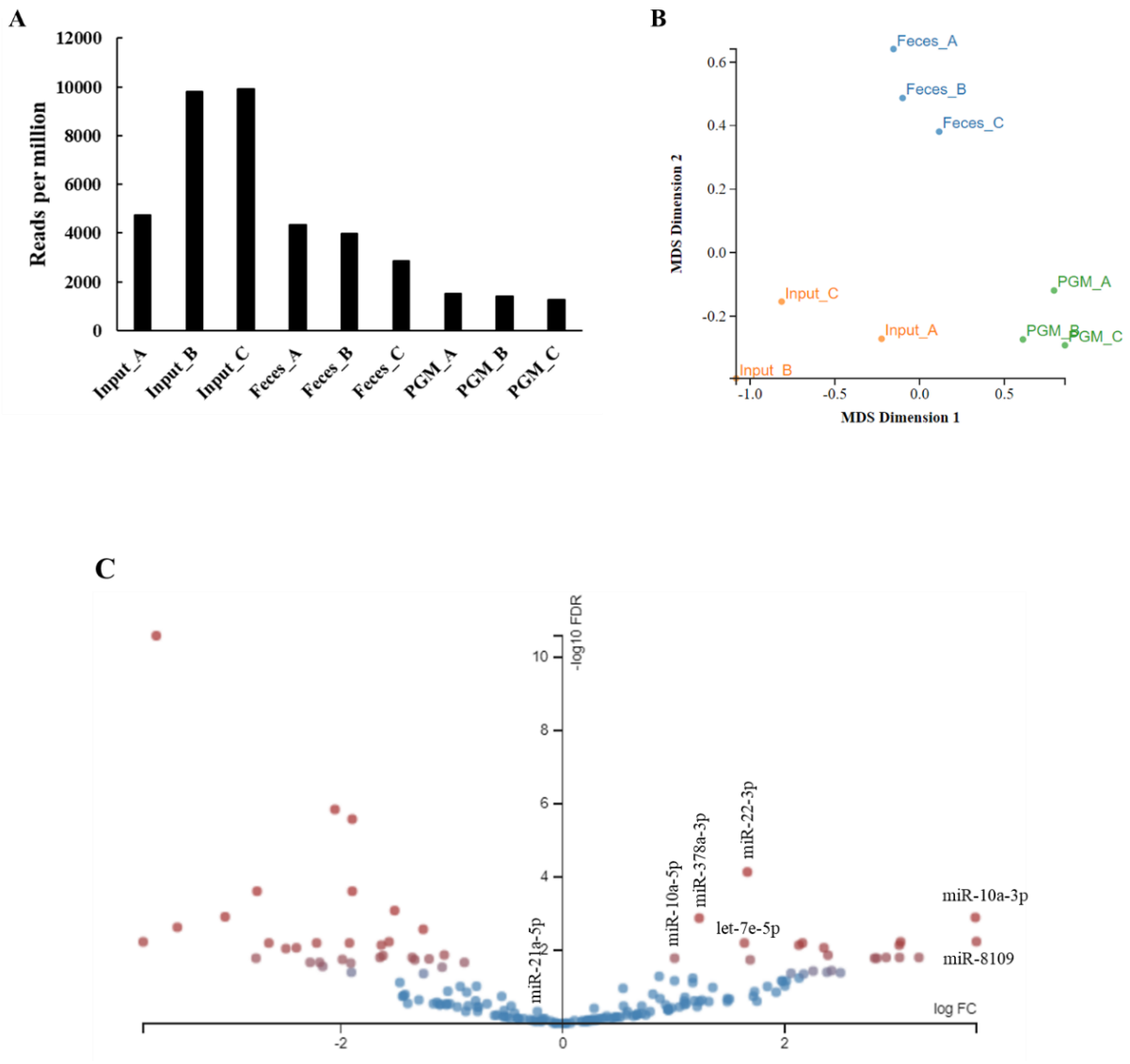


Figure 3.6 The composition of miRNAs in Input, Faeces, and PGM. A) The number of reads mapping to miRNA loci in Input (Input_A, Input_B, and Input_C), Faeces (Faeces_A, Faeces_B, and Faeces_C), and PGM (PGM_A, PGM_B, and PGM_C). B) Multidimensional scaling plot of miRNA contents in each library. C) Volcano plot of differentially expressed miRNAs in PGM compared to Input. Differential expression was analysed in Degust (Version 4.1.3) using Voom/Lima method. Fold change was plotted against -log10 false discovery rate (FDR). Red dots indicate FDR < 0.05 and blue dots indicate FDR > 0.05.

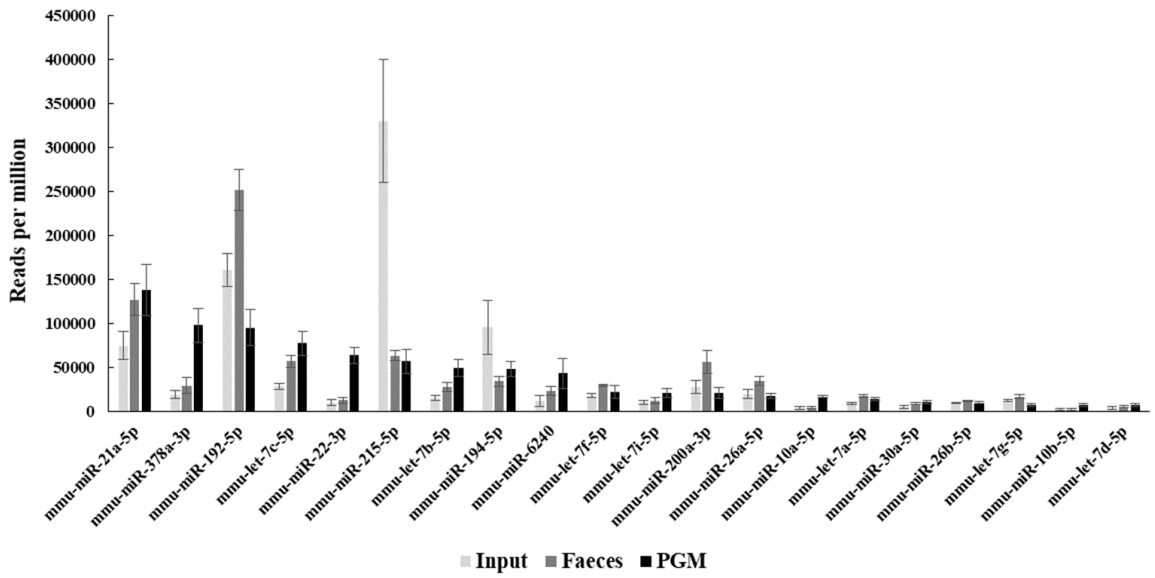


Figure 3.7 The top 20 abundant miRNAs in Input, Faeces, and PGM. miRNAs were sorted in descending order based on the number of reads in PGM samples. Data represent the average of 3 biological replicates ± S.D.

3.2.4 miRNA-21a-5p co-localizes with *Lactobacillus* in PGM

To provide further evidence for the presence of host miRNA in PGM, miRNA FISH was applied to visualize miRNAs within gut bacteria. A previous report suggested that miR-21a-5p could directly suppress the growth of *Lactobacillus* (Santos et al., 2020). miR-21a-5p is the most abundant miRNA in PGM based on sequencing (Figure 3.7) and I also see *Lactobacillus* is enriched in PGM. Therefore, I focused on testing whether miR-21a-5p can be specifically detected inside of any bacteria in PGM and whether this co-localizes with *Lactobacillus*.

I used LNA probes tagged with digoxigenin (DIG) and combined tyramide signal amplification (TSA) to detect miR-21a-5p in PGM (Figure 3.8). In this process, LNA oligonucleotides with Watson-Crick complementary sequences to the mature miRNA served as probes. The 5' end of the probe was labelled with DIG. Following successful binding of these probes to their target sequences and subsequent stringent washing steps to remove any excess or nonspecifically bound probes, horseradish peroxidase (HRP)-anti-digoxigenin antibody was added, followed by the HRP substrate Cyanine 5-conjugated tyramide. HRP can enzymatically convert Cyanine 5-conjugated tyramide to bind tyrosine residues on and surrounding the antibodies and HRP to achieve signal amplification. This sequential process allowed for the amplification and

visualization of the positive signal via fluorescence microscopy. The samples treated with a scrambled LNA probe were used as controls.

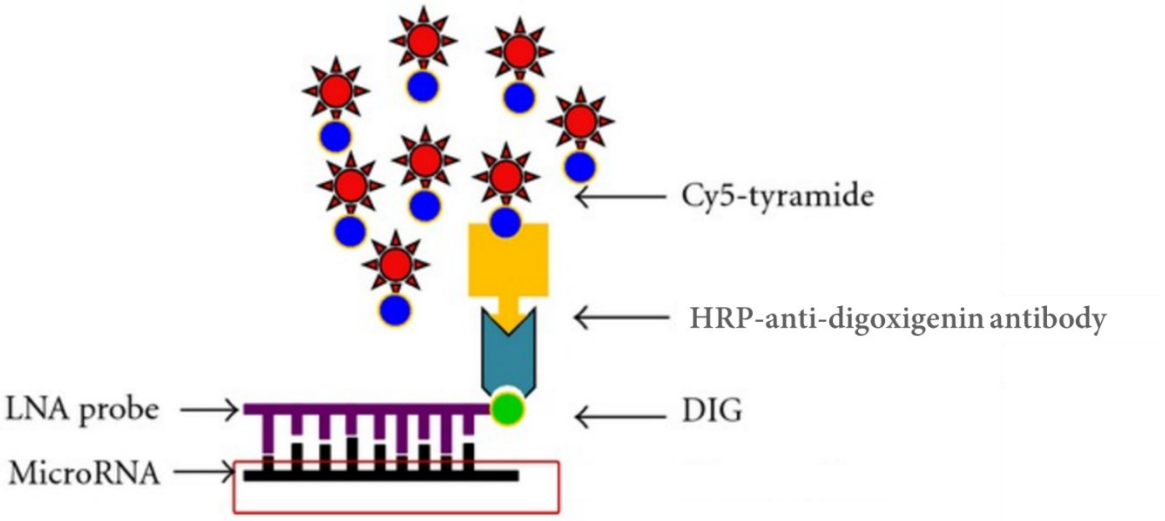


Figure 3.8 Simplified scheme of miRNA FISH using LNA probe and tyramide signal amplification technology. Figure adapted from Shi et al., 2012.

I employed two DNA probes for bacterial visualization in PGM. The universal DNA probe Eub338 (Akhtar et al., 2021) was utilized to detect a wide array of bacterial species by targeting a conserved region within the 16S rRNA gene. Additionally, the species-specific DNA probe Lab158 (Olszewska et al., 2016) was applied to specifically identify *Lactobacillus* species, targeting a unique sequence within their 16S rRNA gene. Both DNA probes were labelled with the fluorescent dye Cy3 for visualization. By performing FISH using the DNA probes Eub338 and Lab158 in comparison to the miR-21a-5p LNA probe, I expect that co-localisation would provide evidence that miR-21a-5p is taken up by *Lactobacillus*. The scrambled LNA probe and scramble Eub338 DNA probe served as controls to ensure specificity.

Figure 3.9 shows that miR-21a-5p is detected in some bacteria (green) of PGM. These bacteria can also be illuminated by universal DNA probe Eub338 and display a short rod shape. I then use Lab158 probe to confirm whether these short rod-shaped bacteria are *Lactobacillus*. Figure 3.10 shows that there is significant overlap (white colour in overlay image) between the miR-21a-5p LNA probe signal and the Lab158 probe signal, indicating that the bacteria imaged by miR-21a-5p are indeed *Lactobacillus*. This co-localization serves as strong evidence supporting the hypothesis that miR-21a-5p enters *Lactobacillus* and may have functional implications

in these bacteria. There is no signal for the scramble LNA probe and scramble Eub338 probe in the control group (Figure 3.11). This absence of signal in the control group confirms the specificity of the FISH experiment and reinforces the validity of the observed co-localization between miR-21a-5p and *Lactobacillus*.

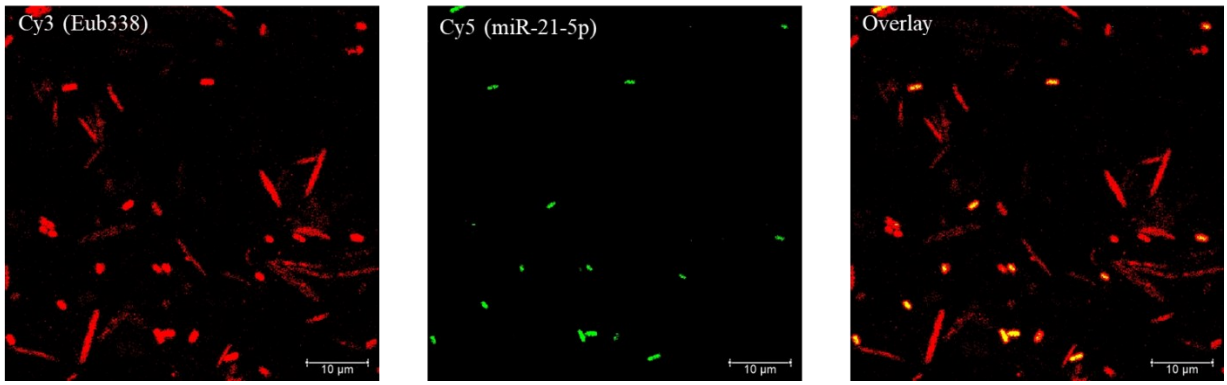


Figure 3.9 Fluorescence in situ hybridization in PGM using the Eub338 probe and miR-21a-5p LNA probe. Representative images were acquired on the Confocal Zeiss LSM880 Airyscan using a 60× objective, with identical settings applied to each probe channel. Cy3 staining (Eub338 probe) of bacterial 16S rDNA in red, Cy5 staining (miR-21a-5p LNA probe) of miR-21a-5p in green, and an overlay of these channels. The scale bar is 10 µm.

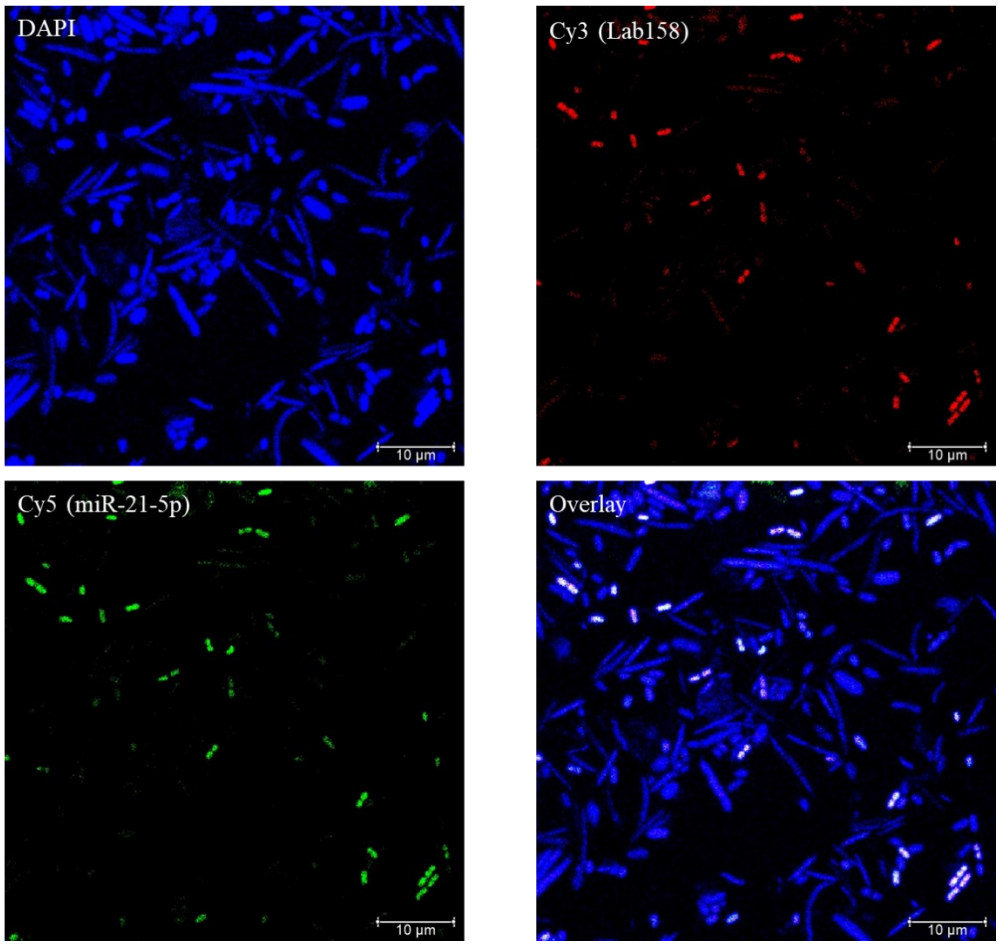


Figure 3.10 Fluorescence in situ hybridization in PGM using the Lab158 probe and miR-21a-5p LNA probe. Representative images were acquired on the Confocal Zeiss LSM880 Airyscan using a 60× objective, with identical settings applied to each probe channel. DAPI staining of cell nuclei in blue, Cy3 staining (Lab158 probe) of *Lactobacillus* in red, Cy5 staining (miR-21a-5p LNA probe) of miR-21a-5p in green, and an overlay of these channels. The scale bar is 10 µm.

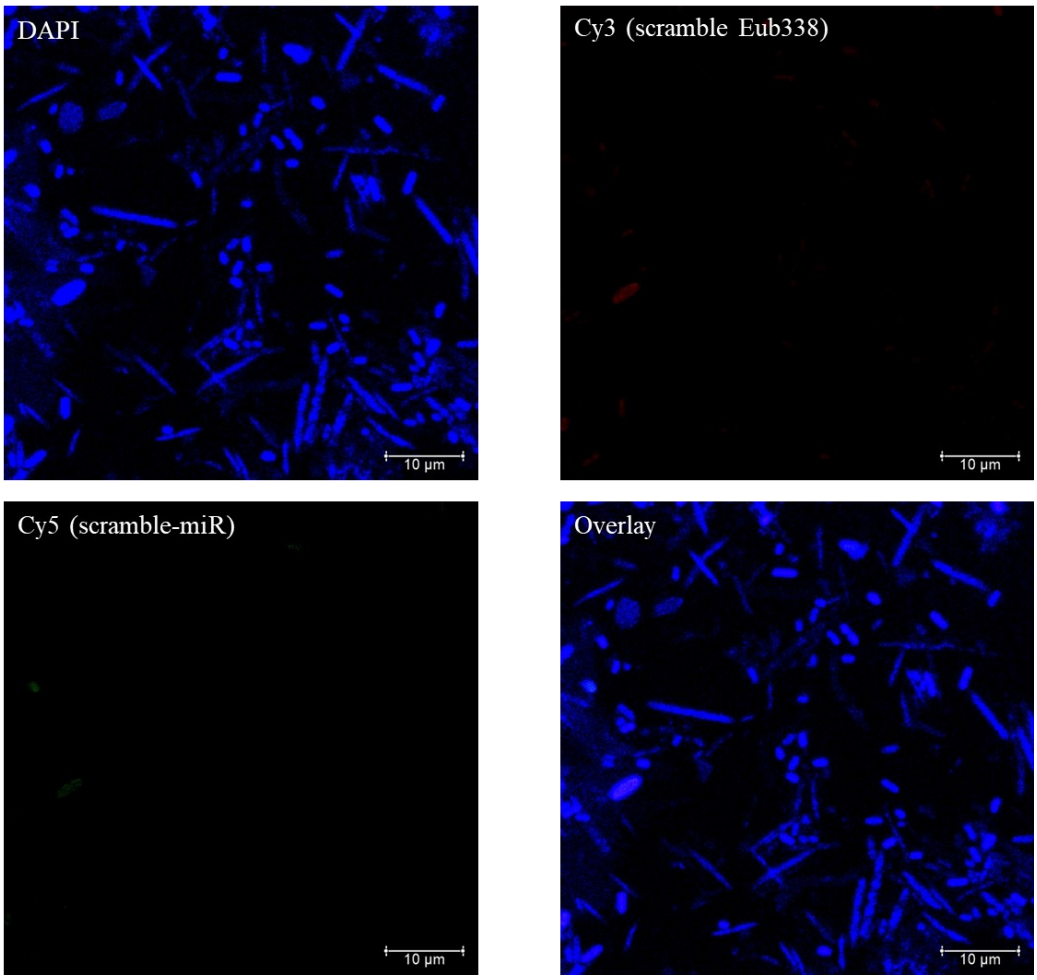


Figure 3.11 Fluorescence in situ hybridization in PGM using the scramble Eub338 probe and scramble-miR LNA probe. Representative images were acquired on the Confocal Zeiss LSM880 Airyscan using a 60× objective, with identical settings applied to each probe channel. DAPI staining of cell nuclei in blue, Cy3 staining (scramble Eub338 probe) in red, Cy5 staining (scramble-miR LNA probe) in green, and an overlay of these channels. The scale bar is 10 µm.

3.3 Discussion

3.3.1 Summary

In this chapter, I developed a purification method to isolate pure gut microbiota from mouse gut contents and profiled transferred host miRNAs in PGM by sRNA sequencing. My data suggest that some host miRNAs are naturally present inside the gut microbiota and the miRNAs within bacteria are a distinct subset compared to the miRNAs found in total gut contents or faeces. miRNA-FISH was applied to visualize host miRNAs within the purified gut bacterial cells and it shows that miR-21a-5p co-localizes with *Lactobacillus,* providing further support for the idea that miR-21a-5p is taken up by *Lactobacillus in vivo*.

3.3.2 Limitations of identifying the transferred host miRNAs by the analysis of sRNA sequencing data in PGM

One of the limitations is the disparity between the PGM and the entirety of the gut microbiota in terms of species that are present. The outer mucus layer of the intestine is known to serve as a habitat for numerous commensal bacteria (Johansson et al., 2013). However, when I collected PGM from mouse gut contents, I unintentionally excluded bacteria residing in the mucus layer. Consequently, these bacteria, which are more likely to interact closely with host cells, have been missed. Additionally, the bacterial 16S rDNA sequence-based survey revealed that certain bacterial species did not survive the purification process and were depleted from the PGM sample (Figure 3.3). Therefore, the miRNA content identified in PGM does not perfectly mirror that of the entire gut microbiota *in vivo*.

The other limitation of the method is that we cannot rule out that trace amounts of mouse contaminants remain in PGM. Assessing the complete removal of mouse cells through PCR analysis of the mouse 18rDNA gene and eEF2 gene by PCR might not be sensitivity enough. Therefore, a small number of miRNAs detected in PGM may still stem from potential mouse contaminants. To identify high confidence transferred miRNAs, I examined both the abundant miRNAs in PGM and the enriched miRNAs in PGM relative to the Input. Among the top 20 most abundant miRNAs in PGM, only mmu-miR-22-3p, mmu-miR-378a-3p, and mmu-miR-10a-5p exhibited enrichment in PGM with an FDR < 0.05 (Figure 3.6 C). miR-21a-5p is the most the abundant miRNA in PGM (Figure 3.7) and I successfully imaged it in gut microbiota by miRNA-FISH

(Figure 3.9). However, miR-21a-5p does not display enrichment in PGM when compared to the Input (Figure 3.6C; Supplementary Table 3.1) which is based on the high quantity of miR-21a-5p present in the Input sample as well. The analysis of sRNA sequencing data in PGM presented high confidence candidates for transferred miRNAs. However, to confirm their presence in the gut microbiota and mitigate the influence of potential trace amounts of mouse contaminants, additional validation methods like miRNA-FISH are necessary.

3.3.3 Future work

sRNA sequencing combined with miRNA-FISH can determine which miRNAs transfer to gut microbiota but this does not identify which bacteria internalize these miRNAs. Using co-staining methods with probes for miR-21a-5p and specific bacterial strains such as *Lactobacillus* I demonstrated co-localization, confirming the internalization of miR-21a-5p by *Lactobacillus*. This is a significant development, and is complementary to a previous study indicating that miR-21a-5p suppresses the growth of *Lactobacillus* (Santos et al., 2020). But that study lacked direct evidence of miR-21a-5p presence within *Lactobacillus* as demonstrated here. This provides an avenue for further exploration into other miRNA-bacteria interactions in the gut. Future work will focus on imaging more miRNAs such as miR-22-3p and miR-378a-3p which are enriched in PGM and identifying the specific bacteria that take up these miRNAs.

This chapter elucidates the natural presence of host miRNAs within gut microbiota, paving the way for investigating the mechanisms behind the transport of host miRNAs into bacteria and their functional roles in subsequent chapters.

Chapter 4: Identification of the role of mouse Ago2 protein in host miRNA-bacteria interaction

4.1 Introduction

A previous study of muscles cells showed that Ago2 enters the mitochondria and the miRNA-Ago2 complexes directly increase mitochondrial translation (Zhang et al., 2014). Since mitochondria resemble bacteria in many aspects this provides some precedent for a possible mechanism by which miRNAs could enter bacteria and elicit gene regulation. Recent results in our lab show that Ago2 is secreted by mouse intestinal epithelial cells predominately in a soluble non-vesicular form and this extracellular Ago2 is associated with specific populations of miRNAs compared to the cellular pool. The presence of non-vesicular Ago2 is consistent with studies in other extracellular contexts. For example, miRNA-Ago2 complexes are also found in human plasma and serum, where Ago2 is proposed to play a critical role in stabilizing extracellular miRNAs (Arroyo et al., 2011; Turchinovich et al., 2011; Geekiyanage et al., 2020). In this chapter, I examine the hypothesis that mouse miRNAs could move with the Ago2 protein into bacteria and thereby mediate miRNA-bacterial RNA interactions. To test this hypothesis, I aim to determine whether mouse Ago2 protein is detected inside purified gut bacteria and to identify bacterial genes directly targeted by Ago2-miRNA complexes.

Several methods for detecting Ago2 protein were considered in this chapter. For example, using antibodies the protein can be assessed by Western blotting and immunofluorescence imaging. Previous study that aimed to detect Ago2 trafficking from the cytoplasm to the mitochondria during muscle differentiation used centrifugation on a discontinuous sucrose gradient to isolate mitochondria from cardiomyocytes. The authors then used an anti-Ago2 western blot analysis to detect Ago2 protein in the highly purified mitochondria (Zhang et al., 2014), providing evidence of transfer. In this chapter, I employ this type of approach to investigate whether the host Ago2 protein is transferred into gut bacteria. Specifically, I purified the gut microbiota from mouse gut contents, followed by the isolation and concentration of Ago2 protein through immunoprecipitation, enabling subsequent anti-Ago2 western blot analysis.

As another method to examine Ago2 presence in bacteria I use immunofluorescence. Immunofluorescence is an immunochemical technique that employs fluorophore-conjugated antibodies to identify and localize proteins within various cells or tissues (Im et al., 2019). As an example, immunofluorescence imaging was used previously to show the co-localization of Ago2 with the multi-vesicular endosome marker CD63 in colon cancer cells (McKenzie et al., 2016). One of the challenges I face in my samples is the uncertainty regarding the expected signal of Ago2 following transfer into recipient bacteria: it is hard to know if the methods will have sufficient sensitivity to definitively answer the question. In line with this, a technique known as tyramide signal amplification (TSA)-based immunofluorescence is of relevance because it has been developed to detect of low-abundance protein targets (Faget and Thomas, 2015). TSA exploits the highly sensitive nature of horseradish peroxidase (HRP) to enzymatically convert tyramide fluorophores into short-lived radical intermediates that bind to tyrosine residues on proteins located near both the antibodies and HRP. An example of the effectiveness of immunofluorescence with TSA is the detection of the low abundant cytokine receptor interleukin-17 receptor B in human osteosarcoma U2OS cells (Wang et al., 2020). I intend to employ the TSA-based immunofluorescence technique with purified gut microbiota to detect mouse Ago2.

Beyond simply detecting Ago2 in gut microbiota, a parallel goal in this chapter was to explore whether Ago2 could directly mediate miRNA-bacterial RNA interactions and identify bacterial genes directly targeted by Ago2-miRNA complexes. At present, the most common approach for interrogating miRNA-target interactions starts with computational target predictions. Numerous computational approaches have been developed to predict miRNA targets, primarily relying on factors such as i) complementarity of the target sequence to the seed region of the miRNA (the first 2-8 nucleotides of the miRNA from the 5’ end), ii) sequence conservation of the miRNA and the target site in the 3’UTR of the transcript, iii) Gibbs free energy of the miRNA-target base pairing and iv) site accessibility of the target site (Peterson et al., 2014). However, computational target predictions can still lead to a substantial number of false positive targets due to the relatively short seed sequence. Furthermore, the rules that would be appropriate for prediction of bacterial targets with host miRNAs are not yet studied. This highlights the challenge and complexity of accurately predicting miRNA-bacterial interactions: there is not a strong foundation on which to build.

To address the limitations of computational predictions and gain more insight into miRNA-bacterial interactions, in this chapter I aimed to test a biochemical approach to capture bacterial targets that might be found associated with Ago2. In recent years biochemical approaches have been developed to directly capture and ligate miRNAs and their targets when bound by Argonaute proteins in cell cultures or tissues. For example, a method named crosslinking ligation and sequencing of hybrids (CLASH) was developed to identify miRNA-mRNA interactions bound by Argonaute proteins in HEK 293 cells (Helwak and Tollervey, 2014). Another modified UV crosslinking and immunoprecipitation strategy termed CLEAR-CLIP (covalent ligation of endogenous Argonaute-bound RNAs–Crosslinking and immunoprecipitation) was applied to mouse brain tissue to discover endogenous *in vivo* miRNA-target interactions mediated by Argonaute proteins (Moore et al., 2015). The basic workflow (Figure 4.1) for CLASH and CLEAR-CLIP is almost the same and this involves a series of key steps, including UV cross-linking of cell cultures or tissues to stabilize the protein and RNA interactions, immunoprecipitation of protein-RNA complexes under stringent conditions, intermolecular ligation of the guide sRNAs to the target RNAs to generate chimeric reads, cDNA library preparation, high-throughput sequencing and bioinformatic analysis. There are a few differences between the CLASH and CLEAR-CLIP protocols CLEAR-CLIP employs more stringent lysis and wash buffers compared to CLASH. Furthermore, in CLEAR-CLIP, the RNase treatment occurs within the cell lysate, whereas in CLASH, this treatment is performed after the immunoprecipitation on beads.

Dr Sarah Ressel in our lab previously optimised and compared CLASH and CLEAR-CLIP protocols with the human bronchial epithelial cell line A549 and her data showed CLEAR-CLIP could generate higher proportion of total miRNA-target chimeras than CLASH. Therefore, I intend to adapt the CLEAR-CLIP protocol in the mouse gut tissue to determine whether I can identify bona fide chimeric reads between transferred host miRNAs and bacterial RNAs associated with Ago2 protein.

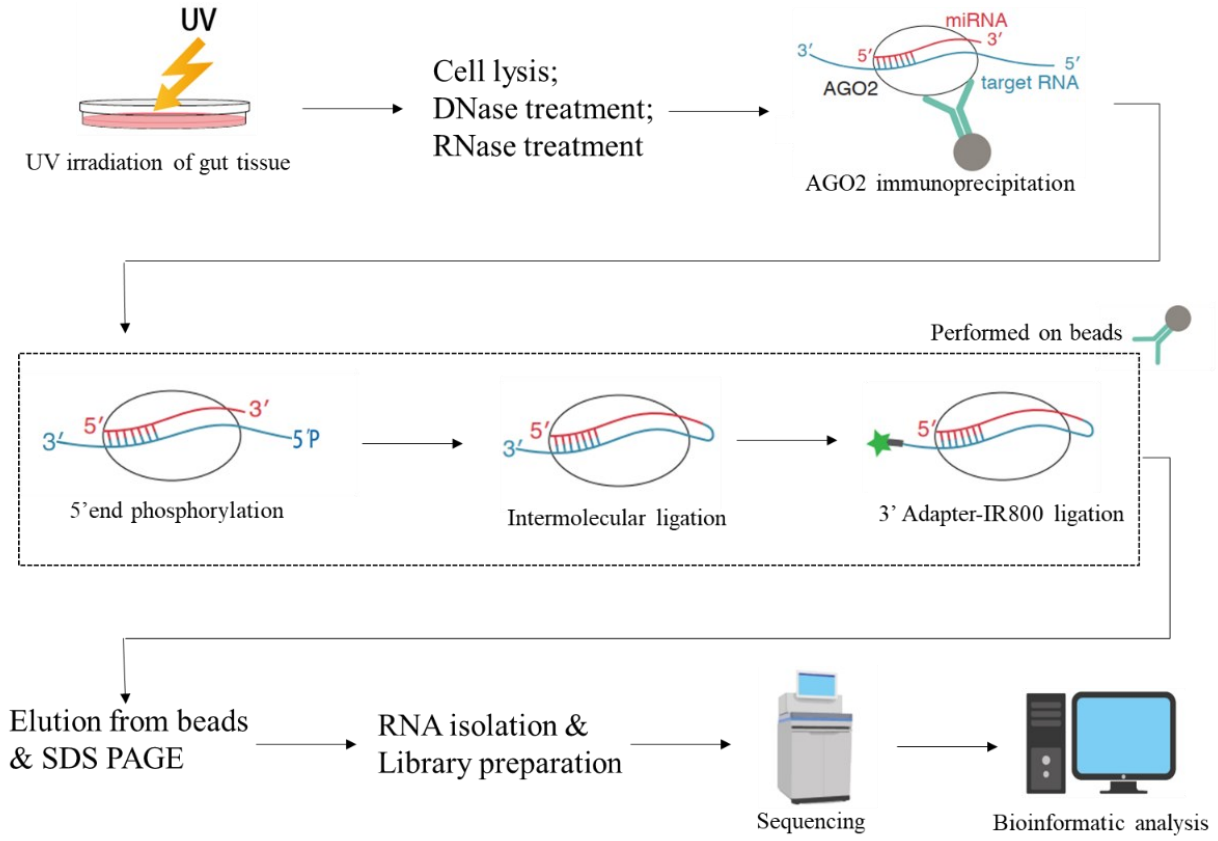


Figure 4.1 Basic workflow for CLASH and CLEAR-CLIP protocols. Cell cultures or tissues are UV cross-linked, lysed and Ago2-RNA complexes are immunoprecipitated using an Ago2 antibody conjugated to magnetic beads. The miRNA was ligated to bound target RNAs to generate chimeric reads for next generation sequencing and bioinformatic analysis.

4.2 Results

4.2.1 Detecting Ago2 in purified gut microbiota

In Chapter 3 I developed a method to purify gut microbiota from mouse gut contents and here the goal is to determine whether mouse Ago2 can be detected in PGM using anti-Ago2 western blot analysis and immunofluorescence imaging.

Firstly, I did anti-Ago2 western blot analysis to detect Ago2 protein in PGM. As the precise quantity of host Ago2 transferred into bacteria remained uncertain, I performed Ago2 protein immunoprecipitation in samples aiming to enrich the Ago2 protein content for anti-Ago2 western blot analysis. To test the efficacy of the Ago2 immunoprecipitation method, I initially applied this technique to mouse gut tissue, gut contents, and faeces collected from C57BL/6J mice from which I collected the gut microbiota in Chapter 3. The gut tissue, specifically from the cecum, was ground into a fine powder under liquid nitrogen and then lysed using lysis buffer in the presence of protease inhibitors. The non-lysed material was spun down and the supernatant was incubated with the beads conjugated to a mouse Ago2-specific antibody. Strong signal at ~97 kDa (which is the expected molecular weight of Ago2) was detected in the western blot analysis, indicating successful immunoprecipitation of Ago2 protein from the gut tissue lysates (Figure 4.2A). For Ago2 immunoprecipitation using mouse gut contents and faeces, the whole intestine was dissected from the sacrificed mouse and faecal pellets were collected from the rectum and the luminal contents were squeezed out from the remaining intestine by tweezers. 100 mg of the gut contents and faecal sample were homogenized and lysed in lysis buffer with 0.1mm zirconium grinding beads using the Tissue Lyser II. Zirconium grinding beads can effectively shear bacterial cell walls, ensuring lysis of all bacterial cells (Vickery et al., 2017). The following steps for immunoprecipitation and western blot were the same as those for gut tissues. A relatively weak band of the size corresponding to Ago2 was observed in the gut content and faecal samples (Figure 4.2A) and a larger band at approximately 140 kDa was also observed in these samples, which could potentially represent nonspecific binding. It is important to note that the presence of weak bands in these samples does not provide direct evidence that Ago2 protein is secreted into the gut lumen and taken up by gut bacteria. There were mouse cell contaminations present in

gut contents and faeces confirmed by the PCR analysis of mouse 18S rDNA gene (Figure 4.2B), which may contribute to the detected signal.

To address this concern the same protocols were applied to the PGM samples. In Chapter 3 I showed that I could purify ~250 mg of PGM from each C57BL/6J mouse and get rid of mouse material with this method. Before conducting the western blot analysis of immunoprecipitated Ago2 in PGM, I performed PCR analysis using 50 mg of each PGM sample to confirm the absence of mouse material. The results indicated that the 18S rDNA gene was not detected in the PGM samples (Figure 4.2B). I therefore lysed 200 mg of each PGM sample in lysis buffer with 0.1mm zirconium grinding beads using the Tissue Lyser II and then carried out an Ago2 immunoprecipitation and anti-Ago2 Western blot. Although I could detect Ago2 in one PGM sample, this result was not reproducible despite multiple experiments. Representative results of anti-Ago2 western blot are showed in Figure 4.2C.

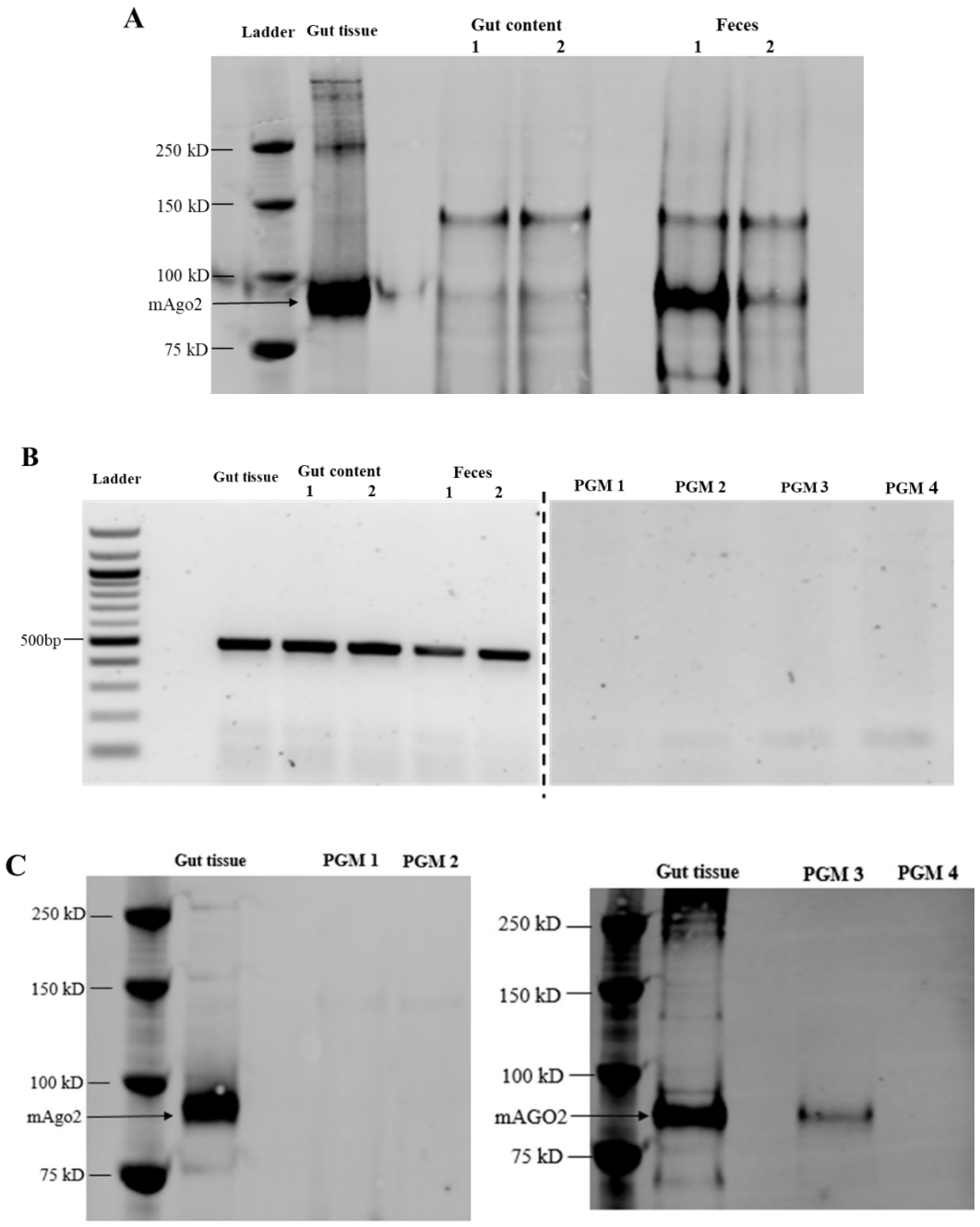


Figure 4.2 Detection of mouse Ago2 in different samples. A) Western blot analysis of immunoprecipitated Ago2 from mouse gut tissues, gut contents, and faeces. 100 mg of the respective materials (gut tissues, gut contents, and faeces) were lysed using appropriate methods. Ago2 was then immunoprecipitated using Protein G beads conjugated with mouse anti-mAgo2 antibody. For gut tissues 10% of the elutant was loaded and for gut contents and faeces, 100% of the elutant was loaded. B) PCR analysis of mouse 18S rDNA gene in gut tissues, gut contents, faeces, and purified gut microbiota (PGM). C) Western blot analysis of immunoprecipitated Ago2 from 200 mg PGM. 100% of the elutant was loaded on the gel.

One limitation with the western blot analysis is that this measures Ago2 in the bulk sample and it is possible that Ago2 might only traffic to specific bacteria. Taking this into account, I utilized immunofluorescence imaging in PGM to determine whether this would detect Ago2 signal within individual bacteria. The workflow of immunofluorescence staining using Tyramide SuperBoost™ Kits is described in Figure 4.3. Briefly, cells were fixed and permeabilized and then incubated with primary and secondary antibodies. For mouse Ago2 detection, the primary antibody was mouse anti-mAgo2 antibody. Mouse IgG was used to control for the specificity of the anti-mAgo2 antibody signal. The secondary antibody was the poly-HRP-conjugated goat anti-mouse IgG antibody. The enzyme horseradish peroxidase (HRP) in the secondary antibody catalyzed the deposition of multiple fluorophore-conjugated tyramide molecules near the antibody-bound target. This resulted in an enhanced fluorescence signal. The cells were then imaged using a fluorescence microscope.

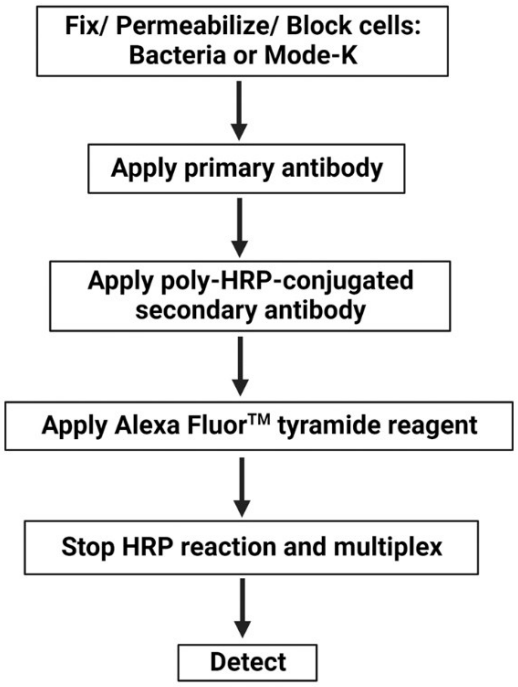


Figure 4.3 Workflow of immunofluorescence staining using Tyramide SuperBoost™ Kits. Figure adapted from the user guide of Tyramide SuperBoost™ Kits.

To assess the effectiveness of this Ago2 detection method, I initially conducted immunofluorescence staining of Ago2 in Mode-K cells which are a mouse intestinal

epithelial cell line. Previous unpublished data from our laboratory showed that Mode-K cells have sufficient levels of Ago2 protein expression for detection by imaging. Mode-K cells were seeded in a 12-well chamber slide, and the complete immunofluorescence staining process was carried out within it. Subsequently, images were captured using the EVOS imaging system. As shown in Figure 4.4, Ago2 protein was successfully detected in Mode-K cells, as evidenced by the presence of Alexa Fluor™ 488 fluorescence signal. In contrast, when using mouse IgG as a negative control, no significant fluorescence signal was observed, confirming the specificity of the Ago2 antibody.

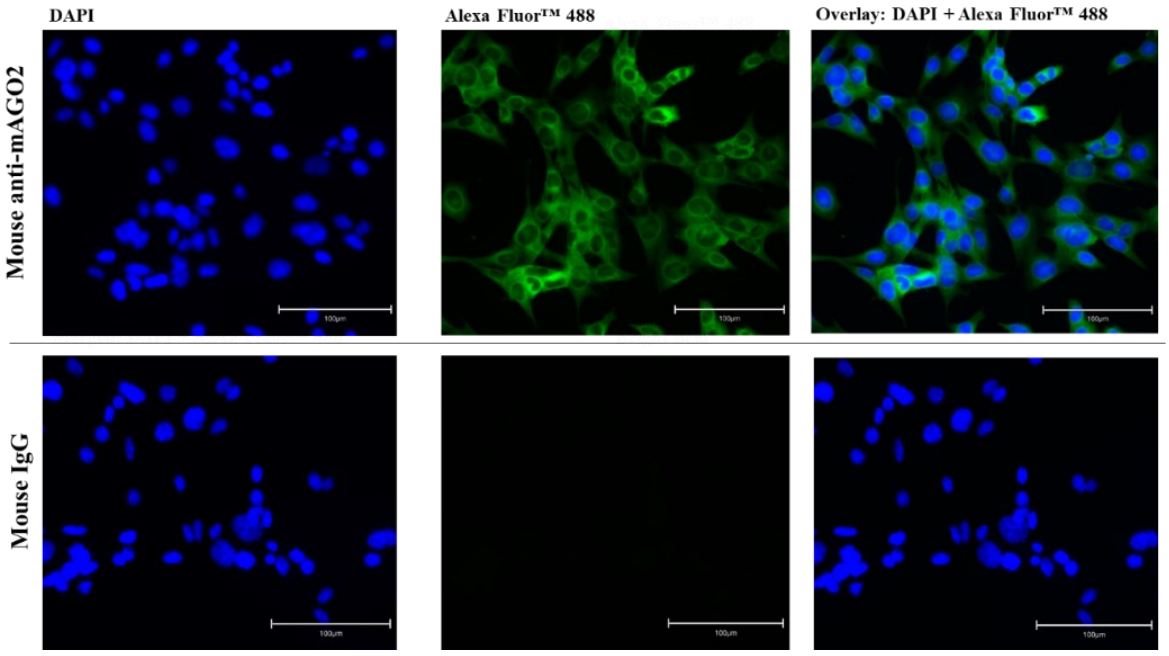


Figure 4.4 Representative immunofluorescence images of Ago2 detection in Mode-K cells. Immunofluorescence staining was performed using Tyramide SuperBoost™ Kits. Mouse anti-mAgo2 antibody and mouse IgG were used as primary antibodies. Mouse IgG served as a control to evaluate the specificity of the anti-mAgo2 antibody signal. The images were captured on the EVOS imaging systems using a 10× objective with identical settings applied to each probe channel. From left to right, the columns represent DAPI staining of cell nuclei in blue, Alexa Fluor™ 488 staining of Ago2 protein in green, and an overlay of these channels. The scale bar is 100 μm.

I then applied this method to the PGM samples. However, it is important to note that the immunofluorescence staining process and microscopy differed somewhat from those used for Mode-K cells. For PGM samples, gut bacteria were suspended in 1.5

ml Eppendorf tubes and required a lysozyme treatment to break down the cell wall before the primary antibody incubation. The gut bacteria were loaded onto a 2% agarose gel on a slide and imaged using the Zeiss LSM880 Airyscan microscopy system. As shown in Figure 4.5, no Alexa Fluor™ 488 fluorescence signal could be found for Ago2 protein in PGM. This suggests that the Ago2 protein is not present in the gut bacteria within the PGM samples or is present at levels below the detection limit of the assay.

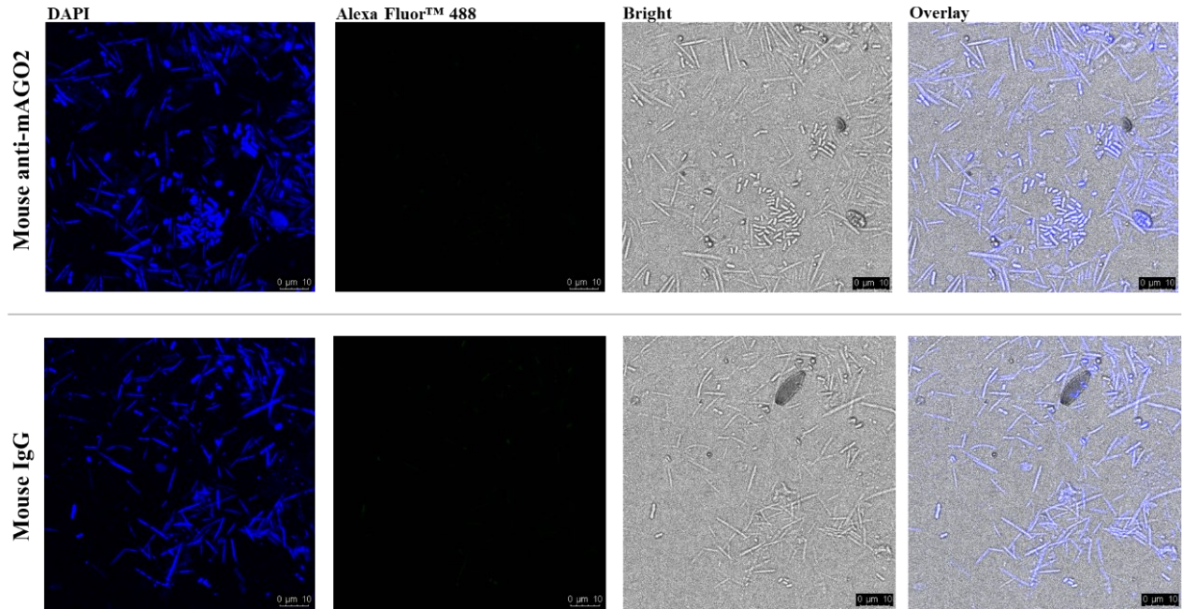


Figure 4.5 Representative immunofluorescence images of Ago2 detection in purified gut microbiota. Immunofluorescence staining was conducted using Tyramide SuperBoost™ Kits with modified protocols. Mouse anti-mAgo2 antibody and mouse IgG were employed as primary antibodies, with the latter serving as a control to assess the specificity of the anti-mAgo2 antibody signal. The images were acquired on the Confocal Zeiss LSM880 Airyscan using a 60× objective, with identical settings applied to each probe channel. The columns, from left to right, depict DAPI staining of cell nuclei in blue, Alexa Fluor™ 488 staining of Ago2 protein in green, brightfield in grey, and an overlay of these channels. The scale bar is 10 µm.

4.2.2 Identification of bacterial targets using Ago2 CLEAR-CLIP in mouse gut

As described above, I was unable to detect Ago2 inside gut bacteria by immunofluorescence imaging and was not able to reliably detect Ago2 within the PGM following immunoprecipitation and anti-Ago2 western blot analysis. However, it needs to be considered that the outer mucus layer of the colon is known to serve as a habitat

for numerous commensal bacteria (Johansson et al., 2013) that were not all present in PGM based on how this is collected. Considering this limitation and with the aim of building on our methods in the lab to identify targets using Ago2 CLEAR-CLIP, I performed CLEAR-CLIP in the whole mouse gut tissue where bacteria from the mucus layer and luminal gut contents would all be included. The goal of this approach is to determine whether we can identify authentic miRNA-bacteria interactions based on chimeric sequencing reads, which would indicate transfer and binding of host miRNAs to bacterial RNAs within Ago2.

When designing the experiment several key factors had to be considered. Firstly, I needed to determine the best approach for conducting UV cross-linking of whole gut tissues. Several studies have presented methods for UV cross-linking in both fresh and frozen tissues. While it is typically essential to triturate the tissue for effective UV penetration, it is not always necessary to create a single-cell suspension. This is because UV light can permeate through multiple cell layers, as established in previous research (Chi et al., 2009). For instance, a study on fresh mouse neocortex tissue involved triturating the tissue to create a tissue suspension, followed by irradiation in culture plates using a Stratalinker with three doses of 400mJ/cm^2 (Chi et al., 2009). In another study (Maurin et al., 2018), a method was outlined for applying UV cross-linking and immunoprecipitations to frozen brain tissues. In this case, approximately 100 mg of brain tissues were ground into a fine powder using liquid nitrogen, and the resulting powders were evenly spread onto 60 mm petri dishes floating in liquid nitrogen. Subsequently, RNAs and proteins were cross-linked through three doses of 254 nm UV irradiation, with each dose being set at 400 mJ/cm^2. Given that our laboratory has limited experience with UV cross-linking in gut tissues (and there is no literature with this tissue), I decided to test both fresh and frozen gut tissues for cross-linking to evaluate the efficiency of chimeric read recovery.

A second factor to consider is the selection of sampling locations within the gut tissues. Given that species richness and the abundance of gut bacteria vary along the gastrointestinal tract (Donaldson et al., 2016), different gut locations exhibit distinct microbial compositions. These variations may significantly influence the interactions between miRNAs and bacteria that could be detected. In order to cast the net wide and compare miRNA-bacteria interactions at different locations along the

gastrointestinal tract, I divided the gut tissue into four segments for CLEAR-CLIP analysis: the duodenum, ileum, cecum, and colon.

The third factor to consider is controls. Commercial enzymes for the CLEAR-CLIP experiments contain some bacterial RNA contaminants whose introduction could then result in chimeric reads with miRNAs (Moore et al., 2015). To establish the baseline for such contamination that could be introduced during the protocol, I used gut tissues from germ-free mice as a control. These samples were kindly provided to us by Prof. Ruth Ley's Department in Germany. It is worth noting that these gut tissues were divided into only two segments, small intestine, and large intestine, and were frozen at -80°C. As a result, I conducted parallel experiments using these frozen tissues. Additionally, to assess the specificity of the Ago2 antibody, gut tissues were subjected to immunoprecipitation using mouse IgG. The detailed sample collection plan is outlined in Table 4.1. Two biological replicates were carried out for each type of sample.

Table 4.1 Samples collected for Ago2 CLEAR-CLIP in mouse gut

	Fresh gut tissue	Sample ID
C57BL/6 mice	duodenum	Fresh_D1, Fresh_D2
	ileum	Fresh_I1, Fresh_I2
	cecum	Fresh_CE1, Fresh_CE2
	colon	Fresh_CO1, Fresh_CO2
	Frozen gut tissue	**Sample ID**
C57BL/6 mice	duodenum	Frozen_D1, Frozen_D2
	ileum	Frozen_I1, Frozen_I2
	cecum	Frozen_CE1, Frozen_CE2
	colon	Frozen_CO1, Frozen_CO2
	Frozen gut tissue	**Sample ID**
Germ-free C57BL/6 mice	Small intestine	GF_S1, GF_S2
	Large intestine	GF_L1, GF_L2

From previous work in our lab, we recognized that if UV crosslinking doses are too high, they can result in protein loss and fewer miRNA-target chimeras (e.g., 400mJ/cm^2 in A549 cells is too high; Ressel et al., 2023). Therefore, prior to Ago2 CLEAR-CLIP experiments, I tested whether there was any loss of Ago2 after three doses of 400mJ/cm^2 UV crosslinking with both fresh and frozen gut tissues. I collected gut tissues from C57BL/6J mice and 120 mg of the fresh or frozen gut tissues

underwent UV cross-linking, lysis, and Ago2 immunoprecipitation as outlined in Chapter 2, Section 2.11. Subsequent anti-Ago2 western blot analysis (Figure 4.6) revealed a similar yield of Ago2 protein from fresh and frozen gut tissues after UV crosslinking, in comparison to the non-crosslinked group. This indicates that the UV crosslinking dose of 400mJ/cm^2 did not result in obvious protein loss for mouse gut tissues.

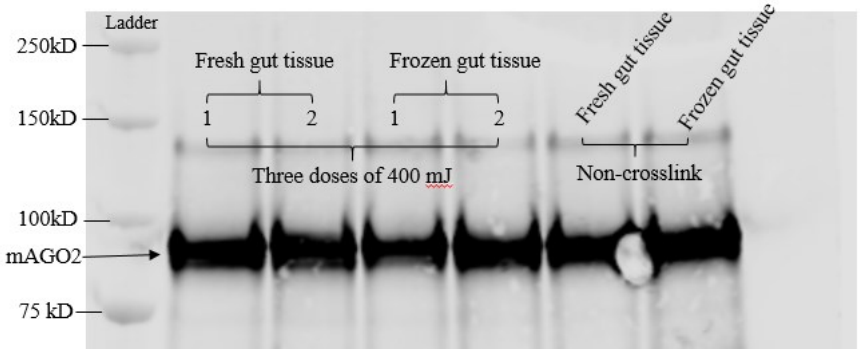


Figure 4.6 Western blot analysis of immunoprecipitated Ago2 from fresh and frozen gut tissue after UV-crosslinking. Fresh and frozen gut tissues were under three doses of 400mJ/cm2 UV crosslinking and two replicates for each. For these IPs, 50 µl of Dynabeads Protein G beads with 10 µl of mouse anti-mAGO2 antibody (1.15 mg/ml) were used per sample. Western blots were first probed with 1:4,000 mouse anti-mAGO2 primary antibody followed by 1:10,000 goat anti-mouse AlexaFluor 680. The arrow indicates the position of the mouse Ago2 protein on the gel.

Next, I prepared samples for Ago2 CLEAR-CLIP experiments with the samples detailed in Table 4.1. Gut tissues were harvested from four C57BL/6J mice, with two for fresh gut tissue and two for frozen gut tissue. Both fresh and frozen gut tissues underwent UV crosslinking and lysis. Subsequently DNase and RNase treatments were performed in the cell lysates. Ago2-RNA complexes were then immunoprecipitated using a mouse Ago2-specific antibody or mouse IgG that was conjugated to magnetic beads. For further processing, I conducted 5' end phosphorylation, intermolecular RNA-RNA ligation, and fluorescent 3' adaptor ligation while the RNA complexes remained attached to the beads. After eluting Ago2-RNA complexes from the beads, I preserved 10% of the eluted sample for anti-Ago2 western blot analysis. Anti-Ago2 Western blots confirmed that mouse Ago2 was successfully pulled down from all the samples, in contrast to the IgG control groups (Figure 4.7).

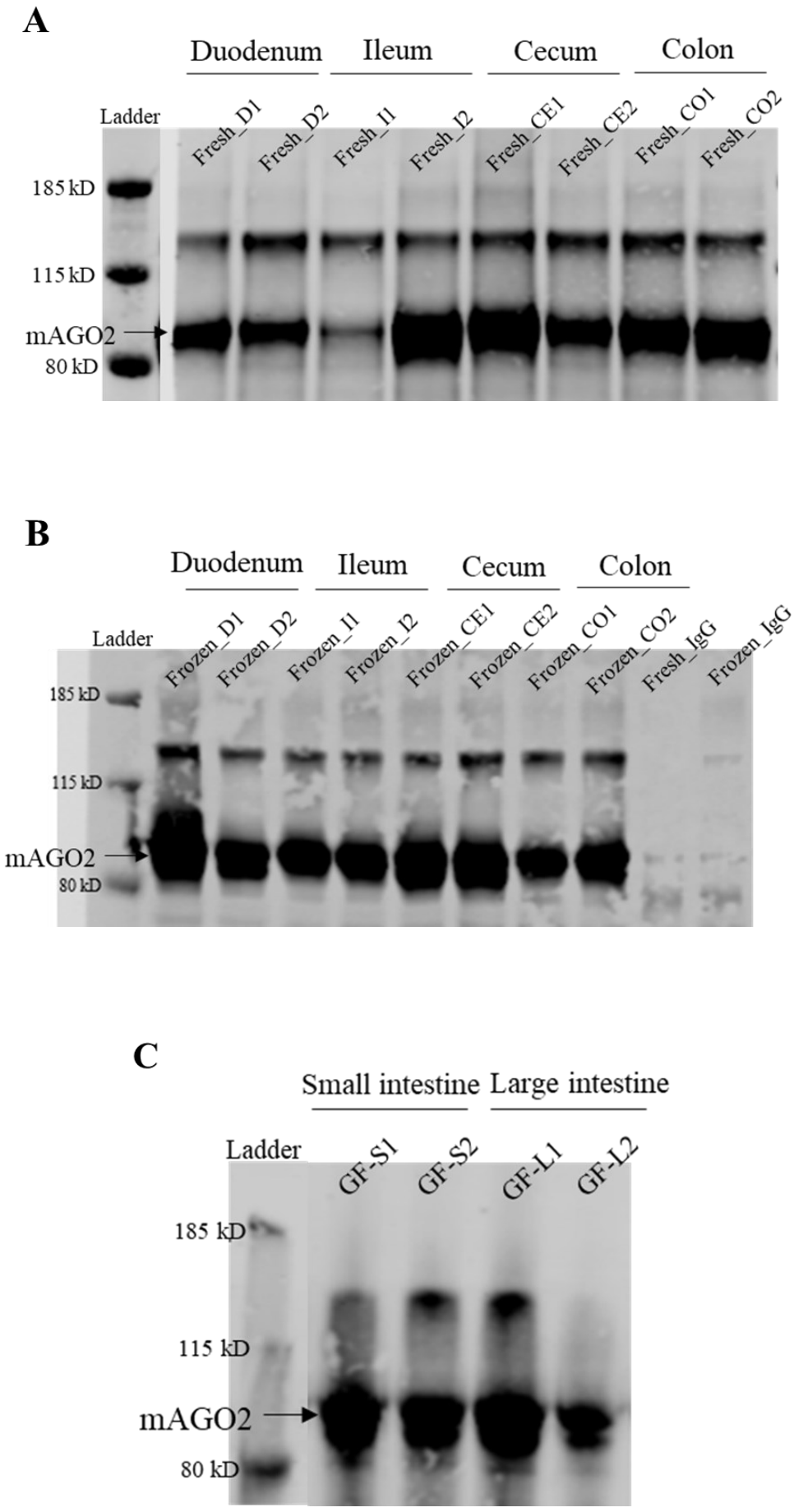
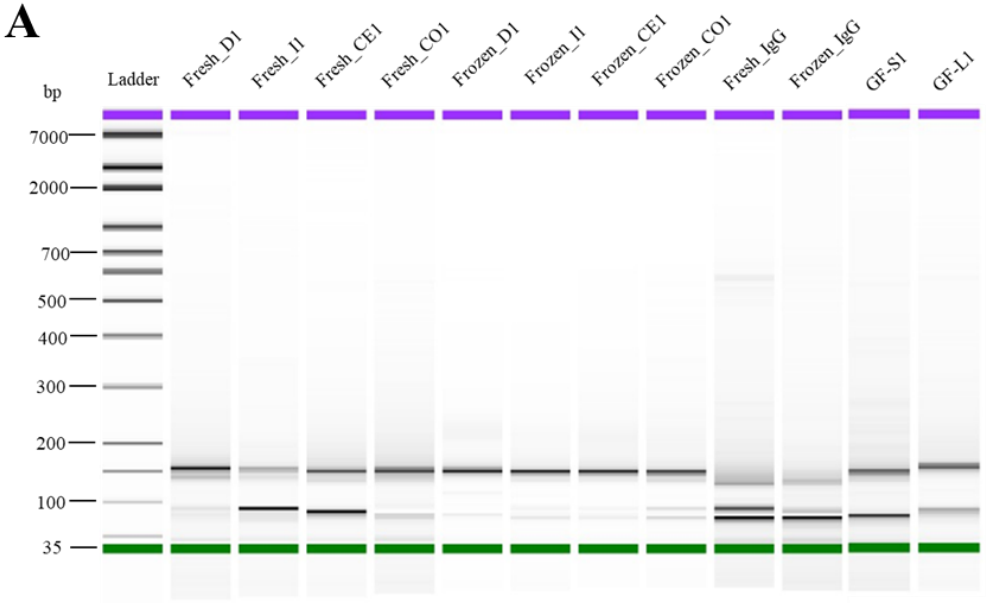

A
Duodenum
Ileum
Cecum
Colon
Ladder
Fresh_D1
Fresh_D2
Fresh_I1
Fresh_I2
Fresh_CE1
Fresh_CE2
Fresh_CO1
Fresh_CO2
185 kD
115 kD
mAGO2
80 kD
B
Duodenum
Ileum
Cecum
Colon
Ladder
Frozen_D1
Frozen_D2
Frozen_I1
Frozen_I2
Frozen_CE1
Frozen_CE2
Frozen_CO1
Frozen_CO2
Fresh_IgG
Frozen_IgG
185 kD
115 kD
mAGO2
80 kD
C
Small intestine
Large intestine
Ladder
GF-S1
GF-S2
GF-L1
GF-L2
185 kD
115 kD
mAGO2
80 kD

Figure 4.7 Quality control analyses of immunoprecipitated Ago2 in CLEAR-CLIP samples. A) Western blot results of immunoprecipitated Ago2 in fresh gut tissues. B) Western blot results of immunoprecipitated Ago2 in frozen gut tissues. The samples labelled “Fresh_IgG” and “Frozen_IgG” represent the fresh or frozen gut tissues that were immunoprecipitated using mouse IgG. C) Western blot results of immunoprecipitated Ago2 in germ-free mouse gut tissues. There were two biological replicates for each type of sample, all of which were immunoprecipitated using a mouse Ago2-specific antibody. Western blots were first probed with 1:4,000 mouse anti-mAgo2 primary antibody followed by 1:10,000 goat anti-mouse AlexaFluor 680. The arrow indicates the position of the mouse Ago2 protein on the gel.

The remaining eluted sample was separated by SDS-PAGE, and Ago2-RNA complexes were visualized using RNA-IR800-3’ adapters using a Li-cor Odyssey CLx. Previous studies have indicated that Ago–miRNA–mRNA complexes typically migrate within the 115-185 kD range (Moore et al., 2015; Ressel et al., 2023). I therefore excised regions within this range from the SDS-PAGE gel to isolate Ago–miRNA–mRNA complexes for RNA extraction. Subsequently, the RNA was reverse transcribed, PCR amplified, and the resulting library sizes were assessed using a Bioanalyzer High Sensitivity DNA chip. The chip results from one replicate of each type of sample (Figure 4.8A) demonstrate that libraries prepared with the Ago2 antibody exhibit a consistent size distribution pattern, distinguishing them from libraries prepared with mouse IgG. Within one library (Frozen_D1) prepared with the Ago2 antibody, the peak table (Figure 4.8B) reveals that a main band around 150 bp corresponding to the size of miRNAs reads and certain peaks exceed 160 bp corresponding to the expected size of miRNA–mRNA chimeric reads. To create the library for sequencing, I pooled 20 µl of each library, followed by size selection between 160-260 bp to select for the miRNA–mRNA chimeric reads’ size range. Sequencing of these libraries was performed using a single-end 100 base pair sequencing run on the Illumina NextSeq 2000 platform by the Edinburgh Clinical Research Facility. The table below (Table 4.2) lists the number of high-quality reads for each library.

A



B

Peak table for sample Frozen_D1

Peak	Size [bp]	Conc. [pg/µl]	Molarity [pmol/l]	Observations
1	35	125.00	5,411.3	Lower Marker
2	45	194.10	6,536.6	
3	48	109.56	3,457.0	
4	79	704.84	13,452.9	
5	94	86.10	1,388.2	
6	114	351.93	4,672.0	
7	152	14,821.09	147,518.5	
8	210	783.49	5,640.3	
9	218	1,241.61	8,623.4	
10	249	120.38	732.8	
11	260	59.41	346.7	
12	10,380	75.00	10.9	Upper Marker

Figure 4.8 Analysis of the profile of the sRNA libraries generated using RNA from CLEAR-CLIP. A) Bioanalyzer High Sensitivity DNA analysis of the libraries (0.25%) of fresh gut tissues (Fresh_D1, Fresh_I1, Fresh_CE1, Fresh_CO1), frozen gut tissues (Frozen_D1, Frozen_I1, Frozen_CE1, Frozen_CO1), germ-free frozen gut tissues (GF_S1, GF_L1) and fresh or frozen gut tissues immunoprecipitated using mouse IgG (Fresh_IgG, Frozen_IgG). Size of DNA is indicated in base pairs (bp) on the left. B) Peak table for the Frozen_D1 library displaying the peak size and its corresponding concentration.

Table 4.2 The number of high-quality reads in each CLEAR-CLIP library.

	Fresh gut tissue	No. of reads with the average quality score > 30*	Frozen gut tissue	No. of reads with the average quality score > 30*
Duodenum	Fresh_D1	10,003,201	Frozen_D1	37,408,373
	Fresh_D2	8,799,351	Frozen_D2	15,573,399
Ileum	Fresh_I1	6,816,601	Frozen_I1	23,864,190
	Fresh_I2	7,853,066	Frozen_I2	22,487,737
Cecum	Fresh_CE1	16,756,655	Frozen_CE1	38,304,697
	Fresh_CE2	6,895,411	Frozen_CE2	31,384,867
Colon	Fresh_CO1	8,833,242	Frozen_CO1	12,211,873
	Fresh_CO2	2,361,912	Frozen_CO2	20,575,757
Small intestine	-	-	GF_S1	4,404,312
	-	-	GF_S2	7,863,645
Large intestine	-	-	GF_L1	10,457,732
	-	-	GF_L2	11,514,888
	Fresh_IgG	2,381,112	Frozen_IgG	7,448,285

***** A quality score exceeding 30 implies that the likelihood of erroneously calling a base is less than 1 in 1000, ensuring an accuracy of over 99.9%.

Bioinformatic analysis of the CLEARP-CLIP data was performed by Dr Jose Roberto Bermudez-Barrientos with valuable input from Dr Cei Abreu-Goodger and Dr Sujai Kumar. A summary of the bioinformatic pipeline is summarised in Figure 4.9. Briefly, after initial processing of the data including quality control and adapter removal, the reads were aligned to the mature mouse miRNAs found in miRBase version 22.1, those reads that mapped to the database and whose remaining right part of the read was ≥18 nt were considered as putative miRNA-target chimeras. The second part of the read, hereto referred to as the target portion, was aligned to the mouse genome with Bowtie2 and the aligning part of the read was extracted and later aligned with Shortstack (allowing up to one mismatch). Those reads whose target portion did not align to the mouse genome were aligned to a collection of mouse gut metagenome genes (Xiao et al., 2015). If the target portion of a read did not align to the mouse nor the metagenome genes it was labelled as an unmapped chimeric read.

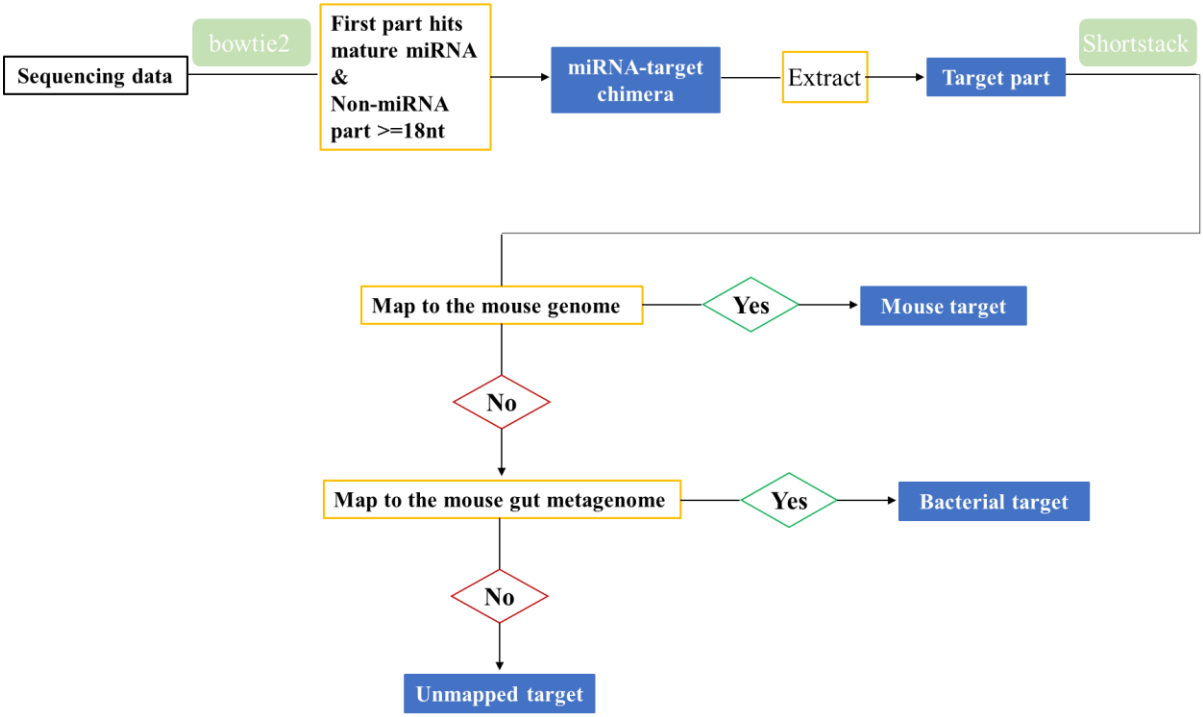


Figure 4.9 Overview of the bioinformatic pipeline to identify bacterial targets. Reads that mapped to the mature mouse microRNAs database and had a remaining right part of the read of ⩾18 nucleotides were considered putative miRNA-target chimeras. Subsequently, the target portion was aligned with the mouse genome and the mouse gut metagenome to identify bacterial target.

We first determined the proportion of chimeric reads (miRNA:target chimeras) and non-chimeric reads in the CLEAR-CLIP datasets (Figure 4.10A). Despite protocol optimization and library size selection efforts, each library contained mostly non-chimeric reads. This aligns with findings from other studies where the proportion of miRNA:target chimeras in reads ranged between 3% to 7% (Moore et al., 2015; Ressel et al., 2023). Libraries prepared with the Ago2 antibody exhibited a significantly higher proportion of chimeric reads, ranging from 2.0% to 8.4%, in contrast to the library prepared with mouse IgG, which contained only 0.1% chimeric reads. We also observed that the percentage of chimeric reads in libraries derived from fresh tissues (ranging from 2.3% to 5.8%) was comparable to those from frozen tissues (ranging from 2.0% to 8.4%).

The chimeric reads were further analyzed and grouped into three categories (Figure 4.10B): miRNA-mouse target chimeras, miRNA-bacteria target chimeras, and miRNA-unmapped target chimeras. Over 80% of the chimeric reads in all libraries are miRNA-

mouse target chimeras. In the libraries from germ-free gut tissues (GF_SI and GF_LI), a small proportion (0.2%-0.5%) of chimeric reads were categorized as miRNA-bacteria target chimeras, suggesting either a background signal of bacterial sequences introduced post-harvest or mis-annotation of some sequences as bacterial. The background signal could originate from the enzymes used in the CLEAR-CLIP protocol or potential laboratory contaminations. Conversely, in the C57BL/6J mice gut tissue libraries, the percentage of miRNA-bacteria target chimeras varied from 0.8% to 5.6%. The libraries from the cecum and colon exhibited a higher proportion of miRNA-bacteria target chimeras compared to those from the duodenum and ileum, aligning with the known distribution of bacteria along the gastrointestinal tract, where cecum and colon harbour more bacteria than the duodenum and ileum. The bacterial target sequences were further categorised into RNA biotypes including rRNA, tRNA, mRNA and other, as shown in Figure 4.10C. We observed a high percentage of reads classified as mRNA across all libraries, ranging from 43% to 89%. Additionally, rRNA was found to be abundant in each library, particularly in the cecum and colon.

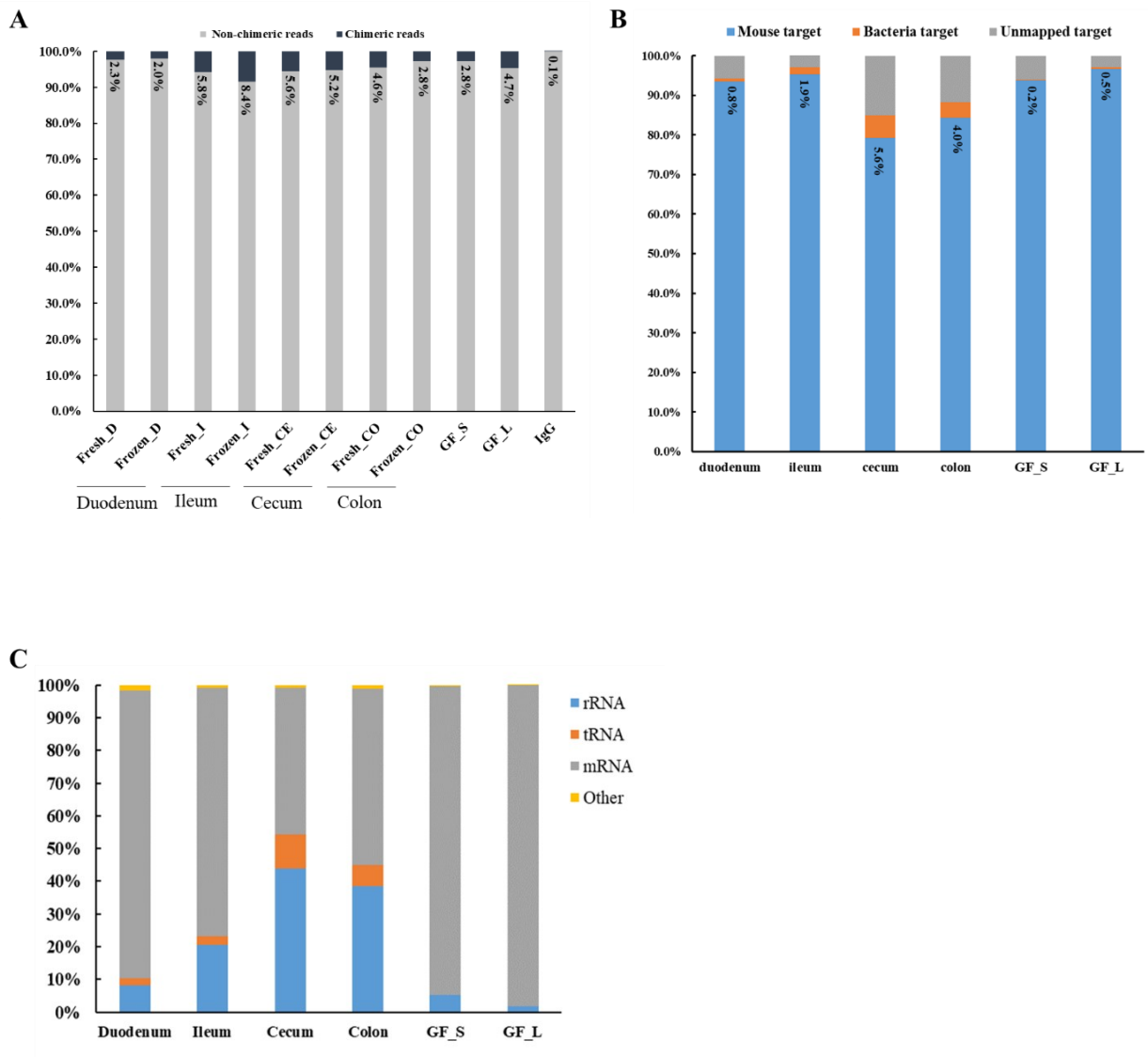


Figure 4.10 Composition of reads sequenced in CLEAR-CLIP libraries. A) Percentage of chimeric and non-chimeric reads. The number in the bar graph is the percentage of chimeric reads. Data represent the average of 2 biological replicates. B) Percentage of miRNA-mouse target chimeras, miRNA-bacteria target chimeras, and miRNA-unmapped target chimeras in each gut segment. The number in the bar graph is the percentage of miRNA-bacteria target chimeras. Data represent the average of replicates from fresh and frozen gut tissues. C) RNA biotypes of the bacterial targets.

To further try to filter for high-confidence targets, we focused on bacterial targets present in two or more replicates of each gut segment. We aimed to distinguish specific targets from potential non-specific background noise, as targets found in only one sample might have a lower chance of representing a real miRNA target. The percentages of reproducible bacterial targets in each gut segment are shown in Figure

4.11A. Across our datasets, the percentage of bacterial targets demonstrating reproducibility varied from 3.3% in the duodenum to 9.1% in the cecum. The RNA biotypes of the reproducible bacterial targets are presented in Figure 4.11B. In the duodenum, ileum, and colon, there are more candidate target reads mapping to mRNA compared to rRNA. However, in the cecum, the trend is reversed, with a lower presence of mRNA compared to rRNA.

Using the reproducible bacterial target reads, I examined the potential origin of the reads in terms of which bacterial species were represented. Figure 4.12 illustrates the bacterial composition in each gut segment, classified at a family-level taxonomy. Across the gut segments, a significant portion of the bacteria is classified as uncategorized. When referring to bacteria as uncategorized, it indicates that these bacteria could only be classified above the family-level taxonomy. Bacterial RNAs from the *Clostridiaceae* and *Bacteroidaceae* families are found in chimeric reads across all gut segments. Additionally, each gut segment has a distinct ranking of bacterial families represented from the chimeric reads.

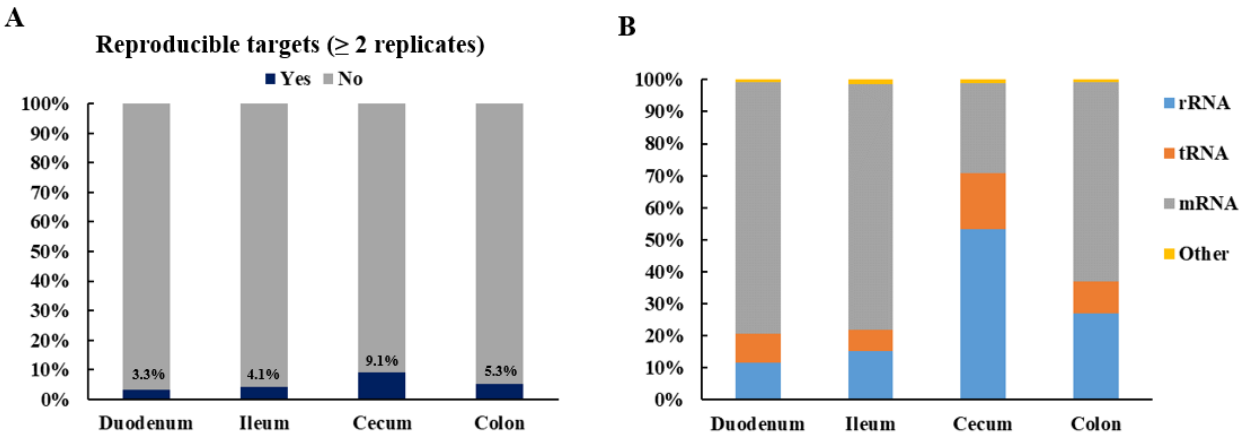


Figure 4.11 Reproducible bacterial targets in each gut segment. A) Percentage of reproducible bacterial targets. No = non-reproducible target (only found in one replicate), Yes = reproducible target (found in ⩾ 2 replicates). Number in bar is the percentage of reproducible targets (found in ⩾2 replicates). B) RNA biotypes of the reproducible bacterial targets. Data represent the average of replicates from fresh and frozen gut tissues.

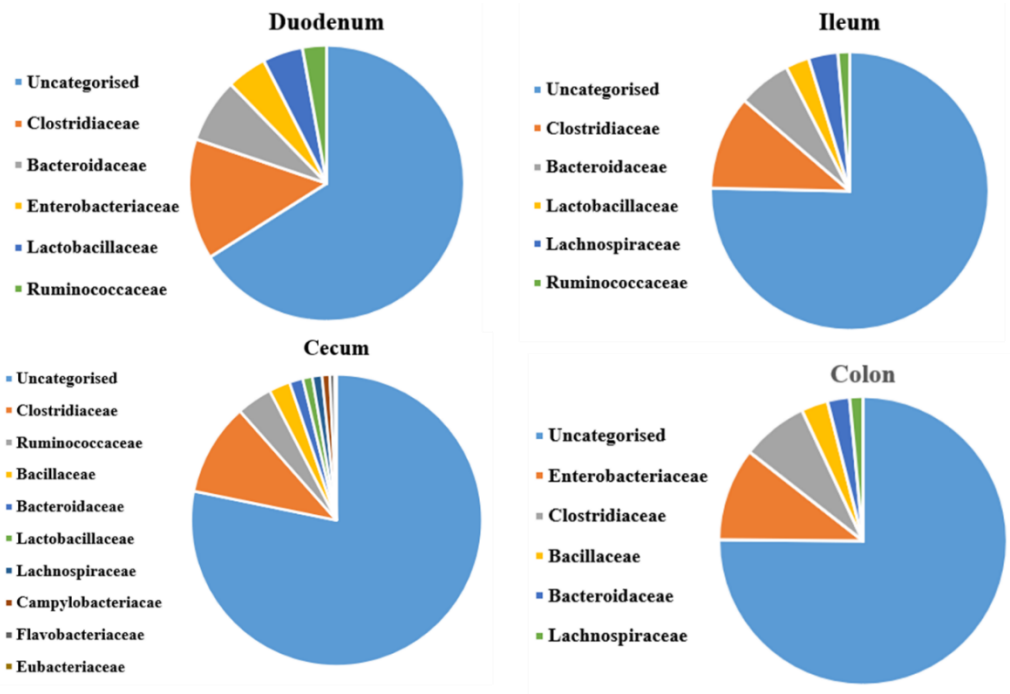


Figure 4.12 Bacterial species identified from the reproducible bacterial targets in each gut segment. Bacteria species are classified at a family-level taxonomy and the order of taxa in the legend corresponds to the sequence of taxa in the pie chart.

4.3 Discussion

4.3.1 Summary

In this chapter, I investigated whether the mouse Ago2 protein is detectable within gut microbiota and adapted a method to identify bacterial genes directly targeted by miRNA-Ago2 complexes. The detection of Ago2 was performed in PGM using western blot analysis and immunofluorescence imaging and my results suggest that there is no consistent Ago2 signal in PGM. For the identification of bacterial target genes, I adapted the CLEAR-CLIP in the gut tissue to capture Ago2 and isolate associated RNAs. I identified potential bacterial targets co-purifying with Ago2 in each gut segment. The analysis of reproducible bacterial targets suggests miRNA-bacteria interactions could occur in each segment.

4.3.2 The limitations of Ago2 detection experiments in PGM

I conducted anti-Ago2 western blot analysis on samples from C57BL/6 mice, including gut tissue, gut contents, faeces, and PGM. The results indicated that Ago2 protein was present in gut tissue, gut contents, and faeces. However, the Ago2 signal in PGM was not consistently detected and was observed only once. To complement these findings, I employed immunofluorescence imaging to visualize Ago2 within individual gut bacterial cells and I did not detect Ago2 signal in PGM. These data suggest that Ago2 protein may be not present in the gut bacteria within the PGM samples or its level is below the detection limit of the assay. However, the limitation of these data is that they were obtained using PGM. During the collection of PGM from gut contents, bacteria residing in the mucus layer, which are more likely to interact with host cells due to their proximity, were not included. Moreover, the 16S rDNA sequence-based survey conducted in Chapter 3 suggested that certain bacterial species may not survive the purification process, leading to alterations in the microbiome composition of PGM compared to gut contents samples. Since PGM does not encompass all bacterial species in the mouse gut, it remains possible that Ago2 protein could transfer into some gut bacteria. To mitigate this limitation, two potential approaches could be considered. Firstly, investigating methods aimed at collecting and purifying gut bacteria specifically from the mucus layer. Secondly, performing immunofluorescence imaging on gut tissue sections that encompass the entire spectrum of gut bacteria.

4.3.3 Identification of bona fide bacterial targets

CLEAR-CLIP was performed in gut tissues that included all the bacteria in the gut. My data reveals fresh and frozen gut tissues for UV crosslinking have similar recoveries of chimeric reads. Additionally, the cecum and colon libraries contain more bacterial targets compared to the duodenum and ileum, consistent with the distribution of gut bacteria. However, it is important to acknowledge the challenges associated with identifying and validating the genuine bacterial targets of mouse miRNAs when using the CLEAR-CLIP method. One limitation of this method is the occurrence of random ligation events during the CLEAR-CLIP process. To assess this, a previous study conducted by Moore et al., 2015 carried out mixing experiments for example where CLEAR-CLIP was done with equal mass amounts of cross-linked mouse cortex and *E. coli* RNA or a large excess (six-fold) of *E. coli* RNA. They found that in mouse-only control samples, 1% of chimeric reads mapped to the *E. coli* genome, while this percentage increased to 1.9% in equal-mixture samples and 5.2% in excess-mixture samples. These results highlighted that some chimeric reads can be generated due to random ligation events between RNAs that do not naturally interact. I did not do mixing experiments in this study and therefore cannot assess the background that could be present due to random ligation events.

To account for some sources of background that could lead to random ligation, I used germ-free mouse gut tissues as controls to assess the background of bacterial contaminants from the commercial enzymes and the lab environment. Only a small proportion (0.2%-0.5%) of miRNA-bacteria target chimeras were found, which is much lower than those of C57BL/6J mice gut tissues (Figure 4.10A). However, these controls do not evaluate the extent of random ligation between miRNA and gut bacterial sequences within the samples during the experimental process. To distinguish specific targets from potential non-specific background noise, we filtered the bacterial targets that were present in two or more replicates. The percentage of bacterial targets that were reproducible ranged from 3.3% in the duodenum to 9.1% in the cecum (Figure 4.11A).

The RNA biotypes analysis of the reproducible bacterial targets shows mRNA and rRNA are dominant across each gut segment. Notably, in the cecum, there is a higher percentage of target reads mapping to rRNA in comparison to mRNA (Figure 4.11B).

In animals, the canonical targets of miRNAs are mRNA targets (3’ UTR and CDS) (Bartel, 2018). CLEAR-CLIP datasets generated from mouse or human cells show that the majority of target reads map to mRNA and only a small proportion of target reads map to rRNA (Moore et al., 2015; Ressel et al., 2023). While it remains unclear whether miRNA-bacterial target interactions follow the same rules as miRNA-host target interactions, it is rational to consider that mRNA reads possess a higher likelihood of being bona fide targets and caution is required when assessing the rRNA reads.

4.3.4 Future work

To identify the genuine bacterial targets, further efforts are necessary from both experimental and computational sides. On the experimental side, we will introduce a new set of controls that involve mixing germ-free mouse gut tissues with some abundant gut bacteria such as *Bacteroides* and *Clostridium*. This refined control group may help distinguish genuine miRNA-bacteria interactions from those that may arise due to random ligation. Canonically, miRNA targeting relies on the base pairing of the seed region, specifically nucleotides 2-7 within the miRNA, to complementary sites located in the 3’ UTRs of mRNA molecules (Lewis et al., 2005). Therefore, on the computational side, our next step should involve filtering the reproducible bacterial targets (especially for the reads mapping to mRNA) based on the presence of a strong miRNA seed site.

Chapter 5: Transport of RNA from mouse to bacteria via extracellular vesicles: probing specificity and function

5.1 Introduction

To explore other miRNA transport methods beyond Ago2, I then investigated extracellular vesicles (EVs). EVs are small, membrane-bound structures secreted by cells into the extracellular space, have important roles in intercellular communication in mammals (Yin et al., 2018). EVs can also mediate inter-organismal communication in various living systems (Buck et al., 2014), and there is some evidence suggesting that mammalian EVs directly interact with certain bacteria (Hendricks et al., 2021). Given that mammalian EVs from different cell types carry diverse RNA species, including miRNAs, mRNAs, and long non-coding RNAs (Cai et al., 2023), there's a promising avenue for considering EVs as a potential transport mechanism for mouse miRNAs in the gut. I hypothesize that EVs derived from mouse intestinal epithelial cells have the capability to transport mouse miRNAs into bacteria.

To overcome the limitations associated with studying host-bacteria interactions within mouse gut tissue, (where there is a complex mix of possible recipient bacteria for miRNAs/EVs), I decided to explore the interaction between EVs and bacteria through *in vitro* models. In our laboratory, we have access to the Mode-K cell line, which is derived from the duodenum of C3H/HeJ mice and serves as a well-established model for enterocytes in studies related to host-bacteria interactions (Campeau et al., 2012; Luongo et al., 2013). Mode-K cells have also been shown to release EVs that can be readily purified in our lab. The goal is to investigate the interaction between Mode-K EVs and bacteria, as well as to assess whether these EVs could transport mouse miRNAs into bacteria. To ensure a suitable control for Mode-K EVs, I considered previous research indicating that the lipid composition of plant exosome-like nanoparticles can influence their uptake by *Lactobacillaceae* (Teng et al., 2018). In our lab, we also have access to EVs from *Heligmosmoides bakeri*, a gastrointestinal nematode parasite that naturally infects mice (Buck et al., 2014). The surface charge, lipid composition and cargo of sRNA and proteins of *H. bakeri* EVs differ from that of mammalian EVs (Simbari et al., 2016; White et al., 2020) and an open question is whether mouse and nematode EVs would show different target specificities. Therefore, I selected *H. bakeri* EVs as a comparison for Mode-K EVs.

In choosing the most relevant bacteria for this study, I focused on species commonly found in or known to infect the small intestine. *Salmonella enterica serovar* Typhimurium (*S.* Typhimurium), an intracellular pathogen that infects the intestinal epithelia of both humans and animals (Andino and Hanning, 2015), was obtained through collaboration with Dr Prerna Vohra's lab. The *S.* Typhimurium strain SL1344 was used in my study, which is a commonly used wild-type strain (Lathrop et al., 2015). For comparison, I also employed a well-established wild-type strain *Escherichia coli* W3110. *S.* Typhimurium *and E. coli* are closely related taxonomically, belonging to the family *Enterobacteriaceae* (Winfield et al., 2003).

Ultracentrifugation has traditionally been the gold standard for isolating EVs from cells, and our lab previously utilized this method to isolate EVs from Mode-K cells (Simbari et al., 2016). Consequently, I employed the same approach to isolate EVs from Mode-K cells in my project and carried out quality control measures included western blot analysis of the EV-associated marker CD9 (Colombo et al., 2014), particle concentration and size distribution analysis using Zetaview (Bachurski et al., 2019), along with morphology visualization using transmission electron microscopy (TEM). In order to assess Mode-K EVs uptake by bacteria, a method for labelling EVs for detection upon uptake was required. There are various EVs labelling strategies, such as the introduction of fluorescent tags on EV-associated proteins via genetic modification of parent cells (Bian et al., 2020; Matinha-Cardoso et al., 2022), and the use of lipophilic dyes like PKH dyes (Li et al., 2020). Genetic modification of cell lines offers certain advantages, such as non-invasiveness, which helps maintain the integrity of EVs. However, for my project, obtaining genetically modified Mode-K cells would be time-consuming and challenging due to the lack of relevant experience in our lab. Thus, I chose to use lipophilic dyes for EV labelling. This method is straightforward, compatible with biomolecules, and allows versatile functionalization. Although PKH dyes are widely used, they have been associated with dye aggregation during the labelling process, and the aggregate's size is similar to that of EVs, potentially leading to false positives in microscopy imaging (Morales-Kastresana et al., 2017; Dominkuš et al., 2018). To mitigate these issues related to PKH dyes, I employed a membrane-binding fluorescent probe called MemGlow dyes, which have shown excellent EV labelling with no aggregation and minimal particle enlargement (Shimomura et al., 2021).

To explore the potential function of mouse EVs in bacteria, one straightforward assay is to measure the effect of EVs on bacterial growth. For example, in previous literature plant exosome-like nanoparticles were shown to directly promote the growth of *Lactobacillus* rhamnosus (Teng et al., 2018). Measurement of growth curves based on optical density (OD) is a widely employed technique in microbiology to monitor bacterial growth in response to different agents. For example, it is commonly used to assess the growth of bacteria and other microorganisms when exposed to antibiotics and various substances (Bollenbach et al., 2009). It is also employed to investigate how microbes respond to different environmental conditions and stressors (De Silvestri et al., 2018; Rolfe et al., 2012). When designing the bacterial growth experiment, the choice of bacterial culture medium is a critical factor to consider. It has been demonstrated that the type of bacterial growth medium can significantly impact the transcriptome and phenotype of *S.* Typhimurium (Blair et al., 2013). Transcriptomics analyses revealed that when *S.* Typhimurium was cultured in MOPS minimal media as opposed to the nutrient-rich Lysogeny broth (LB) medium, there was a notable increase in the expression of 42 genes associated with amino acid synthesis and 23 genes encoding ABC transporters.

In this chapter, I begin by isolating EVs from Mode-K cells and examining whether there is selective miRNA packaging by comparing the miRNA compositions of Mode-K EVs and EV-depleted supernatants. Then I investigate the uptake of EVs derived from Mode-K cells and *H. bakeri* by SL1344 and *E. coli* W3110 to explore their uptake specificity. The role of EVs in bacterial function is assessed by studying their impact on the growth of these bacteria in two different types of growth media.

5.2 Results

5.2.1 Isolation and characterization of Mode-K EVs

Mode-K EVs were isolated using an ultracentrifugation method previously developed and tested in the Buck lab for this cell line (Figure 5.1A). To mitigate the potential contamination from serum proteins and serum-derived EV, I transitioned from the DMEM culture medium supplemented with 10% fetal bovine serum to serum-free advanced DMEM medium when cells reached 70%-80% confluence. After 24-36 hours of incubation, I collected the culture media and centrifuged it to remove cell debris. The resulting supernatants were further filtered through 0.22 µm filter units and concentrated to approximately 12 ml. A small portion of this concentrated material (e.g., 200 µl) was saved as an “Input” sample, while the remaining material was ultra-centrifuged at 100,000g, 4°C, to collect the EV pellet. The EV-depleted supernatant from the first ultracentrifugation was also saved.

To assess the purity of Mode-K EVs, I examined the expression of the EV-associated marker CD9 and the ubiquitous eukaryotic protein β-actin. The total protein content of samples was quantified using the Qubit Protein Assay kit and equal amounts (3 µg) of protein from Mode-K cell lysates, Input, EVs, and EV-depleted supernatants were used for western blot analysis. In Figure 5.1B, I observe that the EV-associated marker CD9 is enriched in EVs compared to the cell lysate, Input, or EV-depleted supernatants. In contrast, β-actin exhibits a strong signal in the cell lysate but no signal in EVs and EV-depleted supernatants. Particle concentration and size distribution of Mode-K EVs were measured using Zetaview (Figure 5.1C). It reveals that the majority of Mode-K EVs are approximately between 90-200 nm in diameter, with a peak size diameter of 123.8 nm. The morphology of Mode-K EVs was visualized using TEM, which display a characteristic cup-shaped appearance (Figure 5.1D).

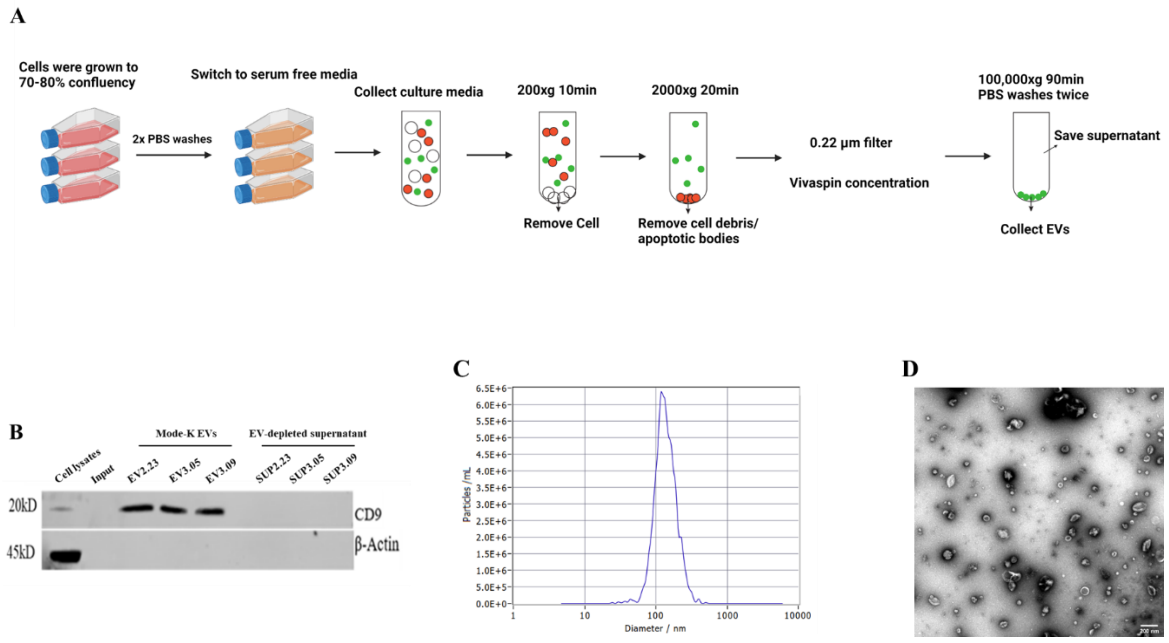


Figure 5.1 Isolation and characterization of Mode-K EVs. A) Workflow illustrating the isolation of extracellular vesicles (EVs) from an *in vitro* Mode-K cell culture using ultracentrifugation. B) Western blot analysis of EVs marker CD9 and β-actin in Mode-K cell lysates, Input, EVs, and EV-depleted supernatant. Equal amounts (3 µg) of protein were loaded onto 4-12% NuPage Bis-Tris polyacrylamide gel. C) Representative size distributions of Mode-K EVs. Zetaview measurement was performed using scattered light. D) A representative TEM image of Mode-K EVs. The scale bar is 200nm.

5.2.2 miRNA profiling of Mode-K EVs

To gain a comprehensive understanding of the small RNA (sRNA) content within Mode-K EVs and examine whether there is selective miRNA packaging within Mode-K EVs when cells secrete miRNAs into extracellular space, I profiled and compared the miRNA content of Mode-K EVs and EV-depleted supernatants.

I isolated total RNA from Mode-K EVs and EV-depleted supernatant using the miRNeasy Serum/Plasma Kit. I attempted to quantify the RNA concentration of EVs using Qubit® RNA HS Assay Kits, but the concentrations were below the limits of detection. To assess RNA quality and quantity, I employed the Agilent RNA 6000 Pico Kit. The results from the PicoChip analysis (Figure 5.2A) revealed the presence of 18S and 28S ribosomal RNA in both cell lysates and Mode-K EVs, with small RNAs (<200 nt) present in all samples. Subsequently, I prepared sRNA libraries from Input,

three independent batches of Mode-K EVs (EV2.23, EV3.05, EV3.09) and EV-depleted supernatants (SUP2.23, SUP3.05, SUP3.09). Additionally, a library prepared from advanced DMEM medium RNA was included to assess the background signal from the culture medium. These libraries were prepared using the CleanTag small RNA library prep kit that is designed to reduce adapter-dimer products and therefore is suitable for low-input material (Olivares et al., 2020). Prior to sequencing, I evaluated the profiles of each sRNA library using the High Sensitivity DNA Bioanalyzer chip. The chip results (Figure 5.2B) indicated that all sRNA libraries displayed a prominent band at approximately 150 bp. In order to achieve uniform sequencing depth coverage across all libraries, I combined equal amounts of each sRNA library based on the High Sensitivity DNA Bioanalyzer chip analysis, creating a final pooled sample for sequencing. Products falling within the 140-180 bp range corresponding to the size of 20-60 bp sRNAs were size-selected, and the libraries were sequenced using a single-end 100-base pair sequencing run on the Illumina NovaSeq 6000 platform by Edinburgh Genomics.

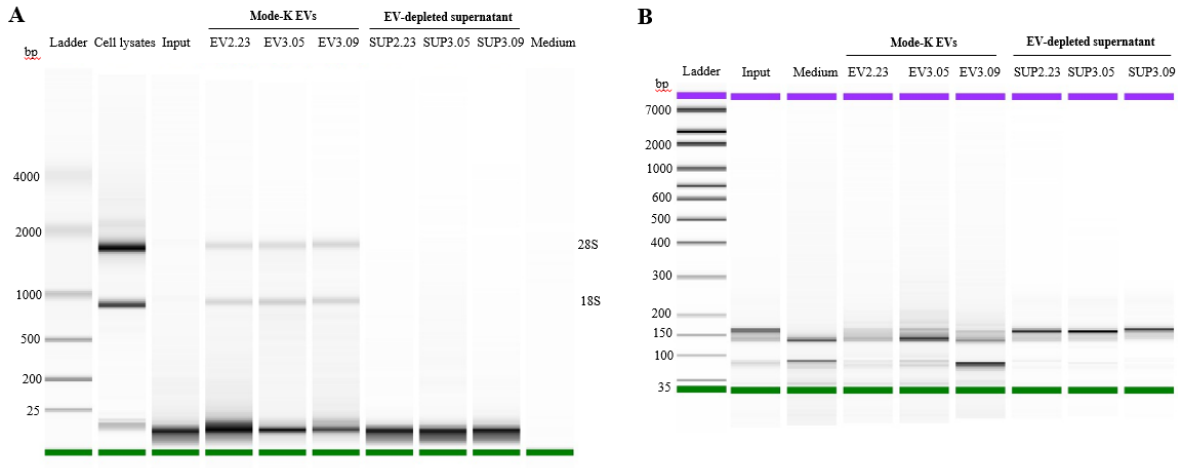


Figure 5.2 Generation of small RNA sequencing libraries of Mode-K EVs. A) Quality of extracted RNA was assessed using the Agilent Bioanalyzer RNA Pico 6000 assay. B) Bioanalyzer High Sensitivity DNA analysis of the libraries (0.25%) of Input, Medium, Mode-K EVs (EV2.23, EV3.05, EV3.09) and EV-depleted supernatants (SUP2.23, SUP3.05, SUP3.09). Size in base pairs (bp) is indicated on the left of the gel image.

Length distribution and first nucleotide analysis (Figure 5.3) reveal significant differences in the sRNA profiles of EVs and EV-depleted supernatants. The peak size of reads in the EVs libraries is 22 nucleotides long, with the majority beginning with uridine (U). This observation aligns with previous reports indicating that most miRNAs are approximately 22 nt in length and start with a 5′-U (Seitz et al., 2011), suggesting an enrichment of miRNA reads in EVs. In contrast, the libraries for EV-depleted supernatants display a 32 nt peak containing U and G-rich sequences, along with a smaller 22 nt peak. The Input library is more akin to the EV-depleted supernatant libraries, but it prominently features a 22 nt peak that starts with U. The Media library exhibits a staircase-like pattern, where the number of reads decreases as the length increases. The bars indicate an even contribution of nucleotides, suggesting a lack of a significant biological signal in this sample. I also observed a much smaller library size in the Medium library (Table 5.1).

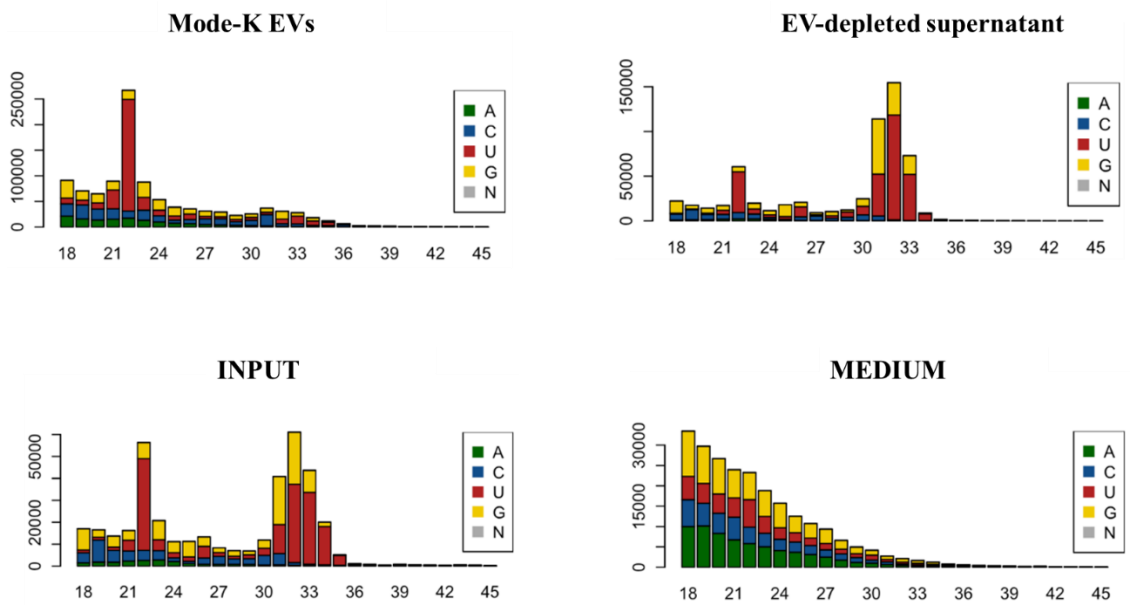


Figure 5.3 Length distribution and first nucleotide plots of the sRNA. The data of Mode-K EVs and EV-depleted supernatant is taken from one representative replicate. The x-axis represents the read length in nucleotides and the y-axis shows the number of reads.

I then analyzed the mouse miRNA content of these samples using the QuickMIRSeq pipeline (Zhao et al., 2017). Adapter sequences were trimmed from the reads and reads shorter than 16 bp were discarded. After trimming and filtering, more than 75%

of the reads from each library (excluding Medium) (Table 5.1) were mapped to miRBase version 22.1 (https://mirbase.org/) to obtain the read counts and RPM values for each miRNA. To illustrate the differences in miRNA composition between Mode-K EVs and EV-depleted supernatants, I did multidimensional scaling (MDS) analysis for these samples (Figure 5.4A). The MDS plot shows that Mode-K EVs cluster together on the left top of the graph while the EV-depleted supernatants cluster on the right top. The clustering of these samples suggests that Mode-K EVs have a distinct miRNA content compared to EV-depleted supernatants. The top 10 abundant miRNAs in Mode-K EVs are displayed in Figure 5.4B. Notably, seven of these miRNAs belong to the let-7 family, with mmu-let-7c-5p and mmu-let-7b-5p being the most abundant. Differential expression analysis of miRNAs in Mode-K EVs and EV-depleted supernatants was performed on Degust (Powell et al., 2019) using Voom/Lima method. False discovery rate (FDR) of less than 0.05 was used to identify significantly changing miRNAs. A total of 72 miRNAs (FDR < 0.05) were differentially expressed in Mode-K EVs compared to EV-depleted supernatants (Supplementary Table 5.1). Of these 43 were upregulated and 29 downregulated in Mode-K EVs. Differential expressions of the top 10 abundant miRNAs were showed in Figure 5.4C and let-7c-5p, let-7b-5p, let-7i-5p, let-7a-5p, let-7e-5p, let-7f-5p, let-7d-5p, miR-221-3p, miR-199b-3p are upregulated in Mode-K EVs, miR-99b-5p is downregulated.

Table 5.1 The number of high-quality reads before and after trimming and filtering.

Sample	No. of reads with the average quality score > 30*	No. of reads after trimming and filtering	Percentage of reads passing trimming and filtering
EV2.23	3,171,373	2,391,215	75.4%
EV3.05	3,784,918	2,986,300	78.9%
EV3.09	2,855,720	2,384,526	83,5%
SUP2.23	1,219,216	1,142,405	93.7%
SUP3.05	1,349,695	1,284,909	95.2%
SUP3.09	1,514,708	1,398,075	92.3%
Input	1,329,885	1,207,535	90.8%
Medium	523,720	254,004	48.5%

***** A quality score exceeding 30 implies that the likelihood of erroneously calling a base is less than 1 in 1000, ensuring an accuracy of over 99.9%.

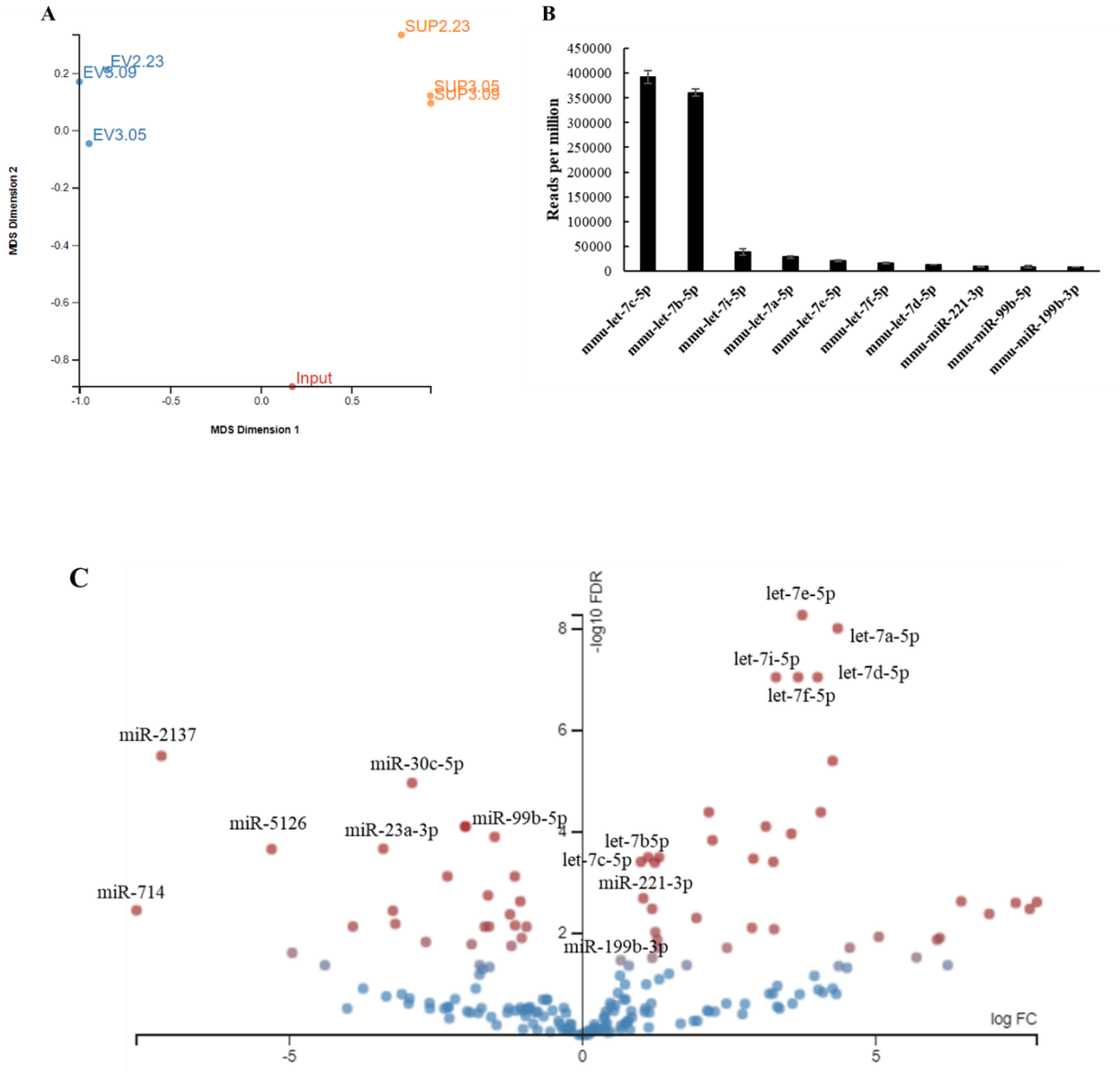


Figure 5.4 Clustering analysis and the composition of miRNAs in Mode-K EVs. A) Multidimensional scaling plot of miRNA contents in Input, Mode-K EVs and EV-depleted supernatants. B) The top 10 abundant miRNAs in Mode-K EVs. Data represent the average of 3 biological replicates. C) Volcano plot of differentially expressed miRNAs in Mode-K EVs compared to EV-depleted supernatants. Differential expression was analysed in Degust (Version 4.1.3). Fold change was plotted against -log10 false discovery rate (FDR). Red dots indicate FDR < 0.05 and blue dots indicate FDR > 0.05.

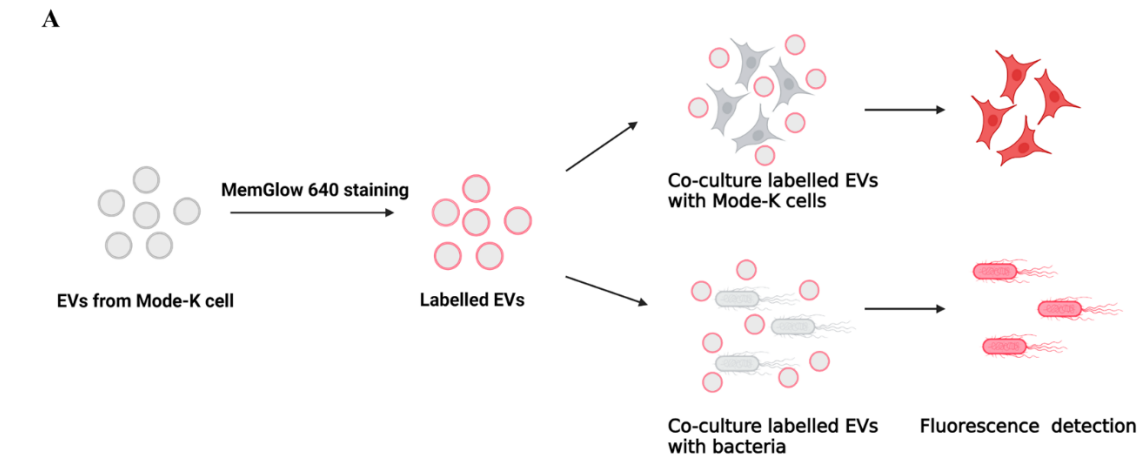
5.2.3 The uptake of Mode-K EVs by SL1344

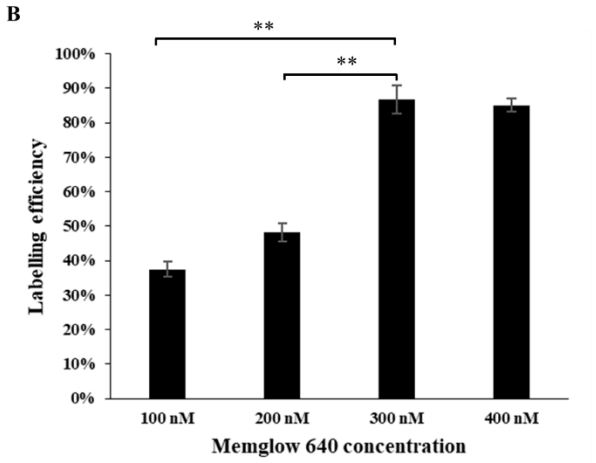
To investigate whether Mode-K EVs could be taken up by bacteria, I utilized MemGlow™ 640 to label EVs for detection upon uptake into bacteria, as outlined in Figure 5.5A. MemGlow™ 640 is a fluorogenic membrane probe (Ilahibaks et al., 2023). Our initial step involved optimizing the dye to EVs ratio to achieve a high labelling efficiency. MemGlow™ 640 to EVs ratios were adjusted by changing the concentration of MemGlow™ 640 while holding the EVs concentration constant. The standard working concentration for MemGlow™ 640 in our lab protocol is 200 nM. To find the ideal ratio, I tested four concentrations around 200 nM, which were 100 nM, 200 nM, 300 nM, and 400 nM. MemGlow™ 640 and Mode-K EVs were initially diluted in PBS and then promptly mixed by pipetting up and down. The particle concentration of EVs in the reaction was $2.2x10^{10}$ particles/ml. Following incubation of MemGlow™ 640 and EVs at room temperature, I separated the labelled EVs from the free dye using ultracentrifugation. This step was crucial because excess dye aggregates could potentially be mistaken for labelled EVs and might interact with recipient cells upon treatment. In order to quantify the EVs, the Zetaview (Particle Metrix) instrument was used. This instrument can directly measure particles based on the scattering of light, and can also measure fluorescence (which allows the measurement of both total particles and the fluorescent particles). Labelling efficiency was calculated as the percentage of fluorescent particles relative to the total particles in the labelled EVs. As illustrated in Figure 5.5B, I observed that the labelling efficiency increased with the concentration of MemGlow™ 640, reaching its highest level (86.8%) at a concentration of 300 nM. Consequently, I utilized 300 nM MemGlow™ 640 for labelling EVs in the subsequent experiments.

I also collected and labelled EVs from *Heligmosmoides bakeri*, a gastrointestinal nematode parasite that naturally infects mice, using the same method. It has been reported that the lipid content of *H. bakeri* EVs differs from that of mammalian EVs (Simbari et al., 2016) which could affect membrane fusion and uptake specificity. Consequently, *H. bakeri* EVs were employed as a comparative reference to understand if there is any specificity in terms of Mode-K EVs interactions with bacteria in subsequent experiments. It was also of interest to understand if the parasite-derived EVs (which are known to interact with mammalian cells) might also have the capacity to target bacteria. To control for background dye signal, PBS was labelled and

subjected to ultracentrifugation alongside EVs. The fluorescent particle concentrations of the labelled EVs for the subsequent co-culture experiments are detailed in Figure 5.5C.

Previous findings from former members of the Buck lab suggested that both Mode-K EVs and *H. bakeri* EVs could be taken up by Mode-K cells. To assess the effectiveness of MemGlow labelled EVs and ability to detect uptake, I treated Mode-K cells with labelled Mode-K EVs, *H. bakeri* EVs, and PBS. Labelled EVs were supplied in the culture medium at the concentration of $1x10^9$ fluorescent particle/ml. To evaluate uptake, I employed fluorescence detection using the EVOS imaging system with a 10× objective and uniform settings. Imaging of Mode-K cells was conducted at 12 hours post-treatment to visualize uptake. The presence of the MemGlow™ 640 fluorescent signal was observed in Mode-K cells treated with both labelled Mode-K EVs and *H. bakeri* EVs, suggesting successful EVs uptake (Figure 5.6).





C

EVs	Fluorescent particles/ml
MK-EV9.15	$2.7x10^{10}$
MK-EV9.21	$3.9x10^{10}$
HB-EV9.17	$4.5x10^{10}$
HB-EV9.22	$3.2x10^{10}$

Figure 5.5 Using fluorescently labelled EVs to study the uptake by bacteria. A) Workflow illustrating the method to study the uptake of EVs by bacteria. Mode-K cells were used to assess the effectiveness of MemGlow labelled EVs and ability to detect uptake. B) Labelling efficiency of Mode-K EVs with MemGlow™ 640. Labelling efficiency was calculated as the percentage of fluorescent particles relative to the total particles in the labelled EVs. Three replicates were performed for each dye concentration. **: p-value < 0.01 (two-tailed t test). C) The concentrations of fluorescent particles in the labelled EVs used for the subsequent co-culture experiments.

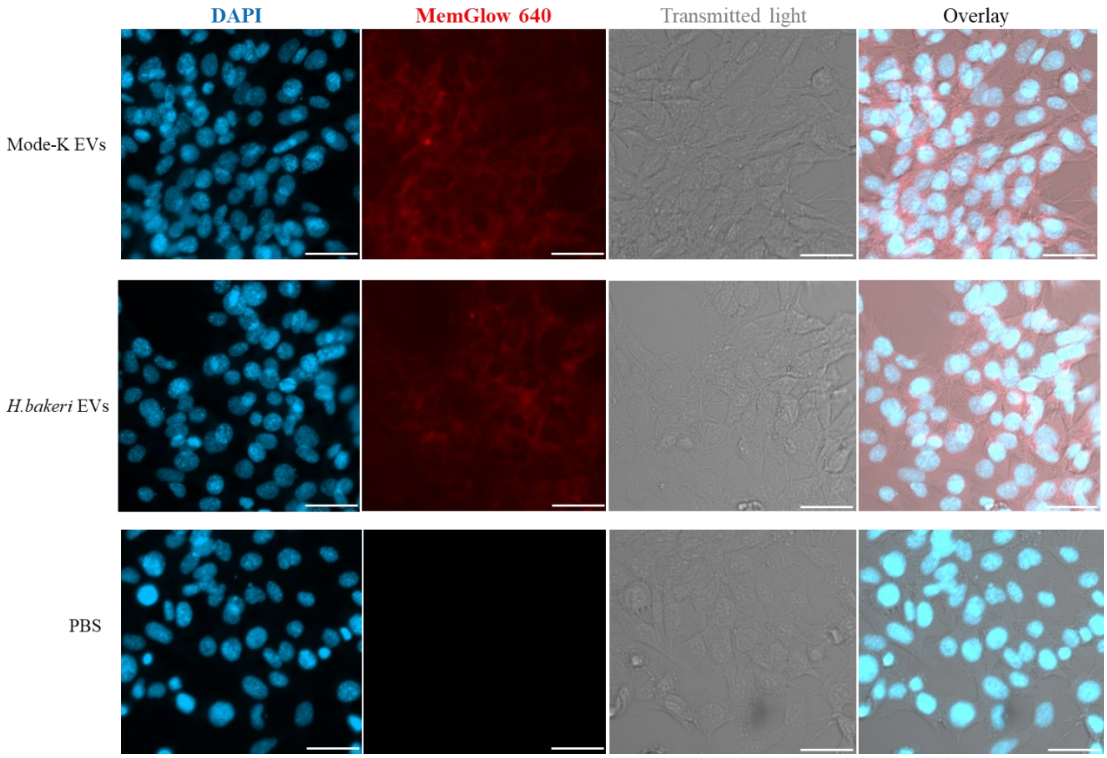


Figure 5.6 The uptake of Mode-K EVs and *H. bakeri* EVs by Mode-K cells. Fluorescence microscopy of Mode-K cells cultured with labelled Mode-K EVs (MK-EV9.15), *H. bakeri* EVs (HB-EV9.17) and PBS for 12 hour. Labelled EVs were supplied in the culture medium at the concentration of $1x10^9$ fluorescent particles/ml. Left to right columns, DAPI staining of cell nuclei (blue), MemGlow™ 640 labelled EVs (red), transmitted light for visualisation of cell boundaries (grey), overlay. The scale bar is 100 µm.

My next objective was to determine whether Mode-K EVs could be taken up by specific bacteria. The bacterial strains used in my study are SL1344 and *E. coli* W3110, which were obtained from Dr Prerna Vohra. An important aspect to consider in the experimental design is the choice of growth media for culturing these bacteria. It is reported that the minimal medium and the nutrient-rich medium can influence *S.* Typhimurium's interactions with host cells (Sridhar and Olivia, 2016). Based on this consideration, I compared two different growth media, M9 minimal medium and the nutrient-rich LB medium, when measuring the effect of EVs on bacterial growth. As SL1344 is a histidine auxotroph due to a mutation in hisG, it is essential to supplement the M9 medium with 40 µg/ml histidine when growing SL1344 (Lathrop et al., 2015). The M9 medium used in my study was supplemented with 40 µg/ml histidine to support the growth of SL1344.

SL1344 and *E. coli* W3110 were cultured in M9 and LB medium and treated with the same batch of labelled Mode-K EVs, *H. bakeri* EVs and PBS. Labelled EVs were supplied in the culture medium at the concentration of $1x10^9$ fluorescent particles/ml. Subsequently, bacteria were imaged using the Zeiss LSM880 Airyscan microscopy at 6 hours post-treatment to visualize uptake. The confocal images (Figure 5.7) show that the MemGlow™ 640 fluorescent signal was only observed in SL1344 treated with labelled Mode-K EVs in M9 medium which suggests that Mode-K EVs were taken up by SL1344. This uptake was not observed when SL1344 was cultured in LB medium or when treated with *H. bakeri* EVs. Furthermore, the uptake of Mode-K EVs and *H. bakeri* EVs was not observed in *E. coli* W3110, regardless of whether they were cultured in M9 or LB medium (Figure 5.8).

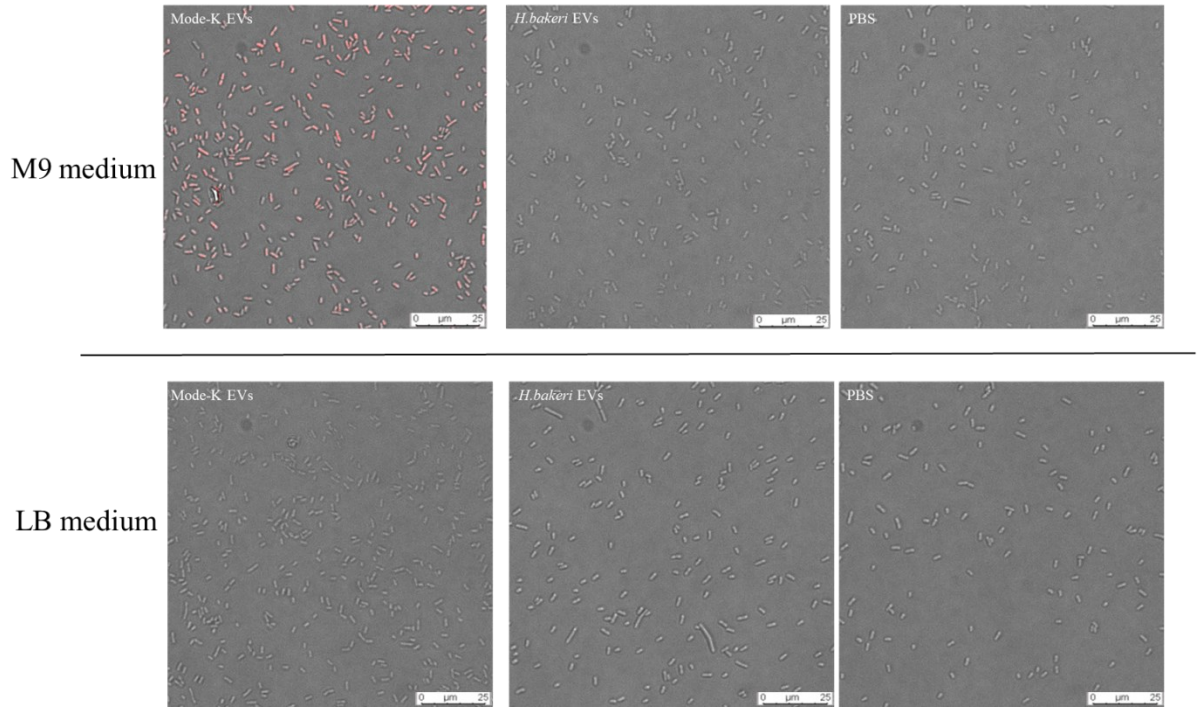


Figure 5.7 Representative confocal microscopy images of SL1344 co-cultured with labelled EVs. SL1344 were cultured in M9 and LB medium and treated with labelled Mode-K EVs (MK-EV9.21), *H. bakeri* EVs (HB-EV9.22), and PBS. Images are overlay of MemGlow™ 640 (red) and transmitted light (grey). The scale bar is 25 µm.

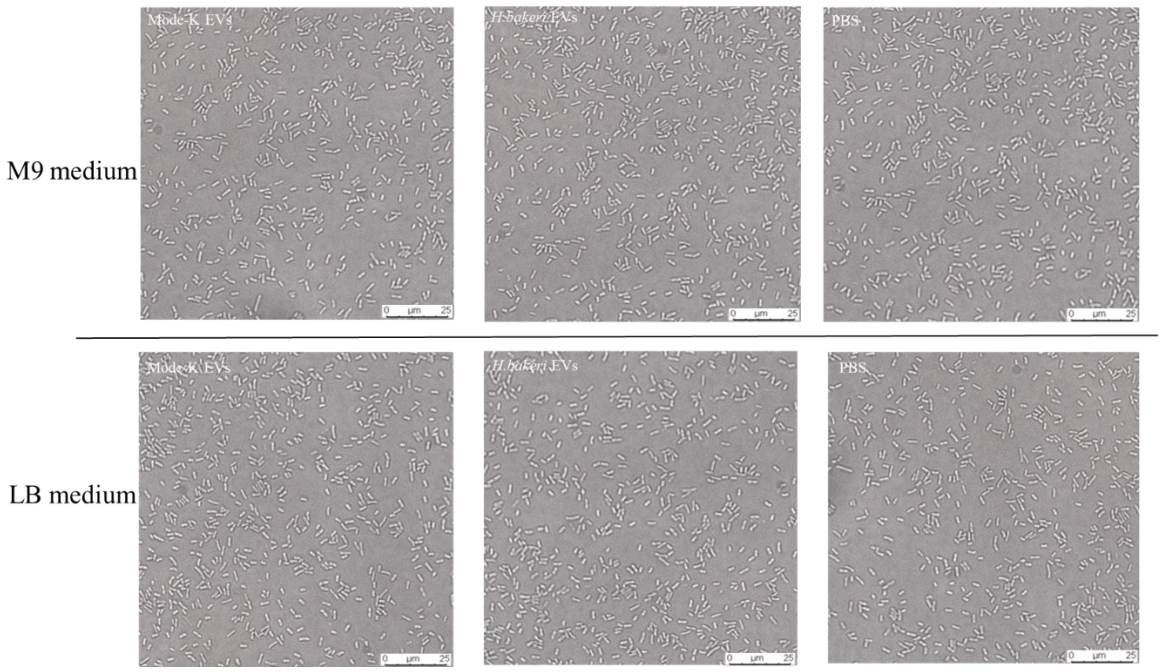


Figure 5.8 Representative confocal microscopy images of *E. coli* W3110 co-cultured with labelled EVs. *E. coli* W3110 were cultured in M9 and LB medium and treated with labelled Mode-K EVs (MK-EV9.21), *H. bakeri* EVs (HB-EV9.22), and PBS. Images are overlay of MemGlow™ 640 (red) and transmitted light (grey). The scale bar is 25 μm.

5.2.4 Mode-K EVs promote the growth of SL1344

After observing the uptake of Mode-K EVs by SL1344 cultured in M9 medium, I began to explore the potential functions of Mode-K EVs on the bacteria. Assessing bacterial growth is a straightforward way to investigate this. In brief, bacteria were cultured in liquid M9 medium, grown aerobically to the log phase at 37°C. These cultures were used to inoculate in 96-well plates and Mode-K EVs and *H. bakeri* EVs were added to the culture at a concentration of 1.5 x10^{10} particles/ml and an equal volume of PBS was used in the control group. Growth was continuously monitored by measuring optical density at 600 nm (OD 600) hourly for up to 16 hours using Tecan Infinite M200 Plate Reader. Optical density values were employed for plotting after blanking with M9 medium. The growth curves (Figure 5.9A) demonstrate that Mode-K EVs promoted the growth of SL1344, while *H. bakeri* EVs and PBS had no such effect. The promotion effect of Mode-K EVs on SL1344 became noticeable in the exponential phase after approximately 6-7 hours of incubation. For comparison, I also assessed the impact of Mode-K EVs and *H. bakeri* EVs on *E. coli* W3110. The growth curves (Figure 5.9B)

revealed that neither Mode-K EVs nor *H. bakeri* EVs influenced the growth of *E. coli* W3110. Although I did not observe the uptake of Mode-K EVs by SL1344 when cultured in LB medium, I conducted a co-culture experiment for SL1344 and Mode-K EVs in LB medium to explore any potential interactions, considering the possibility of limitations in the uptake assay. Figure 5.9C illustrates that Mode-K EVs did not have an impact on the growth of SL1344 in LB medium which suggested that the promotion effect Mode-K EVs on SL1344 may be dependent on the low nutrient condition.

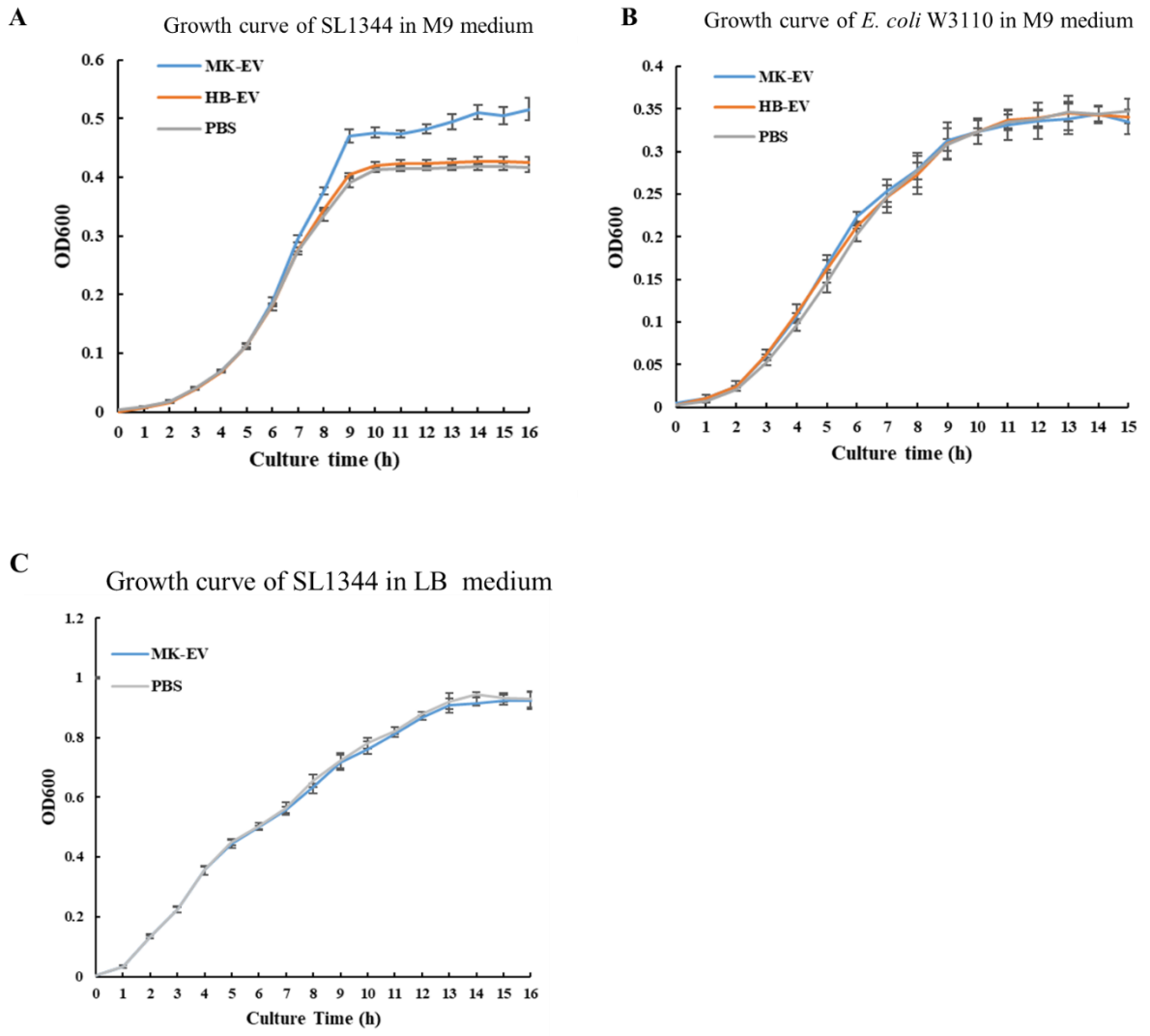


Figure 5.9 Study the functions of EVs on the bacteria by monitoring bacterial growth. A) Growth curve of SL1344 co-cultured with Mode-K EVs (blue), *H. bakeri* EVs (orange) and PBS (grey) in M9 medium. B) Growth curve of *E. coli* W3110 co-cultured with Mode-K EVs (blue), *H. bakeri* EVs (orange) and PBS (grey) in M9 medium. C) Growth curve of SL1344 co-cultured with Mode-K EVs and PBS LB medium. EVs were supplied in the culture at a concentration of 1.5 x10^{10} particles/ml,

and an equal volume of PBS was used in the control group. Bacterial growth was monitored as absorbance at 600 nm (OD600) once per hour for up to 16 hours. OD600 values were employed for plotting after blanking with M9 or LB medium.

Triton X-100 is a common non-ionic surfactant known for its ability to disrupt the phospholipid bilayer of EVs (Tian et al., 2020). To test whether the intact EV is required for the growth promotion of SL1344, I subjected Mode-K EVs to treatment with Triton X-100. I incubated Mode-K EVs with 0.05% Triton X-100 or PBS at 37°C for 30 minutes and then quantified the particle concentrations using Zetaview. The results revealed a significant 271-fold decrease in particle concentration in Triton X-100 treated EVs compared to those treated with PBS (Figure 5.10A). Next, I used these Triton X-100 or PBS-treated Mode-K EVs to co-culture with SL1344. To investigate the role of RNA cargoes within EVs, I added RNase to degrade any free-floating RNAs. The results depicted in Figure 5.10B indicate that Triton X-100-treated Mode-K EVs did not exhibit the same growth-promoting effect on SL1344, and furthermore, RNase treatment did not affect the functionality of the EVs.

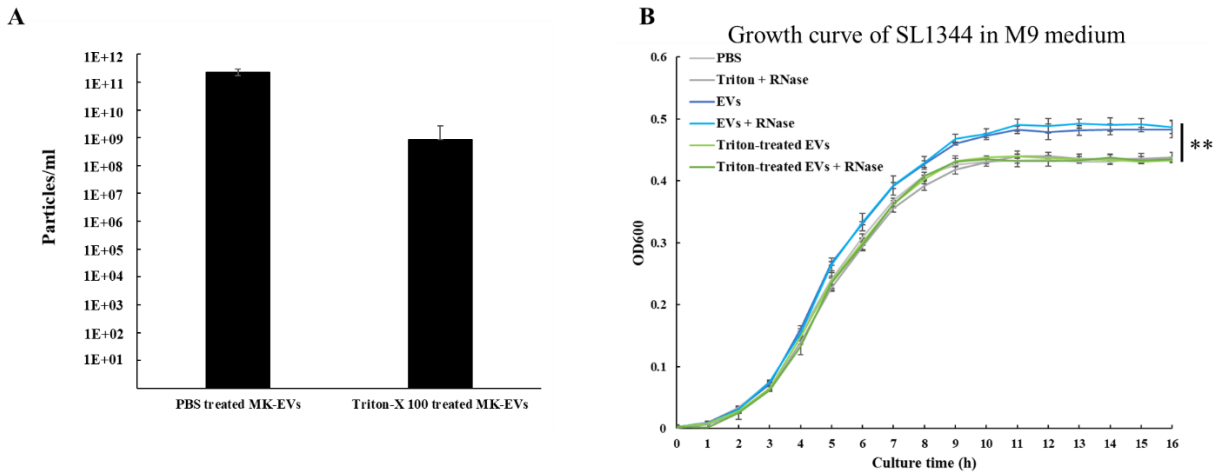


Figure 5.10 The promotion effect of Mode-K EVs was compromised after Triton X-100 treatment. A) Particle concentrations of Triton X-100-treated EVs and PBS-treated EVs measured by Zetaview. B) Growth curve of SL1344 co-cultured with Triton X-100-treated EVs and PBS-treated EVs. EVs were introduced into the culture at a concentration of $1.5x10^{10}$ particles/ml and RNase-IT was added to the culture at a concentration of 0.05 µg/µl. A two-tailed t-test was conducted to analyse the OD600 values between PBS-treated EVs and Triton X-100-treated EVs at hour 16. **: p-value < 0.01.

To investigate whether the promotion effect of Mode-K EVs on SL1344 is dose-dependent, I conducted co-culture experiments using SL1344 and different concentrations of Mode-K EVs. In the previous co-culture experiment, the particle concentration of EVs in the culture medium was 1.5 x 10^{10} particles/ml. For this experiment, I chose to test a higher concentration of 6.5 x 10^{10} particles/ml and a lower concentration of 0.3x10^{10} particles/ml (Figure 5.11A). The growth curves in Figure 5.11B demonstrate a dose-effect relationship. Higher concentrations of EVs corresponded to higher optical density values, indicating a direct correlation between EV concentration and the observed effect.

A

Treatments	Final particles con. of EVs
MK-EV1	6.5x10^{10} particles/ml
MK-EV2	1.5x10^{10} particles/ml
MK-EV3	0.3x10^{10} particles/ml

B

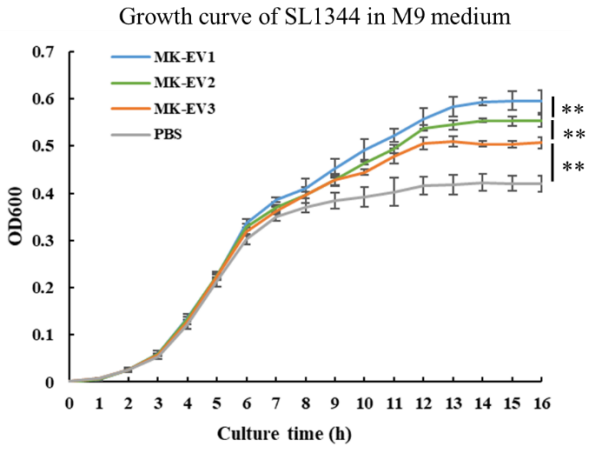


Figure 5.11 Dose effect of Mode-K EVs on the growth of SL1344. A) The dose of Mode-K EVs co-cultured with SL1344. B) Growth curve of SL1344 co-cultured with different dose of Mode-K EVs in M9 medium. EVs were supplied in the culture at a concentration of 6.5x10^{10} particles/ml (blue), 1.5x10^{10} particles/ml (green) and 0.3x10^{10} particles/ml (orange). A two-tailed t-test was conducted to analyse the OD600 values specifically at hour 16. **: p-value < 0.01.

5.2.5 miRNA detection and target prediction in SL1344

From the above work I identified conditions in which Mode-K EVs were internalized by SL1344 and demonstrated an associated effect on bacterial growth that disappears if the EVs are disrupted. To validate the hypothesis that Mode-K EVs transport mouse miRNAs into bacteria, I need to prove that mouse miRNAs can be detected in SL1344 after co-culturing with Mode-K EVs. As showed in Figure 5.4, let-7 family miRNAs are the most abundant miRNAs in the Mode-K EVs (based on sequencing reads). To ascertain the presence of these miRNAs in SL1344, I conducted miRNA RT-qPCR

using let-7 conserved primer. In this context, the presence of a Ct value indicates the target's presence in the reaction, while the absence of a Ct value or an inconclusive Ct value (>34) suggests the absence of the target (Rondaan et al., 2023).

SL1344 were incubated with Mode-K EVs, *H. bakeri* EVs and PBS in M9 medium and after 6 hours bacterial cells were collected and washed with PBS three times to remove any EVs that were not taken up. The total RNA of SL1344 was isolated using Norgen Total RNA Purification Kit and RNA concentration was measured using Qubit RNA HS Assay Kits (Figure 5.12 A). For reverse transcription, 200 ng of RNA was utilized to synthesize cDNA using miScript II RT kit. Quantification of miRNAs was carried out using the miScript SYBR Green PCR kit. miRNAs from the let-7 family were amplified using let-7 primer. Bacterial 16S rRNA gene is commonly used as an internal control for RT-PCR to eliminate sample-to-sample variations due to differences in cell numbers and reverse transcription (Phongsisay et al., 2007). Thus, I also did RT-qPCR with the same amount of sample for 16S rRNA. Figure 5.12B shows the average Ct value of let-7 in SL1344 co-cultured with Mode-K EVs is 25.18, which is much lower than those in SL1344 with *H. bakeri* EVs and PBS (38.41 and 39.68, respectively). For qPCR, Ct-values between 34 and 40 were considered as inconclusive, and over 40 were considered negative (Rondaan et al., 2023). The average Ct values of 16S rRNA are comparable across all groups suggesting the similar amount of bacteria cell used for RT-qPCR. These results indicate the presence of let-7 family miRNAs in SL1344 co-cultured with Mode-K EVs.

A

Samples	RNA concentration ng/µl
SL1344 (MK-EV)	183
SL1344 (HB-EV)	151
SL1344 (PBS)	176

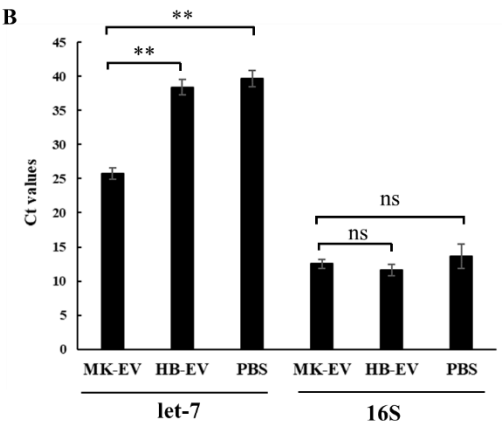


Figure 5.12 The detection of let-7 family miRNAs in SL1344 by RT-qPCR. A) RNA concentrations of SL1344 co-cultured with Mode-K EVs (MK-EV), *H. bakeri* EVs (HB-EV) and PBS. B) RT-qPCR analysis of let-7 family miRNAs and 16S rRNA. Data

represent the average Ct value of three replicates ± S.D. **: p-value < 0.01, ns: not significant (two-tailed t test).

Top 10 abundant miRNA sequences in Mode-K EVs were used to explore potential bacterial gene targets of host miRNA using the reference genome *Salmonella enterica subsp. enterica* serovar Typhimurium str. LT2 (McClelland et al., 2001). Bacterial genes lack introns; therefore, alignment of miRNA sequence to a bacterial gene is predictive of miRNA-mRNA complementary to the bacterial mRNA transcript. miRNA target prediction was performed by Dr Jose Roberto Bermudez-Barrientos using the software TargetFinder (Bo and Wang, 2005) allowing a maximum penalty score (-c) of 6, searching for potential target sites in the sense DNA strand (-r) and writing the output as a table format (-p table). Predicted target description for each miRNA were listed in Table 5.2. Overall, 38 targets of the top 10 abundant miRNA were found in the genome of *S.*Typhimurium str. LT2. Of these targets, 25 targets are let-7 family miRNA targets. mmu-let-7a-5p, mmu-let-7e-5p and mmu-let-7f-5p shared the same target which encode Fis family transcriptional regulator. Fis is one of the most important nucleoid-associated proteins, and functions as a global regulator of transcription in bacteria (Wang et al., 2013).

Table 5.2 Predicted target descriptions in *S.* Typhimurium for the top 10 abundant miRNA of Mode-K EVs.

miRNA	Bacterial targets description	Binding site_5'-3'	Position in the genome	Penalty score
mmu-let-7c-5p	Fimbrial biogenesis outer membrane usher protein	GACCAUGAUGAUCUACUACAUCA	2244011-2244033	5
mmu-let-7c-5p	Unknown	AAUUACACAUCCUACUAUCUCC	553494-553515	5
mmu-let-7b-5p	Unknown	AAUUACACAUCCUACUAUCUCC	553494-553515	4
mmu-let-7b-5p	Diguanylate cyclase	AACCACACUAUCUGUUGCGUCA	2209060-2209081	6
mmu-let-7i-5p	Phosphatidylglycerol--membrane-oligosaccharide Glycerophosphotransferase	AAUGGCCCAAAAUAUUACCUUC	4799776- 4799797	6
mmu-let-7i-5p	iron export ABC transporter permease subunit FetB	AACAGCAUAAAC-GUUACCAUA	563809- 563829	6
mmu-let-7a-5p	Fis family transcriptional regulator	AACUAUAAAAUUU-UUACCUUA	714920-714940	5
mmu-let-7a-5p	Unknown	AAUUACACAUCCUACUAUCUCC	553494-553515	4.5
mmu-let-7a-5p	Putative periplasmic protein	AAUUAUACAUUUUGCUUCCUCA	4222978-4222999	5
mmu-let-7a-5p	YceK/YidQ family lipoprotein	GAAUAUACGCACCUGAUACCUCA	1240541-1240563	5.5
mmu-let-7a-5p	Fimbrial biogenesis outer membrane usher protein	GACCAUGAUGAUCUACUACAUCA	2244011-2244033	6
mmu-let-7a-5p	MFS transporter	UAUUAUCUGACCUACAACCUCG	51008-51029	6
mmu-let-7e-5p	Fis family transcriptional regulator	AACUAUAAAAUUUU-UACCUUA	714920-714940	5.5
mmu-let-7e-5p	Serine hydroxymethyltransferase	AACUGUACAACAUCGUACCUUA	2694502-2694523	4.5
mmu-let-7e-5p	Alanine transaminase AlaA	AACUUUAUUACCUUCUUCCUCA	2442077-2442098	5
mmu-let-7f-5p	Fis family transcriptional regulator	AACUAUAAAAUUU-UUACCUUA	714920-714940	4.5
mmu-let-7f-5p	Putative periplasmic protein	AAUUAUACAUUUUGCUUCCUCA	4222978-4222999	4.5
mmu-let-7f-5p	6-phosphofructokinase	GACUACACUAUCGGCUACUUCA	4272745-4272766	5
mmu-let-7f-5p	Fimbrial biogenesis outer membrane usher protein	GACCAUGAUGAUCUACUACAUCA	2244011-2244033	5.5
mmu-let-7f-5p	Protein for signal transmission	AACUAUGCUAAUUUGCUACUUAA	2511050-2511072	5.5
mmu-let-7f-5p	Phosphoglycerate transport regulator PgtC	AACUGUACAAUUU-UUAUAUCA	1197567-1197587	6
mmu-let-7f-5p	Unknown	AAUUCUUCGGAUCUACUAUCUUA	46145-46167	6
mmu-let-7d-5p	Lipopolysaccharide 3-alpha-galactosyltransferase	CGUUAUGCGCUGUACUACCUCU	3913893-3913914	5
mmu-let-7d-5p	Sgma factor-binding protein Crl	AACUACGCGACCUGCUGAUUUCU	364845-364867	6
mmu-let-7d-5p	Unknown	AGUUAUUGCAACCUACUUCCCCU	3716172-3716194	6
mmu-miR-221-3p	GABA permease	GAAAUUAAGCUGAAAAUGUGGCU	2944196-2944218	5
mmu-miR-221-3p	Dihydrolipoamide succinyltransferase	GAAAUCCUGCCGAUGAUGUACCU	805641-805663	5.5
mmu-miR-221-3p	Aminopeptidase N	CAAGCCCGGCGGAAAAUGUA-CU	1146392-1146413	5.5
mmu-miR-221-3p	FUSC family protein	GAAACCCAGCAGAUGACGCAGCA	2134454-2134476	6
mmu-miR-221-3p	Fe-S cluster assembly protein SufD	GGCGCCCAGCAGAAAACUGUGGCU	1455634-1455657	6
mmu-miR-99b-5p	NADP-dependent succinate-semialdehyde dehydrogenase	CGCGGGGCCGGCUUUACGGGUG	2940626-2940647	3.5
mmu-miR-99b-5p	Nitrate reductase subunit beta	GGCAGGGUACGGUAUUCCGGGUG	1665444-1665466	6
mmu-miR-99b-5p	DNA mismatch repair endonuclease MutL	AGCGAGG-CGAUUUUACUGGUG	4601936-4601956	6
mmu-miR-99b-5p	Uptake [NiFe] hydrogenase, small subunit HyaA	CAUAAGG-CGCUUUUGCGGGUC	1884758-1884778	6
mmu-miR-181b-5p	LPS O-antigen length regulator	GAACAACUGGCGGCAAUGAAUGUG	650511-650534	5
mmu-miR-181b-5p	Glutathione ABC transporter ATP-binding protein GsiA	AACCCACC-ACGGCGCUGGAUGUC	918684-918706	5
mmu-miR-181b-5p	Galactose/proton symporter	AACCAGGCCGAUAGCGAUGAAGGUU	3254254-3254278	5.5
mmu-miR-181b-5p	30S ribosomal protein S1	AACUGGCCGACGGCGUUGAAGGUU	1067199-1067222	6

5.3 Discussion

5.3.1 Summary

In this chapter, I investigated the potential of EVs derived from intestinal epithelial cells to act as a transport mechanism for mammalian miRNAs into specific bacteria. To begin to probe any specificity in the uptake of EVs by bacteria, I compared fluorescently labelled EVs obtained from either Mode-K cells or *H. bakeri*. The results indicate that SL1344 took up Mode-K EVs, but not *H. bakeri* EVs, when cultured in M9 medium. This uptake activity was not observed when SL1344 was cultured in LB medium. The uptake of Mode-K EVs was not observed in *E. coli* W3110, regardless of whether they were cultured in M9 or LB medium. Furthermore, my data revealed that Mode-K EVs had a stimulatory effect on the growth of SL1344 when bacteria were cultured in M9 medium. This growth-promoting effect appeared to be dose-dependent, with a higher concentration of Mode-K EVs resulting in greater growth. The co-culture experiments with Triton X-100-treated Mode-K EVs proved that the intact EVs is required for the growth promotion of SL1344.

Analysis of sRNA sequencing data showed that Mode-K EVs had a distinct miRNA content compared to EV-depleted supernatant, with the let-7 family of miRNAs being the most abundant. Additionally, RT-qPCR confirmed the presence of let-7 family miRNAs in SL1344 after co-culturing with Mode-K EVs. These initial findings suggest that EVs have the capability to selectively deliver RNA cargo to specific bacteria.

5.3.2 The mechanism of uptake specificity by SL1344

From the comparison of two different EVs and two bacterial species, my data suggest there is some specificity in the interaction of Mode-K EVs and SL1344. However, the detailed mechanism behind this interaction remains largely unknown. Specifically, it is unclear how Mode-K EVs are recognized by SL1344 and what the uptake pathway is.

In the context of intercellular communication via EVs, the specificity of cell targeting is typically determined by unique interactions between ligands enriched on the surface of EVs and receptors present on the plasma membrane of recipient cells. For example, EVs released by mesenchymal stem cells express specific lipids and cellular adhesion molecules, enabling them to deliver miRNAs specifically into glioblastoma cells (Sharif and Soleimani, 2018). Once EVs have bound to the surface of recipient cells, their

uptake can occur through various mechanisms, including endocytosis, fusion, and signaling (Liu and Cheng, 2023). Endocytosis is a common mechanism through which EVs transmit signals, and it can take place through various pathways, including clathrin-dependent mechanisms, caveolin-mediated uptake, macropinocytosis, phagocytosis, and lipid raft-mediated internalization (Van et al., 2022). Fusion is another mechanism of internalization, in which the membrane of EVs directly fuses with the cell's plasma membrane (Mulcahy et al., 2014). EVs can also initiate intracellular signalling pathways by directly interacting with the surface receptors or ligands present on target cells. This signalling can significantly impact the cellular phenotype through membrane-bound morphogens, such as Wnt and the Notch ligand DII4 (Luga et al., 2012; Yamamoto et al., 2016).

While the study of inter-organismal communication via EVs is still in its early stages, it is likely that a similar mechanism exists for targeting and uptake between eukaryotic EVs and bacteria. For instance, the selective uptake of ginger exosome-like nanoparticles by *Lactobacillus rhamnosus* is determined by the lipid composition of EVs, which are enriched with PAs (Teng et al., 2018). My data has shown that *H. bakeri* EVs, which have a different lipid and surface protein composition compared to Mode-K EVs (Simbari et al., 2016), do not interact with SL1344. Therefore, future research should focus on identifying which components of the lipid profile and surface proteins of Mode-K EVs determine the uptake specificity. Additionally, the interaction between Mode-K EVs and SL1344 was detected when bacteria were cultured in M9 minimal medium but not in LB medium. This suggests that low-nutrient conditions may influence the expressions of proteins related to EV uptake in SL1344.

EVs are a heterogeneous group of bilayer membrane nanoparticles and Lässer et al., 2018 suggested that EVs can be grouped into subpopulations based on their morphology, density, RNA/DNA/protein cargo, and surface molecule. It has been revealed through cryo-TEM that EVs even released by a single cell type exhibit a highly heterogeneous morphology (Zabeo et al., 2017). Findings have underscored those distinct subpopulations of EVs possess unique RNA cargo compositions (Lunavat et al., 2015). Additionally, investigations by Palma et al., 2012 demonstrated differential packaging of miRNAs and the release of distinct EV subpopulations by tumor cells. Consequently, these EV subpopulations are speculated to serve specific biological roles and potentially exert varying effects on recipient cells. In my study, I

utilized a mixture of Mode-K EV subpopulations to investigate uptake, leaving unexplored whether SL1344 shows a preference for specific subpopulations of Mode-K EVs and how these subpopulations function upon interaction with bacteria.

5.3.3 The mechanism of growth promotion

I have observed that Mode-K EVs promote the growth of SL1344, and this promotion effect is dose-dependent. However, the detailed mechanism behind this effect remains unclear.

One hypothesis is that Mode-K EVs provide nutrients to SL1344, thereby affecting its growth. There is evidence to support this idea in other models, such as EVs released from virus-infected respiratory epithelial cells transferring the nutrient iron to *Pseudomonas aeruginosa* and promoting bacterial biofilm growth. Since the promotion effect of Mode-K EVs on SL1344 was observed when bacteria were cultured in M9 minimal medium, it is reasonable to consider the nutrients hypothesis. During the biogenesis process, EVs acquire important bioactive molecules from their donor cells. These include a diverse array of components such as nucleic acids, lipids, glycans, and proteins (Jurj et al., 2020). Recently, there has been increased interest in the metabolomic analysis of extracellular vesicles (EVs). Metabolites encompass a wide range of biologically relevant molecules, typically smaller than 2 kDa in size. This comprises various molecular species, from steroid hormones and lipids to metabolic intermediates generated during nutrient synthesis and the elemental components of major biopolymer classes (Williams et al., 2019). It was shown that cancer-derived fibroblasts utilize EVs as a means to supply nutrients (amino acids) to cancer cells (Zhao et al., 2016). To validate the nutrients hypothesis, it is crucial to conduct a comprehensive analysis of the cargo carried by Mode-K EVs, specifically focusing on identifying potential nutrient compositions that could influence the growth of SL1344.

Another hypothesis is that some of the cargo in the EVs such as miRNAs could directly influence bacterial gene expression and thereby affect bacterial growth. An increasing body of research suggests that miRNAs derived from plants and mammalian cells can indeed regulate bacterial gene expression and impact bacterial growth (Liu et al., 2016; Teng et al., 2018; Santos et al., 2020). let-7 family miRNAs are abundant in Mode-K EVs and were detected in SL1344 by RT-qPCR after co-culturing with Mode-K EVs. I have predicted the targets of the top 10 abundant miRNAs of Mode-K EVs in the

genome of *S.* Typhimurium to explore this hypothesis further. One of the intriguing targets of let-7 family miRNAs is Fis family transcriptional regulator. Studies have explored the global role of Fis in transcriptional regulation within *S.* Typhimurium, revealing its impact on 291 genes during the exponential growth phase (Kelly et al., 2004). Moreover, Fis-binding sites have been identified on various genes like *rpoS* and *gyrB* in *S.* Typhimurium, suggesting Fis' regulatory influence on these specific genes (Hengen et al., 1997). This potential target Fis may shed light on how the let-7 family miRNAs modulate bacterial gene expression, potentially impacting regulatory pathways in SL1344.

5.3.4 Future work

After detecting mouse miRNAs transported by EVs in SL1334 and discovering potential targets of these miRNAs within the *S.* Typhimurium genome by prediction, my next plan involves examining gene expression changes in SL1344 when co-cultured with Mode-K EVs using RNA-Seq. This will help identify the genuine targets of mouse miRNAs.

Chapter 6: Studying the functional transmission of mammalian EV cargo to SL1344 using a Cre-loxP method

6.1 Introduction

In chapter 5 I demonstrated that Mode-K EVs were taken up by SL1344 and these EVs delivered host miRNAs into bacteria. Though I predicted potential targets of host miRNAs in the SL1344 genome, it is challenging to verify their authenticity and understand their functional implications. An alternative method for tracking whether EV cargo can function in a recipient cell was published previously which involves the Cre-loxP system (Ridder et al., 2014). The Cre-loxP site-specific recombination system was first discovered in bacteriophage P1 (Sternberg and Hamilton, 1981). Cre recombinase (Cre) is a 38-kDa DNA recombinase belonging to the family of tyrosine site-specific recombinases and it was derived from the P1 bacteriophage (Meinke et al., 2016). Cre recognizes specific DNA fragment sequences called loxP (locus of x-over, P1) sites and mediates the site-specific deletion of DNA sequences positioned between two loxP sites (Sauer and Nancy, 1988) (Figure 6.1). The loxP site comprises a 34 bp sequence with two 13 bp inverted and palindromic repeats and an 8 bp core sequence. The Cre-loxP system has been widely used for mammalian gene editing including deletions, insertions, translocations, and inversions at specific sites in the DNA of cells (Thomson and Ann, 2015).

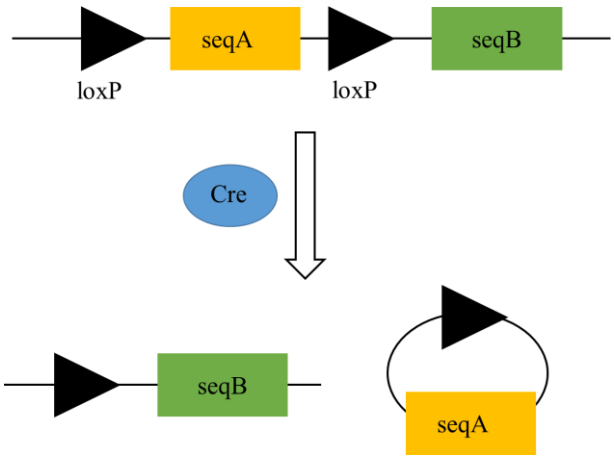


Figure 6.1 An overview of Cre-loxP system. Cre recombinase recognizes the loxP sites and excises flanked loci and deletes the gene seqA.

Several reports in the last decade have used the Cre-loxP system to study the functional transmission of EV cargo within mice (Ridder et al., 2015; Zomer et al., 2016; Rufino-Ramos et al., 2023). This technique involves fluorescently marking Cre-reporter (reporter+) cells that uptake EVs released by Cre-expressing cells (Cre+ cells). Previous work showed that Cre mRNA was reliably detected in EVs by RT-PCR however the levels of Cre protein were below the detection level of western blot (Ridder et al., 2015; Zomer et al., 2015). The schematic of this system is shown in Figure 6.2. The donor Cre+ cells can secrete EVs that carry Cre mRNA. If Cre+ EVs are taken up by reporter+ cells, the mRNA is translated and then protein imported into the nucleus, the Cre recombinase will remove the loxP-flanked DsRed-stop cassette, causing the reporter+ cells to transition from expressing DsRed to enhanced green fluorescent protein (eGFP). This technique can be applied both *in vitro* and *in vivo* to examine functional EV transfer. In an *in vitro* setup, Cre+ and reporter+ cells can either be co-cultured or physically separated by a Transwell membrane that allows EV passage but not cell movement (Zomer et al., 2016). The Transwell assay helps explore EV transfer in a manner independent of direct cell-to-cell contact. Quantification of EV transfer from Cre+ to reporter+ cells in both co-culture and Transwell experiments can be conducted by measuring the percentage of reporter+ cells displaying Cre activity through methods like flow cytometry and fluorescence microscopy. For *in vivo* investigations, having both Cre+ and reporter+ cells within the same animal can be achieved using transgenic Cre+ and reporter+ mouse models or by co-injecting these cell types in the same mouse (Ridder et al., 2014; Ridder et al., 2015). Studies have also examined local and systemic EV exchange between Cre+ and reporter+ breast tumor cells by injecting a mixture of these cells into the mammary gland or by separately injecting the two populations into contralateral mammary glands (Zomer et al., 2015).

This technique was recently applied to explore the transfer of functional biomolecules from outer membrane vesicles (OMVs) of bacteria to mouse cells (Bittel et al., 2021). In this study, they detected both Cre protein and mRNA in OMVs isolated from Cre-expressing *E. coli*. Uptake of *E. coli* derived OMVs was observed to occur by epithelial cells via endocytosis *in vitro and in vivo* leading to the expression of the fluorescent reporter gene in Rosa.tdTomato reporter organoids or mice.

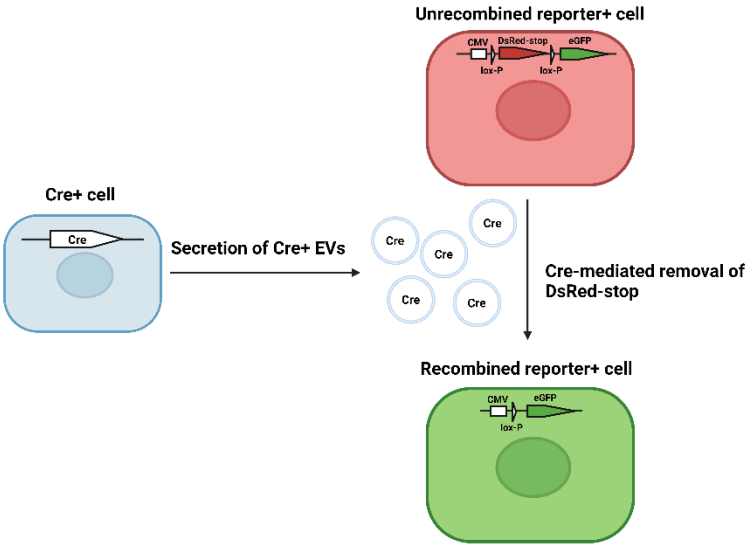


Figure 6.2 Graphical illustration of the Cre-loxP method for studying EV-mediated functional cargo transfer between two experimentally defined cell types. Cre recombinase remove the loxP-flanked DsRed-stop cassette, causing the reporter+ cells to transition from expressing DsRed to eGFP.

This chapter aims to investigate whether Mode-K EVs have the capability to transport functional cargo into SL1344 using the Cre-loxP system. To accomplish this, I plan to create Cre-expressing Mode-K cells with a CFP marker via plasmid transfection to produce Cre+ EVs. Additionally, I will introduce loxP sites into the bacteria to generate reporter+ SL1344. Then I will co-culture Cre+ EVs and reporter+ SL1344. If Cre+ EVs are taken up by reporter+ *Salmonella*, a fluorescent signal will be detected in reporter+ *Salmonella*, indicating the successful transmission of functional cargo.

6.2 Results

6.2.1 Experimental setups

The experimental setups for studying the functional transmission of cargo to SL1344 by Mode-K EVs are described in Figure 6.3, divided into three main objectives. Objective 1 is to establish reporter+ SL1344 by transforming plasmid pBbAW4k-loxP-TT-loxP-mRFP1 and confirm its ability to report Cre activity upon the introduction of Cre protein or mRNA. Objective 2 focuses on establishing Cre+ Mode-K cells to produce Cre+ EVs and analysing EV cargo for the presence of Cre. Objective 3 involves co-culturing Cre+ EVs and reporter+ SL1344 to determine whether Cre is functionally transferred to bacteria using fluorescence as the readout.

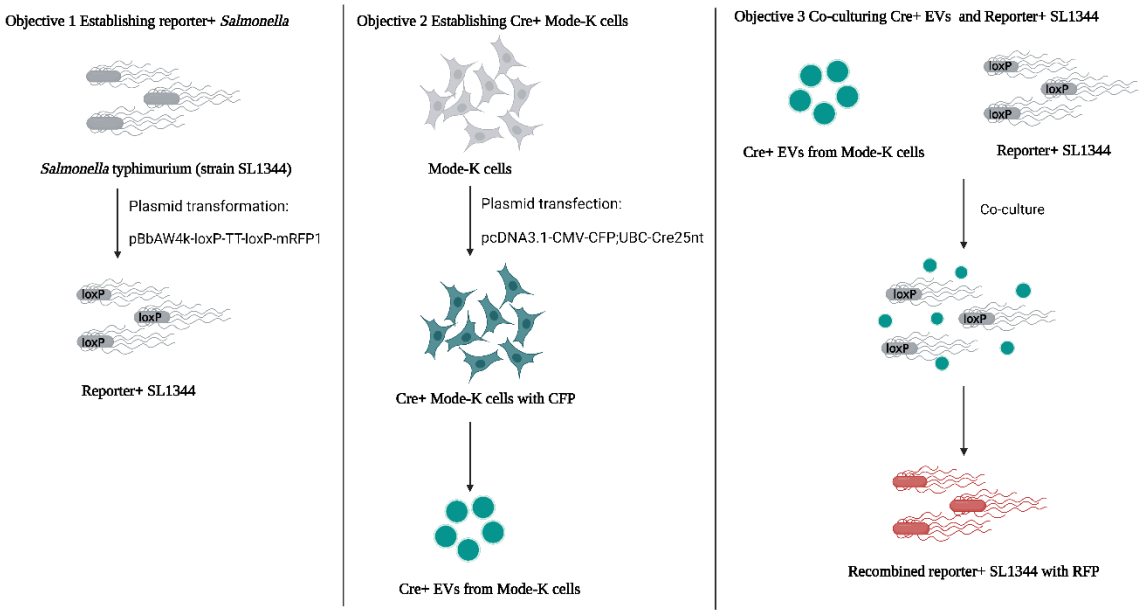


Figure 6.3 Cartoon of experimental setups to study the functional transmission of cargo to SL1344 by Mode-K EVs. The functional transfer can be indicated by the detection of recombined reporter+ SL1344 with red fluorescent protein (rfp) in objective 3 when co-culturing the reporter+ SL1344 obtained from objective 1 and Cre+ EVs obtained from objective 2.

6.2.2 Establishing reporter+ SL1344

To establish the reporter+ SL1344, I used electroporation for transferring plasmid pBbAW4k-loxP-TT-loxP-mRFP1 into bacteria and the transformed cells was enriched by antibiotic selection (Figure 6.4). Since Mode-K EVs were not taken up by *E. coli*

W3110, I also established reporter+ *E. coli* W3110 which was used as a negative control for Cre+ EVs and reporter+ bacteria interaction.

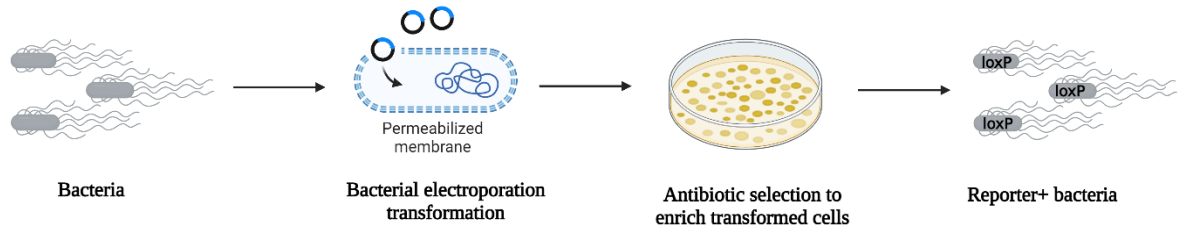


Figure 6.4 Workflow of establishing reporter+ bacteria. Plasmid carrying loxP loci was transformed into bacteria by electroporation and the transformed cells was enriched by antibiotic selection.

The plasmid pBbAW4k-loxP-TT-loxP-mRFP1 (P134405) was ordered from Addgene to generate reporter+ bacteria and the plasmid map and annotation are described in Figure 6.5. This plasmid is a medium-copy p15A origin plasmid with a kanamycin resistance cassette and the gene for red fluorescent protein (rfp) is under the control of a constitutive, medium-strength promoter. There is a transcription terminator flanked by loxP sites between the rfp gene and the promoter. In the absence of Cre, the terminator prevents transcription of rfp. Functional Cre can excise DNA fragments between loxP sites, leading to rfp production.

To test the ability of the reporter to operate in SL1344 and *E. coli*, I introduced into the reporter+ bacteria another plasmid P62730 which can express Cre recombinase protein in bacterial cells (D'Astolfo et al., 2015). Plasmid P62730 has an ampicillin resistance cassette and the Cre gene is under the control of the T7 promoter (Figure 6.6). Bacteria were cultured in M9 medium overnight and analyzed by microscopy. The confocal images show that rfp was detected in both SL1344 (Figure 6.7A) and *E. coli* (Figure 6.7B) that contained both the reporter plasmid P134405 and the Cre plasmid P62730 but was not detected SL1344 and *E. coli* that only contained the reporter plasmid P134405. These data demonstrate that there is no background recombination in reporter+ SL1344 and *E. coli* and that the introduction of Cre recombinase does indeed lead to the expected fluorescence.

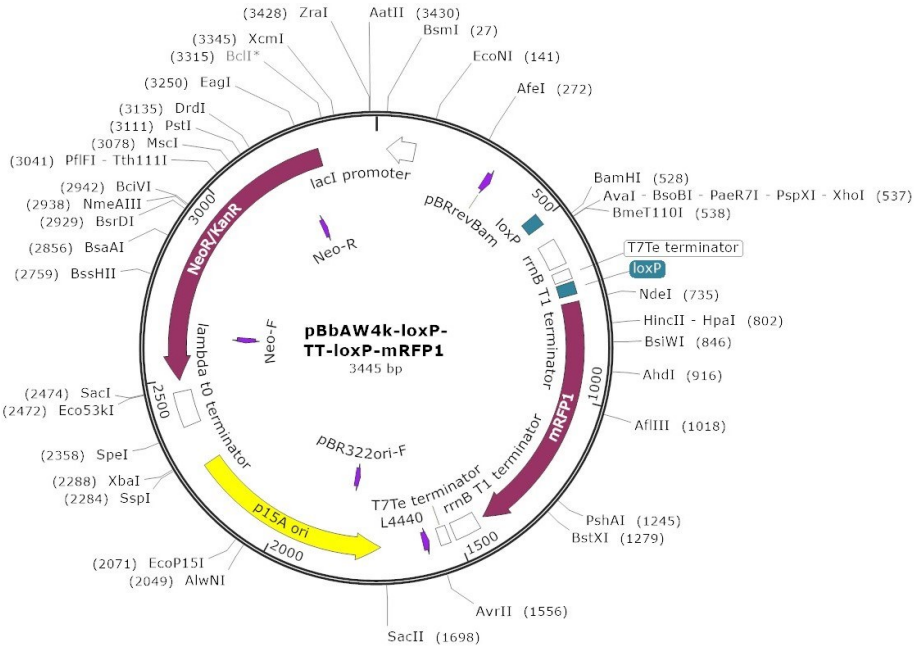


Figure 6.5 Plasmid map and annotation of pBbAW4k-loxP-TT-loxP-mRFP1. The rrnB T1 terminator flanked by loxP sites is positioned upstream of the rfp gene.

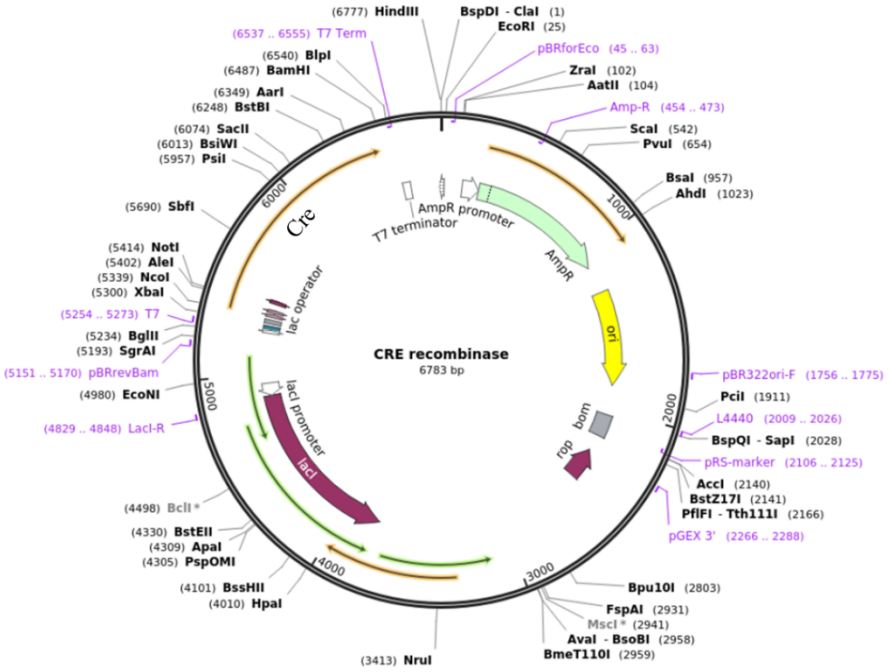


Figure 6.6 Plasmid map and annotation of plasmid P62730. Cre gene is under the control of the T7 promoter for bacterial expression.

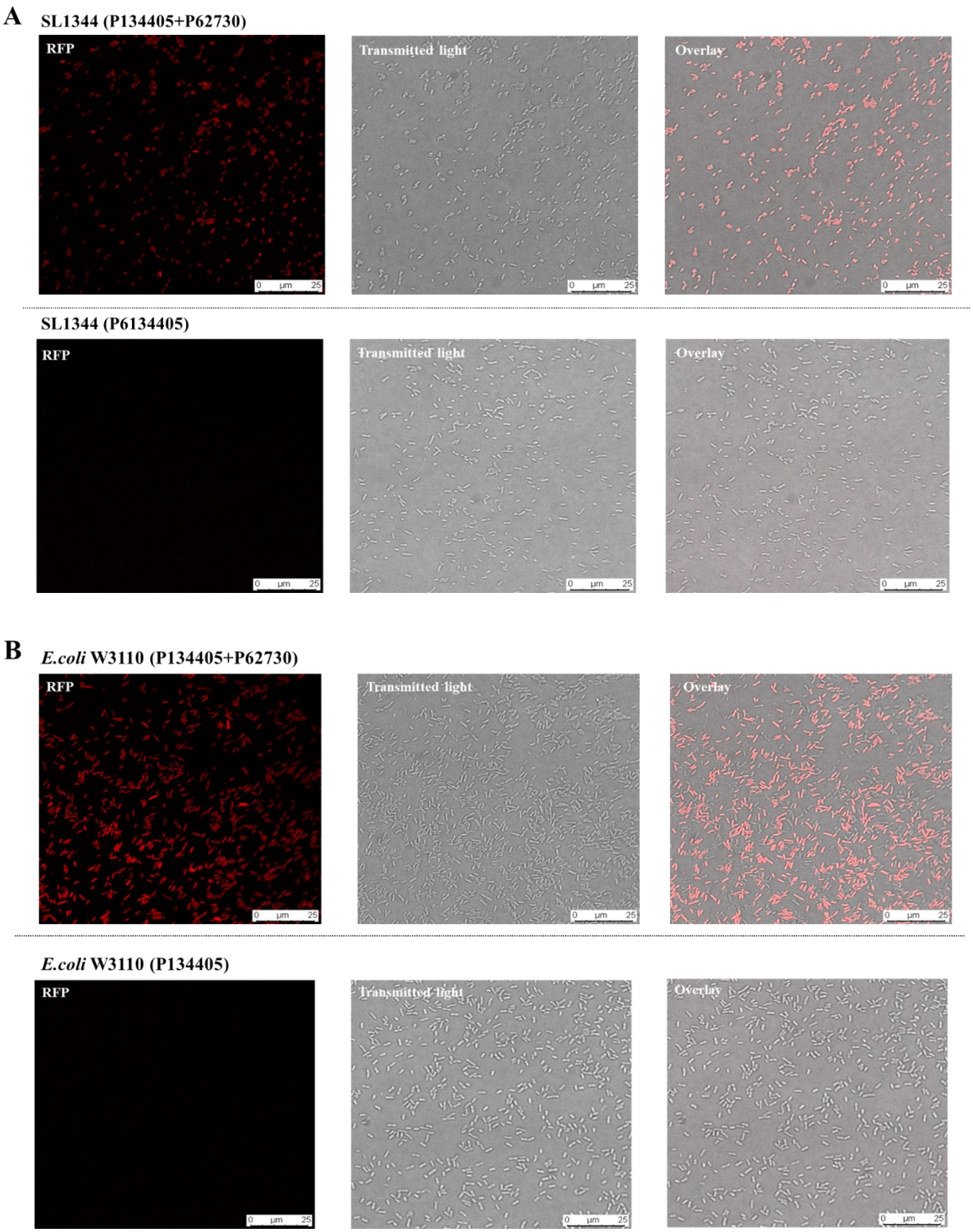


Figure 6.7 Analysis of reporter+ SL1344 and *E. coli* strains with or without introduction of Cre. A) Representative confocal images of SL1344 with reporter plasmid P134405 that has been transformed with the Cre plasmid P62730 in comparison to SL1344 with only the reporter plasmid P134405. B) Representative confocal images of *E. coli* W3110 with reporter plasmid P134405 that has been

transformed with the Cre plasmid P62730 and *E. coli* W3110 with reporter plasmid P134405. Left to right columns: overlay, RFP (red), transmitted light (grey). The scale bar is 25 μm.

6.2.3 Establishing Cre+ Mode-K cells to produce Cre+ EVs

The above work has established the reporter system in the (recipient) bacteria. I then focused on generating the appropriate donor cells that would release Cre in EVs. To do this, I transfected the plasmid pcDNA3.1-CMV-CFP;UBC-Cre25nt (P65727) into Mode-K cells. The Plasmid map of P65727 (Figure 6.8) shows that Cre gene is under the control of ubiquitin C (UBC) promoter which can drive high level constitutive expression in a broad range of tissues and cells types (Makrides, 2003). There is also a 25-nt zip code sequence in the 3' untranslated region of Cre gene that was proposed to increase the loading of Cre mRNA into EVs (Bolukbasi et al., 2012). This plasmid has a cyan fluorescent protein (CFP) marker under the cytomegalovirus (CMV) promoter and a Zeocin selection marker (BleoR) to select and enrich positive clones.

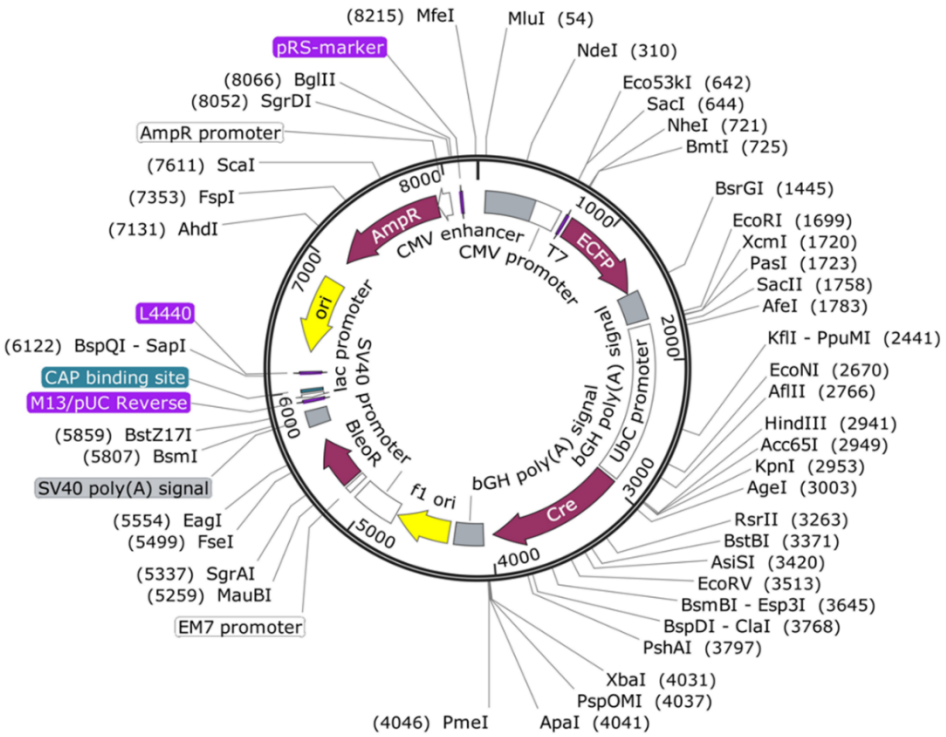


Figure 6.8 Plasmid map and annotation of plasmid pcDNA3.1-CMV-CFP;UBC-Cre25nt (P65727). Cre gene is under the control of the ubiquitin C (UBC) promoter for mammalian expression.

The workflow of establishing Cre+ Mode-K cells to produce Cre+ EVs is described in Figure 6.9. Plasmid transfection was performed using Lipofectamine 2000 (Dalby et al., 2004) and plasmid P65727. Upon inspection the following day, the transfected cells were reseeded in a T75 flask with 150 µg/ml Zeocin (Larmonier et al., 2008) for selection to enrich transfected clones. Cells that had not integrated the pcDNA3.1-CFP;Cre25nt;Zeo construct into their genome will perish during Zeocin selection, leaving only the surviving cells with the integrated construct. CFP+ Mode-K cells were then further sorted using fluorescence-activated cell sorting (FACS) to enrich for cells containing the pcDNA3.1-CFP;Cre25nt;Zeo construct. Because the Cre and CFP genes are under different promoters, polyclonal Cre+ cell lines are often heterogeneous in the expression of Cre and CFP. It is therefore necessary to grow the CFP+ Mode-K cells as monoclonal cultures and select the clones with high Cre and CFP expression by immunofluorescence staining for Cre+ EVs isolation.

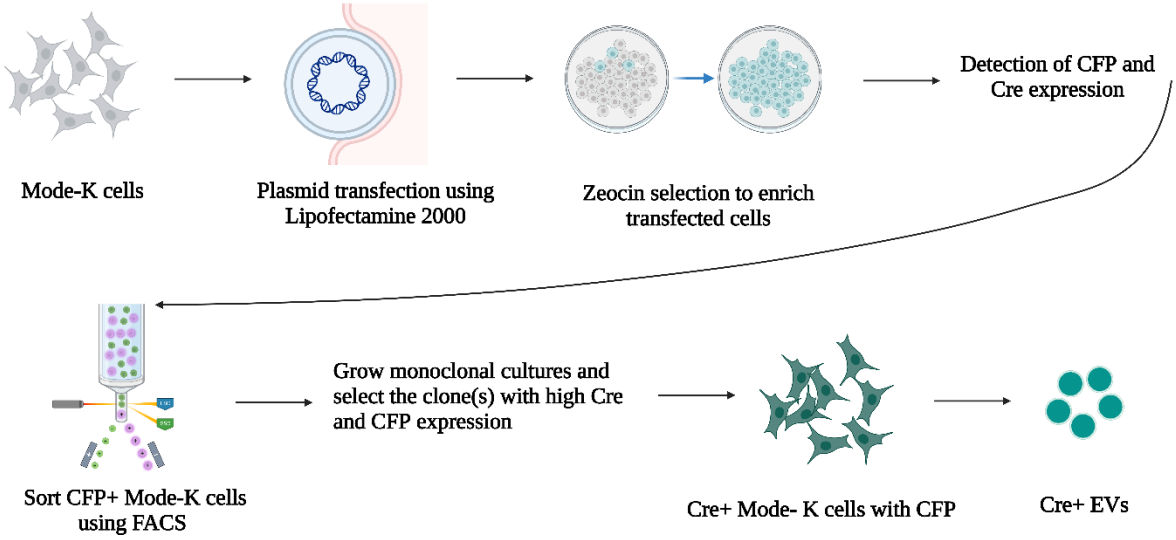


Figure 6.9 Workflow of establishing Cre+ Mode-K cells to produce Cre+ EVs. Plasmid transfection was carried out using Lipofectamine 2000 and the transfected cells were enriched by Zeocin selection. High Cre and CFP expression clones were further selected by FACS and immunofluorescence staining.

Prior to FACS sorting of CFP+ Mode-K cells, the transfected cells are tested for CFP and Cre expression after three weeks' Zeocin selection. I imaged the transfected Mode-K cells using the EVOS imaging system, however a CFP signal was not detected (data not shown). I then designed primers for Cre gene and did PCR using DNA from the transfected cells to confirm the presence of Cre gene in the genome. The PCR results (Figure 6.10A) show that Cre gene was detected in the transfected Mode-K cells (TMK1, TMK2, TMK3) while not in the negative control (untransfected Mode-K cells). I further tested the presence of Cre mRNA by RT-PCR and the presence of Cre protein by Western blot. The RT-PCR data (Figure 6.10B) showed that a weak signal of Cre mRNA was detected in transfected Mode-K cells (TMK1, TMK2, TMK3) and Cre protein was below the detection level of western blot (Figure 6.10C). These data show that the transcription level of Cre mRNA was very low in transfected Mode-K cells. This could suggest that the UBC promoter for the Cre gene is not suitable for protein expression in Mode-K cells. Based on these results, I could not continue to select high Cre and CFP expression clones for Cre+ EVs isolation.

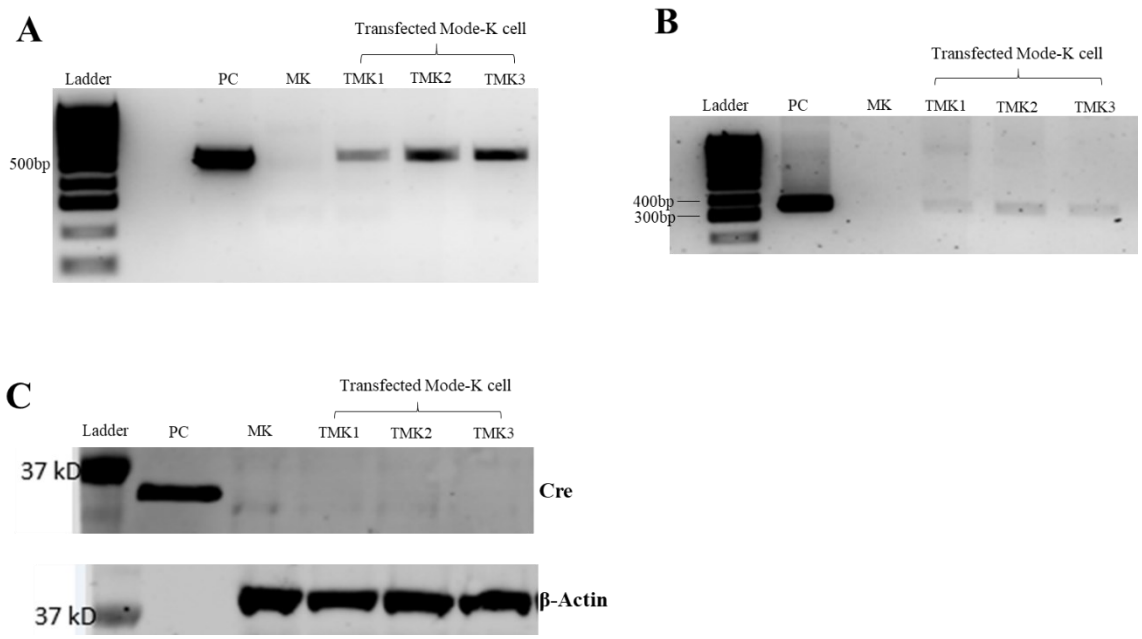


Figure 6.10 Validation of Cre expression in transfected Mode-K cells. A) PCR results for Cre gene in positive control (PC) which is Cre+ *E. coli*, untransfected Mode-K cells (MK), and transfected Mode-K cells (TMK1, TMK2, TMK3). The amplicon length of Cre gene is 525bp. B) RT-PCR results for Cre mRNA. The amplicon length of Cre gene is 353bp. C) Western blot analysis with anti- Cre protein and anti- β-actin antibodies. Equal amounts (3 µg) of protein were loaded onto 4-12% NuPage Bis-Tris polyacrylamide gel.

6.2.4 Uptake of Cre+ EVs from mice serum by reporter+ SL1344

The failure to establish Cre+ Mode-K cells led me to consider alternative approaches to obtain Cre+ EVs. In the mammalian circulatory system, a diverse range of EVs has been identified, originating from blood cells as well as various tissues such as brain, liver, and small intestine (Alberro et al., 2021). Recently, an analysis of 101 human plasma samples revealed that 99.8% of circulating EVs were attributed to hematopoietic cell origins, and 0.2% of the EV population was found to be derived from other tissues (Li et al., 2020). Cre mRNA was found to be present in EVs isolated from the blood plasma of Vav-iCre mice expressing Cre recombinase in the hematopoietic cells (Ridder et al., 2014). In our lab, a postdoctoral researcher is employing Villin-cre transgenic mice for crossing with a strain harbouring a sequence of interest flanked by loxP sites to achieve a tissue-specific deletion of the target sequence. Villin-cre transgenic mice express Cre recombinase under the control of

the mouse villin 1 promoter in villus and crypt epithelial cells of both the small and large intestines (Madison et al., 2002). Given that EVs originating from the small intestine can be detected in the blood, I've considered the possibility of Cre mRNA being present in the blood or within blood-derived EVs from Villin-cre transgenic mice. If Cre mRNA is detected in these samples, it could offer the potential to utilize these Cre+ EVs for co-culture experiments with reporter+ SL1344 cells.

Therefore, I collected blood serum from Villin-cre mice and wild type (WT) C57BL/6 mice. The WT mice served as a negative control for the detection of Cre mRNA. Prior to EVs collection, I did RT-PCR in serum to test the presence of Cre mRNA. I isolated total RNA from Villin-cre mice and wild type mice using miRNeasy Serum/Plasma Kit and the RNA was reverse transcribed using QuantiTect Reverse Transcription Kit. To ensure RNA is free of potential DNA contaminants, I performed PCR using cDNA and RNA from the serum Villin-cre mice and WT mice. The PCR results (Figure 6.11A) showed that RNA is free of Cre DNA contaminants and Cre cDNA was only detected in Villin-cre mice serum suggesting the presence of Cre mRNA in the serum of Villin-cre mice. Next, I isolated EVs from the serum of Villin-cre mice and WT mice using the ultracentrifugation method and the EVs and EVs-depleted supernatant were used for RNA extraction, reverse transcription, and subsequent detection of cDNA using PCR. Cre recombinase cDNA was detected in Villin-cre mice serum EVs but not in wild type mice serum EVs (Figure 6.11B). A weak band was detected in the EVs-depleted supernatant from Villin-cre mice serum which suggests there is still some Cre mRNA in the EVs-depleted supernatant and there also may still be EVs in this sample.

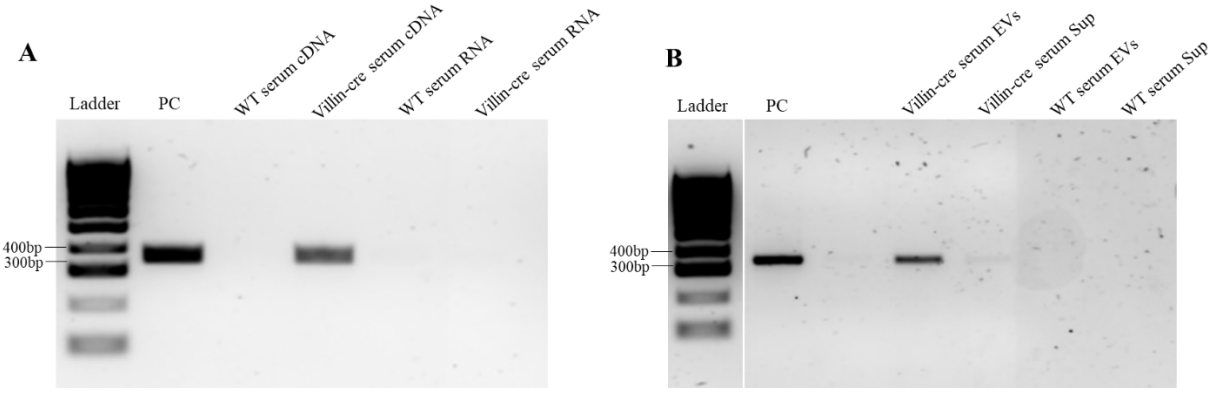


Figure 6.11 Detection of Cre mRNA in serum and serum EVs. A) RT-PCR results for Cre gene in positive control (PC), Villin-cre mice serum and wild type mice. The amplicon length of Cre gene is 353bp. B) RT-PCR results for Cre gene in positive

control (PC), EVs and EVs-depleted supernatant from Villin-cre mice serum and wild type mice serum.

Next, I utilized these serum EVs for co-culture experiments with reporter+ SL1344 cells. When I did the dose experiments in Chapter 5, the lowest dose of EVs I used was were 0.3 $x10^{10}$ particles/ml. However, EVs isolated from blood samples are a complex pool of particles with various origins, the percentage of EVs derived from intestine is not clear but may be less than 0.2% (Li et al., 2020). Therefore, I considered a higher dose of serum EVs for the co-culture experiments. I co-cultured the EVs from Villin-cre mice serum and WT mice serum with reporter+ SL1344 and *E. coli* in M9 medium. EVs were added to the culture at a concentration of 3.0 $x10^{11}$ particles/ml which is 100 time higher than the lowest dose used in the dose experiment. I imaged the reporter+ SL1344 and *E. coli* after overnight incubation. The confocal images (Figure 6.12A) show that RFP was detected in reporter+ SL1344 when co-cultured with EVs from Villin-cre mice serum but not with EVs from WT mice serum. RFP was not detected in reporter+ *E. coli* when co-cultured with EVs from Villin-cre mice serum or WT mice serum (Figure 6.12B).

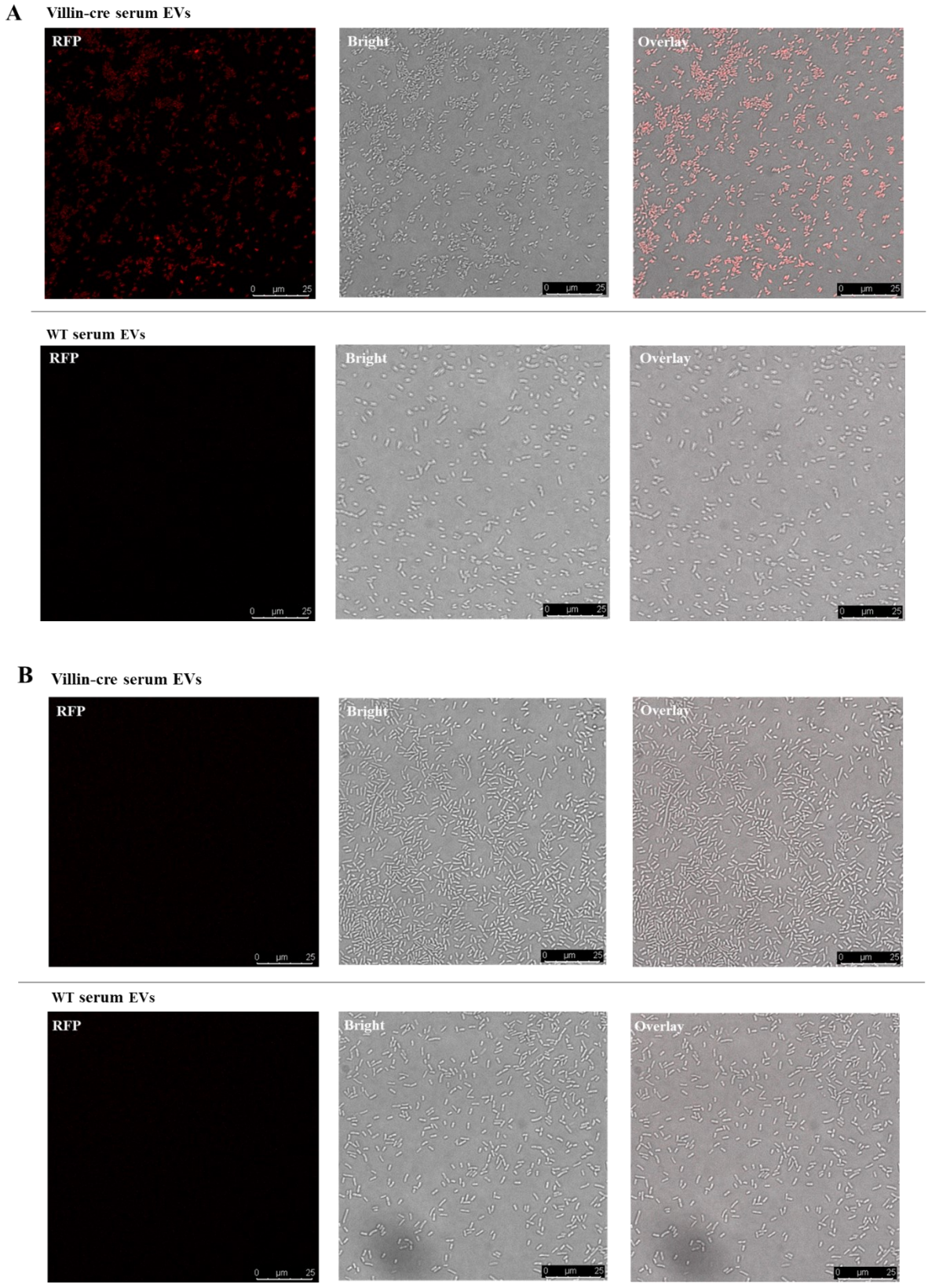


Figure 6.12 Co-culturing mice serum EVs with reporter+ SL1344 and *E. coli*. A) Representative confocal images of reporter+ SL1344 co-cultured with EVs from Villin-cre mice serum and wild type (WT) mice serum. B) Representative confocal images

of reporter+ *E. coli* co-cultured with EVs from Villin-cre mice serum and wild type mice serum. Left to right columns: RFP (red), transmitted light (grey), overlay. The scale bar is 25 μm.

6.3 Discussion

6.3.1 Summary

In this chapter, I investigated whether Mode-K EVs have the capability to transport functional cargo into SL1344 using the Cre-loxP system. I first established a reporter+ SL1344 strain by introducing plasmid pBbAW4k-loxP-TT-loxP-mRFP1 and showed its ability to report Cre activity by introducing another plasmid P62730 that can express Cre recombinase protein in bacterial cells. The attempt to establish Cre+ Mode-K cells to produce Cre+ EVs failed due to the low level of Cre mRNA and the absence of Cre protein in the transfected Mode-K cells. I then collected EVs from Villin-cre mice serum as an alternative approach to obtain Cre+ EVs. Cre mRNA was detected in EVs isolated from Villin-cre mice serum. The co-culture experiments of reporter+ SL1344 and serum EVs demonstrated the functional transmission of cargo to SL1344 by serum EVs. This is a significant finding, for the first time showing functional transmission of molecules from host to bacteria.

6.3.2 Methods to increase gene expression in Mode-K cells

I attempted to establish the Cre+ Mode-K cells by transfection of plasmid pcDNA3.1-CMV-CFP;UBC-Cre25nt. This plasmid was used to establish Cre+ MDA-MB-231 cells successfully by transfection using Lipofectamine 2000 (Zomer et al., 2015). The MDA-MB-231 cell is an epithelial, human breast cancer cell line (Sun et al., 2019). However, this plasmid is not suitable for Mode-K cells. After three weeks' Zeocin selection, I didn't detect CFP signal in transfected Mode-K cells. Figure 6.10 showed that the transcription level of Cre mRNA was very low and Cre protein was not detected in transfected Mode-K cells. The plasmid map (Figure 6.8) shows that CFP is under the cytomegalovirus (CMV) and Cre gene is under the ubiquitin C (UBC) promoter. My data indicate that the CMV and UBC promoter are not suitable for protein expression in Mode-K cells. To solve this problem, I need to construct a vector with Cre under a different promoter that can support gene expression in Mode-K cells. Human β-actin promoter (ACTB), cytomegalovirus (CMV), elongation factor-1α (EF1α), phosphoglycerate kinase (PGK) and ubiquitinC (UBC) are five commonly used constitutive promoters for mammalian studies (Norrman et al., 2010). However, there's a lack of evidence in the literature regarding the use of these promoters to drive gene expression in Mode-K cells. Instead of considering a promoter change, an alternative

strategy to enhance the transcription level of Cre mRNA within Mode-K cells is by exploring the UBC promoter and identifying a suitable transcription factor that could augment its activity. This approach draws inspiration from a study indicating that Mode-K cells exhibit notably lower levels of CCL25 mRNA in comparison to freshly isolated small intestinal epithelial cells and Cdx transcription factors could significantly enhance CCL25 promoter activity, leading to increased endogenous CCL25 mRNA levels in Mode-K cells (Ericsson et al., 2006).

6.3.3 Limitations of using EVs from Villin-cre mice serum

EVs collection from Villin-cre mice serum was an alternative approach to obtain intestinal epithelial Cre+ EVs. I was able to detect Cre mRNA in EVs isolated from Villin-cre mice serum (Figure 6.11B) and the confocal images (Figure 6.12) show that rfp was only detected in reporter+ SL1344 when co-cultured with these EVs. The co-culture experiments of reporter+ SL1344 and serum EVs demonstrated the functional transmission of RNA cargo to SL1344 by EVs. However, serum EVs are a complex pool of EVs with various origins and most of serum EVs are from the hematopoietic cells. Villin-cre transgenic mice express Cre recombinase in villus and crypt epithelial cells of the small and large intestines (Madison et al., 2002). The percentages of EVs derived from intestine and Cre+ EVs are unknown. Therefore, I cannot quantify the amount of Cre+ EVs used in the co-culture experiments. The co-culture experiments of serum EVs and reporter+ bacteria could not provide direct evidence for the functional transmission of cargo to SL1344 by Mode-K EVs because the types of serum Cre+ EVs may be different to those isolated from Mode-K cells.

Chapter 7 Discussion

The goal of this thesis is to directly identify the mouse miRNAs that are naturally transferred to the gut microbiota and explore the specificity and mechanisms involved in their trafficking. In Chapter 3, I developed a purification method to isolate pure gut microbiota from mouse gut contents and profiled the transferred host miRNAs in the purified gut microbiota (PGM) through sRNA sequencing. I then employed miRNA-FISH to visualize the presence of host miRNAs within gut bacterial cells. Chapters 4 and 5 focused on investigating two potential transport mechanisms for host miRNAs into bacteria (which are not mutually exclusive): the Ago2 protein and extracellular vesicles (EVs). Finally, in Chapter 6, I employed a genetic reporter system (Cre-loxP system) to explore whether mouse EVs can deliver functional cargo into bacteria. In this Discussion Chapter, I analyse my findings within the broader context of current research and assess their implications for the field.

7.1 Strategies to identify miRNA-bacteria interactions in the gut

Currently, only a few miRNA-bacteria interactions have been reported and there is an ongoing need to understand exactly how miRNAs traffic to bacteria naturally, which bacteria they traffic to and how they regulate bacterial genes once inside. One strategy to identify specific miRNA-bacteria interactions is based on the analysis of altered miRNAs in faecal samples in certain disease contexts, in an attempt to link miRNA changes to alterations in the microbiome. For example, *A. muciniphila* was found to be increased in the faeces from EAE mice compared with controls, and miR-30d was also enriched in faeces from EAE mice (Liu et al., 2019). The authors showed that miR-30d could directly impact the growth of *A. muciniphila* by culturing the strain in the presence of synthetic miR-30d *in vitro* and by administering synthetic miR-30d orally to mice. miRNA-bacteria interactions can also be inferred based on studies in miRNA knockout mice. In the miR-21 knockout mouse, *Lactobacillus* was found to be significantly increased in the small intestinal lumen when compared with wild-type animals (Santos et al., 2020). The authors showed that synthetic miR-21a-5p directly suppressed the growth of *Lactobacillus in vitro*. However, there is still a lack of strategies to systematically determine or predict which miRNA-bacteria interactions occur under physiological conditions.

In Chapter 3, I developed a purification method to isolate pure gut microbiota from mouse gut contents and profiled transferred host miRNAs in PGM by sRNA sequencing. I aimed to overcome the limitations of previous fecal studies that do not establish a clear association between miRNAs and bacteria in the gut. My data suggest that some host miRNAs are naturally present inside gut microbiota and the miRNAs within bacteria are a distinct subset compared to the miRNAs found in total gut contents or faeces. This miRNA composition difference may arise from the elimination of mouse cells and the alteration of the microbiome composition in the PGM. I then applied miRNA-FISH to visualize the presence of host miRNAs within gut bacterial cells. My data show that miR-21a-5p, identified as the most abundant miRNA in PGM through sequencing, was detected in PGM. My work identifies highly confident candidates for transferred miRNAs, paving the way for further exploration into their transfer mechanisms and potential functions within bacterial species.

Though sRNA sequencing coupled with miRNA-FISH can determine the miRNAs transferred to gut microbiota, it does not specifically identify the bacterial species that internalize these miRNAs. To address this limitation, I proposed a solution: conducting co-staining experiments utilizing probes targeting miR-21a-5p and *Lactobacillus*. My data revealed co-localization between these probes, suggesting the uptake of miR-21a-5p by *Lactobacillus*. This is significant because it suggests enough miR-21a-5p is naturally taken up by *Lactobacillus* to detect. Most other studies in the literature examine uptake with synthetic RNA that is over-expressed *in vitro*. It would be of interest to study biochemically the targets and uptake mechanism of miR-21a-5p in *Lactobacillus,* however this requires further collaboration. However, the co-staining method is constrained by its low throughput and inability to detect unknown bacterial species. A fluorescence in situ hybridization-flow cytometry (FISH/FC)-based method has been used to separate and enrich specific bacterial species in a high-throughput manner (Kalyuzhnaya et al., 2006; Bruder et al., 2016). In these studies, FISH was performed using taxon-specific probes targeting bacterial ribosomal RNA and bacterial cells were separated by fluorescence-activated FC analysis and cell sorting. To identify the bacterial species that internalize specific miRNA, I would adapt a similar strategy to sort bacterial cells following miRNA-FISH and the sorted bacteria can be identified through 16S rDNA sequence analysis. This approach not only boosts

throughput but also facilitates the identification of both known and unknown bacterial species capable of internalizing miRNAs.

In summary, sRNA sequencing can provide candidates of transferred miRNAs which are subsequently validated through miRNA-FISH. Co-staining experiments using probes for miRNA and bacteria can identify the bacterial species that internalize these miRNAs. This methodology provides a promising avenue for further exploration into additional interactions between miRNA and bacteria in the gut. It can also be applied to study simple models like mono-associated mice, which only have a single bacterium in the gut, to explore the transferred host miRNAs in bacteria. The next step involves confirming the role and function of transferred miRNA in bacteria.

7.2 Mechanism of RNA trafficking from mouse to bacteria

The discovery of host miRNAs within gut bacteria lays the foundation for exploring the mechanisms responsible for the transportation of these miRNAs into bacterial cells. In this thesis, I investigated two potential transport mechanisms: the Ago2 protein and EVs. It should be noted that these two mechanisms are not mutually exclusive and may coexist in the gut, and their prevalence may vary depending on the specific miRNAs and gut bacteria involved.

7.2.1 Mouse Ago2 protein

Firstly, I investigated the hypothesis that mouse miRNAs are transported in complex with the Ago2 protein, and this protein then coordinates gene regulation inside the bacteria. To test my hypothesis, I investigated whether the mouse Ago2 protein is detectable within gut microbiota and adapted a method to identify bacterial genes directly targeted by miRNA-Ago2 complexes. Ago2 protein can be secreted to the extracellular space in both the vesicular and non-vesicular forms (Van et al., 2014; Temoche-Diaz et al., 2019) but extracellular miRNAs are primarily associated with Ago2 complexes that are in non-vesicular form (Arroyo et al., 2011; Geekiyanage et al., 2020). In Chapter 4, the methods I used to detect Ago2 protein could not distinguish vesicular and non-vesicular Ago2. The detection of Ago2 was performed in PGM using western blot analysis and immunofluorescence imaging and our results suggest that there is no consistent Ago2 signal in PGM. These data suggest that either the Ago2 protein is not transported into gut bacteria or, if it is, the quantity present falls below the detection limit of the methods used. To enhance the detection capability of

Ago2 protein within PGM samples, it is imperative to explore alternative methodologies beyond western blot analysis and immunofluorescence imaging which are antibody-based methods. Mass spectrometry is a powerful analytical technique capable of identifying and quantifying sample analytes based upon their mass to charge ratio in a vacuum environment and it can provide advantages over antibody-based methods, including greater specificity and higher sensitivity (Jayasena et al., 2016; Noor et al., 2021). Therefore, I can analyse the protein content of PGM using mass spectrometry to determine the presence of Ago2. In addition to detect the uptake of Ago2 protein by gut bacteria in *vivo*, labelled recombinant Ago2 can be used to co-culture with bacteria to explore the uptake *in vitro*. Labelled proteins have proven effective in studying cellular uptake activities. For example, fluorescent-labelled recombinant human alpha-L-iduronidase was used to track the uptake and distribution of this enzyme in human fibroblasts (Tippin et al., 2011). Since it is not clear which bacteria can internalize Ago2, I can start with some common bacterial species such as *E. coli*.

To account for the possibility that Ago2 signal could be lost in the purification process, and to test a new method for directly capturing miRNA-bacteria interactions, I carried out Ago2 CLEAR-CLIP in gut tissues. The goal of this approach is to determine whether I can identify authentic miRNA-bacteria interactions based on chimeric sequencing reads formed by miRNA-bacterial RNAs. This would indicate transfer of host miRNAs and binding with the Ago2 protein which could then in theory regulate gene expression in bacteria. Bioinformatic analysis of the CLEARP-CLIP data indicates the presence of high-confidence bacterial targets in each gut segment and the target reads predominantly map to mRNA and rRNA regions. However, confirming these as genuine targets will require further bioinformatic analysis and experimental validation.

If the miRNA-Ago2 complexes regulate gene expression in bacteria, it is essential to evaluate whether miRNA-bacterial target interactions follow the same rules as miRNA-host target interactions. Firstly, we need to explore the principles of miRNA-target recognition in bacteria. In mammalian cells, the canonical miRNA targeting relies on the base pairing between the seed region, specifically nucleotides 2-7 within the miRNA, and complementary sites in the 3’ UTRs of mRNA molecules (Lewis et al., 2005). To explore whether there exists a similar miRNA-target recognition mechanism, we can filter the reproducible bacterial targets (especially for the reads mapping to

mRNA) based on the presence of a strong miRNA seed site. Secondly, we need to investigate how miRNA-Ago2 complex mediates gene regulation inside of bacteria and whether Ago2 interacts with other bacterial proteins. In mammalian cells, miRNAs function in association with Ago2 and GW182 protein (a component of the processing body), which are the main components of the miRNA-induced silencing complex (miRISC) to repress the translation of the target mRNAs (Krol et al., 2010). However, in the mitochondria miRNA-Ago2 complex stimulates, rather than represses, the translation of specific mRNA independently of its functional partner GW182 (Zhang et al., 2014). Considering that bacteria and mitochondria are similar in the light of genetic elements (Boguszewska et al., 2020), I assume that miRNA-Ago2 complex in bacteria may function more similarly to that in mitochondria.

In summary, our findings indicate the presence of high-confidence bacterial RNAs targeted by miRNA-Ago2 complexes, suggesting the potential transfer of host miRNAs with Ago2 protein and their potential functional role in gut bacteria. However, to substantiate these observations, we need to confirm the existence of Ago2 protein within specific gut bacteria and elucidate the mechanism of how miRNA-Ago2 complex mediates gene regulation inside of bacteria. Moreover, previous work has demonstrated that culturing bacteria with synthetic miRNA mimics *in vitro* directly affects bacterial gene expression (Liu et al., 2019; Santos et al., 2020), it is also crucial to investigate whether and how miRNAs integrate into bacterial sRNA machinery and interact with bacterial protein partners to modulate gene expression. In Gram-negative bacteria, sRNAs usually depend on RNA-binding proteins like Hfq or ProQ to facilitate sRNA-mRNA base pairing. To identify bacterial proteins that might bind to miRNAs in gut bacteria, a biotinylated miRNA pull-down assay can be applied in PGM and pulled-down proteins are then subjected to high-throughput identification by mass spectrometry. Two biotinylated highly exosome-enriched miRNAs, miR-3470a and miR-194-2-3p, have been applied to pull down proteins in hepatocyte extracts and among the RNA-binding proteins precipitated with miRNAs, SYNCRIP is identified to have a role in exosomal sorting of miRNAs (Santangelo et al., 2016). I can use biotinylated miR-21a-5p, which is the most abundant miRNA in PGM, to pull down proteins from PGM lysates and identify the binding proteins by mass spectrometry.

7.2.2 Extracellular vesicles

Another potential mechanism for facilitating miRNA transport could involve EVs. In Chapter 5, I examined the capability of EVs derived from intestinal epithelial cells to transport mammalian miRNAs into specific bacteria. To explore the specificity in the uptake of EVs by bacteria, I conducted a comparison using fluorescently labelled EVs obtained from Mode-K cells and *H. bakeri*. My data show that SL1344 uptake of Mode-K EVs but not *H. bakeri* EVs when cultured in M9 medium. The uptake of Mode-K EVs was not observed in *E. coli* W3110. These data demonstrate a degree of specificity for the interaction between mammalian EVs and SL1344. However, the detailed mechanism behind this interaction remains largely unknown. Firstly, it is unclear how Mode-K EVs are recognized by SL1344. For the intercellular communication via EVs, the specificity of cell targeting relies on unique interactions between ligands enriched on the surface of EVs and receptors present on the plasma membrane of recipient cells (Liu and Cheng, 2023). Since the uptake of Mode-K EVs by SL1344 was observed when bacteria were cultured in M9 medium but not in LB medium, analysing the surface proteins on the plasma membrane of SL1344 when cultured in M9 and LB medium may help identify the receptors associated with EV recognition and uptake. Moreover, over 200 distinct *S.* Typhimurium strains have been characterized, each primarily tailored to specific niches within the environment and the intestinal tracts of various animal species (Luo et al., 2011). It is intriguing to explore whether Mode-K EVs can enter *S.* Typhimurium strains other than SL1344 and exert an influence on their growth. Finally, it is also unclear whether there is cell type specificity concerning the uptake of EVs by SL1344, given that the intestinal epithelium of the small intestine also comprises other cell types, such as Goblet cells, Paneth cells, Microfold cells, and so on (Kong et al., 2018). To address this, employing EVs from other mouse cell types for co-culturing with SL1344 could provide insights into their uptake behaviour.

My data also revealed that Mode-K EVs exerted a stimulatory effect on the growth of SL1344 when bacteria were cultured in M9 medium. This growth-promoting impact seemed to be dependent on the dosage of Mode-K EVs, as higher concentrations correlated with increased growth. Based on previous studies, there are two hypotheses for the mechanism that underlies this growth promotion effect. One hypothesis is that Mode-K EVs provide nutrients to SL1344, thereby affecting its growth. There are some precedents for this in the literature that EVs can provide nutrition for bacteria. For example, EVs released from respiratory epithelium provide

the nutrient iron and promoting bacterial biofilm growth (Hendricks et al., 2021). Given that the promotion effect of Mode-K EVs on SL1344 was observed when cultured in M9 minimal medium, it is reasonable to explore the nutrients hypothesis. During the biogenesis process, EVs acquire crucial bioactive molecules from their donor cells, comprising a diverse array of components such as nucleic acids, lipids, glycans, and proteins (Jurj et al., 2020). Recently, there has been a surge in interest regarding the metabolomic analysis of EVs. Metabolites include various molecular species, from steroid hormones and lipids to metabolic intermediates generated during nutrient synthesis and the elemental components of major biopolymer classes (Williams et al., 2019). To validate the nutrients hypothesis, it is imperative to conduct a comprehensive analysis of the cargo carried by Mode-K EVs, with a specific focus on identifying potential nutrient compositions that could influence the growth of SL1344. However, Triton X-100-treated Mode-K EVs, which are expected to have similar nutrient contents as intact Mode-K EVs, did not stimulate the growth of SL1344. This finding suggests that Mode-K EVs may not supply nutrients to SL1344, thereby influencing its growth, or that the uptake of nutrients is dependent on intact EVs.

The other hypothesis is that some of the cargo in the EVs such as miRNAs could directly influence bacterial gene expression and thereby affect bacterial growth. In Chapter 5, I delved deeper into the investigation of the second hypothesis. I detected let-7 family miRNAs, which are abundant in Mode-K EVs, in SL1344 by RT-qPCR after co-culturing with Mode-K EVs. These data suggest that EVs have the capability to deliver host miRNA to SL1344, at a level that is detectable by qRT-PCR. In addition, the RT-qPCR data of let-7 family miRNAs also provides evidence for the uptake of Mode-K EVs by SL1344. To further confirm that *H. bakeri* EVs are not internalized by SL1344, detection of certain sRNAs from *H. bakeri* EVs, such as Y-RNAs (Buck et al., 2014; White et al., 2020), via RT-qPCR after co-culturing with *H. bakeri* EVs would be beneficial.

I then predicted the targets of the top 10 abundant miRNAs of Mode-K EVs in the genome of *S.* Typhimurium*.* One intriguing target of the let-7 family miRNAs identified is the Fis family transcriptional regulator. Fis is a global family transcriptional regulator and impacts the expression of 291 genes during the exponential growth phase (Kelly et al., 2004). Furthermore, Fis-binding sites have been identified on various genes, such as *rpoS* and *gyrB* in *S.* Typhimurium, suggesting its regulatory impact on these

genes (Hengen et al., 1997). This potential target Fis may provide valuable insights into how the let-7 family miRNAs modulate bacterial gene expression, potentially influencing the growth of SL1344. Additional experiments such as using miRNA inhibitors are required to validate the functionality of let-7 family miRNAs and their genuine targets. miRNA inhibitors are chemically modified, single-stranded oligonucleotides with secondary structure designed to specifically bind to miRNAs and block miRNA regulation of target gene expression (Robertson et al., 2010). I can transfect miRNA inhibitors for let-7 family miRNAs into SL1344 when co-culturing with Mode-K EVs and investigate the growth curve of SL1344 and expression level change of some potential targets.

In summary, my thesis shows for the first time that EVs derived from intestinal epithelial cells can transport mammalian miRNAs into specific bacteria and affect bacterial growth. However, I do not validate that the miRNAs are directly responsible for this effect. Some studies have proposed that EVs carry only a limited number of miRNA molecules, insufficient to cause a biologically significant impact on recipient cells (Chevillet et al., 2014; Albanese et al.,2021). Further work is required to understand exactly how many miRNA molecules enter bacteria and how they can modulate gene expression with or without an Ago protein.

7.3 Mammalian EV-mediated functional cargo transfer to bacteria

EVs were initially described as a mechanism for discarding unwanted cellular debris upon their first observation in the early 1980s (Johnstone et al., 1987). A decade later, EVs derived from B lymphocytes were discovered to have the capability to activate T lymphocytes (Raposo et al., 1996). Subsequently, it was demonstrated that dendritic cells also release EVs (Théry et al., 2001), implying a potential role for EVs in antigen presentation. Further research identified mRNA and miRNA as cargo within EVs, highlighting their transfer into recipient cells and suggesting biological activity (Wysoczynski and Mariusz, 2009; Raposo and Stoorvogel, 2013). Currently, RNA-based therapeutics hold great promise for treating various diseases, due to their ability to address genetic origins through various mechanisms (Kaczmarek et al., 2017). However, RNA molecules usually face a barrier in crossing cell membranes, underscoring the critical need for a safe and efficient delivery vehicle (Dowdy, 2017). Endogenously derived EVs exhibit several characteristics that make them excellent

candidates for therapeutic RNA delivery. These include their inherent ability to selectively transfer RNAs in a functional manner and their superior safety profile compared to synthetic particles (de Voogt et al., 2021). In my thesis, I demonstrated Mode-K EVs can be internalized by SL1344 with a degree of specificity and promote bacterial growth, although the mechanisms remain unclear. This discovery underscores the significant potential of developing mammalian EVs as carriers for drugs targeting bacteria.

In Chapter 6, I used an additional method to validate that Mode-K EVs can deliver functional cargo into *S.* Typhimurium: a genetic reporter system (Cre-loxP system). This technique involves the donor cells secreting EVs containing Cre mRNA, which are then internalized by the reporter cells. If the Cre mRNA is translated after internalization, Cre recombinase can remove the loxP-flanked transcription terminator, leading to the production of fluorescent molecules in the reporter cells. Unfortunately, efforts to generate Cre+ Mode-K cells for producing Cre+ EVs were unsuccessful due to promoter inactivity. As an alternative approach, I collected EVs from the serum of Villin-cre mice to obtain Cre+ EVs. RT-qPCR analysis revealed the presence of Cre mRNA in EVs isolated from Villin-cre mice serum. Co-culture experiments of serum EVs and reporter+ SL1344 demonstrated the functional transmission of mRNA cargo to SL1344 facilitated by serum EVs. However, this did not provide direct evidence for the functional transmission of cargo to SL1344 by Mode-K EVs because the types of serum Cre+ EVs are different to those isolated from Mode-K cells. Therefore, I should continue to generate Cre+ Mode-K cells for producing Cre+ EVs. To address the issue of promoter inactivity, I aim to engineer a vector containing Cre gene under an alternative promoter that is capable of supporting gene expression in Mode-K cells. Except for the Cre-loxP system, another method to detect functional mRNA delivery involves the use of SunTag mRNAs (Yan et al., 2016). It requires genetically engineering 24 SunTag peptide sequences into the transcript of interest in the EV donor cells. If the EVs are taken up by the recipient cells and the SunTag mRNAs are translated, the SunTag sequences can be specifically recognized by fluorescence protein (FP)-coupled single-chain antibodies that are expressed in the EV recipient cells. This method can study dynamics of translation of single mRNAs in the EV recipient cell *in vivo*.

The reporter systems using Cre-loxP system or SunTag mRNAs are limited to the detection of large mRNA, and they could not be used to detect functional transfer of sRNAs. Mammalian EVs contain mainly sRNAs or fragmented mRNAs (Wei et al., 2017) and my PhD project primarily focuses on miRNAs. Therefore, there is a need for a system that can study the EV-mediated transfer of small noncoding RNAs. The CRISPR Operated Stoplight System for Functional Intercellular RNA Exchange (CROSS-FIRE) method offers a promising solution to address this need, which enables the specific detection of functional sRNA transfer *in vitro* (de Jong et al., 2020). CROSS-FIRE operates through a reporter cell line that constitutively expresses mCherry, alongside a linker sequence featuring a target site specific to the single-guide RNA (sgRNA) for CRISPR/Cas9, and two eGFP sequences. These eGFP sequences are purposely designed to be out-of-frame by 1 or 2 nucleotides, respectively. Upon the functional delivery of sgRNA into the recipient cell, CRISPR/Cas9 induces a double-strand break, triggering a frameshift mutation. This mutation leads to the permanent expression of eGFP, alongside mCherry, enabling the detection of functional small RNA transfer. CROSS-FIRE can detect functional transfer of sgRNA at a single-cell resolution using fluorescence microscopy and flow cytometry.

7.4 Concluding remarks

The work presented in this thesis advances our understanding of interactions between host and gut bacteria mediated by host miRNAs and EVs. I demonstrate that host miRNAs are present within gut microbiota under physiological condition and provide an approach to identify miRNA-bacteria interactions in the gut. Furthermore, I show that EVs derived from intestinal epithelial cells can act as a transport mechanism for host miRNAs into specific bacteria and impact bacteria growth. Using the genetic reporter system (Cre-loxP system), I show that mammalian EVs can effectively transfer functional cargo to bacteria. These findings deepen our understanding of the specificity and mechanism of RNA trafficking from mouse to gut bacteria which could hold promising implications for modulating the gut microbiota to treat associated diseases in the future.

References

Adiliaghdam, Fatemeh, et al. "A requirement for Argonaute 4 in mammalian antiviral defense." Cell reports 30.6 (2020): 1690-1701.

Ahmed, Farid E., et al. "Diagnostic microRNA markers for screening sporadic human colon cancer and active ulcerative colitis in stool and tissue." Cancer genomics & proteomics 6.5 (2009): 281-295.

Akhtar, Irfan, et al. "Visualization of endogenous gut bacteria in Drosophila melanogaster using fluorescence in situ hybridization." Plos one 16.2 (2021): e0247376.

Albanese, Manuel, et al. "MicroRNAs are minor constituents of extracellular vesicles that are rarely delivered to target cells." PLoS Genetics 17.12 (2021): e1009951.

Alberro, Ainhoa, et al. "Extracellular vesicles in blood: sources, effects, and applications." International journal of molecular sciences 22.15 (2021): 8163.

Andino, A., and I. Hanning. "Salmonella enterica: survival, colonization, and virulence differences among serovars." The Scientific World Journal 2015 (2015).

Andrews, Simon. "FastQC: a quality control tool for high throughput sequence data." (2010).

Arroyo, Jason D., et al. "Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma." Proceedings of the National Academy of Sciences 108.12 (2011): 5003-5008.

Ayichew, T., et al. "Bacterial probiotics their importances and limitations: a review." J Nutr Health Sci 4.2 (2017): 202.

Bachurski, Daniel, et al. "Extracellular vesicle measurements with nanoparticle tracking analysis–An accuracy and repeatability comparison between NanoSight NS300 and ZetaView." Journal of extracellular vesicles 8.1 (2019): 1596016.

Bartel, David P. "Metazoan micrornas." Cell 173.1 (2018): 20-51.

Bian, Baishijiao, et al. "Exosomes derived from neural progenitor cells preserve photoreceptors during retinal degeneration by inactivating microglia." Journal of Extracellular Vesicles 9.1 (2020): 1748931.

Bittel, Miriam, et al. "Visualizing transfer of microbial biomolecules by outer membrane vesicles in microbe-host-communication in vivo." Journal of Extracellular Vesicles 10.12 (2021): e12159.

Blair, Jessica MA, et al. "Choice of bacterial growth medium alters the transcriptome and phenotype of Salmonella enterica Serovar Typhimurium." PLoS One 8.5 (2013): e63912.

Bo Xiaochen, and Wang Shengqi. "TargetFinder: a software for antisense oligonucleotide target site selection based on MAST and secondary structures of target mRNA." Bioinformatics 21.8 (2005): 1401-1402.

Boguszewska, Karolina, et al. "The similarities between human mitochondria and bacteria in the context of structure, genome, and base excision repair system." Molecules 25.12 (2020): 2857.

Bollenbach, Tobias, et al. "Nonoptimal microbial response to antibiotics underlies suppressive drug interactions." Cell 139.4 (2009): 707-718.

Bolyen, Evan, et al. "Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2." Nature biotechnology 37.8 (2019): 852-857.

Bolukbasi, Mehmet Fatih, et al. "miR-1289 and “Zipcode”-like sequence enrich mRNAs in microvesicles." Molecular Therapy-Nucleic Acids 1 (2012).

Brennecke, Johannes, et al. "High-yield extraction of Escherichia coli RNA from human whole blood." Journal of Medical Microbiology 66.3 (2017): 301-311.

Bruder, Lena M., et al. "Flow cytometric sorting of fecal bacteria after in situ hybridization with polynucleotide probes." Systematic and applied microbiology 39.7 (2016): 464-475.

Buck, Amy H., et al. "Exosomes secreted by nematode parasites transfer small RNAs to mammalian cells and modulate innate immunity." Nature communications 5.1 (2014): 5488.

Buffie, Charlie G., et al. "Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile." Nature 517.7533 (2015): 205-208.

Callahan, Benjamin J., et al. "DADA2: High-resolution sample inference from Illumina amplicon data." Nature methods 13.7 (2016): 581-583.

Cai, Qiang, et al. "Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes." Science 360.6393 (2018): 1126-1129.

Cai, Qiang, et al. "Extracellular vesicles: cross-organismal RNA trafficking in plants, microbes, and mammalian cells." (2023).

Campeau, J. L., et al. "Intestinal epithelial cells modulate antigen-presenting cell responses to bacterial DNA." Infection and immunity 80.8 (2012): 2632-2644.

Cani, Patrice D., et al. "Microbial regulation of organismal energy homeostasis." Nature metabolism 1.1 (2019): 34-46.

Carvalho, Frederic A., et al. "Toll-like receptor–gut microbiota interactions: perturb at your own risk!" Annual review of physiology 74 (2012): 177-198.

Chevillet, John R., et al. "Quantitative and stoichiometric analysis of the microRNA content of exosomes." Proceedings of the National Academy of Sciences 111.41 (2014): 14888-14893.

Chi, Sung Wook, et al. "Argonaute HITS-CLIP decodes microRNA–mRNA interaction maps." Nature 460.7254 (2009): 479-486.

Claycomb, Julie, Cei Abreu-Goodger, and Amy H. Buck. "RNA-mediated communication between helminths and their hosts: the missing links." RNA biology 14.4 (2017): 436-441.

Cleusix, Valentine, et al. "Inhibitory activity spectrum of reuterin produced by Lactobacillus reuteri against intestinal bacteria." BMC microbiology 7.1 (2007): 1-9.

Colombo Marina, Graça Raposo, and Clotilde Théry. "Biogenesis, secretion, and intercellular interactions of exosomes and other extracellular vesicles." Annual review of cell and developmental biology 30 (2014): 255-289.

Connerty Patrick, Alireza Ahadi, and Gyorgy Hutvagner. "RNA binding proteins in the miRNA pathway." International journal of molecular sciences 17.1 (2015): 31.

D'Astolfo, Diego S., et al. "Efficient intracellular delivery of native proteins." Cell 161.3 (2015): 674-690.

Dalby, Brian, et al. "Advanced transfection with Lipofectamine 2000 reagent: primary neurons, siRNA, and high-throughput applications." Methods 33.2 (2004): 95-103.

Dalmasso, Guillaume, et al. "MicroRNAs determine human intestinal epithelial cell fate." Differentiation 80.2-3 (2010): 147-154.

Dalmasso, Guillaume, et al. "Microbiota modulate host gene expression via microRNAs." PloS one 6.4 (2011): e19293.

Davis, Matthew PA, et al. "Kraken: a set of tools for quality control and analysis of high-throughput sequence data." Methods 63.1 (2013): 41-49.

de Jong, Olivier G., et al. "A CRISPR-Cas9-based reporter system for single-cell detection of extracellular vesicle-mediated functional transfer of RNA." Nature communications 11.1 (2020): 1113.

De Silvestri, Andrea, et al. "Determination of temperature dependent growth parameters in psychrotrophic pathogen bacteria and tentative use of mean kinetic temperature for the microbiological control of food." Frontiers in Microbiology 9 (2018): 3023.

de Voogt, Willemijn S., Marvin E. Tanenbaum, and Pieter Vader. "Illuminating RNA trafficking and functional delivery by extracellular vesicles." Advanced drug delivery reviews 174 (2021): 250-264.

Dobin, Alexander, et al. "STAR: ultrafast universal RNA-seq aligner." Bioinformatics 29.1 (2013): 15-21.

Dominkuš, Pia Pužar, et al. "PKH26 labeling of extracellular vesicles: Characterization and cellular internalization of contaminating PKH26 nanoparticles." Biochimica et Biophysica Acta (BBA)-Biomembranes 1860.6 (2018): 1350-1361.

Donaldson, Gregory P., S. Melanie Lee, and Sarkis K. Mazmanian. "Gut biogeography of the bacterial microbiota." Nature Reviews Microbiology 14.1 (2016): 20-32.

Dong H, Lei J, Ding L, et al. MicroRNA: function, detection, and bioanalysis[J]. Chemical reviews, 2013, 113(8): 6207-6233.

Dowdy, Steven F. "Overcoming cellular barriers for RNA therapeutics." Nature biotechnology 35.3 (2017): 222-229.

Du, Xiaochen, Ruth Ley, and Amy H. Buck. "MicroRNAs and extracellular vesicles in the gut: new host modulators of the microbiome?" Microlife 2 (2021): uqab010.

Eissa, Nour, et al. "Appropriateness of reference genes for normalizing messenger RNA in mouse 2, 4-dinitrobenzene sulfonic acid (DNBS)-induced colitis using quantitative real time PCR." Scientific Reports 7.1 (2017): 42427.

Ericsson, Anna, et al. "Functional characterization of the CCL25 promoter in small intestinal epithelial cells suggests a regulatory role for caudal-related homeobox (Cdx) transcription factors." The Journal of Immunology 176.6 (2006): 3642-3651.

Faget, Lauren, and Thomas S. Hnasko. "Tyramide signal amplification for immunofluorescent enhancement." ELISA: Methods and Protocols (2015): 161-172.

Felden, Brice, and Yoann Augagneur. "Diversity and versatility in small RNA-mediated regulation in bacterial pathogens." Frontiers in Microbiology 12 (2021): 719977.

Fernández, Javier, et al. "Colon microbiota fermentation of dietary prebiotics towards short-chain fatty acids and their roles as anti-inflammatory and antitumour agents: A review." Journal of Functional Foods 25 (2016): 511-522.

Frank, Jeremy A., et al. "Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes." Applied and environmental microbiology 74.8 (2008): 2461-2470.

Frankish, Adam, et al. "GENCODE reference annotation for the human and mouse genomes." Nucleic acids research 47.D1 (2019): D766-D773.

García-Martínez, Mariano, et al. "Extracellular vesicles released by J774A. 1 macrophages reduce the bacterial load in macrophages and in an experimental mouse model of tuberculosis." International Journal of Nanomedicine (2019): 6707-6719.

Gaudet, Ryan G., et al. "Cytosolic detection of the bacterial metabolite HBP activates TIFA-dependent innate immunity." Science 348.6240 (2015): 1251-1255.

Geekiyanage, Hirosha, et al. "Extracellular microRNAs in human circulation are associated with miRISC complexes that are accessible to anti-AGO2 antibody and can bind target mimic oligonucleotides." Proceedings of the National Academy of Sciences 117.39 (2020): 24213-24223.

Gil-Zamorano, Judit, et al. "Intestinal miRNAs regulated in response to dietary lipids." Scientific reports 10.1 (2020): 18921.

Hasan, Mohammad R., et al. "Depletion of human DNA in spiked clinical specimens for improvement of sensitivity of pathogen detection by next-generation sequencing." Journal of clinical microbiology 54.4 (2016): 919-927.

Hasan, Nihal, and Hongyi Yang. "Factors affecting the composition of the gut microbiota, and its modulation." PeerJ 7 (2019): e7502.

Helwak, Aleksandra, and David Tollervey. "Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH)." Nature protocols 9.3 (2014): 711-728.

Hendricks, Matthew R., et al. "Extracellular vesicles promote transkingdom nutrient transfer during viral-bacterial co-infection." Cell reports 34.4 (2021).

Hengen, Paul N., et al. "Information analysis of Fis binding sites." Nucleic acids research 25.24 (1997): 4994-5002.

Horne, Rachael, et al. "Microbe and host interaction in gastrointestinal homeostasis." Psychopharmacology 236 (2019): 1623-1640.

Ilahibaks, Nazma F., et al. "TOP-EVs: Technology of Protein delivery through Extracellular Vesicles is a versatile platform for intracellular protein delivery." Journal of Controlled Release 355 (2023): 579-592.

Im, Kyuseok, et al. "An introduction to performing immunofluorescence staining." Biobanking: methods and protocols (2019): 299-311.

Jayasena, Tharusha, et al. "Application of targeted mass spectrometry for the quantification of sirtuins in the central nervous system." Scientific reports 6.1 (2016): 35391.

Jensen, Brian, et al. "High extracellular calcium attenuates adipogenesis in 3T3-L1 preadipocytes." Experimental cell research 301.2 (2004): 280-292.

Johnston, Chris JC, et al. "Cultivation of Heligmosomoides polygyrus: an immunomodulatory nematode parasite and its secreted products." JoVE (Journal of Visualized Experiments) 98 (2015): e52412.

Johnstone, Rose M., et al. "Vesicle formation during reticulocyte maturation. Association of plasma membrane activities with released vesicles (exosomes)." Journal of Biological Chemistry 262.19 (1987): 9412-9420.

Johansson, Malin EV, Henrik Sjövall, and Gunnar C. Hansson. "The gastrointestinal mucus system in health and disease." Nature reviews Gastroenterology & hepatology 10.6 (2013): 352-361.

Johnsen, Peter Holger, et al. "Faecal microbiota transplantation versus placebo for moderate-to-severe irritable bowel syndrome: a double-blind, randomised, placebo-controlled, parallel-group, single-centre trial." The lancet Gastroenterology & hepatology 3.1 (2018): 17-24.

Johnson, Nathan R., et al. "Improved placement of multi-mapping small RNAs." G3: Genes, Genomes, Genetics 6.7 (2016): 2103-2111.

Jurj, Ancuta, et al. "A comprehensive picture of extracellular vesicles and their contents. Molecular transfer to cancer cells." Cancers 12.2 (2020): 298.

Kaczmarek, James C., Piotr S. Kowalski, and Daniel G. Anderson. "Advances in the delivery of RNA therapeutics: from concept to clinical reality." Genome medicine 9 (2017): 1-16.

Kalvari, Ioanna, et al. "Rfam 14: expanded coverage of metagenomic, viral and microRNA families." Nucleic Acids Research 49.D1 (2021): D192-D200.

Kalyuzhnaya, Marina G., et al. "Fluorescence in situ hybridization-flow cytometry-cell sorting-based method for separation and enrichment of type I and type II methanotroph populations." Applied and environmental microbiology 72.6 (2006): 4293-4301.

Kauppinen, Sakari, Birte Vester, and Jesper Wengel. "Locked nucleic acid (LNA): High affinity targeting of RNA for diagnostics and therapeutics." Drug Discovery Today: Technologies 2.3 (2005): 287-290.

Kelly, Arlene, et al. "A global role for Fis in the transcriptional control of metabolism and type III secretion in Salmonella enterica serovar Typhimurium." Microbiology 150.7 (2004): 2037-2053.

Khan, Raees, et al. "Fecal microbiota transplants for inflammatory bowel disease treatment: synthetic-and engineered communities-based microbiota transplants are the future." Gastroenterology Research and Practice 2022 (2022).

Khoruts, Alexander, and Michael J. Sadowsky. "Understanding the mechanisms of faecal microbiota transplantation." Nature reviews Gastroenterology & hepatology 13.9 (2016): 508-516.

Kloosterman, Wigard P., and Ronald HA Plasterk. "The diverse functions of microRNAs in animal development and disease." Developmental cell 11.4 (2006): 441-450.

Kong Shanshan, Yanhui H. Zhang, and Weiqiang Zhang. "Regulation of intestinal epithelial cells properties and functions by amino acids." BioMed research international 2018 (2018).

Kozomara, Ana, Maria Birgaoanu, and Sam Griffiths-Jones. "miRBase: from microRNA sequences to function." Nucleic acids research 47.D1 (2019): D155-D162.

Kristensen, Nadja B., et al. "Alterations in fecal microbiota composition by probiotic supplementation in healthy adults: a systematic review of randomized controlled trials." Genome medicine 8.1 (2016): 1-11.

Krol Jacek, Inga Loedige, and Witold Filipowicz. "The widespread regulation of microRNA biogenesis, function and decay." Nature Reviews Genetics 11.9 (2010): 597-610.

Kutter, Claudia, and Petr Svoboda. "miRNA, siRNA, piRNA: Knowns of the unknown." (2008): 181-188.

Langmead, Ben, and Steven L. Salzberg. "Fast gapped-read alignment with Bowtie 2." Nature methods 9.4 (2012): 357-359.

Larmonier, Claire B., et al. "Limited effects of dietary curcumin on Th-1 driven colitis in IL-10 deficient mice suggest an IL-10-dependent mechanism of protection." American journal of physiology-gastrointestinal and liver physiology 295.5 (2008): G1079-G1091.

Lässer, Cecilia, Su Chul Jang, and Jan Lötvall. "Subpopulations of extracellular vesicles and their therapeutic potential." Molecular aspects of medicine 60 (2018): 1-14.

Lathrop, Stephanie K., et al. "Replication of Salmonella enterica serovar Typhimurium in human monocyte-derived macrophages." Infection and immunity 83.7 (2015): 2661-2671.

Lawrence, Michael, et al. "Software for computing and annotating genomic ranges." PLoS computational biology 9.8 (2013): e1003118.

Lazenby, James J., Erica S. Li, and Cynthia B. Whitchurch. "Cell wall deficiency–an alternate bacterial lifestyle?" Microbiology 168.8 (2022): 001218.

Lewis, Benjamin P., Christopher B. Burge, and David P. Bartel. "Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets." cell 120.1 (2005): 15-20.

Li, Yong-Jiang, et al. "Emerging strategies for labeling and tracking of extracellular vesicles." Journal of Controlled Release 328 (2020): 141-159.

Li, Yuchen, et al. "EV-origin: Enumerating the tissue-cellular origin of circulating extracellular vesicles using exLR profile." Computational and Structural Biotechnology Journal 18 (2020): 2851-2859.

Liu, Jidong, et al. "Argonaute2 is the catalytic engine of mammalian RNAi." Science 305.5689 (2004): 1437-1441.

Liu, Shirong, et al. "The host shapes the gut microbiota via fecal microRNA." Cell host & microbe 19.1 (2016): 32-43.

Liu, Shirong, et al. "Oral administration of miR-30d from feces of MS patients suppresses MS-like symptoms in mice by expanding Akkermansia muciniphila." Cell Host & Microbe 26.6 (2019): 779-794.

Liu Ya-Juan, and Cheng Wang. "A review of the regulatory mechanisms of extracellular vesicles-mediated intercellular communication." Cell Communication and Signaling 21.1 (2023): 1-12.

Lu Jing, and Andrew Tsourkas. "Imaging individual microRNAs in single mammalian cells in situ." Nucleic acids research 37.14 (2009): e100-e100.

Luga, Valbona, et al. "Exosomes mediate stromal mobilization of autocrine Wnt-PCP signaling in breast cancer cell migration." Cell 151.7 (2012): 1542-1556.

Lunavat, Taral R., et al. "Small RNA deep sequencing discriminates subsets of extracellular vesicles released by melanoma cells–Evidence of unique microRNA cargos." RNA biology 12.8 (2015): 810-823.

Luo, Yingqin, et al. "Complete genome sequence of the universal killer Salmonella enterica Serovar Typhimurium UK-1 (ATCC 68169)." (2011): 4035-4036.

Luongo, Diomira, et al. "Differential modulation of innate immunity in vitro by probiotic strains of Lactobacillus gasseri." BMC microbiology 13.1 (2013): 1-12.

Madison, Blair B., et al. "Cis elements of the villin gene control expression in restricted domains of the vertical (crypt) and horizontal (duodenum, cecum) axes of the intestine." Journal of Biological Chemistry 277.36 (2002): 33275-33283.

Makrides, Savvas C. "Vectors for gene expression in mammalian cells." New Comprehensive Biochemistry 38 (2003): 9-26.

Maori, Eyal, et al. "A transmissible RNA pathway in honey bees." Cell reports 27.7 (2019): 1949-1959.

Martin, Fergal J., et al. "Ensembl 2023." Nucleic acids research 51.D1 (2023): D933-D941.

Matinha-Cardoso, Jorge, et al. "Novel protein carrier system based on cyanobacterial nano-sized extracellular vesicles for application in fish." Microbial Biotechnology 15.8 (2022): 219

Maurin, Thomas, et al. "HITS-CLIP in various brain areas reveals new targets and new modalities of RNA binding by fragile X mental retardation protein." Nucleic acids research 46.12 (2018): 6344-6355.

McClelland, Michael, et al. "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2." Nature 413.6858 (2001): 852-856.1-2207.

McKenna, Lindsay B., et al. "MicroRNAs control intestinal epithelial differentiation, architecture, and barrier function." Gastroenterology 139.5 (2010): 1654-1664.

McKenzie, Andrew J., et al. "KRAS-MEK signaling controls Ago2 sorting into exosomes." Cell reports 15.5 (2016): 978-987.

Meinke, Gretchen, et al. "Cre recombinase and other tyrosine recombinases." Chemical reviews 116.20 (2016): 12785-12820.

Moore, Michael J., et al. "miRNA–target chimeras reveal miRNA 3′-end pairing as a major determinant of Argonaute target specificity." Nature communications 6.1 (2015): 8864.

Morales-Kastresana, Aizea, et al. "Labeling extracellular vesicles for nanoscale flow cytometry." Scientific reports 7.1 (2017): 1878.

Mulcahy, Laura Ann, Ryan Charles Pink, and David Raul Francisco Carter. "Routes and mechanisms of extracellular vesicle uptake." Journal of extracellular vesicles 3.1 (2014): 24641.

Murillo, Oscar D., et al. "exRNA atlas analysis reveals distinct extracellular RNA cargo types and their carriers present across human biofluids." Cell 177.2 (2019): 463-477.

Natividad, Jane MM, and Elena F. Verdu. "Modulation of intestinal barrier by intestinal microbiota: pathological and therapeutic implications." Pharmacological research 69.1 (2013): 42-51.

Nawrocki, Eric P., and Sean R. Eddy. "Infernal 1.1: 100-fold faster RNA homology searches." Bioinformatics 29.22 (2013): 2933-2935.

Nishino, Kyohei, et al. "Analysis of endoscopic brush samples identified mucosa-associated dysbiosis in inflammatory bowel disease." Journal of gastroenterology 53 (2018): 95-106.

Noor, Zainab, et al. "Mass spectrometry–based protein identification in proteomics—a review." Briefings in bioinformatics 22.2 (2021): 1620-1638.

Norrman, Karin, et al. "Quantitative comparison of constitutive promoters in human ES cells." PloS one 5.8 (2010): e12413.

Olivares, Dolores, et al. "Optimization of small RNA library preparation protocol from human urinary exosomes." Journal of Translational Medicine 18 (2020): 1-9.

Olszewska, Magdalena A., et al. "Utilization of physiological and taxonomic fluorescent probes to study Lactobacilli cells and response to pH challenge." Microbiological research 192 (2016): 239-246.

Palma, Jaime, et al. "MicroRNAs are exported from malignant cells in customized particles." Nucleic acids research 40.18 (2012): 9125-9138.

Pasquinelli, Amy E. "MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship." Nature Reviews Genetics 13.4 (2012): 271-282.

Peck, Bailey CE, et al. "Functional transcriptomics in diverse intestinal epithelial cell types reveals robust microRNA sensitivity in intestinal stem cells to microbial status." Journal of Biological Chemistry 292.7 (2017): 2586-2600.

Peterson, Sarah M., et al. "Common features of microRNA target prediction tools." Frontiers in genetics 5 (2014): 23.

Phongsisay, Vongsavanh, Viraj N. Perera, and Benjamin N. Fry. "Expression of the htrB gene is essential for responsiveness of Salmonella typhimurium and

Campylobacter jejuni to harsh environments." Microbiology 153.1 (2007): 254-262.

Powell, D. R., A. J. Perry, and M. Milton. "Degust: interactive RNA-seq analysis. Zenodo." (2019).

Quinlan, Aaron R., and Ira M. Hall. "BEDTools: a flexible suite of utilities for comparing genomic features." Bioinformatics 26.6 (2010): 841-842.

Raposo, Graca, et al. "B lymphocytes secrete antigen-presenting vesicles." The Journal of experimental medicine 183.3 (1996): 1161-1172.

Raposo, Graça, and Willem Stoorvogel. "Extracellular vesicles: exosomes, microvesicles, and friends." Journal of Cell Biology 200.4 (2013): 373-383.

Ressel, Sarah, et al. "Respiratory syncytial virus directly binds miR-26, miR-27, and let-7 to de-repress gene targets through canonical and non-canonical mechanisms." bioRxiv (2023): 2023-06.

Ridder, Kirsten, et al. "Extracellular vesicle-mediated transfer of genetic information between the hematopoietic system and the brain in response to inflammation." PLoS biology 12.6 (2014): e1001874.

Ridder, Kirsten, et al. "Extracellular vesicle-mediated transfer of functional RNA in the tumor microenvironment." Oncoimmunology 4.6 (2015): e1008371.

Rinninella, Emanuele, et al. "What is the healthy gut microbiota composition? A changing ecosystem across age, environment, diet, and diseases." Microorganisms 7.1 (2019): 14.

Rolfe, Matthew D., et al. "Lag phase is a distinct growth phase that prepares bacteria for exponential growth and involves transient metal accumulation." Journal of bacteriology 194.3 (2012): 686-701.

Rondaan, Christien, et al. "COVID or no COVID: Interpreting inconclusive SARS-CoV-2 qPCR results in different populations and platforms." Journal of Clinical Virology Plus 3.2 (2023): 100145.

Rossen, Noortje G., et al. "The mucosa-associated microbiota of PSC patients is characterized by low diversity and low abundance of uncultured Clostridiales II." Journal of Crohn's and Colitis 9.4 (2015): 342-348.

Rufino-Ramos, David, et al. "Extracellular communication between brain cells through functional transfer of Cre mRNA mediated by extracellular vesicles." Molecular Therapy (2023).

Saal, Samuel, and Scott J. Harvey. "MicroRNAs and the kidney: coming of age." Current opinion in nephrology and hypertension 18.4 (2009): 317-323.

Santangelo, Laura, et al. "The RNA-binding protein SYNCRIP is a component of the hepatocyte exosomal machinery controlling microRNA sorting." Cell reports 17.3 (2016): 799-808.

Santos, André A., et al. "Host miRNA-21 promotes liver dysfunction by targeting small intestinal Lactobacillus in mice." Gut Microbes 12.1 (2020): 1840766.

Sauer, Brian, and Nancy Henderson. "Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1." Proceedings of the National Academy of Sciences 85.14 (1988): 5166-5170.

Seitz, Hervé, Jogender S. Tushir, and Phillip D. Zamore. "A 5′-uridine amplifies miRNA/miRNA* asymmetry in Drosophila by promoting RNA-induced silencing complex formation." Silence 2 (2011): 1-10.

Serban, Daniela Elena. "Gastrointestinal cancers: influence of gut microbiota, probiotics and prebiotics." Cancer letters 345.2 (2014): 258-270.

Servin, Alain L. "Antagonistic activities of lactobacilli and bifidobacteria against microbial pathogens." FEMS microbiology reviews 28.4 (2004): 405-440.

Sharif, S., M. H. Ghahremani, and M. Soleimani. "Delivery of exogenous miR-124 to glioblastoma multiform cells by Wharton's jelly mesenchymal stem cells decreases cell proliferation and migration, and confers chemosensitivity." Stem cell reviews and reports 14 (2018): 236-246.

Shi, Chenzhang, et al. "MicroRNA-21 knockout improve the survival rate in DSS induced fatal colitis through protecting against inflammation and tissue injury." PloS one 8.6 (2013): e66814.

Shi, Zonggao, Jeffrey J. Johnson, and M. Sharon Stack. "Fluorescence in situ hybridization for microRNA detection in archived oral cancer tissues." Journal of oncology 2012 (2012).

Shimomura, Takashi, et al. "New lipophilic fluorescent dyes for labeling extracellular vesicles: characterization and monitoring of cellular uptake." Bioconjugate Chemistry 32.4 (2021): 680-684.

Silahtaroglu, Asli N., et al. "Detection of microRNAs in frozen tissue sections by fluorescence in situ hybridization using locked nucleic acid probes and tyramide signal amplification." Nature protocols 2.10 (2007): 2520-2528.

Simbari, Fabio, et al. "Plasmalogen enrichment in exosomes secreted by a nematode parasite versus those derived from its mouse host: implications for exosome stability and biology." Journal of extracellular vesicles 5.1 (2016): 30741.

Singh, Natasha, et al. "The murine caecal microRNA signature depends on the presence of the endogenous microbiota." International journal of biological sciences 8.2 (2012): 171.

Sridhar, Sushmita, and Olivia Steele-Mortimer. "Inherent variability of growth media impacts the ability of Salmonella Typhimurium to interact with host cells." PLoS One 11.6 (2016): e0157043.

Sternberg, Nat, and Daniel Hamilton. "Bacteriophage P1 site-specific recombination: I. Recombination between loxP sites." Journal of molecular biology 150.4 (1981): 467-486.

Storz, Gisela, Jörg Vogel, and Karen M. Wassarman. "Regulation by small RNAs in bacteria: expanding frontiers." Molecular cell 43.6 (2011): 880-891.

Stratiki, Z., et al. "The effect of a bifidobacter supplemented bovine milk on intestinal permeability of preterm infants." Early human development 83.9 (2007): 575-579.

Sun, Yang, et al. "Resveratrol inhibits the migration and metastasis of MDA-MB-231 human breast cancer by reversing TGF-β1-induced epithelial-mesenchymal transition." Molecules 24.6 (2019): 1131.

Suzuki, Keiichiro, et al. "Aberrant expansion of segmented filamentous bacteria in IgA-deficient gut." Proceedings of the National Academy of Sciences 101.7 (2004): 1981-1986.

Tabibian, James H., Jayant A. Talwalkar, and Keith D. Lindor. "Role of the microbiota and antibiotics in primary sclerosing cholangitis." BioMed research international 2013 (2013).

Tippin, Brigette L., et al. "Biochemical characterization of fluorescent‐labeled recombinant human alpha‐l‐iduronidase in vitro." Biotechnology and applied biochemistry 58.6 (2011): 391-396.

R Core Team. “R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.” (2021). Available at: https://www.r-project.org/.

Robertson, Barbara, et al. "Specificity and functionality of microRNA inhibitors." Silence 1.1 (2010): 1-9.

Temoche-Diaz, Morayma M., et al. "Distinct mechanisms of microRNA sorting into cancer cell-derived extracellular vesicle subtypes." Elife 8 (2019): e47544.

Teng, Yun, et al. "Plant-derived exosomal microRNAs shape the gut microbiota." Cell host & microbe 24.5 (2018): 637-652.

Théry, Clotilde, et al. "Proteomic analysis of dendritic cell-derived exosomes: a secreted subcellular compartment distinct from apoptotic vesicles." The Journal of Immunology 166.12 (2001): 7309-7318.

Thoendel, Matthew, et al. "Comparison of microbial DNA enrichment tools for metagenomic whole genome sequencing." Journal of microbiological methods 127 (2016): 141-145.

Thursby, Elizabeth, and Nathalie Juge. "Introduction to the human gut microbiota." Biochemical journal 474.11 (2017): 1823-1836.

Thomson, James G., and Ann Blechl. "Recombinase technology for precise genome engineering." Advances in new technology for targeted modification of plant genomes (2015): 113-144.

Tian, Ye, et al. "Quality and efficiency assessment of six extracellular vesicle isolation methods by nano-flow cytometry." Journal of extracellular vesicles 9.1 (2020): 1697028.

Timár, Csaba I., et al. "Antibacterial effect of microvesicles released from human neutrophilic granulocytes." Blood, The Journal of the American Society of Hematology 121.3 (2013): 510-518.

Tsitsiou, Eleni, and Mark A. Lindsay. "microRNAs and the immune response." Current opinion in pharmacology 9.4 (2009): 514-520.

Turchinovich, Andrey, et al. "Characterization of extracellular circulating microRNA." Nucleic acids research 39.16 (2011): 7223-7233.

Valadi, Hadi, et al. "Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells." Nature cell biology 9.6 (2007): 654-659.

Van Deun, Jan, et al. "The impact of disparate isolation methods for extracellular vesicles on downstream RNA profiling." Journal of extracellular vesicles 3.1 (2014): 24858.

Van Niel, Guillaume, et al. "Intestinal epithelial cells secrete exosome–like vesicles." Gastroenterology 121.2 (2001): 337-349.

Van Niel, G., et al. "Intestinal epithelial exosomes carry MHC class II/peptides able to inform the immune system in mice." Gut 52.12 (2003): 1690.

Van Niel, Guillaume, et al. "Challenges and directions in studying cell–cell communication by extracellular vesicles." Nature Reviews Molecular Cell Biology 23.5 (2022): 369-382.

Van Schooneveld, Trevor C., Alan Gross, and Andre C. Kalil. "Duodenal infusion of feces for recurrent Clostridium difficile." N Engl J Med 368.22 (2013): 2143.

Veziroglu, Eren M., and George I. Mias. "Characterizing extracellular vesicles and their diverse RNA contents." Frontiers in Genetics 11 (2020): 700.

Vickers, Kasey C., et al. "MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins." Nature cell biology 13.4 (2011): 423-433.

Vickery, Thad W., Jennifer M. Kofonow, and Vijay R. Ramakrishnan. "Characterization of sinus microbiota by 16S sequencing from swabs." Diagnostic Bacteriology: Methods and Protocols (2017): 23-38.

Vidal, Karine, et al. "Immortalization of mouse intestinal epithelial cells by the SV40-large T gene: phenotypic and immune characterization of the MODE-K cell line." Journal of immunological methods 166.1 (1993): 63-73.

Viennois, Emilie, et al. "Host-derived fecal microRNAs can indicate gut microbiota healthiness and ability to induce inflammation." Theranostics 9.15 (2019): 4542.

Vrieze, Anne, et al. "Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome." Gastroenterology 143.4 (2012): 913-916.

Walters, William A., Zech Xu, and Rob Knight. "Meta-analyses of human gut microbes associated with obesity and IBD." FEBS letters 588.22 (2014): 4223-4233.

Wang, Herui et al. “Detection of Cytokine Receptors Using Tyramide Signal Amplification for Immunofluorescence.” Methods in molecular biology (Clifton, N.J.) vol. 2108 (2020): 89-97.

Wang, Hui, et al. "Genome-wide analysis of the salmonella Fis regulon and its regulatory mechanism on pathogenicity islands." PloS one 8.5 (2013): e64688.

Waters, Lauren S., and Gisela Storz. "Regulatory RNAs in bacteria." Cell 136.4 (2009): 615-628.

Wei, Zhiyun, et al. "Coding and noncoding landscape of extracellular RNA released by human glioma stem cells." Nature communications 8.1 (2017): 1145.

Weiberg, Arne, et al. "Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways." Science 342.6154 (2013): 118-123.

Wen, Li, et al. "Innate immunity and intestinal microbiota in the development of Type 1 diabetes." Nature 455.7216 (2008): 1109-1113.

White, Ruby, et al. "Extracellular vesicles from Heligmosomoides bakeri and Trichuris muris contain distinct microRNA families and small RNAs that could underpin different functions in the host." International Journal for Parasitology 50.9 (2020): 719-729.

Wickham, H., et al. "dplyr: a grammar of data manipulation. R package version 1.1. 2." (2023).

Williams, Charles, et al. "Metabolomics applied to the study of extracellular vesicles." Metabolites 9.11 (2019): 276.

Winfield, Mollie D., and Eduardo A. Groisman. "Role of nonhost environments in the lifestyles of Salmonella and Escherichia coli." Applied and environmental microbiology 69.7 (2003): 3687-3694.

Wu, Chung Wah, et al. "Detection of miR-92a and miR-21 in stool samples as potential screening biomarkers for colorectal cancer and polyps." Gut 61.5 (2012): 739-745.

Wu, Feng, et al. "MicroRNAs are differentially expressed in ulcerative colitis and alter expression of macrophage inflammatory peptide-2α." Gastroenterology 135.5 (2008): 1624-1635.

Wu, Hsin-Jung, et al. "Gut-residing segmented filamentous bacteria drive autoimmune arthritis via T helper 17 cells." Immunity 32.6 (2010): 815-827.

Wysoczynski, Marcin, and Mariusz Z. Ratajczak. "Lung cancer secreted microvesicles: underappreciated modulators of microenvironment in expanding tumors." International journal of cancer 125.7 (2009): 1595-1603.

Xiao, Liang, et al. "A catalog of the mouse gut metagenome." Nature biotechnology 33.10 (2015): 1103-1108.

Yan, Xiaowei, et al. "Dynamics of translation of single mRNA molecules in vivo." Cell 165.4 (2016): 976-989.

Yang, Yongzhi, et al. "Overexpression of miR-21 in patients with ulcerative colitis impairs intestinal epithelial barrier function through targeting the Rho GTPase RhoB." Biochemical and biophysical research communications 434.4 (2013): 746-752.

Yamamoto, Seiji, et al. "Significance of extracellular vesicles: pathobiological roles in disease." Cell Structure and Function 41.2 (2016): 137-143.

Yau, Tung On, et al. "Faecal microRNAs as a non-invasive tool in the diagnosis of colonic adenomas and colorectal cancer: A meta-analysis." Scientific reports 9.1 (2019): 9491.

Yaylak Bilge, and Bünyamin Akgül. "Experimental microRNA detection methods." miRNomics: MicroRNA Biology and Computational Analysis (2022): 33-55.

Yin, Pengbin, et al. "Exchange of genetic material: a new paradigm in bone cell communications." Cellular and Molecular Life Sciences 75 (2018): 1989-1998.

Yin, Yuan, et al. "Tumor-secreted miR-214 induces regulatory T cells: a major link between immune evasion and tumor growth." Cell research 24.10 (2014): 1164-1180.

Young, Alexander P., Daniel J. Jackson, and Russell C. Wyeth. "A technical review and guide to RNA fluorescence in situ hybridization." PeerJ 8 (2020): e8806.

Zabeo, Davide, et al. "Exosomes purified from a single cell type have diverse morphology." Journal of extracellular vesicles 6.1 (2017): 1329476.

Zhang, Xiaorong, et al. "MicroRNA directly enhances mitochondrial translation during muscle differentiation." Cell 158.3 (2014): 607-619.

Zhao, Hongyun, et al. "Tumor microenvironment derived exosomes pleiotropically modulate cancer cell metabolism." elife 5 (2016): e10250.

Zhao, Shanrong, et al. "QuickMIRSeq: a pipeline for quick and accurate quantification of both known miRNAs and isomiRs by jointly processing multiple samples from microRNA sequencing." BMC bioinformatics 18 (2017): 1-14.

Zhou, Cissy Chenyi, et al. "Targeted expression of Cre recombinase provokes placental-specific DNA recombination in transgenic mice." PloS one 7.2 (2012): e29236.

Zhu, Kegan, et al. "Plant microRNAs in larval food regulate honeybee caste development." PLoS genetics 13.8 (2017): e1006946.

Zomer, Anoek, et al. "In vivo imaging reveals extracellular vesicle-mediated phenocopying of metastatic behavior." Cell 161.5 (2015): 1046-1057.

Zomer, Anoek, et al. "Studying extracellular vesicle transfer by a Cre-loxP method." Nature protocols 11.1 (2016): 87-101.

Supplementary Tables

All the Supplementary Tables are found at:

https://drive.google.com/drive/folders/1iBJXHaZH4bCYy5R-sk57jPKXecQzpUH_

Supplementary Table 3.1 Differential expression analysis of miRNAs in PGM compared to Input. The table shows the fold change (FC) of miRNAs found in PGM relative to Input, false discovery rate (FDR), average expression (AveExpr), and p value.

Supplementary Table 5.1 Differential expression analysis of miRNAs in Mode-K EVs compared to EV-depleted supernatants. The table shows the fold change (FC) of miRNAs found in Mode-K EVs relative to EV-depleted supernatants, false discovery rate (FDR), average expression (AveExpr), and p value.