UNIVERSITY OF LATVIA

Faculty of Biology

Laila Silamiķele

DOCTORAL THESIS

THE EFFECT OF METFORMIN TREATMENT AND HOST-RELATED FACTORS ON THE GUT MICROBIOME IN DIET-INDUCED TYPE 2 DIABETES MOUSE MODEL

Promotion to the degree of Doctor of Biology

Molecular Biology

Supervisor: Dr. Biol., Prof. Jānis Kloviņš

Riga, 2022

The doctoral thesis was carried out in Latvian Biomedical Research and Study Centre at the Human Genetics and Molecular Medicine department and Faculty of Biology, Department of Molecular Biology, University of Latvia, from 2014 to 2022.

This research was supported by the European Regional Development Fund projects “Investigation of interplay between multiple determinants influencing response to metformin: search for reliable predictors for efficacy of type 2 diabetes therapy” (Project No. 1.1.1.1/16/A/091); and “Role of miRNAs in the host-gut microbiome communication during metformin treatment in the context of metabolic disorders” (Project No. 1.1.1.1/18/A/092); and by the European Social Fund project “Strengthening of the Capacity of Doctoral Studies at the University of Latvia within the Framework of the New Doctoral Model” (identification No. 8.2.2.0/20/I/006).

Form of the thesis: dissertation in biology, subfield – molecular biology.

Supervisor: Dr. biol., Prof. Jānis Kloviņš

Reviewers:

1) *PhD, Daiva Baltriukiene, The Life Sciences Center at Vilnius University*
2) *Dr. med., Prof., Baiba Jansone, University of Latvia*
3) *Dr. biol. Dace Pjanova, Latvian Biomedical Research and Study Centre*

The thesis will be defended at the public session of the Doctoral Committee of Biology, University of Latvia, at …... on ……... at Latvian Biomedical Research and Study Centre, Ratsupites Str. 1 k-1.

The thesis is available at the Library of the University of Latvia, Kalpaka blvd. 4.

This thesis is accepted for the commencement of the degree of Doctor of Biology on ……by the Doctoral Committee of Biology, University of Latvia.

Chairman of the Doctoral Committee _____/ /

Secretary of the Doctoral Committee _____/ /

ABSTRACT

Type 2 diabetes is a metabolic disease with increasing prevalence and burden on healthcare and the economy. Metformin, a first-line treatment for type 2 diabetes, is a medication that has strong effects on the gut microbiome. A substantial part of metformin users experiences gastrointestinal side effects or intolerance, limiting the use of this effective drug. Several modification strategies for gut microbiome exist, though these methods have limitations and are generally not targeted. Therefore, novel techniques on how to modulate gut microbiome precisely are necessary. These methods would be applicable not only in the context of metformin treatment of type 2 diabetes but other metformin targets as this medication has been shown to have pleiotropic effects, and its application is continuously extended. Furthermore, the targeted microbiome modulation method would be highly valued in numerous other microbiome-related disorders.

This study aimed to investigate the effects of metformin therapy and host-related factors on the gut microbiome in the intestinal tract of a mouse model of type 2 diabetes. First, we performed an animal experiment in which type 2 diabetes was modeled by a high-fat diet feeding in mice of both sexes, followed by a ten-week-long metformin treatment. Using various sequencing-based methods for microbiome composition determination, we characterized the metformin-induced microbiome alterations in fecal samples and various gastrointestinal tract sites. In addition, we sequenced fecal and mucosal miRNAs to determine whether there is a correlation between host miRNAs and the gut microbiome composition. To assess the translatability of the mice study, we evaluated the correlation between host miRNAs and gut microbiome in human subjects undergoing colonoscopy procedure.

We showed that metformin increases the relative abundance of various *Bacilli* and *Bacteroides* representatives in feces. Furthermore, we were able to locate this increase mainly in the distal small intestine, where in general, metformin’s effects on the gut microbiome were more pronounced, corresponding to the absorption pattern of the medication. We observed significant differences in metformin’s effects between sexes in fecal and intestinal microbiome samples and miRNA composition. We found a correlation between host miRNAs and distinct bacterial families and genera in mice and humans. We analyzed the potential biological interaction between host miRNAs and gut bacteria and found several miRNAs with potential bacterial targets. These results provide a systemic insight into metformin’s effects on the gut microbiome and a potential strategy for modulating it in a targeted fashion that should be studied in-depth in further functional studies.

KOPSAVILKUMS

Otrā tipa cukura diabēts ir vielmaiņas slimība ar pieaugošu izplatību un slogu uz veselības aprūpi un ekonomiku. Metformīns, pirmās izvēles līdzeklis 2. tipa diabēta ārstēšanai, ir zāles, kurām ir spēcīga ietekme uz zarnu mikrobiomu. Ievērojamai daļai metformīna lietotāju rodas kuņģa-zarnu trakta blakusparādības vai nepanesība, kas ierobežo šo efektīvo zāļu lietošanu. Pastāv vairākas zarnu mikrobioma modifikācijas stratēģijas, lai gan šīm metodēm ir ierobežojumi un parasti tās nav mērķētas. Ir nepieciešami jauni paņēmieni, kā precīzi modulēt zarnu mikrobiomu. Šīs metodes būtu piemērojamas ne tikai otrā tipa cukura diabēta metformīna terapijas kontekstā, bet arī citiem metformīna mērķiem, jo ir pierādīts, ka šīm zālēm ir pleiotropiska iedarbība, un to lietošana tiek nepārtraukti paplašināta. Turklāt mērķētās mikrobioma modulēšanas metode būtu vērtīga daudzu citu ar mikrobiomu saistītu slimību ārstēšanā.

Šī pētījuma mērķis bija analizēt metformīna terapijas un ar saimniekorganismu saistīto faktoru ietekmi uz zarnu mikrobiomu otrā tipa cukura diabēta peļu modeļa zarnu traktā. Pirmkārt, mēs veicām eksperimentu ar dzīvniekiem, kurā abu dzimumu pelēm ar diētas ar paaugstinātu tauku saturu palīdzību tika ierosināts otrā tipa cukura diabēts, kam sekoja desmit nedēļu ilga metformīna terapija. Izmantojot dažādas uz sekvencēšanu balstītas metodes mikrobioma sastāva noteikšanai, mēs raksturojām metformīna izraisītās mikrobioma izmaiņas fēču un dažādu zarnu trakta vietu paraugos. Papildus tam, mēs sekvencējām fēču un gļotādas miRNS, lai noteiktu, vai pastāv korelācija starp saimniekorganisma miRNS un zarnu mikrobioma sastāvu. Lai novērtētu peļu pētījuma rezultātu pārnesamību, mēs analizējām saimniekorganisma miRNS un zarnu mikrobioma korelāciju cilvēkos, kam veikta rutīnas kolonoskopijas procedūra.

Mēs parādījām, ka metformīns palielina dažādu *Bacilli* un *Bacteroides* pārstāvju relatīvo sastopamību fēcēs. Turklāt mēs varējām lokalizēt šo pieaugumu galvenokārt tievās zarnas distālajā daļā, kur kopumā metformīna ietekme uz zarnu mikrobiomu bija izteiktāka, kas atbilst zāļu absorbcijas raksturojumam. Mēs novērojām būtiskas atšķirības metformīna iedarbībā starp dzimumiem fēču un zarnu mikrobioma paraugos un miRNS sastāvā. Mēs noskaidrojām, ka pastāv korelācija starp saimniekorganisma miRNS un noteiktām baktēriju dzimtām un ģintīm pelēs un cilvēkos. Mēs analizējām iespējamo bioloģisko mijiedarbību starp saimniekorganisma miRNS un zarnu baktērijām un atradām vairākas miRNS ar iespējamiem bakteriāliem mērķiem. Šie rezultāti sniedz sistēmisku ieskatu par metformīna ietekmi uz zarnu mikrobiomu un iespējamo stratēģiju tā mērķtiecīgai modulēšanai, kas būtu padziļināti jāizpēta turpmākos funkcionālajos pētījumos.

TABLE OF CONTENTS

ABBREVIATIONS

16S rRNA	16S ribosomal RNA
3D	three-dimensional
3Rs	replacement, reduction, refinement
Ago2	argonaute 2
AMPK	adenosine monophosphate-activated protein kinase
ANCOM-BC	analysis of compositions of microbiomes with bias correction
ASV	amplicon sequence variant
ATP	adenosine triphosphate
bp	base pair
CCS	continuous culture system
CD	control diet
cDNA	complementary DNA
CEC	cecum
clr	centered log ratio normalization
COL	colon
DBG	de-bruijn graph
DGCR8	DiGeorge critical region 8
DPP4	dipeptidyl peptidase 4
DSI	distal small intestine
ELISA	enzyme-linked immunosorbent assay
ENS	effective number of species
EV	extracellular vesicle
FDR	false discovery rate
FMT	fecal microbiota transplantation
GLP1	glucagon-like peptide 1
HFD	high-fat diet
HOMA-IR	Homeostatic Model Assessment for Insulin Resistance
HuMiX	human-microbial cross talk
IEC	intestinal epithelial cells
IgA	immunoglobulin A
IQR	interquartile range
KEGG	Kyoto Encyclopedia of Genes and Genomes
KK	Kuo Kondo
L	lumen
LDL	low-density lipoprotein
LogFC	log fold change
M	mucosa
MATE1	multidrug and toxin extrusion protein 1
MATE2	multidrug and toxin extrusion protein 2
Met-	without metformin
Met+	with metformin
MFE	minimum free energy
mGPD	mitochondrial glycerophosphate dehydrogenase
miRNA	micro RNA
miRNP	microribonuclear protein
MMC	multiprotein complex
MODY	maturity-onset diabetes of the young
MRE	miRNA response element
mRNA	messenger RNA
M-SHIME	Simulator of the Human Intestinal Microbial Ecosystem with the mucosal compartment
nt	nucleotides
NZO	New Zealand obese
OCT1	organic cation transporter 1
OCT2	organic cation transporter 2
OCT3	organic cation transporter 3
OCTN1	carnitine/ organic cation transporter 1
OLETF	Otsuka Long Evans Tokushima Fatty
PCA	principal component analysis
PCR	polymerase chain reaction
PMAT	plasma membrane transporter
pre-miRNA	precursor micro RNA
pri-miRNA	primary micro RNA
PSI	proximal small intestine
QC	quality control
RISC	RNA-induced gene silencing complex
RNase III	ribonuclease III

SCFA	short-chain fatty acid
SERT	serotonin reuptake transporter
SGLT1	sodium glucose co-transporter 1
SGLT2	sodium glucose co-transporter 2
SHIME	Simulator of the Human Intestinal Microbial Ecosystem
SIDLE	SMURF Implementation Done to acceLerate Efficiency
SIRT1	sirtuin 1
SMURF	Short Multiple Reads Framework
SPF	specific pathogen-free
STZ	streptozotocin
T1D	type 1 diabetes
T2D	type 2 diabetes
THTR2	thiamine transporter 2
TIM1	TNO (The Netherlands Organisation for Applied Scientific Research) Gastro-Intestinal Model 1
TIM2	TNO Gastro-Intestinal Model 2
TZD	thiazolidinedione
UMI	unique molecular index
UTR	untranslated region

INTRODUCTION

Type 2 diabetes is a common metabolic disease characterized by chronic hyperglycemia, increasing prevalence, healthcare burden, and associated costs. Metformin is an antidiabetic drug that, in addition to controlling glucose levels, strongly affects the gut microbiome. Metformin provides effective and inexpensive treatment; however, some patients develop gastrointestinal side effects or drug intolerance. Research into the effects of metformin on the gut microbiome is an active area of biomedical studies, as this knowledge may improve the utility of the drug. In addition, other therapeutic targets for metformin are increasingly being discovered, demanding improved management of its use. Controlling gut microbiome changes associated with metformin therapy may improve metformin efficacy; therefore, targeted microbiome modulation strategies are needed for successful metformin treatment management.

Importance of this work: An in-depth, systematic study of the effects of metformin therapy on the gut microbiome in feces and the entire intestinal tract under controlled conditions provides valuable information on the interaction between metformin and the microbiome. It is also necessary to evaluate other host-related factors that may affect this interaction. Once metformin's interaction with members of the gut microbiome has been characterized, it is necessary to search for targeted microbiome modification methods that could serve as a tool to improve therapy. Such knowledge allows personalized care of patients in the context of various diseases.

Aim of the study: To investigate the effects of metformin therapy and host-related factors on the gut microbiome in the intestinal tract of a mouse model of type 2 diabetes and test those factors for other conditions.

Tasks to achieve the aim:

(1) Determine whether and how metformin therapy affects the gut microbiome and its functions using fecal metagenomic analysis.

(2) Characterize the composition of the gut microbiome in different parts and layers of the intestinal tract and analyze the effect of metformin therapy on it using the 16S rRNA gene sequencing analysis of gut microbiome samples.

(3) Determine whether the host miRNAs are able to affect the gut microbiome using miRNA analysis of fecal and intestinal samples in combination with microbial composition data both in mice and humans.

1 LITERATURE REVIEW

1.1 Type 2 diabetes

Diabetes mellitus is a term for a complex and heterogeneous group of metabolic diseases characterized by chronic hyperglycemia caused by impaired insulin secretion, disturbed insulin effect, or a combination of both (Petersmann *et al.*, 2019). The prevalence of diabetes is constantly increasing – it has been estimated that 9.3% of the adults aged 20–79 years (463 million adults) had diabetes globally in 2019, and it is expected to rise to 578 million by 2030 (Saeedi *et al.*, 2019). The most common form of diabetes mellitus is type 2 diabetes (T2D) which accounts for 90-95% of all diabetes cases (American Diabetes Association, 2019), followed by type 1 diabetes (T1D), occurring in 5-10% of diabetes patients (Mobasseri *et al.*, 2020). According to the current classification, other forms of diabetes mellitus include gestational diabetes (hyperglycemia that develops during pregnancy); monogenic diabetes, including neonatal diabetes and maturity-onset diabetes of the young (MODY); and other rare conditions, for example, pancreatogenic or type 3c diabetes (Redondo *et al.*, 2020).

Type 2 diabetes occurs when insulin resistance is acquired (usually from obesity) on the background of inherited and acquired beta cell dysfunction (Pearson, 2019). T2D results from several disturbances in glucose homeostasis, including impaired insulin secretion; insulin resistance in muscle, liver, and adipocytes; and deviations in glucose uptake in the gut and liver (DeFronzo, 2004).

Glucose homeostasis is maintained by a complex interaction of glucoregulatory hormones, including insulin, glucagon, somatostatin, amylin, and incretins (Campbell, 2011; Andersen and Holst, 2022). Under normal physiology, the increase in plasma glucose concentration caused by glucose ingestion stimulates insulin release. The combination of high insulin and high glucose levels stimulates glucose uptake by the liver, gut, and peripheral tissues, mainly muscle, and suppresses hepatic glucose output (DeFronzo, 2004). The incapacity of insulin to suppress hepatic gluconeogenesis is a major contributor to hyperglycemia characteristic of T2D (Hatting *et al.*, 2019). Sustained hyperglycemia and hyperinsulinemia result in insulin resistance in both human subjects and animal models (Lee *et al.*, 2017). Furthermore, prolonged hyperglycemia triggers beta cell dysfunction, which in turn impairs insulin secretion (Cerf, 2013). As a result, T2D gradually develops, and the chronically increased glucose level is associated with long-term damage,

complications, and failure of various organ systems (American Diabetes Association, 2013). The pathophysiology and main affected organs by T2D have been illustrated in Figure 1.

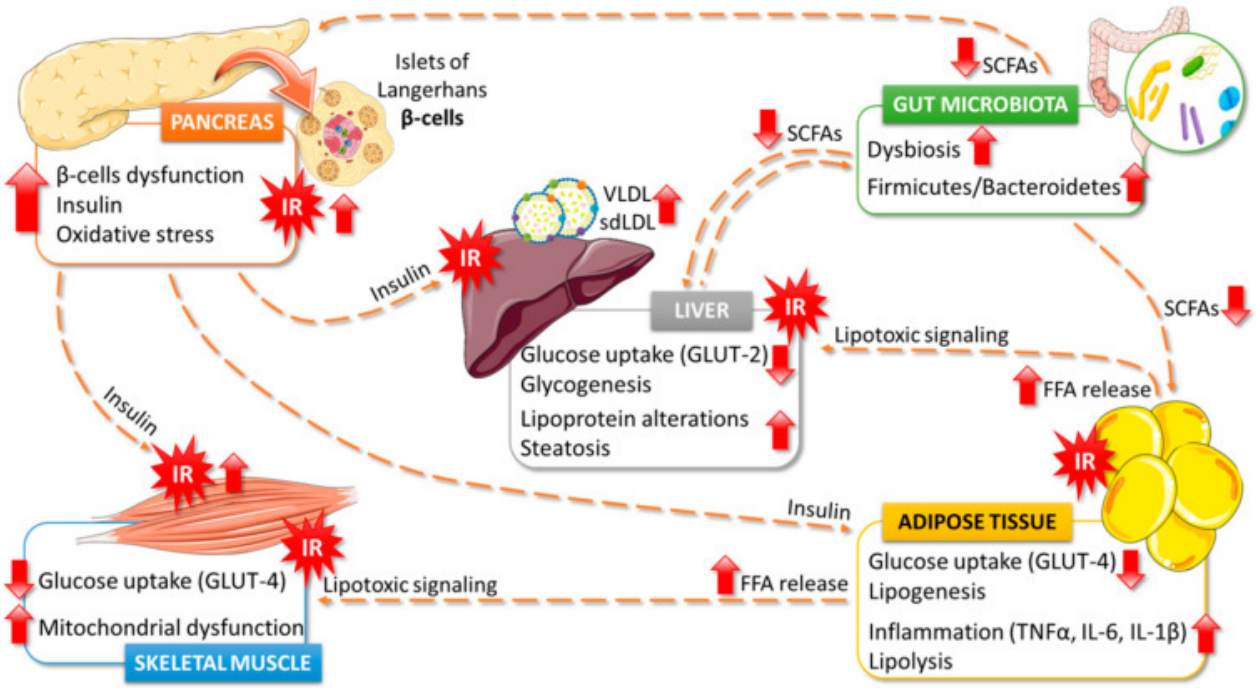

Figure 1. Type 2 diabetes pathophysiology – effects in main affected organs. FFA – free fatty acids; GLUT2 – type 2 glucose transporter; GLUT4 – type 4 glucose transporter; IL-1β – interleukin 1 β; IL-6 – interleukin 6; IR – insulin resistance; sdLDL – small dense low density lipoproteins; SCFAs – short-chain fatty acids; TNFα – tumor necrosis factor α; VLDL – very low density lipoproteins. Adapted from (Bocanegra *et al.*, 2021).

Due to a long asymptomatic period of five to seven years from onset to diagnosis, many patients develop complications before diagnosis. Complications of T2D are separated into two main groups – microvascular and macrovascular. Microvascular complications are more common and include retinopathy, nephropathy, neuropathy, and diabetic foot, while macrovascular complications include cerebrovascular disease, arrhythmias, coronary artery disease, and peripheral artery disease (Viigimaa *et al.*, 2020; Aikaeli *et al.*, 2022).

1.2 Type 2 diabetes treatment

Various international and regional guidelines have been developed for the management of T2D (Mohan *et al.*, 2020; World Health Organization, 2020; Mannucci *et al.*, 2022). A combination of non-pharmacological and pharmacological treatment is most often recommended. Non-pharmacological management includes modifications in risk factors related to lifestyle –

improved diet, reduced calorie intake for overweight patients, physical activities, and avoiding smoking and excessive alcohol use (World Health Organization, 2020).

If lifestyle interventions fail to maintain glycemic control, a variety of medications are available to treat type 2 diabetes. These pharmacological agents can be divided into ten main classes: drugs that reduce hepatic glucose output (biguanides); drugs that promote insulin release from the beta cells (sulphonylureas); drugs that increase incretin levels via blocking dipeptidylpeptidase-4 (DPP4 inhibitors); drugs that promote glucose excretion with urine (sodium-glucose co-transporter-2 (SGLT2) inhibitors) (Whaley *et al.*, 2012); glucagon-like peptide 1 (GLP1) receptor agonists that increase insulin secretion and inhibit glucagon secretion; drugs that delay the digestion and intestinal absorption of complex carbohydrates thus controlling postprandial hyperglycemia (alpha-glucosidase inhibitors) (Hossain *et al.*, 2020); drugs that increase the insulin sensitivity mainly due to changes in body fat metabolism (thiazolidinediones) (Greenfield and Chisholm, 2004); dopamine agonists that decrease insulin resistance; bile acid sequestrants that bind bile acids in the intestine and excrete LDL cholesterol in feces; and insulin secretagogues distinct from sulfonylureas (meglitinides) (Greenfield and Chisholm, 2004; Whaley *et al.*, 2012; Hossain *et al.*, 2020; Padhi, Kumar and Behera, 2020; Taylor, Yazdi and Beitelshees, 2021).

Metformin is recommended as an initial pharmacological treatment to control blood glucose levels, increasing the dosage gradually. However, 20% of patients experience metformin intolerance characterized by gastrointestinal side effects, and 5% cannot continue the treatment due to the severity of intolerance manifestations (McCreight *et al.*, 2018).

Second-generation sulfonylurea is recommended as a first-line treatment when metformin is contraindicated or not tolerated. If metformin and sulfonylurea fail to control glycemia, insulin treatment might be initiated or used in combination with oral agents. However, if insulin is not suitable, another medication might be added, including SGLT2 inhibitors, GLP1 receptor agonists, DPP4 inhibitors, and thiazolidinediones, with each of them having specific efficacy levels, effects on cardiovascular system, weight, and risks for side effects in addition to a high cost (Anderson, 2020; World Health Organization, 2020).

1.3 Metformin and its mechanisms of action

Metformin, a biguanide, is the first-line therapy for the treatment of T2D. According to the American Diabetes Association and European Association for the study of Diabetes guidelines, it

is the preferred option for initiating glucose-lowering due to its efficacy, safety, tolerability, and low cost (Davies *et al.*, 2018). It has been widely used in clinical practice for over 65 years and attenuates not only hyperglycemia but also diabetes mortality and complications (Nasri and Rafieian-Kopaei, 2014; Bailey, 2017). The applicability of metformin is constantly expanding, as evidenced by the growing number of new roles of the medication in the context of numerous cancers (Saraei *et al.*, 2019), cardiovascular diseases (Rena and Lang, 2018), aging (S. Chen *et al.*, 2022), liver diseases (Brandt *et al.*, 2019), obesity (Yerevanian and Soukas, 2019), polycystic ovary syndrome (Zhao *et al.*, 2021), and renal diseases (Ren *et al.*, 2019).

Although metformin has been shown to exert pleiotropic effects on glucose metabolism, general agreement exists that the main activity is mainly due to the inhibition of hepatic gluconeogenesis (Lamoia and Shulman, 2021). However, the precise mechanism is still discussed despite being studied for decades. Several pathways in the liver are potentially involved, including inhibition of mitochondrial complex I that provides indirect AMP-activated protein kinase (AMPK) activation, direct activation of AMPK, suppression of adenylate cyclase function that attenuates glucagon signaling, and inhibition of mitochondrial glycerophosphate dehydrogenase (mGPD) (Wu, Horowitz and Rayner, 2017).

There is increasing evidence that metformin's action mechanism is associated with physiological processes in the gastrointestinal tract. For example, a more pronounced effect of metformin can be observed when the drug is administered orally than intravenously at an equivalent dose (Stepensky *et al.*, 2002). Metformin accumulates in gastrointestinal tissues (Kinaan, Ding and Triggle, 2015); for example, it is 30-300 times more concentrated in the small intestine than in plasma, and 30-50% of the drug reaches the colon and is eliminated with feces.

1.4 Metformin pharmacokinetics

The fate of metformin in the body has been illustrated in Figure 2. Upon oral administration, metformin is absorbed by the enterocytes of the proximal small intestine. Absorption of metformin in the gastrointestinal tract is incomplete, and its oral bioavailability has been estimated to be approximately 55% (Graham *et al.*, 2011). Passive diffusion has been suggested to be responsible for half of the intestinal uptake of metformin (Szymczak-Pajor, Wenclewska and Sliwinska, 2022). The rest of the medication is transported by facilitated diffusion using several transporters. The plasma membrane transporter (PMAT) and the organic cation transporter 3 (OCT3), which has a higher affinity for the drug, are the primary transporters involved in metformin absorption; both

are located on the polarized apical membranes of enterocytes (Szymczak-Pajor, Wenclewska and Sliwinska, 2022). Metformin transport into the portal vein is also mediated by organic cation transporter 1 (OCT1) located in the basolateral membrane of enterocytes. Other transporters potentially involved in the intestinal absorption of metformin, including serotonin re-uptake transporter (SERT), thiamine transporter 2 (THTR2), and carnitine/organic cation transporter 1 (OCTN1), have also been described (Liang and Giacomini, 2017). Peak plasma concentrations of metformin can be detected approximately 3 hours after the intake (Graham *et al.*, 2011).

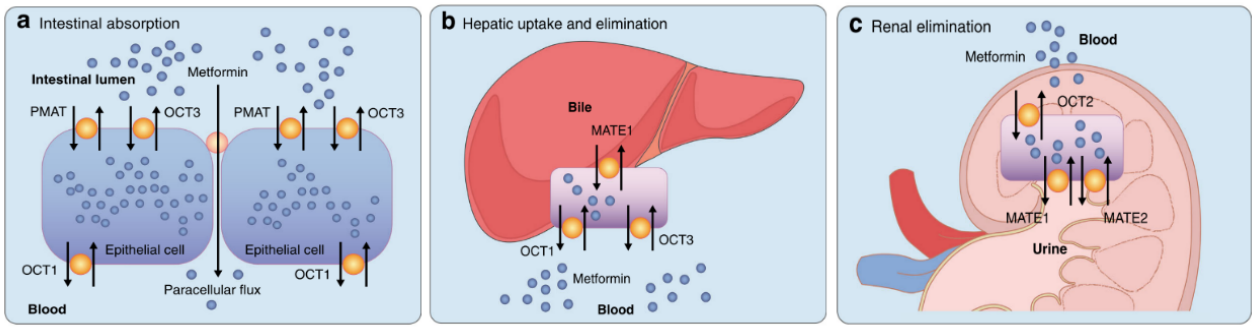

Figure 2. Metformin transporters and the fate of metformin in the body – absorption and elimination. (a) Metformin uptake by enterocytes. (b) Metformin uptake by hepatocytes and disposal into bile. (c) Uptake into renal epithelial cells and excretion of unchanged metformin into the urine. Adapted from (Florez, 2017).

Metformin is not metabolized in the liver; however, the multidrug and toxin extrusion protein 1 (MATE1) found in hepatocytes is involved in the biliary excretion of an unchanged drug or its transport by blood to the kidney (Szymczak-Pajor, Wenclewska and Sliwinska, 2022). The primary route of metformin clearance is renal excretion of the drug in the unchanged form via tubular secretion in the urine. Metformin transport to the kidney is mediated by MATE1, MATE2, and the leading transporter that ensures drug uptake in renal epithelial cells – OCT2. MATE1 has a higher affinity for metformin than MATE2, and its absence causes lactic acidosis, a significant side effect of the drug (Toyama *et al.*, 2012). Up to 30% of an oral dose of metformin is excreted unchanged into feces (McCreight, Bailey and Pearson, 2016).

1.5 Microbiome

The human body is an ecosystem comprised of trillions of microorganisms (bacteria, archaea, microscopic eukaryotes), known collectively as the microbiota (Hou *et al.*, 2022). Older definitions also include viruses, but since they are not living organisms, they have been excluded lately. A microbiome is a collection of microorganisms in a specific site, together with their

genomes and the surrounding environmental conditions, structural elements, and metabolites (Berg *et al.*, 2020). Research of the microbiome and its impact on human health has been a rapidly growing area of biomedical science. The main human body niches include skin, vaginal, gastrointestinal, and oral microbiomes, although the research on other body sites constantly intensifies (Reynoso-García *et al.*, 2022).

It is widely accepted that colonization with the first microbial communities occurs at birth (Houghteling and Walker, 2016). This is supported by differences in the composition of the neonatal microbiome depending on the method of delivery – a vaginally delivered newborn is inoculated with the mother's vaginal and intestinal bacteria, while an infant born by cesarean section is colonized by environmental bacteria with a smaller diversity and lacking *Bacteroidetes* (Dominguez-Bello *et al.*, 2010). Contrary to the existing dogma that the fetus is sterile, *in utero* colonization has been proposed lately (Stinson *et al.*, 2019). Additional factors that affect a child's microbiome in the first years of life include feeding type, use of antibiotics, and household pet keeping (Kim *et al.*, 2020).

The human microbiome is constantly evolving and changing in response to various factors throughout life, including genetic background, age, diet, hormonal changes, underlying diseases, lifestyle, and environmental interactions (Ogunrinola *et al.*, 2020). Nevertheless, each person has distinct and stable microbiota, evidenced by the observation that samples from the same individual are more similar than those from different individuals (Dickson, 2019).

1.6 Gut microbiome

The most diverse and most studied human microbiome subpopulation is the gut microbiome. According to the latest human gut microbiome reference set, almost 3600 species can be found in the gut (Leviatan *et al.*, 2022). At three to five years of age, the establishment of the core gut microbiome appears (Rodríguez *et al.*, 2015). The latest decade has shown that deviations from the core microbiome, collectively named dysbiosis, are associated with a wide variety of diseases, in particular obesity, diabetes, and inflammatory bowel diseases, though it is still unknown what confers these deviations and how the core microbiome is maintained (Singh *et al.*, 2021). The potential for modulating the gut microbiome or host-microbiome interactions offers new therapeutic strategies for many chronic diseases.

The gut microbiome is dominated by four phyla – *Bacteroidetes* (Gram-negative, anaerobic), *Firmicutes* (Gram-positive, anaerobic or obligate or facultatively aerobic), *Actinobacteria* (Gram-

positive, anaerobic), and *Proteobacteria* (Gram-negative, aerobic or facultative anaerobic) in both mice and humans (Belizário and Napolitano, 2015). However, the functional niche for each of these phyla is different as *Bacteroidetes* mainly produces acetate and propionate, but *Firmicutes* – butyrate (Thursby and Juge, 2017).

The gut microbiome exists in close interaction with the host and performs many essential functions, summarized in Figure 3. The main functions of the gut microbiome include digestion of food compounds, synthesis of vitamins and amino acids; immune system development and training; ensuring colonization resistance against opportunistic pathogens; and mechanical strengthening of the intestinal epithelium (Laukens *et al.*, 2015).

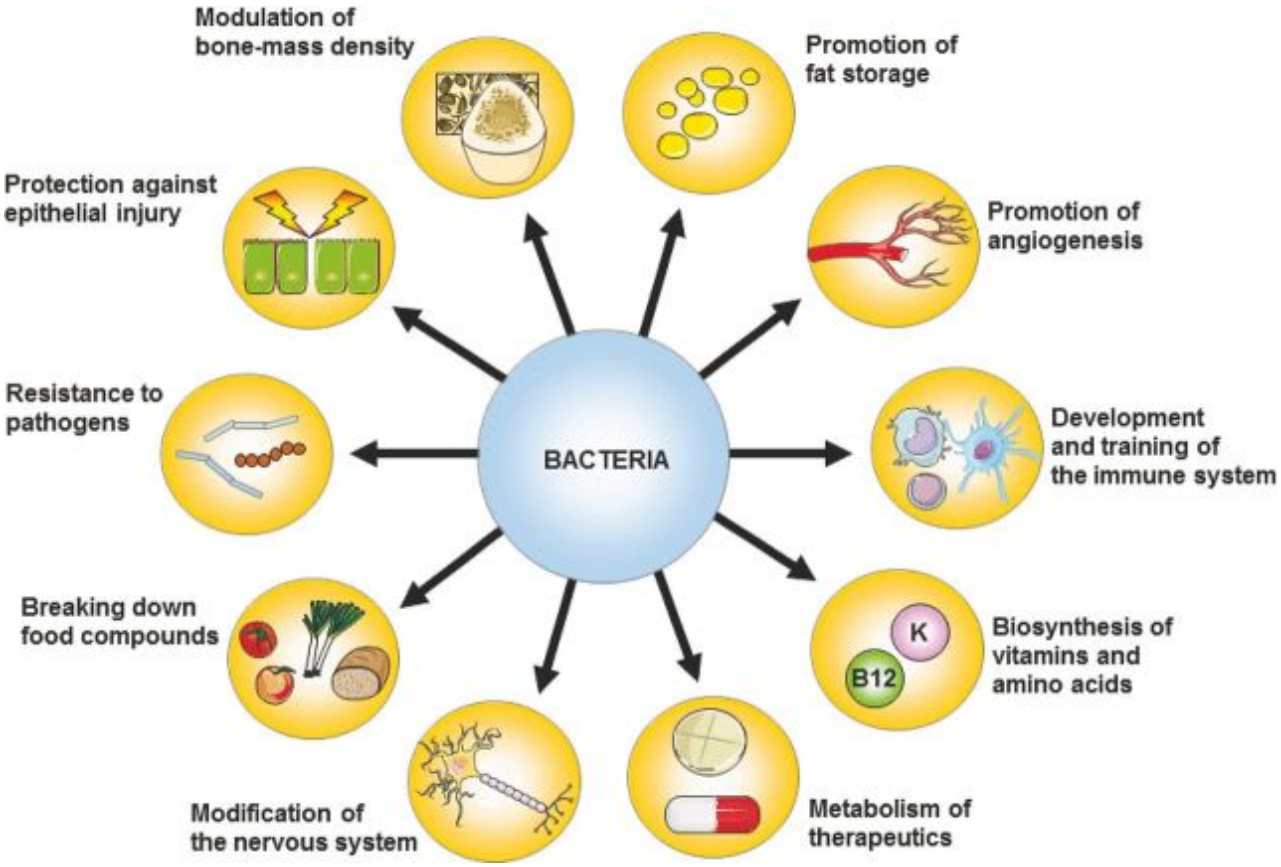

Figure 3. Key functions of the gut microbiome. Adapted from (Laukens *et al.*, 2015).

1.7 Spatial variation of the gut microbiome

Gut microbiome composition and density vary along the gastrointestinal tract longitudinally and cross-sectionally (Tropini *et al.*, 2017). Components contributing to this diversity include nutrient availability, chemical gradients, oxygen levels, mucosal structure, the diameter of the lumen, and cellular composition (Donaldson, Lee and Mazmanian, 2015).

Functional differences between distinct regions of the gut correspond to the distribution of the microbial genera with respect to both identity and abundance. Studies in mice, piglets, and humans have shown that the proximal part of the small intestine – duodenum-jejunum, which is characterized by faster transit and facilitation of simple sugar and amino acid metabolism, is dominated by facultative anaerobes *Lactobacillus*, *Proteobacteria*, and obligate anaerobes *Bacteroides*, distal small intestine – ileum by *Fusobacterium* and *Escherichia* and the cecum and colon where the passage is slower and metabolism advantages fermentation of complex polysaccharides arising from undigested fibers or host mucus is prevailed by saccharolytic *Bacteroidales*, *Clostridiales*, and *Prevotella* (Gu *et al.*, 2013; Tropini *et al.*, 2017; Y. Liu *et al.*, 2019). In healthy mice, *Lactobacillaceae* dominates the stomach and small intestine, while in the large intestine, the most abundant families are *Bacteroidaceae*, *Prevotellaceae*, *Rikenellaceae*, *Lachnospiraceae*, and *Ruminococcaceae* (Yamamoto *et al.*, 2018).

Microbial communities differ on the transverse axis as well. Bacteria associated with the gut mucosa remain located at the same place; thus, these depend on the available substrate, whereas bacteria in the gut lumen can freely associate with various substrates. The mucosa is inhabited by aerotolerant taxa such as *Proteobacteria* and *Actinobacteria*, *Clostridium*, *Lactobacillus*, *Enterococcus*, and *Akkermansia*, while other bacteria predominate towards the lumen – *Bacteroides*, *Bifidobacterium*, *Streptococcus*, *Enterobacteriaceae*, and *Ruminoccoccus* among others (Jandhyala *et al.*, 2015; Tropini *et al.*, 2017). The mucosa-associated microbiota varies in different intestinal parts. Although fecal samples representing luminal content can provide information on global microbiota composition to a certain extent, research on mucosa-associated microbial communities can provide critical information on the interaction between gut microbiota and host (Kirchgessner, 2016).

1.8 Metformin effects on the gut microbiome

Studies in humans have convincingly shown that metformin specifically alters the gut microbiome both in T2D patients (Napolitano *et al.*, 2014; Forslund *et al.*, 2015; Wu *et al.*, 2017) and in healthy subjects (Elbere *et al.*, 2018; Bryrup *et al.*, 2019). A multi-country cross-sectional study employing metformin-untreated T2D patients, metformin-treated T2D patients, and non-diabetic controls has shown depletion of butyrate producers such as *Roseburia*, *Subdoligranulum*, and butyrate-producing *Clostridiales spp*. in the metformin-untreated T2D subset; in turn, metformin treatment significantly increased the abundance of *Escherichia spp*. and decreased the abundance of *Intestinibacter* and, when analyzed functionally, significantly augmented butyrate

and propionate production potential (Forslund *et al.*, 2015). These results have been approved further by a longitudinal double-blind study on individuals with treatment-naive T2D (Napolitano *et al.*, 2014). Another direction of metformin's effects on the gut microbiome is the increased abundance of the mucin-degrading *Akkermansia muciniphila* (De La Cuesta-Zuluaga *et al.*, 2017). Studies in healthy individuals have reported an increase in *Escherichia/Shigella* and a decrease in the abundance of *Intestinibacter spp.* and *Clostridium spp.* (Elbere *et al.*, 2018; Bryrup *et al.*, 2019). It has been shown that metformin exerts its effects on the microbiome already in the first 24 hours (Elbere *et al.*, 2018) and that the alterations in the gut microbiome could be related to the metformin-induced immune response (Ustinova *et al.*, 2020).

A number of studies that have specifically targeted metformin effects in rodents have used 16S rRNA sequencing and have identified that metformin modifies the metabolic profile of high-fat diet-fed animals accompanied by changes in the microbiome (Zhang *et al.*, 2015; Lee *et al.*, 2018; Ji, Wang and Li, 2019). Reduced abundances of *Akkermansia* and *Alistipes* and increases in the proportions of *Anaerotruncus*, *Lactococcus*, *Parabacteroides*, *Odoribacter*, *Lawsonia*, *Blautia*, and *Lactonifactor* have been reported in high-fat diet-fed mice, which were countered by metformin treatment (Shin *et al.*, 2014). The gut microbiome of high-fat diet-fed mice is dominated by *Firmicutes*, in contrast to mice fed the regular diet in which *Bacteroidetes* prevail. Lee and colleagues have demonstrated that metformin treatment increases *Bacteroidetes* abundance to a level similar to that of the control mice and changes the abundances of *Bacteroidaceae*, *Verrucomicrobiaceae*, *Clostridiales family XIII, incertae sedis*, and *Akkermansia muciniphila* (Lee, 2014). Metformin exerts changes in the gut microbiome of healthy mice also by increasing the abundances of *Rikenellaceae*, *Ruminococcaceae*, *Verrucomicrobiaceae*, *Alistipes spp.*, *Akkermansia spp.*, and *Clostridium spp.* (Lee, 2014). A study on the modulation of the gut microbiome in a high-fat diet-fed male rat model in response to metformin treatment has shown increased abundances of *Blautia*, *Bacteroides*, *Butyricoccus*, *Phascolarctobacterium*, *Parasutterella*, *Akkermansia*, *Prevotella*, *Lactobacillus*, and *Allobaculum* (Zhang *et al.*, 2015). A review on the effects of metformin on the gut microbiome in the context of obesity and T2D has summarized the differences in the results obtained in both human and animal studies (Zhang, 2020), showing that certain unclarities still remain, for example, the directions of changes in response to metformin treatment in the abundances of *Prevotella* and *Lactobacillus* differ in various studies. The effect of metformin treatment has mainly been studied in cecal and fecal samples only, providing only limited information about the various aspects of the interaction

between the gut microbiome and metformin. Short-term effects of metformin have been shown in different parts of the small intestine in the luminal layer in a recent study (Bravard *et al.*, 2021).

1.9 Strategies to modulate the gut microbiome

Several strategies to prevent or treat dysbiosis exist. The range of interventions that may modulate the microbiome can be divided into two categories – lifestyle modifications (diet change, caloric restriction, physical activity) and clinical interventions (fecal microbiota transfer, antibiotics, prebiotics, probiotics, postbiotics, non-antibiotic drugs) (Quigley and Gajula, 2020). The strategies from the clinical category are shown to mechanistically change the microbiota composition, modulate innate immunity, improve the intestine's barrier function, prevent pathogen colonization, and display selective cytotoxicity against tumor cells, thus having great potential for successful treatment of microbiome-related diseases (Fong, Li and Yu, 2020). Nevertheless, it should be noted that these approaches are accompanied by risks and controversies that can potentially introduce clinical complications. Most of the studies reported describe efforts to modulate the gut microbiome by using live microorganisms ranging from single strains to non-defined fecal transplants; however, the use of beneficial bacteria fails to live up to its expectations. The main limitations are colonization resistance, inter-individual variation, and the risk of transferring pathogens, suggesting that the mechanisms underlying the modulation of the microbiome are not fully explained and involve host-related factors (Suez and Elinav, 2017).

Probiotics are living microorganisms that benefit the health of the host when administered in adequate amounts (Fong, Li and Yu, 2020). *Lactobacillus*, *Bacillus*, *Bifidobacterium*, *Streptococcus*, and *Enterococcus* are the most prevalent bacterial genera applied as probiotics (Wang and Li, 2022). The main functions of probiotics include immunomodulation, colonization resistance, and enhancing the intestinal barrier (Wang and Li, 2022).

Prebiotics are defined as nondigestible food ingredients that selectively stimulate the growth and/or activity of specific bacteria in the gut and improve host health (Davani-Davari *et al.*, 2019). Prebiotics act in the gut, mainly by stimulating the growth of probiotics; being selectively fermented by probiotics; communicating with pathogens and preventing their colonization; and being absorbed into the intestine and exerting anti-inflammatory effects, although the benefits of prebiotics may not be universal and are dependent from an individual (Fong, Li and Yu, 2020).

Postbiotics, also known as metabiotics, pharmabiotics, or heat-killed probiotics, describe molecules exhausted in the dysbiotic gut, which when supplemented in whole or precursor form,

help to resume the balance in the gut (Chaluvadi, Hotchkiss and Yam, 2016; Diez-Gutiérrez *et al.*, 2020). Postbiotics include all the bioactive functional compounds that can be used or produced by the microbial community to promote health. Vitamin B12, vitamin K, folate, SCFAs, indole produced from amino acids, gamma-aminobutyric acid, and enzymes are typical examples of postbiotics (Diez-Gutiérrez *et al.*, 2020; Thorakkattu *et al.*, 2022). The main functions of postbiotics are exerting selective cytotoxicity against tumor cells and protecting the intestinal epithelium by suppressing apoptosis of epithelial cells and promoting IgA output (Fong, Li and Yu, 2020).

An antibiotic has been defined as a compound made by a microbe to destroy other microbes (Hutchings, Truman and Wilkinson, 2020). Antibiotics are widely used due to their action against pathogenic microbes. However, the use of antibiotics has a negative effect on gut microbiota, including reduced biodiversity of bacteria, alterations in metabolic functions, and selection of antibiotic-resistant bacteria, causing a rapid rise of antimicrobial resistance – a global threat to health (Hutchings, Truman and Wilkinson, 2020; Wang and Li, 2022).

Fecal microbiota transplantation (FMT) is a therapeutic method that is defined as transferring fecal material from a healthy donor into the gastrointestinal tract of the patient (Bibbo *et al.*, 2020). FMT is an approved method for treating *Clostridium difficile* infection in clinical practice (Bibbo *et al.*, 2020). However, it has shown potential for use in managing obesity, inflammatory bowel diseases, diabetes, non-alcoholic fatty liver disease, and cardiovascular diseases (Wang and Li, 2022).

1.10 MiRNAs

MicroRNAs (miRNAs) are small (average 22 nucleotides in length) non-coding RNAs that are involved in regulating gene expression at the post-transcriptional level primarily by interacting with the 3` UTR of target mRNAs; however, binding of miRNAs to 5` UTR, coding regions or within promoter regions of target mRNAs has been described as well (O'Brien *et al.*, 2018). Mature miRNAs repress gene expression through two main mechanisms: binding to the target mRNA molecule, thereby repressing its translation, or promoting target mRNA degradation (Sarshar *et al.*, 2020). Nevertheless, under specific conditions, miRNAs can exhibit gene activation (O'Brien *et al.*, 2018).

MiRNA biogenesis consists of a series of sequential processes to create mature miRNAs (see Figure 4). Half of the described miRNAs are intragenic and processed mainly from introns and

relatively few exons of protein-coding genes, while the others are intergenic and are transcribed independently of the host gene and have their own promoters (O'Brien *et al.*, 2018). Many of the known miRNAs are encoded as gene clusters transcribed on a single polycistronic transcript (Tanzer and Stadler, 2004). These may share similar seed regions, hence forming a miRNA family (O'Brien *et al.*, 2018).

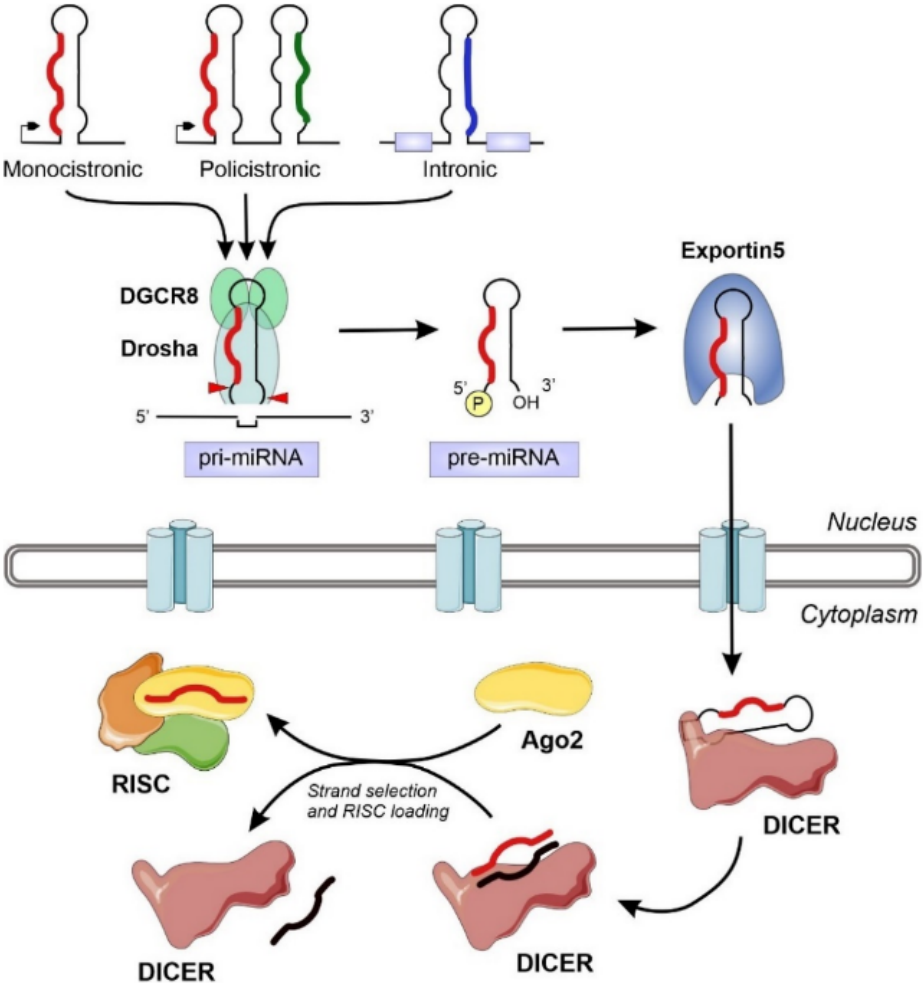

Figure 4. MiRNA biogenesis. Adapted from (Leitao and Enguita, 2022).

Using the canonical biogenesis pathway, the dominant pathway for miRNA processing, the miRNA gene is initially transcribed by RNA polymerase II into a primary precursor, pri-miRNA, which is approximately 200 nucleotides long (Bartel, 2018). The pri-miRNA is then cleaved by the nuclear microprocessor complex consisting of Rnase III group endonuclease Drosha and the DiGeorge critical region 8 (DGCR8) protein (Winter *et al.*, 2009). This results in the formation of a 60-70 nucleotide hairpin precursor miRNA (pre-miRNA) which is exported from the nucleus to the cytoplasm across the nuclear pore by a carrier protein Exportin-5 (Leitao and Enguita, 2022).

The pre-miRNA is further processed by the Rnase III endonuclease Dicer to approximately 22-nucleotide-long, mature, leading- and lagging-strand duplex miRNA (Sarshar *et al.*, 2020). Upon formation, the duplex miRNA is loaded into an Argonaute 2 (Ago2) protein, which then relaxes to its basic conformation thus expelling one of the strands of the duplex miRNA for a following degradation (Bartel, 2018). The mature miRNA strand selected by Ago2 associates with several RNA-binding proteins, including Argonaute, to form a microribonuclear protein (miRNP) complex called the RNA-induced gene silencing complex (RISC), which then binds to the target mRNA to perform one of its gene regulation functions (Sarshar *et al.*, 2020).

The target specificity of RISC is determined by the sequence complementarity between the miRNA and the respective sequence of the target mRNA, called miRNA response element (MRE) (O'Brien *et al.*, 2018). Perfect base pairing of the miRNA with its MRE, frequently found in plants but rarely in animals, causes endonucleolytic cleavage of the target mRNA by Ago2 (Chipman and Pasquinelli, 2020). In animals, most miRNAs exhibit partial complementarity with the MRE and therefore prevent Ago2-mediated cleavage. Instead the expression of the target is regulated through translational repression and destabilization of the mRNA (Chipman and Pasquinelli, 2020). The seed region (nucleotides 2-7 of the 5` side) of the miRNA is generally considered the minimal sequence to bind the MRE principally located at the 3` UTR of the target mRNA (Nair *et al.*, 2020). In addition, the 3` end of the miRNA has been demonstrated to provide additional pairing thus increasing the specificity and the stability of the interaction (O'Brien *et al.*, 2018). MiRNAs can bind to MREs of several different target genes; in turn, a single mRNA can have one or more distinct MRE sites and, therefore, can be targeted by multiple miRNAs (Nair *et al.*, 2020).

Similar to mRNAs, some miRNAs are distributed in specific tissues; for example, the miR-122 family is found in the liver, whereas the miR-124 family is expressed in neural tissues (Weber *et al.*, 2016). The stability of miRNAs is robust compared to mRNA (Jung *et al.*, 2010). Extracellular miRNAs can be secreted and transported both via extracellular vesicles, such as microvesicles, exosomes and in a form associated with high-density lipoproteins or Argonaute protein, thus contributing to extracellular miRNA stability (Liu *et al.*, 2016). Circulating miRNAs can be detected in body fluids, such as blood, saliva, cerebrospinal fluid, breast milk, and urine, making them promising biomarkers (Zhou *et al.*, 2011; Wang *et al.*, 2012; Yoshizawa and Wong, 2013; Kopkova *et al.*, 2018; Shang *et al.*, 2019). Fecal miRNAs have mostly been studied in the context of colorectal cancer (Yuan, M. Burns, *et al.*, 2018) and inflammatory bowel diseases (Ji *et al.*, 2018; Wohnhaas *et al.*, 2020). Despite their known functional role in the intestine, the

involvement of fecal miRNAs in host-microbiome interaction is not yet fully understood (Sarshar *et al.*, 2020).

1.11 MiRNA interaction with gut microbiome

The latest studies have provided evidence that miRNAs not only function in the host cells but also can be transmitted from one species to another, ensuring communication or inducing signal interference between different species or even kingdoms. Thereby the miRNAs endogenous to one species are able to target the biological processes of other species (Liang *et al.*, 2013). Strong evidence exists that a single miRNA can target many mRNAs, and a single mRNA can be targeted by many miRNAs (Taganov *et al.*, 2006). Synthesis of miRNAs is rapid and requires little energy input, being economical in cellular terms; thereby, it is possible for host miRNAs to modulate a large number of members of the gut microbiome (Choi *et al.*, 2017).

Until now, three important interactions between the host and gut microbiome involving miRNAs have been clarified (Figure 5). Firstly, miRNAs regulate host gene expression; secondly, the gut microbiome influences host miRNA expression; and lastly, the host influences the gut microbiome through the release of miRNAs (Williams *et al.*, 2017a).

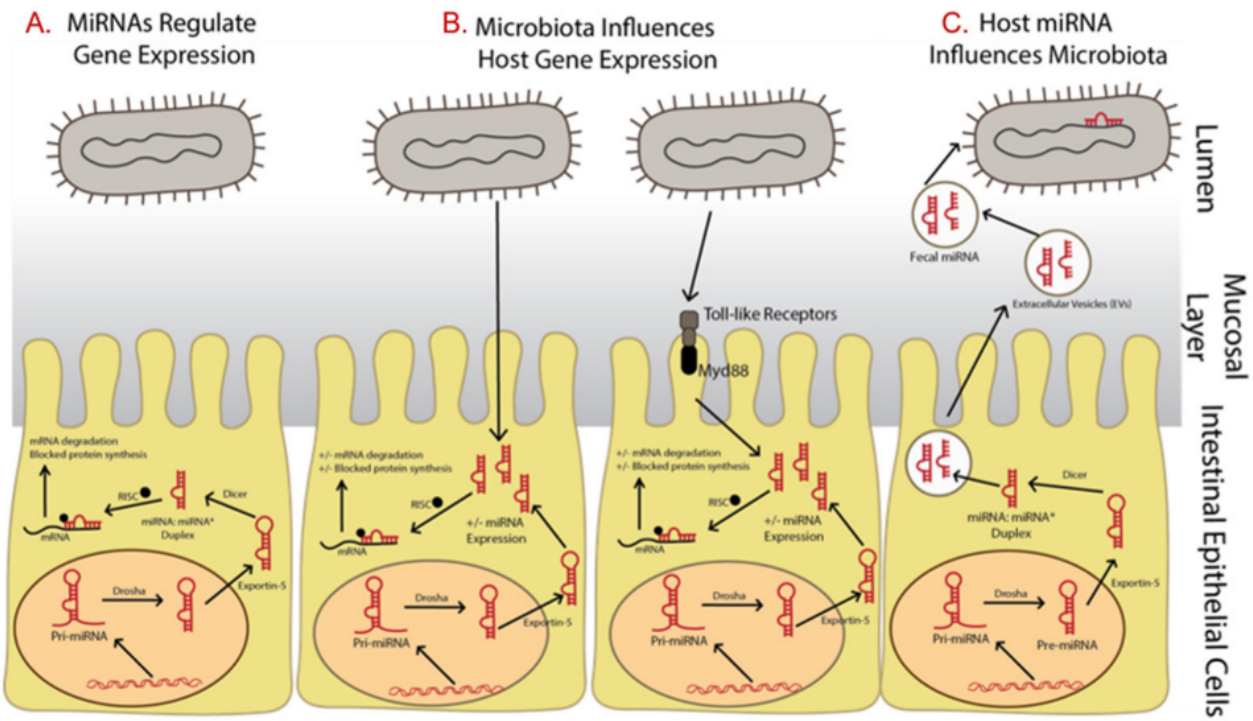

Figure 5. Known interactions between miRNAs and microbiota and the effect on gene expression regulation. (A) MiRNAs regulate gene expression in response to microbiota. (B) Microbiota regulates host

miRNA expression. (C) Host miRNA influences its microbiome through fecal miRNAs, which are introduced in bacteria. Adapted from (Williams *et al.*, 2017b).

Liu and colleagues have shown that the host itself may regulate its microbiome and that the overall abundance of fecal miRNAs is higher in germ-free mice compared to the SPF-colonized mice (Liu *et al.*, 2016). Furthermore, they observed that Dicer1 knockout mice exhibit uncontrolled growth of the gut microbiome, but after fecal miRNA transplantation, homeostasis was restored, suggesting that extracellular miRNAs are involved in controlling bacterial growth. It was observed that intestinal epithelial cells (IECs) are major sources of extracellular fecal miRNAs due to their ability to secrete exosome-like vesicles. Host miRNAs in feces specifically target bacterial genes and thus regulate the gut microbiota (Liu *et al.*, 2016). Though the relationship between miRNA expression and microbiome-associated diseases, as well as the association of gut microbiome dysbiosis to these diseases, has been studied to some extent, studies regarding the communication between the microbiome and miRNAs in the context of gut diseases are not fully investigated (Williams *et al.*, 2017a). Further research regarding host miRNA effects on the microbiome in relation to microbiome-associated diseases such as type 2 diabetes is needed.

1.12 Model systems for gut microbiome research

Model systems are non-human species or non-living systems that replicate essential aspects of the living body and can be manipulated and experimented with to study specific biological phenomena, diseases, molecular pathways, and behaviors. Modeling approaches are often classified into *in vitro*, *ex vivo*, *in vivo*, and *in silico* models. Studying the interaction between a host and its gut microbiome requires complex model systems, as both the host and the microbiome community must be integrated into the model (Elzinga *et al.*, 2019).

In vitro models are characterized by their controlled chemical and physical environment, scalability, high throughput, and being relatively cost-effective (Xu *et al.*, 2021). Limitations of these systems include lack of the representation of the actual anatomy and physiology of the intestine and the lack of interindividual differences (Ashammakhi *et al.*, 2020; Xu *et al.*, 2021). Furthermore, conventional, two-dimensional *in vitro* model systems are not suitable for long-term co-cultivation of mammalian cells and microbiota due to the overgrowth of aerobic bacteria in one day, and cultivation of obligate anaerobic bacteria, which dominate the gastrointestinal tract is not possible (Steinway *et al.*, 2020). Current microbiome culture models do not maintain the functional activities and compositional profiles of the gut microbiome when compared to the

inoculum (L. Li *et al.*, 2019). However, efforts have been invested to increase the maintenance of functional activities of the gut microbiome *in vitro* including the MiPro model which shows improved conservation of the gut microbiome functional profile (L. Li *et al.*, 2019) but still lacks the host component.

Other, more complex *in vitro* models from the host's perspective that can be used for microbiome research include adult stem cell-derived organoids, pluripotent stem cell-derived organoids, and gut-on-a-chip system, among others (Puschhof, Pleguezuelos-Manzano and Clevers, 2021). From the microbe perspective, model systems include continuous culture system (CCS), gastrointestinal tract systems TIM1 (mimics the stomach and the small intestine) and TIM2 (simulates the large intestine), which incorporate peristaltic mixing and absorption of water and fermentation products; simulator of the human intestinal microbial ecosystem (SHIME), which is comprised of five sequentially connected reactors containing luminal microbes; and M-SHIME which in addition includes the mucus layer (Fritz *et al.*, 2013). These model systems do not incorporate host cells or have limited abilities to model the interaction with the host. Studies on the interaction between the host and the microbiome are allowed by several co-cultivation strategies, including 3D microinjection, organoid shearing, organ-on-a-chip co-culture, and explant co-culture (Puschhof, Pleguezuelos-Manzano and Clevers, 2021). Explant cultures are considered *ex vivo* models, as these are taken directly from the living organism and thus represent the mucosal architecture and include a mucus layer. The main limitation of *ex vivo* models is their short lifespan (Xu *et al.*, 2021).

Although more and more complicated alternative methods are being developed including microfluidics-based model systems – gut on a chip (Kim *et al.*, 2016) or the human-microbial cross talk (HuMiX) device (Shah *et al.*, 2016), they currently cannot yet model processes at the level of the whole organism, moreover, these methods use cells originating from intestinal carcinomas Caco-2 or HT-29, which may have altered biological features including miRNA expression profile compared to the T2D background.

In silico models can be used for modeling and predicting various parameters of separate organisms living in a particular host. However, their construction strongly depends on existing experimental data collected in *in vitro* and *in vivo* studies (Fritz *et al.*, 2013). Furthermore, the construction of large-scale, high-quality networks requires vast computational and manual analysis resources, and identified results need experimental validation (Fritz *et al.*, 2013). Other limitations include potential translatability issues if the model is developed based on animal experimental data

and the limited knowledge of the non-bacterial gut microbiome compartment (Molina Ortiz *et al.*, 2021). Nevertheless, *in silico* methods are developing rapidly, and their potential has yet to be fully exploited.

Even though a great effort is invested in developing and advancing various *in vitro* and *in silico* alternative methods to reduce or avoid the use of animals in research according to the widely endorsed 3Rs principle (replacement, reduction, and refinement of animal studies), the use of animal models (*in vivo* model systems) remains the most appropriate approach in many biomedical studies. The use of living organisms helps to investigate and understand the complex etiology and interplay of multiple systems in the context of different diseases or biological conditions (Kottaisamy *et al.*, 2021).

1.13 Type 2 diabetes animal models

Type 2 diabetes is a complex disease affecting various organ systems and processes in the body; therefore, a complex model system is necessary for its research. Rodent models, especially mice, are widely used for gut microbiome studies. Results obtained from such studies can be translated to humans due to a 90% similarity between gut microbiome compositions in mice and humans (Zhang, Franklin and Ericsson, 2021). The main advantages of mice as an *in vivo* model system include a relatively short life span and generation interval, relatively low cost, well-characterized gut microbiome, and miRNAs, and can be maintained and bred under controlled conditions. Although, pigs have been characterized as the closest models to humans regarding both the gastrointestinal tract and gut microbiome composition (Crespo-Piazuelo *et al.*, 2019), these model organisms are not feasible for most studies.

Animal models for studying T2D can be classified into the following categories based on T2D induction methods – spontaneous or congenital diabetes; diet-induced; chemically-induced; surgical; and transgenic models (Srinivasan and Ramarao, 2007). The main manifestations of type 2 diabetes are insulin resistance and beta cell failure; therefore, most T2D animal models involve one or a combination of these characteristics. Obesity is another feature strongly associated with T2D; accordingly, most animal models are obese (King, 2012). Obesity can be caused by spontaneous mutations or genetic manipulation, or it can be induced by a high-fat diet.

Monogenic models of obesity include $Lep^{ob/ob}$ mouse (developed in C57BL/6J strain), which is deficient in leptin, a hormone necessary for the appetite regulation and sensation of satiety, $Lepr^{db/db}$ mouse (developed in C57BL/KsJ strain) and Zucker diabetic fatty ($Lepr^{fa/fa}$) rat, both

deficient in the leptin receptor. Polygenic models include Kuo Kondo (KK) mouse (severe hyperinsulinemia and insulin resistance, alterations in pancreatic islets), Otsuka Long Evans Tokushima Fatty (OLETF) rat (mild obesity and late onset hyperglycemia, inherited in males), New Zealand Obese (NZO) mouse (hyperleptinemia, hyperinsulinemia, impaired glucose tolerance, insulin resistance), TallyHo/Jng mouse (hyperglycemia, hyperinsulinemia, increased adiposity), and NoncNZO10/LtJ mouse (liver and skeletal muscle insulin resistance, hyperglycemia) (Srinivasan and Ramarao, 2007; King, 2012).

High-fat diet induces obesity, hyperinsulinemia, insulin resistance, and impaired glucose homeostasis due to ineffective compensation by pancreatic islets (Srinivasan and Ramarao, 2007). Diet-induced models better correspond to the disease in humans due to their environmental origin rather than the genetic background. C57BL/6 mice are used for this model most widely, with a first report of this model in 1988 by Surwit and colleagues (Surwit *et al.*, 1988). The severity of the disease depends on the duration and composition of the diet. Most popular diet types used for this model are either 45% or 60% kcal from fat. Feeding duration from 4 to 20 weeks old has been proposed as a standard protocol, furthermore, animals of both sexes should be used (Heydemann, 2016).

Chemically-induced T2D models include alloxan and streptozotocin (STZ)-induced diabetes. Alloxan, a uric acid derivative, selectively destroys the pancreatic beta cells causing insulin deficiency, hyperglycemia, glucosuria, hyperlipidemia. It can induce diabetes manifestations in rodents and non-rodent animals and is best used in rabbits due to STZ ineffectiveness in the species (Srinivasan and Ramarao, 2007). STZ is an antimicrobial agent which induces diabetes by its cytotoxic effect on pancreatic beta cells, thus causing hypoinsulinemia and hyperglycemia (Damasceno *et al.*, 2014). STZ is the most popular chemical for the induction of T2D in animals and can be used in many different species including rats, mice, and guinea pigs (Kottaisamy *et al.*, 2021).

Surgical induction of T2D manifestations comprises partial pancreatectomy, as 70-90 percent dissection of pancreas. This approach has been used in dogs, pigs, rabbits, and rats. It causes moderate hyperglycemia (Srinivasan and Ramarao, 2007).

Transgenic and knockout T2D models are used for investigating gene regulation and development, pathogenic mechanisms, and therapeutic targets (Srinivasan and Ramarao, 2007). In the context of T2D many genes and their effects can be altered forming various phenotypic

manifestations ranging from mild to severe hyperglycemia, insulin resistance, hyperinsulinemia (Srinivasan and Ramarao, 2007).

1.14 Translatability of mouse experiments to humans

Remarkable similarity exists between mice and humans in regard to physiology and anatomical structures (Figure 6), which is one of the main reasons why mouse models are widely used in biomedical research (Nguyen *et al.*, 2015). In particular, the gastrointestinal tracts of both species are formed of organs that are anatomically similar.

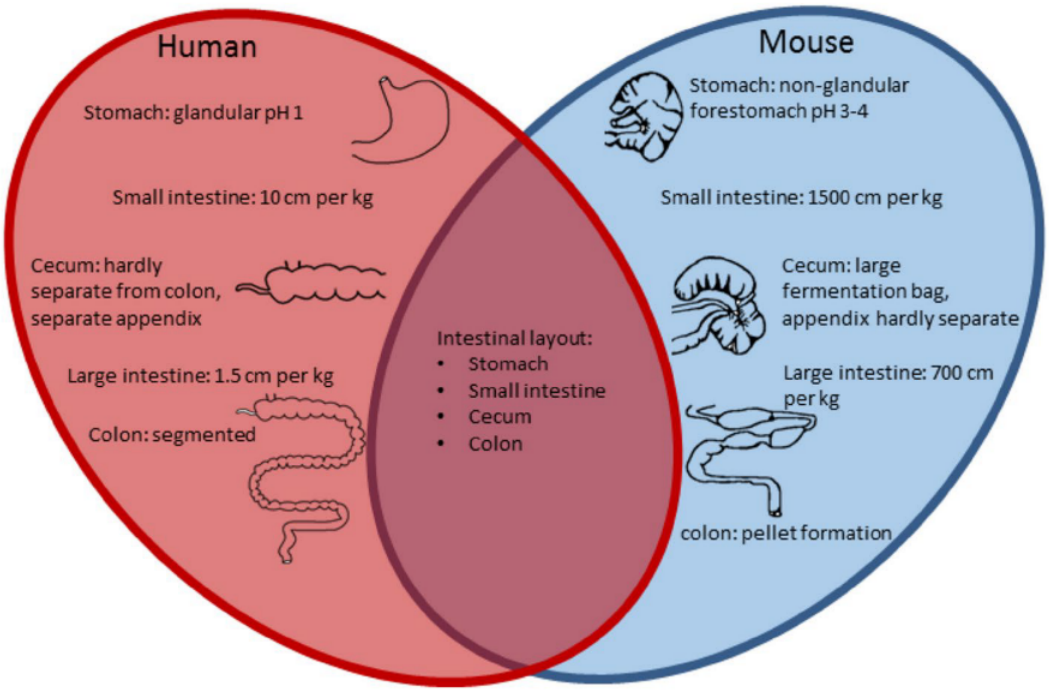

Figure 6. Comparison of human and mouse intestinal tracts. Adapted from (Hugenholtz and Vos, 2018).

Although, it should be noted that significant differences also exist, which might be influenced by their different diets, feeding patterns, body sizes, and metabolic requirements (Nguyen *et al.*, 2015). Mice possess a colon and cecum with a proportionally larger surface area, as well as taller intestinal villi, which could potentially facilitate greater nutrient absorption and essential element synthesis (Park and Im, 2020). Unlike mice, humans have an appendix that serves as a reservoir for microorganisms, while a cecal patch is considered the equivalent of the human appendix in mice (Kooij *et al.*, 2016). Substantial differences in gut microbiome composition have been described between these anatomical structures. The predominant bacterial taxa found in the human appendix are *Firmicutes*, *Proteobacteria*, *Bacteroidetes*, and

Fusobacteria, whereas cecal patches of mice are primarily composed of *Bacteroidetes*, *Firmicutes*, *Actinobacteria*, and *Proteobacteria* (Park and Im, 2020).

Overall, the human and murine microbiomes are approximately 90% similar in terms of phyla and genera content (Krych *et al.*, 2013). The main differences are manifested in the abundance of microorganisms and their proportions in the microbiome composition, with humans having a higher ratio of *Firmicutes* to *Bacteroidetes* – dominant phyla in the gut microbiome of both mice and humans (Park and Im, 2020). At the genera level, *Bacteroides* is the most abundant genus in the human gut microbiome, followed by members of *Ruminococcaceae* and *Clostridiales*, whereas in mice, *S24-7* family members (representatives of *Bacteroidetes* phylum) prevail, followed by *Clostridiales* representatives (Nagpal *et al.*, 2018). The top 15 genera in fecal samples of both mice and humans are illustrated in Figure 7. In turn, *Faecalibacterium*, *Mitsuokellla*, *Megasphera*, *Dialister*, *Asteroleplasma*, *Succinivibrio*, *Sutterella*, *Paraprevotella*, and *Phascolarctobacterium* have been described as found mainly in humans, while *Mucispirillum* generally is rodent-specific (Park and Im, 2020).

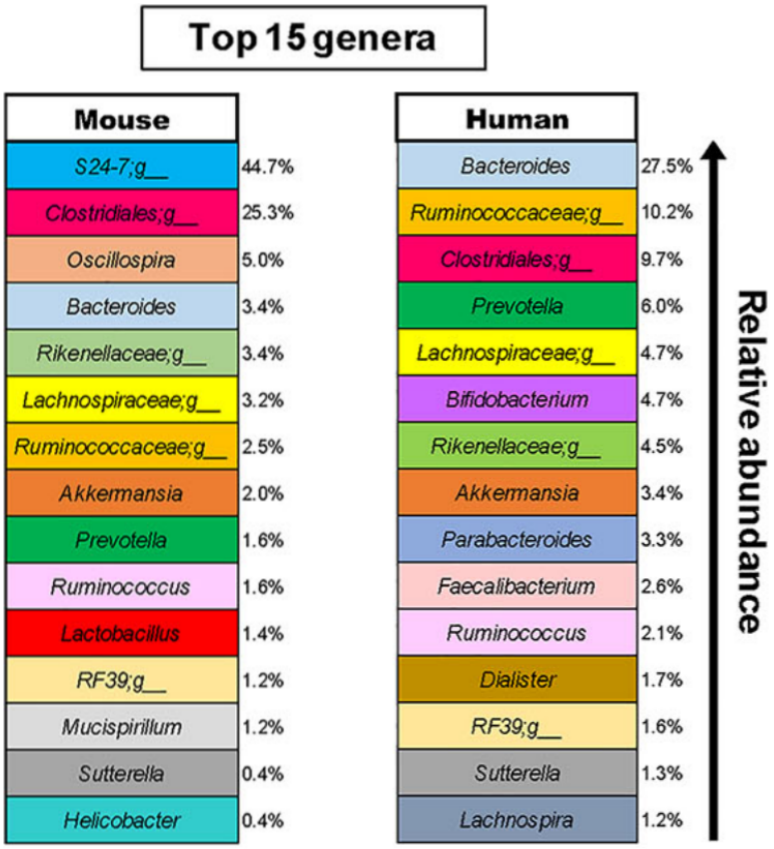

Figure 7. Top 15 genera and their mean relative abundance in fecal samples of mice and humans. Adapted from (Nagpal *et al.*, 2018) .

Despite the limitations, mouse models remain a valuable tool for studying disease mechanisms and testing potential therapies. Animal experimentation allows for manipulations and interventions that are essential for gut microbiome research in the context of metabolic diseases, including dietary interventions, providing the treatment to be studied, and collecting tissue-associated microbiome samples that would not be possible with human subjects. Furthermore, these advantages are accompanied by a controlled environment, known and shared genetic background, and opportunities for the complex experimental setup. Specific pathogen-free status characteristic for many animal facilities allows increased uniformity and reproducibility of the results (Aguanno *et al.*, 2022). Although studies in mice are not fully translatable to humans, currently available model systems do not provide a better alternative (Hugenholtz and Vos, 2018).

1.15 Colorectal adenomas

Neoplastic colorectal polyps, also known as adenomas, are potential predecessors of colorectal adenocarcinoma, the most common form of colorectal cancer (constitutes more than 95% of colorectal cancer cases) (Shussman and Wexner, 2014; Thrumurthy *et al.*, 2016). They are typically small, benign tumors that develop on the inner lining of the colon or rectum, and can often be removed during a routine colonoscopy. Histologically colorectal polyps are classified into two major classes – conventional adenomas and serrated lesions (Rex *et al.*, 2017). Conventional adenomas are further categorized based on the dysplasia grade (high or low) and villousity (tubular, tubulovillous, and villous). Subgroups of serrated lesions include hyperplastic polyps that are not precancerous, sessile serrated polyps (with or without cytologic dysplasia), and traditional serrated adenomas (Rex *et al.*, 2017).

The prevalence of colorectal adenomas varies between 30-50% (Øines *et al.*, 2017). Nonmodifiable risk factors associated with the development of colorectal adenomas include age, sex, family history of colorectal cancer, and hereditary polyposis syndromes. Lifestyle factors, such as smoking, red-meat consumption, alcohol use, physical inactivity among others, may also increase the risk of developing colorectal adenomas (Sninsky *et al.*, 2022). It has been estimated that 5% of adenomas develop into carcinomas, and the adenoma-carcinoma sequence is the widely accepted process by which colorectal cancer arises (J. Li *et al.*, 2019; Nguyen, Goel and Chung, 2020).

Colorectal cancer is the third most common malignancy globally, with 1.93 million new cases and 0.94 million deaths in 2020, and it is projected to progress to over 3 million new cases in

2040 (Xi and Xu, 2021). Due to the heavy financial burden and global healthcare issues, prevention strategies are critical to reducing colorectal cancer risk. The most common prevention and treatment strategy for colorectal adenomas is regular colonoscopy screening to detect and remove any precancerous growths (Herszényi, 2019). Other prevention strategies include regular physical activity, a healthy diet, intake of vitamins, limiting alcohol and red meat consumption, smoking cessation, and aspirin use as a chemoprevention agent among individuals with elevated risk (Hou, Huo and Dignam, 2013).

1.16 MiRNAs and gut microbiome in patients with colorectal adenomas

Individual miRNAs can function as oncogenes or tumor suppressors, though miRNA expression has been shown to be downregulated in tumor cells, suggesting impaired miRNA biogenesis in cancer (Lin and Gregory, 2015). An increasing body of evidence indicates that miRNAs could serve as potential biomarkers for early-stage colorectal cancer detection due to their resistance to degradation and ability to provide insight into the presence of the disease based on their expression levels in colorectal adenomas, blood samples, and feces (Yi *et al.*, 2016). Changes in miRNA profile during the progression from normal colorectal tissue to adenoma and to colorectal cancer have been investigated in several studies (Yin *et al.*, 2016; J. Li *et al.*, 2019). Although existing data on altered miRNA profiles between different tissue states are inconsistent, miRNAs that are upregulated in both colorectal adenoma and carcinoma stages include miR-18a, miR-18b, miR-31, miR-142-5p, miR-212, miR-21, miR-92a, miR-135b, miR-486, miR-1290, miR-15a, miR-17, miR-20a, and miR-31, while miR-320 family, miR-145, miR-451, miR-638, miR-137, miR-30, miR-24, miR-423-5p, and miR-148a are downregulated (Liu and Li, 2019).

There is mounting evidence indicating that the human gut microbiota and microbiota-associated metabolites play a significant role in the early stages of colorectal adenocarcinoma and colorectal adenoma development by disrupting various intestinal functions and contributing to chronic inflammation (Coker *et al.*, 2022). Changes in the microbiome in adenoma and carcinoma stages have been evaluated in fecal samples (Feng *et al.*, 2015) and colonoscopy aspirates (Saito *et al.*, 2019). Colorectal adenocarcinoma has been shown to be associated with an increase of *Fusobacterium nucleatum*, *Peptostreptococcus anaerobius* and *Bacteroides fragilis* in fecal samples (Yuan *et al.*, 2021). In another study, an increase in *Bacteroides dorei* and *Bacteroides massiliensis* has been shown in patients with advanced adenoma compared to healthy subjects, and an increase in *Bacteroides massiliensis*, *Bacteroides ovatus*, *Bacteroides vulgatus*, and *Escherichia coli* in carcinoma patients compared to patients with advanced adenoma (Feng *et al.*, 2015). In

turn, *Ruminococcus*, *Bifidobacterium*, and *Streptococcus* species have been reported to be decreased in patients with adenoma or colorectal cancer (F. Chen *et al.*, 2022).

The interaction between host miRNAs and the gut microbiome have been investigated in the background of colorectal cancer. (Yuan, M. B. Burns, *et al.*, 2018; Tomkovich *et al.*, 2020). Several microbial taxa that are correlated with certain host miRNAs have been identified in tumor tissue samples including *Parabacteroides* and *Blautia* which were positively correlated with miR-129-5p and miR-139-5p, respectively, while the relative abundances of *Rikenellaceae* and *Bacteroidales S24-7* were negatively correlated with the expression of miR-139-3p and miR-143-3p, respectively (Yuan, M. B. Burns, *et al.*, 2018). In mice inoculated with the biofilm-positive human colorectal tumor-derived microbiome showed positive correlation between *Enterococcus* and mmu-miR-103-3p, among others, in turn the abundance of *Roseburia* was negatively correlated with mmu-let-7c-5p (Tomkovich *et al.*, 2020). The inconsistent results between studies necessitate further investigation into the interaction between the host miRNAs and the gut microbiome; furthermore, studies in the context of the colorectal adenoma stage still need to be performed.

2 MATERIALS AND METHODS

2.1 Animal experiment for studies I-III

2.1.1 Ethical approval

Animal procedures were reviewed and approved by the National animal welfare and ethics committee, State Food and Veterinary Service, Riga, Latvia.

2.1.2 Design of animal experiment

This study was designed as a randomized block experiment comprised of three blocks with a three-way factorial treatment arrangement where factors of interest are T2D status induced by high-fat diet (HFD) or control diet (CD) feeding, sex, and metformin therapy status, forming eight different treatment groups – HFD_M_Met-, HFD_F_Met-, HFD_M_Met+, HFD_F_Met+, CD_M_Met-, CD_F_Met-, CD_M_Met+, and CD_F_Met+ (Figure 8). Study's sample size was determined by the Resource Equation method, appropriate for complex designs (Festing *et al.*, 2016). In each of the eight treatment groups, we included three experimental units, 24 in total. As the experimental unit was a cage with three animals – the total number of animals involved in the experiment was 72.

After the adaptation week, all experimental units of the same block were randomly assigned to HFD- or CD-fed groups so that each of the treatment groups would consist of animals with similar body weight, and experimental units of both sexes would be represented in the same number in both types of treatment groups. After the induction of T2D manifestations, experimental units were randomly assigned to receive or not to receive metformin treatment providing that the number of experimental units in each of the groups is equal. During all the procedures, treatments, and measurements, as well as sample collection were performed randomly within the same block. Work with each of the blocks was performed on separate days of the week, keeping the interval between interventions in the same block constant.

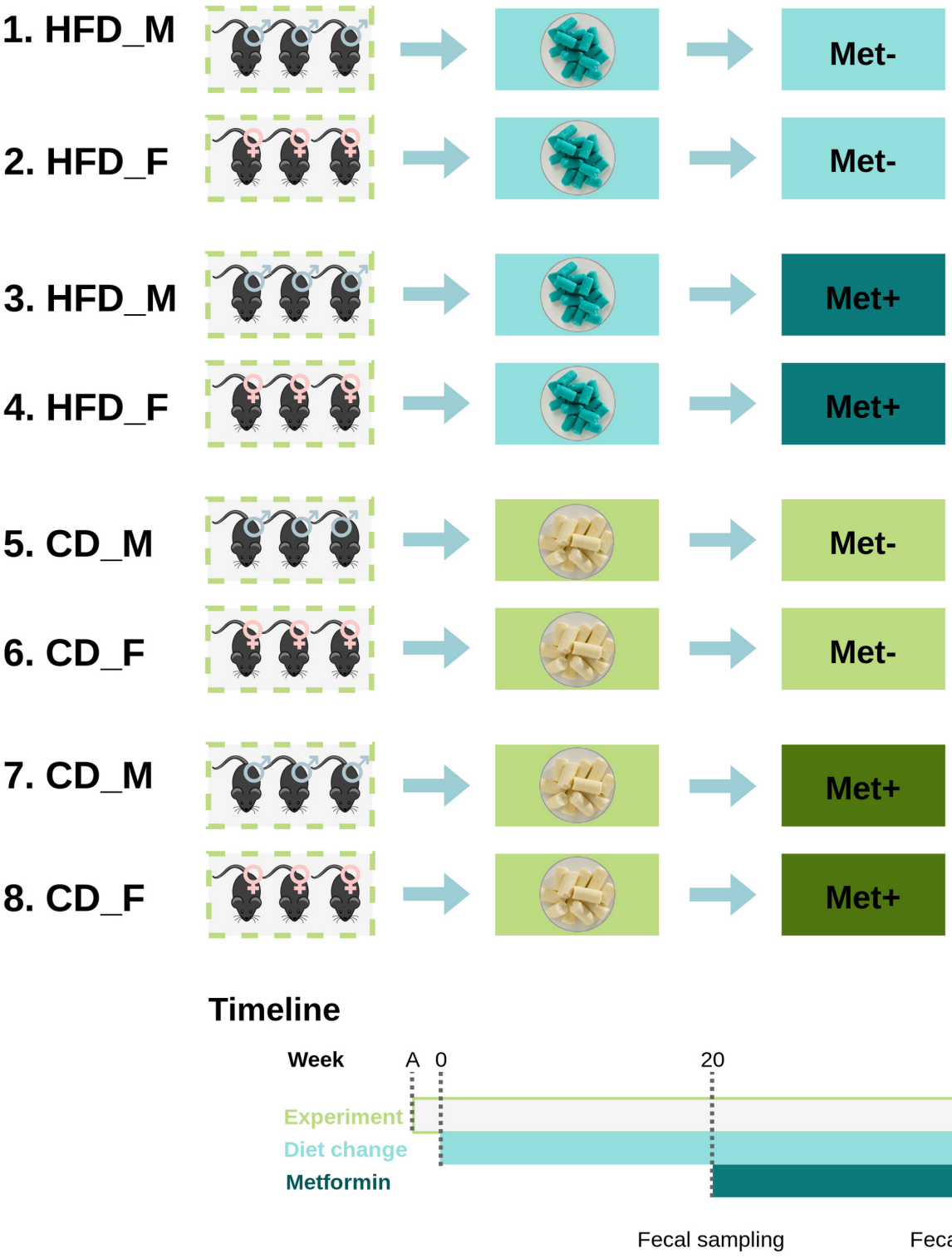

Figure 8. Experimental design and timeline. After the adaptation period (A), the total duration of the study was 30 weeks, including 20 weeks of the induction of type 2 diabetes manifestations followed by 10 weeks-long metformin treatment. Abbreviations: HFD – high-fat diet; CD – control diet; M – male; F – female; Met+ – receiving metformin treatment for 10 weeks; Met- – not receiving metformin treatment.

2.1.3 Experimental animals

Age-matched 4-5-week-old male and female C57BL/6N mice with specific pathogen-free (SPF) status were obtained from the University of Tartu Laboratory Animal Centre. All the animals were housed in SPF conditions, 23 ± 2 ºC, 55% humidity. The light cycle was 12:12 hours, with a light period from 7:00 am to 7:00 pm. All the procedures were performed during the first half of the day in a specially designated procedure room.

Animals were housed in individually ventilated cages (Techniplast) up to three same-sex animals per cage on aspen bedding mixed with ALPHA-dri. All the animals had free access to drinking water. Animals were fed HFD or CD *ad libitum*. All the cages were enriched with cardboard tunnels, plastic shelters, wooden sticks, and nesting material. For the whole duration of the experiment, animals were observed once a day; if we observed any type of suffering that could not be alleviated, the suffering animal was euthanized by cervical dislocation. During the study humane endpoint was implemented for 13 animals mainly due to male fighting wounds. Thus 59 animals completed the study, however, at the end of the experiment each of the experimental units remained represented by at least one animal.

2.1.4 Experimental procedures

After a one to two-weeks long adaptation period during which animals were fed regular chow diet *ad libitum* and received regular drinking water a diet change was initiated. The age of animals at this point was approximately six weeks. All the cages from each of the blocks were randomly assigned to a high-fat diet-fed group or control diet-fed group. Animals were provided with a rodent diet containing 60 kcal% fat (D12492, Research Diets) or rodent diet with 10 kcal% fat (D12450J, Research Diets) *ad libitum*. Both types of diets were sterilized by irradiation. Body weight and food intake (per cage) were measured once a week; water intake (per cage) – twice a week.

The cages were changed on a weekly basis. Upon opening the cage, each mouse was immediately transferred to a clean, separate box in which animals were allowed to defecate voluntarily. Each of the animals was weighted. Feces were collected in sterile tubes in three aliquots. Bottles with drinking water were changed two times a week.

Four weeks after starting the assigned diet, we initiated a measurement of blood glucose levels at two weeks interval using an Accu-Chek Performa glucometer (Roche). Blood for regular glucose level measurements was obtained from the saphenous vein by puncturing the vein with a 25G needle. Accu Chek Performa glucometer with Accu Chek test strips was used to measure the

glucose level in blood samples. Before the procedure animals were fasted for 6 hours starting from 8:00 am to 2:00 pm. The induction of T2D was evaluated by glucose and insulin level measurements in plasma samples at week 20 after starting the assigned diet. The plasma necessary for the analysis was obtained from blood drawn from the saphenous vein. To estimate the insulin resistance, the HOMA-IR index was calculated (by formula $HOMA-IR=\frac{glucose,mmol/L*insulin,mU/L}{180}$) based on fasting plasma glucose and insulin levels determined by the mouse glucose assay and mouse insulin ELISA kit (both from Crystal Chem) following manufacturer's instructions at week 20.

Metformin was provided to mice with drinking water. The concentration of metformin was calculated to correspond to 50 mg/kg body mass/day. During the therapy period, all of the bottles, including those of the control group, were changed every day. Metformin was freshly added to the drinking water every day upon water changing. The duration of metformin therapy was ten weeks.

After the treatment period, at week 30, blood was collected from the saphenous vein for repeated evaluation of fasting glucose and insulin levels in plasma samples. On a day after the blood collection, all of the animals of the corresponding block were sacrificed by cervical dislocation without any other anesthesia, as the effect of other medications would interfere with the study's aims.

2.1.5 Sample collection

Fecal samples

Upon opening the cage, each mouse was immediately transferred to a clean, separate box in which animals were allowed to defecate voluntarily. Each of the animals was weighted. Feces were collected in sterile tubes in three aliquots.

Gut microbiome samples

Luminal and mucosal microbiome samples representing four different intestinal segments – proximal small intestine (duodenum), distal small intestine (ileum), cecum, and colon were collected. First, intestinal segments were rinsed separately with distilled water to collect luminal contents. Second, rinsed tissue samples were put in a tissue dish containing cold PBS and cut longitudinally. The remaining lumen contents were removed by repeated rinsing in separate clean tissue dishes. Mucosa samples containing microbiome were obtained by scraping the inner intestinal surface of the intestinal segment with a cell scraper and collected in sterile tubes. Samples were stored at -80ºC until further analysis.

2.2 Human study (Study IV)

2.2.1 Ethical approval

The study was approved by the Central Medical Ethics committee of Latvia (01-29.1.2/1751) and performed in agreement with the principles of the Declaration of Helsinki. Before the recruitment of donors, written informed consent was obtained from each participant after a thorough explanation of the aim of the study and the nature of each procedure.

2.2.2 Participant engagement criteria

In total, 43 patients undergoing colonoscopy procedures due to various reasons were included in the study. Patients provided informed consent and were recruited based on the following criteria: (1) the participant is of legal age; (2) women are not pregnant or breastfeeding; (3) there are no following health problems: a) acute illness at the time of enrollment; b) oncological diseases during the three years before enrollment in the study (three years after the completion of therapy); c) kidney failure or kidney dysfunction; d) autoimmune diseases; e) HIV, HBV or HCV; f) regular consumption of alcohol; (4) the following medications have not been used in the last two months: antibiotics; probiotics in the form of tablets/capsules; immunosuppressive medications; anticytokines; corticosteroids; proton pump inhibitors; (5) there have not been significant intestinal surgeries (for example, intestinal resection, bariatric surgery), except for appendectomy; (6) no diarrhea in the past week. All of the patients were divided into two groups based on the presence (n = 20) or the absence (n =23) of colorectal adenomas in their intestines.

2.2.3 Sample collection

Intestinal biopsy samples were collected by trained medical personnel during the colonoscopy procedure. Two to three sample aliquots (2 mm each) were collected from each patient in a sterile tube, with RNAlater added. To collect feces, a sterile fecal collection tube was provided to the participant, and the collection procedure was explained. Within the following 24 hours after the collection, the fecal sample was placed and stored at -80^{o}C.

2.3 Laboratory methods

2.3.1 Microbial DNA isolation

The microbial DNA from fecal samples representing each experimental unit at two time points – before and after the metformin treatment (N = 48) (for mouse study) or from patients undergoing colonoscopy procedures (N = 43) was extracted with the FastDNA®SPIN Kit for Soil

(MP Biomedicals) following manufacturer's instructions. DNA yield was determined using the Qubit dsDNA HS Assay Kit on the Qubit® Fluorometer 1.0 (Invitrogen Co.).

2.3.2 Total RNA isolation

Total RNA from fecal samples stored in RNAlater, of mice representing each experimental unit at the two time points – before and after the metformin treatment (N = 48) was isolated using All Prep DNA/RNA/miRNA Universal Kit (QIAGEN) according to the manufacturer's instructions. The same kit was used for total RNA isolation from gut mucosal samples from the same mice at the time point after the metformin treatment collected from two intestinal parts – proximal gut and cecum (N = 48) or intestinal biopsies of patients undergoing colonoscopy procedures (N = 43). Qubit RNA HS Assay Kit on the Qubit 2.0 (Invitrogen Co.) was used to determine the concentration of isolated RNA. The quality of the extracted RNA samples was analyzed using Agilent 2100 Bioanalyzer (Agilent, USA) and Agilent Small RNA Kit (Agilent, USA).

2.3.3 Metagenomic library preparation and sequencing

Libraries for metagenomic shotgun sequencing were prepared using MGIEasy Universal DNA Library Prep Kit (MGI Tech Co., Ltd.) for both mice and human samples. The input of DNA was 200 or 300 ng, respectively. The construction of libraries representing each study was completed in a single batch. Library preparation was performed according to the Universal DNA Library Prep Set User Manual and spiked with 1% PhiX. Preparation steps briefly: DNA shearing into 300 bp fragments by S220 focused-ultrasonicator (Covaris) followed by size selection using magnetic beads; end repair and A-tailing; adapter ligation followed by magnetic bead-based cleanup of adapter-ligated DNA; PCR and cleanup of the product; quality control; denaturation; single strand circularization; enzymatic digestion; cleanup of enzymatic digestion product; quality control by Qubit dsDNA HS Assay Kit on the Qubit® Fluorometer 1.0 (Invitrogen Co.), and Agilent High Sensitivity DNA Kit (Agilent Technologies) on the Agilent 2100 Bioanalyzer (Agilent Technologies). For metagenomic analysis of the samples collected from mice, each experimental unit was represented by one randomly chosen animal, as some experimental units contained only one animal at the end of the experiment. The microbiome generally was shared between animals in the same cage, indicated by sequencing all of the same cage's microbiome samples for some cages.

Pooled, circularized and barcoded libraries were used as templates for DNA nanoball preparation and further analyzed on DNBSEQ-G400RS next generation sequencing platform

(MGI Tech Co., Ltd.) using DNBSEQ-G400RS High-throughput Sequencing Set (FCL PE100) (MGI Tech Co., Ltd.) according to manufacturer's instructions. Sequencing depth was calculated to achieve at least 20 million paired-end 100 bp reads per sample.

2.3.4 16S rRNA gene library preparation and sequencing

The concentration of DNA before library preparation was normalized to 5 ng/μl for every sample. The V1-V2, V3-V4, and V5-V6 hypervariable regions of the 16S rRNA gene were amplified by PCR using specific primers tagged with Illumina sequencing adapters and sample-specific barcodes according to Illumina's instructions. PCR products were analyzed by 1.2% agarose gel electrophoresis and purified using NucleoMag (Macherey-Nagel) magnetic beads. Purified amplicons were pooled at equimolar concentrations, and sample indexes were added by additional PCR. The quality of libraries was assessed by Agilent High Sensitivity DNA Kit (Agilent Technologies) on the Agilent 2100 Bioanalyzer (Agilent Technologies). Samples were sequenced on Illumina MiSeq (Illumina) platform with MiSeq Reagent Kit V2 (500-cycles) (Illumina) according to the manufacturer's instructions.

2.3.5 Small RNA library preparation and sequencing

Fecal samples of mice

QIAseq miRNA Library QC Spike-ins (QIAGEN, Germany) were added before preparing the libraries. For library preparation with QIAseq miRNA Library Kit (QIAGEN) 100 ng of isolated RNA from each sample was used. Preparation steps briefly: 3' ligation; 5' ligation; reverse transcription, converting miRNAs into cDNA while assigning unique molecular indices (UMIs) to every miRNA molecule followed by cDNA cleanup using the magnetic bead-based method; library amplification, where a universal forward primer is paired with reverse primers to assign to each sample followed by library cleanup using the magnetic bead-based method; quality control by Qubit dsDNA HS Assay Kit on the Qubit 2.0 (Invitrogen Co.) and Agilent High Sensitivity DNA Kit (Agilent Technologies) on the Agilent 2100 Bioanalyzer (Agilent Technologies). The prepared libraries were sequenced on MiSeq System (Illumina) using MiSeq Reagent Kit v3 for 150 cycles (Illumina), following the manufacturer's instructions. Sequencing depth was calculated to achieve at least six million single-end small RNA-seq reads per sample.

Gut mucosal samples of mice or human intestinal biopsies

MiRNA libraries were prepared using MGIEasy Small RNA Library Prep Kit, the input of total RNA was 100 ng per sample. Preparation of the libraries involved: 3' ligation; 5' ligation; reverse transcription, library amplification; cleanup with magnetic beads; quality control by Qubit

dsDNA HS Assay Kit on the Qubit 2.0 (Invitrogen Co.) and Agilent High Sensitivity DNA Kit (Agilent Technologies) on the Agilent 2100 Bioanalyzer (Agilent Technologies); pooling of libraries; denaturation; circularization; enzymatic digestion and cleanup of the digestion product by magnetic beads; and quality control by Qubit ssDNA Assay Kit and Agilent High Sensitivity DNA Kit. The prepared libraries were sequenced on DNBSEQ-G400RS sequencing platform (MGI Tech Co. Ltd) using Small RNA FCL SE50 (MGI Tech Co. Ltd) following the manufacturer's instructions. At least 16 million single-end reads per sample were expected to be obtained.

2.4 Bioinformatics and statistical analysis

2.4.1 Biochemical parameters

Differences in body weights and biochemical parameters between HFD-fed and CD-fed groups were determined by one-way anova and t-test. The normality of measurement distributions was assessed by Shapiro-Wilk test, equality of variances was evaluated by an F-test.

2.4.2 Metagenomic data analysis

Mice fecal samples

Read quality evaluation was performed with FastQC (Andrews, 2010). Adapter clipping was performed with cutadapt v1.16 (Martin, 2011). Reads were trimmed from 5' and 3' end using 5 bp window with quality threshold 20 using Trimmomatic v0.39 (Bolger, Lohse and Usadel, 2014). Paired reads with length 75 bp or longer were retained for further data processing.

Reads originating from the host were removed by mapping reads against mouse reference genome GRCm38 release 96. Taxonomic classification of unmapped reads was performed with Kraken 2.0.8-beta against progenomes database (Mende *et al.*, 2020) with additionally included mouse (GRCm38) and human (GRCh38) reference genomes (Wood, Lu and Langmead, 2019). Only reads with a confidence score of 0.5 or higher were regarded as classified. Abundance reestimation was done with bracken v2.5 at species level (Lu *et al.*, 2017). Reads classified as *Homo sapiens* or *Mus musculus* were removed from subsequent analyses. Taxonomies with low read counts were removed using filterByExpr function implemented in edgeR 3.26.8 (Robinson, McCarthy and Smyth, 2009).

Due to the complex experimental design, the differential abundance analysis was not performed using statistical tools that take into account compositional nature of the data. Differential abundance testing was performed with limma 3.40.6 using voom transformation with

sample quality weights (Ritchie *et al.*, 2015). Differential testing was performed for combinations of multiple factors: metformin usage, time, diet, and sex. Correction for multiple testing was implemented with Benjamini-Hochberg method. Taxa with FDR ≤ 0.05 were regarded statistically significant. The effect of the individual mouse was accounted for using the duplicateCorrelation function in limma.

Alpha diversity was expressed as the exponential of Shannon diversity index resulting in the effective number of species, genera, and phyla at the respective taxonomic levels. To account for the compositional nature of taxonomic data, the imputation of zero values was performed with Bayesian-multiplicative replacement method as implemented in R 3.6.3 package zCompositions 1.3.4 with default parameters (Martín-Fernández *et al.*, 2015). Resulting taxonomic data were subsequently transformed using centered log ratio transformation with scikit-bio 0.5.5 (Aitchison, 1986). Aitchison's distance was used as a beta diversity metric. Principal component analysis biplot was constructed from the transformed compositions with scikit-learn 0.22.

Hypothesis testing for changes in alpha diversity before and after therapy was performed using paired t-test for each metformin and diet group separately. The normality of the diversity differences between time points was assessed with the Shapiro-Wilk test. P-values < 0.05 were considered statistically significant.

Human fecal samples

Read quality evaluation was performed with FastQC. Adapter cutting was performed with cutadapt v1.16. Reads were trimmed with fastp (0.20.0) by using default trimming parameters. Paired reads with a length of 75 bp or longer were retained for further data processing. Reads originating from the host were removed with bmtagger using GRCh38 as a reference. Taxonomic classification was performed with Kraken 2.0.8-beta and Bracken 2.7 against UHGG database (version 2) using Kraken’s confidence threshold value of 0.1.

2.4.3 Functional analysis

Functional analysis was performed by mapping sequencing reads against protein database and annotating matches with functional information of the corresponding protein. Paired-end reads were merged with FLASH v1.2.11. Merged and uncombined read-pairs were aligned against the SWISS-PROT database (release 2020_04) with DIAMOND (version 2.0.4). At least 80% of the read had to be aligned with 80% identity for it to be regarded as a possible hit. For merged read-pairs, a single best hit was retained (using DIAMOND's option –max-target-seqs =1). Uncombined

read-pairs were mapped independently. For each read up to 25 best hits were reported. If there was an overlap of matched UniProt IDs for a read-pair, then the hit with highest bit-score sum was selected from overlapping IDs. Otherwise hit with the highest bit-score from any read in a pair was selected. In both cases multiple best hits were resolved by random selection. Where possible, KEGG orthologies corresponding to UniProt identifiers were used to map KEGG BRITE functional hierarchies to UniProt IDs. Reads assigned to functional hierarchies of carbohydrate metabolism (ID 09101), lipid metabolism (ID 09103) and amino acid metabolism (09105) were counted resulting in a read count per BRITE table. BRITE IDs with median less than 100 reads were removed and differential abundance testing was performed as described in differential abundance analysis of taxonomies.

2.4.4 16S rRNA gene sequencing analysis

Data analysis began with the evaluation of sequencing data quality and read count distribution per sample group using the FastQC (v0.11.9) (Andrews, 2010) and MultiQC (v1.12) (Ewels *et al.*, 2016) tools, to identify possible sample level outliers and adapter contamination. Most of the analysis and diversity index calculation was performed using QIIME2 (v2022.2) (Bolyen *et al.*, 2019) microbiome analysis environment. First, we processed the regional amplicons for each individual, using the Cutadapt plugin (Martin, 2011) to trim the forward and reverse primers for each amplicon, specifying the allowed error rate of 0.1 and allowing for indels or deletions in bases when matching with the primers. In this step, reads which did not match to the primers were discarded.

The demultiplexed regional sequences were denoised with the DADA2 (Callahan *et al.*, 2016) plugin to generate Amplicon Sequence Variants (ASVs). The trim lengths for DADA2 were selected for each region based on the base quality drop-off threshold from the visual inspection of the sample group level sequence quality box and whiskers plots, to maximize the number of merged reads. Subsequently, all merged reads for each region were truncated to the length of 200 base pairs. The SILVA (v128) taxonomic database (Quast *et al.*, 2013) was imported into QIIME 2 using the RESCRIPt (Robeson *et al.*, 2021) plugin; the database was filtered to remove any sequence with more than 5 degenerate nucleotides. Regional database reads were extracted using the q2-feature-classifier plugin (Bokulich *et al.*, 2018). The regional database reads were aligned with the representative ASV sequences, allowing a mismatch of 2 nucleotides. Regional average relative abundances were solved through the Sidle implementation of the Short Multiple Reads Framework (SMURF) algorithm (Fuks *et al.*, 2018; Debelius *et al.*, 2021).

The phylogenetic tree was reconstructed by inserting consensus sequences for reconstructed amplicons into the SILVA (v128) backbone using the SEPP algorithm (Mirarab S, Nguyen N, 2012) phylogenetic reference backbone, while also inserting sequences that did not align to the SILVA taxonomy reference database. To discard low information or artifact sequences, the reconstructed ASV table was frequency filtered for features observed in at least three samples and with a taxonomic classification of genus level or higher. From the resulting table, a random feature subsample of 945 sequences per sample was made, to normalize for the differences in library size, which was then used to calculate Shannon diversity, Faith's phylogenetic diversity and Pielou's evenness alpha diversity indices. Statistically significant differences of the alpha diversity between analyzed groups were identified using the Wilcoxon Rank Sum test and Benjamini-Hochberg's procedure. False discovery rate (FDR) values < 0.05 were regarded as statistically significant.

Further analysis and visualization of results were performed in the RStudio (v2021.09.0) environment. Reconstructed and classified data was then imported into the phyloseq (v1.38.0) (McMurdie and Holmes, 2013) environment. Sample level ordination was calculated on rarefied (945 sequences, seed = 43980), genus-level aggregated data, which was then transformed with the centered log ratio method and reduced with principal components analysis for each intestinal part and layer. Taxonomic distribution bar plot graphs and ordination graphs were created with the microViz (v0.9.0.9009) (Barnett, Arts and Penders, 2021) and Matplotlib (3.5.2) (Hunter, 2007) packages while alpha diversity box plot graphs were created with the ggplot2 (v3.3.6) (Wickham, 2016) package. Finally, we performed a differential abundance test with the ANCOM-BC (v1.4.0) (Lin and Peddada, 2020) package, including independent variables in the formula and excluding features not observed in at least 10% of all samples. Median log fold change values of differentially abundant taxa of the same genera were visualized as bar plots using python libraries Matplotlib (3.5.2) and pandas (1.4.3) (McKinney, 2010).

Analysis of compositions of microbiomes with bias correction (ANCOM-BC) was performed using different combinations of factors: metformin treatment status; sex; diet type; intestinal layer; and intestinal part. Additionally, the effect of metformin was evaluated in each of the subsets formed by the combinations of the levels of the studied factors. Features representing the same genus were combined and medians of the LogFC of abundances were plotted in each of the analyzed contrasts. Individual dots were included in the plots to show the genera consisting of multiple features and the distribution of LogFC for each of the features. To investigate the interaction between metformin treatment and the studied factors, ANCOM-BC was performed

separately for each of the different combinations of factor levels, contrasting Met+ and Met- samples. Features of the same genus were combined and medians of the LogFC of abundances were plotted in each of the analysis subsets: proximal small intestine; distal small intestine; cecum; and colon.

2.4.5 MiRNA sequencing analysis – mice samples

Data were analyzed using CLC Genomics Workbench 20.0.4. and Galaxy Release v21.01. QIAseq Small RNA (Version 1.0) reference data set was used in the analysis. For fecal samples read quality assessment was performed with QC for Sequencing Reads. Reads were trimmed using a quality score of 0,012. 3' and 5' adapter trimming as well as sequence filtering on length (reads below 15 nt and above 55 nt were discarded) was performed. Read quality of gut mucosal samples was evaluated by Galaxy platform using FastQC (v0.11.8). Adapters were removed using Cutadapt (v.1.16), indicating AAGTCGGAGGCCAAGCGGTCTTAGGAAGACAA as a 3` adapter sequence and AAGTCGGATCGTAGCCATGTCGTTCTGTGAGCCAAGGAGTTG as a 5` adapter sequence. Quantify miRNA was used to map the reads against miRBase release v22, pointing out *Mus musculus* as prioritized species. Reads fixed for seed counting had a minimum sequence length of 18 nt and a maximum sequence length of 25 nt. Spike-in controls were enabled and reference from QIAseq Small RNA (Version 1.0) dataset was used. RNA-Seq Analysis tool was used to map the reads against mouse reference genome Ensembl v86. Spike-in controls were also used. Reads were mapped allowing 2 mismatches.

Differential abundance was tested with edgeR package 3.32.1 and limma 3.46.0 using voom transformation with sample quality weights. Differential testing was executed for combinations of multiple factors: sex, metformin treatment, time. MiRNA with p-value ≤ 0.05 were noted as statistically significant.

2.4.6 MiRNA sequencing analysis – human samples

MiRNA data were analyzed with Quantify miRNA on CLC Genomics Workbench 20.0.4. CLC Genomics Workbench in order to map the reads against miRBase release v22, designating *Homo Sapiens* as a prioritized species.

Differential abundance was tested with edgeR package 3.38.1 and limma 3.52.1. False discovery rate (FDR) values < 0.05 were regarded as statistically significant. Samples with less than 2000 assigned metagenomic or miRNA reads were removed. MiRNA features were filtered out by edgeR filterByExpr.

2.4.7 Correlation analysis

Taxa and miRNAs with at least 100 and 10 reads in 10% of samples, respectively, were retained for further analyses. Prior to centered log ratio normalization (clr), an arbitrary constant (value 1) was added to all values in order to enable clr. Correlation analysis of metagenome and miRNA datasets was performed on the obtained clr values using sparse partial least squares regression as implemented in R package mixomics 6.20.0.

2.4.8 Molecular target analysis

Potential binding sites for all miRNAs that had at least 10 reads in 10% of samples were evaluated in the corresponding sample metagenomic sequencing data. Since minimum free energy calculation is a relatively resource-intensive task, we first screened for reads that had at most three mismatches/indels with the evaluated miRNAs as a heuristic in order to speed up the process.

The number of mismatches/indels was calculated using Smith-Waterman's algorithm for sequence alignment (but since we are not interested in alignment itself, the traceback part of the algorithm was not implemented). For candidate reads, the minimum free energy (MFE) was calculated with RNAup 2.5.1. All reads with MFE values ≤ -20.0 kcal/mol were considered hits. Taxonomic classification of hits was then performed with Kraken 2.0.8-beta against the UHGG v2 database, and the number of hits per taxa per miRNA was aggregated as a table.

3 RESULTS

Study I: Metformin strongly affects gut microbiome composition in high-fat diet-induced type 2 diabetes mouse model of both sexes

3.1 Body mass and metabolic parameters of mice

Mice were fed with HFD for 20 weeks in order to induce T2D manifestations. Significant differences in body weights between HFD-fed and CD-fed mice were observed after two weeks with mean body weight 22.89 ± 3.36 and 19.9 ± 3.08 g, respectively (P-value = 0.03) (Figure 9). As expected, body weight was higher in males than in females in each of the diet groups. Metformin treatment had no significant effect on body weight gain in any of the studied groups.

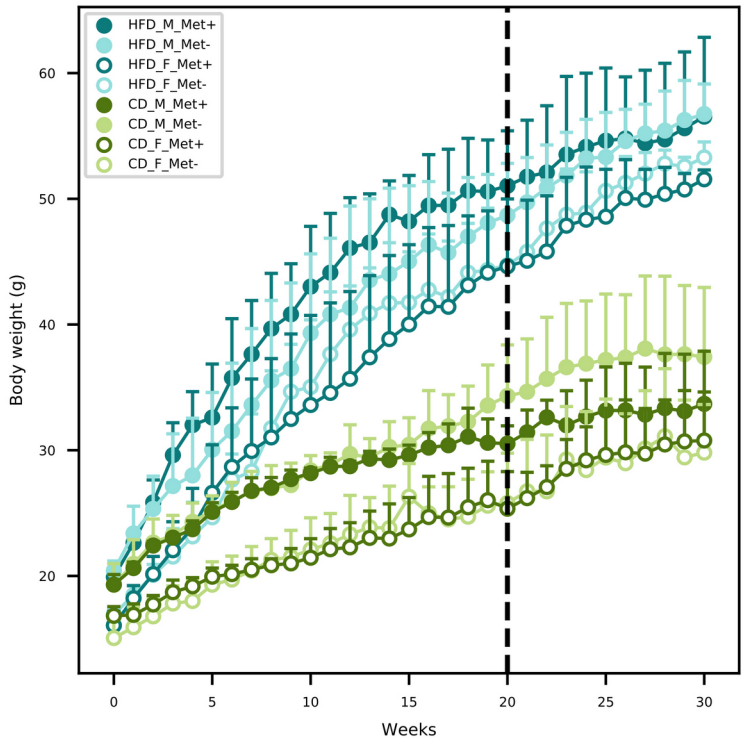

Figure 9. Mean body weight of each of the experimental groups in each of the weeks. Dashed line indicates the beginning of metformin treatment.

Fasting blood glucose level was monitored by glucometer fortnightly. Upon detecting statistically significant differences in blood glucose levels between HFD-fed and CD-fed groups, we initiated regular determination of fasting glucose and insulin levels in plasma samples (Table 1). We calculated the HOMA-IR index, and values above two were considered to correspond to

insulin resistance suggesting the onset of T2D. Before the beginning of the metformin treatment mean HOMA-IR index was above 2 in all of the HFD-fed groups but below 2 in all of the CD-fed groups. After the ten-week-long metformin treatment, fasting glucose and insulin levels in plasma samples were remeasured.

Table 1. Biochemical parameters before and after metformin treatment. n = 3 in each of the studied groups. Significance codes: 0 (***), 0.001 (**), 0.01 (*), >0.05 (NS).

		Before treatment			After treatment		
Group		**Glucose, mmol/L**	**Insulin, pmol/L**	**HOMA-IR**	**Glucose, mmol/L**	**Insulin, pmol/L**	**HOMA-IR**
HFD_M_Met-	Mean	12.21	730.14	7.37	11.11	2079.39	17.91
	SD	1.68	449.58	4.79	4.04	385.61	9.95
HFD_F_Met-	Mean	12.17	309.25	2.89	14.06	661.08	6.71
	SD	1.66	147.92	1.19	1.63	250.64	1.82
HFD_M_Met+	Mean	15.01	1537.10	18.31	11.77	1666.22	14.33
	SD	0.80	526.92	5.92	2.81	525.32	4.59
HFD_F_Met+	Mean	11.46	457.83	4.06	12.00	613.48	5.35
	SD	1.50	256.10	1.91	1.18	358.59	2.76
CD_M_Met-	Mean	11.69	127.82	1.21	12.03	194.80	1.68
	SD	2.89	48.07	0.51	1.01	162.22	1.27
CD_F_Met-	Mean	8.81	96.50	0.67	10.04	127.33	1.01
	SD	1.73	10.09	0.07	2.25	57.42	0.59
CD_M_Met+	Mean	10.15	221.05	1.71	9.23	133.04	0.93
	SD	1.69	234.81	1.75	0.51	131.09	0.96
CD_F_Met+	Mean	8.29	124.45	0.78	9.43	122.71	0.86
	SD	2.24	46.60	0.21	0.25	14.93	0.09
	Group	**	*	*	*	***	***
	Treatment	NS	NS	NS	NS	NS	NS
	Diet	**	**	**	*	***	***
	Sex	*	NS	*	NS	*	NS

3.2 Microbiome composition in experimental groups

We determined the microbiome composition of fecal samples by shotgun metagenomic sequencing. The median value of the obtained paired-end reads was 40044941 (IQR 10957336).

After quality trimming and host removal, a median of 28807601 (IQR 9664060) reads were retained. The median percentage of classified reads after taxonomic classification was 40.33% (IQR 12.24%).

The relative abundances of the top genera in each of the experimental groups are depicted in Figure 10. For comprehensibility reasons, only the genera with a relative abundance of at least 1% are presented.

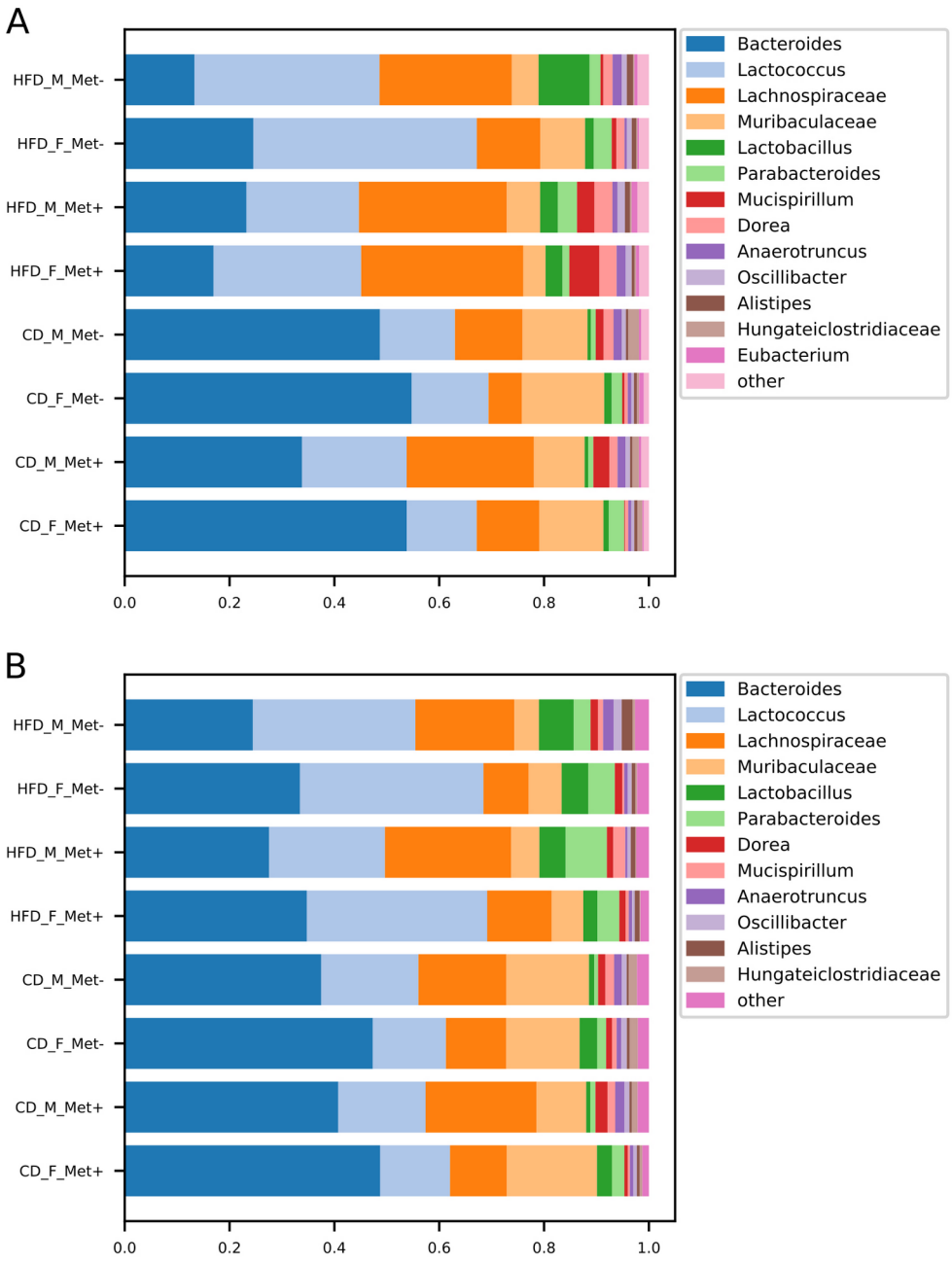

Figure 10. Microbiome composition at genus level in each of the experimental groups before and after the treatment. The abundances of top genera are expressed as proportions, only the genera or other

lowest identified taxa with the relative proportion of at least 1% are shown. (A) time point before the treatment; (B) time point after the treatment.

Before the initiation of the treatment, in all HFD-fed groups, *Lactococcus*, *Bacteroides*, and genera representing *Lachnospiraceae* dominated the gut microbiome composition. Other top taxa represented in all HFD-groups were *Muribaculaceae*, *Lactobacillus*, *Parabacteroides*, *Mucispirillum*, and *Dorea*. *Bacteroides* prevailed in all CD-fed groups, followed by *Lactococcus*, *Lachnospiraceae*, and *Muribaculaceae*, the relative abundance of which was greater in all CD-fed groups compared to HFD-fed mice.

At the time point after the treatment, we identified an increase in *Bacteroides* relative abundance in all HFD-fed groups, which was maintained at the expense of reduced relative abundance of *Lachnospiraceae*, whereas *Bacteroides*, *Lactococcus*, and *Lachnospiraceae* remained top taxa in all HFD-fed groups. All CD-fed groups were dominated by *Bacteroides*, followed by smaller proportions of *Lactococcus*, *Lachnospiraceae*, and *Muribaculaceae*.

3.3 Diversity analysis

3.3.1 Alpha diversity

We identified a trend of decrease in the alpha diversity of the HFD_Met+ groups before and after metformin treatment (Figure 11), though not statistically significant. In CD_Met+ groups and groups that did not receive metformin, no alpha diversity changes were observed before and after metformin treatment.

Overall the effective number of species (ENS) was higher in HFD-fed groups – mean ENS before treatment was 10.99 ± 3.04 and 9.62 ± 2.64 in HFD-fed and CD-fed mice, respectively. The same was observed at the time point after the treatment – mean ENS was 10.62 ± 3.55 and 10.18 ± 1.70 in HFD-fed and CD-fed mice, respectively.

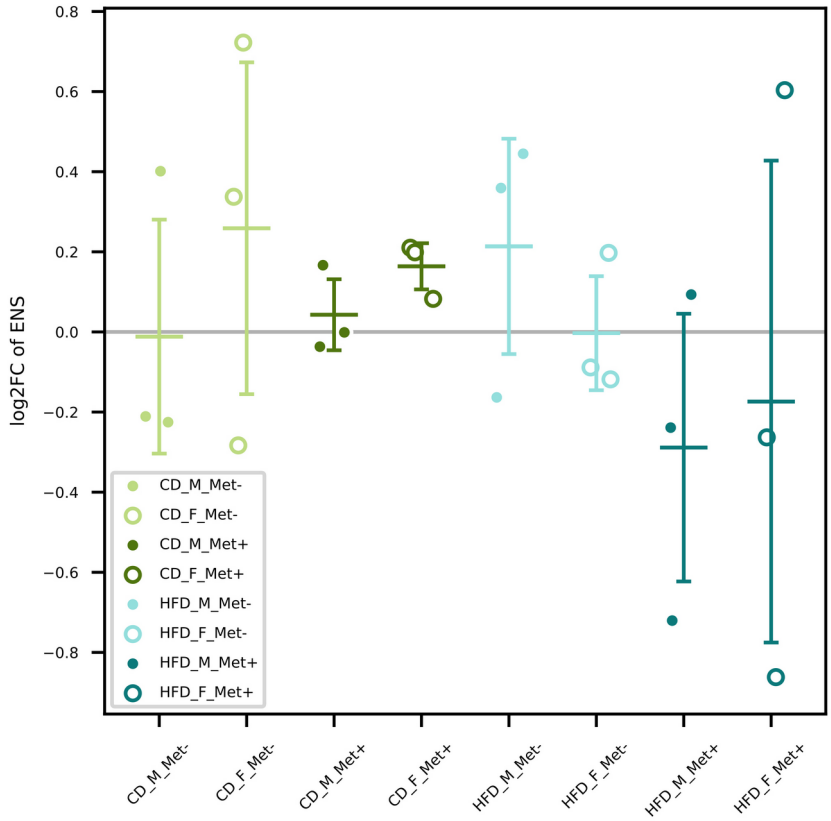

Figure 11. Alpha diversity analysis. Changes in alpha diversity in each of the groups expressed as an effective number of species. Samples representing each experimental unit in each of the experimental groups are shown as dots.

3.3.2 Beta diversity

Beta diversity analysis revealed clustering depending on various factors included in the study (Figures 12A and 12B). Samples obtained from HFD-fed mice clustered apart from those of CD-fed mice at both time points, with the principal microbial identifiers being *Pediococcus acidilactici*, *Lactobacillus spp.*, *Desulfomicrobium orale*, *Desulfovibrio fairfieldensis*, *Fusicatenibacter saccharivorans*, *Proteus spp.* before the beginning of metformin treatment. *Lactospiraceae bacterium 6_1_37FAA*, *Staphylococcus xylosus*, *Intestinimonas massiliensis* supplemented this list of microbial identifiers after the metformin treatment, though some species were represented by < 100 reads in any of the samples. Furthermore, when analyzing both sexes separately, before the treatment, HFD_F groups were directed toward *Lactobacillus* vectors, but HFD_M groups towards *Proteobacteria* members. In both sexes of CD-fed mice, principal identifiers were affiliated to *Bacteroidetes* and *Clostridia*, although the species were different for males and females. After the treatment, samples of HFD-fed mice tended to cluster closer, and

Clostridia members appeared among the most characteristic taxa of these groups. In CD_F groups, representatives of *Bacteroidetes* and *Clostridia* remained the principal identifiers; however, in CD_M_Met+ group, a shift towards *Proteobacteria* was observed. Nevertheless, no apparent clustering regarding metformin treatment status was observed.

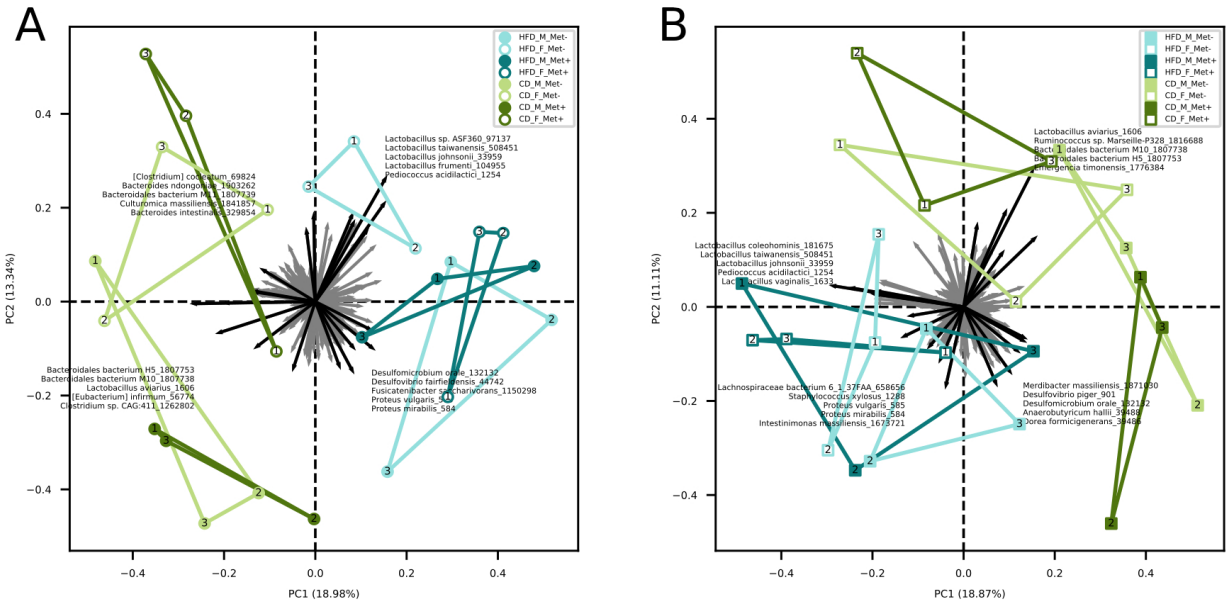

Figure 12. Beta diversity analysis. (A) Beta diversity in each of the groups before the beginning of metformin treatment. (B) Beta diversity in each of the groups after 10 weeks long metformin treatment. Samples representing each experimental unit in each of the experimental groups are shown as dots.

3.4 Differentially abundant species between treatment arms

We evaluated the relative abundance of different species of microbes between groups using various contrasts shown in Figure 13. The analysis revealed significant differences between HFD_Met- and CD_Met- groups and between HFD_Met+ and CD_Met+ groups before the beginning of metformin treatment. We observed 63 and 48 differentially abundant species between HFD- and CD-fed groups with or without metformin treatment, respectively (Figure 13), indicating apparent differences between these experimental units at the time point when experimental units were allocated to receive metformin treatment. At the same time, no differentially abundant taxa were identified between CD-fed groups and between HFD-fed groups. When considering the effect of sex, only one species, *Bacteroides eggerthii* (LogFC = -4.05, FDR < 0.001), was detected to be differentially abundant between HFD-fed groups, indicating that some variability may exist between identically treated groups based on sex differences. No other

substantial differences in microbiome composition between the same groups before the beginning of metformin treatment were found.

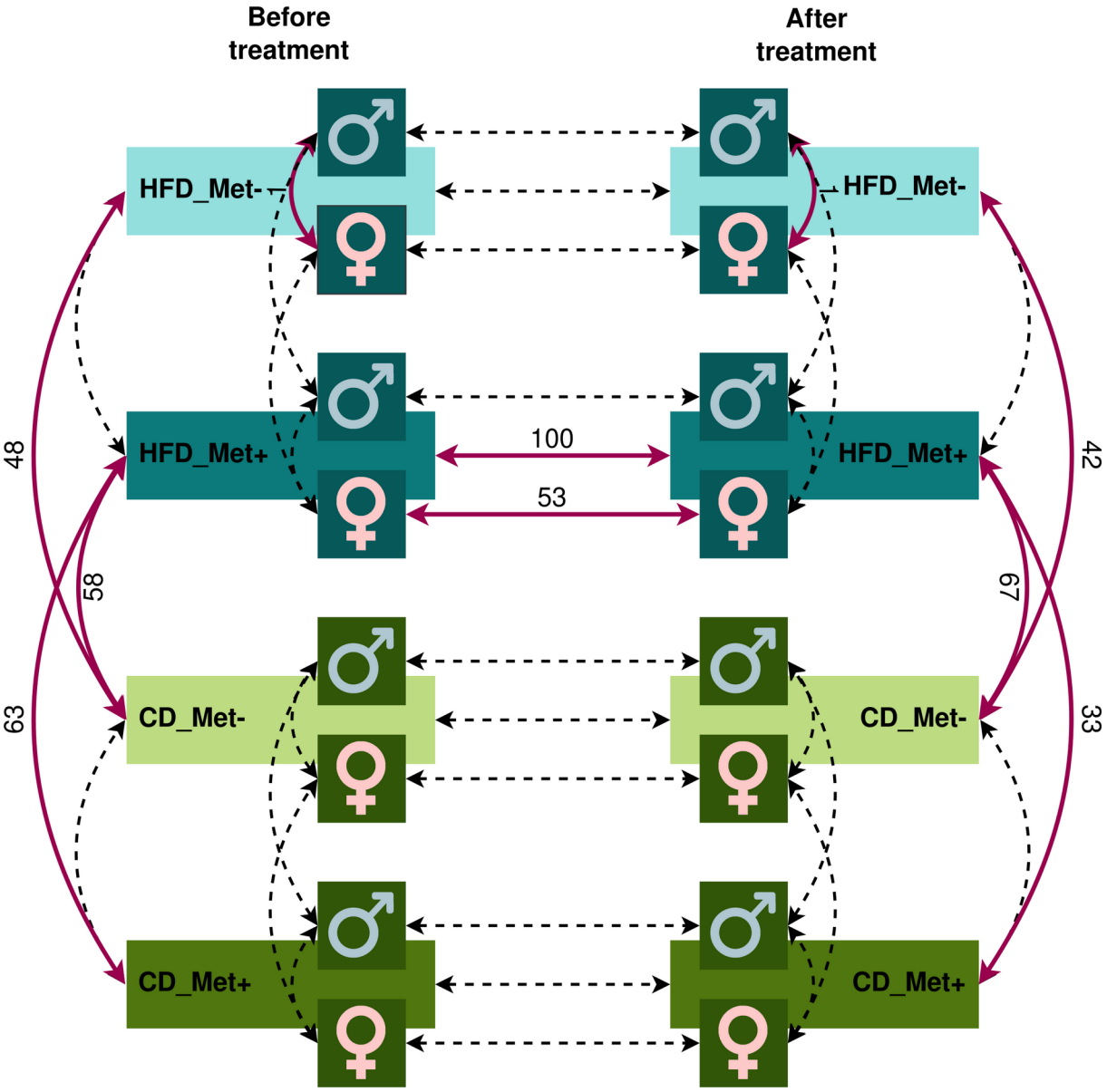

Figure 13. Contrasts in which microbiome compositions were compared. Dashed lines indicate the contrasts between which a comparison was performed. Red bold lines indicate the contrasts between which statistically significant differences in taxa relative abundance at species level were discovered with the numbers of the different species.

We did not find any differences between HFD_Met+ and HFD_Met- groups at the time point after the treatment (Figure 13). The same applies to the contrast between CD_Met+ and CD_Met- groups. When each of the sexes was compared, *Bacteroides eggerthii* remained significantly differentially abundant between HFD_Met- groups (LogFC = -3.64, FDR = 0.007).

We observed a strong effect of diet on microbiome composition both before and after metformin treatment. Before initiating metformin treatment, we observed 48 and 63 differentially abundant species in contrasts between HFD_Met- and CD_Met- groups and between HFD_Met+ and CD_Met+ groups, respectively. The same applied to identical contrasts at the end of the experiment; 42 differentially abundant species were identified between HFD_Met- and CD_Met- groups and 33 – between HFD_Met+ and CD_Met+ groups. Common to all of the contrasts mentioned above, HFD was associated with a lower relative abundance of *Bacteroidales* bacteria, *Prevotella sp.*, *Lactobacillus aviarius*, *Bacteroides helcogenes*, and *Bacteroides oleiciplenus*.

When comparing the differentially abundant taxa between HFD-fed and CD-fed groups representing the same metformin treatment status cross-sectionally, in all contrasts not influenced by metformin (HFD_Met- vs. CD_Met- before treatment, HFD_Met- vs. CD_Met- after treatment, and HFD_Met+ vs. CD_Met+ before treatment) in HFD groups we observed higher relative abundance of *Acetivibrio ethanolgignens* (LogFC values ranging from 3.07 to 3.84, FDR ≤ 0.008) and lower relative abundance of *Prevotella lascolai* (LogFC values ranging from -2.54 to -1.70, FDR ≤ 0.008), *Gabonia massiliensis* (LogFC values ranging from -3.09 to -1.77, FDR ≤ 0.02), *Culturomica massiliensis* (LogFC values ranging from -3.24 to -1.79, FDR ≤ 0.03), and several *Bacteroides* species (LogFC values ranging from -3.72 to -1.82, FDR ≤ 0.03). The pairwise comparison between CD_Met+ groups before and after the treatment showed no differences in taxa relative abundance (Figure 13); the same was observed between CD_Met- groups.

Comparing HFD_Met- groups between the time points of metformin initiation and the end of the experiment, we observed no differentially abundant taxa. However, in HFD_Met+ groups, 100 species were altered (Figures 14 and 15), with most of the species relative abundances being increased due to metformin treatment. The most pronounced changes were in the relative abundance of such *Bacteroidetes* genera as *Bacteroides*, *Parabacteroides*, *Prevotella*, *Paraprevotella*, *Porphyromonas; Firmicutes* genera *Bacillus*, *Butyrivibrio*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, and *Streptococcus* as well as in *Enterorhabdus* representing *Actinobacteria*.

In general, the variability of the magnitude of the differences in the relative abundance of species between the samples of the studied contrast was more pronounced in species representing phylum *Bacteroidetes* (LogFC values ranging from 0.70 to 2.46). In contrast, members of the class *Bacilli – Lactobacillus*, *Lactococcus*, *Enterococcus*, *Leuconostoc*, *Bacillus*, and *Streptococcus*

genera all representing *Firmicutes* were increased (LogFC values ranging from 0.79 to 1.17) in response to metformin in all of the samples relatively uniformly.

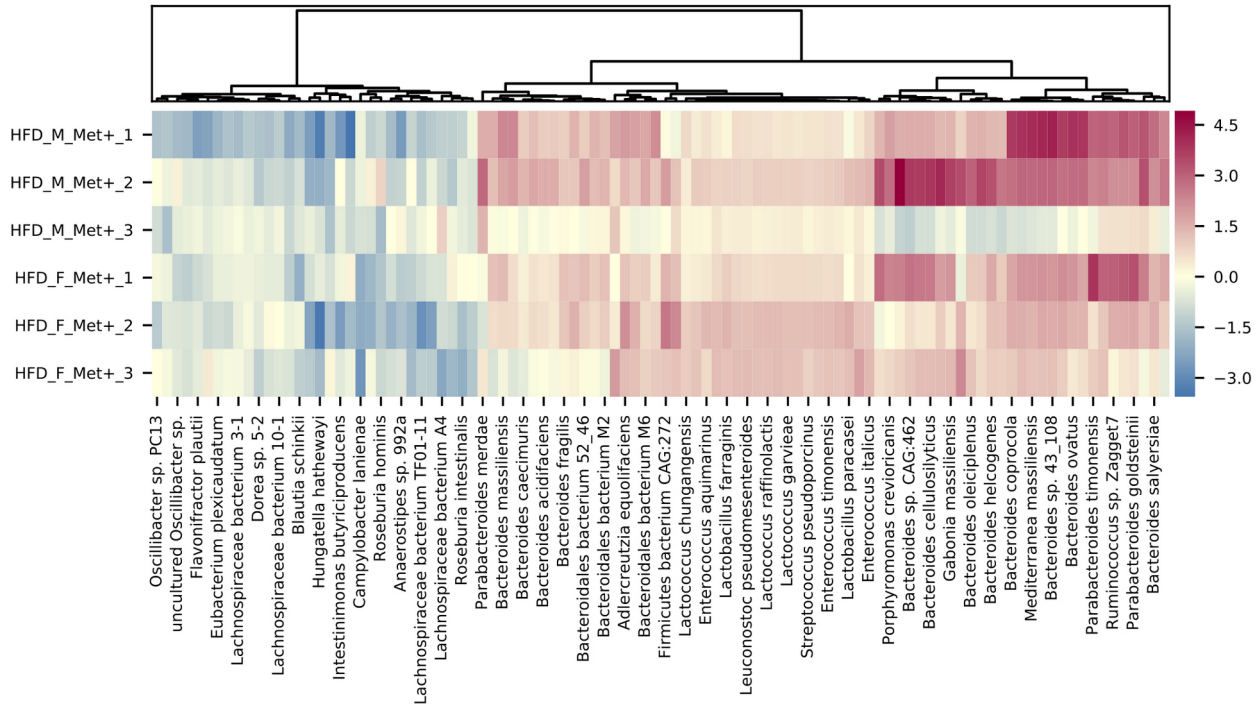

Figure 14. Heatmap showing differentially abundant species in HFD_Met+ groups before and after metformin treatment.

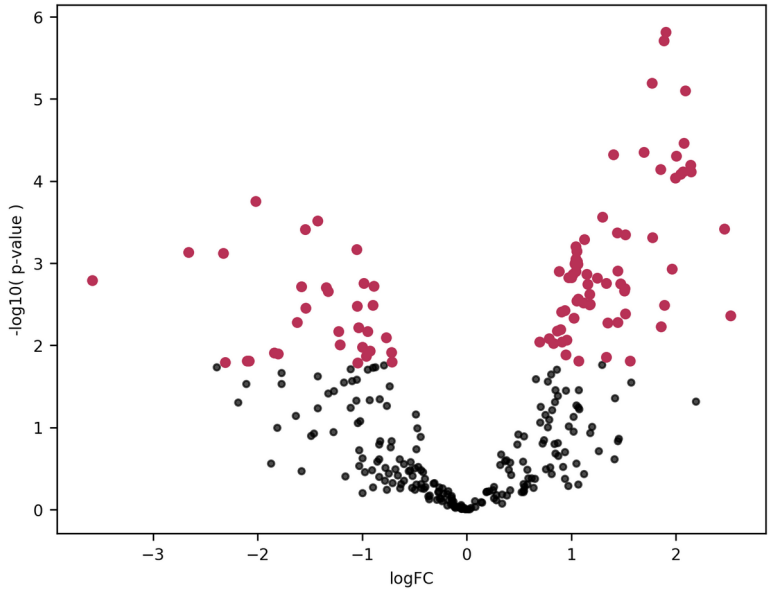

Figure 15. Volcano plot showing P-values of differentially abundant species in HFD_Met+ groups before and after metformin treatment. Red dots represent differentially abundant species with P-values < 0.05.

Analysis of relative taxa abundances in HFD_Met+ groups before and after metformin treatment in each of the sexes separately revealed sex-specific effects of metformin. In males, we did not identify significantly differentially abundant taxa while in females, the relative abundance of 53 species was significantly altered (Figures 16 and 17). A decrease in response to metformin was observed in bacteria of *Clostridia* class including *Faecalibacterium prausnitzii* (LogFC = -2.18, FDR = 0.03), *Enterocloster clostridioformis* (LogFC = -1.74, FDR = 0.03) and *Anaerostipes sp. 992A* (LogFC = -1.64, FDR = 0.04) as well as *Desulfovibrio fairfieldensis* (LogFC = -2.97, FDR = 0.04) of *Deltaproteobacteria* class. We identified an increase in the differentially abundant taxa in response to metformin; a subtle increase in the relative abundance of species representing *Bacilli* was observed (LogFC up to 1.36), while the increase in *Bactoroidia* class representatives was particularly pronounced, for example, *B. ilei* (LogFC = 2.85), *B. vulgatus* (LogFC = 2.45), *B. pyogenes* (LogFC = 2.43). The list of differentially abundant species in females corresponds to the taxa identified in the analysis where samples from both sexes are taken together, although the extent of changes in relative abundances was greater in females.

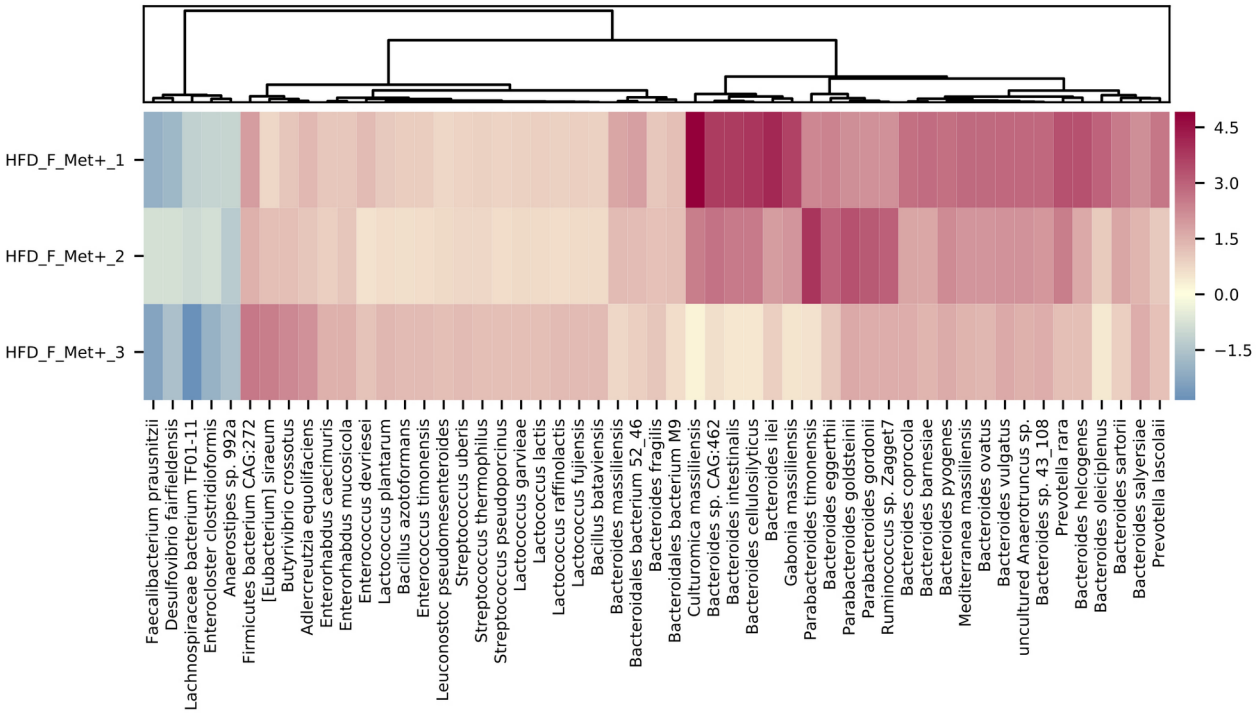

Figure 16. Heatmap showing differentially abundant species in HFD_F_Met+ group before and after metformin treatment.

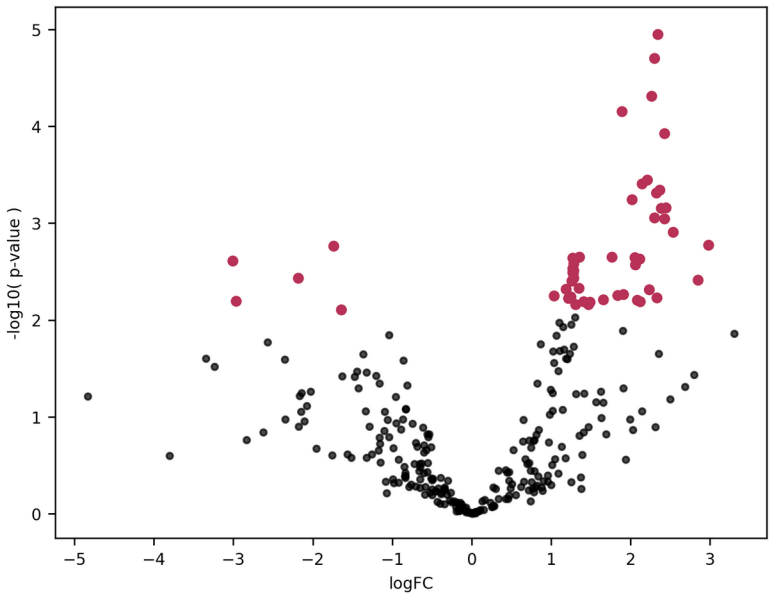

Figure 17. Volcano plot showing P-values of differentially abundant species in HFD_F_Met+ group before and after metformin treatment. Red dots represent differentially abundant species with P-values < 0.05.

To further analyze the effect of metformin treatment we compared microbial compositions between CD_Met- and HFD_Met+ groups. At the time point before the treatment we observed 58 differentially abundant taxa dominated by members of *Clostridia* and *Bacteroidetes* which were increased and decreased in HFD-fed animals, respectively. *Clostridiaceae* were represented by *Hungatella hathewayi* (LogFC = 6.65, FDR < 0.001) and *Acetivibrio ethanolgignens* (LogFC = 4.30, FDR < 0.001). Members of *Lachnospiraceae* including *Roseburia* and *Dorea* followed with LogFC from 0.94 to 3.35, FDR <0.05, as well as non-*Clostridia Mucispirillum schaedleri* (LogFC = 2.89, FDR = 0.02), and *Enterorhabdus mucosicola* (LogFC = 1.27, FDR = 0.02). *Bacteroidetes* were represented by various *Bacteroidales* with the most extreme changes identified in the relative abundance of *Bacteroidales bacterium M10* (LogFC = -10.33, FDR < 0.001), in addition *Parabacteroides timonensis* (LogFC = -2.04, FDR = 0.01), members of *Prevotella* (LogFC from -2.53 to -3.61, FDR < 0.001) and others were decreased.

The same contrast at the time point after the metformin treatment revealed 67 differentially abundant species. *Bacilli* and *Bacteroides* prevailed among the increased taxa, while *Clostridia* and different *Bacteroidetes* represented decreased taxa. *Bacilli* were represented by *Staphylococcus xylosus* (LogFC = 6.40, FDR = 0.02) and members of genera *Enterococcus*, *Streptococcus*, *Leuconostoc*, and *Lactococcus* (LogFC from 1.08 to 1.58, FDR > 0.05). LogFC of *Actinobacteria Enterorhabdus mucosicola* was 1.91, FDR = 0.004). Among increased *Bacteroides*

were such species as *B. salyersiae*, *B. pyogenes*, *B. vulgatus*, *B. ovatus*, *B. coprocola*, *B. eggerthii*, *B. clarus*, *B. congonensis*, *B. caccae*, and *B. acidifaciens* (LogFC from 1.12 to 2.40, FDR < 0.05). Decreased *Bacteroidetes* included *Bacteroidales bacterium M10* (LogFC = -9.68, FDR < 0.001), other *Bacteroidales*, members of *Prevotella*, *Porphyromonas*, and *Bacteroides* species *B. ilei, B. helcogenes, B. gallinarum, B oleiciplenus, B. cellulosilyticus, B. togonis, B. intestinalis,* and *B. coprophilus* (LogFC from -1.19 to -5.62, FDR < 0.05). *Clostridia* were represented by *Dorea formicigenerans* (LogFC = -3.38, FDR = 0.04) and *[Eubacterium] infirmum* (LogFC = -4.65, FDR = 0.009). The 58% and 76% of differentially abundant species were shared between contrasts HFD_Met- vs. CD_Met- and HFD_Met+ vs. CD_Met+, respectively. When the same contrasts were analyzed at the time point after the treatment, 83% and 100% were shared.

3.5 Differentially abundant functional hierarchies between treatment arms

To investigate the changes in the functional profile of the gut microbiome in response to metformin treatment, we evaluated the relative abundance of each Kyoto Encyclopedia of Genes and Genomes (KEGG) BRITE hierarchy in metagenomic data focusing on carbohydrate, lipid, and amino acid metabolism. Statistical analysis of the differentially abundant functional hierarchies between experimental groups in various contrasts corresponding to the ones indicated in Figure 10 showed significant differences in ten occasions mainly shared with the previously identified significant differences in taxa relative abundances.

In HFD_Met+ groups between the time points of metformin treatment initiation and the end of the experiment, 14 metabolic hierarchies were altered. The most markedly decreased functions were phenylalanine metabolism (ko00360) (LogFC -0.69, FDR < 0.001), and glyoxylate and dicarboxylate metabolism (ko00630) (LogFC = -0.66, FDR = 0.002). The top increased hierarchies were arachidonic acid metabolism (ko00590) (LogFC = 0.36, FDR < 0.05), and arginine and proline metabolism (ko00330) (LogFC = 0.32, FDR < 0.05).

Similar to what was observed in the longitudinal contrast between HFD_Met+ groups when analyzing differentially abundant species, functional analysis in each of the sexes separately reinforced sex-related differences. In males, none of the analyzed hierarchies showed alterations in the number of corresponding reads. In females, 5 functions were significantly differentially represented by the reads. The list of altered hierarchies in females coincides with the ones identified in the combined analysis of both sexes, albeit the changes were more pronounced in female mice, for example, phenylalanine metabolism (ko00360) (LogFC =1.07, FDR < 0.001), and

glyoxylate and dicarboxylate metabolism (ko00630) (LogFC = -0.96, FDR = 0.003). In contrast to the combined analysis in both sexes, in females, the relative abundance of only one hierarchy - pyruvate metabolism (ko00620) was slightly increased (LogFC = 0.15, FDR < 0.05).

Study II: Spatial variation of the gut microbiome in response to long-term metformin treatment in high-fat diet-induced type 2 diabetes mouse model of both sexes

3.6 Microbial composition analysis

In total, 192 microbiome samples representing four different intestinal regions at the luminal and mucosal layers collected from 24 mice were sequenced. One of the samples was excluded from further analysis due to possible mislabeling. Mice representing three experimental units were included in each group representing each of the eight treatment arms described in this study (Figure 18).

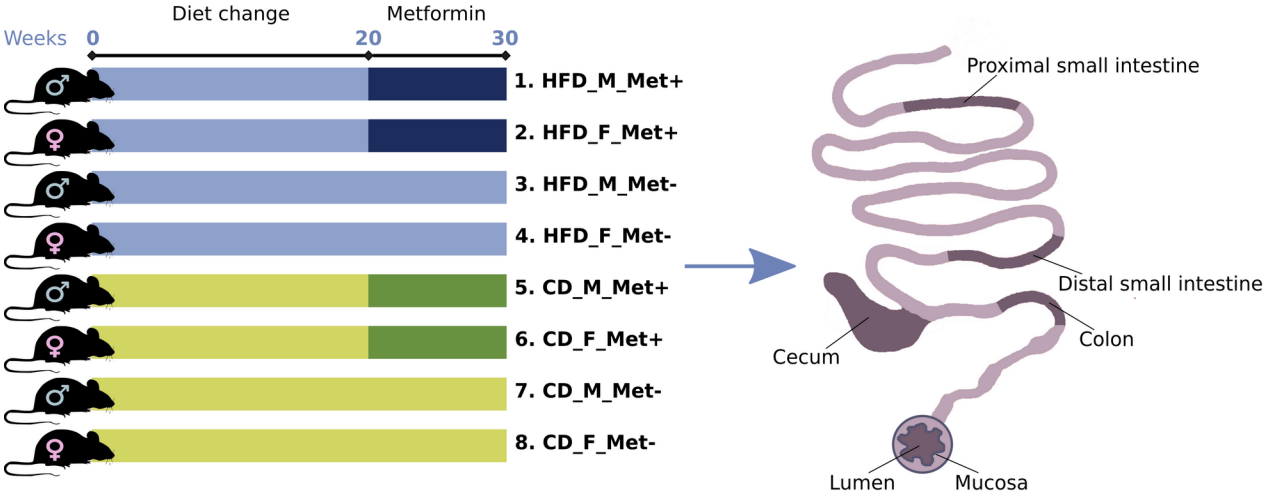

Figure 18. Experimental design of the study (N = 24) **and intestinal sites studied.** Abbreviations: HFD – high-fat diet; CD – control diet; M – male; F – female; Met+ – receiving metformin treatment for 10 weeks; Met- – not receiving metformin treatment.

To summarize the microbiome composition in each analyzed site, we compiled a list of the top 20 location-specific genera, as shown in Figures 19 and 20. We observed variation in microbial composition both longitudinally and cross-sectionally. Both parts of the small intestine were dominated by *Lactobacillus* in high-fat diet (HFD)-fed mice, followed by *Pseudomonas* and *Microbacterium* in the mucosal layer and *Pseudomonas* and *Lactococcus* in the luminal layer (Figures 19A; 19B; 19E; 19F). Likewise, *Lactobacillus* and *Pseudomonas* prevailed in control diet (CD)-fed mice, followed by *Lactococcus* in the mucosal layer (Figures 20A and 20B).

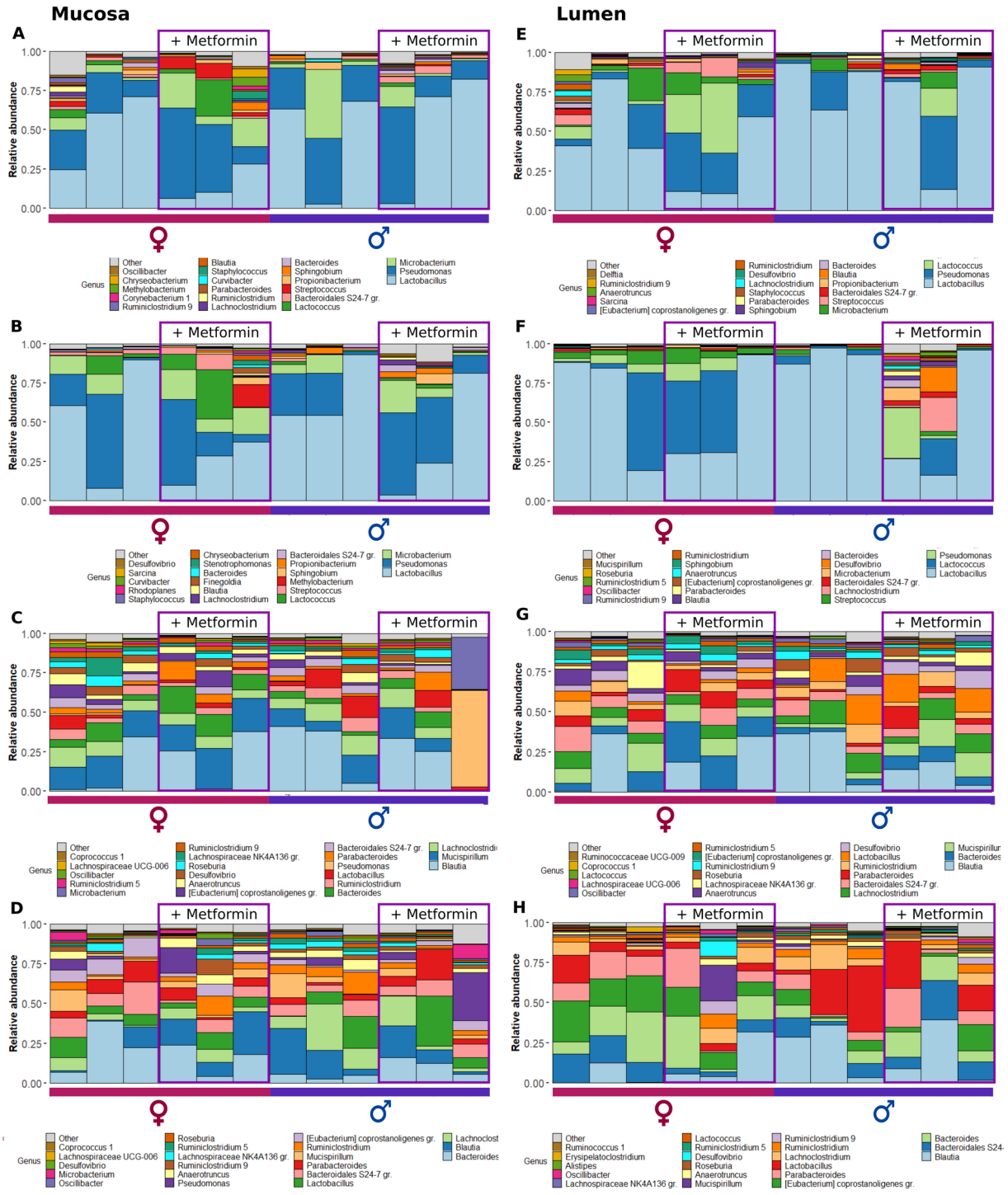

Figure 19. Microbiome composition at different sites of high-fat diet-fed mice at the genus level – top 20 genera are shown. Mucosal layer: (A) proximal small intestine; (B) distal small intestine; (C) cecum; (D) colon. Luminal content layer: (E) proximal small intestine; (F) distal small intestine; (G) cecum; (H) colon. Samples representing mice receiving metformin treatment are highlighted. Red and blue bars under each plot indicate females and males, respectively.

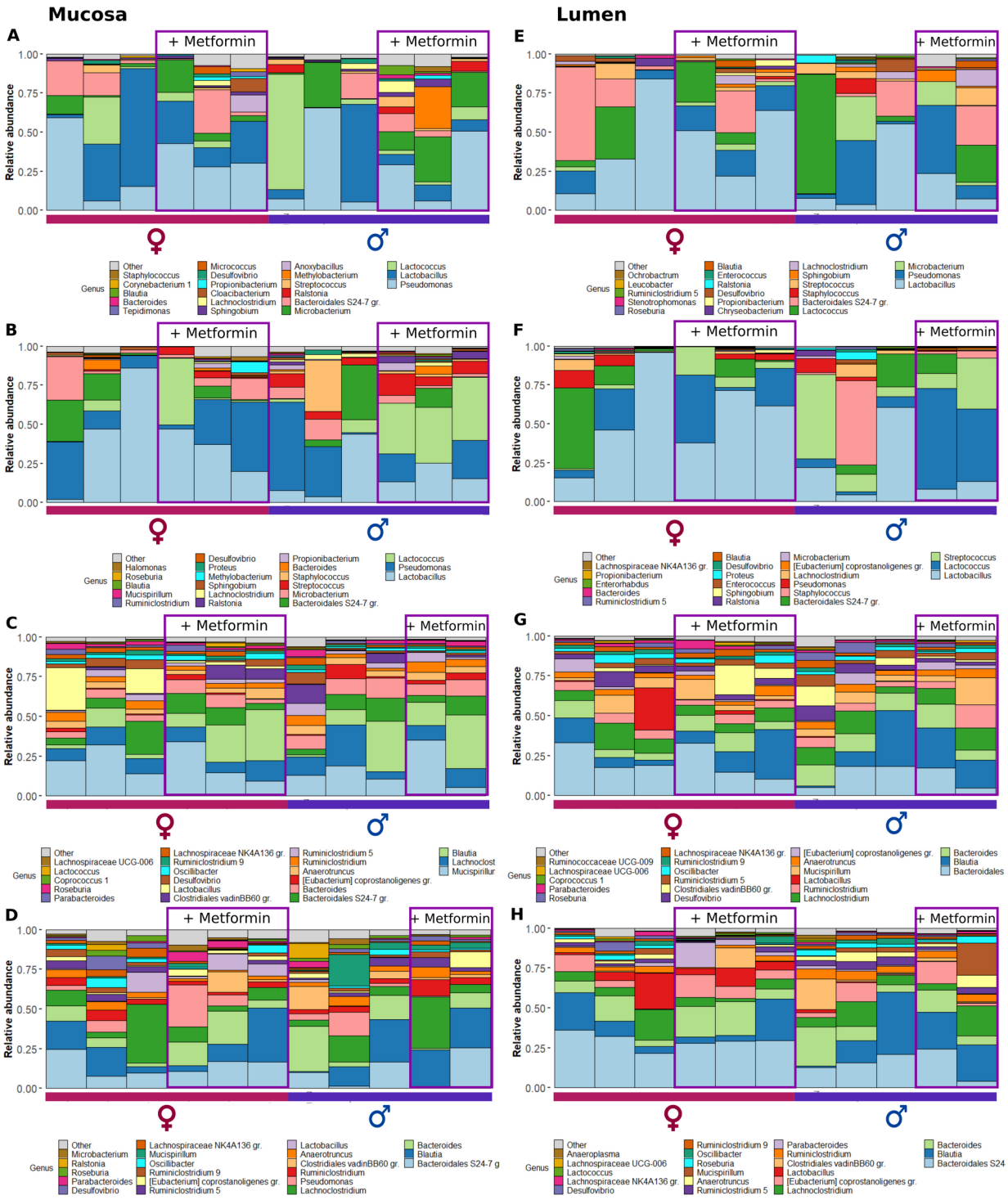

Figure 20. Microbiome composition at different sites of control diet-fed mice at the genus level – top 20 genera are shown. Mucosal layer: (A) proximal small intestine; (B) distal small intestine; (C) cecum; (D) colon. Luminal content layer: (E) proximal small intestine; (F) distal small intestine; (G) cecum; (H) colon. Samples representing mice receiving metformin treatment are highlighted. Red and blue bars under each plot indicate females and males, respectively.

In the luminal layer, *Lactobacillus* was followed by *Pseudomonas* and *Microbacterium* in the proximal part of the small intestine and by *Lactococcus* and *Streptococcus* – in the distal part (Figures 20E and 20F). The relative abundance of *Lactococcus* was increased in the distal small intestine of CD-fed male mice receiving metformin treatment. In contrast, in HFD-fed mice, this effect was more pronounced in female mice in both parts of the small intestine. *Streptococcus* was more abundant in the small intestine of female HFD-fed mice than in males, and its relative abundance was higher in the metformin-treated animals.

In HFD-fed mice, *Blautia* had the highest relative abundance in the cecum at both layers (Figures 19C and 19G). *Mucispirillum*, *Lachnoclostridium*, and *Bacteroides* were among the other top genera in the cecum. In CD-fed mice, *Mucispirillum* showed the highest relative abundance in the mucosa, while luminal content was enriched in *Bacteroidale*s representatives (Figures 20C and 20G). *Blautia* and *Lachnoclostridium* had high relative abundance in both layers. Similar to the cecum, in the colon of HFD-fed mice top bacteria were *Bacteroides* in the mucosa, followed by *Blautia* and *Lachnoclostridium,* and *Blautia* in the luminal layer, followed by *Bacteroides* and other *Bacteroidales* members (Figures 19D and 19H). Similar results were observed in CD-fed mice (Figures 20D and 20H).

3.7 Diversity analysis

3.7.1 Alpha diversity analysis

Shannon diversity index analysis revealed significant differences in alpha diversity between all the studied intestinal segments, with the most pronounced differences present between the distal part of the small intestine and two other locations, cecum, $H = 55.90$, p-value < 0.001 and large intestine, $H = 55.67$, p-value < 0.001 (Figure 21). Similar results were observed for Pielou's evenness and Faith's phylogenetic diversity (data not shown). There were no significant differences in alpha diversity between intestinal layers for all the metrics analyzed. When the effect of metformin treatment on alpha diversity in each intestinal site was evaluated, we did not observe any significant differences after adjustment for multiple testing for all the metrics analyzed.

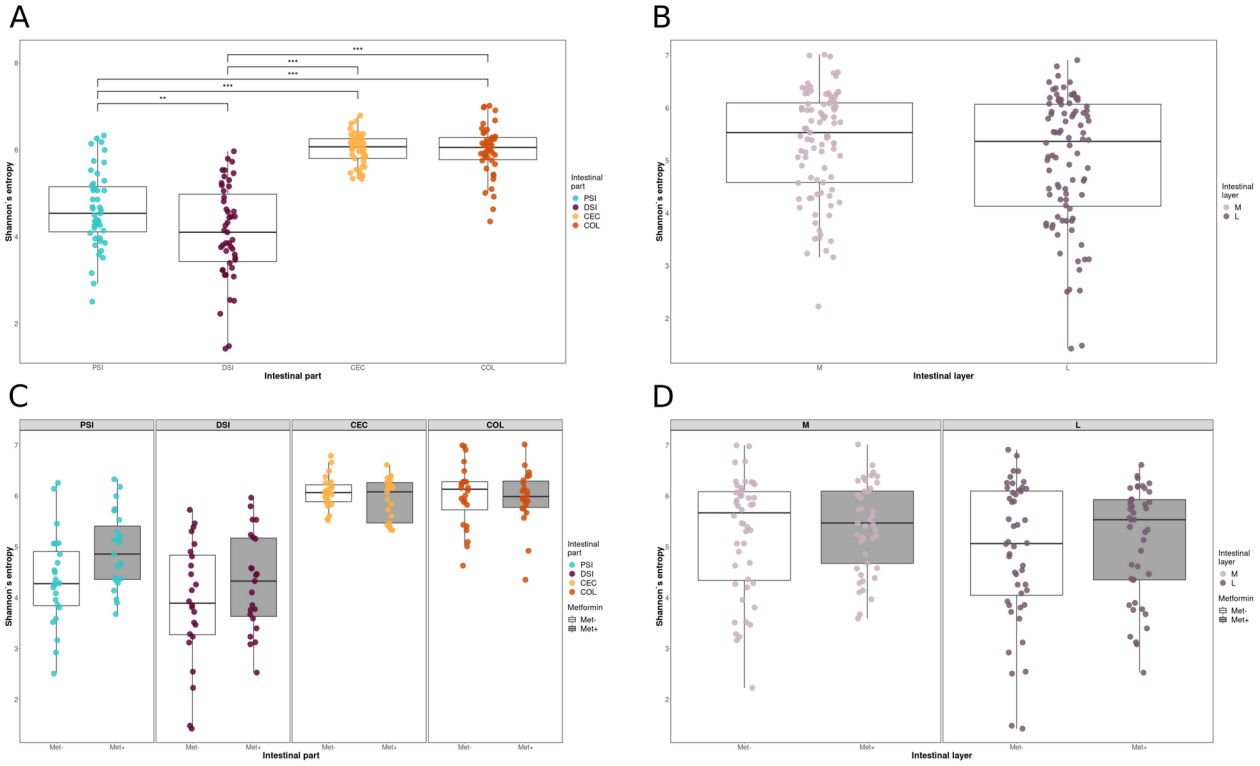

Figure 21. Microbiome alpha diversity analysis expressed as Shannon index: (A) between intestinal parts; (B) between intestinal layers. Microbiome alpha diversity analysis in response to metformin treatment expressed as Shannon index, n = 12 per group: (C) between intestinal parts; (D) between intestinal layers. PSI – proximal small intestine; DSI – distal small intestine; CEC – cecum; COL – colon; M – mucosa; L – lumen. **Adjusted p < 0.05; *** adjusted p < 0.01.

3.7.2 Beta diversity analysis

Beta diversity was evaluated by ordination analysis. Biplots using PCA were generated, including each of the studied groups based on the diet type and metformin treatment status (Figures 22 and 23). When taken together, samples representing both parts of the small intestine clustered separately from the samples of the cecum and colon (Figure 22). *Sphingobium*, *Sarcina*, *Propionibacterium*, *Pseudomonas*, and *Microbacterium* representatives were the principal microbial identifiers of the proximal small intestine. *Lactococcus*, *Lactobacillus*, *Streptococcus*, *Enterococcus*, and *Staphylococcus* were the main drivers of the distal small intestine. In turn, cecum samples were identified by *Mucispirillum*, *Anaerotruncus*, *Blautia*, *Ruminiclostridium*, and *Bacteroides*. Colon was characterized by *Desulfovibrio*, *Parabacteroides*, *Eubacterium_coprostanoligenes_group* members, *Alistipes*, and *Bacteroidales_S24-7_group* (now known as *Muribaculaceae*) representatives.

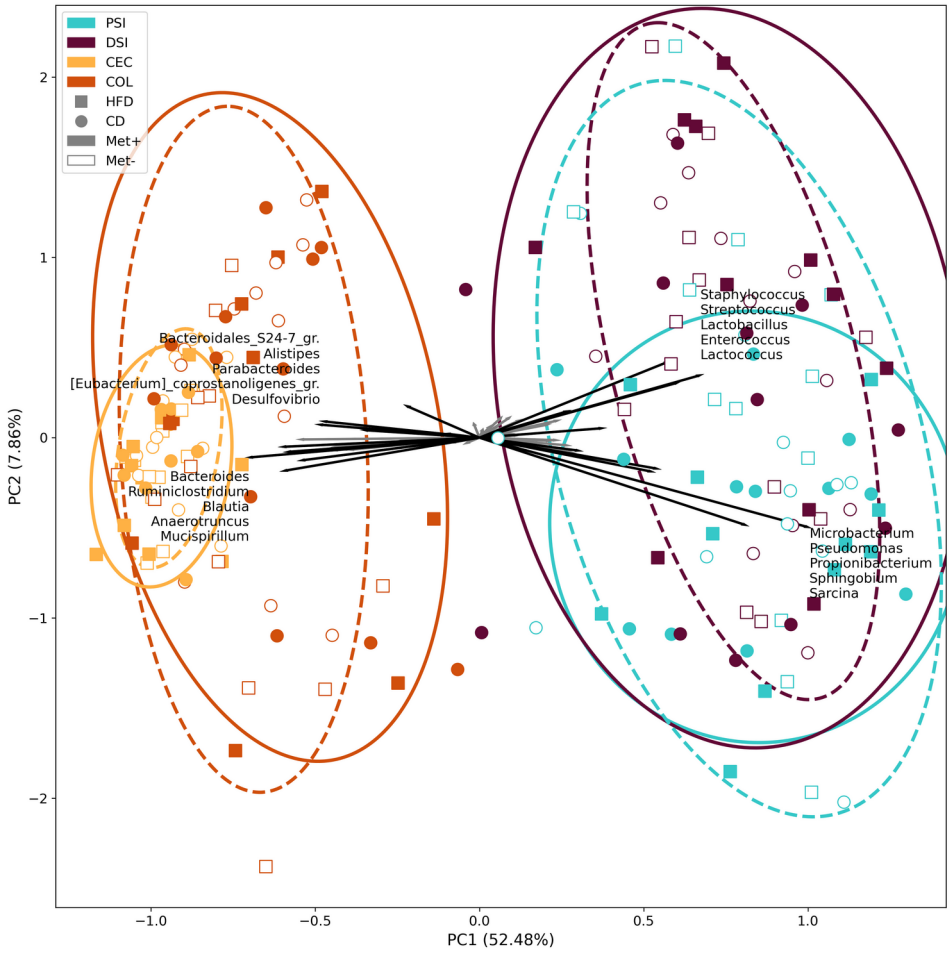

Figure 22. Beta diversity of all samples taken together was estimated using principal components analysis on centered log-ratio transformed values. Intestinal part, diet, and metformin treatment status are indicated. PSI – proximal small intestine; DSI – distal small intestine; CEC – cecum; COL – colon. In the corresponding intestinal part, continuous and dashed ellipses represent Met+ and Met- subsets, respectively.

A separate analysis of each intestinal part (Figure 23) revealed *Chryseobacterium*, *Propionibacterium*, *Corynebacterium_1*, and *Sphingobium* as the principal identifiers of metformin treatment in the proximal small intestine. The distal part was characterized by *Lactococcus*, *Streptococcus*, *Enterococcus*, *Proteus*, and *Staphylococcus* in CD-fed mice and *Blautia*, *Ruminiclostridium_9*, *Sphingobium*, *Bacteroides*, and *Roseburia* in HFD-fed mice. The main identifiers of metformin treatment in the cecum of HFD-fed mice were *Eubacterium* representatives, *Mucispirillum*, *Ruminococcaceae_UCG-003*, and *Streptococcus* and *Lachnospiraceae_FCS020_group*, *Ruminiclostridium* representatives, *Lactococcus*, and *Blautia* in CD-fed mice. At the same time, *Sphingobium*, *Curvibacter*, *Microbacterium*, *Pseudomonas*, and *Mucispirillum* were the strongest identifiers in the colon.

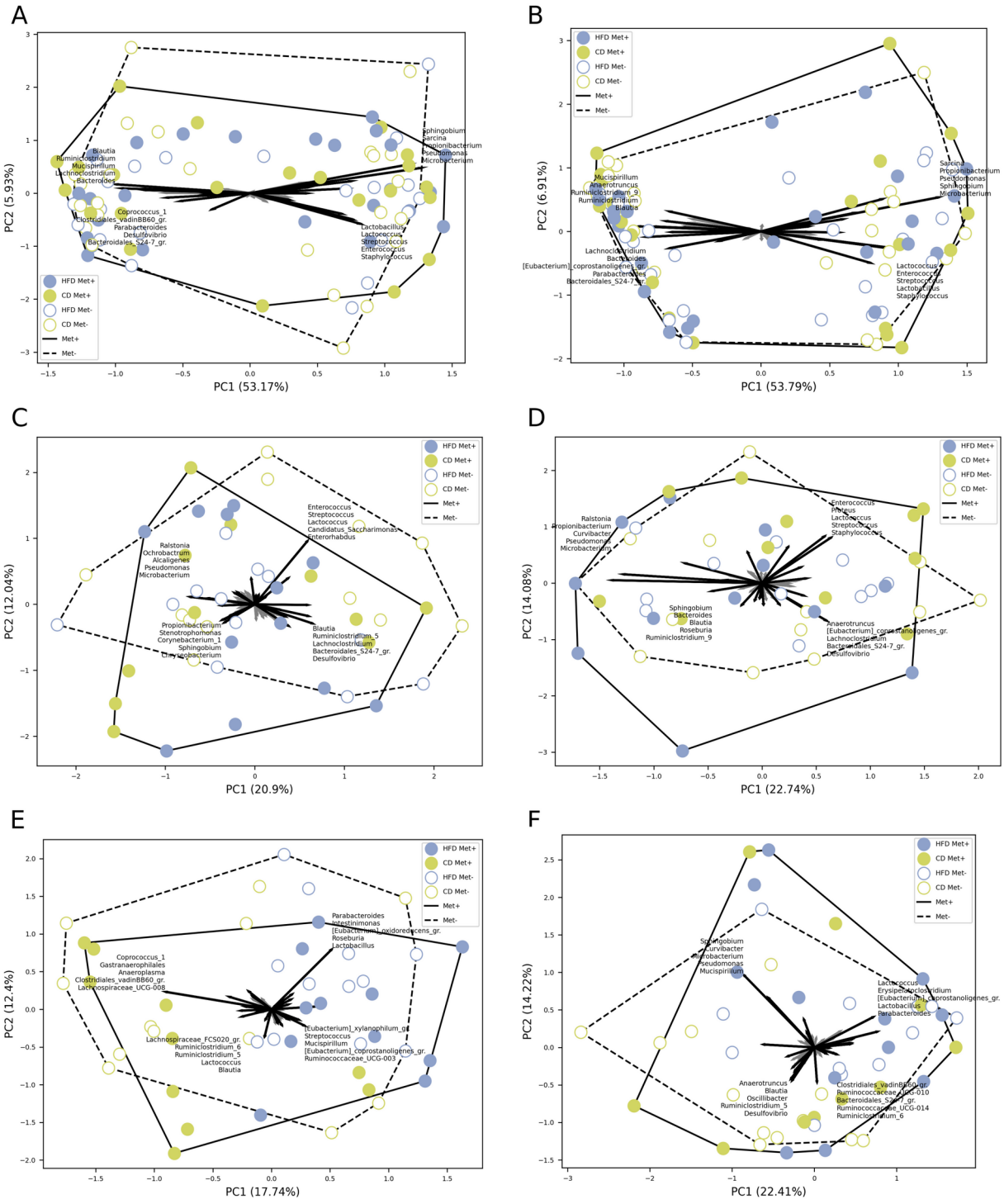

Figure 23. Beta diversity in each of the groups in response to 10 weeks long metformin treatment in each of the intestinal sites, n = 6 per group: (A) mucosa; (B) lumen; (C) proximal small intestine; (D) distal small intestine; (E) cecum; (F) colon.

3.8 Metformin treatment-mediated effects on the abundance of bacteria in different intestinal sites

When samples were contrasted regarding metformin treatment status (Met+ vs. Met- groups), up to 41 different genera were significantly differentially abundant in any of the parts of the small intestine; 28 genera in the cecum and 31 in the colon; and up to 11 genera in each of the intestinal layers studied (Figure 24).

Metformin treatment showed a varied effect in different sites of the intestine. The effect on the abundance of *Bacteroidales_S24-7_group* members was inversed in each layer – metformin increased the abundance of this genus in mucosa samples but decreased it in luminal content samples (Figures 24A and 24B). *Ruminiclostridium* increased in the mucosa and lumen, while *Lachnoclostridium* decreased in both layers. Several genera were significantly affected only in one of the layers. The abundance of *Roseburia* declined in luminal content samples. In contrast, the abundances of *Micrococcus*, *Pseudomonas*, and *Methylophilus* were significantly affected only in the mucosa.

Analysis of each of the intestinal parts separately revealed that metformin had a stronger effect on the abundance of the bacteria in the small intestine. Bacteria with increased abundances in response to metformin treatment only in the proximal small intestine include *Duganella*, *Chryseobacterium*, *Anaeroplasma*, *Undibacterium*, *Corynebacterium*, *Mucispirillum*, and *Methylophilus* (Figure 24C). In turn, *Eubacterium_halli_group* members were increased in the proximal small intestine but decreased in the cecum. Genera augmented uniquely in the distal part of the small intestine, include *Faecalibaculum* and *Tepidimonas*, while *Proteus* was depleted (Figure 24D). The abundance of *Roseburia* and *Micrococcus* was increased in both parts of the small intestine, while *Ochrobactrum* was decreased in these parts. *Pseudomonas* was affected in opposite directions in each part of the small intestine, with the abundance of the genus decreasing in the proximal part and increasing in the distal part. The most pronounced effect of metformin on the increased abundance was detected for *Lactococcus* in the distal small intestine (LogFC = 1.03 ± 0.41, FDR < 0.001), it was also increased in the colon, though to a lesser extent.

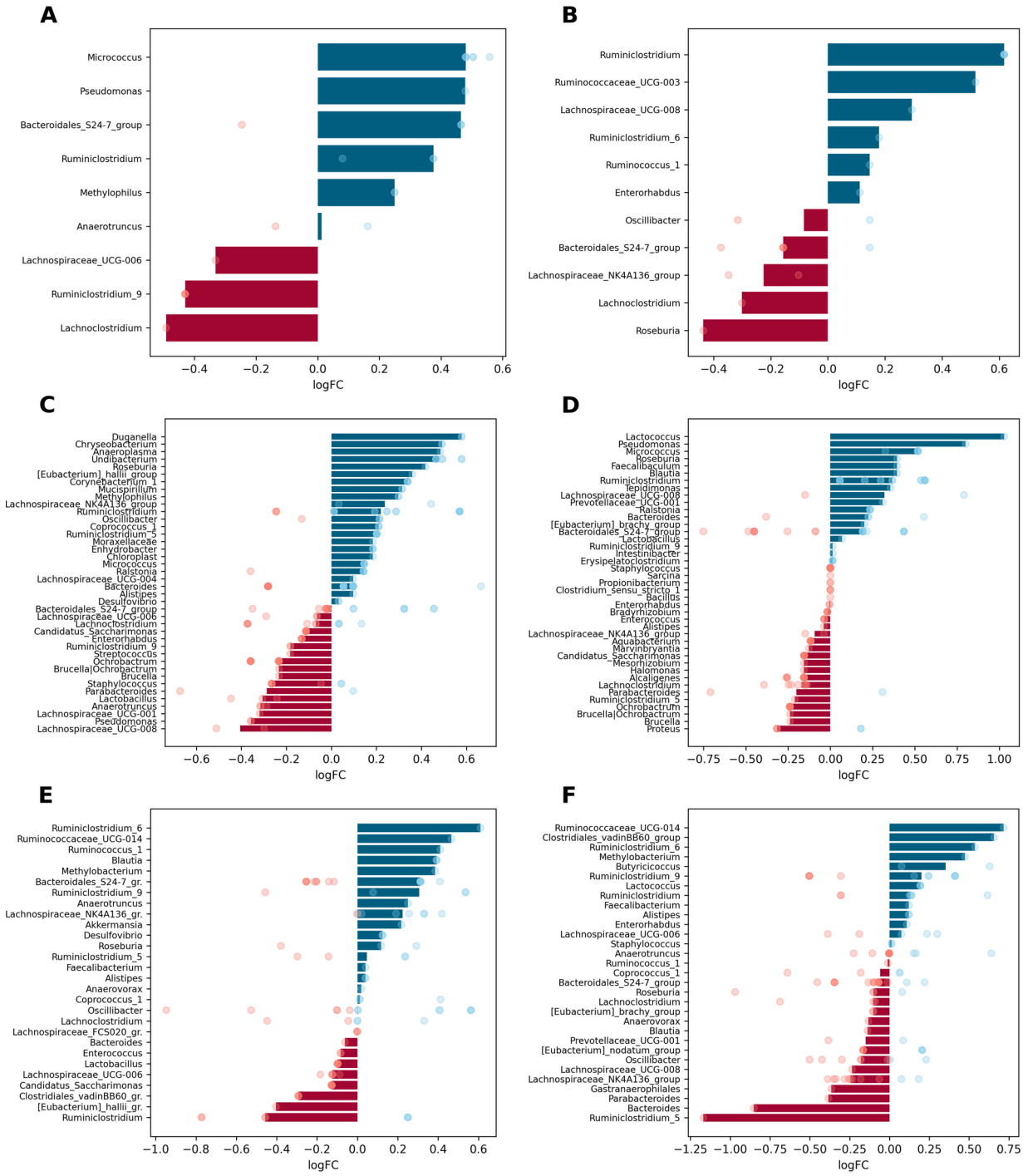

Figure 24. Differentially abundant genera in response to metformin treatment in different intestinal parts and layers (expressed as LogFC), n = 12 per group: (A) mucosa; (B) lumen; (C) proximal small intestine; (D) distal small intestine; (E) cecum; (F) colon. Blue bars represent genera with increased abundance among metformin users, and red bars – with decreased abundance. Dots of the corresponding color indicate all the individual features assigned to the genus.

Blautia was enriched in the distal part of the small intestine and cecum but diminished in the colon samples. *Ruminiclostridium* increased in all sites except cecum, where it was one of the most depleted genera (Figure 24E). Several genera were altered in the cecum and colon in opposite directions. *Methylobacterium* and *Ruminococcaceae_UCG-014* were enriched in both parts, while *Bacteroides* was decreased. In turn, *Clostridiales_vadinBB60_group* members were decreased in the cecum and increased in the colon. *Butyricicoccus* was significantly increased uniquely in the colon, but *Gastranaerophilales* was decreased (Figure 24F). *Ruminiclostridium_5* was the most reduced genus in response to metformin treatment in the colon (LogFC = -1.17 ± 0.40, FDR < 0.001), but it was increased in the proximal small intestine. The abundance of *Parabacteroides* was decreased in both parts of the small intestine and colon but was not affected in the cecum. A similar pattern was detected for *Lachnospiraceae_UCG-008*, the only difference being that it was augmented in the distal part of the small intestine.

3.9 Sex-related differences in the abundance of microbiome members in the intestinal sites studied

The effect of sex was evaluated by contrasting the samples from male mice of all experimental groups with corresponding samples from female mice at each intestinal site separately. In total, 11 genera in mucosal samples and 8 genera in luminal content samples were differentially abundant between males and females. When each of the intestinal parts was analyzed separately, 51 genera in the proximal small intestine; 40 genera in the distal small intestine; 28 genera in the cecum; and 30 genera in the colon were significantly differentially abundant between sexes (Figure 25).

Lactobacillus showed the most pronounced differences between sexes, with being decreased in males both in mucosal and luminal content samples (Figures 25A and 25B). Members of the *Bacteroidales_S24-7_group* were also depleted in both layers in males compared to females. In contrast, *Proteus* was increased in males in both mucosa and lumen, while *Staphylococcus* and *Ruminococcaceae_UCG-003* were increased, and *Ruminiclostridium_5* was decreased only in the lumen.

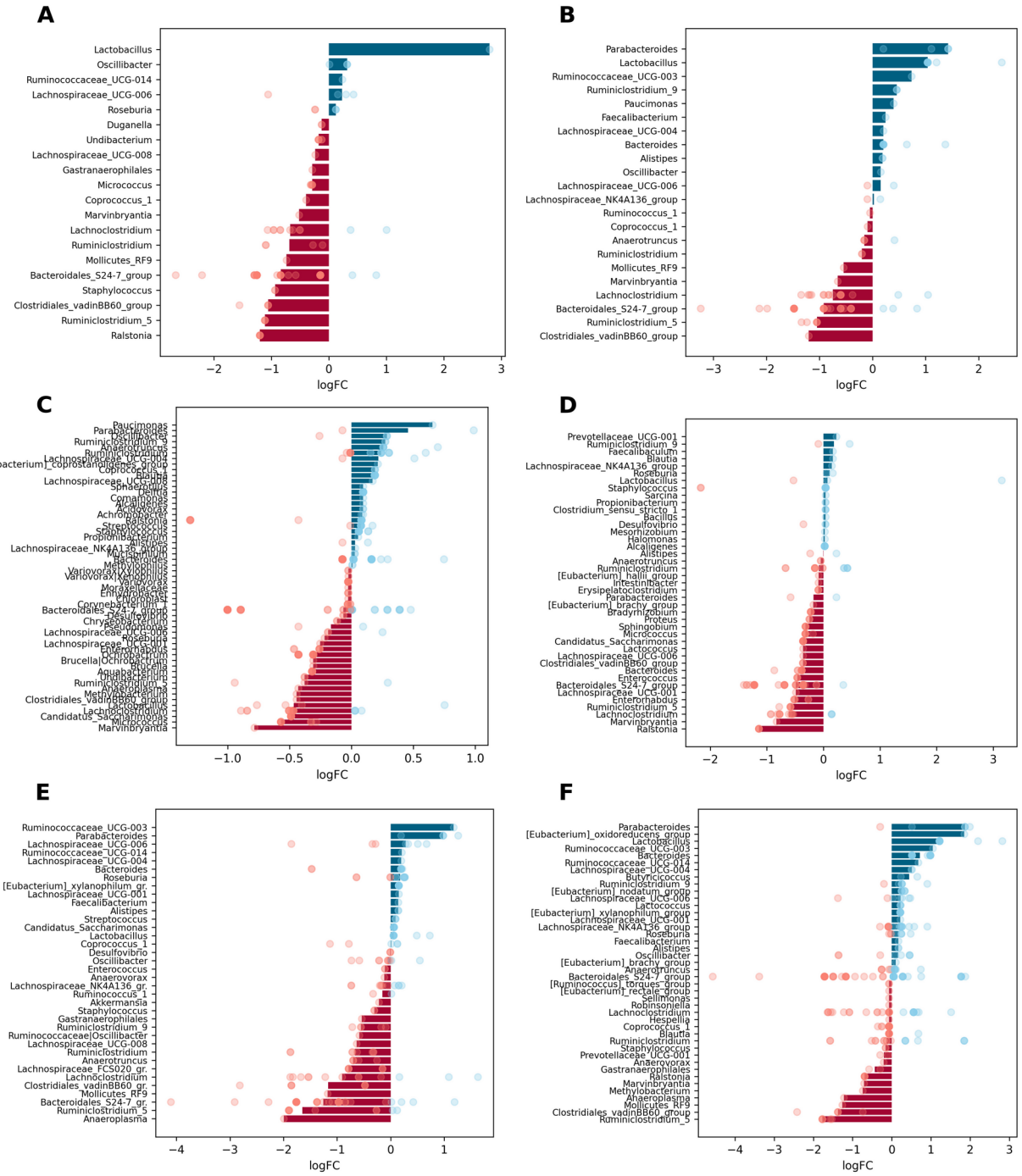

Figure 25. Differentially abundant genera between sexes in different intestinal parts and layers (expressed as LogFC), n = 12 per group: (A) mucosa; (B) lumen; (C) proximal small intestine; (D) distal small intestine; (E) cecum; (F) colon. Blue bars represent genera with increased relative abundance among males, and red bars – with decreased abundance. Dots of the corresponding color indicate all the individual features assigned to the genus.

The abundance of *Lactobacillus* was strongly reduced in all of the intestinal parts of males (distinct features in the cecum). Among the genera with significantly different abundance only in

the proximal small intestine, *Roseburia* was enriched, while *Chryseobacterium*, *Undibacterium*, and *Corynebacterium* were depleted in males relative to females (Figure 25C). *[Eubacterium]_hallii_group* was affected by the sex in the small intestine – in males, it was decreased in the proximal part but increased in the distal part of the small intestine; however, *Enterorhabdus* was enriched in both parts. *Lachnoclostridium* and *Ruminiclostridium_5* were increased in the distal part of the small intestine in males compared to females (Figure 25D). Meanwhile, the abundance of *Proteus* was increased in both the distal small intestine and colon. The abundance of *Anaerotruncus* was markedly decreased in the cecum of males (Figure 25E). *Ruminococcaceae_UCG-003* was increased in males only in the colon. *Parabacteroides*, *Staphylococcus*, and *Prevotellaceae_UCG-001* representatives were enriched in the cecum and colon in males compared to females (Figures 25E and 25F).

3.10 Diet-induced effects on the abundance of intestinal microbiome representatives

To assess dietary effects, samples from all HFD-fed mice were compared with corresponding samples from CD-fed mice at each intestinal site separately. Diet significantly affected the abundance of 20 genera in the mucosa, 22 genera in the lumen, 55 and 41 genera in the proximal and distal small intestine, respectively, 36 genera in the cecum, and 42 genera in the colon (Figure 26).

The abundance of *Lactobacillus* increased in HFD-fed mice in both studied layers LogFC = 2.79 ± 0.48, FDR < 0.001 and LogFC = 2.44 ± 0.53, FDR = 0.002 for mucosa and lumen, respectively (Figures 26A and 26B). In contrast, *Ruminiclostridium_5*, *Mollicutes_RF9*, *Lachnoclostridium*, *Marvinbryantia*, and *Clostridiales_vadinBB60_group* members were significantly lowered in both layers in response to HFD feeding. *Ralstonia* and *Staphylococcus* were decreased in HFD-fed mice solely in mucosa samples. *Parabacteroides* and *Ruminiclostridium_9* were significantly increased in the lumen of HFD-fed mice exclusively.

Uniquely to the proximal small intestine, the abundance of *Paucimonas* and *Oscillibacter* increased, while *Undibacterium* decreased in HFD-fed mice (Figure 26C). Genera, which significantly decreased in both parts of the small intestine in response to HFD feeding, include *Micrococcus*, *Enterorhabdus*, and *Candidatus_Saccharimonas* (Figures 26C and 26D).

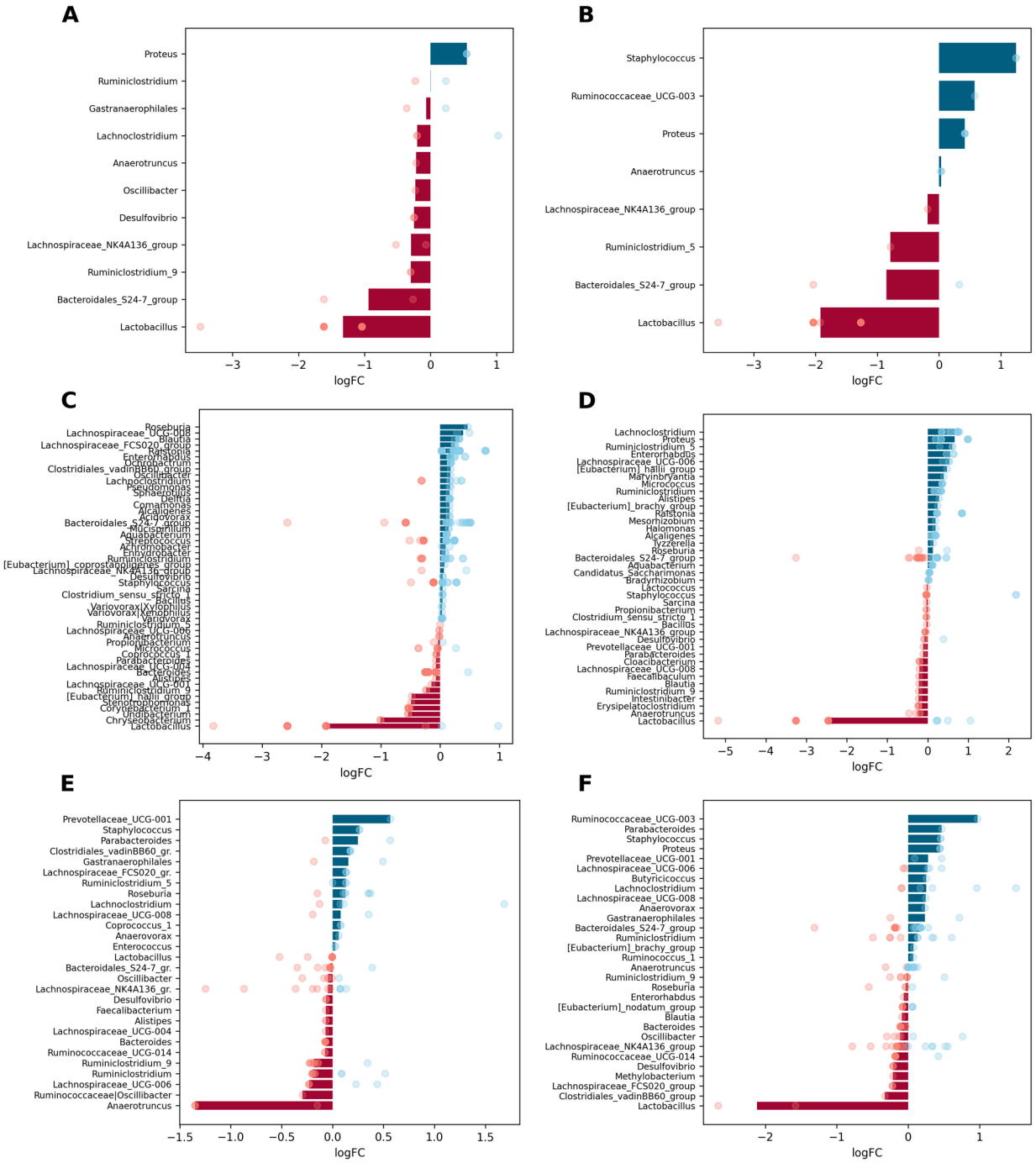

Figure 26. Differentially abundant genera between high-fat diet-fed and control diet-fed mice in different intestinal parts and layers (expressed as LogFC), n = 12 per group: (A) mucosa; (B) lumen; (C) proximal small intestine; (D) distal small intestine; (E) cecum; (F) colon. Blue bars represent genera with increased abundance among HFD-fed mice, and red bars – with decreased abundance. Dots of the corresponding color indicate all the individual features assigned to the genus.

Staphylococcus was depleted in the cecum of HFD-fed mice (Figure 26E). In turn, *Eubacterium_oxidoreducens_goup* members were increased only in the colon (Figure 26F). *Mollicutes_RF9* was decreased in both the cecum and colon of HFD-fed mice, while *Ruminococcaceae_UCG-003* was increased in these parts. The abundance of *Methylobacterium* decreased in the proximal small intestine and colon in response to HFD feeding. *Lactobacillus* was affected similarly in the proximal small intestine but increased in the colon. *Parabacteroides* was augmented mainly in the proximal small intestine and colon. *Ralstonia* and *Bacteroides* were both decreased in the distal small intestine, but in the colon these were decreased and increased, respectively. *Lachnoclostridium* decreased in all the parts up to the cecum. *Marvinbryantia* was reduced in both parts of the small intestine and colon, whereas *Anaeroplasma* was similarly affected in the proximal small intestine, cecum, and colon. In HFD-fed mice, representatives of two genera, *Ruminiclostridium_5* and *Clostridiales_vadin_BB60_group*, were decreased in all intestinal parts compared to CD-fed mice. As described above, these genera were also depleted in both intestinal layers of HFD-fed mice.

3.11 Differentially abundant bacteria between the mucosa and the lumen in all intestinal parts

The luminal and mucosal samples from all experimental groups were contrasted in each of the intestinal parts to investigate layer-related differences in microbiome composition. A total of 41 genera in the proximal small intestine, 48 – in the distal small intestine, 32 – in the cecum, and 38 – in the colon were significantly differentially abundant between the lumen and mucosa (Figure 27).

Differential abundance analysis between mucosa and lumen layers in each of the intestinal parts separately revealed substantial spatial variation of genera. Uniquely to the proximal small intestine, its lumen was depleted of *Corynebacterium_1* and *Micrococcus* (Figure 27A). Genera specifically increased in the lumen of the distal part of the small intestine include *Erysipelatoclostridium*, *Intestinibacter*, and *Marvinbryantia*. *Anaerotruncus* and *Lachnospiraceae_UCG-006* also were enriched in the lumen of the distal small intestine (Figure 27B).

Several genera were increased in the cecum and colon, including *Parabacteroides*, *Prevotellaceae_UCG-001*, *Methylobacterium*, and *Faecalibacterium* (Figures 27C and 27D). The abundance of *Mucispirillum* was lower in the lumen of the proximal small intestine and cecum.

Similarly, *Oscillibacter* was reduced in the lumen of the proximal small intestine but increased in the cecum. *Ruminiclostridium_5* and *Enterococcus* were enriched in the lumen of the distal small intestine and cecum. Genera with altered abundance, specifically in the distal small intestine and colon, include *Microbacterium* and *Propionibacterium*, with reduced abundance in the lumen of both parts; and *Lactococcus*, with increased abundance in the same sites.

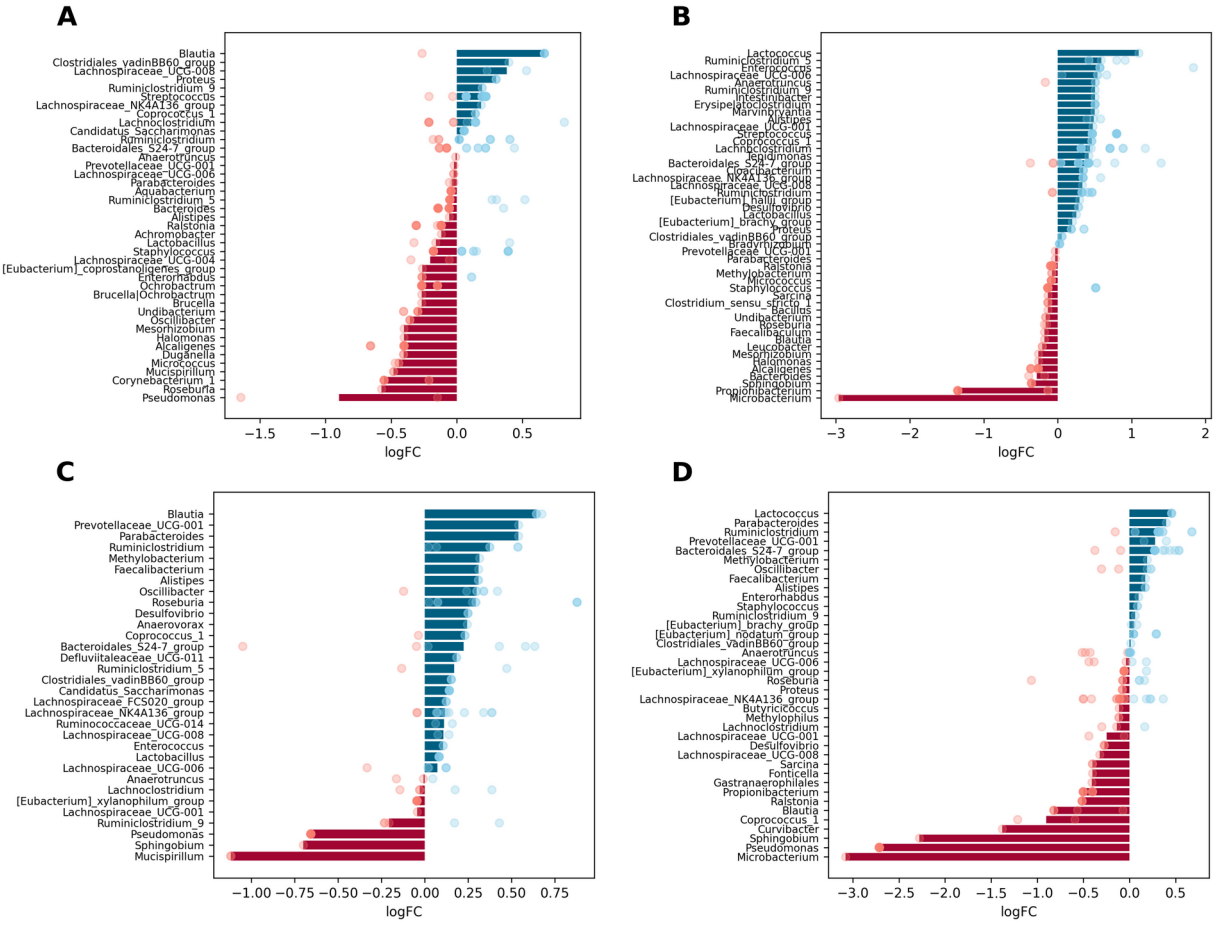

Figure 27. Differentially abundant genera between lumen and mucosa in different intestinal parts (expressed as LogFC), n = 12 per group: (A) proximal small intestine; (B) distal small intestine; (C) cecum; (D) colon. Blue bars represent genera with increased abundance in the lumen, and red bars – with decreased abundance. Dots of the corresponding color indicate all the individual features assigned to the genus.

Curvibacter was depleted solely in the lumen of the colon. *Roseburia* and *Ruminiclostridium_9* were altered in opposite directions in the same sites. *Roseburia* was decreased in both parts of the small intestine and increased in the cecum; in turn, *Ruminiclostridium_9* was depleted in the small intestine and enriched in the cecum. *Ruminiclostridium*, together with *Alistipes* and *Sphigobium*, were oppositely affected, the former

two being enriched in the lumen of the distal small intestine, cecum, and colon, while the latter was reduced in the same sites. *Coprococcus_1* and *Blautia* were differentially abundant between layers in all intestinal parts. *Blautia* was increased in the lumen of the proximal small intestine and cecum but decreased in the distal part of the small intestine and colon. *Coprococcus_1* was enriched in the lumen in all parts of the intestine except the colon, where the abundance of the genus was reduced.

3.12 Interaction of metformin treatment with diet type, sex, and intestinal layer

The genera with the most substantial differences in response to metformin treatment in each of the analyzed subsets were summarized in Figure 28. In total, we found 77 genera with a LogFC of at least 1 in any of the analysis subsets (Figure 28). Metformin affected representatives of all the main phyla found in the intestine – *Bacteroidetes*, *Firmicutes*, *Actinobacteria*, *Proteobacteria*, and *Verrucomicrobia*.

Members of *Betaproteobacteria*: *Paucimonas* and *Alcaligenes* were reduced only in the mucosal layer of both parts of the small intestine, while another representative *Curvibacter* was altered in opposite directions in different sites, including the cecum and colon in the mucosal layer but not in the luminal layer. *Deltaproteobacteria* member *Desulfovibrio* was affected by metformin almost in all studied sites. The most remarkable changes in abundance were observed in the proximal small intestine in both layers and the distal small intestine in the mucosal layer of HFD-fed mice. In addition, this genus showed marked sexual dimorphism in response to metformin treatment, with being increased in HFD-fed males in both parts of the small intestine and decreased in females in the mucosal layer and the proximal small intestine and cecum in the luminal layer. In CD-fed mice, *Desulfovibrio* was reduced in both sexes, though it was more pronounced in females. Sex-related differences in the proximal small intestine were also observed for *Pseudomonas* in mice fed both diet types, with being reduced in males and increased in females. The abundance of *Pseudomonas* was altered only in the mucosal layer, except for an increase in the lumen of the distal small intestine of HFD-fed males.

The most pronounced changes in the abundance of *Actinobacteria* members were found in both layers of the proximal small intestine; however, significant changes in at least one genus were observed in all intestinal parts. *Bacteroides* was increased in both sexes of HFD-fed mice in the lumen of both parts of the small intestine and decreased in the colon. In turn, in the colon of CD-fed mice, *Bacteroides* was increased. *Bacteroidales_S24-7 group* (*Muribaculaceae*) was altered in

almost all studied sites. The strongest reduction in the abundance of the genus was found in the mucosa of the distal small intestine of CD-fed males.

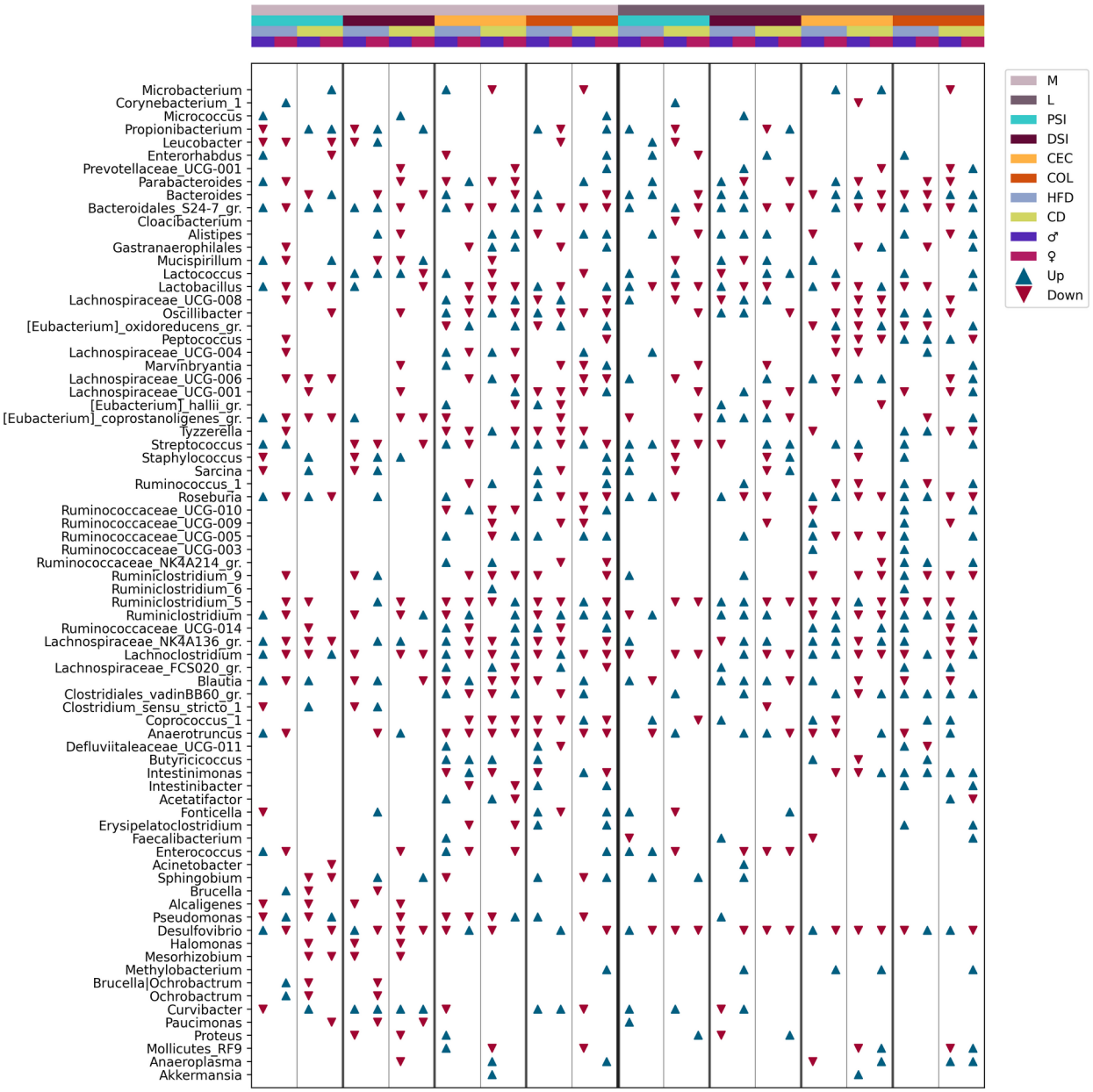

Figure 28. Summary of differentially abundant genera between Met+ and Met- mice in different subsets formed by various combinations of the levels of studied factors (expressed as LogFC), n = 3 per group. Only the genera with an absolute LogFC ≥ 1 in at least one of the subsets are included, indicating the directions of changes (with an absolute LogFC > 0.2), if detected, in all subsets. M – mucosa; L – lumen; PSI – proximal small intestine; DSI – distal small intestine; CEC – cecum; COL – colon. Blue triangle – an increase of the abundance among metformin users; red triangle – a decrease among metformin users.

The abundance of *Akkermansia* was significantly affected exclusively in CD-fed male cecum in both layers. *Mucispirillum* was strongly decreased in the mucosa of the proximal and distal small intestine of HFD-fed females, while it was slightly increased in males. In contrast, CD-fed females had marked increases in the genus at the exact locations. In the luminal layer, *Mucispirillum* was depleted in the proximal small intestine of CD-fed males and the distal small intestine of HFD-fed females, whereas the genus was increased in the distal small intestine of males fed both diet types and in HFD-fed males in the cecum.

A substantial interaction between metformin treatment and sex, diet type, and the intestinal layer was observed regarding the abundance pattern of *Bacilli* members. *Lactobacillus*, *Enterococcus*, *Staphylococcus*, and *Streptococcus*, opposite to *Lactococcus*, were markedly decreased in the lumen of the proximal small intestine of CD-fed males. In contrast, all genera together with *Lactococcus* were increased in HFD-fed males at the same site. *Lactococcus* was substantially increased in the distal small intestine of CD-fed mice of both sexes, while in HFD-fed males, it was reduced. *Lactococcus* was increased in HFD-fed males and CD-fed females in the lumen of the colon. In the mucosa, the abundance of *Lactococcus* was not affected by metformin in the proximal small intestine. However, it was increased in both sexes of HFD-fed mice and CD-fed males in the distal small intestine, whereas in females, it was reduced. The abundance of *Lactobacillus* was mainly reduced in response to metformin, though it was augmented in the mucosa of the proximal small intestine of HFD-fed males.

Intestinimonas and *Clostridiales_vadinBB60_group* members were affected only in the cecum and colon in the mucosa layer. A similar pattern was observed in the lumen, except that *Clostridiales_vadinBB60_group* was strongly increased in CD-fed males in the proximal small intestine and HFD-fed females in the distal small intestine. *Butyricicoccus* was affected similarly, mainly being augmented only in the mucosa of the cecum of male mice and the colon of HFD-fed males. In the lumen of the cecum, HFD-fed male mice were enriched in *Butyricicoccus*, while in CD-fed males genus was reduced. *Butyricicoccus* was increased in HFD-fed females in the lumen of the colon but unchanged in other subsets. Other *Clostridiaceae* representatives were not affected in the cecum and colon in the luminal layer.

Genera representing families *Lachnospiraceae*, *Peptococcaceae*, *Peptostreptococcaceae*, and *Ruminococcaceae* were affected more in the cecum and colon in both layers. In contrast to this observation, *Eubacterium_coprostanoligenes_group* members were more affected in both parts of the small intestine in both layers. In general, members of the families mentioned above were

enriched in the mucosa of the proximal small intestine of HFD-fed males but decreased in HFD-fed females. Other features common to these genera (with a few exceptions) are a decrease in the lumen of the proximal small intestine and cecum of CD-fed mice; and enrichment in the distal small intestine and colon (predominantly males) of HFD-fed mice. *Roseburia* showed strong sex-related differences in the abundance changes in response to metformin treatment. The abundance of the genus was increased in the mucosa of the proximal small intestine of male mice fed both types of diet and decreased in females. *Roseburia* was enriched only in HFD-fed males in the cecum but unchanged in other subsets. Similarly, it was increased in the colon of HFD-fed males, and a decrease was also found in other subsets, CD-fed mice and HFD-fed females. In the lumen, *Roseburia* was increased in HFD-fed mice of both sexes in all intestinal parts (except a decrease in females in the distal small intestine) and decreased in CD-fed mice of both sexes in the cecum and colon and both parts of the small intestine in males only.

Study III: Host miRNAs are associated with gut microbiome composition in high-fat diet-induced type 2 diabetes mouse model of both sexes

3.13 MiRNA composition in feces

A total of 48 fecal samples representing time points before and after ten weeks-long metformin treatment of 24 experimental units included in the animal experiment were used to determine the miRNA profile by miRNA-seq. These samples corresponded to the experimental units used for metagenome sequencing in fecal samples described in Study I of the thesis. The median value of the obtained single-end reads was 2220516. After quality trimming, a median of 443840 reads was retained. The median percentage of annotated reads with miRBase (*Mus musculus*) was 0.61%. The median percentage of mapped reads, when mapped against the *Mus musculus* genome, was 69.53%.

After aligning reads to miRBase, 240 known mature miRNAs were detected by at least one read when samples are pooled together. Of the identified known miRNAs, 49 had only one read count in any of the samples and 47 had at least 100 read counts. The top 20 miRNAs represented 81.7% of the total read counts. The relative abundances of the top 20 miRNAs in each of the experimental groups at the time points before and after the metformin treatment in HFD-fed and CD-fed mice are shown in Figures 29 and 30, respectively.

In all groups mmu-miR-192-5p dominated the miRNA composition, followed by mmu-miR-21a-5p, mmu-miR-200a-3p, and mmu-miR-29a-3p. All of the top 20 miRNAs were represented in each of the experimental groups of each diet type. Differences in the composition of miRNAs depending on the diet status were observed for mmu-let-7f-5p, mmu-let-7c-5p, mmu-miR-26a-5p, mmu-miR-101a-3p, and mmu-miR-30a-5p, which were found among the top 20 miRNAs only in HFD mice. The top 20 miRNAs unique for CD-fed groups include mmu-miR-215-5p, mmu-miR-429-3p, mmu-let-7i-5p, mmu-miR-196b-5p, mmu-miR-26b-5p, and mmu-miR-203-3p.

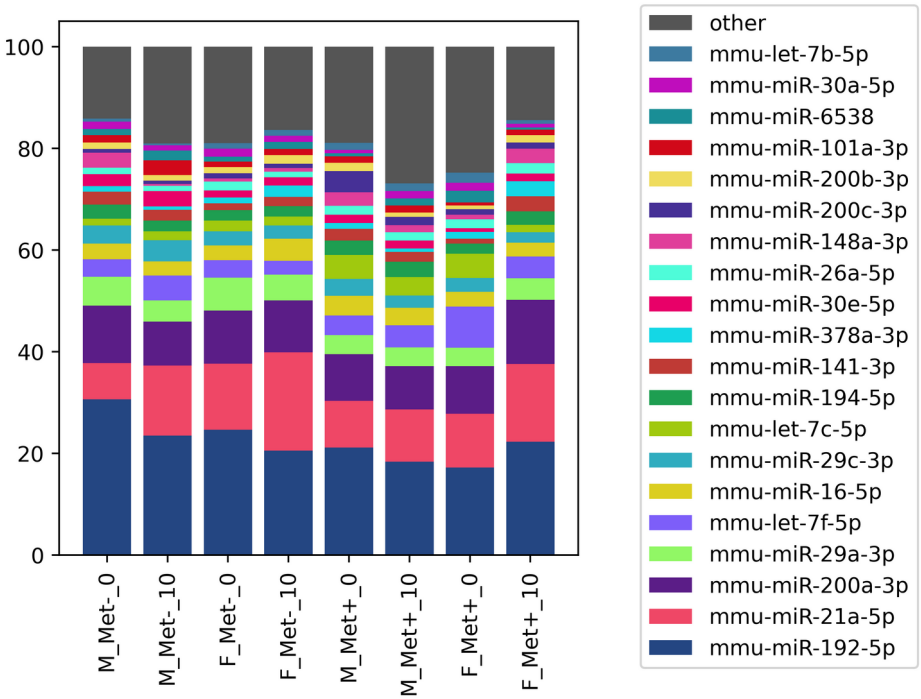

Figure 29. Mean percentage of the 20 most abundant fecal miRNAs in each of the experimental groups of high-fat diet-fed mice before and after metformin therapy. n = 3 in each of the groups.

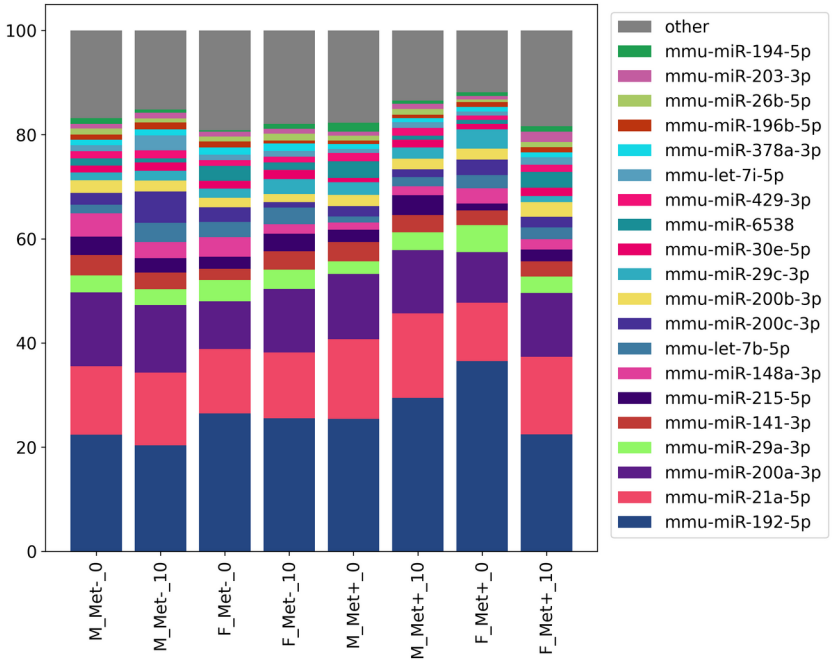

Figure 30. Mean percentage of the 20 most abundant fecal miRNAs in each of the experimental groups of control diet-fed mice before and after metformin therapy. n = 3 in each of the groups.

3.14 MiRNA composition in gut mucosa

In total, 48 samples representing the proximal small intestine and the cecum at the mucosal layer collected from 24 mice were sequenced by miRNA-seq. These samples corresponded to the experimental units used for microbiome determination with the 16S rRNA gene sequencing method in the respective intestinal samples described in Study II of the thesis. The median value of the obtained single-end reads was 20698971. The median percentage of annotated reads with miRBase (*Mus musculus*) and against the *Mus musculus* genome was 0.83% and 42.73%, respectively. After aligning reads to miRBase, 296 known mature miRNAs were detected by at least one read when samples are pooled together in the proximal small intestine, while in the cecum 237 miRNAs were identified.

Of the identified known miRNAs in the proximal small intestine, 117 had less than 100 reads in all samples, and 179 had at least 100 reads in any of the samples. The top 20 miRNAs represented 81.6% of the total read counts. One sample was excluded from the further analysis due to low read count. In the cecum, 115 miRNAs had less than 100 reads in all samples, and 122 had at least 100 reads in any of the samples. The top 20 miRNAs comprised 69.3% of the read counts taken together. Two of this subset of samples were excluded from the future analysis as these had insufficient miRNA read counts.

The relative abundances of the top 20 miRNAs in each of the experimental groups in the mucosa of the proximal gut are shown in Figure 31. In all groups, mmu-miR-215-5p prevailed, except for HFD_F_Met+ and HFD_F_Met-, where mmu-miR-192-5p, followed by mmu-miR-215-5p, dominated the miRNA composition. In all groups, mmu-miR-192-5p and mmu-miR-194-5p were the second and the third most abundant miRNA, respectively (except for HFD_F_Met-, where the third most abundant miRNA was mmu-miR-200c-3p). Overall, the miRNA composition of the proximal gut is relatively similar between experimental groups.

In cecum, all groups were dominated by mmu-miR-192-5p, followed by mmu-miR-143-3p and mmu-miR-21a-5p in different order (Figure 32). In CD_F_Met- group mmu-let-7c-5p, mmu-let-7b-5p, and mmu-let-7f-5p were among the most abundant miRNAs. The composition of the top 20 miRNAs is relatively comparable between the groups. However, it is more diverse than the one observed in the proximal gut, thus each of the miRNAs represents smaller proportion of the whole composition.

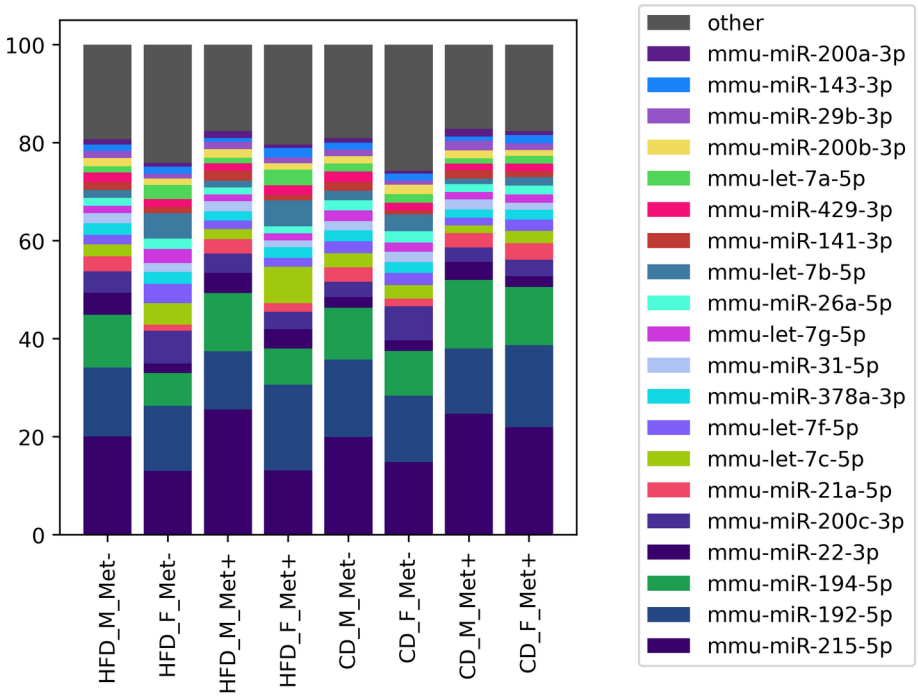

Figure 31. Mean percentage of the 20 most abundant miRNAs in the mucosa of the proximal small intestine. n = 3 in each of the groups.

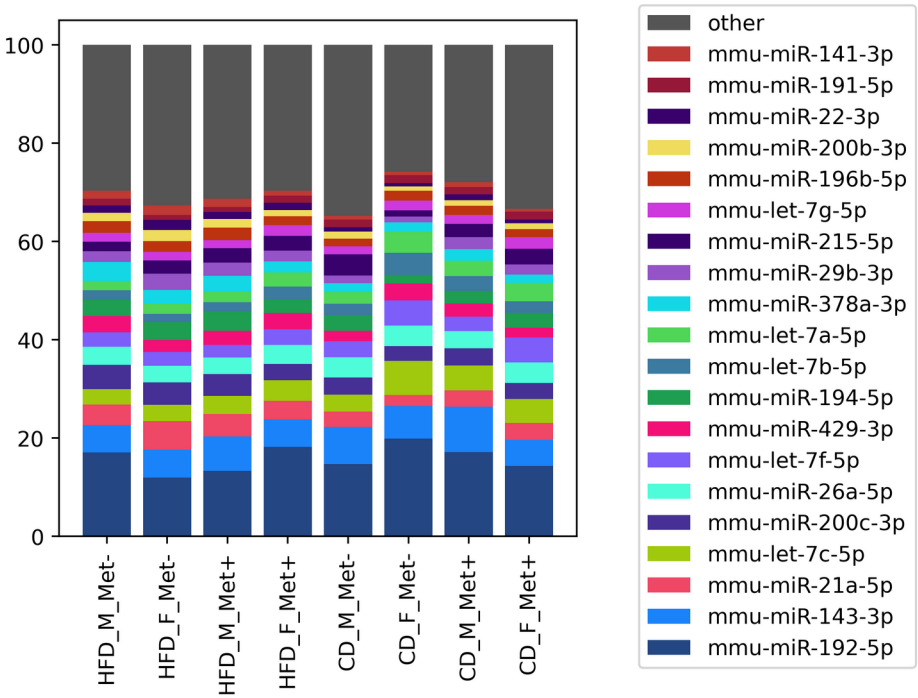

Figure 32. Mean percentage of the 20 most abundant miRNAs in the mucosa of the cecum. n = 3 in each of the groups.

3.15 Differentially expressed miRNAs

We evaluated the relative abundance of different miRNAs between experimental groups using different contrasts in fecal and gut mucosal samples representing the proximal small intestine and cecum. Contrasts for fecal sample analysis and differentially expressed miRNAs are shown in Figure 33 and Table 2.

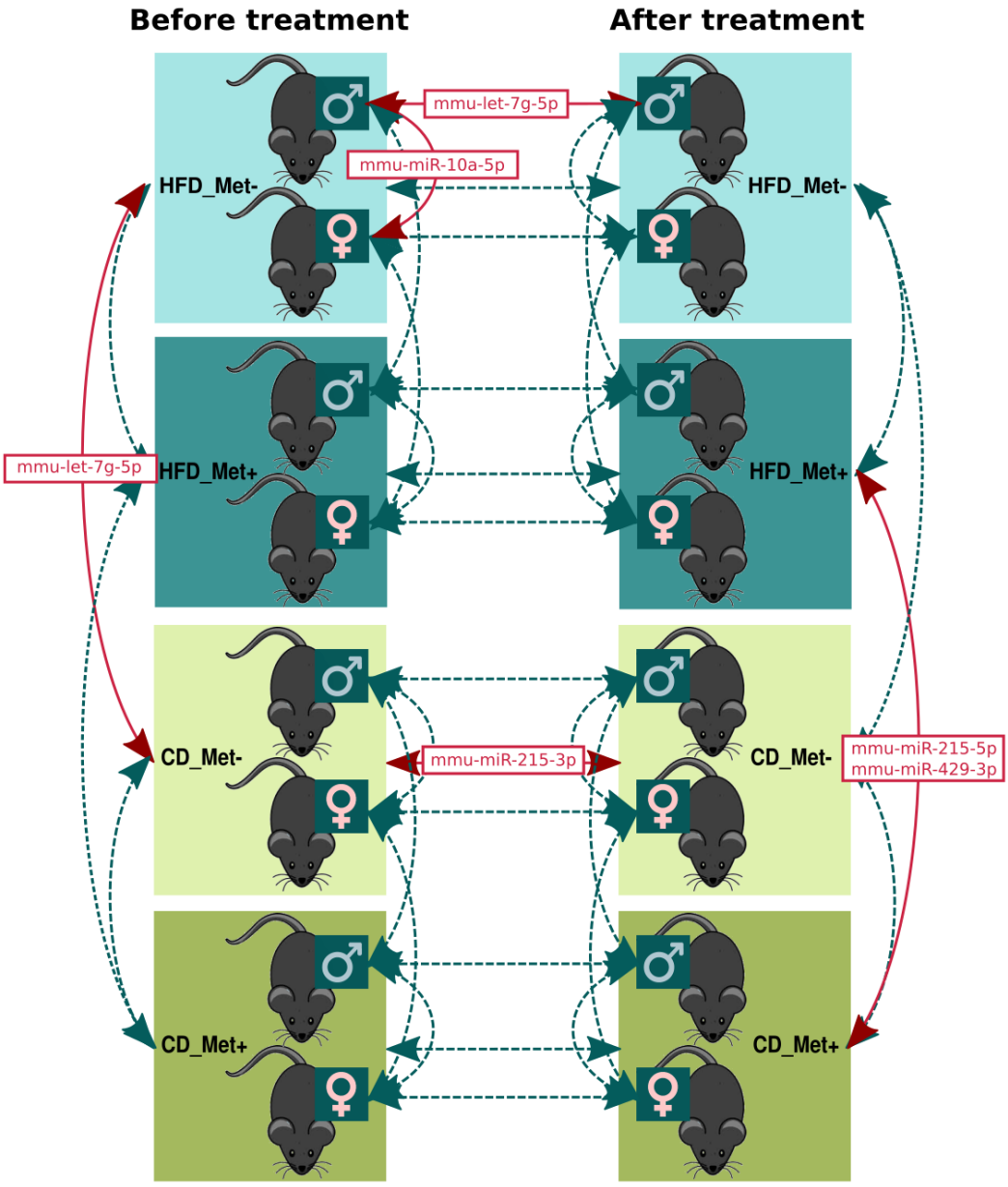

Figure 33. Contrasts in which fecal miRNA compositions were compared. Dashed lines indicate the contrasts between which a comparison was performed. Red bold lines indicate the contrasts between which statistically significant differences in miRNA expression were detected; the significantly differentially expressed (adjusted $p < 0.05$) miRNAs are shown for each comparison.

In fecal samples, mmu-let-7g-5p was depleted in HFD-fed mice compared to CD-fed ones; however, it showed a significant increase in HFD-fed male mice group without metformin treatment over time. Similarly, mmu-miR-215-3p increased over time in CD-fed mice not receiving metformin treatment. Mmu-miR-10a-5p was significantly increased in female HFD-fed mice compared to males at the time point before the initiation of metformin treatment. At the time point after the metformin treatment, mmu-miR-215-5p and mmu-miR-429-3p were significantly depleted in HFD-fed mice compared to CD-fed ones in the groups receiving the treatment. None of the other studied contrasts detected significant differences in miRNA composition, indicating that metformin treatment has no effect on the miRNA composition in studied groups.

Table 2. Differentially expressed miRNAs in fecal samples before (0) **and after** (10) **the metformin treatment**, N = 48. Comparisons in which statistically significant differences have been detected are shown. FDR – false discovery rate.

Compared groups	miRNA	logFC	FDR
HFD_M_Met-_10 vs HFD_M_Met-_0	mmu-let-7g-5p	3.64	0.02
CD_Met-_10 vs CD_Met-_0	mmu-miR-215-3p	3.69	0.03
HFD_M_Met-_0 vs HFD_F_Met-_0	mmu-miR-10a-5p	-4.97	0.02
HFD_Met-_0 vs CD_Met-_0	mmu-let-7g-5p	-4.91	0.01
HFD_Met+_10 vs CD_Met+_10	mmu-miR-215-5p	-3.11	0.02
	mmu-miR-429-3p	-2.78	0.03

As for gut mucosal samples, only cross-sectional contrasts were feasible due to the collection of the samples at one time point – upon the termination of the experiment. Main comparisons aimed to investigate the differences between intestinal parts, the effect of metformin treatment, sex, and diet type, similar to fecal samples. Significantly differentially expressed miRNAs are summarized in Table 3. In the proximal small intestine, we found three differentially expressed miRNAs between HFD-fed males and females after receiving metformin treatment. All of these miRNAs (mmu-miR-217-5p, mmu-miR-99a-5p, and mmu-miR-216a-5p) were depleted in male mice. In the cecum of HFD-fed mice not receiving metformin, mmu-miR-193a-3p was augmented compared to CD-fed mice, while mmu-miR-712-5p and mmu-miR-714 were strongly decreased in mice fed an HFD.

Table 3. Differentially expressed miRNAs in mucosal samples of proximal small intestine (PSI) **or cecum** (CEC)**.** N = 48. Comparisons in which statistically significant differences have been detected are shown. FDR – false discovery rate.

Compared groups	miRNA	logFC	FDR
HFD_M_Met+_PSI vs HFD_F_Met+_PSI	mmu-miR-217-5p	-6.59	0.003
	mmu-miR-99a-5p	-2.89	0.003
	mmu-miR-216a-5p	-5.33	0.01
HFD_Met-_CEC vs CD_Met-_CEC	mmu-miR-712-5p	-7.99	0.003
	mmu-miR-193a-3p	3.45	0.04
	mmu-miR-714	-7.06	0.04

3.16 MiRNA correlation with the gut microbiome

3.16.1 Correlation of fecal miRNAs with the fecal microbiome

We performed a correlation analysis between miRNAs and previously identified gut microbiome at the genus level independent of metformin treatment status using the sequencing data obtained from fecal samples. Due to the significant differences in gut microbiome composition observed between high-fat diet and control diet-fed mice in the previous analysis, correlation analysis was performed for each of these subgroups separately.

Analysis in high-fat diet-fed mice (Figure 34) revealed a very strong positive correlation between miR-8 family members (mmu-miR-200a-3p, mmu-miR-200b-3p, mmu-miR-200c-3p, mmu-miR-141-3p, and mmu-miR-429-3p), miR-378 family members (mmu-miR-378c and mmu-miR-378d), and mmu-miR-23a-3p, and the abundance of *Akkermansia muciniphila*, and *Lachnospiraceae* members (*Acetatifactor*, *Roseburia*, *Eubacterium*), *Erysipelotrichaceae* representative, and *Oscillospiraceae* members (*Oscillibacter* and *Lawsonibacter*). Likewise, the abundance of these genera was very strongly negatively correlated with another set of miRNAs containing mmu-miR-5119, mmu-miR-6239, mmu-miR-6238, and mmu-miR-6240.

A set of miRNAs, including mmu-miR-101a-3p, mmu-miR-194-5p, mmu-miR-30e-5p, and mmu-miR-340-5p showed a strong negative correlation with *Bacteroides*, *Paramuribaculum*, *Christensenella*, *Pseudomonas*, and *Eggerthellaceae* representative. These miRNAs were, in turn, positively correlated with *Duncaniella* and *Schaedlerella*, among others. An opposite correlation pattern with the genera mentioned above was observed for another miRNA set containing mmu-miR-191-5p, mmu-miR-375-3p, and mmu-miR-183-5p as the key representatives.

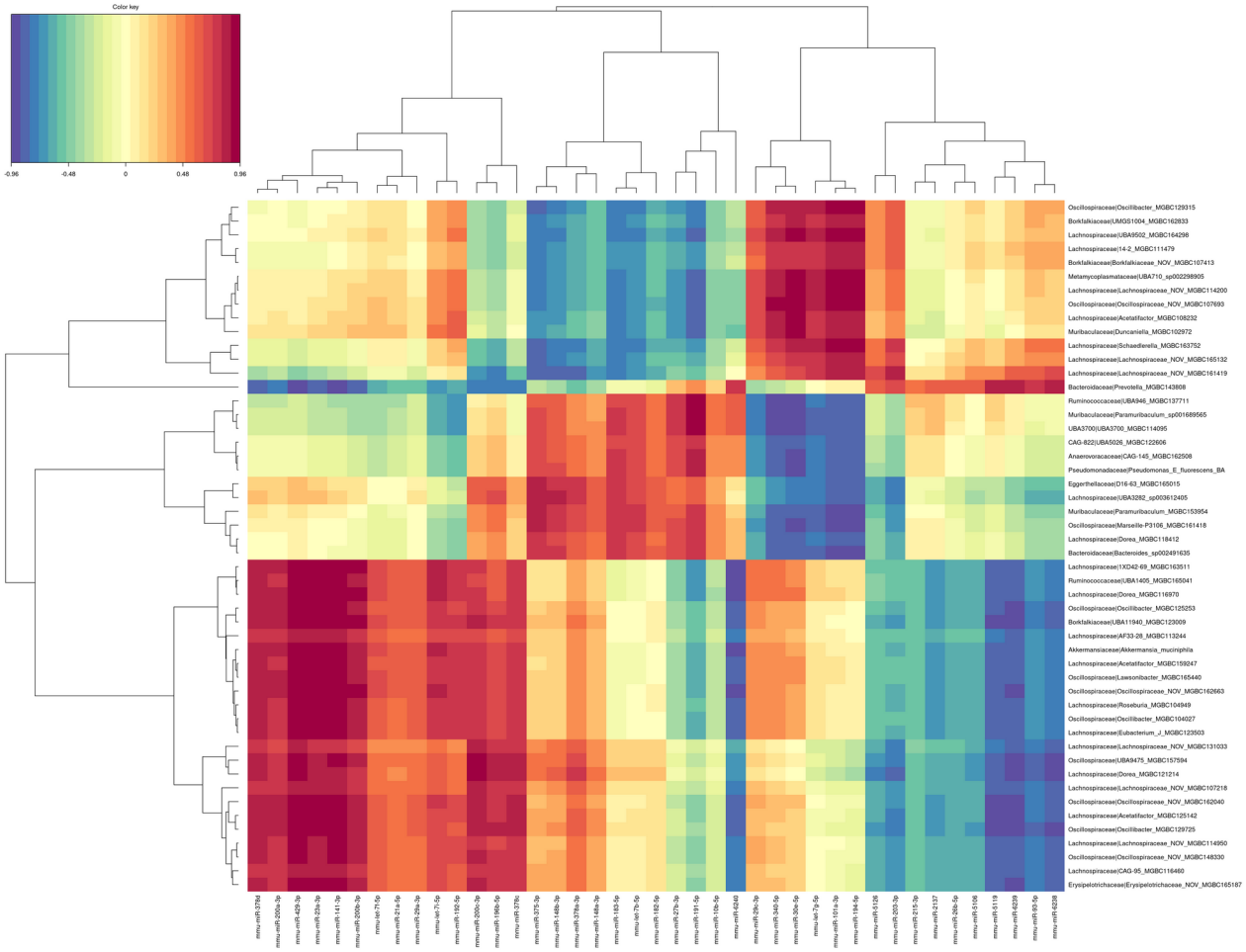

Figure 34. Correlation analysis between fecal miRNAs and gut microbiome representatives of high-fat diet-fed mice identified by miRNA-seq and shotgun metagenomic sequencing, respectively. N = 24 of each of the datasets. The color key indicates the correlation levels, with blue and red color denoting negative and positive correlation, respectively.

In control diet-fed mice (Figure 35), mmu-miR-200c-3p showed the strongest positive correlation with members of *Borkfalkiaceae*, while the correlation was negative with *Roseburia* and several unclassified members of *Lachnospiraceae*. Mmu-miR-194-5p was detected as the most prominent representative with a correlation in the opposite direction.

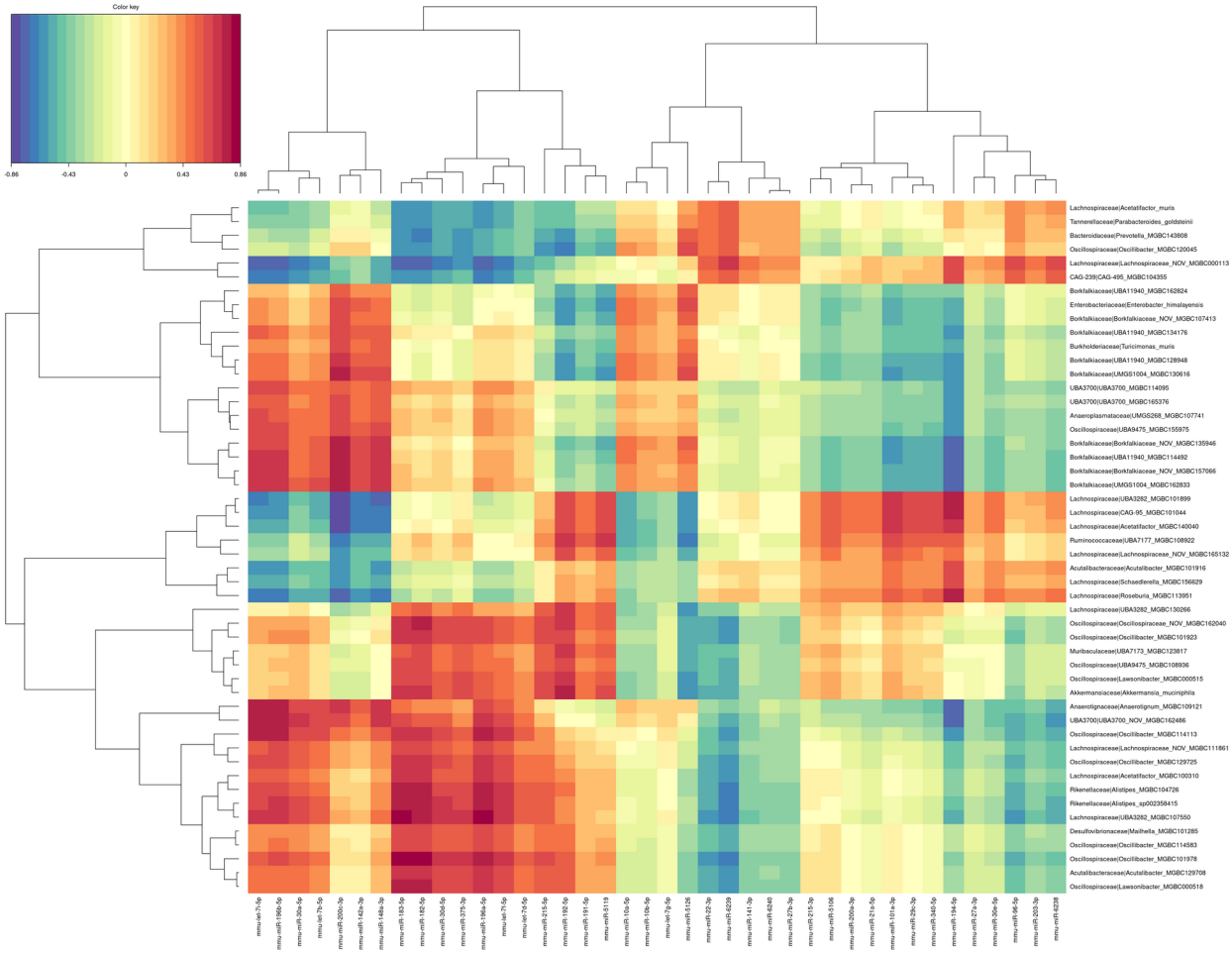

Figure 35. Correlation analysis between fecal miRNAs and gut microbiome representatives of control diet-fed mice identified by miRNA-seq and shotgun metagenomic sequencing, respectively. N = 24 of each of the datasets. The color key indicates the correlation levels, with blue and red color denoting negative and positive correlation, respectively.

3.16.2 Correlation of gut mucosal miRNAs with the gut microbiome

To assess the potential differences in the correlation pattern between various intestinal sites and to expand the range of identified potential correlations, we performed a correlation analysis using miRNAs isolated from gut mucosa samples collected from the proximal small intestine and the cecum and 16S rRNA gene sequencing data obtained both from the mucosal and the luminal microbiome of the respective sites. Due to the 16S rRNA gene sequencing data analysis algorithm, some features can be attributed to different genera with equal probability; therefore, these features show an equal correlation with the same miRNA.

Analysis of samples collected from the proximal small intestine showed a distinct pattern when correlating miRNAs with mucosal or luminal microbiome members. In the analysis subset

with the mucosal microbiome (Figure 36), the strongest negative correlation was observed between mmu-miR-101b-3p, mmu-miR-18a-5p, mmu-miR-19a-3p, mmu-miR-20a-5p, among others, and the abundance of *Lactobacillus*. In turn, mmu-miR-676-3p and mmu-miR-205-5p showed a positive correlation with *Lactobacillus*. *Pseudomonas* was another genus, the abundance of which was strongly correlated with different sets of miRNAs. A set containing mmu-miR-497a-5p, mmu-miR-223-3p, and both strands of mmu-miR-126a-3p as the strongest representatives correlated positively, while a set containing mmu-miR-186-3p as the strongest hit correlated negatively with the abundance of *Pseudomonas*.

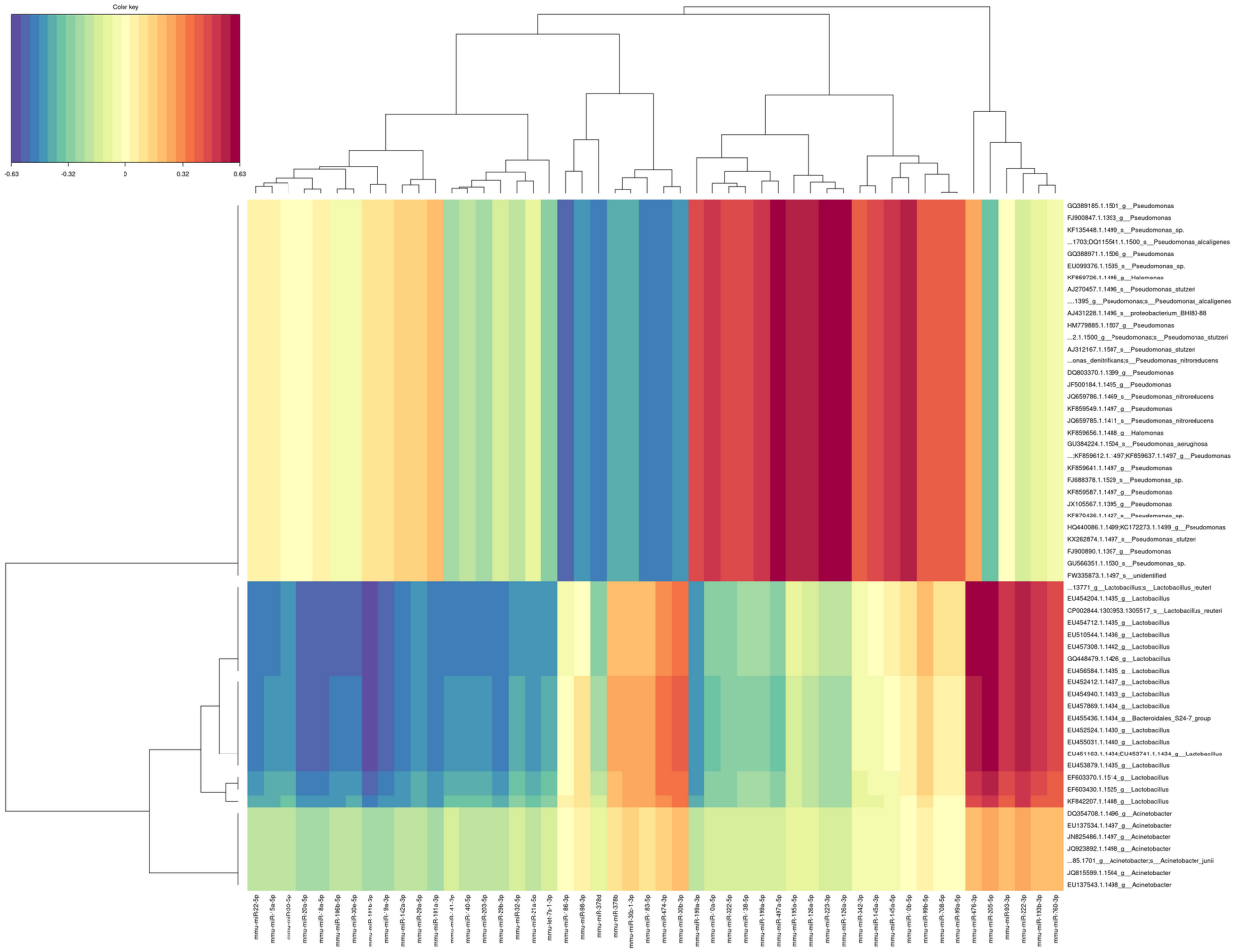

Figure 36. Correlation analysis between mucosal miRNAs and gut microbiome representatives of the mucosa of the proximal small intestine identified by miRNA-seq and 16S rRNA gene sequencing, respectively. N = 24 of each of the datasets. The color key indicates the correlation levels, with blue and red color denoting negative and positive correlation, respectively.

In the analysis subset with the luminal microbiome (Figure 37), the same genera were represented among the correlated ones in addition to *Bacteroidales S24-7 group* members, *Lachnoclostridium*, *Propionibacterium*, and *Microbacterium*. The strongest positive correlation was observed between mmu-miR-222-3p and the abundance of *Lactobacillus* and mmu-miR-136-5p and *Pseudomonas*. In turn, the strongest negative correlation was detected between a set of miRNAs containing mmu-miR-101b-3p and mmu-miR-20a-5p and the abundance of *Lactobacillus*.

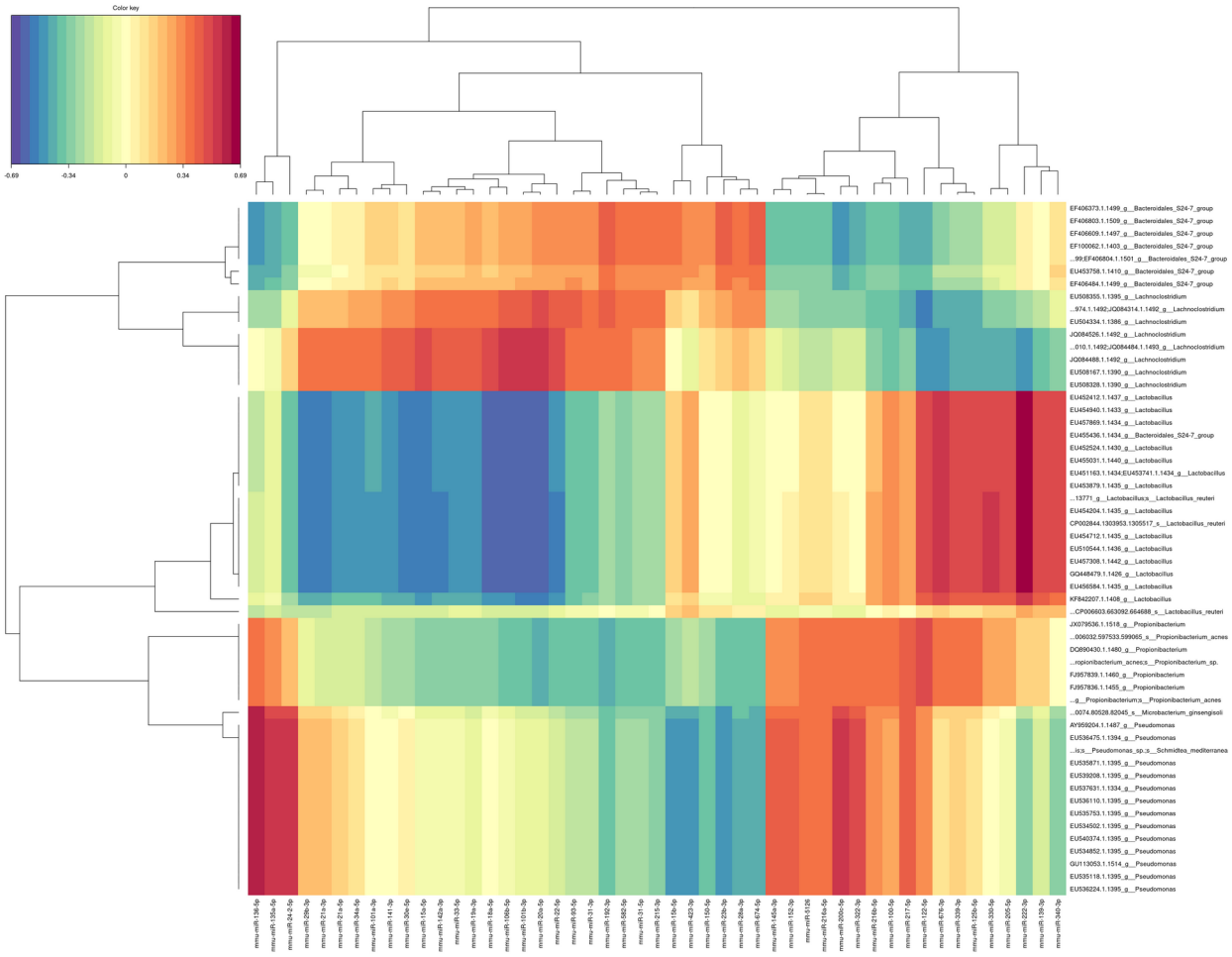

Figure 37. Correlation analysis between mucosal miRNAs and gut microbiome representatives of the lumen of the proximal small intestine identified by miRNA-seq and 16S rRNA gene sequencing, respectively. N = 24 of each of the datasets. The color key indicates the correlation levels, with blue and red color denoting negative and positive correlation, respectively.

Analysis of samples collected from the cecum revealed more similar correlation pictures when analyzing the mucosal and luminal microbiome than in the proximal small intestine. In the

mucosal microbiome analysis subset *Bacteroidales S24-7 group* members showed a strongest correlation with different sets of miRNAs (Figure 38). Two different sets of miRNAs – a set including let-7 family members and a second set containing mmu-miR-203-3p as the strongest representative were negatively correlated with *Bacteroidales S24-7 group*. In turn, miRNAs including mmu-miR-30e-5p and mmu-miR-6236 were correlated with the genus positively. Other bacteria which showed a correlation with miRNAs in this analysis subset showed moderate correlation.

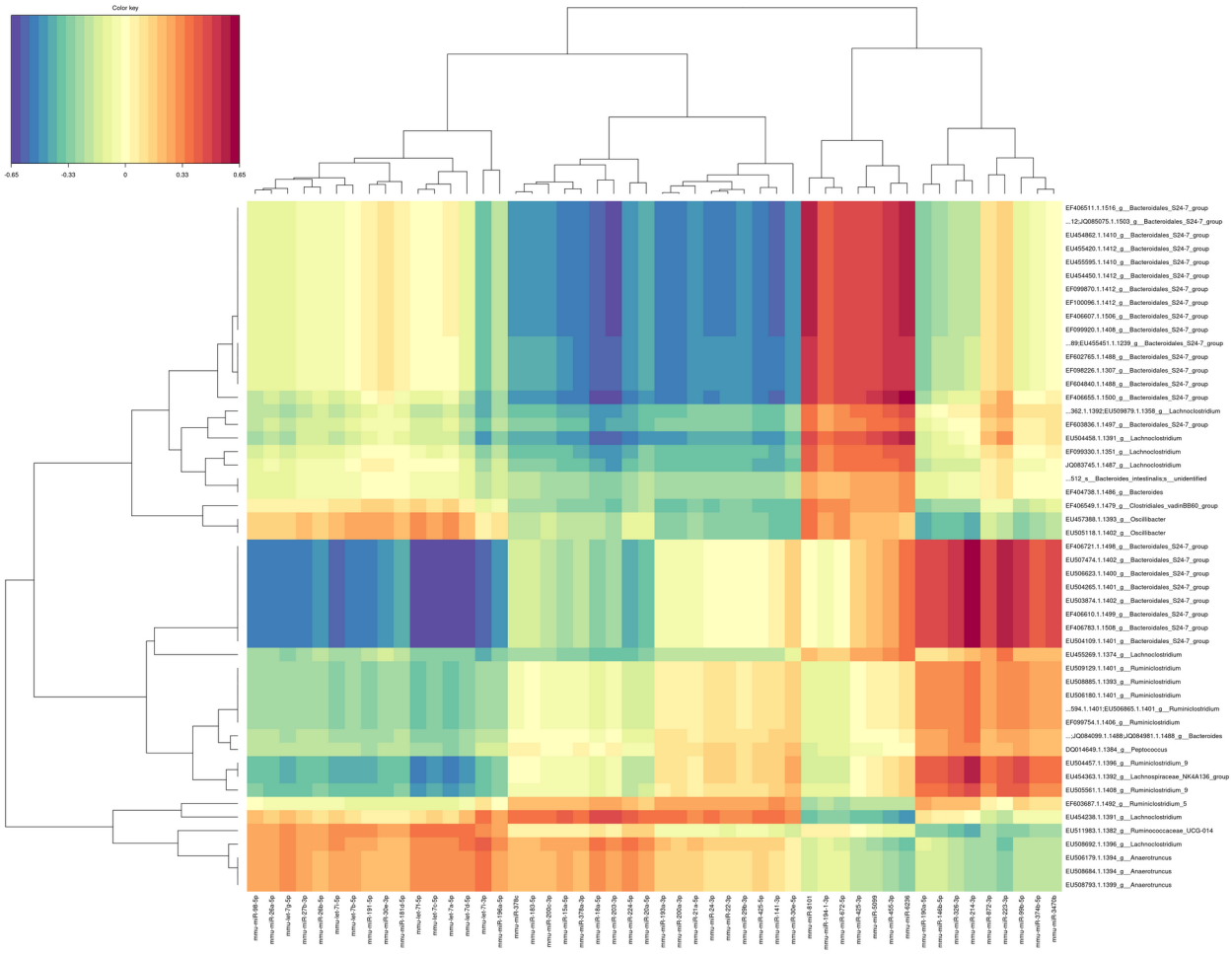

Figure 38. Correlation analysis between mucosal miRNAs and gut microbiome representatives of the mucosa of the cecum identified by miRNA-seq and 16S rRNA gene sequencing, respectively. N = 24 of each of the datasets. The color key indicates the correlation levels, with blue and red color denoting negative and positive correlation, respectively.

Luminal microbiome analysis subset showed two main miRNA sets with a different correlation pattern with bacterial genera (Figure 39). Similar to the analysis with the mucosal

microbiome, the strongest correlation was observed between different sets of miRNAs and *Bacteroidales S24-7 group*. A set of miRNAs including members of miR-8 family (mmu-miR-200a-3p, mmu-miR-200b-3p, mmu-miR-141-3p), miR-30 family (mmu-miR-30b-5p, mmu-miR-30c-5p, mmu-miR-30d-5p and mmu-miR-30e-5p), miR-10 family (mmu-miR-10a-5p, mmu-miR-10b-5p, mmu-miR-99a-5p, and mmu-miR-99b-5p), mmu-miR-21a-5p, mmu-miR-22-3p, and mmu-miR-24-3p, among others showed a negative correlation with *Bacteroidales S24-7 group*, *Bacteroides*, and *Lachnoclostridium*. Other miRNA set containing mmu-miR-714, mmu-miR-6236, mmu-miR-30d-3p, and mmu-miR-8101 was strongly positively correlated with *Bacteroidales S24-7 group* and *Lachnoclostridium*.

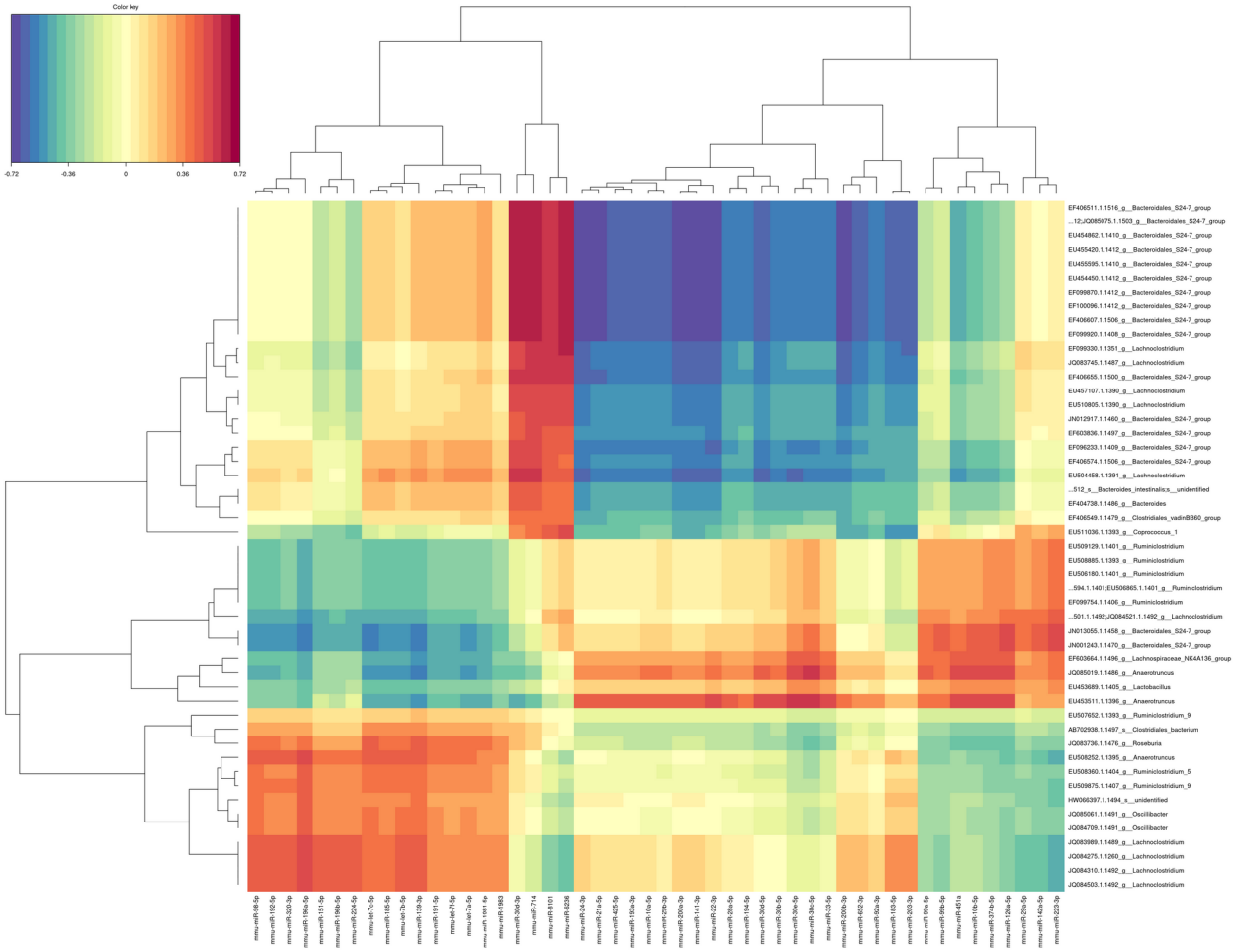

Figure 39. Correlation analysis between mucosal miRNAs and gut microbiome representatives of the lumen of the cecum identified by miRNA-seq and 16S rRNA gene sequencing, respectively. N = 24 of each of the datasets. The color key indicates the correlation levels, with blue and red color denoting negative and positive correlation, respectively.

A moderate positive correlation was observed between a miRNA set including let-7 family members (mmu-let-7a-5p, mmu-let-7b-5p, mmu-let-7c-5p, and mmu-let-7f-5p), mmu-miR-196a-5p, mmu-miR-139-3p, and mmu-miR-192-5p and the abundance of *Roseburia*, *Oscillibacter*, *Ruminiclostridium 5*, and *Ruminiclostridium 9*. In turn, this set was negatively correlated with *Ruminiclostridium* and *Lactobacillus*; however, the correlation with *Anaerotruncus*, *Lachnoclostridium*, and *Bacteroidales S24-7 group* for these miRNAs was not consistent.

3.17 Potential bacterial targets of the correlated miRNAs in mice feces

To investigate the potential of the correlated miRNAs to biologically target the respective bacteria, the homology of all correlated fecal miRNA sequences with the sequenced metagenome data were analyzed. In addition, the thermodynamic stability of the miRNA-mRNA duplex, expressed as the free energy of binding between the miRNA and the target site, was evaluated.

Homology assessment allowing a maximum of three mismatches revealed a median value of 7816.5 (IQR 4217.0) miRNA-read hits per sample (reads from each pair counted independently). This set of pairs was further tested based on the minimum free energy, setting a threshold of minimum free energy (MFE) values ≤ -20.0 kcal/mol. In total, this analysis retrieved 31 miRNAs and 68 microbial genera between which at least one potential interaction exists. The top six miRNAs and their target microorganisms are shown in Figure 40.

Homology and minimum free energy value assessment revealed that for mmu-miR-6538, at least 20 potential targets exist, with the largest number of hits against *Oscillospiraceae_NOV*, *UBA9475* (*Oscillospiraceae* member), and *Oscillibacter*. Other *Oscillospiraceae* members include *Lawsonibacter*, *UMGS1872*, *Evtepia*, and *Flavonifractor*. *Lachnospiraceae* was represented by *Acetatifactor*, *Kineothrix*, and *Lachnospiraceae_NOV*, while *Angelakisella* is affiliated with *Ruminococcaceae*. *Bacteroidales* was represented by members of four families – *Prevotellaceae* (*CAG-873*), *Muribaculaceae* (*UBA7173*, *CAG-485*, *Duncaniella*), *Tannerellaceae* (*Parabacteroides*), and *Rikenellaceae* (*Alistipes*). In turn, phylum *Actinobacteriota* was manifested in a form of *Adlercreutzia*. *Bacteroides* is potentially targeted by mmu-miR-5119 and mmu-miR-215-5p. However, *Lactococcus* was only targeted by mmu-miR-5106.

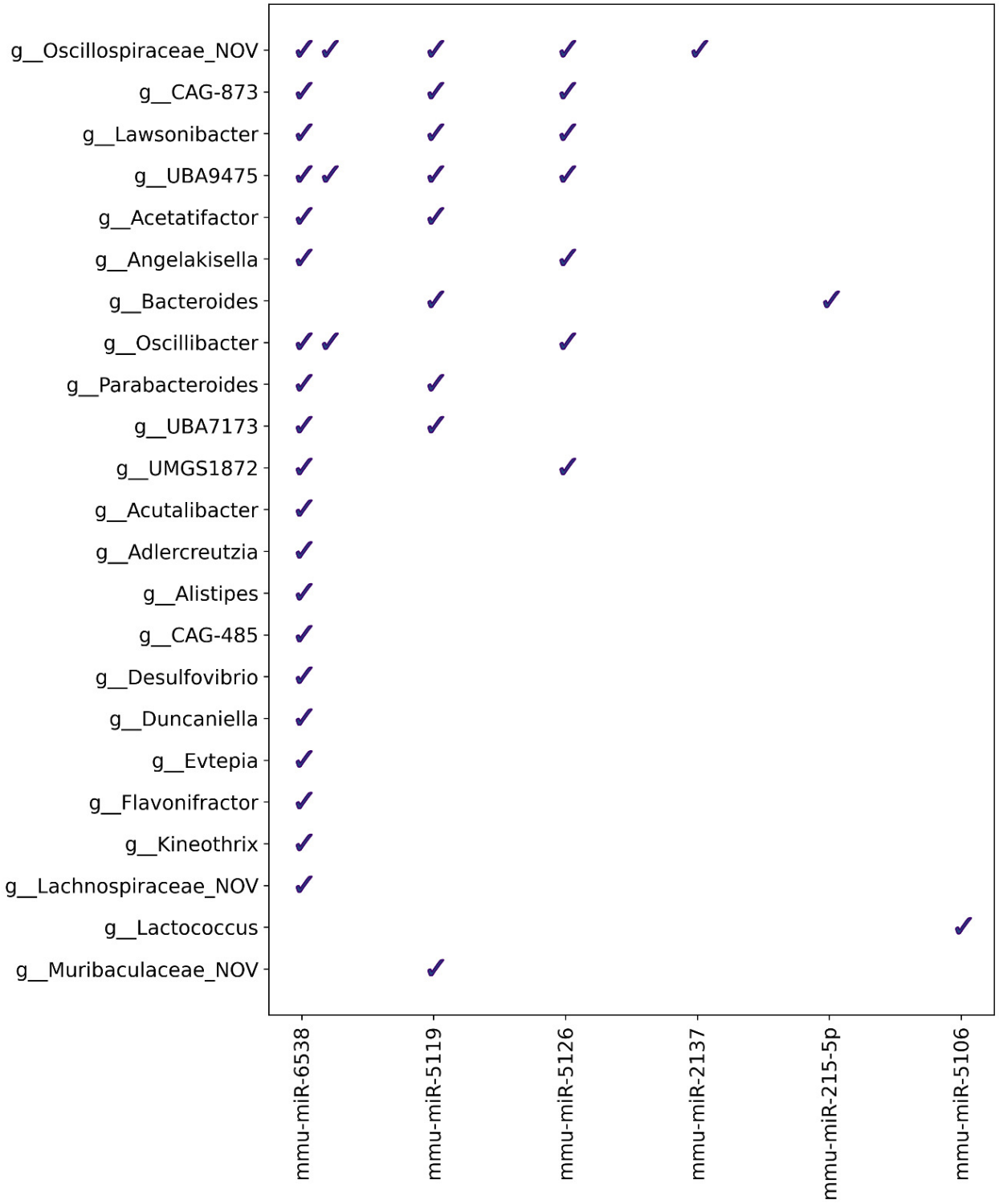

Figure 40. List of fecal miRNAs that have homology with bacterial genomes in mice. One tick indicates that at least 100 reads have been detected for the pair. Two ticks indicate that at least 1000 reads have been found for the pair.

Taken together with a correlation analysis results, *Oscillospiraceae* members (*Clostridia* class), including *Oscillospiraceae_NOV*, *Lawsonibacter*, and *UBA9475*, showed a negative

correlation with mmu-miR-5126 both in high-fat diet- and control diet-fed mice, while with mmu-miR-5119 the correlation was negative in high-fat diet-fed mice and positive in control diet-fed mice for all three bacterial genera. Opposite results were obtained for *CAG-873*, affiliated with *Prevotella*, a genus from *Bacteroidia* class. It was positively correlated with mmu-miR-5126 in mice fed with both types of diet. In contrast, the correlation with mmu-miR-5119 was positive in high-fat diet-fed mice and negative in control diet-fed mice. In the correlation analysis of the control diet-fed mice subset, mmu-miR-5126 represented a distinct miRNA set, which showed an opposite correlation pattern than other above-mentioned miRNAs. In contrast, in the high-fat diet-fed mice, all of these miRNAs were clustered in one miRNA set.

Study IV: Intestinal miRNAs are associated with the presence of colorectal adenomas and are correlated with the abundance of gut bacteria in patients undergoing colonoscopy

3.18 MiRNA composition in different sample types – pilot analysis

First, to investigate the homogeneity of miRNA compositions between different types of samples, a pilot analysis using a small subset of the recruited samples was performed (Figure 41). All analyzed samples were collected from the same three individuals. The study revealed that the most diverse miRNA population can be detected in intestinal biopsies, in addition to its inter-sample homogeneity. Therefore, intestinal biopsies were chosen as a sample type of primary interest for further analysis.

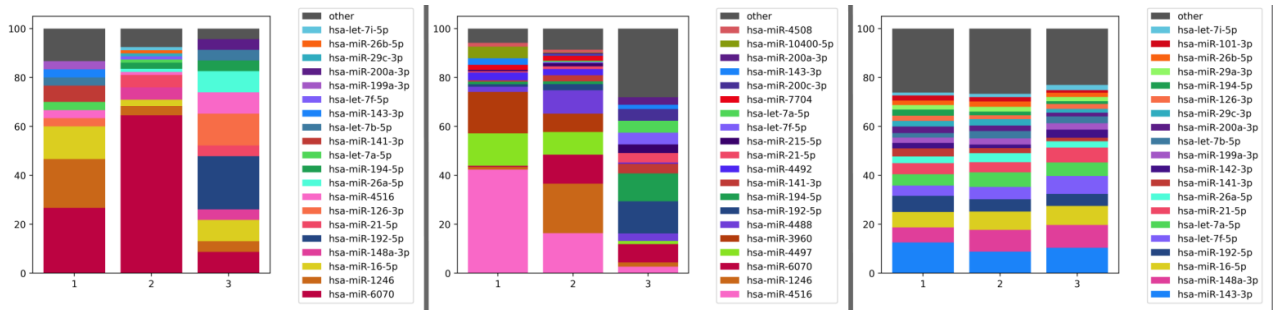

Figure 41. MiRNA composition in different sample types: (A) feces; (B) luminal contents; (C) intestinal biopsies. N = 3 in each dataset.

3.19 Intestinal miRNA composition

Intestinal miRNA composition was determined in a total of 43 patients undergoing routine colonoscopy procedures for various reasons. Patients were divided into two study groups – patients with colorectal adenomas detected by an experienced colonoscopist (n = 20) and patients without colorectal adenomas (n = 23). The main characteristics of the patient cohort are summarized in Table 4.

Table 4. Characteristics of the analyzed patient cohort.

Characteristic	With colorectal adenomas, n = 20	Without colorectal adenomas, n = 23
Males/females, n (%)	11 (55 %) /9 (45 %)	6 (26.1 %) / 17 (73.9 %)
Age (years), mean ± SD	57.1 ± 13.0	47.4 ± 13.2

The median of obtained single-end reads was 3475739. After annotating the reads against the miRBase, 982 mature miRNAs with at least 1 read were found. Of the miRNAs identified, 77 had at least 1 read and 12 had at least 100 reads in any of the samples. The top 20 miRNAs accounted for 71.91% of the total reads.

In both groups, hsa-miR-192-5p was the dominant miRNA, followed by hsa-miR-21-5p and hsa-miR-16-5p (Figure 42). All of the top 20 miRNAs were found in each of the experimental groups, and their relative abundances were similar, except for lower abundances of hsa-miR-126-3p and hsa-miR-101-3p in patients with colorectal adenomas.

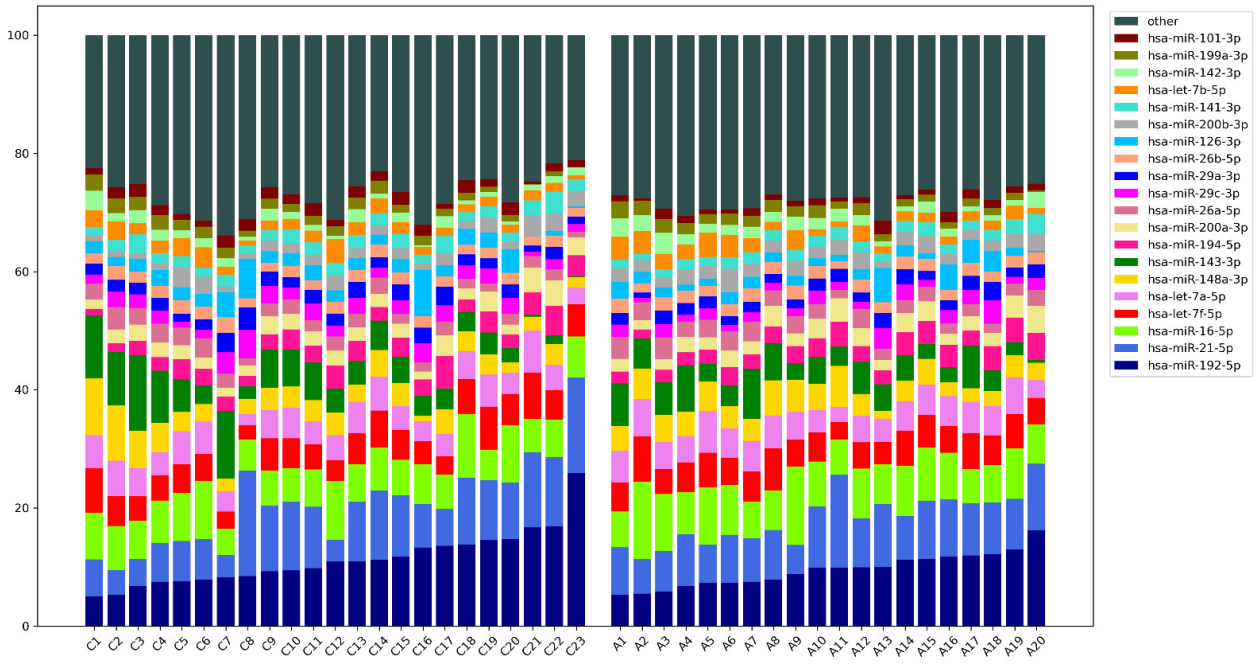

Figure 42. Mean percentage of the 20 most abundant intestinal miRNAs in each of the experimental groups. C – patients without colorectal adenomas, n = 23; A – patients with colorectal adenomas, n = 20.

3.20 Differentially expressed miRNAs

Abundances of miRNAs detected in intestinal biopsy samples were compared between patients with and without colorectal adenomas (N = 43). Analysis revealed significant differences for three miRNAs (Figure 43). Hsa-miR-101-3p was decreased in patients with adenomas (LogFC = -0.66, FDR = 0.005), while hsa-miR-454-3p (LogFC = 0.46, FDR = 0.04) and hsa-miR-132-3p (LogFC = 0.76, FDR < 0.05) were increased.

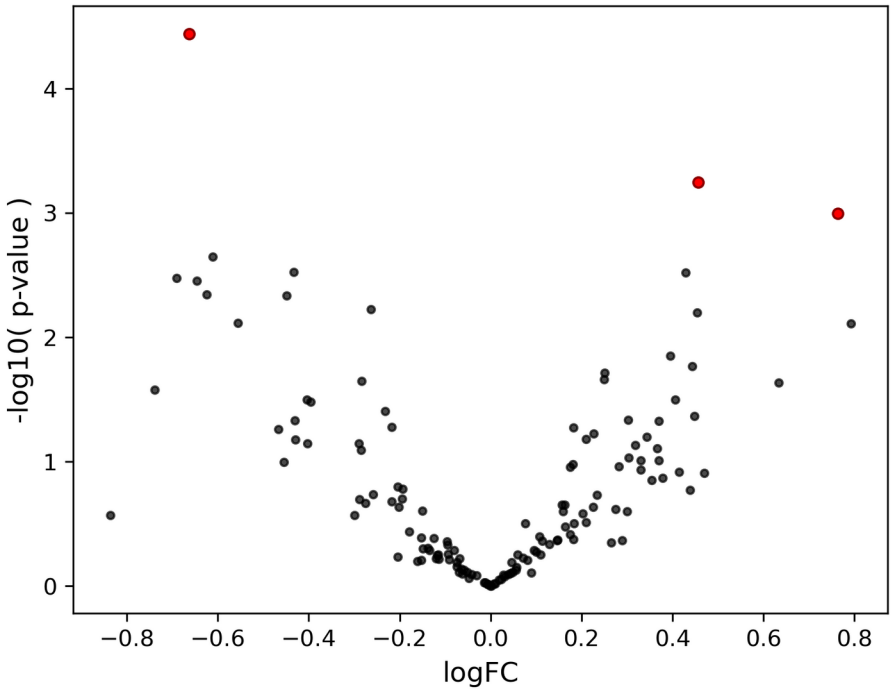

Figure 43. Volcano plot showing P-values of differentially expressed miRNAs between patients with and without colorectal adenomas. Red dots represent differentially expressed miRNAs with P-values < 0.05.

3.21 Fecal microbiome composition in patients with and without colorectal adenomas

Shotgun metagenome sequencing of fecal samples yielded a median of 19473888.5 raw reads. Median value of assigned reads to taxonomy by mOTUs 2 was 6628.5 (interquartile range – 3908). Two of the samples were excluded from the study after the sequencing procedure due to too low read count.

An overview of the gut microbiome composition at the genus level has been shown in Figure 44. Overall, the fecal microbiome compositions of patients in both experimental groups were similar for most of the top 20 genera. However, the relative abundance of the dominant genus *Bacteroides* was lower in patients with colorectal adenomas, while *Prevotella* was more abundant in patients with colorectal adenomas.

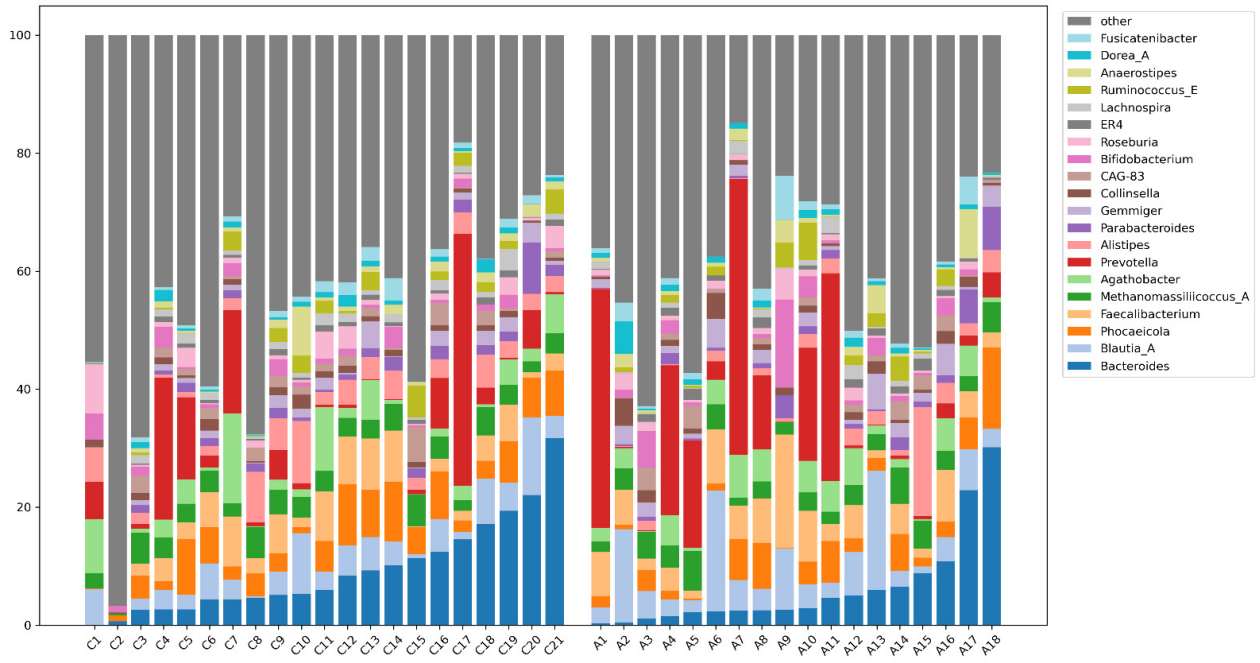

Figure 44. Top 20 microbial genera found in fecal samples from each of the experimental groups. C – patients without colorectal adenomas, n = 21; A – patients with colorectal adenomas, n = 18.

3.22 MiRNA correlation with gut microbiome

Correlation analysis was performed using intestinal miRNA and fecal metagenome datasets obtained from the same patients (N = 41). To detect potential differences in the correlation patterns between patients with colorectal adenomas and patients without due to diverse miRNA compositions, we performed correlation analysis in each of the experimental groups separately (Figures 4 and 5). Taken together, each of the subsets of this study showed a distinct correlation pattern likely due to differences in miRNA and fecal microbiome compositions.

In the group of patients with colorectal adenomas (Figure 45), a set of miRNAs with the strongest representatives affiliated to miR-10 family – hsa-miR-125a-5p, hsa-miR-125b-5p, hsa-miR-99b-5p, and a separate miRNA – hsa-miR-145-5p showed a strong positive correlation with *Actinobacteria* members *Collinsella*, *Pauljensenia*, and *Gordonibacter*. *Firmicutes* members *Peptacetobacter*, *Faecalimonas*, *Streptococcus*, *Bulleidia*, and *Lactiplantibacillus*. In turn, a smaller set, including hsa-miR-552-5p as the strongest representative, correlated with the aforementioned taxa in the opposite direction.

Another set including miR-29 family members positively correlated with *Firmicutes* *Lactiplantibacillus* and *Erysipelotrichaceae*, and *Olsenella* affiliated with *Actinobacteria*, while other *Actinobacteria* members did not show a consistent correlation pattern for this set of

miRNAs. A negative correlation was observed with *Veillonella* and *Gammaproteobacteria* member *Haemophilus*. Correlations in opposite directions with all taxa associated with previous miRNA sets were found for a set of miRNAs that included members of the miR-423 family, among others.

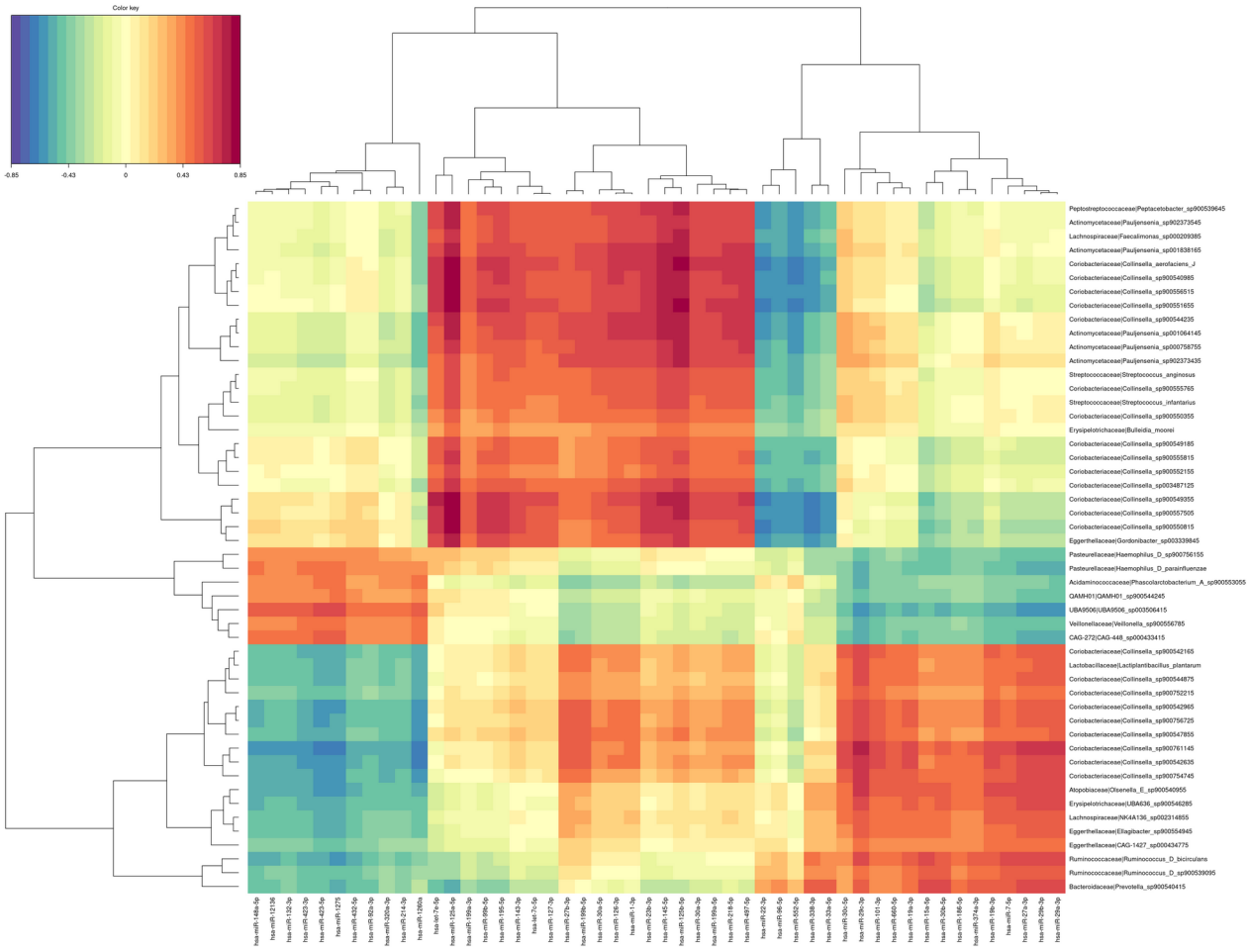

Figure 45. Correlation between intestinal miRNAs and fecal metagenome in patients with colorectal adenomas, N = 18. The color key indicates the correlation levels, with blue and red color denoting negative and positive correlation, respectively.

Overall, the correlation pattern in patients without colorectal adenomas showed distinct clusters of miRNAs with correlations with specific genera in opposite directions (Figure 46). The strongest positive correlation was detected between miR-199 members and *Oscillospiraceae* and *Atopobiaceae* representatives. In addition, *Atopobiaceae* was strongly correlated with both strands of hsa-miR-126, while hsa-miR-30a-5p, hsa-miR-195-5p, and miR-10 family members hsa-miR-125a-5p and hsa-miR-125b-5p among others were associated with *Oscillospirales* representatives including *Acutalibacteraceae* and *CAG-272*. Negative correlation between miRNAs from this set

with the strongest hits observed for hsa-miR-27b-3p, miR-126 family, and miR-199 family and *Lachnospiraceae* members including *Clostridium Q*. In addition, this set of miRNAs was negatively correlated with *Ligilactobacillus ruminis*, *Streptococcus pasteurianus*, *Megasphaera* and *Dielma*.

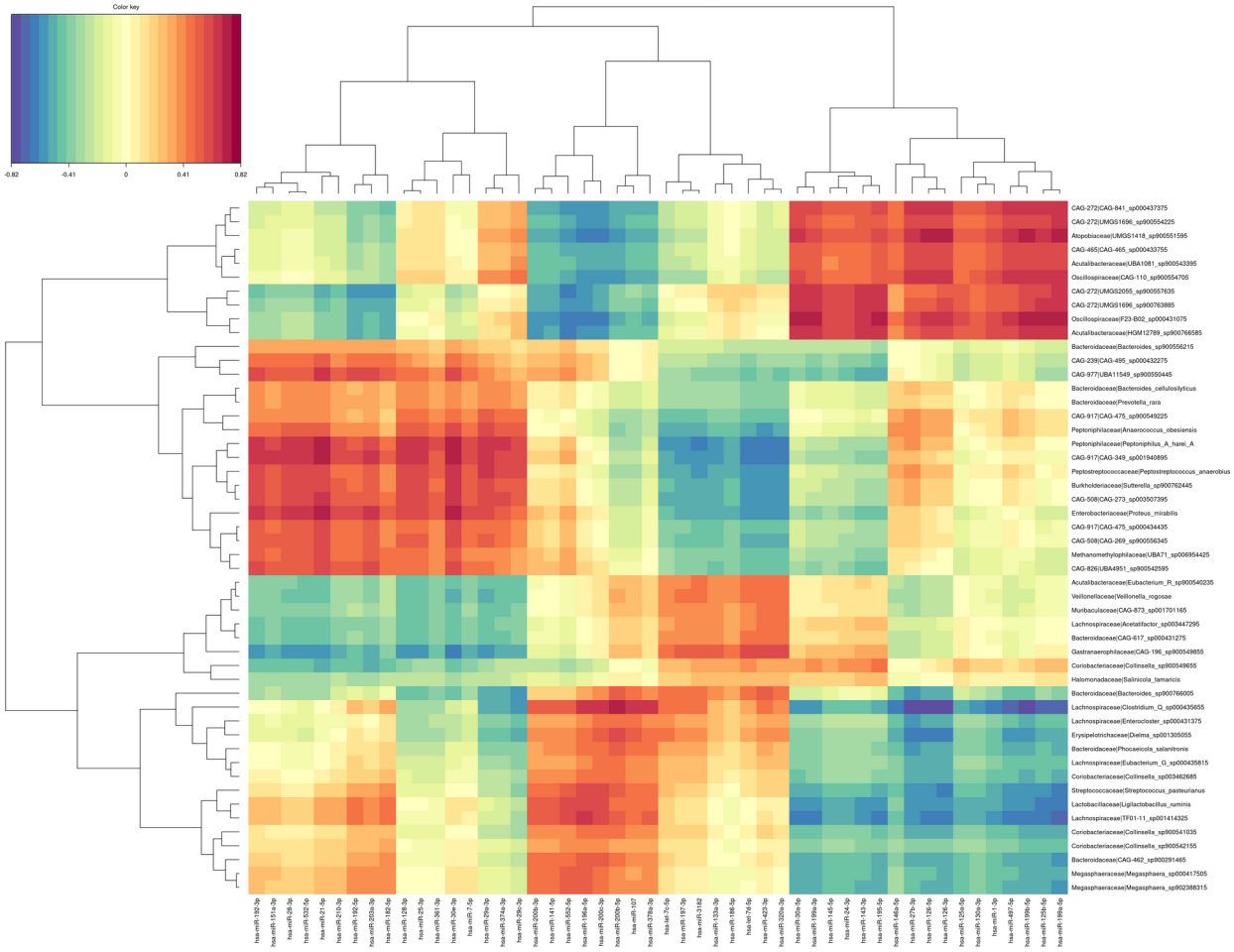

Figure 46. Correlation between intestinal miRNAs and fecal metagenome in patients without colorectal adenomas, N = 21. The color key indicates the correlation levels, with blue and red color denoting negative and positive correlation, respectively.

An opposite correlation pattern was observed for the miRNA set containing miR-8 family members hsa-miR-200b-5p, hsa-miR-200b-3p, hsa-miR-200c-3p, and hsa-miR-141-5p and several other miRNAs and all bacterial taxa mentioned above. The abundance of *Proteobacteria* members with the strongest hit for *Proteus mirabilis* was strongly positively correlated with another set of miRNAs, including hsa-miR-30e-3p, hsa-miR-21-5p, and members of families miR-192, miR-28, and miR-29. Other positively correlated bacterial taxa with this miRNA set include

Peptoniphilaceae and *CAG-917* members, among others, while a negative correlation was observed between these taxa and a set containing let-7 family, hsa-miR-423-3p, and hsa-miR-320a-3p.

3.23 Potential bacterial targets of the correlated miRNAs in humans

To investigate the potential of the correlated miRNAs to biologically target the respective bacteria, the homology of all correlated fecal miRNA sequences with the sequenced metagenome data were analyzed. In addition, the thermodynamic stability of the miRNA-mRNA duplex, expressed as the free energy of binding between the miRNA and the target site, was evaluated.

Homology assessment allowing a maximum of three mismatches, revealed median value of 625062.0 (IQR 304540.0) miRNA-read hits per sample (reads from each pair counted independently). This set of pairs was further tested based on the minimum free energy, setting a threshold of MFE $\leq$ -20.0 kcal/mol. In total, 146 miRNAs had at least one microbial target expressed in terms of at least hundred reads and 240 microbial genera were potential targets for at least one miRNA. For better comprehensibility, only the top 20 miRNAs and their target microorganisms are shown in Figure 47.

The list of top 20 miRNAs which have potential microbial targets includes a row of consecutive miRNAs – hsa-miR-4290, hsa-miR-4291, hsa-miR-4292, hsa-miR-4293, hsa-miR-4294, hsa-miR-4295 (outside the top 20 list), hsa-miR-4296, hsa-miR-4297, hsa-miR-4298, and hsa-miR-4299. Hsa-miR-1275 was found to be the miRNA with the most pronounced potential binding with microbial targets, which was followed by hsa-miR-1260a, hsa-miR-1246, and hsa-miR-12136. Almost all of these miRNAs have potential targets in the following *Bacteroidales* genera – *Prevotella*, *Bacteroides*, *Phocaeicola*, *RC9*, and *Alistipes*. Families, including *Ruminococcaceae* (*Ruminococcus E*, *Ruminococcus D*, *Ruminococcus B*, *Ruminiclostridium E*, *Ruthenibacterium*), *Lachnospiraceae* (*Blautia*, *Coprococcus*, *Hungatella*, *Agathobacter*, *COE1*, *Butyrivibrio A*, *CAG-127*, *Roseburia*, *Acetatifactor*, *Enterocloster*, *Dorea*, *Fusicatenibacter*), *Acutalibacteraceae* (*CAG-177*, *CAG-180*), and *Oscillospiraceae* (*Faecalibacterium*, *Gemmiger*, *CAG-83*, *CAG-170*, *UBA5446*, *CAG-110*), all had potential targets for the above-mentioned miRNAs. Other potentially targeted taxa include *Bacilli* (*Enterococcus* and *Holdemanella*), *Gammaproteobacteria* (*Escherichia* and *Klebsiella*), *Akkermansia*, *Actinobacteria* representative *Bifidobacterium*, and *Archaea* member *Methanomassillicoccus A*.

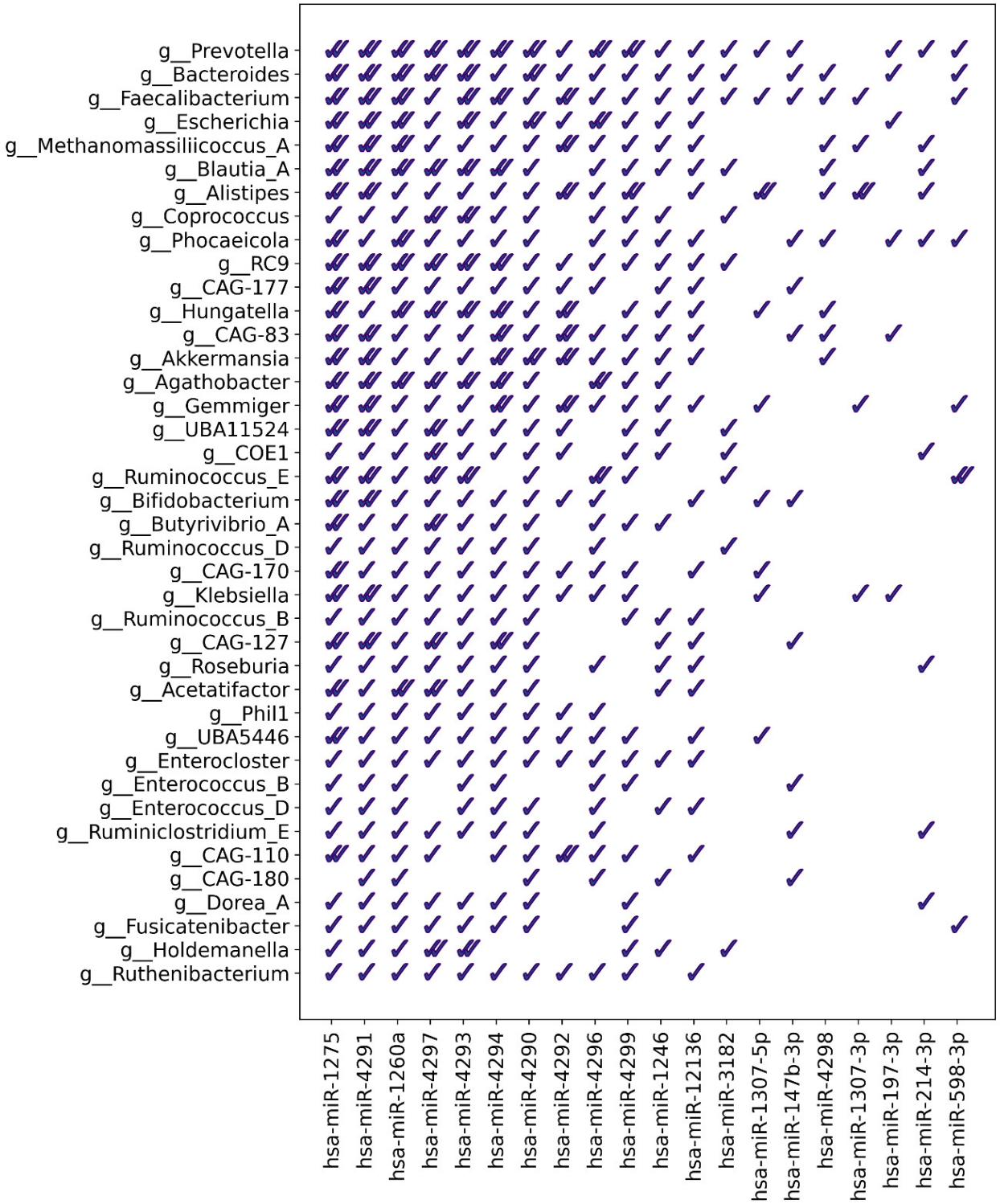

Figure 47. List of miRNAs that have homology with bacterial genomes in humans

3.24 Bacterial targets of the miRNAs identified in different subsets of the correlation analysis in mice and humans

To investigate the translatability of the results between species, correlation patterns were compared. Furthermore, the homology between the sequences of the identified miRNAs and the genomes of bacteria with which they correlate was assessed. In total, we found 39 miRNAs and 95 microbial genera shared between mice and humans in our datasets. Although the exact miRNAs and their microbial targets were not found to be intersected between mice and humans, notable patterns were observed. Both in mice and humans, *Bacteroidales* members were among the top potential microbial targets of miRNAs. In mice, this order was represented by *Muribaculaceae* genera (*UBA9475*, *UBA7173*, *CAG-485*, *Duncaniella*), *Parabacteroides*, and *Alistipes*, while in humans, by *Prevotella*, *Bacteroides*, *Phocaeicola*, *RC9*, and *Alistipes*. In both host species, top microbial targets included members of families *Ruminococcaceae*, *Lachnospiraceae*, *Acutalibacteraceae*, and *Oscillospiraceae*. In addition, phylum *Actinobacteria* was represented by one genus in each host. In contrast to mice, in humans, representatives of separate domain *Archaea* were found to have potential targets for analyzed miRNAs. In each host species, a different set of species-unique miRNAs was found among the miRNAs with the most microbial targets.

4 DISCUSSION

Type 2 diabetes is a widespread metabolic disease, and there is strong evidence that the microbiome plays an important role in its pathogenesis and interaction with therapy (Larsen *et al.*, 2010). The prevalence of the disease is constantly increasing and is recognized as a major concern in public health globally. T2D is associated with the risk of long-term complications that affect patients' functional capacity and quality of life, and potential premature mortality, thus having a significant impact on human lives and the economy (Khan *et al.*, 2020). This global threat demands the optimization of therapeutic strategies and improvements in their efficiency. Metformin is the first-line drug to treat T2D, and it has strong effects on the gut microbiome. Metformin is administered orally, and its main benefits compared with other types of T2D therapy are low risk of hypoglycemia, lowering or neutral effect on body weight, low cost, relative safety, and potential cardiovascular improvements (Dujic *et al.*, 2016; Wu *et al.*, 2017). Nevertheless, there is high variability in the efficacy of metformin therapy, with up to 20-30% of patients receiving metformin developing gastrointestinal side effects and 5% unable to continue the therapy (Dujic *et al.*, 2016). The pharmacodynamic mechanism of metformin action is only partially explained; furthermore, aspects underlying the common gastrointestinal side effects are still unknown. In addition, the pleiotropic effects of metformin have shown promising results in treating many other diseases, thus expanding the drug's applicability and emphasizing the need for successful management of its use.

Studies in both mice and humans have indicated that alterations in the gut microbiome and its metabolic pathways could be related to the antidiabetic effect of metformin (Lee, 2014; Wu *et al.*, 2017). Our previous studies have shown that metformin therapy reduces the diversity of the gut microbiome (Shannon diversity index) for both healthy volunteers and T2D patients (Elbere *et al.*, 2018). The decrease in diversity might be related to the side effects of the therapy. Few studies have been published on the effects of metformin on the gut microbiome in diabetic mice; however, research has tended to focus on fecal samples, which represent the luminal fraction of the large intestine only rather than the microbiome along the entire length of the intestine (Lee, 2014). The same applies to studies involving human subjects to an even greater extent.

We propose this study in order to answer many unclarified issues mentioned above. This thesis aimed to characterize the metformin-induced changes in the gut microbiome composition both in fecal samples and the gastrointestinal tract longitudinally and cross-sectionally. This was proceeded with an analysis of the possible mechanism of how targeted alterations of the gut

microbiome compositions could be introduced. The hypothesis of the study is that the regulation of host miRNA expression in the intestinal cells is one of the mechanisms by which the host can alter the composition of the gut microbiome, and this mechanism is an important part of the T2D pathophysiology. Even more, we hypothesize that metformin may elucidate its effect on the microbiome by changing the miRNA repertoire in intestinal cells.

The effect of metformin on gut microbiome composition and function has been extensively studied both in humans and animal models and even *in vitro*. However, all rodent studies have used 16S rRNA gene sequencing, and most were limited to evaluating metformin effects on only one of the animal sexes. Our study has several strengths. First, we included mice of both sexes in contrast to almost all existing studies in this specific field, providing valuable information on sex-related differences in response to metformin treatment regarding changes in the gut microbiome. Next, we employed a randomized block factorial design, which rendered the study feasible in terms of time and available resources, increased power and reproducibility, and limited the impact of various controllable and uncontrollable sources of bias. We aimed to reliably evaluate metformin's effect on the microbiome rather than reach the maximal effect of metformin response with elevated doses. Therefore, we chose to use a smaller metformin dose (50 mg/kg body mass/day) compared to other studies (200-300 mg/kg body mass/day) to match the concentrations applied to humans. Finally, we fed the mice representing control groups in our study with a well-matched diet and not the regular chow, which is not standardized and has beneficial effects on the gut microbiome *per se* due to its short-chain fatty acid-increasing nutritional composition (Dalby *et al.*, 2017). The use of regular chow is common in previous studies assessing the effect of metformin on the microbiome, which can substantially alter the results.

Moving to separate subsets of the study, fecal samples were analyzed using a shotgun metagenomic sequencing approach, which allows a species- or even strain-level precision in detecting microbiome representatives. Additionally, it enables the determination of the functional profile. Furthermore, we took advantage of using an animal model to collect intestinal samples and their associated microbiome to analyze spatial differences. Our study is the first to report the results of such a systematic assessment of spatial variation in the effects of metformin in animals of both sexes. According to the 3Rs principle, strongly promoted in animal research, we proposed a new hypothesis about the host miRNA-mediated effects of metformin on the gut microbiome, which we investigated using the extensive resources of biological materials collected during the

mice experiment. This approach allows combining different -omics-based data sets, thus enhancing the investigation of the complex picture of metformin effects.

To further test the hypothesis, the interaction between host intestinal miRNAs and gut microbiome was assessed in human subjects as well. As the concept that certain miRNAs molecularly interact with bacterial genomes should be existing in different hosts, available patient cohort (patients undergoing the colonoscopy procedure due to various reasons) was used for the study. In addition, we analyzed the differences in miRNA composition between patients with and without colorectal adenomas, which allowed to determine which miRNAs have biomarker potential for early detection of colorectal cancer risk.

4.1 The effect of metformin treatment on the fecal microbiome

Fecal metagenome analysis revealed a uniform effect of metformin treatment on bacteria representing *Bacilli*. Genus *Lactobacillus*, represented by *L. murinus, L. farraginis, L. paracasei, L. kefiranofaciens*, was increased. This finding is consistent with previous work showing that metformin increases the relative abundance of genus *Lactobacillus* in HFD-fed rats (Bauer *et al.*, 2018). Probiotic supplementation with *Lactobacillus* improves glucose parameters in diabetic rats and prevents insulin resistance and hyperglycemia in HFD-fed mice (Bauer *et al.*, 2018), which is in line with our observations. *Lactococcus*, another genus of lactic acid bacteria *Lactococcus*, including *L. raffinolactis, L. garvieae, L. fujiensis, L. chungangensis, L. lactis, L. plantarum* is increased in response to metformin treatment, which is in agreement with a study evaluating the effect of metformin on the microbiome in the short term HFD-induced obesity (Ji, Wang and Li, 2019).

One of the gut microbiome's main roles is to break down the dietary fiber and starch incompletely hydrolyzed by the host. Short-chain fatty acids (SCFA), including propionate and butyrate, are the main fermentation products of fiber digestion and can be used for lipid or glucose *de novo* synthesis. Changes in the SCFA profiles are associated with changes in the gut microbiome (McKnite *et al.*, 2012). In our study, taxa associated with SCFA production, such as *Bacteroides* and *Enterorhabdus*, which convert amino acid derivatives as the energy source (Clavel *et al.*, 2014), are represented among the species that are increased after metformin treatment.

The relative abundance of *Bacteroides* was significantly higher after the metformin treatment, with the most represented bacteria being *B. intestinalis, B. vulgatus*, and *B. acidifaciens*. However, the effect was heterogeneous, as we observed a decrease in several *Bacteroides* species

in some experimental units. The increased abundance of *Bacteroides* and *Parabacteroides* is in line with existing data (Lee *et al.*, 2018). It has been described that *B. fragilis* colonization aggravates metabolic disorders induced by HFD, whereas metformin inhibits the growth of *B. fragilis* through modification of folate and methionine metabolism (Sun *et al.*, 2018).

In a study on core gut bacteria of healthy mice, a high prevalence (99%) and a relatively high abundance of *Parabacteroides* were observed, suggesting that *Parabacteroides* might be essential to host health (Wang *et al.*, 2019). This fact corresponds to metformin's reported health benefits regarding the microbiome composition shift (Wu *et al.*, 2017). It has been hypothesized that the microbial growth-inhibitory effect of metformin is more pronounced on anaerobic organisms than aerobic organisms, as anaerobic respiration produces less ATP than aerobic respiration (Prattichizzo *et al.*, 2018).

Metformin-induced changes in the abundance of *Akkermansia muciniphila* and metabolic improvement due to these changes have previously been shown in several studies (Lee, 2014; Wu *et al.*, 2017). Instead, our study reports no significant changes in *A. muciniphila* abundance in response to metformin in fecal metagenomic data.

Our study revealed a consistent effect of metformin on the abundance of species representing *Clostridia*, more specifically, members of *Lachnospiraceae* and *Ruminococcaceae*. Thus all of the identified differentially abundant taxa from these dominant butyrate-producing families (except for *Butyrivibrio* and *Rumonicoccus sp. Zagget7*) were decreased after the treatment in the HFD_Met+ group. Dominant butyrate-producing bacteria *Roseburia hominis*, *Roseburia intestinalis, Faecalibacterium prausnitzii* found in the human intestine (La Rosa *et al.*, 2019), *Intestinimonas butyriciproducens* and *Eubacterium plexicaudatum* first found in mouse intestine (Wilkins, Fulghum and Wilkins, 1974; Kläring *et al.*, 2013), and *Anaerotruncus spp.* are among these taxa. Butyrate has been demonstrated to positively impact gastrointestinal tract homeostasis promoting the growth of intestinal epithelial cells, increasing the expression of tight junction proteins, and acting as an anti-inflammatory agent (Kant *et al.*, 2016). Previous studies have reported that obesity and T2D are associated with a reduction of butyrate-producing bacteria and an increase in opportunistic pathogens (Forslund *et al.*, 2015; McCreight, Bailey and Pearson, 2016). A recent randomized pilot study in which the impact of probiotic supplement on metformin's effect on glycemia in prediabetic subjects was assessed has identified an increase in the relative abundance of *Anaerotruncus colihominis (Cluster IV)* only in participants receiving both metformin and the probiotic but not in participants taking either metformin or probiotic alone (Palacios *et al.*, 2020).

Our results show that metformin alone can impact the relative abundance of different *Anaerotruncus* species in opposite directions; *Anaerotruncus sp.* G3(2012) was decreased, while the relative abundance of uncultured *Anaerotruncus sp.* was augmented. Another member of *Lachnospiraceae – Dorea* is shown to be increased in T2D individuals and negatively correlated with the abundance of butyrate-producing bacteria (Qian Li *et al.*, 2020). Gene-targeted approaches to investigate the butyrate-producing bacterial communities of the human gut microbiome have suggested that butyrate-producing colon bacteria form a functional group rather than a monophyletic group (Rivière *et al.*, 2016). This finding suggests that the functional niche of *R. hominis*, *R. intestinalis*, *I. Butyriciproducens*, and *F. prausnitzii* as butyrate-producing bacteria is substituted by other taxa, possibly butyrate-producing species identified with an increased relative abundance in response to metformin treatment which was discovered in our study.

4.2 Functional analysis

We focused the functional analysis on the hierarchies related to the three main macronutrients consumed in the diet – carbohydrates, lipids, and proteins, which can reach the colon due to their structural complexity or ingested amount, which surpasses the possibilities of primary digestion (Oliphant and Allen-Vercoe, 2019). Most of the altered hierarchies in response to metformin treatment in HFD-fed groups represent amino acid metabolism-related functions. Our data show that metformin increases the metagenome functions affiliated with arginine and proline metabolism; glycine, serine, and threonine metabolism; and cysteine and methionine metabolism; in turn, valine, leucine, and isoleucine biosynthesis is hindered.

In contrast to the previously reported results, in our study, metformin affects the functions associated with propionate (propanoate) and butyrate (butanoate) metabolism in opposite directions (Forslund *et al.*, 2015; Wu *et al.*, 2017). When HFD-fed groups were contrasted to CD-fed mice, regardless of the metformin treatment status, propionate metabolism was enriched. An analysis between HFD_Met+ groups before and after the treatment, both in the combined sexes and female subset, highlighted a decrease in butyrate metabolism. This corresponds to a metformin-related decrease in the relative abundances of different butyrate producers representing *Lachnospiraceae* and *Clostridiaceae* observed in our study.

The role of phenylalanine-derived metabolites – phenylethylamine and trans-cinnamic acid in the context of gut microbiome functions is not well described (Oliphant and Allen-Vercoe, 2019). A study of children consuming a low-phenylalanine diet due to phenylketonuria indicated

that the gut microbiome of these children was depleted in butyrate-producing species and enriched in *Blautia spp.* and *Clostridium spp.* (Bassanini *et al.*, 2019), which is in partial agreement with our data. We observed a significant decrease in the abundance of butyrate metabolism hierarchy in HFD_Met+ groups before and after metformin treatment simultaneously with decreased phenylalanine metabolism; however, in contrast, the relative abundance of *Blautia spp.* was not promoted.

4.3 Spatial variation of the effect of metformin on the gut microbiome

The effect of metformin therapy on the composition and function of the gut microbiome has been demonstrated in previous studies (Wu *et al.*, 2017; Lee *et al.*, 2018; Dong *et al.*, 2019; Ahmadi *et al.*, 2020; Qingzhong Li *et al.*, 2020; Jauvain *et al.*, 2021; Liu *et al.*, 2021); however, an in-depth analysis of spatial variation of the effects of long-term metformin treatment on the gastrointestinal tract has not yet been performed *in vivo*. Our study provides novel information on the effects of metformin on both the gastrointestinal mucosa and the luminal layers at four different sites of the gut.

We observed significantly different alpha diversities between all the studied intestinal parts corresponding to anatomical and functional differences of the various sites. The cecum and colon microbiomes were more diverse than those of the small intestine, consistent with a previous report (Lkhagva *et al.*, 2021). Analysis revealed that metformin treatment has the most pronounced effect on the samples representing the proximal small intestine, followed by the distal small intestine. Cecum samples of all experimental groups clustered most closely, showing the least effect of any studied factors on this intestinal part and suggesting its relative stability.

We have shown the effect of metformin treatment, sex, diet type, and intestinal layer on the spatial variation of the gut microbiome by analyzing each of these factors separately. Furthermore, we have investigated the effect of metformin treatment in each of the subsets formed by combinations of levels of the studied factors: intestinal layer, intestinal part, diet type, and sex, and detected substantial variation of metformin’s effects in each of these subsets.

Metformin treatment had a more pronounced effect on the microbiome composition in both parts of the small intestine (indicated by higher absolute LogFC values). This finding confirms the hypothesis that the effect of metformin treatment is not uniform in the whole intestine but rather depends on the absorption characteristics of the medication. When all mice were contrasted based on metformin treatment status, several genera were increased uniquely in the proximal small

intestine, where according to a previous study (Bauer *et al.*, 2018), the absorption of metformin occurs most. Among these genera, many are aerobic bacteria – *Duganella*, *Chryseobacterium*, *Undibacterium*, *Corynebacterium*, *Methylophilus*, and the anaerobes *Anaeroplasma* and *Mucispirillum*.

Mucispirillum is commensal in the microbiota of humans and various vertebrates. The only species of the genus, *Mucispirillum schaedleri*, is a core member of the laboratory mouse microbiota throughout the whole gastrointestinal tract. *Mucispirillum schaedleri* has not been widely identified in human studies due to its low abundance in fecal samples, but it is enriched in the gut mucosa (Herp *et al.*, 2021). Genomic data have indicated that instead of degrading mucosal glycans such as mucin, *M. schaedleri* predominantly processes monosaccharides, oligopeptides, amino acids, glycerol, and short-chain fatty acids produced by other bacteria (Loy *et al.*, 2017).

Metformin treatment substantially increased the abundance of *Lactococcus* in the distal part of the small intestine. This is consistent with a previous study with a similar duration and route of metformin treatment but using only aged male mice, where an increase in *Lactococcus* was observed in fecal samples after treatment with metformin (Ahmadi *et al.*, 2020). Lactic acid bacteria, including *Lactococcus*, produce lactate, a substrate for other members of the microbiota to convert it into butyrate (Thursby and Juge, 2017). We observed a subtle increase in another lactate producer, *Lactobacillus*, in the distal small intestine, whereas the abundance of the genus was reduced in the proximal small intestine and cecum and unchanged in the colon. Previous studies have produced conflicting results regarding the effect of metformin on *Lactobacillus* (Zhang, 2020). This could be explained by the use of fecal samples that do not fully represent the small intestinal microbiota, differences in experimental designs, and species-specific effects not detected in the genus-level analysis (Gurung *et al.*, 2020). *Lactobacillus* has been suggested to help restore glucose sensing in the upper small intestine by increasing *SGLT1* expression (Bauer *et al.*, 2018). The most reduced genus *Proteus* in the distal small intestine was also not significantly affected elsewhere. A previous study showed that metformin reduces the abundance of opportunistic pathogens, including *Proteus*, in db/db mice (Zhang *et al.*, 2015).

Roseburia, another butyrate-producing genus, was increased in both parts of the small intestine. This, together with the intestinal part-unique effects of metformin treatment, further supports the proximal and the distal part of the small intestine as principal sites of metformin-mediated effects on the gut microbiome. Furthermore, we are the first to locate the previously

reported metformin-induced increase in *Parabacteroides* in both parts of the small intestine and colon.

Our results show that metformin mostly targets genera belonging to *Clostridia* in all intestinal sites. Altered genera include the representatives of *Lachnospiraceae* (*Blautia*, *Lachnoclostridium*, *Lachnospira*, *Marvinbryantia*, *Roseburia*), *Clostridiaceae* (*Butyricicoccus*, *Clostridium*), *Ruminococcaceae* (*Anaerotruncus*, *Faecalibacterium*, *Oscillibacter*, *Ruminococcus*), and *Eubacteriaceae* (*Eubacterium*).

Our study, in which we analyzed fecal metagenome similarly to others, has shown that metformin therapy decreases the abundance of members of *Clostridia* families *Ruminococcaceae* and *Lachnospiraceae* – major butyrate producers (Bravard *et al.*, 2021). We found an increase of *Roseburia* in both parts of the small intestine and a depletion of the genus in the cecum and colon in the luminal layer. This could explain the reduced abundance of *Roseburia* in feces after metformin treatment (Ahmadi *et al.*, 2020), as feces mostly correspond to the samples of colonic contents. *Bacteroides* has been shown to be affected by metformin treatment. Our spatial variation study supports these findings as we report a decrease in *Bacteroides* in response to metformin, specifically in the colon (Lee *et al.*, 2018).

4.4 Sex differences in response to metformin treatment and gut microbiome composition

We found substantial sex-related differences in response to metformin treatment. In contrast to another study involving both sexes of mice where serum glucose level decreased after metformin treatment, especially in female mice, we observed an increase in plasma glucose level in females, although male mice responded with a decrease in plasma glucose level in both HFD- and CD-fed groups (Lee, 2014). Also, insulin resistance indicated by the HOMA-IR index declined after the metformin intervention in the HFD_M_Met+ group but, inconsistent with previously reported data, insulin resistance was augmented in the HFD_F_Met+ group. The same pattern was manifested in CD_Met+ groups, thereby strengthening the discrepancies in metformin effects between sexes. These differences might be explained by the fact that we used a substantially lower dose of metformin (50 mg/kg body mass/day), which is considered the maximum dose that the body can utilize efficiently (He and Wondisford, 2015).

In addition, several studies, which have investigated sex-related differences in the gut microbiome in the context of changed diet, have identified hormonal effects on gut microbiome

composition (Org *et al.*, 2016; Kaliannan *et al.*, 2018; Acharya *et al.*, 2019), specifically emphasizing estrogen-induced gut microbiome changes and protection against metabolic syndrome in mice, both of the C57BL/6 and *ob/ob* background. These effects possibly also contribute to the identified differences between sexes.

In addition, we observed a strong sex-specific effect on the spatial variation of the gut microbiome when all mice were contrasted based on sex, as the abundance of *Lactobacillus* was strongly reduced in all intestinal parts and layers of males. This finding is supported by a study in healthy humans, where an increase of *Lactobacillales* in female mucosa-associated microbiota was reported (Borgo *et al.*, 2018). *Lactobacillus* has been shown to be increased in diabetes patients and animals fed with a high-fat diet (Forslund *et al.*, 2015; Sedighi *et al.*, 2017; Wang *et al.*, 2020), consistent with enrichment of the genus in the colon and both intestinal layers of HFD-fed mice compared to CD-fed mice in our study. However, we observed that *Lactobacillus* was decreased in the proximal small intestine of HFD-fed mice, showing the differences between the composition of the fecal microbiome and the microbiome of different intestinal parts.

Previous studies have shown that *Lactobacillus* spp. reduce blood glucose levels in high-fat diet-induced diabetic male mice (Gulnaz *et al.*, 2021), but in turn is positively correlated with blood glucose levels in European women with T2D (Karlsson *et al.*, 2013), and metabolites produced by *Lactobacilli* may contribute to the glucose modulation of metformin (Rooj, Kimura and Buddington, 2010). Sex-related differences in *Lactobacillus* abundance that we observed might partially explain these discrepancies emphasizing the need to perform studies on individuals of both sexes to obtain as complete information as possible. An increase in *Anaerotruncus* has been found in women with metabolic syndrome compared to men (Santos-Marcos *et al.*, 2019). Our data support these sex-related differences, narrowing them down to the cecum. We also detected a higher abundance of opportunistic pathogens *Proteus* in both layers and *Staphylococcus* in the lumen of male mice compared to females. *Proteus* was increased in the distal small intestine, *Staphylococcus* in the cecum, and both genera in the colon of males. This suggests that the microbiota of male mice in this model contains more opportunistic pathogens than females, which could contribute to another layer of phenotypic sexual dimorphism reported in HFD-fed mice (Casimiro *et al.*, 2021).

4.5 The effect of studied factors on the miRNA composition in mice

The study by Liu and colleagues demonstrated for the first time that host-derived fecal miRNAs specifically target bacterial genes and thereby modulate the gut microbiota (Liu *et al.*, 2016). Since then, several studies investigating the effects of host miRNAs on gut microbiota in different disease contexts have been performed, including colorectal cancer (Yuan, M. B. Burns, *et al.*, 2018; Tomkovich *et al.*, 2020), inflammatory bowel diseases (He *et al.*, 2022; Shen *et al.*, 2022), insulin resistance (Guo *et al.*, 2020), liver dysfunction (Santos *et al.*, 2020), neurological disorders (Hewel *et al.*, 2019; S. Liu *et al.*, 2019; H. M. Chen *et al.*, 2022) as well as healthy mice and humans (Horne *et al.*, 2019; Tarallo *et al.*, 2022).

Based on the findings by Liu and colleagues (Liu *et al.*, 2016), we hypothesized that host miRNAs are mediators in metformin-induced changes in the gut microbiome composition and functions in the context of type 2 diabetes. To test this idea, we investigated the effect of ten weeks long metformin treatment on miRNA expression in fecal and gut mucosal samples representing proximal small intestine and cecum of diabetic and control group mice and combined these datasets with metformin-induced gut microbiome changes in the background of T2D described in the previous parts of the thesis. To our best knowledge, this is the first study to address the effect of metformin on fecal miRNA composition and its implications on the gut microbiome.

Our results show that the repertoire of miRNAs in fecal samples is relatively consistent despite the changes in time and treatment status. The fact that at least 20 miRNAs were represented in each of the experimental groups, including longitudinal samples, provides an information on the overall stability of the fecal miRNA content over time, marking them as consistent players of the microbiome ecosystem. All miRNAs from the top 20 list have been shown previously to be present in fecal samples of mice and humans (Liu *et al.*, 2016; Moloney *et al.*, 2018; Horne *et al.*, 2019; S. Liu *et al.*, 2019; Tarallo *et al.*, 2019; Viennois *et al.*, 2019; Tomkovich *et al.*, 2020; Wohnhaas *et al.*, 2020). It should be noted that we have used a sequencing-based approach compared to many of the other studies, thus providing a relatively unbiased view of fecal miRNA content. Intestinal miRNA composition of the proximal small intestine was similar to the one of the cecum; almost all top 20 miRNAs corresponded between sites, though the relative abundances of miRNAs were different.

Differential expression analysis between the studied groups mainly revealed diet-related and some sex-related effects on miRNA composition, while metformin had no significant effect in the different contrasts. Dietary lipid effects on intestinal miRNAs have been characterized previously

and have shown a large number of miRNAs that are influenced in response to high-fat diet feeding (Gil-Zamorano *et al.*, 2020)

Correlation analysis showed the ability of certain miRNAs to distinguish bacteria at the family and genus levels. Thus specific groups of miRNAs were inversely associated with the *Akkermansia* and the genera from *Lachnospiraceae* (with some exceptions), *Oscillospiraceae*, *Erysispelotrichaceae* and the genera from *Bacteroidaceae* and *Muribaculaceae* (both from *Bacteroidales*). Most of the miRNAs grouped based on their affiliation within miRNA families. Strong example of this is miR-8 family with its members – mmu-miR-200a-3p, mmu-miR-200b-3p, mmu-miR-200c-3p, mmu-miR-141-3p, and mmu-miR-429-3p.

The negative correlation between miR-29 family members and *Bacteroides* further supports the relationship between miR-29, T2D status, and specific microbiome composition. It can be hypothesized that the previously shown decrease in miR-29 abundance in response to metformin therapy (Udesen *et al.*, 2020) is the joining factor between the elevated abundance of *Bacteroides* and the reduction of T2D manifestations after the treatment.

When analysis of potential bacterial targets in feces was combined with a respective correlation analysis, several miRNA and bacterial genus pairs were identified. Both mmu-miR-5126 and mmu-miR-5119 potentially target several *Oscillospiraceae* genera and *Prevotellaceae* genus *CAG-873* and affect their abundance in opposite directions. This is consistent with the results of the fecal metagenome analysis, where we have shown a lower relative abundance of *Prevotella* and an increase in *Oscillibacter* in the feces of high-fat diet-fed mice compared to control diet-fed ones both of which were altered by metformin treatment (Silamiķele *et al.*, 2021). *Oscillospira* has been suggested to be a butyrate producer and can use both animal-derived and plant sugars, furthermore, glucose significantly increases its growth (Yang *et al.*, 2021). Furthermore, a study on the interaction between colon mucosal biofilms and mice hosts after their inoculation with colorectal cancer patient-derived microorganisms has identified several miRNAs that mainly target bacterial genes (Tomkovich *et al.*, 2020). These miRNAs include mmu-miR-2137, mmu-miR-5126, and mmu-miR-6538, which completely corresponds to our results. Additionally, we identified mmu-miR-5119, mmu-miR-215-5p, and mmu-miR-5106 as miRNAs that target *Bacteroides*, *Parabacteroides*, *Muribaculaceae* members, and *Lactococcus* not targeted by the first mentioned three miRNAs in our analysis.

4.6 The effect of studied factors on the miRNA composition in humans

Studies profiling miRNA expression in the transition from normal colorectal tissue to adenoma and to carcinoma have identified several potential diagnostic and treatment biomarkers (J. Li *et al.*, 2019). Analysis of the differentially expressed miRNAs between patients with and without colorectal adenomas in the present study revealed three significantly different miRNAs. All of these miRNAs have previously been associated with the development of different cancer types. Hsa-miR-101-3p has been described as a tumor suppressor in various types of cancer by targeting genes involved in the cell growth, apoptosis, migration, and invasion of cancer cells (Liu *et al.*, 2015; Gu *et al.*, 2021). This is consistent with the downregulation of hsa-miR-101-3p in our data from patients with colorectal adenomas, a potential precursor of colorectal cancer. Furthermore, hsa-miR-101-3p has been suggested as a biomarker for colorectal cancer diagnosis and prognosis (Strillacci *et al.*, 2013). In contrast, in another study in which miRNAs involved in colorectal tumorigenesis were investigated, miR-101 has been identified as one of the miRNAs upregulated continuously as the normal tissue progresses to adenoma and to cancerous tissues (Yin *et al.*, 2016). It has been shown that hsa-miR-101-3p and hsa-miR-132-3p both have experimentally validated target genes involved in the KEGG pathway of colorectal cancer (X.-N. Li *et al.*, 2019). These studies and our results suggest that hsa-miR-101-3p may have a prognostic potential for colorectal adenoma progression to carcinoma.

MiRNAs which were increased in patients with colorectal adenomas, have been reported to be involved in carcinogenesis. Compared to sensitive ones, an upregulation of hsa-miR-454-3p has been shown in oxaliplatin (first-line chemotherapy for colorectal cancer)-resistant colorectal cancer cells (Qian *et al.*, 2021). Studies have shown that hsa-miR-454-3p exhibits both tumor suppressive and oncogenic properties in various types of cancer. It promotes breast cancer progression by inducing the epithelial-mesenchymal transition phenotype (Wang *et al.*, 2021), while in glioblastoma and several other types of cancers hsa-miR-454-3p functions as a tumor suppressor (Zuo *et al.*, 2019).

Similarly, hsa-miR-132-3p has been involved in different human tumors (Rafat *et al.*, 2021) and exhibits both oncogene and tumor suppressor functions (Moghbeli *et al.*, 2021). In the context of colorectal cancer, hsa-miR-132-3p has been described as downregulated in tumor tissues compared to paired non-tumoral tissues (Yong *et al.*, 2016). In contrast, in another study using colorectal cancer tissues, the expression of hsa-miR-132-3p, which targets anti-apoptotic protein SIRT1, was augmented compared to the normal colon tissue (Kara *et al.*, 2015).

In contrast to a previous study that evaluated differentially expressed miRNAs in colorectal adenoma and carcinoma patients compared to healthy subjects, and described eight dysregulated miRNAs (Liu and Li, 2019), none of these miRNAs were significantly differentially expressed in our study.

Microbial target analysis showed that the differentially expressed miRNAs in colorectal adenoma patients have several potential microbial targets. For hsa-miR-101-3p, the main targeted bacterial families include *Lachnospiraceae*, *Acutalibacteraceae*, and *Ruminococcaceae*, while hsa-miR-132-3p targets *Bacteroidaceae* members and *Lachnospiraceae* representative *Marvinbryantia*. Hsa-miR-454-3p potentially targets *Acutalibacteraceae* member *CAG-180* and *Methanomassiliicoccus A* affiliated with *Archaea*. Another *Archaea* representative *Methanobrevibacter A* is potentially targeted by hsa-miR-101-3p.

When the results of differential expression analysis, correlation analysis, and analysis of potential microbial targets were combined, some hypotheses can be put forward. Hsa-miR-101-3p, which was decreased in colorectal adenoma patients, potentially targets members of *Lachnospiraceae* (*Blautia A*, *Coprococcus*, *Ruminococcus B*, *Lachnospira*), *Ruminococcaceae*, and *Acutalibacteraceae* families, in addition one *Archaea* target has been shown. In turn, hsa-miR-132-3p, which was increased in colorectal adenoma patients, has potential targets in *Bacteroidaceae* members *Bacteroides* and *Phocaeicola* as well as *Marvinbryantia* affiliated with *Lachnospiraceae*. Correlation analysis in the colorectal adenoma patient subset is in agreement with the opposite roles for these miRNAs, as each of them corresponds to a different miRNA set. Furthermore, hsa-miR-132-3p shows a negative correlation with *Prevotella* (*Bacteroidaceae* member) and hsa-miR-101-3p a positive correlation with *Lachnospiraceae* and *Ruminococcaceae* representatives.

Another finding originating from the target analysis relates to the widely described differentially abundant genus *Fusobacterium*, enriched in colorectal adenomas and adenocarcinomas (Kostic *et al.*, 2013; Xue *et al.*, 2021). We found that this genus is potentially targeted by miRNAs, including hsa-miR-4291, hsa-miR-1260a, hsa-miR-4299, hsa-miR-190a-5p, and hsa-miR-29a-3p. MiR-29 family members have been described as potential biomarkers of colorectal cancer (Nguyen *et al.*, 2022), suggesting a possible interaction between miR-29 and *Fusobacterium* in the development of colorectal cancer.

4.7 Bacterial targets for miRNAs – similarities and differences between species and backgrounds

The microbial taxa identified as potential targets of host miRNAs in both mice and humans correspond to the described microbiome composition in both species. As illustrated before (Nagpal *et al.*, 2018), mouse fecal microbiome is dominated by the genus of *S24-7* (now known as *Muribaculaceae*), followed by *Clostridiales* members and *Oscillospira*. In humans, *Bacteroides* is the top genus, and *Ruminococcaceae*, *Clostridiales*, and *Prevotella* are the taxa that follow. All of these taxa were represented among the ones with the most bacterial targets for host miRNAs in each of the respective host species. Possible explanation of this is that the count of microbial reads directly correlates to the relative abundance of the species in the microbiome.

According to miRBase, most of the top miRNAs with potential microbial targets are species-specific for both mice and humans. This could be explained by the miRNA nomenclature and naming system in which a novel miRNA is given an identification number consequently. This means that, for example, miR-4290 to miR-4299 found only in humans or miR-5106, miR-5119, and miR-5126 found only in mice are not described in other species yet. As most of existing studies have focused on searching the miRNA targets in host itself, miRNAs widely characterized in different species tend to have lower numbers in their names.

Another possible explanation for the differences between species was the miRNA source for correlation and target analysis; in mice fecal miRNAs were analyzed together with fecal microbiome, while in humans fecal microbiome data were combined with intestinal miRNAs. The fact that we identified the same miRNAs which have bacterial targets in mouse host as the ones in another study using mouse model inoculated with microbiome from colorectal cancer patients (Tomkovich *et al.*, 2020) suggests that host miRNAs which have bacterial targets are rather host species-specific than disease-specific. Furthermore, miRNAs seem to specialize to some extent regarding their potential targets – part of the identified miRNAs target host genes, while another part mainly targets microbial genes; however, it should be noted that these interactions are not exclusive.

Our results, together with previous studies in different conditions, including multiple sclerosis, inflammatory bowel diseases, and colorectal cancer (S. Liu *et al.*, 2019; Tarallo *et al.*, 2019; Viennois *et al.*, 2019) demonstrate that host miRNAs found in fecal and intestinal samples potentially mediate the changes in the gut microbiome composition also in the context of T2D. *Prevotella*, *Oscillospiraceae* members (represented by *Faecalibacterium* in humans), and

Bacteroides were among the bacterial taxa with the most targets for miRNAs in both host species. This supports the previously reported ability of several different miRNAs to target the same mRNA target (Taganov *et al.*, 2006). Furthermore, in agreement with the existing literature (Choi *et al.*, 2017), we identified that several bacterial nucleic acid sequences are potentially targeted by the same miRNAs, thus showing their potential to serve as a tool for microbiome composition interventions.

4.8 Limitations of the study

Though the study was designed considering almost all the possible feasibly-controllable aspects, several limitations were still present. The sample size for the study was determined by the resource equation method, which is appropriate for such complex designs as our animal study. Despite theoretical knowledge that the microbiome should be shared between animals housed in the same cage due to the characteristic coprophagy of mice, we did not observe this phenomenon to the full extent. Animals were bred in specific pathogen-free conditions, and our facility corresponds to this microbiological status. Despite the assumptions that the microbiome composition between litter mates coming from the same institution should be relatively similar, we observed a substantial variation in microbiome compositions between different mice. To overcome this, the sample size in future studies should be increased. In addition, we were not able to measure the intake of food and water and thus metformin dosage for each mouse individually. Because of this, the precise information on metformin dosage was not available, and it could differ between mice.

Some of the animals were single-housed for a substantial duration of the study mainly due to the euthanasia of their cage mates because of substantial fighting wounds. This possibly led to a decreased exchange of gut microbiome between mice and potentially influenced the microbiome composition dynamics in these cages, which is another possible source of the microbiome variation between animals. In the future, samples representing all of the animals included in the study could be processed and sequenced to obtain knowledge and microbiome variation between cages and its dependence on the cage's social structure and the number of mice in a cage.

Another limitation is related to data processing and analysis methods. Existing databases and tools often do not provide complete information on microbiome abundance or classification; hence we were not yet able to explain a substantial part of the data. Though this deficiency could potentially be eliminated by re-analyzing the data with state-of-the-art tools and techniques as they

emerge. MiRNA sequencing data showed low mapping to the miRBase and the respective host genome. Nevertheless, for miRNA analysis three different library preparation kits were used and libraries were sequenced on two different platforms. When the obtained data were compared between different methods, the results were comparable, showing the robustness of the obtained results.

The low sequencing depth of miRNAs potentially did not allow enough power to detect metformin-related changes in miRNA expression, as these changes could be pronounced in the miRNAs with a relatively low expression. Therefore, although we did not observe statistically significant changes in miRNA expression, these could in fact exist. Re-sequencing of the samples could be necessary to eliminate this doubt.

Limitations of the human study include the relatively small sample size and small sequencing depth of miRNA sequencing, thus not allowing to characterize the full scope of miRNAs present. Although this was not the aim of the human study, it would benefit from the inclusion of the analysis between patients with colorectal adenomas and colorectal carcinomas as this would allow to fully assess the role of miRNAs in the adenoma-carcinoma sequence, furthermore, paired samples (normal tissue, adenoma, carcinoma) from the same patients would be optimal for such an analysis.

4.9 Future perspectives

The interaction between host miRNAs and the gut microbiome should also be investigated in a group of type 2 diabetic patients to strengthen the translatability of miRNA analysis results in mice. However, obtaining intestinal biopsy samples from such patients is challenging; therefore, fecal samples should be considered a potential alternative.

An analysis of the exact genes targeted by miRNAs in the microbial targets would provide a valuable continuation of the study. After combining the data obtained in mice with the data from humans, candidate miRNAs for further validation of their effects *in vivo* should be selected. Similar to the existing studies investigating the potential of host miRNAs to alter the gut microbiome, synthetic miRNA mimics could be utilized for oral administration in mice to assess the changes induced in the microbiome composition and the functional implications of such treatment as well as to functionally validate specific miRNA-microbial interactions.

We have identified three differentially expressed miRNAs in intestinal biopsies of patients with colorectal adenomas, though the analysis should be repeated in a larger cohort. To investigate and advance the applicability of these results as potential biomarkers to be used for the prediction of the risk of colorectal cancer development, studies in other, less invasive sample types should be performed, for example, blood or feces. We have collected such samples from all recruited study participants, in addition to luminal content samples, allowing us to proceed with further analysis. Future studies could assess the relationship between host miRNAs and the gut microbiome in the context of various microbiome-related diseases using different types of biological material.

5 CONCLUSIONS

1. Metformin treatment changes the relative abundance of 100 bacterial species in fecal samples of high-fat diet-fed mice.
2. Metformin-induced changes in fecal microbiome composition are more pronounced in female mice than in males.
3. Most altered functional hierarchies in response to metformin treatment in high-fat diet-fed groups represent amino acid metabolism-related functions.
4. Metformin mainly exerts its microbiome-modulating effects in the small intestine; however, significant changes are induced in all intestinal sites.
5. The strongest effect of metformin is observed in the distal small intestine on the relative abundance of *Lactococcus*.
6. The relative abundance of *Lactobacillus* in the male intestinal microbiome composition is significantly lower than in females.
7. Significant differences in microbiome composition between mucosal and luminal layers of the intestinal tract exist.
8. Fecal and intestinal miRNA compositions are similar, and mice fecal miRNA content is relatively stable over time. No metformin treatment-related changes in miRNA expression have been detected.
9. Three differentially expressed miRNAs, previously associated with the development of different cancer types, have been identified between patients with and without colorectal adenomas.
10. The abundances of multiple host miRNAs and bacterial taxa are correlated in fecal and intestinal samples, and bacteria that are potential microbial targets of host miRNAs have been identified.

6 THESIS

1. Metformin changes fecal microbiome composition and substantial sex-related differences in metformin’s effects on the gut microbiome exist.
2. The effect of metformin treatment on the gut microbiome is not uniform throughout the gastrointestinal tract.
3. Host fecal and intestinal miRNAs have the potential to target specific bacterial taxa found in the respective microbiome samples in both mice and humans, suggesting a targeted microbiome modulation strategy to be studied in further functional studies.

7 PUBLICATIONS

1. **Silamiķele L**, Silamiķelis I, Ustinova M, Kalniņa Z, Elbere I, Petrovska R, Kalniņa I, Kloviņš J. Metformin Strongly Affects Gut Microbiome Composition in High-Fat Diet-Induced Type 2 Diabetes Mouse Model of Both Sexes. Front Endocrinol (Lausanne). 2021 Mar 19;12:626359. doi: 10.3389/fendo.2021.626359. PMID: 33815284; PMCID: PMC8018580.

2. **Silamiķele L**, Saksis R, Silamiķelis I, Kotoviča PP, Brīvība M, Kalniņa I, Kalniņa Z, Fridmanis D, Kloviņš J. Spatial variation of the gut microbiome in response to long-term metformin treatment in high-fat diet-induced type 2 diabetes mouse model of both sexes. Gut Microbes. 2023 Jan-Dec;15(1):2188663. doi: 10.1080/19490976.2023.2188663. PMID: 36927522; PMCID: PMC10026874.

8 APPROBATION OF RESEARCH

1. **Laila Silamikele**, Rihards Saksis, Ivars Silamikelis, Patricija Pauline Kotovica, Monta Briviba, Ineta Kalnina, Zane Kalnina, Davids Fridmanis, Janis Klovins. *Spatial variation of the gut microbiome in response to metformin treatment in high-fat diet-induced type 2 diabetes mouse model of both sexes*, Diabetes Day 2022, Symposium organized by Karolinska Institute, Stockholm, Sweden, 14.11.2022. Poster presentation.

2. **Laila Silamikele**, Patricija Kotovica, Rihards Saksis, Monta Briviba, Ineta Kalnina, Janis Klovins. *Spatial variation of gut microbiome in response to metformin treatment in high-fat diet-induced type 2 diabetes mouse model*, Inter-organ communication in physiology and disease, Symposium organized by EMBL, online, 21.-23.03.2022. Poster presentation.

3. **Laila Silamiķele**, Patrīcija Kotoviča, Ivars Silamiķelis, Jānis Kloviņš. *Saimniekorganisma miRNS un zarnu mikrobioma mijiedarbība otrā tipa cukura diabēta peļu modelī*, 80th international scientific conference of the University of Latvia, Section of Molecular Biology, Riga, Latvia, 18.02.2022. Oral presentation.

4. **Laila Silamiķele**, Patrīcija Paulīne Kotoviča, Ivars Silamiķelis, Jānis Kloviņš. *Host miRNAs are associated with gut microbiome composition in high-fat diet-induced type 2 diabetes mouse model of both sexes*, 2022 Virtual Microbiome Symposium organized by The University of Alabama at Birmingham, online, 10.01.2022. Oral presentation.

5. **Laila Silamiķele**. *Saimniekorganisma miRNS ietekme uz zarnu mikrobiomu,* X Latvian Gastroenterology Congress, Riga, Latvia, 3.-4.12.2021. Oral presentation.

6. **Laila Silamiķele**, Patrīcija Paulīne Kotoviča, Ivars Silamiķelis, Monta Ustinova, Rihards Saksis, Zane Kalniņa, Ramona Petrovska, Ilze Elbere, Ineta Kalniņa, Jānis Kloviņš. *Zarnu mikrobioma kompozīcija izmainīta metabolisma peļu modelī metformīna terapijas ietekmē*, 78th international scientific conference of the University of Latvia, Section of Molecular Biology, Riga, Latvia, 31.01.2020. Oral presentation.

7. **Laila Silamiķele**, Monta Ustinova, Ineta Kalniņa, Jānis Kloviņš. *Gut microbiome in diet-induced type 2 diabetes mouse model – Does the sex matter?*, 14th FELASA congress 2019, Prague, Czech Republic, 10.-13.06.2019. Poster presentation.

9 ACKNOWLEDGMENTS

This research was supported by the European Regional Development Fund projects “Investigation of interplay between multiple determinants influencing response to metformin: search for reliable predictors for efficacy of type 2 diabetes therapy” (Project No. 1.1.1.1/16/A/091); and “Role of miRNAs in the host-gut microbiome communication during metformin treatment in the context of metabolic disorders” (Project No. 1.1.1.1/18/A/092); and by the European Social Fund project “Strengthening of the Capacity of Doctoral Studies at the University of Latvia within the Framework of the New Doctoral Model” (identification No. 8.2.2.0/20/I/006).

I thank my supervisor Prof. Jānis Kloviņš for his guidance and trust throughout my doctoral studies. I am also grateful to my colleagues at the Latvian Biomedical Research and Study Centre for their help and practical suggestions, especially Zane Kalniņa, for her immense support and for the pleasure of exploring and advancing the field of animal science together. I thank Ineta Kalniņa and Monta Brīvība for their help with performing the animal experiment. I thank Ilona Vilkoite for the recruitment of patients undergoing colonoscopy and sample collection. I acknowledge all the patients participating in the research and the Genome Database of the Latvian Population for their help in organizing patient recruitment and sample processing. I thank all my co-authors and students for their contribution.

Finally, special thanks go to my husband (and valued co-author without whom this would not be possible), Ivars Silamiķelis, for his endless patience and responsiveness, and to my family and friends for their understanding and support.

REFERENCES

Acharya, K.D. *et al.* (2019) ‘Estradiol and high fat diet associate with changes in gut microbiota in female ob/ob mice’, *Scientific Reports*, 9(1), pp. 1–13. Available at: https://doi.org/10.1038/s41598-019-56723-1.

Aguanno, D. *et al.* (2022) ‘Modeling microbiota-associated human diseases : from minimal models to complex systems’, *Microbiome Res Rep*, 1(17), pp. 1–22. Available at: https://doi.org/10.20517/mrr.2022.01.

Ahmadi, S. *et al.* (2020) ‘Metformin reduces aging-related leaky gut and improves cognitive function by beneficially modulating gut microbiome/goblet cell/mucin axis’, *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, 75(7), pp. E9–E21. Available at: https://doi.org/10.1093/gerona/glaa056.

Aikaeli, F. *et al.* (2022) ‘Prevalence of microvascular and macrovascular complications of diabetes in newly diagnosed type 2 diabetes in low-and- middle-income countries : A systematic review and meta-analysis’, *PLOS Glob Public Health*, 2(6), pp. 1–21. Available at: https://doi.org/10.1371/journal.pgph.0000599.

Aitchison, J. (1986) *The Statistical Analysis of Compositional Data*. GBR: Chapman & Hall, Ltd.

American Diabetes Association (2013) ‘Diagnosis and Classification of Diabetes Mellitus’, *Diabetes Care*, 36(Suppl 1), pp. S67–S74. Available at: https://doi.org/10.2337/dc13-S067.

American Diabetes Association (2019) ‘2 . Classification and Diagnosis of Diabetes : Standards of Medical Care in Diabetes - 2019’, *Diabetes Care*, 42(Suppl. 1), pp. S13–S28. Available at: https://doi.org/10.2337/dc19-S002.

Andersen, D.B. and Holst, J.J. (2022) ‘Peptides Peptides in the regulation of glucagon secretion’, *Peptides*, 148(October 2021), p. 170683. Available at: https://doi.org/10.1016/j.peptides.2021.170683.

Anderson, J.E. (2020) ‘Combining Glucagon-Like Peptide 1 Receptor Agonists and Sodium – Glucose Cotransporter 2 Inhibitors to Target Multiple Organ Defects in Type 2 Diabetes’, *Diabetes Spectrum*, 33(2), pp. 165–174. Available at: https://doi.org/10.2337/ds19-0031.

Andrews, S. (2010) ‘FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]’. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Ashammakhi, N. *et al.* (2020) ‘Gut-on-a-chip: Current Progress and Future Opportunities’, *Biomaterials*, 255. Available at: https://doi.org/10.1016/j.biomaterials.2020.120196.Gut-on-a-chip.

Bailey, C.J. (2017) ‘Metformin : historical overview’, *Diabetologia*, 60(9), pp. 1566–1576. Available at: https://doi.org/10.1007/s00125-017-4318-z.

Barnett, D., Arts, I. and Penders, J. (2021) ‘microViz: an R package for microbiome data visualization and statistics’, *Journal of Open Source Software*, 6(63), p. 3201. Available at: https://doi.org/10.21105/joss.03201.

Bartel, D.P. (2018) ‘Metazoan MicroRNAs’, *Cell*, 173(1), pp. 20–51. Available at: https://doi.org/10.1016/j.cell.2018.03.006.

Bassanini, G. *et al.* (2019) ‘Phenylketonuria diet promotes shifts in firmicutes populations’, *Frontiers in Cellular and Infection Microbiology*, 9(MAR), pp. 1–9. Available at: https://doi.org/10.3389/fcimb.2019.00101.

Bauer, P. V. *et al.* (2018) ‘Metformin Alters Upper Small Intestinal Microbiota that Impact a Glucose-SGLT1-Sensing Glucoregulatory Pathway’, *Cell Metabolism*, 27(1), pp. 101-117.e5. Available at: https://doi.org/10.1016/j.cmet.2017.09.019.

Belizário, J.E. and Napolitano, M. (2015) ‘Human microbiomes and their roles in dysbiosis, common diseases, and novel therapeutic approaches’, *Frontiers in Microbiology*, 6(OCT), pp. 1–16. Available at: https://doi.org/10.3389/fmicb.2015.01050.

Berg, G. *et al.* (2020) ‘Microbiome definition re-visited: old concepts and new challenges’, *Microbiome*, 8, pp. 1–22.

Bibbo, S. *et al.* (2020) ‘Fecal Microbiota Transplantation : Screening and Selection to Choose the Optimal Donor’, *Journal of Clinical Medicine*, 9(6).

Bocanegra, A. *et al.* (2021) ‘Whole Alga , Algal Extracts , and Compounds as Ingredients of Functional Foods : Composition and Action Mechanism Relationships in the Prevention and Treatment of Type-2 Diabetes Mellitus’, *International Journal of Molecular Sciences*, 22(8).

Bokulich, N.A. *et al.* (2018) ‘Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2’s q2-feature-classifier plugin’, *Microbiome*, 6(1), pp. 1–17. Available at: https://doi.org/10.1186/s40168-018-0470-z.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) ‘Trimmomatic: A flexible trimmer for Illumina sequence data’, *Bioinformatics*, 30(15), pp. 2114–2120. Available at: https://doi.org/10.1093/bioinformatics/btu170.

Bolyen, E. *et al.* (2019) ‘Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2’, *Nature Biotechnology*, 37(8), pp. 852–857. Available at: https://doi.org/10.1038/s41587-019-0209-9.

Borgo, F. *et al.* (2018) ‘Body mass index and sex affect diverse microbial niches within the gut’, *Frontiers in Microbiology*, 9(FEB), pp. 1–12. Available at: https://doi.org/10.3389/fmicb.2018.00213.

Brandt, A. *et al.* (2019) ‘Metformin attenuates the onset of non-alcoholic fatty liver disease and affects intestinal microbiota and barrier in small intestine’, *Scientific Reports*, 9(1), pp. 1–14. Available at: https://doi.org/10.1038/s41598-019-43228-0.

Bravard, A. *et al.* (2021) ‘Metformin treatment for 8 days impacts multiple intestinal parameters in high-fat high-sucrose fed mice’, *Scientific Reports*, 11(1), pp. 1–12. Available at: https://doi.org/10.1038/s41598-021-95117-0.

Bryrup, T. *et al.* (2019) ‘Metformin-induced changes of the gut microbiota in healthy young men: results of a non-blinded, one-armed intervention study’, *Diabetologia*, 62(6), pp. 1024–1035. Available at: https://doi.org/10.1007/s00125-019-4848-7.

Callahan, B.J. *et al.* (2016) ‘DADA2: High resolution sample inference from Illumina amplicon data’, *Nature Methods*, 13(7), pp. 581–583. Available at: https://doi.org/10.1038/nmeth.3869.

Campbell, R.K. (2011) ‘Clarifying the Role of Incretin-Based Therapies in the Treatment of Type 2 Diabetes Mellitus’, *CLITHE*, 33(5), pp. 511–527. Available at: https://doi.org/10.1016/j.clinthera.2011.04.015.

Casimiro, I. *et al.* (2021) ‘Phenotypic sexual dimorphism in response to dietary fat manipulation in C57BL/6J mice’, *Journal of Diabetes and its Complications*, 35(2), p. 107795. Available at: https://doi.org/10.1016/j.jdiacomp.2020.107795.

Cerf, M.E. (2013) ‘Beta cell dysfunction and insulin resistance’, *Frontiers in Endocrinology*, 4(March), pp. 1–12. Available at: https://doi.org/10.3389/fendo.2013.00037.

Chaluvadi, S., Hotchkiss, A.T. and Yam, K.L. (2016) *Gut Microbiota: Impact of Probiotics, Prebiotics, Synbiotics, Pharmabiotics, and Postbiotics on Human Health*, *Gut Microbiota. Probiotics, Prebiotics, and Synbiotics*. Elsevier Inc. Available at: https://doi.org/10.1016/B978-0-12-802189-7.00036-8.

Chen, F. *et al.* (2022) ‘Integrated analysis of the faecal metagenome and serum metabolome reveals the role of gut associated metabolites in the detection of colorectal cancer and adenoma’, *Gut*, 71, pp. 1315–1325. Available at: https://doi.org/10.1136/gutjnl-2020-323476.

Chen, H.M. *et al.* (2022) ‘Exploration of the relationship between gut microbiota and fecal microRNAs in patients with major depressive disorder’, *Scientific Reports*, 12(1), pp. 1–10. Available at: https://doi.org/10.1038/s41598-022-24773-7.

Chen, S. *et al.* (2022) ‘Metformin in aging and aging-related diseases : clinical applications and relevant mechanisms’, *Theranostics*, 12(6), pp. 2722–2740. Available at: https://doi.org/10.7150/thno.71360.

Chipman, L.B. and Pasquinelli, A.E. (2020) ‘MiRNA Targeting – Growing Beyond the Seed’, *Trends Genet.*, 35(3), pp. 215–222. Available at: https://doi.org/10.1016/j.tig.2018.12.005.MiRNA.

Choi, J.W. *et al.* (2017) ‘Tiny RNAs and their voyage via extracellular vesicles: Secretion of bacterial small RNA and eukaryotic microRNA’, *Experimental Biology and Medicine*, 242(15), pp. 1475–1481. Available at: https://doi.org/10.1177/1535370217723166.

Clavel, T. *et al.* (2014) ‘Intestinal microbiota in metabolic diseases: From bacterial community structure and functions to species of pathophysiological relevance’, *Gut Microbes*, 5(4), pp. 544–551. Available at: https://doi.org/10.4161/gmic.29331.

Coker, O.O. *et al.* (2022) ‘Altered gut metabolites and microbiota interactions are implicated in colorectal carcinogenesis and can be non-invasive diagnostic biomarkers’, *Microbiome*, 10(1), pp. 1–12. Available at: https://doi.org/10.1186/s40168-021-01208-5.

Crespo-Piazuelo, D. *et al.* (2019) ‘Association between the pig genome and its gut microbiota composition’, *Scientific Reports*, 9(1), pp. 1–11. Available at: https://doi.org/10.1038/s41598-019-45066-6.

Dalby, M.J. *et al.* (2017) ‘Dietary Uncoupling of Gut Microbiota and Energy Harvesting from Obesity and Glucose Tolerance in Mice’, *Cell Reports*, 21(6), pp. 1521–1533. Available at: https://doi.org/10.1016/j.celrep.2017.10.056.

Damasceno, D.C. *et al.* (2014) ‘Streptozotocin-Induced Diabetes Models: Pathophysiological Mechanisms and Fetal Outcomes’, *BioMed Research International*, 2014. Available at: https://doi.org/10.1155/2014/819065.

Davani-Davari, D. *et al.* (2019) ‘Prebiotics: Definition, Types, Sources, Mechanisms, and Clinical Applications’, *Foods*, 8(3), p. 92. Available at: https://doi.org/10.3390/foods8030092.

Davies, M.J. *et al.* (2018) ‘Management of hyperglycemia in type 2 diabetes, 2018. A consensus report by the American Diabetes Association (ADA) and the european association for the study of diabetes (EASD)’, *Diabetes Care*, 41(12), pp. 2669–2701. Available at: https://doi.org/10.2337/dci18-0033.

Debelius, J.W. *et al.* (2021) ‘A comparison of approaches to scaffolding multiple regions along the 16S rRNA gene 1 for improved resolution 2’, *bioRxiv*, p. 2021.03.23.436606. Available at: https://doi.org/10.1101/2021.03.23.436606.

DeFronzo, R.A. (2004) ‘Pathogenesis of type 2 diabetes mellitus’, *Med Clin N Am*, 88, pp. 787–835. Available at: https://doi.org/10.1016/j.mcna.2004.04.013.

Dickson, I. (2019) ‘Stability and individuality of adult microbiota’, *Nature*, 108(June), p. S11.

Diez-Gutiérrez, L. *et al.* (2020) ‘Gamma-aminobutyric acid and probiotics : Multiple health benefits and their future in the global functional food and nutraceuticals market’, *Journal of Functional Foods*, 64(August 2019), p. 103669. Available at: https://doi.org/10.1016/j.jff.2019.103669.

Dominguez-Bello, M.G. *et al.* (2010) ‘Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns’, *Proc Natl Acad Sci USA*, 107(26), pp. 11971–11975. Available at: https://doi.org/10.1073/pnas.1002601107.

Donaldson, G.P., Lee, S.M. and Mazmanian, S.K. (2015) ‘Gut biogeography of the bacterial microbiota’, *Nature Reviews Microbiology*, 14(1), pp. 20–32. Available at: https://doi.org/10.1038/nrmicro3552.

Dong, T.S. *et al.* (2019) ‘Metformin alters the duodenal microbiome and decreases the incidence of pancreatic ductal adenocarcinoma promoted by diet-induced obesity’, *American Journal of Physiology - Gastrointestinal and Liver Physiology*, 317(6), pp. G763–G772. Available at: https://doi.org/10.1152/ajpgi.00170.2019.

Dujic, T. *et al.* (2016) ‘Organic cation transporter 1 variants and gastrointestinal side effects of metformin in patients with Type 2 diabetes’, *Diabetic Medicine*, 33(4), pp. 511–514. Available at: https://doi.org/10.1111/dme.13040.

Elbere, I. *et al.* (2018) ‘Association of metformin administration with gut microbiome dysbiosis in healthy volunteers’, *PLoS ONE*, 13(9), pp. 1–17. Available at: https://doi.org/10.1371/journal.pone.0204317.

Elzinga, J. *et al.* (2019) ‘The Use of Defined Microbial Communities To Model Host-Microbe Interactions in the Human Gut’, *Microbiol Mol Biol Rev*, 83(2), pp. e00054-18.

Ewels, P. *et al.* (2016) ‘MultiQC: Summarize analysis results for multiple tools and samples in a single report’, *Bioinformatics*, 32(19), pp. 3047–3048. Available at: https://doi.org/10.1093/bioinformatics/btw354.

Feng, Q. *et al.* (2015) ‘Gut microbiome development along the colorectal adenoma-carcinoma sequence’, *Nature Communications*, 6. Available at: https://doi.org/10.1038/ncomms7528.

Festing, M.F.W. *et al.* (2016) *The Design of Animal Experiments: Reducing the use of animals in research through better experimental design (Laboratory Animal Handbooks) Second Edition*. SAGE Publications Ltd.

Florez, J.C. (2017) ‘The pharmacogenetics of metformin’, *Diabetologia*, 60, pp. 1648–1655. Available at: https://doi.org/10.1007/s00125-017-4335-y.

Fong, W., Li, Q. and Yu, J. (2020) ‘Gut microbiota modulation : a novel strategy for prevention and treatment of colorectal cancer’, *Oncogene*, pp. 4925–4943. Available at: https://doi.org/10.1038/s41388-020-1341-1.

Forslund, K. *et al.* (2015) ‘Disentangling the effects of type 2 diabetes and metformin on the human gut microbiota’, *Nature*, 528(7581), pp. 262–266. Available at: https://doi.org/10.1038/nature15766.Disentangling.

Fritz, J. V *et al.* (2013) ‘From meta-omics to causality: experimental models for human microbiome research’, *Microbiome*, 1(14), pp. 1–15.

Fuks, G. *et al.* (2018) ‘Combining 16S rRNA gene variable regions enables high-resolution microbial community profiling’, *Microbiome*, 6(1), pp. 1–13. Available at: https://doi.org/10.1186/s40168-017-0396-x.

Gil-Zamorano, J. *et al.* (2020) ‘Intestinal miRNAs regulated in response to dietary lipids’, *Scientific Reports*, 10(1), pp. 1–15. Available at: https://doi.org/10.1038/s41598-020-75751-w.

Graham, G.G. *et al.* (2011) ‘Clinical pharmacokinetics of metformin’, *Clinical Pharmacokinetics*, 50(2), pp. 81–98. Available at: https://doi.org/10.2165/11534750-000000000-00000.

Greenfield, J.R. and Chisholm, D.J. (2004) ‘Thiazolidinediones – mechanisms of action’, *Australian Prescriber*, 27(3), pp. 67–70.

Gu, S. *et al.* (2013) ‘Bacterial Community Mapping of the Mouse Gastrointestinal Tract’, *PLoS ONE*, 8(10). Available at: https://doi.org/10.1371/journal.pone.0074957.

Gu, Z. *et al.* (2021) ‘Inhibition of MicroRNA miR-101-3p on prostate cancer progression by regulating Cullin 4B (CUL4B) and PI3K / AKT / mTOR signaling pathways’, *Bioengineered*, 12(1), pp. 4719–4735. Available at: https://doi.org/10.1080/21655979.2021.1949513.

Gulnaz, A. *et al.* (2021) ‘Lactobacillus Sps in reducing the risk of diabetes in high-fat diet-induced diabetic mice by modulating the gut microbiome and inhibiting key digestive enzymes associated with diabetes’, *Biology*, 10(4). Available at: https://doi.org/10.3390/biology10040348.

Guo, Y. *et al.* (2020) ‘miRNA-10a-5p Alleviates Insulin Resistance and Maintains Diurnal Patterns of Triglycerides and Gut Microbiota in High-Fat Diet-Fed Mice’, *Mediators of Inflammation* [Preprint].

Gurung, M. *et al.* (2020) ‘Role of gut microbiota in type 2 diabetes pathophysiology’, *EBioMedicine*, 51, p. 102590. Available at: https://doi.org/10.1016/j.ebiom.2019.11.051.

Hatting, M. *et al.* (2019) ‘Insulin regulation of gluconeogenesis’, 1411(1), pp. 21–35. Available at: https://doi.org/10.1111/nyas.13435.Insulin.

He, L. *et al.* (2022) ‘Fecal miR-142a-3p from dextran sulfate sodium- challenge recovered mice prevents colitis by promoting the growth of Lactobacillus reuteri’, *Molecular Therapy*, 30(1), pp. 388–399. Available at: https://doi.org/10.1016/j.ymthe.2021.08.025.

He, L. and Wondisford, F.E. (2015) ‘Metformin action: Concentrations matter’, *Cell Metabolism*, 21(2), pp. 159–162. Available at: https://doi.org/10.1016/j.cmet.2015.01.003.

Herp, S. *et al.* (2021) ‘The human symbiont Mucispirillum schaedleri: causality in health and disease’, *Medical Microbiology and Immunology*, 210(4), pp. 173–179. Available at: https://doi.org/10.1007/s00430-021-00702-9.

Herszényi, L. (2019) ‘The “ Difficult ” Colorectal Polyps and Adenomas: Practical Aspects’, *Digestive Diseases*, 37, pp. 394–399. Available at: https://doi.org/10.1159/000495694.

Hewel, C. *et al.* (2019) ‘Common miRNA Patterns of Alzheimer’s Disease and Parkinson’s Disease and Their Putative Impact on Commensal Gut Microbiota’, *Frontiers in Neuroscience*, 13(March), pp. 1–13. Available at: https://doi.org/10.3389/fnins.2019.00113.

Heydemann, A. (2016) ‘An Overview of Murine High Fat Diet as a Model for Type 2 Diabetes Mellitus’, *Journal of Diabetes Research*, 2016. Available at: https://doi.org/10.1155/2016/2902351.

Horne, R. *et al.* (2019) ‘Microbe and host interaction in gastrointestinal homeostasis’, *Psychopharmacology*, 236(5), pp. 1623–1640. Available at: https://doi.org/10.1007/s00213-019-05218-y.

Hossain, U. *et al.* (2020) ‘An overview on the role of bioactive α-glucosidase inhibitors in ameliorating diabetic complications’, *Food and Chemical Toxicology* [Preprint], (Nov).

Hou, K. *et al.* (2022) ‘Microbiota in health and diseases’, *Sig Transduct Target Ther*, 135. Available at: https://doi.org/10.1038/s41392-022-00974-4.

Hou, N., Huo, D. and Dignam, J.J. (2013) ‘Prevention of colorectal cancer and dietary management’, *Chin Clin Oncol*, 2(2), pp. 1–20. Available at: https://doi.org/10.3978/j.issn.2304-3865.2013.04.03.

Houghteling, P.D. and Walker, W.A. (2016) ‘Why is initial bacterial colonization of the intestine important to the infant’s and child’s health?’, *J Pediatr Gastroenterol Nutr.*, 60(3), pp. 294–307. Available at: https://doi.org/10.1097/MPG.0000000000000597.Why.

Hugenholtz, F. and Vos, W.M. De (2018) ‘Mouse models for human intestinal microbiota research : a critical evaluation’, *Cellular and Molecular Life Sciences*, 75(1), pp. 149–160. Available at: https://doi.org/10.1007/s00018-017-2693-8.

Hunter, J.D. (2007) ‘Matplotlib: A 2D graphics environment’, *Computing in Science and Engineering*, 9(3), pp. 90–95. Available at: https://doi.org/10.1109/MCSE.2007.55.

Hutchings, M.I., Truman, A.W. and Wilkinson, B. (2020) ‘Antibiotics : past , present and future’, *Current Opinion in Microbiology*, 51(Figure 1), pp. 72–80. Available at: https://doi.org/10.1016/j.mib.2019.10.008.

Jandhyala, S.M. *et al.* (2015) ‘Role of the normal gut microbiota’, *World Journal of Gastroenterology*, 21(29), pp. 8836–8847. Available at: https://doi.org/10.3748/wjg.v21.i29.8787.

Jauvain, M. *et al.* (2021) ‘Metformin modifies the gut microbiota of mice infected with helicobacter pylori’, *Pharmaceuticals*, 14(4). Available at: https://doi.org/10.3390/ph14040329.

Ji, S., Wang, L. and Li, L. (2019) ‘Effect of Metformin on Short-Term High-Fat Diet-Induced Weight Gain and Anxiety-Like Behavior and the Gut Microbiota’, *Frontiers in Endocrinology*, 10(October), pp. 1–11. Available at: https://doi.org/10.3389/fendo.2019.00704.

Ji, Y. *et al.* (2018) ‘Faecal microRNA as a biomarker of the activity and prognosis of inflammatory bowel diseases’, *Biochemical and Biophysical Research Communications*, 503(4), pp. 2443–2450. Available at: https://doi.org/10.1016/j.bbrc.2018.06.174.

Jung, M. *et al.* (2010) ‘Robust MicroRNA stability in degraded RNA preparations from human tissue and cell samples’, *Clinical Chemistry*, 56(6), pp. 998–1006. Available at: https://doi.org/10.1373/clinchem.2009.141580.

Kaliannan, K. *et al.* (2018) ‘Estrogen-mediated gut microbiome alterations influence sexual dimorphism in metabolic syndrome in mice’, *Microbiome*, 6(1), pp. 1–22. Available at: https://doi.org/10.1186/s40168-018-0587-0.

Kant, R. *et al.* (2016) ‘Genome sequence of the butyrate-producing anaerobic bacterium Anaerostipes hadrus PEL 85’, *Genome Announcements*, 3(2), pp. 3–4. Available at: https://doi.org/10.1128/genomeA.00224-15.

Kara, M. *et al.* (2015) ‘Differential expressions of cancer-associated genes and their regulatory miRNAs in colorectal carcinoma’, *Gene* [Preprint]. Available at: https://doi.org/10.1016/j.gene.2015.04.065.

Karlsson, F.H. *et al.* (2013) ‘Gut metagenome in European women with normal, impaired and diabetic glucose control’, *Nature*, 498(7452), pp. 99–103. Available at: https://doi.org/10.1038/nature12198.

Khan, M.A.B. *et al.* (2020) ‘Epidemiology of Type 2 Diabetes – Global Burden of Disease and Forecasted Trends’, *Journal of Epidemiology and Global Health*, 10(1), pp. 107–111.

Kim, H. *et al.* (2020) ‘Birth mode, breastfeeding, pet exposure, and antibiotic use: associations with the gut microbiome and sensitization in children’, *Curr Allergy Asthma Rep.*, 19(4), pp. 17–19. Available at: https://doi.org/10.1007/s11882-019-0851-9.Birth.

Kim, H.J. *et al.* (2016) ‘Contributions of microbiome and mechanical deformation to intestinal bacterial overgrowth and inflammation in a human gut-on-a-chip’, *Proceedings of the National Academy of Sciences of the United States of America*, 113(1), pp. E7–E15. Available at: https://doi.org/10.1073/pnas.1522193112.

Kinaan, M., Ding, H. and Triggle, C.R. (2015) ‘Metformin: An Old Drug for the Treatment of Diabetes but a New Drug for the Protection of the Endothelium’, *Medical Principles and Practice*, 24(5), pp. 401–415. Available at: https://doi.org/10.1159/000381643.

King, A.J.F. (2012) ‘The use of animal models in diabetes research’, *British Journal of Pharmacology*, 166(3), pp. 877–894. Available at: https://doi.org/10.1111/j.1476-5381.2012.01911.x.

Kirchgessner, A. (2016) *The Role of the Microbiota and Potential for Dietary Intervention in Chronic Fatigue Syndrome, The Gut-Brain Axis Dietary, Probiotic, and Prebiotic Interventions on the Microbiota*. Elsevier Inc. Available at: https://doi.org/10.1016/B978-0-12-802304-4.00021-9.

Kläring, K. *et al.* (2013) ‘Intestinimonas butyriciproducens gen. nov., sp. nov., a butyrate-producing bacterium from the mouse intestine’, *International Journal of Systematic and Evolutionary Microbiology*, 63(PART 12), pp. 4606–4612. Available at: https://doi.org/10.1099/ijs.0.051441-0.

Kooij, I.A. *et al.* (2016) ‘The immunology of the vermiform appendix: a review of the literature’, *Clinical ans Experimental Immunology*, 186(1), pp. 1–9. Available at: https://doi.org/10.1111/cei.12821.

Kopkova, A. *et al.* (2018) ‘MicroRNA isolation and quantification in cerebrospinal fluid: A comparative methodical study’, *PLoS ONE*, 13(12), pp. 1–12. Available at: https://doi.org/10.1371/journal.pone.0208580.

Kostic, A.D. *et al.* (2013) 'Fusobacterium nucleatum potentiates intestinal tumorigenesis and modulates the tumor immune microenvironment', *Cell Host Microbe*, 14(2), pp. 207–215. Available at: https://doi.org/10.1016/j.chom.2013.07.007.Fusobacterium.

Kottaisamy, P.C.D. *et al.* (2021) 'Experimental animal models for diabetes and its related complications — a review', *Laboratory Animal Research*, 37(1), pp. 1–14. Available at: https://doi.org/10.1186/s42826-021-00101-4.

Krych, L. *et al.* (2013) 'Quantitatively Different, yet Qualitatively Alike: A Meta- Analysis of the Mouse Core Gut Microbiome with a View towards the Human Gut Microbiome', *PLoS ONE*, 8(5). Available at: https://doi.org/10.1371/journal.pone.0062578.

De La Cuesta-Zuluaga, J. *et al.* (2017) 'Metformin is associated with higher relative abundance of mucin-degrading akkermansia muciniphila and several short-chain fatty acid-producing microbiota in the gut', *Diabetes Care*, 40(1), pp. 54–62. Available at: https://doi.org/10.2337/dc16-1324.

Lamoia, T.E. and Shulman, G.I. (2021) 'Cellular and Molecular Mechanisms of Metformin Action', *Endocrine Reviews*, 42(1), pp. 77–96. Available at: https://doi.org/10.1210/endrev/bnaa023.

Larsen, N. *et al.* (2010) 'Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults', *PLoS ONE*, 5(2). Available at: https://doi.org/10.1371/journal.pone.0009085.

Laukens, D. *et al.* (2015) 'Heterogeneity of the gut microbiome in mice: Guidelines for optimizing experimental design', *FEMS Microbiology Reviews*, 40(1), pp. 117–132. Available at: https://doi.org/10.1093/femsre/fuv036.

Lee, H. (2014) 'Effect of Metformin on Metabolic Improvement and Gut Microbiota', *Applied and Environmental Microbiology*, 80(19), pp. 5935–5943. Available at: https://doi.org/10.1128/AEM.01357-14.

Lee, H. *et al.* (2018) 'Modulation of the gut microbiota by metformin improves metabolic profiles in aged obese mice', *Gut Microbes*, 9(2), pp. 155–165. Available at: https://doi.org/10.1080/19490976.2017.1405209.

Lee, Y. *et al.* (2017) 'Hyperglycemia- and hyperinsulinemia-induced insulin resistance causes alterations in cellular bioenergetics and activation of inflammatory signaling in lymphatic muscle', *FASEB Journal*, 31(7), pp. 2744–2759. Available at: https://doi.org/10.1096/fj.201600887R.

Leitao, A.L. and Enguita, F.J. (2022) 'A Structural View of miRNA Biogenesis and Function', *Non-coding RNA*, 8(10).

Leviatan, S. *et al.* (2022) 'An expanded reference map of the human gut microbiome reveals hundreds of previously unknown species', *Nature Communications*, 13. Available at: https://doi.org/10.1038/s41467-022-31502-1.

Li, J. *et al.* (2019) 'MicroRNA expression profiling in the colorectal normal-adenoma-carcinoma transition', *Oncology Letters*, 18(2), pp. 2013–2018. Available at: https://doi.org/10.3892/ol.2019.10464.

Li, L. *et al.* (2019) 'An in vitro model maintaining taxon-specific functional activities of the gut microbiome', *Nature Communications*, 10, p. 4146. Available at: https://doi.org/10.1038/s41467-019-12087-8.

Li, Qingzhong *et al.* (2020) 'Combination of Oligofructose and Metformin Alters the Gut Microbiota and Improves Metabolic Profiles, Contributing to the Potentiated Therapeutic Effects on Diet-Induced Obese

Animals', *Frontiers in Endocrinology*, 10(February), pp. 1–14. Available at: https://doi.org/10.3389/fendo.2019.00939.

Li, Qian *et al.* (2020) 'Implication of the gut microbiome composition of type 2 diabetic patients from northern China', *Scientific Reports*, 10(1), pp. 1–8. Available at: https://doi.org/10.1038/s41598-020-62224-3.

Li, X.-N. *et al.* (2019) 'Circular RNA circVAPA is up-regulated and exerts oncogenic properties by sponging miR-101 in colorectal cancer', *Biomedicine & Pharmacotherapy*, 112(January), p. 108611. Available at: https://doi.org/10.1016/j.biopha.2019.108611.

Liang, H. *et al.* (2013) 'New roles for microRNAs in cross-species communication', *RNA Biology*, 10(3), pp. 373–376. Available at: https://doi.org/10.4161/rna.23663.

Liang, X. and Giacomini, K.M. (2017) 'Transporters Involved in Metformin Pharmacokinetics and Treatment Response', *Journal of Pharmaceutical Sciences*, 106(9), pp. 2245–2250. Available at: https://doi.org/10.1016/j.xphs.2017.04.078.

Lin, H. and Peddada, S. Das (2020) 'Analysis of compositions of microbiomes with bias correction', *Nature Communications*, 11(1), pp. 1–11. Available at: https://doi.org/10.1038/s41467-020-17041-7.

Lin, S. and Gregory, R.I. (2015) 'MicroRNA biogenesis pathways in cancer', *Nat Rev Cancer*, 15(6), pp. 321–333. Available at: https://doi.org/10.1038/nrc3932.MicroRNA.

Liu, G. and Li, B. (2019) 'Role of miRNA in transformation from normal tissue to colorectal adenoma and cancer', *Journal of Cancer Research and Therapeutics*, 15(2), pp. 278–285. Available at: https://doi.org/10.4103/jcrt.JCRT.

Liu, S. *et al.* (2016) 'The Host Shapes the Gut Microbiota via Fecal microRNA', *Cell Host Microbe*, 19(1), pp. 32–43. Available at: https://doi.org/10.1016/j.chom.2015.12.005.

Liu, S. *et al.* (2019) 'Oral Administration of miR-30d from Feces of MS Patients Suppresses MS-like Symptoms in Mice by Expanding Akkermansia muciniphila', *Cell Host and Microbe*, 26(6), pp. 779-794.e8. Available at: https://doi.org/10.1016/j.chom.2019.10.008.

Liu, X.-Y. *et al.* (2015) 'MicroRNA-101-3p suppresses cell proliferation , invasion and enhances chemotherapeutic sensitivity in salivary gland adenoid cystic carcinoma by targeting Pim-1', *Am J Cancer Res*, 5(10), pp. 3015–3029.

Liu, Y. *et al.* (2019) 'Examination of the temporal and spatial dynamics of the gut microbiome in newborn piglets reveals distinct microbial communities in six intestinal segments', *Scientific Reports*, 9(1), pp. 1–8. Available at: https://doi.org/10.1038/s41598-019-40235-z.

Liu, Z. *et al.* (2021) 'Metformin Affects Gut Microbiota Composition and Diversity Associated with Amelioration of Dextran Sulfate Sodium-Induced Colitis in Mice', *Frontiers in Pharmacology*, 12(May), pp. 1–9. Available at: https://doi.org/10.3389/fphar.2021.640347.

Lkhagva, E. *et al.* (2021) 'The Regional Diversity of Gut Microbiome Along the GI Tract of male C57BL/6 mice', *BMC Microbiology*, 21(44), pp. 1–13.

Loy, A. *et al.* (2017) 'Lifestyle and Horizontal Gene Transfer- Mediated Evolution of Mucispirillum schaedleri, a Core Member of the Murine Gut Microbiota', *mSystems*, 2(1), pp. e00171-16.

Lu, J. *et al.* (2017) ‘Bracken: estimating species abundance in metagenomics data’, *PeerJ Computer Science*, 3, p. e104. Available at: https://doi.org/10.7717/peerj-cs.104.

Mannucci, E. *et al.* (2022) *Italian guidelines for the treatment of type 2 diabetes*, *Acta Diabetologica*. Springer Milan. Available at: https://doi.org/10.1007/s00592-022-01857-4.

Martín-Fernández, J.A. *et al.* (2015) ‘Bayesian-multiplicative treatment of count zeros in compositional data sets’, *Statistical Modelling*, 15(2), pp. 134–158. Available at: https://doi.org/10.1177/1471082X14535524.

Martin, M. (2011) ‘Cutadapt removes adapter sequences from high-throughput sequencing reads’, *EMBnet.Journal*, 17(1), pp. 10–12. Available at: https://doi.org/https://doi.org/10.14806/ej.17.1.200.

McCreight, L.J. *et al.* (2018) ‘Pharmacokinetics of metformin in patients with gastrointestinal intolerance’, *Diabetes, Obesity and Metabolism*, 20(7), pp. 1593–1601. Available at: https://doi.org/10.1111/dom.13264.

McCreight, L.J., Bailey, C.J. and Pearson, E.R. (2016) ‘Metformin and the gastrointestinal tract’, *Diabetologia*, 59(3), pp. 426–435. Available at: https://doi.org/10.1007/s00125-015-3844-9.

McKinney, W. (2010) ‘Data Structures for Statistical Computing in Python’, *Proceedings of the 9th Python in Science Conference*, 1(Scipy), pp. 56–61. Available at: https://doi.org/10.25080/majora-92bf1922-00a.

McKnite, A.M. *et al.* (2012) ‘Murine gut microbiota is defined by host genetics and modulates variation of metabolic traits’, *PLoS ONE*, 7(6). Available at: https://doi.org/10.1371/journal.pone.0039191.

McMurdie, P.J. and Holmes, S. (2013) ‘Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data’, *PLoS ONE*, 8(4). Available at: https://doi.org/10.1371/journal.pone.0061217.

Mende, D.R. *et al.* (2020) ‘ProGenomes2: An improved database for accurate and consistent habitat, taxonomic and functional annotations of prokaryotic genomes’, *Nucleic Acids Research*, 48(D1), pp. D621–D625. Available at: https://doi.org/10.1093/nar/gkz1002.

Mirarab S, Nguyen N, W.T. (2012) *SEPP: SATé-enabled phylogenetic placement*, *Pac Symp Biocomput.* Available at: https://doi.org/10.1142/9789814366496_0024.

Mobasseri, M. *et al.* (2020) ‘Prevalence and incidence of type 1 diabetes in the world : a systematic review and meta-analysis’, 10(2), pp. 98–115. Available at: https://doi.org/10.34172/hpp.2020.18.

Moghbeli, M. *et al.* (2021) ‘Molecular mechanisms of the microRNA-132 during tumor progressions’, *Cancer Cell International*, 21(439), pp. 1–11. Available at: https://doi.org/10.1186/s12935-021-02149-7.

Mohan, V. *et al.* (2020) ‘Management of Type 2 Diabetes in Developing Countries : Balancing Optimal Glycaemic Control and Outcomes with Affordability and Accessibility to Treatment’, pp. 15–35. Available at: https://doi.org/10.1007/s13300-019-00733-9.

Molina Ortiz, J.P. *et al.* (2021) ‘Enabling rational gut microbiome manipulations by understanding gut ecology through experimentally-evidenced in silico models’, *Gut Microbes*, 13(1), pp. 1–19. Available at: https://doi.org/10.1080/19490976.2021.1965698.

Moloney, G.M. *et al.* (2018) ‘Faecal microRNAs: Indicators of imbalance at the host-microbe interface?’, *Beneficial Microbes*, 9(2), pp. 175–183. Available at: https://doi.org/10.3920/BM2017.0013.

Nagpal, R. *et al.* (2018) ‘Comparative Microbiome Signatures and Short-Chain Fatty Acids in Mouse, Rat, Non-human Primate, and Human Feces’, *Frontiers in Microbiology*, 9(November), pp. 1–13. Available at: https://doi.org/10.3389/fmicb.2018.02897.

Nair, A.A. *et al.* (2020) ‘Frequency of MicroRNA Response Elements Identifies Pathologically Relevant Signaling Pathways in Triple-Negative Breast Cancer’, *iScience*, 23(6), p. 101249. Available at: https://doi.org/10.1016/j.isci.2020.101249.

Napolitano, A. *et al.* (2014) ‘Novel gut-based pharmacology of metformin in patients with type 2 diabetes mellitus’, *PLoS ONE*, 9(7), pp. 1–14. Available at: https://doi.org/10.1371/journal.pone.0100778.

Nasri, H. and Rafieian-Kopaei, M. (2014) ‘Metformin : Current knowledge’, *Journal of Research in Medical Sciences*, 19(7), pp. 658–664.

Nguyen, L.H., Goel, A. and Chung, D.C. (2020) ‘Pathways of Colorectal Carcinogenesis’, *Gastroenterology*, 158(2), pp. 291–302. Available at: https://doi.org/10.1053/j.gastro.2019.08.059.Pathways.

Nguyen, T.L.A. *et al.* (2015) ‘How informative is the mouse for human gut microbiota research?’, *Disease Models & Mechanisms*, 8(1), pp. 1–16. Available at: https://doi.org/10.1242/dmm.017400.

Nguyen, T.T.P. *et al.* (2022) ‘The Role of miR-29s in Human Cancers — An Update’, *Biomedicines*, 10, pp. 1–17.

O’Brien, J. *et al.* (2018) ‘Overview of microRNA biogenesis, mechanisms of actions, and circulation’, *Frontiers in Endocrinology*, 9(AUG), pp. 1–12. Available at: https://doi.org/10.3389/fendo.2018.00402.

Ogunrinola, G.A. *et al.* (2020) ‘The Human Microbiome and Its Impacts on Health’, *Int J Microbiol.* [Preprint], (Jun).

Øines, M. *et al.* (2017) ‘Epidemiology and risk factors of colorectal polyps’, *Best Practice and Research: Clinical Gastroenterology*, 31(4), pp. 419–424. Available at: https://doi.org/10.1016/j.bpg.2017.06.004.

Oliphant, K. and Allen-Vercoe, E. (2019) ‘Macronutrient metabolism by the human gut microbiome: Major fermentation by-products and their impact on host health’, *Microbiome*, 7(1), pp. 1–15. Available at: https://doi.org/10.1186/s40168-019-0704-8.

Org, E. *et al.* (2016) ‘Sex differences and hormonal effects on gut microbiota composition in mice’, *Gut Microbes*, 7(4), pp. 313–322. Available at: https://doi.org/10.1080/19490976.2016.1203502.

Padhi, S., Kumar, A. and Behera, A. (2020) ‘Type II diabetes mellitus: a review on recent drug based therapeutics’, *Biomedicine & Pharmacotherapy*, 131, p. 110708. Available at: https://doi.org/10.1016/j.biopha.2020.110708.

Palacios, T. *et al.* (2020) ‘Targeting the Intestinal Microbiota to Prevent Type 2 Diabetes and Enhance the Effect of Metformin on Glycaemia: A Randomised Controlled Pilot Study’, *Nutrients*, 12(7), pp. 1–15. Available at: https://doi.org/10.3390/nu12072041.

Park, J.C. and Im, S.-H. (2020) ‘Of men in mice : the development and application of a humanized gnotobiotic mouse model for microbiome therapeutics’, *Experimental & Molecular Medicine*, 52, pp. 1383–1396. Available at: https://doi.org/10.1038/s12276-020-0473-2.

Pearson, E.R. (2019) ‘Type 2 diabetes : a multifaceted disease’, *Diabetologia*, 62, pp. 1107–1112.

Petersmann, A. *et al.* (2019) ‘Definition, Classification and Diagnosis of Diabetes Mellitus’, *Exp Clin Endocrinol Diabetes*, 127(Suppl 1), pp. S1–S7.

Prattichizzo, F. *et al.* (2018) ‘Pleiotropic effects of metformin: Shaping the microbiome to manage type 2 diabetes and postpone ageing’, *Ageing Research Reviews*, 48, pp. 87–98. Available at: https://doi.org/10.1016/j.arr.2018.10.003.

Puschhof, J., Pleguezuelos-Manzano, C. and Clevers, H. (2021) ‘Organoids and organs-on-chips: Insights into human gut-microbe interactions’, *Cell Host and Microbe*, 29(6), pp. 867–878. Available at: https://doi.org/10.1016/j.chom.2021.04.002.

Qian, X.-L. *et al.* (2021) ‘MiR-454-3p Promotes Oxaliplatin Resistance by Targeting PTEN in Colorectal Cancer’, *Frontiers in Oncology*, 11(May), pp. 1–12. Available at: https://doi.org/10.3389/fonc.2021.638537.

Quast, C. *et al.* (2013) ‘The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools’, *Nucleic Acids Research*, 41(D1), pp. 590–596. Available at: https://doi.org/10.1093/nar/gks1219.

Quigley, E.M.M. and Gajula, P. (2020) ‘Recent advances in modulating the microbiome’, *F1000Res*, 9(F1000 Faculty Rev-46).

Rafat, M. *et al.* (2021) ‘The outstanding role of miR-132-3p in carcinogenesis of solid tumors’, *Human Cell*, 34(4), pp. 1051–1065. Available at: https://doi.org/10.1007/s13577-021-00544-w.

Redondo, M.J. *et al.* (2020) ‘The clinical consequences of heterogeneity within and between different diabetes types’, *Diabetologia*, 63(10), pp. 2040–2048. Available at: https://doi.org/10.1007/s00125-020-05211-7.The.

Ren, H. *et al.* (2019) ‘Metformin alleviates oxidative stress and enhances autophagy in diabetic kidney disease via AMPK/SIRT1-FoxO1 pathway’, *Molecular and Cellular Endocrinology*, p. 110628. Available at: https://doi.org/10.1016/j.mce.2019.110628.

Rena, G. and Lang, C.C. (2018) ‘Repurposing Metformin for Cardiovascular Disease’, *Circulation*, 137(5), pp. 422–424. Available at: https://doi.org/10.1161/CIRCULATIONAHA.117.031735.

Rex, D.K. *et al.* (2017) ‘Colorectal Cancer Screening: Recommendations for Physicians and Patients from the U . S . Multi-Society Task Force on Colorectal Cancer’, *The American Journal of Gastroenterology*, 112(July), pp. 1016–1030. Available at: https://doi.org/10.1038/ajg.2017.174.

Reynoso-García, J. *et al.* (2022) ‘A complete guide to human microbiomes: Body niches, transmission, development, dysbiosis, and restoration’, *Frontiers in Systems Biology*, (July), pp. 1–22. Available at: https://doi.org/10.3389/fsysb.2022.951403.

Ritchie, M.E. *et al.* (2015) ‘Limma powers differential expression analyses for RNA-sequencing and microarray studies’, *Nucleic Acids Research*, 43(7), p. e47. Available at: https://doi.org/10.1093/nar/gkv007.

Rivière, A. *et al.* (2016) ‘Bifidobacteria and butyrate-producing colon bacteria: Importance and strategies for their stimulation in the human gut’, *Frontiers in Microbiology*, 7(JUN). Available at: https://doi.org/10.3389/fmicb.2016.00979.

Robeson, M.S. *et al.* (2021) *RESCRIPt: Reproducible sequence taxonomy reference database management*, *PLoS Computational Biology*. Available at: https://doi.org/10.1371/journal.pcbi.1009581.

Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2009) ‘edgeR: A Bioconductor package for differential expression analysis of digital gene expression data’, *Bioinformatics*, 26(1), pp. 139–140. Available at: https://doi.org/10.1093/bioinformatics/btp616.

Rodríguez, J.M. *et al.* (2015) ‘The composition of the gut microbiota throughout life, with an emphasis on early life’, *Microbial Ecology in Health & Disease*, 26(0), pp. 1–17. Available at: https://doi.org/10.3402/mehd.v26.26050.

Rooj, A.K., Kimura, Y. and Buddington, R.K. (2010) ‘Metabolites produced by probiotic Lactobacilli rapidly increase glucose uptake by Caco-2 cells’, *BMC Microbiology*, 10. Available at: https://doi.org/10.1186/1471-2180-10-16.

La Rosa, S.L. *et al.* (2019) ‘The human gut Firmicute Roseburia intestinalis is a primary degrader of dietary β-mannans’, *Nature Communications*, 10(1), pp. 1–14. Available at: https://doi.org/10.1038/s41467-019-08812-y.

Saeedi, P. *et al.* (2019) ‘Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition’, *Diabetes Research and Clinical Practice*, 157, p. 107843. Available at: https://doi.org/10.1016/j.diabres.2019.107843.

Saito, K. *et al.* (2019) ‘Metagenomic analyses of the gut microbiota associated with colorectal adenoma’, *PLoS ONE*, 14(2), pp. 1–18.

Santos-Marcos, J.A. *et al.* (2019) ‘Sex Differences in the Gut Microbiota as Potential Determinants of Gender Predisposition to Disease’, *Molecular Nutrition and Food Research*, 63(7), pp. 1–11. Available at: https://doi.org/10.1002/mnfr.201800870.

Santos, A.A. *et al.* (2020) ‘Host miRNA-21 promotes liver dysfunction by targeting small intestinal Lactobacillus in mice’, *Gut Microbes*, 12(1), pp. 1–18. Available at: https://doi.org/10.1080/19490976.2020.1840766.

Saraei, P. *et al.* (2019) ‘The beneficial effects of metformin on cancer prevention and therapy: a comprehensive review of recent advances’, *Cancer Management and Research*, 11, pp. 3295–3313.

Sarshar, M. *et al.* (2020) ‘Fecal micrornas as innovative biomarkers of intestinal diseases and effective players in hostmicrobiome interactions’, *Cancers*, 12(8), pp. 1–21. Available at: https://doi.org/10.3390/cancers12082174.

Sedighi, M. *et al.* (2017) ‘Comparison of gut microbiota in adult patients with type 2 diabetes and healthy individuals’, *Microbial Pathogenesis*, 111, pp. 362–369. Available at: https://doi.org/10.1016/j.micpath.2017.08.038.

Shah, P. *et al.* (2016) ‘A microfluidics-based in vitro model of the gastrointestinal human-microbe interface’, *Nature Communications*, 7. Available at: https://doi.org/10.1038/ncomms11535.

Shang, J.W. *et al.* (2019) ‘Expression and significance of urinary microRNA in patients with chronic hepatitis B’, *Medicine (United States)*, 98(37). Available at: https://doi.org/10.1097/MD.0000000000017143.

Shen, Q. *et al.* (2022) ‘Extracellular vesicle miRNAs promote the intestinal microenvironment by interacting with microbes in colitis’, *Gut Microbes*, 14(1), pp. 1–21. Available at: https://doi.org/10.1080/19490976.2022.2128604.

Shin, N.R. *et al.* (2014) ‘An increase in the Akkermansia spp. population induced by metformin treatment improves glucose homeostasis in diet-induced obese mice’, *Gut*, 63(5), pp. 727–735. Available at: https://doi.org/10.1136/gutjnl-2012-303839.

Shussman, N. and Wexner, S.D. (2014) ‘Colorectal polyps and polyposis syndromes’, *Gastroenterology Report*, 2(1), pp. 1–15. Available at: https://doi.org/10.1093/gastro/got041.

Silamiķele, L. *et al.* (2021) ‘Metformin Strongly Affects Gut Microbiome Composition in High-Fat Diet-Induced Type 2 Diabetes Mouse Model of Both Sexes’, *Frontiers in Endocrinology*, 12(March). Available at: https://doi.org/10.3389/fendo.2021.626359.

Singh, R. *et al.* (2021) ‘Gut Microbial Dysbiosis in the Pathogenesis of Gastrointestinal Dysmotility and Metabolic Disorders’, *J Neurogastroenterol Motil*, 27(1), pp. 19–34.

Sninsky, J.A. *et al.* (2022) ‘Risk Factors for Colorectal Polyps and Cancer’, *Gastrointest Endoscopy Clin N Am*, 32, pp. 195–213. Available at: https://doi.org/10.1016/j.giec.2021.12.008.

Srinivasan, K. and Ramarao, P. (2007) ‘Animal models in type 2 diabetes research: An overview’, *Indian J Med Res*, 125(3), pp. 451–472.

Steinway, S.N. *et al.* (2020) ‘Human Microphysiological Models of Intestinal Tissue and Gut Microbiome’, *Frontiers in Bioengineering and Biotechnology*, 8(July). Available at: https://doi.org/10.3389/fbioe.2020.00725.

Stepensky, D. *et al.* (2002) ‘Pharmacokinetic-pharmacodynamic analysis of the glucose-lowering effect of metformin in diabetic rats reveals first-pass pharmacodynamic effect’, *Drug Metabolism and Disposition*, 30(8), pp. 861–868. Available at: https://doi.org/10.1124/dmd.30.8.861.

Stinson, L.F. *et al.* (2019) ‘The Not-so-Sterile Womb: Evidence That the Human Fetus Is Exposed to Bacteria Prior to Birth’, *Frontiers in Microbiology*, 10(June), pp. 1–15. Available at: https://doi.org/10.3389/fmicb.2019.01124.

Strillacci, A. *et al.* (2013) ‘Loss of miR-101 expression promotes Wnt / β -catenin signalling pathway activation and malignancy in colon cancer cells’, *Journal of Pathology*, 1(November 2012), pp. 379–389. Available at: https://doi.org/10.1002/path.4097.

Suez, J. and Elinav, E. (2017) ‘The path towards microbiome-based metabolite treatment’, *Nature Microbiology*, 2(May), pp. 1–5. Available at: https://doi.org/10.1038/nmicrobiol.2017.75.

Sun, L. *et al.* (2018) ‘Gut microbiota and intestinal FXR mediate the clinical benefits of metformin’, *Nature Medicine*, 24(December). Available at: https://doi.org/10.1038/s41591-018-0222-4.

Surwit, R.S. *et al.* (1988) ‘Diet-Induced Type II Diabetes in C57BL/6J Mice’, *Diabetes*, 37(9), pp. 1163–1167.

Szymczak-Pajor, I., Wenclewska, S. and Sliwinska, A. (2022) ‘Metabolic Action of Metformin’, *Pharmaceuticals*, 15(7), p. 810.

Taganov, K.D. *et al.* (2006) 'NF- B-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses', *Proceedings of the National Academy of Sciences*, 103(33), pp. 12481–12486. Available at: https://doi.org/10.1073/pnas.0605298103.

Tanzer, A. and Stadler, P.F. (2004) 'Molecular Evolution of a MicroRNA Cluster', *Journal of Molecular Biology*, 339(2), pp. 327–335. Available at: https://doi.org/10.1016/j.jmb.2004.03.065.

Tarallo, S. *et al.* (2019) 'Altered Fecal Small RNA Profiles in Colorectal Cancer Reflect Gut Microbiome Composition in Stool Samples', *mSystems*, 4(5), pp. 1–16. Available at: https://doi.org/10.1128/msystems.00289-19.

Tarallo, S. *et al.* (2022) 'Stool microRNA profiles reflect different dietary and gut microbiome patterns in healthy individuals', *Gut*, 71, pp. 1302–1314. Available at: https://doi.org/10.1136/gutjnl-2021-325168.

Taylor, S.I., Yazdi, Z.S. and Beitelshees, A.L. (2021) 'Pharmacological treatment of hyperglycemia in type 2 diabetes', *The Journal of Clinical Investigation*, 131(2).

Thorakkattu, P. *et al.* (2022) 'Postbiotics : Current Trends in Food and Pharmaceutical Industry', pp. 1–29.

Thrumurthy, S.G. *et al.* (2016) 'Colorectal adenocarcinoma: risks, prevention and diagnosis', *BMJ*, pp. 1–12. Available at: https://doi.org/10.1136/bmj.i3590.

Thursby, E. and Juge, N. (2017) 'Introduction to the human gut microbiota', *Biochemical Journal*, 474(11), pp. 1823–1836. Available at: https://doi.org/10.1042/BCJ20160510.

Tomkovich, S. *et al.* (2020) 'Human Colon Mucosal Biofilms and Murine Host Communicate via Altered mRNA and microRNA Expression during Cancer', *mSystems*, 5(1), pp. 1–20. Available at: https://doi.org/10.1128/msystems.00451-19.

Toyama, K. *et al.* (2012) 'Loss of multidrug and toxin extrusion 1 (MATE1) is associated with metformin-induced lactic acidosis', *British Journal of Pharmacology*, 166(3), pp. 1183–1191. Available at: https://doi.org/10.1111/j.1476-5381.2012.01853.x.

Tropini, C. *et al.* (2017) 'The Gut Microbiome: Connecting Spatial Organization to Function', *Cell Host and Microbe*, 21(4), pp. 433–442. Available at: https://doi.org/10.1016/j.chom.2017.03.010.

Udesen, P.B. *et al.* (2020) 'Metformin decreases mir-122, mir-223 and mir-29a in women with polycystic ovary syndrome', *Endocrine Connections*, 9(11), pp. 1075–1084. Available at: https://doi.org/10.1530/EC-20-0195.

Ustinova, M. *et al.* (2020) 'Whole-blood transcriptome profiling reveals signatures of metformin and its therapeutic response', *PLoS ONE*, 15(8 August). Available at: https://doi.org/10.1371/journal.pone.0237400.

Viennois, E. *et al.* (2019) 'Host-derived fecal microRNAs can indicate gut microbiota healthiness and ability to induce inflammation', *Theranostics*, 9(15), pp. 4542–4557. Available at: https://doi.org/10.7150/thno.35282.

Viigimaa, M. *et al.* (2020) 'Macrovascular Complications of Type 2 Diabetes Mellitus', *Current Vascular Pharmacology*, 18(2). Available at: https://doi.org/10.2174/1570161117666190405165151.

Wang, B. *et al.* (2020) 'A high-fat diet increases gut microbiota biodiversity and energy expenditure due to nutrient difference', *Nutrients*, 12(10), pp. 1–20. Available at: https://doi.org/10.3390/nu12103197.

Wang, C.-P. *et al.* (2021) ‘MicroRNA- 454-5p promotes breast cancer progression by inducing epithelial-mesenchymal transition via targeting the FoxJ2/E - cadherin axis’, *Oncology Reports*, 46(1), pp. 1–11. Available at: https://doi.org/10.3892/or.2021.8078.

Wang, J. *et al.* (2019) ‘Core gut bacteria analysis of healthy mice’, *Frontiers in Microbiology*, 10(APR), pp. 1–14. Available at: https://doi.org/10.3389/fmicb.2019.00887.

Wang, K. *et al.* (2012) ‘Comparing the MicroRNA spectrum between serum and plasma’, *PLoS ONE*, 7(7). Available at: https://doi.org/10.1371/journal.pone.0041561.

Wang, Y. and Li, H. (2022) ‘Gut microbiota modulation : a tool for the management of colorectal cancer’, *Journal of Translational Medicine*, 20, pp. 1–14. Available at: https://doi.org/10.1186/s12967-022-03378-8.

Weber, J.A. *et al.* (2016) ‘The MicroRNA Spectrum in 12 Body Fluids’, *Clin Chem*, 56(11), pp. 1733–1741. Available at: https://doi.org/10.1373/clinchem.2010.147405.The.

Whaley, J.M. *et al.* (2012) ‘Targeting the kidney and glucose excretion with dapagliflozin : preclinical and clinical evidence for SGLT2 inhibition as a new option for treatment of type 2 diabetes mellitus’, *Diabetes Metab Syndr Obes.*, 5, pp. 135–148.

Wickham, H. (2016) ‘ggplot2: Elegant Graphics for Data Analysis’, in. Springer-Verlag New York.

Wilkins, T.D., Fulghum, R.S. and Wilkins, J.H. (1974) ‘Eubacterium plexicaudatum sp. nov., an anaerobic bacterium with a subpolar tuft of flagella, isolated from a mouse cecum’, *International Journal of Systematic Bacteriology*, 24(4), pp. 408–411. Available at: https://doi.org/10.1099/00207713-24-4-408.

Williams, M.R. *et al.* (2017a) ‘MicroRNAs-based inter-domain communication between the host and members of the gut microbiome’, *Frontiers in Microbiology*. Frontiers Media S.A. Available at: https://doi.org/10.3389/fmicb.2017.01896.

Williams, M.R. *et al.* (2017b) ‘MicroRNAs-based inter-domain communication between the host and members of the gut microbiome’, *Frontiers in Microbiology*, 8(SEP), pp. 1–10. Available at: https://doi.org/10.3389/fmicb.2017.01896.

Winter, J. *et al.* (2009) ‘Many roads to maturity: microRNA biogenesis pathways and their regulation’, *Nature Cell Biology*, 11(3), pp. 228–234.

Wohnhaas, C.T. *et al.* (2020) ‘Fecal MicroRNAs Show Promise as Noninvasive Crohn’s Disease Biomarkers’, *Crohn's & Colitis 360*, 2(1), pp. 1–12. Available at: https://doi.org/10.1093/crocol/otaa003.

Wood, D.E., Lu, J. and Langmead, B. (2019) ‘Improved metagenomic analysis with Kraken 2’, *Genome Biology*, 20(1), pp. 1–13. Available at: https://doi.org/10.1186/s13059-019-1891-0.

World Health Organization (2020) ‘Diagnosis and Management of Type 2 Diabetes’.

Wu, H. *et al.* (2017) ‘Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug’, *Nature Medicine*, 23(7), pp. 850–858. Available at: https://doi.org/10.1038/nm.4345.

Wu, T., Horowitz, M. and Rayner, C.K. (2017) ‘New insights into the anti-diabetic actions of metformin: from the liver to the gut’, *Expert Review of Gastroenterology and Hepatology*, 11(2), pp. 157–166. Available at: https://doi.org/10.1080/17474124.2017.1273769.

Xi, Y. and Xu, P. (2021) ‘Global colorectal cancer burden in 2020 and projections to 2040’, *Translational Oncology*, 14(10). Available at: https://doi.org/10.1016/j.tranon.2021.101174.

Xu, Y. *et al.* (2021) ‘An overview of in vitro, ex vivo and in vivo models for studying the transport of drugs across intestinal barriers’, *Advanced Drug Delivery Reviews*, 175. Available at: https://doi.org/10.1016/j.addr.2021.05.005.

Xue, J.-H. *et al.* (2021) ‘Fecal Fusobacterium nucleatum as a predictor for metachronous colorectal adenoma after endoscopic polypectomy’, *Journal of Gastroenterology and Hepatology*, 36(10), pp. 1–9. Available at: https://doi.org/10.1111/jgh.15559.

Yamamoto, Y. *et al.* (2018) ‘A Metabolomic-Based Evaluation of the Role of Commensal Microbiota throughout the Gastrointestinal Tract in Mice’, *Microorganisms*, 6(4), p. 101. Available at: https://doi.org/10.3390/microorganisms6040101.

Yang, J. *et al.* (2021) ‘Oscillospira - a candidate for the next-generation probiotics’, *Gut Microbes*, 13(1), p. e1987783. Available at: https://doi.org/10.1080/19490976.2021.1987783.

Yerevanian, A. and Soukas, A.A. (2019) ‘Metformin: Mechanisms in Human Obesity and Weight Loss’, *Current Obesity Reports*, 8(2), pp. 156–164. Available at: https://doi.org/10.1007/s13679-019-00335-3.

Yi, R. *et al.* (2016) ‘MicroRNAs as diagnostic and prognostic biomarkers in colorectal cancer’, *World Journal of Gastrointestinal Oncology*, 8(4), pp. 330–340. Available at: https://doi.org/10.4251/wjgo.v8.i4.330.

Yin, Y. *et al.* (2016) ‘Systematic analysis of key miRNAs and related signaling pathways in colorectal tumorigenesis’, *Gene*, 578(2), pp. 177–184. Available at: https://doi.org/10.1016/j.gene.2015.12.015.

Yong, C. *et al.* (2016) ‘Decreased expression of miR-132 in CRC tissues and its inhibitory function on tumor progression’, *Open Life Sciences*, 11(1), pp. 130–135. Available at: https://doi.org/10.1515/biol-2016-0018.

Yoshizawa, J.M. and Wong, D.T.W. (2013) ‘Salivary micrornas and oral cancer detection’, *Methods in Molecular Biology*, 936(14), pp. 313–324. Available at: https://doi.org/10.1007/978-1-62703-83-0_24.

Yuan, B. *et al.* (2021) ‘Fecal Bacteria as Non-Invasive Biomarkers for Colorectal Adenocarcinoma’, *Frontiers in Oncology*, 11(August), pp. 1–9. Available at: https://doi.org/10.3389/fonc.2021.664321.

Yuan, C., Burns, M.B., *et al.* (2018) ‘Interaction between Host MicroRNAs and the Gut Microbiota in Colorectal Cancer’, *mSystems*, 3(3). Available at: https://doi.org/10.1128/msystems.00205-17.

Yuan, C., Burns, M., *et al.* (2018) ‘Interaction Between Host MicroRNAs and the Gut Microbiota in Colorectal Cancer’, *mSystems*, 3(3), pp. 1–13. Available at: https://doi.org/10.1101/192401.

Zhang, C., Franklin, C.L. and Ericsson, A.C. (2021) ‘Consideration of gut microbiome in murine models of diseases’, *Microorganisms*, 9(5). Available at: https://doi.org/10.3390/microorganisms9051062.

Zhang, Q. (2020) ‘Effects of Metformin on the Gut Microbiota in Obesity and Type 2 Diabetes Mellitus’.

Zhang, Xu *et al.* (2015) ‘Modulation of gut microbiota by berberine and metformin during the treatment of high-fat diet-induced obesity in rats’, *Scientific Reports*, 5(August), pp. 1–10. Available at: https://doi.org/10.1038/srep14405.

Zhao, H. *et al.* (2021) ‘Comparative efficacy of oral insulin sensitizers metformin , thiazolidinediones , inositol , and berberine in improving endocrine and metabolic profiles in women with PCOS: a network meta - analysis’, *Reproductive Health*, 18(1), pp. 1–12. Available at: https://doi.org/10.1186/s12978-021-01207-7.

Zhou, Q. *et al.* (2011) ‘Immune-related microRNAs are abundant in breast milk exosomes’, *International Journal of Biological Sciences*, 8(1), pp. 118–123. Available at: https://doi.org/10.7150/ijbs.8.118.

Zuo, J. *et al.* (2019) ‘miR-454-3p exerts tumor-suppressive functions by down-regulation of NFATc2 in glioblastoma’, *Gene*, 710(December 2018), pp. 233–239. Available at: https://doi.org/10.1016/j.gene.2019.06.008.