# PROJECT DELIVERABLE REPORT

## Project Title:

Advanced personalised, multi-scale computer models preventing osteoarthritis

SC1-PM-17-2017 - Personalised computer models and in-silico systems for well-being

| | |
|---|---|
| **Deliverable number** | **4.2** |
| **Deliverable title** | Qualification of OA-related exosomal and microbiome biomarkers |
| **Submission month of deliverable** | 28 |
| **Issuing partner** | LEITAT |
| **Contributing partners** | LEITAT, HULAFE, UNIC, ANIMUS |
| **Dissemination Level (PU/PP/RE/CO):** | CO |
| | |
| **Project coordinator** | University of Nicosia (UNIC) |
| **Tel:** | **+357 22 841 528** |
| **Fax:** | **+357 22 357481** |
| **Email:** | **felekkis.k@unic.ac.cy & giannaki.c@unic.ac.cy** |
| **Project web site address** | www.oactive.eu |

**Revision History**

| Version | Date | Responsible | Description/Remarks/Reason for changes |
|---|---|---|---|
| 1.0 | 01/04/2021 | LEITAT | Initial draft of D4.2 |
| 2.0 | 03/04/2021 | LEITAT | Second draft of D4.2 |
| 3.0 | 07/04/2021 | LEITAT | Final version |
| | | | |

## 1. Table of Contents

## 1 List of Figures



## 2 List of Tables

## 3 Summary

Osteoarthritis (OA) is one of most common types of musculoskeletal diseases worldwide and is categorized by the European Commission as one of the twenty-four priority diseases. The OActive project aims to improve the life's quality of people who suffer or would probably suffer from OA through personalized prevention and intervention. Over the last few years, new evidence has emerged suggesting that changes in the gut microbiome may also be linked to developmental factors in the onset of musculoskeletal disorders such as OA, but the possible role of the mechanisms involved remains unclear.

Computer modelling is going to be used for integrating various datasets such as lifestyle, organ and tissue level mechanistic, environmental and biochemical biomarkers among other. The miRNA exosome content and microbiome data generated during this task will be included in the computer modelling algorithm in order to help to improve the personalized prediction of OA onset.

Contributions from partners

For the fulfilment of this task, technical inputs have been received from the following partners:

- LEITAT – will use the knowledge attained during the project in order to strengthen their research activities in preclinical assays, such as exosome studies and biomarkers investigation on miRNA and microbiome. Moreover, the project will enable to enlarge their expertise in microbiome techniques, expanding thus their portfolio in Technical Services. This will enable them to provide greater technological value long-term to their customers, maintain their principles and values and expand their business network.
- UNIC – Plasma samples from OA patients
- HULAFE - Plasma and stool samples from healthy volunteers and early OA patients.
- ANIMUS – Plasma samples from high-risk OA athletes.

This report refers to Deliverable 4.2, which relates to the OACTIVE WP 4, "Biochemical modelling and inflammation biomarkers" led by LEITAT and specifically Task 4.3 "Qualification of OA-related exosomal and microbiome biomarkers", also led by LEITAT.

## 4 Introduction

Osteoarthritis (OA) is one of the world's most common types of musculoskeletal diseases and its prevalence is increased, especially since the middle of the 20$^{th}$ century (Biver et al., 2019). In Europe is the most common form of chronic pain (34%), representing a great economic and social cost to society. At present there are no recognized drugs to modify the disease state, the only alternative being palliation of symptoms. Different risk factors may play an important role in OA, including obesity, joint trauma, age, hormonal disorders, and genetic background (Abramoff and Caldera, 2020). There is increasing evidence for an inflammatory component to OA linked with joint tissue damage (Robinson *et al.*, 2016)Scanzello, 2017).

Knee osteoarthritis (OA), the most common arthritis, is a leading cause of disability in the elderly population mainly due to pain, which is the primary symptom of the disease (Neogi, 2013). It has been recently estimated that more than 250 million people suffer from knee OA (Ro, D.H., et al.,2019). Knee OA is characterized by structural modifications to primarily articular cartilage and the subchondral bone, but also Hoffa's fat pad, synovia, ligaments, and muscles, suggesting that knee OA should be observed as a whole joint disease (Loeser et al. 2012). Moreover, OA is a complex disease involving all the tissues of the joint as well as inflammation and thus it is a collection of heterogeneous pathologies that result in a

common outcome (pain and joint destruction), rather than being one homogeneous disease (Laslett., et al., 2016). Furthermore, the Osteoarthritis Research Society International (OARSI) has recently endorsed a new definition of OA: "Osteoarthritis is a disorder involving movable joints characterized by cell stress and extracellular matrix degradation initiated by micro-and macro-injury that activates maladaptive repair responses including pro-inflammatory pathways of innate immunity. The disease manifests first as a molecular derangement (abnormal joint tissue metabolism) followed by anatomic, and/or physiologic derangements (characterized by cartilage degradation, bone remodeling, osteophyte formation, joint inflammation, and loss of normal joint function), that can culminate in illness" (Kraus, V.B., et al., 2015). From this definition, and as aforementioned, can be easily deduced that OA is a heterogeneous disease with a variety of pathophysiologic drivers leading to multiple phenotypes.

Currently, there are no therapies and/or medication approved by regulatory authorities that alter the onset or progression of OA structural damage while the available treatments have only moderate effects, and therefore, patients are not satisfied with their efficacy (Zhang, et al., 2010). Due to the failure of the current medications to treat knee OA, the number of joint replacement surgeries is increasing by ~10% annually (Hunter, D.J., et al., 2014). The anatomical severity of OA is usually assessed by joint imaging using standard radiographs, (x-Ray, and magnetic resonance imaging), however, by the time OA is detected on the radiograph, significant cartilage degradation has already occurred (Ishijima, M., et al., 2011). Therefore, there is an urgent need for the identification of alternative methods for early detection and monitoring OA while the identification of OA before it becomes evident on radiographs remains a challenge (Cibere et al., 2009).

In the past few years, the prognostic and diagnostic utility of biomarkers in OA has been investigated. Specific molecules of the cartilage and bone turnover that are released in the blood or other biological samples could be used as biomarkers for identifying patients at greater risk for developing OA earlier than the time required to reveal radiographic changes and also to monitor the progression of the disease (Cahue, S., et al., 2007). Biomarkers are candidates that are now being used to detect and monitor cartilage, bone turnover, and synovial metabolism for the critical assessment of the pathophysiological processes that lead to joint failure and pain in OA patients (Ishijima, et al., 2011). Several biomarkers have been tested in samples from patients with varying severity of OA during the last year. Interestingly, specific biomarkers have been associated with OA presence and/or severity in cross-sectional studies; some markers have also been found to predict OA progression in longitudinal studies (Cahue, S., et al., 2007). This work aimed to validate characteristic a new source of biomarkers, as exosomes content and gut microbiome in samples from OA patients. The main objectives of this work were to:

- Perform an exploratory analysis to determine the potential use of microRNA content of exosomes data as an OA prognosis biomarker.
- Perform an exploratory analysis to determine the potential use of microbiome data as OA prognostic biomarkers for OA.
- Generate the data needed in the correct format to be integrated into the computer models designed for personalized prevention of OA.

## 5 Exosome content - miRNA studies

The evaluation of new markers for clinical practice requires standardized pre-analytical, analytical, and post-analytical protocols aimed at providing accurate, reproducible, and consistent, preferably non-invasive, measures by controlling for all variables that may introduce biases in the detection of biomarkers (type of sample, manipulation, method of measurement). In the case of the analysis of circulating miRNAs as biomarkers in clinical practice, and research they show several advantageous features as they can be evaluated in human biofluids (e.g., urine, plasma, and serum). However, it is essential to standardize sampling and manipulation protocols to reduce the bias that may affect the pre-analytical phase of miRNA

validation. The selection of reference genes along with the standardization strategy applied for miRNA quantification represents the most difficult problem due to the absence of a standardized methodology. RT-qPCR data for miRNA expression can be normalized using single or multiple endogenous and/or exogenous reference genes, as well as by the mean expression value of all measured miRNAs. Taking all these aspects into account, it is necessary to carry out a careful validation of the results and a comparative work with the public databases in the studies of this field, in our case, of OA. For this reason, a data collection procedure has been developed that includes quality controls from each step, as expressed in the workflow diagram in Figure 1, to obtain robust results that can be applied to the early diagnosis and prognosis of OA in clinical practice and can serve also to contributing to the computer modelling of the OActive project.

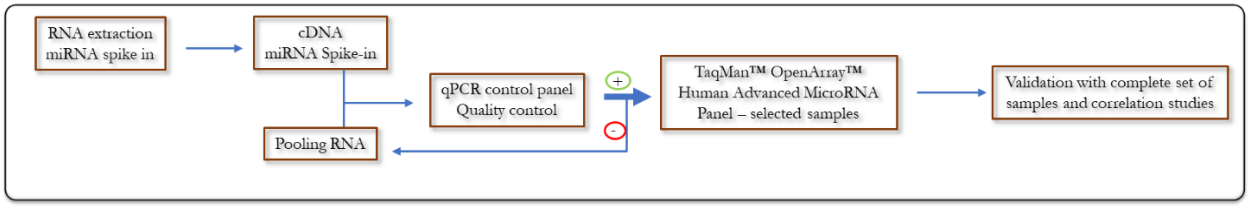



***Figure 1:*** *Diagrammatic representation of the workflow. The thick arrow shows where is now the research.*

### 5.1 Relation of plasma and urine samples received

HULAFE: 84 plasma and urine samples were received during 2019 and 2020.

UNIC: 51 plasma samples. July 2019.

ANIMUS: 25 plasma samples. November 2020.

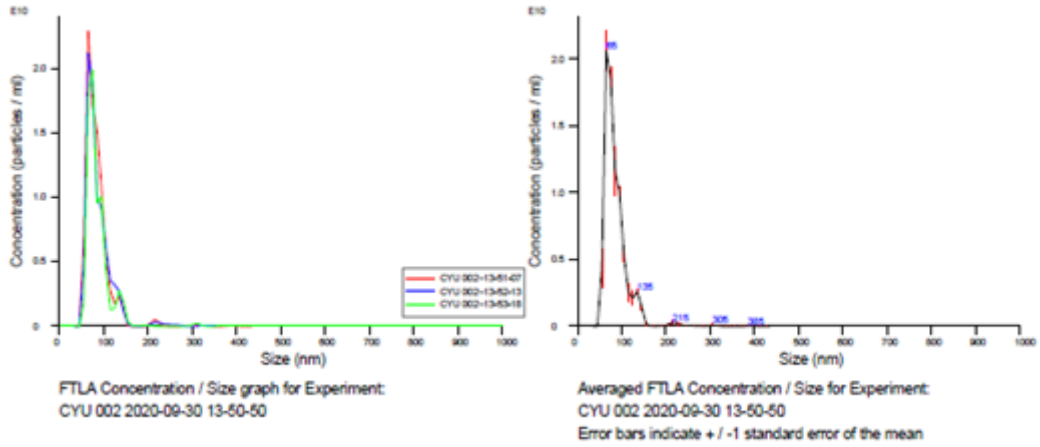
Most of the plasma samples were 5 ml or less, and urine samples from 4 to 10 ml. HULAFE samples were split to leave 2ml of plasma and urine for ELISA assays. This small volume forces us to limit the study to miRNA analysis.

### 5.2 Exosome extraction

Exosomes from plasma samples were extracted with the Plasma/Serum Exosome Purification kit from Norgen Biotech. Plasma volume for all samples was from 2 to 4 ml. Quality control -size particle and concentration- was done with 12 samples in a Nanosight N300 nanoparticle analyser. As it is shown in figure 2, exosome suspension was homogenous and had the expected reported size, around 100 nm. Then, it was assumed that all samples extracted will have similar quality.

Exosomes from urine samples were extracted using an adapted protocol in Leitat that was the one that gave higher yields. Briefly, 5 ml of urine sample was centrifuged at 2000rcf, 30min at 4ºC. The supernatant (sn) was recovered and centrifuged at 16000 rcf for 1h at 4ºC. Then, the sn was filtered through a 0.22 µm filter and the Total Exosome Isolation Reagent (for urine) (Invitrogen) solution was added at 1:1 and incubated for 1h at RT. The mixture was centrifuged at 10000 rcf for 1h at 4ºC, and the pellet was resuspended in 500 µl of PBS. Quality control was done in the same way as plasma samples, but with fewer samples because it was done before with volunteers' samples to set up the method. As it can be shown in Figure 2, there was more heterogenous composition and less concentrated samples.

A)




B)

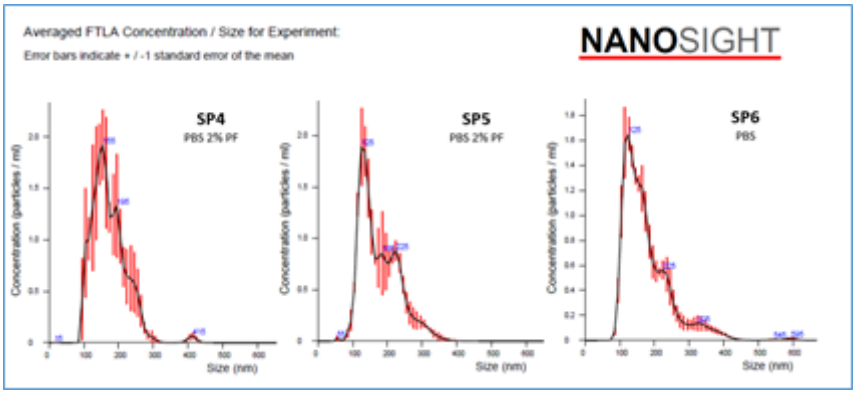



***Figure 2.*** *Examples of quantification and size analysis of microvesicles extracted from A) plasma and B) urine samples. The analysis was done with Nanosight N300*

### 5.3 miRNA purification and cDNA synthesis

RNA from all exosomes was extracted using the RNA extraction kit from Norgen. For quality control of the miRNA yield, spike miRNA (RNA Spike-In Kit Qiagen) was added to the lysis buffer. These spike miRNAs were used as a control of RNA purification in the miRCURY LNA SYBR® Green PCR amplification system using the QC PCR Panel, miRCURY LNA miRNA QC PCR Panel from Qiagen. It was not possible to quantify RNA concentration due to the low yield obtained in exome extraction. That is why it was necessary to include spike miRNA for quality control of the RNA extractions.

Two methods of cDNA synthesis will be used. The miRCURY LNA RT Kit method for quality control adding spike miRNA, and the Taqman Adv miRNA cDNA synthesis (Applied Biosystems) to use in the TaqMan™ OpenArray™ Human Advanced MicroRNA Panel, QuantStudio™ 12K Flex (Applied Biosystems).

### 5.4 Quality control analysis

#### 5.4.1 cDNA synthesis

cDNA was synthetized with the miRCURY LNA RT Kit. As it was pointed out, it was not possible to measure RNA concentration, so for the quality control assay, it was necessary to synthetize cDNA using increasing volumes of the RNA samples. Table 1 shows the reaction settings.

***Table 1*** *cDNA synthesis using different amounts of RNA to determine the right amount for the qPCR analysis.*

| | RNA reaction (µl) |
|---|---|
| Sample RNA | 6 |
| 5X miRCURY SYBR Green RT reaction | 2 |
| RNAse free H2O | 14 |
| 10x miRCURY RT enzyme mix | 4 |
| miR-39-3p | 2 |
| Template RNA | 12 |
| TOTAL reaction volume | 40 |

### 5.4.2 qPCR Quality control

The miRCURY LNA miRNA QC PCR panel was used. Table 2 shows the reaction settings and Table 3 the composition of the qPCR panel. The qPCR panel includes all the Spike miRNA added in the purification and cDNA synthesis.

***Table 2 –*** *Reaction mix for qPCR control panel*

| COMPONENT | QC PCR PANEL, µl |
|---|---|
| 2x miRCURY SYBR Green Master Mix | 70 |
| cDNA template | 1.4 |
| RNAse free water | 68.6 |
| TOTAL | 140 |

The qPCR settings for a LightCycler® 480 Instrument II (Roche Life Science) were:

Activation 95°C 2 min

Amplification 45 cycles 95 °C, 10 sec

56 °C, 60 sec

***Table 3*** *miRNA sequences included in the miRCURY LNA miRNA QC PCR 384 panel. Up to 32 samples can be run. These panels include the spike-in sequences included in the RNA extraction and cDNA synthesis and known miRNA sequences expressed in different tissues.*

| miRname (human) | microRNA target sequence | Corresponding LNA™ microRNA PCR primer set (Prod No) | Sample | Assay type |
|---|---|---|---|---|
| hsa-miR-103a-3p | AGCAGCAUUGUACAGGGCUAUGA | YP00204063 | S1 | GOI |
| hsa-miR-191-5p | CAACGGAAUCCCAAAAGCAGCUG | YP00204306 | S1 | GOI |
| hsa-miR-451a | AAACCGUUACCAUUACUGAGUU | YP02119305 | S1 | GOI |
| hsa-miR-23a-3p | AUCACAUUGCCAGGGAUUUCC | YP00204772 | S1 | GOI |
| UniSp6 | | YP00203954 | S1 | Spike |
| UniSp2 | | YP00203950 | S1 | Spike |
| UniSp4 | | YP00203953 | S1 | Spike |
| UniSp5 | | YP00203955 | S1 | Spike |
| cel-miR-39-3p | | YP00203952 | S1 | Spike |
| UniSp3 IPC | | YP02119288 | S1 | IPC |

The quality analysis of samples was performed only in 96 samples (green box before the number) out of the 102 predicted by the analysis with the TaqMan™ OpenArray™ Human Advanced MicroRNA Panel. Table 4. The criteria for HULAFE and ANIMUS samples was to select by sex and body mass index (BMI) because all patients were early OA for HULAFE and risk volunteers for ANIMUS. The criteria for UNIC were sex and Kellgren-Lawrence (K&L) scores for the left (L) and right knee (R). Controls from HULAFE were also included.

***Table 4 –*** *Relation of samples included in the miRNA analysis. Green boxes indicate the samples that were selected for the TaqMan™ OpenArray™ Human Advanced MicroRNA Panel*

| **HULAFE** | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Selected | Patient | Sex | Status | BMI | COMP | HA | PIICP | Age | IL1β | TNFα | IL6 | K&L L | K&L R |
| | SPH089 | Female | Early OA | 18.62 | 161.9 | 155.9 | 438.648 | 53 | <1.95 | 21.7 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH069 | Female | Early OA | 19.492 | 172.2 | 83.3 | 257.484 | 46 | <1.95 | 30.8 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH082 | Female | Early OA | 20.69 | 266.5 | 66.3 | 666.066 | 54 | <1.95 | 22 | <10,24 | - | - |
| | SPH067 | Female | Early OA | 21.855 | 208.8 | 31.9 | 525.76 | 47 | <1.95 | 34.5 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH019 | Female | Early OA | 23.01 | 198.177 | 65.382 | 589.123 | 58 | 2.81892 | 14.588 | 1.164607 | 0= 0 points | 0= 0 points |
| | SPH072 | Female | Early OA | 23.524 | 120.1 | 49.7 | 326.095 | 49 | <1.95 | 52.7 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH017 | Female | Early OA | 23.533 | 135.88 | 65.212 | 834.375 | 44 | 2.423282 | 5.736855 | 4.054927 | 0= 0 points | 0= 0 points |
| | SPH047 | Female | Early OA | 23.533 | 382.7 | 150.8 | 625.208 | 55 | <1.95 | 28.7 | <10,24 | 0= 0 points | 0= 0 points |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SPH071 | Female | Early OA | 23.627 | 214 | 67.5 | 466.4 | 43 | <1.95 | 33.9 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH022 | Female | Early OA | 23.147 | 232.034 | 94.033 | 526.451 | 60 | 3.066193 | 21.80005 | 0.465843 | 2= 3-4 points | 2= 3-4 points |
| | SPH044 | Female | Early OA | 24.281 | 292.8 | 73.2 | 670.691 | 57 | <1.95 | 26.6 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH058 | Female | Early OA | 24.747 | 252.6 | 77 | 683.026 | 53 | <1.95 | 27.1 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH091 | Female | Early OA | 24.805 | 215.2 | 70 | 253.629 | 46 | 0.2 | 29.6 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH053 | Female | Early OA | 24.883 | 456.7 | 226.8 | 998.328 | 57 | <1.95 | 31.8 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH001 | Female | Early OA | 25.097 | 363.15 | 117.586 | 680.409 | 65 | <1.95 | <23.4 | 0.815225 | 0= 0 points | 0= 0 points |
| | SPH078 | Female | Early OA | 25.796 | 178.9 | 80.2 | 653.731 | 51 | <1.95 | 24.7 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH_95 | Female | Early OA | 26.504 | 64 | 0.6 | 89 | 1166.2 | 2 | 35.1 | <10,24 | 1= 1-2 points | 1= 1-2 points |
| | SPH_94 | Female | Early OA | 26.915 | - | 0.4 | 34.1 | 753.6 | 2 | 34.2 | <10,24 | 1= 1-2 points | 0= 0 points |
| | SPH054 | Female | Early OA | 27.716 | 194.1 | 201.9 | 607.477 | 56 | <1.95 | 27.7 | <10,24 | - | - |
| | SPH033 | Female | Early OA | 27.84 | 276.9 | 58.9 | 434.793 | 56 | <1.95 | 30.8 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH062 | Female | Early OA | 28.172 | 180.6 | 132.5 | 313.76 | 54 | <1.95 | 28.7 | <10,24 | - | - |
| | SPH016 | Female | Early OA | 28.65 | 270.857 | 38.632 | 833.03 | 60 | 1.730916 | 13.44063 | 0.193092 | 0= 0 points | 0= 0 points |
| | SPH096 | Female | Early OA | 29.261 | - | - | - | 53 | - | - | - | 0= 0 points | 0= 0 points |
| | SPH064 | Female | Early OA | 29.456 | 257.8 | 86 | 728.51 | 51 | <1.95 | 29.8 | <10,24 | - | - |
| | SPH038 | Female | Early OA | 29.469 | 270.9 | 70.7 | 346.138 | 57 | <1.95 | 29.2 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH040 | Female | Early OA | 31.398 | 355.6 | 81.8 | 856.481 | 52 | <1.95 | 37.6 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH002 | Female | Early OA | 31.445 | 253.603 | 93.129 | 1168.528 | 52 | <1.95 | <23.4 | 1.902191 | 0= 0 points | 0= 0 points |
| | SPH004 | Female | Early OA | 32.447 | 639.223 | 43.603 | 1345.354 | 55 | <1.95 | 5.822925 | 3.998484 | 0= 0 points | 0= 0 points |
| | SPH023 | Female | Early OA | 34.844 | 187.343 | 35.746 | 578.801 | 42 | 3.461831 | 34.74895 | 1.242247 | 0= 0 points | 0= 0 points |
| | SPH031 | Female | Early OA | 35.339 | 90.3 | 94.5 | 430.6 | 46 | 2.9 | 12.5 | 0.2 | 0= 0 points | 0= 0 points |
| | SPH075 | Female | Early OA | 35.725 | 202.8 | 147.9 | 421.688 | 46 | <1.95 | 22.4 | <10,24 | 1= 1-2 points | 1= 1-2 points |
| | SPH048 | Female | Early OA | 38.83 | 183.8 | 46.2 | 450.982 | 60 | <1.95 | 28.7 | <10,24 | 1= 1-2 points | 0= 0 points |
| | SPH030 | Female | Healthy | 25.089 | 179.2 | 75.9 | 481.5 | 42 | 3.4 | 17.2 | 0.6 | 0= 0 points | 0= 0 points |
| | SPH028 | Female | Healthy | 25.113 | 92.092 | 48.578 | 321.474 | 47 | 3.214558 | 13.27672 | 0.776405 | - | - |
| | SPH087 | Female | Healthy | 25.496 | 247.4 | 73.1 | 732.364 | 47 | <1.95 | 21.7 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH041 | Female | Healthy | 25.97 | 305.9 | 163.9 | 633.688 | 53 | <1.95 | 31.3 | <10,24 | - | - |
| | SPH027 | Female | Healthy | 26.026 | 184.183 | 39.412 | 367.926 | 42 | 2.225463 | 10.81807 | 1.552809 | 0= 0 points | 0= 0 points |
| | SPH080 | Female | Healthy | 26.272 | 335 | 43.2 | 776.306 | 59 | <1.95 | 20.1 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH083 | Female | Healthy | 26.396 | 270.2 | 115.9 | 471.026 | 53 | <1.95 | 20.5 | <10,24 | 0= 0 points | 0= 0 points |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SPH079 | Female | Healthy | 26.573 | 204.5 | 92.8 | 640.626 | 58 | <1.95 | 19.4 | 6.9 | 0= 0 points | 0= 0 points |
| | SPH018 | Female | Healthy | 26.814 | 218.04 | 72.409 | 633.363 | 58 | 2.522191 | 0.901506 | 0.85404 5 | 0= 0 points | 0= 0 points |
| | SPH008 | Female | Healthy | 27.074 | 339.926 | 49.682 | 1002.46 | 57 | <1.95 | 6.961634 | <10.24 | 0= 0 points | 0= 0 points |
| | SPH045 | Female | Healthy | 28.043 | 154.7 | 237.8 | 272.902 | 48 | <1.95 | 31.8 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH065 | Female | Healthy | 28.125 | 244.6 | 197.8 | 909.673 | 60 | <1.95 | 30.3 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH026 | Female | Healthy | 28.214 | 190.503 | 101.23 | 968.109 | 54 | 3.65965 | 41.46927 | 0.85404 5 | 0= 0 points | 0= 0 points |
| | SPH025 | Female | Healthy | 28.353 | 187.343 | 54.112 | 754.284 | 56 | 3.461831 | 16.39101 | 0.15528 1 | 0= 0 points | 0= 0 points |
| | SPH035 | Female | Healthy | 30.022 | 255 | 88.2 | 770.139 | 53 | <1.95 | 32.4 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH097 | Female | Healthy | 30.11 | - | - | - | 43 | - | - | - | 0= 0 points | 0= 0 points |
| | SPH052 | Female | Healthy | 30.408 | 276.1 | 199.7 | 818.706 | 62 | <1.95 | 34.5 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH093 | Female | Healthy | 30.664 | - | - | - | 62 | - | - | - | 0= 0 points | 0= 0 points |
| | SPH057 | Female | Healthy | 31.033 | 179.4 | 67.2 | 578.182 | 60 | <1.95 | 30.3 | 10.5 | | |
| | SPH060 | Female | Healthy | 31.056 | 177.8 | 62.7 | 713.091 | 49 | <1.95 | 30.8 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH036 | Female | Healthy | 31.229 | 175.8 | 51.6 | 1299.754 | 52 | <1.95 | 27.1 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH086 | Female | Healthy | 32.456 | 191.7 | 29 | 659.899 | 41 | <1.95 | 25.8 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH092 | Female | Healthy | 34.865 | - | - | - | 47 | - | - | - | - | - |
| | | | | | | | | | | | | | |
| | SPH039 | Male | Early OA | 22.14 | 249.4 | 133.7 | 467.171 | 58 | <1.95 | 26.6 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH042 | Male | Early OA | 24.49 | 247 | 216.2 | 896.568 | 55 | <1.95 | 25.1 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH007 | Male | Early OA | 25.136 | 140.846 | 16.464 | 607.796 | 46 | <1.95 | 27.61448 | <10.24 | 1= 1-2 points | 0= 0 points |
| | SPH063 | Male | Early OA | 25.295 | 234.7 | 57 | 834.895 | 45 | <1.95 | 27.1 | <10,24 | - | - |
| | SPH009 | Male | Early OA | 25.69 | 376.492 | 68.729 | 821.601 | 54 | <1.95 | 8.121906 | <10.24 | 0= 0 points | 0= 0 points |
| | SPH021 | Male | | 25.617 | 401.32 | 124.925 | 523.501 | 51 | 4.154198 | 15.24364 | 3.260899 | 0= 0 points | 3= 5-9 points |
| | SPH051 | Male | Early OA | 26.908 | 193.7 | 85.6 | 532.699 | 57 | <1.95 | 29.8 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH077 | Male | Early OA | 26.919 | 207.4 | 77.9 | 816.393 | 48 | <1.95 | 23.2 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH032 | Male | Early OA | 26.99 | 92.1 | 100 | 492.5 | 48 | 2.9 | 15.4 | 0.5 | 0= 0 points | 0= 0 points |
| | SPH006 | Male | Early OA | 27.745 | 274.469 | 16.98 | 805.464 | 46 | <1.95 | 11.37067 | <10.24 | 0= 0 points | 0= 0 points |
| | SPH074 | Male | Early OA | 27.835 | 220.6 | 137.9 | 366.182 | 59 | <1.95 | 28.1 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH073 | Male | Early OA | 28.076 | 449 | 120.4 | 569.702 | 48 | <1.95 | 25.8 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH_98 | Male | Early OA | 28.998 | 52 | 0.2 | 45.5 | 962.8 | 2 | 32.9 | <10,24 | - | - |
| | SPH076 | Male | Early OA | 31.161 | 292.5 | 27.6 | 947.448 | 52 | <1.95 | 22 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH068 | Male | Healthy | 25.04 | 215.2 | 68.3 | 707.695 | 53 | <1.95 | 29.2 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH088 | Male | Healthy | 25.112 | 208.6 | 72.8 | 506.488 | 53 | <1.95 | 22 | <10,24 | 0= 0 points | 0= 0 points |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SPH090 | Male | Healthy | 25.436 | 163.6 | 57.6 | 738.531 | 44 | <1.95 | 25.8 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH003 | Male | Healthy | 25.912 | 344.691 | 9.898 | 474.673 | 46 | <1.95 | 0.253171 | 5.16309 | 0= 0 points | 0= 0 points |
| | SPH050 | Male | Healthy | 26.15 | 222 | 144.2 | 625.979 | 46 | <1.95 | 25.6 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH013 | Male | Healthy | 26.504 | 194.114 | 48.986 | 1048.179 | 50 | 2.373827 | 32.25557 | 2.12401 | 0= 0 points | 0= 0 points |
| | SPH066 | Male | Healthy | 27.668 | 198.9 | 146.6 | 986.764 | 47 | <1.95 | 26.6 | <10,24 | - | - |
| | SPH070 | Male | Healthy | 27.932 | 158.7 | 69.6 | 739.302 | 41 | <1.95 | 27.7 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH020 | Male | Healthy | 27.933 | 221.2 | 72.85 | 597.971 | 56 | 2.72001 | 1.393236 | 1.164607 | 0= 0 points | 0= 0 points |
| | SPH034 | Male | Healthy | 28.025 | 305.9 | 68.1 | 1106.255 | 47 | <1.95 | 27.7 | <10,24 | 0= 0 points | 0= 0 points |
| | SPH011 | Male | Healthy | 28.713 | 193.663 | 82.997 | 986.324 | 51 | 2.967284 | 7.657797 | <10.24 | 0= 0 points | 0= 0 points |
| | SPH005 | Male | Healthy | 29.297 | 280.789 | 84.611 | 656.205 | 52 | <1.95 | 13.92327 | <10.24 | 0= 0 points | 0= 0 points |
| | SPH085 | Male | Healthy | 29.584 | 173.9 | 32.2 | 638.313 | 50 | <1.95 | 24.7 | <10,24 | - | - |
| | SPH024 | Male | Healthy | 31.028 | 177.366 | 49.868 | 682.026 | 48 | 3.461831 | 8.359417 | 1.086966 | - | - |
| | SPH056 | Male | Healthy | 35.834 | 171.5 | 112.1 | 1088.524 | 39 | <1.95 | 54.8 | <10,24 | - | - |

| UNIC | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Selected | SECOND SORTING FEMALES | | | | | | | | | | | |
| | | | | | 1 | 2 | 3 | 4 | 5 | | | |
| | Sample | Sex | K&L L | K&L R | BMI | COMP | HA | PCIIP | Age | IL1b | TNFa | IL6 |
| | CYU161 | Female | 2= 3-4 points | 2= 3-4 points | 24.61 | 240.34 | 52.8 | 100.62 | 79 | <3.9 | <23.4 | <10.2 |
| | CYU129 | Female | 2= 3-4 points | 2= 3-4 points | 27.04 | 229.19 | 21.03 | 26.5 | 50 | <3.9 | <23.4 | <10.2 |
| | CYU004 | Female | 2= 3-4 points | 2= 3-4 points | 27.68 | 122.94 | 16.87 | 179.11 | 81 | <3.9 | <23.4 | <10.2 |
| | CYU171 | Female | 2= 3-4 points | 2= 3-4 points | 28.52 | 142.96 | 212.59 | 152.07 | 69 | <3.9 | <23.4 | <10.2 |
| | CYU172 | Female | 2= 3-4 points | 2= 3-4 points | 29.3 | 227.94 | 255.66 | 168.62 | 80 | <3.9 | <23.4 | <10.2 |
| | CYU163 | Female | 2= 3-4 points | 2= 3-4 points | 29.38 | 255.32 | 43.45 | 95.61 | 73 | <3.9 | <23.4 | <10.2 |
| | CYU118 | Female | 2= 3-4 points | 2= 3-4 points | 29.72 | 342.64 | 40.29 | 214.22 | 66 | <3.9 | <23.4 | <10.2 |
| | CYU168 | Female | 2= 3-4 points | 2= 3-4 points | 31.22 | 282.46 | 20.35 | 17.88 | 63 | <3.9 | 14.6 | <10.2 |
| | CYU175 | Female | 2= 3-4 points | 2= 3-4 points | 39.54 | 206.54 | 122.09 | 168.62 | 74 | <3.9 | 45.43 | <10.2 |
| | CYU020 | Female | 2= 3-4 points | 2= 3-4 points | 40 | 187.5 | 32.51 | 176.56 | 64 | <3.9 | <23.4 | <10.2 |
| | | | | | | | | | | | | |
| | CYU009 | Female | 1= 1-2 points | 3= 5-9 points | 20.78 | 123.45 | 52 | 163.16 | 65 | <3.9 | <23.4 | <10.2 |
| | CYU174 | Female | 3= 5-9 points | 3= 5-9 points | 24.24 | 162.81 | 29.13 | 124.48 | 60 | <3.9 | 33.7 | 168 |
| | CYU127 | Female | 3= 5-9 points | 3= 5-9 points | 24.97 | 393.31 | 285.96 | 28.47 | 79 | <3.9 | <23.4 | <10.2 |
| | CYU110 | Female | 3= 5-9 points | 3= 5-9 points | 25 | 291.36 | 88.74 | 271.6 | 76 | <3.9 | <23.4 | <10.2 |
| | CYU167 | Female | 3= 5-9 points | 2= 3-4 points | 25.39 | 384.62 | 69.92 | 86.58 | 73 | <3.9 | <23.4 | <10.2 |
| | CYU173 | Female | 3= 5-9 points | 3= 5-9 points | 26.67 | 226 | 147.16 | 224.72 | 71 | <3.9 | <23.4 | <10.2 |
| | CYU002 | Female | 3= 5-9 points | 3= 5-9 points | 26.95 | 265.46 | 176.5 | 220.56 | 70 | <3.9 | <23.4 | <10.2 |
| | CYU013 | Female | 3= 5-9 points | 3= 5-9 points | 26.95 | 191.37 | 20.86 | 140.79 | 51 | <3.9 | <23.4 | <10.2 |
| | CYU102 | Female | 3= 5-9 points | 3= 5-9 points | 27.34 | 557.99 | 158.38 | 110.69 | 84 | <3.9 | <23.4 | <10.2 |
| | CYU023 | Female | 0= 0 points | 3= 5-9 points | 28.25 | 403.57 | 52.59 | 146.14 | 64 | <3.9 | <23.4 | <10.2 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CYU006 | Female | 3= 5-9 points | 2= 3-4 points | 28.65 | 333.75 | 81.12 | 189.23 | 74 | <3.9 | <23.4 | <10.2 |
| | CYU176 | Female | 3= 5-9 points | 2= 3-4 points | 28.72 | 248.37 | 93.54 | 146.55 | 66 | <3.9 | <23.4 | <10.2 |
| | CYU012 | Female | 1= 1-2 points | 3= 5-9 points | 29.14 | 420.4 | 66.21 | 177.38 | 77 | <3.9 | <23.4 | <10.2 |
| | CYU113 | Female | 3= 5-9 points | 3= 5-9 points | 30.49 | 326.39 | 16.84 | 38.97 | 65 | <3.9 | <23.4 | <10.2 |
| | CYU166 | Female | 3= 5-9 points | 3= 5-9 points | 31.14 | 410.62 | 7.13 | 88.22 | 51 | <3.9 | <23.4 | <10.2 |
| | CYU162 | Female | 3= 5-9 points | 3= 5-9 points | 31.22 | 210.05 | 52.56 | 94.46 | 58 | <3.9 | <23.4 | <10.2 |
| | CYU119 | Female | 3= 5-9 points | 3= 5-9 points | 31.63 | 104.03 | 14.73 | 8.77 | 80 | <3.9 | <23.4 | <10.2 |
| | CYU112 | Female | 1= 1-2 points | 3= 5-9 points | 32.44 | 344.45 | 28.33 | 67.62 | 75 | <3.9 | <23.4 | <10.2 |
| | CYU114 | Female | 3= 5-9 points | 3= 5-9 points | 32.81 | 334.47 | 19.74 | 177.12 | 77 | <3.9 | <23.4 | <10.2 |
| | CYU024 | Female | 1= 1-2 points | 3= 5-9 points | 33.2 | 478.98 | 81.26 | 175.82 | 67 | <3.9 | <23.4 | <10.2 |
| | CYU157 | Female | 3= 5-9 points | 3= 5-9 points | 33.2 | 402.48 | 124.31 | 102.67 | 66 | <3.9 | <23.4 | <10.2 |
| | CYU160 | Female | 3= 5-9 points | 3= 5-9 points | 33.98 | 421.83 | 54.32 | 11.23 | 79 | <3.9 | 22.9 | <10.2 |
| | CYU137 | Female | 3= 5-9 points | 3= 5-9 points | 34.96 | 315.97 | 172.72 | 168.5 | 74 | <3.9 | <23.4 | <10.2 |
| | CYU125 | Female | 3= 5-9 points | 1= 1-2 points | 44.44 | 184.55 | 86.08 | 38.32 | 74 | <3.9 | <23.4 | 10.6 |
| | CYU003 | Female | 3= 5-9 points | 3= 5-9 points | 46.88 | 539.1 | 37.03 | 164.39 | 72 | <3.9 | <23.4 | 78.4 |
| | | | | | | | | | | | | |
| | CYU159 | Female | 0= 0 points | 4= 10 points | 29.38 | 519.48 | 154.04 | 14.59 | 73 | <3.9 | <23.4 | <10.2 |
| | CYU165 | Female | 4= 10 points | 3= 5-9 points | 31.22 | 727.33 | 127.94 | 179.5 | 67 | <3.9 | <23.4 | <10.2 |
| | CYU136 | Female | 4= 10 points | 3= 5-9 points | 33.3 | 786.52 | 201.68 | 266.03 | 84 | <3.9 | <23.4 | <10.2 |
| | CYU104 | Female | 4= 10 points | 4= 10 points | 37.04 | 492.09 | 147.57 | 104.61 | 60 | <3.9 | <23.4 | <10.2 |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| | SECOND SORTING MALES | | | | | | | | | | | |
| | | | | | 1 | 2 | 3 | 4 | 5 | | | |
| | Sample | Sex | K&L L | K&L R | BMI | COMP | HA | PCIIP | Age | IL1b | TNFa | IL6 |
| | | | | | | | | | | | | |
| | CYU170 | Male | 2= 3-4 points | 3= 5-9 points | 27.47 | 289.89 | 120.17 | 484.96 | 72 | <3.9 | <23.4 | <10.2 |
| | CYU005 | Male | 1= 1-2 points | 3= 5-9 points | 28.09 | 586.12 | 131.7 | 212.99 | 81 | <3.9 | <23.4 | <10.2 |
| | CYU106 | Male | 3= 5-9 points | 3= 5-9 points | 29.41 | 281.41 | 46.67 | 130.59 | 75 | <3.9 | <23.4 | <10.2 |
| | CYU117 | Male | 2= 3-4 points | 3= 5-9 points | 30.93 | 293.03 | 78.92 | 274.31 | 50 | <3.9 | <23.4 | <10.2 |
| | CYU111 | Male | 3= 5-9 points | 2= 3-4 points | 31.14 | 546.05 | 142.11 | 58.92 | 84 | <3.9 | <23.4 | <10.2 |
| | CYU109 | Male | 2= 3-4 points | 3= 5-9 points | 31.96 | 323.62 | 43.72 | 194.2 | 59 | <3.9 | <23.4 | <10.2 |
| | CYU124 | Male | 1= 1-2 points | 3= 5-9 points | 36.9 | 556.97 | 95.49 | 43.24 | 66 | <3.9 | <23.4 | <10.2 |
| | CYU019 | Male | 3= 5-9 points | 3= 5-9 points | 37.98 | 232.62 | 54.57 | 173.6 | 66 | <3.9 | <23.4 | <10.2 |
| | | | | | | | | | | | | |
| | CYU123 | Male | 3= 5-9 points | 4= 10 points | 25.95 | 995.47 | 57.05 | 21.82 | 71 | <3.9 | <23.4 | 24.94 |
| | CYU156 | Male | 4= 10 points | 4= 10 points | 33.06 | 510.81 | 171.61 | 83.05 | 80 | <3.9 | <23.4 | <10.2 |
| | CYU105 | Male | 4= 10 points | 4= 10 points | 36.72 | 612.82 | 128.99 | 427.38 | 80 | <3.9 | <23.4 | <10.2 |

| ANIMUS | | | | |
|---|---|---|---|---|
| Selected | Number | Sex | BMI | Age |
| | GRA007 | Female | 20.1 | 19 |
| | GRA033 | Female | 20.3 | 22 |

| | | | | |
|---|---|---|---|---|
| | GRA069 | Female | 20.8 | 21 |
| | GRA028 | Female | 20.8 | 20 |
| | GRA019 | Female | 21.5 | 15 |
| | GRA043 | Female | 21.8 | 21 |
| | GRA056 | Female | 23.0 | 21 |
| | GRA031 | Female | 23.3 | 33 |
| | GRA039 | Female | 24.3 | 18 |
| | Number | Sex | BMI | Age |
| | GRA004 | Male | 20.9 | 25 |
| | GRA027 | Male | 21.1 | 24 |
| | GRA052 | Male | 22.3 | 38 |
| | GRA021 | Male | 22.9 | - |
| | GRA037 | Male | 24.5 | 21 |
| | GRA034 | Male | 24.9 | 27 |
| | GRA030 | Male | 25.0 | 31 |
| | GRA067 | Male | 26.0 | 37 |
| | GRA074 | Male | 26.0 | 32 |
| | GRA046 | Male | 26.6 | 29 |
| | GRA063 | Male | 26.6 | 26 |
| | GRA075 | Male | 26.7 | 43 |
| | GRA024 | Male | 27.8 | 29 |
| | GRA026 | Male | 31.1 | 26 |

For example, Table 5 shows the quality control Cq results from HULAFE plasma samples. miRNA control from RNA extraction, UniSp2, UniSp4, and UniSp5, that were at 160 fmol, 1.6 fmol, and 0.016 fmol concentrations, showed an increase of mean of Cq values of 15.1, 21.7, and 29.6, respectively, that was equivalent to a 64-fold increase among UniSp2 and UniSp4, and 234-fold among UniSp4 and UniSp5. Similar results were obtained for all plates (results not shown). Also, cDNA synthesis control, UniSp6 (concentration not supplied), and cel-miR-39-3p (0.16 fmol) gave a mean Cq of 16.3 and 24.7, respectively. Comparing values from the former UniSp4, UniSp5 and cel-miR-39-3p, the concentration was, as expected, 7.8-fold among UnisP4 and cel-miR-39-3p, and 29-fold among cel-miR-39-3p and UniSp5.

***Table 5*** *– Partial set of* ***S****amples included in the miRNA quality analysis.*

| Sample | mir103 | mir191 | mir451 | mir23a | UniSp6 | UniSp2 | UniSp4 | UniSp5 | cel-miR-39-3p | UniSp3 | mir124 | miR30c | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 31.7 | 30.6 | 24.52 | 30.64 | 16.26 | 14.81 | 21.55 | 29.3 | 25.11 | 19.66 | 29.12 | 32.14 | most tissues |
| 2 | 30.66 | 26.89 | 24.09 | 28.14 | 16.57 | 15.03 | 21.86 | 29.16 | 25.1 | 18.56 | 33.05 | 31.67 | brain |
| 3 | 29.01 | 29.2 | 21.46 | 28.49 | 16.49 | 16.16 | 22.95 | 30.04 | 25.36 | 15.79 | 24.32 | 29.83 | kidney |
| 4 | 30.36 | 30.51 | 24.68 | 28.97 | 16.34 | 14.84 | 21.5 | 29.3 | 25.12 | 19.73 | 32.47 | 32.24 | blood |
| 6 | 26.7 | 25.46 | 20.02 | 25.76 | 16.32 | 14.89 | 21.51 | 28.64 | 19.5 | 9.79 | 21.15 | 26.28 | mRNA extraction |
| 7 | 30.02 | 27.53 | 22.68 | 29.19 | 16.65 | 14.83 | 21.49 | 29.14 | 25.04 | 19.78 | 33.33 | 32.12 | cDNA synthesis |
| 9 | 31.52 | 30.22 | 23.16 | 23.82 | 15.89 | 12.88 | 21.44 | 28.77 | 24.48 | 11.48 | 25.1 | 27.87 | PCR efficiency |
| 10 | 28.32 | 28.78 | 23.95 | 30 | 16.32 | 14.74 | 21.47 | 29.17 | 25.01 | 19.31 | 32.85 | 33.62 | |
| 11 | 30.62 | 24.76 | 23.87 | 23.96 | 15.5 | 14.7 | 22.07 | 28.45 | 24.8 | 10.96 | 26.27 | 28.94 | |
| 12 | 27.73 | 25.82 | 22.24 | 26.71 | 16.45 | 14.53 | 21.09 | 28.94 | 25.12 | 19.53 | 33.08 | 31.61 | |
| 13 | 30.56 | 29.01 | 23.56 | 24.23 | 16 | 14.72 | 16.01 | 28.95 | 23.06 | 11.08 | 25.12 | 31.89 | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 16 | 22.36 | 16.85 | 21.88 | 27.32 | 16.53 | 15 | 21.7 | 29.67 | 25.15 | 19.68 | 29.04 | 31.31 |
| 17 | 28.73 | 29.14 | 22.28 | 26.89 | 15.11 | 8.99 | 21.42 | 28.9 | 25.44 | 19.91 | 22.14 | 30.35 |
| 18 | 25.32 | 17.47 | 23.59 | 29.43 | 16.55 | 15.03 | 21.79 | 29.57 | 25.32 | 19.61 | 30.25 | 31.76 |
| 19 | 28.12 | 28.8 | 21.84 | 26.97 | 12.63 | 15.23 | 21.65 | 29.11 | 20.64 | 10.78 | 19.37 | 21.59 |
| 20 | 28 | 21.14 | 21.79 | 26.95 | | 15.27 | 21.68 | 29.56 | 25.53 | 19.67 | 28.03 | 29.78 |
| 21 | 29.73 | 29.05 | 22.83 | 23.07 | 16.55 | 15.83 | 18.98 | 29.12 | 24.31 | 8.7 | 31.5 | 31.44 |
| 22 | 28.79 | 31.99 | 22.48 | 27.65 | 16.76 | 15.28 | 21.69 | 29.66 | 25.03 | 19.79 | 29.3 | 31.3 |
| 23 | 29.8 | 30.95 | 22.27 | 26.53 | 16.62 | 14.73 | 20.74 | | 21.87 | 19.59 | 31.96 | 27.45 |
| 24 | 28.19 | 26.89 | 22.09 | 27.9 | 16.89 | 15.24 | 21.77 | 29.7 | 25.17 | 19.91 | 31.57 | 30.86 |
| 25 | 28.35 | 29.5 | 21.5 | 27.53 | 16.44 | 14.54 | 21.3 | 28.92 | 24.35 | 19.86 | 23.56 | 30.64 |
| 26 | 27.88 | 26.66 | 21.68 | 26.93 | 16.48 | 15.14 | 21.88 | 29.49 | 25.27 | 19.86 | 31.21 | 29.05 |
| 27 | 29.04 | 26.65 | 21.91 | 26.83 | 16.51 | 16.51 | 21.47 | 28.17 | 25.69 | 12.78 | 30.22 | 29.89 |
| 28 | 27.12 | 30.68 | 23.3 | 27.55 | 16.56 | 16.81 | 23.07 | 30.48 | 25.25 | 19.56 | 31.79 | 31.99 |
| 29 | 29.61 | 28.04 | 26.84 | 28.14 | 16.56 | 15.85 | 22.81 | 30.75 | 24.8 | 19.54 | 25.91 | 30.77 |
| 30 | 28.14 | 25.19 | | 26.64 | 16.73 | 14.83 | 21.79 | 29.51 | 25.77 | 19.47 | 33.3 | 31.46 |
| 31 | 30.61 | 30.6 | 23.76 | 29.29 | 16.8 | 16.44 | 23.69 | 31.42 | 25.26 | 18.14 | 27.76 | 31.16 |
| 32 | 27.91 | 27.42 | 19.75 | 24.98 | 16.69 | 15.47 | 22 | 29.63 | 25.86 | 18.71 | 32.1 | 31.23 |
| 33 | 28.92 | 24.96 | 21.58 | 26.88 | 16.75 | 16.17 | 22.92 | 30.34 | 25.35 | 20.36 | | 37.68 |
| 34 | 28.61 | 29.33 | 21.18 | | 15.89 | 16.09 | 23.14 | 29.86 | 25.22 | 19.72 | 31.21 | 32.35 |
| 35 | 28.46 | 29.18 | 21.94 | 27.44 | 16.62 | 16.26 | 22.93 | 31.23 | 25.12 | 19.62 | 29.32 | 29.8 |
| 36 | 29.75 | 30.49 | 22.22 | 27.92 | 16.76 | 16.58 | 23.11 | 31.8 | 25.37 | 19.51 | 31.18 | 31.84 |
| Mean | 28.77 | 27.49 | 22.61 | 27.19 | 16.30 | 15.11 | 21.70 | 29.57 | 24.67 | 17.51 | 28.92 | 30.68 |
| SD | 1.8 | 3.6 | 1.4 | 1.8 | 0.8 | 1.4 | 1.4 | 0.9 | 1.4 | 3.7 | 3.9 | 2.6 |

miR-451 and miR-23a-3p are found in plasma and serum and serve as haemolysis and internal control markers, respectively. Even more, they have been described also as cancer biomarkers (Karimi et al., 2017; Bai H et al., 2019) and miR-451 was increased in the OA articular cartilage (BD, Schwartz Z. Regulation of inflammatory and catabolic responses to IL-1β in rat articular chondrocytes by microRNAs miR-122 and miR-451 and miR-23a promotes OA (Kang L et al., 2016). Then, it is expected to detect both miRNAs in serum samples and urine samples at the same levels as indicated by the commercial supplier (Qiagen – miRCURY LNA miRNA QC PCR Panel Handbook). Taken all values together (males and females in all samples, and control vs early OA in HULAFE samples), no differences were observed for plasma samples, but in urine samples, miR-451 Cq was much lower and miR-23a slightly higher (Table 6).

***Table 6*** *– Whole Cq mean values of miR-451 and miR-23a from the three sample origins. Male and females were pooled. Values from Qiagen were taken from the miRCURY LNA miRNA QC PCR Panel Handbook*

| | | miR-451 | miR-23a |
|---|---|---|---|
| QIAGEN | Urine | 23 | 30 |
| | Plasma | 21 | 28 |
| O'Active | HULAFE urine | 30.1 ± 3.8 | 26.7 ± 2.9 |
| | HULAFE plasma | 22.0 ± 1.4 | 29.0 ± 1.8 |
| | CYU plasma | 22.8± 1.6 | 29.5 ± 1.1 |
| | Animus plasma | 24.2 ± 0.7 | 29.5 ± 0.6 |

As a preliminary analysis using values of the quality control analysis, miR-451 and miR-23A values from HULAFE healthy and Early OA, UNIC K&L L/K&L R 2, K&L L/K&L R 3 and K&L L/K&L R 4, and ANIMUS were compared. Males and females were analyzed separately. Figure 3 shows the results from males and females. Mir-23a was significantly overexpressed in males and females K&L L/K&L R 2 and K&L L/K&L R 3 compared to healthy volunteers and patients with early OA, and in males, also ANIMUS samples were overexpressed. However, miR-451 is underexpressed in K&L L/K&L R 4 in female patients, but not in male patients, where ANIMUS athletes showed an overexpression. Results from miR-23a are expected according to the literature, but not for miR-451, that it has been described its overexpression in OA (Bottani M et al. 2020). However, the type of samples -synovial liquid vs plasma in this study- could affect our results. Also, this is a preliminary study but a limited number of samples, so it is necessary to have complete data from all participants included in this study.

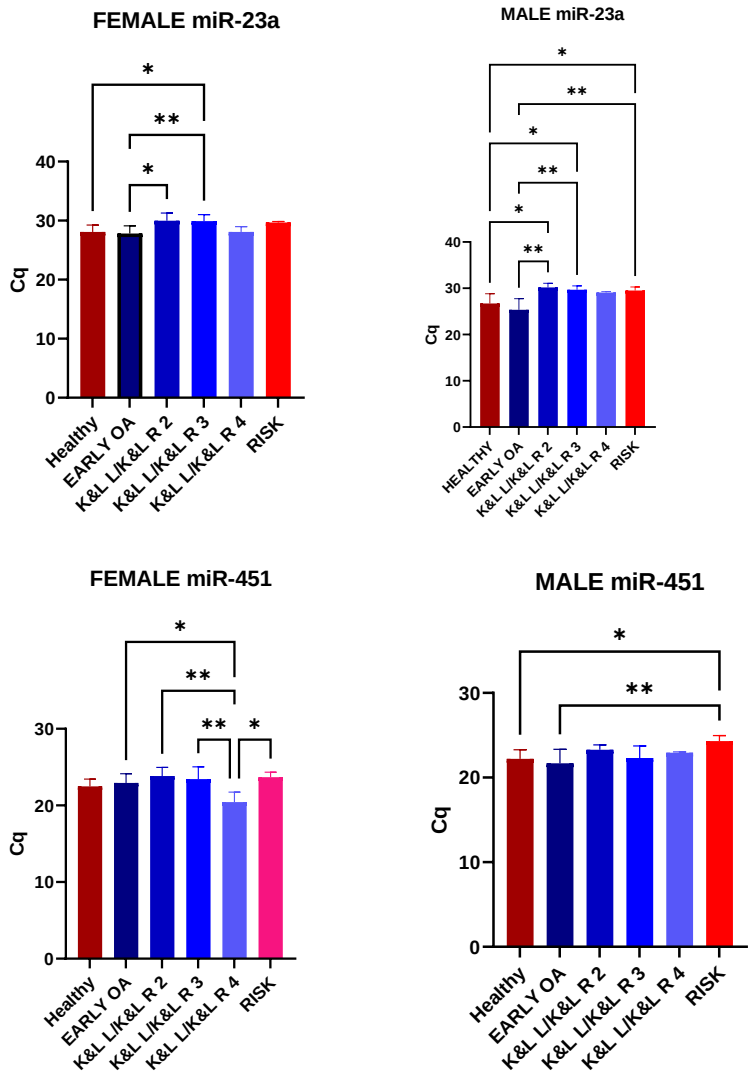



***Figure 3*** *– Graphical representation of Cq values of miR-23a and miR-451 from healthy volunteers and patients with early OA (HULAFE), UNIC patients, and athletes with risk of OA (ANIMUS). Ordinary one-way ANOVA statistical analysis was done with Graph Prism 9.0.*

### 5.5 TaqMan™ OpenArray™ miRNA analysis

This assay is designed for the analysis of miRNA expression levels by real-time PCR. Each array contains 754 different miRNAs. Three samples/array were run per plate and a total of 102 samples are under analysis. At this time, 90 samples were run and are under analysis. Figure 4 outlines the procedure. The protocol consists of 102 samples, representing males and females, all OA stages, and K&L L/K&L R scores, together with the BMI. miRNA that correlated with OA stages combined with BMI for males and females, will be selected for further validation with the rest of the samples.

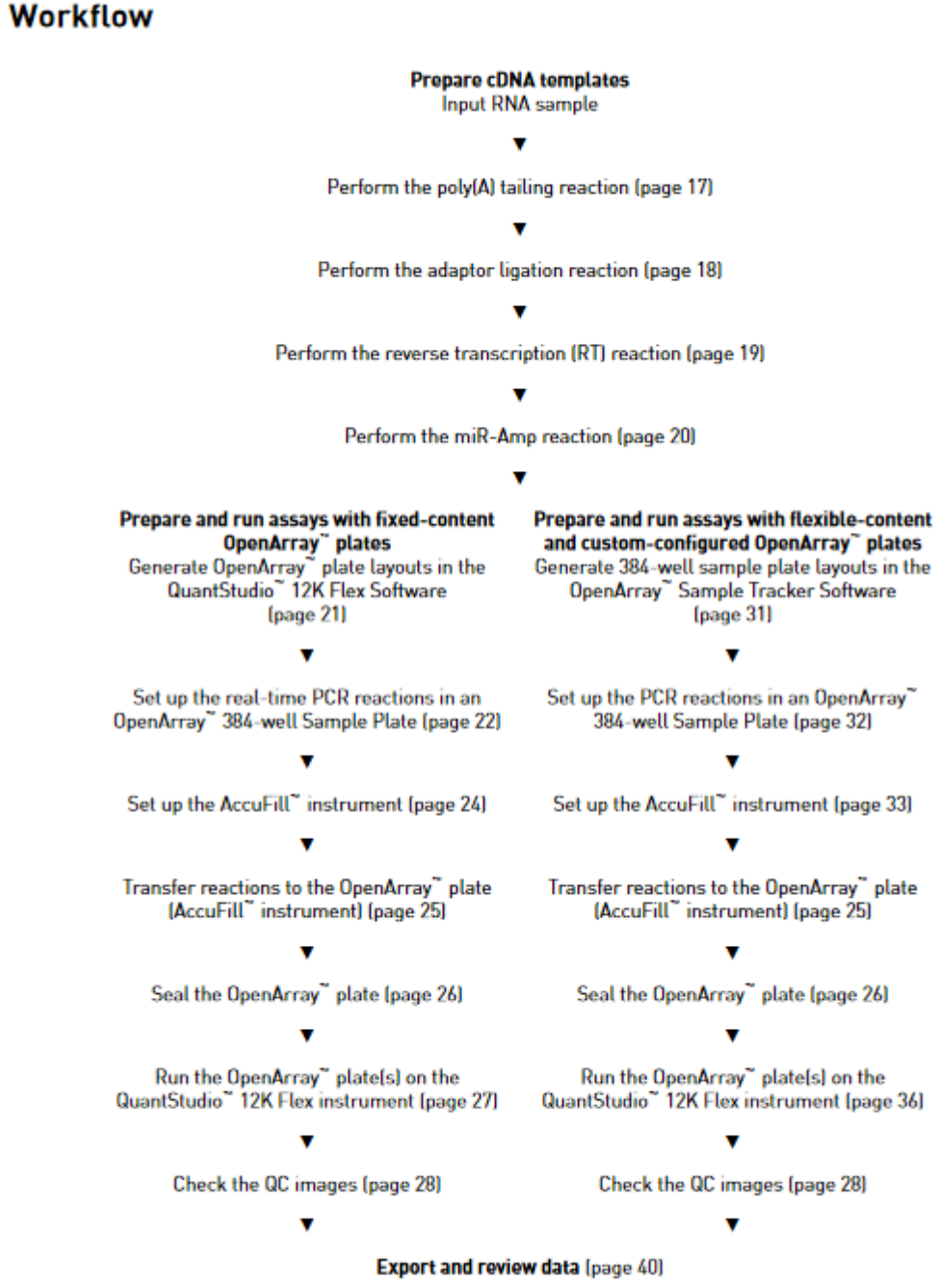


**Figure 4.** Workflow of miRNA analysis. From the handbook "TaqMan™ Advanced miRNA Assays – User guide" Applied Biosystems – Thermo Fisher scientific.

### 5.6 Deviation from the expected work

There have been deviations from the proposed work in Task 4.3.

- It was expected to include in the study the protein and miRNA exosome content. However, due to the small amount of plasma and urine samples sent, it was decided to concentrate the study on miRNA, together with the standard biomarkers included in Task 4.2 for HULAFE samples.
- COVID-19 restrictions, mainly during the lockdown on March-June 2020 and mobility restrictions from the rest of the year, forced a delay both in recruiting patients and laboratory work, so it was not possible to follow the expected schedule. For this reason, at this time we have run all samples and we are in the process of analyzing the Taqman OpenArray results. During this second quarter, we plan to validate the selected miRNA and run all samples for the final correlation analysis with clinical and biomarkers data generated and transfer the whole data to the consortium partners involved in computer modelling.

## 6 Microbiome analysis

### 6.1 Introduction

The gut microbiota has a critical role in the immune system's development, function, and inflammatory reactions (Rooks and Garrett, 2016). Perturbations in the microbiome can activate the innate immune system and may lead to enhanced production of pro-inflammatory cytokines which could affect multiple organs, including the joint (Boer et al., 2019). The intestinal microbiome composition is also highly associated with a large number of pathological disorders like obesity (Isolauri, 2017), inflammatory bowel diseases (Serban, 2015), aging (Aleman and Valenzano, 2019) and metabolic syndrome (Collins *et al.*, 2015). Many of the pathologies listed are shared between OA and microbiota imbalance. Consequently, there is growing evidence suggesting that changes in the gut microbiome may also be linked to developmental factors in the onset of musculoskeletal disorders such as OA.

The study of the gut microbiome and its relationship with the development of OA is an emerging area of study. The few published research papers to date include mostly murine models fed with a high-fat diet in order to induce obesity and to affect joint integrity. Very few studies involving human microbiome intestinal samples have been carried out to date, and all of them have been conducted in subjects diagnosed with OA. For this reason, there is an interest in exploring the relationship between the gut microbiome and the onset of OA in order to investigate whether there are existing early indicators/markers for the initiation of the OA process. In the course of this experimental research, 82 stool samples were provided by volunteers from La Fe University Hospital, in Valencia. Twenty of them are part of the control group without OA and 62 samples belong to volunteers with an initial OA.

### 6.2 Material and methods

#### 6.2.1 DNA extraction

Fecal human samples were collected in Stool Nucleic Acid Collection and Preservation Tubes (Cat. 45630, 45660) from Norgen Biotek. The samples were stored at room temperature and shipped to LEITAT facilities. Total genomic DNA was extracted from a 150 mg fecal sample aliquot using a repeated bead beating method from ZymoBIOMICS DNA/RNA Miniprep Kit following the supplier instructions. The DNA extraction concentration was measured using Qubit 4 Fluorometer from Thermo Fisher.

#### 6.2.2 16S rRNA V3-V4 metagenomic sequencing

Metagenomic analysis was performed using Illumina Miseq sequencing platform according to standard protocols devised by Illumina. The V3-V4 region of 16S rRNA with 515F and 806R primers designed for

dual indexing genes was targeted to characterize and estimate bacterial communities present in faecal samples. Polymerase chain reaction (PCR) was performed in 25µl reaction volume containing 1µl of template DNA (5ng/µl), 5µl of each primer (1µM) and 12.5 µl of 2x KAPA HiFi HotStart ReadyMix. Miseq amplicons library preparation. PCR was carried out under the following conditions: initial denaturation for 47 min at 94°C, followed by 25 cycles of denaturation for 45 sec at 94°C, annealing for 60 sec at 52°C and elongation for 90 sec at 72°C, and a final elongation step for 10 min at 72°C. Duplicates were combined, purified with the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel), and quantified using the Quant-iT PicoGreen dsDNA kit (Invitrogen). Equimolar concentrations along with Illumina sequence adapter and index primers were utilized in emulsion PCR to generate amplicon libraries followed by PCR clean up.
For bioinformatic analysis, sequences were imported into QIIME2 q2cli v2019.10 (Bolyen et al., 2019). The overall quality was inspected with the demux plugin (https://github.com/qiime2/q2-demux), and the total of reads was processed with DADA2 (Callahan et al., 2016) in order to remove trim reads, correct errors, merge reads pairs and remove PCR chimeras and to finally obtain representative ASV sequences and the abundances of each ASV.

The representative sequences were classified taxonomically using the q2-feature-classifier classify-conseunsus-vsearch(https://github.com/qiime2/q2-feature-classifier). SILVA v132 16S-only reference sequences and taxonomy were downloaded from the SILVA website for QIIME. SILVA v132, clustered at 99% database was used as reference reads.

Finally, the feature table obtained was normalized with q2-gcn-norm (https://github.com/Jiung-Wen/q2-gcn-norm). This plugin normalizes sequences by 16S rRNA gene copy number (GCN) based on rrnDB database.

### 6.2.3 Statistical analysis

Statistical analysis was performed using R 4.0.3 (*R: The R Project for Statistical Computing*,) and QIIME2 (Bolyen *et al.*, 2019).

## 6.3 Results

This Chapter outlines all the results obtained regarding the OActive deliverable 4.2.

### 6.3.1 OA pilot volunteers profile

A total of 82 volunteers agreed to provide a stool sample for the OActive pilot study. The samples were collected at LaFe University Hospital in Valencia and were shipped to LEITAT's facilities where they were properly stored and processed. The female group is higher than the male group 67% vs 33%. These results are consistent with the higher incidence of OA in women compared to men.

*Table 7 Depicted are the mean and ± the Standard deviation. P-values were obtained using the Mann-Whitney U test for comparison between groups. BMI – Body Mass index, WOMAC - Western Ontario and McMaster Osteoarthritis Index, LDL-C – Low-density lipoprotein cholesterol.*

| | *Females* | | | *Males* | | |
|---|---|---|---|---|---|---|
| | Initial OA | Healthy | p-value | Initial OA | Healthy | p-value |
| *Cohort participants* | 39 | 16 | - | 23 | 4 | - |
| *Age (years)* | 52,17 ± 5,71 | 52,75 ± 6,32 | 0,75 | 49,69 ± 3,30 | 49,69 ± 4,76 | 0,86 |
| *BMI (kg/m2)* | 26,30 ± 4,96 | 28,32 ± 4,22 | 0,194 | 26,80 ± 2,95 | 26,34 ± 3,88 | 0,78 |
| *WOMAC (pain score)* | 2,05 ± 1,86 | 1,21 ± 1,39 | **0,032** | 1,29 ± 1,18 | 0,22 ± 0,33 | **0,0022** |
| *LDL-C (mg/dL)* | 106,13 ± 25,27 | 99,68 ± 23,82 | 0,37 | 111,34 ± 28,14 | 113,25 ± 23,04 | 0,95 |
| *Triglycerides (mg/dL)* | 103,02 ± 42,72 | 118,43 ± 63,27 | 0,38 | 121,3 ± 54,58 | 83,75 ± 37,17 | 0,11 |

No differences were observed between the Healthy and Initial OA groups in age, BMI, lipoprotein levels, and blood triglycerides. However, a significant difference in the WOMAC score was observed. WOMAC is one of the scores used to diagnose OA. The difference in the male group is highly significant ($p=0.0022$) while in the female group, due to high variability, this difference is lower, but still significant ($p=0.032$).

### 6.3.2 Bacterial communities taxonomic profile

For the 82 faecal samples, genomic DNA extractions were performed to amplify the V3 -V4 hypervariable regions from the 16S rRNA gene. After quality control, 81 samples were included in the post-analysis and one sample was discarded due to not reaching the minimum number of sequences per sample required. 3,145,458 sequences were obtained after quality filtering, corresponding to 6195 unique ASVs (Amplicon Sequence Variant). The results of the bacterial taxonomy assignation classified 5945 ASVs to the genus level and 5188 to the species level.

Figure 4 illustrates the relative abundances at the phylum level for each of the individual samples as well as the mean composition for the Healthy and Initial OA groups. Section A of figure 5 shows the inter-individual differences between the volunteers. The bacterial taxonomic composition obtained for this cohort is very similar to other studies with cohorts of European Caucasians. (Boer et al., 2019)(Gupta, Paul, and Dutta, 2017) (Jackson *et al.*, 2018).

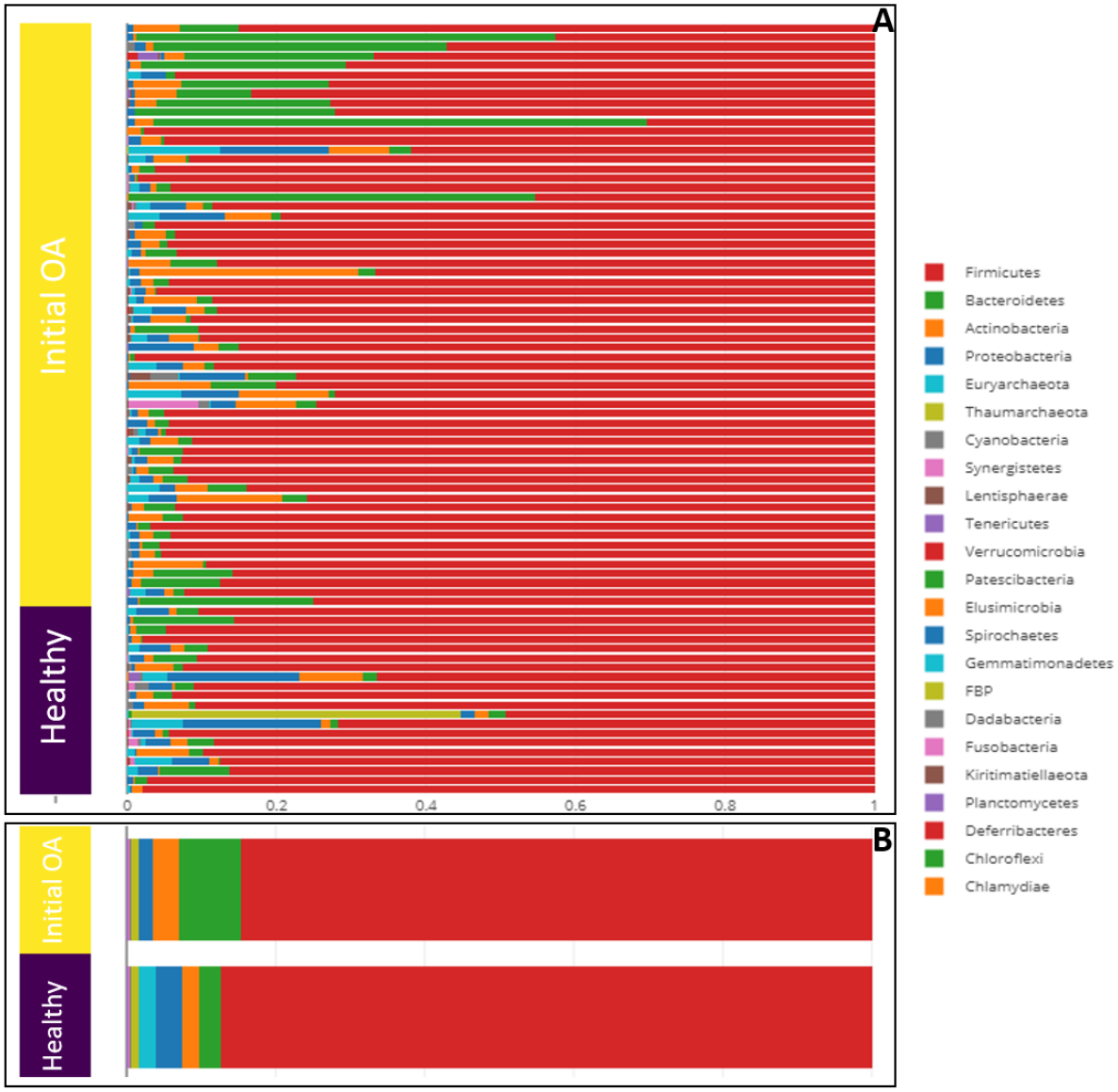



*Figure 5 Taxa plot at phylum level for A) each individual sample and B) collapsed for all samples.*

The most abundant phylum for each group is Firmicutes with 86.66% of the total abundance for the Initial OA group and 90.29% for the Healthy group, followed by Bacteroidetes 7.7% and 2.7% respectively. These two phyla are the most abundant in faecal samples extracted from adult individuals.

At the genus level, Figure 5 highlights the 20 most abundant genera representing about 85% of the relative abundance of the bacterial communities. The two most abundant genera found are *Faecalibacterium* and *Blautia*. Both genera are widely described in the state of the art and their variation is associated with different pathologies such as Crohn's disease. (Fujimoto *et al.*, 2013) Colorectal Cancer (Tilg *et al.*, 2018) for *Faecalibacterium* and obesity (Ozato *et al.*, 2019) for *Blautia*. Nevertheless, no differences in the abundance of these two genera are observed between the Healthy and Initial OA groups.

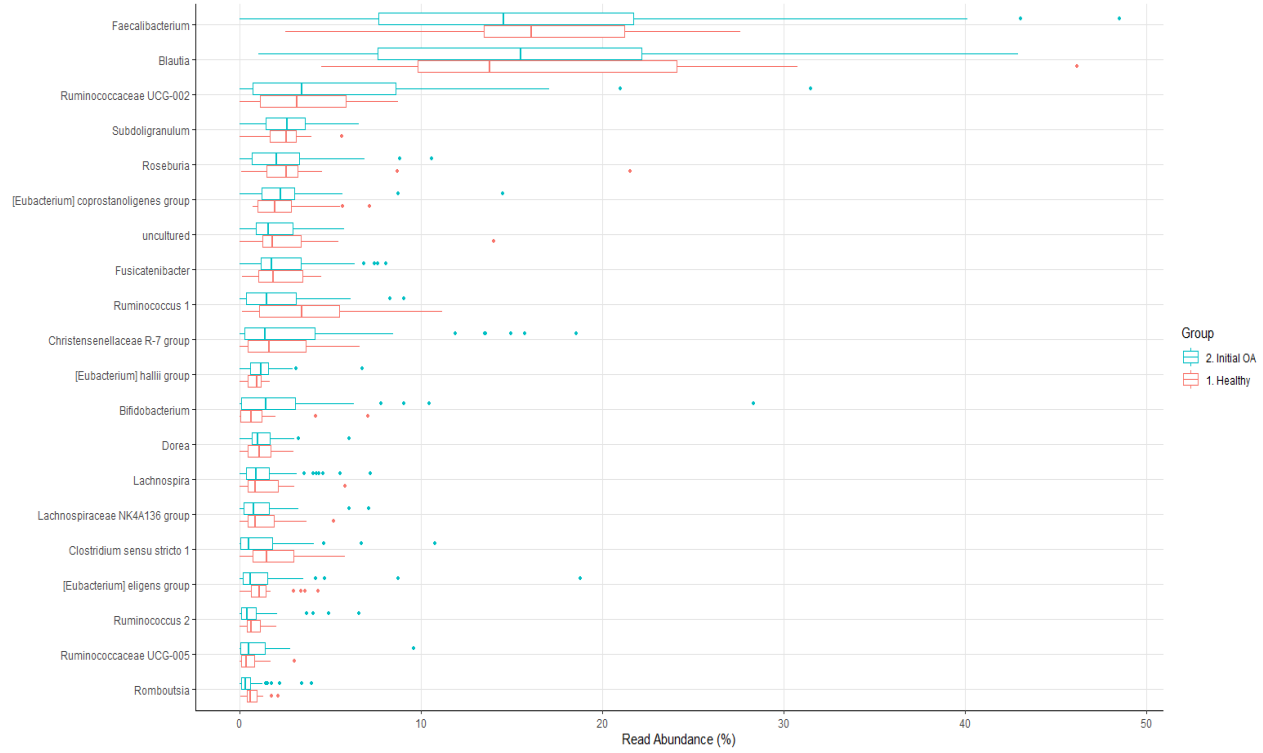



*Figure 6 Boxplot for the 20 most abundant genera detected.*

### 6.3.3 Alpha and Beta diversity

Alpha diversity can be used to calculate the diversity of each sample in terms of the number of different ASVs and beta diversity calculate the similarity between two samples or a group of samples, taking into account the phylogenetic distance between the different ASVs obtained.

Figure 7 shows the alpha diversity (Shannon index) of the bacterial communities from the 82 fecal samples (OA and Control groups). No significant differences in alpha diversity were detected between the Healthy and Initial OA groups. These results are consistent with those obtained in other similar studies involving patients diagnosed with OA and Control groups (Boer *et al.*, 2019) (Coulson *et al.*, 2013).

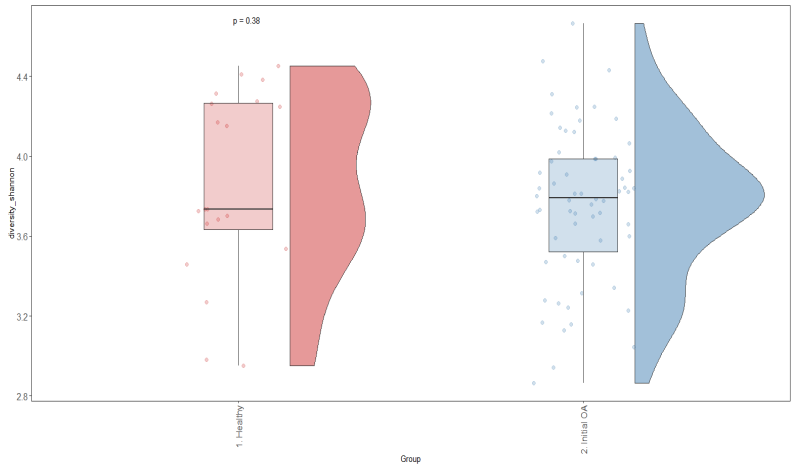



*Figure 7 Box plot with sample distribution for Healthy and Initial OA group. Mann-Whitney U test was used to compare the differences between groups.*

Beta-diversity analyses were also performed to investigate composition differences among the fecal bacterial communities from the samples corresponding to the two groups under study (Initial OA and Healthy). Comparisons between the Initial OA and Healthy group were made using 4 different beta-diversity metrics. Figure 8 plots the beta-diversity distance between the various microbiome samples. For none of the four metrics chosen a separation into distinct clusters is apparent, as both Initial OA and Healthy group bacterial communities are largely mixed.

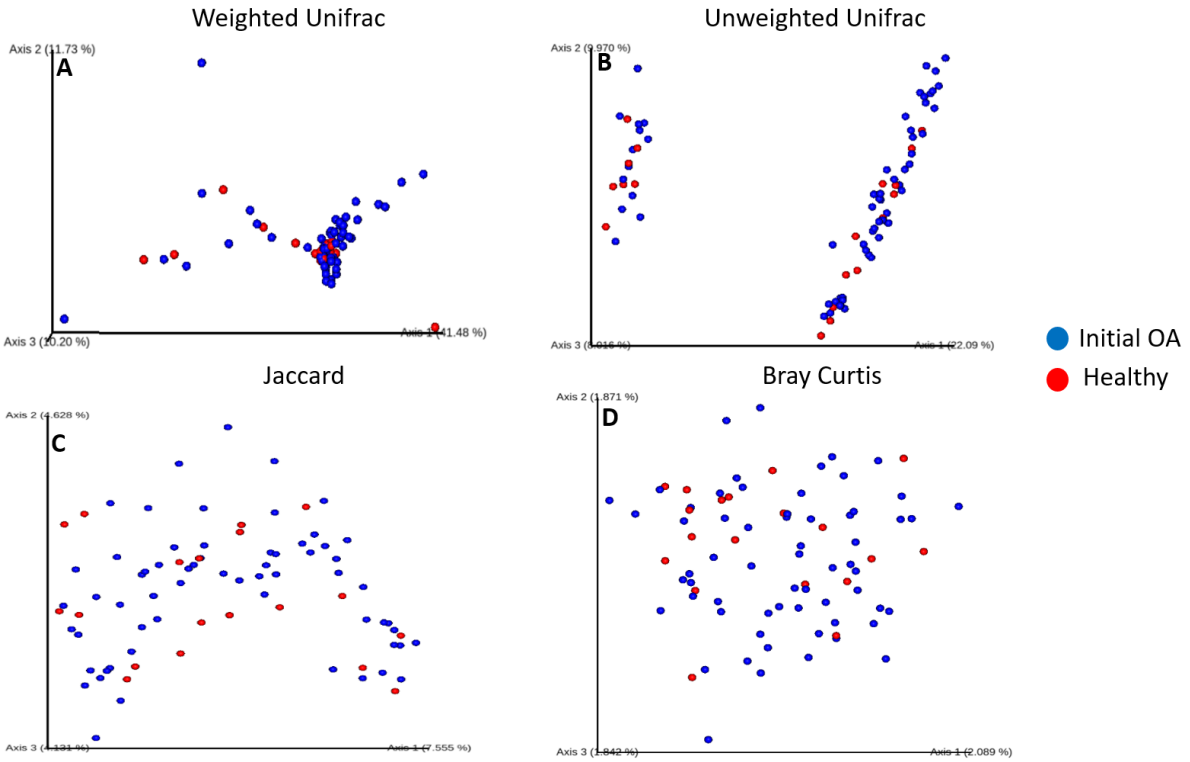



*Figure 8 Beta-diversity plots labelled by Initial OA and Healthy subjects. A. Weighted Unifrac, B. Unwaghted Unifrac, C. Jaccard and D. Bray Curtis*

To confirm the hypothesis of no compositional differences between the two groups, a PERMANOVA test was performed with 999 permutations for each of the beta-diversity metrics. The results can be found in table 8. None of the 4 metrics obtains a significant p-value result. This result confirms that there are no significant differences in the composition of the bacterial microbiome between the Healthy group and the Initial OA group.

*Table 8 PERMANOVA results after 999 permutations for the four betadiversity metrics analysed.*

*PERMANOVA RESULTS*

| | Sample size | Number of groups | Test statistic | p-value | Number of permutations |
|---|---|---|---|---|---|
| *Weighted Unifrac* | 81 | 2 | 1,366 | 0,219 | 999 |
| *Unweighted Unifrac* | 81 | 2 | 1,4192 | 0,127 | 999 |
| *Jaccard* | 81 | 2 | 0,954 | 0,604 | 999 |
| *Bray Curtis* | 81 | 2 | 0,965 | 0,861 | 999 |

### 6.3.4 Biomarker analysis

The last analysis was focused on the search for possible biomarkers using the variables generated from the volunteers' microbiome. A good biomarker should be more abundant in the Initial OA group over the Healthy group and should be able to predict the onset of OA. This is one of the main objectives of the project. Two approaches have been used to carry out this study. On the one hand a machine learning algorithm (Random Forest) and on the other hand a multivariate statistical comparison (Welch t-test). (Alekseyenko, 2016). The relative abundances between the Healthy group and the Initial OA group were compared for the 240 bacterial genera identified.

Random Forest (Breiman, 2001) is a machine learning method based on a decision tree algorithm, that can be used for classification and regression-based analysis. Random Forest is widely used in the discovery of new microbiome-associated biomarkers (Li et al., 2020) (Ren et al., 2019) (Wang, Xu and Xia, 2016)(Sprockett et al., 2019). Typically, around two-thirds of a particular study sample is used for the model fitting or training while the remaining one-third is used for model testing.

Receiver Operating Characteristic (ROC) curves are a graphical representation of the classification accuracy of a machine-learning model. The top-left corner of the plot represents the "optimal" performance position, indicating an FPR of zero and a TPR of one. A greater area under the curve (AUC) indicates better performance. This can be compared to the error rate achieved by random chance, which is represented in figure 9B as a diagonal line extending from the lower-left to upper-right corners. Additionally, the "steepness" of the curve is important, as a good classifier should maximize the TPR while minimizing the FPR.

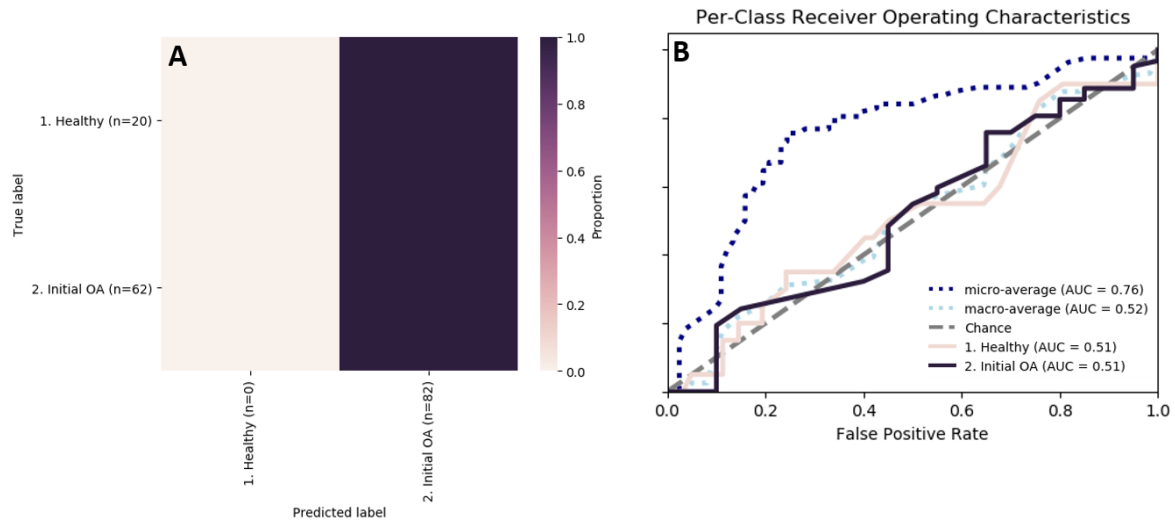



*Figure 9 A. Random forest classification model accuracy. B. The ROC curve plots showing the relationship between the true positive rate (TPR, on the y-axis) and the false positive rate (FPR, on the x-axis) "Micro-averaging" calculates metrics globally by averaging across each sample; hence class imbalance impacts this metric. "Macro-averaging" is another average metric, which gives equal weight to the classification of each sample.*

Figure 9A illustrates the classification by the Random Forest machine learning model. After using the microbiome variables as input, the model is not able to correctly classify the different samples into the group to which each of them belongs. None of the 20 samples belonging to the Healthy group were classified as Healthy, all of them were included in the Initial OA group. The same tendency is observed in figure 5B, where the Area under the curve (AUC) is 0.51 for both groups, very close to 0.5 which is the value obtained when the samples are randomly classified (discontinuous diagonal line).

After analyzing the Random Forest model results, we can conclude that assessing complete bacterial microbiomes does not allow us to differentiate specifically between the two groups.

Finally, to further investigate possible genera that may play a role as indicators of the onset of OA, we have used the Welch multivariate test, specifically designed to be applied to unbalanced data sets, as in this case. This test reduces type I error.

After comparing the relative abundances across the ASVs belonging to 240 different genera, a total of 9 markers with a p-value lower than 0.05 have been detected. Eight of these markers are enriched in the Initial OA group and only one in the Healthy group. However, after adjusting the p-values obtained by Benjamini & Hochberg's false discovery rate (FDR) the p-values become non-significant.

The relative abundance across the 9 potential biomarker candidates is shown in Figure 10.

*Table 10 Welch test results between Healthy and Initial OA group. P values and p adjusted with False Discovery rate Benjamini & Hochberg for all the genera detected are highlighted.*

| | *feature* | *enrich_group* | *diff_mean* | *pvalue* | *Padj (fdr BH)* |
|---|---|---|---|---|---|
| *marker1* | [Eubacterium] hallii group | 2. Initial OA | -0.00426 | 0.0146 | Ns |
| *marker2* | Alistipes | 2. Initial OA | -0.0044 | 0.0370 | Ns |
| *marker3* | Anaerostipes | 2. Initial OA | -0.00306 | 0.000940 | Ns |
| *marker4* | Barnesiella | 2. Initial OA | -0.000475 | 0.0141 | Ns |
| *marker5* | Coprococcus 3 | 2. Initial OA | -0.00357 | 0.00480 | Ns |
| *marker6* | Dialister | 2. Initial OA | -0.00186 | 0.005881 | Ns |
| *marker7* | Parabacteroides | 2. Initial OA | -0.000551 | 0.00365 | Ns |
| *marker8* | Paraprevotella | 2. Initial OA | -0.000396 | 0.0107 | Ns |
| *marker9* | Ruminococcus 1 | 1. Healthy | 0.0191 | 0.0297 | Ns |

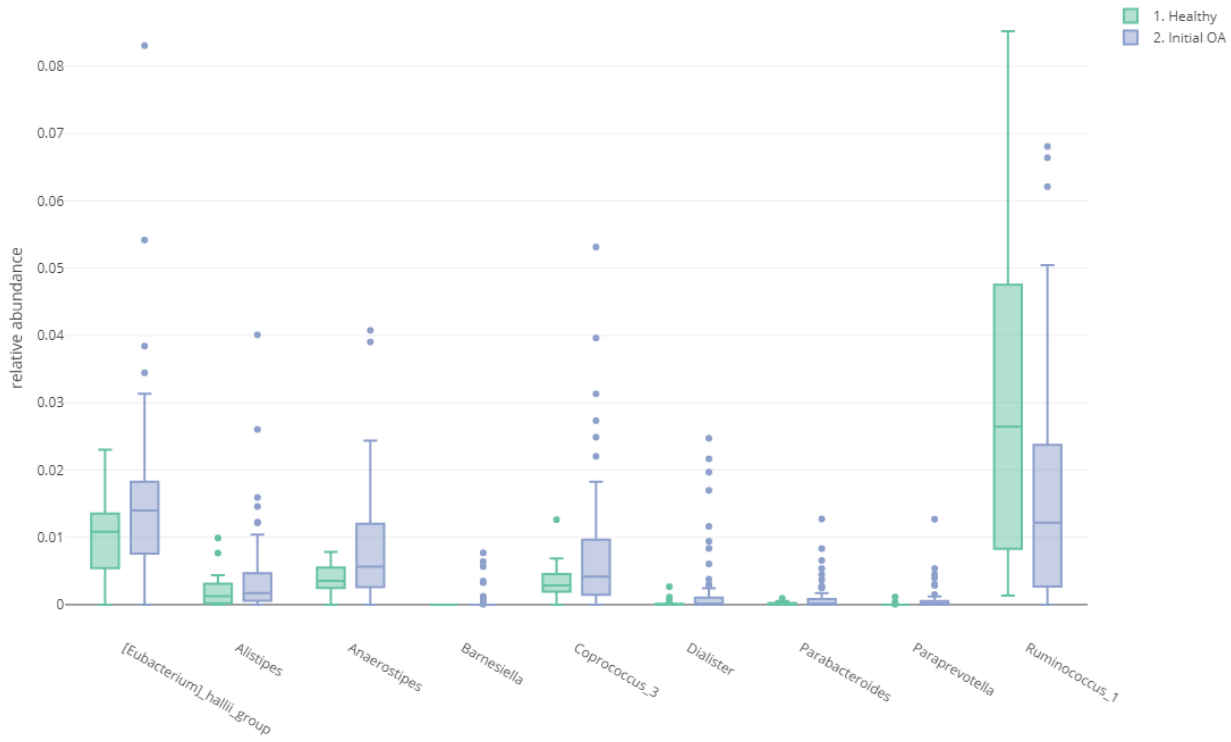



*Figure 10 Box plot with relative abundance for the 9 marker genus detected after Multivariate Welch t-test.*

### 6.3.5 Discussion

Microbiome studies about OA with samples from real individuals have a high degree of novelty. This is highlighted because most of the studies pointing to a possible cause-effect relationship between the microbiome and OA are conducted in mice, either germ-free (Ulici *et al.*, 2018) or mice fed with a high-fat diet (Rios *et al.*, 2019)(Collins *et al.*, 2015)(Schott *et al.*, 2018). Moreover, the studies correlating variation in the microbiome with OA using human samples include patients with advanced stages of OA. (Boer *et al.*, 2019) (Coulson *et al.*, 2013). Some other papers do not analyze the microbiome through 16S rRNA gene sequencing but establish possible correlations through LPS detected in blood as a metabolite generated by the microbiome and the WOMAC pain index. (Huang *et al.*, 2016).

The microbiome variables obtained were analyzed to identify differences between the Healthy and Initial OA groups. The microbiome diversity measured by the Shannon alpha diversity index shows no significant difference between the Healthy and Initial OA groups. These results are aligned with other studies that have looked at the alpha diversity obtained from OA patients. (Boer *et al.*, 2019).

Regarding beta diversity measures, none of the 4 indices studied (Weighted and Unweighted Unifrac, Jaccard, and Bray Curtis) differentiate the composition between the Healthy and Initial OA groups. This indicates therefore that the microbiome composition as a whole is not sufficiently distinct to differentiate between the Healthy and Initial OA groups. The same hypothesis is confirmed by applying the Random Forest classification model, yielding an 0,51 AUC for both groups. Such results of no significant difference in beta-diversity values are like those obtained by one of the most exhaustive studies performed to date in the Rotterdam cohort. In this study, 1444 participants were enrolled in the Rotterdam study-III with hip and/or knee osteoarthritis (Boer *et al.*, 2019). At first, no significant differences in beta diversity indices were detected, however, after choosing a subset of 256 taxa, they found an association between increased WOMAC score and the relative abundance of ASVs corresponding to the genus *Streptococcus*, widely reported as proinflammatory.

The study leads in the direction of not only looking for differences at the community level, but also for taxa that may be valid predictors of disease development. Following this approach, the 240 genera detected in the OActive volunteers' samples were compared using the Welch multivariate test. This model returned a series of 9 markers with predictive potential in the development of OA, 8 of them being enriched in the Initial OA group. None of these 9 markers corresponds to the *Streptococcus* genus detected in the Rotterdam cohort. (Boer *et al.*, 2019). However, after adjusting the p-values obtained by the false Discovery rate, significant differences were found.

Looking at the relative abundance distribution for the 9 genera detected by the model, a high proportion of outliers can be identified, mainly in the Initial OA group. Due to the high inter-individual sample variability, these variations in the abundances of the detected genera lead to a decrease in the discovery of new potential biomarkers, it is expected that, by increasing the N with more volunteers, these differences should decrease, more genera will be differentially detected as potential biomarkers and the adjusted p-values after FDR will continue to be significant.

Nevertheless, based on data analyzed so far, we are unable to assert that the microbiome, or taxa belonging to the microbiome, in our database alone can reliably predict the development of OA disease.

However, once the data generated will be introduced together with the other set of variables obtained during the project into the combined database, perhaps some of the microbiome taxa variables in combination with the other data sets can help the predictive algorithm to make a better decision in predicting the disease development.

Further research is needed to confirm the possible relationship between new biomarker candidates and the onset of OA.

## 7 References

Abramoff, B. and Caldera, F. E. (2020) 'Osteoarthritis: Pathology, Diagnosis, and Treatment Options', *Medical Clinics of North America*. W.B. Saunders, pp. 293–311. doi: 10.1016/j.mcna.2019.10.007.

Alekseyenko, A. V. (2016) 'Multivariate Welch t-test on distances', *Bioinformatics*. Oxford University Press, 32(23), pp. 3552–3558. doi: 10.1093/bioinformatics/btw524.

Aleman, F. D. D. and Valenzano, D. R. (2019) 'Microbiome evolution during host aging', *PLOS Pathogens*. Edited by J. M. Leong. Public Library of Science, 15(7), p. e1007727. doi: 10.1371/journal.ppat.1007727.

Bai H, Wu S. miR-451: A Novel Biomarker and Potential Therapeutic Target for Cancer. Onco Targets Ther. 2019

Biver, E. *et al.* (2019) 'Gut microbiota and osteoarthritis management: An expert consensus of the European society for clinical and economic aspects of osteoporosis, osteoarthritis and musculoskeletal diseases (ESCEO)', *Ageing Research Reviews*. Elsevier Ireland Ltd, p. 100946. doi: 10.1016/j.arr.2019.100946.

Boer, C. G. *et al.* (2019) 'Intestinal microbiome composition and its relation to joint pain and inflammation', *Nature Communications*. Nature Publishing Group, 10(1), pp. 1–9. doi: 10.1038/s41467-019-12873-4.

Bolyen, E. *et al.* (2019) 'Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2', *Nature Biotechnology*. Nature Publishing Group, pp. 852–857. doi: 10.1038/s41587-019-0209-9.

Breiman, L. (2001) 'Random forests', *Machine Learning*. Springer, 45(1), pp. 5–32. doi: 10.1023/A:1010933404324.

Callahan, B. J. *et al.* (2016) 'DADA2: High-resolution sample inference from Illumina amplicon data', *Nature Methods*. Nature Publishing Group, 13(7), pp. 581–583. doi: 10.1038/nmeth.3869.

Cahue S, Sharma L, Dunlop D, Ionescu M, Song J, Lobanok T, King L, Poole AR. The ratio of type II collagen breakdown to synthesis and its relationship with the progression of knee osteoarthritis. Osteoarthritis Cartilage. 2007 Jul;15(7):819-23

Cibere J, Zhang H, Garnero P, Poole AR, Lobanok T, Saxne T, Kraus VB, Way A, Thorne A, Wong H, Singer J, Kopec J, Guermazi A, Peterfy C, Nicolaou S, Munk PL, Esdaile JM. Association of biomarkers with pre-radiographically defined and radiographically defined knee osteoarthritis in a population-based study. Arthritis Rheum. 2009 May;60(5):1372-80

Collins, K. H. *et al.* (2015) 'Relationship between inflammation, the gut microbiota, and metabolic osteoarthritis development: Studies in a rat model', *Osteoarthritis and Cartilage*. W.B. Saunders Ltd, 23(11), pp. 1989–1998. doi: 10.1016/j.joca.2015.03.014.

Coulson, S. *et al.* (2013) 'Green-lipped mussel extract (Perna canaliculus) and glucosamine sulphate in patients with knee osteoarthritis: Therapeutic efficacy and effects on gastrointestinal microbiota profiles', *Inflammopharmacology*, 21(1), pp. 79–90. doi: 10.1007/s10787-012-0146-4.

Fujimoto, T. *et al.* (2013) 'Decreased abundance of Faecalibacterium prausnitzii in the gut microbiota of Crohn's disease', *Journal of Gastroenterology and Hepatology (Australia)*. Blackwell Publishing, 28(4), pp. 613–619. doi: 10.1111/jgh.12073.

Gupta, V. K., Paul, S. and Dutta, C. (2017) 'Geography, ethnicity or subsistence-specific variations in human microbiome composition and diversity', *Frontiers in Microbiology*. Frontiers Media S.A., p. 1162. doi: 10.3389/fmicb.2017.01162.

Huang, Z. Y. *et al.* (2016) 'Both systemic and local lipopolysaccharide (LPS) burden are associated with

knee OA severity and inflammation', *Osteoarthritis and Cartilage*. W.B. Saunders Ltd, 24(10), pp. 1769–1775. doi: 10.1016/j.joca.2016.05.008.

Hunter DJ, Nevitt M, Losina E, Kraus V. Biomarkers for osteoarthritis: current position and steps towards further validation. Best Pract Res Clin Rheumatol. 2014 Feb;28(1):61-71

Ishijima M, Watari T, Naito K, Kaneko H, Futami I, Yoshimura-Ishida K, Tomonaga A, Yamaguchi H, Yamamoto T, Nagaoka I, Kurosawa H, Poole RA, Kaneko K. Relationships between biomarkers of cartilage, bone, synovial metabolism and knee pain provide insights into the origins of pain in early knee osteoarthritis. Arthritis Res Ther. 2011 Feb 14;13(1):R22

Isolauri, E. (2017) 'Microbiota and Obesity', 88, pp. 95–105. doi: 10.1159/000455217.

Jackson, M. A. *et al.* (2018) 'Detection of stable community structures within gut microbiota co-occurrence networks from different human populations', *PeerJ*. PeerJ Inc., 2018(2), pp. e4303–e4303. doi: 10.7717/peerj.4303.

Kang L, Yang C, Song Y, Liu W, Wang K, Li S, Zhang Y. MicroRNA-23a-3p promotes the development of osteoarthritis by directly targeting SMAD3 in chondrocytes. Biochem Biophys Res Commun. 2016 Sep 9;478(1)

Karimi N, Ali Hosseinpour Feizi M, Safaralizadeh R, Hashemzadeh S, Baradaran B, Shokouhi B, Teimourian S. Roufayel R, Kadry S. Expression of miR-23a by apoptotic regulators in human cancer: A review. Cancer Biol Ther. 2017 May 4;18(5):269-276

Kraus VB, Blanco FJ, Englund M, Karsdal MA, Lohmander LS. Call for standardized definitions of osteoarthritis and risk stratification for clinical trials and clinical use. Osteoarthritis Cartilage. 2015 Aug;23(8)

Laslett, L.L., et al., *Measuring Disease Progression in Osteoarthritis.* Current Treatment Options in Rheumatology, 2016. **2**(2): p. 97-110.

Li, Z. *et al.* (2020) 'Gut Microbiota and Liver Fibrosis: One Potential Biomarker for Predicting Liver Fibrosis', *BioMed Research International*. Hindawi Limited, 2020. doi: 10.1155/2020/3905130.

Loeser RF, Goldring SR, Scanzello CR, Goldring MB. Osteoarthritis: a disease of the joint as an organ. Arthritis Rheum. 2012 Jun;64(6):1697-707

Neogi, T., *The epidemiology and impact of pain in osteoarthritis.* Osteoarthritis Cartilage, 2013. **21**(9): p. 1145-53.

Ozato, N. *et al.* (2019) 'Blautia genus associated with visceral fat accumulation in adults 20–76 years of age', *npj Biofilms and Microbiomes*. Nature Publishing Group, 5(1), pp. 1–9. doi: 10.1038/s41522-019-0101-x.

*R: The R Project for Statistical Computing* (no date). Available at: https://www.r-project.org/ (Accessed: 25 March 2021).

Ren, Z. *et al.* (2019) 'Gut microbiome analysis as a tool towards targeted non-invasive biomarkers for early hepatocellular carcinoma', *Gut*. BMJ Publishing Group, 68(6), pp. 1014–1023. doi: 10.1136/gutjnl-2017-315084.

Rios, J. L. *et al.* (2019) 'Protective effect of prebiotic and exercise intervention on knee health in a rat model of diet-induced obesity', *Scientific Reports*. Nature Publishing Group, 9(1). doi: 10.1038/s41598-019-40601-x.

Ro DH, Lee J, Lee J, Park JY, Han HS, Lee MC. Effects of Knee Osteoarthritis on Hip and Ankle Gait Mechanics. Adv Orthop. 2019 Mar 24;2019:9757369

Robinson, W. H. *et al.* (2016) 'Low-grade inflammation as a key mediator of the pathogenesis of

osteoarthritis', *Nature Reviews Rheumatology*. Nature Publishing Group, pp. 580–592. doi: 10.1038/nrrheum.2016.136.

Rooks, M. G. and Garrett, W. S. (2016) 'Gut microbiota, metabolites and host immunity.', *Nature reviews. Immunology*. Nature Publishing Group, 16(6), pp. 341–52. doi: 10.1038/nri.2016.42.

Scanzello, C. R. (2017) 'Role of low-grade inflammation in osteoarthritis', *Current Opinion in Rheumatology*. Lippincott Williams and Wilkins, pp. 79–85. doi: 10.1097/BOR.0000000000000353.

Schott, E. M. *et al.* (2018) 'Targeting the gut microbiome to treat the osteoarthritis of obesity', *JCI insight*. NLM (Medline), 3(8). doi: 10.1172/jci.insight.95997.

Serban, D. E. (2015) 'Microbiota in inflammatory bowel disease pathogenesis and therapy: Is it all about diet?', *Nutrition in Clinical Practice*. SAGE Publications Inc., pp. 760–779. doi: 10.1177/0884533615606898.

Sprockett, D. *et al.* (2019) 'Treatment-specific composition of the gut microbiota is associated with disease remission in a pediatric crohn's disease cohort', *Inflammatory Bowel Diseases*, 25(12), pp. 1927–1938. doi: 10.1093/ibd/izz130.

Thaiss, C. A. *et al.* (2016) 'The microbiome and innate immunity', *Nature*. Nature Publishing Group, pp. 65–74. doi: 10.1038/nature18847.

Tilg, H. *et al.* (2018) 'The Intestinal Microbiota in Colorectal Cancer', *Cancer Cell*. Cell Press, 0(0). doi: 10.1016/j.ccell.2018.03.004.

Ulici, V. *et al.* (2018) 'Osteoarthritis induced by destabilization of the medial meniscus is reduced in germ-free mice', *Osteoarthritis and Cartilage*. W.B. Saunders Ltd, 26(8), pp. 1098–1109. doi: 10.1016/j.joca.2018.05.016.

Wang, X., Xu, X. and Xia, Y. (2016) 'Further analysis reveals new gut microbiome markers of type 2 diabetes mellitus.', *Antonie van Leeuwenhoek*. Springer International Publishing. doi: 10.1007/s10482-016-0805-3.

Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, Bierma-Zeinstra S, Brandt KD, Croft P, Doherty M, Dougados M, Hochberg M, Hunter DJ, Kwoh K, Lohmander LS, Tugwell P. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage. 2008 Feb;16(2):137-62

## 8 Acknowledgment

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 777159.

**Disclaimer**