UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA VEGETAL

Macrophages in M1/M2 states and exposed to amyloid-beta: a RNA-seq analysis in human cells

Inês de Almeida Lopes

Mestrado em Bioinformática e Biologia Computacional
Especialização em Bioinformática

Dissertação orientada por:
Doutor João Pedro de Magalhães e Professor Francisco Dionísio

2017

Resumo

A Doença de Alzheimer (DA) é uma doença associada à idade e é uma das doenças neurodegenerativas com maior prevalência mundial. Esta doença causa a destruição progressiva de funções cognitivas e memória, levando à eventual morte do indivíduo. Um dos principais achados patológicos, para esta doença, são as placas extracelulares constituídas por beta amiloide (Aβ) que são encontradas frequentemente no cérbero de pacientes com DA. Enquanto que os mecanismos envolvidos na progressão desta doença permanecem desconhecidos, algumas hipóteses foram propostas envolvendo estas placas amiloides. Uma dessas hipótese é conhecida como a hipótese cascata de amiloide, e esta propõe que, através de uma cascata de eventos, a acumulação das placas de Aβ eventualmente resulta em disfunção neuronal e eventual morte celular. Microglia, os macrófagos residentes no cérbero, foram encontrados perto destas placas e já foi provado que eles têm capacidade de fagocitar estas placas.

Neste trabalho foi feita uma análise de RNA-seq em dados humanos, com o objectivo de comparar os perfis genéticos e de miRNA de um grupo de controle, de macrófagos polarizados para os fenótipos M1 e M2a e de macrófagos estimulados com Aβ. Adicionalmente, também se pretendia descobrir novas vias ou genes expressos em macrófagos presentes num ambiente com Aβ. Para tal, uma análise de expressão diferencial foi feita, tanto para os dados relativos ao mRNA como para os dados relativos ao miRNA. Uma análise do enriquecimento funcional foi feita para ajudar na compreensão dos genes diferencialmente expressos no mRNA. Por fim, uma série de testes de correlação foram feitos de modo a perceber o quão influente é a expressão de miRNA sobre o seu mRNA alvo.

Através da utilização de marcadores de expressão de mRNA e miRNA, foi possível verificar a polarização dos macrófagos para M1 e M2a. Além disso, foram descobertos alguns miRNA com alta expressão nos estados de polarização, que poderiam ajudar no estudo destes estados de polarização. Relativamente ao teste de correlação, os resultados estavam enviesados para correlações positivas devido a um miRNA em particular. Uma hipótese foi criada para tentar explicar esse resultado, mas nenhuma conclusão sólida foi alcançada neste projecto.

Palavras Chave: Envelhecimento; Doença de Alzheimer; beta amiloide; Polarização de macrófagos; RNA-seq

Abstract

Alzheimer's Disease (AD) is an age-related disease and it is one of the most prevalent neurodegenerative diseases in the world. This disease causes a progressive destruction of cognitive functions and memory, leading to the eventual death of the individual. One of the main pathological findings, for this disease, are the extracellular plaques of amyloid-beta (Aβ) often found in the brain of AD patients. Whilst the mechanisms involved in the progression of AD remain unknown, some hypotheses have been proposed surrounding the amyloid plaques. One such hypothesis is the Amyloid cascade hypothesis, and it purposes that, through a cascade of events, the accumulation of Aβ plaques leads to neuronal dysfunction and eventual cell death. Microglias, brain macrophages, have been found around such plaques, and it has been proven that they can phagocyte said plaques.

In this present work, an analysis of human RNA-seq data was done, aiming to compare the genetic and miRNA profiles of a control group, macrophages polarized towards the M1 and M2a phenotype, and of macrophages stimulated with Aβ. Additionally, another aim was to discover of novel pathways or genes expressed in macrophages in an Aβ environment. To do so, a differential expression analysis was done for both the mRNA and miRNA data. Functional enrichment analysis was done to sort the differential expressed genes for the mRNA. Lastly, a series of correlation tests were done in order to access the amount of influence done by the expression of miRNA to their mRNA targets.

It was possible to attest for the polarization of the macrophages into M1 and M2a, through the usage of mRNA and miRNA markers. Furthermore, some over-expressed miRNAs were discovered for the polarized states that could provide insight in to the study of this polarization states. For the correlation test, the results were skewed towards a positive correlation due to a single miRNA. While a hypothesis was constructed as to why it may have happened, no solid conclusion was achieved in this project thus far.

Keywords: Ageing; Alzheimer's disease; amyloid-beta; macrophage polarization; RNA-seq

Resumo Alargado

O envelhecimento e as doenças associadas ao envelhecimento têm sido uma grande fonte de estudos, devido aos seus papeis nas taxas de mortalidade na população humana. Estas doenças associadas ao envelhecimento podem ser cancro, doenças cardiovasculares, doenças neurodegenerativas, entre outras.

Neste trabalho, o foco principal foi o estudo da Doença de Alzheimer (DA), uma das principais doenças neurodegenerativas. Esta doença é caracterizada pela progressiva destruição de funções cognitivas e da memória, eventualmente resultando na morte do indivíduo. Actualmente, existem dois achados patológicos principais encontrados com frequência no cérebro de pacientes com DA. Um deles trata-se das placas senis, constituídas por agregados extracelulares de péptidos de amiloide beta (Aβ), e foi sobre estas placas que este estudo se realizou. Já foi demonstrado, em outros estudos, que existe uma acumulação de macrófagos residentes do cérbero, microglia, perto destas placas, despontando uma reacção inflamatória que inclui a tentativa de fagocitar estas placas e os seus resíduos, por parte destas células do sistema imunitários. No entanto, existem resultados conflituosos sobre o grau de digestão que as células microglia conseguem fazer deste péptido. Alguns estudos apontam para uma digestão apenas parcial de Aβ, enquanto que outros mostram uma digestão total. Em relação a macrófagos resultantes de células em circulação, estes parecem ter a capacidade de degradar Aβ fibrilar.

Macrófagos são importantes constituintes do sistema imunitário, responsáveis por fagocitar agentes patogénicos, produção de citocinas e quimiocinas, apresentação de antigénio, entre muitas outras funções. Estas células podem-se adaptar ao ambiente em que estão através da sua polarização. Sucintamente, a polarização pode ser realizada em dois grupos distintos, M1 ou M2. Os macrófagos com polarização em M1 são pro-inflamatórios, produzindo citocinas, quimiocinas e moléculas efectoras, recrutando outras células, entre outras funções. Células com polarização em M2 são anti-inflamatórias, tendo um grande papel na resolução da inflamação, na reparação de tecido, na remoção de parasitas, entre outras.

Para este estudo, pretendia-se descobrir qual o efeito que o péptido Aβ tem nos macrófagos que o fagocitam, e como esse efeito se compara aos estados de polarização presentes nos macrófagos (M1/M2). Para tal, foi feita uma análise de RNA-seq sobre amostras já previamente preparadas e sequenciadas. Estas amostras foram preparadas através da remoção de monócitos, células percursoras de macrófagos, de indivíduos saudáveis. De seguida, estas foram divididas de modo a obter-se quatro grupos. No grupo de controlo estão presentes macrófagos que não receberam qualquer tipo de estímulo, no grupo M1 as células foram estimuladas pela combinação de LPS e INF-γ para obterem polarização M1, no grupo M2 as células foram estimuladas com IL-4 de modo a obterem a polarização M2a e, finalmente, o grupo Ab onde os macrófagos foram estimulados pela presença de Aβ. De seguida, foi retirado o mRNA e miRNA destas amostras, efectuando-se um passo de sequenciamento. O presente trabalho começou com um passo de alinhamento, onde se alinhou as *reads* obtidas a um genoma de referência. Depois foi realizado um passo de contagem, onde cada alinhamento é contado, construindo-se uma tabela que inclui o nome do gene e o número de *reads* que alinharam para este gene. Em todos estes passos, foi feita uma análise de qualidade, com o fim de se detectar algum erro no mapeamento e contagem. Finalmente, a análise estatística para a expressão diferencial (ED) foi feita. Adicionalmente, como se pretendia entender o efeito que os miRNAs têm nos respectivos genes alvos, realizaram-se testes de correlação entre os miRNAs detectados como expressos

diferencialmente e os seus genes alvos. No caso particular do mRNA, foi feita uma análise de enriquecimento funcional, de modo a obter informações biológicas relevantes, a partir das listas de genes resultante de análise de ED.

Em relação aos resultados obtidos para o mRNA, através da análise de qualidade feita nos passos de mapeamento e contagem, foi possível verificar que as bibliotecas preparadas antes de sequenciamento estavam contaminadas, notando-se a presença de diversas *reads* mapeadas para regiões intrónicas e não-genómicas. Assim, muitas *reads* acabaram por ser descartadas antes da análise de ED. Aquando da análise de ED, foram obtidas listas de genes considerados diferencialmente expressos quando se compararam todas as condições entre si. Enquanto que, as comparações com as amostras de controle resultaram em muitos genes com expressão diferencial para os grupos M1 e M2, o mesmo não pode ser dito para a comparação com o grupo Ab. Os perfis genéticos entre o grupo de controle e o grupo Ab foram demasiado semelhantes, proporcionando apenas um gene com expressão elevada nos macrófagos activados com Aβ, *SOD2*. Conseguiu-se, no entanto, comprovar a polarização bem-sucedida dos macrófagos a M1 ou M2 através da identificação de genes presentes nestes estados de polarização na literatura. Adicionalmente, os genes *ACOD1*, *ANKRD22*, *SNX10* e *RP11-44K6.2* para a polarização em M1 e o gene *CTNNAL1* para a polarização em M2 foram detectados com altos níveis de expressão.

A análise de enriquecimento funcional mostrou, para os macrófagos polarizados para M1, que a expressão de certos receptores e de gene associados ao processamento e apresentação de antigénio foi aumentada. Adicionalmente, houve o aumento de genes que participam na via de sinalização NOD, responsável pelo reconhecimento de estruturas bacterianas e produção de citocinas. Quando foi considerado o grupo M2, esta análise demonstrou o aumento de expressão de genes associados a receptores e imunidade adaptativa. Adicionalmente, foram detectados diversos genes com expressão alta para células M2 associados a *DNA housekeeping* e controlo de expressão de mRNA, possivelmente devido as mudanças epigenéticas causadas pelo processo de polarização, e outros associados a interacção viral, onde esta polarização parece ter um papel mais importante que as células polarizadas para M1. Para os macrófagos estimulados com Aβ, estes aparentam ter uma maior quantidade de genes associados ao metabolismo do colesterol, quando comparados com os macrófagos com polarização M1, e genes associados a glicoproteínas e péptido sinais, quando comparado com os macrófagos com polarização M2.

Em relação aos resultados para o miRNA, o passo de contagem também mostrou alguns problemas. Neste caso, os problemas causados foram devido à natureza do miRNA. Estes pequenos RNA têm como função o controlo de expressão de certos genes e, para isso, a sequência de miRNA é semelhante a esse alvo. Assim, no passo de mapeamento, houve o mapeamento deste miRNA a diversos genes ao mesmo tempo, e o programa usado na contagem não consegue lidar com esses alinhamentos múltiplos. Em termos da análise de ED, foi obtida uma lista de miRNAs diferencialmente expressos tanto para o grupo M1 como para o grupo M2. Para o grupo Ab, não foram obtidos nenhuns miRNAs diferencialmente expressos, quando comparado com o grupo Controlo. Verificou-se a presença de vários miRNA, já anteriormente detectados em outros trabalhos para estes tipos de polarização, M1 ou M2. Adicionalmente, outros miRNAs foram detectados para estas polarizações, fazendo deles objectos de estudo interessantes no estudo destes estados de polarização de macrófagos.

Um *script* foi criado, com o objectivo de realizar testes de correlação entra a expressão de miRNA diferencialmente expressos e seus respectivos mRNA alvos. Os resultados mostraram o enviesamento dos resultados do teste para correlações positivas, sendo a origem deste enviesamento

um miRNA em particular. Uma vez que os valores de expressão deste miRNA eram máximos para os grupos Ab e Controle e mínimos para os grupos M1 e M2, foi criada a hipótese que a difusão fosse resultado da polarização dos macrófagos e não do efeito do miRNA em específico. Esta hipótese foi testada olhando para as vias em que os genes alvos com correlação positiva participam. No entanto, não foi possível provar completamente esta teoria e nenhuma conclusão sólida, no que respeita a este enviesamento, foi alcançada neste trabalho.

Em conclusão, o presente projecto permitiu a identificação do gene SOD2, que poderá ter um papel central na resposta imunitária a Aβ. Adicionalmente, não só foi possível verificar a polarização bem-sucedida das células em M1 e M2, como também a identificação de genes que não são usualmente associados a estes estados de polarização nestas condições. Foram também identificados alguns miRNAs novos nestas condições de polarização, que poderão ser interessantes alvos de estudo.

Acknowledgements

First, I would like to thank my supervisor at the Integrative Genomics of Ageing Group, João Pedro de Magalhães, for providing me with the opportunity to work on his group and for accepting me in this project. His guidance was essential in this project.

Secondly, I would like to thank my supervisor at the Faculty of Sciences of the University of Lisbon, Professor Francisco Dionísio for accepting to be my supervisor and for his help on several instances.

In addition, I would like to thank Shah, for his help and immense support while I worked on this project.

I would also like to thank the rest of Pedro's lab, especially Gianni, for their help, support and advice.

To all my friends, I want to thank for helping and supporting me.

Finally, I would like to thank my family, especially my parents for the great amount of unconditional support, whether be it monetary or emotional, and their push towards me taking on new and scary challenges.

Contents

List of Figures

List of Tables

List of abbreviations

Alzheimer’s disease == AD

Amyloid beta == Aβ

Amyloid precursor protein == APP

Central nervous system == CNS

Differential expression == DE

False Discovery Rate == FDR

Familial Alzheimer’s Disease == FAD

Interferon gamma == IFN-γ

Lipopolysaccharide == LPS

Major histocompatibility complex == MHC

Messenger RNA == mRNA

MicroRNA == miRNA

Reactive oxygen intermediate == ROI

Reactive oxygen species == ROS

Replicative senescence == RS

Small nuclear RNA == snRNA

Sporadic Alzheimer’s Disease == SAD

T helper cells type 1 == Th1

T helper cells type 2 == Th2

Chapter 1: Introduction

1.1 Immune systems

The immune system is a crucial part of the human body. It allows us to be protected against foreign bodies known as pathogens that threaten our normal bodily functions. As a whole, the immune system can be divided into two sub systems: the innate immune system and the adaptive immune system.

Innate responses are non-specific but are the faster of the two systems. It includes body barriers (like mucus layers over epithelia), neutrophils, monocytes, macrophages, cytokines (regulate function of other cells), chemokines (help recruit other cells of the immune system) among others. On the other hand, adaptive responses are more precise in action and include T lymphocytes and B lymphocytes. The antigen receptors on these cells are able to identify foreign bodies activating the activity of both B and T cells. The adaptive response also has the capability of retaining the "memory" of a previous attack, making the response in later infections faster.

While it may sound that both these systems are opposites, their cooperation is essential for the protection of the human body. The innate immune system acts the first line of defence (due to its fast action time) and activates the cell of the adaptive immune system, allowing for a more specific and effective response [1,2].

1.1.1 Macrophages

Macrophages were first identified by Ilya Metchnikoff, who observed the phagocytic activity of these cells and highlighted its importance in immunity, a fact that would earn him a Nobel Prize in 1908 [3]. These cells are important components of our immune system, aiding in the defence against foreign microorganisms. However, it's often forgotten that these cells also have a role in homeostasis, wound repair and embryonic development [4].

In the presence of an inflammation, macrophages have multiple responses: production of cytokines and chemokines, phagocytic clearance of the pathogens [5], and the presentation of antigens in order to activate the T-cells, following said phagocytic clearance [4], to name a few.

Macrophages that reside in specific tissues have major differences, especially when it comes to their function and phenotype. For example, the microglial cells that reside in the brain are very different when compared to Kupffer cells that reside in the liver [6]. In regards to tissue macrophages, these cells can have two origins: the majority are prenatal and are independent from hematopoietic input, whilst the minority of them result from the infiltration of monocytes into tissue. This is usually because of an infection, where these monocytes are often recruited by tissue macrophages, and they subsequently differentiate into macrophages. This however is not always the case, seeing as different tissue handle their quantity of embryonic and monocyte-derived differentially. Examples of this are the intestines and the skin, which have a great deal of non-embryonic macrophages; and the microglia, which have a large quantity of embryonic cells. In some cases, monocyte-derived macrophages are able to replace the previous tissue macrophages [7,8].

1.1.2 Macrophage polarization

When certain environmental cues are present, macrophages can differentiate and undergo phenotypic change to better respond to said cues. This process is known as macrophage polarization, and the most common designation for these cells is M1, classically activated, and M2, alternatively activated. This is mirroring the polarization of T helper cells into type 1 and type 2 [9]. However, this definition is not fundamentally true, considering that there are macrophages with similar gene expression to M2 macrophages but not entirely alike. These macrophages are known as M2-like macrophages. M2a are the ones corresponding to the original definition and original stimuli used, M2b, M2c etc. are the M2-like macrophages. M2-like macrophages and their polarizing cues are different than that used to obtain M2 macrophages [10]. Yet, the usage of the M1/M2 paradigm should be done with some caution. With the existence of M2-like phenotypes, there is a possibility of a spectrum of polarizations, with M1 and M2 being the terminals of said spectrum [11], and so the categorization of only M1 or M2 ,or known M2-like phenotypes, might not be entirely correct. Additionally, it should be noted that this polarization, or differentiation as many authors call it, is not permanent and can be reversed. An example of this is the possible reprogramming of M1 macrophages to M2a macrophages [12].

Table 1.1: ***Macrophage polarization expression patterns.*** *Genes often found expressed in M1 and M2 macrophages. FR – folate receptor; GR – galactose receptor; MR – mannose receptor; SR – scavenging receptor; RNI – reactive nitrogen intermediate; ROI – Reactive oxygen intermediate.*[9,10,13–16]

	Cytokines	**Chemokines**	**Receptors**	**Others**
M1	TNF-α, IL-1β, IL-6, IL-12α, IL-12β, IL-15, IL-18, IL-23α, type I IFN	CXCL1, CXCL2, CXCL3, CXCL5, CXCL8, CXCL9, CXCL10, CXCL11, CXCL16, CCL2, CCL3, CCL4, CCL5, CCL8, CCL11, CCL15, CCL20, CX3CL1	CD40, CD80, CD86, IL1R1, TLR2, TLR4, CCR7	MHC I/II, STAT1, IRF1, IRF5, IDO1, KYNU, GBP1, ROI, RNI
M2 (M2a)	TGFβ, IL-10, IL-1Ra	CCL1, CCL2, CCL4, CCL13, CCL17, CCL18, CCL22, CCL24	CD23, CD163, CD206, CD200R1, CD301, IL1R2, IL17RB, STAB1, MARCO, ADORA3, TGFBR2, LYVE1, SR, GR, FR, MR	MHC II, Ym1, Fizz1, ARG1, IRF4, SOCS1, GATA3, FN1, TGFB1, MMP1, MMP12, TG, F13A1, TGM2, ALOX15, MSF

Table 1.1 provides an example of some genes expressed in M1 and M2 cells. The aforementioned genes are heterogeneously expressed both within and between M1 and M2 populations, thus the term "marker" was not used when referencing them. The existence of these

problematic “markers” has been noticed before [13], and it is mainly due to the fact that macrophage polarization is a spectrum depending on stimuli. This may cause M1 cells to have different gene expression due to the fact that different stimuli were used, or that a M1 cell activated with a specific stimulus may share some “markers” with M2 cells while the rest of the family does not. Additionally, some of these genes are murine-only (like *Ym1* and *Fizz1*) with no human orthologs [9] or may be differentially expressed when comparing two different species [13].

Upon the start of an infection caused by intracellular pathogens, the initial line of defence includes the M1 macrophages. This type of polarization is pro-inflammatory, meaning that it will work to heighten the inflammatory response. The presence of IFN-γ by itself, or in conjugation with certain products of this pathogens (LPS) or pro-inflammatory cytokines (*TNF*), can cause this type of activation. This results in cells that produce large amounts of pro-inflammatory cytokines (IL-1β, IL-6, IL-12, IL-18, IL-23 and *TNF* for example), and chemokines (*CXCL1*, *CXCL2*, *CXCL3*, *CXCL5*, *CXCL9*, *CXCL10*, to name a few) that aid in the recruitment of other immune cells. This includes the recruitment and polarization of Th1 cells, and effecter molecules such as reactive nitrogen and oxygen intermediates, that have strong microbicidal and tumoricidal activity. Additionally, these cells also have a much bigger expression of MHC II, which shows the important role of antigen presentation that falls upon these macrophages as well [9–11,14,17].

After the infection is under control, macrophages will start to attain an M2 type of polarization. This anti-inflammatory polarization has a role in the resolution of inflammation (immune regulation), tissue remodelling and repair, angiogenesis, tumour progression, in parasite clearance and apoptotic cell internalization. The original M2 polarization, now classified as M2a, was discovered as a response to the cytokine IL-4. This polarization type has up-regulation of scavenging (*CD163*), mannose (*CD206*) and galactose receptors, and the arginine metabolism produces ornithine and polyamines. These cells have a different chemokine profile, producing *CCL17*, *CCL22*, and *CCL24*, which are involved in the recruitment of Th2 cells. Phagocytic activity is increased when compared with M1 cells [9–11,14,17–19].

1.2 Ageing

The one process of human biology, that is the most observed and yet not completely understood is ageing. We all experience it, being it by observing others or by experiencing it ourselves. This process is characterized by a functional decline, reduced homeostasis and eventual death that happens with age [20,21]. There are many theories that attempt to pinpoint the reason why this phenomenon happens. One theory is genomic instability, where external factors challenge DNA stability, which in return causes DNA damage. Another theory is telomere shortening, where progressive loss of telomeres after each cellular replication prevents further replication. Finally, a theory that includes cellular senescence, where cells enter permanent cell cycle arrest, also exists. [20,22].

It is hypothesised that one of the major contributors to ageing and age-related diseases is cell senescence. Senescent cells tend to accumulate in aged-tissue [20,23,24]. As previously mentioned, this phenomenon is characterized by stable arrest of the cell cycle and this cells are phenotypically different than normal cells [20,23–25]. Replicative senescence (RS) was first documented by Leonard Hayflick and Paul Moorhead, and they observed that human cells in vitro could only divide a finite amount of times, resulting in the Hayflick limit [26]. Telomere shortening is considered by many to be the main cause for RS. The ends of the chromosome are called telomeres and they protect the cells'

DNA during replication. However, due to the nature of DNA replication and the fact that most cells do not possess telomerase, the enzyme that elongates telomeres, telomeres erode with each cell division. When a critical minimal length is achieved, it provokes a DNA damage response, which will eventually lead to senescence. However, there are other stimuli than can induce senescence in cells. DNA damage due to ROS for example, could induce senescence before even the critical shortening of the telomeres [20,23,24,27].

Whilst senescence is thought to be one of the major players in ageing, it should not be forgotten that its main role is a different one. Not a malignant role that would cause the death of an individual, but rather one that would benefit the individual. Senescence exists to stop the propagation of cells that are damaged and to signal the immune system to eliminate said cells. If said cells were allowed to live, there would be an accumulation of mutations, which would eventually lead to the formation of cancerous cells. The signalling by senescence cells towards the immune system is accomplished due to the fact that senescence cells are capable of producing pro-inflammatory molecules. In aged individuals, this system may be faulty, thus resulting in the accumulation of senescence cells [20,25,28].

Aging can affect many human systems, including the immune system. Many changes occur in the immune system as human beings age. In general, the adaptive response seems to be more affected, but both the adaptive and innate responses suffer the effects of ageing [29,30]. Macrophages in older individuals, for example, produce lower amounts of cytokines and have an altered expression of MHC II, which in turn may contribute for a poorer T cell response. Additionally, in mice, the expression of MHC molecules is lower, the phagocytoses process is impaired and the production of ROS is also lower [29,31]. Other changes can include, the smaller number of T and B cells, especially those that are naïve, the increase of myeloid and NK cells, a decrease in antibody production, among others [29,32].

One consequence of ageing in immune systems is known as inflamm-aging. This term refers to the low levels of chronic inflammation that affects aged individuals. This phenomenon is characterized by the increased levels of circulating cytokines, chemokines and pro-inflammatory markers, and it can be associated with many age-related diseases including diabetes, osteoporosis, atherosclerosis, cardiovascular and neurodegenerative diseases. The cytokines IL-6, IL-1beta and TNF-α seem to have a major role in inflamm-aging [29,31,32]. The reason why there is such an accumulation of cytokines and chemokines is not fully understood. However, one hypothesis could be associated with the accumulation of senescent cells that happens with age. More importantly, the acquisition of a senescence-associated secretory phenotype (SASP), which would account for an increased secretion of pro-inflammatory cytokines and chemokines [31].

1.2.1 Ageing related diseases and Alzheimer's disease

Ageing itself is not the major cause of death, but rather a risk factor for age-related diseases, the true cause of death in human populations. Diseases like cancer, cardiovascular diseases, metabolic syndrome, atherosclerosis and neurodegenerative diseases are taking a more prominent role as the leading cause of death in industrialized countries [23,24]. As a risk factor, it is possible that targeting ageing can possibly lead to a lower mortality rate by these aforementioned diseases.

One of the most important neurodegenerative diseases is Alzheimer's disease (AD). This disease is characterized by a progressive destruction of cognitive functions and memory. The two

classical, and main, pathological findings for AD are senile plaques (SP), that are extracellular aggregates composed of β-amyloid (Aβ) peptides, and neurofibrillary tangles (NFTs) in the cortex, that are intracellular aggregates composed of Tau protein [21,33,34].

Whilst the mechanisms involved in the progression of this disease remain unknown, several hypothesises have been proposed, one of which is the Amyloid cascade hypothesis. This hypothesis states that the accumulation of the previously mentioned Aβ plaques, through a cascade of events, lead to the neuronal dysfunction and cell death, making it the main cause behind AD. This Aβ plaques are composed of Aβ peptides and, in turn, these peptides are obtained through the enzymatic cleavage of Amyloid precursor protein (*APP*) [35–38]. These peptides can have different aggregation states, like monomers, dimers, oligomers, and fibrils. The most common formation of these peptides is soluble oligomers and fibrils. The majority of these aggregates are associated with neurodegeneration [18,37–39]. It should be noted that Aβ is produced normally and that, under normal circumstances, there is equilibrium between the production and elimination of this peptide. The disturbance of the clearing process leads to the formation of plaques that will also lead to AD [33,37].

Genetics can also play a role in AD. Familial AD (FAD) is a specific form of early-onset AD. This specific type of AD occurs in a minority of cases and is associated with the mutation of APP and the presenilins 1 and 2 (PSEN1 and PSEN2) genes. Both of these mutations lead to an increase in the production of Aβ. On the other hand, sporadic AD (SAD), which is another name for late-onset AD, is the most common form of this disease found. For SAD, the biggest genetic risk is the *APOE* gene, which is involved in cholesterol metabolism, immune signalling and synaptic plasticity. Specifically, the allele ε4 for this gene seems to be the bigger risk factor when comparing it with the other alleles (ε2 and ε3) [34,36,39]. Alternatively, hypotheses that include cellular senescence have also been made to explain this disease [18,24,34,39,40].

1.2.2 Macrophages in the context of AD

Microglia, as it was mentioned before, are the brain's resident macrophages and have a major role in inflammatory response associated with neurological diseases [41]. Besides the expected role in inflammation and immune responses, these cells also participate in the maintenance of neuronal tissue, especially when it comes to synapses [18,41].

Under normal circumstances, microglias are found in a resting state, which is often characterized by their small cell body, scanning the environment for possible changes. Upon the detection of changes in the central nervous system (CNS), these cells are able to take an *ameboid* shape, allowing them free movement through neural tissue. Upon arrival to the injury site, these cells are activated, which means that they are able to phagocytose pathogens, present antigens and produce pro-inflammatory molecules [33,42,43]. Additionally, these cells seem to be able to attain a polarization state, similar to the previously described M1/M2 paradigm, with an added deactivated state. A mixture of all this polarization states has been observed in AD [18,41].

An accumulation of microglia cells around Aβ plaques has been seen in AD. These cells are able to bind themselves to Aβ soluble oligomers or fibrils, through cell-surface and Toll-like receptors, starting an inflammatory reaction and leading to the production of pro-inflammatory cytokines and chemokines. After the binding of this receptors, microglial cells can phagocyte the Aβ and degrade it and the degradation of fibrils or soluble Aβ is possibly done by different mechanisms [18,33,39,41]. However, the effectiveness of said degradation, after internalization, has presented some mixed

results, with some groups finding microglia cells to be able to fully digest these molecules and others finding poor degradation of Aβ [41,44]. Under normal circumstances, after inflammation and pathogen clearing, there would be a cessation of inflammation and normality would be restored. That is not the case for AD, since there are several mechanisms in play that prevent that, like a positive-feedback loop between inflammation and the processing of the Aβ precursor (*APP*). As such, there is severe inflammation which, in turn, impairs the function of the microglial cells, diminishing the amount of Aβ that is cleared from the system [18,33].

Microglia and their role in AD tends to be considered a "double-edged sword". On one side, and as previously mentioned, these cells have an important role in neural and immune defence. On the other hand, and as it is seen in many neurodegenerative diseases, the excessive presence of these activated cells can cause or aggravate neuron damage [41–43].

In terms of non-tissue specific macrophages, these seem to be able degrade fibrillar Aβ [44].

1.3 RNA-seq

A fundamental component of cells is DNA, and its capacity to encode all the properties and functions of a single cell. With it, cells can access that information and adapt to their environment. This information, which is encoded in genes, can be transcribed into RNA molecules, to be later translated into proteins, via mRNA, or to be used to control the expression of other genes, via miRNA. As such, the set of RNAs expressed when certain conditions are present or at a certain time, can tell us a lot about the current status of a cell and provide insight in to diseases and the mechanisms in which they operate. This is called differential gene expression, and can be used to compare between different conditions (Control vs Condition or ConditionA vs ConditionB etc.) or between different tissues, based on what needs to be achieved [45].

Currently, the most popular method for studying gene expression is RNA-seq. This technique replaced microarrays, and the advantages of using RNA-seq over microarrays are: RNA-seq can still be performed even when there is no mapped genome, whilst microarray depends on that knowledge for probe design; it can be used to characterize exon junctions and to detect non-coding RNA; high levels of reproducibility, higher resolution and a much lower limit of detection [45–47].

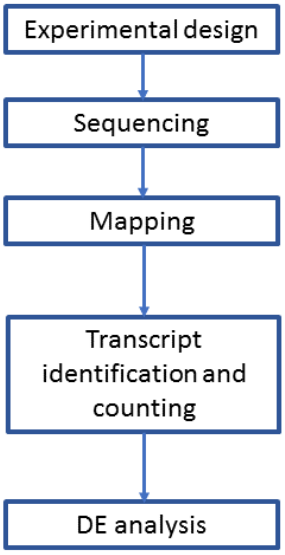

Figure 1.1: ***Example of an RNA-seq experiment****. While many of the steps remain similar to the ones in this figure, it is possible to adapt these steps, in order to answer the biological question required.*

Typically, this technique (Figure 1) starts with the extraction of the population of RNA to study, which is then converted into a cDNA library of fragments. Each fragment can have an adaptor attached to one of its' ends or to both, and they are later sequenced in a high-throughput machine, for example an Illumina machine. If the fragment only has an adaptor on one end, the sequencing is called single-end; on the other hand, if both ends have an adaptor, it is called a paired-end. Reads can be 30-400bp depending on the machine used [47,48].

Afterwards, the data obtained must be analysed. Normally, the pipeline (Figure 1.1) starts by mapping the reads, where the raw reads, obtained from the sequencing, are mapped against a reference genome for example. Following that there is usually a counting step, in which the reads are counted, based on the gene that they mapped to. Finally, the analysis of the Differential expression (DE) is made, where a conjugation of statistical methods is used, in order to determine which genes have a significant different expression between two compared conditions. However, this is but a general overview, depending on the objective of the analysis or the organism studied, several steps may be added or changed from it [49].

One problem present in the analysis of this data is the fact that there are many methods and tools that can be applicable for this kind of analysis, but no standard protocol on how to do so. This happens mainly because the determination of tools performance is hindered, due to the lack of gold-standard measures and the frequent update of these tools [45,46].

1.4 Motivation

As previously mentioned before, Alzheimer's disease is one of the most important neurodegenerative diseases, especially in older populations. As such, it is important to understand its players and mechanisms. A lot of components, in the relationship between macrophages and Aβ

remain unknown, and their discovery could possibly answer a lot of questions relating to inflammation in the AD brain.

1.5 Objectives

Whilst there are many studies done involving macrophage polarization [12,50,51], the comparison between macrophage polarization and macrophages stimulated with Aβ does not have as much of a focus. As such, this project was developed with following objectives in mind:

Objective 1: Comparison of the mRNA and miRNA profiles between the control group and the polarized macrophages (M1/M2), or the macrophages exposed to Aβ;

Objective 2: Comparison of the mRNA and miRNA profiles between the polarized macrophages (M1 vs M2);

Objective 3: Comparison of the mRNA and miRNA profiles between the polarized macrophages (M1/M2) and the cells exposed to Aβ;

Objective 4: Identification of key molecular players in the innate immune response against Aβ.

1.6 Contributions

With the motivation and objectives mentioned before in mind, the following contributions are available in this work:

Contribution 1: Comparison of gene and miRNA expression between the control group and the polarized macrophages (M1/M2) or the macrophages exposed to Aβ;

Contribution 2: Comparison of gene and miRNA expression between the polarized macrophages (M1 vs M2);

Contribution 3: Comparison of gene and miRNA expression between the polarized macrophages (M1/M2) and the cells exposed to Aβ;

Contribution 4: Functional expression analysis of the differential expressed genes for all the previously mentioned comparisons;

Contribution 5: Evaluation of the influence of miRNA over their respective gene targets, miRNA-Target correlation analysis.

1.7 Overview

This work is divided into five chapters. The first chapter corresponds to the Introduction, were some key concepts are explained in order to better understand this project. The second chapter corresponds to the methodology, where a description of the data, databases and tools used can be found, in addition to a description of the methods used to achieve the desired objectives. Chapter three as the presentation of all results found in this work, while chapter four discusses said results based on

what has been previously done by other groups. Finally, the fifth chapter presents the final conclusion of this project and provides some future studies that can also be done.

Chapter 2: Methodology

2.1 Data and Databases

This section explains how the original data used for the sequencing was obtained and gives an introduction to other data or databases used in this work.

2.1.1 Experimental data

To achieve the objectives proposed, monocytes were extracted from nine age-matched healthy individuals. These cells were differentiated into macrophages and later on, were divided in four groups, based on the stimuli that they received. Cells stimulated with LPS and INF-γ achieved a M1 activation profile and are, from this moment on, referred to as the M1 group. Cells stimulated with IL-4 were able achieve the M2 activation profile, from now on were referred to as the M2 group. Cells exposed to Aβ fibrils and oligomers, were referred to in this work as the Ab group. Cells that received no stimuli were considered the control group.

After the establishment of these groups, mRNA and miRNA were extracted from the cells of each group in all individual and three pools of three individuals were created for each condition, resulting in three biological replicates for each condition (Table 2.1).

These samples were prepared by the PhD student Ana Viegas. This work concerns only the computational analysis of the data obtained from those samples.

Table 2.1: ***Pool creation representation.***

Pools	Patients	Condition
Pool 1	E + J + RC	Control
		Aβ
		IL-4
		LPS + INF-γ
Pool 2	I + H + GI	Control
		Aβ
		IL-4
		LPS + INF-γ
Pool 3	Is + C + R	Control
		Aβ
		IL-4
		LPS + INF-γ

2.1.2 Ensembl Genome

One important step in RNA-Seq is the mapping step, followed by a counting step. Two files are needed for this. The first is a genome file, which has the complete sequenced genome of an organism. The second is an annotation file, that properly identifies regions of the genome, be it a region associated with a gene, splicing region, among others. For this work, both the genome file and the annotation file were obtained from Ensembl.

Ensembl [52] started off as a joint project between the EMBL European Bioinformatics Institute and the Wellcome Trust Sanger Institute and it aims to provide a centralized resource of genome information. To that end, it provides its users with genome assemblies for several organisms, for example human and model organisms such as mouse and fruitfly; as well as their respective genome annotations. This annotation is automated using a pipeline of Perl scripts whose output is later stored in databases. Ensembl gets updated frequently adding new organisms and updating genes, in order to provide its users with high quality information. The release of Ensembl used in this work was 86. Ensembl is available at http://www.ensembl.org/index.html.

2.1.3 miRBase

For the analysis of the miRNA data, an annotation file containing only miRNA would be ideal since, due to the nature and function of miRNAs, that is, to modulate the expression of certain genes, by ways of degradation and the repression of translation of mRNA [53], their reads could map to other genes, providing an identification problem later on in the differential expression analysis.

With the growing research on miRNA, there was a need to compile their sequences into a shared platform. Currently managed by researchers at the University of Manchester, miRBase [54–56] not only provides the sequence for miRNAs of several species (that can be user submitted), but also an annotation file compatible with the Ensembl genome files. The version 21 of this database was used in this work. This platform is available at http://www.mirbase.org/.

2.1.4 miRTarBase

In order to better understand the effect that the miRNAs have on their respective target a correlation analysis was performed for this project. As such, there was a need for a database that has validated miRNA-Target interactions. This was done to investigate, for example, if the lowering of expression on an mRNA is due to a specific miRNA that targets it or due to something else entirely.

For this investigation, the miRTarBase [57] database was used. This database's data is collected by the manual surveying of articles that studied the functionality of miRNAs. This collected data does also have to be validated experimentally by a variety of methods. Two different classification methods for the type of evidence in each pair are used by miRTarBase. Strong evidence is classified as experimental methods that show, more strongly than others, an association between miRNA and the respective target that is being studied. "Strong evidence" includes methods like Reporter assay, Western blot and qPCR and "Less strong evidence" includes methods like Microarray, NGS, psILAC, HITS-CLIP among others. This classification was created and is used by the miRTarBase. The release of miRTarBase used in this work was 6.0. This database is available at http://mirtarbase.mbc.nctu.edu.tw/.

2.2 Tools

This section describes and explains some of the inner-workings of the tools used in this project. There is no mention of the tools used in the data sequencing part of the protocol, since that part of the experiment was made by a third-party.

2.2.1 FastQC

FastQC [58] provides a simple way of doing quality control checks on raw or filtered high throughput sequence data. Its output provides an HTML report with summary graphs and tables that help in the assessment of data quality. The 0.11.5 version of this tool was used in this project. FastQC is available at https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

2.2.2 MultiQC

MultiQC [59] is a tool that compiles the output report files of multiple programs into a single unified report, and provides a graphical visualization of the statistics of each program, making it easy to evaluate the behaviour of the data at each step. For this work, MultiQC was used to observe graphically the statistics associated with the STAR mapping step and the HTSeq-count counting step. For this work, the version of MultiQC used was 1.0. This tool is available at http://multiqc.info/.

2.2.3 STAR

An important part of any RNA-Seq analysis is the mapping step. A STAR [60] workflow normally consists of two steps: an indexing step, where both the genome file and the annotation of the genome file are taken into account to generate indexes that will be used in the following step. Indexes will contain information about genes and splice junction for example, arranged in a way that it is easy and fast to be accessible by the search mechanisms of this program. Next is a mapping step, which maps each FASTQ file (each sample) to the genome and outputs a SAM/BAM file. The file contains information about the positioning and the scoring of the mapping. This aligner has been shown to be extremely accurate and fast, when compared with other tools [61]. The version of this aligner used in this work was 2.5.2b. This widely used program is available at https://github.com/alexdobin/STAR.

2.2.4 SAMtools

The aforementioned SAM files are files that store large nucleotide sequence alignments. SAMtools [62] provide a variety of programs that allow the interaction with the data stored in those files. This tool was used for sorting the alignments via the -n option so they could be used as input for HTSeq-count and to index the associated BAM file to allow the visualization of the data through the Integrative Genomics Viewer [63,64]. The 1.3.1 version of SAMtools was used in the present work. These batches of tools are available at http://www.htslib.org/.

2.2.5 HTSeq-count

Counting the aligned reads is another important step in the analysis of RNA-seq data. One of the most commonly used programs for this, with a great amount of compatibly with other programs for the rest of the downstream analysis, is HTSeq-count [65]. In this project, most of the default definitions were used for both the analysis of mRNA and miRNA, except for the -t and -I option for the miRNA, that were changed to a more appropriate option, when the annotation file from miRBase was used. When compared with other counting tools like featureCounts, HTSeq-count is much slower and can sometimes wield a lesser gene count [66], mainly since it is more conservative. However, HTSeq is the most widely used tool, with 1599 articles cited compared to featureCounts' 481, and it is highly compatible with a lot of differential expression tools, like DESeq and DESeq2. The 0.6.1 version of this software was used in this project. The HTSeq documentation, which includes HTSeq-count is available at https://htseq.readthedocs.io/en/release_0.9.1/index.html.

2.2.6 R and R packages

R [67] is a language and environment mainly created for to aid statistical computing and graphical construction. This free software is available at https://www.r-project.org/. Due to some technical problems, the R versions for the Differential expression analysis of the mRNA (version 3.2.4) and miRNA (version 3.3.3) data were different (as seen in sections 2.3.1.4 and 2.3.2.4). This however is insignificant for the comparison of the data since the different versions yield the exact same results.

For this work, one of the main packages used for the study of the Differential Expression was DESeq2 [68]. The steps performed by this program are: estimation of size factors, estimation of dispersion, negative binomial generalized linear model (GLM) fitting, and hypothesis testing for differential expression using the Wald test. The Negative Binomial distribution is used in this package due to its superior performance when compared with Poisson-based methods [45,68]. Additionally, two extra steps are made automatically: independent filtering and the adjustment of the obtained p-value. Unlike in other packages, DESeq2 uses independent filtering to remove weakly expressed genes, which will not present any differential expression, due to the fact that the low read count in combination with sample noise will hide any biological effect that these genes may have. These genes will influence the multiple testing procedures and, as such need to be removed. The filtering criterion is the average expression strength of each gene, across all samples, and it omits all the genes detected with mean normalized counts below the filtering threshold. This, by default, maximizes the number of gene found at a certain FDR specified by the user [68,69]. For the creation of the adjusted p-value for each gene, the Benjamini-Hochberg (BH) adjustment is done, and the value is calculated as an answer to the following question: "if one would consider all genes as significant when their adjusted p-value is less than or equal to this gene's adjusted p-value threshold, what would be the fraction of false positives (also know as false discovery rate, FDR) among them?" For experiments with less than 12 replicates, both edgeR (exact) and DESeq2 are recommended [70]. On top of that, DESeq2 was chosen for this project also due to its' higher accessibility, when compared with edgeR. The version of this package used in this work was 1.10.1.

In case of the presence of batch effects, the package sva [71] was used to try to diminuish said effects in the data. Batch effects introduce a degree of variability to samples that is not related to a

biological variable in study and this effect becomes rather problematic when the variation induced by batch effects overpowers that induced due to a biological variable, putting in question the veracity of any conclusion derived from that data [72]. This package tries to account for these effects, by first identifying the part of the data that is affected by the batch effects and later estimating values for said effects to be added to the design formula. The design formula informs DESeq2 on how to treat the samples, based on their grouping [71]. The version 3.18.0 of the sva package was used in this project.

For the creation of heatmaps shown in this work, the pheatmap package was used. For this work, the version of this package used was 1.0.8.

The biomart package grants acess to a great amount of databases that are compatible with the BioMart software suite, allowing the retrievel of large amounts of data. This package was used in the present work in order to get the gene names corresponding to the ensembl IDs that were outputed in the counting step (later mentioned in section 2.3.1.3). In the present work, the version of this package used was 2.26.1.

For the correlation teste (see section 2.3.3) a specific built-in function of R was used, called cor.test(). This function was used to determine the correlation coefficient, through the Pearson method, and the respective probability value (p-value). The test statistic (t) is based on the correlation coefficient (r) and follows a t distribuition with n-2 degrees of freedom (n represents the number of element in each vector analysed). The formula for this test statistic is as follows:

$$t = \frac{r * \sqrt{n-2}}{\sqrt{1-r^2}} \quad (2.1)$$

The value obtained as a result is then used in the determination of the p-value.

2.2.7 DAVID

High-throughput experiments provide substantial amounts of data as a result of its analysis, and a lot of methods and tools have been created to extract the biological meaning behind those results. One such tool is the Database for Annotation, Visualization and Integrated Discovery (DAVID) [73,74], a functional enrichment tool, that possesses several analytical modules. For the functional enrichment analysis, DAVID uses a plethora of resources, including ontologies (GO terms), protein domains, pathway (KEGG and BioCarta), among others [73]. This tool can be used from its website: https://david.ncifcrf.gov/home.jsp.

2.3 Methods

Here are described all the steps taken in this work, for both the analysis of the data associated with the extracted mRNA and miRNA. The data sequencing step is referred here, even though it was not done by the author of this thesis, as a way to help understand the steps taken to obtain the sequenced data that was later used for this analysis. All the R code used for the differential analysis is available at https://github.com/InesALopes/RNA-seq-analysis .

2.3.1 Analysis of the mRNA

This section describes the methods used for the analysis of the mRNA data (Figure 2.1).

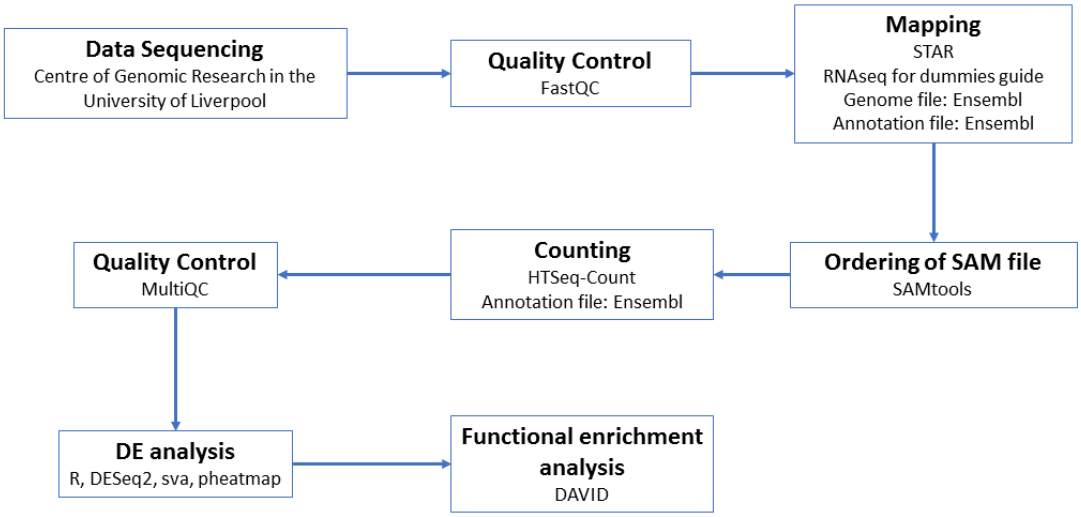

Figure 2.1: ***Flowchart for the analysis of the mRNA data.***

2.3.1.1 Data sequencing

The samples were sequenced at the Centre for Genomic Research in the University of Liverpool (available at https://www.liverpool.ac.uk/genomic-research/). The machine used was Illumina HiSeq and the sequencing was done as paired-end (maximum read length of 125 bp) with no technical replicates and three biological replicates for each condition. After obtaining the raw reads, a trimming step was performed, to exclude the adaptors from the sequence, using Cutadapt (version 1.2.1) [75] and the -O 3 option. For quality trimming, Sickle (version 1.200) [76] was used. Read with a score less than 20 and shorter than 10 bp after trimming were removed. Both the reads in a pair were discarded when either both or only one of them survived the trimming.

2.3.1.2 Quality Control

FastQC (version 0.11.5) [58] was used to do the quality control of the samples. MultiQC (version 1.0) [59] was also used to visualize the quality of the mapping and counting steps.

2.3.1.3 Mapping and Counting

The human genome for the mapping was obtained from Ensembl (release 86) [52]. To perform the mapping of the data, STAR (version 2.5.2b) [60] was used following the instructions of the RNAseq for dummies guide (available at http://www.genefriends.org/RNAseqForDummies/). SAMtools (version 1.3.1) [62] was used to order the samples, and the counting was done using HTSeq-count (version 0.6.1) [65].

2.3.1.4 Differential Expression Analysis

For the DE analysis, the DESeq2 (version 1.10.1) package [68] for R (version 3.2.4) [67] was used. Unwanted variation or non-recorded batch effects were examined and removed with the sva (version 3.18.0) [71] package. The heatmap were produced with the pheatmap package (version 1.0.8) and the PCA plots with a function provided by the DESeq2. The adjusted p-value (FDR) for which the significant genes were selected was equal to 0.05. The selection of genes that were considered down-regulated was done at log2 Fold Change inferior or equal to -1 and for up-regulated was done at log2 Fold Change superior or equal to 1.

2.3.1.5 Enrichment analysis

DAVID [73,74] was used for the enrichment analysis. The clustering mode was used and from each cluster the term with highest count was selected. In case of a tie, the term with the lowest FDR (maximum FDR accepted was 0.2) was chosen from each cluster. Only the top 5 clusters have been reported as long as their enrichment score was bigger or equal to 2. The Benjamini method was used to control FDR, therefore all mentioned FDR are calculated by this method. The genes that belong to the background were selected if they had some expression (5 mapped counts) in, at least, one of the samples.

2.3.2 Analysis of the miRNA

This section describes the methods used for the analysis of the miRNA data (Figure 2.2).

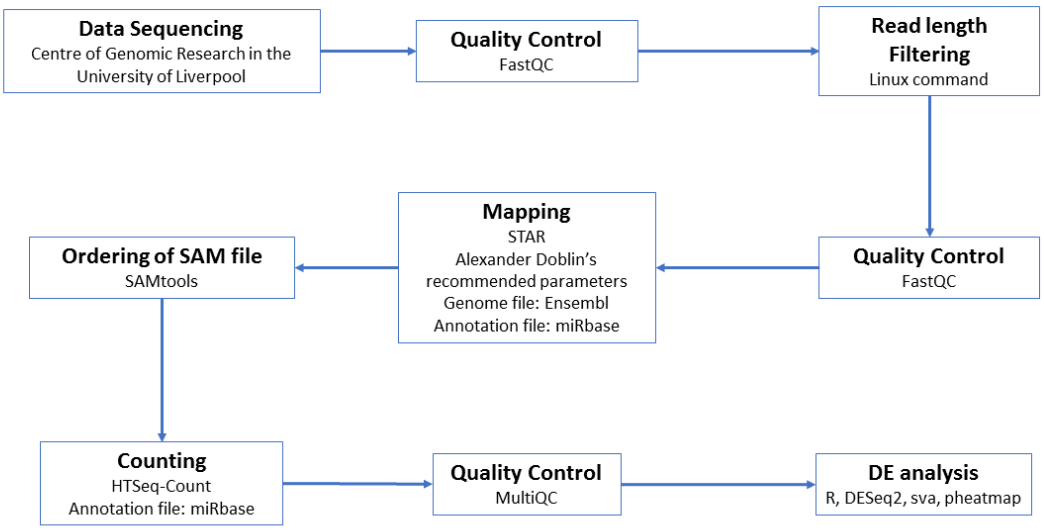

Figure 2.2: ***Flowchart for the analysis of the miRNA data.***

2.3.2.1 Data sequencing

The samples were sequenced at the Centre for Genomic Research in the University of Liverpool (available at https://www.liverpool.ac.uk/genomic-research/). The following description matches the steps that they followed to obtain the trimmed sequencing data. The machine that was used was Illumina HiSeq and the data for each sample was single-end (with the maximum read length at 50 bp) with two technical replicates (per sample) and three biological replicates (per condition). After obtaining the raw reads, a trimming step was performed, to exclude the adaptors from the sequence, using Cutadapt (version 1.2.1) [75] and the -O 3 option. Finally, a quality trimming step was done to prevent the inclusion in our data of low quality reads (reads with scores less than 20 and reads shorter that 10 bp were removed). For this the program Sickle (version 1.200) [76] was used.

2.3.2.2 Quality Control

FastQC (version 0.11.5) [58] was used to do the quality control of the samples. A Linux command (awk) was used for the selection of reads based on their size. MultiQC (version 1.0) [59] was also used to visualize the quality of the mapping and counting steps.

2.3.2.3 Mapping and Counting

The human genome for this step was obtained from Ensembl (release 86) [52]. The genome annotation file was obtained from a public miRNA database called miRBase (version 21) [54–56]. The mapping of the data was done using STAR (version 2.5.2b) [60] following some recommended parameters (available at https://groups.google.com/d/msg/rna-star/RBWvAGFooMU/yi9tlK-eVVsJ), given by the author (Alexander Dobin) of this tool, the following being: splicing was turned off by not including the annotation file in the indexation step, outFilterMismatchLmax has value 0.05, outFilterMatchNmin has value 16, outFilterScoreMinOverLread has value 0, outFilterMatchNminOverLread has value 0 and alignIntronMax has value 1. SAMtools (version 1.3.1) [62] was used to order the mapped sequences and HTSeq-count (version 0.6.1) [65] to do the counting step, only taking into account the mature sequences of miRNA rather than their precursor sequences.

2.3.2.4 Differential Expression Analysis

For the DE analysis, the DESeq2 package (version 1.10.1) [68] for R (version 3.3.3) [67] was used. The sva package (version 3.18.0) [71] assisted in the removal of unwanted variations or possible, non-detected, batch effects. The production of the heatmap was done with pheatmap package (version 1.0.8) and, in case of the PCA plots, the function for it in the DESeq2 package was used. The adjusted p-value (FDR) for which the significant miRNAs were selected was equal to 0.05, the selection of miRNA that were considered down-regulated was for genes miRNA log2 Fold Change was inferior or equal to -1 and for up-regulated was for miRNA whose log2 Fold Change was superior or equal to 1.

2.3.3 Correlation analysis

To check for a correlation between the level of expression of miRNA and the level of expression of their respective targets, a custom R script was built. In it, the following packages were used: biomart (version 2.26.1), readxl (version 1.0.0), plyr (version 1.8.4) and DESeq2 (version 1.10.1).

First, the number of reads was extracted for miRNA that were considered differentially expressed. The normalization of these values was done with the help of the normalization function integrated in the DESeq2 package. The same was done to all the mRNA data. Using the confirmed miRNA:Target pairs from the miRTarBase (release 6.0) [57], it was possible to assign the expression values for each pair and, with that, calculate the correlation value and p-value (using the cor.test function) for the miRNA:Target pairs. A visual representation of this process is given by the Figure 2.3.

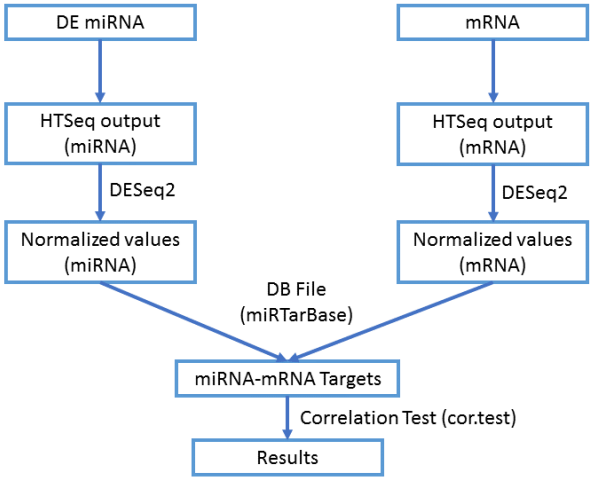

Figure 2.3: ***Flowchart for the correlation analysis.***

Chapter 3: Results

3.1 Analysis of the mRNA data

The main focus of this work was to twofold. Firstly, to understand the changes in macrophages when they are exposed to Aβ. Secondly, to understand said changes when they are compared with the different macrophage polarization states. Additionally, and beyond the initially proposed objectives, the comparison of the expression profiles of both polarization stages was done as well.

To do so, samples were obtained from healthy age-matched individuals and stimulated with Aβ (the Ab group), LPS and INF-γ (the M1 group) and IL-4 (the M2 group). Cells that did not receive any stimuli were considered controls. These groups had had their mRNA and miRNA sequenced and analysed. The following results pertain to the analysis of the mRNA data.

3.1.1 Quality Control

As previously noted, several steps of quality control were done in order to ensure that the sequenced data had good quality, that the analysis was well done and, in case of something going amiss, to figure out the best way to proceed in the analysis of the data.

The first quality control step was done on the reads provided by the sequencing facility. Overall, the reads provided had good quality. Only one issue was found, in sample 9 (Control), characterized by an overrepresented sequence (around 0.18% of the total of reads). However, since it is present in low percentages, this issue was ignored. The read lengths were also not consistent (as seen in Figure 3.1), but that was probably due to the fact that a quality trimming was done on the data.

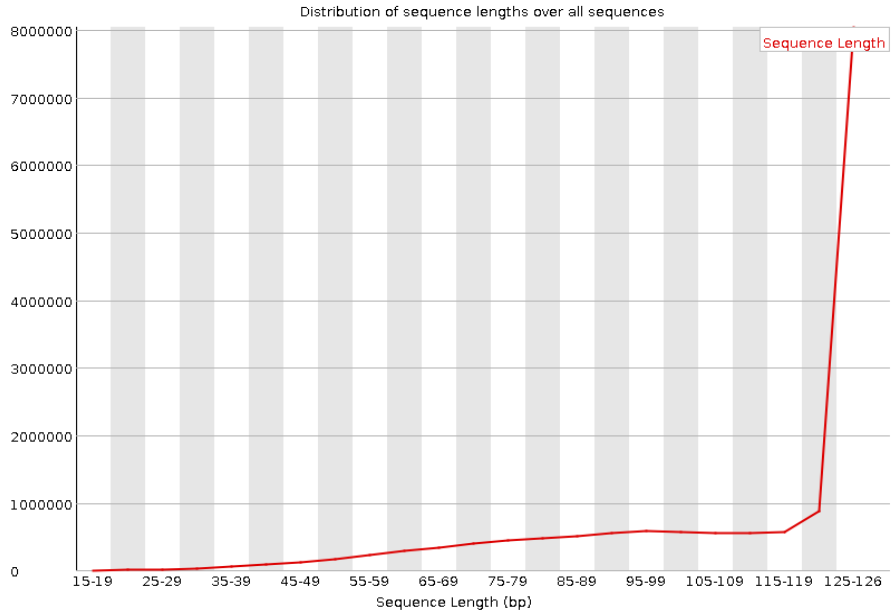

Figure 3.1: ***Plot of the distribution of the sequence length.*** *This plot refers to the first read of the first sample. Sequence length (bp) is plotted against the number of sequences that possess that sequence length. Plot generated by FastQC* [58].

The second quality control step was done after the mapping and counting step, with the help of the MultiQC software [59]. In terms of mapping, on average, 90.42% of the reads were successfully mapped (Figure 3.2), resulting in an average of 13583683 reads mapped per sample.

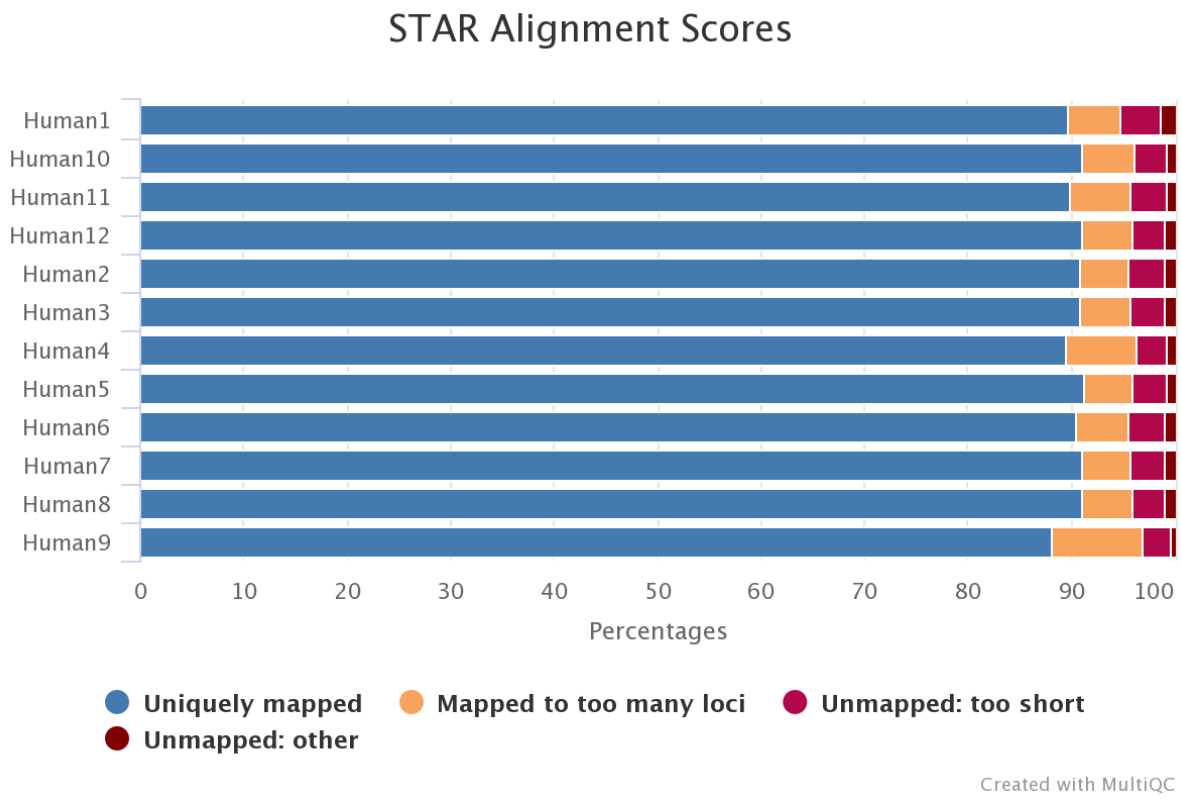

Figure 3.2: ***Percentage distribution across all samples for the mRNA STAR mapping.*** *Each bar corresponds to one of the samples (names ranging from Human1 to Human12) and each colour matches the STAR output statistics. Plot generated by MultiQC* [59].

However, when the same analysis is done to the results of the counting step, less positive results are seen. There is a big percentage of reads (on average, 85.54% of reads mapped) that were classified as "No Feature" in this step (Figure 3.3). When a counting program is used, like HTSeq-Count, there is usually an option to specify the type of feature that are meant to be considered (by default this option is exon, but one could change it for intron if they so wished), and when a specific read is not mapping to a location of that feature, that read is considered in the "No Feature" group. On average, 1792094 reads were successfully assigned, which is not an ideal number for an analysis like this.

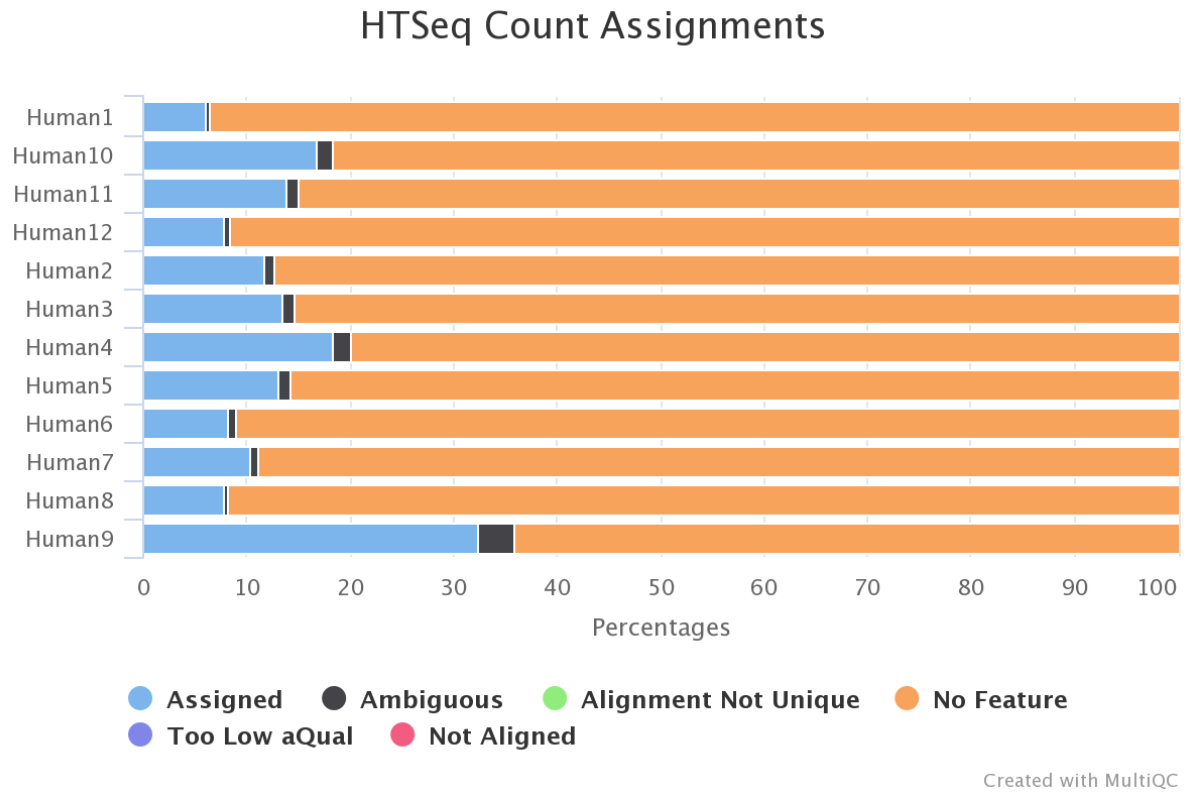

Figure 3.3: ***Percentage distribution across all samples for the mRNA HTSeq-Count counting step.*** *Each bar corresponds to one of the samples (names ranging from Human1 to Human12) and each colour matches the HTSeq-Count output statistics. Plot generated by MultiQC* [59].

To try to understand the existence of this problem, the Integrative Genomics Viewer (IGV) [63,64] was used. This program allows for the visualization of alignments, and it was used on the BAM file produced after the mapping step of the first sample. SeqMonk [77] was also used to help better visualize this over the course of several genes (Figure 3.6). Has it can be seen in Figures 3.4, 3.5 and 3.6, there was a great amount of background reads for non-exonic regions. The most likely explanation for this is that there was DNA contamination in the samples.

Figure 3.4: ***First example of possible mapped reads associated with contamination.*** *Each coloured strip corresponds to a read and the different colours are to help distinguishing between pairs (Each read in a pair has a different colour). The lower tab shows the gene distribution in the genome (the fuller line represents the exon and the arrowed line is the intronic region). Image obtained with IGV* [63,64].

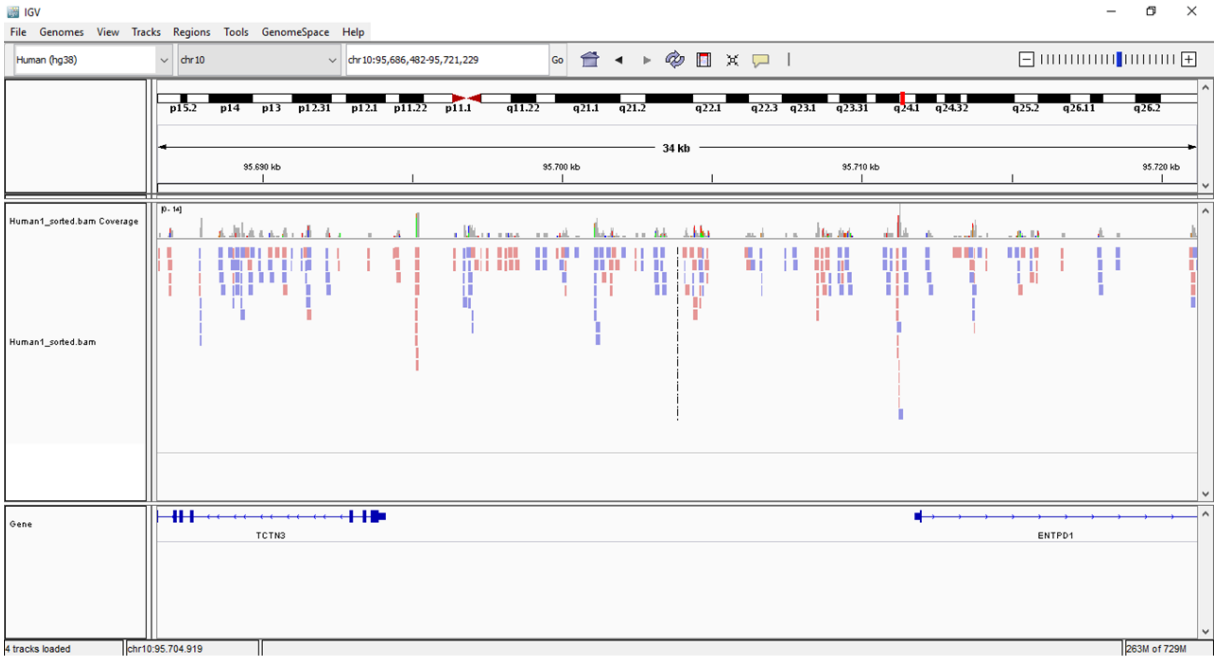

Figure 3.5: ***Second example of possible mapped reads associated with contamination.*** *Each coloured strip corresponds to a read and the different colours are to help distinguishing between pairs (Each read in a pair has a different colour). The lower tab shows the gene distribution in the genome (the fuller line represents the exon and the arrowed line is the intronic region). Image obtained with IGV* [63,64].

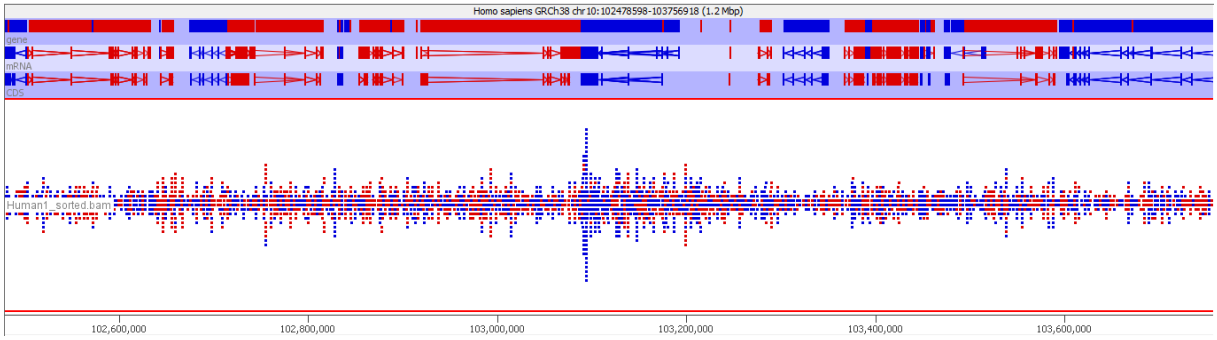

Figure 3.6: ***Third example of possible mapped reads associated with contamination.*** *Each coloured strip corresponds to a read and the different colours are to help distinguishing between pairs (Each read in a pair has a different colour). The top tab shows the gene distribution in the genome. Image obtained with SeqMonk* [77].

3.1.2 Differential expression analysis

One of the objectives of this project was to compare the genetic profiles of macrophages stimulated with different conditions. One of the first steps in a differential expression analysis is the creation of a PCA plot. These plots allow the user to visualize sample-to-sample distances, giving a general idea of the similarity between samples. What samples are similar to each other and which are different? This is important because it helps in the evaluation of the data, especially in the detection of possible batch effects.

The first PCA plot created is coloured by condition (M1, M2, Ab and Control) and shows that, when comparing all of the conditions studied, they do somewhat cluster together, but there is still a very high variation between samples of the same group (Figure 3.7). However, even with this variation, it is possible to see that there is a great difference between the M1 samples and all the others. We can also see that there is not a lot of variation between the control and Ab (Aβ) conditions, so one could expect a small number of genes differentially expressed between these two groups.

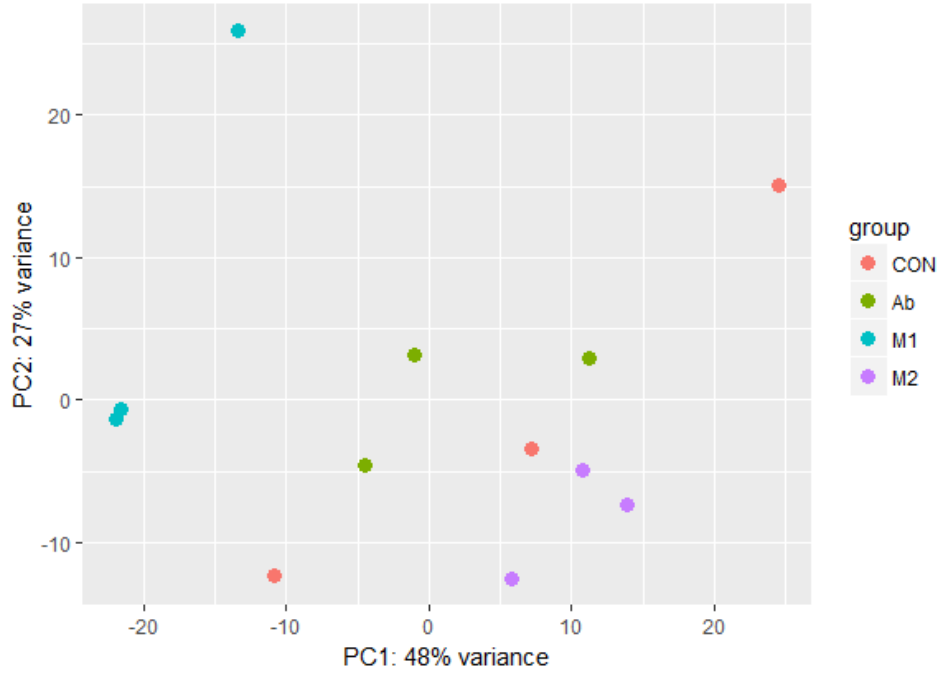

Figure 3.7: ***PCA plot for the different conditions (mRNA data).*** *Each dot represents one of the samples. Key: CON – Control, Ab – cells treated with Aβ, M1 – cells treated to achieve the M1 polarization state, M2 – cells treated to achieve the M2 polarization state. Plot generated in R* [67] *with the DESeq2* [68] *package.*

When the colouring of the PCA plot is changed to show the variation among the pools (Figure 3.8), that are constituted of the individuals whose samples were used for this work, it is possible to observe a somewhat clustering dependent of pools, with some outliers for all pools (which correspond to the M1 samples).

Optimally, it is expected from a PCA graph to have each condition clustered together in different places and each pool would be equally distributed between those. However, the fact that different conditions cluster somewhat together is not a major reason for concern. It only means that there is not a great difference in the genetic profiles between some of the conditions, this is especially prevalent when comparing the Ab group with the control group. However, and as stated before, normally a graph like this would show a greater clustering of samples with the same treatment, which does not occur in the data analysed. This could be due to two possible reasons: a great difference between pools of individuals, even when the treatment of each condition is applied or another unknown source of variation could be present in the data. The previously mentioned sva [71] package is capable of handling both of those cases and, as such, it was used to try to account for this sources of variation in the data and to dissipate their effects throughout the rest of the analysis.

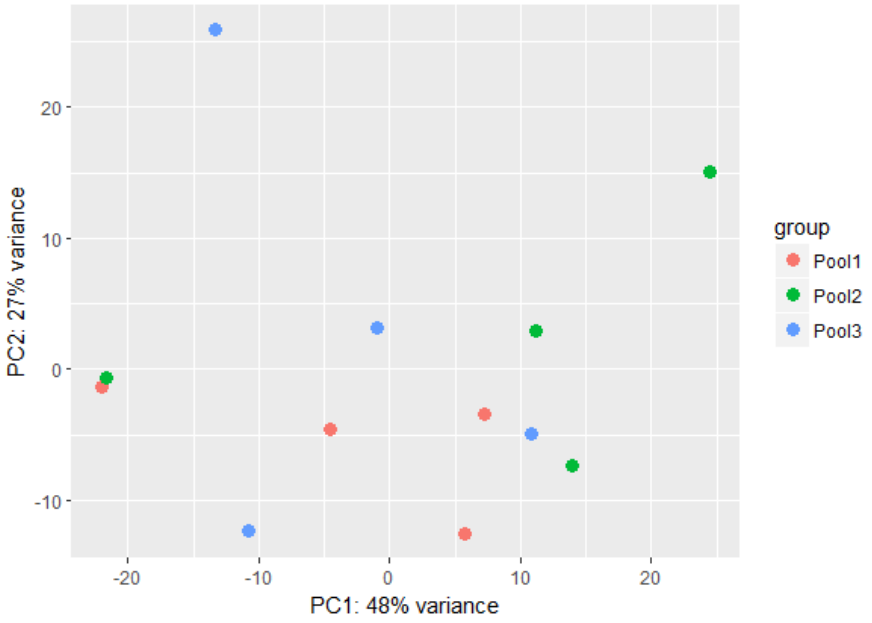

Figure 3.8: ***PCA plot for the different pools of individuals (mRNA data).*** *Each dot represents one of the samples. Plot generated in R* [67] *with the DESeq2* [68] *package.*

In a comparison of genetic profiles like this one, the packages used for this have a very specific way of representing that comparison. The user can select contrasts (A vs B, CON vs M1, etc.) to specify the comparison that they are looking for, and the output will give a list of genes "up-regulated" and "down-regulated". Genes that appear as up-regulated mean that their expression is higher in the first condition of the comparison. For example, in CON vs M1, genes up-regulated would be the ones with a higher expression in the Control group when this is compared with the M1 group. On the other hand, genes that are classified as down-regulated have a higher expression in the second condition of the comparison. For example, in CON vs M1, genes down-regulated would be the ones with a higher expression in the M1 group when compared with the CON group.

After accounting for the possible batch effects, the rest of the DESeq2 pipeline was ran, outputting a list of genes differentially expressed for each comparison of the conditions (Table 3.1). The comparison with the biggest number of differential genes was the M1 vs M2 comparison, as expected due to the great distance between them seen in the PCA plot (Figure 3.7). On the other hand, the comparison between the Control and the Ab group yielded the lowest number of differential genes; with only one gene with high expression in the Ab conditions.

Table 3.1: ***Genes up-regulated and down-regulated in the different comparisons tested.***

	Number of genes up-regulated	**Number of genes down-regulated**
Control vs M1	237	436
Control vs M2	88	164
Control vs Aβ	0	1
M1 vs M2	563	573
M1 vs Aβ	308	226
M2 vs Aβ	153	148

A heatmap was also produced, to help visualize the difference between the genetic profiles in the different conditions (Figure 3.9). Like in Table 3.1, it is possible to observe that the M1 samples are the most different when compared with all other conditions.

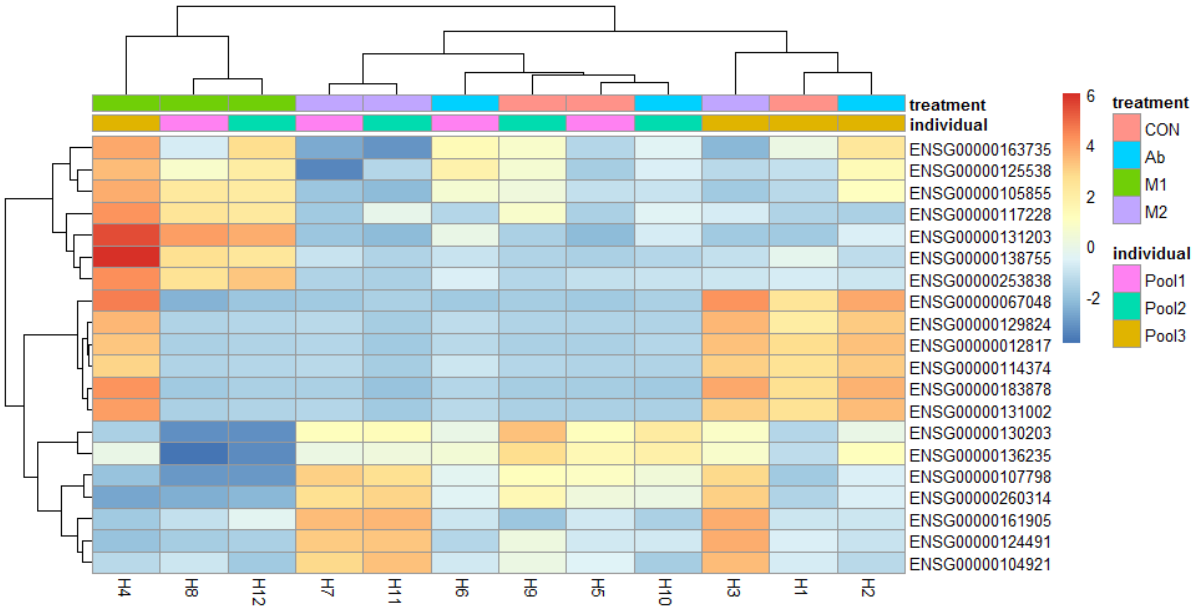

Figure 3.9: ***Heatmap of relative rlog-transformed values across samples (mRNA).*** *Treatment and pool identification are shown with coloured bars at the top of the heatmap. The values of expression here presented correspond to the amount by which each gene deviates in each sample, from that gene's average across all samples. Plot generated in R* [67] *using the pheatmap package.*

As mentioned before, the comparison between the Control group and the group of macrophages stimulated with Aβ only had one gene differentially expressed. This could possibly mean that the gene expression profile between these two conditions is too similar. To test this, a Venn diagram (Figure 3.10) was built, comparing the genes enriched in the Control and Ab groups when they are both compared to M1 or M2.

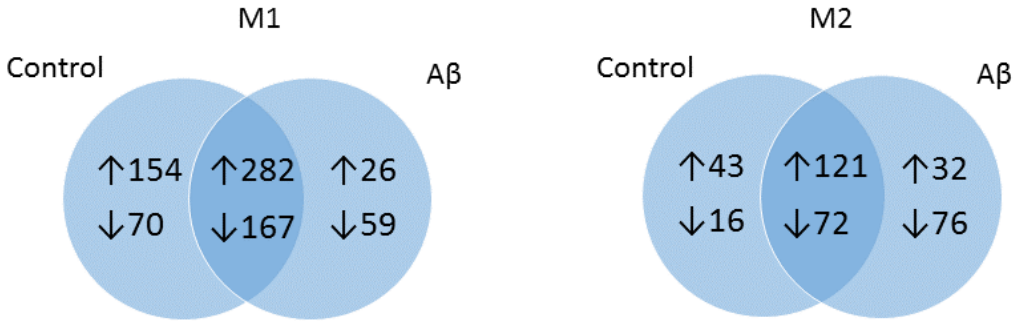

Figure 3.10: ***Venn diagrams for the different genes expressed when comparing different conditions against Control and Ab.*** *Arrow up (↑) represents the genes with great expression in the M1 or M2 conditions. Arrow down (↓) represents genes with a lesser expression in the M1 or M2 conditions (so a greater expression in either the control or Aβ conditions).*

As seen in Figure 3.10, while there are a lot of genes similar between the Control and Ab groups (in the context of the comparisons explained above), there are still a large group that are not shared between these two. However, this difference is just not significant enough to be detected when these two groups are compared. It can also be seen in the PCA plot, this allows us to conclude that, while they share a similar genetic profile, the similarity is not 100%.

The genes that had a bigger expression for the M1 condition were *IDO1*, *RP11-44k6.2*, *ACOD1*, *SNX10*, *ANKRD22*, *CXCL9* and *CCR7*, for the M2 condition they were *ALOX15*, *FCER2*, *TGM2*, *F13A1*, *CTNNAL1*, *MRC1* and *CLEC10A* and for the macrophages stimulated with Aβ they were *SOD2*, *FUCA1*, *FCN1*, *CD9*, *APOE*, *JAML*, *ADAMDEC1*, *ADAM28*, *PID1*, *RBMS1P1* (Table 3.2 and Table 3.3).

Table 3.2: ***Genes up-regulated is all comparisons.*** *The genes are ordered by Log2 Fold change. A gene was considered up-regulated if the Log2 Fold change value was equal or bigger than 1.*

Ensembl ID	Gene symbol	Gene name	Log2 Fold change	FDR
Control vs M1				
ENSG00000179163	FUCA1	fucosidase, alpha-L- 1, tissue	2.946591409	1.50E-26
ENSG00000085265	FCN1	ficolin 1	2.912948058	2.22E-41
ENSG00000010278	CD9	CD9 molecule	2.894277105	8.03E-23
ENSG00000160593	JAML	junction adhesion molecule like	2.558564814	5.41E-29
ENSG00000244682	FCGR2C	Fc fragment of IgG receptor IIc (gene/pseudogene)	2.516087535	7.79E-25
Control vs M2				
ENSG00000153823	PID1	phosphotyrosine interaction domain containing 1	2.339103132	4.91E-15
ENSG00000042980	ADAM28	ADAM metallopeptidase domain 28	2.197107203	8.00E-37
ENSG00000112303	VNN2	vanin 2	2.00163365	3.03E-12
ENSG00000166527	CLEC4D	C-type lectin domain family 4 member D	1.912522128	2.31E-11
ENSG00000153208	MERTK	MER proto-oncogene, tyrosine kinase	1.88033909	1.31E-14
M1 vs M2				
ENSG00000131203	IDO1	indoleamine 2,3-dioxygenase 1	5.555731558	1.20E-48
ENSG00000253838	RP11-44K6.2	Known sense intronic	4.892022974	2.45E-33
ENSG00000102794	ACOD1	aconitate decarboxylase 1	4.174092941	2.11E-41
ENSG00000086300	SNX10	sorting nexin 10	4.066583176	3.26E-65
ENSG00000152766	ANKRD22	ankyrin repeat domain 22	3.904260736	1.99E-39
M1 vs Aβ				
ENSG00000253838	RP11-44K6.2	Known sense intronic	4.328718146	3.97E-27
ENSG00000131203	IDO1	indoleamine 2,3-dioxygenase 1	4.199295252	1.68E-28
ENSG00000152766	ANKRD22	ankyrin repeat domain 22	3.778570628	7.64E-36
ENSG00000138755	CXCL9	C-X-C motif chemokine ligand 9	3.728045725	3.00E-15
ENSG00000102794	ACOD1	aconitate decarboxylase 1	3.685220254	5.77E-34
M2 vs Aβ				
ENSG00000161905	ALOX15	arachidonate 15-lipoxygenase	4.832113287	2.39E-57
ENSG00000104921	FCER2	Fc fragment of IgE receptor II	4.632259988	1.14E-46
ENSG00000198959	TGM2	transglutaminase 2	4.604702767	7.09E-82
ENSG00000124491	F13A1	coagulation factor XIII A chain	4.530684650	3.65E-78
ENSG00000119326	CTNNAL1	catenin alpha like 1	4.379555413	1.47E-74

Table 3.3: ***Genes down-regulated in all comparisons.*** *The genes are ordered by Log2 Fold change. A gene was considered down-regulated if the Log2 Fold change value was equal or smaller than -1.*

Ensembl ID	Gene symbol	Gene name	Log2 Fold change	FDR
Control vs M1				
ENSG00000131203	IDO1	indoleamine 2,3-dioxygenase 1	-5.471323144	9.67E-45
ENSG00000253838	RP11-44K6.2	Known sense intronic	-4.549378138	3.71E-29
ENSG00000102794	ACOD1	aconitate decarboxylase 1	-4.217541399	2.13E-39
ENSG00000126353	CCR7	C-C motif chemokine receptor 7	-3.921260488	1.93E-27
ENSG00000152766	ANKRD22	ankyrin repeat domain 22	-3.893712369	4.48E-36
Control vs M2				
ENSG00000161905	ALOX15	arachidonate 15-lipoxygenase	-4.734390296	3.95E-57
ENSG00000198959	TGM2	transglutaminase 2	-4.118587938	3.23E-72
ENSG00000119326	CTNNAL1	catenin alpha like 1	-3.939063746	5.95E-68
ENSG00000124491	F13A1	coagulation factor XIII A chain	-3.912919004	2.10E-63
ENSG00000104921	FCER2	Fc fragment of IgE receptor II	-3.725438299	2.16E-36
Control vs Aβ				
ENSG00000112096	SOD2	superoxide dismutase 2, mitochondrial	-1.033955307	0.014046843
M1 vs M2				
ENSG00000124491	F13A1	coagulation factor XIII A chain	-5.155662297	2.82E-89
ENSG00000260314	MRC1	mannose receptor, C type 1	-5.025939591	2.74E-101
ENSG00000161905	ALOX15	arachidonate 15-lipoxygenase	-4.812643403	1.33E-58
ENSG00000104921	FCER2	Fc fragment of IgE receptor II	-4.519275764	1.11E-45
ENSG00000132514	CLEC10A	C-type lectin domain family 10 member A	-4.356718267	1.70E-38
M1 vs Aβ				
ENSG00000179163	FUCA1	fucosidase, alpha-L- 1, tissue	-3.582691646	3.13E-42
ENSG00000085265	FCN1	ficolin 1	-3.050057878	1.13E-48
ENSG00000010278	CD9	CD9 molecule	-2.913355760	8.73E-25
ENSG00000130203	APOE	apolipoprotein E	-2.762910780	4.01E-11
ENSG00000160593	JAML	junction adhesion molecule like	-2.428973637	5.12E-28
M2 vs Aβ				
ENSG00000134028	ADAMDEC1	ADAM like decysin 1	-2.423522870	1.71E-28
ENSG00000112096	SOD2	superoxide dismutase 2, mitochondrial	-2.337175287	6.69E-29
ENSG00000042980	ADAM28	ADAM metallopeptidase domain 28	-2.311064358	3.77E-41
ENSG00000153823	PID1	phosphotyrosine interaction domain containing 1	-2.272832379	2.31E-14
ENSG00000225422	RBMS1P1	RNA binding motif single stranded interacting protein 1 pseudogene 1	-2.132746399	1.88E-07

3.1.3 Functional enrichment analysis

In order to analyse the great amount of gene information obtained from the Differential expression analysis, a functional enrichment analysis was done for all comparison, excluding CON vs Ab, since that one only yielded one gene. To do so, DAVID was used in the clustering mode. This mode clusters terms that are similar in function, allowing for the analysis of groups of annotations, rather than individual ones.

For the M1 group of genes, one of the most relevant term was immunity, which was expected due to the nature of the cells analysed. Other terms included were: signal peptide, immune response, glycoprotein, lipopolysaccharide-mediated signalling pathway, membrane, host-virus interactions, inflammatory response, immunoglobulin-like fold, cellular response to interleukin-1, NOD-like receptor signalling pathway, positive regulation of interleukin-1 beta secretion, and antigen processing and presentation. In terms of the M2 group of genes, the terms found were: membrane, immunoglobulin-like fold, MHC classes I/II-like antigen recognition protein, adaptive immunity, immunity, host-virus interactions, poly(A) RNA binding, methylation, nuclear nucleosome, and host cell receptor for virus entry. Finally, for the Ab group, the terms outputted by DAVID were: immunity, glycoprotein, protein heterodimerization activity, lectin, cholesterol metabolic process, signal peptide, response to lipopolysaccharide, and chemotaxis.

The most relevant member of a cluster (chosen based on the rules set on section 2.3.1.5) for all comparisons is shown in tables 3.4 and 3.5.

Table 3.4: ***Functional enrichment results for the up-regulated genes for each comparison.*** *The top five clusters were selected (ordered by enrichment score), and only one term for each cluster was included. The term must have the highest count and its FDR value must not be bigger than 0.2.*

Comparison	Term	Count	Benjamini (FDR)	Enrichment Score
CON vs M1	Immunity	26	2.00E-08	7.46
	Signal-anchor	14	5.70E-02	3.29
	receptor-mediated endocytosis	13	1.20E-04	3.24
	Membrane	102	1.40E-01	3.04
	domain:DAPIN	4	1.70E-01	2.08
CON vs M2	Glycoprotein	42	6.80E-07	7.09
	extracellular space	17	2.30E-03	6.42
	Immunity	12	4.60E-05	5.63
	Chemotaxis	6	8.80E-04	3.94
	Lysosome	6	5.40E-02	2.35
M1 vs M2	Immunity	51	1.90E-18	12.62
	signal peptide	115	3.80E-04	7.83
	immune response	45	3.20E-16	7.82
	Glycoprotein	144	1.40E-04	7.2
	lipopolysaccharide-mediated signalling pathway	11	1.10E-06	5.52
M1 vs Ab	Membrane	142	1.00E-05	4.1
	Host-virus interaction	17	3.30E-03	2.52
	Inflammatory response	12	2.30E-04	2.4
	Immunoglobulin-like fold	22	9.00E-02	2.14
	cellular response to interleukin-1	6	1.50E-01	2.05
M2 vs Ab	Membrane	90	1.30E-06	6.62
	Immunoglobulin-like fold	23	3.70E-06	3.98
	MHC classes I/II-like antigen recognition protein	8	3.50E-06	3.25
	Adaptive immunity	7	6.90E-03	2.03

Table 3.5: ***Functional enrichment results for the down-regulated genes for each comparison.*** *The top five clusters were selected (ordered by enrichment score), and only one term for each cluster was included. The term must have the highest the highest count and its FDR value must not be bigger than 0.2.*

Comparison	Term	Count	Benjamini (FDR)	Enrichment Score
CON vs M1	Membrane	188	2.60E-05	3.95
	immune response	37	2.30E-12	3.9
	NOD-like receptor signalling pathway	10	1.30E-03	3.52
	positive regulation of interleukin-1 beta secretion	7	1.50E-03	3.38
	Antigen processing and presentation	7	1.00E-01	2.93
CON vs M2	Membrane	95	5.70E-06	5.63
	Immunity	17	3.50E-05	4.31
	Host-virus interaction	12	8.20E-03	2.85
	Adaptive immunity	7	9.20E-03	2
M1 vs M2	poly(A) RNA binding	57	2.50E-02	11.41
	Methylation	50	8.00E-04	8.92
	Immunity	39	2.70E-08	6.11
	nuclear nucleosome	13	6.30E-09	4.37
	Host cell receptor for virus entry	10	1.50E-03	3.38
M1 vs Ab	Immunity	20	6.10E-05	5.17
	Glycoprotein	81	5.30E-05	3.3
	protein heterodimerization activity	16	4.40E-02	3.3
	Lectin	9	1.00E-02	2.69
	cholesterol metabolic process	7	4.00E-02	2.33
M2 vs Ab	Glycoprotein	56	2.20E-04	5.93
	signal peptide	47	2.60E-04	5.07
	Immunity	14	6.60E-04	4.44
	response to lipopolysaccharide	11	2.80E-04	3.04
	Chemotaxis	6	1.60E-02	2.81

3.2 Analysis of the miRNA data

The following results pertain to the analysis of the miRNA data.

3.2.1 Quality Control

Like for the analysis of the mRNA data, a study in the quality control of both the samples and the pipeline steps was done.

Again, the first step in quality control was done on the raw reads obtained from the sequence facility. Like the reads of mRNA, these reads were also of good quality. The read length graph (as

exemplified by Figure 3.11A for the P1_1 sample) for the unfiltered samples shows a great variety of peaks at different lengths. A lot of these reads were probably small nuclear RNAs (snRNA) or degraded mRNA and, as such, read length was selected to be between 18 and 26 [49]. To achieve this, a custom awk command was used on the Linux terminal. After the read length filtering, the number of peaks diminished a lot, seeing as now most samples possess 3 peaks (Figure 3.11B).

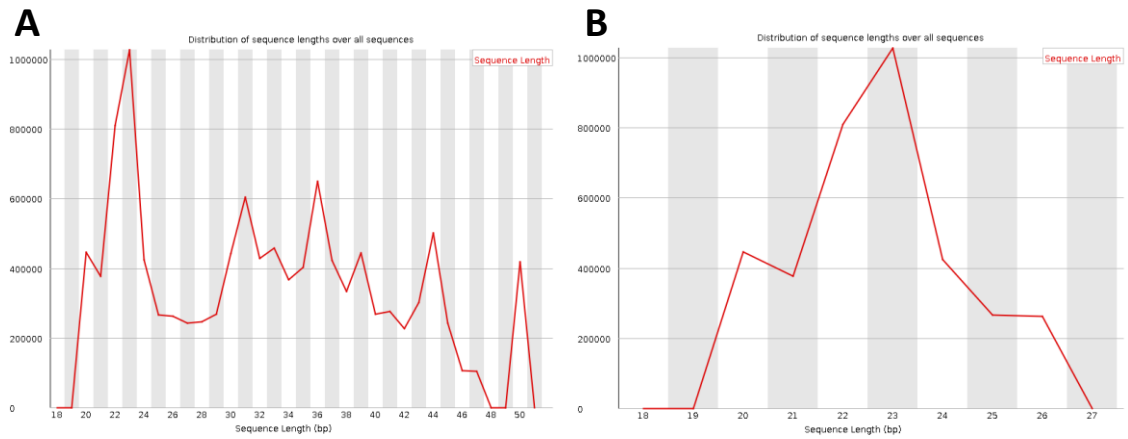

Figure 3.11: ***Plot of the distribution of the sequence length before and after the read filtering.*** *Sequence length (bp) plotted against the number of sequences that possess that sequence length. The sample used to build this plot was P1_1 (Sample 1, technical replicate 1). Plot generated by FastQC* [58]. *A – Before length selections; B – After length selections.*

There were a lot more sequences overrepresented, when compared with the mRNA analysis. This could be due to the nature of the samples, particularly the fact that this analysis is for miRNA and some miRNA are simply over expressed. For the raw reads, the most highly expressed sequences were miRNA, snRNA, or mRNA among others. After filtering the reads, the top overrepresented sequences seem to only be miRNA.

In terms of the mapping step, there is no consensus on what percentage of alignment of the reads should be seen. All the samples had 50% to 60% of their reads uniquely mapped (Figure 3.12). Sample 7, however, had a much lower number of reads when compared with all the other samples (around 600000). This had already been noted by the sequencing facility, so it might be attributed to the preparation of the samples. On average, 2110541 reads were uniquely mapped for each sample.

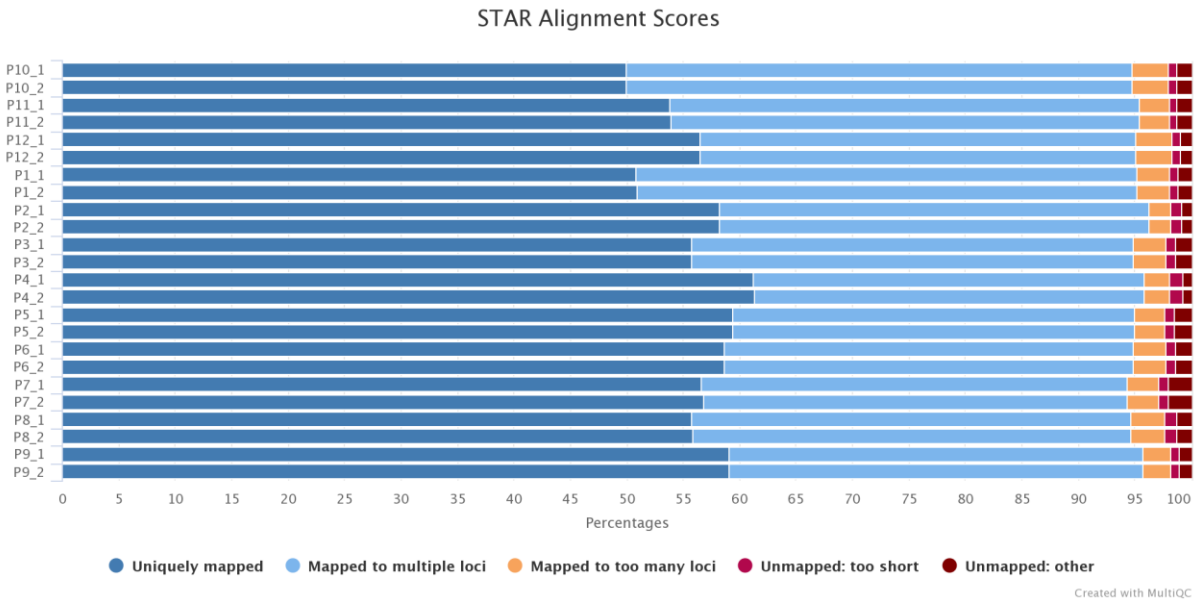

Figure 3.12: ***Percentage distribution across all samples for the miRNA STAR mapping.*** *Each bar corresponds to one of the samples (names ranging from P1_1 to P12_2) and each colour matches the STAR output statistics. Plot generated by MultiQC* [59].

There was a great number of mapped reads that showed as alignment not unique (80%-87%) in HTSeq-count (Figure 3.13). This could be explained by the fact that miRNA can align with multiple places in the genome and most counting software's are not equipped to deal with this disparity. Actual reads assigned varied between 9% and 18%, leading an average of 1617686 reads per sample that can be used for the rest of the analysis. This number of reads is not ideal. It should be noted that there is a discrepancy between the percentage of assigned reads and the actual number of assigned reads, when these are compared with the number of uniquely mapped reads obtained in STAR. This happened due to the fact that STAR counts each multi-mapped read once, while HTSeq-Count counts them as many times as they are mapped, in other words, HTSeq treats the value "alignment not unique" as an alignment count rather than a read count. The values of "Assigned", "Ambiguous" and "No Feature" are considered read counts and their added values correspond approximately to the number of uniquely mapped reads given by STAR. Taking this into account, a more accurate average of the percentage of assigned reads would be around 77% per sample.

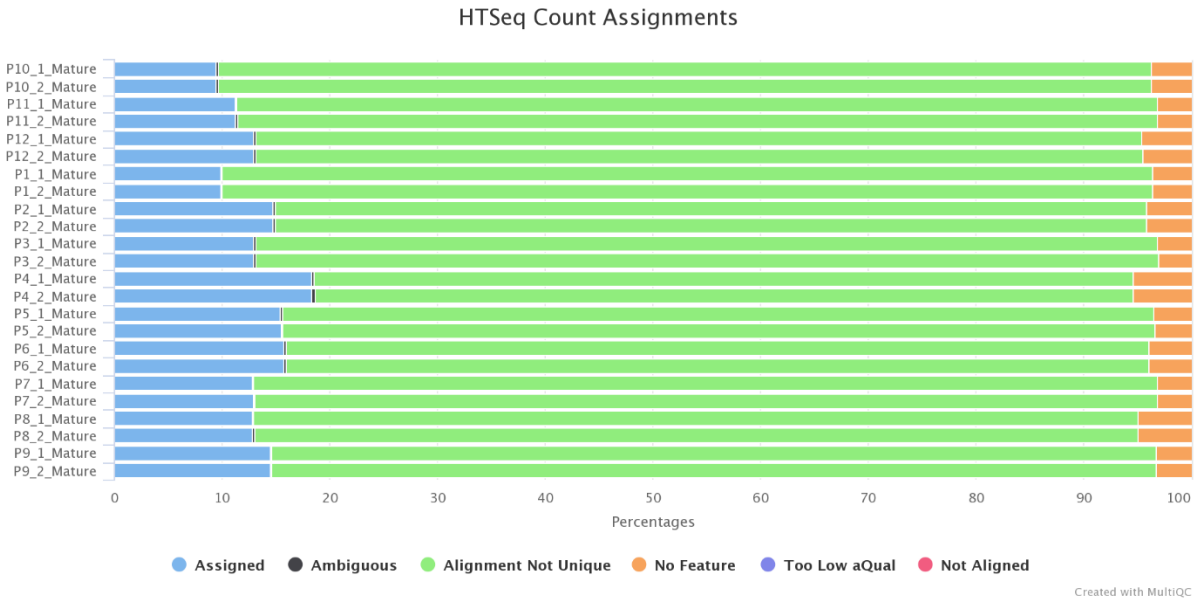

Figure 3.13: ***Percentage distribution across all samples for the miRNA HTSeq-Count counting step.*** *Each bar corresponds to one of the samples (names ranging from P1_1 to P12_2) and each colour matches the HTSeq-Count output statistics. Plot generated by MultiQC* [59].

3.2.2 Differential expression analysis

Similarly to what was done for the analysis of the mRNA data, a differential expression analysis was also done to the miRNA reads, following the mapping and counting steps.

Starting with the construction of the PCA plot, it is possible to see that each condition seems to somewhat cluster together and, like in the mRNA data, there is a great variation between samples of each group (Figure 3.14). In a similar fashion to what was observed for the mRNA data, the Control and Ab samples seem to cluster together, whilst M1 has the biggest variation from all the other samples, seeing as it clusters in its own side of the graph. Through the observation of the distances between samples mapped by this graph, there is a great likelihood that the comparison between the Ab and Control groups will not yield a significant amount of differential expressed miRNAs and that any comparison made against the M1 group will entail in a great number of differential expressed miRNAs.

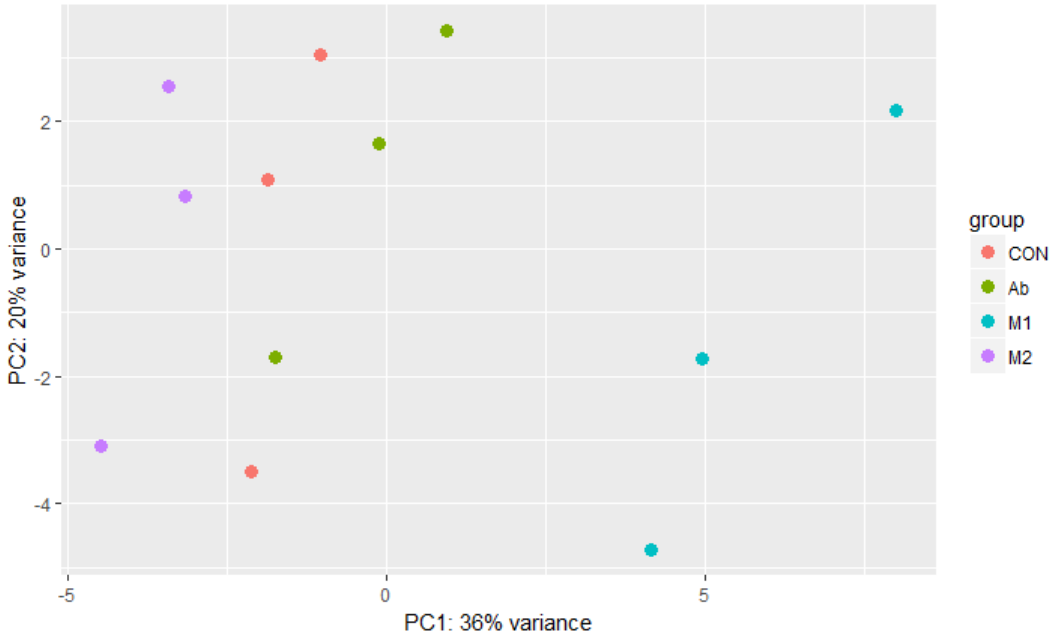

Figure 3.14: ***PCA plot for the different conditions (miRNA data).*** *Each dot represents one of the samples. Key: CON – Control, Ab – cells treated with Aβ, M1 – cells treated to achieve the M1 polarization state, M2 – cells treated to achieve the M2 polarization state. Plot generated in R* [67] *with the DESeq2* [68] *package.*

Changing the colours of the PCA plot to be dependent on the pool of individuals (Figure 3.15), it is possible to see that the similarities between the mRNA and miRNA data do not end there. Just like before, each pool of individuals tends to cluster together. Again, there seems to be a somewhat greater variability between individuals (or due to other unknown conditions) than variability induced by the treatment. As such, the sva package [71] was used again to account for that.

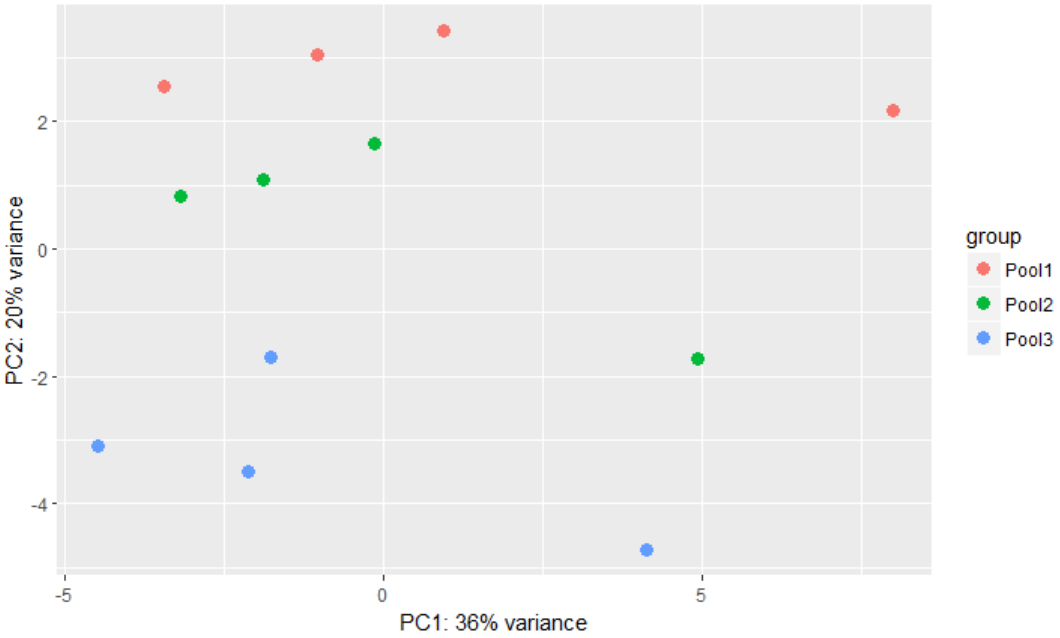

Figure 3.15: ***PCA plot for the different pools of individuals (miRNA data).*** *Each dot represents one of the samples. Plot generated in R* [67] *with the DESeq2* [68] *package.*

In terms of the differential expressed miRNA obtained from this analysis (Table 3.6), the comparison between the M1 and M2 groups yielded the biggest amount of miRNA, while the comparison between the Control and Ab groups had none, returning poorer results than the mRNA analysis.

Table 3.6: ***miRNA up-regulated and down-regulated in the different comparisons tested.***

	Number of miRNA up-regulated	Number of miRNA down-regulated
Control vs M1	6	18
Control vs M2	1	6
Control vs Aβ	0	0
M1 vs M2	23	11
M1 vs Aβ	14	7
M2 vs Aβ	5	3

A heatmap was also constructed (Figure 3.16), and it is possible to see that, for both the M1 and M2 groups, the samples cluster together, creating condition-specific clusters. Additionally, due to the similarity between the Control and Ab groups, they somewhat cluster by pools rather than condition.

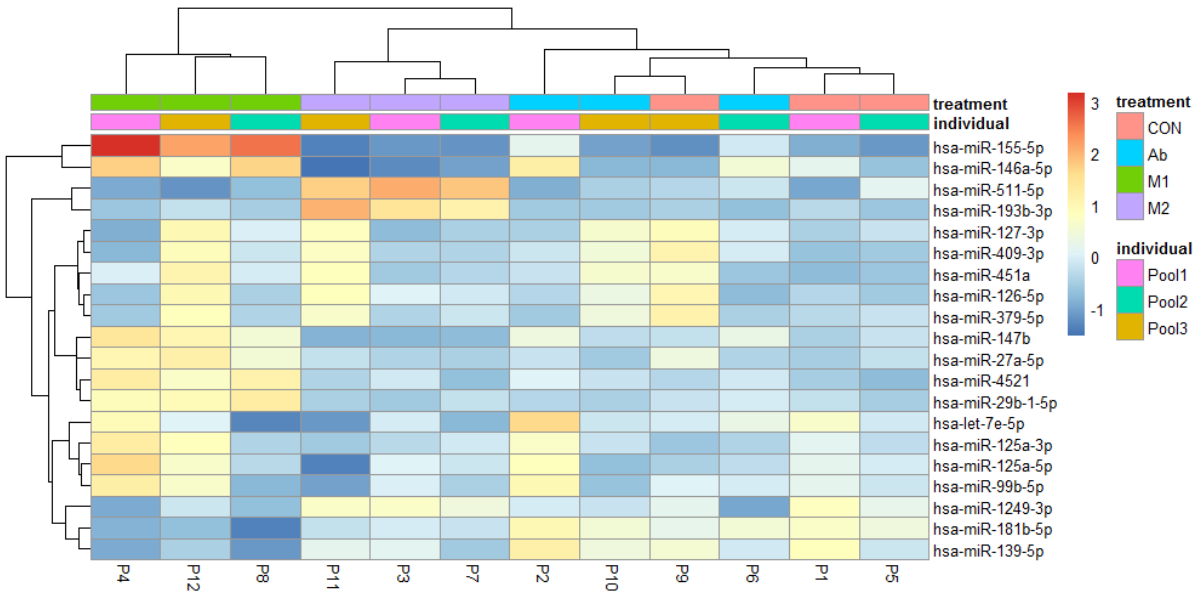

Figure 3.16: ***Heatmap of relative rlog-transformed values across samples (miRNA).*** *Treatment and pool identification are shown with coloured bars at the top of the heatmap. The values of expression here presented correspond to the amount by which each gene deviates in each sample, from that gene's average across all samples. Plot generated in R* [67] *using the pheatmap package.*

For the M1 group, the miRNAs up-regulated for this condition were: hsa-miR-155-5p, -146a-5p, -147b, -29b-1-5p, -449c-5p, -155-3p, -27a-5p and -4521. For M2, they were: hsa-miR-193b-3p, -511-5p, -365a-5p, -193b-5p, -1249-3p, 132-3p and -511-3p. For the macrophages stimulated with Aβ (and only when comparing this group with the M1 and M2 groups), they were: hsa-miR-181b-5p, -139-3p, -139-5p, -1296-5p, -30c-1-3p, -146a-5p, -34a-5p and -147b. All of this can be seen in Table 3.7.

Table 3.7: ***miRNA up-regulated and down-regulated in all comparisons.*** *The miRNAs were ordered by Log2 Fold change. A miRNA was considered up-regulated if the Log2 Fold change value was equal or bigger than 1 and down-regulated if the Log2 Fold change value was equal or smaller than -1.*

Up-regulated			**Down-regulated**		
miRBase ID	Log2 Fold change	FDR	miRBase ID	Log2 Fold change	FDR
Control vs M1			**Control vs M1**		
hsa-miR-181b-5p	1.629769665	2.70E-10	hsa-miR-155-5p	-4.066773084	2.45E-71
hsa-miR-139-5p	1.468294398	5.63E-05	hsa-miR-29b-1-5p	-1.938073642	5.22E-07
hsa-miR-139-3p	1.459452018	0.005482	hsa-miR-449a	-1.93708571	1.22E-05
hsa-miR-1296-5p	1.192301092	0.012025	hsa-miR-146a-5p	-1.891727137	6.78E-10
hsa-miR-1249-3p	1.140984413	0.001978	hsa-miR-4521	-1.788082598	4.80E-18
Control vs M2			**Control vs M2**		
hsa-miR-34a-5p	1.297099033	0.000228	hsa-miR-193b-3p	-3.059871191	2.23E-14
M1 vs M2			hsa-miR-511-5p	-2.45466273	6.61E-10
hsa-miR-155-5p	4.180201927	5.26E-64	hsa-miR-365a-5p	-1.817508105	0.000830503
hsa-miR-146a-5p	2.75441328	6.07E-19	hsa-miR-193b-5p	-1.498430645	0.014157368
hsa-miR-147b	2.306068728	1.07E-11	hsa-miR-132-3p	-1.130910242	2.72E-05
hsa-miR-29b-1-5p	2.218501363	2.39E-08	**M1 vs M2**		
hsa-miR-449c-5p	2.076994461	5.14E-06	hsa-miR-511-5p	-3.315567904	1.67E-18
M1 vs Ab			hsa-miR-193b-3p	-3.010147137	6.02E-15
hsa-miR-155-5p	3.223816935	1.40E-44	hsa-miR-511-3p	-1.951053277	0.000108616
hsa-miR-29b-1-5p	1.894398319	9.28E-07	hsa-miR-365a-5p	-1.82477785	0.000108616
hsa-miR-449c-5p	1.872718584	3.43E-05	hsa-miR-193b-5p	-1.502226198	0.001423178
hsa-miR-155-3p	1.784102728	0.000799	**M1 vs Ab**		
hsa-miR-27a-5p	1.458239627	2.28E-07	hsa-miR-181b-5p	-1.870304816	1.82E-13
M2 vs Ab			hsa-miR-139-3p	-1.544910453	0.0031307
hsa-miR-193b-3p	3.330747534	7.36E-17	hsa-miR-139-5p	-1.544451836	2.09E-05
hsa-miR-511-5p	2.650828705	1.36E-11	hsa-miR-1296-5p	-1.404584343	0.002581245
hsa-miR-365a-5p	1.83107499	0.000645	hsa-miR-30c-1-3p	-1.109161231	0.024410352
hsa-miR-193b-5p	1.421936112	0.022828	**M2 vs Ab**		
hsa-miR-1249-3p	1.092993088	0.031129	hsa-miR-146a-5p	-1.687812812	2.34E-06
			hsa-miR-34a-5p	-1.459717931	1.16E-05
			hsa-miR-147b	-1.206874985	0.018740345

In terms of only the M1 and M2 results, there were a total of 24 and 12 miRNAs with greater expression in these groups, respectively (Table 3.8).

Table 3.8: ***miRNA up-regulated for the M1 and M2 samples, across all comparisons.***

M1	M2
hsa-miR-155-5p	hsa-miR-193b-3p
hsa-miR-29b-1-5p	hsa-miR-511-5p
hsa-miR-449a	hsa-miR-365a-5p
hsa-miR-146a-5p	hsa-miR-193b-5p
hsa-miR-4521	hsa-miR-132-3p
hsa-miR-147b	hsa-miR-132-5p
hsa-miR-146a-3p	hsa-miR-511-3p
hsa-miR-449c-5p	hsa-miR-1296-5p
hsa-miR-155-3p	hsa-miR-1249-3p
hsa-miR-365b-5p	hsa-miR-3613-5p
hsa-miR-29a-3p	hsa-miR-342-5p
hsa-miR-4645-3p	hsa-miR-139-5p
hsa-miR-125a-3p	
hsa-miR-27a-3p	
hsa-miR-27a-5p	
hsa-miR-130b-3p	
hsa-miR-130a-3p	
hsa-miR-187-3p	
hsa-miR-29a-5p	
hsa-miR-181d-5p	
hsa-miR-125a-5p	
hsa-miR-1277-3p	
hsa-miR-22-5p	
hsa-miR-200a-3p	

3.3 Correlation analysis

In order to understand better the relationship between the miRNA levels of expression and target mRNA level of expression, a correlation analysis was done. As described in section 2.3.3, the database miRTarBase was used. This database compounds the miRNA:Target relationships that have been proven in other works. As such, there was no use of predicted targets. This database also distinguishes between what is called “Strong evidence” and “Less strong evidence” as was explained in section 2.3.3. When all of the evidence is taken into account, there is a noticeable skewering of the amount of positive correlations caused by one miRNA in particular, miR-34a-5p (Table 3.9). However, when only the “strong evidence” is selected there is a much smaller skewing of the data (Table 3.10).

Table 3.9: ***Results for the correlation using all targets.*** *If the value of the correlation coefficient was equal or bigger than 0, the correlation was considered to be positive. If the value of the correlation coefficient was smaller than 0, then the correlation was considered negative. All correlations considered had p-value lesser than 0.05.*

miRNA	Positive Correlations	Negative Correlations	Total Correlations	Percentage of Positive correlations	Percentage of Negative correlations
hsa-miR-1249-3p	0	1	1	0	100
hsa-miR-125a-3p	10	1	11	90.91	9.09
hsa-miR-125a-5p	12	0	12	100	0
hsa-miR-1296-5p	1	1	2	50	50
hsa-miR-130a-3p	5	1	6	83.33	16.67
hsa-miR-130b-3p	10	13	23	43.48	56.52
hsa-miR-132-3p	3	0	3	100	0
hsa-miR-139-3p	13	2	15	86.67	13.33
hsa-miR-146a-3p	4	18	22	18.18	81.82
hsa-miR-146a-5p	8	5	13	61.54	38.46
hsa-miR-147b	1	0	1	100	0
hsa-miR-155-3p	1	0	1	100	0
hsa-miR-155-5p	28	7	35	80	20
hsa-miR-181b-5p	3	1	4	75	25
hsa-miR-181d-5p	11	7	18	61.11	38.89
hsa-miR-193b-3p	15	3	18	83.33	16.67
hsa-miR-193b-5p	3	0	3	100	0
hsa-miR-200a-3p	3	1	4	75	25
hsa-miR-22-5p	5	0	5	100	0
hsa-miR-27a-3p	12	2	14	85.71	14.29
hsa-miR-27a-5p	3	1	4	75	25
hsa-miR-29a-3p	6	3	9	66.67	33.33
hsa-miR-29a-5p	2	0	2	100	0
hsa-miR-29b-1-5p	1	0	1	100	0
hsa-miR-30c-1-3p	5	3	8	62.5	37.5
hsa-miR-342-5p	1	1	2	50	50
hsa-miR-34a-5p	193	10	203	95.07	4.93
hsa-miR-3613-5p	1	1	2	50	50
hsa-miR-365a-5p	2	1	3	66.67	33.33
hsa-miR-365b-5p	6	1	7	85.71	14.29
hsa-miR-449a	3	2	5	60	40
hsa-miR-449c-5p	2	0	2	100	0
hsa-miR-4521	1	0	1	100	0
hsa-miR-505-3p	0	5	5	0	100
hsa-miR-511-3p	2	0	2	100	0
hsa-miR-511-5p	0	1	1	0	100
Total	376	92	468	80.34	19.66

Table 3.10: ***Results for the correlation test using only targets with "strong evidence".*** *If the value of the correlation coefficient was equal or bigger than 0, the correlation was considered to be positive. If the value of the correlation coefficient was smaller than 0, then the correlation was considered negative. All correlations considered had p-value lesser than 0.05. See section 2.3.3 for the definition of types of evidence.*

miRNA	Positive Correlations	Negative Correlations	Total Correlations	Percentage of Positive Correlations	Percentage of Negative Correlations
hsa-miR-125a-5p	1	0	1	100	0
hsa-miR-130b-3p	0	1	1	0	100
hsa-miR-146a-5p	5	1	6	83.33	16.67
hsa-miR-155-5p	6	3	9	66.67	33.33
hsa-miR-181d-5p	1	0	1	100	0
hsa-miR-200a-3p	1	1	2	50	50
hsa-miR-27a-3p	3	0	3	100	0
hsa-miR-29a-3p	5	1	6	83.33	16.67
hsa-miR-34a-5p	17	6	23	73.91	26.09
hsa-miR-449a	3	0	3	100	0
Total	42	13	55	76.36	23.64

Chapter 4: Discussion

4.1 mRNA data analysis

As mentioned before, there was a DNA contamination in the samples. This could happen when the digestion of the DNA, by the DNAse, is ineffective, for example. Other works don't often mention the rate of non-exonic mapping, either because the contamination is detected in the library preparation step or because it is simply not detected at all. DNA contamination tends to not raise any flags in standard analysis pipelines and they don't affect the rate of mapping considerably, making it hard to detect. On a work done for human samples, the rate of non-exonic mapping was found to be around 9% [78], and on another done in mice, the rate detected was 7% [79]. This rate can be slightly bigger, depending on how well annotated the genome is for that organism. Contamination of samples by DNA can also bolster the non-exonic mapping rate, which is what most likely happened in this case.

One of the main objectives of this work was to compare the expression profiles between the conditions proposed: macrophages induced into an M1 or M2 polarization state or stimulated with Aβ. In addition to that, a control sample was also designed. The differential expression analysis allows for that gene profile comparison and its results can be analysed through other tools (like functional enrichment tools), to help answer a biological question.

In terms of the top 5 genes, across all comparisons (Tables 3.2 and 3.3), with greater expression in the M1 group of samples, three (*IDO1*, *CCR7* and *CXCL9*) of them are identified as markers for this polarization type (Table 1.1). The gene *IDO1* codes for an enzyme that catalyses tryptophan and has a possible role in depression and other diseases of the brain, like Huntington's disease [80]. Its activity is increased in neuroinflammation upon the release of cytokines [81], and that might be why this gene is over-expressed for this condition, since the M1 polarization produces a lot of cytokines (Tables 1.1). Another gene often found expressed in these cells is *CCR7*, coding to a chemokine receptor whose expression has been shown to be associated with the number of activated T cells in circulation [82]. This gene is a characteristic marker of pro-inflammatory macrophages [83]. *CXCL9* codes to a chemokine that is often expressed in M1 polarized macrophages and it has a role in the recruitment and polarization of Th1 cells [14].

However, other non-marker genes were identified with high expression for this polarization state. The *ACOD1* gene (Immune-responsive gene 1, *IRG1*) codes for an enzyme that produces itaconic acid, used in immune defence. This gene is expressed in macrophages especially under pro-inflammatory conditions. i.e. presence of LPS [84]. When up-regulated, this gene can act as a negative regulator for TLR by increasing the expression of *TNFAIP3* or A20 using ROS [85]. For the *ANKRD22* gene, a study has shown that, in T cell-mediated rejection (TCMR) conditions, this gene has a great expression in macrophages that were treated with IFN-γ [86]. Since IFN-γ was one of the inducers used in this work to polarize macrophages into the M1 phenotype, it is possible that the expression of this gene is in response to IFN-γ, independent of the TCMR conditions. *SNX10* is a type of sorting nexin that has vacuolizing activity [87]. This type of activity is crucial in macrophages due to many of its functions being dependent on it, such as phagocytosis, digestion of pathogens, among others. In case of *SNX10* Knock Out, the polarization process in macrophages will tend more towards M2. As such, *SNX10* is important for the phagocytosis done in M1 cells (clearing of pathogens and antigen presenting) but not that relevant for the one done by M2 cells (clearance of apoptotic cells,

debris, tissue repair and remodelling) and, as such, can control the polarization of macrophages to M2 [88]. *RP11-44K6.2* (as of the most recent Ensembl update, known as *AC007991.2*) is an intronic sequence that is within the intronic region of the gene *IDO1*. As previously discussed, there was a great number of reads mapping to introns and, as such it is possible that the differential expression of this gene is a consequence of reads mapping to that region.

For the M2 polarized macrophages, six genes (*ALOX15*, *FCER* also known as *CD23*, *TGM2*, *F13A1*, the mannose receptor *MRC1* and *CLEC10A*) were identified as markers for this polarization, from the group of genes in Tables 3.2 and 3.3. *ALOX15* is a lipoxygenase, an enzyme that is responsible for the synthesis of biological active lipids [89]. This gene has been found to be up-regulated in M2 polarized cells that can do efferocytosis; non-inflammatory removal of apoptotic cells in adipose tissue [90,91]. It also appears to be modulated by the activation of macrophages with IL-4 [92], which would account for its greater expression in M2a cells, that are induced using IL-4. One of the contributors to an allergic reaction is immunoglobulin E. *FCER2* (or *CD23*) codes for a low affinity receptor for this kind of immunoglobulin that may inhibit the production of IgE through a negative feedback mechanism, and one of its forms can be found in monocytes. Its expression at the cell surface can be induced through various mechanisms including cytokines (IL-4 and IL-13) [93], which may be the reason for its high levels of expression in the M2a cells. The gene *TGM2* codes for an enzyme that belongs to the transglutaminase family. This family is one responsible for the catalytic function that makes post-translational modification of proteins [94]. This enzyme participates in the clearance of apoptotic cells, a function in which the M2 cells are also involved [19]. Interestingly enough, M1 cells also had great levels of expression for this enzyme when compared with the Control and Ab groups, but not when compared to M2 cells. So, perhaps, the definition of this gene as a marker should be accompanied with a warning that it only applies for comparisons against other polarization states, and that both M1 and M2 cells express this gene to a greater extent than un-stimulated cells. *F13A1* codes to a protein called Coagulation factor XIII A chain, that is able to crosslink with fibrin and play a role in inflammation [95]. The mannose receptor coded by *MRC1* (*CD206*) has a role in the regulation of phagocytosis [96] and its levels of expression are high in M2 polarized cells due to the activation with IL-4 [92]. *CLEC10* (also known as *CD301*) is a galactose-type C-type lectin that acts as a receptor and its expression is increased in the presence of IL-4 [97].

The only gene not identified as marker of M2 polarization, obtained from Tables 3.2 and 3.3, was *CTNNAL1* (α-catenin). This catenin belongs to the Rho family, which is often involved in cell mobility and cell shape remodelling [98]. This gene could be up-regulated in the M2 cells, since, as mentioned before, these cells have more phagocytic activity [14] which would require the remodelling of the cell shape, a function that α-catenin takes part of.

The results for the samples stimulated with Aβ leave a lot to be desired. Almost all of the genes up-regulated in the Tables 3.2 and 3.3 were also present as up-regulated in Control comparisons. The only exception was the sole gene *SOD2*, which resulted from the comparison between the Control and Ab groups. The elimination of ROS (catalysing it into hydrogen peroxide and oxygen) from the system can be done through the SOD (superoxidase dismutase) family, making the members of this family antioxidant agents [99,100], and with one of its members being the mitochondrial *SOD2* [101]. Oxidative stress markers have been detected in brains affected with AD (post-mortem) [102] and the reduction of *SOD2* in animal models has been found to increase the levels of Aβ [103]. A link between the expression of Aβ and the high levels of oxidative stress exists, where Aβ induces the generation of ROS through several different reactions and *SOD2* has been shown to increase the resistance of cells towards this induced oxidative stress [104]. The fact that this is the only gene that has a higher level of expression in the Ab group of cells when compared with the Control group is to be expected, since the

cells of the Ab group had Aβ in their environment and, as stated before, *SOD2* tends to be the "go to" enzyme when it comes to the attainment of a resistance by cells to the oxidative stress caused by this oligomer. It has been shown that macrophages are capable of digesting Aβ fully [44], so one would expect a bigger percentage of Differential expressed genes related to this process of degradation. Curiously enough, M1 cells had a bigger expression of this gene when compared with all other conditions, including Aβ induced macrophages. This could be due to the fact that this cells produce ROI [14] and, as such, would require some sort of protection for themselves. An interesting thing should be noted, about the Ab group. The gene profile for this group was more similar to the M2 polarization, than for the M1 polarization, when comparing them through the PCA plot (Figure 3.6), and their number of genes differentially expressed (Table 3.1). This might indicate that these macrophages, in an Aβ environment, have more of a similarity with an anti-inflammatory profile than to a pro-inflammatory profile.

Due to the significant volume of data from the differential expression analysis, a functional enrichment analysis was done to try to determine the main functions different between each condition (Tables 3.4 and 3.5).

For the M1 cells, when compared with the Control group, there is the presence of the membrane cluster in both up and down-regulated groups. Usually this group is associated with receptors located in the membrane and it is possible to see that there is a greater expression of only some receptors, while others get less expressed. A few genes related to the DAPIN domain have also been detected as down-regulated in M1. Members of this domain include some interferon-induced proteins and others may play a role in apoptosis or the activation of MHC class II [105], so there seems to be a lack of expression of a small quantity of these genes in the M1 cells. As mentioned before, M1 cells are responsible for the production of pro-inflammatory cytokines and, as such, it is normal for some of the genes with greater expression in these cells to be related with the NOD-like signalling pathway. This pathway is responsible for the recognition of bacterial structures leading to the production of pro-inflammatory cytokines, by the processing of Interleukin-1 beta (IL-1β) and others into their mature form [106,107]. Finally, there is also the increased expression of some genes related to antigen processing and presentation, a function that is more efficient in this cells [14].

When the comparison is made with the M2 cells, there seems to be a significant up-regulation of genes associated with signalling. There is also the greater expression of gene associated with response to lipopolysaccharides, which is reasonable since the M1 cells were activated with LPS.

For the comparison between M1 and Aβ cells, the M1 cells seem to express more genes related with the interaction of viruses. There are some studies showing that viruses like HIV can promote the M1 polarization [108,109] and this cells have antiviral activity [110]. M1 cells have a higher expression of proteins with immunoglobulin-like fold, probably due to the high number of receptors and MHC class II antigens expressed in these cells. Cells from the Ab group also had some over-expressed genes associated with lectins. Lectins are glycan-binding receptors that have a role in some immune functions, for example, some C-type lectins participate in cell adhesion and pathogen recognition, initiating intracellular signalling processes that can result in the production of inflammatory cytokines and a phagocytic response [111]. It is possible, that these genes have a lower expression in the M1 due to the lack of usefulness of these specific lectins to the M1 cells, while other lectins retain normal expression levels. Finally, for this comparison, there seems to be a lower expression of genes associated with the metabolism of cholesterol, which might indicate that the M1 cells, when compared with the cells treated with Aβ, have a much smaller capacity for this process; or that the cells that processed Aβ have a much greater expression of genes related to this function. The

APOE gene is over-expressed in the Aβ treated cells when compared with M1, and this gene is related to the transport of cholesterol [112]. It is possible that a high level of cholesterol might cause a higher level of accumulation of the Aβ plaques [112] and, as such, the cells treated with Aβ started to adapt to process the cholesterol as an attempt to decrease the formation of the Aβ plaques.

In terms of the M2 cells, when the comparison is done against the Control condition, there is a low expression of a small number of genes related with chemotaxis and that are a part of the lysosome. Chemotaxis is done by the M2 cells as this cells recruit others through the presence of chemokines and M2 cells, due to their high phagocytic activity, will need good lysosomal activity [14]. The genes that have a lower expression in these M2 polarized macrophages might not be specialized or related to the specific functions that M2 polarized cells can do. They promote the chemotaxis of some cells that are of no interest to M2 cells for example. On the other hand, there are a high number of genes widely expressed in these cells with function in the membrane (most likely receptors, which these cells possess in great quantities) and a smaller number of genes related to adaptive immunity and host-virus interaction. Regarding the latter, it has been shown, that in the presence of a virus, there will be a bigger polarization towards M2 cells rather than M1 cells [113,114], so this cells may play a much bigger role in the immune response to virus than the M1 polarized cells, especially when it comes to fixing the damage done in tissue by this virus. These proteins may also hint to a bigger susceptibility towards the infection of this M2 cells by virus.

There seems to be a big expression of genes associated with DNA housekeeping and mRNA expression control in the M2 cells (when compared with M1). The reason for this can be the fact that, under polarizing conditions, certain epigenetic changes are done within the cell.

When comparing the M2 cells with the Aβ stimulated cells, there is, like in the M1 cells, a bigger expression of a lot of receptors (Membrane and Immunoglobulin-like fold). There is also some expression of MHC class II proteins that, as noted in the Table 1.1, is also present in this type of polarization. The Aβ treated cells seem to have a bigger expression of glycoproteins and signal peptides than the M2 cells. The signal peptide term probably means that there are a lot of proteins that assign this signal and are involved in the transportation of the proteins towards the secretory pathway, meaning that there might be a bigger amount of production of proteins in these cells when compared to the M2 cells.

4.2 miRNA data analysis

In order to compare the miRNA expression profiles between all of the conditions, a differential expression analysis was done.

As shown before (Table 3.6), there were no miRNA detected when comparing the Control and Ab conditions. However, the number of miRNA up-regulated in Control or Ab is different when this two are compared with the M1 or M2 polarized cells. Like stated beforehand, for the mRNA data, these two conditions are slightly different, but the difference is not significant enough to show up when comparing them directly.

Three miRNAs (hsa-miR-505-3p, hsa-miR-146a-5p and hsa-miR-147b) were uniquely up-regulated for the Ab conditions. These miRNAs were identified by comparing the results for gene up-regulated in CON or Ab when compared with M1 and M2 conditions. Of these three, only hsa-miR-505-3p was not present in other comparisons, meaning that it's less likely that its up-regulation is due

to a low expression in the M1 or M2 samples. This miRNA has been identified in several studied related to cancer [115–117], but nothing that would show an association of this miRNA with Alzheimer's disease. Unfortunately, as stated before, while this miRNA has a different expression for the Ab group, this difference is not significant enough to show up in the statistical testing against the Control group and, as such, no tangible conclusions can be made from focusing on this miRNA.

Due to the poor results found for the Ab group, the rest of this analysis was focused in the polarization states M1/M2. Table 3.8 represents all of the up-regulated genes for the M1 and M2 groups.

Of the 24 up-regulated miRNAs for the M1 condition across all comparisons, 9 (miR-155-5p, -146a-5p, -147b, -146a-3p, -155-3p, -125a-3p, -130a-3p, -187-3p, -181d-5p and -125a-5p) have been identified as present in this cells in previous studies [51,118–123]. Others were not directly associated with the polarization into M1 but with other factors: miR-29b and miR-22 may have elevated expression due to their role in the differentiation of monocytes into macrophages [119,124]; miR-29a has been found to be enriched in hematopoietic cells [125]; miR-200a-3p expression can be high following the activation of the *TLR1* signalling pathway [126]. For macrophages treated with LPS, the expression of miR-27a as often been found to be lower in this cells [127,128], but the opposite was seen in the data that was analysed for this work. It was up-regulated in the M1 cells and the reason for this is yet to be determined. In mice on a high-fat diet, miR-130b has been shown to increase the polarization of macrophages towards M1, dampening M2 polarization [129] but, under LPS conditions, the expression of this miRNA is decreased [130], which is the opposite of what happened in this case.

For the M2 condition, of the 12 miRNAs with high levels of expression across all comparisons, 4 (miR-193b-3p, -193b-5p, -511-3p and -511-5p) have been associated, in other studies, with this polarization state [118,119]. Other miRNAs have been detected without association with the M2 polarization state. Ones includes miR-132, that was found to have increased expression in M2b macrophages activated using IgG-coated wells and LPS [119]. The opposite of which happens in the data analysed for this work, seeing as miR-132-3p and -132-5p have high expression for the M2a polarized macrophages activated with IL-4; miR-342-5p has a role in the regulation of the polarization into the M1 state [131].

It should be noted that, these discrepancies between miRNAs detected in each polarization states and the defined markers or experimental observations can possibly be caused by a lot of factors. Different species, conditions, differences in what was used to cause polarization, among other reasons. There is also the fact that miRNAs form a complicated web of interactions and, as such, whilst some experiments may have the expression of a particular miRNA, others may have a much lower expression and as a result of, for example, a negative feedback loop that had time to take effect due to the difference in timing of the acquisition of the samples.

Either way, the miRNAs with the highest expression are the ones previously associated as markers. Taking the comparison M1vsM2 as an example, the most expressed gene for M1 is miR-155-5p (4.18 log2 Fold Change) and for M2 are miR-511-5p (-3.32 log2 Fold Change) and 193b-3p (-3.01 log2 Fold Change). As such, this could be taken as proof of a successful polarization of the cells, making the miRNAs that have not been studied thoroughly (like miR-449a, -4521, -449c-5p, etc.), interesting subjects in the study of polarization and immune systems. Especially if later their over-expression can be validated by techniques such as RT-qPCR.

4.3 Correlation test analysis

A correlation analysis was done to evaluate the effect of miRNA expression over their target mRNA expression. However, this analysis was plagued with the skewing of the number of positive correlations by miR-34a-5p (Tables 3.9 and 3.10). One would normally expect the prevalence of negative correlations, due to the nature of miRNA.

In terms of the abundance of positive correlations in the miR-34a-5p miRNA, it's possible to assess that its maximum values are associated with the Control and Ab conditions, and its minimum values are associated with the M1 and M2 condition (data not shown). So, this miRNA has a smaller expression in the polarization states, when compared with the other conditions.

For a gene to have a positive correlation, it means that it will follow the same trend. The gene will also have a bigger expression in the CON and Ab conditions, and a smaller expression in the M1 and M2 conditions. As seen, there are a great number of gene targets with positive correlation for miR-34a-5p. When taking into account the difference between polarization states and non-polarization states, it was hypothesised that this could be due to the fact that, the effect of the polarization on the genes will be greater than the effect of the regulation by the miRNA. If that was the case, then majority of positive correlations could be associated with a massive amount of target genes identified that take a role in macrophage polarization.

To test this hypothesis, all of the genes associated with positive correlations for this miRNA were ran through Reactome [132,133] to check for any common pathways, shared by the majority of these genes. If a common pathway was involved in the polarization of the macrophages, then the hypothesis would be proven to be correct.

When analysing the data in which all the evidence were used, the one with the bigger skewing, there was no common pathway shared by the majority of the genes. However, even when the top pathways have such a lower number of genes associated with them than expected, the conjugation of some of them may be a by-product of macrophage polarization. The top pathways for this case are associated with cellular senescence, the wnt pathway, reactions involving histones, chromatin organization, and DNA methylation. The chromatin organization and the methylation of DNA could be caused by the polarization, due to the epigenetic changes that happens in these macrophages and the wnt pathway has been proven to be active in M2 macrophages [134]. Whilst this does not prove the proposed hypothesis, especially because it does not justify the low expression in M1/M2 that the genes involved in this pathway must have, it does seem to show the slight trend of the difference between the Con/Ab and polarized conditions.

In regards of the data with "Strong evidence", there are 8 of the 17 genes with positive correlations, which participate in the NOTCH pathway. This specific pathway has been found to be inhibited in M2 macrophages and, when active, induces and increases the number of M1 macrophages [135,136]. It's possible that, while this difference in activation exists, M1 macrophages can possibly have lower levels of expression for genes associated with this pathway, when compared with its Con/Ab counterparts, yet still be active.

In conclusion, the purposed hypothesis was not proven to be correct. However, there are certain interesting trends in the genes that are positively correlated with miR-34a-5p. There is also the possibility of other unknown factors being in play. As stated before, miRNA form complex networks of interaction, to the point that often, when a majority of negative correlations is expected, the result ends up being evenly distributed between positive and negative. There is also the possibility that a lot

of targets for this miRNA may be regulated by other miRNA, thus influencing its expression, and creating a similar trend of expression between miR-34a-5p and its targets. Regardless, the skewing is due to only this particular miRNA, which is quite unusual.

Chapter 5: Conclusion

The main conclusion of this work is that, the polarization of the macrophages into M1 and M2 was successful as shown by the detection of gene and miRNA markers. Additionally, some new miRNAs were identified for these conditions and they could be some interesting targets for study of inflammation and therapeutics.

Unfortunately, the same could not be said for the macrophage samples stimulated with Aβ. These samples were too similar to the Control group for any relevant and new information to be found. Whilst for the mRNA, there was one gene outputted, the lack of miRNAs differentially expressed was not ideal for this line of study. It is possible that this lack of results may be a consequence of the strong variability caused either by human error in sample preparation, or by some other unknown factor. The sva package improved the results greatly, but it was still insufficient.

In regards to the correlation test, the biggest surprise was the skewing towards positive correlation due to one single miRNA, rather than the majority of positive correlations. Some other works have shown some strong and abundant positive correlations, contrary to what it is usually expected, but they are not usually caused by a single miRNA. It just goes to show, if everything done in this work was correct, that at the end of the day, these miRNAs are much more complex than what is now known in the literature.

Future Work

The present work was a good foundation to try to identify and compare the gene profiles of macrophages stimulated with Aβ against polarized macrophages/Control samples. However, as previously mentioned, the work was plagued with the fact that the Ab group and the Control group were too similar. As such, the experiment should be repeated, with especial caution towards batch effects and contamination in the preparation of the libraries.

In addition to that, and to expand on this study, analyses like RT-qPCR could be done in both the miRNA and mRNA data, to verify the values of expression.

References

1. Parkin J, Cohen B. An overview of the immune system. Lancet. 2001;357:1777–89.

2. Chaplin DD. Overview of the immune response. J. Allergy Clin. Immunol. [Internet]. Elsevier Ltd; 2010;125:S3–23. Available from: http://dx.doi.org/10.1016/j.jaci.2009.12.980

3. Tauber AI. Metchnikoff and the phagocytosis theory. Nat. Rev. Mol. Cell Biol. 2003;4:897–901.

4. Hume DA. Differentiation and heterogeneity in the mononuclear phagocyte system. Mucosal Immunol. [Internet]. 2008;1:432–41. Available from: http://www.nature.com/doifinder/10.1038/mi.2008.36

5. Dunster JL. The macrophage and its role in inflammation and tissue repair: Mathematical and systems biology approaches. Wiley Interdiscip. Rev. Syst. Biol. Med. 2016;8:87–99.

6. Zhang L, Wang CC. Inflammatory response of macrophages in infection. Hepatobiliary Pancreat. Dis. Int. [Internet]. The Editorial Board of Hepatobiliary & Pancreatic Diseases International; 2014;13:138–52. Available from: http://dx.doi.org/10.1016/S1499-3872(14)60024-2

7. Varol C, Mildner A, Jung S. Macrophages: Development and Tissue Specialization [Internet]. Annu. Rev. Immunol. 2015. Available from: http://www.annualreviews.org/doi/10.1146/annurev-immunol-032414-112220

8. Epelman S, Lavine KJ, Randolph GJ. Origin and Functions of Tissue Macrophages. Immunity [Internet]. Elsevier Inc.; 2014;41:21–35. Available from: http://dx.doi.org/10.1016/j.immuni.2014.06.013

9. Labonte AC, Tosello-Trampont A-C, Hahn YS. The Role of Macrophage Polarization in Infectious and Inflammatory Diseases. Mol. Cells [Internet]. 2014;37:275–85. Available from: http://www.molcells.org/journal/view.html?doi=10.14348/molcells.2014.2374

10. Mantovani A, Biswas SK, Galdiero MR, Sica A, Locati M. Macrophage plasticity and polarization in tissue repair and remodelling. J. Pathol. 2013;229:176–85.

11. Liu Y-C, Zou X-B, Chai Y-F, Yao Y-M. Macrophage Polarization in Inflammatory Diseases. Int. J. Biol. Sci. [Internet]. 2014;10:520–9. Available from: http://www.ijbs.com/v10p0520.htm

12. Freilich RW, Woodbury ME, Ikezu T. Integrated Expression Profiles of mRNA and miRNA in Polarized Primary Murine Microglia. Khoury J El, editor. PLoS One [Internet]. 2013;8:e79416. Available from: http://dx.plos.org/10.1371/journal.pone.0079416

13. Murray PJ, Allen JE, Biswas SK, Fisher EA, Gilroy DW, Goerdt S, et al. Macrophage Activation and Polarization: Nomenclature and Experimental Guidelines. Immunity [Internet]. Elsevier; 2014;41:14–20. Available from: http://dx.doi.org/10.1016/j.immuni.2014.06.008

14. Biswas SK, Mantovani A. Macrophage plasticity and interaction with lymphocyte subsets : cancer as a paradigm. Nat. Immunol. [Internet]. Nature Publishing Group; 2010;11:889–96. Available from: http://dx.doi.org/10.1038/ni.1937%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/20856220

15. Chung Y, Hong JY, Lei J, Chen Q, Bentley JK, Hershenson MB. Rhinovirus Infection Induces Interleukin-13 Production from CD11b-Positive, M2-Polarized Exudative Macrophages. Am. J. Respir. Cell Mol. Biol. [Internet]. 2015;52:205–16. Available from:

http://www.atsjournals.org/doi/abs/10.1165/rcmb.2014-0068OC

16. Shen B, Liu X, Fan Y, Qiu J. Macrophages Regulate Renal Fibrosis Through Modulating TGFβ Superfamily Signaling. Inflammation [Internet]. 2014;37:2076–84. Available from: http://link.springer.com/10.1007/s10753-014-9941-y

17. Ju C, Mandrekar P. Macrophages and Alcohol-Related Liver Inflammation. Alcohol Res. [Internet]. 2015;37:251–62. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4590621&tool=pmcentrez&rendertype=abstract

18. Heneka MT, Carson MJ, Khoury J El, Landreth GE, Brosseron F, Feinstein DL, et al. Neuroinflammation in Alzheimer's disease. Lancet Neurol. [Internet]. 2015;14:388–405. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1474442215700165

19. Röszer T. Transcriptional control of apoptotic cell clearance by macrophage nuclear receptors. Apoptosis [Internet]. Springer US; 2017;22:284–94. Available from: http://link.springer.com/10.1007/s10495-016-1310-x

20. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. Cell. 2013;153.

21. Hung CW, Chen YC, Hsieh WL, Chiou SH, Kao CL. Ageing and neurodegenerative diseases. Ageing Res. Rev. [Internet]. Elsevier B.V.; 2010;9:S36–46. Available from: http://dx.doi.org/10.1016/j.arr.2010.08.006

22. Kirkwood TBL. Understanding the odd science of aging. Cell. 2005;120:437–47.

23. de Magalhães JP. The biology of ageing: A primer. In: Stuart-Hamilton I, editor. An Introd. to Gerontol. [Internet]. Cambridge: Cambridge University Press; 1994. p. 21–47. Available from: http://ebooks.cambridge.org/ref/id/CBO9780511973697A009

24. Sikora E, Bielak-Zmijewska A, Mosieniak G. Cellular senescence in ageing, age-related disease and longevity. Curr. Vasc. Pharmacol. [Internet]. 2014;12:698–706. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24350932

25. Childs BG, Durik M, Baker DJ, van Deursen JM. Cellular senescence in aging and age-related disease: from mechanisms to therapy. Nat. Med. [Internet]. 2015;21:1424–35. Available from: http://www.nature.com/doifinder/10.1038/nm.4000

26. Hayflick L, Moorhead PS. The serial cultivation of human diploid cell strains. Exp. Cell Res. [Internet]. 1961;25:585–621. Available from: http://linkinghub.elsevier.com/retrieve/pii/0014482761901926

27. Ladislas R. Cellular and molecular mechanisms of aging and age related diseases. Pathol. Oncol. Res. [Internet]. 2000;6:3–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10749581

28. Oishi Y, Manabe I. Macrophages in age-related chronic inflammatory diseases. npj Aging Mech. Dis. [Internet]. Nature Publishing Group; 2016;2:16018. Available from: http://www.nature.com/articles/npjamd201618

29. Castelo-Branco C, Soveral I. The immune system and aging: a review. Gynecol. Endocrinol. [Internet]. 2014;30:16–22. Available from: http://www.tandfonline.com/doi/full/10.3109/09513590.2013.852531

30. Montecino-Rodriguez E, Berent-Maoz B, Dorshkind K. Causes, consequences, and reversal of immune system aging. J. Clin. Invest. [Internet]. 2013;123:958–65. Available from: http://www.jci.org/articles/view/64096

31. Pinti M, Appay V, Campisi J, Frasca D, Fülöp T, Sauce D, et al. Aging of the immune system: Focus on inflammation and vaccination. Eur. J. Immunol. [Internet]. 2016;46:2286–301. Available from: http://doi.wiley.com/10.1002/eji.201546178

32. Salvioli S, Monti D, Lanzarini C, Conte M, Pirazzini C, Giulia Bacalini M, et al. Immune System, Cell Senescence, Aging and Longevity - Inflamm-Aging Reappraised. Curr. Pharm. Des. [Internet]. 2013;19:1675–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23061725%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/23589904

33. Cai Z, Hussain MD, Yan L-J. Microglia, neuroinflammation, and beta-amyloid protein in Alzheimer's disease. Int. J. Neurosci. [Internet]. 2014;124:307–21. Available from: http://www.tandfonline.com/doi/full/10.3109/00207454.2013.833510

34. Amemori T, Jendelova P, Ruzicka J, Urdzikova L, Sykova E. Alzheimer's Disease: Mechanism and Approach to Cell Therapy. Int. J. Mol. Sci. [Internet]. 2015;16:26417–51. Available from: http://www.mdpi.com/1422-0067/16/11/25961/

35. Hardy J, Higgins G. Alzheimer's disease: the amyloid cascade hypothesis. Science (80-.). [Internet]. 1992;256:184–5. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.1566067

36. Hardy J, Selkoe DJ. The Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics. Science (80-.). [Internet]. 2002;297:353–6. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.1072994

37. Sadigh-Eteghad S, Sabermarouf B, Majdi A, Talebi M, Farhoudi M, Mahmoudi J. Amyloid-Beta: A Crucial Factor in Alzheimer's Disease. Med. Princ. Pract. [Internet]. 2015;24:1–10. Available from: https://www.karger.com/?doi=10.1159/000369101

38. Barage SH, Sonawane KD. Amyloid cascade hypothesis: Pathogenesis and therapeutic strategies in Alzheimer's disease. Neuropeptides [Internet]. Elsevier Ltd; 2015;52:1–18. Available from: http://dx.doi.org/10.1016/j.npep.2015.06.008

39. Hunter S, Arendt T, Brayne C. The Senescence Hypothesis of Disease Progression in Alzheimer Disease: an Integrated Matrix of Disease Pathways for FAD and SAD. Mol. Neurobiol. [Internet]. 2013;48:556–70. Available from: http://link.springer.com/10.1007/s12035-013-8445-3

40. Golde TE, Miller VM. Proteinopathy-induced neuronal senescence: a hypothesis for brain failure in Alzheimer's and other neurodegenerative diseases. Alzheimers. Res. Ther. [Internet]. 2009;1:5. Available from: http://alzres.biomedcentral.com/articles/10.1186/alzrt5

41. Wyss-Coray T, Rogers J. Inflammation in Alzheimer Disease--A Brief Review of the Basic Science and Clinical Literature. Cold Spring Harb. Perspect. Med. [Internet]. 2012;2:a006346–a006346. Available from: http://perspectivesinmedicine.cshlp.org/lookup/doi/10.1101/cshperspect.a006346

42. Gold M, El Khoury J. β-amyloid, microglia, and the inflammasome in Alzheimer's disease. Semin. Immunopathol. [Internet]. 2015;37:607–11. Available from: http://link.springer.com/10.1007/s00281-015-0518-0

43. Li Y, Tan M-S, Jiang T, Tan L. Microglia in Alzheimer's Disease. Biomed Res. Int. [Internet]. Hindawi

Publishing Corporation; 2014;2014:1–7. Available from: http://www.hindawi.com/journals/bmri/2014/437483/

44. Majumdar A, Chung H, Dolios G, Wang R, Asamoah N, Lobel P, et al. Degradation of fibrillar forms of Alzheimer's amyloid β-peptide by macrophages. Neurobiol. Aging [Internet]. 2008;29:707–15. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0197458006004593

45. Finotello F, Di Camillo B. Measuring differential gene expression with RNA-seq: challenges and strategies for data analysis. Brief. Funct. Genomics [Internet]. 2015;14:130–42. Available from: https://academic.oup.com/bfg/article-lookup/doi/10.1093/bfgp/elu035

46. Mantione KJ, Kream RM, Kuzelova H, Ptacek R, Raboch J, Samuel JM, et al. Comparing Bioinformatic Gene Expression Profiling Methods: Microarray and RNA-Seq. Med. Sci. Monit. Basic Res. [Internet]. 2014;20:138–41. Available from: http://www.basic.medscimonit.com/abstract/index/idArt/892101

47. de Magalhães JP, Finch CE, Janssens G. Next-generation sequencing in aging research: Emerging applications, problems, pitfalls and possible solutions [Internet]. Ageing Res. Rev. 2010. p. 315–23. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1568163709000749

48. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. [Internet]. 2009;10:57–63. Available from: http://www.nature.com/doifinder/10.1038/nrg2484

49. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of best practices for RNA-seq data analysis. Genome Biol. [Internet]. 2016;17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26813401

50. ZHANG Y, ZHANG M, ZHONG M, SUO Q, LV K. Expression profiles of miRNAs in polarized macrophages. Int. J. Mol. Med. [Internet]. 2013;31:797–802. Available from: https://www.spandidos-publications.com/

51. Lu L, McCurdy S, Huang S, Zhu X, Peplowska K, Tiirikainen M, et al. Time Series miRNA-mRNA integrated analysis reveals critical miRNAs and targets in macrophage polarization. Sci. Rep. [Internet]. Nature Publishing Group; 2016;6:37446. Available from: http://www.nature.com/articles/srep37446

52. Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, et al. Ensembl 2016. Nucleic Acids Res. [Internet]. 2016;44:D710–6. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkv1157

53. Karnati HK, Panigrahi MK, Gutti RK, Greig NH, Tamargo IA. miRNAs: Key Players in Neurodegenerative Disorders and Epilepsy. J. Alzheimer's Dis. [Internet]. 2015;48:563–80. Available from: http://www.medra.org/servlet/aliasResolver?alias=iospress&doi=10.3233/JAD-150395

54. Griffiths-Jones S. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res. [Internet]. 2006;34:D140–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16381832

55. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. miRBase: tools for microRNA genomics. Nucleic Acids Res. [Internet]. 2008;36:D154–8. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkm952

56. Kozomara A, Griffiths-Jones S. miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. [Internet]. 2011;39:D152–7. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkq1027

57. Chou C-H, Chang N-W, Shrestha S, Hsu S-D, Lin Y-L, Lee W-H, et al. miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. Nucleic Acids Res. [Internet]. 2016;44:D239–47. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkv1258

58. Simon Andrews. FastQC: a quality control tool for high throughput sequence data [Internet]. 2010. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

59. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics [Internet]. 2016;32:3047–8. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btw354

60. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics [Internet]. 2013;29:15–21. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/bts635

61. Engström PG, Steijger T, Sipos B, Grant GR, Kahles A, Alioto T, et al. Systematic evaluation of spliced alignment programs for RNA-seq data. Nat. Methods [Internet]. 2013;10:1185–91. Available from: http://www.nature.com/doifinder/10.1038/nmeth.2722

62. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics [Internet]. 2009;25:2078–9. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp352

63. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat. Biotechnol. [Internet]. 2011;29:24–6. Available from: https://www.broadinstitute.org/igv/node/250

64. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. [Internet]. 2013;14:178–92. Available from: https://academic.oup.com/bib/article-lookup/doi/10.1093/bib/bbs017

65. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics [Internet]. 2015;31:166–9. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btu638

66. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics [Internet]. 2014;30:923–30. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btt656

67. R Core Team. R: a language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2016. Available from: https://www.r-project.org/

68. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. [Internet]. 2014;15:550. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8

69. Bourgon R, Gentleman R, Huber W. Independent filtering increases detection power for high-throughput experiments. Proc. Natl. Acad. Sci. [Internet]. 2010;107:9546–51. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.0914005107

70. Schurch NJ, Schofield P, Gierliński M, Cole C, Sherstnev A, Singh V, et al. How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? RNA [Internet]. 2016;22:839–51. Available from: http://rnajournal.cshlp.org/lookup/doi/10.1261/rna.058339.116

71. Leek JT. svaseq: removing batch effects and other unwanted noise from sequencing data. Nucleic Acids Res. [Internet]. 2014;42:e161–e161. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gku864

72. Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE, et al. Tackling the widespread and critical impact of batch effects in high-throughput data. Nat. Rev. Genet. [Internet]. 2010;11:733–9. Available from: http://www.nature.com/doifinder/10.1038/nrg2825

73. Huang DW, Sherman BT, Lempicki R a. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. [Internet]. 2008;4:44–57. Available from: http://www.nature.com/doifinder/10.1038/nprot.2008.211

74. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. [Internet]. 2009;37:1–13. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkn923

75. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal [Internet]. 2011;17:10–2. Available from: http://journal.embnet.org/index.php/embnetjournal/article/view/200

76. Joshi N, Fass J. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files [Internet]. 2011. p. 2011. Available from: https://github.com/najoshi/sickle

77. Simon Andrews. SeqMonk [Internet]. Available from: http://www.bioinformatics.babraham.ac.uk/projects/seqmonk/

78. Ponting CP, Belgard TG. Transcribed dark matter: meaning or myth? Hum. Mol. Genet. [Internet]. 2010;19:R162–8. Available from: https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddq362

79. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods [Internet]. 2008;5:621–8. Available from: http://www.nature.com/doifinder/10.1038/nmeth.1226

80. Mazarei G, Leavitt BR. Indoleamine 2,3 Dioxygenase as a Potential Therapeutic Target in Huntington's Disease. J. Huntingtons. Dis. [Internet]. 2015;4:109–18. Available from: http://www.medra.org/servlet/aliasResolver?alias=iospress&doi=10.3233/JHD-159003

81. Gonzalez-Pena D, Nixon SE, Southey BR, Lawson MA, McCusker RH, Hernandez AG, et al. Differential Transcriptome Networks between IDO1-Knockout and Wild-Type Mice in Brain Microglia and Macrophages. Harrison JK, editor. PLoS One [Internet]. 2016;11:e0157727. Available from: http://dx.plos.org/10.1371/journal.pone.0157727

82. Singh N, Sundar S. Inflammatory chemokines and their receptors in human visceral leishmaniasis: Gene expression profile in peripheral blood, splenic cellular sources and their impact on trafficking of inflammatory cells. Mol. Immunol. [Internet]. Elsevier Ltd; 2017;85:111–9. Available from: http://dx.doi.org/10.1016/j.molimm.2017.02.008

83. Pinto ML, Rios E, Silva AC, Neves SC, Caires HR, Pinto AT, et al. Decellularized human colorectal cancer matrices polarize macrophages towards an anti-inflammatory phenotype promoting cancer cell invasion via CCL18. Biomaterials [Internet]. 2017;124:211–24. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0142961217300728

84. Michelucci A, Cordes T, Ghelfi J, Pailot A, Reiling N, Goldmann O, et al. Immune-responsive gene 1 protein links metabolism to immunity by catalyzing itaconic acid production. Proc. Natl. Acad. Sci.

[Internet]. 2013;110:7820–5. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3651434&tool=pmcentrez&rendertype=abstract%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/23610393%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3651434

85. Li Y, Zhang P, Wang C, Han C, Meng J, Liu X, et al. Immune Responsive Gene 1 (IRG1) Promotes Endotoxin Tolerance by Increasing A20 Expression in Macrophages through Reactive Oxygen Species. J. Biol. Chem. [Internet]. 2013;288:16225–34. Available from: http://www.jbc.org/lookup/doi/10.1074/jbc.M113.454538

86. Venner JM, Famulski KS, Badr D, Hidalgo LG, Chang J, Halloran PF. Molecular Landscape of T Cell-Mediated Rejection in Human Kidney Transplants: Prominence of CTLA4 and PD Ligands. Am. J. Transplant. [Internet]. 2014;14:2565–76. Available from: http://doi.wiley.com/10.1111/ajt.12946

87. Qin B, He M, Chen X, Pei D. Sorting Nexin 10 Induces Giant Vacuoles in Mammalian Cells. J. Biol. Chem. [Internet]. 2006;281:36891–6. Available from: http://www.jbc.org/lookup/doi/10.1074/jbc.M608884200

88. You Y, Zhou C, Li D, Cao Z-L, Shen W, Li W-Z, et al. Sorting nexin 10 acting as a novel regulator of macrophage polarization mediates inflammatory response in experimental mouse colitis. Sci. Rep. [Internet]. Nature Publishing Group; 2016;6:20630. Available from: http://www.nature.com/articles/srep20630

89. Joshi YB, Giannopoulos PF, Praticò D. The 12/15-lipoxygenase as an emerging therapeutic target for Alzheimer's disease. Trends Pharmacol. Sci. [Internet]. 2015;36:181–6. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0165614715000152

90. Kwon H-J, Kim S-N, Kim Y-A, Lee Y-H. The contribution of arachidonate 15-lipoxygenase in tissue macrophages to adipose tissue remodeling. Cell Death Dis. [Internet]. Nature Publishing Group; 2016;7:e2285. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27362803

91. Lee Y-H, Kim S-N, Kwon H-J, Maddipati KR, Granneman JG. Adipogenic role of alternatively activated macrophages in β-adrenergic remodeling of white adipose tissue. Am. J. Physiol. - Regul. Integr. Comp. Physiol. [Internet]. 2016;310:R55–65. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26538237%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4796635

92. Gordon S, Martinez FO. Alternative Activation of Macrophages: Mechanism and Functions. Immunity [Internet]. Elsevier Inc.; 2010;32:593–604. Available from: http://dx.doi.org/10.1016/j.immuni.2010.05.007

93. Conrad DH, Ford JW, Sturgill JL, Gibb DR. CD23: An overlooked regulator of allergic disease. Curr. Allergy Asthma Rep. [Internet]. 2007;7:331–7. Available from: http://link.springer.com/10.1007/s11882-007-0050-y

94. Eligini S, Fiorelli S, Tremoli E, Colli S. Inhibition of transglutaminase 2 reduces efferocytosis in human macrophages: Role of CD14 and SR-AI receptors. Nutr. Metab. Cardiovasc. Dis. [Internet]. Elsevier B.V; 2016;26:922–30. Available from: http://dx.doi.org/10.1016/j.numecd.2016.05.011

95. Hoppe B, Häupl T, Skapenko A, Ziemer S, Tauber R, Salama A, et al. Fibrinogen and factor XIII A-subunit genotypes interactively influence C-reactive protein levels during inflammation. Ann. Rheum. Dis. [Internet]. 2012;71:1163–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22267327

96. Stein M, Keshav S, Harris N, Gordon S. Interleukin 4 potently enhances murine macrophage mannose receptor activity: a marker of alternative immunologic macrophage activation. J. Exp. Med.

[Internet]. 1992;176:287–92. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2119288&tool=pmcentrez&rendertype=abstract

97. Raes G, Brys L, Dahal BK, Brandt J, Grooten J, Brombacher F, et al. Macrophage galactose-type C-type lectins as novel markers for alternatively activated macrophages elicited by parasitic infections and allergic airway inflammation. J. Leukoc. Biol. [Internet]. 2005;77:321–7. Available from: http://www.jleukbio.org/cgi/doi/10.1189/jlb.0304212

98. Wiesner C, Winsauer G, Resch U, Hoeth M, Schmid JA, van Hengel J, et al. α-Catulin, a Rho signalling component, can regulate NF-κB through binding to IKK-β, and confers resistance to apoptosis. Oncogene [Internet]. 2008;27:2159–69. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17952117

99. Fridovich I. Biological effects of the superoxide radical. Arch. Biochem. Biophys. [Internet]. 1986;247:1–11. Available from: http://linkinghub.elsevier.com/retrieve/pii/0003986186905266

100. Abreu IA, Cabelli DE. Superoxide dismutases—a review of the metal-associated mechanistic variations. Biochim. Biophys. Acta - Proteins Proteomics [Internet]. Elsevier B.V.; 2010;1804:263–74. Available from: http://dx.doi.org/10.1016/j.bbapap.2009.11.005

101. Fukai T, Ushio-Fukai M. Superoxide Dismutases: Role in Redox Signaling, Vascular Function, and Diseases. Antioxid. Redox Signal. [Internet]. 2011;15:1583–606. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21473702%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3151424

102. Flynn JM, Melov S. SOD2 in mitochondrial dysfunction and neurodegeneration. Free Radic. Biol. Med. [Internet]. 2013;62:4–12. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0891584913002372

103. Li F, Calingasan NY, Yu F, Mauck WM, Toidze M, Almeida CG, et al. Increased plaque burden in brains of APP mutant MnSOD heterozygous knockout mice. J. Neurochem. [Internet]. 2004;89:1308–12. Available from: http://doi.wiley.com/10.1111/j.1471-4159.2004.02455.x

104. Dumont M, Wille E, Stack C, Calingasan NY, Beal MF, Lin MT. Reduction of oxidative stress, amyloid deposition, and memory deficit by manganese superoxide dismutase overexpression in a transgenic mouse model of Alzheimer's disease. FASEB J. [Internet]. 2009;23:2459–66. Available from: http://www.fasebj.org/content/23/8/2459%5Cnhttp://www.fasebj.org/content/23/8/2459.full.pdf%5Cnhttp://www.fasebj.org/content/23/8/2459.short%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/19346295%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=271778

105. Staub E, Dahl E, Rosenthal A. The DAPIN family: a novel domain links apoptotic and interferon response proteins. Trends Biochem. Sci. [Internet]. 2001;26:83–5. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0968000400017175

106. Futosi K, Fodor S, Mócsai A. Neutrophil cell surface receptors and their intracellular signal transduction pathways. Int. Immunopharmacol. [Internet]. The Authors; 2013;17:638–50. Available from: http://dx.doi.org/10.1016/j.intimp.2013.06.034

107. Griebel T, Maekawa T, Parker JE. NOD-like receptor cooperativity in effector-triggered immunity. Trends Immunol. [Internet]. Elsevier Ltd; 2014;35:562–70. Available from: http://dx.doi.org/10.1016/j.it.2014.09.005

108. Cassol E, Cassetta L, Alfano M, Poli G. Macrophage polarization and HIV-1 infection. J. Leukoc. Biol. [Internet]. 2010;87:599–608. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20042468

109. Cassetta L, Cassol E, Poli G. Macrophage Polarization in Health and Disease. Sci. World J. [Internet]. 2011;11:2391–402. Available from: http://www.hindawi.com/journals/tswj/2011/213962/

110. Zdrenghea MT, Makrinioti H, Muresan A, Johnston SL, Stanciu LA. The role of macrophage IL-10/innate IFN interplay during virus-induced asthma. Rev. Med. Virol. [Internet]. 2015;25:33–49. Available from: http://doi.wiley.com/10.1002/rmv.1817

111. Drickamer K, Taylor ME. Recent insights into structures and functions of C-type lectins in the immune system. Curr. Opin. Struct. Biol. [Internet]. Elsevier Ltd; 2015;34:26–34. Available from: http://dx.doi.org/10.1016/j.sbi.2015.06.003

112. Reiss AB, Voloshyna I. Regulation of Cerebral Cholesterol Metabolism in Alzheimer's Disease. J. Investiagative Med. 2012;60:576–82.

113. Ohtsuki T, Kimura K, Tokunaga Y, Tsukiyama-Kohara K, Tateno C, Hayashi Y, et al. M2 Macrophages Play Critical Roles in Progression of Inflammatory Liver Disease in Hepatitis C Virus Transgenic Mice. Diamond MS, editor. J. Virol. [Internet]. 2016;90:300–7. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4702575&tool=pmcentrez&rendertype=abstract

114. Bility MT, Cheng L, Zhang Z, Luan Y, Li F, Chi L, et al. Hepatitis B Virus Infection and Immunopathogenesis in a Humanized Mouse Model: Induction of Human-Specific Liver Fibrosis and M2-Like Macrophages. Walker CM, editor. PLoS Pathog. [Internet]. 2014;10:e1004032. Available from: http://dx.plos.org/10.1371/journal.ppat.1004032

115. Jiang Y, Gao H, Liu M, Mao Q. Sorting and biological characteristics analysis for side population cells in human primary hepatocellular carcinoma. Am. J. Cancer Res. [Internet]. 2016;6:1890–905. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27725897

116. Ninomiya M, Kondo Y, Funayama R, Nagashima T, Kogure T, Kakazu E, et al. Distinct MicroRNAs Expression Profile in Primary Biliary Cirrhosis and Evaluation of miR 505-3p and miR197-3p as Novel Biomarkers. Ansari AA, editor. PLoS One [Internet]. 2013;8:e66086. Available from: http://dx.plos.org/10.1371/journal.pone.0066086

117. Fricke A, Ullrich P V, Heinz J, Pfeifer D, Scholber J, Herget GW, et al. Identification of a blood-borne miRNA signature of synovial sarcoma. Mol. Cancer [Internet]. Molecular Cancer; 2015;14:151. Available from: http://dx.doi.org/10.1186/s12943-015-0424-z

118. Cobos Jimenez V, Bradley EJ, Willemsen AM, van Kampen AHC, Baas F, Kootstra NA. Next-generation sequencing of microRNAs uncovers expression signatures in polarized macrophages. Physiol. Genomics [Internet]. 2014;46:91–103. Available from: http://physiolgenomics.physiology.org/cgi/doi/10.1152/physiolgenomics.00140.2013

119. Graff JW, Dickson AM, Clay G, McCaffrey AP, Wilson ME. Identifying Functional MicroRNAs in Macrophages with Polarized Phenotypes. J. Biol. Chem. [Internet]. 2012;287:21816–25. Available from: http://www.jbc.org/lookup/doi/10.1074/jbc.M111.327031

120. Su S, Zhao Q, He C, Huang D, Liu J, Chen F, et al. miR-142-5p and miR-130a-3p are regulated by IL-4 and IL-13 and control profibrogenic macrophage program. Nat. Commun. [Internet]. Nature Publishing Group; 2015;6:8523. Available from: http://www.nature.com/doifinder/10.1038/ncomms9523

121. LIN L, LIN H, WANG L, WANG B, HAO X, SHI Y. miR-130a regulates macrophage polarization and is associated with non-small cell lung cancer. Oncol. Rep. [Internet]. 2015;34:3088–96. Available from: https://www.spandidos-publications.com/

122. Moschos SA, Williams AE, Perry MM, Birrell MA, Belvisi MG, Lindsay MA. Expression profiling in vivo demonstrates rapid changes in lung microRNA levels following lipopolysaccharide-induced inflammation but not in the anti-inflammatory action of glucocorticoids. BMC Genomics [Internet]. 2007;8:240. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1940008&tool=pmcentrez&rendertype=abstract

123. Su R, Lin H-S, Zhang X-H, Yin X-L, Ning H-M, Liu B, et al. MiR-181 family: regulators of myeloid differentiation and acute myeloid leukemia as well as potential therapeutic targets. Oncogene [Internet]. 2015;34:3226–39. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25174404

124. Shen C, Chen M-T, Zhang X-H, Yin X-L, Ning H-M, Su R, et al. The PU.1-Modulated MicroRNA-22 Is a Regulator of Monocyte/Macrophage Differentiation and Acute Myeloid Leukemia. Grimes HL, editor. PLOS Genet. [Internet]. 2016;12:e1006259. Available from: http://dx.plos.org/10.1371/journal.pgen.1006259

125. Roy S. miRNA in Macrophage Development and Function. Antioxid. Redox Signal. [Internet]. 2016;25:795–804. Available from: http://online.liebertpub.com/doi/10.1089/ars.2016.6728

126. Wang Y, Xu G, Han J, Xu T. miR-200a-3p regulates TLR1 expression in bacterial challenged miiuy croaker. Dev. Comp. Immunol. [Internet]. Elsevier Ltd; 2016;63:181–6. Available from: http://dx.doi.org/10.1016/j.dci.2016.06.004

127. Xie N, Cui H, Banerjee S, Tan Z, Salomao R, Fu M, et al. miR-27a Regulates Inflammatory Response of Macrophages by Targeting IL-10. J. Immunol. [Internet]. 2014;193:327–34. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24835395

128. Ma S, Liu M, Xu Z, Li Y, Guo H, Ge Y, et al. A double feedback loop mediated by microRNA-23a/27a/24-2 regulates M1 *versus* M2 macrophage polarization and thus regulates cancer progression. Oncotarget [Internet]. 2016;7:13502–19. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=prem&NEWS=N&AN=26540574

129. Zhang M, Zhou Z, Wang J, Li S. MiR-130b promotes obesity associated adipose tissue inflammation and insulin resistance in diabetes mice through alleviating M2 macrophage polarization via repression of PPAR-γ. Immunol. Lett. [Internet]. Elsevier B.V.; 2016;180:1–8. Available from: http://dx.doi.org/10.1016/j.imlet.2016.10.004

130. Zheng H, Dong X, Liu N, Xia W, Zhou L, Chen X, et al. Regulation and mechanism of mouse miR-130a/b in metabolism-related inflammation. Int. J. Biochem. Cell Biol. [Internet]. Elsevier Ltd; 2016;74:72–83. Available from: http://dx.doi.org/10.1016/j.biocel.2016.02.021

131. Zhou D, Yang K, Chen L, Zhang W, Xu Z, Zuo J, et al. Promising landscape for regulating macrophage polarization: epigenetic viewpoint. Oncotarget [Internet]. 2017; Available from: http://www.oncotarget.com/fulltext/17027

132. Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, et al. The Reactome pathway Knowledgebase. Nucleic Acids Res. [Internet]. 2016;44:D481–7. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkv1351

133. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, et al. The Reactome pathway knowledgebase. Nucleic Acids Res. [Internet]. 2014;42:D472–7. Available from:

https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkt1102

134. Cosín-Roger J, Ortiz-Masiá D, Calatayud S, Hernández C, Álvarez A, Hinojosa J, et al. M2 Macrophages Activate WNT Signaling Pathway in Epithelial Cells: Relevance in Ulcerative Colitis. Mizoguchi E, editor. PLoS One [Internet]. 2013;8:e78128. Available from: http://dx.plos.org/10.1371/journal.pone.0078128

135. Wang Y-C, He F, Feng F, Liu X-W, Dong G-Y, Qin H-Y, et al. Notch Signaling Determines the M1 versus M2 Polarization of Macrophages in Antitumor Immune Responses. Cancer Res. [Internet]. 2010;70:4840–9. Available from: http://cancerres.aacrjournals.org/cgi/doi/10.1158/0008-5472.CAN-10-0269

136. Zhang J, Zhou Q, Yuan G, Dong M, Shi W. Notch signaling regulates M2 type macrophage polarization during the development of proliferative vitreoretinopathy. Cell. Immunol. [Internet]. Elsevier Inc.; 2015;298:77–82. Available from: http://dx.doi.org/10.1016/j.cellimm.2015.09.005